# THE SOCIAL, CULTURAL, EPISTEMOLOGICAL AND TECHNICAL BASIS OF THE CONCEPT OF "PRIVATE DATA."

A thesis submitted to the University of Manchester for the degree of

PhD

in the Faculty of Humanities

2011

Karen Marie Mc Cullagh

School of Social Sciences

# Declaration

A declaration stating:

that no part of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

## Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on Presentation of Theses.

The University of Manchester

Karen Marie Mc Cullagh

Doctor of Philosophy
(PhD)

# THE SOCIAL, CULTURAL, EPISTEMOLOGICAL AND TECHNICAL BASIS OF THE CONCEPT OF "PRIVATE DATA."

(2011)

Abstract

In July 2008, the UK Information Commissioner launched a review of EU Directive 95/46/EC on the basis that:

> "European data protection law is increasingly seen as out of date, bureaucratic and excessively prescriptive. It is showing its age and is failing to meet new challenges to privacy, such as the transfer of personal details across international borders and the huge growth in personal information online. It is high time the law is reviewed and updated for the modern world."[1]

Legal practitioners such as Bergkamp have expressed a similar sense of dissatisfaction with the current legislative approach:

> "Data Protection as currently conceived by the EU is a fallacy. It is a shotgun remedy against an incompletely conceptualised problem. It is an emotional, rather than rational reaction to feelings of discomfort with expanding data flows. The EU regime is not supported by any empirical data on privacy risks and demand…A future EU privacy program should focus on actual harms and apply targeted remedies."[2]

Accordingly, this thesis critiques key concepts of existing data protection legislation, namely 'personal' and 'sensitive' data, in order to explore whether current data protection laws can simply be amended and supplemented to manage privacy in the information society. The findings from empirical research will demonstrate that a more radical change in EU law and policy is required to effectively address privacy in the digital economy. To this end, proposed definitions of data privacy and private data was developed and tested through semi-structured interviews with privacy and data protection experts. The expert responses indicate that Bergkamp *et al*[3] have indeed identified a potential future direction for privacy and data protection, but that further research is required in order to develop a coherent definition of privacy protection based on managing risks to personal data, and harm from misuse of such information.

[1] ICO Press Release "UK privacy watchdog spearheads debate on the future of European privacy Law" <http://www.ico.gov.uk/upload/documents/pressreleases/2008/ico_leads_debate_070708.pdf > (Last accessed:07.07.08)

[2] Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects of Europe's Data protection Policy in an Information Driven Economy" *Computer Law & Security Report*, Vol. 18 No.1 pp. 31-14, p. 31

[3] Other proponents of reform include industry experts such as Google's Global Privacy Counsel, Peter Fleischer <http://peterfleischer.blogspot.com/2007/09/eric-schmidt-on-global-privacy.html> (Last accessed 26.03.09)

# Table of Contents

Thesis wordcount 75,691

# Tables & Figures

# Cases and Legislation

**Legislation**

APEC Privacy Framework (APEC, 2005)

Charter of Fundamental Rights of the European Union (2000)

Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data (CoE, 1981)

Data Protection Act, (DPA,1998)

Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Directive 95/46/EC)

European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR, 1950)

International Covenant on Civil and Political Rights (ICCPR, 1966)

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (OCED, 1980).

UN Guidelines for the Regulation of Computerized Personal data files (UN Guidelines1990)

Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community (2007)

Universal Declaration of Human Rights (UNDHR, 1948)

**Cases**

*Bodil Lindqvist v Kammaraklagaren* (2003) C-101/01

*Caparo Industries plc v Dickman* [1990] 2 AC 605

*Donoghue v Stevenson* [1932] UKHL 100

*Durant v FSA* [2003] EWCA Crim 1746

*Hedley Byrne & Co Ltd v Heller & Partners Ltd* [1964] AC 465

*Janvier v Sweeney* [1919] 2 KB 316

Klass and Others v Federal Republic of Germany [1979] 2 EHRR 214:

*Leander v Sweden* [1987] 9 EHRR 433

*Malone v. Commissioner of Police*, (1979) 2 All E.R. 620

*Willcock v Muckle* [1951] 2 The Times LR 373

*Wilkinson v Downton* [1897] 2 QB 57

*X v Iceland* (1976) 5 DR 86

# Chapter 1

# Introduction

## 1. Introduction

This thesis examines how privacy is understood and defined both in legal terms and how it operates in peoples' lives. It considers whether existing privacy and data protection laws adequately protect individuals' privacy, or whether fundamental reform of EU data protection laws may be required to effectively address privacy in the Information Society. A mixed method approach to data collection was employed; including two surveys and a series of semi- structured interviews with key stakeholders both in the UK and internationally. The data findings are used to support recommendations for legislative reform, so that data protection laws meet the challenges posed by the data disclosure-privacy conundrum in the information society. [4]

### *1.1 Rationale for privacy protection*

Karvalics[5] asserts that during the 20th century most developed countries in the world gradually developed into information societies and that within the coming decade the majority of the world's population will be living and working in a global information society.[6] Marshall defines an information society as:

> "A society in which low cost information technology, computers, and telecommunications are widely used to facilitate communication nationally and internationally, and to promote access to libraries, data archives, and other stores of information held by private organizations or in the public domain."[7]

[4] This research study is prescient, as in 2009 the EU Commission announced a review the Data Protection Directive 95/46/EC. European Commission, "Review of the data protection legal framework," <http://ec.europa.eu/justice/policies/privacy/review/index_en.htm>

[5] Karvalics, L. (2007) "Information Society – what is it exactly? (The meaning, history and conceptual framework of an expression)" <http://www.ittk.hu/netis/doc/ISCB_eng/02_ZKL_final.pdf> , p.21

[6] *Ibid*, p. 15

[7] Marshall, G. (1998) "Information society," A Dictionary of Sociology, <http://www.encyclopedia.com/doc/1O88-informationsociety.html>

In an information society an enormous amount of data about individuals is generated, disclosed, collected and processed. Marshall argues that:

> "this greater facility of communication and increased access to information creates a qualitatively different society with attendant new problems, such as information overload, and the need for new forms of regulation to control information flows between persons, companies, and countries."[8]

Similarly, the House of Lords reported that:

> "The combination of information technology and high speed communications is breaking down the traditional barriers to the movement of information (distance, location, time and volume) at an unprecedented rate."[9]

The beneficial purposes for which access to data is sought in an information society are many and diverse: establishing and maintaining intimate relationships, selling or providing products or services, healthcare and welfare provision, security and public safety, crime prevention and investigation, the administration of justice, the war against terrorism, and so on. Indeed, Rule *et al* contend that no area of human life is inherently too private to attract the application of bureaucratic surveillance. Rather, most sensitive and personal aspects of life are most associated with social uncertainties that render systematic monitoring and control attractive.

> "People yield all sorts of embarrassing or otherwise sensitive information to medical personnel as one of the costs of modern medical care."[10]

Many of these data disclosures lead to positive outcomes for the individual data subjects; for instance benefits can be achieved in government activity (e.g. tax and social welfare) when the general populace supports administrations that make decisions based on a wealth of discriminatory personal information which allows it to:

> "Render to each person his or her 'due,' that is, the correct form of bureaucratic action in light of all relevant information on that person's history and current status."[11]

---

[8] *Ibid*

[9] House of Lords Select Committee on Science and Technology (1996) Information Society: Agenda for Action in the UK, 5th Report, HL Paper 77, <http://www.parliament.the-stationery-office.co.uk/pa/ld199596/ldselect/inforsoc/ch1.htm>

[10] Rule, J. B., Mc Adam, D., Stearns, L.D. in Johnson, D. G. & Nissenbaum, H. (1995) *Computers, Ethics & Social Values*, (Prentice Hall), p.318

[11] *Ibid*, p.315

Similarly, it allows businesses to offer convenient, customised, personalised services, and financial rewards for customer loyalty. Yet, irrespective of the beneficence of such outcomes they also create privacy concerns. Strahilevitz encapsulated the data disclosure-privacy conundrum, when he stated:

> "When asked to imagine the most private facts about ourselves, we will typically think of sexual encounters and bodily functions, sensitive medical information, shameful past misdeeds, unfavourable opinions about peers...most of us would regard the disclosure of these details to our entire circle of acquaintances, let alone the public at large, as a personal disaster. At the same time, no one among us has guarded that embarrassing information with maximum diligence. Certain indubitably "private" acts, such as sexual intercourse, necessarily take place in the presence of at least one other person. Other facts might be created in solitude, but remain, by common parlance, "private" even when shared to some extent. We all tell some people about our medical ailments...We are, in short, constantly disclosing embarrassing information about ourselves to third parties, yet we often harbour strong subjective expectations of privacy when doing so."[12]

A wide range of privacy concerns about the disclosure and processing of individuals data can be espoused. These include physical harm, e.g. parents do not want information about their children to be freely available online in case it is viewed by paedophiles. Also, the misuse of information can cause emotional or psychological harm in the form of annoying, irritating, and unwanted intrusions in daily lives. These include the unwanted phone calls and spam emails or cyberstalking.[13] The misuse of an individual's data also can cause economic harm, such as denial of credit, or even a job, based on inaccurate or incomplete information. In extreme cases, the misuse of information also can lead to identity theft. Moreover, there are also concerns about loss of rights with respect to autonomy, and about the conversion of individuals into 'data subjects' to be used for purposes or ends that they have not set for themselves. It is the potential for such adverse consequences from personal information misuse that drives concerns about privacy. The term privacy is defined the Oxford English Dictionary as:

> "The state or condition of being alone, undisturbed, or free from public attention, as a matter of choice or right; seclusion; freedom from interference or intrusion…absence or avoidance of publicity or display; secrecy, concealment, discretion; protection from public knowledge or availability."[14]

---

[12] Strahilevitz, L. J. (2005) "A Social Networks Theory of Privacy," *The University of Chicago Law Review*, Vol. 72, No. 3, pp. 919-988

[13] Wright, S. (2007) "Cyber stalker left me living in terror, says victim of 7/7," <http://www.dailymail.co.uk/news/article-458109/Cyber-stalker-left-living-terror-says-victim-7-7.html> <http://www.dailymail.co.uk/news/article-461450/Bloggers-help-track-cyberstalker-harassed-7-7-survivor.html> (Last accessed 16.07.09)

[14] Oxford English Dictionary, <http://www.oed.com/>

Likewise, information to which privacy obligations attach, is described as 'private,' which is defined as:

> "Kept or removed from public view or knowledge; secret; concealed… Of a conversation, communication, etc.: intended only for or confined to the person or persons directly concerned; confidential...Of a person: intimate or confidential (*with* a person); sexually intimate...Relating to or connected with activities restricted to one person or a few people."[15]

Thus, private data is information which an individual would seek to limit or control disclosure of. Accordingly, the challenge for law makers is to enact conceptually coherent laws for regulating data disclosures that also afford privacy protection, since without a meaningful legislative framework for understanding privacy and private data, decision makers will have great difficulty identifying and protecting individuals from deleterious incursions on privacy.

## *1.2 Subject matter and aims of the thesis*

The subject matter of this thesis is the collection of laws which seek to allow the processing of personal information whilst safeguarding the privacy of individuals' information. Collectively, the laws are commonly referred to as 'data protection' in European jurisdictions, whereas the term 'privacy protection' is employed in other jurisdictions such as Australia, Canada, New Zealand and the USA.[16] However, a brief overview of the development of these laws will serve the current discussion.[17] Such laws are comparatively new additions to the global legal landscape. The first privacy and data protection laws were not introduced at international level until the 1940's and at national level until the 1970's.[18] At the International level, in the post World War II era, the United Nations established the right to privacy in Art 12 of the Universal Declaration of Human Rights (1948):

> "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

[15] Oxford English Dictionary, <http://www.oed.com/>
[16] These designations are sometimes problematic since data protection law is often confused with copyright protection of databases in the EU Database Directive 96/9/EC
[17] See Chapter 3 for a detailed discussion.
[18] The first national Data Protection Act was enacted in Hesse, Germany in 1970.

Since then, a variety of international, European and national data protections laws have been enacted. The OECD Guidelines of 1980 and the Council of Europe Convention (hereafter CoE Convention) of 1981,[19] established standards among European member countries to ensure the free flow of information among them without infringing personal privacy. The CoE convention states this objective in Art 1:

> "The purpose of this convention is to secure in the territory of each party for every individual, whatever his nationality or residence, respect for his rights and fundamental freedoms, and in particular his right to privacy, with regard to automatic processing of personal data relating to him ("data protection")."

At the national level, the UK's first Data Protection Act was introduced in 1984. The Act required public and private organisations with access to computer-held personal data to register with a Data Protection Registrar.[20] It did not, however, explicitly recognise an individual's right to privacy. That changed with the enactment of the Data Protection Act 1998, which built on an EC Directive of 1995 and was introduced with the explicit aim of protecting the right to privacy. Thus, the focus of this thesis is an analysis of Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data[21] (hereafter Directive 95/46/EC) and, as a corollary, the Data Protection Act 1998. The Directive sought to harmonize data protection laws throughout the EU member states, and was implemented in response to the development of a frontier-free Internal market and of the so-called 'information society' which resulted in an increase in the flow of personal data between Member States and beyond. As stated previously, the main aim of the Directive is to safeguard the privacy of an individual when information about them is processed by others. Hence, Art 1(1) states:

> "In accordance with this Directive, Member States shall protect the fundamental rights and freedoms of natural persons, and in particular their right to *privacy*, with respect to the processing of personal data (emphasis added)."

The Directive has its origins in the European Convention on Human Rights (ECHR) 1950, Article 8 (1) of which states:

---

[19] Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, No. 108, 1981 <http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm> (Last accessed 24.07.09)

[20] In January 2001, the Data Protection Registrar's Office was given the added responsibility of the Freedom of Information Act 2000 and changed its name to the Information Commissioner's Office. <http://www.ico.gov.uk/about_us/who_we_are/history_of_ico_page.aspx> (Last accessed 23.07.09)

[21] Directive 95/46/EC is a European Union directive legislating protection of data pertaining to individuals. It was implemented in 1995 by the European Commission.

> "Everyone has the right to respect for his *private* and family life, his home and his correspondence."

Thus, *prima facie,* data protection laws claim to protect privacy and, as a corollary, private data, i.e. pertaining to an individual's private life, even though neither of the terms is defined in the ECHR (1950) or Directive 95/46/EC. Consequently, this thesis will contend that the privacy protection rationale of data protection laws has been too readily accepted without analysis of its conceptual adequacy, or its ability to aid in the generation of harmonized regulatory measures. Indeed, effective data protection is only possible if the terms that are defined in the laws are conceptually certain, interpreted uniformly, and applied equally; yet the term privacy is not defined in the Directive or other legislative measures.[22] Thus, the central aim of this thesis is to explore how the terms found in these laws embody or intersect with a concern for privacy protection (however defined),[23] and as a corollary of this, whether private data is adequately protected through current data protection laws.

## *1.3 Rationale for review of Directive 95/46/EC*

The rationale for reviewing the provisions of current laws stems from survey evidence that despite the existence of Directive 95/46/EC for the last sixteen years, public concern about privacy protection in the UK has increased, which *prima facie* suggests that privacy concerns are not allayed by the existing laws. A Eurobarometer survey[24] commissioned by the European Commission indicated that despite the introduction of the Directive in 1995, the level of concern about data protection has only changed slightly since the early 1990s. Two-thirds of respondents were concerned about this in 1991. The level of concern decreased between 1991 and 1996 from 66% to 58%. However, it increased insignificantly in 2003 to 60%. Yet, by 2008, 68% of respondents were concerned, which is similar to pre-Directive levels of concern. Also, the ICO Annual Track survey of UK citizens[25] indicates that the percentage of respondents who consider that protecting people's personal information is an issue of social importance has increased from 70% in 2004 to 94% in 2009. Additionally, complex, inter-related

[22] Napier, B. (1992) "International Data Protection Standards and British Experience," *Informatica e drittio*, Vols. 1-2, pp.83-100.
[23] See Chapter 2 for a detailed analysis of privacy conceptions.
[24] Gallup Organisation, (2005) Flash Eurobarometer No 225: Data Protection in the European Union - Citizens' Perceptions <http://ec.europa.eu/public_opinion/flash/fl_225_en.pdf> (Last accessed: 27.07.09)
[25] SMSR (2009) Report on the Findings of the Information Commissioner's Office Annual Track 2009 Individuals <http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/ico_annual_tracking_individuals_final_report2009.pdf >

developments during the past fifteen years, (since the introduction of Directive 95/46/EC) make a review of data protection laws both warranted and critical. According to the Canadian Privacy Commissioner, Dr Cavoukian:

> "The world has less than a decade to make the protection of personal information and online privacy a priority before the concepts are lost forever." She contends that reform is needed because "legislation meant to safeguard privacy already can't keep pace with the flow of information and advances in technology." She also claims that "current legislation, which reacts to the problem after it has presented itself, doesn't work and needs to be overhauled."[26]

Broadly, these developments can be grouped into three sets: *social & economic changes, legislative changes* and *technological & computing advancements.*

### *1.3.1 Social & economic changes*

Firstly, it is appropriate to give a historical overview of the social and political triggers that underpinned the development of data protection measures in Europe in order to illustrate how changes in social and economic conditions necessitate a review of the provisions of Directive 95/46/EC. Historically, nations within Europe had, over the centuries, engaged in acts of war with each other. This peaked with the battles of World War II and the persecution of human beings with particular characteristics e.g. individuals who were physically or mentally disabled because of their medical condition, ethnic Poles, Soviet civilians or political prisoners, Romany gypsies, Jews and people of black skin colour because of their racial or ethnic origin, communists, socialists and political prisoners because of their philosophical beliefs or political opinions, homosexuals because of their sexual orientation, and Jehovah Witnesses because of their religious beliefs.[27] Nevertheless, Davie observes that although:

> "In 1945 Europe had come close to self-destruction for the second time in a century. The idea of European unity was barely conceivable as individual nations struggled to rebuild the fabric of their devastated societies. Surprisingly quickly, however, the seeds of a European Community began to germinate in the form of Coal and Steel Agreements, which embodied the principle that the weapons of war should themselves be subject to supranational, if not international, control. Since then (the mid 1950's) Europe has moved inexorably, if not very steadily, towards a greater common identity."[28]

---

[26] Pilieci, V. (2010) "World is losing grip on privacy: watchdog - Next decade will be crucial in protecting personal data," Ottawa Citizen, (17th August)

[27] Niewyk, D. & Nicosia, F. (2000) *The Columbia Guide to the Holocaust*, (Columbia University Press: New York), p.45

[28] Davie, G. (1994) *Religion in Britain Since 1945: Believing without Belonging*, (Wiley Blackwell: UK), p.2

As well as economic agreements, European countries became signatories of a legislative measure which explicitly protected human rights, namely the European Convention on Human Rights (ECHR) 1950, Article 8 (1) of which states:

> "Everyone has the right to respect for his *private* and family life, his home and his correspondence."

Furthermore, the characteristics associated with persecution during World War II formed the basis of special/sensitive data classifications in subsequent data protection laws, including Directive 95/46/EC. In parallel with this, the welfare state in the UK and similar schemes in other countries were created. With these developments it became routine for individuals to volunteer data in order to access societal or economic benefits. It has been a natural extension for Governments to seek to use this data about individuals to improve public services such as social security provision and tax administration. So much so, that processing of such data is now a key activity for government services, and this trend is being further accelerated by e-Government initiatives.[29] For instance, during the last decade, the UK government has sought to initiate data sharing[30] between various administrative departments in order to decrease data redundancy, reduce the risk of inconsistencies, and provide personalised 'one-stop-shops' for citizens. Thus, individuals have become accustomed to providing the public sector with personal data in return for economic or welfare benefits.

With the development of free-market economies, European countries experienced a further wave of significant economic change. For example, as Gillespie[31] reports, Trade Union membership in the UK was at its highest in the late 1970's when over half the workforce belonged to a union, but, since the Thatcher years (when legislative measures were introduced to weaken the power of trade unions) membership has dropped to around 26% of the UK workforce. Additionally, economic migrants have changed the population of the UK and other European countries. Indeed, the 2001 census, recorded a minority ethnic population of 4.6 million or 7.9 per cent of the total population of the United Kingdom.[32] Concomitantly, the UK has witnessed changes in the religious make up of its population. For instance, the percentage of Muslims has increased from

[29] See for instance, the European Commission's Interchange of Data between Administrations, Citizens and Consumers Programme: <http://ec.europa.eu/idabc/> (Last accessed: 27.07.09)
[30] Walport, M. & Thomas, R. (2008) Data Sharing Review Report <http://www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf> (Last accessed 27.07.09)
[31] Gillespie, A. (2007) *Foundations of Economics*, (OUP: UK)
[32] ONS (2003) Ethnicity Nugget <http://www.statistics.gov.uk/CCI/nugget.asp?ID=273>

0.4% in 1975 to 1.5% in 2001.[33] Accordingly, changes in social and economic conditions necessitate a review of the provisions of Directive, to assess whether the types of data listed in Arts 8(1) and Art 8(5) as *a priori* sensitive, remain sensitive, or whether changes in social and economic conditions have lessened the privacy sensitivity of such types of data.

### *1.3.2 Legislative developments*

Changes in post World War II UK society have been underpinned by a raft of legislative measures which prevent discrimination on the basis of the categories of sensitive data listed in the Directive. For instance, the Civil Partnership Act 2004 affords same sex couples the same property rights as heterosexual couples; the Disability Discrimination Act 1995 (as amended by the Equality Act 2006), made it unlawful to discriminate against disabled persons in connection with employment, the provision of goods, facilities and services or the disposal or management of premises; The Employment Equality (Religion or Belief) Regulations 2003 made it unlawful to discriminate against workers because of religion or similar belief; The Employment Equality (Sexual Orientation) Regulations 2003 made it unlawful to discriminate against workers because of their sexual orientation; whilst the Race Relations Act 1976 made it unlawful to treat a person less favourably than another on racial grounds in the areas of employment, education, and the provision of goods, facilities, services and premises. These cover grounds of race, colour, nationality (including citizenship), and national or ethnic origin. In the UK such measures have culminated in the establishment of the Equality and Human Rights Commission, a statutory body charged with the responsibility of protecting, enforcing and promoting equality across the seven "protected" grounds - age, disability, gender, race, religion and belief, sexual orientation and gender reassignment.[34] Consequently, in the 21st century, there is less scope for discrimination on the basis of the categories listed in Directive 95/46/EC, and as a result, one of the aims of this thesis it to review the categories of sensitive data set out in Art 8 (1) and Art 8 (5), in order to assess their utility and continuing relevance in data protection legislation sixty one years after the ECHR (1950) was drafted and sixteen years after Directive 95/46/EC was introduced.

[33] Crabtree, V. (2007) Religion in the United Kingdom: Diversity, Trends and Decline <http://www.vexen.co.uk/UK/religion.html> (Last accessed: 15.01.10)
[34] Equality & Human Rights Commission (2008) "Who we are and what we do," <http://www.equalityhumanrights.com/uploaded_files/who_we_are.pdf> (Last accessed: 15.01.10)

Also, Directive 95/46/EC has, to date, been the dominant legislative measure for data protection in Europe and acted as a prominent marker for data protection norms around the world; however, a new contender has recently emerged, namely the APEC Privacy Framework. The APEC Framework contains a set of principles for privacy protection agreed by twenty one, non-European, member states. Whilst the principles are broadly similar to Directive 95/46/EC, there are some important differences. Significantly, the APEC Privacy Framework contains a *harm* principle which is absent from the Directive. The concept of harm has gained currency among data protection and privacy experts including Solove[35] and Calo[36] who assert that harm is an under-theorized aspect of such laws; and the absence of it in current EU legislation has been criticized by Berkgamp, who asserts that:

> "Data Protection as currently conceived by the EU is a fallacy. It is a shotgun remedy against an incompletely conceptualised problem. It is an emotional, rather than rational reaction to feelings of discomfort with expanding data flows. The EU regime is not supported by any empirical data on privacy risks and demand… A future EU privacy program should focus on actual *harms* and apply targeted remedies"[37] (emphasis added)

Accordingly, it is appropriate to consider the merits of a harm principle in order to determine whether Directive 95/46/EC should be revised to incorporate it. Also, although it is too early to predict whether the APEC Framework will supplant Directive 95/46/EC as the pre-eminent data protection measure, it is worthy of consideration since arguably this development undermines the potential of the Directive, in that the goal of harmonized legal measures which facilitate transnational data transfers may not be realised through it.

### *1.3.3 Technological & computing advancements*

Since the enactment of Directive 95/46/EC there have been many advances in technologies that impact on privacy, for instance, the development of biometric technologies poses major challenges for the protection of privacy of the human body itself as they allow individuals to be identified by finger or iris scanning and facial, voice and gait recognition. These technologies are still in their infancy, but in the

[35] Solove, D. J. (2008) *Understanding Privacy*, (Harvard University Press)
[36] Calo, R. M. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp. 1132-1161
[37] Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects of Europe's Data protection Policy in an Information Driven Economy," *Computer Law & Security Report*, Vol. 18, No.1, pp.31-14, at p. 31

medium to long term, they could have profound implications for an individual's sense of privacy. Thus, Vitalis claims:

"At the dawn of the twenty-first century, there is no point in hiding from the fact that there are hard times ahead for data protection. It must face accelerating innovation with the emergence of new and, from a risk assessment standpoint, poorly understood control technologies, such as biometrics or RFID chips, which make it possible to track not only objects, but also their owners."[38]

Additionally, the development of the Internet[39] has generated new and difficult privacy issues. For instance, internet search engines (e.g. Google, Bing etc.) collect information about the terms users enter when conducting searches. As Bankston, staff attorney at the Electronic Frontier Foundation, remarked, this poses privacy implications since:

> "Your search history shows your associations, beliefs, perhaps your medical problems. The things you Google for define you. [...] data that's practically a printout of what's going on in your brain: What you are thinking of buying, who you talk to, what you talk about."[40]

A high profile instance of privacy invasion occurred in 2006, when AOL released a list of the Web search inquiries of 658,000 users on a website to academic researchers. Although the users were not personally identified in the data (search data was released using unique identity numbers rather than names), the logs contained enough information in some cases to discern an individual's identity. It culminated in the cause célèbre data trail revelation that AOL Searcher No. 441774 was Thelma Arnold, a 62-year-old widow who lived in Lilburn, Ga., frequently researched her friends' medical ailments and loved her three dogs.[41]

Similarly, information can be collected about how individuals interact with websites, and which other website they linked from (i.e. clickstream data). For example, clickstream recordings can collect information about a user's internet service provider

---

[38] Vitalis, A. (2008) "France" in Rule, J. & Greenleaf, G. *Global Privacy Protection: The First Generation*, (Edward Elgar: Cheltenham, UK), p. 140

[39] It is a type of super network; a worldwide collection of interconnected computer networks based on a set of standard communication protocols. The opening of the network to commercial interests began in 1988. The internet has several characteristics that make it difficult to control, or to trace the flow of data within it: it has no borders – it is not physically located in any one state and can be accessed from anywhere; It is not centrally owned or controlled; It is interactive and dynamic.

[40] Quoted in Mills, E. (2005) "Google Balances Privacy, Reach" C|Net News.com <http://news.com.com/Google+balances+privacy%2C+reach/2100-1032_3-5787483.html>

[41] Barbaro, M., Zeller, T., & Hansell, S. (2006) "A Face Is Exposed for AOL Searcher No. 4417749," New York Times, (9th August) <http://query.nytimes.com/gst/fullpage.html?res=9E0CE3DD1F3FF93AA3575BC0A9609C8B63&sec=&spon=&pagewanted=1>

(ISP), internet protocol (IP) address, computer software and hardware as that user navigates the site. It poses a threat to privacy because some ISPs sell users' clickstream data to companies who mine it and use it for profiling purposes, often without the knowledge or consent of individuals. Thus, Wong asserts that the Directive has several weaknesses because:

> "it was implemented at a time with no consideration of major technological developments such as the use of personal data on webpages and possible profiling techniques used and aggregated by companies such as clickstream data."[42]

The development of the internet has also resulted in individual citizens themselves collecting, managing and using personal data, e.g. through social networks, online dating sites and blogs. In particular, the phenomenon of blogging, which can involve individuals publishing information about themselves and others on the Internet poses new issues for the protection of privacy. A blog[43] is a personal online journal that is frequently updated. A key feature of a blog is that it may be available for general public consumption; unlike traditional diaries or journals, where the only anticipated reader was the writer of the diary. Zuckerberg, CEO of Facebook, perhaps the best known social networking site, has claimed that the rise of social media reflects changing privacy attitudes, in that:

> "When I got started in my dorm room at Harvard, the question a lot of people asked was, 'why would I want to put any information on the internet at all? Why would I want to have a website?' ...Then in the last 5 or 6 years, blogging has taken off in a huge way... People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people... [The] social norm [privacy] is just something that has evolved over time."[44]

As well as blogging about themselves, bloggers may post about third parties e.g. employers, co-workers, family members or friends. The third parties mentioned in blogs may not always be aware that material about them has been posted, and are unlikely to have given their consent. Hence, when they become aware that information about them

[42] Wong, R. (2005) "Privacy: charting its developments and Prospects," in Klang, M. & Murray, A. *Human Rights in the Digital Age*, (Glasshouse Press: UK) p. 159

[43] Search Oriented Architecture "Weblog Definition" Blogs are defined by their format: a series of entries posted to a single page in reverse-chronological order. <http://searchsoa.techtarget.com/sDefinition/0,,sid26_gci213547,00.html> (Last accessed: 20.07.09)

[44] Kirkpatrick, M. (2010) "Facebook's Zuckerberg Says The Age of Privacy is Over," Read Write Web, <http://www.readwriteweb.com/archives/facebooks_zuckerberg_says_the_age_of_privacy_is_ov.php> (Last accessed 11.01.10)

is accessible on the internet, they may feel that their privacy has been invaded, as information on blog posts can be stored permanently, viewed repeatedly and used for data profiling purposes.[45] It can also be digitally altered or taken out of context and given new meanings, including embarrassing or derogatory meanings. Thus, Floridi asserts that:

> "Because digital ICTs are radically modifying our informational environments, ourselves and our interactions, it would be naive to expect that informational privacy in the future will mean exactly what it meant in the industrial Western world in the middle of the last century."[46]

Given that the aforementioned technologies did not exist when Directive 95/46/EC was enacted it is critical to investigate whether the categories of sensitive data should be revised to include new categories such as clickstream and biometric data, or, whether an alternative approach to privacy protection should be adopted in the internet era.

## *1.4 Research questions*

The aim of this thesis is to investigate the concept of private data since Directive 95/46/EC claims to protect privacy, yet it defines neither privacy in general nor private data in particular.[47] Accordingly, the overarching research question in this thesis is *'what is private data?'* A four-fold approach is adopted to answer this question.

Firstly, the thesis explores the nature of the privacy interests and values which data protection laws promote. This requires two investigations:

(i) A literature review to identify the varied conceptions of privacy, followed by an analysis of whether any of these privacy conceptions are or indeed, should be, embodied in data protection laws. It will be shown that privacy is a generic term, and that the focus of data protection laws should be informational/data privacy, a narrower conception of privacy.

[45] Unless the blogger/social networker has taken specific steps to 'lock down' their privacy settings.
[46] Floridi, L. (2005) "The ontological interpretation of informational privacy," *Ethics and Information Technology*, Vol.7, No. 4, pp. 185–200, at p. 191
[47] It is important to note that this work is concerned with the private data of individuals, but not the data of private institutions. A church may regard the tithes it collects as a private matter, and the intellectual property of a pharmaceutical company may be considered private, but issues such as these, are not the focus of this research. This distinction reflects the provisions and approach of current legislation.

(ii) An investigation of whether privacy is still valued in the information society given the huge advances in technology and concomitant social and behavioural changes described above.

Secondly, since Directive 95/46/EC seeks to protect privacy by regulating the processing of *personal data*, this thesis explores whether the two concepts of *personal* and *private* data are synonymous or distinct. Personal data is defined in Art 2 as 'data relating to and permitting identification of individual natural, living persons' (henceforth referred to as individuals).[48] Current data protection laws consist of rules that regulate the various stages of processing of personal data. That is, these laws regulate the manner in which personal data is collected, registered, stored, disseminated, and used. Thus the second research question is: *is personal data synonymous with private data?*

Thirdly, the thesis examines the relationship between private data and the notion of sensitivity in order to answer the third research question: *'is sensitive data is synonymous with private data?*' Directive 95/46/EC contains provisions regulating the processing of an exhaustive list of special categories of data (commonly referred to as *sensitive* data), set out in Art 8 (1) and (5) on the basis that such categories of data inherently merit stricter processing conditions as they pose a greater privacy risk if misused. The exhaustive list approach is distinct from the approach taken in other legislative measures; for instance the OECD Guidelines do not include a concept of sensitive data, whilst the CoE Convention (108) contains a list which is intended to be exemplary rather than exhaustive. Accordingly, the thesis explores the utility of an exhaustive list, and examines the current categories of sensitive data in order to assess their continuing relevance and effectiveness for determining the conditions of data processing and affording privacy protection. This is important, as advances in science e.g. genetics, biometrics etc. and developments in society e.g. secularization, may influence perceptions of sensitivity of any given type of data. In concurrence with Raab & Bennett,[49] the findings from this research will demonstrate that an exhaustive list of sensitive data is a fallacy as *any* data may be considered sensitive, and present a privacy risk, depending on whose data it is, how it is used, and the context of that use.

[48] Directive 95/46/EC of the European Parliament and of the Council (24th October 1995) Official Journal of the European Communities, No 1, 281/31<http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf>; <http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part2_en.pdf> (Last accessed: 06.05.2011)

[49] Raab, C. & Bennett, C. (1998) "The Distribution of Privacy Risks: Who needs protection?," *The Information Society*, Vol. 14, No. 4, pp. 263-74, at p. 266

The final research question is: *Would a harm-based definition of private data be appropriate?* On the basis that the research findings indicate that neither personal nor sensitive data are completely synonymous with private data, the thesis investigates whether Directive 95/46/EC is still an effective tool for the protection of privacy, and explores what advantages could be gained through adopting an alternative approach. In particular, the thesis investigates whether it would be appropriate to develop a harm-based conception of private data, in line with the APEC Privacy Framework, which is emerging as a popular alternative legislative measure to Directive 95/46/EC amongst countries that have recently implemented privacy laws Accordingly, the thesis tests a harm-based definition of private data in order to propose and evaluate potential legislative reform proposals for the protection of privacy in the developing socio-technical context.

## *1.5 Research Approach & Methods*

The thesis is a primarily a piece of socio-legal scholarship. It is not a standard jurisprudential thesis, for it is not solely concerned with examining the basic nature of legal reasoning or legal concepts. Similarly, it is not solely a work of legal sociology, for it does not seek to systematically study the way in which data protection laws are actually practices. Rather, the thesis draws upon jurisprudence and legal sociology, as well as elements of other fields of study, including computer science, sociology, philosophy, psychology and political science.

Although the focus of the analysis is Directive 95/46/EC, and concomitantly, the Data Protection Act 1998 which implemented the Directive into domestic UK law, they are not the only legislative measures considered in this thesis; rather this thesis considers a range of international data protection measures as well as a variety of national laws. Two justifications for reviewing a variety of legislative measures in this research are offered. Firstly, a cross-national perspective is useful because the data protection laws of all EU member states are based upon a shared set of principles. Secondly, as the processing of personal data is increasingly a global occurrence, the way in which data processing is regulated should be considered on an international basis.

Accordingly, the thesis seeks to explore whether existing privacy and data protection laws protect private data or whether more radical change in EU data protection law may be required to effectively address privacy in the Information Society. Such an

assessment can only be made after empirical data has been collected and analyzed. In this thesis several different data collection methods are employed. Firstly, the literature review draws from three disciplines: law, sociology, and philosophy, whilst doctrinal analysis is employed to critique the development of case law. Secondly, both qualitative and quantitative research methods are used to gain primary, empirical research data in the form of a telephone survey of the British public, and an online survey of bloggers which examines their privacy attitudes and expectations, and semi-structured interviews with forty privacy and data protection experts. Thus, the thesis also seeks to discover the attitudes and expectations of potential data subjects on the basis that laws are, in practice, only successful if they have the support of the public, and meet societal needs.

## *1.6 Thesis Structure*

The remainder of this chapter describes how the thesis is set out and provides a description of the purpose of each chapter. It is important to note that this work is concerned with the private data of individuals, but not the data of private institutions. A church may regard the tithes it collects as a private matter, and the intellectual property of a pharmaceutical company may be considered private, but issues such as these, are not the focus of this research. Also, the thesis is not concerned with anonymised data about individuals, since this is not considered personal information.[50] This distinction reflects the provisions and approach of current legislation.

Chapter 2: One of the key objectives of data protection laws is the protection of privacy. In order to evaluate the effectiveness of data protection laws meeting this objective, it is first necessary to explore and unpack the concept of privacy, as only by determining what it is, and why it is valued, will it be possible to review and critique the adequacy of existing legislative measures. The discussion will reveal that there is no universally accepted definition of privacy or private data. The chapter concludes by drawing on the work of a number of academics, including the definition of privacy offered by Rossler, the framework of 'contextual integrity' developed by Nissenbaum,[51] the pragmatic, harm-based approach' developed by Solove,[52] and the categories of privacy harm

[50] There are arguments in the literature concerning when data is truly anonymous, but this is not relevant to this thesis.
[51] Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, Vol. 79, No. 1, pp. 119-158
[52] Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4. pp. 1087-1155.

developed by Calo[53] in order to develop definitions of '*data privacy'* and '*private data'* that will be used in the remainder of this thesis to evaluate the adequacy of current data protection laws and develop reform proposals.

Chapter 3: This chapter begins by exploring the rationale for the introduction of privacy and data protection laws; in particular, it discusses the historical catalysts for the enactment of such laws. Thereafter, it considers the relationship between privacy protection and data protection. This chapter will demonstrate that although one of the key objectives of current privacy and data protection laws is the protection of privacy, the term privacy is not defined within such laws. It will draw upon the literature in chapter two to illustrate that most human rights based privacy laws seek to protect a broad concept of privacy, whereas data protection laws seek to protect a narrower element of the concept, namely informational or data privacy. The chapter will conclude that the failure to define privacy in existing data protection laws is a major weakness, and draw upon the proposed definition of 'data privacy' and 'private data' offered in chapter two, in order to generate reform proposals.

Chapter 4: This chapter examines the terms personal and sensitive data, since data protection laws seek to protect privacy by imposing conditions on the processing of these types of data. It will begin by examining the term personal data, and illustrate that there is a broad consensus regarding the definition of personal data in supranational legislative measures, but that is it not synonymous with private data. Also, this chapter will examine the term sensitive data in supranational legislative measures, demonstrating that some supranational legislative measures specifically enumerate categories of sensitive data meriting special protection, whilst others advocate a context-based approach to privacy protection. It will explore the continuing relevance of legally recognized categories of sensitive data and suggest that changes in society and technological developments may influence the sensitivity of data, and give rise to new types of sensitive data. It will also explore whether sensitive data is synonymous with private data, and suggest that the two terms are not synonymous. The chapter will conclude that the conceptual adequacy of existing data protection laws has been too readily accepted, and further that technological developments warrant a review of the 'fitness for purpose' of the existing legislation.

[53] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp. 1132-1161

Chapter 5: This chapter discusses the research approach of the thesis; the researcher's position, and the models chosen for data generation. This includes a review of the innovative qualitative and quantitative data collection methods selected and the rationale behind those methods. The various stages of data collection are described as well as the methods of recruitment used to select data subjects, the sampling strategy, and the types of questions asked. Overall, a socio-legal approach was adopted. The research questions were framed and refined during a one week placement at the Office of the UK Information Commissioner. Thereafter, semi-structured interviews were conducted with forty privacy and data protection expert stakeholders, including: privacy commissioners, lawyers, corporate privacy officers, consultants, computer scientists, and academics from sociology, politics, market research, statistics and law, from a variety of continents and countries, including: Australia, Belgium, Canada, France, Germany, Italy, New Zealand, UK and the USA. The interview responses led to the development of a telephone survey question to test the attitudes of the British public regarding categories of sensitive data. The survey was a unique collaboration, supported and sponsored by the Office of the Information Commissioner, which culminated in the insertion of a survey question on a nationally representative telephone survey. Finally, an innovative online survey was employed to capture the privacy attitudes and expectations of bloggers.

Chapter 6: This chapter reports the findings of an online survey of bloggers from around the world. It explores the privacy attitudes and expectations of bloggers by examining their blogging practices and their expectations of privacy when publishing online. The main aim of the chapter is to examine whether, in the information society, privacy is still valued since industry experts such as Zuckerberg, CEO of Facebook, claim that as a result of individuals acting as information producers and processors, social norms are changing. Zuckerberg claims that the 'age of privacy is over' that is, individuals no longer value or seek to protect their personal privacy; rather they desire opportunities for maximum openness and disclosure.[54] The chapter concludes that privacy is valued by bloggers and that they actively take steps to protect their privacy.

Chapter 7: This chapter begins by examining the empirical data collected on sensitive data, in order to answer questions on three key areas of enquiry; questions about satisfaction and continuing relevance of existing categories in data protection laws, with

[54] Kirkpatrick, M. (2010) "Facebook's Zuckerberg Says The Age of Privacy is Over," Read Write Web, <http://www.readwriteweb.com/archives/facebooks_zuckerberg_says_the_age_of_privacy_is_ov.php> (Last accessed 11.01.10)

a particular focus on Directive 95/46/EC and the Data Protection Act 1998; questions about potential new categories of sensitive data; and questions about the relationship between the terms sensitive and private data. It will demonstrate that the legally recognised classifications of sensitive data are somewhat outdated and ineffective for determining the conditions of data processing and affording privacy protection. It will further demonstrate that technological developments are potentially giving rise to new categories of sensitive data. The chapter concludes that the terms sensitive and private are not synonymous and that the focus of the Directive and other data protection laws, namely the UN Guidelines (1990) and the CoE Convention 108, has erroneously been on classifying data as sensitive instead of regulating harm arising from data uses.

Chapter 8: This chapter begins by examining the empirical data collected on personal data, in order to answer questions on three key areas of enquiry: questions about interpretation and application of the term personal data in data protection laws; questions about the relationship between the terms personal and private data; and questions about the merits of a harm-based definition of private data. It will demonstrate that the term personal data has been widely defined in data protection laws and interpreted broadly since it is the subject matter of data processing activities, not all of which concern privacy protection, and, accordingly, the terms personal and private data are not synonymous. This chapter also tests the proposed definition of 'private' data offered in chapter two, and reports the responses of expert interviewees to the definition. The findings indicate that a harm-based conception of private data has merit and appeal, and that future amendments of Directive 95/46/EC and other data protection laws should focus on the risk of *unreasonable harm* posed by data processing.

Chapter 9: This chapter offers conclusions and recommendations informed by analysis of the empirical data collected during this research project. These contribute to the wider, ongoing legislative and policy debates surrounding data protection and privacy, at UK, EU and International level. In particular, this thesis recommends the assumptions and concepts underpinning current legislation are reviewed, and that any future global privacy and data law should focus on actual harms caused by the misuse of personal data and apply targeted remedies to such misuses.

## *1.7 Thesis Contributions*

This thesis makes several contributions to knowledge. Firstly, it enhances existing knowledge on 'privacy' by demonstrating that it is still valued in the information society. Thereafter, it demonstrates that current data protection laws refer to privacy, but this concept is too wide. Accordingly, this thesis contends that the focus of data protection laws should be data privacy. At present, the concept of data privacy is not explicitly stated in the legislation, and so one recommendation of this thesis is that Art 1 of Directive 95/46/EC should be amended to refer to data privacy instead of a generic concept of privacy. Secondly, the thesis makes a novel contribution through the collection and analysis of empirical data on the concepts of *personal* and *sensitive* data. The findings indicate that the current categories of sensitive data, drafted in a post-World War II era, are now in need of revision as some of the categories are no longer considered as sensitive as they once were, whilst changes in technology and computing have generated potential new types of sensitive data. Notwithstanding this, the thesis contends that classifying data as sensitive (or not) is a fallacy as all personal data can pose a privacy risk, dependent on the context of the data processing. Finally, the thesis contributes to the debate on private data and explores legislative reform proposals. The findings indicate that a harm-based conception of private data has merit and appeal, and that future amendments of the Directive should focus on the risk of *unreasonable harm* posed by data processing, as opposed to a personal–sensitive dichotomy.

# Chapter 2

## Privacy: a conceptual analysis

### *2.1 Introduction*

One of the key objectives of data protection laws is the protection of privacy. In order to evaluate the effectiveness of data protection laws meeting this objective, it is first necessary to explore and unpack the concept of privacy, as only by determining *w*hat it is, and why it is valued, will it be possible to review and critique the adequacy of existing legislative measures. Thus, this chapter commences by exploring the concept of privacy from a multidisciplinary perspective, drawing on work from sociology and philosophy and law, to examine what it is and why it is valued.

The chapter will demonstrate that, despite the importance attributed to privacy by the general public and many legal scholars, a universally accepted theory of privacy has yet to emerge. Instead, privacy scholars have offered numerous conceptualizations of privacy including the right to be let alone; limited access to the self; intimacy; secrecy; and control of personal information. These conceptualisations and their inadequacies will be examined. Thereafter, the chapter will conclude by drawing on the work of a number of academics, including the definition of privacy offered by Rossler, the framework of 'contextual integrity' developed by Nissenbaum,[55] the pragmatic, harm-based approach' developed by Solove,[56] and the categories of privacy harm developed

[55] Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, Vol. 79, No. 1, pp.119-157

[56] Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4, pp. 1087-1155

by Calo[57] in order to develop a conceptual approach to privacy (and in particular a definition of data privacy) that will be used in the remainder of this thesis to evaluate the adequacy of current data protection laws and develop reform proposals.

## *2.2. The value of privacy*

In the 19th century, philosopher John Stuart Mill outlined the value of privacy when he stated:

> "there is a sphere of action in which society, as distinguished from the individual, has, if any, only an indirect interest: comprehending all that portion of a person's life and conduct which affects only himself or, if it also affects others, only with their free, voluntary, and undeceived consent and participation. When I say only himself, I mean directly and in the first instance; for whatever affects himself may affect others through himself."[58]

In his view, privacy supports individuals' autonomy, and informed, uncoerced freedom to act; in the sense that individuals cannot act truly freely, unless they have some measure of knowledge of, and control over, interlocutors' knowledge of them. Accordingly, he contends that autonomous individuals require a private sphere to facilitate self-development, engage in intimate relationships, and express themselves freely. This view is supported by Moore, a sociologist, who postulates that privacy is valuable because it regulates individuals' relations with wider society:

> "the need for privacy is socially created. Without society there would be no need for privacy"[59]

Equally, Rubenfeld, a legal scholar, considers privacy valuable because it supports autonomy, that is:

> "the fundamental freedom not to have one's life too totally determined by a progressively more normalising state"[60]

Recently, these views regarding the value of privacy have been cogently supported by the philosopher Rossler, who asserts that individuals regard privacy as valuable because they regard autonomy as valuable, since autonomy allows individuals:

---

[57] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal,* Vol. 86, No. 3, pp.1132-1161
[58] Mill, J. S. (1859) *On Liberty And Other Essays*, Ch. 1 (World Classics Series, OUP)
[59] Moore, B. (1984) *Privacy: Studies in Social and Cultural History* (M E Sharpe, Armonk, NY), p.73
[60] Rubenfeld, J. (1989) "The Right of Privacy," in Slobogin, C. (2007) *Privacy at Risk: The New Government Surveillance and the Fourth Amendment* (University of Chicago Press) p.105

> "control over their self-presentation, that is, control of how they want to present or stage themselves, to whom they want to see themselves and how they want to be seen."[61]

Yet, for individuals to be autonomous they must have privacy since:

> "autonomy can only be lived out in all its aspects and articulated in all its senses with the help of the conditions of privacy and by means of rights and claims to privacy."[62]

In particular, she contends that the self or identity of an individual is constituted dialogically, in the sense that an individual is reliant upon symbolic interactions with other people in order to develop a sense of self-identity. This view is shared by the sociologist, Mead who proposed a theory of the self which distinguished between what he called the *'I'* and what he called the '*me.'* Mead viewed each as a fundamental and indispensable aspect of the self. He associated the *'me'* with 'social control,' and the '*I'* with 'self-expression.' In particular, he claimed that the '*I'* is the response of the organism to the attitudes of the others; the '*me'* is the organized set of attitudes of others which one himself assumes:

> "Taken together, they constitute a personality as it appears in social experience. The self is essentially a social process going on with these two distinguishable phases."[63]

Accordingly, people depend upon relationships that convey the love and self esteem essential to self development, for it is only in such relationships that the self-confidence necessary for self-identities to flourish actually be attained. She contends that such relationships can only occur within the protective sphere of privacy, since privacy allows an individual to be vulnerable or behave more vulnerably in relationships of their choosing. Similarly, Bloustein, a philosopher, claims that privacy is valuable because it facilitates personhood and autonomous action by protecting against conduct that is demeaning to individuality, an affront to personal dignity or an assault on human personality.[64] Rossler further suggests that:

> "Privacy can thus be seen as a protective shield allowing the individual to act towards all possible third parties, whether individual persons or institutions, in

[61] Rossler, B. (2005) *The Value of Privacy*, (Polity Press) p. 116
[62] *Ibid* p. 129
[63] Mead, G. H. (1964) *On Social Psychology* (University of Chicago Press), p. 238
[64] Bloustein, E. J. (1964) "Privacy as an Aspect of human Dignity: An Answer to Dean Prosser" *New York University Law Review*, p. 974.

> accordance with their expectations concerning the 'level of information' they each have."[65]

Rossler contends that individuals regulate their behaviour and social relations on the basis of the information they give other people about themselves or that they know other people have about them and that without this ability to exercise:

> "'controlled self-disclosure'- the self-chosen diversity in one's relations would not be possible. Nor, therefore, would self-determined, context-dependent, authentic behaviour towards others...Nor would it be possible to find an answer authentically to the question of how one wants to live."[66]

Similarly, Westin, a legal scholar, asserts that privacy provides the opportunity for individuals to autonomously regulate their behaviour and self-disclosures by allowing individuals to remove their social mask:

> "On any given day a man may move through the roles of stern father, loving husband, car-pool comedian, skilled lathe operator, union steward, water-cooler flirt, and American Legion committee chairman – all psychologically different roles that he adopts as he moves from scene to scene on the individual stage...Privacy...gives individuals, from factory workers to Presidents, a chance to lay their masks aside for rest. To be always 'on' would destroy the human organism."[67]

The important elements of choice and control regarding self-presentation are supported by the philosopher Benn when he states:

> "[R]espect for someone as a person, a chooser, implie[s] respect for him as one engaged on a kind of self-creative enterprise, which could be disrupted, distorted or frustrated even by so limited an intrusion as watching"[68]

Likewise, Rossler claims that if people have to assume on a structural and systematic basis that they will be observed, supervised and controlled, then:

> "this may produce a change in their interpretation of what self-created, self-determined behaviour, or in other words, autonomy can be in the first place."[69]

She observes that a lack of privacy would be highly problematic in normative terms, not only because of the connection between autonomy, authenticity, and fulfilling life, but also because liberal democracies are reliant upon subjects who are autonomous and who

---

[65] Rossler, B. (2005) *The Value of Privacy*, (Polity Press: Cambridge), p. 129
[66] *Ibid*, p. 116
[67] Westin, A. F. (1967) *Privacy and Freedom* (New York: Atheneum), pp. 34–35
[68] Benn, S. I. (1971) "Privacy, Freedom, and Respect for Persons," in Pennock, J. & Chapman, R. (Eds.), Nomos XIII: *Privacy*, (New York: Atherton Press), p. 26
[69] Rossler, B. (2005) *The Value of Privacy*, (Polity Press: Cambridge), p. 129

see themselves as such. Thus, liberal democracies necessarily have a substantial interest in self-determination, since they would otherwise be jeopardised in their very function; individuals need social and political privacy in order to be free to behave, and to associate with others, without the continual threat of being observed, since surveillance would chill behaviour and speech, and undermine democracy.

Therefore, privacy is considered valuable, and worthy of protection in western liberal democracies because it facilitates the autonomy of individuals and affords the psychological freedom of thought, presentation and action necessary for individuals to create and sustain both intimate and professional relationships. In particular, privacy offers individuals an environment in which they can share confidences and intimacies, and engage in limited and protected communication. In short, privacy is valuable for both individuals and wider society, and so, worthy of protection.

## *2.3 The Scope of privacy: public v private spheres*

What then is privacy? The terms 'privacy' and 'private' are widely used in everyday language as well as in philosophical and legal discussions, yet confusion remains over meaning of the terms. Accordingly, it is appropriate to examine the scope of these terms. Also, it is appropriate to examine how the term "private" may be distinguished from the term "public" because privacy scholars have long drawn a distinction between public and private spheres of activity. For instance, public and private spheres were a feature of Greek society by the time of Aristotle (384-322 BC) who acknowledged a boundary between affairs of the State *(polis)* and household affairs *(oikos)*. Recounting the work of Aristotle, Habermas noted:

> "In the fully developed Greek city-state the sphere of the *polis*, which was common to the free citizens, was strictly separated from the sphere of the *oikos;* in the sphere of the *oikos*, each individual is in his own realm."[70]

Habermas further contended that the private sphere is important because distance, space, and solitude from public life facilitates contemplative activity, which he regarded as necessary for human flourishing and self-fulfilment.

More recently, Rossler[71] has observed that there are two distinct semantic models underlying use of the terms 'private' and 'public.' One way in which the scope of

[70] Habermas, J. (1962) *The Structural Transformation of the Public Sphere,* (MIT Press, Cambridge), translated by Burger, T. & Lawrence, F. p. 10

privacy can be demarcated is to portray it using an 'Onion model'[72] which distinguishes between different layers or spheres of privacy. In this model, the centre layer of the onion is the realm of personal or bodily intimacy and privacy (including an individual's private diary) as opposed to which everything else is public. The second, middle layer is that concerning family/intimate relationships. Rossler opines that the second layer of privacy is regarded as the classic realm of privacy, since it juxtaposes the home which constitutes the private realm against the outside world of society and the state which constitute the public realm. The third, outer layer conceives of privacy as the realm of economic structures or public civil society, whilst intervention by the state forms the public realm.

Alternatively, Rossler asserts that the scope of privacy can be determined by reference to 'protected dimensions of actions.'[73] This model of privacy can only be described in terms of dimensions of action and responsibility, dimensions of interest and concern. Using this model, the term 'private' is predicated of actions or decisions that an individual may perform no matter where they happen to physically be located. For instance, when or where an individual goes to church would be regarded as a private matter. Similarly, comments made in public as a private person could be regarded as private comments, since it is the context of the comments, not the location which is relevant when considering whether or not they are private. Therefore, Rossler observes that the predicate 'private' is ascribed to actions, situations, states of mind, places and objects in both the onion model and the dimensions model. She further synthesises the two models in order to develop her own definition of privacy which focuses on control of unwanted access. Rossler contends that:

> "Something counts as private if one can oneself control the access to this 'something.' ...The term access can have both direct, concrete, physical meaning e.g. as when I demand to be able myself to control the access to my home, but it can also mean metaphorically. Metaphorical refers both to the control I have over who has what access to knowledge about me, for example, who knows which (relevant) data about me and the control I have over which people have access in the form of the ability to interfere or intervene when it comes to decisions that are relevant to me."[74]

She asserts that privacy can be classified into three basic dimensions, namely: decisional, information and local privacy as set out below:

---

[71] Rossler, B. (2005) *The Value of Privacy*, (Polity Press: Cambridge)
[72] *Ibid*, p. 5
[73] *Ibid*, p. 6
[74] Rossler, B. (2005) *The Value of Privacy*, (Polity Press: Cambridge), p.8

Firstly, *decisional* privacy: this dimension serves to secure the scope for an individual to make decisions and take action in all their social relations. It allows an individual to claim the right to protection form of unwanted access in the sense of unwanted interference in decisions and actions. For instance, whether an individual attends church, and if so, which one, is a private mode of action (in public) as is which school an individual's children attend, or the clothes an individual chooses to wear to a wedding.

Secondly, *Informational* privacy: this dimension serves to secure a horizon of expectations regarding what others know about him that is necessary for his autonomy. It allows an individual to claim the right to protection against unwanted access in the sense of interference in personal data about themselves i.e. access to information about individuals that they have no desire to see in the wrong hands.

Thirdly, *Local* privacy: serves to protect the possibilities for spatial withdrawal upon which a subject is dependent for the sake of their autonomy. It allows an individual to claim the right to privacy against the admission of people to spaces or areas. For instance, dwellings or rooms that are places of restricted access to the general public are considered private by Rossler. Having mapped out the broad dimensions of privacy, it is appropriate to examine how scholars have attempted to define privacy.

## *2.4 Definitions of privacy: Limited access v Control of personal information*

A number of different definitional approaches are evident in philosophical writings on privacy. For the purpose of this discussion they are divided into two groups, firstly definitions concerned with regulating the degree of (primarily physical) access third parties have to an *individual*, and secondly, definitions concerned with regulating access to an *individual's information*. Below, each definitional approach will be considered in turn, along with criticisms specific to each conception in the context of data protection laws.

### *2.4.1 Limited Access*

The first group, concerned with regulating the degree of access third parties have to individuals comprises three subgroups, namely: the right to be let alone, limited access

and intimacy. A common theme of the three sub-groups is that an individual has the right to determine the extent of their interaction with third parties on the basis that is it necessary to ensure their self-fulfillment as autonomous beings.

#### *2.4.1.1 The right to be let alone*

In their seminal article, 'The Right to Privacy,' Warren and Brandeis described the right to privacy as "the next step which must be taken for the protection of the person and for securing to the individual what Judge Cooley calls the '*right to be let alone*.'"[75] They argued that 19th century common law should expand its recognition of an individual's interest in his "inviolate personality" and so protect the privacy of an individual's 'thoughts, emotions, and sensations.' According to Warren & Brandeis, privacy rights are one aspect of a broader interest in being left alone, which in turn finds its justification in an individual's inviolate personality."[76] However, Allen, a legal scholar, observes:

> "If privacy simply meant "being let alone", any form of offensive or harmful conduct directed toward another person could be characterized as a violation of personal privacy. A punch in the nose would be a privacy invasion."[77]

Similarly, Sparkes, a philosopher, observes that:

> "If I hide in a bush in order to watch what goes on in your bedroom, I am acting with intent to invade your privacy, but also with intent to let you alone (if I don't let you alone, my snooping project will be frustrated)."[78]

Also, another legal scholar, Gavison recognises that whilst Warren & Brandeis were claiming a right to non-interference by the State, this conception is inadequate as the right to be left alone could be considered a negative civil liberty:

> "the typical privacy claim is not a claim for non-interference by the state at all. It is a claim for state interference in the form of legal protection against other individuals."[79]

At this juncture it is appropriate to examine the relationship between this definition of privacy and data protection laws. Arguably this definition is inadequate since data

---

[75] Justice Cooley had used this phrase in relation to attempted physical touching as a tort injury in his treatise on torts: Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4, p. 1100

[76] Warren, S. & Brandeis, L. (1890) "The Right to Privacy" *Harvard Law Review*, Vol. 4, p. 207

[77] Allen, A. (1988) *Uneasy Access: Privacy for Women in a Free Society, (*Rowman & Littlefield, Totowa, NJ), p.7

[78] Sparkes, A. W. (1981) "The right to be let alone: a violation of privacy," *Bulletin of the Australian Society of Legal Philosophy*, p. 254

[79] Gavison, R. (1980) "Privacy and the Limits of the Law," *Yale Law Journal*. Vol. 89, p. 438

protection laws seek to balance privacy against other interests, yet this definition fails to provide guidance about how privacy should be valued vis-a-vis other interests, such as free speech or law enforcement. As the Younger Committee noted, the right to be let alone would result in a very strange law:

> "the formula that privacy is the 'right to be let alone' turns out on closer examination to go so far beyond any right which the individual living in an organised society could reasonably claim, that it would be useless as a basis for the granting of legal protection. Any law which proclaimed this as a general right would have to qualify the right in so many ways that the generality of the concept would be destroyed." [80]

Thus, in concurrence with Solove[81] it is suggested that whilst a definition of privacy based on a 'right to be let alone,' has an intuitive appeal in that it relates to the local privacy dimension outlined by Rossler, it is incomplete as it is too broad and vague to form the basis of data protection laws.[82] From a data protection perspective it is particularly weak, since it is primarily concerned with controlling physical access by third parties to an individual, rather than with personal information protection.

#### *2.4.1.2 Limited access to the self*

Another definition of privacy which to some extent overlaps with the right 'to be let alone' definition offered by Warren & Brandeis is that of 'limited access to the self.' Allen, a legal scholar, opines that privacy covers not just restricted physical access, but also mental and informational access, describing privacy as:

> "a condition of inaccessibility of the person, his or her mental states, or information about the person to the senses and surveillance devices of others"[83]

that facilitates being apart from others and the ability to control the disclosure of information about oneself. Likewise, Bok, a philosopher, defines privacy as the "condition of being protected from unwarranted access by others – either physical access, personal information or attention."[84] This definitional approach satisfies all three dimensions of privacy outlined by Rossler, namely: decisional, informational and local privacy. However, scholars differ on the degree of control and choice that an individual

[80] Younger, K.. (1972) Report of the Committee on Privacy, Cmnd 5012, p. 10
[81] Solove, D. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, p. 1087
[82] Arguably, this conception of privacy underpins the Protection from Harassment Act 1997. However, this legislation is not the focus of this thesis, and so it will not be explored further.
[83] Allen, A. (1988) *Uneasy Access: Privacy for Women in a Free Society* (Random and Littlefield, Totowa), p.15
[84] Bok, S. (1983) *Secrets: On the Ethics of Concealment and Revelation"* (Pantheon Books: New York), p.10

should have when limiting access by third parties. Some scholars such as Gavison claim that: "an individual enjoys perfect privacy when he is completely inaccessible to others,"[85] whereas, other scholars such as Rossler contend that the inaccessibility definition is incomplete unless it is joined with the factor of control. She provides the example:

> "a crevasse into which I have fallen is clearly not private even though it does not comply with the condition of inaccessibility, if the state of isolation, seclusion, or secrecy is enforced and not freely chosen, i.e. if the individual in question has no control over it, then one would not describe it as private."[86]

Therefore, this conception is too weak since not all situations of limited access are private because they have not been chosen e.g. an individual on a deserted island is completely isolated, but arguably is not private since privacy denotes control over access to others, so if there is no-one else around, then an individual is not in a position to seek or control privacy.

At this juncture it is appropriate to examine the relationship between this definition of privacy and data protection laws. Bygrave[87] asserts that concerns about non-interference inform data protection measures, which include provisions restricting the amount of personal information that can be collected, the secondary uses to which the information can be put, and to whom the information can be disclosed. According to Bygrave:

> "Implementation of such provisions lessens the risk of a decision being made about a person on the basis of inaccurate or irrelevant information. This, in turn, lessens the risk of the decision maker then taking, say, unwarranted investigative action which interferes with or disturbs that person."[88]

Nevertheless, this definitional approach is incomplete since it fails to specify the degree of access necessary to an individual to constitute a privacy invasion, nor does it indicate what types of information are private. Accordingly, it is too weak to form the primary basis of data protection laws.

#### *2.4.1.3 Intimacy*

According to this conception of privacy, private is defined as information that falls within the sphere of the household, of reproduction, of biological necessities and

[85] Gavison, *R.* (1980) *"*Privacy and the Limits of the Law*," Yale Law Review*, Vol. 89, pp. 421-471 p.428
[86] Rossler, B. (2005) *The Value of Privacy*, (Polity Press: Cambridge), p. 7
[87] Bygrave, L. (2001) "The Place of Privacy in Data Protection Law," *University of New South Wales Law Journal*, Vol. 6 <http://www.austlii.edu.au/au/journals/UNSWLJ/2001/6.html>
[88] *Ibid*, para 10

intimate relationships. Further, this theory recognises that individuals form relationships with differing degrees of intimacy and self revelation, and require privacy so that the varying levels of intimacy for different relationships can be maintained. Indeed, Rosen, a legal scholar, states, that:

> "In order to flourish, the intimate relationships on which true knowledge of another person depends need space as well as time: sanctuaries from the gaze of the crowd in which slow mutual self-disclosure is possible."[89]

Accordingly, Rachels, a philosopher, conceives of privacy as being:

> "based on the idea that there is a close connection between our ability to control who has access to us and to information about us, and our ability to create and maintain different sorts of social relationships with different people."[90]

By emphasising the value of relationship-orientated privacy this conception of privacy tries to define what aspects of life an individual should be able to control, keep secret or restrict access to. Thus, this definitional approach has an intuitive appeal in that it satisfies the decisional, informational and local dimensions of privacy outlined by Rossler. Innes claims that:

> "Privacy is the state of the agent having control over a realm of intimacy, which contains her decisions about intimate access to herself (including intimate informational access) about her decisions about her own intimate actions"[91]

However, Solove contends that this conception is problematic as although relationships based upon trust, love and intimacy are facilitated by privacy they are not the sole ends of privacy.[92] Indeed, it is possible to have private relationships without intimacy and to perform private acts that are not intimate. For instance, DeCew,[93] a philosopher, observes that a discussion between an individual and their bank manager about an individual's financial information may not be considered intimate but it would be considered to be private. Equally, the information shared between a psychotherapist and a patient would be considered private, but not intimate. Consequently, Weinstein, a legal scholar, observes that:

> "There is a wide range of instances where to speak of something as private is not to imply intimacy. Individuals not intimately related may nevertheless assert that

[89] Rosen, J. (2000), *The Unwanted Gaze: The destruction of Privacy in America*, (Random House), p. 8
[90] Rachels, J. (1975) "Why Privacy is Important?" *Philosophy and Public Affairs*, Vol. 4, p.292
[91] Inness, J. C. (1992) *Privacy, Intimacy, and Isolation*, (Oxford University Press, New York), p.56
[92] Weinstein, W. L. (1971) "The Private and the Free: A Conceptual Inquiry," in Pennock J. R. & Chapman, J. W. (eds), Nomos XIII *Privacy* (New York), p.33
[93] DeCew Wagner, J. (1997) *In Pursuit of Privacy: Law, Ethics and the Rise of Technology*, (Cornell University Press, Ithaca & London), p.56

> their relation or activity is a private one in the sense that it is not the proper concern of the community or some institution, such as the state, a church, or a business firm."[94]

Likewise, Regan, a political scientist, observes, computer databases pose a significant threat to privacy but:

> "do not primarily affect . . . relationships of friendship, love, and trust. Instead, these threats come from private and governmental organizations-the police, welfare agencies, credit agencies, banks, and employer."[95]

Examining the relationship between this definition of privacy and data protection laws, Bygrave claims that this conception of privacy is relatively unpopular because:

> "intimacy-oriented definitions of privacy are unable to anticipate and capture the process by which detailed personal profiles of individuals are created through combining disparate pieces of ostensibly innocuous information."[96]

He cites the Data Protection Directive 95/46/EC as an example of a legislative measure containing provisions specifically incorporating an intimacy-oriented conception of privacy:

> "in the provisions that place extra restrictions on the processing of certain categories of especially sensitive, personal data."[97]

However, data protection laws cannot be based solely on an intimacy orientated notion of privacy since individuals routinely share information and seek privacy protection in non-intimate relationships and have an expectation of privacy when doing so e.g. when they talk to their bank manager about their financial affairs. Indeed, a particular problem with this conception of privacy is that it is based on a private realm, yet data protection laws generally not concerned with regulating intimate or domestic sphere relationships, as these are generally considered beyond the scope of such laws.

