

Response to the Data Sharing Review carried out by Richard Thomas and Dr Mark Walport

Use and Sharing of personal information in the public and private sectors.

Introduction

This response is from Professor Brian Collins in a personal capacity. He holds two official positions, the Chief Scientific Adviser at the Department for Transport and Professor of Information Systems at Cranfield University. He has seen and commented on the response that the Department for Transport has submitted to the consultation paper and senior officials at the Department for Transport have seen and commented on this response. However this response is essentially his own personal views and should not be taken as representative of those of the Department for Transport. The commentary is organised as requested against the structure of the questionnaire.

Section 1: Background

Question 1: My interest in information sharing started when I held the position of technical director of GCHQ when the issues of control of information sharing in an extremely sensitive part of UK government needed to be managed not only within the UK government itself but between the UK government and a very selected group of allies. Indeed within those allied nations there were very selected organisations which were selected for sharing of information, but in few cases was personal information involved in the sharing process at that time.

Since then my interests have broadened to a wider range of subjects to do with information management, personal and public information sharing, identity management, consent, privacy and security and it is in the context of those interests that my comments below should be considered.

I do not have any active involvement in personal information sharing except as a working member in the organisations mentioned above. The information I deal with is essentially professional personal information, names and professional addresses with appropriate contact details such as email addresses and telephone numbers.

Section 2: Scope of personal information sharing including benefits barriers and risk of data sharing.

Question 2: The key benefit to individuals of personal information being shared is essentially that more value to the individual can be delivered by service providers the more information they have at their disposal. This is in either a personal way for them as individuals in society or as members of anonymised groups which are categorised against some criteria such as where they live, what professional category they are classified as being in, what their age is etc, etc. The benefits accrue from the *quality* of the services provided to them as individuals. The differentiation between the public and private sector with regards to those services and the underlying processes I think is marginal in that most of the services that are provided by both public and private sectors are essentially tailored towards them as individuals. Whilst the

legalities of the differences might be significant the practicalities and the benefits are very similar.

The benefit to society that comes from sharing of data is to make the services more efficient and better quality and hence consume less resource and achieve greater benefits. It will also provide better information infrastructure for the delivery of information based services whether commercial or social. The barriers to sharing in both cases are essentially to do with trust. This is because people limit their trusting behaviour because they either do not trust the system or services that are provided to them or because the trustworthiness of the implementation is either in fact or perceived to be inadequate for the purpose. Data protection legislation in its current form would seem to provide some measure of safeguard based on legislation and regulation for limiting errors and criminal activity. However the balancing activity of providing educational services and improved awareness of government services and structures would seem to be less mature and should be in my view attract considerable attention and investment in the light of recent events.

Question 3: The key risks of sharing information are that the information is lost, abused, misused or at worst destroyed. These risks are managed by both the individual who is giving consent for their information to be shared and by the organisations between which the information is shared. There are some subtle issues which arise in this space which are to do with how information is labelled in such a way that in the future it is governed and managed according to a set of rules. Also it is vital to understand in what way information is aggregated with information of a similar categorisation for either statistical purposes or to plot trends for individuals. The risk management of information is a well developed subject; however that body of knowledge is only now becoming enshrined in good business practice, particularly in the areas of workflow database management and online services delivery. It is therefore my view that the risks at the moment are going up with regards to the sharing of personal information both for individuals and for society and the potential economic and social benefits that could be realised by better sharing information are not being achieved due to the lack of trust and inadequate service design and provision.

Question 4: The concept of standardisation of methods of personal information is attractive in that it would potentially reduce the risks of sharing and would possibly provide fewer opportunities for illegal activities such as identity theft and fraud. However the wide variety of contexts in which personal information is shared and used make it difficult to conceive of how a standard way of *carrying out* these processes could be articulated. However *principles of information sharing* could well be articulated. Indeed in certain standards and guidelines, such as the ISO 27000 family, the ITIL codes of practice for service delivery and the BERR publications on Information Security Management and Information Asset Management, attempts are being made to improve the quality of the principles from which business practices are derived.

Question 5: Public authorities need to hold personal information in order to provide appropriate levels of service. It is my view that there is relatively little guidance of what is the appropriate and in some cases minimum amount of data about people that should be held for the service. That is because the service structures and mechanisms

are insufficiently well articulated for the data architectures that are used within them to be described appropriately. There is a body of academic literature on data architecture and on service quality which could be exploited in these circumstances but to my knowledge it is only used rarely. It is therefore my view that greater coupling between the academic expertise that exists in this country, which it is worthy of note has been partially funded by the UK Government, to service designers and operators would be beneficial.

