



Big Data

Public views on the collection, sharing and use of personal data by government and companies



Version 1, April 2014

Table of Content

1	Executive Summary	1
2	Introduction	2
3	Skimming the surface of public views	3
	3.1. What data does the public consider to be personal?	4
	3.2. What are the public’s views on the use of personal data by government and companies?	
	3.3. What advantages do the public perceive there to be for the use of their personal data by government and companies?	7
	3.4. What disadvantages do the public perceive there to be for the use of their personal data by government and companies?	8
	3.5. How do public’s views compare with actual behaviour?	10
	3.6. Conclusion	11
4	Delving the depths of public views	11
	4.1. What trade-offs are the public willing to make when faced with the advantages and disadvantages of use of their personal data?	12
	4.2. Is there a consistent public view on the use of personal data by government and companies?	14
	4.3. Do the public trust some organisations more than others to collect and use their data?	15
	4.4. How does the public think that the use of personal data should be controlled?	16
	4.5. Conclusion	18
5	Gap analysis	18
	5.1. How do public views change over time?	18
	5.2. What effect do current and emerging hot topics have on public views?	18
	5.3. What are public views on specific Big Data innovations?	19
	5.4. What factors affect how the public makes trade-offs?.....	19
6.	Appendix A: Case study: Personal data use for medical research	20
	Conclusion.....	23
7.	Appendix B: Summary of sources of information	24

This report is a review of information on the views and values of the public on big data available at time of writing, January to April 2014. It provides a snapshot of public views and is a live document, open to comments and additions.

The report will be useful to those interested in the public's views on new and emerging areas of science and technology and is particularly targeted to assist those involved in policy involving science and technology as they provide a background to what is already known about public views.

It is worth noting that this report focuses on a high-level account of public views summarised across various different engagement methodologies. It does not intend to capture the nuances or reasoning behind those methodologies.

The views and values of the public will change and new information will become available. Hence, we welcome your views, insights or comments.

- Do you know of further evidence that we should include?
- Do you have any comments or suggestions to improve the report?

You can comment [here](#).

This report, and the others in the series, has been produced by Sciencewise.

Sciencewise is a BIS funded programme to encourage the more widespread use of public dialogue in policy involving science and technology. Sciencewise provides advice and guidance to help those involved in the development of policy to understand and to take into account the views and values of the public in the development of policy involving science and technology. Sciencewise is able to provide:

- Advice and guidance on public dialogue and engagement.
- Assistance with the implementation of engagement as appropriate
- Financial support for the implementation of selected public dialogue projects
- Training and mentoring to assist those involved in policy development to build their understanding of the benefits and their confidence around engagement with the public

1 Executive Summary

Ever increasing amounts of data are being generated, at a faster pace and in more formats than ever before. The growing power to analyse vast and complex datasets can offer great insight into complicated issues, improving the quality of decision-making, delivery of public services, scientific research and many other areas.

Business and government are united in their belief in the potential of 'Big Data' to drive economic growth, scientific innovation and service efficiency. The views of the public, on the other hand, are much more varied, complex and nuanced.

This review focuses on the public's views on the collection, sharing and use of personal data by governments and companies. Though Big Data does not necessarily involve the collection or use of personal data, where Big Data is in tension with personal data privacy, ownership and control is where it is likely to cause most concern for the public.

This review finds that:

1. Data about who you are (i.e. personal information) is generally considered by the public to be more personal than data about what you do (i.e. behavioural data), though this distinction is likely to become increasingly spurious. Awareness of data collection and use by government and companies is quite high, but the level of understanding of what this means in practice is much lower. This suggests that individuals need to be engaged on the issue as citizens (deliberating on the conditions and safeguards), as well as consumers (agreeing or disagreeing to terms of service).
2. When asked, the public are ostensibly opposed to any form of data use and collection by government and companies, but in practice the public consider there to be no alternative to sharing personal information with government and companies in the modern world and expect it to increase in future.
3. Personal benefit is the strongest incentive for being in favour of the collection and use of personal data by government and companies, but the public report currently seeing little benefit from sharing their data and little confidence that they will see benefits in future. The public also identifies public goods (e.g. health research, prevention and detection of crime, and unearthing of dishonesty or fraudulent behaviour) as potential benefits of personal data use.
4. The public is particularly concerned about losing control of their personal data, with fear that they will become a victim of fraud or identity theft, and that their data will be shared with others without their knowledge or agreement.
5. Offering a specific personal or public benefit can significantly increase the general public's acceptance of the collection, sharing and use of their data by government and companies, but even when a specific benefit is offered, the public remain concerned about the collection, sharing and use of particular types of personal data (e.g. bank account, savings and pension details).
6. There is no consistent "public view" on what constitutes personal data, the benefits of sharing personal information and behavioural data, and comfort levels with different uses of data. The public can be segmented into a number of groups sitting along a continuum between pro- and anti-sharing.
7. The public thinks that personal data should only be used by government and companies for their personal benefit. People are keen to have more control over the use of their personal data and want stronger safeguards towards its use, and there is strong support from the public for more information on how government and companies collect, share and use data.
8. It would be beneficial to gather more evidence around how public views change over time, the effect of media attention, what public views are on specific data technologies and what factors affect how the public makes trade-offs.

The combination of high complexity, low public understanding and high public interest, mark the issue of personal data use out as one that requires much greater public deliberation.

2 Introduction

'Every day, we create 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. This data is big data.' (IBM)¹

In the past couple of years “Big Data” has become big news around the world, moving into the mainstream of public, economic, business and science discourse. The term, which originated among developers in Silicon Valley,² has been used since the 1990s, but it was 2012 when it came to the forefront of the minds of many politicians, policy makers, business leaders and journalists.

As illustrated by the quote above, ever-increasing amounts of data are being generated, at a faster pace and in more formats than ever before. Technological developments in phones, computers, websites, telescopes, video cameras, biological and chemical environmental monitoring equipment, among many others, are all generating new streams of digital data.

This expansion of the volume, velocity and variety of data³ has reached the point that standard software tools and statistical skills are no longer sufficient for managing the size and complexity. Big Data, therefore, refers to data that requires new technologies to make sense of it.

'Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.' (Gartner)⁴

The growing power to analyse vast and complex datasets can offer significant insight into complex issues, improving the quality of decision-making, public services, scientific research and many industries. In 2013, David Willets (Minister for Universities and Science) identified Big Data as one of Eight Great Technologies⁵ to propel the UK to future growth⁶, and in October 2013 the UK Government published its strategy for UK data capability:

'The UK government is determined to position the UK to make the most of the data revolution. Big data is one of our Eight Great Technologies, and for good reason – its potential impact is so significant that it could transform every business sector and every scientific discipline. We are supporting our data infrastructure, most recently with £189 million of funding for big data in last year's Autumn Statement. We have also established the E-infrastructure Leadership Council to advise the government on the computing infrastructure and skills we need to take advantage of this opportunity.' (David Willets MP & Matthew Hancock MP)⁷

Though the opportunity offered by Big Data is great, there is also significant potential for Big Data to be misused and/or have unintended negative consequences for individuals and society. In particular, the explosion of data and the increasingly sophisticated way it is analysed has raised concerns about how, when and why personal information is collected, shared and used. Two recent examples of the tension between the possible benefits and concerns with Big Data are the mass collection and use of communications data by security services, and the UK Government's “care.data” initiative to share patient data within the NHS and with some third parties.

