BILETA
British & Irish Law, Education and Technology Association
2007 Annual Conference
Hertfordshire 16-17 April
UH

Data Sensitivity: resolving the conundrum[1]

Karen Mc Cullagh
PhD candidate, CCSR, University of Manchester
Email: Karen.mccullagh@postgrad.manchester.ac.uk

Abstract:

Legislators of EU Directive 95/46/EC felt it necessary to specifically demarcate a category of sensitive data meriting further protection. It is important to review the continuing relevance of existing categories of sensitive data in light of changes in societal structures, working conditions 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 and also on findings from a survey of bloggers in order to explore satisfaction with the current categories of sensitive data. It will be shown that the current classification of sensitive data appears a somewhat outdated and ineffective criterion for determining the conditions of data processing. Finally, a number of possible reform proposals will be reviewed, including a purpose-based approach and context-based approach.

Keywords: sensitive data, data protection, reform.

1. Origins of Sensitive Data

The concept of 'sensitive' data was first considered for introduction into international law[2] by the Expert Group drafting the **OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (1980).**[3] Ultimately they decided not to include extra safeguards for designated categories of sensitive data. The absence of safeguards seems to be partly due to a failure by the Expert Group to achieve a consensus on which categories of data deserve special protection, and partly to a belief that sensitivity of personal data is not an *a priori* given but dependent upon the context in which the data are used

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

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)**.[4] Whilst the Explanatory Report[5] advocates a context rather than definitional approach to determining risk of harm from personal data processing, it recognises exceptional cases where the processing of certain categories of data is likely to lead to encroachments on individual rights and interest. [6] These 'sensitive' categories are listed in Article 6 as

---

[1] 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] Swedish and German Hesse Laws had incorporated the concept in national laws.
[3] http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html
[4] European Treaty Series - No. 108, (28.I.1981), http://conventions.coe.int/Treaty/EN/Treaties/Html/108.htm
[5] http://conventions.coe.int/Treaty/EN/Reports/HTML/108.htm
[6] Paragraph 43.

> "Personal data revealing racial origin, political opinions or religious or other beliefs,[7] as well as personal data concerning health[8] 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."

The drafters of the convention asserted that the list is not meant to be exhaustive. Rather, they advocate that a Contracting State should be free to include other categories of sensitive data. They reasoned that the degree of sensitivity of categories of data 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 **UN** issued **Guidelines for the Regulation of Computerized Personal Data Files (1990)[9]** which addressed the issue of sensitive data under the *Principle of non-discrimination*

> "…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."[10]

This international treaty is broader than the Council of Europe convention 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. Through the aforementioned convention and guidelines, States were given 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.

1.2 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).[11] The Directive includes a provision, which states that sensitive data must be treated in a clearly distinct way.[12] It is defined in Article 8 (1) as

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

Also, Art 8(5) makes special provision for criminal records and the like

---

[7] Explanatory Report, Paragraph 44 indicates that "revealing ... political opinions, religious or other beliefs" covers also activities resulting from such opinions or beliefs.

[8] Explanatory Report, 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.

[9] Adopted by General Assembly resolution 45/95 of 14 December 1990

[10] http://ec.europa.eu/justice_home/fsj/privacy/instruments/un_en.htm Principle 5

[11] 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

[12] 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.

> "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, regarded as especially sensitive for data subjects, should be subject to more stringent controls than other types of personal data. The Directive differs from the Council of Europe 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 criminal convictions. A more radical difference exists between the Directive and the OECD guidelines in which the legislators specifically chose not to demarcate special categories of sensitive data.

It is important to review the continuing relevance of existing categories of sensitive data in the Directive in light of changes in societal structures, working conditions and advances in technology. 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, a telephone survey of the British public and an online survey of bloggers from around the world.

2. Current categories of sensitive data:

a) Responses from expert interviewees

Semi-structured interviews were conducted with a range of privacy and data protection experts, including privacy commissioners, lawyers, industry experts, statistical methodologists, computer scientists, and academics from a variety of disciplines e.g. sociology, market research and law.[13]
Some respondents were happy with the existing definition and the types of data covered

> "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)

[13] 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). 37 interviews were conducted.

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 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)

It is important to ascertain if the legal definitions accord 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 have the support of the public.

b) Findings from ICO Annual Track telephone survey of British public.