Question 6: In essence my answer to question 5 also answers question 6 in that whilst that I do not believe private sectors are any worse or any better I have no evidence to suggest that they are more rigorous in minimising the amount of information held and indeed may feel less inclined to consider this in that they unlike the government have no restrictions for minimising data holdings except regulatory and commercial ones.

Question 7: The sharing of information between two or more bodies in order to achieve benefit where it is not currently taking place suggests that we could identify a number of business processes where information about people could be used to good effect. Some of these of course would be hypothetical and would not only have potential benefits for the individual as perceived by the delivery organisation but might have dis-benefits from the point of the individuals with regards to their privacy being further eroded and the possibility of major identity theft being increased. The barriers to such personal information sharing I believe at the moment to be largely legal and financial in that organisations do not share information where it is not deemed necessary and even if it is possible they do not share because they consider that the possible risks are too high and that the benefits maybe too low. There are some really good case studies in the public sector of information that might be regarded as personal being shared to good effect. One is in Transport, the vehicle excise duty payment systems. This connects people's details to insurance companies and to the vehicle's MOT record. The process contains the name of the registered owner of the vehicle as well as information about the vehicle. The system provides benefits to the individual and seems to be acceptable. So the barriers I believe to the sharing of personal information are not only the four mentioned in the question that is legal, cultural, financial and institutional. What must be also considered is the perception that unless services that demand shared information that overcome these barriers is actually required, innovating in this space is seen as too risky. So the barrier I believe is not so much the *processes* of sharing in any of the categories identified more it is the *perceptions* of risk by all parties that will come from actually attempting to do so.

Question 8: The sharing of information where it should not be taking place is even more problematic. The possibility that information is being shared inadvertently rather than deliberately of course is a possibility and one can take that to a further extreme, that data is mislaid or lost but is inadvertently shared between two or more bodies where it certainly shouldn't be. So my conclusion to this question is that I am unaware of where personal information is being shared two or more bodies *deliberately* for services that are inappropriate but I am aware of course that personal information is mislaid or lost through *human error or bad systems design*. This raises the interesting question of how systems and processes are designed and operated in such a way that such losses occur with much lower probability than currently is the case. I believe this to be a very fundamental issue which I will come back to later.

Section 3 The legal framework

This section of the questionnaire is one which I will comment on less than others in that I am not a legal expert but will nevertheless I will comment from a technological and systems engineering point of view as well as a business governance view point where appropriate.

Question 9: The DPA would appear to be working reasonably well in that it provides a framework within which personal data is managed. However it would appear that the guidance for how to use and interpret the DPA is not widely understood. Making the consequences of compliance with the DPA for an organisation more approachable will seem to be a useful step forward and ensuring that individuals within organisations understand their roles and responsibilities would also seem to be important. The training and education of people I believe to be absolutely core to achieving an enhancement in our current position of compliance with the DPA. All of these issues are ones of governance.

Question 10: The second principle of the DPA as outlined in the questionnaire implies that there is a mechanism by which the purpose to which any piece of personal data can be put is identified and this is bound to the data such that it is clear what the purpose was when the data was collected and that can be compared with the purpose to which it is potentially being put. This process should be complied with when any given purpose is thought about for the use of some personal data. It would then seem that we do not categorise the purpose of personal data accurately enough when it is collected and furthermore we do not either put either procedural or system checks in place to ensure compliance with the second principle such that both are visible. Were personal information to be categorised in such a way that the purpose was part of the categorisation then it would become clear whether any given purpose to which a piece of personal data was about to be put was compatible with the purpose for which it was originally collected. The second principle as outlined suggests "incompatible with" and I wonder whether that allows some loopholes in interpretation to creep in. Without a categorisation which is generally agreed "compatible with" can be interpreted loosely enough that almost any purpose could be instantiated even though it was not the one for which the data was originally collected.

Question 11: I will only comment on the technical barriers with regards to the effectiveness of the DPA. It does seem to me that some measure of codification of the DPA in certain circumstances would enhance compliance with it. This would allow systems rules, workflow rules and storage rules to be better implemented within organisations so that compliance with the DPA was seen to be in place and could be audited regularly both in real time and periodically to ensure that was the case. This would make the DPA more effective. In order to do that both institutional and societal if not cultural factors would need to change significantly.

Question 12: The further powers that the Information Commissioner with respect to the DPA might be invested with in order to achieve greater effectiveness seem to me to be the more in the area of education and training and audit, both real time and periodic, rather than tighten up the legislation in such a way that penalties for non

compliance were seen to be greater. My perception is that organisations respond better to persuasion and to better and more regulation where appropriate than to the threat of legislation. Therefore the DPA itself may need little change whereas the improvement of its context in so far as it is in part a framework which provides guidance and regulation might make it much more effective.