### *2.4.2 Privacy theories concerned with protection of individuals' information*

The second group, concerned with regulating access to an individual's information comprises two subgroups: firstly, those based on secrecy, and, secondly, those based on

---

[94] Weinstein, W.L. (1971) "The Private and the Free: A Conceptual Inquir*y,"* in Pennock, J. R & Chapman, J. W. (eds.), Nomos XIIL *Privacy* (New York), p. 27

[95] Regan, P. (1995) *Legislating Privacy: Technology, Social values and Public Policy*, (Chapel Hill: University of North Carolina Press), p. 213

[96] Bygrave, L. (2001) "The Place of Privacy in Data Protection Law," *University of New South Wales Law Journal*, Vol. 6 <http://www.austlii.edu.au/au/journals/UNSWLJ/2001/6.html> para 11

[97] *Ibid*, para 12

information control, though it is a truism to state that the sub-groups are inter-related and overlap, since a common theme of the two sub-groups is that an individual has the right to determine the nature and extent of the information they disclose to third parties.

#### *2.4.2.1 Secrecy/concealment of discreditable information*

This conception views privacy as the right to conceal information. Judge Posner, an eminent legal scholar, regards privacy as a form of self-interested economic behaviour, which allows individuals to conceal true but harmful facts about themselves for financial gain. Thus, he claims that:

> "[W]hen people today decry lack of privacy," ... "what they want, I think, is mainly something quite different from seclusion; they want more power to conceal information about themselves that others might use to their disadvantage."[98]

Posner asserts that individuals use privacy selectively to mislead. He claims that they do this in a number of ways: firstly, true facts are selectively revealed in order to create different perceptions that people have of them. For instance, an individual reveals different information to friends than they reveal to their employer. Secondly, individuals are only reticent about disclosing negative information about themselves, such as an employee conceals a serious health problem from his employer or a prospective husband conceals his sterility from his fiancée.[99] By ascribing an economic value to privacy, Posner contends that it would be financially inefficient for privacy laws to permit anything other than full disclosure of information. Solove,[100] contends that this definition is too restrictive and narrow as it focuses only on discreditable facts, which is erroneous, since not all information that an individual might want to keep secret is necessarily discreditable or misleading. For instance, prospective parents may want to keep secret or control disclosure of a pregnancy until after the first semester has passed. Also, Posner's conception fails to consider the importance that an individual's control over information plays in privacy. Whilst individuals are willing to disclose some information, they often expect to maintain some control over the information. For instance, whilst individuals are likely to fully disclose details of medical ailments to their doctor in order to receive medical treatment, they would expect that the information would not be shared with others e.g. their local baker.

---

[98] Posner, R. (1983) *The Economics of Justice*, (Harvard University Press) p. 271
[99] Posner, R. (1978) "John A. Sibley Lecture: The Right to Privacy," *Georgia Law Review*, Vol. 12, p. 399
[100] Solove, D. J. (2007) "I've Got Nothing to Hide and Other Misunderstandings of Privacy," *San Diego Law Review*, Vol. 44, p. 562

In a less normatively charged manner, Etzioni, a sociologist, defines privacy as: "the realm in which an actor...can legitimately act without disclosure and accountability to others."[101] Etzioni's conception can be understood as a subset of the limited access conception since secrecy of personal information is a way to limit access to the self. However, this conception is narrower than the limited access conception, as secrecy only concerns the concealment of personal facts. Etzioni's conception is problematic because it doesn't cater for the fact that an individual may want to keep certain information from some people but not from others. For instance, Solove claims that:

> "Criticizing a boss to a coworker does not mean that the employee desires that her boss know her comments."[102]

Accordingly, some theorists claim that this approach is incomplete unless it is coupled with the concept of control. For instance, Inness, a philosopher, observes:

> "Privacy might not necessarily be opposed to publicity; its function might be to provide the individual with control over certain aspects of her life."[103]

This aspect was also recognized by another philosopher, Benn, who observed that privacy is not that one's private affairs:

> "are kept out of sight or from the knowledge of others that makes them private. Rather, [one's private affairs] are matters that it would be inappropriate for others to try to find out about, much less report on, without one's consent."[104]

So, Etzioni's[105] definition of privacy as 'selective' secrecy is problematic as privacy concerns an individual's ability to control information usage as well as its disclosure, and, this theory does not explain what matters are private or what degree of access would constitute a privacy violation. Also, this conception is inadequate as secret information is often not private e.g. military plans, whilst private information is often not secret e.g. an individual's bankruptcy declaration.[106] Moreover, the secrecy conception cannot accommodate "decisional privacy," for instance the decision by consenting adults whether or not to use contraceptive devices when engaging in sexual intercourse is an example of decisional privacy. Therefore, while most theorists would recognize the disclosure of certain secrets to be a violation of privacy, many commonly

---

[101] Etzioni, A. (1999) *The Limits of Privacy*, (Basic Books, NY) p. 196
[102] Solove, D. (2002) "Conceptualizing Privacy," *California Law Review*, Vol. 90, p.1108
[103] Innes, J. (1992) *Privacy, Intimacy and Isolation,* (OUP) p. 6
[104] Benn, S. (1971) "*Privacy, Freedom, and Respect for Persons,"* in Pennock, J. & Chapman, R. (Eds.), Nomos XIII: *Privacy*, (New York: Atherton Press), p. 2
[105] Etzioni, A. (1999) *The Limits of Privacy* (Basic Books, New York) p.196
[106] DeCew, J.W. (1997) *In Pursuit of Privacy: Law, Ethics and the Rise of Technology* (Cornell University Press, Ithaca) 81-94, p. 48

recognized privacy invasions do not involve the loss of secrecy. Thus, secrecy as the common denominator of privacy makes the conception of privacy too narrow to form the basis of data protection laws.

#### *2.4.2.2 Control over personal information (Informational self-determination)*

Bygrave observes that definitions of privacy framed in terms of 'information control' are the most popular in data protection discourse[107] and that the most influential of such definitions was offered by Westin, a legal scholar who 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...[It is] the desire of people to choose freely under what circumstances and to what extent they will expose themselves, their attitude and their behaviour to others."[108]

That is, individuals must be allowed to choose what information is made available about them, and under which circumstances. Similarly, another legal scholar, Miller opines that "the basic attribute of an effective right of privacy is the individual's ability to control the circulation of information relating to him."[109] Likewise, legal scholar, Fried asserts that:

> "Privacy is not simply an absence of information about us in the minds of others; rather it is the *control* we have over information about ourselves."[110]

According to Bygrave, informational self-determination definitions of privacy are popular in data protection discourse because:

> "they appear directly applicable to the issues raised by the data-processing practices of organisations. They also harmonise fairly well with, and build upon, many of the basic rules of data protection law, particularly those rules that enable persons to participate in, and influence, the processing of information about them."[111]

However, whilst this conception of privacy has an intuitive appeal, in the sense that it confers on individuals informational self-determination rights, which accords with the philosophical concepts of autonomy and personhood, it is not without problems.

---

[107] See, United Kingdom, Committee on Data Processing, (1978) *Report of the Committee on Data Protection* (Cmnd 7341) 10, [2.04]; Blekeli, R.D. (1980) "Framework for the Analysis of Privacy and Information Systems" in Bing, J. & Selmer, K. S. (eds), *A Decade of Computers and Law* (Universitetsforlaget: Oslo) p. 24

[108] Westin, A. (1967) *Privacy and Freedom*, (New York: Atheneum) p.7

[109] Miller, A. R. (1971) *The Assault on Privacy: Computers, Data banks and Dossiers*, (Ann Arbor: University of Michigan Press) p. 25

[110] Fried, C. (1968) "Privacy," *Yale Law Journal* Vol. *77*, pp. 482-483

[111] Bygrave, L. (2001) "The Place of Privacy in Data Protection Law," *University of New South Wales Law Journal*, Vol. 6 <http://www.austlii.edu.au/au/journals/UNSWLJ/2001/6.html> para 7

One problem with this approach is the difficulty of defining its terms. For instance, what type of information is covered? Is it all personal information, or only personal information that is privacy sensitive, on the basis that not all personal data is equally sensitive from the point of view of their dissemination causing harm or embarrassment to an individual.[112]

Some legal theorists define the type of information over which individuals should exercise control extremely broadly. For instance, Fried defines privacy as "control over knowledge about oneself"[113] that is necessary to protect "fundamental relations" of "respect, love, friendship and trust."[114] This definition is too narrow because it excludes important information such as medical or financial records. Another problem with this privacy conception is the difficulty of defining what is meant by control. Frequently, control is understood as a form of ownership. For example, Westin concludes that "personal information thought of as the right of decision over one's private personality, should be defined as a property right."[115] By this, he means that people should own information about themselves, and, as owners of property, be entitled to control what is done with it.

However, this approach is often inadequate because information can be possessed by several individuals simultaneously, which leads to problems regarding the commodification of information, as noted by Solove, who states that:

> "[T]here are problems with viewing personal information as equivalent to any other commodity. Personal information is often formed in relationships with others, with all parties to that relationship having some claim to that information. For example, individuals are not the lone creators of their web browsing information, for most of that information is created from the interaction between the user and websites."[116]

Similarly, it is unclear who owns an individual's medical records? Is it the individual patient or the doctor? Thus, in concurrence with Bygrave, it is suggested that individual control definitions of privacy does not adequately address ownership issues, since they fail to recognise that:

---

[112] Turn, R. (1976) "Classification of personal information for privacy protection purposes," AFIPS National Computer Conference, pp. 301-307, at p. 301
[113] Fried, C. (1968*) "*Privacy," *Yale Law Journal*, Vol. 77, p.483
[114] *Ibid* p. 477
[115] Westin, A. (1967) *Privacy and Freedom*, (New York: Atheneum), p. 7
[116] Solove, D. J (2002) "Conceptualizing Privacy," *California Law Review*, Vol. 90, p.1113 (footnotes omitted)

> "data protection laws rarely give persons an absolute right to dispense with data about themselves as they see fit. Thus, the laws are better viewed as manifestations of an interest in informational *co*-determination as opposed to self-determination."[117]

Other theorists critique the control over personal information conception as being too narrow because it focuses too heavily on individual choice. For instance, Schwartz, a legal scholar, argues that this conception wrongly assumes that individuals have the autonomy to exercise control over their personal data in all situations; an assumption that fails to recognize "that individual self-determination is itself shaped by the processing of personal data."[118] Schwartz also questions the assumption that individuals are able to exercise meaningful choices with regard to their information, given disparities in knowledge and power when bargaining over the transfer of their information. For instance, if an individual wants to purchase an item from an online shopping website they will often have to consent to the collection and processing of their personal details in order to complete the transaction. Indeed, an individual wishing to buy a book from Amazon.com may be required to supply details of their gender, age etc. even though such extraneous details are not terms of the contract. Rather, the website will gather such information in order to profile their customers, and potentially sell such profile data to third party organisations.

## *2.5 An alternative, contextual, harm-based approach*

As indicated above, Solove and Nissenbaum claim that orthodox conceptions of privacy are inadequate. Nissenbaum[119] claims that three features of modern technology threaten the traditional conceptions of privacy. Firstly, there are virtually no limits on the amount of information that can be collected; secondly, there are no limits on the level of data analysis that can be conducted, and thirdly, there are virtually no limits on the indefinite storage of information, due to ever increasing capacity and decreasing storage costs.[120] Also, she asserts that the 'permanence, malleability and transportability' of information raises questions regarding the collection and uses of information that is not generally understood to be intimate or sensitive, and which individuals were not concerned about when, prior to the advent of information technology, they were assured a kind of

[117] Bygrave, L. (2001) "The Place of Privacy in Data Protection Law," *University of New South Wales Law Journal, Vol.* 6 <http://www.austlii.edu.au/au/journals/UNSWLJ/2001/6.html>, para 8
[118] Schwartz, P. (1999) "Privacy and Democracy in Cyberspace," *Vanderbilt Law Review*, Vol. 50, p.1661
[119] Nissenbaum, H. (1998) "Protecting Privacy in an Information Age: The Problem of Privacy in Public," *Law & Phil.* Vol. 17, p.576
[120] A study, published in *Science*, calculated the amount of data stored in the world by 2007 as 295 exabytes, i.e. the equivalent of 1.2 billion average hard drives. Stewart, J. (2011) 'Global data storage calculated at 295 exabytes,' <http://www.bbc.co.uk/news/technology-12419672>

practical obscurity.[121] Similarly, Solove suggests that as a result of technological developments the landscape of privacy is constantly changing and that accordingly privacy scholars, legislators and judges may be misled by trying to fit new problems into old conceptions.[122] Thus, he asserts that:

> "the [traditional] top down approach [to conceptualising privacy] of beginning with an overarching conception of privacy designed to apply in all contexts often results in a conception that does not fit well when applied to the multitude of situations and problems involving privacy."[123]

He advocates a pragmatic approach to privacy, advising that two questions should be asked in each particular context: Firstly, is there a privacy interest at stake? and, secondly, is there a fundamentally important competing right that justifies overriding the privacy interest in the case? He claims that this approach:

> "turns away from universals and focuses on specific situations…by examining specific problematic situations rather than trying to fit each situation into a rigid predefined category."[124]

Thus, Solove's approach foregoes a search for a universally agreed concept of privacy in favour of a pragmatic approach that focuses specifically on privacy problems and their resulting harms to individuals and society. He summarizes his position as follows:

> "my approach is from the bottom up rather than the top down because it conceptualizes privacy within particular contexts rather than in the abstract."[125]

Building on the work of Post, who claims that privacy:

> "cannot be reduced to objective facts like spatial distance or information or observability; it can only be understood by reference to norms of behaviour"[126]

Solove contends that privacy problems involve disruptions to 'practices,' that is, activities, customs, norms, and traditions, such as writing letters, talking to one's psychotherapist, engaging in sexual intercourse, and making decisions regarding medical treatment, over which individuals traditionally assert privacy claims. Adopting a similar approach, Nissenbaum argues that there are norms specific to particular

---

[121] Nissenbaum (1998) p. 562
[122] Solove, D. J (2002) "Conceptualizing Privacy," *California Law Review*, Vol. 90, p.1146.
[123] *Ibid*, p. 1099
[124] Solove, D. J. (2002) "Conceptualizing Privacy," *California Law Review*, Vol. 90, p. 1126
[125] Solove, D. J. (2002) "Conceptualizing Privacy," *California Law Review*, Vol. 90, p. 1125
[126] Post, R. C. (1989) "The Social Foundations of Privacy: Community and Self in the Common Law Tort," *California Law Review* Vol. 77, p. 969

relationships and situations that govern what kind of information, and how much information it is appropriate and relevant to share.[127] She further asserts that:

> "These contextual norms explain the boundaries of our underlying entitlements regarding personal information; our privacy is invaded when these contextual boundaries are violated."[128]

This 'contextual integrity' approach involves respecting the norms relating to personal information disclosure applicable to particular contexts, on the basis that privacy invasions occur when these boundary norms are violated. Thus, as stated earlier, Nissenbaum asserts that traditional theories of privacy which focus on the protection of the private sphere of individuals' thoughts and emotions by protecting sensitive information about individuals are too narrow. Accordingly, instead of focusing on the classification of data as privacy sensitive or non-sensitive, her approach argues that it is particular contexts that make information privacy sensitive, and thus explains why individuals may seek to claim privacy entitlements over non-sensitive information. Similarly, Solove contends that:

> "We should conceptualize privacy by focusing on the specific types of disruption and the specific practices disrupted rather than looking for the common denominator that links all of them."[129]

Calo[130] contends that whilst Solove's focus on harms in the form of disruption of specific practices lends itself well to a legal and policy centered discussion focused on the prevention or remedying of harms, the harm element is under-theorized. Calo contends that privacy harms fall into two categories: *subjective* and *objective*. He claims that the two categories of privacy harm are distinct but not entirely separate, since they are both concerned with the loss of control over personal information. He defines the subjective category of privacy harm as the perception of unwanted observation, and asserts that it is 'subjective' in the sense of being internal to the individual harmed. He further claims that periodic absence from the perception of observation is a necessary element of the human condition, by drawing upon the work of Solove, who claims that:

"People need solitude for comfort, curiosity, self-development, even mental health."[131]

---

[127] Nissenbaum, H. (1998) "Protecting Privacy in an Information Age: The Problem of Privacy in Public," *Law & Phil*. Vol. 17, p. 582
[128] *Ibid*, p.586
[129] *Ibid*, 1130
[130] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp.1132-1161
[131] Solove, D. J. (2008) *Understanding Privacy* (Harvard University Press: Boston), pp. 163-163

Importantly, he states that actual observation need not occur to cause harm; perception of observation can suffice, as it can result in fear or discomfort regarding the loss of control over information. He also defines term observation broadly, stating that it includes:

> "watching a person directly—their body, brain waves, or behaviour—is observation. So, too, is reading a report of their preferences, associations, and whereabouts. Observation can also include inference, as when we make "an observation" about someone on the basis of what we know about them."[132]

He further stipulates that the observation at issue must be "unwanted" to constitute a harm; otherwise almost any interaction could be considered a privacy problem. Also, he claims that the underlying cause of subjective privacy harm can be acute or ongoing. For instance, a person may feel embarrassed by a single act of observation, such as when they walk through a back-scatter device in airport security that creates a picture of their naked body.[133] Or, an individual may feel an ongoing sense of regret about an embarrassing revelation they have posted in an online blog.[134]

> "This category describes unwelcome mental states—anxiety, embarrassment, fear—that stem from the belief that one is being watched or monitored." [135]

Moreover, he states that aversion to observation accommodates degrees of harm, as subjective privacy harms can range in severity from mild discomfort at the presence of a security camera to "mental pain and distress far greater than could be inflicted by mere bodily injury."[136] The second category of harm identified by Calo is "objective" in the sense of being external to the individual harmed. He defines this category of privacy harm as:

> "the unanticipated or coerced use of personal information concerning a person against that person."[137]

Accordingly, it is generally not a privacy harm to use an individual's information if they publicized it or where they understood and agreed to the use.[138] Thus, it is not

---

[132] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal,* Vol. 86, pp.1132-1161 at p. 1144

[133] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal,* Vol. 86, p. 1144. For a discussion of backscatter devices, see Rosen, J. (2010) "Nude Awakening," *The New Republic*, Feb. 10, p. 8

[134] *Ibid*, p.1144

[135] *Ibid*, p. 1133

[136] *Ibid*, p. 1142

[137] *Ibid*, p. 1143

[138] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal,* Vol. 86, No. 3, pp.1132-1161 at p. 1148

necessarily a privacy harm to trade an email address for a chance to win a sweepstakes where both parties understand that the email will be used for marketing purposes.[139] Rather, the problem arises where an individual has no idea that the information was even collected or, if she does, how it will be used. This fundamental tension is evident in the context of online privacy, where bloggers may not realise that the content of their blog posts is collected and mined by third party companies for the purposes of customer profiling and creation of personalised targeted adverts. Indeed, Wortham reports that:

> "a person may share information on a social networking website and not realize that it could be used to deny her a job or admission to college."[140]

Calo also defines objective privacy harm as occurring where there is a forced or coerced use of personal information against an individual's best interests. It is important to note that the coerced or forced action justified by reference to personal information must be adverse; otherwise, it is likely not a "harm" in the common understanding of word.[141] For instance, medical care is premised upon giving up information or revealing one's body in potentially embarrassing and uncomfortable ways, yet there may be little alternative to such surveillance in daily life, since "Doctors look at our bodies not to harm us but to protect our health."[142]

Finally, Calo asserts that no human being actually needs to see the personal information itself for it to be misused against an individual. Machines can analyse personal information and use it to make automatic decisions that affect individuals in tangible and negative ways. As Citron observes:

> "In the past, computer systems helped humans apply rules to individual cases. Now, automated systems have become the primary decision makers. These systems often take human decision making out of the process of terminating individuals' Medicaid, food stamp, and other welfare benefits…Computer programs identify parents believed to owe child support and instruct state agencies to file collection proceedings against those individuals. Voters are purged from the rolls without notice, and small businesses are deemed ineligible for federal contracts."[143]

---

[139] *Ibid*, p. 1149

[140] Wortham, J (2009). *More Employers Use Social Networks to Check Out Applicants*, N.Y. Times (Aug. 20, 2009). A 2009 study by Harris Interactive showed that forty-five percent of employers surveyed used social networks to vet potential hires., <http://bits.blogs.nytimes.com/2009/08/20/more-employers-use-social-networks-to-check-out-applicants/>

[141] *Ibid*, p. 1150

[142] *Ibid*, p.1150

[143] Citron, D. K.. (2008) "Technological Due Process," *Washington University Law Review*, Vol. 85, p.1252

Calo acknowledges that the exact source of the information these systems rely upon is not always known, but claims that there is every indication that it includes personal information not supplied by citizens for this purpose.

In concurrence with Calo, it is contended that delineating the specific boundaries of privacy harm offers a number of advantages. Firstly, this combined approach captures the full range of harms from observation. For instance, this approach acknowledges that the perception of observation can still be harmful even if no human being ever sees the information.[144] Secondly, this approach acknowledges that machines are clearly capable of collecting, processing, and acting upon personal information in harmful ways without any human being ever seeing it.[145] Another advantage of this approach is that the two components of privacy harm are testable:

> "Courts and regulators are capable of investigating—particularly with the help of experts—whether a person felt observed, whether she consented to observation or collection, and whether she anticipated a given use of her information."[146]

Also, the approach also provides criteria for "sizing" privacy harms and ranking their relative severity. In the case of subjective privacy harms, legislators could assess subjective privacy harms by reference to the degree of aversion to any observation or by reference to the amount of observation experienced, since Calo's approach indicates that high degrees of both translate into the greatest harm, but harm is also possible if either is very high.[147]

Similarly, legislators and decision makers could assess objective privacy harms by reference to the degree of knowledge or consent, as distinct from the severity of the information use. Finally, the categories are also capable of flexible interpretation so they could be applied to novel technological developments and situations in the future.[148]

---

[144] *Ibid*, p. 1154
[145] Calo, R. M. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp.1132-1161, at p. 1154
[146] *Ibid*, p. 1154
[147] *Ibid*, p. 1154. Calo notes that the two are, obviously, related. Extensive surveillance can breed greater aversion. The idea here is that each context, state, or activity may be attended by a specific level of aversion to observation that can in turn be invaded to a lesser or greater degree.
[148] *Ibid*, p. 1154

### *2.5.1 A proposed definition of data privacy & private data*

The foregoing analysis leads me to claim that the most conceptually coherent approach to privacy protection emerges when Nissembaum and Solove's contextual approaches to privacy protection are combined with Calo's categories of subjective and objective harms, and accordingly, this combined approach should underpin privacy and data protection laws. From this analysis, I have developed a proposed definition of data privacy which will be tested on privacy and data protection experts, in order to critique existing legislative measures and assess the adequacy of this definition as a reform measure. Below is the proposed definition:

*Data privacy concerns the legal regulation of the boundary between personal and private data. Private data is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual.*
*It is the legal right of an individual to withhold consent to the collection, processing, communication or usage of personal data, the disclosure of which could cause unreasonable harm to that individual. Disclosure of private data should not be compulsory except where it is in the public interest, for instance if disclosure is in the interests of national security, public safety, the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health, or for the protection of the rights and freedoms of others or society as a whole.*

The proposed definition begins by defining the term 'data privacy.' Many scholars, e.g. Rossler use the terms data and information privacy interchangeably. However, in light of the fact that the nomenclature of such laws is 'data protection' in Europe, I elected to utilise the term 'data privacy,' to ensure conceptual consistency.

A second element of the definition is that it draws a distinction between personal and private data. The aims of this distinction are twofold: firstly, to recognise that not all personal data warrants privacy protection. The next chapter will reveal that contextual factors influence privacy perceptions, e.g. an individual seeking to avoid an abusive ex-partner may consider their name and address private, and restrict their listing in a public telephone directory. Secondly, the term private is used to denote that this concept is distinct from the term 'sensitive,' as defined in Arts 8(1) and 8 (5) of Directive 95/46/EC, since the discussion in chapter 4 will demonstrate that *a priori* classification of data as privacy sensitive is fallacious.

A third element of the definition is 'unreasonable harm.' The phrase 'unreasonable' is used, as opposed to the phrase 'any harm' to indicate that not all privacy invasions will be considered adverse. Thus, the definition is not drafted as broadly as the 'right to be left alone' definition offered by Warren & Brandeis. Rather, it focuses on 'unreasonable harm,' which acknowledges that not all uses of personal data could be considered unreasonable e.g. sharing of medical data between doctors would normally be considered a 'reasonable' activity, whereas the disclosure of medical information to a local baker would normally be considered 'unreasonable', irrespective of the consequences of either.

A fourth element of the definition is that disclosure should not be 'compulsory except.' The phrase 'compulsory' indicates a degree of choice and control on the part of the individual data subject. However, the phrase 'except' indicates that wider societal interests must be considered, and where appropriate, given priority over an individual's privacy claim. For instance, it might be appropriate for security officers to invade an individual's privacy by monitoring their emails, if such activity would disrupt the activities of a terror cell, and protect national security. Thus, this definition recognises that individuals do not have an *absolute* right of control over their information.

## *2.6 Summary*

In conclusion, analysis of the various conceptions of privacy revealed that whilst privacy is difficult to define it has, and continues to play, an important role in the formation of Western liberal democracies. A review of the privacy theories revealed that they are classified in two main ways: those which are access based, and those which are information control based. Also, the analysis indicated that current data protection laws are heavily influenced by the informational control conception of privacy espoused by Westin. However, the discussion above revealed that Westin's theory is incomplete since it leads to difficulties classifying the types of data that are to be protected and over-emphasises the roles of control and consent in data protection. Thereafter, the analysis revealed that it would be appropriate to develop the pragmatic, contextual, harm-based approaches to privacy espoused by Nissenbaum and Solove by recognising subjective and privacy harms identified by Calo, since this approach would provide a flexible and responsive framework for legislators and decision makers to determine what amounts to a privacy harm.

Accordingly, in the next chapter I will review the conceptual adequacy of the terms ‘privacy’ and ‘private’ in current data protection legislation before testing the harm-based definition of data privacy and private data in subsequent chapters in order to generate legislative reform proposals.

# Chapter 3

## Critique of Privacy & Data Protection Laws

### *3.1 Introduction*

A review of the literature in chapter two indicated that although privacy is a protean concept with contested meanings, it is worthy of protection. Accordingly, this chapter will consider its recognition as a legal right, and explore the adequacy of existing legislative measures. The chapter begins by exploring the rationale for the introduction of privacy and data protection laws; in particular, it discusses the historical catalysts for the enactment of such laws. Thereafter, it considers the relationship between privacy protection and data protection.

This chapter will demonstrate that although one of the key objectives of current privacy and data protection laws is the protection of privacy, the term privacy is not defined within such laws. It will draw upon the literature in chapter two to illustrate that most human rights based privacy laws seek to protect a broad concept of privacy, whereas data protection laws seek to protect a narrower element of the concept, namely informational or data privacy.

The chapter will conclude that the failure to define privacy in existing data protection laws is a major weakness, and draw upon the proposed definition of 'data privacy' and 'private data' offered in chapter two, in order to generate reform proposals.

The importance of this underlined by the recent upsurge in calls for a global legal framework for privacy and data protection For instance, in 2005, a call for global harmonization of the protection of information privacy was launched by the world's privacy and data protection commissioners at their annual international conference. They adopted the Montreux Declaration entitled 'The protection of personal data and privacy in a globalized world: a universal right respecting diversities.' In this text, the privacy commissioners stated that:

> "It is necessary to strengthen the universal character of this right in order to obtain a universal recognition of the principles governing the processing of personal data whilst respecting legal, political, economical and cultural diversities."[149]

In addition, the Commissioners committed themselves to working with governments as well as international and supranational organisations with a view to drafting a global legal instrument on data protection to be submitted to the United Nations to:

> "prepare a binding legal instrument which clearly sets out in detail the rights to data protection as privacy as enforceable human rights."[150]

Meanwhile, private sector organisations also called for harmonised global privacy standards. For instance, on 14th September 2007 Peter Fleischer, Google's global privacy counsel, pleaded at a UNESCO conference for the creation of global international privacy standards. He also posted the following text on Google Public Policy Blog:

> "Google is calling for a discussion about international privacy standards which work to protect everyone's privacy on the Internet. These standards must be clear and strong, mindful of commercial realities, and in line with oftentimes divergent political needs. Moreover, global privacy standards need to reflect technological realities, taking into account how quickly these realities can change."[151]

However, Fleischer recommends that a global data protection measure be based, not on Directive 95/46/EC, but instead on the APEC Privacy Framework as:

> "The APEC framework already carefully balances information privacy with business needs and commercial interests. And unlike the OECD Guidelines and the European Directive, it was developed in the Internet age."[152]

These calls for global standardisation in data protection law, make it ever more important that: the provisions of existing laws are critically analysed, definitional deficiencies are highlighted and reform proposals are generated.

---

[149] Data Protection & Privacy Commissioners, (2005) "Montreux Declaration: The protection of personal data and privacy in a globalised world: a universal right respecting diversities," <http://www.libertysecurity.org/IMG/pdf/montreux_declaration_eng.pdf> (Last accessed 20.02.11)
[150] *Ibid*
[151] Fleischer, P. (2007) "Call for Global Privacy Standards," Google Public Policy Blog, 14th Sept 2007 <http://googlepublicpolicy.blogspot.com/2007/09/call-for-global-privacy-standards.html> (Last accessed 20.02.11)
[152] *Ibid*

## *3.2 Impetus for enactment of privacy and data protection laws*

The main driver for the introduction of privacy laws was a concern for protecting human rights, prompted by World War II atrocities, whereas the main drivers for the introduction of data protection laws were: a concern for allaying citizens fears about government surveillance capabilities, and, a concern for facilitating transborder personal data transfers in order to promote economic gain. These will be explored in more detail below.

### *3.2.1 Impetus for enactment of privacy laws*

The genocidal atrocities perpetrated by the Nazis during World War II provided an impetus for the introduction of privacy and data protection laws in Europe, since the holocaust was inadvertently facilitated by the data collection practices employed by European Governments. For instance, in the 1930's, the Netherlands implemented a 'cradle to grave' population registration system that included the collection of comprehensive data on citizens. By accessing these registers, the Nazi regime systematically identified members of religious and ethnic groups with ease and subjected them to persecution. According to Seltzer & Anderson, Dutch Jews had the highest death rate of Jews residing in all occupied countries in Western Europe during this period.[153] Naturally, these experiences imprinted deep and resonating privacy concerns in the countries affected by the Nazi regime, and provided the impetus for recognition of privacy as a fundamental human right in the UN Declaration of Human Rights, the International Covenant on Civil and Political Rights and in the European Convention on Fundamental Freedoms and Human Rights. Thus, governments sought to allay, at an international legislative level, widespread public concerns about privacy through the General Assembly of the United Nations adoption of the Universal Declaration of Human Rights (UNDHR, 1948).[154] It includes *Article 12*:

> "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

[153] Seltzer, W. & Anderson, M. (2001) "The Dark side of Numbers: The role of Population Data Systems in Human Rights Abuses," *Social Research* , pp. 486-488

[154] Universal Declaration of Human Rights (UNDHR, 1948) <http://www.un.org/Overview/rights.html> (Last accessed 20.02.11) The Universal Declaration was adopted by the General Assembly on 10 December 1948 by a vote of 48 in favour, 0 against, with 8 abstentions (all Soviet Bloc states [i.e., Byelorussia, Czechoslovakia, Poland, Ukraine and The USSR],Yugoslavia, South Africa and Saudi Arabia

The Declaration was followed at the international level by the adoption of the International Covenant on Civil and Political Rights (ICCPR, 1966),[155] Article 17 of which is identical to UNDHR article 12.

At the European Level, Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR, 1950),[156] stipulates a right to respect for private and family life:

> "(1) Everyone has the right to respect for his private and family life, his home and his correspondence.
>
> (2) There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."

The ECHR created the European Commission of Human Rights and the European Court of Human Rights (ECtHR) to oversee enforcement. Both bodies have consistently viewed Article 8's protections broadly and interpreted the restrictions narrowly.[157] For instance, in the case of *X v Iceland*, the Commission determined that:

> "For numerous Anglo-Saxon and French authors, the right to respect "private life" is the right to privacy, the right to live, as far as one wishes, protected from publicity...In the opinion of the Commission, however, the right to respect for private life does not end there. It comprises also, to a certain degree, the right to establish and develop relationships with other human beings, especially in the emotional field for the development and fulfilment of one's own personality."[158]

The ECtHR has also reviewed member states' laws and imposed sanctions on numerous countries for failing to regulate wiretapping by both governments and private individuals.[159] It has also reviewed cases of individuals' access to their personal information in government files to ensure that adequate procedures exist,[160] and expanded the protections of Article 8 beyond government actions to those of private persons where it appears that the government should have prohibited those actions.[161]

---

[155] International Covenant on Civil and Political Rights (ICCPR, 1966) <http://www.unhchr.ch/html/menu3/b/a_ccpr.htm> (Last accessed 20.02.11)
[156] Entered into force September 1953
[157] Strossen, N. (1990) "Recent United States and International Judicial Protection of Individual Rights: A comparative Legal Process Analysis and Proposed Synthesis," *Hastings Law Journal*, Vol. 41, p. 816
[158] X v Iceland, 5 European Commission on Human Rights 86.87 (1976).
[159] European Court of Human Rights, Case of Klass and Others: Judgement of 6 September 1978, Series A No. 28 (1979); Malone v. Commissioner of Police, 2 All E.R. 620 (1979). Burke, K. C., (1981) "Secret Surveillance and the European Convention on Human Rights," *Stanford Law Review*, Vol. 33, No. 6, p. 1122
[160] Leander v Sweden, Judgement of 26 March 1987, 9 EHRR 433
[161] *Ibid*, pp. 848-49.

More recently, at the European Level, the European Commission, Council and Parliament developed the Charter of Fundamental Rights of the European Union,[162] which came into force in 2009, alongside the Treaty of Lisbon.[163] Article 7 of the charter states that:

> "Everyone has the right to respect for his or her private and family life, home and communications."

Additionally, Article 8 of the Charter states that:

> "Everyone has the right to the protection of personal data concerning him or her."

Thus, the charter conceives of two separate fundamental rights; Art 7 is concerned with a broad notion of *limited physical access* privacy, whilst Art 8 is concerned with *individual data* privacy. The European Union (EU) is obliged to act and legislate consistently with the Charter and the EU's courts will strike down EU legislation which contravenes it. However, in the negotiations leading up to the signing to the Lisbon Treaty, the United Kingdom secured a protocol to the treaty relating to the application of the Charter. Of particular note is that Article 1(1) which precludes both the domestic courts in the UK and the EU's courts from finding that "laws, regulations or administrative provisions, practices or action" in the countries to which it applies are inconsistent with the Charter. However, Pernice asserts that the protocol is an interpretative protocol which will either have limited or no legal consequence.[164]

All four post World War II legislative measures recognise a need for a mechanism to protect the privacy of individuals. However, a comparative textual analysis reveals interesting differences in wording and content between the International and European legislative provisions. In both international treaties a right to privacy is recognised on an indirect basis, "a consequence of the duty of every person not to intrude in the private life of another" whereas, in both the European Convention and the European Charter there is specific recognition of a right to privacy, "every person has a right to the respect for his familial and private life, for his home and correspondence." These textual differences are important as, arguably, they give rise to two slightly different

---

[162] Charter of Fundamental Rights of the European Union (2000/C, 364/01) <http://www.europarl.europa.eu/charter/pdf/text_en.pdf> (Last accessed 21.0611)
[163] Official Journal C 306 , 17 December 2007
[164] Pernice, I. E. A., (2009) "The Treaty of Lisbon: Multilevel Constitutionalism in Action," *Columbia Journal of European Law*, Vol. 15, Vol. 3, pp. 349- 407

rights. The Universal Declaration and International Covenant give rise to a derivative right, that is, a protection by default, as evidenced by the wording 'consequence.' Under these treaties there is an obligation of non-intrusion, which by default gives rise to a right to privacy. Further, the right to privacy is expressed on the one hand in terms of banning attacks or interference and on the other in terms of a right to legal protection against such interferences. In contradistinction, protection for privacy is directly recognised by the European Convention and the European Charter, as privacy is the legal principle to be protected and the limits are the exceptions. Thus, while the protection under the international treaties is against an attack on the right to privacy it is underpinned by a duty not to intrude; whereas under the terms of the European Convention and European Charter, a right to a private sphere is envisaged; a sphere, which contains private and familial life, home, and correspondence.

There are also differences between the European Convention and European Charter rights. In particular, the Convention explicitly states that the protection afforded to privacy by the Convention is not absolute, whereas the Charter is silent on this issue. Article 8 (2) of the Convention provides explicit criteria that delineate the circumstances in which the imposition of restrictions on privacy are legitimate. These are: (i) the restriction of privacy must be foreseen by law; (ii) the restriction can (iii) it can only be used to achieve one of the specific and limited goals set out in Art 8 ECHR, including public security and the safeguarding of rights and freedoms of others (legitimacy criterion); (iv) the ECJ has added the condition that any action must be useful, indispensible and proportional to achieve the set goal (proportionality criterion).
Also, unlike the Convention, the Charter gives rise to two separate rights, one concerned with privacy protection, and the other with data protection. These separate rights mirror the distinction between physical access privacy and access to an individual's information discussed in chapter two.

To summarise, the normative basis for privacy laws derives from fundamental human rights set out in post World War II multilateral legislative instruments, which expressly recognise privacy as a fundamental human right, and interpret this right to privacy broadly in line with both the physical access definitions of privacy (the right to be let alone; limited access to the self; personhood and intimacy) and the informational access definitions of privacy (secrecy and control of personal information) outlined in chapter two. The Charter is unique, in that it makes an explicit distinction between data protection and privacy rights. At this juncture it is appropriate to examine the historical

catalysts for the introduction of data protection laws, before proceeding to analyse the relationship between privacy and data protection.

### *3.2.2 Impetus for enactment of data protection laws*

The main drivers for the introduction of data protection laws were: technological and organisational developments, a concern for allaying citizens' fears about surveillance capabilities, and a concern for facilitating transborder personal data transfers in order to promote economic gain.

In the pre-computing era, there were natural barriers for data protection, namely the cumbersome and expensive methods involved in data collection and processing. Indeed:

> "[U]p until the 1960's, most surveillance was low-tech and expensive since it involved following suspects around from place to place and could use up to 6 people in terms of two working 3 eight hour shifts. All of the material and contacts gleaned had to be typed up and filed away with little prospect of rapidly cross checking. Even electronic surveillance was highly labour intensive. The East German police, for example, employed 500,000 secret informers, 10,000 of which were needed just to listed and transcribe citizen's phone calls."[165]

However, in the 1960's the development of mainframe computing technology facilitated and accelerated data flows, making information available virtually anywhere in the world. Also, during this period national governments started to adopt mainframe computers to process an ever increasing volume of information generated by their social welfare systems.[166] For instance, in Sweden in the second half of the 1960s the government proposed to merge data concerning taxation with census data should be in a national databank. Similarly, in Germany there were plans for the connection of databanks on local, state and federal levels, and there was a scheme to centralize data procession in administration on state level (for example in Hesse and Bavaria).[167] Similarly, businesses began to harness computer technology in the 1960's to assist their daily operations, for instance, proposals emerged to link databanks in order to help organisations reduce fraud, better target customer needs an increase their efficiency. However, the databank proposals were quickly met with opposition from groups in societies that feared the network of databanks could be used for nefarious purposes and permit governments and organisations to gain excessive informational power over

---

[165] Scientific and Technological Options Assessment Unit of the European Parliament (STOA) (1998) "An Appraisal of Technologies of Political Control," point 4. <http://cryptome.org/stoa-atpc.htm> (Last accessed 20.02.11)

[166] Mayer–Schönberger, V. (1997) "Generational Development of Data Protection in Europe," in Agre, P. & Rotenberg, M. (eds) *Technology and Privacy: The New Landscape*, (Cambridge, Massachusetts, The MIT Press) p. 222.

[167] *Ibid*, p. 222

individual citizens.[168] Thus, the increasing portability of data and the ease with which information could be aggregated, mined and exchanged triggered an intensification of privacy concerns. In response to these concerns, legislatures passed data protection laws. Initially, data protection laws were developed at regional or national level.

### *3.3 Development of national data protection laws*

The first data protection law was enacted in the German Land of Hesse in 1970.[169] It was quickly followed by legislative measures in Sweden[170] and France,[171] though not all counties were persuaded of the merits of enacting data protection laws.[172] However, by the beginning of the 1980s it had become apparent to legislators that personal data processing was not restricted to mainframe computers, but rather was increasingly conducted on networked computers. Moreover, some of these networks were transnational structures e.g. it was becoming common for financial services data and airline passenger data to be collected, transferred and processed across a number of national borders. Such transborder data flows generated challenges for national regulators, as they could not ensure compliance with national legislation once the data had left their jurisdiction. According to Bainbridge:

> "To prohibit transborder data flows to countries without data protection laws in any kind of systematic way would have enormous economic and therefore political consequences."[173]

Unsurprisingly, when national data protection authorities began to use their legislative power to restrict the flow of information within Europe to countries that lacked data protection laws it led to tensions. For instance, in 1989, the French data privacy authority (CNIL) blocked the transfer of personal data about employees and customers from Fiat France to the parent company, Fiat Italy, arguing that the absence of data protection laws in Italy rendered the transfer illegal.[174] As a temporary solution Fiat

---

[168] Newman A, "Protecting Privacy in Europe: Administrative Feedbacks and regional Policies," in Meunier, S. & McNamara, K. R. (2007) (eds) *Making history: European Integration and Institutional change at fifty* (OUP)

[169] Hessisches Datenschutzgesetz of 20th September 1970 (Data Protection Act of the German Federal State of Hessen)

[170] Sweden introduced its first Data Protection statute in 1973

[171] France, Austria, Norway and Denmark enacted data protection laws in 1978.

[172] Notably, the UK was slow to legislate; it did not enact a Data Protection Act until 1984. British firms were the most adamant in their opposition, warning that European intervention would have severe implications for industry; Dwek, R (1990) "EC Scheme for Data Protection Stuns UK," *Marketing*, 12th July, p. 3

[173] Bainbridge, D. (1996) '*EC Data Protection Directive*,' (Butterworths) p. 15

[174] Fauvet, J. (1989) "Privacy in the New Europe," Transnational Data & Communications Report, Nov 17-18 cited in Newman A, "Protecting Privacy in Europe: Administrative Feedbacks and regional

was required to sign a contract with CNIL that it would guarantee privacy protection for any information transferred from France.[175] Such data transfer disputes had a political and economic impact and led to calls for supranational laws.

## *3.4. Development of supranational data protection laws*

To date five supranational legislative measures have been developed; at the international level: the OECD Guidelines (1980), Council of Europe Convention 108 (1980) and UN Guidelines (1990), at the European Level: the EU Directive 95/46/EC, and most recently, the Asian Pacific Rim countries have become signatories of the APEC Privacy Framework (2005). They will be discussed in turn, as the rationale, remit and scope of each of these legislative measures differs. It is important to critically analyse the provisions of these existing laws in order to explicate the deficiencies in existing law, particularly in light of the claim by Cate that:

> "Modern privacy law is often expensive, bureaucratic, burdensome, and offers surprisingly little protection for privacy. It has substituted individual control of information, which it in fact rarely achieves, for privacy protection." [176]

### *3.4.1 The OECD Guidelines*

Firstly, in 1980, in response to calls for international data protection measures, the Organization for Economic Cooperation and Development (OECD), which is an international economic organisation of 34 countries (founded in 1961 to stimulate economic progress and world trade), formulated the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.[177]

Since their development, despite not being legally binding, they have been adopted by 30 countries around the world, including the UK, Australia and the USA. The Guidelines apply to "personal data, whether in the public or private sectors, which, because of the manner in which they are processed, or because of their nature or the

---

Policies," in Meunier, S. & McNamara, K. R. (2007) (eds) *Making history: European Integration and Institutional change at fifty* (OUP)

[175] Transnational Data and Communications Report, (1989) "No Fiat for Fiat," November, p.10

[176] Cate, F. (2006) "The Failure of Fair Information Practice Principles," in Winn, J. K. (ed) *Consumer Protection in the Age of the Information Economy,(Ashgate, UK)*

[177] Text of the Guidelines available at: <http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html> (Last accessed 20.02.11)

context in which they are used, pose a danger to privacy and individual liberties."[178] The rationale for their introduction was set out in the preface which cautions that:

> "The development of automatic data processing, which enables vast quantities of data to be transmitted within seconds across national frontiers, and indeed across continents, has made it necessary to consider privacy protection in relation to personal data. Privacy protection laws have been introduced, or will be introduced shortly, in approximately one half of OECD Member countries…to prevent what are considered to be violations of fundamental human rights, such as the unlawful storage of personal data, the storage of inaccurate personal data, or the abuse or unauthorised disclosure of such data."[179]

The preface to the Guidelines also indicates that concerns surrounding the processing of data using computing technology and the related ability to transfer personal data beyond national borders were key drivers:

> "there is a danger that disparities in national legislation [sic] could hamper the free flow of personal data across frontiers; these flows have greatly increased in recent years and are bound to grow further with the widespread introduction of new computer and communications technology. Restrictions on these flows could cause serious disruption in important sectors of the economy, such as banking and insurance."[180]

However, Gutwirth observes that the OECD's main concern is economic; it views personal data as an asset:

> "So it is only natural the OECD guidelines serve the interest of the economic forces, which need a profitable fast and free international flow of data."[181]

Similarly, Kirby[182] asserts that the aim of the OECD Guidelines is to create an international legal area within which personal data can freely circulate, by developing an agreed set of guidelines for transborder data flows. Likewise, Gutwirth[183] claims that the OECD wanted to erase any differences between legislation in member states since divergences in the protection of privacy could be used by OECD states to regulate and limit data flows. For instance, a state which offered a high privacy protection threshold could ban the export of data to a nation with a low threshold. Thus, Kirby acknowledges that:

---

[178] *Ibid*
[179] *Ibid*
[180] *Ibid*
[181] Gutwirth, S. (2002) *Privacy and the Information Age*, (Rowman & Littlefield), p. 88
[182] Kirby, M. (2011) "The history, achievement and future of the 1980 OECD guidelines on privacy," *International Data Privacy Law*, Vol. 1, No. 1, pp. 6-14
[183] Gutwirth, S. (2002) *Privacy and the Information Age*, (Rowman & Littlefield), p. 88

> "The suspicion that several non-European countries had was that the European treaty approach to protecting privacy was heavy-handed with bureaucracy; potentially expensive to implement; insufficiently sensitive to the values of TBDF; and (even possibly) motivated by economic protectionism so as to strengthen the European technology of informatics behind legally established data protection walls. The suspicion of Europeans was that the non-European member states would insist on a 'toothless tiger'. They would give the appearance of agreement; but without any real or practical effectiveness."[184]

Therefore, the primary aim of the OECD Guidelines is to avoid the creation of unjustified data protection obstacles to the transborder flow of personal data, since this would impede economic development. These principles were designed to "represent a consensus on basic principles which can be built into existing national legislation"[185] and to "serve as a basis for legislation in those countries which do not yet have it."[186] However, Gutwirth is critical of the Guidelines, questioning whether they:

> "are little more than a privacy friendly front for hiding the true purpose of promoting an economic policy which puts personal data on the same level as any other economic product. The establishment of a minimum level of legal protection of personal data to compensate for their effective international traffic is a disappointing result."[187]

Indeed, the purpose of the OECD Guidelines is not to define privacy or private data and although they suggest that equal weight is afforded to privacy protection and the free flow of information, privacy protection is a really just an ill-defined consequence of data flow controls rather than an end in itself. Gutwirth[188] asserts that this approach is deficient because more controls on data flows may, in certain circumstances, improve the protection of privacy. Overall, the Guidelines represent a minimum level of privacy protection for personal data necessary to compensate for the effective international flow of data traffic.

### *3.4.2 Council of Europe Convention*

The Council of Europe's adoption of the 1981 Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data[189] was the next legislative measure introduced at the international level. The convention was produced by the Council of Europe, a body established after the Second World War for the

[184] Kirby, M. (2011) "The history, achievement and future of the 1980 OECD guidelines on privacy," *International Data Privacy Law,* Vol. 1, No. 1, pp. 8-9
[185] Gutwirth, S. (2002) p.88
[186] O.ECD. Doc. (C 58 final) (Oct. 1, 1980).
[187] Gutwirth, S. (2002) p. 89
[188] Gutwirth, S. (2002) p. 89
[189] ETS No. 108, 28.01.1981 The Convention was adopted in 1980 and opened for ratification in January 1981 <http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm> (Last accessed 27.02.11)

purpose of achieving greater unity of democratic countries in Europe through the promotion of human rights in its forty six member states.[190] Accordingly, Gutwirth opines that:

> "It is logical that the CoE puts more weight on the protection of private life in comparison with the substantially similar initiative of the OECD. Convention 108 is explicitly in line with Articles 8 and 10 of the ECHR. It refers to the balancing of interests between privacy, freedom of information, and freedom of speech, a common practice in the ECHR framework."[191]

Article 1 of the Convention states that its purpose is to protect privacy through the regulation of the processing of personal data:

> "to secure in the territory of each Party for every individual, whatever his nationality or residence, respect for his rights and fundamental freedoms, *and in particular his right to privacy with regard to automatic processing of personal data relating to him* ('data protection') [emphasis added]."[192]

Thus, the Convention does not define privacy or private data. Instead, it imposes restrictions on the processing of certain types of personal data that are considered more privacy sensitive. The explanatory report to the Convention noted that such measures were required in view of the increasing use of computers for administrative purposes. Convention 108 obliges signatory member states to implement the legislative measures a national level and, because of this, Convention 108 has had a major impact on the laws of many European nations. As Gutwirth asserts:

> "Convention 108 also leads to the isolation of countries without privacy rules, or 'data havens.' So these countries too are forced to adopt privacy legislation."[193]

However, Gutwirth is critical of the provisions of the Convention, stating:

> "the states commit themselves not to check and restrict the export of personal data to other signatory states. Instead of protectionism, we get deregulation. The free flow of personal information within international networks is given an official government escort."[194]

---

[190] Tomkins, A. (1997) "Civil Liberties in the Council of Europe: A Critical Survey" in Gearty, C. (ed) *European Civil Liberties and the European Convention on Human Rights: A comparative study*, (The Hague, Martinus Nijhoff Publishers) pp. 2-4

[191] Gutwirth, S. (2002) p. 90

[192] ETS No. 108, 28.01.1981 <http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm> (Last accessed 27.02.11)

[193] Gutwirth, S. (2002) p. 90

[194] Gutwirth, S. (2002) p. 90

### *3.4.3 UN Guidelines (1990)*

Thereafter, the UN,[195] an international organization founded in 1945 to stop wars between countries and to provide a platform for dialogue, issued Guidelines for the Regulation of Computerized Personal data files (1990)[196] that concern all member countries.[197] These guidelines do not offer a definition of 'privacy' or 'private' data, simply stating that privacy concerns "information about persons" or "personal data." The UN Guidelines are similar to the principles contained in the OECD Guidelines 1980 regarding the processing of personal data. Additionally, they mirror the provisions of the CoE Convention 108 in that they impose specific restrictions on the processing of seven categories of sensitive data in order to prevent discrimination, but include slightly different categories of sensitive data. However, the UN Guidelines are not binding, but rather serve as recommendations for member countries when implementing national privacy and data protection legislation.

In spite of the efforts by the UN, Council of Europe and OECD at the international level, harmonised privacy and data protection law did not occur, since national states were influenced by their particular economic and societal circumstances.

### *3.4.4 The EU Data Protection Directive 95/46/EC*

Conscious of internal market goals, the European Commission reviewed the impact of the OECD Guidelines, UN Guidelines and Convention 108, and found that they had not produced the desired result, that is, harmonized legislative measures facilitating transborder data flows. Indeed, by the early 1990s, some member states (e.g. Belgium) were becoming data havens,[198] while in others the legislative process had stalled.[199] The European Commission deemed that incidents such as that of the French subsidiary of the Italian car manufacturer Fiat described above[200] had a negative impact on the core principles of the EU: the free market; the free traffic of goods, people, services and capital; free enterprise; and a series of economic activities at the EU level. Accordingly,

---

[195] The UN's stated aims include facilitating cooperation in international law, international security, economic development, social progress, human rights, and achievement of world peace.

[196] United Nations High Commissioner for Human Rights, Guidelines for the Regulation of Computerized Personal Data Files, Resolution 45/95 of December 14th, 1990 <http://www.unhcr.org/refworld/docid/3ddcafaac.html> (Last accessed 27.02.11)

[197] The General Assembly of the United Nations is the main deliberative, policymaking and representative organ of the United Nations. It currently comprises 192 Members

[198] Gutwirth, S. (2002) *Privacy and the Information Age*, (Rowman & Littlefield p. 91

[199] *Ibid*

[200] *Ibid*

in 1990, it started laying the groundwork for a Directive on the processing of personal data. The Bangemann report recognised that:

> "the demand for protection of privacy will rightly increase as the potential of new technologies to secure (even across national frontiers) and to manipulate detailed information on individual from data, voice and image sources is realised…Europe leads the world in the protection of the fundamental right of the individual with regard to data processing."[201]

It took five years and an immense amount of lobbying and politicking in the European Parliament before the Directive came to fruition. Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data (hereafter 'The Directive')[202] was finally adopted in 1995, and member states had to implement its provisions by 1998.[203] Although, the goals of the European Union institutions are primarily economic, it was nevertheless, forced to impose a high level of privacy protection because it had to take existing laws and international legal instruments into account. It has since become the most important legislative measure relating to privacy protection standards for European Union member states, though Bennett and Raab have pointed out that:

> "The Directive was only possible because of prior agreement on data protection principles within the OECD and the Council of Europe. It attempts to rectify some of the perceived weaknesses within these instruments, especially with regard to the enforceability of data protection rules in a global economy."[204]

Its twin objectives are set out in Article 1:

> "(1) In accordance with this Directive, Member States shall protect the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of personal data. (2) Member States shall neither restrict nor prohibit the free flow of personal data between Member States for reasons connected with the protection afforded under paragraph 1."

Thus, one objective of the Directive was to ensure the free flow of data among member states and beyond, and a second objective was to create a unified level of privacy protection across member states. The harmonization of privacy protection is addressed by point 10 of the preamble:

---

[201] Bangemann Report, (1994) Europe and the Global Information Society <http://ec.europa.eu/archives/ISPO/infosoc/backg/bangeman.html> (last accessed: 20.02.11)
[202] The text of the Directive can be found here: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML> (Last accessed 20.02.11)
[203] The United Kingdom implemented the Directive through the enactment of the Data Protection Act 1998.
[204] Bennett, C. J., & Raab, C., D. (2003) *The governance of privacy: policy instruments in global perspective* (Ashgate, UK)

> "[w]hereas the object of the national laws on the processing of personal data is to protect fundamental rights and freedoms, notably the right to privacy, which is recognized both in Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms and in the general principles of Community law; whereas, for that reason, the approximation of those laws must not result in any lessening of the protection they afford but must, on the contrary, seek to ensure a high level of protection in the Community."[205]

Once again, although the Directive explicitly establishes a link between data protection and personal privacy, it fails to define privacy or private data. It does, however, state that certain types of data are considered more privacy sensitive, and thus requires explicit consent from the data subject to process such data.[206] Nevertheless, Gutwirth[207] opines that the European Commission intends to place the privacy protection bar high by specifically referring to Art 8 of the ECHR and to the fact that privacy is recognised in the general principles of community law. Similarly, the Directive proclaims that it will 'give substance to and amplify' the principles of Convention 108. However, Gutwirth claims that:

> "The establishment of a high level of protection must nevertheless be seen in light of the political impossibility of setting a low harmonized threshold. The main aim of this effort is to erase barriers limiting the free flow of personal data traffic though the establishment of an equivalent level of privacy protection. Of course, this cannot imply that some member states should reduce their established level of protection. The reduction of privacy to the lowest common denominator not only would make a mockery of the national efforts to protect privacy but also of Convention 108."[208]

The European legislator had thus been forced to impose a high level of privacy protection because it had to take existing laws and international legal instruments into account, and one of the objectives of the Directive is that it should lead to no diminution in the level of protection already provided in any existing national law.

Consequently, when the Court was asked in the *Bodil Lindqvsit case*[209] to determine whether a member state violates the Directive if it sets stricter data protection regulations compared to the ones stipulated by the Directive, the European Court of

---

[205] Directive 95/46/EC preamble Directive can be found here: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML> (Last accessed 20.02.11)
[206] Subject to legitimate exceptions set out in the Directive.
[207] Gutwirth, S. (2002) p. 91
[208] *Ibid* p. 92
[209] Bodil Lindqvist v Sweden, Case C-101/01, <http://curia.europa.eu/jurisp/cgi-bin/gettext.pl?lang=en&num=79968893C19010101&doc=T&ouvert=T&seance=ARRET> (Last accessed 20.02.11)

Justice ruled that the measures of member states have to comply both with the privacy measure of the Directive and the objective of free flow of data, but that "nothing prevents a Member State from extending the scope of the national legislation implementing the provisions of Directive 95/46 to areas not included in the scope thereof provided that no other provision of Community law precludes it."[210] Overall, whilst it has, to some extent, succeeded in harmonizing data protection laws among EU member states and the European Economic Area signatory countries, it did not lead to a global framework of data protection. Instead, it has resulted in protracted battles with other countries such as the USA and Australia, who were displeased at their companies and organisations being subjected to investigations regarding the 'adequacy' of their data protection measures. Indeed, rather than become a global gold standard of data protection, the Directive has found itself challenged by other regulatory measures, in particular the APEC Privacy Framework.

### *3.4.5 The APEC Privacy Framework (2005)*

The most recent legislative measure was introduced by the Asia-Pacific Economic Co-operation (APEC) in 2005. Like the OECD, the APEC's activities focus solely on facilitating economic development,[211] as is demonstrated by its Three Pillars: Trade and Investment Liberalisation Business Facilitation, and Economic and Technical Cooperation,[212] but its focus is the twenty one member countries situated in the Asia-Pacific rim. The APEC Privacy Framework is largely based upon the OECD Guidelines,[213] thus, it seeks to develop appropriate privacy protections whilst also preventing the creation of unnecessary barriers to personal information flow.[214] Indeed, Para 8 of the preamble states that the Framework was developed in recognition of the importance of:

---

[210] *Ibid* Para, 98

[211] APEC's focus on economic concerns has given rise to criticism when contrasted with human rights approach of European organizations such as the European Union or the Council of Europe. Academics such as McCormick have identified ASEM, the Asia-Europe Meeting, as an organization with greater potential to create change in Asia. McCormick, S. (2004) "ASEM A Promising Attempt to Overcome Protective Regionalism and Facilitate the Globalization of Trade," *Annual Survey of International & Comparative Law*, Vol. 10, p. 233.

[212] APEC Secretariat, (2005) "APEC at a Glance," <http://www.apec.org/About-Us/~/media/Files/AboutUs/About%20APEC/210secAPECGlance.ashx> (Last accessed 25.02.11)

[213] *Ibid*, p. 2

[214] Greenleaf, G., (2003) *"*Australia's APEC privacy initiative: the pros and cons of 'OECD Lite'*,"* *Privacy Law & Policy Reporter*, Vol. 10, pp. 1–6; Greenleaf describes the Privacy Framework's Principles as 'OECD Lite' on the basis that they are diluted versions of the main principles of the OECD Guidelines

> "Developing appropriate privacy protections for personal information, particularly from the harmful consequences of unwanted intrusions and the misuse of personal information;
> Recognizing the free flow of information as being essential for both developed and developing market economies to sustain economic and social growth;
> Enabling global organizations that collect, access, use or process data in APEC member economies to develop and implement uniform approaches within their organizations for global access to and use of personal information."

Bulford claims that:

> "The intended purpose of the APEC Framework, however, is to permit implementation of the Principles in a manner that effectuates privacy protection in commerce without deeming privacy an absolute right."[215]

Thus, like the foregoing legislative measures, it does not define privacy or private data. Moreover, it does not confer on an individual a right to data privacy, but it is noteworthy since it is the first legislative measure to introduce a concept of 'harm' from the misuse of personal data; principle 1, para 14 states:

> "Recognizing the interests of the individual to legitimate expectations of privacy, personal information protection should be designed to prevent the misuse of such information. Further, acknowledging the risk that harm may result from such misuse of personal information, specific obligations should take account of such risk, and remedial measures should be proportionate to the likelihood and severity of the harm threatened by the collection, use and transfer of personal information."

Unlike the EU Data Protection Directive, the APEC Framework is not binding upon member countries, and accordingly, does not require treaty obligations from its participants or in any way limit their behaviour through its actions.[216] Thus, the form of national laws may vary widely,[217] since the Framework explicitly contemplates that member economies will vary their implementation of the Principles, based upon:

> "differences in [their] social, cultural, economic, and legal backgrounds."[218]

Moreover, although the Privacy Framework has been adopted by APEC in principle, it will not be fully implemented until all member economies have implemented privacy

---

[215] Bulford, C. (2007-08) "Between East and West: The APEC Privacy Framework and the Balance of International Data Flows," *I/S: A Journal of Law & Policy* Vol. 3. No. 3, p.720
[216] *Ibid*
[217] Guyon, R. (2006) "Outline of Privacy and Spain Laws in Japan and Australia (From a Company Perspective) and APEC Privacy Framework Brief Overview*,"* 865 PLIIP, p. 616 (June/July 2006), cited in Bulford, C. (2007-08) "Between East and West: The APEC Privacy Framework and the Balance of International Data Flows," *I/S: A Journal of Law and Policy*, Vol. 33, p.719
[218] APEC Secretariat, (2005) APEC Privacy Framework, <http://www.ag.gov.au/www/agd/rwpattach.nsf/VAP/(03995EABC73F94816C2AF4AA2645824B)~APEC+Privacy+Framework.pdf/$file/APEC+Privacy+Framework.pdf> (Last accessed: 25.02.11) s 12

policies built around its Principles, yet the Framework has no legislative timeline, nor does it create a penalty for non-compliance. At this stage, given the voluntary nature of the APEC Privacy Framework, it is not a competitor to Directive 95/46/EC. However, if widely adopted, it could, over time become as influential as the Directive, given that the EU has 27 member states, while the APEC Framework has 21 member states and the economies in some member countries e.g. China are growing rapidly and becoming increasingly influential in trade relations. Thus, Fleischer predicts that the APEC framework could form the basis of a global privacy and data protection law:

> "the APEC Framework is the most promising foundation on which to build, especially since competing models are flawed (the USA model is too complex and too much of a patchwork, the EU model is too bureaucratic and inflexible)."[219]

## *3.5 Relationship between privacy and data protection*

The foregoing analysis indicated that all of the data protection measures claim to protect privacy, but none of them define 'privacy' or 'private' data. In concurrence with Bygrave, it is submitted that one of the key problems with current data protection measures is that they utilise the terms data protection and privacy protection interchangeably, which is inappropriate since:

"it would be wrong to assume that the concepts of "data protection" and "privacy" are completely synonymous. While closely linked, they are not identical." [220]

For instance, as acknowledged in the Lindop report,[221] the use of inaccurate or incomplete information for taking decisions about people is properly a subject for data protection, but it may not always raise questions of privacy. Wacks opines that:

> "Data protection statutes are not fashioned to provide comprehensive protection for individual privacy, but they routinely stipulate that personal data must be collected by means that are both lawful and fair. Such legislation thus affords incidental protection to privacy."[222]

Thus, data protection laws seek to protect many interests of the data subject, of which his privacy is only one, as noted by Bygrave when he stated:

---

[219] Fleischer, P. (2007) "The Need for global privacy standards," UNESCO Conference, Ethics and Human Rights in the Information Society, (13-14 September), <http://portal.unesco.org/ci/en/files/25452/11909026951Fleischer-Peter.pdf/Fleischer-Peter.pdf>

[220] Bygrave, L. A. (2010) "Privacy and Data Protection in an International Perspective," *Scandinavian Studies in Law* p. 168

[221] Lindop, N. (1978) *Report of the Committee on Data Protection* (Chairman: Sir Norman Lindop), (Cmnd 7341), Home Office

[222] Wacks, R. (2010) *Privacy: A very Short Introduction,* (OUP, UK) p.124

> "data protection instruments are expressly concerned with setting standards for the quality of personal information. While adequate information quality can serve to secure the privacy of individuals, it breaks down into a multiplicity of interests (including concern for, inter alia, the validity, integrity, availability, relevance and completeness of data) that have little *direct* connection to privacy-related values."[223]

Moreover, although the term 'privacy' is included in all the data protection laws, they do not seek to protect the broad notion of privacy found in human rights based privacy laws. As observed in the Lindop report:

> "There are aspects of privacy which have no immediate connection with the handling of personal data in information systems, such as intrusion into the home, power of entry and search, and embarrassing publicity in the media. There are also aspects of data protection which have no immediate connection with privacy. For example, the use of inaccurate or incomplete information for taking decisions about people is properly a subject for data protection, but it may not always raise questions of privacy."[224]

Thus, data protection laws have, as their focus, a narrower conception of privacy, which relates to Westins's conception of informational/data privacy, that is, they protect data about people, rather than a private sphere of action. Clarke opines that this approach is:

> "justified on the pragmatic grounds that it is an operational concept more easily coped with by business and government agencies than the abstract notion of privacy, and it is therefore easier to produce results."[225]

However, he is critical of this legislative approach because he argues "it's not what humans actually need."[226] Similarly, Rule has criticised the provisions of existing data protection laws on the basis that:

> "data protection regimes so far have tended to operate with largely procedural rules that do not seriously challenge established patterns of information use but seek merely to make such use more efficient, fair, and palatable for the general public." [227]

Thus, he voices concerns that legislators have introduced such laws in order to promote public acceptance of new forms of data processing technologies, rather than with the

---

[223] Bygrave, L. (2001) "The Place of Privacy in Data Protection Law" *UNSWLawJ*, Vol. 24, No. 1, para 15

[224] Lindop, N. (1978) *Report of the Committee on Data Protection,* (Chairman: Sir Norman Lindop), (Cmnd, 7341) Home Office, p. 9

[225] Clarke, R. (1998) "A History of Privacy in Australia: Context" <http://www.anu.edu.au/people/Roger.Clarke/DV/OzHC.html#ContI> (Last accessed 20.02.11)

[226] Clarke, R. (2006) "Introduction to Dataveillance and Information Privacy, and Definitions of Terms" <https://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html> (Last accessed 20.02.11)

[227] Rule, J., McAdam, D., Stearns, L., & Uglow, D. (1980) *The Politics of Privacy: Planning for Personal Data Systems as Powerful Technologies*, (Elsevier, New York)

intention of assuaging legitimate concerns about privacy protection. Additionally, Cate asserts that compliance with data protection laws is increasingly focused on providing required notices in proper form and at the right time, rather than on ensuring that the privacy of personal information is protected. Accordingly, he asserts that:

> "Of the hundreds of enforcement actions brought in Europe, the United States, and other countries, few have involved allegations of *substantive harms* to individuals, while most have alleged failures to comply with procedural requirements. Meanwhile, serious risks to consumers, such as the apparent widespread insecurity of personal data, have gone largely unexamined. This is a powerful indictment of modern data protection law, and it requires not just tinkering with notice and choice requirements or rethinking enforcement strategies. It requires rethinking the purpose of data protection law and re-examining the principles on which that law is based."[228] (emphasis added)

Consequently, Cate calls for reform on the basis that existing data protection laws place burdens upon entities that must comply with those laws, whilst at the same time failing to provide adequate protection to individuals whose privacy they are supposed to be protecting. To this end, this thesis proposes that if a global privacy framework is drafted it should explicitly state that it is concerned with the protection of 'data privacy.' The terms 'informational' and 'data' privacy are sometimes used interchangeably, but in this thesis the term *data privacy* is used. This ensures consistency with current data protection legislation and also with the Lindop report which defined *data privacy* as:

> "the individual's claim to control the circulation of data about himself,"[229]

This approach will more clearly draw attention to the personal information privacy protection element of legal instruments. It would also accord with the recognition in the European Union Charter on Fundamental Rights that data protection should be regarded as a legal obligation that is distinct from privacy protection. Accordingly, this chapter proposes that the merits of the definition of 'data privacy' and 'private data' offered in chapter two, and repeated below:

> ***Data privacy*** *concerns the legal regulation of the boundary between personal and private data.* ***Private data*** *is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to the individual.*

[228] Cate, F. (2006) "The Failure of Fair Information Practice Principles," in Winn, J. K. (ed) *Consumer Protection in the Age of the Information Economy*,(Ashgate, UK) p. 369
[229] Lindop, N. (1978) *Report of the Committee on Data Protection* , (Cmnd 7341)

be empirically tested in order to provide evidence to drive forward legislative reform proposals.