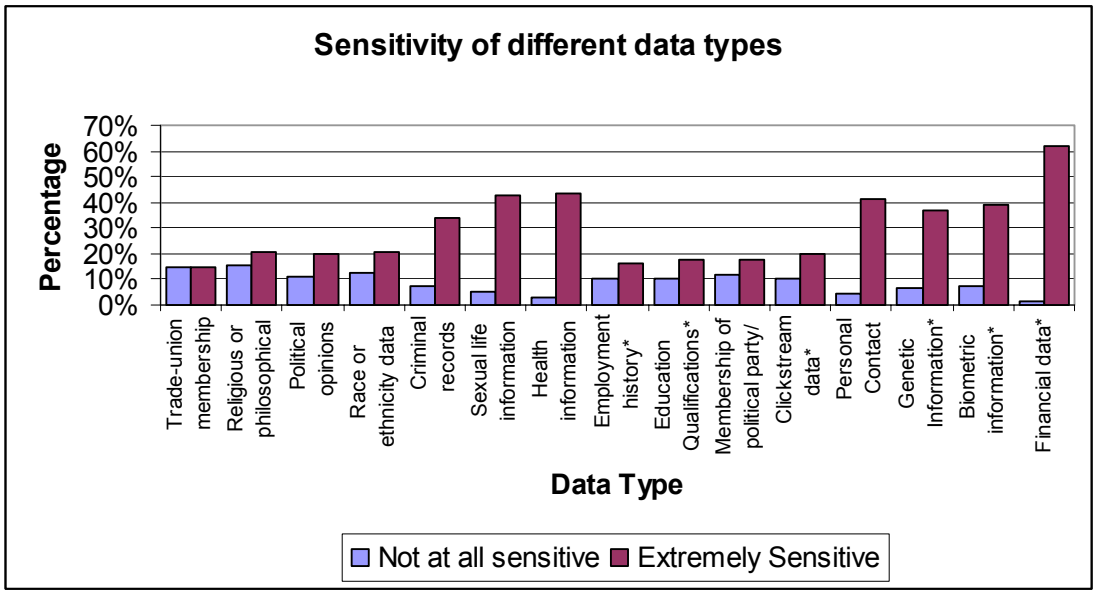
The views of UK citizens regarding the concept of sensitive data were sought through the ICO Annual Track (Individual survey 2006).[14] The question was designed to examine public perceptions of sensitive data. Firstly, it was used to test sensitivity ratings of seven categories of data which are currently 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 1.

Table 1: *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

[14] 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.[14] 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* - ICO survey



(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 not legally recognised categories were considered to be more sensitive than the legally recognised types of sensitive data mentioned in Art 8 of EU Directive 95/46/EC. For instance, financial data was considered extremely sensitive by most respondents (62.1%), whilst 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 2). 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 2: 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 3: *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%	26.8%	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 3 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 4: *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)

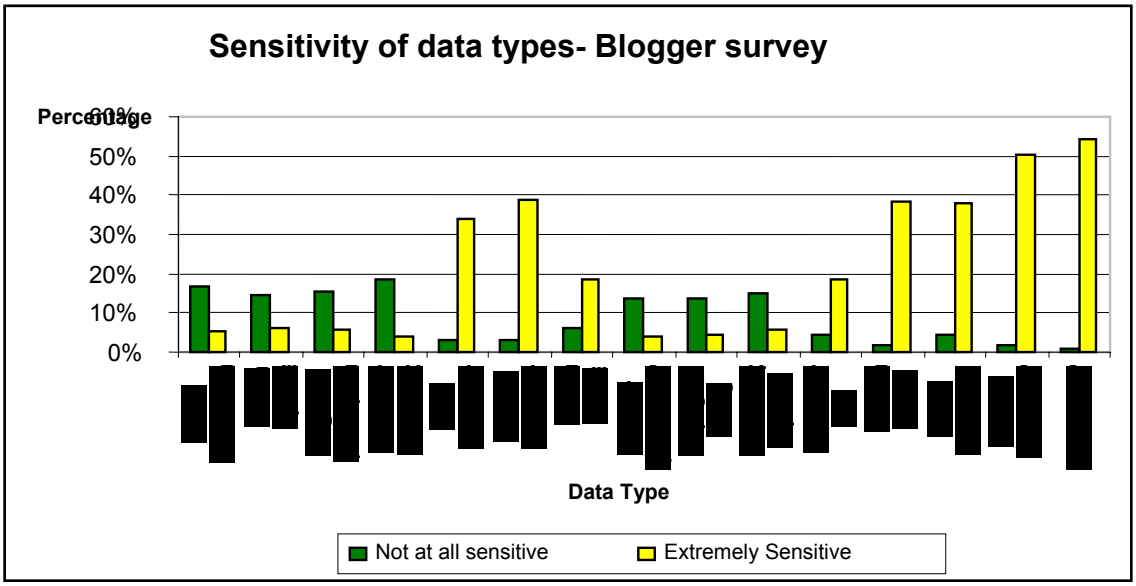
Table 4 displays categories of sensitive data that are not legally recognised. The table indicates that financial data was considered extremely sensitive by over 70% of respondents, and just under half (46.4%) considered their personal contact details extremely sensitive, whereas only 21.3% of respondents considered employment history data to be extremely

sensitive. This table indicates that some not legally recognised categories of sensitive data are considered more sensitive than others.

c) Findings from blog survey

The views of bloggers from around the world regarding the concept of sensitive data were sought through an online survey. The respondents to this survey 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 several 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 generalized to the entire blogging population. Rather, the findings are representative of certain niches of the educated, English-speaking blogging world. The question replicated the one on the Information Commissioner's survey. Out of the total number (1314) of responses received, 1258 were selected for data analysis; the remainder responses were incomplete and were disregarded.

