



The Royal Academy  
of Engineering

## Data Sharing Review

Richard Thomas and Dr Mark Walport



February 2008

## Question 1

**Please explain what your interest in information sharing is.**

The Royal Academy of Engineering produced a report in March 2007 entitled 'Dilemmas of Privacy and Surveillance'. The report investigated recent and likely future developments in technologies for the collection, storage and sharing of electronic personal information and the societal impacts of those technologies. The responses to the questions below are taken from the conclusions of that report and the considerations of the working group that developed the report. The Royal Academy of Engineering continues to have an interest in the use and impact of technologies for data collection and the responses below are from the point of view of this interest.

### **Section 2: Scope of personal information sharing, including benefits, barriers and risks of data sharing and data protection**

## Question 2

**What in your view are the key benefits of sharing personal information to a) individuals and b) society? Please provide examples.**

a) The key benefits to individuals of data sharing lie in the greater convenience and easier access to services (both public and private) that data sharing can offer. Not having to give the same information to different bodies in the pursuit of a single issue would be of great benefit to individuals if it can be achieved without significant risks, and many people are likely to be content for their data to be shared in order to make administrative tasks easier. There are also great benefits promised by sharing medical data. If a person's medical records are available to any medic treating them, it is more likely that their condition will be accurately diagnosed and that they will receive appropriate treatment; and it is less likely that they will receive treatments that could be dangerous – eg, drugs to which they are allergic or which interact with medication that they take. In the private sector, data sharing will allow more services to be tailored to the individual with people being offered only those options that are suitable or of interest to them rather than blanket advertising of goods.

b) Perhaps the most persuasive examples of the benefit to society of information sharing revolve around the sharing of information about criminals. The results of the Bichard inquiry show that better use of information might have prevented the Soham murders (though it must be kept in mind that the barriers to information sharing in this case were not legislative but lay in poor organization and that Ian Huntley, the perpetrator, had never previously been convicted of a crime). Other criminal behaviour might also be obstructed by greater data sharing, including between public and private sectors. Benefit fraud, for example, involving lies about personal circumstances, may be hampered if these personal circumstances can easily be checked against data held by other organisations.

In general, collecting and sharing comprehensive data can help to better shape public services, by providing a basis for predicting demand. Collecting and sharing data on income and state benefits; educational achievements and social background; immigration statistics; epidemiological patterns; age of residents in a locality and similar issues can support the fairer and more effective provision of public services.

### Question 3

**What in your view are the key risks of sharing personal information to a) individuals and b) society? Please provide examples.**

a) Most generally, the loss of privacy is the greatest risk that is posed to individuals by increased data sharing. People value privacy in the form of anonymity – the ability to engage in activities without anyone knowing who they are; privacy as confidentiality – the ability to keep certain facts about oneself concealed from selected others; privacy as control of one's personal data – ie, the ability to determine what information is revealed to whom and under which conditions. As data is more routinely shared the possibility of these forms of privacy diminishes. Privacy is recognized as a right in European and UK law and the loss of one's personal rights is of great detriment to any individual.

Many categories of individuals have particular threats to fear from loss of control over their personal data. People may have good reasons to conceal their age, HIV status, addictions, mental illness, religion, politics, past traumas (e.g. rape), race/ethnicity, previous gender, sexual orientation, disability, employer, criminal record, previous identities, or address; for reasons for example of discrimination, escaping abusive relationships, witness protection, avoiding ID fraud, concealing pre-take-over company investigations, protecting celebrities, or statutory requirements (e.g. protecting the identity of children in court cases or following adoption).

Another significant threat to individuals is the loss of opportunity or prejudice that might arise through greater data sharing. For example, if companies that manage supermarket loyalty schemes share the data generated by a person's use of the card with insurance companies, individuals may find that their premiums are affected by the insurance company's knowledge that they spend money on alcohol or junk food. Data about individuals is often 'profiled' to assign people to specific groups, and such assignments are used to make automated decisions about an individual – for example, if they apply for credit or insurance. The more data that can be so profiled, the more limiting the group to which that individual is assigned might be. This can mean a restriction in the services offered to and received by individuals and it is often difficult for an individual to challenge decisions based on automated processes (this is the now-familiar 'computer says no' scenario).