## *3.6 Summary*

This chapter revealed that privacy laws such as the UNDHR (1948), ICCPR (1966), ECHR (1950) and EU Charter expressly recognise privacy as a fundamental human right, and define and interpret this right to privacy broadly in line with both the physical access definitions of privacy (the right to be let alone; limited access to the self; personhood and intimacy) and the informational access definitions of privacy (secrecy and control of personal information) outlined in chapter 2. In contrast, none of the five supranational data protection measures (i.e. OECD Guidelines, UN Guidelines, CoE Convention, Directive 95/46/EC or APEC Framework) define 'privacy' or 'private' data. This chapter also revealed that although the terms data protection and privacy protection are often used interchangeably, they are not synonymous, since data protection is also concerned with the protection of a narrow, personal information conception of privacy, and simultaneously with the protection of other interests such as integrity, accuracy and security of data. Consequently, privacy protection is often incidental to other data processing activities, rather than a central concern of data protection laws.

A better approach would be based on explicit definitions of 'data privacy' and 'private data' such as those outlined in chapter two of this thesis and therefore it is the aim of this thesis test such definitions in order to determine whether they should form the basis of a future global privacy and data protection framework. To do so, it is necessary to first consider the adequacy of the terms personal and sensitive data, as they are the key terms in existing laws.

# Chapter 4

## Conceptual (In)adequacy

### *4.1 Introduction*

A review of the literature in chapter two indicated that although privacy is a protean concept, with contested meanings, it is worthy of legal protection. Thereafter, chapter three indicated that existing data protection laws do not define 'privacy' or 'private' data. Rather, they seek to achieve privacy protection by imposing conditions on the processing of personal and sensitive data. Thus, this chapter begins by exploring the terms 'personal' and 'sensitive' data. It will illustrate that there is a broad consensus regarding the definition of personal data in supranational legislative measures, and proceed to examine whether personal data is synonymous with private data. Also, this chapter will examine the term sensitive data in supranational legislative measures, demonstrating that the term is not universally defined in the legislative provisions; in fact, is not included in some legislative provisions. It will also explore whether sensitive data is synonymous with private data. The analysis will surmise that that the conceptual adequacy of existing data protection laws has been too readily accepted, and further that technological developments warrant a review of the 'fitness for purpose' of the existing legislation. This review is timely as Viviane Reding, the Vice-President of the European Commission and Commissioner responsible for Justice, Fundamental Rights and Citizenship, recently observed that the way individuals' data is used and shared in the information society is constantly changing. Reding asserts that the challenge posed to legislators is to:

> "establish a legislative framework that will stand the test of time... [also]... No matter how complex the situation or how sophisticated the technology, clarity must exist on the applicable rules and standards that national authorities have

> to enforce and that businesses and technology developers must comply with. Individuals should also have clarity about the rights they enjoy."[230]

Accordingly, this chapter will illustrate how technological advances are challenging the conceptual foundations of existing data protection laws; in particular sensitive personal data, since laws developed in an era of isolated mainframe computers are unable to respond to the needs of new technologies and the privacy expectations of individuals. Indeed, Tene claims that:

> "Although modelled to be technologically neutral and apply across industries, the Current Framework is in danger of being unravelled by a new generation of users utilizing a new generation of technologies. The fundamental concepts underlying the Current Framework, including basic terms such as 'personal data'... have been disrupted by shifting technological realities."[231]

The chapter will conclude by suggesting that empirical research be conducted to test the terms contained in data protection laws, and the merits of a harm based definition of 'data privacy' and 'private data,' outlined in chapter two.

## *4.2 Definitions of personal data*

Chapter three indicated that all data protection laws seek to achieve privacy protection through the imposition of rules regarding the processing of personal data, i.e. consent of the data subject must be obtained. In common parlance, the words 'personal' and 'private' are often used interchangeably. Indeed, the House of Commons recently produced a report entitled "Protection of Private data"[232] which examined the data security failures at HM Revenue & Customs (HMRC) that led to the loss of personal data. Within the report, no mention was made of the term private data. Such interchangeable uses of the terms 'personal' and 'private' give rise to the premise that the terms 'personal' and 'private' data are synonymous, and concomitantly, that privacy protection may be achieved though the consent based control of personal data processing. Thus, it is appropriate to examine how the term **'**personal**'** data is defined in existing data protection laws, particularly since the

[230] Reding, V. (2011) "Tomorrow's Privacy The upcoming data protection reform for the European Union," *International Data Privacy Law*, Vol. 1, No. 1, p. 3
[231] Tene, O. (2011) "Privacy: the new generations," *International Data Privacy Law*, Vol. 1, No. 1, p. 15
[232] House of Commons: Justice Committee (2008) "Protection of Private Data," First Report of Session 2007–08 HC 154.

Introduction of the Article 29 Working Party Opinion[233] on the concept of personal data in Directive 95/46/EC states:

> "Information about current practice in EU Member States suggests that there is some uncertainty and some diversity in practice among Member States as to important aspects of this concept which may affect the proper functioning of the existing data protection framework in different contexts."[234]

Identical definitions of personal data are offered in part 1 of the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (1980), [235] and Art 2 of the Council of Europe Convention For The Protection of Individuals With Regard To Automatic Processing Of Personal Data (1981);[236] personal data is defined as:

> "any information *relating to* an identified or identifiable individual (data subject),"

The term personal data is not defined in the UN Guidelines for the Regulation of Computerized Personal Data Files (1990),[237] presumably because it refers to and builds upon the previous international legislative instruments. In contrast, a comprehensive definition is offered in Art 2 (a) of Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data,[238] which states that personal data means:

> "any information *relating to* an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity." (emphasis added)

Recital 26 of the Directive adds the following clarification:

> "to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person."

---

[233] Article 29 Working Party, (2007) Opinion 4/2007 on the concept of personal data (WP136 of 20 June 2007).
[234] Ibid, p 3
[235] OECD Guidelines (1981) <http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.htm l > (Last accessed 01.03.09)
[236] European Treaty Series (28.I.1981), <http://conventions.coe.int/Treaty/EN/Treaties/Html/108.htm >(Last accessed: 01.03.09)
[237] Adopted by General Assembly resolution 45/95 of 14 December 1990
[238] The text of the Directive can be found here: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML> (Last accessed 20.02.11)

Directive 95/46/EC was transposed into UK law through the enactment of the Data Protection Act 1998, which phrases the definition of personal data slightly differently in section 1(1), stating that it is:

> "data which relate to a living individual who can be identified:
> (a) from those data; or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller; and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual."

A differently phrased definition is found in the APEC Privacy Framework (2005), which defines personal data in part ii, section 9 as:

> "any information about an indentified or identifiable individual."

It further states that the Framework is intended to apply to:

> "information about natural living person, not legal persons. The APEC Privacy Framework applies to personal information, which is information that can be used to identify an individual. It also includes information that would not meet this criteria alone, but when put together with other information would identify an individual."

Although there appears to be a high degree of commonality between all the above mentioned definitions, Booth *et al* observe that:

> "a widespread, definitive understanding of the concept of 'personal data' has been assumed by commentators and policy makers alike. However, when one examines the debates and questions which have started to emerge following the implementation of Directive 95/46/EC, it becomes apparent that there is a strong case for a rigorous re-consideration of the conceptual foundations of 'personal data.'" [239]

Korff[240] conducted a review of how the provisions of Directive 95/46/EC had been implemented into the national laws of member states. His comparative textual analysis reported that the definitions of personal data in the UK Data Protection Act 1998 (the 'Act') and Directive 95/46/EC are consistent in their use of the phrase 'relate to,' but, under the Directive, consideration is first directed to whether the information relates to an identifiable individual and then whether it is processed, whereas the definition of personal data in the Data Protection Act approaches the

[239] Booth, S., Jenkins, R., Moxon, D., Semmens. N. Spencer, C., Taylor, M. & Townend, D. (2004) "What are 'Personal Data'?" A study conducted for the UK Information Commissioner

[240] Korff, D. (2002) EC Study on the Implementation of Data Protection Directive, Comparative study of national laws <http://www.eu.int/comm/justice_home/fsj/privacy/docs/lawreport/consultation/univessex-comparativestudy_en.pdf> (Last accessed: 23.06.07)

concept in the reverse order, as the Act focuses on the issue from a processing view first and then moves on to whether or not there is an identifiable individual. The Directive and Act also differ with respect to when an individual should be considered as 'identifiable.' Booth *et al*[241] observed that the way that the phrase 'relate to' is interpreted has major implications regarding what is or is not classed as personal data. If it is interpreted very narrowly, the term personal data could be restricted to data which is capable of identifying an individual, either by itself or in combination with other data. Identification, in this context, could be direct or indirect. In contradistinction, if the term 'relating to' is interpreted broadly it could conceivably include any data which may 'affect' the individual in some way, regardless of its capacity to identify. The consequences of a narrow interpretation of 'relating to' will be explored in an analysis of the *Durant* decision.

## *4.3 Interpretation of 'personal' data by UK Courts*

In the case of *Durant v FSA*,[242] Mr Durant had lodged a complaint with the Financial Services Authority (FSA) following a legal dispute with Barclays bank. The FSA dismissed his complaint. He then made a subject access request for information held manually and electronically by the FSA on his complaint. The FSA released the information held in computerised form, but refused to disclose the information held on manual files. Mr Durant applied to the Court under s 7(9) of the DPA 1998 for an order requiring the FSA to comply with the subject access request. The Court of Appeal was asked to decide: was the information held by the FSA relating to the investigation of Mr Durant's complaint 'personal' data under the Data Protection Act 1998? The definitional issue which arose concerned whether the data could be said to 'relate to' Mr Durant.[243] Mr Auld LJ referred to Directive 95/46/EC and ruled that the statutory right of access under the DPA is designed to enable the data subject to:

> "check whether the data controller's processing of it unlawfully infringes his privacy and, if so, to take such steps as the Act provides…to protect it."[244]

---

[241] Booth, S., Jenkins, R., Moxon, D., Semmens, N., Spencer, C., Taylor, M. & Townend, D. (2004) "What are 'Personal Data'? : A study conducted for the UK Information Commissioner" <http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/final_report_21_06_04.pdf>

[242] [2003] EWCA Crim 1746

[243] Identifiabililty was not an issue because the information in the manual files essentially comprised letters of complaint written by Mr Durant and material generated in response to his complaint.

[244] [2003] EWCA Crim 1746, [27]

From this the Court concluded that the relevant information is:

> "information that affects [the data subject's] privacy, whether in his personal or family life, business or professional capacity."[245]

This interpretation of personal data means that not all identifying information will fall within the scope of 'personal' data. Rather, only information that is capable of adversely affecting the privacy of the data subject will be considered personal. In order to determine whether or not data 'relates to' the data subject, Auld LJ proposed two tests. The first test is:

> "whether the information is *biographical* in a significant sense, that is, going beyond the recording of the putative data subject's involvement in a matter or an event that has no personal connotations, a life event in respect of which his privacy could not be said to be compromised."[246] (emphasis added)

The second test is whether:

> "the information has the putative *data subject as its focus* rather than some other person with whom he may have been involved or some transaction or event in which he may have figured or have had an interest, for example, …an investigation into the some other person's or body's conduct that he may have instigated."[247] (emphasis added)

Buxton LJ agreed, stating that the potential effect of processing of particular data on an individual's privacy was the guiding principle. The court also drew support for a narrow interpretation of the term personal data from the wording of the DPA 1998. Auld LJ asserted that the DPA's definition of personal data extends to expressions of opinion about an individual which would be otiose if the words 'relate to' were construed broadly. Thus, the Court of Appeal ruled that the information about Mr Durant's complaints to the FSA or about their investigation of his complaint were not 'personal data' as the data did not relate to Mr Durant in the requisite sense, that is, the court decided that the information sought by Mr Durant was information about his complaints, as opposed to data relating to him. Furthermore, the court ruled that the mere fact that a document is retrievable by reference to the name of the data subject does not render the information personal data:

> "Whether it does so in any particular instance depends on where it falls in a continuum of relevance or proximity to the data subject."[248]

---

[245] *Ibid*, [28]
[246] *Ibid*, [28]
[247] *Ibid*
[248] (2003) EWCA Civ 1746

Thus, when an individual's name appears on a document, the information contained in that document will not necessarily be personal data about the named individual. Rather, it is more likely that an individual's name will be 'personal data' where the name appears together with other information about the named individual such as address, telephone number[249] or information regarding his hobbies.[250] This interpretation of the term personal data is very narrow and if this decision were to be followed, only information that is capable of adversely affecting the privacy of the data subject would be considered personal data.[251]

## *4.4 Relationship between personal and private data*

The interpretation of personal data offered in the *Durant* case has been criticised by many scholars for being too narrow, and misconceiving the purpose of a personal data definition.[252] The term personal data has been widely defined in data protection laws and interpreted broadly since it is the subject matter of data processing activities, not all of which concern privacy protection. For instance, the term personal is used to indicate that the legislation is concerned with the data of individuals, as opposed to corporate or legal entities e.g. the tithes collected by a church would not be considered personal data; neither would a company's accounts. Once it has been established that the data in question is personal data, then the data subject gains data protection rights e.g. the right of access enables the data subject to check the accuracy of the information and to seek to have any inaccuracies rectified. From an individual's perspective, inaccurate information may be more privacy protective than accurate information; however this example serves to reinforce the observations in chapter three that privacy protection is not the sole objective of data protection laws. Accordingly, Wacks asserts that it would be wrong to describe

[249] See European Court of Justice decision in *Bodil Lindqvist v Kammaraklagaren* (2003) C-101/01, paragraph 27, as referred to in paragraph 28 of the Durant judgment

[250] See Lindqvist case, paragraph 27, and Durant at paragraph 28

[251] Subsequently, the Art 29 Working Group issued an opinion on the concept of personal data, which contains a broader notion of personal data. Art 29 Data Protection Working Party 'Opinion 4/2007 on the concept of personal data' (2007), <http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf)> (Last accessed 01.03.09) Thereafter, the Office of the UK Information Commissioner issued a technical guidance note to the effect that Durant is relevant to the question of whether data "relates" to a living individual only in difficult cases where the information in question is not "obviously about" someone. However, the Durant decision has not been overruled, and neither the Art 29 WG Opinion, nor the ICO technical guidance note are legally binding. ICO, (2007) 'Data Protection Technical Guidance - determining what is personal' (v1.0 21.08.07) <http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/personal_data_flowchart_v1_with_preface001.pdf> (Last accessed 01.03.09)

[252] Lindsay, D. (2004) "Misunderstanding 'personal information': Durant v Financial Services Authority," *Privacy Law & Policy Reporter*, Vol. 10, No. 10, No. 13

every instance of the dissemination of personal information as a loss of privacy since:

> "it is not always individual privacy that is violated by the collection, use, storage, or transfer of personal data..."[253]

Indeed, the Lindop report contended that:

> "Privateness" is clearly not an attribute of [personal] data themselves, for the data may be regarded as very private in one context and not so private, or not private at all, in another. Equally, when data are regarded as private, that does not mean that they are, or should be, known only to the individual to which they refer: rather it means that he wants them to be known only to him and those others who he agrees should know them.[254]

Thus, this thesis contends that in locating the protection of data privacy at the level of personal information, two questions arise: firstly, what is to be understood by personal? Secondly, under what circumstances is a data to be regarded as private? This thesis contends that individuals seek to limit the disclosure of personal data which would cause unreasonable harm. For instance, employees are required to provide personal information such as name, address, national insurance number, bank account details etc. to employers, so that they can receive their wages. However, in some contexts such personal information can gain a private quality. For instance, whilst an employee may have provided their home address and phone number to the Human Resources Department, they would consider such personal data to be private, in the sense that it should not be provided to an abusive ex partner. Thus, this thesis suggests that the attempt by Auld LJ to conflate the terms 'personal' and 'private' is a fallacy, and that data protection laws would be conceptually coherent if they distinguished between personal and private data. However, at present there is a lack of empirical research regarding the merits of a private data definition. Accordingly, in later chapters of this thesis the definition of private data offered below will be tested:

> *Private data is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual*

Before that, it is necessary to examine the term sensitive data and consider whether the terms sensitive and private data are synonymous, since a number of data

[253] Wacks, R (2010) "*Privacy: A very short introduction*," (OUP, UK) pp.122-123
[254] Lindop, N. (1978) *Report of the Committee on Data Protection,* (Chairman: Sir Norman Lindop), (Cmnd, 7341) Home Office, p. 9

protection laws include the term 'sensitive' data and impose stricter consent requirements to process such data on the basis that it poses a greater privacy risk.

## *4.5 Definitions of sensitive data*

Chapter three indicated that some data protection laws seek to achieve privacy protection through the imposition of rules regarding the processing of sensitive personal data (i.e. explicit consent must be obtained from the data subject). Thus, it is appropriate to examine how the term **'**sensitive**'** data is defined in existing data protection laws. The term 'sensitive' data was first considered for introduction into international law by the expert group drafting the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (1980).[255] Sweden and the German state of Hesse had already incorporated the concept into national and state law.[256] Ultimately the drafters of the Guidelines (and more recently the drafters of the APEC Privacy Framework (2005)) decided not to include extra safeguards for designated categories of sensitive data. The absence of safeguards seems to be partly due to a failure to achieve consensus on which categories of data deserve special protection, as the guidelines state:

> "...it is probably not possible to define a set of data which are universally regarded as being sensitive." (para 19 (a)).

This approach may also reflect the belief that personal data does not need protection in the absolute sense and that appropriateness of protection is dependent upon the context in which the data are used. In contrast, specific categories of sensitive data were introduced into International law through the Council of Europe Convention For The Protection of Individuals With Regard To Automatic Processing Of Personal Data (1981).[257] Although the Explanatory Report[258] advocates a context based approach to determining risk of harm from personal data processing, it recognises

[255] OECD Guidelines (1981) <http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.htm l > (Last accessed 01.03.09)

[256] The emergence of data protection laws, starting in Hessen (Hesse is English translation) 1970 and Sweden 1973, was closely linked to use of computer technology as a tool for collecting and distributing personal data. Sieghart, P., (1976) *Privacy and Computers*, (Latimer, London). Simitis reports that Germany consistently rejected all abstract categorisations of personal data and instead focussed on a **context-orientated appreciation** of the data as sensitive or non-sensitive. Simitis, S. (1999), "Revisiting Sensitive Data," <http://www.coe.int/t/dghl/standardsetting/dataprotection/Reports/Report_Simitis_1999.pdf>

[257] European Treaty Series - No. 108, (28.I.1981), <http://conventions.coe.int/Treaty/EN/Treaties/Html/108.htm > (Last accessed: 01.03.09)

[258] <http://conventions.coe.int/Treaty/EN/Reports/HTML/108.htm > (Last accessed: 01.03.09)

exceptional cases where the processing of certain categories of data may encroach on individual rights and privacy interests.[259] These 'sensitive' categories are listed in Article 6 as:

> "Personal data revealing racial origin, political opinions or religious or other beliefs, as well as personal data concerning health or sexual life, may not be processed automatically unless domestic law provides appropriate safeguards. The same shall apply to personal data relating to criminal convictions."

Paragraph 44 of the Explanatory Report states that "revealing ... political opinions, religious or other beliefs" also covers activities resulting from such opinions or beliefs. Paragraph 45 indicates that "personal data concerning health" includes information concerning the past, present and future, physical or mental health of an individual. The information may refer to a person who is sick, healthy or deceased. This category of data also covers those relating to abuse of alcohol or the taking of drugs. The categories listed in Article 6 are not meant to be exhaustive. Rather, the Convention provides that a Contracting State should be free to include other categories of sensitive data, as data sensitivity depends on the legal and sociological context of the country concerned:

> "Information on trade union membership for example may be considered to entail as such a privacy risk in one country, whereas in other countries it is considered sensitive only in so far as it is closely connected with political or religious views." (para 48)

Subsequently, the United Nations issued Guidelines for the Regulation of Computerized Personal Data Files (1990)[260] which addressed the issue of sensitive data under a *Principle of non-discrimination.* The Guidelines defined such data as:

> "…data likely to give rise to unlawful or arbitrary discrimination, including information on racial or ethnic origin, colour, sex life, political opinions, religious, philosophical and other beliefs as well as membership of an association or trade union, should not be compiled."[261]

This international treaty is broader than the Council of Europe Convention (discussed above) as it includes the categories, ethnic origin and colour. In addition, it includes membership of trade unions or other associations. However, it does not include criminal convictions or health data. Both the Convention and the Guidelines provide opportunities for States to regulate risks stemming from the processing of personal

[259] Paragraph 43.
[260] Adopted by General Assembly resolution 45/95 of 14 December 1990
[261] <http://ec.europa.eu/justice_home/fsj/privacy/instruments/un_en.htm > Principle 5(Last accessed: 01.03.09)

data by applying an internationally approved regulatory model. Indeed, they remain free to enact rules that better fulfilled their requirements, or even to abstain from any legislative action.

In contrast, Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data, (the Directive)[262] sought to achieve harmonization of laws across member states by providing an exhaustive list of categories of sensitive data. Such data is defined in Article 8 (1) as:

> "Personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and … data concerning health or sex life."

Also, Article 8(5) also makes special provision for criminal records and the like:

> "Processing of data relating to offences, criminal convictions or security measures may be carried out only under the control of official authority, or if suitable specific safeguards are provided under national law, subject to derogations which may be granted by the Member State under national provisions providing suitable safeguards…"

The principle of sensitivity outlined in Arts 8(1) and 8(5) hold that the processing of seven types of data should be subject to stricter controls than other types of personal data. [263]

The Directive differs from the Council of Europe approach in two main respects: firstly, it includes the trade union membership as a specific category of sensitive data; secondly, the list is considered exhaustive, whereas the Council of Europe list is merely indicative. The Directive differs from the UN Guidelines as it lacks a category of data on colour or membership of association, but includes a category of criminal convictions. A more radical difference exists between the Directive and the OECD Guidelines, in which drafters adopt a contextual approach and do not specifically enumerate special categories of sensitive data.

---

[262] <http://europa.eu.int/comm/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf <http://europa.eu.int/comm/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part2_en.pdf >(Last accessed: 01.03.09)

[263] In principle, such data cannot be processed. Derogation is permitted under very specific circumstances. These circumstances include the data subject's explicit consent, processing mandated by employment law, where it may be impossible for the data subject to consent (e.g. blood test to the victim of a road accident), processing of data has been publicly announced by the data subject or processing of data about members by trade unions, political parties or churches. Member states may provide for additional exceptions for reasons of substantial public interest.

Given the differing approaches to defining sensitive data outlined above, it is appropriate to examine the relationship between sensitive and personal data.

## *4.6 The relationship between sensitive and private data*

As outlined in chapter one, the characteristics associated with individuals persecuted during World War II formed the basis of special/sensitive data classifications in current data protection laws. For instance, the persecution of individuals who were physically or mentally disabled, prompted the inclusion of a *health category*, whilst the persecution of ethnic Poles, Soviet civilians, Romany gypsies, Jews and people of black skin colour led to the inclusion of a category of *racial or ethnic origin*. Similarly, the persecution of communists, socialists and political prisoners resulted in the inclusion of three categories, namely: *philosophical beliefs, political opinions* and *criminal convictions.* Furthermore, persecution of homosexuals led to the inclusion of a *sex life* category, whilst the persecution of Jehovah Witnesses culminated in a *religious beliefs* category. Accordingly, legislators sought to give increased privacy protection to categories of personal data that were pre-determined as posing a greater privacy risk, by imposing more stringent requirements when processing such data. Turn asserts that legislators implicitly recognised that:

> "items of personal information...are not equally sensitive from the point of view of their dissemination causing harm or embarrassment to an individual."[264]

Thus, *prima facie*, the terms private and sensitive data are synonymous. Moreover, the decision to classify these types of data as sensitive on an *a priori* basis suggests that the types of data classified as sensitive are fixed and constant, that is, they are always equally privacy sensitive, and not subject to change.

However, this approach is not without problems. Indeed, Simitis reports that the classification of certain types of data as sensitive was contested at the time the Council of Europe Convention and Directive 95/46/EC were drafted. For instance, in relation to trade union membership:

[264] Turn, R. (1976) "Classification of personal information for privacy protection purposes," AFIPS National Computer Conference, pp.301-307, at p. 301

> "The Convention did not include them, the Directive cites them expressly. The divergence is all but accidental. At the time of the Convention especially the Scandinavian States saw no reason to mention them. In their view such a reference is simply superfluous once collective bargaining functions in a both efficient and frictionless way. For those however who advocated the inclusion, the decisive argument was that in their experience unionised workers were still discriminated and should consequently be protected by inhibiting the collection of data revealing trade union membership."[265]

Thus, Norway and Sweden were compelled to amend their laws, classifying trade union membership data as sensitive, even though this did not reflect their societal or economic conditions. On the other hand, Simitis notes that 'financial data' was not included in the list of sensitive data of either the CoE Convention or the Directive:

> "As to the elimination of the data regarding financial situation, the abandonment of their special status is not least a result of the growing impact of sunshine laws, whose main purpose is to increase the transparency of financial activities. That the access to information regarding the financial situations, and consequently the creditworthiness of the data subjects must nonetheless be limited is best illustrated by the crucial role of data protection laws in connection with the processing of consumer data."[266]

Consequently, the inclusion of a list of sensitive data in legislation leads to questions regarding the continuing relevance of existing categories; in particular, whether the list can be revised. This is an issue of great importance in the Information Society, since the list reflects data considered sensitive in a post World War II society. However, in the intervening period a raft of legislative measures that prevent discrimination on the basis of the categories of sensitive data listed in the Directive, CoE Convention and UN Guidelines have been introduced. For instance, in the UK, the Civil Partnership Act 2004 affords same sex couples the same property rights as heterosexual couples, whilst the Disability Discrimination Act 1995 (as amended by the Equality Act 2006), made it unlawful to discriminate against disabled persons in connection with employment, the provision of goods, facilities and services or the disposal or management of premises. The Employment Equality (Religion or Belief) Regulations 2003 made it unlawful to discriminate against workers because of religion or similar beliefs; whilst the Employment Equality (Sexual Orientation) Regulations 2003 made it unlawful to discriminate against workers because of their sexual orientation, and the Race Relations Act 1976 made it unlawful to treat a

[265] Simitis, S. (1999) "Revisiting Sensitive Data," <http://www.coe.int/t/dghl/standardsetting/dataprotection/Reports/Report_Simitis_1999.pdf>
[266] Simitis, S. (1999) "Revisiting Sensitive Data," <http://www.coe.int/t/dghl/standardsetting/dataprotection/Reports/Report_Simitis_1999.pdf>

person less favourably than another on racial grounds in the areas of employment, education, and the provision of goods, facilities, services and premises. Consequently, in the 21st century, there is less scope for discrimination on the basis of the categories listed in Directive 95/46/EC, the UN Guidelines and CoE Convention (108) and, as a result, the categories may no longer be relevant.

Also, the list reflects data that were considered sensitive in a in pre-internet era, but advances in technology may mean that the list does not include new types of sensitive data. Wong asserts that:

> "The question that has arisen is whether the categorisation falls short of the dangers highlighted in recent technological developments?"[267]

Indeed, Tene[268] notes that the Internet was developed when Directive 95/46/EC was just in its infancy, and so some types of data did not exist, e.g. genetic data, biometric data and clickstream data.[269] Recently, Poullet *et al*[270] recommend that *identification numbers* that enable many databases or data to be connected together and *profiles*[271] should be classified as sensitive data. However, Korrf[272] reports that creating new categories raises difficulties, for instance, Luxembourg, and the Netherlands define 'genetic' data as data on *health*, whilst Portugal defines 'genetic' data as data on *health and sex life*, whereas in Sweden the processing of such data is not formally regarded as falling within the specific category to which the rules on '*sensitive data'* apply. Recently, biometric data[273] has come to the fore as a privacy issue, yet, whilst it shares some similarities with genetic data, it is often used for different purposes, so

---

[267] Wong, R. (2007) "Data Protection Online: Alternative Approaches to Sensitive Data?" *Journal of International Commercial Law and Technology,* Vol. 2 No.1 pp. 9-16

[268] Tene, O. (2011) "Privacy: the new generations," *International Data Privacy Law,* Vol. 1, No. 1, p. 15

[269] the generic name given to the information a website can know about a user simply because the user has browsed the site."[269] Internet companies rely heavily on tracking clickstream data to profile user preferences in order to deliver customized services and advertisements to Internet users.[269]

[270]Poullet, Y., Dinant, J-M., de Terwange, C. & Perez-Asinari, M. V. (2004) "Report on the application of data protection principles to the worldwide telecommunication networks: Informational self-determination in the internet era," <http://www.coe.int/t/dghl/standardsetting/dataprotection/TPD%20documents/T-PD%20_2004_%2004%20E%20final%20Report%20POULLET.pdf> (Last accessed: 29.04.11)

[271] Bygrave defines profiling as "the process of inferring a set of characteristics (typically behavioural) about an individual person or collective entity and then treating that person or entity (or other persons or entities) in the light of these characteristics." Bygrave, L. (2000) "Minding the machine: art 15 of the EC Data Protection Directive and automated profiling," *Privacy Law and Policy Reporter,* Vol. 7, No. 1, Art.10

[272] Korff, D. (2002) EC Study on the Implementation of Data Protection Directive, Comparative study of national laws <http://www.eu.int/comm/justice_home/fsj/privacy/docs/lawreport/consultation/univessex-comparativestudy_en.pdf> (Last accessed: 23.06.07)

[273] the generic name given to the process of identifying people using unique bodily features e.g. including iris and retina scans, hand geometry, ear shape, facial recognition, gait recognition, and recently voice, odour, scent, and sweat pore analysis.

it is not clear whether it may be subsumed within the category of health data. Likewise, it is not readily apparent how profiles, identification numbers, or clickstream data could be subsumed within existing categories of sensitive data.

Furthermore, Simitis claims that a top-down, *a priori* classification of data as sensitive does not recognise that contextual factors influence the privacy sensitivity of information:

> "Sensitivity is no more perceived as an a priori given attribute. On the contrary, any personal datum can, depending on the purpose or the circumstances of the processing be sensitive. **All data must consequently be assessed against the background of the context that determines their use**. The specific interests of the controller as well as of the potential recipients of the data, the aims for which the data are collected, the conditions of the processing and its possible consequences for the persons concerned are factors that, put together, allow both the range and the effects of the processing to be discerned and thus to determine its degree of sensitivity."[274] (emphasis added)

An example of how innocuous personal information can become sensitive is offered by Turn:

> "while a person's name is usually public information, it becomes sensitive when associated with a system of psychiatric treatment records."[275]

Moreover, the sensitivity principle implies that all data concerned should be subject to the same degree of processing restrictions, but this approach is arguably too broad. For instance, data concerning an employee's absence from employment due to hay fever is considered as sensitive as that person's HIV status.[276] Accordingly, Turn asserts that sensitivity is a highly subjective and context-dependent property of personal information:

> "what one individual may consider very sensitive may be regarded with indifference by many others, and it is likely that there is a large range of sensitivity assessments for every information item."[277]

This 'context-dependent' approach is identical to the 'contextual integrity' approach advocated by Nissenbaum in chapter two of this thesis. As discussed earlier,

[274] Simitis, S. (1999), "Revisiting Sensitive Data," <http://www.coe.int/t/dghl/standardsetting/dataprotection/Reports/Report_Simitis_1999.pdf>
[275] Turn, R. (1976) "Classification of personal information for privacy protection purposes," AFIPS National Computer Conference, pp.301-307, at p. 301
[276] Lindqvist: the decision to upload a photo of church volunteers to a website amounted to processing of sensitive data, on the basis that the photo showed that one of the people in the photo had a broken leg.
[277] Turn, R. (1976) "Classification of personal information for privacy protection purposes," AFIPS National Computer Conference, pp.301-307, at p. 306

Nissenbaum asserts that instead of focusing on the classification of data as privacy sensitive or non-sensitive, it is particular contexts that make information privacy sensitive, and thus explains why individuals may seek to claim privacy entitlements over non-sensitive information. Thus, this thesis will test Poullet *et al's* assertion that:

> "the extremely broad definition of sensitive data ... makes it absolutely necessary to abandon the approach based on a definition of the actual nature of data."[278]

Accordingly, in later chapters of this thesis the effectiveness of the *a priori* classification of data as sensitive will be tested. Also, this thesis will empirically examine whether conflation of the terms 'sensitive' and 'private' is a fallacy, by testing the contextual, harm-based definition of private data outlined below:

> *Private data is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual*

## *4.7 Summary*

The foregoing analysis suggests that personal data is data concerning an individual that should be processed in accordance with the eight data protection principles,[279] which include: accuracy, secure storage etc. On the other hand, as established in chapter one, private data is a subset of personal data, and it is information that an individual would seek to limit or control disclosure of on the basis that it might cause them unreasonable harm. Thus, this thesis contends that the attempt by Auld LJ in *Durant v FSA* to conflate the terms 'personal' and 'private' is a fallacy, and data protection laws would be conceptually coherent if they distinguished between personal and private data.

In addition, this chapter revealed that in the absence of a universally agreed definition of private data, legislators used sought to protect 'sensitive data' as a proxy. However, the analysis of the literature indicated that contextual factors influence privacy perceptions, e.g. an individual seeking to avoid an abusive ex-partner may consider their name and address private, and restrict their listing in a

[278]Poullet, Y., Dinant, J-M., de Terwange, C. & Perez-Asinari, M. V. (2004) "Report on the application of data protection principles to the worldwide telecommunication networks: Informational self-determination in the internet era," <http://www.coe.int/t/dghl/standardsetting/dataprotection/TPD%20documents/T-PD%20_2004_%2004%20E%20final%20Report%20POULLET.pdf> (Last accessed: 29.04.11)

[279] They are set out in schedule one of the Data Protection Act 1998

public telephone directory. Accordingly, this thesis contends that *a priori* classification of data as privacy sensitive is fallacious.

Later chapters of this thesis will collect and analyse empirical data to explore satisfaction with the current categories of sensitive data, as well as testing potential new categories of sensitive data. It will demonstrate that conflation of the terms sensitive and private is legislatively convenient, but conceptually problematic.

The foregoing analysis leads me to claim that the most conceptually coherent approach to privacy protection in data protection laws emerges when Nissembaum and Solove's contextual approaches to privacy protection are combined with Calo's categories of subjective and objective harms. From this analysis, I have developed a proposed definition of private data that will be tested on privacy and data protection experts, in order to review existing legislative measures and assess the adequacy of this definition as a reform measure. Below is the proposed definition:

> *Private data is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual*

# Chapter 5

# Research Design and Methodology

## *5.1 Introduction*

This chapter begins by briefly re-stating the research questions formulated in chapters two, three and four, and suggesting that empirical research should be conducted to test the concepts contained in data protection laws; in particular whether privacy continues to be valued in the information society, and the adequacy of the terms of personal and sensitive data, since Bennett and Raab remarked that:

> "[u]nfortunately, we have little systematic cross-national survey evidence about attitudes to privacy with which to investigate the nature and influence of wider cultural attributes. Much of th[e] argumentation tends, therefore, to invoke anecdotes or cultural stereotypes: 'the Englishman's home is his castle', and so on."[280]

It will also reiterate a contextual, harm-based definition of private data, outlined in chapter 2, which will be tested in subsequent chapters, and used to develop a legislative reform proposal. It then describes the research approach and the researcher's position, as well as models chosen for empirical data generation. An innovative mixed methods approach was designed including: quantitative data collection in the form of a telephone survey and also an online survey, whilst qualitative data was collected through a series of semi-structured interviews with key stakeholders, both in the UK and internationally. The various stages of data collection are described below, as well as the participant recruitment methods used, the sampling strategy employed and the types of questions asked. The chapter also reports responses to challenges that arose during the data collection phase (for

[280] Bennett, C.J. & Raab, C.D., (2003) *The Governance of Privacy. Policy instruments in global perspective*, (Ashgate, Aldershot), p. 15

instance, the necessity of conducting email and telephone instead of face-to-face interviews, to accommodate language and time constraints of the data protection and privacy experts), the innovative use of *Youtube* as a promotion tool for the survey of bloggers, and the response to encountering distressing information on blogs and in the responses of blog respondents) and reflects on necessary resultant deviations from the original research design.

## *5.2 Recap of Research Questions*

An analysis of the literature in preceding chapters indicated that that the conceptual adequacy of existing data protection laws has been too readily accepted. In particular, current legislation is based on a number of premises which should be tested through the collection and analysis of empirical data.

The first premise is that privacy has a continuing value in the information society. The data collected and analysed in subsequent chapters will explore whether privacy valued by individuals in the information society; in particular whether privacy is valued by bloggers, that is, individuals who act as information processors.

A second premise is that privacy protection can be achieved through regulation of the processing of personal data. This suggests that personal data may be synonymous with private data. Accordingly, this thesis will test whether personal data is synonymous with private data.

A third premise is that some types of data are considered 'special' or more privacy sensitive, in that, stricter conditions must be adhered to in order to process such data. This leads to two further research questions: one, whether the current categories of sensitive data reflect post World War II concerns, and are in need of review in order to assess their continuing relevance, and two, whether sensitive data may is synonymous with private data.

The final research question is, whether the proposed definitions of 'data privacy' and 'private data' which draw upon the harm based contextual integrity approach

advocated by Nissenbaum,[281] and Solove,[282] have merit as a legislative reform proposal.

Below, I will outline the methods employed to collect and analyse primary data in order to assess the adequacy of the assumptions which underpin existing privacy and data protection laws, and suggest legislative reform proposals.

## *5.3 Researcher's position*

I am aware that I have grown up in a society which traditionally values and seeks to protect privacy, and so, I am undoubtedly influenced by my cultural and personal experiences. Also, earlier in my career I worked as a Solicitor, so I was familiar with the concepts from a legal perspective. However, my legal training and practice taught me to conduct research and present arguments in a factual, dispassionate manner. Accordingly, this research benefits from the legal and technical expertise I have gained as a socio-legal researcher. This allowed me to conduct the research as a detached observer,[283] viewing my role as that of an independent gatherer and reporter of results.

## *5.4 Research methodology: A socio-legal approach*

Chapter 1 announced that this thesis is not solely an exercise in classic, black-letter doctrinal scholarship since the primary aim is not to analyse in minute detail the contents of data protection laws, provision by provision; rather the approach adopted in this thesis is socio-legal, that is, it studies, through the collection of primary empirical data, the concepts embedded in data protection laws, with a view to understanding how they operate and what effects they have from the perspective of both experts and potential data subjects. A justification of this approach was offered by Twining, when he stated that doctrinal analysis is sometimes inadequate because:

> "Black-letter analysis that remains fixated mainly on points of law as contested in appeal courts remains excessively abstract and distant from how law is actually used in society as a matter of empirical fact, and how it is experienced differently by members of various social groups."[284]

[281] Nissenbaum, H. (2004) Privacy as Contextual Integrity, *Washington Law Review*, pp.101-139
[282] Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4. pp. 1087-1155.
[283] Blaikie, N. (1993) *Approaches to Social Enquiry* (Cambridge: Polity Press), p.52
[284] Twining, W. (1985) *Karl Llewellyn and the Realist Movement*, (London: Weidenfeld) p.32

Thus, data protection concepts cannot be understood adequately when viewed only from the standpoint of a lawyer[285] arguing over the meaning and scope of points of law. On the contrary, they must be studied from the perspective of those whose actions are most likely to be affected by legal decision making as this represents law in action within a societal context. The rationale for this approach is further derived from the recognition by Salter and Mason that:

> "One defining feature of sociolegal studies....is the shared belief that the assumptions of black letter agenda provide a totally inadequate, or at least an insufficient basis for conducting a viable form of legal research. This is because its focus upon describing doctrinal rules and principle 'out of context' ignores the derivation, practical operation and social impact of specific legal measures."[286]

Accordingly, the approach of this thesis mirrors that of Harris who contends that:

> "Empirically, law is a component part of the wider social and political structure, is inextricably related to it in an infinite variety of ways, and can therefore only be properly understood if studied in that context."[287]

Further, Bradney *et al*[288] assert that to understand how the legal system works, it is necessary to make use of the materials and techniques of the social scientist[289] whilst Bradshaw[290] argues that socio-legal research requires the researcher to gather 'data wherever appropriate to the problem' using whatever methods are most likely to generate such data. Similarly, Campbell & Wiles contend that:

> "The problems raised by socio-legal studies are clearly suited to empirical research methods, and the traditional range of social surveys, questionnaires, formal interviews and standard quantitative techniques are widely employed."[291]

Thus, when analyzing the conceptual adequacy of the terms 'personal,' 'sensitive' and 'private' data, I was free to choose from the full spectrum of qualitative and quantitative data collection methods, constrained only by factors such as time and

---

[285] Either academic or practitioner
[286] Salter, M. & Mason, J. (2007) *Writing Law Dissertations: An introduction and Guide to the Conduct of Legal Research*, (1st edn) (Pearson Longman) p.129
[287] Harris, P. (1986) "Curriculum Development in Legal Studies," *Law Teacher*, Vol. 20, No. 2, pp 110-123 at p.112
[288] Bradney, A., Cownie, F., Masson, J., Neal, A. & Newell, D. (2000) "Law in action/law in books*,"* 'in *How to study the law* (4th edn) (London: Sweet & Maxwell), p. 20
[289] Bradney, A. *et al*, (2000) "Law in action/law in books," in: *How to study the law* (London: Sweet & Maxwell), p. 20
[290] Bradshaw, A. (1997) "Sense and Sensibility: Debates and Development in Socio-Legal Research Methods," in Thomas, P. (1997) (ed) *Socio-Legal Studies*, (Aldershot: Ashgate-Dartmouth), p.99
[291] Campbell, C. & Wiles, P. (1976) "The Study of Law in Society in Britain," *Law & Society Review*, Vol. 10, pp. 547- 578, at p.547.

money. As discussed in Chapter 2, privacy is a complex high-level phenomenon, which is currently not well understood, so a traditional scientific approach using only quantitative measurements and calculations would be inadequate, as although it could lead to generalisable results, they would lack a richly descriptive interpretation. The concepts of privacy and private data require an explanatory research approach, and so imply that a qualitative approach also be utilised. Also, in addition to seeking the views of data protection and privacy experts, it was important to ascertain if the legal definitions accorded with the views of the public, who often play the role of data subject, as government legislative initiatives are intended to give effect to the legal requirements of a society, and will only be successful if they are valued and supported by the public.

## *5.5 Research design*

The research employed socio-legal research methods which involved doctrinal analysis as well as empirical data collection, as outlined below.

### *5.5.1 Legal analysis*

In a classic law thesis, doctrinal analysis would be the only method employed. However, in this socio-legal thesis, whilst analysis of the law is pervasive, it is accompanied by the collection and analysis of empirical data. An initial literature review (see chapters two, three and four) analysed the adequacy of existing data protection laws and identified their inherent problems, allowing the research questions to be framed. Moreover, both legislation and case law were analysed in the results chapters (see chapters five, six, seven and and) of this thesis in order to situate the empirical data within the critique of the existing approach to privacy and data protection and to develop the arguments for reform. Further, a comparative analysis of other legislative approaches e.g. the APEC privacy framework was conducted in order to explore potential reform proposals, including the testing of a harm-based approach to data protection. Nevertheless, the focus of this chapter is the empirical data collection methods employed to collect primary data in this research study.

### *5.5.2 Empirical data collection: Mixed Methods*

The socio-legal approach employed involves a mixed-method design, that is, one in which both quantitative and qualitative methods are used to answer research

questions in a single study. As Denzin[292] noted, no single method could never adequately solve the problem of rival causal factors and each method reveals different aspects of empirical reality. In contrast, a mixed methods approach is expected to create reliable explanation through triangulation, that is, the use of more than one approach to the investigation of a research question in order to enhance confidence in the ensuing findings.[293] Mixed methods have particular value when a researcher is trying to solve a problem that is present in a complex social context.[294] Morse described the advantages of mixed methods:

> "by combining and increasing the number of research strategies used within a particular project, we are able to broaden the dimensions and hence the scope of our project. By using more than one method within a research study, we are able to obtain a more complete picture of human behaviour and experience. Thus we are better able to hasten our understanding and achieve our research goals more quickly."[295]

Qualitative and quantitative data collection can occur in parallel form or sequential form. The research design in this study is a combination of parallel mixed-methods design and a pragmatic sequential mixed model. A parallel mixed-methods model is one in which qualitative and quantitative data are collected and analysed to answer a single study's research questions. The final inferences are based on both data analysis results. The two types of data are collected independently at the same time, or with a time lag.[296] In contrast, in the pragmatic sequential mixed-methods/models design, one type of data (e.g. qualitative) provides a basis for the collection of another type of data (e.g. quantitative). It answers one type of question by collecting and analysing two types of data. Inferences are based on the analysis of both types of data.[297] Also, a sequential mixed-models design is one in which the conclusions that are made on the basis of the first strand lead to formulation of questions, data collection and data analysis for the next strand. The final inferences are based on the results of both strands of the study. In some cases the second strand/phase of the

---

[292] Denzin, N. K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods,* (New York: McGraw-Hill).
[293] Bryman, A. in Lewis-Beck, M.S., Bryman, A. & Liao, T. F. (2004) *The Sage Encyclopaedia of Social Science Research Methods,* (Sage) p.1142
[294] Leech, N. L, & Onwuegbuzie, A.J. (2009) "A typology of mixed methods research designs." *Quality and Quantity, Vol.* 43, pp. 265-275.
[295] Morse, J. (2002) "Principles of mixed and multi-methods research design" (pp. 189- 208) in Tashakkori, A. & Teddlie, C. (eds) *Handbook of mixed methods in social research and behavioural research,* (Thousand Oaks: Sage) p.189
[296] Onwuegbuzie, A. & Teddlie, C (2002) "A framework for analysing data in mixed methods research," (pp.351-384) in Tashakkori, A. & Teddlie, C. (eds*) Handbook of mixed methods in social research and behavioural research,* (Thousand Oaks: Sage), p. 438
[297] Onwuegbuzie, A. & Teddlie, C. (2002) in Tashakkori, A. & Teddlie, C. (eds) *Handbook of mixed methods in social research and behavioural research,* (Thousand Oaks: Sage), p. 439

study is used to confirm or disprove the inferences from the first strand or to provide further explanation for unexpected findings in the first strand. This approach can be used to generate hypotheses to be explored in more depth, or to develop surveys that use correct language for the population. In this study a mix of both sequential and parallel data collection methods was used to collect data in order to answer the research questions. A sequential data collection model was employed when the data collected from the ICO placement interviews were analysed and used to inform the questions asked during the semi-structured interviews with the privacy and data protection experts, and also when the data collected during the expert interviews was used to develop a question for insertion on the nationally representative telephone survey. A parallel data collection model was employed when a question on data sensitivity was asked in both the telephone survey and the online survey, so that the responses from the two different survey samples could be compared. The individual data collection methods, which comprised four key components, are illustrated by Table. 5.1 below.

Table 5.1 *Summary Table of Empirical Data Collection Methods*

| Method | Population | Number of Respondents | Data collection period |
|---|---|---|---|
| Semi-structured Interviews | ICO personnel | 8 | 7th – 11th March, 2005 |
| Semi-structured Interviews | Privacy & Data protection Experts | 40 | April – August 2006 (* piloted on 10 ICO personnel, academics and lawyers Sept-Dec 2005) |
| Telephone Survey | UK citizens | 1066 (*nationally representative of landline telephone owners) | April – May 2006 (*piloted on 10 lawyers and 10 non-legal participants Jan-Feb 2006) |
| Online survey | Bloggers from around the world | 1258 | (25th Sept - 20th Nov 2006) (*piloted (15-20th Sept, 2006) on 10 bloggers/blog readers) |

## *5.6. Component 1 - Policy & Practice case study*

A one-week placement (7th – 11th March, 2005) at the Office of the UK Information Commissioner was used to gain information on how the Data Protection Act 1998 is operationalised in the UK. During the placement eight interviews were conducted

with key personnel (See Appendix A for questions). From these interviews an insight was gained into the key issues facing the Information Commissioner in interpreting the provisions of the Data Protection Act 1998. I also gained an insight into the compliance and enforcement roles of the Office of the Information Commissioner, and how they seek the co-operation of data controllers in a proactive, educational manner. The responses to the interview questions were written up and used to inform subsequent information gathering and research.

## *5.7 Component 2 - Interviews with privacy and data protection experts*

The issues raised during the placement with the ICO formed the basis of the questions developed for use in semi-structured interviews with privacy and data protection experts. During the planning phase, consideration was given to several possible data collection methods and their associated benefits and limitations, including a quantitative survey of data protection experts or qualitative interviews. Whilst a quantitative survey would offer the advantage of structured responses, thereby maximising the reliability of measurement of key concepts, this data collection method was rejected in favour of a qualitative approach. Given the novelty of the issues being discussed, and the linguistic and interpretative difficulties[298] associated with discussing concepts, the flexibility of qualitative interviews outweighed the resultant limitations on statistical analysis. The decision to use qualitative interviews as a data collection method was influenced by Ely *et al* who state that:

> "qualitative researchers want those who are studied to speak for themselves, to provide their perspectives in words and other actions."[299]

The rationale for conducting interviews with experts was that any attempt to understand the conceptual underpinnings of private data necessitated an exploration of the views and interpretations of the views of those charged with operationalising it in the course of their employment as data controllers and those charged with interpreting it on behalf of the general public.

---

[298] English was not the first language of all interviewees. Also, when words are used in legislation, they often have a precise legal meaning which differs from their meaning in spoken language e.g. the phrase 'reasonable person' implies an objective test or standard of behaviour in common law jurisdictions, but this concept is not similarly understood or utilised in civil law jurisdictions.

[299] Ely, M., Anzul, M., Freidman, T., Garner, D., & McCormack-Steinmetz, A. (1991) *Doing Qualitative research: circles within circles*, (London, England: Falmer Press), p. 4

### *5.7.1 Interview methods*

Bryman[300] identified several different interview methods, namely: standardised (structured), unstandardised (informal) and semi-standardised (semi-structured) interviews. Unstandardised interviews were immediately discounted as a data collection method due to the potential for such interviews to become unfocused and lack critical analysis or consistency. A key motivation for deciding to conduct semi-structured interviews, as opposed to structured interviews, was the degree of response freedom afforded to respondents; allowing them to explain their thoughts and expertise in depth, and in particular to draw out and discuss the legislative complexities and contradictions they have experienced. Of the forty interviews conducted, sixteen were conducted face-to-face, sixteen via telephone and eight via email.[301] Qualitative researchers generally advocate face-to-face interviewing when conducting semi-structured and in-depth interviews, as it allows a rapport to build between the interviewer and interviewee, and allows the interviewer to flexibly alter the questions on the interview guide in response to visual clues offered by the respondent. The original research design called for all face-to-face interviews, and in this study I (the researcher) personally conducted sixteen face-to-face interviews and recorded the responses through contemporaneous note-taking, after seeking consent from the respondents. Each interview lasted between 30 minutes and one hour. (See Appendix B for the questions). However, McCracken has noted that participation in qualitative interviewing can be:

> "time consuming, privacy endangering, and intellectually and emotionally demanding."[302]

Therefore, Sturges & Hanrahan[303] advise researchers to do whatever is possible to maximize data quality while minimizing imposition on respondents. Indeed, due to time, language and geographical differences and interviewee availability clashes, it became necessary to conduct sixteen of the interviews by telephone and eight via

---

[300] Bryman, A. (2004) Chapter 15: "Interviewing in qualitative research," *Social Research Methods* (2nd edition), (Oxford University Press, Oxford); p. 314

[301] Time differences and transport costs made it impractical to interview each respondent on a face-to-face basis. Also, a few respondents who were not native English speakers requested that interviews be conducted through a series of email exchanges, so that they could reflect and formulate their responses.

[302] Mc Cracken, G. (1988), *The Long Interview. Qualitative Research Methods,* (Newbury Park, Sage). p. 27

[303] Sturges, J.E. & Hanrahan, K. J. (2004) "Telephone and face-to-face interviewing," *Qualitative Research,* 4 (1) 107- 118, p.109

email. Creswell[304] acknowledges that use of a telephone to conduct an interview is potentially problematic because it deprives the researcher of visual clues, that is, the opportunity to observe the respondents' informal, non-verbal communication during the interview. Nevertheless, he asserts that it is possible to conduct interviews by telephone when the researcher does not otherwise have access to the respondent. Moreover, Sturgess and Hanrahan[305] conducted research by collecting data though a mix of telephone and face-to-face interviews. They reported no significant differences in the quality or quantity of interview data collected using either method. Similarly, in relation to email interviews, Selwyn & Robson found that:

> "Using e-mail as an interview tool eschews the conventional constraints of spatial and temporal proximity between interviewer and respondent and offers the considerable practical advantage of providing 'ready-transcribed' data. However, e-mail interviews suffer from a lack of tacit communication."[306]

Thus, in those instances where it was not possible to conduct a face-to-face interview, potential respondents were offered the opportunity to complete the interview via telephone or email, as this allowed the researcher to gain responses from hard to access respondents.

### *5.7.2 Advantages of the interviews*

Bryman[307] notes that a semi-structured interview is conducted by the interviewer asking questions from an interview schedule and recording the respondent's responses. Accordingly, in this study, the interview schedule questions were piloted (Sept - Dec 2005) on personnel from the UK Information Commissioner's Office and personnel from CCSR,[308] as well as on lawyers and academics[309] so that ambiguous and poorly worded questions could be identified and improved.[310] This allowed an interview guide (see Appendix B) to be prepared before-hand, and used as an *aide memoir* in interviews to ensure that all relevant topics were covered. As stated in Chapter two of this thesis, some academics assert that current data protection laws

---

[304] Creswell, J.W. (1998) *Qualitative Inquiry and Research Design: Choosing among five traditions*, (Thousand Oaks: Sage) p.86

[305] Sturges, J.E. & Hanrahan, K. J. (2004) "Telephone and face-to-face interviewing," *Qualitative Research*, 4 (1) 107- 118, p.108

[306] Selwyn, N. & Robson, K. (1998), "Using e-mail as a research tool," *Social Research Update*, Issue 21 <http://sru.soc.surrey.ac.uk/SRU21.html>, p.1

[307] Bryman, A. (2004) Chapter 15: "Interviewing in qualitative research," *Social Research Methods* (2nd edition), (Oxford University Press, Oxford); p. 314

[308] Staff and colleagues at CCSR provided both informal and formal feedback at various stages of questionnaire piloting.

[309] The lawyers were personal acquaintances of the researcher, as were the academics, from diverse disciplines including computing, marketing, politics and pharmacy.

[310] See Appendix **B** for a full version of the Interview Guide

are inadequate and in need of reform. Accordingly, it was appropriate to ask the experts questions which probed their satisfaction with the current legislative provisions, in particular, the concepts of personal and sensitive data. Also, a second purpose of the interviews was to test the merits of a harm based approach to privacy and data protection and collect responses to a proposed definition of private data. Therefore, the proposed definition of private data was sent to all the interviewees in advance of the interview.[311] The overall aim of the interview process was to have a discussion with the respondent so that all the themes mentioned in the interview guide were covered. Thus, a key feature of the interviews was that the structure of the interviews was not fixed, and indeed, the flexible structure of the interviews allowed deviation according to the respondent's area of expertise, or in response to questions which arose naturally in the flow of the interview on the basis of respondent's responses.

Also, the interview questions comprised both closed and open-ended questions. The open-ended questions were important for the experts to express their views and experiences as freely as possible. Indeed, a key reason for using semi-structured interviews with elite respondents is that:

> "the investigator is willing and often eager to let the interviewee teach him what the problem, the question, the situation is – to the limits, of course, of the interviewer's ability to perceive relationships to his basic problem, whatever these may be..."[312]

Probing questions were also utilised to elicit further information from the respondents, particularly when it was felt that further elaboration was necessary. For instance, statisticians would initially hesitate in discussing the deficiencies of existing laws with the interviewer because they were aware that the interviewer was a qualified lawyer. However, when it was made clear that the interviewer would particularly welcome the responses they could offer based on their experience of interpreting and applying the legislation in the context of statistical analysis, then they offered useful insights into the challenges and interpretation difficulties experienced by non-lawyers in interpreting and complying with the legislative measures. Overall, the degree of flexibility afforded the by the semi-structured

[311] The definition is lengthy and a considered response could only be given if the interviewees had the opportunity to read and reflect on the details.
[312] Dexter, L. A. (1970) *Elite and Specialised Interviewing*, (Northwestern University Press) p.19; republished with an introduction by Ware, A. & Sánchez-Jankowski, M., (2006) European Consortium for Political Research.

approach was beneficial because it allowed the interviewer to elicit rich, thick descriptive responses from the experts.

### *5.7.3 Sample selection*

In order to obtain a range of responses from individuals with a variety of disciplinary backgrounds, a respondent matrix was created using quota and snowball sampling.[313] This is an approach supported by Denscombe who asserts that snowball sampling is an effective technique for building up a reasonable sized sample, especially when used as part of a small-scale research project.[314] Although this method did not produce a representative sample, it was an appropriate strategy to use when seeking to interview elite[315] respondents, who, because of their busy work schedules, are often hard to access by researchers. Indeed, the UK Information Commissioner was part of a network of professionals who meet regularly,[316] and they provided introductions to other specialists whom they interact with, as they were keen to include wider societal and business viewpoints and experiences in the debate, even when such respondents did not share similar viewpoints. During the period April 2006 – August 2006/7, forty, semi-structured interviews were conducted with privacy and data protection experts, namely: privacy commissioners, lawyers, corporate privacy officers, consultants, computer scientists, and academics from sociology, politics, market research, statistics and law, from several continents and countries, namely: Australia, Belgium, Canada, Czech Republic, Finland, France, Germany, Iceland, India, Ireland, Italy, Netherlands, New Zealand, Poland, Spain, UK and the USA.

### *5.7.4 Limitations of the interviews*

Those opposed to the use of interviews for data collection could argue that they may be criticised for three main reasons. Firstly, that they are potentially 'unreliable;' critics could question whether interviews with the same set of respondents over and over again (by the same or a different researcher) would yield the same or similar responses. Secondly, they could argue that interviews do not produce generalisable results because the sample was not random and only a small number of interviews

---

[313] Denscombe, M. (1998) *The Good Research Guide for Small Scale Social Research Projects* (Open University Press: Milton Keynes, UK) p. 18

[314] Denscombe, M. (1998) pp. 18-19

[315] Manheim & Rich assert that "Their elite status depends not on their role in society but on their access to information that can help answer a given research question" Manheim, J. B. & Rich, R. C. (1986) *Empirical Political Analysis* (New York: Longman) p.132

[316] e.g. at their Annual Conference <http://www.mfa.gov.il/MFA/Government/Communiques/2009/Israel_host_2010_international_conference_data_protection_8-Dec-2009.htm>

were conducted with individuals who could have anomalous views or experiences that are not normatively representative. Thirdly, critics could complain that interviews require the reader to place too much 'faith' in the interviewer, by unquestioningly accepting the researcher's claim that they have followed prescribed techniques, have asked the 'right' questions and have accurately recorded responses.

Nevertheless, Epstein offers a rebuttal of these criticisms by contending that 'if proper procedures are followed, interviews are just as reliable as other forms of data collection'[317] and that reliability is a necessary but insufficient indicator of good operational measures, and further that, if a researcher follows the prescribed guidelines for conducting semi-structured interviews then, arguably the reader can have faith in the results as this form of data collection is as systematic as other forms of data collection

Overall, this data collection method was appropriate as Kerlinger astutely observed that an interview 'can be an exploratory device to help identify variables and relations, and to guide other phases of the research.'[318] Thus, in the context of this research, the responses of experts regarding current and potential new categories of sensitive data were used to form variables in the telephone and blog surveys. Additionally, although the semi-structured interviews solicited responses that were more difficult to analyse, this data collection method was appropriate because it allowed the interviewer to elicit rich, thick descriptive responses from the experts. Indeed, it facilitated the collection of primary data in the form of unique and valuable responses from experts to the proposed definition of private data. Thus, this data collection method fulfilled a key objective of the research study namely to map out various viewpoints and probe key concepts, so that a fuller appreciation of the conceptual difficulties could be distilled.

## *5.8 Component 3 – Nationally representative Telephone survey of UK Citizens*

The Office of the UK Information Commissioner conducts an annual survey of UK citizens by telephone. This 'Annual Track (Individual)' survey measures awareness,

[317] Epstein, L. (1986) "Strategies of Judicial Research: Interviewing U.S. Supreme Court Justices and Interest Group Attorneys," Paper Presented at Annual meeting of Southern Political Science Association, Atlanta, Georgia, <http://epstein.law.northwestern.edu/research/conferencepapers.1986SPSA.pdf>, p.16

[318] Kerlinger, F. N. (1973) *Foundations of Behavioural Research*, (New York: Holt, Reinhart & Winston) p.480

understanding, relevance and perceptions of the Data Protection Act by UK citizens. It also measures the importance of specific matters relating to personal information and their use of it.[319]

To support the collaborative PhD CASE partnership which they part sponsored, they agreed to help develop my research strategy by offering an opportunity to insert an additional question on their 'Annual Track (Individual)' survey of 2006. This offer was accepted because it afforded a unique opportunity to collect data via a national survey.

As stated earlier, the findings from semi-structured interviews with privacy and data protection experts led to the development of a question suitable for insertion on the quantitative survey. The experts had indicated that, in their opinion, some of the legally recognised categories are not considered at all sensitive by them (e.g. in Iceland trade union data is not considered at all sensitive since the country is so small that everyone knows where everyone lives and works), and also that in their opinion advances in technology and computing had given rise to new types of data which attracted high levels of privacy concern among data subjects.

Accordingly, a multi-part Likert scale question was constructed to test sensitivity perceptions of fifteen categories of data on a scale of one to ten, with one labelled not at all sensitive and ten extremely sensitive. (See Appendix C for questions). The question was piloted on ten lawyers[320] and ten non-legal participants.[321] Thereafter, the telephone survey was conducted by SMSR's, a research company employed by the ICO to conduct the survey using their in-house telephone interviewing team. This allowed a large number (1066) of telephone surveys to be conducted in a short time period (a 3 week period April – May 2006).

### *5.8.1 Advantages of telephone survey*

Telephone surveying has a number of advantages over other data collection methods. For example, when compared to face-to-face interviewing, Boland *et al* assert that a telephone survey allows a large, geographically dispersed sample to be easily reached with the result that interview travel time and associated costs are reduced as

[319] SMSR Ltd (2006) 'Report on Information Commissioner's Office Annual Track,'

[320] The term 'lawyer' is used in a generalist sense to refer to my personal contacts who work either as Solicitors, Barristers or Academics.

[321] These participants were friends are from a variety of countries including: UK, Ireland, Malaysia, Hong Kong, Spain, China, Jordan, Greece and Italy. They also worked in a variety of professions, including: Sociology, Pharmacy, Computing, Accounting, Dentistry, Philosophy, and Mathematics.

interviewers do not have to physically visit all respondents.[322] Moreover, Mc Givern asserts[323] that respondents surveyed via telephone perceive a greater degree of anonymity, and, as a result, are more forthcoming in answering questions of a sensitive nature. This advantage influenced the decision to insert a question on the survey testing the sensitivity rating of different data types. The telephone survey question on sensitivity of data types was designed to answer two research questions. Firstly, it was used to test attitudes of the UK public towards the seven legally recognised categories of sensitive data, as the interviews with experts revealed that not all of the legally recognised categories of sensitive data are considered to be sensitive. Further, during the course of interviews with privacy and data protection experts, eight potential new categories of sensitive data emerged. Thus, the second purpose of the telephone survey was to test sensitivity perceptions of potentially new categories of sensitive data.

### *5.8.2 Limitations of telephone survey*

Those opposed to the use of interviews for data collection report that:

> "While resulting in a higher response rate than postal surveys, telephone surveys often attract a higher level of refusals than face-to-face interviews as people feel less inhibited about refusing to take part when approached over the telephone."[324]

Accordingly, to compensate for this potential limitation, quotas were set by the team employed to conduct the survey, with England divided into nine regions, whilst national samples were collected from Scotland, Wales and Northern Ireland, in order to survey a representative sample of UK citizens.[325] Also, quotas were set on age, sex and social grade to ensure a nationally representative sample was achieved.

However, subject contact data was accessed using the Names and Numbers ADF software. Thus, the survey cannot claim to be nationally representative of all households in the UK as some households are ex-directory, or do not have a landline.

[322] Consequently, concerns about interviewer safety are reduced.
[323] McGivern, Y. (2003) *The Practice of Market and Social Research: An Introduction,* (Pearson Education, Harlow), p.256.
[324] Kelley, K., Clark, B., Brown, V. & Sitzia, J. (2003) "Methodology Matters: Good Practice in the conduct and reporting of survey research," *International Journal for Quality in Health Care*, Vol. 15 (3) pp 261-266
[325] A booster sample of 100 interviews in Northern Ireland was collected so that this cell was large enough to analyse without compromising the representative main sample.

For instance, in 2006, 87% of UK households had a fixed line telephone.[326] Importantly, 12% of households in the UK now rely on mobile telephony, that is, they do not have landlines installed. This is significant, as mobile phone numbers are not available in directory format, and the increasing number of households who rely solely on mobile phones reduces the representativeness of any survey conducted using a simple random sample of phone numbers in a directory. Nevertheless, the sample of 1066 interviews conducted is nationally representative of households in the UK that have a landline and have permitted their landline telephone number to be listed in phone directories, and so, notwithstanding the above limitations the insertion of this question into the ICO survey was a significant value for this research project.

## *5.9 Component 4 – Global Online survey of Bloggers*

As stated in chapters one and three of this thesis, some academics[327] assert that existing data protection laws are in need of reform because they do not address the challenges posed by advances in computing and technology. Indeed, one of the key criticisms is that the Directive was drafted in a pre-internet era, and is not fit for purpose as advances such as the development of web 2.0 social media, for instance, online diaries or blogs have given internet users the ability to participate in exchanges with other individuals and concomitantly diminished their desire for informational privacy. In particular, the act of keeping an online diary, or blog, detailing personal thoughts and experiences appears oxymoronic, since the act of keeping a diary was traditionally considered a private act; diaries were historically often kept under lock and key and therefore not available for public consumption. Indeed, as Serfaty observes:

> "In social representations, diaries are first and foremost intimate writings and making them available online therefore appears to raise intractable privacy issues: diaries are believed to be basically private documents that should never get public exposure."[328]

Accordingly, this research sought to investigate whether privacy is still valued in the Internet era by examining whether bloggers take any action, for instance, by limiting the subject matter on which they post, or by restricting who can read their posts, in order to protect their privacy. Only by answering such questions would this study be

[326] Ofcom (2008) Nations and Regions CMR UK summary <http://www.ofcom.org.uk/research/cm/cmrnr08/uksummary.pdf>
[327] Tene, O. (2011) "Privacy: the new generations," *International Data Privacy Law*, Vol. 1, No. 1, p. 15
[328] Serfaty, V. (2004) *The mirror and the veil: An overview of American online diaries* (Rodopi: New York), p.1

able to determine whether the provisions of the Data Protection Directive (drafted in a pre-Internet era), are still fit for purpose, or in need of reform. To meet this aim, an online survey of bloggers was conducted.