¹ <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html>

² Lohr, S. (2013) 'The Origins of “Big Data”: An Etymological Detective Story.' *The New York Times*. <http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/>

³ Laney, D. (2001) 3D Data Management: Controlling Data Volume, Velocity, and Variety. META Group. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

⁴ <http://www.gartner.com/it-glossary/big-data/>

⁵ Willets, D. (2013) *Eight Great Technologies*. London: Policy Exchange

⁶ <https://www.gov.uk/government/speeches/eight-great-technologies>

⁷ HM Government (2013) *Seizing the data opportunity: A strategy for UK data capability*. London: Department for Business, Innovation and Skills. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/254136/bis-13-1250-strategy-for-uk-data-capability-v4.pdf

While business and government are united in their belief in the potential of 'Big Data' to drive economic growth, scientific innovation and service efficiency, the views of the public are much more varied, complex and nuanced.

This review focuses on the public's views on the collection, sharing and use of *personal data* by governments and companies. Big Data does not necessarily involve the collection or use of personal data; the Large Hadron Collider, for example, generates Big Data (30 petabytes, produced at one gigabyte a second, in fact⁸) with no obvious implications for personal privacy. However, many of the applications of Big Data do require the collection, sharing and use of data relating to the personal characteristics and behaviour of individuals in their day-to-day lives – including, for example, when accessing public services, buying goods and services, communicating with others, surfing the internet and being present in public and private spaces. Where Big Data is in tension with personal data privacy, ownership and control is where it is likely to cause most concern for the public.

The answers to each of the following questions are, therefore, of critical importance when considering the applications of Big Data:

1. What data does the public consider to be personal?
2. What are the public's views on the use of personal data by government and companies?
3. What advantages do the public perceive there to be of the use of personal data by government and companies?
4. What disadvantages do the public perceive there to be for the use of personal data by government and companies?
5. What trade-offs are the public willing to make when faced with the advantages and disadvantages of the use of their personal data?
6. Is there a consistent public view on the use of personal data by government and companies?
7. To what extent does the public trust government and companies to use their data responsibly?
8. How does the public think that the use of personal data should be controlled?

This review summarises the existing evidence of public views on these questions. It was collated through online research and focuses on the last five years. A full list of sources is contained in Appendix 1, including details of their methodological approach.

3 Skimming the surface of public views

The public have been asked about their views on data on a number of occasions in recent years. For the most part, this has taken the form of opinion surveys by public, private and civil society organisations, exploring the public's understanding and views on the collection and use of personal data by government and companies. Other studies have sought to delve into these views in more detail through focus groups and deliberative methodologies.

The heavy reliance on opinion surveys to understand public views on such a complex issue is far from ideal. As Hallinan and Friedewald (2012) highlight in their review of the evidence on the European public's views on the use of personal data:

'Public opinion is a notoriously difficult substance to judge – not least owing to the nuance and constant shift of individuals' opinions. Public opinion is particularly difficult to judge in relation to complex, value laden and abstract issues such as individuals' conceptions [of] data flows and the data environment.

In this analysis, certain issues should be specifically borne in mind as problematic. Firstly, surveys are an imprecise tool in the creation of an image of a diverse public. Secondly, with each survey, a number of

⁸ <http://www.marketingmagazine.co.uk/article/1185012/big-data-issues-try-coping-large-hadron-collider>

methodological flaws can influence the reality and truth of eventual findings. Thirdly, considering the abstract nature of the subject matter, it is difficult to gauge whether answers to survey questions genuinely reflect individuals' perceptions related to the question topic or how much they tell us about how strongly convictions indicated in an answer are actually held.' (Hallinan & Friedewald, 2012)

As this review will outline, the combination of high complexity, low public understanding and high public interest, mark the issue of personal data use out as one that requires much greater public deliberation.

3.1. What data does the public consider to be personal?

Summary of public views

- Data about *who you are* (i.e. personal information) is generally considered by the public to be more personal than data about *what you do* (i.e. behavioural data).
- Awareness of data collection and use by government and companies is quite high, but the level of understanding of what this means in practice is much lower.
- The public's awareness and understanding of data collection and use by government and companies is not necessarily improving.

Recent studies of public views on data suggest that **understanding of what constitutes personal data is not fixed**; it varies from person to person and is changing over time. That said, the general view appears to be that **data about who you are is more personal than data about what you do**. Demos (2012), for example, found that 'the public tends to consider information that might allow someone to be personally identifiable or details about their personal lives - such as phone number or how many children one has - as personal', whereas people view 'information about behaviour - often generalisable or aggregatable - as less personal' (Demos, 2012). This is supported by a Eurobarometer (2011) poll which found that data such as financial information (87%), medical information (83%), passport number (77%), fingerprints (73%) and address (71%) were considered to be personal by the vast majority of respondents. On the other hand, less than half of respondents considered the websites you visit (41%), your tastes and opinions (32%) and things you do (32%) to be personal. However, the collection and use of behavioural data at scale is a relatively recent development and is likely less obvious to the public than the collection and use of their personal information. **As the processing of behavioural data by companies and governments increases, and the public's understandings of its implications for privacy become more developed, public opinion could shift.** The example of the US shopping chain Target, which can predict whether a customer is pregnant based on their consumption habits, illustrates how the distinction between behavioural data and personal information can be quite spurious; behavioural data, combined with the right predictive model, can reveal a considerable amount about an individual, and their family and friends.⁹ The distinction between personal and behavioural data is likely to become increasingly fuzzy.

The Wellcome Trust (2013) found that their focus group participants distinguished between types of personal data according to:

- 'The degree of seriousness/risk if the data were misused or stolen;
- The perceived level of security of the data;
- Anonymous vs. personally identifiable data;
- Recognition of the value of data collection (to self vs. to others) vs. unclear benefit;
- Free choice to create data vs. enforced/necessary existence of the data;
- Government and non-government data.' (Wellcome Trust, 2013)

The public's awareness **of the use of personal data** by companies and public bodies is on the face of it quite high. Deloitte (2012), for example, found that 82 per cent of the public report having some

⁹ Duhigg, C (2012) 'How companies learn your secrets'. *New York Times*. <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>

degree of awareness that private sector companies and public sector bodies collect data on people and their activities. Likewise, Ipsos MORI (2014b) found that participants were ‘familiar with the importance of data in modern society’ and spontaneously mentioned its worth to individuals and organisations.