Fig 2. *Sensitivity of different data types- Blogger survey*



(Source: Blog Survey 2006) (n=1258)

Fig 2 shows how blogger 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 not legally recognised categories were considered to be more sensitive than the legally recognised types of sensitive data mentioned in Art 8 of EU Directive 95/46/EC. For instance, financial data was considered extremely sensitive by most respondents (54.3%), whilst religious opinions were considered to be not at all sensitive by 15.3% of respondents. Likewise more than one half of respondents rated biometric data as extremely sensitive, while more than a third of respondents rated genetic and contact details as extremely sensitive, whereas fewer than one tenth 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 2). 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 5: *Sensitivity of legally recognised data types* – All Bloggers

	Trade-union membership	Religious or philosophical beliefs	Political opinions	Data concerning race or ethnic origin	Criminal records	Sexual life information	Health information
No Answer	12.7%	12.1%	12.2%	12.4%	12.6%	12.4%	12.6%
Not at All Sensitive	28.5%	24.1%	21.6%	23.0%	9.1%	4.1%	4.4%
A little Sensitive	16.7%	16.9%	16.8%	15.7%	9.1%	4.0%	4.1%
Sensitive	24.0%	22.6%	26.4%	23.4%	20.9%	9.9%	12.8%
Very Sensitive	11.8%	15.3%	14.5%	16.4%	22.7%	19.4%	20.6%
Extremely Sensitive	6.4%	9.1%	8.5%	9.1%	25.6%	50.3%	45.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

(Source: Blog Survey 2006) (n=1258)

Table 5 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 6: *Sensitivity of not legally recognised data types* - All Bloggers

	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
No Answer	12.3%	12.4%	12.4%	12.2%	12.2%	12.7%	12.9%	12.2%
Not at All Sensitive	22.3%	22.7%	22.8%	7.9%	3.2%	7.2%	3.1%	1.6%
A little Sensitive	16.1%	17.1%	16.5%	11.0%	4.5%	5.2%	4.1%	1.8%
Sensitive	26.2%	25.5%	25.6%	18.4%	10.6%	11.0%	6.9%	3.7%
Very Sensitive	15.5%	15.2%	14.2%	24.7%	22.0%	16.0%	12.9%	13.5%
Extremely Sensitive	7.6%	7.1%	8.4%	25.8%	47.5%	47.9%	60.2%	67.2%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

(Source: Blog Survey 2006) (n=1258)

Table 6 indicates that financial data was considered extremely sensitive by 67.2% of respondents, and just 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.

Of the 1258 blogger respondents 497 were from the UK. Direct comparison cannot be made between the ICO UK telephone respondents and UK blogger respondents, because the blog sample was not a random sample and is not necessarily representative. However, from an illustrative perspective it is interesting to examine their attitudes towards sensitivity of data types.

Table 7: *Sensitivity of legally recognised data types* – UK Bloggers

	Trade-union membership	Religious or philosophical beliefs	Political opinions	Data concerning race or ethnic origin	Criminal records	Sexual life information	Health information
No Answer	1.8%	1.2%	1.2%	1.4%	1.6%	1.6%	1.8%
Not at All Sensitive	28.4%	28.2%	22.1%	26.0%	9.3%	4.0%	5.2%
A little Sensitive	17.9%	18.1%	19.5%	18.1%	11.5%	4.4%	4.4%
Sensitive	30.0%	27.6%	31.0%	26.6%	20.7%	9.7%	13.9%
Very Sensitive	14.9%	14.1%	16.3%	18.3%	26.4%	24.3%	26.8%
Extremely Sensitive	7.0%	10.9%	9.9%	9.7%	30.6%	55.9%	47.9%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

(Source: Blog Survey 2006) (n=497)

Table 7 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 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 3).

Table 8: *Sensitivity of not legally recognised data types* – UK Bloggers

	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
No Answer	1.6%	1.6%	1.4%	1.2%	1.2%	2.0%	2.0%	1.4%
Not at All Sensitive	26.4%	25.6%	24.9%	9.5%	3.6%	7.6%	3.6%	1.6%
A little Sensitive	16.1%	18.1%	18.1%	13.1%	3.8%	6.2%	4.8%	2.8%
Sensitive	30.2%	28.8%	29.6%	17.5%	13.5%	12.5%	8.7%	3.6%
Very Sensitive	18.5%	17.9%	14.9%	28.8%	25.8%	21.5%	15.5%	15.3%
Extremely Sensitive	7.2%	8.0%	11.1%	30.0%	52.1%	50.1%	65.4%	75.3%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

(Source: Blog Survey 2006) (n=497)

Table 8 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.