Finally, individuals are at risk if data is shared more extensively because of the threat that the data will not properly be managed. One organization may keep data up-to-date and relevant but organizations that they have previously shared that data with might not be so diligent. The more bodies, public and private, amongst which data are shared, the more difficult it is to ensure the quality of that data over time. The very process of passing data from one organization to another also puts data at risk of loss, theft or misuse. This was shown by HMRC's attempts to share the data that it held with the National Audit Office.

b) The risks to society of greater sharing of information are the same as those for individuals. Again, the main concern is the effect of a real or perceived loss of privacy. This may lead to a loss of trust, as society loses faith in organizations to support their rights to the forms of privacy listed above. Failures to properly protect data and to share it securely will also certainly lead to a loss of trust, again as evidenced by the loss of data by HMRC and the DVLA. This is problematic because loss of trust poses a threat to the functional aspects of data collection as it can lead to a lack of cooperation with the bodies that collect data.

#### Question 4

**As mentioned in the introduction, there are wide variations in the scope and methods of personal information sharing. What scope and what methods, in your view, pose the greatest opportunities or risks? Please explain the reasoning behind your response.**

One of the least secure methods for sharing information would be to simply hold all information on a single database accessible to a wide range of organizations. If a large amount of data is held on one database that becomes a honeypot – a target for information fraud – and the whole database is at risk. A better method is to store only the requisite data on carefully managed databases so that data can be kept up to date and is not so vulnerable to attack. Data should also always be strongly encrypted and should be moved between organizations only in its encrypted form.

There should be a mandatory requirement that only the minimum subset of data required for the purposes at hand is shared between organizations (this is precisely what did *not* happen when HMRC shared their data on child benefits with the National Audit Office – despite the latter's request). The data held on a subject should be organized into distinct, separable fields so that only the relevant data fields are shared.

It is also important that individuals have the right and the power to withhold some or all of their personal data from being shared; that automatic audit trails are created whenever data is shared and accessed; and that these audit trails are readily available to the data subject, who can easily challenge the legitimacy of any access and obtain compensation if the data has been improperly accessed.. In particular, it is important to be able to have a clear record of when data has been changed and by whom, to guard against mistaken or malicious alterations in data held. Without these safeguards, there is little motivation for data controllers to treat personal data appropriately particularly as the cost of handling highly confidential data appropriately is much higher than for non-confidential data.

#### Question 5

**Please provide examples of where, in your view, the public authorities hold too much data or not enough personal information, and the reasoning behind your response.**

It is very difficult to judge, as there is no central record of what data is held, by whom, for what purposes, and how widely it is shared. To find this information, an individual needs to write to each organisation and will usually need to pay a fee. That is a very effective deterrent and a barrier to transparency. This is problematic because not knowing what data is held about oneself or by whom is a further threat to trust.

#### Question 6

**Please provide examples of where, in your view, private sector organisations hold too much personal information or not enough personal information, and the reasoning behind your response.**

Companies routinely collect large amounts of personal data about their customers. For example, mail order or internet-based companies seem routinely to hold the details of customers, including their purchase history and the credit card details they

have supplied. It is rare that the customer is informed about this and given the option to withhold permission, and it is rare that the customer is informed if their data is mislaid or lost. The organisations may make decisions about the customers on the basis of this data, without customers' knowledge, and the problems of profiling discussed under question 3 apply as a result.

Data processing may be distributed between many computers and databases, between different companies in a group, and outsourced to suppliers. In this context, it is simply too difficult for a data subject to know who is holding personal data about them, and where they should apply to see it. The aggregate fees payable may be very high, and there is no way to tell whether or not the supplied data is complete.

The following restrictions on private sector organizations' handling of personal data are recommended, with powers given to the Information Commissioner to ensure that they are respected (including sanctions for failure):

- 1) Companies should specify how long they will retain personal data.
- 2) Restrictions should be placed on the length of time companies can retain data for, and they should be required to renew permission from data subjects after that period has elapsed.
- 3) Companies should make it possible for data subjects to request complete removal of their details from databases and should give clear instructions on how to do this.

#### **Question 7**

**Please provide examples of cases where you believe the sharing of personal information between two or more bodies would be beneficial, but where it is not currently taking place.**

**Please explain as fully as possible why information is not being shared, detailing what the barriers to the sharing of personal information are – e.g. legal, cultural, financial, institutional – and how these barriers can be overcome.**

No specific examples to contribute.