However, under closer scrutiny, **awareness appears to be broad but shallow**; individuals for the most part are aware that personal data collection and use takes place, but do not understand the methods or impact. For example, Ipsos MORI found that while dialogue participants would speak spontaneously about the collection of personal data through filling out forms or surveys, they had to ‘be prompted to start thinking about more passive forms of data collection, for example cookies or data on travel or purchasing patterns’ (2014b). Likewise, Hallinan and Friedewald (2012) report that:

‘Whilst the public places significant weight on the values of privacy and data protection and has a formal understanding of the features of the data environment, there is a significant knowledge deficit relating to the specifics of data flows and processing. Although the public felt that they were being forced into engaging in an ever increasing number of data transactions, they lacked the clarity and understanding to evaluate the significance of these transactions either at the individual or social level.’ (Hallinan & Friedewald, 2012)

Likewise, Demos (2012) found that while ‘the public is aware that personal information and behavioural data are used for commercial purposes, [...] understanding what this means in practice is limited’ and Deloitte (2012) concluded that ‘such a high level of awareness among the public signifies neither understanding nor consent’. **The public’s superficial awareness appears to mask a great deal of uncertainty and confusion below the surface.**

Despite the media attention in the past couple of years it appears that **public awareness of personal data collection and use is not necessarily increasing**. In fact, Deloitte (2013) found that the percentage of people who say they are “fully aware” that organisations collect data about them and their activities fell by 10 per cent (from 45% to 35%) between 2012 and 2013. Deloitte suggest that this is a sign that public understanding is not keeping pace with the development of new ways of collecting and using data. Alternatively, it could be the case that some segments of the population are becoming more aware of how unaware they are of the full extent of data collection and use by governments and companies.

3.2. What are the public’s views on the use of personal data by government and companies?

Summary of public views

- The public consider the collection and use of personal data to be a big issue
- When asked, the public are ostensibly opposed to any form of data use and collection by government and companies.
- In practice, the public consider there to be no alternative to sharing personal information with government and companies in the modern world and expect it to increase in future.
- A significant proportion of the public expects to feel less comfortable about sharing personal data in future.

Although public understanding is quite shallow, **the collection and use of personal data by government and companies is considered by a significant majority of the public to be a big issue** for them. When Eurobarometer (2011) asked, two-thirds of the public (67%) said that disclosing personal information is a big issue, compared with less-than-a-third (30%) who said it is not.

The **public’s immediate reaction to the collection and use of their data by companies or government appears to be one of opposition**. Deloitte (2012) found that the majority of the public oppose the use of their data by companies providing goods and services in almost all circumstances. Opposition ranges from 82 per cent for use of personal finance details and 76 per cent for information on health, medical history and past illness, to 68 per cent for goods and services purchasing history and 59 per cent for social networking posts. **Rejection of personal data storage and use also**

applies to government, as well as companies; IIPS (2008), for example, found that 64 per cent of the public disagreed with the statement: "I don't mind what information government departments hold on me." However, findings from Ipsos MORI's (2014b) public dialogue appear, at least in part, to contradict this. The authors report that few participants saw compulsory collection or data linking by government as a problem, with some even suggesting that giving data to government is 'an act of citizenship which allows the government to make better decisions and improve the way things work' (2014b):

'Participants commonly assumed that governmental administrative data is already linked and shared across departments, and supported this for operational uses.' (Ipsos Mori, 2014b)

A significant majority of **the public see the sharing of personal information as a necessary part of being a consumer** in the modern world. Eurobarometer (2011), for example, found that 65 per cent of the public agreed with the statement that "There is no alternative to disclose personal information if one wants to obtain products or services". The vast majority of **the public also recognise that data collection and use will increase** in the future. Many members of the public appear to take a rather fatalistic view - considering it a foregone conclusion - and perhaps as a result consider opting out of sharing their data an extreme choice (Wellcome Trust, 2013). The vast majority (82%) of the public agree that 'Disclosing personal information is an increasing part of modern life' (Eurobarometer, 2011). Ipsos MORI, similarly, report that their public dialogue participants:

'[D]escribed how data is collected from them all the time, for example when using companies' services, interacting with the government, and making applications for jobs or courses. In general, they were either uninterested in this, or resigned to it, seeing the modern world as one in which people collecting data from or about you on a regular basis as "just part of life".' (Ipsos MORI, 2014b)

This is reflected in the finding from Demos (2012) that, in 10 years' time, 48 per cent of the public expect to be sharing more information with companies, compared with 19 per cent expecting to be sharing less. During the same period, 47 per cent of people expect to be sharing more information with government, compared with 15 per cent expecting to be sharing less. Hallinan and Friedewald (2012) suggest that:

'The deterministic approach to obligatory information disclosure can arguably also be seen as a practical coping mechanism. The necessity to act in an environment which the individual does not understand, and in which there is an awareness of risk, still needs justification in the individual's mind.' (Hallinan & Friedewald, 2012)

While the public might expect to be sharing more data in the future, **they also expect to not feel comfortable about it;** according to Demos (2012), 45 per cent and 39 per cent say they expect to feel less comfortable with sharing information with government and companies respectively.

IIPS (2008) found a distinction between public views on data collection – which people were generally accepting of – and data use – which people became more concerned about:

'People are happy to give personal information: it's seen as a necessary part of modern life, and you get something in exchange for it, such as money off, or a quicker service. However, the worry comes later about how it's used, or could be used. A lot of discomfort was expressed about predictive profiling and social engineering. It's seen as acceptable to get loyalty points and the benefits they bring, but not if the information is then used to track your cake consumption and suggest you go on a diet: that's seen as an intrusion into privacy'. (IIPS, 2008)

This interest in the personal benefits but concern about the wider use of data is borne out across the research on public views of personal data collection and use.

3.3. What advantages do the public perceive there to be for the use of their personal data by government and companies?

Summary of public views

- Personal benefit is the strongest incentive for being in favour of the collection and use of personal data by government and companies
- The public report currently seeing little benefit from sharing their data and little confidence that they will see benefits in future
- The public also identifies public goods (e.g. health research, prevention and detection of crime, and unearthing of dishonesty or fraudulent behaviour) as benefits of personal data use
- The collection, sharing and use of personal data to improve public services is seen by some as a benefit

The Wellcome Trust (2013) found the **main benefits identified by members of the public for the collection and use of personal data** by government and companies to be:

- ‘the Government identifying needs, planning resources and services, and allocating funds’;
- ‘prevention and detection of crime and, including terrorism’;
- ‘tailored marketing’;
- ‘identifying social/population trends and statistics’;
- ‘convenience and time-saving when shopping and doing other transactions, if personal data were already held’;
- ‘unearthing dishonesty (e.g. fraudulent benefit claimants and tradesmen)’; and
- ‘availability of vital medical information in a medical emergency.’

These benefits are a **mixture of both personal and public goods**; indeed, among those in favour of organisations using personal data, the top two reasons for approval were found by Deloitte (2012) to be receiving more tailored and personalised services or recommendations (29%) and seeing it as a public good or thinking it would benefit society as a whole (15%). In contrast, helping companies to do better, make more profit or be more efficient (4%) came bottom. Similarly, Ipsos MORI (2014a) found significantly more support for the use of personal data for public benefits than for commercial benefits, concluding that, ‘[p]eople on balance oppose personal data being used for commercial gain’. **Support for the collection and use of personal data therefore relies on individuals believing that they or wider society, not just companies, will derive some benefit from it.**

Though personal benefit is reported to be the strongest incentive (among those in favour of personal data collection and use) for sharing personal information, **the public currently do not see - or have confidence that they will see - benefits from governments and companies using their data.** Demos (2012), for example, found that ‘in general, **the public sees only limited benefits of sharing personal information and behavioural data**’, with less than half of respondents saying they could see the benefits of data being used for any of the purposes tested. Similarly, Deloitte (2012) found that the public reports having relatively **little confidence that giving their data will result in a personal benefit**; 62 per cent of the public are not confident that sharing their data with companies or public sector bodies will result in better services or more relevant products (Deloitte, 2012).