The blog survey was piloted (15-20th September, 2006) on ten respondents who were either bloggers themselves or regular readers of blogs.[329] Thereafter, the survey was conducted (25th September - 20th November 2006). The question on sensitivity of data types included in the telephone survey was replicated in the online survey. Additionally, participants answered questions about their blogging practices and their expectations of privacy and accountability when publishing online. They also answered questions about whether they had experienced any problems due to materials posted on their blogs. The survey focused on four key areas: *Blogging practices*, *Privacy Expectations, Blog content and Privacy Attitudes* and *Questions about other people's privacy*. (See Appendix D for survey questions)

### *5.9.1 Advantages of Online survey*

This method of data collection was chosen because Van Selm & Jankowski[330] assert that conducting online surveys is appropriate where the objective is to reach a population with internet usage experience. Similarly, Sills & Song[331] claim that, for particular populations that are 'connected and technologically savvy,' the low cost, ease and speed of delivery and response, ease of data cleaning and analysis weigh in favour of the internet as a delivery and collection method for survey research. Medlin et al[332] identified a number of other advantages offered by web survey programs, including the ability to: check for non-completion of questions, require completion of all questions before allowing respondents to proceed, and automatically control for branching according to respondent answers.

Vehovar and Manfreda[333] elaborated that self-administered surveys are advantageous both for the researcher and the respondent. For instance, respondents can complete the survey questionnaire at a time, place and pace of their convenience. Also, self-

[329] They provided feedback on survey layout, colours, question navigation and download speeds.
[330] Van Selm, M. & Jankowski, N. W. (2006) "Conducting Online Surveys," *Quantity & Quality*, 40: 435-456, p.436
[331] Sills, S. J. & Song, C. (2002) "Innovations in Survey Research: An application of web-based surveys," *Social Science Computer Review*, 20 (1) 22-30, p. 28
[332] Medlin, C., Roy, S. & Ham Chai, T. "World Wide Web Versus Mail Surveys: A Comparison And Report," Paper presentation at ANZMAC99 Conference, Marketing in the Third Millennium, Sydney, Australia, <http://smib.vuw.ac.nz:8081/www/ANZMAC1999/Site/M/Medlin.pdf> p.3
[333] Vehovar, V. & Manfreda, K.L. (2008) "Overview: Online Surveys," in Fielding, N. *et al The Sage Handbook of Online Research Methods*, (Sage: London), p. 179

completion facilitates an increased sense of privacy which may improve the accuracy of responses. Equally, self-administration reduces the costs of administering the survey, and may reduce interviewer-respondent related biases, thereby leading to an improvement in data quality. From the researcher's perspective this method is advantageous because answers collected from respondents are immediately stored in a computer database ready for analysis. A particular advantage is the automatic coding of closed questions by the computer, leaving only open-ended questions to be manually coded. This reduces time, costs and the errors associated with traditional surveys during the data entry phase. Also, responses are typically received much faster than with mail surveys. For instance, Van Selm & Jankowski[334] reported that in a web survey of a women's magazine, nearly 2500 responses were received within two weeks of posting the questionnaire behind a banner of the magazine's website.

### *5.9.2 Online survey design*

Recommendations identified in the literature review were implemented when designing the survey, namely:

1. a plain questionnaire style as Dillman *et al*[335] discovered that a plain questionnaire design provided a better response rate, in terms of total number of surveys, and individual question completion than more elaborate questionnaire design.
2. the survey was displayed using a series of screens with a progress bar at the top, as Couper *et al*[336] reported that the presence of a progress indicator reduces respondent loss and that the use of multiple screens to display the questionnaire generates faster completion times and reduces item non-response.
3. radio buttons were used as Couper *et al*[337] found that these reduced item non-response,
4. text boxes were included to allow respondents to explain or elaborate on their responses, as academics had reported that this increased the quality of responses.

---

[334] Van Selm, M. & Jankowski, N. W. (2006) "Conducting Online Surveys," *Quantity & Quality*, 40: 435-456, p.499

[335] Dillman, D. A., Tortora, R.D., Conradt, J. & Bowker, D. (1998) "Influence of plain vs. fancy design on response rates for web surveys," Paper presented at Joint Statistical Meetings, Dallas, Texas <http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf>, p.4

[336] Couper, M. P., Traugott, M., & Lamias, M. (2001) "Web surveys: Perception of burden," *Social Science Computer Review*, 19 (2), pp. 146-162

[337] Couper, M. P., Traugott, M., & Lamias, M. (2001) "Web surveys: Perception of burden," *Social Science Computer Review*, 19 (2), pp. 146-162

Thus, the blog survey consisted of 4 Likert-scale questions (to measure attitudinal questions), 15 closed questions (with radio buttons) and 17 open-ended questions which allowed respondents to give expansive answers, as Sheehan & McMillan[338] reported that respondents appeared to be more willing to reply to open-ended questions in an online format than in traditional paper surveys.

### *5.9.3 Sample recruitment*

A key issue was how to invite bloggers to participate in the survey. Kay and Johnson claim that:

> "by its nature, the internet poses a unique set of problems in guaranteeing a random sample of respondents."[339]

They argue that since internet usage is not universal among the general population, obtaining a random sample through directories of postal addresses or telephone numbers would generate a large number of respondents who do not have access to the internet. It is also problematic because the absence of a central registration directory of internet users means that it is impossible to construct a sampling frame for email addresses or urls of blog sites. To deal with this, the respondents to this survey were not randomly selected but were selected through a range of snowball sampling strategies:

1. Announcements for the online survey were posted to mailing lists of three universities in the UK.[340]

2. the online survey was forwarded to a number of bloggers whom the researcher knows personally and it was also distributed by posting invitations (with a URL link to the survey) on a number of blogs which were known to attract a large volume of readers,[341] and on those blogs which had drawn media attention in recent months.

[338] Sheehan, K. B., & McMillan, S. J., (1999) "Response variation in e-mail surveys: An exploration," *Journal of Advertising Research*, 39 (4), pp. 45-54.
[339] Kay, B. & Johnson, T.J. (1999) "Research methodology: taming the cyber frontier. Techniques for improving online surveys," *Social Science Computer Review*, 17 (3): 323-337, p. 325
[340] Details of the blog survey were forwarded to students at Manchester University, Manchester Metropolitan University and to IT Law students at Queen's University, Belfast.
[341] Identified through Blog aggregators e.g. Technorati (Blog aggregators are websites that collect, organize, and publish links to large numbers of blogs by using web crawlers to locate blogs or by receiving pings from blogs. Powered by mathematical algorithms, they can calculate and rank a blog's

3. the survey was drawn to the attention of vloggers[342] on *YouTube*, some of whom had been involved in a debate on privacy and anonymity following the British media's attempt to uncover the identity of a popular vlogger, "Geriatric 1927."[343] The viral nature of blogs meant that links to the survey quickly spread to many other blogs generating a non-random sample.

Although there was no way of measuring the response rate to this online survey, due to the snowball distribution methods employed, the number of responses received was high. Of the 1314 responses received, 1258 (95.7%) were selected for analysis. The other 56 (4.3%) were disregarded as they were either incomplete or duplicate responses. A two-pronged approach was used to detect and filter out duplicate questionnaire completion by a single respondent. Firstly, the survey software recorded the IP address[344] of each respondent, and secondly, where an IP number completed a questionnaire within 2 minutes,[345] the multiple responses were manually checked. On its own, recording of the IP address would not be sufficient to identify and delete true duplicates, or multiple single respondents, as some computers e.g. those in internet cafes and libraries often have a multi-user function, hence different users may legitimately have the same IP address. The two-pronged approach filtered out responses where a respondent hit the send button twice or more in quick succession e.g. because there was a server delay in displaying the completion 'Thank you' message.

### *5.9.4 Distressing disclosures: ethical obligations*

During the survey distribution and data analysis phases occasional distressing disclosures were encountered. Stern defines a distressing disclosure as:

popularity or authority. Blog aggregator services also monitor blogging activity, displaying a dynamic list of recently updated blogs).

[342] Vlogger is the term used to refer to a blogger who communicate via video or audio posts on internet sites such as YouTube.

[343] Geriatric 1927 "Telling it All 7," <http://www.youtube.com/watch?v=muSZuSj6w3Y>; A response to Geriatric 1927's privacy concerns <http://www.youtube.com/watch?v=muSZuSj6w3Y>

[344] An IP address is a unique ID number by which a computer connected to the internet can be identified.

[345] The survey could not physically be completed in this time period – as it would take an average internet user 15mins to complete, and an experienced, computer literate user 10mins to complete the survey.

> "information that indicates an online communicant is considering harming him/herself or another/others (e.g. online users' announcements of suicide intentions, threats to kill another person, etc.)"[346]

However, I followed the advice of Stern who cautioned that:

> "researchers would wisely remember that behind every online communication if a real, living, breathing person."[347]

Accordingly, where distressing self disclosures were observed on blogs, or made through the online survey, I endeavoured to check that the postings were recent, and appeared to be authentic, based on the other contextual information on the blog e.g. the tone of other posts. In making such assessments I made case-by-case decisions, fully aware that I was not in a position to verify the facts of the post, or the true identity of the author. Where the distressing disclosure appeared recent and authentic, I directed the respondent to an appropriate agency or source of information e.g. Samaritans, or gave advice to contact the police, or a lawyer where the respondent was the alleged victim of cyberstalking. Further, although romantic proposals made by respondents could not be classified as 'distressing disclosures,' they were politely, but firmly rejected, on the basis of ethical and personal safety concerns!

### *5.9.5 Limitations of online survey*

As stated above, snowball sampling is a non-probability sampling method. One disadvantage of such a sample method is that it can lead to sample selection bias which is beyond the researcher's control. A particular risk is that respondents who opt to complete the survey may not be representative of the general population. For instance, lack of equal access to the internet could result in significant selection and non-response biases. Further, it is impossible to calculate non response rate as:

> "There is no way in which to know how many individuals might have seen the survey or its link but declined to participate. Only the number of completed surveys is known and not the number of refusals."[348]

Consequently, even though a range of measures were taken to encourage high response rates, and to filter out duplicate responses, the sample collected could nevertheless, suffer from some limitations. For instance, in common with traditional

[346] Stern, S. R. (2003) "Encountering distressing information in online research: a consideration of legal and ethical responsibilities," *New Media & Society*, Vol. 5, No.2, p.249
[347] *Ibid*, p.257
[348] Kay, B. & Johnson, T.J. (1999) "Research methodology: taming the cyber frontier. Techniques for improving online surveys," *Social Science Computer Review*, Vol.17, No.3, pp. 323-337, at p. 326

methods of survey e.g. mail or telephone, participants could decline to respond completely, withdraw at any stage during the survey, or selectively answer questions. Also, in this survey respondents were asked to anonymously self-report on their blogging practices and their privacy attitudes and expectations. This self-disclosure approach has two important implications: firstly, there could be disparities between stated privacy attitudes and actions, and secondly, respondents' perceptions of their blogs might differ from those of outside observers and researchers. It is well documented that people's perceptions of their own behaviour can differ from how they actually behave.[349] However, as with any anonymous study, it was impossible to verify whether the responses were misrepresented or exaggerated in any way.
Moreover, Fricker states that:

> "Unrestricted self-selected surveys are a form of convenience sampling and, as such, the results cannot be generalised to a larger population."[350]

Nevertheless, Berson *et al*[351] opine that the fact that the results are not generalisable to a larger population does not automatically detract from the value of the research; indeed Plous reported that 'most studies on the representation of web-study participants suggest that, if anything, those populations are more representative of the public than samples from more traditional lab experiments using college students.' [352] Thus, whilst the results obtained in this research highlight the privacy attitudes and expectations of bloggers are not generalisable, they are valuable as an exploratory study which furthers understanding of blogging and aids in the development of future research.

## *5.10 Data analysis*

Given the multiple data collection methods employed in this research, it was appropriate to use a combination of qualitative and quantitative data analysis techniques namely: qualitative analysis of interview data and both exploratory statistical analysis and a multiple linear regression model of the survey data. Details are provided below.

---

[349] Whyte, W. H. (1990) *City: Rediscovering the Center.* (New York, N.Y.: Anchor.)
[350] Fricker, R.D.(2008) "Sampling Methods for Web and Email Surveys," in Fielding, N. *et al The Sage Handbook of Online Research Methods*, (Sage: London), p. 205
[351] Berson, I.R., Berson, M.J. & Ferron, J.M. (2002) "Emerging risks of violence in the digital age: lessons for educators from an online study of adolescent girls in the United States*," Meridian, A Middle School Computer Technologies Journal*, <http://www.ncsu.edu/meridian/sum2002/cyberviolence/cyberviolence.pdf>
[352] Plous, S. quoted Azar, B. (2000) "A Web of Research: They're fun, they're fast and they save money, but do Web experiments yield quality results?," *Monitor on Psychology,* Vol. 31, pp.42-47, at p.42

### *5.10.1 Qualitative data*

After transcription of the interview responses, all the interviews were collated according to question responses. The responses were then read and coded according to themes that emerged both from the investigator's prior theoretical understanding of the phenomenon under study (an *a priori* approach) and from the data (an inductive approach).[353] The *a priori* themes came from several sources, including professional definitions identified in the literature review phase. They also came from common-sense constructs, my research values, and personal experience with the subject matter.[354] However, most of the themes were induced from empirical data, in line with Dey's argument that:

> "Even with a fixed set of open-ended questions, one cannot anticipate all the themes that arise before analyzing the data."[355]

Thus, the analytic process for qualitative data was 'systematic and comprehensive, but not rigid.'[356] In concurrence with the recommendations of Lincoln & Guba[357] analysis of the interview data was stopped when no new themes or information emerged from analysis. Of course, this was a subjective decision, for as Dey notes:

> "there is no single set of categories [themes] waiting to be discovered. There are as many ways of 'seeing' the data as one can invent."[358]

The output from the analysis is a descriptive picture of the themes that emerged from the data which is rich in detail. A potential criticism of this qualitative approach is that it is not possible to claim that the findings are valid. However, in concurrence with Bernard it is argued that the validity of a concept depends on the utility of the device that measures it and the collective judgment of the scientific community that a construct and its measure are valid:

---

[353] Ryan, G. W. & Bernard, H. R. (2003) "Techniques to Identify Themes," *Field Methods*, Vol. 15, No.1, pp. 85–109, at p.88

[354] Bulmer, M. (1979) "Concepts in the analysis of qualitative data," *Sociological Review Vol.* 27, No. 4, pp. 651–77; Strauss, A. (1987) *Qualitative analysis for social scientists*, (Cambridge, UK: Cambridge University Press); Maxwell, J. (1996) *Qualitative research design: An interactive approach* (Thousand Oaks, CA: Sage).

[355] Dey, I. (1993) *Qualitative data analysis: A user-friendly guide for social scientists* ( London: Routledge Kegan Paul), pp.97–98

[356] Tesch, R. (1990) *Qualitative research: Analysis types and software tools* (New York: Falmer Press), p. 95

[357] Lincoln, Y. S. & Guba, E.G. (1985) *Naturalistic Inquiry* (Newbury Park, CA: Sage Publications).

[358] Dey, I. (1993) *Qualitative data analysis: A user-friendly guide for social scientists* ( London: Routledge Kegan Paul), pp.110-111

> "we are left to deal with the effects of our judgments, which is just as it should be. Valid measurement makes valid data, but validity itself depends on the collective opinion of researchers."[359]

Likewise, Denzin claims that rules for establishing a valid sample 'are only symbolic - they have no meaning other than that given by the community of scientists.'[360] Thus, the validity of these findings will be determined over time as they are digested, interpreted and debated by peers working in this field of research.

### *5.10.2 Quantitative data*

The telephone data was supplied as a dataset suitable for analysis using the statistical software package, SPSS. The online survey data consisted of a mixture of answers in both numerical and word format. The responses to open-ended questions were coded using the techniques that were employed to thematically code the expert interview responses. Thereafter, the themed responses were assigned numerical codes so that the word data could be converted into numerical format for analysis using SPSS. Several quantitative data analysis techniques were used to report the results of the telephone and online surveys. These included descriptive statistics and a regression model.

#### *5.10.2.1 Multiple linear regression model*

A multiple linear regression model was produced to test for any significant differences in the sensitivity perceptions between the blog survey respondents and telephone survey respondents. The regression analysis was run on a merged dataset generated by combining the ICO telephone survey respondents with the UK blog survey respondents, after 64 respondents under eighteen years of age were excluded from the blog survey, as all respondents in the telephone survey were over this age. The model examined the residual effects of being a respondent in the blog survey as opposed to being a respondent in the telephone survey, once socio-demographic (age, gender, relationship status, parenting status, employment status and earning status) factors had been accounted for. By examining the residual effects, comparisons could be made between the perceptions of blog respondents and telephone survey respondents regarding the sensitivity of different data types. This, in effect, allowed the assessment of the differences between bloggers and the general population.

---

[359] Bernard, H. R. (1994) *Research methods in anthropology: Qualitative and quantitative Approaches*, (2d ed.) (Walnut Creek, CA: Alta Mira) p. 43

[360] Denzin, N. K. (1970) *The research act: A theoretical introduction to sociological methods*. (Chicago: Aldine) p.106

## *5.11 Summary*

In summary, a mixed methods approach comprising qualitative and quantitative primary data collection methods was employed in this study in order to capture the complexity of the concepts investigated. Firstly, eight semi-structured interviews were conducted with key personnel during a case study placement at the Office of the UK Information Commissioner. The interview responses were used to refine the research questions. Secondly, forty semi-structured interviews were conducted with privacy and data protection experts from around the world. The interviews had two main purposes; firstly, to test satisfaction with current legislation and secondly, to test responses to a proposed definition of private data. The responses to questions on satisfaction with current legislation resulted in potential new categories of sensitive data being identified. The interview data collected was further used to generate and refine survey questions. Thirdly, a question on data sensitivity was prepared and included on the ICO national telephone survey. The question focused on capturing the attitude of UK citizens towards legally recognised categories of sensitive and potential new categories of sensitive data which emerged during the interviews with experts. Fourthly, an innovative online survey tested the privacy attitudes and expectations of bloggers when they posted information online i.e. whether they consciously decided not to post certain types of information or whether they took steps to limit who could access such information because of privacy concerns. Thereafter, the responses of UK bloggers were compared with the responses of UK telephone respondents, in order to determine whether the use of internet technology is changing privacy attitudes and expectations.

It is important to acknowledge the limitations of the data collection methods, i.e. that the findings from the small number of interviews conducted with experts and the findings from the blog survey cannot be generalised. Nevertheless, the data collected does make a valuable contribution to knowledge as it is the first study which has attempted to gather empirical data regarding the conceptual adequacy of the terms contained in existing legislative measures. Also, another key feature of the data is that it provides information on the perspectives of both data protection and privacy experts and potential data subjects, regarding existing concepts and the potential merits of a harm based approach to privacy protection.

# Chapter 6

## The continuing value of privacy

### *6.1 Introduction*

A study of privacy requires an understanding of the changing nature of technology, and the social world created by that technology. One internet user group in particular appears to face unique privacy challenges: *bloggers*. However, the emerging media of blogs has not been fully explored to date. Accordingly, this chapter begins by explaining the technological phenomenon known as blogging. It will demonstrate that blogs, by their very nature, raise a number of privacy issues, since they permeate most niches of social life, addressing a range of topics from scholarly[361] and political issues to family and children's daily lives.[362] This is important, as industry experts such as Zuckerberg, CEO of Facebook, claim that as a result of individuals acting as information producers and processors, social norms are changing:

> "in the last 5 or 6 years, blogging has taken off in a huge way and all these different services that have people sharing all this information. People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time."[363]

He claims that the 'age of privacy is over,' that is, individuals no longer value, or seek to protect, their personal privacy; rather they desire opportunities for maximum openness and disclosure. Yet, Bodil Lindqvist,[364] a Swedish church volunteer, was

[361] Glenn, D. (2003) "Scholars who blog," *The Chronicle of Higher Education*, Vol. 49, A14.
[362] Turnbull, G (2004) "The seven-year-old bloggers," <http://news.bbc.co.uk/2/hi/uk_news/magazine/3804773.stm> (Last accessed 16.03.07).
[363] Kirkpatrick. M. (2010) "The Age of Privacy is Over," ReadWriteWeb, The article reports an on stage between Mark Zuckerberg, CEO of Facebook and the founder of TechCruch , in which Zuckerberg claims that 'the social norms of privacy are changing.' <http://www.readwriteweb.com/archives/facebooks_zuckerberg_says_the_age_of_privacy_is_ov.php> (Last accessed 27.02.11)
[364] Judgment of the Court of 6 November 2003 in Case C-101/01 (Reference for a preliminary ruling from the Göta hovrätt): Bodil Lindqvist, OJ 2004 C7/3 [hereafter *Lindqvist Judgment*],

prosecuted for posting personal and sensitive information about others, on her personal website. Her actions mirror those of a typical blogger, so it is important to assess whether privacy is still valued and protected by individual bloggers, or as Bygrave ponders:

> "whether indications exist of an opposite development – i.e., increasing *acclimatisation* of people to situations in which they are required to divulge personal information and a concomitant adjustment of what they perceive as problematic for their privacy. Unfortunately, there seems to be little survey evidence addressing this point."[365]

In order to assess the continuing value of privacy, an online survey of bloggers from around the world was conducted. The survey explored bloggers' subjective sense of privacy by examining their blogging practices and their expectations of privacy when publishing online. The findings discussed below will indicate that whilst bloggers were not always cognisant of the fact that the Internet and blogs are public spaces, they were generally concerned about privacy, aware of the privacy risks posed by blogging, and actively employed mechanisms to try to protect privacy.

## *6.2 Blogs: An overview*

A blog is a frequently updated website consisting of personal observations, excerpts from other sources, etc., typically run by a single person, and usually with hyperlinks to other sites; an online journal or diary.[366] A fundamental difference between blogs and other web-based publishing sites such as personalised homepages is that, rather than substituting new materials for old ones, a blogger simply adds new posts, creating an ever-growing compilation of entries and an archive of previous posts. Accordingly, Walker has defined a blog as:

> "a frequently updated website consisting of dated entries arranged in reverse chronological order."[367]

Compilations of posts serve as context for readers of blogs. Thus, regular readers can get a sense of the identifying 'voice' or 'persona' behind the posts. Over time, a blog archive can become a detailed portrait of the blogger's interests and experiences. Hence, by their very nature, blogs raise a number of privacy issues as they are persistent and cumulative, resulting in large amounts of sometimes personal

[365] Bygrave. L. A. (2004) "Privacy Protection in a Global Context – A Comparative Overview," *Scandinavian Studies in Law*, Vol. 47, pp. 319–348, at pp. 330-31

[366] OED definition. A blog may alternatively be referred to as a weblog.

[367] Walker, J. (2003) Weblog. In *Definition for the Routledge Encyclopaedia of Narrative Theory*

information being broadcast across the Internet.[368] Blogging poses new opportunities for privacy violations to occur, as individuals discuss personal matters and provide opinions openly in a format that can be easily accessed by anyone with an internet connection and that is, furthermore, archived indefinitely. Indeed, accounts of bloggers hurting friends' feelings, being sued[369] or losing their jobs[370] because of materials published on their sites are becoming more frequent.[371] Additionally, the issue has come before the ECJ in the *Lindqvist* case,[372] in which the Court ruled that merely posting personal information about individuals on a website constitutes automatic processing of personal data within the meaning of the Directive, thereby triggering obligations under the EU's privacy protection regime. Accordingly, it is important to explore the implications of this decision before examining why, in the face of these risks, bloggers appear to have chosen to forego some of their, and others,' privacy.

## *6.3 The Lindqvist decision*

Mrs Bodil Lindqvist was a church maintenance worker and volunteer, who took a computer class, and created some web pages with a variety of information about herself, her husband, and other church volunteers without their permission.[373] The

[368] Viégas, F. B. (2005) "Bloggers' expectations of privacy and accountability: An initial survey" *Journal of Computer-Mediated Communication,* Vol. 10, No.3, article 12

[369] For instance, during the week 19- 25 Feb, 2007 bloggers featured on the BBC news website twice. Firstly, a Hollywood film studio launched legal action against www.perezhilton.com, a popular celebrity gossip blog site, for publishing topless photographs of the actress, Jennifer Aniston. The studio claimed the images on the site were stolen and illegally copied; BBC (2007) "Blogger sued over topless Aniston," (22nd Feb)
<http://news.bbc.co.uk/1/hi/entertainment/6385677.stm> (Last accessed: 01.03.09). Secondly, an Egyptian blogger was sentenced to four years' prison for insulting Islam and the President. BBC, (2007) "Egyptian blogger jailed for insult," (22nd Feb)
<http://news.bbc.co.uk/1/hi/world/middle_east/6385849.stm?ls>. Meanwhile, Amnesty International was continuing to collect signatures for their campaign for freedom of speech on the Internet. Smith, D. (2007) <http://observer.guardian.co.uk/amnesty/story/0,,2000545,00.html> (Last accessed: 01.03.09)

[370] Petite Anglaise was scheduled to attend a *prud'hommes* (French Industrial Tribunal) hearing to contest her employment dismissal. She was sacked for blogging about work, even though she did not mention the name of her employer; Petite Anglaise (2007) "Wrong footed,"
<http://news.bbc.co.uk/1/hi/world/middle_east/6385849.stm?ls> (Last accessed: 01.03.09).

[371] Bray, H. (2004) "Job blogs hold perils, opportunities," The Boston Globe, p1, in Viegas, F. (2005) "Bloggers expectations of privacy and accountability: an initial survey," *Journal of Computer- Mediated Communication*, Vol. 10, No. 3, No. 12. <http://jcmc.indiana.edu/vol10/issue3/viegas.html> (Last accessed 27.02.08); Lichtenstein, S. D., Darrow, J. (2006) "Employment Termination for Employee Blogging: Number One Tech Trend for 2005 and Beyond, or a Recipe for Getting Dooced?" *UCLA Journal of Law and Technology (JOLT),* Vol. *10*, No.2, pp. 1-49.

[372] (C-101/01) [2004] 1 C.M.L.R. 20
<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62001J0101:EN:HTML> (Last accessed: 01.03.09)

[373] Judgment of the Court of 6 November 2003 in Case C-101/01 (Reference for a preliminary ruling from the Göta hovrätt): Bodil Lindqvist, OJ 2004 C7/3 [hereafter *Lindqvist Judgment*], paras 12–14 (noting

web pages "included some full names, telephone numbers and references to hobbies and jobs held by her colleagues,"[374] and included information about "preparing to take Communion at the church"[375] as well as information about "one volunteer's foot injury.[376] When asked to remove the web pages she acquiesced, but the Swedish Data Protection Inspectorate charged her with having:

> [1] processed personal data by automatic means without giving prior written notification to the data subjects or the Data Protection Inspectorate, and,
> [2] processed sensitive personal data (i.e. information about the volunteer's injured foot) without authorisation,

and she was fined SEK 4,000 (approximately £390) by the Swedish District Court. Mrs Lindqvist appealed against the decision to the Swedish Court of Appeal, which referred the case to the European Court of Justice. The ECJ ruled that: the act of referring, on an internet page, to various persons and identifying them by name or by other means, for instance by giving their telephone number or information regarding their working conditions and hobbies, constituted processing of personal data wholly or partly by automatic means. The Court rejected Mrs Lindqvist's argument that her actions fell within the Directive's Art 3(2) exemption for personal or domestic activities. The Court called Lindqvist's activities "charitable and religious,"[377] but said that the exceptions did not apply to the "charitable and religious" activities, but rather applied to the "exercise of activities which are exclusively personal or domestic, correspondence and the holding of records of addresses."[378] The ECJ also held that the exception of Article 3 applied to "activities...carried out in the course of private or family life of individuals"[379] and not to "publication on the internet so that those data are made accessible to an indefinite number of people."[380]

Additionally, the ECJ had to decide whether reference to the fact that one volunteer had injured her foot constituted sensitive personal data concerning health within the meaning of Article 8(1) of Directive 95/46/EC which states that "special categories" of data, namely "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership and the processing of data

---

that Lindqvist originally set up her web page, which was linked to the Church's website, to provide information for parishioners making Confirmation).

[374] *Lindqvist Judgment*, para 13
[375] *Ibid*, para 86
[376] *Ibid*, para 13
[377] *Ibid*, para 39
[378] *Ibid*, para 46
[379] *Ibid*, para 47
[380] *Ibid*, para47

concerning health or sex life." The Court held that information regarding health—both mental and physical—should be given a "wide interpretation," and that the reference was clearly health information under the Directive.[381] Interestingly, the Swedish authorities did not raise the question of whether naming individuals who acted as church volunteers, constituted processing of sensitive data regarding "religious or philosophical beliefs." Garcia postulates that this may be because Art 8 (2) (d) provides an exception for:

> "processing...in the course of its legitimate activities with appropriate guarantees by a...non-profit-seeking body with a political, philosophical, religious or trade-union aim and on condition that the processing relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes and that the data are not disclosed to a third party without the consent of the data subjects."[382]

Prior to Lindqvist, the Court had not ruled on the scope of the Directive in the context of the Internet. After Lindqvist, any individual or entity that posts personal information concerning EU citizens on the Internet may be subject to the Directive's privacy restrictions, regardless of whether the information is part of a commercial endeavour or non-commercial activity. The implications for bloggers are considerable, since posting information, including the names of individuals, on an internet page, constitutes processing which is, at least in part, automatic, and may result in liability if the Directive's requirements of prior notification and consent are not met. Accordingly, it was appropriate to examine the privacy attitudes and expectations of bloggers in order to investigate the continuing value of privacy in the information society.

## *6.4 Survey of Bloggers*

A global, online survey of bloggers was conducted, in which participants answered questions about their blogging practices and their expectations of privacy and accountability when publishing online. [383] This involved asking a question regarding the relative social importance of protecting personal information; questions regarding posting personal information about themselves; questions about posting personal information about others; and questions regarding any actions they take to protect their privacy.

---

[381] *Lindqvist Judgment*, para 50

[382] Garcia, F. J. (2005) "*Bodil Lindqvist*: A Swedish Churchgoer's Violation of the European Union's Data Protection Directive Should Be a Warning to U.S. Legislators," *Fordham Intellectual Property Media & Entertainment Law Journal*, pp. 1205-1239, at p. 1225.

[383] The data collection method was discussed in greater detail in chapter 5.

In this survey,[384] 49.1% of respondents were female, 35.4% were male, whilst 15.5% did not disclose their sex. Over half of the respondents were between 19 and 34 years of age (54.9%). Over one third of participants were from the UK (39.5%) which is not surprising, given that the survey questionnaire was available only in English and that announcements for the survey were posted to email lists in three UK universities.[385] These demographic characteristics contrast with findings from other blog surveys, in which participants in these spaces tended to be "young adult males residing in the United States."[386]

### *6.4.1 Bloggers value privacy*

One objective of the survey was to investigate whether privacy is valued in the Internet era. Accordingly, respondents were asked to rate a list of issues that could be considered of social importance on a scale of 1 to 5, where 1 is not at all concerned and 5 is very concerned.

Table 6.1 *Social importance of issues*

| Socially Important? | 1 Not at all concerned | 2 | 3 | 4 | 5 Very Concerned | No answer | Total |
|---|---|---|---|---|---|---|---|
| Preventing Crime | 5.2% | 7.9% | 22.7% | 27.8% | 22.5% | 13.9% | 100% |
| Improving standards in education | 3.3% | 3.5% | 11.9% | 26.3% | 41.3% | 13.7% | 100% |
| Protecting people's personal information | 2.8% | 6.1% | 17.4% | 27.6% | 33.3% | 12.8% | 100% |
| Protecting freedom of speech | 3.0% | 2.3% | 8.6% | 21.5% | 51.9% | 12.7% | 100% |
| Equal rights for everyone | 3.1% | 2.6% | 7.2% | 21.0% | 53.2% | 12.9% | 100% |
| Unemployment | 4.7% | 13.0% | 29.6% | 25.3% | 13.4% | 14.0% | 100% |
| Environmental issues | 4.0% | 7.7% | 18.7% | 27.4% | 28.4% | 13.8% | 100% |
| Access to information held by public authorities | 4.1% | 11.2% | 23.2% | 22.0% | 25.4% | 14.1% | 100% |

[384] Socio-demographic statistics are set out in detail in Appendix E

[385] Manchester University, Manchester Metropolitan University (student mail list), and Queen's University Belfast (Computer & Law students only)

[386] Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). "Bridging the gap: A genre analysis of weblogs," *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)* Los Alamitos: IEEE Computer Society Press. <http://www.blogninja.com/DDGDD04.doc> p. 5

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Providing health care | 4.5% | 6.7% | 16.0% | 27.1% | 31.6% | 14.1% | 100% |
| National security | 9.4% | 14.9% | 25.4% | 19.8% | 15.4% | 14.9% | 100% |

(Blog Survey 2006, n = 1258)

Of the issues listed, Table 6.1 indicates that the highest percentage (over half) of respondents were very concerned about equal rights for everyone, whilst over 1/3rd of respondents were very concerned with the protection of personal information. A higher percentage of respondents were very concerned with the protection of personal information than with the issues relating to preventing crime, unemployment, environmental issues, and access to information, or national security. Only 2.8% were not at all concerned with protecting personal information. The responses prima facie indicate that bloggers do value privacy.

### *6.4.2 Blogs as public or private spaces*

In order to test whether bloggers conceived of blogs as public spaces or private spaces, they were asked whether they kept a traditional diary or journal (i.e. a written book that is not shared online). Nardi *et al* report that, traditionally, diaries were written and stored in secret:

> "the classic diary is a volume whose privacy is secured by lock and key."[387]

In contrast, online blogs which serve as journals or diaries are usually publicly available.

Table 6.2 *Traditional Diary*

| Keep a traditional Diary? (i.e. not shared online) | Percentage |
|---|---|
| Yes | 21.2% |
| No | 66.3% |
| No Answer | 12.5% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

[387] Nardi, B., Schiano, D. J & Gumbrecht, M. (2004) "Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary?" Proceedings Conference on Computer-Supported Cooperative Work, (New York: ACM Press) pp. 222-231, at p. 222.

One fifth of respondents kept a traditional diary as well as an online blog. This lead to the question whether bloggers decide to write something in your diary/ journal but not to write about it online.

Table 6.3 *Diary Not Blog*

| Decide to write in Diary not blog? | Percentage |
|---|---|
| Yes | 21.5% |
| No | 23.2% |
| No Answer | 55.3% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

21.5% of respondents indicated a decision to write about it in their offline diaries as opposed to their blog. This suggests that the question of where to draw the boundaries between publishable and non-publishable materials is a matter of concern to bloggers. Below are illustrations:

> "My traditional diary is for my eyes only. It's more personal to me. I can put what I like without worrying about being read. I can [write] about people by their real names. (Female, 35-44, UK) [sic]
>
> If I'm particularly embarrassed about something I'm more likely to put it in my private diary than my blog, even though my blog is anonymous. Some things you just don't share." (Female, 19-24, UK)

These responses indicate that a minority of bloggers conceive of a traditional diary as a private space, since it is not accessible by anyone, whereas, they consider an online diary to be a more public space, and they restrict the information they place on the blog as a result.

### *6.4.3 Blogging about self*

In a previous study, Herring *et al*[388] coded a random sample of blogs for overall use based on the nature of the content posted; they found that the overwhelming majority of blogs (70.4%) were of the personal journal type: "in which authors report on their lives and inner thoughts and feelings." Thus, this study examined the main topic of blogs and reasons for blogging.

[388] Herring, S. Scheidt, L. Bonus, S. & Wright, E. (2004) "Bridging the Gap: A Genre Analysis of Weblogs," Proceedings of the 37th Hawaii International Conference on System Sciences, p. 10

Table 6.4 *Main Blog Topic*

| Main Topic of Blog | Percentage |
|---|---|
| My life (personal diary/journal) | 60.1% |
| Politics and government | 4.8% |
| Entertainment (movies, music, MP3's) | 6.0% |
| Sport | 0.6% |
| News and current events | 4.8% |
| Business | 1.3% |
| Technology (computers, internet, programming) | 8.5% |
| Religion/Spirituality/Faith | 0.7% |
| A particular hobby | 2.1% |
| Health (general health, an illness) | 1.1% |
| Gossip | 0.4% |
| Other | 7.6% |
| Prefer not to answer/No answer | 1.8% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

In keeping with previous studies, this survey found that most respondents (60.1%) said their entries could be characterized as "My life (personal diary/journal)". Many respondents indicated that their blogs featured a combination of different topics. Respondents were also asked to specify their reasons for blogging. The responses are outlined in Table 6.5 below.

Table 6.5 *Reasons for Blogging*

| Activity | Main Reason | Minor reason | Not a reason | Prefer not to answer | Total |
|---|---|---|---|---|---|
| To document your personal experiences and share them with others | 62.6% | 27.2% | 8.1% | 2.1% | 100% |
| To express yourself creatively | 50.9% | 35.9% | 10.8% | 2.4% | 100% |
| To influence the way other people think | 12.0% | 31.2% | 53.0% | 3.8% | 100% |
| To motivate other people to action | 10.8% | 30.0% | 54.8% | 4.3% | 100% |
| To share practical knowledge or skills with others | 16.5% | 36.8% | 42.4% | 4.3% | 100% |
| To network or to meet new people | 18.1% | 41.1% | 37.0% | 3.8% | 100% |
| To entertain people | 31.9% | 42.3% | 21.6% | 4.2% | 100% |
| To discuss problems with others | 18.9% | 39.3% | 37.4% | 4.3% | 100% |
| To stay in touch with friends and family | 31.0% | 24.2% | 41.0% | 3.8% | 100% |
| To make money | 1.6% | 4.7% | 88.3% | 5.4% | 100% |
| To store resources or information | 14.5% | 35.2% | 46.3% | 4.1% | 100% |

(Blog Survey 2006, n = 1258)

When asked to select reasons for blogging, the highest percentage (62.6%) of respondents indicated that their main reason for blogging was to document their personal experiences and share them with others, whereas the lowest percentage (1.6%) indicated that their main reason for blogging was to make money. Indeed, 88.3% percent indicated that making money was not a reason for them to blog.

In order to examine whether bloggers actually value privacy, a number of questions regarding their blogging behaviour were asked, since research by Acquisiti & Grossklags[389] indicates that there may be disparities between stated privacy attitudes and actual behaviour. Firstly, bloggers were asked questions regarding the type of personal information they revealed about themselves in blog posts.

Bloggers were asked if they identified themselves on their blog. In a small, qualitative study, Nardi *et al*[390] found that most bloggers provided accurate identity information and even contact information (via links to personal home pages). These findings echoed those of Herring *et al* [391] where 92.2% of the sampled blogs included explicit personal information on their first pages. Full names were found in 31.4% of sites, first names on 36.2%, and pseudonyms appeared on 28.7% of blogs. More than half of the blogs in the sample provided other kinds of explicit personal information such as age, occupation, and geographic location.

Table 6.6 *Self identification*

| Do you identify yourself on your blog (i.e. is your real name on the site)? | Percentage |
|---|---|
| Yes | 42.4% |
| No | 32.1% |
| On Some, not others | 7.7% |
| Prefer not to answer | 0.3% |
| It's More complicated than that | 16.3% |
| No Answer | 1.2% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

[389] Acquisiti, A. & Grossklags, J. (2004) "Privacy Attitudes and Privacy Behaviour: Losses, Gains, and Hyperbolic Discounting," in Camp, L. J. & Lewis, S. *The Economics of Information Security* (Kluwer), pp. 165-178

[390] Nardi, B., Schiano, D. J & Gumbrecht, M. (2004) "Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary?" Proceedings Conference on Computer-Supported Cooperative Work, (New York: ACM Press) pp. 222-231

[391] Herring, S. Scheidt, L. Bonus, S. & Wright, E. (2004) "Bridging the Gap: A Genre Analysis of Weblogs," Proceedings of the 37th Hawaii International Conference on System Sciences, p.3

In this survey, respondents for the most part (42.4%) identified themselves on their blogs by posting their real name. This tendency for self-identification *prima facie*, suggests that bloggers are not overly concerned with protecting their privacy. However, 16.3% of respondents indicated that they exercised some restraint in revealing personal name details or other identifying information. The types of behaviour are listed below:

1) *First name only, or first name and a surname initial, or maiden name instead of legal surname*
2) *Pseudonym, a nickname, penname or alias*
3) *First name and geographical data e.g. State or town*
4) *Of those who reveal only their first name on the blog many stated that their full name details could be found in the URL, (or via their email address which they use to respond to posts on others blogs)*
5) *No name, but photograph*
6) *Full name – only because it is very common e.g. Mike Martin (many Google hits – so effective anonymity)*

Bloggers were also asked how often they post personal information on their blogs.

Table 6.7 *Frequency of posting personal information*

| Frequency of posting personal information | Percentage |
|---|---|
| All the Time | 24.8% |
| Most of the Time | 29.2% |
| Some of the Time | 25.0% |
| Rarely Ever | 7.6% |
| Never | 2.0% |
| No Answer | 11.4% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

When asked how often they had posted personal information on their blogs, 24.8% of respondents said they had done so "All the time." Only 2% of respondents said they had "never" posted anything highly personal on their blogs. So, most blogs contain personal information.

Also, bloggers were asked whether they ever considered some information 'too personal' to post on their blog.

Table 6.8 *'too personal' to post on blog*

| Too Personal to post? | Percentage |
|---|---|
| Yes | 65.6% |
| No | 18.3% |
| No Answer | 16.1% |
| Total | 100% |

Blog Survey 2006, n = 1258)

Most respondents (65.6%) said they had considered certain topics were too personal to write about on their blogs. This suggests that the question of whether certain materials were too personal to blog about weighed heavily in their decisions of what to publish. Chapter 8 of this thesis will explore in more detail the comments made by bloggers regarding types of information they considered too personal to post; in particular, it will examine the relationship between 'personal' and 'private' data.

### *6.4.4 Blogging about others*

Additionally, bloggers were asked questions regarding posting personal information about third parties. This is important since the Lindqvist decision determined that any person who publishes personal information without seeking consent may be prosecuted for breach of data protection laws.

Table 6.9 *Identification of others*

| When you write things about people you know personally in your blog, do you reveal their full names? | Percentage |
|---|---|
| Yes | 7.4% |
| No | 71.1% |
| I never write about people I don't personally know | 7.0% |
| Prefer not to answer | 14.5% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

When asked whether they identified other people by name on their blogs, 7.4 % of respondents said they revealed full names, whereas 71.1% said they did not. This indicates that most bloggers are sensitive to issues of privacy when blogging about

people they know. Indeed, Table 6.9 below indicates that over half (51.8%) of bloggers used an identifier instead of name when blogging about someone they know personally.

Table 6.10 *Use of Identifiers*

| When you write things about people you know personally in your blog, do you use an identifier instead of name? | Percentage |
|---|---|
| Yes | 51.8% |
| No | 26.7% |
| I never write about people I don't personally know | 6.5% |
| Prefer not to answer | 14.9% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

### *6.4.5 Privacy invasions*

Bloggers were also asked to reflect on whether their privacy had ever been invaded through the processing of personal information on third party blogs.

Table 6.11 *Privacy invasion*

| Has anyone ever invaded your privacy by mentioning you on their blog? | **Percentage** |
|---|---|
| Yes | 11.2% |
| No | 76.8% |
| Prefer not to answer | 12.0% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

Table 6.11 indicates that more than one in ten bloggers had experienced privacy invasion through the activities of other bloggers. When asked to explain the ways in which their privacy had been invaded, respondents described the following situations:

*Revelations by a former relationship partner:*

> "After a break-up, my ex decided to use her 'blog' to write an open letter to me."(Male, 25-34, UK)

> "Personal health information was babbled to the world by an ex. Not appreciated!"(Female, 25-34, Canada)

The responses revealed a desire for the details of intimate relationships to remain undisclosed, protected by a sense of trust and implicit confidentiality or secrecy that is normally associated with such relationships.

*Friends revealing identifiers*

> "A friend linked to me and when she talked about me in her blog, she used my real name." (Male, 25-34, USA)
>
> "I have friends who sometimes mention where I live, and I don't really like that."(Female, 35-44, UK)
>
> "Only in a small way - some of my friends call me by my first name when they comment." (Female, 45-54, USA)
>
> "Putting my picture on their blog without asking me, though I have done the same." (Female, 19-24, England)

The disclosure of personal identifiers e.g. name or photo by friends reveals that some bloggers perceive blogs as an extension of their personal relationships – forgetting that information revealed there can be publicly available.

Thereafter, bloggers were asked whether they had ever invaded a third party's privacy by omitting to seek consent before posting personal information about them.

Table 6.12 *Invasion of other people's privacy*

| When you write things about **people you know personally** in your blog, do you ask their **permission?** | **Percentage** |
|---|---|
| Yes | 15.4% |
| No | 61.8% |
| I never write about people I don't personally know | 8.3% |
| Prefer not to answer | 14.5% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

Table 6.12 indicates that 61.8% of respondents did not seek other people' permission before writing about others. Thus the great majority of respondents write about people they know, but most of them never ask their permission to do so. If the decision in the *Lindqvist* case were followed, many bloggers could be prosecuted for posting personal information about other individuals on their blogs without permission.

Table 6.13 *Gotten into trouble*

| You or friends gotten into trouble? | Percentage |
|---|---|
| Yes | 19.6% |
| No | 66.1% |
| No Answer | 14.4% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

When asked if they or their friends had ever gotten in trouble because of things they had written on their blogs, Table 6.13 indicates that almost one in five (19.6%) of respondents said they had. When asked to explain the kinds of problems they had encountered because of materials published on their blogs, respondents described the following situations:

*Blogging about family/relationships*

> "In my youth, I made a statement of my desire for my own life to end; this upset my mother more than a little." (Male 19-34, Country Unspecified)
>
> "My boyfriend read my blog and although I hadn't written anything bad and had concealed both of our identities, he still felt violated."(Female, 25-34, Canada)
>
> "My friend posted rather scathing comments about her family at Thanksgiving one year; a cousin found it and was incredibly hurt and upset. This same friend also posted about an ex boyfriend and he read her posts and didn't speak to her for a few weeks." (Female, 19-34, Canada)
>
> "A friend totally moved her blog after a family member read one of her posts and she was upset by it. There were no identifying names or information that would make sense to anyone other than those involved. Another friend promised her partner that she wouldn't write about their relationship after she posted some negative comments about him and he saw them."(Female, 25-34, Australia)
>
> "I have a friend who went through a divorce and her blog proved to be sticky in the proceedings and angered her ex-in-laws."(Female, 25-34, USA)
>
> "My 19 year old daughter (living with her father) found my blog and was truly hurt by quite explicit personal things mentioned about her father and his physical abuse towards me during our marriage. (We are divorced). I decided to delete and move my blog over to another host and rename it." (Female, 35-44, USA)
>
> "I upset my then boyfriend, now husband, when he found my blog soon after we met because I had said some unflattering things about his family. I subsequently took that blog down and started a new one on which I do not

write anything inflammatory about anyone I know personally." (Female, 25-34, USA)

*Blogging about friends*

"A flatmate had been reading my blog for a significant period of time; under the impression I didn't know he had access to it. Due to a statcounter - which identified his IP - I was in fact well aware he read frequently. I would very occasionally mention him - using an identifier - and after a period of time (after he had moved out) he confronted me, saying he was concerned that I had been secretly writing about him on the internet. I pointed out that I hadn't mentioned him by name and that I was well aware he had been reading, and thus perhaps he should get his own house in order first because as far as he was knew he had been reading without my knowledge. He didn't agree, and we no longer speak."(Male, 25-34, UK)

"I wrote something personal which my best friend did not know about when I first started my anonymous blog. A while later she found my blog and realised it was me and got me into trouble with her."(Female, 19-24, UK)

*Blogging about work*

"Upset a colleague - removed the post, vowed never to blog about work again." (Male, 35-44, UK)

"I did blog about work, once, publicly; I was asked to stop; I stopped." (Male, 25-34, UK)

"Yes, my fiancée was sacked from his job at Lambeth Palace after writing about the change of Archbishops online." (Female, 19-24, UK)

"Yes, one of my acquaintances has gotten into legal trouble for mentioning his workplace. Another of my acquaintances is an author, and someone used his semi-private blog as a source of information for an article."(Female, 25-34, Sweden)

"I almost got fired from my last job, so I deleted it and started a new one. I work at home now, so whatever I say is only about how much I work because I cannot divulge any information on a public (even password protected) forum or blog. I signed a contract, and to do so, and get caught would be breach of contract and termination."(Female, 25-34, USA)

"I posted something about a terrible boss I had and he found the blog and threatened legal action even though I hadn't mentioned his name or the name of his business. I removed the post and found another job but it taught me not to give a wider berth to other people's stories on my blog." (Female, 35-44, USA)

"A friend of mine who teaches high school has had fake Myspace profiles created by his students which identified him as sexually involved with them; he was investigated by the authorities (both school and police) and found to have had nothing to do with the acts." (Female, 25-34, USA)

*Blogging about strangers*

> "I made fun of a professional ping-pong player's death and was spammed by members of the professional ping-pong community. I am not making this up. Before then, I had no idea such a community existed. Eventually, I pulled the comment from my blog." (Male, 25-34, UK)

> "Early on, I misinterpreted something someone said (a stranger) and commented on it. He was furious. I started a dialog with him and we both apologized."(Female, 25-34, Canada)

*Disputes about intellectual property or third parties*

> "A blog friend tried to go after some jerk for copyright infringement because he'd been stealing posts from her blog and publishing them on his own site with Google ads to make money. This same jerk, in retribution for her pursuing the case, posted her full name and address on the internet. She filed her complaint with Google so they would remove their ads from his site, and to file the complaint she had to provide her name and address and the geniuses at Google forwarded her info to the evil jerk. Nice, huh?" (Female, 25-34, USA)
> "I was once asked to remove an MP3 I posted in advance of its release date and to replace it with another song from the same album, which had been cleared for release for promotional purposes. (And I did so, and received a very friendly email thanking me) I wasn't threatened with legal action or anything of the sort."(Female, 25-34, USA)

> "A friend of mine posted about very bad service she has received from a company, and described it using a profanity as a nickname. The CEO found out and invited her to visit the company and express her frustrations in person. She was very humiliated to have been found out in such a public way, and ended up apologizing profusely on a later post."(Female, 25-34, Jordan)

### *6.4.6 Bloggers Employ Mechanisms to Protect Privacy*

The bloggers' responses indicated that they had gotten into trouble over their blog posts. Accordingly, it was appropriate to examine whether they employed any privacy protection measures**.**

Table 6.14 *Restriction of Access to Content*

| Do you do anything to limit who can read your blog? | Percentage |
|---|---|
| Yes | 25.6% |
| No | 72.3% |
| No Answer | 2.1% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

Bloggers were asked whether they do anything to limit who gets to read what they post, 72.3% of respondents said no. However, in the open-ended responses it became clear that access control is of major importance to bloggers. Of those who revealed detailed identifying information, many stated that they were happy to do so because they controlled audience access e.g. limited to friends only – and as they know such people in real life they are happy to continue sharing such information on the internet. These limitations included:

*Leaving out key details:*

> "I made the title and address completely unconnected to me and don't use my surname so a Google search of my name wouldn't flash it up."(Male, 19-24, UK)

> "I do not document everything that has happened to me on my site. While my blog is predominantly a personal one (i.e. 'What I did today', 'What I learned today', 'who I spoke to today', etc.) I prefer having a sense of anonymity." (Female, 19-24, Singapore)

*Using passwords*

> "To avoid Spammers I have put a computer word verification in place. I only did this after receiving some questionable responses to my posts."(Female, 35-44, USA)

> "Some entries are completely private and require personal login. The rest are completely public."(Male, 25-34, UK)

> "Password for blog is only provided for friends."(Male, 19-24, UK)

*Keeping the fact that they blog secret*

> "Most people who know me personally will never know about my personal blog. Only folks I totally trust and already share all my stories and inner issues with were invited."(Female, 35-44, Country Unspecified)

> "I use a pseudonym so that close family members about whom I may write (using an initial, not their full name) cannot easily come across my blog if

> they were to search one day. I do not want to have to explain myself if what I write is 'injurious'…" (Female, 35-44, USA)
>
> "Other than my partner, my family does not know that my blog exists. Two real-life friends know about it. All other readers do not know me personally."(Female 25-34)

*Editing robot.text file*

> "Edit the robots.txt file which controls whether search engines are allowed to crawl your site."(Male, 25-34, Australia)
>
> "I use blogger and I blocked the search engine option. So, only if you click on a link from someone else's blog can you stumble across mine." (Female, 35-44, USA)

*Blocking IP addresses*

> "I ban IP addresses of people who come to my site just to insult me. I also use private categories to tuck away posts that I'd rather not have the general public read."(Female, 19-24, USA)
>
> "I block the IP address of my sister-in-law as well as the IP [address] of my company, but I turn that on and off, as sometimes I update at work."(Female, 25-34, USA)
>
> "I block a lot of spambots, and the occasional troll - but the main thing I do is block any IP addresses and domains that I know my parents use, to stop them accidentally coming across it."(Male, 25-34, UK)

*Using privacy filters*

> "Most of the posts are public; these include ones about what I do, my fandoms etc. I have several custom friends groups to discuss things I think some people on the list would disapprove of (for example my religious and spiritual beliefs)" (Female, 35-44, UK)
>
> "Generally I make my entries open for all to see and respond to. However, if I'm discussing something either particularly personal or something that involves people who might be reading the blog, I make my entries friends-only."(Female, 19-24, USA)
>
> "Some of the content of my blog is lightly filtered to avoid spoiling surprises for people or to talk about work or to preserve other people's privacy or to send a message or invitation to a certain section of people."(Female, 19-24,UK)

The comments indicate that bloggers used a variety of mechanisms to restrict public access to their blogs, such as locks, password access, and friends only filter. Such behaviour indicates that bloggers negotiate a boundary between self and society that they feel comfortable with, yet at the same time they are able to interact socially with

their readers. In this way they are able to define and maintain the desired level of public accessibility or privacy that they wish to achieve through the level of personal exposure that they allow. There is evidence of a growing concern about protecting anonymity among some respondents. A common reason for limiting details was to prevent it from leading 'Google' searches by employers to their personal blogs.

> "I use my first name, but always leave out my surname. I also try not to mention by name where I work or where I grew up. This isn't so much because I don't want my audience knowing these details, but rather that I am aware that including such details makes it much more likely an employer, former acquaintance or anyone I wouldn't want reading might accidentally 'Google' their way onto my site. Despite these safeguards, some friends have still managed to Google their way to my blog, so I think my concerns are well founded. If I were to start blogging afresh, I would give serious consideration to adopting a pseudonym."(Male, 25-34, UK)

> "I don't have my full name on the blog about page, but I have mentioned it many times. I want my day job work to be my primary Google search result for my name."(Anonymous)

> "Like to keep work and home life separate (I'm a social worker) so using my real name is not a good idea in case a client did an internet search."(Female, 25-34, UK)

Some bloggers initially preserved their anonymity, but are aware that the reasons for their initial behaviour are changing, e.g. because they have changed jobs, or are comfortable sharing the information they post:

> "I use a fake name, which I originally assumed to keep my blogging completely separate from my work life. I've since left that job, and now am a lot more forthcoming with personal details, but I've kept the name, partly because I've become fond of it, and partly because other bloggers now know me as Ben."(Male, 25-34, UK)

> "I do not openly list my name on my blog, however I do reference my family members by first name, have listed my last name on occasion and list the city in which I live. I suppose that I had at one time planned to remain an unknown, however have not found that to be as important as time has gone on."(Female, 35-44, USA)

> "I started the blog anonymously and posted under a pseudonym. Then I launched a web site with my name in the URL. The blog is now hosted on the website, but I still post under the pseudonym. … Obviously we're one and the same, but I still find the alter ego a useful psychological and literary device."(Male, 25-34, USA)

In contrast, some bloggers are moving towards anonymity:

> "First my name was public, now...my name is hidden." (Female, 19-24, Germany)

> "...after a while realized that I didn't want to have my name be so easy to Google…I don't really have an issue if people know who I am and where I live, but since the purpose of my blog is to keep my family updated on the lives of my kids, it doesn't seem necessary. The blogs that I keep for my children's birth families do NOT have names."(Female, 25-34, USA)

> "It was on all my blogs until a few months ago. My main blog has a pseudonym while I apply for jobs, and will revert back to real name afterwards. On my secondary blogs, I use my real name, but these are primarily professional/hobby-related. My name is unique, which makes me quite careful."(Female, 35-44, Australia/UK)

Some bloggers noted that they rely on anonymity in their blogging activities, particularly when posting information they considered to be more personal or private by traditional standards – the anonymity of blogging made them more likely to post such information.

These findings suggest that privacy norms are emerging among bloggers. For instance, as the comments above illustrate, some bloggers are beginning to create informal guidelines for publishing the names of people and employers in their blog entries. It is suggested that there is evidence of bloggers altering their behaviour according to employment prospects as the comments indicate that some bloggers are wary of revealing personal information to prospective future employers, and of revealing such information to potential clients. The degree of accessibility is a major part of what makes a blog public or private. In this regard, the more accessible or visible a blog is, the more it is considered to be public. Some bloggers opined that anonymity or privacy is not possible on the internet.

> "Anyone publishing anonymously in any medium must accept the risk of being 'outed'. Though I deplore the gratuitous and often destructive identification of anonymous bloggers, it would be foolish for anyone to assume anonymity is a right."(Anonymous)

> "Blogs are a public thing. Some might think that they are private like emails, but should realise that both emails and Blogs are public in the sense that they can be found by someone who wants to find them. It's like paparazzi taking a photo of a famous person topless on a public beach." (Male, 25-34, Japan)

> "Many bloggers know that there is no real privacy and that anonymity is just a temporary matter. If someone wants to find out the person behind the blog, it would be quite easy to do so."(Female, 35-44, Kuwait)

> "It's very hard to have a totally anonymous blog. People who know you may reveal who you are. There is also a chance that you can make a slip that reveals who you are." (Male, 35-44, UK)

The degree of accessibility is a major part of what makes a blog a public or private space – the more accessible or visible a space is the more it is considered to be public. Also, level of familiarity/audience knowledge is a key determinant of whether a blog is considered a public space. Hence, the fact that a blog could be read by strangers could mean that the space is considered as public by many bloggers.

## *6.5 Discussion*

Most respondents in this study described their blogs as the personal diary/journal type, and the majority of bloggers post personal information on their blogs which indicates that blogs are used to provide opportunities for self expression and communication with others. However, bloggers face unique privacy concerns because, on blog posts, meaningful and personal information is often shared, yet, most blogs exist in a fully public arena, which means that, once published, entries are readable by anyone on the Web. Yet, bloggers comments indicate that they are aware of a risk posed to their personal privacy by external parties who might be interested in collecting or collating the information they post and thus they seek to restrict their blog readership and content.

In addition, respondents indicated that they often post personal information about others on their blogs, and do so without seeking the consent of the third party. On the one hand, this type of behaviour could be regarded as evidence of a change in social norms, i.e. as suggested by Zuckerberg, that it is becoming increasingly normal for individuals to share large volumes of personal information about themselves and others because they no longer value privacy. On the other hand, a significant portion (19%) of respondents stated having gotten in trouble because of things they had written on their blogs. Indeed, the results from this survey suggest that bloggers are starting to encounter a range of privacy-related issues varying from minor embarrassments with family and friends to termination of their employment, as a result of their online disclosures. However, the findings also indicate that many respondents are developing strategies for minimizing potential problems with others

when posting their entries online. They reported employing mechanisms and changes in their behaviour to avoid such problems in the future. This finding concurs with an assertion by Palen and Dourish[392] that privacy in networked environments is a dynamic, dialectic process of negotiation that is conditioned by people's own expectations and experiences and by those of others with whom they interact.

The findings in this study also support assertions by Rosen,[393] Grudin,[394] and Palen and Dourish[395] that bloggers consciously and intentionally negotiate the boundary between public and private. They take responsibility to ensure that their posts are in line with their desires as to how public or private they want to be at that specific time - a process that may shift from day-to-day, and from topic-to-topic.

## *6.6 Summary*

Over 1/3rd of bloggers who participated in the survey indicated that protection of personal information was an issue of social importance. Also, although the very act of blogging appears to fly in the face of privacy, the bloggers' comments indicated that they do continue to value privacy, and actively seek mechanisms to protect their individual privacy. Given that the internet usage by the general public did not become commonplace until 1996, and easy to use blogging technology did not emerge until 1999, the norms of behaviour when using such technology to interact are still in the process of being negotiated and settled. As a result bloggers are learning to manage both their own privacy preferences, and those of others whom they blog about. The findings suggest that they often breach the privacy of others by posting information about third parties without seeking prior consent. However, there is evidence of changing practices and behaviour, in that bloggers often modify their practices when they are informed that their activities have caused harm e.g. upset a third party.

---

[392] Palen, L. & Dourish, P. (2003) Unpacking "privacy" for a networked world. Proceedings of the ACM CHI. < http://delivery.acm.org/10.1145/650000/642635/p129-palen.pdf?key1=642635&key2=7273414711&coll=&dl=GUIDE&CFID=15151515&CFTOKEN=6184618> (Last accessed: 01.03.09)

[393] Rosen, J. (2000) *The Unwanted Gaze: The Destruction of Privacy in America.* (Vintage Books: New York)

[394] Grudin, J. (2001). "Desituating action: Digital representation of context." *Human-Computer Interaction, 16* (2-3), pp.269-286.

[395] Palen, L., & Dourish, P. (2003). Unpacking "privacy" for a networked world. *Proceedings of the ACM CHI.* Ft. Lauderdale, FA <http://delivery.acm.org/10.1145/650000/642635/p129-palen.pdf?key1=642635&key2=7273414711&coll=&dl=GUIDE&CFID=15151515&CFTOKEN=6184618> (Last accessed: 01.03.09)

# Chapter 7

## Critique of sensitive data

### *7.1 Introduction*

This chapter begins by examining empirical evidence on the continuing relevance of existing categories of sensitive data, since chapter four indicated that changes in society and technological developments may influence the sensitivity of data. It will demonstrate that the current classifications of sensitive data are somewhat outdated and ineffective for determining the conditions of data processing and affording privacy protection. Further, this chapter investigates whether the term 'sensitive' is synonymous with the term 'private' data, and concludes that the terms sensitive and private are not synonymous, and that the focus of the Directive has erroneously been on classifying data as sensitive instead of regulating harm arising from data uses.

#### *7.1.1 Aim*

The main aim of this chapter is to critique the term 'sensitive' data and explore the relationship between the protection of sensitive data and the protection of privacy. This produced three key areas of questions: questions about satisfaction and continuing relevance of existing categories in Directive 95/46/EC and the Data Protection Act 1998; questions about potential new categories of sensitive data; and questions about the relationship between the terms sensitive and private data.

#### *7.1.2 Approach*

A two-fold approach was required to answer the questions: an analysis of how the term 'sensitive' data is understood by Privacy and Data Protection experts from around the world and, an analysis of how the term 'sensitive' data is understood and

perceived by potential data subjects. This was achieved through a four-phase study: Phase 1 – an analysis of the term 'sensitive' in legislation; Phase 2 – semi-structured interviews with privacy and data protection experts; Phase 3 – telephone survey of UK respondents; Phase 4 - Internet survey of bloggers.[396]

## *7.2 Recap of definitions of Sensitive Data*

Chapter four critiqued the term 'sensitive' data. It explained that a number of data protection measures seek to protect privacy by stipulating that the processing of certain types of data should be subject to stricter controls than other types of personal data.

Table 7.1 *Categories of sensitive data in International Legislation*

| OECD Guidelines (1980) | Council of Europe Convention (1981) | UN Guidelines (1990) | Directive 95/46/EC | APEC Privacy Framework (2005) |
|---|---|---|---|---|
| None | Racial origin | Racial or Ethnic origin | Data concerning race or ethnic origin | None |
| | Political opinions | Political opinions | Political opinions | |
| | Religious or other beliefs | Religious/philosophical/other beliefs | Religious or philosophical beliefs | |
| | Sexual life data | Sex life | Sexual life information | |
| | Health data | Membership of a trade union | Health information | |
| | Criminal convictions | Membership of an association | Trade-union membership | |
| | | Colour | Criminal records | |

The textual analysis conducted in chapter four, and briefly summarised in Table 7.1 above, indicates that two radically different legislative approaches have been adopted. On the one hand, the OECD guidelines and APEC Privacy Framework do not single out specific categories of personal data as having a 'sensitive' quality which merits extra legal protection. In contradistinction, the CoE Convention 108, UN Guidelines (1990) and Directive 95/46/EC enumerate categories of data classified as sensitive. Even so, there is a lack of uniformity regarding the categories of sensitive data. For instance, Directive 95/46/EC includes the trade union membership as a specific category of sensitive data. Similarly, Directive 95/46/EC

[396] The data collection methods were discussed in detail in chapter 5.

differs from the UN Guidelines as it lacks a category of data on colour or membership of association, but includes a category of criminal convictions.

## *7.3 Satisfaction with sensitive data classification*

The first objective of this chapter is to critique the term 'sensitive' data by examining whether the current categories are still considered sensitive, and also whether new categories of sensitive data have emerged. Accordingly, semi-structured interviews were conducted with forty privacy and data protection experts, whilst potential data subjects' views were sought through the Information Commissioner's annual telephone survey of the British public and an online survey of bloggers.[397]

### *7.3.1 Responses from expert interviewees*

#### *7.3.1.1 Satisfaction with current categories of sensitive data*

Some respondents were happy with the existing definition and the types of data covered. For example, one respondent stated:

> "In the UK existing categories of sensitive data have merit in that they are associated with a right to human dignity/freedom of political activity. The difficulty with the current provisions is the overriding public interest tests, in the EU Directive there is a categorical prohibition on the processing of certain data – but it is subject to higher public interest tests…Existing categories of sensitive data are sensible." (UK)

Likewise, an Irish expert interviewee stated:

> "I'm broadly happy with existing definitions in Ireland. The approach taken in the Directive is correct. Sensitive data is an arbitrary list." (Ireland)

Others did not agree with all classifications. For instance, an expert from Iceland commented that:

> "We had to introduce the concept of sensitive data in Iceland but we don't agree with all the categories e.g. according to the Directive data on trade unions is considered sensitive, but in Iceland such information is not as everyone knows where you work and what unions you belong to and they don't care about these things…" (Iceland)

*b) Potential new categories:*

Also, some Interviewees suggested new categories of sensitive data. Below are some illustrations:

[397] Data collection methods are detailed in Chapter 5.

