DECLARATION OF INTERESTS
Dr Kevin Macnish worked at GCHQ and the US Department of Defense for six years. He now works as a Teaching Fellow at the University of Leeds. His research interests are in philosophical ethics in the areas of surveillance, security and technology. Dr Macnish is the author of several articles on surveillance ethics and has worked as an ethics advisor to a number of European FP7 projects. This contribution is his own and is not submitted on behalf of any organisation.

WHAT ARE THE ETHICAL CONCERNS OF USING PERSONAL DATA?

EXECUTIVE SUMMARY

Ethical concerns regarding the collection and use of personal data include:
• The purpose of collecting the data
• Informed consent of those to whom the data pertains
• Security of the data
• Abuse of the data
• Function creep of the data

This statement recommends:
• Regulation of the collection, maintenance and end use of personal data
• Efforts be made to ensure social discrimination does not occur or is not exacerbated by the collection and use of personal data

TEXT

1. Ethical concerns regarding the collection and use of personal data can be summarised as issues in the purpose of collecting the data, informed consent (concerning the gathering of the data), security (concerning the maintenance of the data), the eventual use/abuse of the data, and the function creep of the data to ends other than those originally envisaged. There are also ethical concerns regarding not using personal data as opportunities for advancement of public health, for example, could be missed. Note also that there is ‘personal’ data which is about me but not exclusively mine (in the sense of owned by me) because inevitably in the public domain, e.g. how I look, how I walk, etc. There is other data pertaining to my person which is in principle private and owned by me. This submission concentrates on the latter.

2. The harms involved in the collection and use of personal data include privacy violations (with individual and social consequences), chilling effects on behaviour, an increased sense of vulnerability in society, and a decreased sense of trust in society. The above mentioned concerns contribute to the risk of these harms occurring.

Purpose

3. Personal information is collected in a variety of ways by a variety of entities for a number of reasons. Some of these reasons are ethically legitimate while others are not. Some serve the interests of a small group of entities, others serve the public interest. Some may serve both. Collecting data with the purpose of blackmail, for example, is unethical. Collecting data on a suspected terrorist with the purpose of apprehending the terrorist is more likely to be ethical. Collecting data on personal health records with the full knowledge and informed consent of those whose data it is for the sole purpose of improving public health is ethical.
4. Beyond interests, there are different ways of distinguishing legitimate from illegitimate reasons to collect data. If the subject of the data is liable (i.e. through involvement in a criminal act) then apprehending the subject is a legitimate reason to collect personal data on that subject. Also, if the subject gives informed consent then the data may be collected and used within the parameters set by that consent. In the absence of liability or consent, though, the collection of data is ethically problematic.

**Consent**

5. Informed consent is a standard requirement when an organisation seeks to make use of a person or their belongings in most contemporary settings. If you enter my house without consent it is trespass, if you take my property without my consent it is theft. Overriding or ignoring the requirement for consent involves harm, demonstrates a lack of respect, and requires justification.

6. Key aspects of an individual’s personal data are created and shared in many different contexts. Medical records, comments on social networks, and shopping preferences are a few examples. In each case, the data is shared by a particular person for a particular reason and with a particular understanding of the context and implications. If this data is then pooled and used for purposes other than those to which the participant has agreed, that data is being used without informed (or possibly any) consent.

7. It may be objected that, in some cases, consent has been given through the signing of a contract (e.g. to sign up for a shop loyalty card). However, it is widely known that these forms are rarely read in detail or understood. The print is small and the eventual use often obscure. Even if signing the form could be considered an act of consent, it is often not an act of informed consent. *Caveat emptor* is an unfair response if a large number are signing up for the service on a (widely recognized) limited understanding of the future use of the information they will provide.

8. A second objection is that a company owns the data and so can do with it as they please. This is weak. Owning something does not give the owner *carte blanche* to use it as they choose (ownership of medical records does not permit a GP to sell those records to the press). Possible harm must be taken into account, both on the part of those to whom the data pertains and of those who may be affected by the data.

9. It is also disputable that companies do own the data they receive from individuals. If I share an original idea with you, it is still “my” idea. This is widely recognized in professional ethics, academic practice and copyright law. While you now have access to my idea, you do not own it and are not justified in using it at will without reference to me or gaining my permission.

10. A third objection is that large public benefit can come from the use of data, e.g. medical records. However, if these benefits rely on unethically sourced data then they are problematic. A more ethical approach is to seek permission from patients in order to build a database to which those contributing have consented. This may take longer to compile, but it will meet the ethical concerns raised above.

11. Hence in the absence of liability or informed consent, the collection and use of personal data is ethically problematic. It should be noted, though, that neither liability nor consent gives *carte blanche* to the collection of relevant personal data. The data may not be held securely, it may be subject to abuse, or it may be used for purposes other than those initially intended.
Security

12. Once collected, the data must be held securely. Data can be made secure, but true security is extremely difficult to achieve. Any retention of big data must therefore be acknowledged to carry with it associated risks to those whose data is held. This does not mean that such retention cannot be justified. But the justification will involve a weighing of risks and benefits and it would be wrong to mislead the public with suggestions that the holding of big data can be made perfectly secure.