While the general “public view” is quite negative on the benefits of companies collecting and using their data, this masks **significant variation across segments of the population**; for example, Demos (2012) found that 71 per cent of “enthusiastic sharers”, compared with only 25 per cent of “non sharers”, say they can see the benefit of online purchasing data being used to suggest future purchases.

In the public sector, some participants in the Wellcome Trust’s (2012) focus groups stressed the **importance of the collection of health data when it benefits the individual**. Data sharing within the NHS was considered by those in these focus groups to be positive, with the perception that more data sharing could be done within the NHS (Wellcome Trust, 2012).

However, Deloitte's (2012) opinion survey found the public to be **split on whether public bodies should share more data about people between themselves**, with 32% agreeing and 38% disagreeing that public sector organisations should share more data about people to improve the services they provide. Similarly, when IIPS (2008) asked, "How do you feel generally about the idea of public services sharing personal information that you provide with other public services?", 9 per cent and 42 per cent of respondents said they would be happy for "all information" and "some information" to be shared respectively, with 47 per cent saying they would not be happy for any information to be shared.

As will be outlined in section 4.1, when offered a specific public good, acceptance of data sharing can increase significantly. Ipsos MORI, for example, found that:

'[P]eople are most supportive of individuals' data being used when there are tangible public service benefits. Nine-in-ten (88%) support the use of people's data to help develop treatment for cancer, three-quarters (73%) support data being used to improve the scheduling of transport services and seven-in-ten (70%) support data use to prevent crimes.' (Ipsos MORI, 2014a)

Ipsos MORI's (2014b) public dialogue on the use of government administrative data found that the:

- 'Use of administrative data to improve the way services are run or to increase national security was uncontroversial for most.
- Those who were more trusting of government generally tended to think that any government use of data would be fairly benign and benefit the general public, or a subsection of the population.
- There was strong support for the use of administrative data in planning for future service provision.
- Participants wanted data to be used to reduce fraud in government services.
- Many were open to the idea of data being used more efficiently to avoid repeat collection of the same information from individuals.' (Ipsos MORI, 2014b)

As well as to accrue benefits to individuals and the public, **the use of personal data for enforcement is also considered by some participants in opinion polls and public dialogues to be a benefit**. For example, the detection of fraud is mentioned by participants in a number of studies as being a benefit of data collection and use by government and companies, for example in identifying benefit cheats. The Wellcome Trust (2013) report that some think more could be done to catch those who flout the system by linking data between organisations (e.g. linking Facebook data and benefits payments). However, **others are concerned about data being used by government to punish or withdraw a benefit or service from individuals** (Wellcome Trust, 2013). When talked through a range of scenarios of how data could be linked between organisations (public and private) for a potential personal or public benefit, participants in the Wellcome Trust's (2013) focus groups were typically concerned about data being used to target specific individuals or groups of individuals.

The public's default position when it comes to the sharing and linking of data within government and public services is one of significant caution, for many of the reasons outlined in the following section. However, while the public is not willing to give government a free pass to collect, share and use personal data, it is (as will be discussed in section 4.1) willing to give conditional consent in particular circumstances when there are clear personal and/or public benefits on offer.

3.4. What disadvantages do the public perceive there to be for the use of their personal data by government and companies?

Summary of public views

- The public is particularly concerned about losing control of their personal data, with fear that they will become a victim of fraud or identity theft, and that their data will be shared with others without their knowledge or agreement
- Linked to concerns about loss of control, the public is worried about how personal data will be used by government and companies, with concerns that it will be used to identify or

target particular individuals

- Lower social grades are particularly concerned about the sharing of personal data between public services

The Wellcome Trust (2013) found **the main drawbacks identified by members of the public for the collection and use of personal data by government and companies** to be:

- ‘the potential for data to be lost, stolen, hacked and leaked, and shared without consent, leading to security concerns;
- invasion of privacy, with a sense of Big Brother watching;
- unsolicited marketing and advertising, with a high nuisance factor; incorrect/inaccurate data collection, which would be hard to correct and undo;
- potential discrimination (e.g. data falling into the hands of a prospective employer).’

Similarly, the top five reasons for opposing data use found by Ipsos MORI (2014a) were:

- ‘Abuse of personal information/identity theft (40%)’
- ‘People have a right to privacy (32%)’
- ‘Don’t trust companies/don’t want them to profit (18%)’
- ‘Being sent junk mail/spam (18%)’

Loss of control over personal data and information is found to be a significant concern of the public across surveys, interviews and focus groups. The Wellcome Trust, for example, found that:

‘People acknowledge that their personal data are often held by others such as public officials and customer service personnel, and while it is recognised that this might be inevitable, it still leaves people with a sense of powerlessness that they do not control their own data.’ (Wellcome Trust, 2013)

Likewise, the top five risks identified by respondents to the Eurobarometer (2011) survey were all linked to losing control of their data, including being a victim of fraud (65%), being at risk of identity theft (56%), information being used without their knowledge (34%), information being shared with third parties without their agreement (33%) and information being used in different contexts from the ones where it was disclosed (23%). Similarly, Demos (2012) found people’s principle concerns to be companies using their data without their permission (80%), companies losing their personal data (76%), data being shared with third parties (76%) and ID theft (70%).

Specifically, the risk of personal data theft and misuse is consistently found to be at the top of the public’s list of concerns. For example, Ipsos MORI report from their public dialogue on the use of government administrative data for research, that:

‘Personal data security was very important to participants, and this framed much of the discussion. They were particularly concerned about identity theft, and personal data being sold on to other organisations.’ (Ipsos MORI, 2014b)

The **sharing of personal data between organisations is of particular concern to the public**, in part due to concern over loss of control. IIPS, example, found that:

‘People are generally happy for data to stay with one organisation, but are concerned when it’s shared. They expressed a fear of the “master file” and the “data lock nightmare”, where errors are perpetuated and you’re trapped in a cycle of data you’ve given. There is a desire to have an element of control over their own personal information.’ (IIPS, 2008)

That said, Ipsos MORI (2014b) found that participants in their public dialogue on the use of governmental administrative data commonly assumed that it was already linked and shared across departments.

The concern over losing control of data is not simply a question of principle, but linked to a real **worry about what personal data could be used for**, particularly where individuals can be identified. Wellcome Trust, through their focus groups, found there to be ‘fairly widespread wariness about being watched / snooped on by the government, corporations and criminals’ (2013).

The **public appears significantly less concerned about the collection and use of aggregated and anonymised data, than data that can be used to identify and target particular individuals**. The Wellcome Trust found there to be particular concern with individual level data being used to target individuals, with the expectation being ‘that blame and desired behaviour change will be implicit, which fundamentally threatens the concept of free will.’ For example, focus group participants were asked their opinions on ‘anonymous loyalty card data being used to influence/gauge public health programmes *and* loyalty card data being linked with personal patient health records’. They found that, ‘the first part of this linkage example was found broadly acceptable, because anonymous. But the second part sparked lively debate and strong reactions. The idea of the government snooping on people’s purchasing behaviour in order for their GP to tell them later they are buying the ‘wrong’ food or eating things that are bad for them was considered to be outrageous!’ (2013). However, the public’s understanding of what is meant by aggregated and anonymised data is likely to be only partial. Indeed, aggregated and anonymised data can also be used to target individuals based on predictive models.¹⁰

The collection and sharing of personal data across public services is a particular concern among lower social grades. IIPS (2008) found that 69 per cent of social grades E and 59 per cent of D would not be happy with their data being shared with other public services, compared with 36 per cent of ABs, 42 per cent of C1s and 49 per cent of C2s.