Question 13: I do not feel that I can comment on the legal aspects of either UK or European Union law with regards to data sharing or data protection.

Question 14: The issues of identity authentication is core to the sharing of the personal information in that without proper rigorous identity authentication operation of consent mechanisms would seem to be very fragile. So consent in my view is the major factor in providing better and more secure personal information. Making sure that consent mechanisms are transparent to all people and that they are secure is vital. The means by which they can be overridden should be under appropriate oversight with appropriate secure audit so that after the event it can be very clear how and where overrides of personal consent or limiting personal consent have been obtained. Therefore the development of personal identity management service which provide those authentication mechanisms is an absolutely core piece of infrastructure and services should be developed rapidly in order that personal information can be appropriately protected.

In some ways some elements of the national identity scheme are in the category of personal identity management services but have been confused with other major purposes for immigration, access to other public services and for law enforcement purposes whereas a personal identification service that supports authentication and supports rights for organisations to share to personal information would seem to me to be equally important and valuable and could draw on same technological and systems base. These much less politically charged as a concept.

Question 15: I do not feel that I can qualify to comment on the legal framework.

Section 4: Consent and Transparency

Question 16: It is not clear at all whether and when you need individuals consent to share information. This is a particularly cloudy area I believe for the future of the DPA and for the use and exploitation of personal information in both the public and private sectors. The confusion is the basis for how effective phishing attacks by organisations by a whole range of mechanisms have now become widespread. People need to be educated in how to look after their own privacy and their own information in such a way that consent is only given for the information that is absolutely needed for the service they are attempting to use and that they know when they share that information that it will only be used for that purpose and it will not be stored in the way in which compromises their privacy for any future use and that they will continue to have access to that information where it is needed to be changed. There are a number of initiatives that are current in this area. In particular there is the development of research programmes (EPAC) by the technology strategy board. I have been involved as a mentor to the development of those programmes and may possibly be involved as an adviser to one of the proposals were it to be funded and successful. This initiative for improving and understanding of how consent fits into

the social fabric and the technological architecture is critical to the success of sharing of information in a transparent way in both public and private sectors and in my view should be given considerable visibility when the results start to appear.

Question 17: Gaining consent would appear to be the main barrier to improving the situation for sharing of personal information and it is usually commented that without very rapid consent being available the value of the service that comes from sharing of information is thoroughly degraded. However modern technology and techniques allow consent to be given much more rapidly than some people realise and I believe this to be an area where further research would show that the requirement for gaining consent is fundamental to assessing the success of services. People, individuals and organisations could then exploit modern technology in such a way that consent can be given in a timely way and be combined with the trustworthiness that all parties would need to achieve reliable services.

Question 18: Making the sharing of information more transparent is achieved by allowing individuals access to not only their own information in the place that it is stored but by anybody to whom they have given consent for storage under appropriate authentication controls. They also have rights to modify certain elements of that data, such as address, marital status etc. Other fields which are essentially part of their identity authentication should not be changed and indeed should be stored in such a way that if anyone attempts to change them alerts and alarms are set off. Organisations clearly can explain the use and sharing of personal information to all stakeholders both public and others by online notices, declarations by other forms of awareness, publication and by having appropriate online help menus. There are some subtleties to do with transparency however which come from people having access to a number of places in which information about them is held. One such is the aggregation of those data fields which inadvertently might occur in a machine back up system which contains all the searches on sets of personal information which in turn might originally have been held on separate sources. The process of reviewing them may cause them to all end up in one place and make them more vulnerable than they might otherwise have been. Ensuring that individuals understand how systems which are designed to be reliable and resilient against failure also produce vulnerabilities of this nature needs to be achieved so that they can delete those aggregated file sets or understand that they have to be very careful in the use of the machines which have been used for that purpose.

Question 19: The development of a code of practice by the information commissioner for information sharing policy is of course a necessary part of the armoury of the public sector in achieving what they are trying to go about by way of delivery of online services. Such a code will certainly contribute to providing proper transparency scrutiny and accountability. However the question remains as to its sufficiency. It is my view that further guidance needs to be developed which is around the areas of categorisation of personal information, an understanding of how to present information about service quality and having some mechanism by which the trustworthiness of the integrated set of activities which make up a good experience of a data owner is better articulated. It is in my view that this could best be done a combination of the Central Sponsor for Information Assurance (CSIA) and CESG within government in combination with appropriate professional bodies. The professional bodies I have in mind are the British Computer Society, the Institute for

Information Systems professionals and the Institute Engineering and Technology. However the sociological and business aspects of such a development would also need development and it is possible these three organisations should combine with the CBI and the IoD to cover off those aspects and ensure that a compatible set of guidelines is produced. I have had in other contexts preliminary conversations with all of these organisations and all are keen to help in order that a holistic view of how society, business and individuals can coexist in such a way that information sharing of personal information in the public and private sectors is improved in a way which satisfies all stakeholders.