> "Some regard or suggest financial data to be sensitive – in this regard the categorisation of it as non-sensitive is clearly arbitrary – it may be worthwhile amending the legislation to make it sensitive." (Ireland)

Interviewees indicated that technological developments are generating potential new categories of sensitive data, for example:

> "They could be expanded e.g. to include financial data. They could be ramified. E.g. for health data a biometric template[398] should probably be considered personal data but probably isn't sensitive data. Whereas, genetic information could be regarded as sensitive because of the potential for prejudice and unfairness of inappropriate disclosure." (UK)

Ten potential new categories of sensitive data emerged from the semi-structured interviews. They are listed in Table 7.2 below

Table 7.2 *Potential new categories of sensitive data*

| **Potential new categories of sensitive data** |
|---|
| *Employment history* |
| *Education Qualifications* |
| *Membership of political party / organisation* |
| *Clickstream data (e.g. record of web pages visited)* |
| *Personal Contact Details* |
| *Genetic Information* |
| *Biometric information (e.g. iris scans, facial scans and finger prints)* |
| *Financial data* |
| *Data relating to children* |
| *Email address* |

Using these categories, a question was designed to examine public perceptions of sensitive data.[399] Firstly, it was used to test sensitivity ratings of seven categories of data which are currently recognised in the Directive as sensitive. Also, it was used to test perceptions of sensitivity towards eight not legally recognised categories of sensitive data which emerged in interviews with data protection and privacy experts. A decision was made to subsume the potential new category of email address into the

[398] Biometrics comes from the Greek words *bios* (life) and *metrikos* (measure). The term refers to any specific and uniquely identifiable physical human characteristic, *e.g.*, of the retina, iris, acoustic spectrum of the voice (*i.e.*, voiceprint), fingerprint(s), handwriting, pattern of finger lengths, *etc.*, that may be used to validate the identity of an individual.

[399] Two versions of the question were utilised. Firstly, a question was inserted in the ICO telephone survey. This question was modified slightly when included in the online survey questionnaire, to reflect the fact that non-EU bloggers might respond to the survey. The questions can be found in Appendices C & D.

category of contact details, as the concept of contact details was sufficiently broad to encompass email address. Also, the experts expressed a need to offer a high level of protection to all data relating to children. However, the health data of a child is not necessarily more sensitive than the health data of an adult. Rather, the experts were expressing a view that such data should be processed properly so that children are not unfairly discriminated against, or slip through the system. These concerns could fit within existing categories of sensitive data e.g. health, criminal record or education categories, and so a decision was made not to include a separate category of data relating to children, as the concerns related to the nature of the data subject rather than the sensitivity of any particular data type . The 15 categories of sensitive data tested are displayed in Table 7.3.

Table 7.3 *Classification of sensitive data*

| Art 8 Legally recognised categories | Not legally recognised categories |
|---|---|
| *Trade-union membership* | *Employment history* |
| *Religious or philosophical beliefs* | *Education Qualifications* |
| *Political opinions* | *Membership of political party / organisation* |
| *Data concerning race or ethnic origin* | *Clickstream data (e.g. record of web pages visited)* |
| *Criminal records* | *Personal Contact Details* |
| *Sexual life information* | *Genetic Information* |
| *Health information* | *Biometric information (e.g. iris scans, facial scans and finger prints)* |
| | *Financial data* |

## *7.3.2 Findings from ICO Annual Track telephone survey of British public.*

The views of UK citizens regarding the concept of sensitive data were sought through the ICO Annual Track (Individual survey 2006) telephone survey.[400]

### *7.3.2.1 Exploratory Analysis*

An exploratory analysis of the data was conducted in order to assess whether there were any differences in sensitivity perceptions between legally recognised categories of sensitive data, and the potential, new categories of sensitive data which emerged during the course of interviews with privacy and data protection experts.

[400] The survey was conducted by telephone. All the interviews were conducted in house by SMSR's telephone interviewing team. The total sample was 1,066 interviews. Quotas were set on age, sex, region and social grade to ensure a nationally representative sample was achieved. The research method is discussed in greater detail in Chapter 5. The survey questions are in Appendix C.

Table 7.4 *Sensitivity of different types of personal information - ICO Survey*

| | Don't Know | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data concerning race or ethnic origin** | **1.3%** | **12.7 %** | **6.8 %** | **6.6%** | **4.6 %** | **21.9 %** | **5.0%** | **7.2%** | **9.4%** | **3.8%** | **20.7 %** | 100% |
| **Political opinions** | **0.9%** | **10.7 %** | **5.3 %** | **6.8%** | **6.3 %** | **21.7 %** | **6.5%** | **7.4%** | **9.9%** | **4.5%** | **20.0 %** | 100% |
| **Religious or philosophical beliefs** | **0.9%** | **15.3 %** | **6.1 %** | **6.8%** | **5.3 %** | **21.4 %** | **6.8%** | **5.3%** | **8.1%** | **3.5%** | **20.5 %** | 100% |
| **Trade-union membership** | **1.4%** | **14.6 %** | **6.9 %** | **6.5%** | **7.4 %** | **23.5 %** | **7.1%** | **8.0%** | **7.1%** | **2.7%** | **14.7 %** | 100% |
| **Health information** | **0.9%** | **3.3%** | **.6%** | **2.6%** | **2.4 %** | **12.8 %** | **5.4%** | **8.5%** | **12.1 %** | **8.0%** | **43.3 %** | 100% |
| **Sexual life information** | **1.6%** | **4.8%** | **2.1 %** | **3.8%** | **2.9 %** | **13.4 %** | **4.6%** | **5.9%** | **11.2 %** | **6.8%** | **43.1 %** | 100% |
| **Criminal records** | **1.1%** | **7.7%** | **3.5 %** | **3.7%** | **3.6 %** | **18.9 %** | **4.0%** | **7.5%** | **10.0 %** | **6.0%** | **34.1 %** | 100% |
| **Education Qualifications*** | **1.4%** | **10.4 %** | **5.4 %** | **6.1%** | **5.6 %** | **20.5 %** | **9.0%** | **8.0%** | **11.5 %** | **4.6%** | **17.4 %** | 100% |
| **Employment history*** | **1.1%** | **10.1 %** | **5.8 %** | **6.1%** | **6.0 %** | **23.2 %** | **7.0%** | **8.2%** | **11.2 %** | **4.8%** | **16.5 %** | 100% |
| **Membership of political party/ organisation*** | **1.5%** | **11.9 %** | **5.5 %** | **7.8%** | **5.6 %** | **22.7 %** | **7.4%** | **6.6%** | **8.9%** | **4.0%** | **18.0 %** | 100% |
| **Clickstream data*** | **2.5%** | **10.5 %** | **5.4 %** | **5.7%** | **5.7 %** | **20.6 %** | **6.8%** | **8.3%** | **9.9%** | **4.2%** | **20.2 %** | 100% |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Personal Contact Details*** | **0.4%** | **4.1%** | **2.9 %** | **3.4%** | **3.8 %** | **13.2 %** | **4.6%** | **7.5%** | **13.9 %** | **5.3%** | **40.9 %** | 100% |
| **Genetic Informatio n*** | **1.6%** | **6.3%** | **2.4 %** | **3.5%** | **2.8 %** | **15.1 %** | **5.7%** | **8.1%** | **11.2 %** | **6.8%** | **36.5 %** | 100% |
| **Biometric information *** | **2.2%** | **7.3%** | **3.1 %** | **3.2%** | **3.7 %** | **13.4 %** | **4.1%** | **6.3%** | **11.4 %** | **6.5%** | **38.9 %** | 100% |
| **Financial data*** | **0.6%** | **1.2%** | **0.7 %** | **1.1%** | **1.4 %** | **5.0%** | **2.2%** | **5.5%** | **11.9 %** | **8.3%** | **62.1 %** | 100% |

(Source: ICO Annual Track Survey 2006) (n=1066) (*= Not legally recognised category)

Table 7.4 illustrates that whilst all the legally recognised categories of sensitive data are still considered sensitive, there are variations in the level of sensitivity among the different types of data. Of the seven legally-recognised categories of sensitive data, the highest percentage (15.3%) of respondents considered Religious or Philosophical data to be not at all sensitive, whilst the lowest percentage (3.3%) of respondents considered health information to be not at all sensitive. Of the eight categories of sensitive data that are not currently legally recognised the lowest percentage (1.2%) of respondents considered Financial data not at all sensitive, whilst the highest percentage (11.9%) considered Membership of a political party not at all sensitive.

Of the legally-recognised types of sensitive data, health and sex life information were considered extremely sensitive by the highest percentage of respondents. Interestingly, some of the not legally recognised categories were considered to be more sensitive than the legally-recognised types of sensitive data. For example, financial data was considered extremely sensitive by most respondents (62.1%) and had a higher sensitivity rating than any of the legally recognised categories. Likewise, more than one third of respondents rated biometric, genetic and contact details as extremely sensitive, whereas only one fifth of respondents rated data concerning race or ethnic origin, political opinions or data concerning religious or philosophical beliefs as extremely sensitive. The results indicate that new categories of sensitive data have emerged as a result of technological developments, for instance in the areas of biometrics and clickstream data.

### *7.3.2.2 Recoding & Analysis*

For the remainder of the analysis, the 10 scale data rating was re-coded into five categories (see Table 7.5).

Table 7.5 *Recoding of data sensitivity from 10 point scale into 5 categories*

| Original value | Recode value | Category Label |
|---|---|---|
| 1, 2 | 1 | Not at all Sensitive |
| 3, 4 | 2 | A little Sensitive |
| 5, 6 | 3 | Sensitive |
| 7, 8 | 4 | Very Sensitive |
| 9, 10 | 5 | Extremely Sensitive |

The data was analysed and is displayed in tables according to whether it is classified as legally recognised or not legally recognised as a category of sensitive data.

Table 7.6: *Sensitivity of legally recognised data types - ICO Survey*

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| **Trade Union Membership** | **1.4%** | **21.6%** | **13.9%** | **30.6%** | **15.1%** | **17.4%** | **100 %** |
| **Religious or Philosophical beliefs** | **0.9%** | **21.4%** | **12.1%** | **28.2%** | **13.3%** | **24.0%** | **100 %** |
| **Political Opinions** | **0.9%** | **15.9%** | **13.1%** | **28.1%** | **17.4%** | **24.5%** | **100 %** |
| **Data concerning race or ethnic origin** | **1.3%** | **19.5%** | **11.2%** | **26.8%** | **16.6%** | **24.6%** | **100 %** |
| **Criminal records** | **1.1%** | **11.2%** | **7.2%** | **22.9%** | **17.5%** | **40.1%** | **100 %** |
| **Sexual life Information** | **1.6%** | **6.8%** | **6.7%** | **18.0%** | **17.1%** | **49.8%** | **100 %** |
| **Health information** | **0.9%** | **3.8%** | **5.1%** | **18.2%** | **20.6%** | **51.3%** | **100 %** |

(Source: ICO Annual Track Survey 2006) (n=1066)

Table 7.6 displays the legally recognised categories of sensitive data and indicates that Health data was considered extremely sensitive by over half of the respondents (51.3%), and almost half considered sexual life information to be extremely sensitive, whereas fewer respondents considered religious or philosophical beliefs to be extremely sensitive (24%) and only 17.4% considered trade union membership data to be extremely sensitive. Thus, some of the legally recognised categories of sensitive data are considered less sensitive than others.

Table 7.7 *Sensitivity of not legally recognised data types- ICO Survey*

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| **Employment History** | **1.1%** | **15.9%** | **12.1%** | **30.2%** | **19.3%** | **21.3%** | **100%** |
| **Education Qualifications** | **1.4%** | **15.9%** | **11.7%** | **29.5%** | **19.5%** | **22.0%** | **100%** |
| **Membership of political party /organisation** | **1.5%** | **17.4%** | **13.4%** | **30.1%** | **15.5%** | **22.0%** | **100%** |
| **Clickstream data** | **2.5%** | **15.9%** | **11.4%** | **27.5%** | **18.2%** | **24.4%** | **100%** |
| **Personal Contact Details** | **0.4%** | **7.0%** | **7.1%** | **17.8%** | **21.4%** | **46.2%** | **100%** |
| **Genetic Information** | **1.6%** | **8.7%** | **6.3%** | **20.8%** | **19.2%** | **43.3%** | **100%** |
| **Biometric Information** | **2.2%** | **10.4%** | **6.8%** | **17.5%** | **17.6%** | **45.4%** | **100%** |
| **Financial data** | **0.6%** | **1.9%** | **2.5%** | **7.1%** | **17.4%** | **70.5%** | **100%** |

(Source: ICO Annual Track Survey 2006) (n=1066)

Table 7.7 displays categories of sensitive data that are not legally recognised. The table indicates that financial data was considered extremely sensitive by over 70% of respondents, and just under half (46.4%) considered their personal contact details extremely sensitive, whereas only 21.3% of respondents considered employment history data to be extremely sensitive. The finding from the survey indicates that one fifth of telephone respondents considered trade union membership, religious/philosophical beliefs or data concerning racial/ethnic origin to be not at all sensitive.

### *7.3.3 Findings from online survey of bloggers*

The views of bloggers from around the world regarding the concept of sensitive data were sought through an online survey. The research question was modified slightly, to reflect the fact that potential respondents may not work or reside in the EU and thus be unfamiliar with the 'sensitivity' classification of certain types of personal data in Directive 95/46/EC.[401]

Table 7.8 *Sensitivity ratings of legally recognised data types* - All Bloggers

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| Trade Union Membership | 12.7% | 28.5% | 16.7% | 24.0% | 11.8% | 6.4% | 100% |
| Religious or Philosophical beliefs | 12.1% | 24.1% | 16.9% | 22.6% | 15.3% | 9.1% | 100% |
| Political Opinions | 12.2% | 21.6% | 16.8% | 26.4% | 14.5% | 8.5% | 100% |
| Data concerning race or ethnic origin | 12.4% | 23.0% | 15.7% | 23.4% | 16.4% | 9.1% | 100% |
| Criminal records | 12.6% | 9.1% | 9.1% | 20.9% | 22.7% | 25.6% | 100% |
| Sexual life Information | 12.4% | 4.1% | 4.0% | 9.9% | 19.4% | 50.3% | 100% |
| Health information | 12.6% | 4.4% | 4.1% | 12.8% | 20.6% | 45.5% | 100% |

(Source: Blog Survey 2006) (n=1258)

[401] *Online Survey question:* In Europe some types of personal information are considered 'sensitive' and given extra protection in law. Please read the list below and indicate on a scale of 1 to 10, how sensitive you consider each one to be. 1 means not at all sensitive and 10 is extremely sensitive.

Table 7.8 displays the legally recognised categories of sensitive data and indicates that sexual life data was considered extremely sensitive by over half of the respondents (50.3%), and 45.5% considered health information to be extremely sensitive, whereas fewer respondents considered religious or philosophical beliefs to be extremely sensitive (9.1%) and only 6.4% considered trade union membership data to be extremely sensitive. Thus, some of the legally recognised categories of sensitive data are considered less sensitive than others.

Table 7.9 *Sensitivity ratings of not legally recognised data types* - All Bloggers

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| **Employment History** | 12.3% | 22.3% | 16.1% | 26.2% | 15.5% | 7.6% | **100%** |
| **Education Qualifications** | 12.4% | 22.7% | 17.1% | 25.5% | 15.2% | 7.1% | **100%** |
| **Membership of political party /organisation** | 12.4% | 22.8% | 16.5% | 25.6% | 14.2% | 8.4% | **100%** |
| **Clickstream data** | 12.2% | 7.9% | 11.0% | 18.4% | 24.7% | 25.8% | **100%** |
| **Personal Contact Details** | 12.2% | 3.2% | 4.5% | 10.6% | 22.0% | 47.5% | **100%** |
| **Genetic Information** | 12.7% | 7.2% | 5.2% | 11.0% | 16.0% | 47.9% | **100%** |
| **Biometric Information** | 12.9% | 3.1% | 4.1% | 6.9% | 12.9% | 60.2% | **100%** |
| **Financial data** | 12.2% | 1.6% | 1.8% | 3.7% | 13.5% | 67.2% | **100%** |

(Source: Blog Survey 2006) (n=1258)

Table 7.9 indicates that financial data was considered extremely sensitive by 67.2% of respondents, and almost two-thirds considered Biometric data extremely sensitive, whilst (47.5%) considered their personal contact details extremely sensitive, whereas only 7.2% of respondents considered education qualifications to be extremely sensitive. This table indicates that some categories of sensitive data that are not legally recognised are considered more sensitive than others.

### *7.3.3.1 Online survey of UK Bloggers*

Of the 1258 blogger respondents 497 were from the UK. Tables 7.10 and 7.11 report the sensitivity perceptions of UK bloggers who responded to the survey.

Table 7.10 *Sensitivity ratings of legally recognised data types* – UK Bloggers

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| Trade Union Membership | 1.8% | 28.4% | 17.9% | 30.0% | 14.9% | 7.0% | 100% |
| Religious or Philosophical beliefs | 1.2% | 28.2% | 18.1% | 27.6% | 14.1% | 10.9% | 100% |
| Political Opinions | 1.2% | 22.1% | 19.5% | 31.0% | 16.3% | 9.9% | 100% |
| Data concerning race or ethnic origin | 1.4% | 26.0% | 18.1% | 26.6% | 18.3% | 9.7% | 100% |
| Criminal records | 1.6% | 9.3% | 11.5% | 20.7% | 26.4% | 30.6% | 100% |
| Sexual life Information | 1.6% | 4.0% | 4.4% | 9.7% | 24.3% | 55.9% | 100% |
| Health information | 1.8% | 5.2% | 4.4% | 13.9% | 26.8% | 47.9% | 100% |

(Source: Blog Survey 2006) (n=497)

Table 7.10 displays the legally recognised categories of sensitive data and indicates that Health data was considered extremely sensitive by almost half of the respondents (47.9%), and over half considered sexual life information to be extremely sensitive, whereas fewer respondents considered religious or philosophical beliefs to be extremely sensitive (10.9%) and only 7.0% considered trade union membership data

to be extremely sensitive. Thus, some of the legally recognised categories of sensitive data are considered less sensitive than others. Interestingly, approximately 10% fewer bloggers rated trade union membership, religious or philosophical beliefs, political opinions, data concerning race or ethnic origin and criminal records extremely sensitive than ICO respondents (Table 7.6).

Table 7.11 *Sensitivity ratings of not legally recognised data types* – UK Bloggers

| | Don't Know | Not at All Sensitive | A Little Sensitive | Sensitive | Very Sensitive | Extremely Sensitive | Total |
|---|---|---|---|---|---|---|---|
| **Employment History** | **1.6%** | **26.4%** | **16.1%** | **30.2%** | **18.5%** | **7.2%** | **100%** |
| **Education Qualifications** | **1.6%** | **25.6%** | **18.1%** | **28.8%** | **17.9%** | **8.0%** | **100%** |
| **Membership of political party /organisation** | **1.4%** | **24.9%** | **18.1%** | **29.6%** | **14.9%** | **11.1%** | **100%** |
| **Clickstream data** | **1.2%** | **9.5%** | **13.1%** | **17.5%** | **28.8%** | **30.0%** | **100%** |
| **Personal Contact Details** | **1.2%** | **3.6%** | **3.8%** | **13.5%** | **25.8%** | **52.1%** | **100%** |
| **Genetic Information** | **2.0%** | **7.6%** | **6.2%** | **12.5%** | **21.5%** | **50.1%** | **100%** |
| **Biometric Information** | **2.0%** | **3.6%** | **4.8%** | **8.7%** | **15.5%** | **65.4%** | **100%** |
| **Financial data** | **1.4%** | **1.6%** | **2.8%** | **3.6%** | **15.3%** | **75.3%** | **100%** |

(Source: Blog Survey 2006) (n=1258)

Table 7.11 indicates that financial data was considered extremely sensitive by over 75.3% of respondents, and just over half (52.1%) considered their personal contact details extremely sensitive, whereas only 7.2% of respondents considered employment history data to be extremely sensitive. This table indicates that some categories of sensitive data that are not legally recognised are considered more sensitive than others.

### *7.3.3.2 Multiple Regression Model*

A multiple regression model was produced to test for any significant differences in the sensitivity perceptions between the blog survey respondents and telephone survey respondents. The regression analysis was run on a merged dataset generated by combining the ICO telephone survey respondents with the UK blog survey respondents, after under 18year olds were excluded from the blog survey, as all respondents in the telephone survey were over this age.

Table 7.12 *Reference categories:*

| **Variable** | **Reference Category** |
|---|---|
| Gender | Female |
| Age Band | 18-24yrs |
| Relationship Status | Married |
| Child | Do Not Have Child |
| Status Working | Yes, Working |
| Main Earner | Yes, Main Earner |
| File | ICO Survey |

The model examined the residual effects of being a respondent in the blog survey as opposed to being a respondent in the telephone survey, once socio-demographic factors had been accounted for. By examining the residual effects, comparisons could be made between the sensitivity perceptions of blog respondents and telephone survey respondents regarding the sensitivity of different data types. This in effect allowed the assessment of the differences between bloggers and the general population

Table 7.13 *Multiple regression of data sensitivity*

| **Dependent Variable** | Unstandardized Coefficients | | | |
|---|---|---|---|---|
| | **B** | Std. Error | t | **Sig.** |
| **Personal Contact Details** | .462 | .175 | 2.642 | .008** |
| **Financial data** | .095 | .135 | .706 | .480 |
| **Race or Ethnicity** | -.562 | .197 | 2.848 | **.004**** |
| **Criminal records** | -.020 | .199 | .102 | .919 |
| **Biometric Information** | 1.188 | .198 | 5.998 | **.000***** |
| **Political Opinions** | -.620 | .193 | 3.217 | **.001**** |
| **Membership of Political party or Organisation** | -.437 | .194 | 2.255 | .024* |
| **Clickstream data** | .624 | .200 | 3.124 | .002** |
| **Religious or Philosophical Beliefs** | -.551 | .199 | 2.764 | .006** |
| **Genetic Information** | .360 | .195 | 1.841 | .066 |
| **Health Information** | .022 | .172 | .126 | .899 |
| **Sex Life Information** | .540 | .186 | 2.911 | .004** |
| **Education Qualifications** | -1.152 | .188 | 6.133 | **.000***** |
| **Employment History** | -1.099 | .186 | 5.916 | **.000***** |
| **Trade Union Membership** | -.652 | .189 | 3.450 | .001** |
| | | | | |
| **Mean** | -0.120 | | | |

* $P$ <0.05. ** $P$ <0.01. *** $P$ <0.001.

The Multiple regression results (Table 7.13) show that significant differences were found in the sensitivity perceptions of blog and telephone respondents in relation to twelve of the fifteen categories of sensitive data. Bloggers indicated higher sensitivity regarding Sex life information, Clickstream data, Biometric information, and financial data. The mean score (-0.120) indicates that blog respondents had slightly lower perceptions of sensitivity than telephone survey respondents. Indeed, blog respondents were less sensitive about race or ethnicity data, criminal records, political opinions, membership of political party or organization, religious or philosophical beliefs, education qualifications, employment history and trade union membership data.

No significant differences were found in relation to health information, genetic information and financial data.

## *7.4 Relationship between sensitive and private data*

Chapter four indicated that there is a lack of empirical data on whether the terms 'sensitive' and 'private' data are synonymous. This research seeks to remedy this deficiency by reporting the findings of interviews with privacy and data protection experts and an online survey of bloggers.[402]

### *7.4.1 Confluence of sensitive and private data: views of data protection and privacy experts*

The privacy and data protection experts offered a range of responses. A few respondents indicated that the terms private and sensitive are synonymous:

> "Yes, I think private is synonymous with sensitive data." (UK)
>
> "Sensitive has two connotations: private and harmful." (UK)
>
> However, some respondents commented that the current categories of sensitive data do not completely capture the concept of private data:
>
> "Sensitive data e.g. income mostly covers private data. However not all private data is sensitive e.g. religion in Australia is not considered sensitive, as we do not have religious persecution in our recent history." (Australia)
>
> "The fact I had flu last week is less sensitive or private than how much pay I take home. Yet, my health data is sensitive but my salary is not, even though I am more concerned about the privacy of my salary." (Iceland)
>
> "Private is defined by the individual whereas sensitive is defined by the polity e.g. legislative process. Children's data is sensitive in the US. It depends on the country you live in. In the US, political party membership/registration is not sensitive e.g. publicly available electoral roll– but it can be an intimidating piece of information." (USA)
>
> "Yes. Many sensitive data may be regarded as private data. The problem is that not all data in each category of sensitive data may be under regime of private data. For example, the category of "medical data;" some of them shall be disclosed to competent institutions regardless of wish of data subject. Others, one may be strictly keep in full decision of data subject." (Czech Republic)

The responses indicate that classification of data as 'sensitive' is not a necessary precondition for classifying information as 'private.' Rather, the responses indicate that data is classified as private if is poses a risk of harm to an individual e.g. if an

[402] Data collection methods are discussed in chapter 5

individual's data is misused they could suffer identity fraud, financial harm, or a personal or proprietary security risk.

#### *7.4.1.1 Utility of distinction between sensitive and private data in legislation*

Respondents were asked whether it would be useful if legislation drew a distinction between 'sensitive' and 'private data':

> "Not sure. Can't think of a situation where I would need a distinction." (Italy)

> "We should have differentiations in the definition of sensitive data e.g. some health data is less sensitive than financial data." (Iceland)

> "I remain to be convinced that we should have more subcategories. We have a broad definition of personal information, then privacy principles, then what is classified as sensitive – it is harder to collect and disclose sensitive information - but we don't have different tests for storage and usage." (Australia)

> "I suppose to some extent the need for sensitive data is reduced if you have a concept of private data. What do you include in the concept of sensitive data is a big debate in the literature – particularly whether it should be a list or based on circumstances. Enhanced protection is a good idea, but it is difficult to know what data are private or sensitive." (Belgium)

The responses indicate that the term sensitive does not always encapsulate the range of privacy concerns associated with a particular type of data. For instance, in law, all health data is considered equally sensitive, but the disclosure of the fact that an individual has the flu virus has less serious discriminatory/harm implications than the disclosure of the fact that an individual has the HIV virus. Secondly, it fails to recognise the effect of contextual factors. For instance, in certain circumstances e.g. health, employment or relationship settings an individual may not have privacy concerns when revealing that they have the HIV virus.

#### *7.4.1.2 Risk based approach to private data*

Instead of distinguishing between sensitive and private data, respondents indicated that a distinction should be drawn between information which has harmful privacy implications if misused, and information which does not pose a privacy risk.

> "Sensitive data is private data which would make an individual uncomfortable if disclosed. Some sensitive data might not be private. Some might be sensitive about age/income category but might not think it harmful if disclosed (the 3rd category of individuals – who participate and reveal information but who are wary of doing so)

> Sensitive data is defined in the legislation – but the definition is quite unhelpful. It depends on the circumstances e.g. you might be happy to reveal political information to a political organisation." (UK)
>
> "It depends what you mean by private – something about a person that requires more protection, but that isn't sensitive. The concept is not bad as we need more differentiation in order to tell priorities for enforcement e.g. protection of names and addresses rather than IP address. "(Belgium)

The responses indicate that the focus of legislative attention should not be on the classification of data into specific categories e.g. personal, sensitive or private. Rather, legislation should be cognisant of the tension between facilitating information flows for legitimate purposes and preventing privacy harm when data is misused. These responses indicate that the focus of legislation should shift towards preventing misuse of data so that potential privacy harms are minimised.

### *7.4.2 Views of bloggers*

Bloggers were not asked a direct question about the relationship between sensitive and private data. Instead, they were asked to explain what private information meant to them. Table 7.14 indicates that some of the types of data considered private overlap with legally recognised categories of sensitive information, whilst other types of information e.g. contact details, biometric information are not encapsulated by the term sensitive data.

Table 7.14 *Classification of private data types by Bloggers*

| Art 8 Legally recognised categories of sensitive data | Private categories synonymous with sensitive data (Bloggers) | Private categories distinct from sensitive data (Bloggers) |
|---|---|---|
| *Political opinions* | *Political beliefs* | *Contact details e.g. my address, family name, names and addresses of family and friends.* |
| *Health information* | *Health e.g. medical records, blood type* | *Biometric information* |
| *Trade-union membership* | | *Financial data* |
| *Religious or philosophical beliefs* | | *Product plans* |
| *Trade-union membership* | | *Trade secrets* |
| *Data concerning race or ethnic origin* | | *Passwords and PINs* |
| *Criminal records* | | *Employer's name* |
| *Sexual life information* | | |

The responses mirror the responses of the privacy and data protection experts in that the bloggers classified data as 'private' on the basis of risk of harm e.g. identity fraud if contact details were revealed online. The responses reveal that bloggers are aware

that disclosure of information through blog posts can lead to a variety of privacy risks e.g. identity fraud, personal security, financial risk, health, religious or political discrimination, as well as an unemployment risk if information were disclosed to employers. The blogger responses mirror the responses of the experts in that they emphasise the risk of harm from information misuse.

## *7.5 Discussion of findings*

The findings suggest that the current list needs to be reformed, as *prima facie* it doesn't reflect the sensitivity perceptions of data subjects. Moreover, the findings suggest that new categories of sensitive data are emerging due to changes in society and technological developments. For instance, amidst post-September 2001 security concerns the UK government proposed the introduction of passports which rely on biometrics.[403] Such technology did not exist during the World-War II era when the UK previously utilised personal identifiers in the form of identity cards, and indeed identity cards were removed from circulation in 1952 amid widespread public resentment.[404] This raises the issue of whether the current list of sensitive data could or should be amended.

## *7.6 Criticisms of sensitivity classification*

Korrf[405] conducted a comparative textual analysis of legislation. He found that the French, Austrian, British, Czech, Estonian, Finnish, Greek, Hungarian, Italian, Spanish, and Swiss laws state that the list their legislation contains is exhaustive, whilst some countries (for instance, Denmark and Iceland) consider their lists as merely indicative. Nevertheless, all laws provide ways and means to reopen the apparently closed list, so *prima facie* the list of sensitive data categories could be amended.[406]

[403] Home Office (2005) "UK Passport Service: Improving Passport Security and Tackling ID Fraud," Reference: UKPS001/2005 (Press Release: 24th Mar 2005). In 2006 ePassports were introduced. They feature an electronic biometric chip which contains the holder's personal details. This security measure meets the demands of the US visa waiver programme. These biometric e-passports are also embedded with microchips that automatically flash their data onto border security officers' screens.

[404] *Willcock v Muckle*, [1951] 2 The Times LR 373 The judge in the case said that the cards were an "annoyance" and "tended to turn law-abiding subjects into law breakers".

[405] Korrf, D. (2002) EC Study On Implementation Of Data Protection Directive (Study Contract ETD/2001/B5-3001/A/49)

[406] For instance, the Estonian Act states that the list can be modified by law. In the UK the Westminster Parliament could update the list by amending the Data Protection Act 1998

However, creating new categories could give rise to difficulties, for instance, Luxembourg, and the Netherlands define genetic data as data on *health*, whilst Portugal defines it as data on *health and sex life*, whereas, in Sweden the processing of such data not formally regarded as falling within the specific category to which the rules on "*sensitive data*" apply.

Likewise, Hungary and Poland have included "details of addictions" in their list of sensitive data. Many addictions would clearly fall within the health related category set out in the Directive already: for example drug addiction and alcoholism. Others, such as gambling or computer games, might not. It remains to be seen how regulators will interpret this additional restriction.[407] Thus, the first problem associated with amending the current categories of sensitive data is that any attempts to modify or extend the current list would require transnational agreement otherwise a lack of harmonization will occur, and defeat the objective of the Directive.

A second, related problem is that a definition-based approach would require a casuistic form of regulation, which is more complex and lengthy to administer. Indeed, Bing[408] attempted to categorise all personal data according to their sensitivity. However, this approach was quickly abandoned, because it failed to delineate clearly the boundaries of the various spheres of an individual's life, why these exist and what might constitute a breach of these.

Moreover, Simitis[409] asserts that classification of data as sensitive or not, does not address the privacy problem. For instance, detailed profiles could be created through the aggregation of ostensibly innocuous personal information, which could nevertheless have a detrimental impact upon a person's privacy. Thus, interviewees raised the importance of extraneous information, rather than simply relying on a definitional approach to sensitive data. For instance:

> "I've never made much use of the concept, e.g. your postcode and newspaper preference both appear to be innocuous information. However, if you work for Experian (a credit score, credit report and credit reference agency), you can draw inferences about a person simply based on those two prices of information – that settles the point. How can you define what is sensitive?

[407] Linklaters (2004) Hot Topic: History repeats itself: the implementation of EU data protection legislation in the accession countries. http://www.linklaters.com/pdfs/briefings/040517_DP.pdf
[408] Bing, J. (1972) "Classification of Personal Information with respect to the Sensitivity Aspect" Proceedings of the First International Oslo Symposium on Data Banks and Society, 98-141
[409] Simitis, S. (1973) cited in Bygrave, L. (2002) *Data Protection Law: Approaching its Rationale, Logic and Limits*, Kluwer, 132

> E.g. if you can work out my political views from my newspaper preference, then arguably my postcode and newspaper preference should be considered sensitive information." (UK)
>
> "Definitions of sensitive data are very subjective e.g. where you live is sensitive if you have an estranged violent husband." (UK)

Likewise, another respondent opined:

> "I don't like the idea of sensitive data. All data is potentially sensitive, depending upon the context." (UK)

Accordingly, Simitis contends that personal data becomes sensitive according to its context. This mirrors the approach formerly adopted by countries such as Austria and Germany, which, prior to the introduction of the data protection directive had consistently rejected all abstract categorisations of personal data and instead focussed on a context-orientated consideration of the data. He asserts that:

> "Sensitivity is no more perceived as an *a priori* given attribute. On the contrary, any personal datum can, depending on the purpose or the circumstances of the processing be sensitive."[410]

This approach reflects the opinions of some of the interviewees, for instance:

> "Another example is related to the employment code we have drafted. Health is regarded as sensitive data. All employers keep records of sickness leave, but the issue is: does self-certified sick notes require the same level of protection as a medical note from a GP? Arguably a self-certified note is less sensitive, particularly given that the individual may have told colleagues the reason for their absence…yet no distinction is drawn in the law – but we would advise employers that they should take a common sense approach." (UK)
>
> "The idea that all health information is sensitive is too restrictive in some instances e.g. it can cause difficulties between two contracting parties such as an insurance company and an individual. We need safeguards to protect sensitive uses of sensitive data." (Spain)

Simitis reasoned that it is vital to consider contextual information when determining the sensitivity of data. Contextual information includes: the interests of the data controller as well as the potential recipients of the data, the aims for which the data are collected, the conditions of the processing and its possible consequences for the individual and others. An evaluation of the sensitivity requires hence more than a

[410] Simitis, S. (1999) "Revisiting Sensitive data"< http://www.coe.int/t/e/legal_affairs/legal_co-operation/data_protection/documents/reports%20and%20studies%20by%20experts/1Z-Report_Simitis_1999.asp>

definitional approach to sensitive data. Furthermore, Simitis advocates that sensitivity lists should be purely exemplary, and:

> "Only where the legislators can fully concentrate on a specific context, are they also able to reach a degree of precision that appropriately responds to the particularities of the processing circumstances."[411]

Moreover, Wacks[412] asserts that what changes from context to context is not the degree of sensitivity of information, but the extent to which one is prepared or required to allow it to be disclosed or used. Should the context change, it is not the nature of the information that changes, but an individual's *attitude* towards its use. An individual is likely to have considerably different views about the purposes for which sensitive data is used, for instance, an expert interviewee responded:

> "I think it will be extremely important to regulate who can access what information and for what reason e.g. whilst it may be acceptable to allow police to deduce racial information through DNA profiling it would not be appropriate to allow a security guard to have access to this type of information when simply determining if an individual should have permission to enter a building." (UK)

## *7.7 Summary*

The findings suggest that the time is ripe to reconsider the inclusion of categories of sensitive data in data protection laws since the expert interviewee responses and the findings from both surveys indicate that whilst not all of the legally recognised categories of data continue to be perceived as sensitive, some not legally recognised categories of data are emerging which are considered extremely sensitive. However, a decision to simply include new categories, or delete existing categories should not be taken lightly. Any attempt to grade data according to their sensitivity would be fraught with difficulties, as it would require a casuistic form of regulation, which is more complex and lengthy. Indeed, the findings indicate that a more radical approach should be adopted; one which recognises that the concept of sensitivity is outdated. An expert interviewee opined that:

> "The concept of sensitive is a failed attempt to capture something, which isn't a natural kind. By saying something is sensitive you are attempting to treat something to do with claim for making different reasons in a single manner.

[411] Simitis, S. (1999)"Revisiting Sensitive data"< http://www.coe.int/t/e/legal_affairs/legal_co-operation/data_protection/documents/reports%20and%20studies%20by%20experts/1Z-Report_Simitis_1999.asp>

[412] Wacks, R. (1989) *Personal Information: Privacy and the Law*, (Oxford: Clarendon Press), 23, 181

> Whereas, life is not reducible to a single algorithm – so you should be wary of this approach." (UK)

Therefore, instead of trying to resolve the sensitivity conundrum, it would be prudent to consider the proposed definitions of data privacy and private data outlined in chapter two, which advocate a harm-based approach to data protection. These alternatives are explored in the next chapter.

# Chapter 8

## A harm-based definition of 'private' data

### *8.1 Introduction*

An underpinning principle of all data protection laws is the protection of privacy. Yet, as outlined in chapter three, none provide a conclusive understanding of the terms 'privacy' or 'private' data. Rather, privacy protection is to be achieved through the regulation of the conditions under which personal and sensitive data may be processed. Following the findings in chapter 7 (critique of sensitive data); this chapter investigates whether the term personal is synonymous with the term private data, and determines that the terms are not synonymous. Thereafter, it tests the merits of proposed definitions of data privacy and private data, outlined in chapter two. The chapter concludes that a review of the assumptions and concepts underpinning the current data protection laws is necessary, and that future legislation should focus on unreasonable harm caused by the misuse of private data.

#### *8.1.1 Aims*

This chapter has two main aims. Firstly, it seeks to explore whether the concept of 'private' data is synonymous with the term 'personal' which is legally recognized and protected in data protection laws. Secondly, this chapter seeks to contribute to the debate surrounding calls by regulators for reform of Directive 95/46EC[413] and calls by industry experts for the introduction of a global privacy framework,[414] by testing a proposed definition of 'private' data which is harm-based in nature. In other words, the chapter aims to cover three key areas of enquiry: questions about interpretation and application of the term personal data; questions about the

[413] European Commission, (2009) "Review of the data protection legal framework," <http://ec.europa.eu/justice/policies/privacy/review/index_en.htm> (Last accessed: 20.05.11)
[414] Fleischer, P. (2007) "Call for Global Privacy Standards," Google Public Policy Blog, 14th Sept 2007 <http://googlepublicpolicy.blogspot.com/2007/09/call-for-global-privacy-standards.html> (Last accessed 20.02.11)

relationship between the terms personal and private data; and questions about the merits of a harm-based definition of private data.

### *8.1.2 Approach*

A two-fold approach was required to answer the questions: an analysis of how the term 'private' data is understood by data protection experts and potential data subjects; and an evaluation of the utility of a harm-based legislative approach. This was achieved through a four-phase study: Phase 1- analysis of legislation and case law; Phase 2 - semi-structured interviews with privacy and data protection experts; Phase 3 – Internet survey of bloggers and Phase 4 - testing a harm-based definition of private data.[415]

## *8.2 Relationship between personal and private data: views of experts*

Analysis of the term personal data in chapter two indicated that the term has, for the most part, been broadly defined and widely interpreted. The discussion also revealed that the term personal data was interpreted narrowly in the case of *Durant v FSA*.[416] The narrow interpretation was criticised by academics,[417] but there was a lack of empirical evidence to support either approach. This research seeks to remedy that deficiency by collecting and analysing data on the terms personal and private data.

### *8.2.1 Confluence of personal and private data*

Semi-structured interviews were conducted with forty privacy and data protection experts. The interview questions sought to explore whether the terms private and personal are synonymous, or whether private data could be conceived of as a subset of personal data. When asked if the terms private and personal are synonymous a few respondents stated that they are not. They used the term 'private' to denote information that is not governed by data protection legislation:

> "No the terms are not synonymous; we use the word private to mean that it is outside the scope of the Data Protection legislation." (Germany)

> "We don't have a concept of private except with regards to the 1983 legislation concerning protection of reputation or honour from interference." (Spain)

---

[415] The data collection methods were discussed in greater detail in chapter 5.

[416] *Michael John Durant v Financial Services Authority* [2003] EWCA Civ 1746, Court of Appeal (Civil Division) decision of Lord Justices Auld, Mummery and Buxton dated 8th December 2003. <http://www.bailii.org/ew/cases/EWCA/Civ/2003/1746.html> (Last accessed: 27.05.11)

[417] Lindsay, D. (2004) "Misunderstanding 'personal information': Durant v Financial Services Authority," *Privacy Law & Policy Reporter*, Vol. 10, No. 10, p. 13

However, some respondents indicated that the terms are synonymous on an informal, non-legal basis:

> "Yes, but not on a legal basis." (Canada)

> "In our law the word private isn't even used, so it doesn't have a legal meaning. The general population take them to mean the same thing."(Australia)

> "I think in a legal sense – in a data protection sense, yes. However, privacy and data protection are different, but in a colloquial sense they are synonymous."(Belgium)

Others contended that the terms private and personal are not synonymous. Such respondents frequently preferred to draw a distinction between 'private' information that is not publicly available and 'personal' information which is readily observable and in the public domain:

> "No, they're not [synonymous]. My address/height/hair colour is personal but not private, whereas my sexual preferences are personal and private." (USA)

> "Private data is the part of the personal data that the respondent does not want to make public." (Spain)

> "I would draw a distinction between private and public data." (Italy)

Thus, expert respondents drew a distinction between the terms personal and private data. They favoured a broad interpretation of the term personal, that is information relating to a person e.g. address. Interestingly, one respondent asserted that the terms were not synonymous, since the label private was one used by data subjects on an *ex post facto* basis when seeking to restrict access to particular data:

> "They are not synonymous. Private is an *ex post facto* term used mainly to label those claims for non-disclosure that we've accepted on other contextual grounds. Whereas the term personal concerns information about which less contested claims are made e.g. the personal fact that I'm bald and short-sighted is personal but hardly a private fact." (UK)

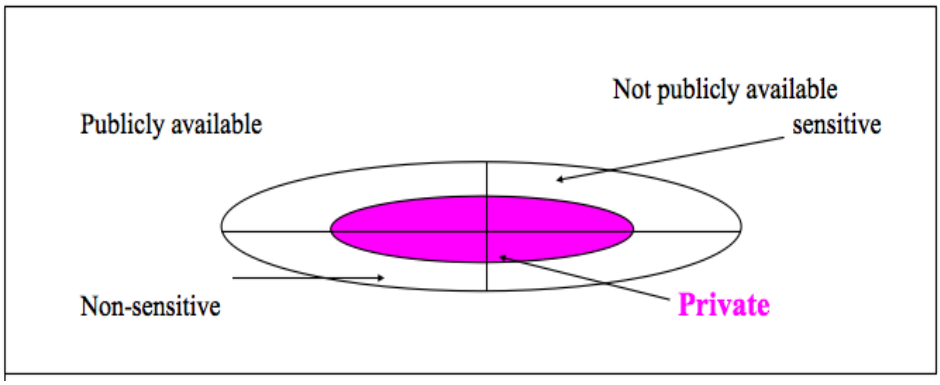
These respondents used the term 'private' to denote information that an individual would want to restrict access to, or disclosure of, in order to protect their privacy. This response mirrors the contextual integrity approach advocated by Nissenbaum,[418] since it states that privacy claims may be made on contextual grounds.

[418] Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, pp.101-139

### *8.2.2 Private data is a subset of personal data*

Some respondents stated that private data was not synonymous with personal data; instead they classified private data as a subset of personal information. One respondent drew a diagram to represent their conception of private data as a subset of personal data. (See Fig 8.1)

Fig 8.1 *Expert respondent diagram of private data as a subset of personal data*



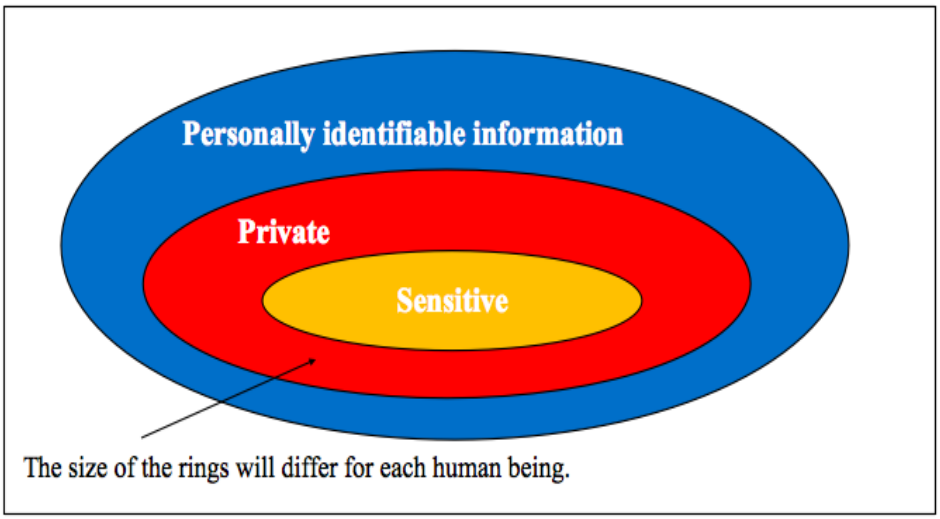

In Fig 8.1 the respondent illustrated that private data is totally enclosed by personally identifiable information (that is, there is no private data that is not also personally identifiable information). (Australia)

> "Personal relates to a person, an individual. Private is a subset of personal information. It is not used in our Act. It is something not revealed to others, or only revealed to a select group. It is a concept close to confidentiality but without the legal connotations e.g. disclosure to a family member or bank staff or personnel office e.g. my salary would be considered private." (Australia)

Thus, the respondent viewed private information as a subset of personal information, and stated that it is personal information which an individual would wish to exert greater control over.

Another respondent drew a different diagram to illustrate the relationship between personal and private data (See Fig 8.2):

Fig 8.2 *Expert respondent diagram of private data as a subset of personal data*




When discussing the diagram Fig 8.2, the expert respondent stated that it indicates that:

> "private data is a subset of personally identifiable information. Within that set we have a subset of sensitive data. Sensitive data is a subset, the misuse/invalid use/non-consensual use of which could cause unreasonable harm. It doesn't have to be disclosed to cause unreasonable harm e.g. talk about ethnicity. Lots of information is not sensitive. The problem is to define where the boundary should lie." (UK)

> "Private – something about a person that requires more protection, but that isn't sensitive. The concept is not bad as we need more differentiation in order to tell priorities for enforcement e.g. protection of names and addresses rather than IP address." (Belgium)

These respondents acknowledged that the concepts of personal and private data overlap, but are not synonymous. Rather, private data is a subset of personal data for which an individual seeks greater privacy protection.

### *8.2.3 Risk based approach to private data*

Some of the data protection and privacy experts asserted that some types of data are not classified as personal, but could be used in a privacy invasive way e.g. if collated with other information to build a profile on an individual:

***Profile data***

> "From the value of my property inferences can be drawn. Also, e.g. Ipod tunes – personal preferences. These examples are not personal information but they reveal private information about a person, as a profile can be built of an individual from such data." (USA)

"An example is car registration numbers. They are not personal data, but if combined with data such as car mileage and linked to car MOT details e.g. vehicle mileage and ownership details and used by a direct marketing firm to make statements e.g. your car has too many miles and is too old, you need to buy a new one…understandably the car owner may become upset and insulted. So, my answer is no, data does not have to be personal to be private." (Netherlands)

***Security data***

"Private data is possibly not always personal data – e.g. shared secret that you and I know that no one else knows (e.g. password) – it is not personal but it is private even though it is shared." (Australia)

***Family data***

"For example, information that describes life inside of family or a group of relatives." (Czech Republic)

"It is important to explore the concept of family data. In Asian cultures it is normal for all family members to know information about each other to a greater extent that you find in British culture. Their family knows everything about them. Family has a higher status than the individual. An example is a Pakistani Muslim immigrant in the UK who falls pregnant at 15yrs of age. Her family believe they have the right to know all the details surrounding the pregnancy e.g. from medical sources as the rights of the individual are subjugated to wider family." (UK)

"Information about my legal status and legal affairs is not normally considered personal, but it is often considered private, that is, it needn't be disclosed e.g. administrative fact that I'm British, but for some transactions e.g. if I want to buy a daily newspaper I would be offended if I were asked to confirm my nationality." (UK)
"Also, with the internet information on your neighbours- it is often published on the internet but it is still private as far as they are concerned." (Finland)

"Information about where you keep personal phone numbers."(Australia)

The responses appear to suggest that classification of data as personal is not a necessary precondition for classifying information as private. Rather, the responses indicate that data is classified as private if it poses a privacy risk to an individual e.g. if an individual's data is misused they could suffer identity fraud, financial harm, or a personal or proprietary security risk.

### *8.2.4 Utility of distinction between personal and private data in legislation*

Respondents were asked whether it would be useful if legislation drew a distinction between 'personal' and 'private' data. Some respondents indicated that a distinction would be useful:

> "Yes: private data is the part of the personal data that the respondent does not want to make public." (Spain)

> "It depends what you mean by private – something about a person that requires more protection, but that isn't sensitive. The concept is not bad as we need more differentiation in order to tell priorities for enforcement e.g. protection of names and addresses rather than IP address." (Belgium)

Others opined that, in theory it would be useful, but in reality the concepts overlap to the extent that they cannot be distinguished:

> "From a DP authority point of view it would be helpful if a distinction could be made, but I'm pessimistic because in different circumstances people see the same data differently, therefore it is very difficult to define this kind of data e.g. If we approach our bank manager for a loan, then we will be willing to discuss our salary but in other cases you won't tell someone your salary...One more example, if I may, is the question whether an employer should be allowed to read employee emails – our legislation is very good as it has a specific provision about this activity. However, we had a case where a prosecutor dealt with a case involving an employer and two employees whose emails were read. The prosecutor didn't prosecute for reading 1st employee's emails because she said that the messages were not very 'private.' I use this as an example when I give presentations to make people think about how do we know what is private? How did the prosecutor know if it was private or not? I never give verdicts by deciding if this is or is not very private. Instead, I take the approach – is this legal, based on consent and for a good purpose?" (Finland)

No, a distinction would not be useful:

> "In Germany in the 1970's we discussed if it is possible to differentiate between personal and private data – with a higher level of protection for private data. We decide it was impossible to differentiate, and so we decided that we have to protect all personal data." (Germany)

The responses indicate that the focus of legislative attention should not be on the classification of data into specific categories e.g. personal, sensitive or private. Rather, legislation should be cognisant of the tension between facilitating information flows for legitimate purposes and preventing privacy harms when data is misused. These responses indicate that the focus of legislation should shift towards preventing misuse of data so that potential privacy harms are minimised.

## *8.3 Views of bloggers*

The views of bloggers were sought through an online survey.[419] They were asked whether they posted personal information on their blogs, and further whether they classified information as too personal or private to post.

Table 8.1 *Classification of 'private' and 'too personal' data*

| Does **'private'** information **mean the same** to you **as** information which is **'too personal'** to write about in your blog? | **Percentage** |
|---|---|
| Yes | 40.7% |
| No | 44.5% |
| Prefer not to answer | 14.8% |
| Total | 100% |

(Blog Survey 2006, n = 1258)

Table 8.1 indicates that 44.5% of blogger respondents drew a distinction between personal and private data. Those who asserted that the terms 'too personal' and 'private' were not synonymous were asked to explain. Below are some illustrative comments:

### *a) Private relates to factual identifiers, whereas personal relates to emotions*

A number of bloggers used the term private when referring to factual information that they would not disclose on their blog:

> "Private information is my name, address, etc." (Female, 25-34, France)
>
> "Private information is rational, factual, numerical-esq information, like bank details or blood type, personal is more emotional. (Male, 19-24,UK)
>
> "I tend to think of 'private information' as more practical (i.e. facts and figures such as addresses, phone numbers, etc) and information which is 'too personal' as being to do with personal emotions."(Female, 25-34, UK)
>
> "To me, "too personal" implies emotional discomfort with revealing too much of oneself. "Private" information feels more like personal facts, e.g., earnings. But I think there is overlap between the two terms."(Female, 35-34, UK)
>
> "All information about my name, age, address, etc. is private. I am theoretically anonymous on the internet, and would consider obviously

[419] See chapter 5 for details of data collection methods. The privacy attitudes and expectations of bloggers were explored in chapter 6.

identifying information to be private. Too personal is usually too emotionally fraught."(Female, 19-24, Spain)

"I view private information more as data about me, rather than personal experiences which might be too personal to post about."(Female, 25-34, USA)

***b) Not disclosed to anyone***

Some bloggers indicated that 'private' data is information which is not disclosed to anyone:

"Private information is something you want to keep to yourself regardless of how "personal" you rate it. It is something that does not need to be shared with a wider audience."(Male, 19-24, UK)

"Private is not necessarily too personal. It's simply information that "belongs" to me, information that no one else needs to know. It's sometimes personal, but it's also sometimes about someone else."(Female, 25-44, USA)

"Private is just that it's for me only."(Male, 35-44,UK)

"Private: it's what I don't want others to know."(Female, 35-44, UK)

***c) Controlled disclosure to trusted parties e.g. friends***

For other bloggers, 'private' data is information that an individual may choose to disclose under controlled circumstances:

"Private info is info that I don't mind sharing, and in some cases, would like to share with friends who for example could help give advice and/or encouragement. Too personal information I see as nobody else's business but my own, maybe sharing with very close friends but not by the medium of a blog."(Female, 25-34, UK)

"Things 'too personal' are intended for a limited audience, while 'private' information will rarely every be shared with anyone."(Male, 19-24, Germany)

The responses reveal that some bloggers equated 'private' data with information that is only disclosed in the context of relationships based on trust.

***d) Risk based: Potentially harmful consequences of disclosure***

Bloggers also distinguished 'personal' and 'private' data on the basis of potential harms that could arise from the disclosure of information:

"Private information, when it comes to an on-line environment, refers to any data which a third party having knowledge of could cause me actual harm, whether financial, or by restrictions of civil liberties. This therefore refers to

my financial details, address and contact details (though to a lesser degree).”(Male, 25-34, UK)

A range of different types of harms/risks can be identified from the responses:

***Identity risk***

“'Private' is details of my life that could identify me or others, or could give away information that could be used against me. ‘Personal' is just stuff that should be kept between close friends and not shared with the world.”(Female, 25-34, UK)

***Physical location and safety***

“Private information to me means anything which could be used to determine my location from the internet, or other aspects that could compromise my safety if disclosed with others.”(Female, 19-24, UK)

“Private information is information which can be used to hurt me (for instance, my precise geographical location) or information about other people which they have deemed private and have shared with me under condition that I don't share it with other people.” (Female, 19-24, UK)

***Legal risk***

“Something that posting it could be used against me in a serious way (court etc)”

***Security risk***

“Private = things controlled personally e.g. Bank account details (logins); things where there would be a security risk if released Personal = things not everyone wants to hear about; things my mother would turn puce if she knew.”(Female, 25-34, UK)

***Employment risk***

“Information that could identify a patient or colleague - this could have far-reaching consequences for colleagues (if I was to criticise) or for myself (if I was recognised: "dooced"[420] would take on a whole new meaning if it meant I lost my place in medical school, or later, had my right to practice medicine revoked.”(Female, 19-24, Australia)

The responses reveal that bloggers are aware that disclosure of information through blog posts can lead to a variety of privacy risks e.g. identity fraud, personal security, financial risk, health, religious or political discrimination, as well as an unemployment risk if information were disclosed to employers. The blogger responses mirror the responses of the experts in that they emphasise the risk of harm from information misuse.

[420] As explained in Chapter 6, the term ‘dooced’ refers to an employee who is dismissed for blogging about their work. Lichtenstein, S. D., Darrow, J. (2006) “Employment Termination for Employee Blogging: Number One Tech Trend for 2005 and Beyond, or a Recipe for Getting Dooced?” *UCLA Journal of Law and Technology (JOLT),* Vol. 10, No.2, pp. 1-49

## *8.4 A harm-based alternative*

The responses to the questions exploring the relationship between private and personal data indicate that data protection laws are misconceived and misdirected. Their focus has erroneously been on the classification of data as such, instead of the harms arising from data misuses. These findings support observations by academics and experts. For instance, Bergkamp[421] contends that Directive 95/46/EC regulates at the wrong level and fails to balance competing interests properly, since it regulates the collection and processing of data upstream, as opposed to regulating specific harmful uses downstream:

> "Data Protection as currently conceived by the EU is a fallacy. It is a shotgun remedy against an incompletely conceptualised problem. It is an emotional, rather than rational reaction to feelings of discomfort with expanding data flows. The EU regime is not supported by any empirical data on privacy risks and demand… A future EU privacy program should focus on *actual harms and apply targeted remedies*"[422] (emphasis added)

Bergkamp's assessment appears to be gaining support. In November 2006, the UK ICO called for a review of the effectiveness of the work of each national privacy regulator

> "We must all prioritise, especially by reference to the *seriousness and likelihood of harm*. We must primarily concentrate on the main risks which individuals are now facing and be careful not to be excessively rigid or purist on issues which do not deserve it. We must be ready for more pragmatism and more flexibility."[423] (emphasis added)

Thereafter, in September 2007, Google's Global Privacy Counsel questioned how to "update privacy concepts for the Information Age" and called for the creation of "minimum standards of privacy protection that meet the expectations and demands of consumers, businesses and governments."[424] He rejected the US approach as too

---

[421] Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects of Europe's Data protection Policy in an Information Driven Economy" *Computer Law & Security Report*, Vol. 18 No.1 pp31-47, p. 42

[422] Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects of Europe's Data protection Policy in an Information Driven Economy" *Computer Law & Security Report*, Vol. 18 No.1 pp31-47, p. 31

[423] Turk, A. & Thomas, R. (2006) "Communicating Data Protection and Making it More Effective." The London Initiative, presented at 28th International Conference of Data Protection and Privacy Commissioners," <http://www.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Cooperation/Conference_int/06-11-03_London_initiative_EN.pdf>

[424] Fleischer, P. (2007) "The Need for Global Privacy Standards" <http://peterfleischer.blogspot.com/2007/09/eric-schmidt-on-global-privacy.html> (Last accessed: 20.03.09)

fragmented and the EU model as too bureaucratic. He suggested that "the most promising foundation" would be the Privacy Framework adopted by the members of the Asia-Pacific Economic Cooperation Forum (APEC). The APEC Privacy Framework contains nine information privacy principles which overlap to a large extent with those in the OECD Guidelines, the Council of Europe Convention and EU Directive. Uniquely, the APEC Framework contains a harm prevention principle. Thus, it is important to test whether the term harm should be included in any proposed revisions of Directive 95/46/EC or form the basis of any future global data protection framework.

### *8.4.1 The APEC Privacy Framework*

The harm prevention principle is found in Principle I:

> "Recognizing the interests of the individual to legitimate expectations of privacy, personal information protection should be designed to prevent the misuse of such information. Further, acknowledging the risk that harm may result from such misuse of personal information, specific obligations should take account of such risk, and remedial measures should be proportionate to the likelihood and severity of the harm threatened by the collection, use and transfer of personal information."[425]

Opinion is divided as to whether the APEC initiative is a positive development. Both Pounder[426] and Greenleaf[427] are critical of the elevation of a 'harm' test to the status of a principle. They fear that the statement in the Principle that 'specific obligations should take account of such risk [of harm].' will be interpreted as a threshold test for other rights or obligations. In contrast, Waters[428] opines that the commentary on this principle does not provide any support for the fear. He asserts that the statement about risk is no different from the 'such steps as are reasonable in the circumstances' qualifier found in many principles in most laws, which allow data users to base their compliance on their own risk assessment – subject to independent judgement in case of complaints or audits.

---

[425] APEC Privacy Framework (2005), Principle I <http://www.apec.org/apec/enewsletter/jan_vol7/onlinenewsd.primarycontentparagraph.0001.LinkURL.Download.ver5.1.9> (Last accessed: 20.03.09)

[426] Pounder, C. (2007) "Why the APEC Privacy Framework is unlikely to protect privacy." <http://www.out-law.com/default.aspx?page=8550 > (Last accessed: 20.03.09)

[427] Greenleaf, G (2009) "Five years of the APEC Privacy Framework: failure or Promise?" Computer Law & Security Review Vol. 25, pp 28-42, p. 30

[428] Waters, N (2008) "The APEC Asia-Pacific Privacy Initiative – a new route to effective data protection or a Trojan horse for self-regulation?" UNSWLRS, Paper 59

### *8.4.2 A proposed definition*

In order to test the utility of the APEC approach, a proposed definition of 'private' data was developed which is harm-based in nature. The definition does not claim to be a comprehensive alternative to existing legislative approaches. Rather, the purpose of the definition is to probe key concepts, so that a fuller appreciation of the conceptual difficulties can be distilled. The proposed definition in Figure 8.3 was tested during the course of the semi-structured interviews with the forty privacy and data protection experts.[429]

*Fig 8.3 Proposed definition of private data*

| |
|---|
| *"**Data privacy** concerns the legal regulation of the boundary between personal and private data. **Private data** is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual.*<br>*It is the legal right of an individual to withhold consent to the collection, processing, communication or usage of personal data, the disclosure of which could cause unreasonable harm to an individual. Disclosure of private data should not be compulsory except where it is in the public interest, for instance if disclosure is in the interests of national security, public safety, the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health, or for the protection of the rights and freedoms of others or society as a whole."* |

Rather than engage in an *a priori* classification of data as privacy sensitive or not, (which was shown to be ineffective in chapters four and seven) this definition seeks to draw upon Nissenbaum's[430] contextual integrity theory that the same information can be considered more or less private dependent upon contextual factors, and accordingly, these factors will influence an individual's decision to give consent. Consent is a feature of all data protection laws. For instance, in Directive 95/46/EC, Art 2 (h) states that the data subject should have given their 'unambiguous' consent to the processing of personal data, whilst Art 8 (2) (a) states that the data subject should give 'explicit' consent for the processing to be legitimate.[431] However, Austin asserts that:

> "an appeal to the value of control alone cannot determine why different kinds of information require different levels of protection"[432]

[429] Methodology discussed in detail in Chapter 5. The definition was supplied in advance of the interviews, so that respondents could consider it and formulate a response.
[430] Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, pp.101-139
[431] Exceptions to the general consent rule do exist – for instance, Art 8 (2) (c) processing is necessary to protect the vital interests of the data subject or of another person where the data subject is physically or legally incapable of giving his consent.
[432] Austin, L. (2003) "Privacy and the Question of Technology," *Law and Philosophy*, Vol. 22, No. 2, pp. 119-166, p. 128

Consequently, by coupling the term 'consent' with the term 'unreasonable harm,' this definition allows individuals to choose to give consent to the processing of personal data that does not pose a privacy risk, and withhold consent to the processing of data that could cause them harm. When privacy and data protection experts were asked to discuss the merits of this decision, some respondents favoured adopting a harm-based approach:

> "This is the kind of thing that is in our mind when we practice. It is the approach we take when it comes to a person whose has experienced social difficulties e.g. alcoholism. They (alcoholics or drug users) deserve protection where the disclosure could cause them harm." (Iceland)

In particular, one respondent commented that the current approach which seeks to prohibit the collection of sensitive data, unless it fits within one of the exceptions, e.g. the data subject has given explicit consent is misconceived, on the basis that:

> "Regulation of collection is a losing battle – instead *ensure it is not used malevolently* – information will always need to be collected so you need to focus work on how it is used." (Australia)

Thus, given that a contextual, ham-based definition appears to have some merit, it was appropriate to examine the concept of harm.

### *8.4.3 Harm*

The concept of 'harm' is not defined in Principle I of the APEC Framework, nor are any indications given regarding the type of harm that is contemplated. Accordingly, some expert respondents queried what is meant by harm, and whether it should be included in the definition:

> "What is harm? US business interests – harm to be limited to measurable (proportionality of harm), compensatable harm. In my opinion the definition needs to be softer e.g. non-compensatable loss of reputation." (Australia)

> "Harm – this raises the issue of whether it should be subjective or objective test of harm? Also, should the harm be foreseeable? ... I'm biased towards a subjective test of harm, but this doesn't solve the problem of foreseeability." (UK)

> "In an English common law environment, we view harm in terms of physical /property characteristics; although in the last 30 yrs we've developed an award for pain, suffering and emotional distress – this is a relatively recent development. Historically, we tended to look for physical harm. However, privacy must cover notions of offence being caused – we're not good at applying those notions e.g. how humiliated I am if you 'out' me as being gay? Am I more humiliated if you 'out' me for being gay than if you reveal that I have been cheating on my wife?" (Canada)

> "There are two problems with this approach: Firstly, such an approach would call for the commissioner to be incredibly empathetic! We would need to have the French concept of 'le patrimone' – intangible things that accrue to you as a human being – based on a core notion of integrity or dignity. Secondly, the complaint process would need to be modified. At present, we often conduct an inquiry in writing, therefore it is very difficult to assess harm as we get submissions in writing – it is likely that complainants will vary in their abilities to express harm in writing (in Alberta we have more oral enquiries than other provinces – so I have the advantage of being able to watch facial expressions)." (Canada)

Several respondents asserted that a weakness of this approach lies in defining and testing the concept of harm. However, it is suggested that the concept of 'harm' is recognised in law e.g. tort of negligence[433] as including: physical,[434] mental[435] or economic harm.[436] Also, future research could examine the merits of the definition of the types of informational harm offered by Beales.[437] He defined harm according to five factors, namely: i) using information to make inappropriate decisions, ii) using information deceptively, iii) using information to commit fraud and ID theft, iv) using information to intrude without value and v) not protecting information adequately, i.e. information security.

### *8.4.4 Level of harm*

Principle I of the APEC does not give any indication of the threshold level of harm that is contemplated. Thus, the proposed definition sought to test whether a *de minimis* ('any effect') or a high threshold of harm approach would be favourable, by exploring whether a definition of private data should protect against *any effect* on an individual, or, only against *unreasonable harm* to an individual.

#### *8.4.4.1 Any effect*

A *de minimis* approach is encapsulated by the phrase 'any effect.' A number of the expert respondents advocated this approach:

> "It shouldn't just be unreasonable harm; a small effect upon someone should be sufficient. The breach may not be harmful but it may be upsetting." (Australia)

---

[433] *Caparo Industries plc v Dickman* [1990] 2 AC 605
[434] *Donoghue v Stevenson* [1932] UKHL 100
[435] *Mental* harm is typically referred to as 'nervous shock' in tort cases; *Wilkinson v Downton* [1897] 2 QB 57 and *Janvier v Sweeney* [1919] 2 KB 316
[436] *Hedley Byrne & Co Ltd v Heller & Partners Ltd* [1964] AC 465
[437] Beales, H. (2003) Remarks of J. Howard Beales, III, Director, Bureau of Consumer Protection, Federal Trade Commission, Before the 2003 Symposium on the Patriot Act, Consumer Privacy, and Cybercrime, North Carolina Journal of Law Technology, Vol. 5, Iss 1: FALL, pp. 1-32

> "I choose 'any effect' because it is more effective. The term 'unreasonable harm' builds in a weaker effect of the law. The definition should be sharp and clear." (Netherlands)

> "I would agree with 'any effect.' I would say 'harm' is a limited category of data and 'unreasonable' is a threshold which is difficult to define." (Belgium)

> "I prefer 'any effect'. My view is that private data is absolutely closed against disclosure (expect very special circumstances) regardless of whether such disclosure creates unreasonable harm to an individual or not." (Czech Republic)

> "'Any effect:' The data subject must be protected against any kind of invasion to his fundamental right of data protection, being considered private or not." (Spain)

> "In a commercial world there is discussion regarding the scope of personal data and many are seeking to weaken the extent of the concept. In my opinion, coded health information should still be regarded as personal information. So I choose 'any effect' to give as much protection as possible to the individual." (Spain)

The responses indicate that some experts advocate a *de minimis* approach to harm assessment. These respondents were concerned that if the threshold test for engaging privacy protection was based on reasonableness it would be difficult for an individual to discharge the burden of proving unreasonableness.