This public concern over how personal data, once collected, will be used is not peculiar to the government and public sector. 80 per cent of the public report being concerned about companies using information they hold about them for a different purpose to one it was collected for, without informing them (Eurobarometer, 2011).

3.5 How do public’s views compare with actual behaviour?

Summary of public views

- The majority of the public (knowingly or unknowingly) hands over personal data on a regular basis
- There is a significant discrepancy between the public’s stated preferences and their actual behaviour
- The discrepancy in views and actions is in part due to not having the information required to rationally make the necessary trade-offs on a day-to-day basis

The general public’s real world approach to sharing their personal data with government and companies suggests that they are much less concerned about the risks and much more sold on the benefits than they make out. As found by Deloitte (2012), the majority of individuals create a ‘digital footprint’, with 61 per cent regularly pay for goods and services by credit or debit card, 51 per cent have one or more shop loyalty card(s), and 36 per cent have one or more online shopping account(s). By contrast, only 27 per cent have an ex-directory phone number, 23 per cent have opted out of having their details available for marketing from the electoral register and 17 per cent have a separate email account which they use when they think they will get spam (Deloitte, 2012). Similarly, Ipsos MORI found that:

¹⁰ Duhigg, C (2012) ‘How companies learn your secrets’. *New York Times*. <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>

'[W]hile people may have concerns about how their data are used, most do not act on these concerns, especially if they are already signed up to a particular service. Over two-fifths (46%) say they have decided not to take up (...) services (...) because of concerns about how their data might be used. Under a fifth (16%) say they have stopped using any of these services because of this.' (Ipsos MORI, 2014a)

However, this discrepancy between the public's stated preferences and their actual behaviour appears to be, at least in part, due to individuals not having the information required to rationally make the necessary trade-offs on a day-to-day basis; as Hallinan and Friedewald (2012) conclude:

'The lack of understanding of the data environment generally, translates onto the evaluation of each data transaction individually. The holes in comprehension impact on certain key factors in the decision making process to reduce the ability for the individual to 'rationally' balance each action. Thus, whilst abstractly the individual may attach high importance to personal data and be aware of certain risks in releasing this data, this does not transfer onto each specific transaction.' (Hallinan & Friedewald, 2012)

An alternative way of expressing this might be that while public concerns typically exist at a fairly general level, benefits are often a lot more specific, apparent and tangible, at the time a decision is made. Therefore, general concerns are not necessarily applied in a particular scenario, and if they are might be trumped by more immediate interests. This suggests that **implied consent is not a reliable indicator of actual consent, but neither necessarily is stated preference.**

Public opinions can differ significantly depending upon how a question is asked and a topic is framed. As Singleton and others (2007) highlight, 'people will express concerns if questioned about 'concerns', but will readily trade these 'concerns' for health or other benefits, even altruistic ones. "Real world" choices can be very different (and constrained) from those offered in opinion surveys where costs and trade-offs may not appear.'

3.6. Conclusion

When asked their views on the collection, sharing and use of personal data – despite having limited understanding of the scale, collection methods and potential use – the public typically expresses concern and opposition, combined with a fatalistic view that it is an unavoidable part of modern life.

In part, opposition appears due to limited understanding of how the use of their personal data can bring benefit to them or the wider public, coupled with scepticism that any potential benefits will be realised. But more significant than this is the unease over how their personal data might be used in ways that are not of benefit to them and, therefore, a concern to maintain some control over its use.

The apparent unconcern with data privacy displayed by many in their day-to-day lives stands in stark contrast with their stated preferences. This suggests that government and companies should be careful not to take the public's seeming willingness to part with their personal data for granted or as a sign of consent, but neither should opposition be taken at face value or applicable to all circumstances. Rather, it is **important to dig deeper into the trade-offs that the public is willing to make, under what circumstances and with what conditions.**

4 Delving the depths of public views

This review has so far examined evidence of public views at a relatively basic level. This is not to discount these views or the importance of understanding them. However, the complexity of the issue of personal data use demands a deeper dive into public views to understand the trade-offs that the public are willing to make, how public views are distributed (i.e. the extent to which opinions are bunched at certain points or spread along a spectrum), whether public acceptance depends on the organisation in question, and what safeguards the public think should be adopted.

4.1. What trade-offs are the public willing to make when faced with the advantages and disadvantages of use of their personal data?

Summary of public views

- Offering a general personal or public benefit slightly increases the general public's acceptance of the collection, sharing and use of their data by government and companies
- Offering a specific personal or public benefit significantly increases the general public's acceptance of the collection, sharing and use of their data by government and companies
- Even when a specific benefit is offered, the public remain concerned about the collection, sharing and use of particular types of personal data (e.g. bank account, savings and pension details).
- The public is significantly more comfortable with anonymised and aggregated personal data being shared among public services than identifiable and individual level data.

Opinion surveys suggest **some willingness on the part of the public to trade-off their concerns against the potential benefits** to themselves or the wider public. Deloitte (2013), for example, found that some individuals say they would be happy for organisations to share their data with other organisations, where otherwise they wouldn't have been, if they were informed of how their data would be used for their or the public benefit. Indeed, a number of participants in Demos' (2012) workshops are reported to have noted that one of the reasons for low acceptance of data collection and use by companies are that the benefits are often not clearly stated:

"It just feels like I don't gain anything from letting companies use my data, none of them have ever told me how I benefit." (Participant - Demos, 2012)

Similarly, IIPS (2008) found that people who reject data sharing between public services often do so because of a lack of information about the purpose behind it: '53% of rejectors, as compared to 35% of acceptors, 'Don't know' why the government is keen for public services to share information about citizens.'

The evidence suggests that there is a **greater willingness to give data to government departments when general benefits are offered** (IIPS, 2008). 68 per cent of respondents said that they would be happy to provide details to government departments if it meant that they would provide a better service, compared to 25 per cent who would not. The case is particularly stark for explaining the benefits of handing over personal details to government when looking at those who would typically be in opposition; if it means a better service, 49 per cent of this group would shift from being "rejectors" to "acceptors". **When asked to make some specific trade-offs (with a clear personal or public benefit), the public appears to be much more comfortable with their data being used.**

While people seem more willing to share data with public services if a benefit is offered, some opinion survey findings suggest that the effect can be quite modest when it comes to the **sharing of an individual's personal data between organisations**. For example, Deloitte (2013) found that 10 per cent of people are typically happy with organisations sharing personal data with another company, rising to 13% if they know the data is used to provide tailored offers, while 22 per cent of people are happy with organisations sharing personal data with public sector bodies, rising to 31 per cent if they know the data is used for good, such as medical research. Similarly, IIPS (2008) found that only a quarter of respondents would be happy for public services to share their personal data with other public services if it meant that they would not have to provide the same details to other parts of government more than once.