#### **Question 8**

**Please provide examples of cases where you believe that personal information is being shared between two or more bodies, but where this should not be taking place.**

**Please describe the information-sharing concerned and why you believe it should not be taking place, including the risks involved in such information-sharing.**

No specific examples to contribute – again, this is an example of how lack of transparency means that individuals rarely know who shares data with whom. However it is felt that there should be recognition of the fact that personal data is the property of the data subject and that therefore non-essential data-sharing should be

on the basis of permission by the data subject. In the commercial sector individuals have the opportunity to express a preference as to whether data is shared with other companies for marketing purposes; this right to choice should extend also to the public sector.

### **Section 3: The legal framework**

**The Data Protection Act (DPA) regulates the processing of information, including its obtaining, holding, use and disclosure.**

**The second principle of the DPA is as follows:**

**“Personal data shall be obtained only for one or more specified and lawful purposes and shall not be further processed in any manner incompatible with that purpose or those purposes.”**

#### **Question 9**

**In your view, how well does the DPA work? Please outline the DPA’s main strengths and weaknesses and any proposals for changes you would like to see made, including suggestions for their implementation.**

The DPA ‘s strength is that it provides a framework within which some attention must be paid to data protection. The DPA needs to be strengthened in the following ways:

- it should require that the data processed should be the *minimum* that is needed to fulfill a particular transaction, and that it should be retained for the *minimum* practical period unless the data subject has explicitly opted in to further processing;
- it should require that data losses are disclosed to data subjects promptly, with advice on whether action is needed on the part of the subject and what that action is;
- it should provide a mechanism whereby compliance with the Act is routinely audited (for example, as part of the statutory audit of public companies);
- It should give the Information Commissioner the power, with proper notice, to enter premises to check compliance;
- it should provide a framework for assessing the appropriate security measures that should be applied to maintain the confidentiality of classes, or quantities, of personal data;
- it should require rapid, straightforward compensation for data subjects whose data is processed unlawfully;

The DPA was drafted before large-scale distributed processing of personal data, including outsourcing, became commonplace. The Act should be reconsidered in the light of technological changes that have taken place since 1998 to establish whether the Act still applies given the ways that data are currently processed. For example,

the development of image recognition technology used in conjunction with CCTV may affect whether the DPA still adequately covers the processing of CCTV images.

#### **Question 10**

**In your view, how well do public authorities and private organisations adhere to the second principle of the DPA? How valuable do you believe the second principle is? Please provide examples and the reasoning behind your response.**

The second principle of the DPA is valuable and should be adhered to. It allows individuals privacy in the form of a right to control who can access their personal information and what it is used for. Without this principle and its proper enforcement individuals' data, once obtained, can be used for purposes that the data subject might never have assented to.

However, 'lawful processing' has to be defined carefully. The purposes that a data controller declares may be very broad (e.g. "marketing" or "to provide a more personal service") so currently the DPA is not effective in preventing processing of data beyond the informed consent given by the data subject. Clearer limits need to be set on what counts as lawful processing, again taking into consideration whether developments in technology have affected the definition of lawful processing since the DPA became law in 1998.

#### **Question 11**

**What technical, institutional or societal barriers stand in the way of the effectiveness of the DPA? Please provide examples.**

Greater weight is placed on efficiency or convenience than on confidentiality, across most public and private organisations. As long as this attitude prevails it is a barrier to the effectiveness of the DPA.

#### **Question 12**

**What further powers, safeguards, sanctions or provisions do you believe should be included in the DPA.**

One of the most useful changes to the DPA would be to allow the Information Commissioner's Office powers of audit, so that the ICO could perform spot checks to ascertain whether organizations process data securely and in accordance with the DPA. This has been considered for government departments in the wake of recent data losses but should extend to all organizations that process personal data. In addition, company auditors should be required to audit (and report on) the strength of a company's data controls.

There should be a statutory duty to inform data subjects of any loss of their personal data, and statutory compensation. If compensation has to be paid to data subjects for every breach of the DPA, even if the sum is small, this will be a significant deterrent to careless handling of data – especially in cases where such carelessness might result in compromising the data of a very large group of individuals (as in the recent HMRC and DVLA cases).