#### *8.4.4.2 Unreasonable harm*

Other experts indicated that more than trivial harm should be required to engage privacy protection. These responses tended to balance the positive benefits of data processing against potential dangers from misuse:

> "I agree with 'unreasonable harm.' Data is necessary to the efficient operations of both government and business. There needs to be a balancing test in regards to data usage. The 'unreasonable harm' concept reflects this position." (USA)

> "A definition of private data should only protect against 'unreasonable harm' to an individual. I think that protection against 'any' harm to an individual is not feasible. Any activity has some risk and in my opinion "privacy" is not different than the others. So, limits are required to what "protection" means." (Spain)

> "I chose 'unreasonable harm' because 'any effect' could potentially prevent a lot of useful treatments which do not really harm anyone." (Spain)

> "I prefer 'unreasonable harm.' It is more strict. According to your definition, if we use the definition then we should protect against unreasonable harm.

> The 'any effect' would not permit exceptions e.g. those mentioned in the definition." (Italy)
>
> "I choose 'unreasonable harm.' 'Any effect' is too much. The important thing is to protect individuals against harm. If you try to protect against 'any effect' (negative and positive) then you will not publish nothing. Data protecting against 'unreasonable harm' will have more utility than data protecting against 'any effect.' In other words, 'any effect' is overprotection." (Spain)

The responses indicate that advocates of the unreasonable harm approach recognised the necessity of information flows in today's information economy and asserted that a *de minimis* approach would place an unduly restrictive burden on data processors.

#### *8.4.4.3 Objective v subjective test of unreasonableness*

Several respondents indicated that they preferred a harm-based approach to an 'any effect' model. However, they expressed dissatisfaction with the term 'unreasonable' as it is not precisely defined:

> "The term unreasonable is not one that I am familiar with. In Spain you only have to show harm to get legal redress. I don't know how you would interpret and apply the term unreasonable." (Spain)
>
> "Unreasonable harm – what does it mean? It depends upon the person – is a little bit of embarrassment ok or too much?" (USA)
>
> "The word unreasonable leaves room for manoeuvre e.g. it could be claimed that data processors do not have to seek the consent of the data subject because they believe that the data processing will not cause unreasonable harm. The phrase "the disclosure of which would cause unreasonable harm to an individual" would cause confusion – See Arts 6-8 of EU Directive. Also, technology cannot measure the concept of unreasonable harm – does the war on terrorism make certain data disclosure reasonable? Consider using more precise terms and objectifying the concept. ... Every country could have its own unreasonableness test! E.g. regarding anonymisation – some countries could take the approach that if one person could be identified then the data is not anonymised, whereas other countries could use statistical models to determine at what point reasonable identification disappears." (Netherlands)
>
> "Data protection authorities lack resources to become involved where there is no harm, but you need to look further at defining the term 'unreasonable'" (Belgium)

Some respondents suggested that foreseeable harm would be a more appropriate test. Generally, a foreseeable harm test is an objective test. However, the expert respondents qualified the test by suggesting that foreseeable harm should be subjectively determined by the data subject:

> "Harm – this raises the issue of whether it should be subjective or objective test of harm? Also, should the harm be foreseeable? ... I'm biased towards a subjective test of harm, but this doesn't solve the problem of foreseeability." (UK)

> "I prefer 'harm,' but I prefer the concept of foreseeable, serious harm according to the preference of the individual. Also need to determine if harm should include emotional harm and economic disadvantage." (UK)

Several respondents asserted that a weakness of this approach lies in defining and testing the concept of 'unreasonable' harm. However, the discussion in chapter two revealed that Calo advocates that the test for privacy harm should be both objective and subjective. Calo asserts that legislators, regulators and decision-makers could assess subjective privacy harms by reference to the degree of aversion[438] to any observation, or by reference to the amount of observation[439] experienced, since his approach indicates that high degrees of both translate into the greatest harm, but harm is also possible if either are very high.[440] Similarly, legislators, regulators and decision makers could assess objective privacy harms by reference to the degree of knowledge or consent,[441] as distinct from the severity of the information use. Accordingly, future research could examine the subjective and objective elements of privacy harm.

## *8.5 Summary*

The foregoing analysis indicates that the decision in *Durant v FSA*[442] is at odds with the general principles of data protection. It attempts to limit the scope of personal data. Whilst this approach is *prima facie* useful from privacy perspective, it fails to recognise that data protection legislation also serves other interests, and that a broader interpretation of personal data is necessary to achieve these purposes. The responses indicate that the terms personal and private data are not synonymous. Private data is data that may, or may not, be personal, but it is data over which an individual wishes to exert a privacy claim, on the basis that disclosure or processing

---

[438] It must be unwanted; Calo, R. M (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal,* Vol. 86, No. 3, pp.1132-1161, at p. 1154

[439] It can be acute or ongoing; Calo, R. M (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp.1132-1161, at p. 1144

[440] *Ibid*, p. 1154. Calo notes that the two are, obviously, related. Extensive surveillance can breed greater aversion. The idea here is that each context, state, or activity may be attended by a specific level of aversion to observation that can in turn be invaded to a lesser or greater degree.

[441] Forced or coerced consent must be adverse; Calo, R. M (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp.1132-1161, p. 1150

[442] *Michael John Durant v Financial Services Authority* [2003] EWCA Civ 1746, Court of Appeal (Civil Division) decision of Lord Justices Auld, Mummery and Buxton dated 8th December 2003. <http://www.bailii.org/ew/cases/EWCA/Civ/2003/1746.html> (Last accessed: 27.05.11)

may cause harm. The responses further indicate that a harm based definition of private is appealing at an intuitive level, since concentrating on harm from misuses of data reflects what concerns individuals the most, while not unduly restricting the free flow of information that underpins the economy. The findings from the data protection and privacy experts to the harm-based definition of private data mirrored the responses of individuals in the online survey of bloggers, regarding the concerns about the risk of harm from misuse of personal data. The results indicate that Bergkamp *et al*[443] have indeed identified a potential future direction for privacy and data protection, but that further research is required in order to develop a coherent definition of privacy protection based on managing risks to personal data, and 'unreasonable' harm from misuse of such information. Future research could specifically focus on the APEC countries approaches to implementing the harm principle, and examine how the term 'harm' should be tested, that is, whether it should be an objective or subjective, or combined test of harm, and whether the term foreseeable would be better than unreasonable.

[443] Other proponents of reform include industry experts such as Google's Global Privacy Counsel, Peter Fleischer <http://peterfleischer.blogspot.com/2007/09/eric-schmidt-on-global-privacy.html> (Last accessed 26.03.09)

# Chapter 9

## Conclusions & Recommendations

### *9.1. Introduction*

A major challenge of the information society is to establish whether privacy is still valued and worthy of legal protection. If privacy has an enduring value, then a concomitant challenge is to establish a proper balance between individual privacy and freedom of information; between individual autonomy and societal interest in information. This has significant implications for data protection laws regulating the processing of personal information, as, on one hand, for a modern society to function, personal information must flow; whilst on the other hand, societies which value individual liberty, autonomy and individuality must also ensure that 'private' data about an individual is not shared with others. That is, it requires that individual rights to privacy be established and enforced. Not only does this require a society to regulate how the state collects and processes information on its citizens; it also necessitates that it regulate the activities of private commercial companies which provide many services in today's society; and increasingly it necessitates the regulation of how individuals interact with each other via internet-based social media such as blogs. This thesis sought to address the question of what private data is, by examining various conceptions of privacy in the literature, and in particular, examining legislative measures currently employed to protect privacy. Also, since privacy is a socially determined construct that is influenced by peoples' perceptions of it, this thesis sought to explore individuals' perceptions of privacy as well as the views of privacy and data protection experts. A mix of qualitative and quantitative data collection methods were employed to generate empirical data, in order to examine whether existing privacy and data protection laws adequately protect individuals' privacy, or whether fundamental reform of data protection laws may be required to effectively address privacy in the Information Society.

## *9.2 Key findings*

A number of key findings have emerged from this research. Firstly, Zuckerberg's claim that social norms are changing, with the effect that privacy is no longer valued,[444] and Scott McNealy's assertion that: "you already have zero privacy – get over it"[445] are unwarranted. An analysis of data revealed that privacy has an enduring value in the information society. Even bloggers, who disclose personal information via a public medium, continue to value privacy; chapter 6 indicated that over 1/3rd of respondents of an online survey of bloggers were very concerned with the protection of personal information. Moreover, the analysis of bloggers' practices and expectations revealed that they actively take steps to protect their privacy and mitigate risks of privacy harm when blogging.

Secondly, a review of different theories of privacy in chapter two revealed that a universally accepted definition of privacy does not exist. That does not mean, however, that the concept of privacy is meaningless, rather, scholars have different goals in mind when they are defining the concept. In particular, some theorists adopt a broad definitional approach to privacy, which seeks to protect, not just an individual's personal information, but also their private sphere of action. However, this thesis determined that the focus of data protection laws should be on a narrower conception of privacy; one that has personal information protection as its focus. Accordingly, this thesis drew upon the work of a number of academics, including the definition of privacy offered by Rossler, the framework of 'contextual integrity' developed by Nissenbaum,[446] the pragmatic, harm-based approach' developed by Solove,[447] and the categories of privacy harm developed by Calo[448] in order to develop a harm-based definition of data privacy and private data.

[444] Kirkpatrick. M. (2010) "The Age of Privacy is Over," ReadWriteWeb, The article reports an on stage between Mark Zuckerberg, CEO of Facebook and the founder of TechCruch , in which Zuckerberg claims that 'the social norms of privacy are changing.' <http://www.readwriteweb.com/archives/facebooks_zuckerberg_says_the_age_of_privacy_is_ov.php> (Last accessed 27.02.11)

[445] Sprenger, P. (1999) "Sun on Privacy: 'Get Over It'" <http://www.wired.com/news/politics/0,1283,17538,00.html> (26.01.1999) (Last accessed 20.05.11)

[446] Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, pp.119-157

[447] Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4. pp. 1087-1155

[448] Calo, M. R. (2011) "The Boundaries of Privacy Harm," *Indiana Law Journal*, Vol. 86, No. 3, pp. 1132-1161

Thereafter, a review of the literature in chapter three on the relationship between privacy and data protection revealed that current data protection laws refer to a generic concept of 'privacy,' but this concept is not defined. It is a truism to state that effective data protection is only possible if the terms that are defined in the laws are conceptually certain, interpreted uniformly, and applied equally; yet the terms privacy or private data are not defined in the Directive or other legislative measures.[449] The analysis concluded that the failure to define privacy and private data in existing data protection laws is a major weakness. Moreover, this thesis demonstrated that although human rights-based privacy laws seek to protect a broadly defined concept of privacy, the focus of data protection laws should be narrower, on the basis that they are primarily concerned with the processing of personal data. Accordingly, a narrow definition of 'data privacy,' was developed and found to be conceptually coherent when tested on privacy and data protection experts.

Furthermore, analysis of the literature and primary data on data protection revealed that such laws seek to achieve privacy protection through the imposition of rules on the processing of personal and sensitive data. Thus, an aim of this thesis was to explore how the terms found in these laws embody or intersect with a concern for privacy protection[450] and as a corollary of this, whether private data is adequately protected through current data protection laws. The analysis revealed that the terms personal and private are not synonymous. Rather, personal data is a broader concept. It is defined and interpreted broadly since the purpose of definition of 'personal data' is to distinguish identifying information from anonymous information, as personal data is subject to the provisions of the Directive, whilst, anonymous data is regarded as non-personal data, and not subject to the provisions of the Directive. In contrast, 'private' is a term used to denote information that an individual seeks to make greater privacy claims in respect of, on the basis that it may result in harm to the individual if disclosed.

Additionally, this thesis examined the relationship between sensitive and private data. Analysis of the literature and primary data on sensitive data revealed that in the absence of a universally agreed definition of private data, legislators used sought to protect 'sensitive data' as a proxy. However, analysis of the literature indicated that

[449] Napier, B. (1992) "International Data Protection Standards and British Experience," *Informatica e drittio*, Vols. 1-2, pp.83-100
[450] See Chapter 2 for a detailed analysis of privacy conceptions

the term is not universally defined in the legislative provisions; in fact, is not included in some legislative provisions.

In those legislative measures which do contain the term sensitive data, i.e. CoE Convention (108), UN Guidelines, and Directive 95/46/EC, the definitions reflect post World War II concerns regarding discrimination and protection of human dignity, e.g. health, religious and racial data are considered sensitive. However, analysis of the survey data from telephone respondents and blog respondents indicated that not all of the legally recognised categories of data continue to be perceived as sensitive, e.g. trade union membership and political or philosophical opinions. Moreover, analysis of the literature, and interviews with expert respondents suggested that developments in technology are raising new potential categories of sensitive data. Indeed, findings from interviews and the survey indicate that some not legally recognised categories of data are emerging, which are considered extremely sensitive e.g. biometric, and clickstream data.

However, this thesis suggests a decision to simply include new categories, or delete existing categories should not be taken lightly. An attempt to grade new categories of data according to their sensitivity would be fraught with difficulties, as it would require a casuistic form of regulation, which is more complex and lengthy. Moreover, the analysis of the literature indicated that contextual factors influence privacy perceptions, e.g. an individual seeking to avoid an abusive ex-partner may consider their name and address private, and restrict their listing in a public telephone directory. Accordingly, this thesis contends that *a priori* classification of data as privacy sensitive is erroneous. It is a fallacy since the privacy sensitivity of data cannot be pre-determined; rather it is influenced by contextual factors, and so, should be determined on *a posterio* basis. Thus, although conflation of the terms sensitive and private is legislatively convenient, it is conceptually problematic and an inadequate approach to privacy protection.

During the literature review it was noted that some privacy advocates such as Nissenbaum and Solove do not seek to engage in an *a priori* classification of data as personal or private, rather they asserted that contextual factors influence the privacy quality of information. Similarly, the APEC Privacy Framework (2005) does not contain the term sensitive; instead it contains a harm principle, which stipulates that:

> "personal information protection should be designed to prevent the misuse of such information...specific obligations should take account of such risk, and remedial measures should be proportionate to the likelihood and severity of the harm threatened by the collection, use and transfer of personal information."

Building on this approach, this thesis sought to test the merits of a harm-based definition of **data privacy** and **private data.** The defintion states that:

> *"**Data privacy** concerns the legal regulation of the boundary between personal and private data. **Private data** is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual. It is the legal right of an individual to withhold consent to the collection, processing, communication or usage of personal data, the disclosure of which could cause unreasonable harm to an individual. Disclosure of private data should not be compulsory except where it is in the public interest, for instance if disclosure is in the interests of national security, public safety, the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health, or for the protection of the rights and freedoms of others or society as a whole."*

The definition is technologically neutral and therefore flexible, which is an advantage, as it is difficult to prospectively regulate for advances in technology or information uses that have not been invented or even considered. On a positive note, the expert responses to the definition indicate that a harm based definition of private is appealing at an intuitive level, since concentrating on harm from misuses of data reflects what concerns individuals the most, while not unduly restricting the free flow of information that underpins the economy. The findings from the data protection and privacy experts to the harm-based definition of private data were mirrored the responses of individuals in the online survey of bloggers, regarding the concerns about the risk of harm from misuse of personal data. The results indicate that Bergkamp *et al*[451] have indeed identified a potential future direction for privacy and data protection, but that further research is required in order to develop a coherent definition of privacy protection based on managing risks to personal data, and 'unreasonable' harm from misuse of such information.

Based on the analysis in this thesis, Directive 95/46/EC appears to be misconceived and mis-focused. The focus has erroneously been on classification of data as such,

[451] Other proponents of reform include industry experts such as Google's Global Privacy Counsel, Peter Fleischer <http://peterfleischer.blogspot.com/2007/09/eric-schmidt-on-global-privacy.html> (Last accessed 26.03.09)

instead of the harms arising from such data uses. The Directive regulates at the wrong level and fails to balance competing interests properly. It regulates the collection and processing of data upstream, when it should regulate specific harmful uses downstream.[452] The real tension in the current privacy debate is between an individual (data subject's) desire for greater privacy and their wish for the many benefits that flow from readily available personal information. As with all other information, personal data can be used or misused. Future legislation should focus on actual harms caused by the misuse of personal data and apply targeted remedies to such misuses. Thus, it is suggested that Directive 95/46/EC and all other data protection laws should be reviewed, and updated to reflect changes in society and technology. Future legal provisions should be attuned to a risk of harm continuum, rather than a dichotomy between personal and sensitive data.

## *9.3 Recommendations:*

1) The key terms of Data Protection Directive 95/46/EC should be reviewed. Establishing a consensus over the interpretation of key terms e.g. personal data would improve harmonization of privacy protection in member states.

2) A clause should be inserted in Directive 95/46/EC and any future global data protection measure explaining that the legislation is concerned with *data privacy*. 'Data protection' is a misnomer, and generates confusion with the EU Database Directive.

3) Directive 95/46/EC and any future global data protection framework should include the term personal data. A defintion of personal data should be retained, since the purpose of definition of 'personal data' is to distinguish identifying information from anonymous information, as personal data is subject to the provisions of the Directive, whilst, anonymous data is regarded as non-personal data, and not subject to the provisions of the Directive.

4) Directive 95/46/EC and any future global privacy data protection law should not contain the term 'sensitive' data. The current approach of listing certain types of personal data as sensitive engages in an *a priori* classification exercise which is flawed. It is a fallacy. The privacy sensitivity of data cannot be pre-

[452] Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects o Europe's Data Protection Policy in an Information Driven Economy," *Computer Law & Security Report*, Vol. 18. No. 1, pp. 31-37, p. 42

determined; rather it is influenced by contextual factors, and so, should be determined on *a posterio* basis.

5) Directive 95/46/EC and any future global data protection law should focus on actual harms caused by the misuse of personal data and apply targeted remedies to such misuses.

## *9.4 Thesis contributions & further research*

In conclusion, this thesis makes several contributions to knowledge.[453] Firstly, it enhances existing knowledge on 'privacy' by demonstrating that it is still valued in the information society. Secondly, the thesis makes a novel contribution through the collection and analysis of empirical data on the concepts of *personal* and *sensitive* data. The findings indicate that the current categories of sensitive data, drafted in a post-world war II era, are now in need of revision as some of the categories are no longer considered as sensitive as they once were, whilst changes in technology and computing have generated potential new types of sensitive data. Notwithstanding this, the thesis contends that classifying data as sensitive (or not) is a fallacy as all personal data can pose a privacy risk dependent on the context of the data processing. Finally, the thesis contributes to the debate on private data and explores legislative reform proposals. The findings indicate that a harm-based conception of private data has merit and appeal, and that future amendments of the Directive should focus on the risk of *unreasonable harm* posed by data processing, as opposed to a personal–sensitive dichotomy. However, further research is required in order to develop a coherent definition of privacy protection based on managing risks to personal data, and 'unreasonable' harm from misuse of such information. Future research could specifically focus on the APEC countries approaches to implementing the harm principle, and examine how the term 'harm' should be tested, that is, whether it should be an objective or subjective, or combined test of harm, and whether the term 'foreseeable' would be better than 'unreasonable.'

[453] Including 3 peer-reviewed journal articles, copies of which can be found in Appendix F

# Bibliography


Acquisiti, A. & Grossklags, J. (2004) "Privacy Attitudes and Privacy Behaviour: Losses, Gains, and Hyperbolic Discounting," in Camp, L. J. & Lewis, S. The Economics of Information Security (Kluwer), pp. 165-178

Allen, A. (1988) Uneasy Access: Privacy for Women in a Free Society (Rowman & Littlefield, Totowa, NJ)

Art 29 Data Protection Working Party 'Opinion 4/2007 on the concept of personal data' (2007), <http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf )> (Last accessed 01.03.09)

Austin, L. (2003) "Privacy and the Question of Technology," *Law and Philosophy*, Vol. 22, No. 2, pp. 119-166

Bainbridge, D. (1996) 'EC Data Protection Directive,' (Butterworths)

Bangemann M. (1994) Europe and the Global Information Society <http://ec.europa.eu/archives/ISPO/infosoc/backg/bangeman.html> (Last accessed: 20.02.11)

Barbaro, M., Zeller, T., & Hansell, S. (2006) 'A Face Is Exposed for AOL Searcher No. 4417749,' New York Times, (9th August) <http://query.nytimes.com/gst/fullpage.html?res=9E0CE3DD1F3FF93AA3575BC0A9609C8B63&sec=&spon=&pagewanted=1>

BBC (2007) "Blogger sued over topless Aniston," (22nd Feb) <http://news.bbc.co.uk/1/hi/entertainment/6385677.stm> (Last accessed: 01.03.09)

BBC, (2007) "Egyptian blogger jailed for insult," (22nd Feb) <http://news.bbc.co.uk/1/hi/world/middle_east/6385849.stm?ls>

Beales, H. (2003) Remarks of J. Howard Beales, III, Director, Bureau of Consumer Protection, Federal Trade Commission, Before the 2003 Symposium on the Patriot Act, Consumer Privacy, and Cybercrime, North Carolina Journal of Law Technology, Vol. 5, Iss 1: FALL, pp. 1-32

Benn, S. (1971) "Privacy, Freedom, and Respect for Persons," in Pennock, J. & Chapman, R. (Eds.), Nomos XIII: *Privacy*, (New York: Atherton Press)

Bennett, C. J., & Raab, C., D. (2003) The governance of privacy: policy instruments in global perspective (Ashgate, UK)

Bergkamp, L. (2002) "EU Data Protection Policy: The Privacy fallacy: Adverse Effects of Europe's Data protection Policy in an Information Driven Economy" *Computer Law & Security Report*, Vol. 18 No.1 pp. 31-47

Bernard, H. R. (1994) Research methods in anthropology: Qualitative and quantitative Approaches, (2d ed.) (Walnut Creek, CA: Alta Mira) p. 43

Berson, I.R., Berson, M.J. & Ferron, J.M. (2002) 'Emerging risks of violence in the digital age: lessons for educators from an online study of adolescent girls in the United States,' *Meridian, A Middle School Computer Technologies Journal*, <http://www.ncsu.edu/meridian/sum2002/cyberviolence/cyberviolence.pdf>

Bing, J. (1972) "Classification of Personal Information with respect to the Sensitivity Aspect" Proceedings of the First International Oslo Symposium on Data Banks and Society, 98-141

Bing, J & Selmer, K. S. (1980) (eds), *A Decade of Computers and Law* (Universitetsforlaget: Oslo)

Blaikie, N. (1993) Approaches to Social Enquiry (Cambridge: Polity Press)

Blekeli, R.D. (1980) 'Framework for the Analysis of Privacy and Information Systems' in Bing, J & Selmer, K. S. (eds), *A Decade of Computers and Law* (Universitetsforlaget)

Bloustein, E.J. (1964) "Privacy as an Aspect of human Dignity: An Answer to Dean Prosser" *New York University Law Review*. Vol. 39, pp.962-1007

Bok, S. (1983) Secrets: On the Ethics of Concealment and Revelation" (Pantheon Books, New York)

Booth, S., Jenkins, R., Moxon, D., Semmens, N., Spencer, C., Taylor, M. & Townend, D. (2004) "What are 'Personal Data'? : A study conducted for the UK Information Commissioner" <http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/final_report_21_06_04.pdf> (Last accessed: 20.03.09)

Bradney, A., Cownie, F., Masson, J., Neal, A. & Newell, D. (2000) "Law in action/law in books", 'in *How to study the law* (4th edn) (London: Sweet & Maxwell)

Bradshaw, A. (1997) "Sense and Sensibility: Debates and Development in Socio-Legal Research Methods," in Thomas, P. (1997) (ed) *Socio-Legal Studies*, (Aldershot: Ashgate-Dartmouth), p.99

Bray, H. (2004) "Job blogs hold perils, opportunities," The Boston Globe, p1, in Viegas, F. (2005) "Bloggers expectations of privacy and accountability: an initial survey," *Journal of Computer- Mediated Communication*, Vol. 10, No. 3, No. 12. <http://jcmc.indiana.edu/vol10/issue3/viegas.html> (Last accessed 27.02.08)

Bryman, A. (2004) Chapter 15: 'Interviewing in qualitative research,' *Social Research Methods* (2nd edition), (Oxford University Press, Oxford)

Bryman, A. in Lewis-Beck, M.S., Bryman, A. & Liao, T. F. (2004) *The Sage Encyclopedia of Social Science Research Methods*, (Sage)

Bulford, C. (2007-08) "Between East and West: The APEC Privacy Framework and the Balance of International Data Flows," *I/S: A Journal of Law & Policy* Vol. 3. No. 3, pp.705-722

Bulmer, M. (1979) “Concepts in the analysis of qualitative data,” *Sociological Review* 27 (4) pp. 651–77

Burke, K. C., (1981) “Secret Surveillance and the European Convention on Human Rights,” *Stanford Law Review,* Vol. 33, No. 6, pp. 1113-1140

Bygrave, L. (2000) “Minding the machine: art 15 of the EC Data Protection Directive and automated profiling,” *Privacy Law and Policy Reporter,* Vol. 7, No. 1, Art.10

Bygrave, L. (2001) 'The Place of Privacy in Data Protection Law', University of New South Wales Law Journal, Vol. 24, No. 1, Art. 6 <http://www.austlii.edu.au/au/journals/UNSWLJ/2001/6.html>

Bygrave, L. (2002) *Data Protection Law: Approaching its Rationale, Logic and Limits*, (Kluwer)

Bygrave. L. A. (2004) “Privacy Protection in a Global Context – A Comparative Overview,” *Scandinavian Studies in Law*, Vol. 47, pp. 319–348, at pp. 330-31

Bygrave, L. A. (2010) “Privacy and Data Protection in an International Perspective,” *Scandinavian Studies in Law,* pp. 165-200

Campbell, C. & Wiles, P. (1976) “The Study of Law in Society in Britain,” *Law & Society Review*, Vol. 10, pp. 547–578

Calo, R.M. (2011) ‘The Boundaries of Privacy Harm,’ *Indiana Law Journal*, Vol. 86, No. 3, pp. 1132-1161

Cate, F. (2006) “The Failure of Fair Information Practice Principles,” in Winn, J. K. (ed) *Consumer Protection in the Age of the Information Economy, (Ashgate, UK)*

Citron, D. K., (2008) “Technological Due Process,” *Washington University Law Review,* Vol. 85, pp. 1249-1313

Clarke, R. (1998) “A History of Privacy in Australia: Context” <http://www.anu.edu.au/people/Roger.Clarke/DV/OzHC.html#ContI> (Last accessed 20.02.11)

Clarke, R. (2006) “Introduction to Dataveillance and Information Privacy, and Definitions of Terms” <https://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html> (Last accessed 20.02.11)

Couper, M. P., Traugott, M., & Lamias, M. (2001) “Web surveys: Perception of burden,” *Social Science Computer Review*, Vol. 19, No.2, pp. 146-162

Crabtree, V. (2007) Religion in the United Kingdom: Diversity, Trends and Decline <http://www.vexen.co.uk/UK/religion.html> (Last accessed: 15.01.10)

Creswell, J.W. (1998) Qualitative Inquiry and Research Design: Choosing among five traditions (Thousand Oaks: Sage)

Data Protection & Privacy Commissioners, (2005) "Montreux Declaration: The protection of personal data and privacy in a globalised world: a universal right respecting diversities," <http://www.libertysecurity.org/IMG/pdf/montreux_declaration_eng.pdf> (Last accessed 20.02.11)

Davie, G. (1994) Religion in Britain Since 1945: Believing without Belonging, (Wiley Blackwell: UK)

DeCew Wagner, J. (1997) In Pursuit of Privacy: Law, Ethics and the Rise of Technology, (Cornell University Press, Ithaca & London)

Denscombe, M. (1998) The Good Research Guide for Small Scale Social Research Projects (Open University Press: Milton Keynes, UK)

Denzin, N. K. (1970) The Research Act: A theoretical introduction to sociological methods. (Chicago: Aldine)

Denzin, N. K. (1978). The Research Act: A Theoretical Introduction to Sociological Methods, (New York: McGraw-Hill).

Dexter, L. A. (1970) *Elite and Specialised Interviewing*, (Northwestern University Press); republished with an introduction by Ware, A. & Sánchez-Jankowski, M., (2006) European Consortium for Political Research.

Dey, I. (1993) Qualitative data analysis: A user-friendly guide for social scientists (London: Routledge Kegan Paul)

Dillman, D. A., Tortora, R.D., Conradt, J. & Bowker, D. (1998) "Influence of plain vs. fancy design on response rates for web surveys," Paper presented at Joint Statistical Meetings, Dallas, Texas <http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf>

Dwek, R (1990) "EC Scheme for Data Protection Stuns UK," *Marketing*, 12th July

Ely, M., Anzul, M., Freidman, T., Garner, D., & McCormack-Steinmetz, A. (1991) *Doing Qualitative research: circles within circles*, (London, England: Falmer Press)

Epstein, L. (1986) "Strategies of Judicial Research: Interviewing U.S. Supreme Court Justices and Interest Group Attorneys," Paper Presented at Annual meeting of Southern Political Science Association, Atlanta, Georgia, <http://epstein.law.northwestern.edu/research/conferencepapers.1986SPSA.pdf >, p.16

Equality & Human Rights Commission (2008) 'Who we are and what we do,' <http://www.equalityhumanrights.com/uploaded_files/who_we_are.pdf> (Last accessed: 15.01.10)

Etzioni, A. (1999) The Limits of Privacy, (Basic Books, NY)

European Commission, (2009) "Review of the data protection legal framework," <http://ec.europa.eu/justice/policies/privacy/review/index_en.htm>

Fauvet, J. (1989) "Privacy in the New Europe," Transnational Data & Communications Report, Nov 17-18 cited in Newman A, "Protecting Privacy in Europe: Administrative Feedbacks and regional Policies," in Meunier, S. & McNamara, K. R. (2007) (eds) *Making history: European Integration and Institutional change at fifty* (OUP)

Fleischer, P. (2007) "Call for Global Privacy Standards," Google Public Policy Blog, 14th Sept 2007 <http://googlepublicpolicy.blogspot.com/2007/09/call-for-global-privacy-standards.html> (Last accessed 20.02.11)

Fleischer, P. (2007) "The Need for global privacy standards," UNESCO Conference, Ethics and Human Rights in the Information Society, (13-14 September), <http://portal.unesco.org/ci/en/files/25452/11909026951Fleischer-Peter.pdf/Fleischer-Peter.pdf>

Fleischer, P. (2007) "Call for Global Privacy Standards," Google Public Policy Blog, 14th Sept 2007 <http://googlepublicpolicy.blogspot.com/2007/09/call-for-global-privacy-standards.html> (Last accessed 20.02.11)

Floridi, L. (2005) "The ontological interpretation of informational privacy," *Ethics and Information Technology*, Vol.7, pp. 185–200, at p. 191

Fricker, R.D. (2008) "Sampling Methods for Web and Email Surveys," in Fielding, N. *et al The Sage Handbook of Online Research Methods*, (Sage: London)

Fried. C. (1968) "Privacy," *Yale Law Journal* Vol. 77, pp. 482-483

Gallup Organisation, (2005) Flash Eurobarometer No 225: Data Protection in the European Union - *Citizens' Perceptions* <http://ec.europa.eu/public_opinion/flash/fl_225_en.pdf>

Garcia, F. J. (2005) "*Bodil Lindqvist*: A Swedish Churchgoer's Violation of the European Union's Data Protection Directive Should Be a Warning to U.S. Legislators," *Fordham Intellectual Property Media & Entertainment Law Journal*, pp. 1205-1239

Gavison, R. (1980) "Privacy and the Limits of the Law" *Yale Law Journal*. Vol. 89, No. 3, pp. 421-471

Gillespie, A. (2007) Foundations of Economics, (OUP: UK)

Glenn, D. (2003) "Scholars who blog," The Chronicle of Higher Education, Vol. 49, A14.

Greenleaf, G., (2003) "Australia's APEC privacy initiative: the pros and cons of 'OECD Lite'," *Privacy Law & Policy Reporter*, Vol. 10, pp. 1–6

Greenleaf, G (2009) "Five years of the APEC Privacy Framework: failure or Promise?" Computer Law & Security Review Vol. 25, pp 28-42

Grudin, J. (2001). "Desituating action: Digital representation of context." Human-Computer Interaction, 16 (2-3), pp.269-286.

Gutwirth, S. (2002) *Privacy and the Information Age*, (Rowman & Littlefield), p. 88

Guyon, R. (2006) "Outline of Privacy and Spain Laws in Japan and Australia (From a Company Perspective) and APEC Privacy Framework Brief Overview*,"* 865 PLIIP, p. 616 (June/July 2006), cited in Bulford, C. (2007-08) "Between East and West: The APEC Privacy Framework and the Balance of International Data Flows," *I/S: A Journal of Law and Policy*, Vol. 3, p.719

Habermas, J. (1962) *The Structural Transformation of the Public Sphere,* (MIT Press, Cambridge), translated by Burger, T. & Lawrence, F.

Harris, P. (1986) 'Curriculum Development in Legal Studies,' *Law Teacher*, Vol. 20, No.2, pp. 110-123

Herring, S. C., Scheidt, L. A., Bonus, S., & Wright, E. (2004). "Bridging the gap: A genre analysis of weblogs," *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)* Los Alamitos: IEEE Computer Society Press. <http://www.blogninja.com/DDGDD04.doc>

Home Office (2005) "UK Passport Service: Improving Passport Security and Tackling ID Fraud," Reference: UKPS001/2005 (Press Release: 24th Mar 2005).

House of Commons: Justice Committee (2008) "Protection of Private Data," First Report of Session, HC 154.

House of Lords Select Committee on Science and Technology (1996) "Information Society: Agenda for Action in the UK," 5th Report, HL Paper 77, <http://www.parliament.the-stationery-office.co.uk/pa/ld199596/ldselect/inforsoc/ch1.htm>

ICO, (2007) 'Data Protection Technical Guidance - determining what is personal' (v1.0 21.08.07)<http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/personal_data_flowchart_v1_with_preface001.pdf> (Last accessed 01.03.09)

Inness, J.C. (1992) *Privacy, Intimacy, and Isolation* (Oxford University Press, New York)

Karvalics, L. (2007) "Information Society – what is it exactly? (The meaning, history and conceptual framework of an expression)" <http://www.ittk.hu/netis/doc/ISCB_eng/02_ZKL_final.pdf>

Kay, B. & Johnson, T.J. (1999) "Research methodology: taming the cyber frontier. Techniques for improving online surveys," *Social Science Computer Review*, Vol. 17, No. 3, pp. 323-337

Kelley, K., Clark, B., Brown, V. & Sitzia, J. (2003) 'Methodology Matters: Good Practice in the conduct and reporting of survey research, *International Journal for Quality in Health Care*, Vol. 15, No. 3, pp. 261-266

Kerlinger, F. N. (1973) *Foundations of Behavioural Research*, (New York: Holt, Reinhart & Winston)

Kirby, M. (2011) "The history, achievement and future of the 1980 OECD guidelines on privacy," *International Data Privacy Law*, Vol. 1, No. 1, pp. 6-14

Kirkpatrick. M. (2010) "The Age of Privacy is Over," ReadWriteWeb, <http://www.readwriteweb.com/archives/facebooks_zuckerberg_says_the_age_of_privacy_is_ov.php> (Last accessed 27.02.11)

Korff, D. (2002) EC Study on the Implementation of Data Protection Directive, Comparative study of national laws <http://www.eu.int/comm/justice_home/fsj/privacy/docs/lawreport/consultation/univessex-comparativestudy_en.pdf> (Last accessed: 23.06.07)

Leech, N. L, & Onwuegbuzie, A.J. (2009) "A typology of mixed methods research designs." *Quality and Quantity,* Vol. 43, pp. 265-275

Lichtenstein, S. D., Darrow, J. (2006) "Employment Termination for Employee Blogging: Number One Tech Trend for 2005 and Beyond, or a Recipe for Getting Dooced?" *UCLA Journal of Law and Technology (JOLT), Vol. 10*, No. 2, pp. 1-49

Lincoln, Y. S. & Guba, E.G. (1985), *Naturalistic Inquiry* (Newbury Park, CA: Sage Publications)

Lindop, N. (1978) *Report of the Committee on Data Protection* (Chairman: Sir Norman Lindop), (Cmnd 7341) United Kingdom

Lindsay, D. (2004) "Misunderstanding 'personal information': Durant v Financial Services Authority," Privacy Law & Policy Reporter, Vol. 10, No. 10, No. 13

Linklaters (2004) Hot Topic: History repeats itself: the implementation of EU data protection legislation in the accession countries. <http://www.linklaters.com/pdfs/briefings/040517_DP.pdf>(Last accessed: 20.05.11)

Manheim, J. B. & Rich, R. C. (1986) *Empirical Political Analysis* (New York: Longman)

Marshall, G. (1998) 'Information society,' *A Dictionary of Sociology*, <http://www.encyclopedia.com/doc/1O88-informationsociety.html>

Maxwell, J. (1996) Qualitative research design: An interactive approach (Thousand Oaks, CA: Sage)

Mayer–Schönberger, V. (1997) "Generational Development of Data Protection in Europe," in Agre, P. & Rotenberg, M. (eds) *Technology and Privacy: The New Landscape*, (Cambridge, Massachusetts, The MIT Press)

McCormick, S. (2004) "ASEM: A Promising Attempt to Overcome Protective Regionalism and Facilitate the Globalization of Trade," *Annual Survey of International & Comparative Law*, Vol. 10, No. 10, Article 10

McGivern, Y. (2003) The Practice of Market and Social Research: An Introduction, (Pearson Education, Harlow)

Mead, G. H. (1964) *On Social Psychology* (University of Chicago Press)

Medlin, C., Roy, S. & Ham Chai, T. (1999) "World Wide Web Versus Mail Surveys: A Comparison And Report," Paper presentation at ANZMAC99 Conference, Marketing in the Third Millennium, Sydney, Australia, <http://smib.vuw.ac.nz:8081/www/ANZMAC1999/Site/M/Medlin.pdf>

Mill, J. S. (1859), *On Liberty And Other Essays*, Ch. 1 (World Classics Series, OUP)

Miller, A. R. (1971) The Assault on Privacy: Computers, Data banks and Dossiers, (Ann Arbor: University of Michigan Press)

Mills, E. (2005) 'Google Balances Privacy, Reach' C|Net News.com <http://news.com.com/Google+balances+privacy%2C+reach/2100-1032_3-5787483.html>

Moore, B. (1984) Privacy: Studies in Social and Cultural History (M E Sharpe, Armonk, NY)

Morse, J. (2002) "Principles of mixed and multi-methods research design" in Tashakkori, A. & Teddlie, C. (eds) *Handbook of mixed methods in social research and behavioural research*, (Thousand Oaks: Sage)

Napier, B. (1992) "International Data Protection Standards and British Experience," *Informatica e drittio*, Vols. 1-2, pp. 83-100.

Nardi, B., Schiano, D. J & Gumbrecht, M. (2004) "Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary?" Proceedings Conference on Computer-Supported Cooperative Work, (New York: ACM Press) pp. 222-231

Newman A, "Protecting Privacy in Europe: Administrative Feedbacks and regional Policies," in Meunier, S. & McNamara, K. R. (2007) (eds) *Making history: European Integration and Institutional change at fifty* (OUP)

Niewyk, D. & Nicosia, F. (2000) *The Columbia Guide to the Holocaust*, (Columbia University Press: New York)

Nissenbaum, H. (1998) "Protecting Privacy in an Information Age: The Problem of Privacy in Public," *Law & Phil.* Vol.17, No.5/6, pp. 559-596

Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review*, Vol. 79, No. 1, pp.119-157

Ofcom (2008) Nations and Regions CMR UK summary <http://www.ofcom.org.uk/research/cm/cmrnr08/uksummary.pdf>

Office of National Statistics (2003) Ethnicity Nugget <http://www.statistics.gov.uk/CCI/nugget.asp?ID=273>

Onwuegbuzie, A. & Teddlie, C (2002) "A framework for analysing data in mixed methods research" in Tashakkori, A. & Teddlie, C. (eds*) Handbook of mixed methods in social research and behavioural research*, (Thousand Oaks: Sage)

Palen, L., & Dourish, P. (2003). Unpacking "privacy" for a networked world. *Proceedings of the ACM CHI.* Ft. Lauderdale, FA <http://delivery.acm.org/10.1145/650000/642635/p129-palen.pdf?key1=642635&key2=7273414711&coll=&dl=GUIDE&CFID=15151515&CFTOKEN=6184618> (Last accessed: 01.03.09)

Pernice, I. E. A., (2009) "The Treaty of Lisbon: Multilevel Constitutionalism in Action," *Columbia Journal of European Law,* Vol. 15, No. 3, pp. 349-407

Petite Anglaise (2007) "Wrong footed," <http://news.bbc.co.uk/1/hi/world/middle_east/6385849.stm?ls> (Last accessed: 01.03.09).

Pilieci, V. (2010)"World is losing grip on privacy: watchdog - Next decade will be crucial in protecting personal data," Ottawa Citizen, (17th August)

Plous, S. quoted Azar, B. (2000) "A Web of Research: They're fun, they're fast and they save money, but do Web experiments yield quality results?," *Monitor on Psychology,* Vol.31, pp.42-47

Posner, R. (1978) "John A. Sibley Lecture: The Right to Privacy," *Georgia Law Review*, Vol. 12, No. 3, pp. 393

Posner, R. (1983) *The Economics of Justice,* (Harvard University Press)

Post, R. C. (1989) "The Social Foundations of Privacy: Community and Self in the Common Law Tort," *California Law Review,* Vol. 77, pp. 957- 1010

Poullet, Y., Dinant, J-M., de Terwange, C. & Perez-Asinari, M. V. (2004) "Report on the application of data protection principles to the worldwide telecommunication networks: Informational self-determination in the internet era," <http://www.coe.int/t/dghl/standardsetting/dataprotection/TPD%20documents/T-PD%20_2004_%2004%20E%20final%20Report%20POULLET.pdf> (Last accessed: 29.04.11)

Pounder, C. (2007) "Why the APEC Privacy Framework is unlikely to protect privacy." <http://www.out-law.com/default.aspx?page=8550 > (Last accessed: 20.03.09)

Raab, C. & Bennett, C. (1998) 'The Distribution of Privacy Risks: Who needs protection?' *The Information Society*, Vol. 14 (4) pp. 263-274

Rachels, J. (1975) "Why Privacy is Important?" *Philosophy and Public Affairs*, Vol. 4, No. 4, pp. 323-333

Reding, V. (2011) "Tomorrow's Privacy The upcoming data protection reform for the European Union," *International Data Privacy Law*, Vol. 1, No. 1

Regan, P. (1995) Legislating Privacy: Technology, Social values and Public Policy, (Chapel Hill: University of North Carolina Press)

Rosen, J. (2000), The Unwanted Gaze: The destruction of Privacy in America, (Random House)

Rosen, J. (2010) "Nude Awakening," *The New Republic*, Feb. 10, p.8

Rossler, B. (2005) The Value of Privacy, (Polity Press)

Rubenfeld, J. (1989) "The Right of Privacy," in Slobogin, C. (2007) Privacy at Risk: The New Government Surveillance and the Fourth Amendment (University of Chicago Press)

Rule, J., McAdam, D., Stearns, L., & Uglow, D. (1980) The Politics of Privacy: Planning for Personal Data Systems as Powerful Technologies, (Elsevier, New York)

Rule, J. B., Mc Adam, D., Stearns, L.D. in Johnson, D. G. & Nissenbaum, H. (1995) *Computers, Ethics & Social Values*, (Prentice Hall)

Ryan, G. W. & Bernard, H. R. (2003) "Techniques to Identify Themes," *Field Methods*, Vol.15, No.1, pp. 85–109

Salter, M. & Mason, J. (2007) Writing Law Dissertations: An introduction and Guide to the Conduct of Legal Research, (1st edn) (Pearson Longman)

Schwartz, P. (1999) "Privacy and Democracy in Cyberspace," *Vanderbilt Law Review*, Vol. 52, pp. 1609-1701

Scientific and Technological Options Assessment Unit of the European Parliament (STOA) (1998) "An Appraisal of Technologies of Political Control," point 4. <http://cryptome.org/stoa-atpc.htm> (Last accessed 20.02.11)

Search Oriented Architecture, (2000) "Weblog definition," Blogs are defined by their format: a series of entries posted to a single page in reverse-chronological order. <http://searchsoa.techtarget.com/sDefinition/0,,sid26_gci213547,00.html> (Last accessed: 20.07.09)

Seltzer, W. & Anderson, M. (2001) "The Dark side of Numbers: The role of Population Data Systems in Human Rights Abuses," *Social Research* , pp. 486-488

Selwyn, N. & Robson, K. (1998) "Using e-mail as a research tool," *Social Research Update*, No. 21 <http://sru.soc.surrey.ac.uk/SRU21.html>

Serfaty, V. (2004) The mirror and the veil: An overview of American online diaries (Rodopi: New York)

Sheehan, K. B., & McMillan, S. J., (1999) "Response variation in e-mail surveys: An exploration," *Journal of Advertising Research*, Vol. 39, No. 4, pp. 45-54.

Sieghart, P. (1976) *Privacy and Computers*, (Latimer, London).

Sills, S. J. & Song, C. (2002) "Innovations in Survey Research: An application of web-based surveys," *Social Science Computer Review*, Vol. 20, No.1, pp. 22-30

Simitis, S. (1999), "Revisiting Sensitive Data," <http://www.coe.int/t/dghl/standardsetting/dataprotection/Reports/Report_Simitis_1999.pdf> (Last accessed 20.05.11)

Slobogin, C. (2007) *Privacy at Risk: The New Government Surveillance and the Fourth Amendment* (University of Chicago Press)

Smith, D. (2007) <http://observer.guardian.co.uk/amnesty/story/0,,2000545,00.html> (Last accessed: 01.03.09)

SMSR Ltd (2006) "Report on Information Commissioner's Office (Annual Track)

SMSR (2009) Report on the Findings of the Information Commissioner's Office Annual Track (Individuals) <http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/ico_annual_tracking_individuals_final_report2009.pdf >

Solove, D. J. (2002) "Conceptualising Privacy," *California Law Review*, Vol. 90, No. 4. pp. 1087-1155.

Solove, D. J. (2007) "I've Got Nothing to Hide and Other Misunderstandings of Privacy," *San Diego Law Review*, Vol. 44, pp.745-772

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

Sparkes, A. W. (1981) "The right to be let alone: a violation of privacy," *Bulletin of the Australian Society of Legal Philosophy* Vol. 58

Stern, S. R. (2003) 'Encountering distressing information in online research: a consideration of legal and ethical responsibilities,' *New Media & Society*, Vol. 5, No. 2 pp. 249-266

Stewart, J. (2011) 'Global data storage calculated at 295 exabytes,' <http://www.bbc.co.uk/news/technology-12419672>

Strahilevitz, L. J. (2005) "A Social Networks Theory of Privacy," The University of Chicago Law Review, Vol. 72, No. 3, pp. 919-988

Strauss, A. (1987) *Qualitative analysis for social scientists*, (Cambridge, UK: Cambridge University Press)

Strossen, N. (1990) "Recent United States and International Judicial Protection of Individual Rights: A comparative Legal Process Analysis and Proposed Synthesis," *Hastings Law Journal*, Vol. 41, pp. 805 - 904

Sturges, J.E. & Hanrahan, K. J. (2004) "Telephone and face-to-face interviewing," *Qualitative Research*, Vol. 4, No. 1, pp. 107- 118

Tashakkori, A. & Teddlie, C. (eds) Handbook of mixed methods in social research and behavioural research, (Thousand Oaks: Sage)

Tene, O. (2011) "Privacy: the new generations," *International Data Privacy Law*, Vol. 1, No. 1

Tesch, R. (1990) Qualitative research: Analysis types and software tools (New York: The Falmer Press).

Thomas, P. (1997) (ed) *Socio-Legal Studies*, (Aldershot: Ashgate-Dartmouth)

Tomkins, A. (1997) "Civil Liberties in the Council of Europe: A Critical Survey" in Gearty, C. (ed) *European Civil Liberties and the European Convention on Human Rights: A comparative study*, (The Hague, Martinus Nijhoff Publishers)

Transnational Data and Communications Report, (1989) "No Fiat for Fiat," November, p.10

Turnbull, G (2004) "The seven-year-old bloggers," <http://news.bbc.co.uk/2/hi/uk_news/magazine/3804773.stm> (Last accessed 16.03.07).

Twining, W. (1985) Karl Llewellyn and the Realist Movement, (London: Weidenfeld)

Van Selm, M. & Jankowski, N. W. (2006) 'Conducting Online Surveys,' *Quantity & Quality*, Vol.40, pp. 435-456

Vehovar, V. & Manfreda, K.L. (2008) 'Overview: Online Surveys,' in Fielding, N. *et al The Sage Handbook of Online Research Methods*, (Sage: London)

Viégas, F. B. (2005) "Bloggers' expectations of privacy and accountability: An initial survey" *Journal of Computer-Mediated Communication*, Vol. 10, No.3, article 12

Vitalis, A. (2008) 'France' in Rule, J. & Greenleaf, G. *Global Privacy Protection: The First Generation*, (Edward Elgar: Cheltenham, UK)

Wacks, R. (2010) Privacy: A very Short Introduction, (OUP, UK)

Walker, J. (2003) Weblog. In Definition for the Routledge Encyclopaedia of Narrative Theory

Walport, M. & Thomas, R. (2008) Data Sharing Review Report <http://www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf> (Last accessed 27.07.09)

Warren, S. & Brandeis, L. (1890), "The Right to Privacy" *Harvard Law Review* Vol. 4, No. 5, pp. 193-220

Weinstein, W.L. (1971) The Private and the Free: A Conceptual Inquiry, in Pennock J. R. & Chapman, J. W. (eds) Nomos XIII *Privacy* (New York)

Westin, A. F. (1967) *Privacy and Freedom* (New York: Atheneum)

Whyte, W. H. (1990). City: Rediscovering the Center (New York, N.Y.: Anchor.)

Wortham, J. (2009) More Employers Use Social Networks to Check Out Applicants, N.Y. Times (Aug. 20th) <http://bits.blogs.nytimes.com/2009/08/20/more-employers-use-social-networks-to-check-out-applicants/>

Wong, R. (2005) "Privacy: charting its developments and Prospects," in Klang, M. & Murray, A. *Human Rights in the Digital Age*, (Glasshouse Press: UK)

Wright, S. (2007) "Cyber stalker left me living in terror, says victim of 7/7." <http://www.dailymail.co.uk/news/article-458109/Cyber-stalker-left-living-terror-says-victim-7-7.html> <http://www.dailymail.co.uk/news/article-461450/Bloggers-help-track-cyberstalker-harassed-7-7-survivor.html> (Last accessed 16.07.09)

Younger, K. (1972) *Report of the Committee on Privacy*, (Chairman: Sir Kenneth Younger) (Cmnd 5012) United Kingdom

# Appendices

# Appendix A

## ICO placement interviews

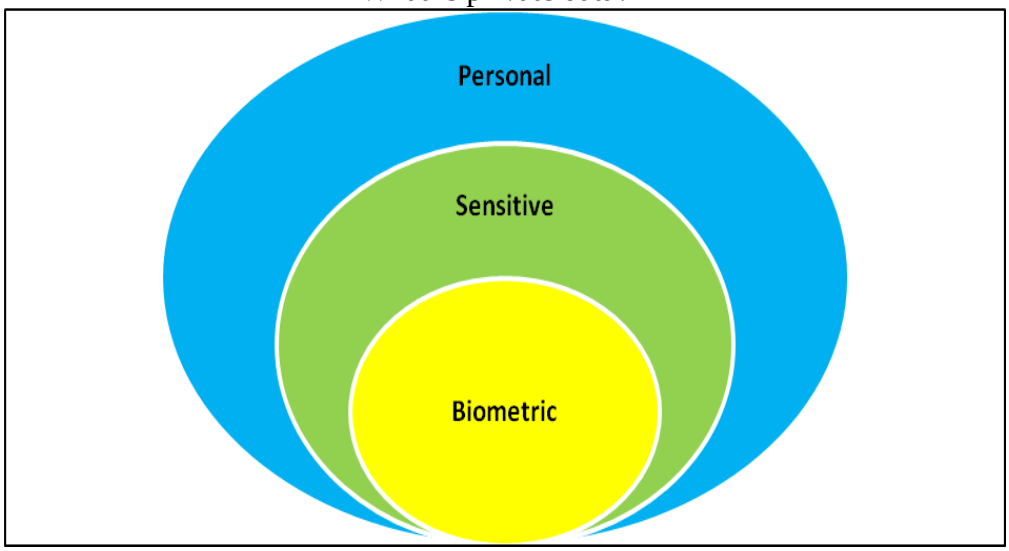
**Interview schedule:** the interview schedule comprises 5 main topic areas, namely general questions, previous research, biometrics, freedom of information and complaints and enquiry handling.

### General Questions:

What is privacy?

Is the model below an accurate representation of privacy?

What is private data?




Privacy in this country is currently protected by a combination of common law and statutory remedies? Is this appropriate?

The Data Protection Act 1998 and Human Rights Act 1998 afford protection to the private lives of individuals? Do you think this legislation is adequately protects privacy?

Are there gaps in the privacy protection available to individuals in this country?

How should these be remedied?

Are there any legislative reforms which you think should be introduced to protect privacy?

Is there a distinction between personal and private data?

It is arguable that Germany has interpreted the EU Directive 95/46 more strictly than the UK in implementing national data protection legislation? Why do you think this has occurred?

It is claimed that there are four types of privacy: (1) informational, (2) communication (3) territorial and (4) bodily privacy? Does UK law protect all these types of privacy?

Is the IC a reactive or proactive organisation?

Are the powers of the IC adequate?

Is the IC adequately funded?

Is the IC adequately staffed?

Will there be increased or diminished privacy for citizens in the future?

Will there be more privacy concerns in the future?

What issues do you expect to generate privacy concerns in the future?

Does the IC take a proactive or reactive approach to privacy protection?

Do you think the public are fully aware of privacy issues?

Is education of the public on privacy issues a priority for the IC?

## Previous research:

What qualitative research has been conducted by the IC?

Have you conducted concept based research?

Has the IC conducted quantitative research into privacy issues?

Has the IC conducted longitudinal research into privacy issues?

Has the IC been involved in trans-national research?

Do privacy concerns of UK citizens differ from their European counterparts? If so, why?

Are there reports available regarding previous studies?

Have you found socio-demographic differences regarding privacy concerns?

Have you found socio-demographic differences regarding levels of privacy concern?

How is this research linked to government policy?

Does this research inform government policy?

Does the government heed the opinion of the IC regarding privacy issues? Does the government adhere to recommendations?

Does the government try to influence the IC?

Does the government consult on privacy issues for the sake of consulting, rather than to gain insight into privacy issues?

## Biometrics:

Do biometrics raise new privacy issues?

Do different biometric techniques raise different privacy issues?

Are some biometric variables more sensitive than others?

Should a distinction be drawn between non-digital and digitised photographs?

Should a distinction be drawn between photographs and facial scans?

Is there a privacy issue regarding storage of biometric templates?

Do cryptographic algorithms offer adequate privacy protection to biometric data?

Under what conditions would you permit collection of biometrics by government?

Under what conditions would you permit collection of biometrics by the private sector?

Have you conducted research on public attitudes towards biometrics?

The IC has concerns regarding the proposed IC card and national identity register. In March 2003, 8 out of 10 respondents claimed to be in favour of ID cards? Does this worry you, if so, why?

In March 2003 67% of respondents claimed to have no knowledge of ID card proposals. Does this worry you?

Who should bear the responsibility of educating the public on biometric privacy issues?

Should biometric policy decisions be made at an international level, if so why?

## Freedom of Information:

Has the introduction of FOI legislation impacted on privacy?

Is it leading to changes in the concept of privacy?

## Complaint handling and Helpline enquiry process:

How are data protection enquiries/complaints generated?

How many complaints/enquiries are received on a daily basis?

What is the procedure for investigating a data protection complaint?

How does the IC respond to a complaint/enquiry?

Are there time limits for responding?

What issues generate most complaints?

Is there a record of all complaints?

How has the type of complaints changed over a period of years?

What organisations or types of industries are the subject of frequent complaints?

Do public or private sector organisations generate different levels of complaints?

If there is ambiguity in the legislation, how is this resolved?

What policy decisions underpin the handling of complaints?

May I review a case study – i.e. find out how a complaint is dealt with from beginning to completion?

Organisations seeking to process personal data are supposed to register with the IC. How many companies are on the list? Is it possible to see the list?

How often do people ask to consult the list?

How often is the list updated?

What new issues are the public raising through the helpline?

Do you often encounter organisations which are not on the list but which process personal data?

Do organisations respond positively to guidance issued by your office?

# Appendix B

## Interviews of Privacy & Data Protection Experts

**Interview schedule:**
The interview schedule comprises 3 main topic areas: firstly, discussions of the concepts of private data, data privacy, anonymity, personal and sensitive data; secondly, responses to a proposed definition of private data; and thirdly, questions on the limits of legal protections of privacy.

***Interview questions:***
1. a) Is the term private data formally defined in your country?

If so, please provide the formal definition below.

If defined, have there been any problems regarding interpretation or definition of the term private data?

1. b) If the term private data is not defined, what do you understand it to mean?

2. a) Is Data Privacy defined in your country? If so, please provide definition

2. b)If Data Privacy is NOT defined in your country, would it be useful to have a definition?

3. If Data Privacy is NOT defined in your country, please review the definition below:

Data privacy concerns the legal regulation of the boundary between personal and private data. Private data is regarded as a subset of personal data the disclosure of which could cause unreasonable harm to an individual.
It is the legal right of an individual to withhold consent to the collection, processing, communication or usage of personal data, the disclosure of which could cause unreasonable harm to an individual. Disclosure of private data should not be compulsory except where it is in the public interest, for instance if disclosure is in the interests of national security, public safety, the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health, or for the protection of the rights and freedoms of others or society as a whole.

Is this an appropriate approach to defining data privacy?

If this approach is appropriate, how should it be developed to provide a definition of privacy as the basis of law reform in this area?

If you do not consider it to be appropriate, how do you think data privacy should be defined?

4.Does anonmyisation (the removal of unique identifiers) of data alter its status from personal to not personal?

Is a distinction drawn between different methods of anonymisation?

If so, why?

5. Are the terms private and personal synonymous?

Should a distinction be drawn between personal and private data?

Does information have to be personal to be considered private?

If not, please give examples of private data that would not be classified as personal

Should a distinction be drawn between sensitive and private data?

6. Please indicate, giving reasons which statement you agree with most:

a) A definition of private data should only protect against unreasonable harm to an individual

Or,
b) A definition of private data should protect against any effect on an individual

Why did you choose A or B?

7. Should data privacy be a tradable commodity? For instance, should a person be able to sell their data to an organisation?

8. Should a person be able to waive their right to data privacy?

9. However defined, should the right to data privacy be an absolute right or subject to limitations?

What limitations do you think are most important?

10. Is there anything further you would like to add?

# Appendix C

## Telephone survey

**Methodology:** *ICO Annual track Individual survey question to test sensitivity of data*
The survey was conducted by telephone. All the interviews were conducted in house by SMSR's telephone interviewing team (a third party employed by the ICO). Subject contact data was collected via the Names and Numbers ADF software. The total sample was 1,066 interviews. Quotas were set on age, sex, region and social grade to ensure a nationally representative sample was achieved. The quotas were set as below:

Area Sampled

| **Area** | **Quota** | **Achieved** |
|---|---|---|
| North West | 11.4% | 10.9% |
| North East | 4.3% | 4.9% |
| Yorkshire and Humber | 8.4% | 8.7% |
| East Midlands | 7.1% | 7.8% |
| West Midlands | 9.0% | 9.1% |
| East | 9.2% | 8.8% |
| London | 12.2% | 11.6% |
| South East | 13.6% | 12.1% |
| South West | 8.4% | 8.4% |
| Wales | 4.9% | 5.4% |
| Scotland | 8.6% | 8.7% |
| Northern Ireland | 2.9% | 3.5% |

Age categories

| **Age** | **Quota** | **Achieved** |
|---|---|---|
| 18 - 24 | 11.0% | 10.7% |
| 25- 34 | 18.4% | 17.2% |
| 35 – 44 | 19.3% | 17.2% |
| 45 – 54 | 17.1% | 17.2% |
| 55 – 64 | 13.7% | 14.5% |
| 65 and over | 20.5% | 21.3% |

Social Economic Status of respondents

| **Social Economic Status** | **%** |
|---|---|
| AB | 21% |
| C1 | 26% |
| C2 | 21% |
| DE | 23% |
| Refused | 9% |

**Survey Questionnaire:**
The survey question was developed in a Likert-scale format to ensure compatibility with the other questions on the survey.

**Survey question:**
***Some types of personal information are considered 'sensitive' and given extra protection in law.***

***I am going to read out a list and I would like you to tell me, on a scale of 1 to 10, how sensitive you consider each one to be. 1 means not at all sensitive and 10 is extremely sensitive.***

READ OUT ONE AT A TIME

Personal contact details (e.g. home address, phone number)
Financial data (e.g. income and savings)
Data concerning race or ethnic origin
Criminal record
Biometric information (e.g. iris scans, facial scans and finger prints)
Political opinions
Membership of political party/organisation
Clickstream data (e.g. record of web pages visited)
Religious or philosophical beliefs
Genetic information
Health information
Sexual life information
Education qualifications
Employment history
Trade-union membership

**Purpose of question:**
The proposed question was designed to test both satisfaction with current definitions of sensitive data and also attitudes towards new categories of sensitive data raised in recent studies. Also, because the Annual Track survey is conducted on a yearly basis it could be used in the future to monitor changes over time.

Analysis of the survey responses allows an examination of any differences in attitudes towards sensitive personal information in relation to key demographics such as age, gender, marital status and employment status.

# Appendix D

## Online Survey of Bloggers

**Survey questions:**

1. What is the MAIN topic of your blog?

- ☐ My life (personal diary/journal)
- ☐ Politics and government
- ☐ Entertainment (movies, music, MP3 blogs)
- ☐ Sport
- ☐ News and current events
- ☐ Business
- ☐ Technology (computers, internet, programming)
- ☐ Religion/Spirituality/Faith
- ☐ A particular hobby
- ☐ Health (general health, an illness)
- ☐ Gossip
- ☐ Prefer not to answer
- ☐ Other, please explain:

2. Below are some reasons a person might blog. Please indicate if each one is a reason for YOUR blog, or not. If YES: Is it a MAJOR reason or only a MINOR reason?

| | **Main reason** | **Minor reason** | **Not a reason** | **Prefer not to answer** |
|---|---|---|---|---|
| To document your personal experiences and share them with others | ☐ | ☐ | ☐ | ☐ |
| To express yourself creatively | ☐ | ☐ | ☐ | ☐ |
| To influence the way other people think | ☐ | ☐ | ☐ | ☐ |
| To motivate other people to action | ☐ | ☐ | ☐ | ☐ |
| To share practical knowledge or skills with others | ☐ | ☐ | ☐ | ☐ |

| | | | | |
|---|---|---|---|---|
| To network or to meet new people | ☐ | ☐ | ☐ | ☐ |
| To entertain people | ☐ | ☐ | ☐ | ☐ |
| To discuss problems with others | ☐ | ☐ | ☐ | ☐ |
| To stay in touch with friends and family | ☐ | ☐ | ☐ | ☐ |
| To make money | ☐ | ☐ | ☐ | ☐ |
| To store resources or information | ☐ | ☐ | | |

3. On your blog, do you have?

| | **Yes** | **No** | **Prefer not to answer** |
|---|---|---|---|
| Text, e.g. written entries, articles or essays | ☐ | ☐ | ☐ |
| Photos | ☐ | ☐ | ☐ |
| Video | ☐ | ☐ | ☐ |
| Audio | ☐ | ☐ | ☐ |
| Images other than photos, e.g. drawings, graphs or clipart | ☐ | ☐ | ☐ |

4. Do you identify yourself on your blog (i.e. is your real name on the site)?

- ☐ Yes
- ☐ No
- ☐ On some blogs, but not all
- ☐ Prefer not to answer
- ☐ It is more complicated. Please explain:

5. How well do you feel you know your blog's audience?

- ☐ Extremely well
- ☐ Very well
- ☐ Quite well
- ☐ A little
- ☐ Not at all
- ☐ Prefer not to answer
- ☐ It is more complicated. Please explain:

6. Do you do anything to limit who can read your blog?

- ☐ Yes
- ☐ No
- ☐ Don't Know
- ☐ Prefer not to answer
- ☐ If yes, please explain:

7. How often do you

| | Always | Often | Sometimes | Hardly ever | Never | Prefer not to answer |
|---|---|---|---|---|---|---|
| Quote other people or media sources directly | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Post corrections to something you have written | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Take content from other sources and remix it into something new | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Discuss current events or news | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Include links to original source material you have cited or used | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

| | | | | | | |
|---|---|---|---|---|---|---|
| Spend time verifying facts you want to include in your post | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Respond to posts or comments from others | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Get permission to post copyrighted material | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

8. Below are a list of issues that could be considered of social importance. Please indicate how concerned you are about each issue by rating it on a scale of 1 to 5, where 1 is not at all concerned and 5 is very concerned.

| | **1** | **2** | **3** | **4** | **5** | **Don't know** | **Prefer not to answer** |
|---|---|---|---|---|---|---|---|
| Preventing crime | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | |
| Improving standards in education | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Protecting people's personal information | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Protecting freedom of speech | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Equal rights for everyone | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Unemployment | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Environmental issues | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Access to information held by public authorities | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Providing health care | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| National security | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

9. Below are a list of possible consequences that could result from organisations not handling your information responsibly. Please indicate how concerned you are about each issue by rating it on a scale of 1 to 5, where 1 is not at all concerned and 5 is very concerned.

| | 1 | 2 | 3 | 4 | 5 | Don't know | Prefer not to answer |
|---|---|---|---|---|---|---|---|
| Threat to personal safety | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Threat to your health | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Financial loss | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Indignity | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Loss of liberty | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Annoyance or inconvenience | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Invasion of privacy | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Personal distress | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

10. In Europe some types of personal information are considered 'sensitive' and given extra protection in law. Please read the list below and indicate on a scale of 1 to 10, how sensitive you consider each one to be. 1 means not at all sensitive and 10 is extremely sensitive.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Personal contact details (e.g. home address, phone number) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Financial data (e.g. income and savings) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Data concerning race or ethnic origin | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Criminal record | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Biometric information (e.g. iris scans, facial scans and fingerprints) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Political opinions | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Membership of political party/organisation | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Clickstream data (e.g. record of web pages visited) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Religious or philosophical beliefs | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Genetic information | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Health information | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Sexual life information | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Education qualifications | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Employment qualifications | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Trade-union membership | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

11. How often have you written about personal things in your blog?

- ☐ All the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ Rarely ever
- ◉ Never
- ☐ Prefer not to answer
- ☐ If Rarely ever or Never, please explain why:

12. Have you ever considered something 'too personal' to write about in your blog?

- [ ] Yes
- [x] No
- [ ] Prefer not to answer
- [ ] If yes, please explain what kind of information is 'too personal' to write about in your blog

13. Do you keep a traditional diary or journal (i.e. a written book that is not shared online?)

- [ ] Yes
- [x] No
- [ ] Prefer not to answer

14. If you have a traditional diary/journal and an online blog, do you decide to write something in your diary/journal but not to write about it on your online blog?