Section 5 Technology

Question 20: Technological advances have increased the ease with which personal information can be shared and have actually also improved the mechanisms for the protection of personal information. Examples of this are the collaboration software that now comes bundled with a number of office packages which allows information to be shared very easily due to the use of underlying representational standards such as XML. The protection has been improved by the use of embedded encryption techniques such as SSL at the internet level and SOAP encryption for web services which largely prevents eavesdropping and by the use of encryption techniques for stored information to prevent static data theft. However the development of robust and well designed business processes and robust audit and governance procedures would seem to be lagging behind the technological advances so the framework within which a lot of these technologies are being used is immature compared with them. I believe that there is a considerable body of work that needs to be done in governance and management in order to improve matters.

Question 21: Whether or not the law should be the mechanism by which technical safeguards are improved or whether use of a slightly less strong mechanism of regulation is more appropriate I think it worthy of debate. It is arguable that the law should only be used in those instances where misuse or abuse of personal information would produce catastrophic consequences. A graded set of measures may well be worthy of consideration partly because if it is made illegal for personal information to be held in certain instances then enforcement becomes expensive and one has to look at how the law could actually be implemented. A regulatory framework might be more effective and much cheaper to implement and provide a proportionate and more agile response. The extreme cases maybe for certain classes of health records or for details of children who are at risk or for people who are in extremely trusted positions within national security as examples would need to be covered by legislation.

Question 22: Privacy enhancing techniques are certainly part of the armoury of the technological shopping basket and should certainly be used in an embedded way where appropriate in facilitating activities such as performing medical research as illustrated in the question but other statistical based social science activities which are increasingly going to be important as society reacts to the complexities of the context in which it finds itself.

Section 6: International Comparisons

Question 23: The legal framework for sharing and protecting personal information outside the UK it is not something I would regard as being expert in so I will not provide any comment to this question.

Question 24: In a similar vein I do not have expertise in the context of international examples of good practice.

Question 25: Similarly I do not have any comment of jurisdictions that have adopted any particular approach except to say that I am aware of differences one has to understand in the different coding of the legal structures and also the different cultural heritage in which they work so the translation of any experience from any other country to our own is dangerous because of the contextual differences.

Question 26: We do seem in the UK to be considerably more conservative about our private information and its use than in some countries where it is expected that disclosure of personal information happens and that services are derived as a result of it. This could come from people's cultural attitude to the legal system but it could also come from people's trust of each other in the societies they live in.

Section 7 Additional questions

Questions 27: The sharing of personal information is complex and this review I believe is a very worthwhile and laudable process for gathering a whole range of opinions on what is important, what techniques are available, what the government, the information commissioner and the regulatory bodies should do with regard to the sharing and the protecting of personal information. One aspect is that technological underpinnings are changing extremely rapidly and I believe the review should take cognisance of that fact. A number of information management metaphors are changing, particularly on the internet and in the use of the web; examples include the use of mobile and portable devices, surveillance both online and in the streets where the concept of personal information starts to become something which is essentially disclosed by a human being by being in a societal context and over which they have little control. The disclosure of the information is vital if they want to have any form of interaction with a society which has the internet and the services derived from it embedded as an integral part of that society. The issues apply explicitly to almost all sectors of government, both central and local, and therefore I think the review should recommend that a piece of horizon scanning work is justified that looks at how technological and other social contextual changes that are predictable will affect the recommendations of this review so that they are embedded in a forward looking context rather than one which could be seen as reactive to a current situation.

Question 28: Outside the context of personal information in the public and private sectors there is the bigger issue of the sharing of information in general between and within the public and private sectors whether it is personal information or not and there is an extremely complicated and difficult issue to face up to which is the aggregation of personal information with non personal information which is then shared between public and private sector organisations. The reality of any services that use these two types of information as we currently understand them is that they

will be aggregated and are being aggregated. This review should not turn a blind eye to those factors and should take account of the fact that it is very difficult to separate one from the other in any real service and that the reality is complex. I am not offering any solutions to that issue except to say that a much better understanding both as an academic subject and as a set of business metaphors of information as an asset, information processing as a discipline and information management and governance as part of good business practice need to be much more thoroughly researched, many more businessmen and public servants need to be educated in and the factors that affect the benefits that can be derived and the risk to be mitigated need to be much better understood.

I would be of course be delighted to meet with the review team to discuss further in the all of the points made above and I welcome the opportunity to have been able to comment.

Professor Brian Collins

5 February 2008