However, when faced with specific and tangible scenarios and benefits of data being shared across public services, respondents seem much more comfortable with their data being shared. For example, 91 per cent of IIPS (2008) respondents agreed with the proposition that medical staff across the country should have access to their GP medical records, meaning that their medical history would be available to services if they needed medical care outside of their area. Almost as

many (89%) agreed that medical records should also be accessible to the police and emergency services in order that they could be accessed if they were involved in an accident. And four fifths agreed with the Department of Work and Pensions having access to their medical records to verify whether they were entitled to disability benefits (IIPS, 2008).

While the public might be encouraged to accept the sharing of medical records across public services where it delivers a benefit to them or society, they are significantly less accepting of some other types of personal data being shared. There is **significant nuance in the public's views depending upon the type of personal data in question**. Even those who would typically accept their data being shared between public services, in order that they did not have to give it more than once, are extremely cautious about some types of data being shared. Bank account details and information about savings and pensions were found by IIPS (2008) to be particularly sensitive, with an even majority (74% and 63% respectively) of those who would typically be happy for their data to be shared, saying they were not happy for these types of data to be shared. There are also some important nuances within data types; for example, certain types of health information, such as mental health data, are considered by some to be particularly sensitive (Wellcome Trust, 2013).

A significant majority of the public appear willing to trade-off their concerns about the collection of many – if not all – types of personal data (including sensitive data such as medical records) if there is a strong enough personal benefit in return. The Wellcome Trust (2013), for example, found that:

'With personal health data representing protection, care and cure at individual level, especially in circumstances when health is failing, it was reassuring, comforting and necessary for personal health data to be collected by and held within the NHS. (Wellcome Trust, 2013)

The same also appears to be true, perhaps to a lesser extent, for the public good. For example, **the majority of the public appear ready to trade-off their concerns about the collection and use of their medical data if there is a clear benefit to medical research.** For example, Ipsos MORI (2007) found that:

'More than two-thirds (circa 69%) [of the public] say they are 'likely', including just over one in ten (14%) who say they are 'certain' to allow their personal health information to be used for medical research purposes, compared with around a quarter who feel they would 'not be likely' (including 7% certain not to). This suggests both positive feeling towards the use of personal health information for medical research, and perhaps some caution, or desire for more information before any firm commitment is given.' (Ipsos MORI, 2007)

Further findings from this study suggest that a key factor in determining whether an individual agrees to the use of personal health information for medical research is whether they believe the advantages outweigh the disadvantages:

'Perhaps not surprisingly, perception that advantages of medical research outweigh its disadvantages has a key influence on likelihood of allowing personal health information to be used for medical research. Specifically, 45% of those who feel advantages outweigh disadvantages are certain or likely to allow their information to be used compared with just one in five (20%) of those who feel the disadvantages outweigh the advantages.' (Wellcome Trust, 2013)

As well as the framing of a topic, the evidence suggests that the specific context has a significant effect on how people view and make trade-offs. **Public opinions on one trade-off cannot necessarily be generalised and applied to other apparently similar trade-offs.** Differences in the type of data, format of data, proposed use of data, type of organisation, and possible advantages and disadvantages, can all affect the outcome of a trade-off.

4.2. Is there a consistent public view on the use of personal data by government and companies?

Summary of public views

- There is no consistent “public view” on what constitutes personal data, the benefits of sharing personal information and behavioural data, and comfort levels with different uses of data
- Age and social class are both linked to an individual’s views on the collection and use of their data, with younger generations typically sharing more but being less aware than older generations, and higher social classes being more comfortable with sharing their personal data than lower social classes
- The public can be segmented into a number of groups sitting along a continuum between pro- and anti-sharing, including non-sharers (30%), sceptics (22%), pragmatists (20%), value hunters (19%) and enthusiastic sharers (8%)

Though this review as referred to “public views”, there is in fact no one “public” view on the collection and use of personal data. **Public opinion is stretched across a range of positions, with the balance of opinion falling in different places depending upon the particular circumstances.**

An individual’s age and social class both appear to have some bearing on their views on data, with younger generations typically sharing more but being less aware and older generations sharing less but being more aware, and higher social classes being more comfortable with sharing their personal data than lower social classes. Rather than being polarised between pro-privacy and pro-sharing, public views on the use and collection of data have been found to sit along a spectrum (Demos, 2012; Wellcome Trust, 2013).

Demos (2012) segment the public into five groups according to their how personal or impersonal they regard their data to be, how comfortable they are in sharing their data, whether they believe they gain something from sharing their data, and how receptive they are to ideas that increase data safeguards:

Segment	% of population	Characteristics
Non-sharers	30%	<ul style="list-style-type: none"> • Very cautious about technology and sharing their personal data, and tend not to be experienced at using modern technology. • View their data as personal, and take proactive measures to keep them private: unsubscribing, deleting their browsing history, and alerting companies to possible violations. • This attitude towards privacy is not just internet specific: non-sharers often list their number an ex-directory and place a ‘no junk mail’ sign outside their door. • Knowledgeable about data protection and receptive to ideas that allow them to withdraw their data.
Sceptics	22%	<ul style="list-style-type: none"> • No single view about whether information is personal or impersonal in principle, but are sceptical about whether or not government and companies can be trusted. • Unlike the non-sharers they do not use online services much, and tend to be older – so there is little scope for them to build up trust through experience. • Cynical about a range of issues –including the benefits of sharing data. • Although they sometimes buy into ‘value exchange’ transactions when the personal benefits are clear, they would welcome measures to give

		them simple, direct and regular control over their data.
Pragmatists	20%	<ul style="list-style-type: none"> Do not know all the details of how their data are used, but take small measures to protect their privacy, such as retaining ownership over their data even after they have been shared with third parties. Prefer efficient services to complete privacy – seeing benefits from the sharing of personal information – so their trust in the companies or institutions that hold their data is key.
Value hunters	19%	<ul style="list-style-type: none"> Understands the financial value of their personal data, and recognises that sharing them can save money and time, and thus have a beneficial outcome. Tend to be young, and are often early adopters of technology and on the lookout for new advances that can make a practical difference to their lives Are not overly concerned about risks to personal information being shared, and are reasonably comfortable with their data being used. Are knowledgeable and understand the value of the data they are giving up.
Enthusiastic sharers	8%	<ul style="list-style-type: none"> Categorise a lot of the information about them as impersonal, and subsequently are comfortable with sharing it. Understand 'value exchange' transactions, seeing the benefits of sharing information, and are amenable to sharing even more in the future. Have some concerns about the ways in which their data might be misused, but are comfortable if data use is specified.

Views of what constitutes personal data, the benefits of sharing personal information and behavioural data, and comfort levels with different uses of data were all found by Demos (2012) to differ significantly between these groups.

The Future Foundation (2012), similarly, categorise consumers into a three way segmentation model: privacy pragmatists (53%), who will make trade-offs on a case by case basis as to whether the service or enhancement of service offered is worth the information requested; Privacy fundamentalists (31%), those who are unwilling to provide personal information even in return for service enhancement; and 'privacy unconcerned' (16%), those who are unconcerned about the collection and use of personal information. Interestingly, despite significant changes in how data is exchanged, the Future Foundation report that, **the size of each privacy segment remains largely unchanged since a similar study was conducted in 1997**: pragmatists (60%), fundamentalists (25%) and unconcerned (15%) (Future Foundation, 2012).