- [ ] Yes
- [x] No
- [ ] Prefer not to answer
- [ ] If yes, please explain:

15. Does 'private' information mean the same to you as information which is 'too personal' to write about in your blog?

- [ ] Yes
- [ ] No
- [ ] Prefer not to answer
- [x] If no, please explain what ‘private’ information means to you:

16. Has anyone ever invaded your privacy by mentioning you on their blog?

- [ ] Yes
- [ ] No

If yes, please explain what happened:

17. Generally, when you write things about people you know personally in your blog:

| | Always | Most of the time | Some of the time | Rarely ever | Never | I never write about people I personally know | Prefer not to answer |
|---|---|---|---|---|---|---|---|
| do you ask them permission to do so? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| do you reveal their full names? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| do you use an identifier instead of a name? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

18. Generally, when you write things about people (e.g. celebrities) you don't know personally in your blog:

| | Always | Most of the time | Some of the time | Rarely ever | Never | I never write about people I don't personally know | Prefer not to answer |
|---|---|---|---|---|---|---|---|
| do you ask their permission to do so? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| do you reveal their full names? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| do you use an identifier instead of name? | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

19. Recently daughters of Bobby Brown and Marie Osmond had their anonymous blogs identified. In your opinion, did this breach their privacy?

- ☐ Yes
- ☐ No
- ☐ Don’t Know
- ☐ Prefer not to answer

Comments:

20. Do you think MySpace.com was right to remove the blogs of Bobbi Kristina Brown and Jessica Osmond when their true identities were revealed?

- ☐ Yes
- ☐ No
- ☐ Don’t Know
- ☐ Prefer not to answer

Please explain why you said Yes or No:

21. What advice would you give to famous people who want to blog?

22. Do you ever blog about your work?

- ☐ Yes
- ☐ No
- ☐ Prefer not to answer

23. Have you or your friends ever gotten into trouble (e.g. upset family or been threatened with legal action) for anything posted on a blog?

- ☐ Yes I have
- ☐ Yes, my friends have
- ☐ No, neither I nor my friends have gotten into trouble
- ☐ Prefer not to answer

If yes, please explain:

24. Several people have been 'dooced' i.e. sacked from their job becasue they referred to work in their blog? Do you agree with this?

- ☐ Yes
- ☐ No
- ☐ Prefer not to answer
- ☐ Please explain why you said yes or no:

25. What gender are you?

☐ Male ☐ Female ☐ Prefer not to answer

26. In which age range are you?

Under 16
16- 18
19 - 24
25 – 34
35 – 44
45 – 54
55 – 64
65+
Prefer not to answer

27. What is your marital status?
Single
Married
Living with partner
Widowed /Divorced/Separated
Other
Prefer not to answer

28. Do you have any children under 18 living at home?
☐ Yes ☐ No ☐ Prefer not to answer

29. If you have children at home, what age are they?
☐ 0-3 ☐ 4-6 ☐ 7-10 ☐ 11-13 ☐ 14-15 ☐ 16-17 ☐ Prefer not to answer

30. Are you working? (If yes go to Q30, If No, go to Q31)
☐ Yes ☐ No ☐ Prefer not to answer

31. If you are working, is it full time or part time?
☐ Part time ☐ Full time ☐ Prefer not to answer

32. What is your job status?
Retired
At home raising family/house wife/house husband
Registered unemployed
Student in full time education
Other
Prefer not to answer
Don’t know

33. Are you the main income earner in your household?
☐ Yes ☐ No ☐ Prefer not to answer ☐ Don't know

34. What is your occupation?

35. What country do you live in?

36. Any other comments you'd like to make:

# Appendix E

## Socio-demographics: Blog Survey

Table: 1 - Sex Distribution

| Sex distribution of respondents | Percentage |
|---|---|
| Female | 49.1% |
| Male | 35.4% |
| Undisclosed | 15.5% |

(Blog Survey 2006, n = 1258)

Table: 2 - Age Bands

| Age distribution of bloggers | Percentage |
|---|---|
| Under 18 | 5.3% |
| 18-24 | 22.1% |
| 25-34 | 32.8% |
| 35-44 | 16.8% |
| 45-54 | 5.4% |
| 55-64 | 2.1% |
| 65+ | 0.6% |
| Refused | 14.9% |

(Blog Survey 2006, n = 1258)

Table: 3 – Relationship Status

| Relationship status | Percentage |
|---|---|
| Single | 39.7% |
| Married | 25.0% |
| Living with partner | 12.1% |
| Widowed / separated / divorced | 3.2% |
| Other | 2.5% |
| Refused | 17.6% |

(Blog Survey 2006, n = 1258)

Table: 4 - Country

| Country | Percentage |
|---|---|
| UK | 39.5% |
| Canada | 3.2% |
| Other EU & EEA | 8.4% |
| USA | 19.5% |
| Australia & NZ | 3.9% |
| Middle East | 2.9% |
| Asia | 2.7% |
| S. America | 0.8% |
| Africa | 0.8% |
| No Answer | 18.3% |

(Blog Survey 2006, n = 1258)

The country variable was recoded so that regions with different legislative approaches could be contrasted. Although a small response rate was recorded for Canada it was not appropriate to merge this country with the USA as they have very different approaches to privacy and data protection.

Table: Country recode

| Original Country | Recode Label |
|---|---|
| UK, Other EU & EEA | EU |
| Canada | Canada |
| USA | USA |
| Australia & NZ, Middle East, Asia, S. America and Africa | Other countries |

(Blog Survey 2006, n = 1258)

Table: 5 Country/ Region recode

| Country/Region | Percentage |
|---|---|
| EU | 48.3% |
| Canada | 6.5% |
| USA | 19.3% |
| Other Countries | 7.6% |
| No Answer | 18.3% |

(Blog Survey 2006, n = 1258)

Table: 6 Work Status

| Are you working? | Percentage |
|---|---|
| Yes | 60.7% |
| No | 21.8% |
| Refused | 17.5% |

(Blog Survey 2006, n = 1258)

Table: 7 Income Earner Status

| Are you the Main Earner? | Percentage |
|---|---|
| Main Earner | 31.5% |
| Not Main Earner | 44.0% |
| Refused | 23.6% |

(Blog Survey 2006, n = 1258)

# Appendix F

## Publications from thesis research

Mc Cullagh, K. (2009) "Protecting 'privacy' through control of personal data collection: a flawed approach," *International Review of Law, Computers & Technology*, Vol. 23, Nos 1-2, pp. 47-58

Mc Cullagh, K. (2008) "Blogging: Self-presentation and privacy," *Information & Communications Law*, Vol. 17, No.1, pp. 3-23

Mc Cullagh, K. (2007) "Data Sensitivity: proposals for resolving the conundrum," *Journal of International Commercial Law & Technology*, Vol. 4, No. 2, pp.190-201

*International Review of Law, Computers & Technology*
Vol. 23, Nos. 1–2, March–July 2009, 13–24

Routledge
Taylor & Francis Group

# Protecting 'privacy' through control of 'personal' data processing: A flawed approach

Karen McCullagh*

*Salford Law School, University of Salford, UK*

The development of a frontier-free internal market and of the so-called 'information society' have resulted in an increase in the flow of personal data between EU member states. To remove potential obstacles to such transfers, data protection legislation was introduced. One of the underpinning principles of Directive 95/46/EC is the protection of privacy. Yet, the legislation does not provide a conclusive understanding of the terms 'privacy' or 'private' data. Rather, privacy protection is to be achieved through the regulation of the conditions under which personal data may be processed. An assessment of whether, 10 years after the enactment of the Data Protection Act 1998 (DPA 1998), a coherent understanding of the concept of personal data exists, necessitated an analysis of the decisions in *Durant* v. *FSA* ([2003] EWCA Civ 1746) and *CSA* v. *SIC* ([2008] 1 WLR 1550, [2008] UKHL 47). Furthermore, in order to examine the effectiveness of the legislation, this article examines whether the term 'personal' is synonymous with the term 'private' data and whether control over processing of personal information protects privacy. By drawing on interviews with privacy and data protection experts, and from the findings of a survey of bloggers, it will be shown that a review of the assumptions and concepts underpinning the legislation is necessary.

**Keywords:** data protection; personal; private

## Introduction

As IT usage and processing capabilities evolve, regulators, privacy practitioners and citizens are increasingly questioning the suitability and adequacy of data protection legislation to allow the effective processing of personal data while simultaneously safeguarding the privacy of individuals. Indeed, the Office of the UK's Information Commissioner (ICO) commissioned research into how the EU Directive 95/46/EC should be updated, because

> We want to generate new thinking. European data protection law is increasingly seen as out of date, bureaucratic and excessively prescriptive. It is showing its age and is failing to meet new challenges to privacy, such as the transfer of personal details across international borders and the huge growth in personal information online.[1]

This article begins by exploring the relationship between privacy and data protection at EU and UK level. It will be demonstrated that the concepts of 'privacy' and 'private' data remain nebulous as they are not defined in the Directive. Instead, the Directive provides

*Email: k.mccullagh@salford.ac.uk

ISSN 1360-0869 print/ISSN 1364-6885 online

DOI: 10.1080/13600860902742562
http://www.informaworld.com

a definition of 'personal' data and stipulates the conditions under which such data may be processed. Thus, this research explores the meaning of the term 'personal' data by reviewing the cases of *Durant* v. *FSA*[2] and *CSA* v. *SIC*.[3] Also, the article assesses whether control over personal information protects privacy by drawing on interviews with privacy and data protection experts from a range of countries and disciplines. Furthermore, the views of potential data subjects are explored through a survey of bloggers, which reports their conceptions of the terms 'private' data. The article concludes that a review of data protection legislation is necessary, and that, in particular, the assumptions and concepts underpinning the term 'personal' and 'private' need to be revised.

## Privacy in Directive 95/46/EC

The goal of privacy protection is expressly stated in the opening provisions EU Directive 95/46/EC, wherein Art. 1 states that the objective is:

> to protect the fundamental rights and freedoms of natural persons and in particular their right to *privacy*, with regard to the processing of personal data. (Emphasis added)

However, the term privacy is not defined in the Directive. Rather, the Directive seeks to achieve privacy protection through regulation of the processing of personal data. This is understandable because no adequate definition of privacy has ever been produced.[4]

## Personal data defined

The Directive prohibits, subject to exhaustively listed exceptions, the collection and processing of personal data. In the DPD, personal data is defined in Art. 2 (a) as:

> Any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, psychological, mental, economic, cultural or social identity.

Whereas, personal data is defined in the Data Protection Act 1998, as:

> Data which relate to a living individual who can be identified:
>
> (a) from those data, or
> (b) from those data and other information which is in the possession of, or is likely to come into the possession of the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

Korff[5] noted that the DPA 1998 makes a formal distinction between 'data' and 'information' but that in practice, it produced no material differences. The definitions in the Act and Directive are consistent in their use of the phrase 'relate to' but, under the Directive, consideration is first directed to whether the information relates to an identifiable individual and then whether it is processed. Whereas, the definition of personal data in the Act approaches the concept in the reverse order, as the Act focuses on the issue from a processing view first and then moves on to whether or not there is an identifiable individual. The Directive and Act also differ with respect to when an individual should be considered as 'identifiable'.

The term is wider under the DPA as it includes information that 'may' come into the possession of the controller. Korff argues that the phrase may mean that the definition of personal data in the Directive can be read as being 'relative', because potentially 'any data that can conceivably be linked to an individual (in whatever way, by whoever) [can] be regarded as personal'.[6] Booth et al.[7] observed that the way that the phrase 'relate to' is interpreted has major implications regarding what is or is not classed as personal data. If it is interpreted very narrowly, the term personal data could be restricted to data which is capable of identifying an individual, either by itself or in combination with other data. Identification, in this context, could be direct or indirect. In contradistinction, if the term 'relating to' is interpreted broadly it could conceivably include any data which may 'affect' the individual in some way, regardless of its capacity to identify. The consequences of a narrow interpretation of 'relating to' will be explored in an analysis of the Durant decision.

**UK interpretation of 'personal' data**

In the case of *Durant* v. *FSA*,[8] Mr Durant had lodged a complaint with the Financial Services Authority following a legal dispute with Barclays bank. The FSA dismissed his complaint. He then made a subject access request for information held manually and electronically by the FSA on his complaint. The FSA released the information held in computerised form, but refused to disclose the information held on manual files. Mr Durant applied to the Court under s 7(9) of the DPA 1998 for an order requiring the FSA to comply with the subject access request. The Court of Appeal was asked to decide: was the information held by the FSA relating to the investigation of Mr Durant's complaint 'personal' data under the Data Protection Act 1998? The definitional issue which arose concerned whether the data could be said to 'relate to' Mr Durant.[9]

Mr Auld LJ referred to Directive 95/46/EC and ruled that the statutory right of access under the DPA is designed to enable the data subject to:

> check whether the data controller's processing of it unlawfully infringes his privacy and, if so, to take such steps as the Act provides . . . to protect it.[10]

From this the Court concluded that the relevant information is:

> information that affects [the data subject's] privacy, whether in his personal or family life, business or professional capacity.[11]

This interpretation of personal data means that not all identifying information will fall within the scope of 'personal' data. Rather, only information that is capable of adversely affecting the privacy of the data subject will be considered personal. In order to determine whether or not data 'relates to' the data subject, Auld LJ proposed two tests.The first test is:

> whether the information is *biographical* in a significant sense, that is, going beyond the recording of the putative data subject's involvement in a matter or an event that has no personal connotations, a life event in respect of which his privacy could not be said to be compromised.[12] (Emphasis added)

The second test is whether:

> the information has the putative *data subject as its focus* rather than some other person with whom he may have been involved or some transaction or event in which he may have

> figured or have had an interest, for example, . . . an investigation into the some other person's or body's conduct that he may have instigated.[13] (Emphasis added)

Buxton LJ agreed, stating that the potential effect of processing of particular data on an individual's privacy was the guiding principle. The Court also drew support for a narrow interpretation of the term personal data from the wording of the DPA 1998. Auld LJ asserted that the DPA's definition of personal data extends to expressions of opinion about an individual which would be otiose if the words 'relate to' were construed broadly. Thus, the Court of Appeal ruled that the information about Mr Durant's complaints to the FSA or about their investigation of his complaint were not 'personal data' as the data did not relate to Mr Durant in the requisite sense. Rather, the Court decided that the information sought by Mr Durant was information about his complaints, as opposed to data relating to him. Furthermore, the Court ruled that the mere fact that a document is retrievable by reference to the name of the data subject does not render the information personal data:

> Whether it does so in any particular instance depends on where it falls in a continuum of relevance or proximity to the data subject.[14]

Thus, simply because an individual's name appears on a document, the information contained in that document will not necessarily be personal data about the named individual. Rather, it is more likely that an individual's name will be 'personal data' where the name appears together with other information about the named individual such as address, telephone number[15] or information regarding his hobbies.[16]

This conception of the term personal data is very narrow. If this decision were to be followed, only information that is capable of adversely affecting the privacy of the data subject would be considered personal data. Subsequently the Art. 29 Working Group issued an opinion on the concept of personal data,[17] which contains a broader notion of personal data. Thereafter, the Office of the UK Information Commissioner issued a technical guidance note[18] to the effect that Durant is relevant to the question of whether data 'relates' to a living individual only in difficult cases where the information in question is not 'obviously about' someone. However, the ICO guidance note is not legally binding, as the Durant decision has not been overruled.

In the case of *CSA* v. *SIC*,[19] a researcher submitted a request under the Freedom of Information (Scotland) Act 2002 (or 'FOISA')[20] to the Common Services Agency (the 'CSA'),[21] for details of the recorded incidence of childhood leukaemia for certain years in a geographical area, broken down by census ward. The researcher wanted to explore a suspected risk to public health arising from the Ministry of Defence's operations, a decommissioned nuclear reactor and an operational nuclear processing facility. The CSA refused to disclose the information on the grounds that it was personal data, the disclosure of which would breach the data protection principles. On application to the Scottish Information Commissioner (the 'SIC'), the SIC ordered the CSA to disclose the information sought in an anonymised form using a technique called 'barnardisation' which perturbs the dataset in order to substantially reduce the risk that individual data subject could be identified from it. The case raised the importance of whether or not the barnardised information was 'personal data' within the meaning of the DPA 1998.

The Lords ruled that the barnardised data was information about the health of the children involved. It therefore obviously related to the children and there was therefore no need to turn to the Durant decision and its concepts of 'focus' and 'significant biographical data', to decide whether the definition of 'personal data' was satisfied.

The second issue which arose was whether any of the children could be identified from the barnardised information (either alone or taken together with other information in the possession, or likely to come into the possession, of the CSA). The Court unanimously ruled the fact that the CSA continued to hold 'other information' which would ultimately have allowed it to 'decode' the barnardised information to identify each of the children to whom it related, did not necessarily mean that the barnardised information was still personal data. However, several different rationales can be identified from the judgment.

Lord Hope took the view that data can be 'fully anonymised' in the hands of the data controller and thereby cease to be personal data, even where the data controller does have information which would theoretically allow it to unlock the identities of the subjects of that data, but did not explain exactly how, or, in what circumstances that anonymisation might be achieved. Lord Rodger thought that data would remain personal data in the hands of the data controller provided that the data controller could identify the subjects of that data using 'reasonable means'. However, the practical implications of that reasoning are not clear. In marked contrast, Baroness Hale focused instead on the proposed recipient of the data, and whether he or should could identify the subject(s) of that data from that data alone (given that he or she would not have access to any of the 'other information' in the hands of the disclosing data controller). This lack of unanimity appears to have arisen from the difficulty which their Lordships faced in reconciling the definition of 'personal data' in the DPA with the spirit of Directive 95/46/EC and in particular with Recital 26 of the Directive which states that 'the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable'. Indeed, Baroness Hale stated that '[while their Lordships would] all like the legal position to be that, if the risk of identification [of the children] can indeed be eliminated, the Agency is obliged to provide [the information requested]',[22] in line with the 'expectation in Recital 26', she had 'much more difficulty in spelling out [that conclusion] from the definition of "personal data" in section 1(1) of the Act'.[23] The foregoing analysis indicates that the attempt to protect privacy through regulation of processing of personal data is fraught with difficulties, due to the confusion surrounding the concept. The decision does not clarify how the 'identifiability' requirement should be interpreted and applied in future cases. Also, questions remain as to precisely what factors are to be taken into account in determining when data can be said to be 'fully anonymised' and as such, no longer personal data.

## Relationship between 'personal' and 'private' data

Moreover, the absence of a concept of 'private' data in the Directive and DPA 1998 and the fact that privacy protection is to be achieved through the regulation of the conditions under which personal data may be processed, *prima facie* suggests that the terms 'personal' and 'private' are synonymous, or alternatively, that protection of personal data effectively protects privacy. Yet, there is a lack of research data the effectiveness of the model of privacy protection advocated in the Directive. This article seeks to remedy that deficiency by reporting the findings of interviews with privacy and data protection experts, and also the responses from a survey of bloggers.

## Research method

To answer the questions posed, data was collected in two phases. First, semi-structured interviews[24] were conducted with forty privacy and data protection experts, namely: privacy commissioners, lawyers, corporate privacy officers, consultants, computer

scientists and academics from sociology, politics, market research, statistics and law.[25] The second phase of the data collection process consisted of an online survey of bloggers from around the world.[26] Out of the total number (1314) of responses received, 1258 were selected for data analysis; the remainder of responses were incomplete and were disregarded. However, the resulting population of participants does not qualify as a random sample and, accordingly the results from this survey cannot be generalised to the entire blogging population.[27] Rather, the findings are representative of certain niches of the English-speaking blogging world.

## Research findings

### *(1) Views of data protection and privacy experts*

All respondents were familiar with the concept of personal data as they had knowledge of the EU Directive 95/46/EC or it had been implemented into their national legislation, but indicated difficulties in drawing the lines between 'personal' and 'not personal' data. Some discussed the fact that technological developments are causing difficulties, e.g. advances in genetics are leading to greater pressure to collect health data and, while this is often stored and processed in the form of 'coded' data, there is a lack of clarity whether such data should be considered personal data. Another example cited was transaction data/behavioural data on the internet, e.g. clickstream data can lead to a profile being created which may, or may not, be considered personal data. When asked whether the concepts of personal and private data are synonymous, a range of responses were recorded. They have been classified under four broad headings:

(i) *Private concept not legally recognised. Informally it is synonymous with personal data*

> I think in a legal sense – in a data protection sense, yes (the terms private and personal are synonymous). However, privacy protection and data protection are different, but in a colloquial sense they are synonymous. (Belgium)

> In our law the word 'private' isn't even used, so it doesn't have a legal meaning. The general population take them to mean the same thing. (Canada)

The experts drew a distinction between personal data that is protected in law and private data that is not legally recognised, but which in the mind of the general public, is a synonymous term. When asked to elaborate on the concept of private data, they stated:

(ii) *Private data is a subset of personal data that the individual wants to keep absolutely secret*

> Private data is the part of the personal data that the respondent does not want to make public. (Spain)

(iii) *Private data is a subset of personal data that the individual wants to control access to or reveal in limited circumstances*

> It is something not revealed to others, or only revealed to a select group. It is a concept close to confidentiality but without the legal connotations, e.g. disclosure to a family member/bank/personnel office e.g. my salary would be considered private. (New Zealand)

> Personal data is data relating to an individual . . . . Private data is something you want to keep to yourself or something that people need to seek your permission to give out. (Australia)

These responses imply that individuals will face choices regarding disclosure of information and that the individual providing the data should decide the nature and extent of disclosure. These responses reflect the informational control conception of privacy espoused by Westin, who 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.[28]

Similarly, in the German census case[29] the notion of 'informational self-determination' was advocated. The German Constitutional Court ruled that

> This basic right warrants in this respect the capacity of the individual to determine in principle the disclosure and use of his/her personal data. Limitations to this informational self-determination are allowed only in case of overriding public interest.[30]

However, while the experts identified private data as a subset of personal data that merits extra legal safeguards in order to protect privacy, they did not offer clearly delineated boundaries for this term, which could be of general application. Instead, the comments below illustrate that, in their experience, claims that information is 'private' arise on an *ex post facto* as opposed to *ab initio* basis.

(iv) *All personal data can be private, depending on contextual factors*

> They are not synonymous. Private, is an *ex post facto* term, used mainly to label those claims for non-disclosure that we've accepted on other contextual grounds. Whereas, the term personal, concerns information about which less contested claims are made, e.g. the personal fact that I'm bald and short-sighted is personal data but it is hardly a private fact. (UK)

> Personal data can become private. . . . Some pieces of data we don't want to go elsewhere are what we consider private – but it depends on the company, e.g. happy for A to know but not B to know. (UK)

> If data is generally personal, it may become private depending upon place, time and circumstances . . . in different circumstances people see the same data differently, therefore, it is very difficult to define this kind of data. For example, if we approach our bank manager for a loan then we will be willing to discuss our salary but, in other circumstances you won't tell someone your salary. (Finland)

The responses from the data protection and privacy experts embody the philosophical ideals of autonomy and dignity through 'informational self-determination'. The experts recognise that a data subject should have the right to a degree of control over information that identifies them or relates to them, since control over disclosure identifying information is necessary for the development of autonomous individuals. They further acknowledged that it is not possible to predict in advance what personal information will be claimed as private by a data subject, since such claims are usually made on an *ex post facto* basis, depending on contextual factors.

### *(2) Views of bloggers*

In the survey, almost one-quarter (24.8%) of respondents said that they had posted personal information on their blog 'all the time'.[31] However, bloggers seem to reflect regularly on the content of posts when deciding whether or not to post personal information online. Most

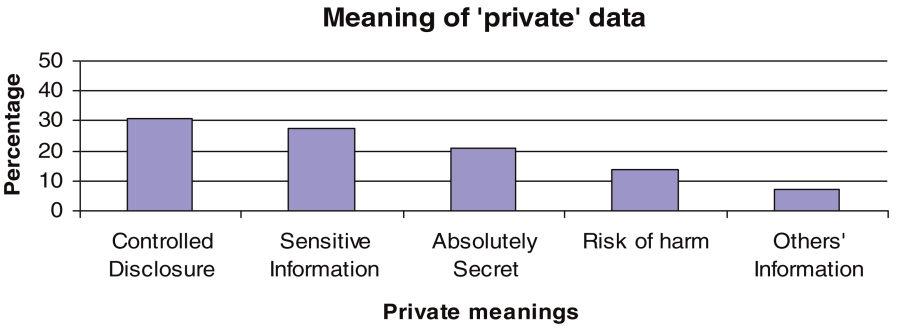



Figure 1. Graph illustrating meaning of 'private' data (source: Blog Survey 2006, $n = 101$).

respondents (65.6%) said they had considered some information 'too personal' or 'private' to write about on their blogs.[32] Respondents who stated that 'private' data did not mean the same as information which is 'too personal' (36.5%) to write about in their bogs were asked to explain what private information meant.

The largest percentage (30.7%) equated private with '*controlled disclosure*'. It is information which the individual wants to control access to, or disclose in limited circumstances. Below are some illustrative comments:

> Some information that I might want to discuss with only a select few .... Then I would make that entry a secure one, so that only those people belonging to that group could read and comment. It is not information that I consider public, but neither is it too personal to share.

> Private' information varies – there's stuff you'd share with friends, then only close friends, or nobody at all.

These responses mirror the responses of the experts and reflect the informational control conception of privacy, as they indicate that bloggers are aware that they constantly face choices about the nature and extent of information disclosures they make on their blog posts.

Private data was equated with either legally recognised or potential new categories of '*sensitive data*' by just over one-quarter (27.6%) of respondents, as illustrated by comments

> Private information is data about me as an individual such as biometrics, financial, political beliefs etc. Things which are too personal are to do with relationships with other people, etc.

> Private information, to me, describes data (financial information, phone number, etc.), whereas 'too personal' describes emotional information (how I felt about something my friend said last week).

These responses encapsulate the definition of privacy offered by Innes who stated that it is 'the state of possessing control over a realm of intimate decisions, which includes decisions about intimate access, intimate information and intimate actions'.[33] According to this view of privacy, not every disclosure of information about a person will amount to a loss of privacy; there will only be a loss when 'sensitive' or 'intimate' personal information is disclosed.[34] The responses indicate that bloggers are aware of such distinctions and actively limit disclosure of such information on their blogs.

It is information which the individual wants to keep '*absolutely secret*' was the meaning attributed to 'private' information by one fifth (20.8%) of respondents, as illustrated by comments:

> Private is information that I don't want to discuss with anyone.
>
> Information I would not trust other people with. It doesn't have to be 'personal' in the sense of intimate. It could be financial details, for instance.

A minority of bloggers (13.9%) equated 'private' data with the '*risk of harm*' arising from data uses:

> Private information, when it comes to an on-line environment, refers to any data which a third party having knowledge of could cause me actual harm, whether financial, or by restrictions of civil liberties. This therefore refers to my financial details, address and contact details (though to a lesser degree).
>
> Private information is information I don't want total strangers to have (e.g. home address), or that could be used to cause me a harm (monetary or otherwise, e.g. credit card and PIN numbers).

'*Information about others*' was the meaning attributed to private data by the lowest number of respondents (6.9%):

> This is information like the names of people I am writing about if they are not in the public arena or making them identifiable in some other way if I haven't asked.
>
> I don't write things that are too personal about my friends and family. I don't paraphrase them or post IM [Instant Messenger] conversations without running it by them first. I don't use other people's first names unless they have their own blog that they have given me permission to link to where they use their own first names.

These responses also indicate that bloggers are aware that the notion of privacy hinges on the concepts of control and consent regarding disclosure. Thus, each individual should decide for themselves the nature and extent of information which is disclosed. Also, there may be circumstances in which an individual does not have direct control over their personal information, but their privacy is nonetheless respected, e.g. a blogger does not post information about friends without express or implied consent. Indeed, the responses indicate that bloggers are aware that some information is shared in the context of a variety of relationships and that maintaining confidentiality and secrecy in respect of such information is a necessary requisite for healthy functioning relationships. These responses fall within the 'intimacy' conception of privacy espoused by Fried who, defined privacy as the 'control over knowledge about oneself'.[35] He based his definition of private information on intimate relationships by asserting that privacy should be valued because it is necessary to protect 'fundamental relations' of 'respect, love, friendship and trust'.[36]

### Criticisms of approach in Directive 95/46/EC

The complexity surrounding the concepts of privacy, private and personal data is evidenced by the responses from experts and bloggers. The responses reveal that there are no consistently agreed meanings attached to any of the terms, and indeed, these terms are often used interchangeably and in an overlapping fashion. One reason for this is that the data

protection principles are stated in broad, general terms, rather than in specific terms capable of precise legal delineation. However, Art. 22 of Directive 95/46/EC requires EU member states to provide a right to a judicial remedy for a breach of any of the rights guaranteed by the Directive. This means that UK Courts must necessarily confer some precise meaning on the general principles. The *Durant* decision illustrates the difficulties posed by this requirement. In that case the Court attempted to apply a purposive approach – asserting that because the purpose of the access right is to protect the privacy of the data subject, it is only information that is relevant to that purpose which can be subject to the access right. However, this approach is artificial and unhelpful, as it misconceives the role of personal data in determining the scope of privacy within data protection legislation. It also fails to recognise that data protection legislation serves other interests, e.g. data accuracy and data quality.

**An alternative 'harm' based approach to privacy protection**

The responses from the data protection and privacy experts embody the philosophical ideals of autonomy and dignity through 'informational self-determination'. The experts recognise that a data subject should have the right to a degree of control over information that identifies them or relates to them, since control over disclosure identifying information is necessary for the development of autonomous individuals. However, although control may in fact protect privacy in many circumstances, equating control with privacy is not always effective. For instance, individuals may be provided with control and subsequently decide to give up their privacy. Alternatively, once information is shared with another, e.g. in the course of a friendship or business transaction, an individual no longer has exclusive control over the disclosure of the information. Yet, the individual's privacy may (as a matter of good customer relations, or in the interests of sustaining a friendship), or may not, be protected in the absence of direct control over the information. Moreover, the responses by bloggers and experts acknowledge that it is not possible to predict in advance what personal information will be claimed as private by a data subject, since such claims are usually made on an *ex post facto* basis, depending on contextual factors. Accordingly, some of the experts were critical of the underlying approach of the Directive, claiming that this model of privacy protection which is based on collection limitation principles is outdated. They assert that it regulates at the wrong level and fails to balance competing interests properly. The Directive regulates the collection and processing of data, as opposed to regulating specific harmful uses of the data:

> There is a realisation that information is gathered, collected. It [data collection] is ubiquitous and [it is] impossible to chase wrongful collection; therefore, the focus has shifted towards a harm-based approach. (USA)

Accordingly, they contend the focus of regulatory activity should shift. It should centre on harm related to the misuse of personal information.

> Regulation of collection is a losing battle – instead ensure it is not used malevolently – information will always need to be collected, so you need to focus work on how it is used. (Australia)

Both bloggers and privacy experts recognised that personal data can be used or misused. The interchangeable and overlapping uses of the terms personal and private data by bloggers indicate that the focus of data protection legislation has erroneously been on the

categorisation of data into personal or sensitive data, and the limitation of collection of such data, instead of the harm arising from data uses.

## Conclusion

The foregoing analysis indicates that the decision in *Durant* v. *FSA* is at odds with the general principles of data protection. It attempts to limit the scope of personal data. While this approach is *prima facie* useful from privacy perspective, it fails to recognise that data protection legislation also serves other interests and that a broader interpretation of personal data is necessary to achieve these purposes. This failure undoubtedly reflects the notorious difficulties that have plagued attempts to give privacy a precise, analytically serviceable and universally accepted meaning. The failure to define 'private' data in data protection laws has a cost, in so far as it detracts from the capacity of those laws to offer prescriptive guidance. A further cost is that it perpetuates the vulnerability of the privacy concept to the criticisms that it is incapable of definition, has no independent, coherent meaning and should be subsumed by other concepts.

It is suggested that the time is ripe to review the provisions of the Directive, as the focus of the current legislative model is erroneously on the categorisation of data into personal and sensitive data, and the application of different levels of privacy protection to the different categories of data. The responses of the experts and blogggers indicate that, in the information society, the notion of privacy has changed. In this era, privacy is the absence of harmful use and application of information about an individual. As the UK Information Commissioner stated, the Directive 'out of date, bureaucratic and excessively prescriptive. It is showing its age and is failing to meet new challenges to privacy, such as . . . the huge growth in personal information online.' This paper echoes the Commissioner's call for a review of the legislation. In particular the interpretation of the concept of personal data should be reviewed. It is suggested that it should receive a 'broad' interpretation and the question of when information is 'identifiable' should be answered using a risk of re-identification approach. Also, the concepts of consent and control should be revisited. Further research is needed on the concept of consent. It may be worthwhile developing a test for implied consent in order to achieve a balance between privacy interests and the legitimate interests of others. Also, future legislation could focus on regulation of specific harmful uses of personal data and the availability of appropriate remedies.

## Notes

1. Information Commissioner's Office (ICO), 'UK Privacy Watchdog Spearheads Debate on the Future of European Privacy Law', ICO, Wilmslow, Cheshire, 2008. Available at http://www.ico.gov.uk/upload/documents/pressreleases/2008/ico_leads_debate_070708.pdf (accessed February 25, 2009)
2. *Durant* v. *FSA* [2003] EWCA Crim 1746.
3. *CSA* v. *SIC* [2008] 1 WLR 1550, [2008] UKHL 47.
4. The issue has occupied the minds of scholars and jurists alike for decades. It is a common feature of any privacy analysis to start with a disclaimer about the inherent difficulty or impossibility of defining exactly what privacy is or, of dissecting the concept into its various components. While the definitions espoused by Judge Cooley, Samuel D. Warren and Louis D. Brandeis, Alan Westin have a certain intuitive appeal, none have become universally accepted.
5. D. Korff, 'Comparative Study of National Laws', EC Study on the Implementation of Data Protection Directive, 2002, http://ec.europa.eu/justice_home/fsj/privacy/docs/lawreport/consultation/univessex-comparativestudy_en.pdf
6. Ibid.

7. S. Booth, R. Jenkins, D. Moxon, N. Semmens, C. Spencer, M. Taylor and D. Townend, 'What are "Personal Data"? A Study Conducted for the UK Information Commissioner', 2004, http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/final_report_21_06_04.pdf
8. *Durant* v. *FSA* [2003] EWCA Crim 1746.
9. Identifiabililty was not an issue because the information in the manual files essentially comprised letters of complaint written by Mr Durant and material generated in response to his complaint.
10. *Durant* v. *FSA* [2003] EWCA Crim 1746, [27].
11. Ibid., [28]
12. Ibid.
13. Ibid.
14. Ibid.
15. See European Court of Justice decision in *Bodil Lindqvist* v. *Kammaraklagaren* (2003) C-101/01, para. 27, as referred to in para. 28 of the Durant judgment.
16. See *Bodil Lindqvit* v *Kammaraklagaren* (2003) C-101/01, para 27, ECJ Judgment, http://eur-lex.europa.eu/LexuriServ.do?uri=CELEX:62001J0101:EN:HTML (accessed February 25, 2009).
17. Art. 29 Data Protection Working Party, 'Opinion 4/2007 on the Concept of Personal Data', 2007, http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf)
18. ICO, 'Data Protection Technical Guidance – Determining What is Personal', v1.0, 21 August 2007, http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/personal_data_flowchart_v1_with_preface001.pdf
19. *CSA* v. *SIC* [2008] 1 WLR 1550, [2008] UKHL 47.
20. The corresponding provisions of the Freedom of Information Act 2000 are in materially the same terms and the judgment is therefore of relevance throughout the UK.
21. A specialist health board in Scotland which collects statistical information from other health boards.
22. *CSA* v. *SIC* [2008] UKHL 47 [91].
23. Ibid., [92].
24. In the interviews semi-structured questions were used. The aim was to have a discussion with the respondent so that all the themes in the interview guide were covered. Some of the themes in the interview guide were too complex for a few of the participants. For instance, statistical methodologists were not comfortable when answering questions about the specific detail of the legislation in their country.
25. A respondent matrix was created using quota and snowball sampling. Snowballing is an effective technique for building up a reasonable sized sample, especially when used as part of a small-scale research project (M. Denscombe, *The Good Research Guide for Small Scale Social Research Projects* (Milton Keynes, UK: Open University Press, 1998).
26. The respondents were not randomly selected but were found through a variant of the snowball-sampling strategy. Announcements for the online survey were posted to mailing lists in three universities in the UK area as well as on a few high-traffic blogs. The viral nature of blogs meant that the links to the survey page quickly spread to many other blogs and YouTube.
27. It is the bloggers' subjective sense of privacy and liability that is revealed. This self-disclosure approach has three important implications: (1) There can be disparities between stated privacy attitudes and actions; (2) Participants' perceptions of their blogs might differ from those of outside observers and researchers; (3) accuracy is difficult to verify, e.g. no external validation was conducted.
28. A.F. Westin, *Privacy and Freedom* (New York: Atheneum Press, 1967), 7.
29. BVerfGe 65, 1. 'Volkszählungsurteil', 1983, http://www.datenschutz-berlin.de/gesetze/sonstige/volksz.htm
30. BVerfGE 65, 1. 'Volkszählung', 1983.
31. Only 2% of respondents said they had 'never' posted anything highly personal on their blogs.
32. It is important to note is that a consensus of definition does not exist regarding the terms private or personal. Indeed many used the terms interchangeably.
33. J.C. Innes, *Privacy, Intimacy, and Isolation* (New York: Oxford University Press, 1992), 140.
34. Ibid., 58.
35. C. Fried, 'Privacy', *Yale Law Journal* 77, no. 3 (1968): 483.
36. Ibid., 477.

*Information & Communications Technology Law*
Vol. 17, No. 1, March 2008, 3–23





# Blogging: self presentation and privacy

Karen McCullagh*

*Salford Law School, Salford, UK*

Blogs are permeating most niches of social life, and addressing a wide range of topics from scholarly and political issues[1] to family and children's daily lives. By their very nature, blogs raise a number of privacy issues as they are easy to produce and disseminate, resulting in large amounts of sometimes personal information being broadcast across the Internet in a persistent and cumulative manner. This article reports the preliminary findings of an online survey of bloggers from around the world. The survey explored bloggers' subjective sense of privacy by examining their blogging practices and their expectations of privacy when publishing online. The findings suggest that blogging offers individuals a unique opportunity to work on their self-identity via the degree of self-expression and social interaction that is available in this medium. This finding helps to explain why bloggers consciously bring the 'private' to the public realm, despite the inherent privacy risks they face in doing so.

**Keywords:** blogging; personal information; privacy; private data; survey

## Introduction

In this article, I begin by explaining the technological phenomenon known as blogging. I then provide background information on privacy issues in relation to blogging, as well as exploring a number of conceptions of privacy, before electing to use DeCew's cluster concept of privacy as a framework for testing the subjective privacy attitudes and expectations of bloggers. Thereafter, the findings of a survey that explores bloggers' privacy attitudes and expectations are presented. Finally, the concluding remarks summarize the major findings and point to the need for further work in this area.

## Blogs from a privacy perspective

A fundamental difference between blogs and other web-based publishing sites, such as personalised home pages, is that rather than substituting new materials for old ones, a blogger simply adds new posts, creating an ever-growing compilation of entries and an archive of previous posts. Compilations of postings serve as context for readers of blogs. Thus regular readers can get a sense of the identifying 'voice' or 'persona' behind the posts. Over time, a blog archive can read very much like an evolving portrait of the blogger's interests and experiences. Thus, by their very nature, blogs raise a number of privacy issues. On the one hand, they are persistent and cumulative. On the other hand, they are easy to produce and disseminate which results in large amounts of sometimes personal

*Email: k.mccullagh@salford.ac.uk

ISSN 1360-0834 print/ISSN 1469-8404 online

DOI: 10.1080/13600830801886984
http://www.informaworld.com

information being broadcast across the Internet. The survey explores the tension between the freedom of expression experienced by bloggers and the potentially problematic privacy consequences of public, persistent blog entries.

## Defining privacy

Privacy is an elusive concept.[2] Numerous different definitions of privacy have been offered.[3] For instance, Warren and Brandeis defined it as 'the right to be let alone'.[4] Westin defined it 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'.[5] Bloustein claims that privacy protects against conduct that is demeaning to individuality, an affront to personal dignity or an assault on human personality.[6] Likewise, Reiman asserts that privacy 'protects the individual's interest in becoming, being and remaining a person'.[7] Benn recognises the important element of choice in this conception. He states 'respect for someone as a person, a chooser, implie[s] respect for him as one engaged on a kind of self-creative enterprise, which could be disrupted, distorted or frustrated even by so limited an intrusion as watching'.[8] Van Hove argues that privacy means two things: (1) that 'a person has the right to a private sphere, and (2) that a person has 'the right to control the flow of information about his private life'.[9] Clarke offers a broader definition, stating that it is 'the interest that individuals have in sustaining a ''personal space'', free from interference by other people and organizations'.[10] Van Der Haag defines it as 'the exclusive access of a person to a realm of his own. The right to privacy entitles an individual to exclude others from (a) watching, (b) utilizing, (c) invading his private [personal] realm.'[11] Rachels sees privacy as being 'based on the idea that there is a close connection between our ability to control who has access to us and to information about us, and our ability to create and maintain different sorts of social relationships with different people'.[12] By emphasising the value of relationship-orientated privacy, Rachels' conception of privacy tries to define what aspects of life an individual should be able to control, keep secret or restrict access to. These definitions differ greatly in the fundamental way that they approach privacy, with some referring to physical aspects of privacy, others to personal information, and still others to issues of autonomy. Privacy encompasses a variety of different issues and is important for a number of reasons. Therefore, a single definition that adequately incorporates all the subtle differences that privacy evokes has so far proven impossible. However, a comprehensive and useful framework for the purposes of this study is offered by DeCew's cluster concept of privacy.

### *A broad conception of privacy: A cluster concept*

DeCew argues that privacy is 'a broad and multifaceted cluster concept',[13] which encapsulates 'our ability to control information about ourselves, our ability to govern access to ourselves, and our ability to make self-expressive autonomous decisions free from intrusion or control by others'.[14] Thus, DeCew envisages privacy as a 'complex of three related clusters of claims concerning information about oneself, physical access to oneself, and decision making and activity that provide one with the independence needed to carve out one's self-identity through self-expression and interpersonal relationships'.[15] The cluster includes three aspects of privacy: (1) informational privacy; (2) accessibility privacy; and (3) expressive privacy.

### *Informational privacy*

Informational privacy centres on the notion of control over one's information. Informational privacy considers the arguments that much information about oneself 'need not be available for public perusal'.[16] The importance of informational privacy lies in its ability to shield individuals from intrusions, as well as from threats of intrusions. It also affords individuals control in deciding who should have access to the information and for what purposes.[17]

### *Accessibility privacy*

The second aspect of the cluster concept of privacy concerns physical privacy 'focus[ing] not merely on information or knowledge but more centrally on observations and physical proximity'.[18] It protects against traditional privacy violations, such as a house being wiretapped, or a family consistently being watched via a neighbour's telescope, or a 'peeping Tom' creeping around a house. Such examples indicate the importance of accessibility privacy so people can dictate who has access to them, and to what degree. Implicit in the discussion of accessibility privacy is also the degree to which someone is aware of the accessibility violation, on the basis that being wiretapped without one's knowledge intuitively feels like a privacy violation. Accessibility privacy, while perhaps the most narrow understanding in the cluster, is nonetheless an important aspect.

### *Expressive privacy*

Expressive privacy concerns an individual's ability to freely choose, act, self-express and socially interact. It is closely linked with intimacy, emotional vulnerability, autonomy, and social roles. It is integral to protecting individual autonomy.[19] For instance, if an individual is aware that they are under constant observation and have no privacy in a domain where they would normally have 'wide discretion concerning how to behave',[20] they will presumably structure their actions not just according to their own will or intention, but will also try to keep them in line with what they envisions their observers would like to, or expect to, see. In this way, issues of expressive privacy and autonomy are also inherently intertwined with the social pressure that results from social judgments and norms. Thus, in a transparent society where all are visible to everyone, we would be completely subject to public scrutiny and would likely conform to societal norms for fear of being ostracized. This would have serious repercussions for autonomy. First, individuals would no longer be able to play with, and test, social norms backstage, which is a crucial act to forming self-identities.[21]

Secondly, under such constant social scrutiny, individuals would be implicitly forced to conform to societal norms. Thus, society could quickly become an undifferentiated mass where everyone says and does the same things in order to be deemed socially acceptable. In this situation, there would no longer be any room for individual thoughts, feelings, or emotions – our self-expression would be seriously limited. And even if such individual thoughts could continue to occur undetected, the actions that would normally correspond would likely cease to exist due to fear of social judgment.

Thirdly, expressive privacy also plays a crucial role in developing social roles and relationships[22] as it works to protect and maintain intimacy. This would inhibit intimate relationships because intimacy is premised on the fact that individuals know particularly

personal, or otherwise unknown, information about each other. Without privacy, an individual's relationship with their mother would be no different from their relationship with their employer, as both could know exactly the same amount and degree of information about them. Thus, without expressive privacy, social relations could not be as varied and social interaction would be seriously diminished.

Finally, expressive privacy also enables work on self-identity to proceed as it regulates and allows social interaction to occur. Since the self can only be developed via social interaction with others, expressive privacy and the reflexive formation of self-identity are closely connected.[23]

### *Self-reflexive identity and privacy*

According to Giddens, self-identity in late modernity is highly reflexive so that sustaining a coherent yet continuously revised biographical narrative is key. First, the self is a reflexive project for which the individual is responsible. In this way, individuals are what they make of themselves. Self-identity is routinely created and sustained in everyday activities, via the routines of practical consciousness, so that it is continuously revised. Self-identity is the self as reflexively understood by the person in terms of their biography. A person with a stable self-identity has a feeling of biographical continuity so that they can grasp and communicate it. In this respect, a person's identity is really about the capacity to keep a particular narrative going. Moreover, a person's biography can not be wholly fictive and must continuously integrate events into the ongoing story of the self. In essence, 'In order to have a sense of who we are, we have to have a notion of how we have become and of where we are going.'[24]

This reflexivity of the self is thus continuous and all-pervasive and the narrative of the self is made explicit in an attempt to sustain an integrated sense of self. As Giddens makes clear, 'in the reflexive project of the self, the narrative of self-identity is inherently fragile'.[25] Making a coherent and continuous narrative amidst a constantly changing life experience is a continuous burden for the individual in modernity. One's self-identity 'must be continually reordered against the backdrop of shifting experiences of day to day life and the fragmenting tendencies of modern institutions'.[26] Overall, Giddens concludes that the difficulty in sustaining a coherent narrative is because of modernity's dynamism and reflexivity. In this way, the relationship between self and society can be understood as follows, 'The self establishes a trajectory which can only become coherent through the reflexive use of the broader social environment. The impetus towards control, geared to reflexivity, thrusts the self into the outer world in ways which have no clear parallel in previous times.'[27] In this way, the abstract systems of high modernity allow the self more mastery over the social relations and contexts incorporated into self-identity.

Giddens' insights into the role of self-identity and society in late modernity are particularly useful for the purposes of this study because they may aid thinking and understanding of the ways that bloggers negotiate the boundary between public and private and, hence, the society and the self.

## The study

### *Methodology*

The findings presented here are from an online survey of bloggers from around the world. Participants answered questions about their blogging practices and their expectations

of privacy and accountability when publishing online. The questionnaire consisted of multiple-choice Likert-scale questions, and open-ended essay questions to allow for further qualification of answers. Out of the total number (1,314) of responses received, 1,258 were selected for data analysis; the remainder of the responses were incomplete and were disregarded. The respondents were not randomly selected, but were found through a variant of the snowball-sampling strategy. Announcements for the online survey were posted to mailing lists in three universities in the UK as well as on a few high-traffic blogs. The viral nature of blogs meant that the links to the survey page quickly spread to many other blogs and to YouTube. However, the resulting population of participants does not qualify as a random sample and, accordingly, the results from this survey cannot be generalised to the entire blogging population. Rather, the findings are representative of certain niches of the English-speaking blogging world.

### *Study population*

The majority of respondents in this study (49.1%) were female. Even though some of the popular blogging sites attract mostly teenagers, the respondents tended to be older, with over half of them between 19 and 34 years of age (54.9%). Over one-third of participants were from the UK (39.5%) which is not surprising, given that the survey questionnaire was available only in English and that announcements for the survey were posted to email lists in three UK universities.[28] These demographic characteristics contrast with findings from other blog surveys in which participants in these spaces tended to be 'young adult males residing in the United States'.[29] Almost equal percentages of respondents were single (39.7%) or living with someone (37.1%),[30] and the majority (60.7%) were working though only a minority claimed to be the main earner (31.5%).

### *Limitations of study*

The respondents were not randomly selected, but were found through a variant of the snowball-sampling strategy. Announcements for the online survey were posted to mailing lists in three universities in the UK as well as on a few high-traffic blogs. The viral nature of blogs meant that the links to the survey page quickly spread among many other blogs and YouTube. However, the respondents were not randomly selected and, accordingly, the results from this survey cannot be generalised to the entire blogging population. Rather, the findings are representative of certain niches of the English-speaking blogging world. Indeed, the demographic profile of respondents here differs from what is known from studies and popular blogging sites in relation to age. On popular blogging sites, such as LiveJournal, where usage statistics are available, teenagers account for the majority of the blogger population, whereas most of the respondents in this survey were between 19 and 34 years old. The results from this survey might thus have been different had the pool of respondents been randomly sampled.

In the survey, respondents were asked to self-report on their blogging practices, and their privacy attitudes and expectations. Unlike other studies where researchers accessed participants' blogs and conducted content analysis of posts,[31] here it is the bloggers' subjective sense of privacy and liability that is revealed. This self-disclosure approach has two important implications: (1) there can be disparities between stated privacy attitudes and actions; and (2) participants' perceptions of their blogs might differ from those of outside observers and researchers. It is well documented that people's perceptions of their own behaviour can differ from how they actually behave.[32] In addition, because of the

self-reporting nature of this study, accuracy is difficult to verify. For example, no external validation was conducted on the sites of participants who described their blogs as being mostly 'My life (personal diary/journal)'. Therefore, comparisons between the present and previous findings should be made with these caveats in mind.

## *The findings*

### *1. Bloggers value self expression and social interaction*

In a previous study Herring et al.[33] coded a random sample of blogs based on the nature of the content posted. They found that the majority of blogs (70.4%) were of the personal journal type: 'in which authors report on their lives and inner thoughts and feelings'.[34] In a subsequent study Herring et al.,[35] looked at gender and age-based differences in the content of blog sites. They found that women and teenagers tended to write personal journal-style blogs.

Table 1. Main blog topic.

| Main topic of blog | Percentage |
|---|---|
| My life (personal diary/journal) | 58.4% |
| Politics and government | 4.7% |
| Entertainment (movies, music, MP3s) | 5.7% |
| Sport | 0.6% |
| News and current events | 4.0% |
| Business | 0.5% |
| Technology (computers, Internet, programming) | 8.3% |
| Religion/spirituality/faith | 0.7% |
| A particular hobby | 1.7% |
| Health (general health, a specific illness) | 1.2% |
| Gossip | 0.4% |
| Other | 11.9% |
| Prefer not to answer/No answer | 1.8% |
| Total | 100% |

(Blog Survey 2006, n = 1258).

In agreement with previous studies, Table 1 indicates that most respondents (58.4.1%) said their entries could be characterised as 'My life (personal diary/journal)'. Thus, bloggers value self-expression and use blogging as a medium for self-reflection. Furthermore, because the self is only developed through interactions with others and because the reflexive project of the self is a characteristic of late modernity, the opportunity to continuously work on the project of the self via the interaction on blogs and comments to posts was the main reason why the majority of bloggers engaged in blogging.

Table 2 illustrates that when asked to select reasons for blogging, the highest percentage (62.6%) of respondents indicated that their main reason for blogging was to document their personal experiences and share them with others, whereas the lowest percentage (1.6%) indicated that their main reason for blogging was to make money. Indeed, 88.3% indicated that making money was not a reason for them to blog. This confirms that bloggers value the self-reflection and social interaction features of blogging.

Table 2. Reasons for Blogging.

| Reason for blogging | Main reason | Minor reason | Not a reason | Prefer not to answer | Total |
|---|---|---|---|---|---|
| To document your personal experiences and share them with others | **62.6%** | 27.2% | 8.1% | 2.2% | 100% |
| To express yourself creatively | 50.9% | 35.9% | 10.8% | 2.4% | 100% |
| To influence the way other people think | 12.0% | 31.2% | 53.0% | 3.8% | 100% |
| To motivate other people to action | 10.8% | 30.0% | 54.8% | 4.4% | 100% |
| To share practical knowledge or skills with others | 16.5% | 36.8% | 42.4% | 4.3% | 100% |
| To network or to meet new people | 18.1% | 41.1% | 37.0% | 3.8% | 100% |
| To entertain people | 31.9% | 42.3% | 21.6% | 4.2% | 100% |
| To discuss problems with others | 18.9% | 39.3% | 37.4% | 4.4% | 100% |
| To stay in touch with friends and family | 31.0% | 24.2% | 41.0% | 3.8% | 100% |
| To make money | 1.6% | 4.7% | 88.3% | 5.4% | 100% |
| To store resources or information | 14.5% | 35.2% | 46.3% | 4.0% | 100% |

(Blog Survey 2006, n = 1258).

*Knowledge of audience*. Audience knowledge potentially raises serious implications for privacy as bloggers who do not have good knowledge of their audience may decide to refrain from publishing details about their private life. Table 3 indicates widespread variation in levels of audience knowledge.

Table 3. Knowledge of audience.

| How well do you feel you know your blog's audience? | Percentage |
|---|---|
| Extremely well | 8.1% |
| Very well | 23.1% |
| Quite well | 32.2% |
| A little | 18.4% |
| Not at all | 7.6% |
| Prefer not to answer | 2.1% |
| It is more complicated | 8.5% |
| Total | 100% |

(Blog Survey 2006, n = 1258).

A range of reasons for the widespread variation in levels of audience knowledge was expressed:

- new – audience unknown
- in real life only
- online only
- different audiences for different blogs, and
- mix of real-life and online audience.

These responses concur with the findings of Nardi et al. who characterised blogs as a 'studied minuet between blogger and audience'[36] and distinguished two kinds of audiences: the bloggers' own, known social networks, and a larger audience beyond the author's friends and family. Of those bloggers who formed online friendships, trust and sharing of information appears to be an issue:

> It really does vary. I have made good friends with a handful of people through blogging who I have gone on to meet. In fact, I had a year long relationship with someone who 'met' me by initially reading my blog. And there are other people who I have had a degree of contact with for 18 months or so who I may not have met, emailed or spoken to, but over such a time it is hard not to form some sort of bond – real or imagined – with such people. However, there is another section of my audience who I don't know much about. Some people read regularly, and from reading their comments and blogs I decide I don't want to get to know them any further and pay them scant attention, and yet they continue to return, getting to know me better by the day whilst I remain purposely oblivious to them. Finally, there is the section of readers who never interact, and yet return on a frequent basis.
>
> The audience changes frequently. Some remain faithful readers and some drift away from you. Some you 'know' better than others.
>
> You don't really 'KNOW' your audience; it could be anyone, preacher, teacher, convict, sexual predator, or anyone in between. You never truly 'know' who is watching or what their motives are.

Some bloggers have different identities for different blogs, on which they disclose different types of information to different audiences:

> My first name is publicly available on my blogs, as I don't believe in pseudonymous blogging, I believe it leads to bullying and arrogance. However, hypocritically, as I used to work in local government, I never include my last name in any of my blogs, and I guard the links between all four blogs jealously so that anyone who identifies me as the author of one would find it difficult to identify me as the author of another ... as one of my blogs deals with work, one with factual information about my life, one with family, and another with the more traditional personalised style of blogging ...
>
> I have a major blog, with a large readership. It's pretty personal, but because it is associated with my real name and if you Google me you find it, I limit what I write about dating, sex, and money. I have a 'secret' blog, with a small (but growing) readership, where I write about dating. That is done with a pseudonym. I've revealed the existence of my other blog to three readers who read both blogs (and who I don't know in person) but I don't know of any other crossover.

Nardi et al. also found that, even though bloggers delighted in their audiences, there was a clear desire to keep the audience at arm's length: 'interactivity was valued, but only in controlled small doses.'[37] Likewise, Gumbrecht characterised blogs as 'protected space'.[38] He asserts that because readers' comments are subservient to the blogger's posts, blogs create a protected arena in which authors feel safe to express emotions and experiences. When asked whether they do anything to limit who gets to read what they post, 72.3% of respondents said no. However, in the open-ended responses it became clear that access control is of major importance to a minority of bloggers. These respondents reported a using variety of differing privacy settings to limit audience access. For instance, they would only reveal identifying information if they were satisfied that their blogs were only accessible by a trusted audience. Of those who revealed detailed identifying information, many stated that they were happy to do so because they controlled audience access, e.g. limited to friends only – and as they know such people in real life they are happy to continue sharing such information on the internet:

> In 'Friends only' entries I sometimes reveal my full name.
>
> My blog is password-protected, so even though I identify myself by name on the blog, only about a dozen people even have the URL and a password to see it.
>
> Mine is a locked Livejournal blog, so I trust the people on there with my real name. But from the outside, no, you can't identify me.
>
> I have mentioned my first name, and the names of most people around me, so technically I could be identified ... but really only by someone who already knows me. Does that make

sense? I mean you couldn't look me up in the phone book from the information on my blog. But if you were my Mother, you'd work out it was me pretty quick.

Thus, the importance of expressive privacy to bloggers is evident because 62.6% of respondents claimed that social interaction through sharing of personal experiences was their main reason blogging, despite the privacy risks. One of the most important functions of expressive privacy is its ability to allow meaningful relationships to develop. Without expressive privacy regulating social interaction and preventing the overreaching of others, such relationships would be impossible and a major reason for the popularity of blogging would cease. Nevertheless, when faced with different people who hold different values, beliefs, lifestyles etc., bloggers were forced to continuously factor in the shifting circumstances caused by different levels of knowledge and trust in their online relationships and adjust their narrative of the self accordingly.

### 2. *Bloggers value privacy*

It was clear that bloggers value privacy. Respondents were asked to rate a list of issues that could be considered of social importance on a scale of 1 to 5, where 1 is not at all concerned and 5 is very concerned.

Table 4. Social importance of issues.

| Socially important? | 1 Not at all concerned | 2 | 3 | 4 | 5 Very Concerned | No answer | Total |
|---|---|---|---|---|---|---|---|
| Preventing crime | 5.2% | 7.9% | 22.7% | 27.8% | 22.5% | 13.9% | 100% |
| Improving standards in education | 3.3% | 3.5% | 11.9% | 26.3% | 41.3% | 13.7% | 100% |
| Protecting people's personal information | 2.8% | 6.1% | 17.4% | 27.6% | 33.3% | 12.8% | 100% |
| Protecting freedom of speech | 3.0% | 2.3% | 8.6% | 21.5% | 51.9% | 12.7% | 100% |
| Equal rights for everyone | 3.1% | 2.6% | 7.2% | 21.0% | 53.2% | 12.9% | 100% |
| Unemployment | 4.7% | 13.0% | 29.6% | 25.3% | 13.4% | 14.0% | 100% |
| Environmental issues | 4.0% | 7.7% | 18.7% | 27.4% | 28.4% | 13.8% | 100% |
| Access to information held by public authorities | 4.1% | 11.2% | 23.2% | 22.0% | 25.4% | 14.1% | 100% |
| Providing health care | 4.5% | 6.7% | 16.0% | 27.1% | 31.6% | 14.1% | 100% |
| National security | 9.4% | 14.9% | 25.4% | 19.8% | 15.4% | 14.9% | 100% |

(Blog Survey 2006, n = 1258).

Of the issues listed, Table 4 indicates, the highest percentage (over half) of respondents were very concerned about equal rights for everyone, whilst over one-third of respondents were very concerned with the protection of personal information. A higher percentage of respondents were very concerned with the protection of personal information than with the issues relating to preventing crime, unemployment, environmental issues, access to information, or national security. Only 2.8% were not at all concerned with protecting personal information.

Almost one-quarter (24.8%) of respondents said that they had posted personal information on their blog 'all the time'. Only 2% of respondents said they had 'never' posted anything highly personal on their blogs. However, bloggers seem to reflect regularly on the content of posts when deciding whether or not to post personal information online. Most respondents (65.6%) said they had considered some information 'too personal' to write

about on their blogs. When asked to elaborate on the types of information that would be too private, the comments of one blogger encapsulate the sentiments expressed by others:

> Basically anything negative in your personal life. You may not mind sharing it at the time, or may want to talk to people about it, but you probably don't really want it indexed forever or available to copy and paste or otherwise spread through gossip.

Several categories of information were considered too private to post in blogs, namely:

*a) Personal information that could identify the blogger.*

> Anything that can identify me, my personal life, as opposed to my online life.
>
> My last name. Anything too specific about where I live. Anything I would mind my mother and my boss reading.
>
> As a Jordanian belonging to a very small community (we're a small country), I find that I need to maintain a certain level of disclosure because everyone knows everyone.

- Information regarding others.

> Anything that could adversely affect people I love, e.g. talking about a friend's bereavement, discussing my or someone else's sex life.
>
> Things concerning the people in real life, like if I have feelings about them they don't know about. Or personal events that happen in OTHER people's lives that I know off, cause it's not my information to tell.
>
> Personal information told to me by others who do not know I have a weblog.
>
> My barometer is whether the information involves someone else and could embarrass/upset them. I'm fairly open about myself but wouldn't dare force someone else to share my standards of openness.
>
> I won't post about anything personal that is not about me exclusively.
>
> I refrain from talking about what happens in school or making opinions about my lecturers or other students in school because I am practising for future employment. My blog remains personal, but I only make posts about things that I have done outside.

*b) Emotions.*

> A break up between myself and my partner, or family arguments. They are personal information and it is very rude to go broadcasting it to all your friends and potential strangers.
>
> Stuff that I've tried to talk about but couldn't explain coherently without sounding like a desperate/emo/completely crazy person.
>
> Many of my innermost feelings and experiences.
>
> I'd find it difficult to write about my late wife's death in any detail, because it's a painful topic (though I have alluded to her long illness and death a few times).
>
> I tend not to write about my sex life, or blather on about my angst. I don't want to be looking back at what I write and cringing because it's so 'teenage diary'. I tend to use it to try and look at the bright side of life, so I can laugh at it.
>
> Anything considering other people where I might have to call names. How I REALLY feel about some things or someone. What is REALLY affecting me or is my intention behind certain actions. Sexuality, religious beliefs, political affiliations etc.
>
> Information related to the divorce that I am currently negotiating with my husband. I'd love to be able to vent about it, but it doesn't seem prudent.

*c) Sex/relationships.*

> My sex life is personal and now that my daughters are getting older (17 at present), I write less and less about them. I no longer write about work either, although I used to 2 years ago.
>
> I don't give specific details and names. I don't talk about sex. I am describing events and my feelings about them. There is a difference between personal and intimate.
>
> Details of sexual life; not because I wouldn't write it up online, but because I prefer to separate it from my other activity; more as a consideration to others/my parents than anything else.
>
> Stuff about my wife: arguments, sex life, her weight, her behaviour towards me.
>
> Mostly stuff about sex. For example, if I've slept with someone who is, along with their family or friends, on my friends list.
>
> Things like sexual activity (when it actually happens), I won't discuss because a lot of times, it involves someone my other readers know. Also, masturbation. I'm not going into the details about that no matter how hard they beg.
>
> This has shifted over time – I used to write explicitly about my sex life and my work as an escort, these days I keep my sex life private, though I still write about sexuality in a more general sense.
>
> My sex life. Because most of my audience knows my girlfriend. Otherwise there'd be no problem.

Decisions to disclose information about sex life or relationship were influenced by how the other people involved would feel and whether the other person was known to the blog audience. In many instances the bloggers were aware of the need to protect the other party's personal information.

*d) Arguments.*

> I once wrote half a post about rowing with my husband, then thought better of it.
>
> Relationship related stuff, arguments with other people who are not 'online' (hence usually don't know I blog) would be unfair.
>
> Fights with my husband, our sex life, issues with my extended family that really aren't anyone else's business.
>
> Anything negative relating to anyone but myself also does not get blogged about; my blog is not my personal grievance platform.

*e) Financial information.*

> Financial and health issues.
>
> Money, family.
>
> Financial problems, personal relationships and sometimes work situations.

*f) Work.*

> At least one of my employees was made aware of my blog before I became his supervisor. Since becoming his supervisor, I've been leery of writing about anything that might undermine my credibility in his eyes.
>
> I know some of my friends read my blog, and some people from where I used to work. I have posted things that I now consider 'too personal' and it got me into some trouble at work. These would be: sexual practices and with whom, blatant one-sided opinions.
>
> Mostly if it is work-related, as I work with some pretty sensitive information.

> Several incidents that might be recognisable to someone reading my blog, most of them concerning patients in hospitals. I tend not to write about my partner; he doesn't even know I blog, so I respect his privacy by not discussing him in detail.

*g) Health information.*

> Health/personal thoughts/mental issues.
>
> My abortion, my eating disorder. I try to keep it light.
>
> Both times I miscarried I had to wait a few weeks before posting about it, mainly because I needed to be distanced from it slightly before I could start to write about it. Otherwise I don't really limit what I write. I don't write about my sex life.

*h) Other issues e.g. illegal activities, political beliefs, religious views.*

> ... the fact that I am an atheist and I don't want family members who read my blog to know, the fact that I am planning on voting Liberty.
>
> Because I blog about and in dialogue with Muslims I prefer to keep the fact that I am gay to myself because it can get in the way of an otherwise positive dialogue with some people.
>
> Last time I've been high on weed, if I enjoy anal sex, sexual acts I performed with my [partner]... Some of my experiences growing up.

These comments indicate a desire to protect informational and expressive privacy. Bloggers are aware of a risk posed by external parties who might be interested in collecting or collating the information, they post; thus they seek to restrict their blog readership and content. Also, the comments reveal that bloggers were likely not to blog about controversial social, moral or philosophical issues which would draw negative responses or criticism from readers or members of wider society. This suggests that bloggers consciously and intentionally negotiate the boundary between public and private. They take responsibility to ensure that their posts are in line with their desires as to how public or private they want to be at that specific time – a process that may shift from day-to-day, and from topic-to-topic. This finding concurs with an assertion by Palen and Dourish[39] that privacy in networked environments is a dynamic, dialectic process of negotiation that is conditioned by people's own expectations and experiences and by those of others with whom they interact. This author concurs with Rosen,[40] and Grudin,[41] and Palen and Dourish[42] in noting that these negotiation processes are fundamentally dependent on people having control over their information and over the contexts in which that information is presented. Thus, bloggers strive to negotiate a boundary between self and society that they feel comfortable with, yet at the same time they are able to interact socially with their readers. In this way they are able to define and maintain the desired level of publicness or privateness that they wish to achieve through the level of personal exposure that they allow. By controlling their information disclosure, bloggers are able to decide where to draw the boundary between themselves and others. Thus, managing their participation via that shifting public/private continuum is an important part of bloggers' experience.