Almost three times more ICO respondents (Table 4) rated employment history and education qualifications as extremely sensitive than UK bloggers, whilst 20% more UK bloggers rated biometric data extremely sensitive than ICO respondents. Direct comparisons between ICO survey and blog survey are not possible because the blog survey is not random, and has not been weighted for harmonisation purposes.

The findings from the surveys indicate 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. Similarly, over one quarter of blog respondents considered trade union membership, religious/philosophical beliefs or data concerning racial/ethnic origin to be not at all sensitive. These findings suggest that the current list is in need of reform, as 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. 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:

Korff (2002)[15] 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 e.g. Denmark and Iceland consider their lists as merely indicative. However, all laws provide ways and means to reopen the apparently definitely closed list. For instance, the Estonian act

[15] Korff, D. (2002) EC Study On Implementation Of Data Protection Directive (Study Contract ETD/2001/B5-3001/A/49)

states that the list can be modified by law, so *prima facie* the list of sensitive 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. 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.

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[16] 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)[17] 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[18] 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

> "Canadian law has a concept of sensitive data, but the approach taken differs from the EU concept of sensitive data. It can be used as an interpretative aid to explain why it is necessary to reveal certain information. Sensitivity depends upon 2 criteria: 1) the point of view of the data subject, 2) the context in which a 3rd party uses the data. Very little data has an absolute character. The Canadian definition can be adapted to reflect information that the data subject wishes to control." (Canada)

At this juncture it is appropriate to review alternative approaches.

4. Reform proposals: resolving the sensitivity conundrum

4.1 Context-based approach:

[16] 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
[17] Simitis, S. (1973) cited in Bygrave, L.A., (2002) "Data Protection Law: Approaching its Rationale, Logic and Limits," Kluwer, 132
[18] A credit score, credit report and credit reference agency.

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)

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)

It 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[19] 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 my *attitude* towards its use. I

[19] Wacks, R. (1989) Personal Information: Privacy and the Law, Oxford: Clarendon Press, 23, 181

am likely to have considerably different views about the purposes for which sensitive data is used, for instance

> "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[20] 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 questions, namely, who should decide what purpose is sensitive? It is submitted that this decision could be made by the Art 29 Data Protection Working Party using an objective approach[21] and the vast experience they have accumulated over the last decade to balance competing interests. An unresolved difficulty is how to decide whether the purpose for which the data is processed is 'sensitive'?

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 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 demarcating categories of data as sensitive and therefore deserving of 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[22] 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)[23]

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, it is opined that whilst it might be reasonable for an employer to conduct alcohol tests on employees for the purpose of ensuring work safety, it would not be reasonable for 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. Concluding remarks

It is suggested that the time is ripe to review the provisions of the directive. The current definitions reflect post World War II concerns regarding discrimination and protection of human dignity, but in the 21st century new concerns are arising due to technological

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

[21] That of the reasonable person

[22] 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.

[23] < http://www.oipc.ab.ca/pipa/>

developments. Indeed, the findings from interviews and 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. It is suggested that 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[24] 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. In concurrence with Simitis[25] it is suggested that sensitivity of data varies from context to context. This 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[26] 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. 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, lessons may be learned from reviewing the approach taken by other legislatures who advocate a 'reasonable' approach to data protection.

[24] 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
[25] Simitis, S. (1973) cited in Bygrave, L. (2002) "Data Protection Law: Approaching its Rationale, Logic and Limits," Kluwer, 132
[26] Wacks, R. (1989) Personal Information: Privacy and the Law, Oxford: Clarendon Press, 23, 181

Bibliography

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

Bygrave, L.A. (2003) Data protection law: approaching its rationale, logic and limits, The Hague: Kluwer Law International.

Denscombe, M. (1998) The Good Research Guide. Open University Press.

Korrf, D. (2002) EC Study On Implementation Of Data Protection Directive (Study Contract ETD/2001/B5-3001/A/49)

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) <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

Simitis, S. (1999) Revisiting sensitive data, <http://www.coe.int/T/E/Legal%5Faffairs/Legal%5Fco%2Doperation/Data%5Fprotection/Documents/Reports/W-Report%20Simitis.asp#TopOfPage> Last accessed February 2007

Wacks, R. (1989) Personal Information: Privacy and the Law, Oxford: Clarendon Press

Wong, R. (2007) “Data Protection Online: Alternative Approaches to Sensitive Data?” Journal of International Commercial Law and Technology, Vol 2, No 1