4.3. Do the public trust some organisations more than others to collect and use their data?

Summary of public views

- The public does not trust government and companies to keep and use their data securely and appropriately
- Public confidence in the use of personal data by companies is particularly low

There is a significant **lack of trust in government and companies to keep and use their data securely and appropriately**. Deloitte, for example, found confidence in companies and public sector bodies to keep data secured (40%) and not to share or sell data without their knowledge (29%) to be

low. However, **findings are mixed on whether the public trusts government or companies least with their personal data.** On the one hand, polling by Ipsos MORI (2013) found that less than a third (32%) of the public trusted government to use their data appropriately, with slightly more (43%) trusting companies to do so. On the other hand, IIPS (2007) found government to be more trusted than the private sector, with 20 per cent of people agreeing that they “do not mind which government departments hold information on me or what they hold” compared to 10 per cent who agreed with the equivalent statement for companies. Likewise, Ipsos MORI’s recent public dialogue found that:

‘Overall, participants trusted government’s intentions more than commercial companies on data security and protection from data misuse (such as selling information on to third parties) or fraud.’ (Ipsos MORI, 2014b)

However, considering public and private as two separate homogeneous blocks appears to mask some considerable differentiation within them. Eurobarometer (2011), for example, found that the organisations most trusted to protect personal information included both public and private organisations; Health and medical institutions (83%) came top, followed by banks and financial institutions (75%) and national public authorities (63%). The same was the case for those scoring lowest, with shops and department stores (48%), phone companies, mobile phone companies and internet service providers (43%), European institutions (38%) and internet companies (30%) all being trusted by less than half of the public.

That said, **confidence in the use of personal data by companies does appear to be particularly low.** For example, only 22 per cent of the public say they feel confident that companies do not sell details about them to other companies without their knowledge, 38 per cent that companies will keep their data safe and 24 per cent that companies will tell them how they use their data (Deloitte, 2013). The stakes, however, seem to be very high for companies, with **the loss of personal data by a company being considered by the public to be the ‘worst corporate sin’**, with 70 per cent saying that the loss of their personal data would lead them to seriously consider not using the company again. By comparison, exploiting overseas workers (53%), damaging the environment (49%) and paying senior executives a large bonus or salary (40%) were considered by significantly less people to be reasons not to use a company again (Deloitte, 2012). Deloitte (2013) conclude that

‘The majority of the British population is still not confident in the way that companies collect, use, handle and share data [...] This level of uncertainty and negative sentiment is not sustainable if businesses are to continue using data for commercial gain’. (Deloitte, 2013)

Demos (2012) reports that people ‘are sharing more than ever, but there is a **public “crisis of confidence” in the way that personal information and behavioural data are being used**’ with the public being uncomfortable about every type of information and data use they are asked about.

4.4. How does the public think that the use of personal data should be controlled?

Summary of public views

- The public thinks that personal data should only be used by government and companies for their benefit
- The public are keen to have more control over the use of their personal data and want stronger safeguards towards its use
- There is strong support from the public towards more information on how government and companies collect, share and use data
- There is strong support from the public for the anonymisation of personal data

As the previous sections have shown, while the public is opposed at first to the collection and use of their data by government and companies, when required to trade-off their concerns against specific and tangible personal and/or public benefits they are much more willing to agree to the collection,

sharing and use of their personal data. That said, the public has some strong opinions on how personal data collection, sharing and use by government and companies should be controlled.

A significant portion (61%) of the population are of the view that **government and companies should only use information about them for their benefit**, while almost two thirds (63%) think organisations should collect less information about them (Deloitte, 2012). There is a particular concern with the over collection of information; for example, 80 per cent of people are concerned about the unnecessary disclosure of personal information in order to obtain an online service (Eurobarometer, 2011).

In line with findings about people's concerns with the loss of control of their personal data, Demos (2012) found that **the public welcomed measures to give them more control over personal information and behavioural data**, especially knowing what is held about them, and to withdraw it if they wish'; for example, 73 per cent want the ability to withdraw data, 70 per cent want to see what information is held on them, and 69 per cent want to know exactly what data is held on them.

Next in the list of the public's priorities, as identified by Demos (2012), were legal protections (68%). Deloitte (2012), likewise, found that the **public want stronger safeguards towards the usage of data**, with 54 per cent of people saying they would be more comfortable with companies and public sector bodies using their information in future if there were stronger laws and safeguards for keeping their data safe; 38 per cent if the default option does not allow the organisation to use their data without their permission; and 27 per cent if there was a dedicated information charter so they know how their information is being used.

One of the key concerns of the public, as discussed earlier, is the use of personal data without an individual's awareness. Half of the public think that public authorities should tackle this by imposing a fine, followed by: Banning them from using such data in the future (41%); Giving people more direct control over their own personal data (33%); Compelling them to compensate the victims (31%) (Eurobarometer, 2011). Demos (2012) also found significant support (62%) for a system of fines for the misuse of data.

Demos (2012) found significant support for more information from companies and public bodies, with 64 per cent wanting a **clear statement on how information is used**. 66 per cent of respondents in Deloitte's (2012) study wanted an online dashboard to control their data. In support of the case for more transparency, Deloitte (2013) found that, when compared to the national average, people who believe that companies tell them how their personal data is used are twice as confident that companies keep personal data safe, twice as confident that companies use personal data to offer better services, three times more confident that companies do not sell personal details to other companies without their knowledge, and three times more confident that companies always remove personal details when passing on personal data to other organisations.

The Wellcome Trust (2013) found that 'clarity, transparency and reassurance' are required, and suggest that:

'The key fears and scepticisms to address in a positive communication of data linkage to the public are: what's the point? Is it a waste of resources? How might I and others benefit? What might leak out that could be harmful to someone? What care is going to be taken with the data? What money interests might be involved? (Wellcome Trust, 2013)

However, Demos (2012) found that **while overall the explanation of various types of data use slightly increases comfort levels, the effect is not even across groups**. For the enthusiastic sharers, value-hunters and pragmatist segments in the Demos (2012) survey, confidence in fact fell after specific methods and techniques were explained

Deloitte (2012) found **strong support for anonymisation**, with 62 per cent of respondents agreeing that information about them should only be shared after having been anonymised. Significantly, Ipsos MORI (2014a) found that 61 per cent of the public do not care how their personal data is used as long as it is anonymised and cannot be linked back to them. Likewise, the Wellcome Trust (2013) found a strong preference for aggregated and anonymised level data; while there were concerns that individual level data could be used to target individuals, data at an aggregate and anonymised level was viewed differently by focus group participants, with the perception that its use for medical research was for the greater good of society.

4.5. Conclusion

While on the face of it the public's immediate reaction to personal data collection and use is concern and opposition, this masks a significant degree of nuance below the surface. The findings of surveys and focus groups both suggest that the public is often willing to trade-off their concerns for personal and public benefits, though this is heavily dependent upon the individual and particular context.