13. There have been a number of high-profile data security breaches in recent years. US company Target was hacked in December 2013 leading to the loss of data pertaining to 40m credit cards, and up to a further 70m phone numbers and zip codes of customers. High end retailer Neiman Marcus, a number of well-known hotel chains, and software company Adobe were also hacked in 2013, leading to the loss of millions more credit card details.

14. These companies had made a considerable investment in security, and yet still fell victim to hackers. In the case of Target the breach was down to human error. As such, any security system will only be as good as the people who use it. This is a truth long recognized in intelligence, where otherwise securely maintained Top Secret documents have been left on trains, or leaked to the press.

15. One method of securing data involves anonymising the data: removing aspects of the data with which the originating source can be identified. Satisfactory procedures for anonymisation are particularly difficult to achieve with big data. “Anonymised” data has in the past been released to the public only to be identified within days. This is not surprising when the amount of data provided about an anonymised person reduces the number of people to whom the data could refer. Even when data within a system has been anonymised, it may be combined with information outside the system which allows for the identity of the individual in question to be discovered.

Abuse

16. Once legitimate uses and collection of data have been established further concerns remain. People are wont to use databases for personal ends. It has recently transpired that several hundred US intelligence analysts have inappropriately accessed national security databases to review information on current or former love interests. Similar revelations occur regarding the UK police on a not-infrequent basis. Use of, as well as access to, databases needs to be controlled and monitored to limit abuses.

17. Secondly, large data sets can be used on “fishing trips” to locate particular information. Hacking voicemail accounts is one (illegal) example. There are also legal examples. For instance, statistical correlations have been demonstrated between freely available information and “private” information (e.g. homosexuality). This suggests that fishing trips could uncover legitimate but embarrassing information which could harm an individual or group.

18. Thirdly, companies using large data sets generally do so for economic advantage such as targeting advertising, sales, and offers to certain societal groups. This often leads to the neglect of other societal groups, typically those at the lower end of the socio-economic scale. This “social sorting” may be unjust if it exacerbates divisions in society. This may happen if reduced prices are offered to regular customers while those who cannot afford to shop at an establishment on a regular basis end up paying more for the same items.
Function Creep

19. Data may be used for the purpose for which it was initially collected. It may, though, also be used for other purposes. Medical data, for example, may be collected for the benefit of serving the individual. It could also be used, though, for the benefit of public health. A further use could be to aid insurance companies in determining premiums on an individual or group membership basis. This is known as function creep. Further uses of data are not necessarily wrong. However, if informed consent has not been given to these uses and the subjects are not liable to be subjected to surveillance, then these further uses are not ethically legitimate. New consent must be sought, or grounds for liability established.

20. It may be objected that the intention behind the collection of the data is pure. While the collection avoids concerns outlined above this is good. However, vigilance should be paid to ensure that the data is not subject to function creep if the collected data appears to have benefits for other entities or other purposes than those initially envisaged.

CONCLUSION

21. The collection and use of personal data holds tremendous advantages for society, not least in the areas of public health, security, and the economy. This statement has focussed on concerns rather than benefits, though. Those advocating the use of personal data usually stress the benefits at the expense of the risks, and so this statement has sought to stress the latter.

22. It is both unrealistic and short-sighted to suggest that personal data not be collected. The collection and use of personal data is broadly accepted for the benefits that can ensue. However, the risks associated with those benefits must not be underplayed. The harms to individuals and society are significant, and the possibility of those harms occurring uncertain. Security is never perfect: hackers, abuses, leaks and loss of information occur. Data is abused by those to whom it is entrusted. Function creep occurs as data moves across boundaries. It is essential that the reality and likelihood of these risks is recognized.

RECOMMENDATIONS

23. The collection of personal data should be closely regulated. Informed consent by those to whom the data pertains should be sought as a matter of course in all instances of collection (except those, such as law enforcement, in which the person or people in question are liable to be subject to surveillance). Collection of data without informed consent should be heavily penalised. The communication of purposes for and reassurance of the security of the data should be clear in all cases to ensure to the greatest extent possible that consent is informed.

24. The maintenance of personal data on databases should also be closely regulated. The more intimate the data, the greater the security should be. Penalties should be established for the hacking, abuse, leaking and loss of personal data.

25. Finally, the use of personal data should be closely regulated. Penalties should be established for the transfer or use of data which occurs without the explicit and informed consent of the person to whom the data pertains. Efforts should be made to ensure that harmful social sorting does not occur or at the very least is not exacerbated through the use of personal data.
Written evidence submitted by Dr Kevin Macnish (SMD0001)

March 2014