### *3. Bloggers are aware of privacy risks*

More than one in ten bloggers had experienced privacy invasion through the activities of other bloggers. When asked to explain the ways in which their privacy had been invaded, respondents described the following situations:

*Some bloggers had their identity exposed by others or are aware of a risk of exposure due to web interconnectivity.*

> Some of my friends call me by my first name when they comment.
>
> A friend linked to me and when she talked about me in her blog, she used my real name.
>
> Personal health information was babbled to the world by an ex.
>
> A white supremacist named me and gave sufficient details about my home address.
>
> Putting my picture on their blog without asking me, though I have done the same.
>
> A friend's partner's family were very upset when they discovered her religious beliefs and made accusations against her and paid a private detective to investigate her.

*Changes in behaviour over time*. There is evidence of a growing concern about protecting anonymity among some respondents

(1) A common reason for limiting details was to prevent it from leading 'google' searches by employers to their personal blogs.

> I use my first name, but always leave out my surname. I also try not to mention by name where I work or where I grew up. This isn't so much because I don't want my audience knowing these details, but rather that I am aware that including such details makes it much more likely an employer, former acquaintance or anyone I wouldn't want reading might accidentally 'google' their way onto my site. Despite these safeguards, some friends have still managed to google their way to my blog, so I think my concerns are well founded. If I were to start blogging afresh, I would give serious consideration to adopting a pseudonym.
>
> I don't have my full name on the blog about page, but I have mentioned it many times. I want my day job work to be my primary Google search result for my name.
>
> Like to keep work and home life separate (I'm a social worker) so using my real name is not a good idea in case a client did an internet search.

Restricting personally identifiable information is another mechanism bloggers employed to lessen the likelihood of privacy risks, here implicating more of the traditional understanding of privacy, one where control of personal information can protect the individual.

(2) Some bloggers initially preserved their anonymity, but are aware that the reasons for their initial behaviour are changing.

> I use a fake name, which I originally assumed to keep my blogging completely separate from my work life. I've since left that job, and now am a lot more forthcoming with personal details, but I've kept the name, partly because I've become fond of it, and partly because other bloggers now know me as Ben.
>
> I do not openly list my name on my blog, however I do reference my family members by first name, have listed my last name on occasion and list the city in which I live. I suppose that I had at one time planned to remain an unknown, however have not found that to be as important as time has gone on.
>
> I started the blog anonymously and posted under a pseudonym. Then I launched a web site with my name in the URL. The blog is now hosted on the website, but I still post under the pseudonym. ... Obviously we're one and the same, but I still find the alter ego a useful psychological and literary device.

(3) some bloggers are moving towards anonymity.

> First my name was public, now I changed that and my name is hidden.
>
> I did at first, but then after a while realized that I didn't want to have my name be so easy to google ... I don't really have an issue if people know who I am and where I live, but since the purpose of my blog is to keep my family updated on the lives of my kids, it doesn't seem necessary. The blogs that I keep for my children's birth families do NOT have names.
>
> It was on all my blogs until a few months go. My main blog has a pseudonym while I apply for jobs, and will revert back to my real name afterwards. On my secondary blogs, I use my real name, but these are primarily professional/hobby-related. My name is unique, which makes me quite careful.

Some bloggers noted that they rely on anonymity in their blogging activities, particularly when posting information they considered to be more personal or private by traditional standards – the anonymity of blogging made them more likely to post such information.

These findings suggest that privacy norms are emerging among bloggers. For instance, as the comments above illustrate, some bloggers are beginning to create informal guidelines for publishing the names of people and employers in their blog entries. It is suggested that there is evidence of bloggers altering their behaviour according to employment prospects as the comments indicate that some bloggers are wary of revealing personal information to prospective future employers, and of revealing such information to potential clients.

### 4. *Blogs are perceived to be public spaces*

The degree of accessibility is a major part of what makes a blog public or private. In this regard, the more accessible or visible a blog is, the more it is considered to be public. Some bloggers opined that anonymity or privacy is not possible on the Internet.

> Anyone publishing anonymously in any medium must accept the risk of being 'outed'. Though I deplore the gratuitous and often destructive identification of anonymous bloggers, it would be foolish for anyone to assume anonymity is a right.
>
> Blogs are a public thing. Some might think that they are private like emails, but should realise that both emails and Blogs are public in the sense that they can be found by someone who wants to find them. It's like paparazzi taking a photo of a famous person topless on a public beach ...
>
> The web is a public place, anything you write is not private. Being identified is a known risk of any web activity, either reading or writing. Same as sitting in a library reading or writing.
>
> Many bloggers know that there is no real privacy and that anonymity is just a temporary matter. If someone wants to find out the person behind the blog, it would be quite easy to do so.
>
> It's very hard to have a totally anonymous blog. People who know you may reveal who you are. There is also a chance that you can make a slip that reveals who you are.
>
> When you made the conscious decision to put a blog on the internet, you know that it is not at all private. Blogs are not protected the same way that bank accounts are. Anyone can find them and they are easily hacked into. Just because you blog anonymously does not mean that your privacy is protected.
>
> I feel the nature of blogs is public ... by writing and publishing online one is exposing one's text to public scrutiny (unless a privacy lock has been placed on the blog). I feel that people who publish in blog format are basically asking to be read, hoping to be noticed.

The degree of accessibility is a major part of what makes a blog a public or private space – the more accessible or visible a space is the more it is considered to be public. Also, level of familiarity/audience knowledge is a key determinant of whether a blog is considered a public space. Hence, the fact that a blog could be read by strangers could mean that the space is considered as public by many bloggers.

*5. Bloggers employ mechanisms to protect privacy*

Blogging software allows differing levels of privacy. The most private blog is password-protected. The most public blog is listed by the user's blog service and will be easily found by search engines. An unlisted blog is less likely to be found, but is not fully private; it is unlisted by the blogging service's directory (similar to an unlisted phone number). Such an unlisted blog cannot be found without knowing the URL, although there is a way such blogs can become public, namely, if the blog contains a link to a webpage that a viewer could click on, then the new webpage will receive the URL as the 'referrer', and it is possible for the 'unlisted' blog to be picked up by search engines. Since most blogs contain links that anyone might click on, unlisted blogs are not secure, although they may remain relatively invisible if they link to sites that few people access and if the links are not activated often. The survey respondents reported a variety of differing privacy settings and approaches, for instance, over a quarter of bloggers took action to limit who could read their blogs. These limitations included:

*1) Leaving out key details.*

> I made the title and address completely unconnected to me and don't use my surname so a google search of my name wouldn't flash it up.
>
> I do not document everything that has happened to me on my site. While my blog is predominantly a personal one (i.e. 'What I did today', 'What I learned today', 'Who I spoke to today', etc.) I prefer having a sense of anonymity.

*2) Using passwords.*

> To avoid Spammers I have put a computer word verification in place. I only did this after receiving some questionable responses to my posts.
>
> Some entries are completely private and require personal login. The rest are completely public.
>
> Password for blog is only provided for friends.

*3) Keeping the fact that they blog secret.*

> Most people who know me personally will never know about my personal blog. Only folks I totally trust and already share all my stories and inner issues with were invited.
>
> I use a pseudonym so that close family members about whom I may write (using an initial, not their full name) cannot easily come across my blog if they were to search one day. I do not want to have to explain myself if what I write is 'injurious' ...
>
> Other than my partner, my family does not know that my blog exists. Two real-life friends know about it. All other readers do not know me personally.

*4) Editing robot.text file.*

> Edit the robots.txt file which controls whether search engines are allowed to crawl your site.
>
> I use blogger and I blocked the search engine option. So, only if you click on a link from someone else's blog can you stumble across mine.

*5) Blocking IP addresses.*

> I ban IP addresses of people who come to my site just to insult me. I also use private categories to tuck away posts that I'd rather not have the general public read.
>
> I block the IP address of my sister-in-law as well as the IP of my company, but I turn that on and off, as sometimes I update at work.
>
> I block a lot of spambots, and the occasional troll – but the main thing I do is block any IP addresses and domains that I know my parents use, to stop them accidentally coming across it.

*6) Using privacy filters.*

> I tend to make more personal posts friends only; sometimes even limiting posts to people I don't know in real life as I sometimes prefer to limit my depressed/suicidal musings to people who can't do much about them.
>
> Most of the posts are public; these include ones about what I do, my fandoms etc. I have several custom friends groups to discuss things I think some people on the list would disapprove of (for example my religious and spiritual beliefs).
>
> Generally I make my entries open for all to see and respond to. However, if I'm discussing something either particularly personal or something that involves people who might be reading the blog, I make my entries friends-only.
>
> Some of the content of my blog is lightly filtered to avoid spoiling surprises for people or to talk about work or to preserve other people's privacy or to send a message or invitation to a certain section of people.

The bloggers who responded to the survey indicated an awareness of a range of privacy risks which fall under De Cew's cluster dimensions of accessibility privacy. The degree of accessibility is a major part of what makes a blog a public or private space – the more accessible or visible a space is, the more it is considered to be public. Inherent in this concept of accessibility of public spaces is the idea of restricted access. Respondents used a variety of mechanisms to restrict public access to their blogs, such as locks, password access, friends only. Such behaviour indicates that bloggers negotiate a boundary between self and society that they feel comfortable with, yet at the same time they are able to interact socially with their readers. In this way they are able to define and maintain the desired level of publicness or privateness that they wish to achieve through the level of personal exposure that they allow.

## Discussion

Bloggers face unique privacy concerns because, on blogs, meaningful and private information is often shared. Thus, the type of 'information' that could be collected, because it is more closely tied to one's self-identity and self-expression, poses a serious privacy risk to bloggers. While the storage of information is a concern for all Internet users, the archiving of blog posts poses threat. For instance, I may have posted on my blog ten years ago my teenage opinions on animal testing. As humans, we tend to change and grow over time. My opinions on animal testing may be different now. However, if I now apply for a job at a pharmaceutical company, the hiring committee could unearth my opinions through a simple Google search. In this case, the storage of my opinions and personal experiences for such a long time poses a serious threat to my privacy and presents consequences that I presumably did not anticipate at the time of disclosure. The storage of information that was initially public presents a unique privacy risk for bloggers because they often assume that the posts are private. It is the presumption that the blogs are

private that ensures that the storage of information can be a serious privacy risk for participants as private information is shared more easily.

Additionally, the collation of all of my posts could easily paint a picture of me as a person. If this picture was then used in a different context, it could pose a serious privacy risk. For example, let's say that my teenage years were emotionally difficult for me and that I frequently blogged about my low moods when I was experiencing mild depression and generally was not happy with myself. Now, however, I am a well-adjusted and happy family lawyer who is up for partnership. What would happen if my appointment committee dredged up my blog posts from ten years ago and made the decision that my emotional well-being is questionable at best because I was quite emotionally unstable ten years ago. There is an increased chance, they argue, that I will become so again and hence am not a 'desirable' candidate. The fact that I disclosed this information could now have potentially serious effects on my career and life plans. Arguably, blog sites serve as the context for the entries they contain. There is no guarantee, however, that individual entries will not be extracted from their original context and exposed in radically different forums in the future. Grudin[43] refers to this 'loss of control' as the steady erosion of *clearly situated action*. 'We are losing control and knowledge of the consequences of our actions, because if what we do is represented digitally, it can appear anywhere, and at any time in the future. We no longer control access to anything we disclose.'[44] Future employers, insurance companies, police investigators or even a future spouse could locate decontextualised, and possibly damaging, statements. Rosen highlights the contextual basis of privacy violations when stating that disclosure of personal information is a highly circumstance-sensitive matter.[45] When taken out of context, the same information can be severely misjudged by others.

Therefore, the disclosure of personal information by bloggers appears to pose very unique privacy threats as expressive privacy plays a fundamental role in our lives. It enables us to choose and dictate the way that we will live, it promotes the creation of our self-identity, and it allows us to enjoy a wide variety of social relationships and roles, including intimate relationships. Expressive privacy sets the stage for social interaction to occur and additionally enables the creation of one's identity by preventing other people's social overreaching throughout this interaction. Furthermore, the degree of accessibility to others and the amount of information one wants others to have are all connected to privacy. It is suggested that acknowledgement of the social dimension of privacy is crucial to understanding how bloggers perceive and negotiate privacy. Expressive privacy protects people from the overreaching social control of others that would inhibit self-expression and freedom of association. In this way, the disclosure of personal information by bloggers appears to pose very unique privacy threats. While blog posts may appear to be posted in public arena, it does not negate the fact that the information they share is often intimately tied to a person. As a result, violations would likely have serious repercussions for bloggers' self-expression and thus their ability to socially interact and develop meaningful relationships.

## Conclusion

Most respondents in this study described their blogs as the personal diary/journal type which indicates that blogging provides a unique opportunity for expressive privacy and furthermore allows bloggers to work out their reflexive project of the self in new ways, despite the inherent privacy risks posed by this medium. Whilst, blog posts may appear public, it does not negate the fact that the information they share is often intimately tied to

a person. Blogging poses new opportunities for privacy violations to occur, as individuals discuss personal matters and provide opinions openly in a format that can be archived indefinitely and easily accessed by anyone with an Internet connection. As a result, violations would likely have serious repercussions for bloggers' self-expression and thus their ability to socially interact and develop meaningful relationships. Participants in this study described tactics for keeping certain information private even when it is publicly published. Despite the emerging privacy strategies described in this study, bloggers reported having difficulty negotiating privacy boundaries under certain circumstances. The workplace is one setting where such problematic situations regularly occur. Bloggers' privacy boundaries in the workplace have yet not been clearly established, either socially or legally. Accordingly, one recommendation that emerges from the findings of this study is that organisations should provide blogging guidelines for employees. A few companies have posted written policies concerning personal blogs on their websites, including clear, point-by-point suggestions that address issues that are sensitive to the company, but that may not occur to employee bloggers when they choose to discuss matters related to the company's technology or business. Such policies could serve as the first step in a broader process of negotiation between employers and employees as blogging practices continue to evolve.

## Notes

1. Glenn, D (2003) Scholars who blog, *The Chronicle of Higher Education*, *49*(39), A14, in F B Viegas (2005) Bloggers expectations of privacy and accountability: an initial survey, *Journal of Computer-Mediated Communication*, *10*(3), article 12. Available at <http://jcmc.indiana.edu/vol10/issue3/viegas.html>. Accessed 27 February 2008.
2. Epstein, R A (2000) Deconstructing privacy: and putting it back together again, *Social Philosophy & Policy*, *17*(2), 1; Frey, R G (2000) Privacy, control, and talk of rights, *Social Philosophy & Policy*, *17*(2), 45; Rosenberg, A (2000) Privacy as a matter of taste and right, *Social Philosophy & Policy*, *17*(2), 68; Weinreb, L L (2000) The right to privacy, *Social Philosophy & Policy*, *17*(2), 25.
3. The literature on privacy is vast. Legal, historical, sociological and policy-centred approaches are all available. This discussion centres on the function of privacy and therefore focuses on the philosophical debate.
4. Warren, S D and Brandeis, L (1890) The right to privacy, *Harvard Law Review*, *4*(5), 193. Available at <http://library.louisville.edu/law/brandeis/privacy.html>. Accessed 27 February 2008.
5. Westin, A F (1967) *Privacy and Freedom*, New York: Atheneum Press, 7.
6. Bloustein, E (1964) Privacy as an aspect of human dignity, in F D Schoeman (ed) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 973, 974.
7. Reiman, J H (1978) Privacy, intimacy, and personhood, *Philosophy and Public Affairs*, *6*(1), 26, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 300–316.
8. Benn, S I (1971) Privacy, freedom, and respect for persons, in R Pennock and J Chapman (eds) *Privacy Nomos XIII*, New York, NY: Atherton, 26.
9. Van Hove, E (1995) The legislation on privacy protection and social research, *Computers in Human Services*, *12*(2), 53.
10. Clarke, R (1999) Internet privacy concerns confirm the Ccse for intervention, *Communications of the ACM 60*, *42*(2). Available at <http://www.anu.edu.au/people/Roger.Clarke/DV/CACM99.html>. Accessed 27 February 2008.
11. Van Den Haag, E (1971) On privacy, *Nomos*, *13*, 149 cited in L D Introna (1997) Privacy and the computer: why we need privacy in the information society, *Metaphilosophy*, *28*(3), 262.
12. Rachels, J (1975), Why privacy is important, *Philosophy and Public Affairs*, *4*, 292.
13. DeCew, J (1997) *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology*, Ithaca: Cornell University Press, 61.
14. *Ibid*., 62.

15. *Ibid*., 78.
16. *Ibid*., 75.
17. *Ibid*., 75.
18. *Ibid*., 76.
19. Benn, S I (1984) Privacy, freedom, and respect for persons, in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*: New York: Cambridge University Press, 223–244; Prosser, W L (1960) Privacy, *California Law Review*, *48*(3), 338, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 104–155; Reiman, J H (1978) Privacy, intimacy, and personhood, *Philosophy and Public Affairs*, *6*(1), 26, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 300–316.
20. Schoeman, F D *Privacy and Social Freedom*, op cit, fn 21, 15.
21. Goffman, E (1963) *Behaviour in Public Places: Notes on the Social Organization of Gatherings*, New York: The Free Press; Goffman, E (1969) *The Presentation of Self in Everyday Life*, London: Penguin.
22. Fried, C (1968) Privacy, *Yale Law Journal*, *77*(3), 475, reprinted in F D Schoeman (ed) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 203–223; Gerstein, R G (1978) Intimacy and privacy, *Ethics*, *89*(1), 76, in F D Schoeman, (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 265–272; Reiman, J H, *op cit*, fn 19.
23. Giddens, A (1994) *Modernity and Self-Identity: Self and Society in the Late Modern Age*, Palo Alto: Stanford University Press.
24. *Ibid*., 54.
25. *Ibid*., 185.
26. *Ibid*., 186.
27. *Ibid*., 148.
28. Manchester University, Manchester Metropolitan University (student mail list), and Queen's University Belfast (computer and law students only).
29. Herring, S C, Scheidt, L A, Bonus, S. and Wright, E (2004a) Bridging the gap: a genre analysis of weblogs, *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)*, Los Alamitos: IEEE Computer Society Press. Available at <http://www.blogninja.com/DDGDD04.doc>, 5. Last accessed 27 February 2008.
30. Combination of married and living together.
31. Herring, S C, Scheidt, L A, Bonus, S and Wright, E (2004). Bridging the gap: a genre analysis of weblogs, *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)*, Los Alamitos: IEEE Computer Society Press. Available at <http://www.blogninja.com/DDGDD04.doc>. Accessed 27 February 2008; Nardi, B, Schiano, D and Gumbrecht, M (2004) Blogging as social activity, or, would you let 900 million people read your diary?, *Proceedings of Computer Supported Cooperative Work 2004*. Available at <http://home.comcast.net/%7Ediane.schiano/CSCW04.Blog.pdf>. Accessed 27 February 2008.
32. Whyte, W H (1990) *City: Rediscovering the Center*, New York, NY: Anchor.
33. Herring, S C, Scheidt, L A, Bonus, S. and Wright, E (2004a) Bridging the gap: a genre analysis of weblogs, *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)*, Los Alamitos: IEEE Computer Society Press. Available at <http://www.blogninja.com/DDGDD04.doc>. Accessed 27 February 2008.
34. *Ibid*., 10.
35. Herring, S C, Kouper, I, Scheidt, L A. and Wright, E (2004b) Women and children last: the discursive construction of weblogs, in L Gurak, S Antonijevic, L Johnson, C Ratliff and J Reyman (eds), *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs*. Available at <http://blog.lib.umn.edu/blogosphere/women_and_children.html>. Last accessed 27 February 2008.
36. Nardi, B, Schiano, D, and Gumbrecht, M (2004) Blogging as social activity, or, would you let 900 million people read your diary?, *Proceedings of Computer Supported Cooperative Work 2004*. Available at <http://home.comcast.net/%7Ediane.schiano/CSCW04.Blog.pdf>, 4. Accessed 27 February 2008.
37. Nardi, B, Schiano, D, and Gumbrecht, M (2004) Blogging as social activity, or, would you let 900 million people read your diary?, *Proceedings of Computer Supported Cooperative Work 2004*. Available at <http://home.comcast.net/%7Ediane.schiano/CSCW04.Blog.pdf>, 6. Accessed 27 February 2008.

38. Gumbrecht, M (2004) Blogs as 'protected space', *WWW2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*. Available at <http://www.blogpulse.com/papers/www2004gumbrecht.pdf>. Accessed 27 February 2008.
39. Palen, L and Dourish, P (2003) Unpacking 'privacy' for a networked world, Proceedings of the ACM CHI. Available at <http://delivery.acm.org/10.1145/650000/642635/p129-palen.pdf?key1=642635&key2=7273414711&coll=&dl=GUIDE&CFID=15151515&CFTOKEN=6184618>. Accessed 27 February 2008.
40. Rosen, J (2000) *The Unwanted Gaze: The Destruction of Privacy in America*, Vintage Books: New York.
41. Grudin, J (2001) Desituating action: digital representation of context, *Human-Computer Interaction, 16*(2–3), 269–286.
42. *Op cit*, fn 41.
43. Grudin, J (2001) Desituating action: digital representation of context, *Human-Computer Interaction, 16*(2–3), 269–286.
44. Grudin, J (2001) Desituating action: digital representation of context, *Human-Computer Interaction*, *16*(2–3), 11.
45. Rosen, J (2000) *The Unwanted Gaze: The Destruction of Privacy in America*, New York: NY: Vintage Books.

## References


Benn, S I (1984) Privacy, freedom, and respect for persons, in F D Schoeman (ed) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 223–244.

Bloustein, E J (1984) Privacy as an aspect of human dignity, in F D Schoeman (ed) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press.

Bray, H (2004) Job blogs hold perils, opportunities, The Boston Globe, p1, in F B Viegas (2005) Bloggers expectations of privacy and accountability: an initial survey, *Journal of Computer-Mediated Communication*, *10*(3), article 12. Available at http://jcmc.indiana.edu/vol10/issue3/viegas.html. Accessed 27 February 2008.

Clarke, R (1999) Internet privacy concerns confirm the case for intervention, *42(2) Communications of the ACM 60*. Available at <http://www.anu.edu.au/people/Roger.Clarke/DV/CACM99.html>. Accessed 27 February 2008.

DeCew, J W (1997) *In Pursuit of Privacy: Law, Ethics & the Rise of Technology*, Ithaca: Cornell University Press.

Epstein, R A (2000) Deconstructing privacy: and putting it back together again, *Social Philosophy and Policy 17*(2), 1.

Frey, R G (2000) Privacy, control, and talk of rights, *Social Philosophy & Policy 17*(2), 45.

Fried, C (1968) Privacy, *Yale Law Journal 77*(3), 475, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 203–223.

Gavison, R (1984) Privacy and the limits of law, in F D Schoeman (ed) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 346–403.

Gerstein, R S (1978) Intimacy and privacy, *Ethics 89*(1), 76, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 265–72.

Giddens, A (1994) *Modernity and Self-Identity: Self and Society in the Late Modern Age*, Palo Alto: Stanford University Press.

Glenn, D (2003) Scholars who blog, *The Chronicle of Higher Education*, *49*(39), A14.

Goffman, E (1963) *Behavior in Public Places: Notes on the Social Organization of Gatherings*, New York: The Free Press.

Goffman, E (1969) *The Presentation of Self in Everyday Life*, London: Penguin.

Gumbrecht, M (2004) Blogs as 'protected space', *WWW2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics*. Available at <http://www.blogpulse.com/papers/www2004gumbrecht.pdf>. Accessed 5 January 2005.

Grudin, J (2001) Desituating action: digital representation of context, *Human-Computer Interaction*, *16*(2–3), 269–286.

Herring, S C, Scheidt, L A, Bonus, S and Wright, E (2004a) Bridging the gap: a genre analysis of weblogs, *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS-37)*, Los Alamitos: IEEE Computer Society Press. Available at <http://www.blogninja.com/DDGDD04.doc>. Accessed 26 February 2008.

Herring, S C, Kouper, I, Scheidt, L A and Wright, E (2004b) Women and children last: the discursive construction of weblogs, in L Gurak, S Antonijevic, L Johnson, C Ratliff, and J Reyman (eds) *Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs*, University of Minnesota. Available at <http://blog.lib.umn.edu/blogosphere/women_and_children.html>. Accessed 26 February 2008.

Herring, S C, Scheidt, L A, Bonus, S and Wright, E (2005) Weblogs as a bridging genre, *Information, Technology & People*, *18*(2), 142–171.

Nardi, B, Schiano, D and Gumbrecht, M (2004) Blogging as social activity, or, would you let 900 million people read your diary?, *Proceedings of Computer Supported Cooperative Work 2004*. Available at <http://delivery.acm.org/10.1145/1040000/1031643/p222-nardi.pdf?key1=1031643&key2=7505414711&coll=Portal&dl=ACM&CFID=13863829&CFTOKEN=98798918>. Accessed 27 February 2008.

Nussbaum, E (2004) My so-called blog, *The New York Times*, Late Edition, Section 6, 33.

O'Shea, W (2003) The sharer of secrets, *Village Voice*, 55.

Palen, L and Dourish, P (2003) Unpacking 'privacy' for a networked world, *Proceedings of the ACM CHI*, Ft Lauderdale, FA. Available at <http://delivery.acm.org/10.1145/650000/642635/p129-palen.pdf?key1=642635&key2=7273414711&coll=&dl=GUIDE&CFID=15151515&CFTOKEN=6184618>. Accessed 27 February 2008.

Pax, S (2003 September 9) I became the profane pervert Arab blogger, *The Guardian*, 2.

Prosser, W L (1960) Privacy, *California Law Review 48*(3), 338, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 104–155.

Rachels, J (1975) Why privacy is important, *Philosophy and Public Affairs 4*, 323–333.

Reiman, J H (1978) Privacy, intimacy, and personhood, *Philosophy and Public Affairs 6*(1), 26, reprinted in F D Schoeman (ed) (1984) *Philosophical Dimensions of Privacy: An Anthology*, New York: Cambridge University Press, 300–316.

Rosen, J (2000 April 20) The Eroded Self, *The New York Times* 46. Available at <http://people.brandeis.edu/~teuber/rosen1.html>. Accessed 27 February 2008.

Rosenberg, A (2000) Privacy as a matter of taste and right, *Social Philosophy & Policy 17*(2), 68.

Schoeman, F D (1988) *Privacy and Social Freedom*, Cambridge: Cambridge University Press.

Turnbull, G (2001) The state of the blog. Interview with Evan Williams, *Writing the Web*. Available at <http://writetheweb.com/Members/gilest/old/106/view>. Accessed 5 January 2005.

Turnbull, G (2004) The seven-year-old bloggers, *BBC News World Edition*. Available at <http://news.bbc.co.uk/2/hi/uk_news/magazine/3804773.stm>. Accessed 16 March 2007.

Van Den Haag, E (1971) On Privacy, in L D Introna (1997) Privacy and the computer: why we need privacy in the information society, *Metaphilosophy 28*(3), 259.

Van Hove, E (1995) The legislation on privacy protection and social research, *Computers in Human Services 12*(2), 53.

Viegas, F B (2005) Bloggers' expectations of privacy and accountability: an initial survey, *Journal of Computer-Mediated Communication*, *10*(3), article 12. Available at <http://jcmc.indiana.edu/vol10/issue3/viegas.html>. Accessed 27 February 2008.

Warren, S D and Brandeis, L (1890) The right to privacy, *Harvard Law Review 4*(5), 193. Available at <http://library.louisville.edu/law/brandeis/privacy.html>. Accessed 27 February 2008.

Weinreb, L L (2000) The right to privacy, *Social Philosophy & Policy 17*(2), 25.

Westin, A F (1967) *Privacy and Freedom*, New York: Atheneum Press.

Whitworth, D (2003 June 30) The (not so) secret diary of a blogger, *The Times*, 4.

Whyte, W H (1990) *City: Rediscovering the Center*, New York, NY: Anchor.

# Data Sensitivity: Proposals for Resolving the Conundrum

**Karen McCullagh**[1]

**Abstract:** The EU Directive 95/46/EC specifically demarcates categories of sensitive data meriting special protection. It is important to review the continuing relevance of existing categories of sensitive data in the light of changes in societal structures and advances in technology. This paper draws on interviews with privacy and data protection experts from a range of countries and disciplines and findings from the Information Commissioner's annual telephone survey of the British public in order to explore satisfaction with the current categories of sensitive data. It will be shown that the current classification of sensitive data appears somewhat outdated and thus ineffective for determining the conditions of data processing. Finally, possible reform proposals will be reviewed, including a purpose-based approach and context-based approach.

## 1. Origins of Protections for Sensitive Data

The concept of 'sensitive' data was first considered for introduction into international law by the expert group drafting the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (1980).[2] Sweden and the German state of Hesse had already incorporated the concept into national and state law.[3] Ultimately the drafters of the Guidelines decided not to include extra safeguards for designated categories of sensitive data. The absence of safeguards seems to be partly due to a failure to achieve consensus on which categories of data deserve special protection, as the guidelines state:

> ...it is probably not possible to define a set of data which are universally regarded as being sensitive. (para 19 (a)).

Moreover this approach may also reflect the belief that personal data is not categorically deserving of protection, but instead that appropriate protection is dependent upon the context in which the data are used.

Although the Guidelines are not binding on OECD Member States, they have influenced the enactment of data protection legislation in both EU and non-EU member countries, such as Australia, New Zealand and Hong Kong. Recently, the twenty one Asia-Pacific Economic Cooperation (APEC) member economies[4] adopted the *APEC Privacy Framework*, which claims that its Framework is 'consistent with the core values' of the Guidelines.[5] However, since the guidelines were not legally binding on any of the member countries, they did not serve as the international data protection law that they were intended to be (Walczuch & Steeghs, 2001). Indeed, experts opined that the guidelines overemphasised the principle of unrestricted trans-border data flows at the expense of the privacy interest of the data subjects (Ellger, 1987).[6] Furthermore, Kirby[7] conducted a review of the Guidelines and suggested that they need to be updated to include new privacy principles appropriate for contemporary technology, such as internet based automatic profiling.

---

[1] PhD candidate, CCSR, University of Manchester, Email: <Karen.mccullagh@postgrad.manchester.ac.uk> This researcher is sponsored by the ESRC and Office of The Information Commissioner, UK. All views expressed in this article are those of the author and do not necessarily represent the views of, and should not be attributed to either of the Sponsors.

[2] http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html

[3] The emergence of data protection laws, starting in Hessen (Hesse is English translation) 1970 and Sweden 1973, was closely linked to use of computer technology as a tool for collecting and distributing personal data. See Sieghart, P. (1976), *Privacy and Computers*, Latimer, London.

[4] There are 21 member economies. See: http://www.mapsofworld.com/apec-member-economies.htm

[5] Greenleaf claims that The Framework is in fact weaker in significant respects than the OECD Guidelines, to some extent in its principles but particularly in its implementation requirements. Greenleaf, G. (2005) "APEC's Privacy Framework: A new low standard," *Privacy Law & Policy Reporter* Vol. 11 No 5, 121- 125

[6] Ellger, R. (1987), "European data protection laws as non-tariff barriers to the transborder flow of information," in Mestmaecker, E.-J. (Ed.), *The Law and Economics of Transborder Telecommunications, A Symposium,* Nomos Verlagsgesellschaft, Baden-Baden, 121-43.

[7] Kirby, M. (1999) "Privacy protection, a new beginning: OECD principles 20 years on," *Privacy Law & Policy Reporter*, Vol. 6 No 3, 25-30

Thereafter the concept of sensitive data was introduced into international law through the Council of Europe Convention For The Protection of Individuals With Regard To Automatic Processing Of Personal Data (1981).[8] Although the Explanatory Report[9] advocates a context based approach to determining risk of harm from personal data processing, it recognises exceptional cases where the processing of certain categories of data may encroach on individual rights and privacy interests.[10] These 'sensitive' categories are listed in Article 6 as:

> Personal data revealing racial origin, political opinions or religious or other beliefs, as well as personal data concerning health or sexual life, may not be processed automatically unless domestic law provides appropriate safeguards. The same shall apply to personal data relating to criminal convictions.

Paragraph 44 of the Explanatory Report states that "revealing ... political opinions, religious or other beliefs" also covers activities resulting from such opinions or beliefs. Paragraph 45 indicates that "personal data concerning health" includes information concerning the past, present and future, physical or mental health of an individual. The information may refer to a person who is sick, healthy or deceased. This category of data also covers those relating to abuse of alcohol or the taking of drugs.

The categories listed in Article 6 are not meant to be exhaustive. Rather, the Convention provides that a Contracting State should be free to include other categories of sensitive data. Data sensitivity depends on the legal and sociological context of the country concerned:

> Information on trade union membership for example may be considered to entail as such a privacy risk in one country, whereas in other countries it is considered sensitive only in so far as it is closely connected with political or religious views. (para 48)

The Council of Europe Convention, in contrast with the OECD guidelines, had to be incorporated into domestic law by the countries that acceded to it. However, not all the Member States passed data protection laws and in those which did, the laws were not all consistent with one another. For instance, the UK law did not cover any manual data , whereas the Hesse data protection law did. The UK had a detailed system of registration, whereas others did not.[11] Hence, the Convention did not succeed in bringing about the full harmonization of data protection laws.

Subsequently, the United Nations issued **Guidelines for the Regulation of Computerized Personal Data Files** (1990)[12] **,** which addressed the issue of sensitive data under a *Principle of non-discrimination.* The Guidelines defined such data as:

> …data likely to give rise to unlawful or arbitrary discrimination, including information on racial or ethnic origin, colour, sex life, political opinions, religious, philosophical and other beliefs as well as membership of an association or trade union, should not be compiled.[13]

This international treaty is broader than the Council of Europe convention (discussed above), as it includes the categories ethnic origin and colour. In addition, it includes membership of trade unions or other associations. However, it does not include criminal convictions or health data. Both the convention and the guidelines provided for States provide opportunities to regulate risks stemming from the processing of personal data by applying an internationally approved regulatory model. Indeed, they remained free to enact rules that better fulfilled their requirements, or even to abstain from any legislative action. Table 1 displays the categories of data listed as sensitive in the three international legislation discussed in the preceding section.

[8] European Treaty Series - No. 108, (28.I.1981), http://conventions.coe.int/Treaty/EN/Treaties/Html/108.htm
[9] http://conventions.coe.int/Treaty/EN/Reports/HTML/108.htm
[10] Paragraph 43.
[11] Jay, R. (2004) "The Data Protection Act 1998 (DPA)" JISC Legal Information Service Briefing Paper
[12] Adopted by General Assembly resolution 45/95 of 14 December 1990
[13] http://ec.europa.eu/justice_home/fsj/privacy/instruments/un_en.htm Principle 5

*Table 1: Categories of sensitive data in International Legislation*

| OECD Guidelines (1980) | Council of Europe Convention (1981) | UN Guidelines (1990) |
|---|---|---|
| None | Racial origin | Racial or Ethnic origin |
| | Political opinions | Political opinions |
| | Religious or other beliefs | Religious/philosophical/other beliefs |
| | Sexual life data | Sex life |
| | Health data | Membership of a trade union |
| | Criminal convictions | Membership of an association |
| | | Colour |

As time passed, an increasing number of countries introduced data protection laws and tighter restrictions on trans-border data flows across national borders were implemented. Many countries with strong data protection interdicted the transfer of protected data to countries with less strong or no data protection measures. This severely impeded the business of some multinational companies. An example of this occurred in 1989, when French authorities halted the transfer of personnel records from Fiat's French office to the Italian base office because Italy had no data protection legislation at that time, while France had high levels of protection (Mei, 1993).[14]

### 1.1 Current EU definition of sensitive data

In order to remove obstacles to the free movement of data without diminishing the protection of personal data, the European Commission decided to harmonize data protection and proposed Directive 95/46/EC (the Directive).[15] The Directive includes a provision that sensitive data must be more stringently protected.[16] Such data is defined in Article 8 (1) as:

> Personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and … data concerning health or sex life.

Article 8(5) also makes special provision for criminal records and the like:

> Processing of data relating to offences, criminal convictions or security measures may be carried out only under the control of official authority, or if suitable specific safeguards are provided under national law, subject to derogations which may be granted by the Member State under national provisions providing suitable safeguards…

Thus, the principle of sensitivity holds that the processing of eight types of data should be subject to stricter controls than other types of personal data. The Directive differs from the Council of Europe's approach in two main respects: 1) it includes the trade union membership as a specific category of sensitive data; 2) the list is considered exhaustive, whereas the Council of Europe list is merely indicative. The Directive differs from the UN Guidelines as it lacks a category of data on colour or membership of association, but includes a category of

[14] Mei, P. (1993), `"The EC proposed data protection law," *Law and Policy in International Business*, Vol. 25, 305-34.

[15] http://europa.eu.int/comm/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf http://europa.eu.int/comm/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part2_en.pdf

[16] In principle, such data cannot be processed. Derogation is permitted under very specific circumstances. These circumstances include the data subject's explicit consent, processing mandated by employment law, where it may be impossible for the data subject to consent (e.g. blood test to the victim of a road accident), processing of data has been publicly announced by the data subject or processing of data about members by trade unions, political parties or churches. Member states may provide for additional exceptions for reasons of substantial public interest.

criminal convictions. A more radical difference exists between the Directive and the OECD guidelines, in which drafters adopt a contextual approach and do not specifically enumerate special categories of sensitive data.

It is important to review the continuing relevance of existing categories of sensitive data in the Directive in the light of changes in societal structures and advances in technology. In the pre-computer era, data processing was not automatic and large-scale, uncontrolled surveillance was costly, thus providing natural barriers for privacy protection. These natural barriers disappeared gradually in the mid 1960s because computerized technology for processing an increasing amount of information needed to develop social welfare-states was available at faster speeds and lower costs.[17] Also, business organizations owning large amounts of records started to use computers. By the 21st-century, businesses are such that customers expect them to operate at all times. It is not only the e-commerce world that experiences this situation. All types of organizations - including health care, financial, manufacturing, and services operate around the clock, or at least their computer systems do. Even when no humans are around, computers are available to take and place orders, send orders to the warehouse, and manage financial transactions, all involving the processing of potentially sensitive personal information.

Several issues arise: firstly, are the current categories still considered sensitive? Secondly, have new categories of sensitive data emerged? If new categories have emerged, can the current legislation incorporate them? Should the list be extended or should an alternative approach be adopted? These issues were explored through semi-structured interviews with experts and findings from the Information Commissioner's annual telephone survey of the British public.

## 2. Current categories of sensitive data

### 2.1 Responses from expert interviewees

Interviews were conducted with thirty seven privacy and data protection experts from a range of disciplines, including privacy commissioners, lawyers, industry experts, statistical methodologists, computer scientists, and academics from a variety of disciplines including sociology, market research and law.[18] In the interviews semi-structured questions were used. The aim was to have a discussion with the respondent so that all the themes in the interview guide were covered. Some of the themes in the interview guide were too complex for a few of the participants. For instance, statistical methodologists were not comfortable when answering questions about the specific detail of the legislation in their country. Accordingly, the researcher was creative and aware of the need to see the issue from the interviewee's position and asked the questions in an appropriate but not leading way. The advantage of this approach is that it allowed a cognitive process to emerge, so that the information obtained from respondents provided not just answers, but reasons for the answers.

Some respondents were happy with the existing definition and the types of data covered. For example, one respondent stated:

> In the UK existing categories of sensitive data have merit in that they are associated with a right to human dignity/freedom of political activity. The difficulty with the current provisions is the overriding public interest tests, in the EU Directive there is a categorical prohibition on the processing of certain data – but it is subject to higher public interest tests…Existing categories of sensitive data are sensible. (UK)

Likewise,

> I'm broadly happy with existing definitions in Ireland. The approach taken in the Directive is correct. Sensitive data is an arbitrary list. (Ireland)

Others did not agree with all classifications,

> We had to introduce the concept of sensitive data in Iceland but we don't agree with all the categories e.g. according to the Directive data on trade unions is considered sensitive, but in Iceland such information is not as everyone knows where you work and what unions you belong to and they don't care about these things… (Iceland)

[17] Mayer–Schönberger 1997, 222, For a discussion of the connection between large databases and social welfare state. Mayer-Schönberger, V. (1997) Generational Development of Data Protection in Europe. *In* Agre, P E. & Rotenberg, M. (eds.) 1997. *Technology and Privacy: The New Landscape*. Cambridge, Massachusetts, MIT Press.

[18] A respondent matrix was created using quota and snowball sampling. Snowballing is an effective technique for building up a reasonable sized sample, especially when used as part of a small-scale research project (Denscombe 1998).

Also, some Interviewees suggested new categories of sensitive data. Below are some illustrations:

> Some regard or suggest financial data to be sensitive – in this regard the categorisation of it as non-sensitive is clearly arbitrary – it may be worthwhile amending the legislation to make it sensitive. (Ireland)

Interviewees indicated that technological developments are generating potential new categories of sensitive data, for example

> They could be expanded e.g. to include financial data. They could be ramified. E.g. for health data a biometric template[19] should probably be considered personal data but probably isn't sensitive data. Whereas, genetic information could be regarded as sensitive because of the potential for prejudice and unfairness of inappropriate disclosure. (UK)

The concept of data protection through legislation is essentially an issue of formal public policy recognition and protection being accorded to values and ideologies that are held to be important by individuals and that are institutionalized in any individual culture (Ajami,1990).Thus, it is important to ascertain if the legal definitions accord with the views of the public, who often play the role of a data subject, as government legislative initiatives are intended to give effect to the legal requirements of a society, and will only be successful if they are valued and supported by the public.

### 2.2 Findings from ICO Annual Track telephone survey of British public

The views of UK citizens regarding the concept of sensitive data were sought through the ICO Annual Track (Individual survey 2006).[20] The survey was designed to examine public perceptions of sensitive data. Firstly, it was used to test sensitivity ratings of seven categories of data which are currently recognised in the Directive as sensitive. Also, it was used to test perceptions of sensitivity towards eight not legally recognised categories of sensitive data which emerged in interviews with data protection and privacy experts. The 15 categories of sensitive data tested are displayed in Table 2.

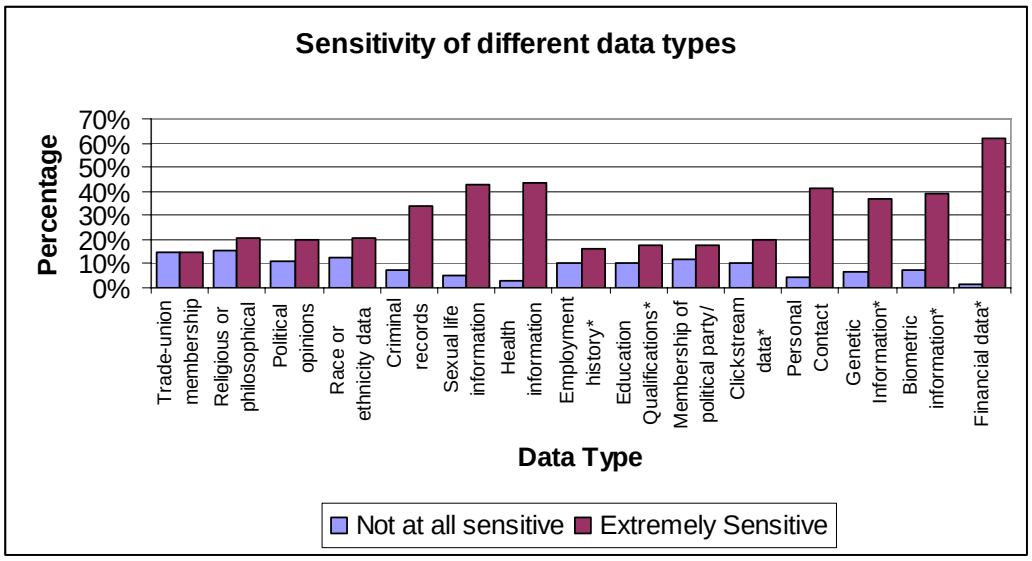
Table 2: *Classification of sensitive data*

| Art 8 Legally recognised categories | Not legally recognised categories |
|---|---|
| *Trade-union membership* | *Employment history* |
| *Religious or philosophical beliefs* | *Education Qualifications* |
| *Political opinions* | *Membership of political party / organisation* |
| *Data concerning race or ethnic origin* | *Clickstream data (e.g. record of web pages visited)* |
| *Criminal records* | *Personal Contact Details* |
| *Sexual life information* | *Genetic Information* |
| *Health information* | *Biometric information (e.g. iris scans, facial scans and finger prints)* |
| | *Financial data* |

[19] Biometrics comes from the Greek words *bios* (life) and *metrikos* (measure). The term refers to any specific and uniquely identifiable physical human characteristic, *e.g.*, of the retina, iris, acoustic spectrum of the voice (*i.e.*, voiceprint), fingerprint(s), handwriting, pattern of finger lengths, *etc.*, that may be used to validate the identity of an individual.

[20] The survey was conducted by telephone. All the interviews were conducted in house by SMSR's telephone interviewing team. The total sample was 1,066 interviews.[20] Quotas were set on age, sex, region and social grade to ensure a nationally representative sample was achieved.

Figure 1: *Sensitivity of different data types*




(Source: ICO Annual Track Survey 2006) (n=1066)

Fig. 1 shows how respondents rated different types of data on a scale of 1 to 10 with 1 being not at all sensitive and 10 extremely sensitive. The results indicate that of the legally-recognised types of sensitive data, health and sex life information were considered extremely sensitive by the highest percentage of respondents. However, some of the other categories were considered to be more sensitive than the legally-recognised types of sensitive data. For instance, financial data was considered extremely sensitive by most respondents (62.1%), while religious opinions were considered to be not at all sensitive by 15.3% of respondents. Likewise, more than one third of respondents rated biometric, genetic and contact details as extremely sensitive, whereas only one fifth of respondents rated data concerning race or ethnic origin, political opinions or data concerning religious or philosophical beliefs as extremely sensitive.

The 10 scale data rating was recoded into five categories (see Table 3). The data was analysed and is displayed in tables according to whether it is classified as legally recognised or not legally recognised as a category of sensitive data.

*Table 3: Recoding of data sensitivity scale from 10 scale into 5 categories*

| Original value | Recode value | Category Label |
|---|---|---|
| 1, 2 | 1 | Not at all Sensitive |
| 3, 4 | 2 | A little Sensitive |
| 5, 6 | 3 | Sensitive |
| 7, 8 | 4 | Very Sensitive |
| 9, 10 | 5 | Extremely Sensitive |

Table 4: *Sensitivity of legally recognised data types – ICO Survey*

| | Trade-union membership | Religious or philosophical beliefs | Political opinions | Data concerning race or ethnic origin | Criminal records | Sexual life information | Health information |
|---|---|---|---|---|---|---|---|
| Don't Know | 1.4% | .9% | .9% | 1.3% | 1.1% | 1.6% | .9% |
| Not at All Sensitive | 21.6% | 21.4% | 15.9% | 19.5% | 11.2% | 6.8% | 3.8% |
| A little Sensitive | 13.9% | 12.1% | 13.1% | 11.2% | 7.2% | 6.7% | 5.1% |
| Sensitive | 30.6% | 28.2% | 28.1% | 1.3% | 22.9% | 18.0% | 18.2% |
| Very Sensitive | 15.1% | 13.3% | 17.4% | 16.6% | 17.5% | 17.1% | 20.6% |
| Extremely Sensitive | 17.4% | 24.0% | 24.5% | 24.6% | 40.1% | 49.8% | 51.3% |
| | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |

(Source: ICO Annual Track Survey 2006) (n=1066)

Table 4 displays the legally recognised categories of sensitive data and indicates that Health data was considered extremely sensitive by over half of the respondents (51.3%), and almost half considered sexual life information to be extremely sensitive, whereas fewer respondents considered religious or philosophical beliefs to be extremely sensitive (24%) and only 17.4% considered trade union membership data to be extremely sensitive. Thus, some of the legally recognised categories of sensitive data are considered less sensitive than others.

Table 5: *Sensitivity of not legally recognised data types- ICO Survey*

| | Employment history | Education Qualifications | Membership of political party / organisation | Clickstream data (e.g. record of web pages visited) | Personal Contact Details | Genetic Information | Biometric information (e.g. iris scans, facial scans and finger prints) | Financial data |
|---|---|---|---|---|---|---|---|---|
| Don't Know | 1.1% | 1.4% | 1.5% | 2.5% | .4% | 1.6% | 2.2% | .6% |
| Not at All Sensitive | 15.9% | 15.9% | 17.4% | 15.9% | 7.0% | 8.7% | 10.4% | 1.9% |
| A little Sensitive | 12.1% | 11.7% | 13.4% | 11.4% | 7.1% | 6.3% | 6.8% | 2.5% |
| Sensitive | 30.2% | 29.5% | 30.1% | 27.5% | 17.8% | 20.8% | 17.5% | 7.1% |
| Very Sensitive | 19.3% | 19.5% | 15.5% | 18.2% | 21.4% | 19.2% | 17.6% | 17.4% |
| Extremely Sensitive | 21.3% | 22.0% | 22.0% | 24.4% | 46.2% | 43.3% | 45.4% | 70.5% |
| | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |

(Source: ICO Annual Track Survey 2006) (n=1066)

Table 5 displays categories of sensitive data that are not legally recognised. The table indicates that financial data was considered extremely sensitive by over 70% of respondents, and just under half (46.4%) considered their personal contact details extremely sensitive, whereas only 21.3% of respondents considered employment history data to be extremely sensitive. The finding from the survey indicates that one fifth of telephone respondents considered trade union membership, religious/philosophical beliefs or data concerning racial/ethnic origin to be not at all sensitive.

However, further research is needed before imposing cut-off sensitivity points as 'sensitivity' is a value which has both an objective and subjective component. Buchholz comments,

> Values are different between people and reflect individual desires and beliefs. Values are properties that human beings associate with or assign to certain forms of human behavior, institutions, or material goods and services. (1992, p. 118)

From a subjective viewpoint, the sensitivity of a particular value is derived through individuals making personal judgements. In contrast from an objective perspective, the sensitivity of a particular value is derived outside the personal experiences of those individuals faced with choices. In this situation values are part and parcel of the behaviour or object in question. A complex interaction between these two perspectives leads to the creation of commonly held societal values that are believed to produce desirable outcomes for society as a whole (Buchholz, 1992; Daleiden, 1990). The conflict between the objective and subjective viewpoints is resolved through the essence of public policy formulation process, i.e. negotiation and compromise (Rule, 1974; Sieghart, 1984). Thus, further research is needed to test, for example, whether the respondents who indicated that race or ethnicity data was not at all sensitive were drawn from the majority white UK population, or whether similar views were expressed by the minority ethnic population.[21] Also, further research is needed to explain the reduced the sensitivity of such information, for instance, whether the Race Relations Act has been successful in protecting the rights and interests of ethnic minorities to the extent that such information is considered not at all sensitive by the UK population. Likewise, have changes in employment law for example equal opportunities[22] and minimum wage legislation reduced the sensitivity of trade union membership information?

The findings suggest that the current list is in need of reform, as *prima facie* it does not reflect the sensitivity perceptions of data subjects. Moreover, the findings suggest that new categories of sensitive data are emerging due to changes in society and technological developments. For instance, amidst post-September 2001 security concerns the UK government proposed the introduction of identity cards which rely on biometrics.[23] Such technology did not exist during the World-War II era when the UK previously utilised identity cards, and indeed they were removed from circulation in 1952, amid widespread public resentment.[24] This raises the issue of whether the current list of sensitive data could or should be amended? Is it possible to formulate an objective category of sensitive information despite claims that sensitivity is *relative* to the individual; and a function of the *context* in which the information is used rather than the type of information itself?

## 3. Criticisms of current approach

Korrf (2002)[25] conducted a comparative textual analysis of legislation. He found that the French, Austrian, British, Czech, Estonian, Finnish, Greek, Hungarian, Italian, Spanish, and Swiss laws state that the list their legislation contains is exhaustive, while, some countries (for instance, Denmark and Iceland) consider their lists as merely indicative. However, all laws provide ways and means to reopen the apparently closed list. For instance, the Estonian act states that the list can be modified by law, so *prima facie* the list of sensitive data categories could be amended.

However, creating new categories raises difficulties, for instance, Luxembourg, and the Netherlands define genetic data as data on *health*, whilst Portugal defines it as data on *health and sex life*, whereas in Sweden the processing of such data not formally regarded as falling within the specific category to which the rules on "*sensitive data*" apply. Hungary and Poland have added to Art 8 (1) "details of addictions". Many addictions would clearly fall within the health related category set out in the Directive already: for example drug addiction and alcoholism. Others, such as gambling or computer games, might not. It remains to be seen how regulators will interpret this additional restriction.[26] Thus, any attempts to modify or extend the current list would require transnational agreement otherwise a lack of harmonization will occur, and defeat the objective of the Directive.

---

[21] According to the 2001 12.5% of the population census across England and Wales are ethnic minorities. http://www.cre.gov.uk/diversity/ethnicity/index.html

[22] Employment Rights Act 1996, Sex Discrimination Act 1975, Equal Pay Act 1975, National Minimum Wage Act 1998

[23] Identity Cards Act 2006

[24] *Willcock v Muckle*, [1951] 2 The Times LR 373 The judge in the case said that the cards were an "annoyance" and "tended to turn law-abiding subjects into law breakers".

[25] Korrf, D. (2002) EC Study On Implementation Of Data Protection Directive (Study Contract ETD/2001/B5-3001/A/49)

[26] Linklaters (2004) Hot Topic: History repeats itself: the implementation of EU data protection legislation in the accession countries. http://www.linklaters.com/pdfs/briefings/040517_DP.pdf

Moreover, a definition-based approach has been criticised by some, as it would require a casuistic form of regulation, which is more complex and lengthy. Indeed, Bing[27] attempted to categorise all personal data according to their sensitivity. However, this approach was quickly abandoned, because it failed to delineate clearly the boundaries of the various spheres, why these exist and what might constitute a breach of the, for instance:

> Definitions of sensitive data are very subjective e.g. where you live is sensitive if you have an estranged violent husband. (UK)

Likewise, another respondent opined:

> I don't like the idea of sensitive data. All data is potentially sensitive, depending upon the context. (UK)

Simitis (1973)[28] asserts that detailed personal profiles could be created through the aggregation of ostensibly innocuous information, which could nevertheless have a detrimental impact upon a person's privacy. Thus, interviewees raised the importance of extraneous information, rather than simply relying on a definitional approach to sensitive data. The responses of several interviewees are exemplified by the following:

> I've never made much use of the concept, e.g. your postcode and newspaper preference both appear to be innocuous information. However, if you work for Experian (a credit score, credit report and credit reference agency), you can draw inferences about a person simply based on those two prices of information – that settles the point. How can you define what is sensitive? E.g. if you can work out my political views from my newspaper preference, then arguably my postcode and newspaper preference should be considered sensitive information. (UK)

Accordingly, some interviewees criticised the arbitrary nature of the exhaustive list base on definitions. At this juncture, it is appropriate to review alternative approaches.

## 4. Reform proposals: resolving the sensitivity conundrum

### 4.1 Context-based approach:

Simitis contends that personal data becomes sensitive according to its context. This mirrors the approach formerly adopted by countries such as Austria and Germany, which, prior to the introduction of the data protection directive had consistently rejected all abstract categorisations of personal data and instead focussed on a context-orientated consideration of the data. He asserts that

> Sensitivity is no more perceived as an *a priori* given attribute. On the contrary, any personal datum can, depending on the purpose or the circumstances of the processing be sensitive (Simitis,1999).

This approach reflects the opinions of some of the interviewees, for instance,

> Another example is related to the employment code we have drafted. Health is regarded as sensitive data. All employers keep records of sickness leave, but the issue is: does self-certified sick notes require the same level of protection as a medical note from a GP? Arguably a self-certified note is less sensitive, particularly given that the individual may have told colleagues the reason for their absence…yet no distinction is drawn in the law – but we would advise employers that they should take a common sense approach. (UK)

> The idea that all health information is sensitive is too restrictive in some instances e.g. it can cause difficulties between two contracting parties such as an insurance company and an individual. We need safeguards to protect sensitive uses of sensitive data. (Spain)

[27] Bing, J. (1972) "Classification of Personal Information with respect to the Sensitivity Aspect" Proceedings of the First International Oslo Symposium on Data Banks and Society, 98-141

[28] Simitis, S. (1973) cited in Bygrave, L. (2002) *Data Protection Law: Approaching its Rationale, Logic and Limits*, Kluwer, 132

Simitis reasoned that it is vital to consider contextual information when determining the sensitivity of data. Contextual information includes: the interests of the data controller as well as the potential recipients of the data, the aims for which the data are collected, the conditions of the processing and its possible consequences for the individual and others. An evaluation of the sensitivity requires hence more than a definitional approach to sensitive data. Furthermore, Simitis advocates that sensitivity lists should be purely exemplary, and

> Only where the legislators can fully concentrate on a specific context, are they also able to reach a degree of precision that appropriately responds to the particularities of the processing circumstances. (Simitis, 1999)

This approach is more comprehensive than a definition-based approach, and more likely to reflect the concerns of data subjects. However it would be costly and difficult to implement, as Simitis recognises that it would need to be linked with sectoral regulation.

### 4.2 Purpose-based approach

In contrast, The Council of Europe (2005) proposed a purpose-based approach which would consider the purpose underlying the processing of personal data, that is, whether the processing is intended to reveal sensitive data.

> This approach would make it possible to consider the actual processing of data as sensitive rather that the data itself, even if no sensitive data were involved. For example, a search of trips to Rome conducted by a web surfer using Google or his or her purchases of religious books, reading of a papal encyclical, etc, may be treated as revealing a religious opinion. (Poullet *et al* 2004)

Searching for information on a trip to Rome would not in itself constitute processing of sensitive religious information, but when it is combined with searches for Vatican city visiting hours the purpose of the information processing may change. Of course, searches for such information may be purely coincidental, for instance if a person has heard that a restaurant within the Vatican grounds is worth a visit and checks the opening times etc.

The purpose-based approach mirrors the approach advocated by the OECD guidelines, namely that it is not possible to classify data as sensitive on a definitional basis. Instead, the actual processing of data, rather that the data itself could be considered sensitive. Moreover, Wacks[29] asserts that what changes from context to context is not the degree of sensitivity of information, but the extent to which one is prepared or required to allow it to be disclosed or used. Should the context change, it is not the nature of the information that changes, but an individual's *attitude* towards its use. An individual is likely to have considerably different views about the purposes for which sensitive data is used, for instance, an expert interviewee responded

> I think it will be extremely important to regulate who can access what information and for what reason e.g. whilst it may be acceptable to allow police to deduce racial information through DNA profiling it would not be appropriate to allow a security guard to have access to this type of information when simply determining if an individual should have permission to enter a building (UK)

Wong[30] contends that such an approach would reduce the number of trivial cases being brought before the courts, and also reduce the administrative burden placed upon data protection authorities. Additionally, it would shift the focus away from all data processors on to only those who intentionally reveal data of a sensitive nature. In essence, this is a teleological approach which seeks to prevent information being used in an unfair, harmful or discriminatory manner, and thus meets fulfils the original aim of the directive. However, Wong recognises that this approach leave an unanswered question, namely, *who* should decide what purpose is sensitive? Another unresolved difficulty is *how* to decide whether the purpose for which the data is processed is 'sensitive' if a definition-based approach is not used? No clear guidance is offered for data processors. Another undesirable consequence of this approach is that it pushes the decision regarding compliance away from rule makers onto already overburdened judges.

[29] Wacks, R. (1989) *Personal Information: Privacy and the Law*, Oxford: Clarendon Press, 23, 181
[30] Wong, R. (2007) "Data Protection Online: Alternative Approaches to Sensitive Data*?" Journal of International Commercial Law and Technology*, Vol. 2, No 1

### 4.3 A 'reasonable' approach to resolving the sensitivity conundrum:

It is suggested that a more radical approach should be adopted; one which recognises that the concept of sensitivity is an outdated concept. An expert interviewee opined that

> The concept of sensitive is a failed attempt to capture something, which isn't a natural kind. By saying something is sensitive you are attempting to treat something to do with claim for making different reasons in a single manner. Whereas, life is not reducible to a single algorithm – so you should be wary of this approach. (UK)

Instead of defining categories of sensitive data that deserve stricter protection, legislators should focus on the reasonableness of any request to process personal information. For instance, the province of Alberta, Canada has enacted privacy legislation[31] which does not distinguish between personal and sensitive information. Rather, it seeks to regulate the processing the collection, use and disclosure of personal information by private sector organizations

> in a manner that recognizes both the right of an individual to have his or her personal information protected and the need of organizations to collect, use or disclose personal information for purposes that are *'reasonable'*. (Emphasis added)[32]

The reasonable person test is an objective legal test. Thus an organization needs to be able to demonstrate that it considered the circumstances around handling personal information and made a decision on what is reasonable in the circumstances. The advantage of this approach is that it adopts a holistic approach to the contextual and purposive aspects of data protection. For instance, whilst it might be reasonable for a haulage company employer to insist that driver employees will be subjected to random alcohol for the purpose of ensuring work safety, it would not be reasonable for such an employer to require an employee to disclose any and all past alcohol problems. Mandatory disclosure would be unreasonable as it is too broad and intrusive and could have harmful discriminatory consequences for the employee.

## 5. Conclusion

It is suggested that the time is ripe to review the provisions of the Data Protection Directive. The current definitions reflect post World War II concerns regarding discrimination and protection of human dignity. In the 21st century , new concerns have risen ; for example, developments in the fields IT and biometrics are raising new potential categories of sensitive data. Indeed, findings from interviews and the survey indicate that whilst not all of the legally recognised categories of data continue to be perceived as sensitive, some which are not legally recognised categories of data are emerging which are considered extremely sensitive.

However, a decision to simply include new categories, or delete existing categories should not be taken lightly. Any attempt to grade data according to their sensitivity would be fraught with difficulties, as it would require a casuistic form of regulation, which is more complex and lengthy. Indeed, Bing[33] attempted to categorise all personal data according to their sensitivity. However, this approach was quickly abandoned, because it failed to delineate clearly the boundaries of the various spheres, why these exist and what might constitute a breach of them. For instance, detailed personal profiles could be created through the aggregation of ostensibly innocuous information, which could nevertheless have a detrimental impact upon a person's privacy. As Simitis[34] has shown , sensitivity of data varies from context to context. This contextual approach is more comprehensive than the purpose-based approach, as not only does it consider the purpose for which the data is collected, but also the conditions of processing and the possible consequences for the data subject. Moreover, Wacks[35] asserts that what changes from context to context is not the degree of sensitivity of information, but the extent to which one is

---

[31] The Personal Information Protection Act, (PIPA) does not apply to federally-regulated organizations such as banks, airlines, telecommunications companies and railways. Those organizations are governed by federal privacy legislation.

[32] The Personal Information Protection Act (PIPA) is in force as of January 2004 <http://www.oipc.ab.ca/pipa/>

[33] Bing, J. (1972) "Classification of Personal Information with respect to the Sensitivity Aspect" Proceedings of the First International Oslo Symposium on Data Banks and Society, 98-141

[34] Simitis, S. (1973) cited in Bygrave, L. (2002) *Data Protection Law: Approaching its Rationale, Logic and Limits*, Kluwer, 132

[35] Wacks, R. (1989) Personal *Information: Privacy and the Law*, Oxford: Clarendon Press, 23, 181

prepared or required to allow it to be disclosed or used. Thus, whilst categorisation of sensitive data serves a useful purpose of reminding data processors that unfair discrimination is prohibited, it should be understood as an indicative flexible, reference list. Finally, instead of trying to resolve the sensitivity conundrum, it would be prudent to consider the approach taken by other legislatures who advocate a 'reasonable' approach to data protection.

## Bibliography


1. Ajami, R. (1990), "Transborder data flow: global issues of concern, values, and options", in Lundstedt, S.V. (Ed.), *Telecommunications, Values, and the Public Interest*, Ablex Publishing Corporation, Norwood, NJ
2. Bing, J. (1972) "Classification of Personal Information with respect to the Sensitivity Aspect" Proceedings of the First International Oslo Symposium on Data Banks and Society
3. Buchholz, R.A. (1992), *Business Environment and Public Policy: Implications for Management and Strategy*, Prentice-Hall, Englewood Cliffs, NJ.
4. Bygrave, L.A. (2002) *Data protection law: approaching its rationale, logic and limits*, The Hague: Kluwer Law International.
5. Daleiden, J.L. (1990), "Social considerations in the development of telecommunications policies", in Lundstedt, S.B. (Ed.), *Telecommunications, Values, and the Public Interest*, Ablex Publishing Corporation, Norwood, NJ.
6. Denscombe, M. (1998) The Good Research Guide. Open University Press.
7. Ellger, R. (1987), "European data protection laws as non-tariff barriers to the transborder flow of information", in Mestmaecker, E.-J. (Ed.), *The Law and Economics of Transborder Telecommunications*, A Symposium, Nomos Verlagsgesellschaft, Baden-Baden, 121-43.
8. Greenleaf, G. (2005) "APEC's Privacy Framework: A new low standard," *Privacy Law & Policy Reporter*, Vol. 11 No 5, 121- 125
9. Jay, R. (2004) "The Data Protection Act 1998 (DPA)" JISC Legal Information Service Briefing Paper
10. Kirby, M. (1999) "Privacy protection, a new beginning: OECD principles 20 years on," *Privacy Law & Policy Reporter*, Vol. 6 No 3, 25-30
11. Korrf, D. (2002) EC Study On Implementation Of Data Protection Directive (Study Contract ETD/2001/B5-3001/A/49)
12. Linklaters (2004) Hot Topic: History repeats itself: the implementation of EU data protection legislation in the accession countries. <http://www.linklaters.com/pdfs/briefings/040517_DP.pdf> Last Accessed May 2007
13. Mayer-Schönberger, V. (1997) "Generational Development of Data Protection in Europe," in Agre, P E. & Rotenberg, M. (eds.) (1997) *Technology and Privacy: The New Landscape*, Cambridge, Massachusetts, MIT Press.
14. Mei, P. (1993), "The EC proposed data protection law," *Law and Policy in International Business*, Vol. 25, 305-34.
15. Poullet Y., & Dinant, J-M., (2004) Report On The Application Of Data Protection Principles To The Worldwide Telecommunication Networks: Information Self-Determination In The Internet Era, Thoughts On Convention No. 108 For The Purposes Of The Future Work Of The Consultative Committee (T-PD)
16. <http://www.coe.int/t/f/affaires_juridiques/coop%E9ration_juridique/protection_des_donn%E9es/T-PD%282004%29rapport_Poullet.pdf> Last accessed February 2007
17. Rule, J.B. (1974), *Private Lives and Public Surveillance: Social Control in the Computer Age*, Schocken Books, New York, NY.
18. Sieghart, P. (1984), "Information privacy and the data protection bill", in Bourn, C. and Benyon, J. (eds), *Data Protection: Perspectives on Information Privacy*, University of Leicester, Leicester.
19. Sieghart, P. (1976), *Privacy and Computers*, Latimer, London.
20. Simitis, S. (1999) Revisiting sensitive data,
21. <http://www.coe.int/T/E/Legal%5Faffairs/Legal%5Fco%2Doperation/Data%5Fprotection/Documents/Reports/W-Report%20Simitis.asp#TopOfPage> Last accessed February 2007
22. Wacks, R. (1989) *Personal Information: Privacy and the Law*, Oxford: Clarendon Press
23. Walczuch, R. M. & Steeghs, L. (2001) "Implications of the new EU Directive on data protection for multinational corporations," *Information Technology & People*,
24. Vol. 14 No. 2, 2001, pp. 142-162.
25. Wong, R. (2007) "Data Protection Online: Alternative Approaches to Sensitive Data?" Journal *of International Commercial Law and Technology*, Vol. 2, No 1