### Question 13

**Are there any other aspects of UK or EU law (such as EU Directive 95/46/EC) that impact positively or negatively on data sharing or data protection? Please provide examples.**

No examples to contribute

### Question 14

**Are there any statutory powers unavailable that would enable better and more secure sharing of personal information– for example for identity authentication purposes – between a) public authorities and b) public authorities and private organisations? If so, what are they?**

**Please provide examples and any steps you believe could be taken to improve matters.**

See the answers to questions 9 and 13.

### Question 15

**Are there any parts of the legal framework that place an unreasonable burden on business? Please provide examples.**

**Please outline your proposals for streamlining the legislation to ensure that such burdens are minimised.**

In view of our comments under question 10 regarding the broadness of the purposes for processing data that are registered by the data controller, it should be possible to dispense with registration altogether – provided that the DPA is strengthened in the other ways that we have recommended.

## **Section 4: Consent and transparency**

### Question 16

**Is it clear whether and when you need individuals' consent to share information about them? Are you clear about the form that consent should take? Please provide examples.**

**Please provide details of any initiative you have been involved in that has been based on consent.**

See comments relating to renewal of consent in answer to question 6. The Academy has no further comments on this section of the report.

## **Section 5: Technology**

### Question 20

**What impact in your view have technological advances had on the sharing and protection of personal information? Please provide examples.**

Technological advances have in general made it easier to share information than to protect it. In The Royal Academy of Engineering report 'Dilemmas of Privacy and Surveillance' a distinction was made between connection technologies, which allow the sharing, exchange and networking of data, and disconnection technologies, which provide access controls to maintain the security of data. Connection technologies are easier to implement than disconnection technologies – for example it is easy to network computers to share the information on them but relatively difficult to partition the data on a computer so that it can only be accessed by certain users.

Encryption is an important disconnection technology, and there is an “arms race” between those who seek to break encryption and those who seek to strengthen it. The former are greatly helped by Moores' Law (which predicts the doubling of processing power every two years), whereas the latter are hindered by export control laws such as EC Regulation 1334/2000 which (in combination with the general export licences) prohibits the export of cryptographic software with keys longer than 56 bits. 56-bit DES can now be broken in hours, so current best practice is 128-bit or 256-bit AES, which makes a growing number of products that contain cryptographic data protection subject to export licensing.

**Question 21**

**Should the law mandate specific technical safeguards for protecting personal information?**

**For example, should there be an explicit requirement that all personal information held on portable devices be encrypted to a particular standard?**

All sensitive personal information held on portable devices such as laptops or other media should be encrypted. This will mean that if the devices or media are lost that data will not be so easily compromised. This should be made a legal requirement for public and private sector organizations handling personal data and the minimum standard of encryption should be specified (for example, AES with 128-bit keys). However, encryption codes can be broken, particularly if the keys or pass-phrases are weak or become compromised, so the use of encryption should not be relied on totally. Other restrictions should be put in place so that large amounts of valuable data cannot be held on one laptop or disc that can be taken out of secure offices. We would like to see clear guidance on the technical and other security measures that should be in place for particular classes of personal data and particular sizes of data aggregates (such as the personal details of thousands, hundreds of thousands, or millions of data subjects). The Academy would be pleased to help draft such guidance.

## Question 22

**How, in your view, could ‘privacy enhancing techniques’, such as the anonymisation or pseudonymisation of personal information, help safeguard personal privacy, whilst facilitating activities such as performing medical research?**

**Is sufficient advice about the deployment of such techniques available? Are you confident about using them? What are the barriers to using them?**

Privacy enhancing technologies should be promoted and used within all organizations, particularly in the public sector. Pseudonymisation and anonymisation techniques should be used wherever possible to make the likelihood of opportunist theft or use of personal data significantly less likely. However, it should be recognized that such technologies have limitations. For example, data used in medical research is difficult to anonymise, since it will often be possible to trace data which is sufficiently specific for the purposes of medical research to a particular individual.

It should be possible (even encouraged) for individuals to use pseudonyms when providing personal data, wherever there is not an overriding public interest reason for them to be identified.

Wherever *authentication* (of someone’s right to do something – such as enter a building or use a train) rather than *identification* (of their actual identity) is required, then identification should be seen as an extension of the minimum data processing requirement, and be unlawful without explicit and informed consent. (It would need to be unlawful to discriminate in the provision of goods and services against anyone withholding such consent, otherwise the protection would be routinely withdrawn through the use of standard terms and conditions).

**The Academy has no further comments to make on the remaining sections of the consultation.**

---