Discussion of the "public view" also hides significant variation across the public. As is usually the case with such a complex issue with competing considerations, there is not one public opinion but a range of views fairly evenly distributed along a spectrum. The balance of where these views lie (i.e. pro- or anti-personal data use) is similarly dependent upon the context. Likewise, there is significant variation across different types of organisation in terms of the extent to which the public trusts it to collect and use their personal data, with the NHS being one of the most trusted organisations, and some types of commercial organisation (e.g. internet companies) being the least trusted.

These things considered, it is perhaps unsurprising that the public is keen to regain some control over their personal data and ensure that it is only used in situations that they agree to. If government and companies are to build public trust and gain consent for using personal data, it is important that:

- The public has easy access to the information necessary to know how their data is being used and make informed decisions when asked to consent to data collection and use
- The public is actively involved in making the trade-offs between privacy and other personal and public goods
- There are systems through which the users of personal data can be held to account

5 Gap analysis

5.1. How do public views change over time?

There is some evidence, as set out in this review, that public understanding and opinion on the collection and use of personal data have not changed much during the past decade. However, with such rapid developments taking place in the field, tracking the degree to which public views change over coming years to identify overarching trends will be important, particularly considering the emphasis and funding being directed at Big Data by the Government. For example, this review noted that personal information is currently considered by the public to be significantly more personal than behavioural information. However, as the analysis of behavioural data becomes more prevalent and apparent to the public, it will be important to track whether/how public views change.

5.2. What effect do current and emerging hot topics have on public views?

Recent months have seen significant media coverage of the collection and use of data by security organisations (particularly in the UK and US), as well as publicity and coverage of the governments "care.data" initiative to share personal data within the NHS and with some third parties. Ipsos MORI, in the Public Attitudes to Science 2014 study, highlight that:

'As a further context for the findings presented here, it should be noted that two major news stories may recently have played a part in influencing attitudes to data usage. First, the US National Security Agency's collecting of personal data emerged in June 2013, a few months before survey fieldwork. Second, the phone hacking trial in the UK began in October 2013, during fieldwork.

It is also important to note that the PAS [Public Attitudes to Science] 2014 fieldwork took place before the rollout of the NHS Care.data database, and its subsequent delay, were announced (in January and February 2014 respectively). Given the large amount of media coverage this proposed database has received, it is possible that national attitudes to big data have developed even further since PAS 2014.' (Ipsos MORI, 2014a)

Tracking the effect (if any) of these “hot topics” on the public’s awareness and views of the collection and use of personal data will be important as:

- Controversial uses of personal data by government and companies could damage public trust in other Big Data initiatives
- Understanding the degree to which public views on personal data use are determined by issues of the day could help to anticipate issues and develop a deeper understanding of public opinions

5.3. What are public views on specific Big Data innovations?

Big Data innovations are incredibly diverse, with different potential advantages and disadvantages inherent in each. As is evidenced in this review, public views on the collection, sharing and use of personal data can vary considerably depending upon the context. It would, therefore, be beneficial to review public views on specific Big Data innovations to understand the nuance of public opinion in different contexts.

5.4. What factors affect how the public makes trade-offs?

As was explored in this review, the public approaches different trade-offs in different ways depending upon the data in question and the possible advantages and disadvantages. Exploring in more detail what factors affect how these trade-offs are made would be useful for anticipating the public’s response to a particular scenario.

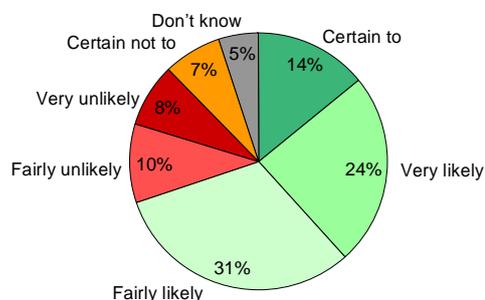
6. Appendix A: Case study: Personal data use for medical research

The collection, sharing and use of personal data for medical research, as discussed in section 4.1, enjoy a significant level of support from the public. A short review of the evidence on why and in what circumstances this is the case provides a useful illustrative example of the complexity and nuance of public views.

As is shown in *figure 1*, 69 per cent of the public say they would be at least 'fairly likely' to allow their personal health information to be used for medical research.

Figure 1: 'Allowing use of personal health information for medical research' (Ipsos MORI, 2007)

Q How likely, if at all, would you be to allow your personal health information to be used for the purposes of medical research?



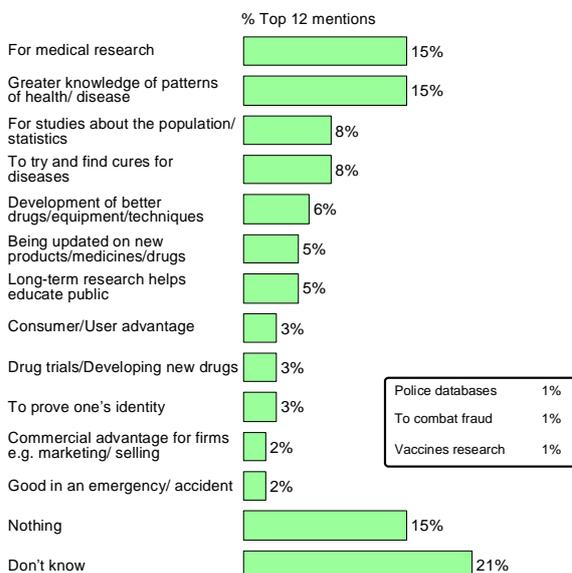
Base: 2,106 UK adults aged 15+. Fieldwork 14-18 September 2006

Source: Ipsos MORI

This willingness appears to be linked to a strong general belief in the advantages to medical research, with 70 per cent of the public feeling that the advantages of data use for this purpose outweigh the disadvantages, compared with just 6 per cent who think the opposite. Indeed, the principle advantages identified by the public of the collection of health information (shown in *figure 2*) are all related to use in medical research.

Figure 2: 'Advantages of collecting personal health information' (Ipsos MORI, 2007)

Q What advantages, if any, are there of collecting and using people's personal health information?



Base: 2,106 UK adults aged 15+. Fieldwork 14-18 September 2006

Source: Ipsos MORI

Where there are such public benefits on offer, the Wellcome Trust (2013) found that:

'There are superficially no/very few objections to medical data being used for the 'general good' (perceived as helping find cures and causes), provided commercial gain is not the priority.' (Wellcome Trust, 2013)

In fact, a majority of the population (60%) go as far as to agree that 'they have a responsibility (as beneficiaries of medical research) to allow their personal health information to be used in approved medical research projects' (Ipsos MORI, 2007).

While the advantages in this particular trade-off may be considered to be greater than the disadvantages, **public concerns for privacy and preferences over how their data will be used are very much still present.** The Wellcome Trust (2013), for example, found that:

'There was also a strong feeling that personal health data are confidential, private and sensitive, and should not be shared outside secure, authorised bodies such as the NHS, and especially not with private companies such as employers, insurance providers and drug manufacturers.' (Wellcome Trust, 2013)

As is shown in *figures 3 and 4*, the public continues to express a number of familiar concerns about the collection and use of their personal health information, with particular worries about privacy, misuse, loss of control and security.

Figure 3: 'Disadvantages of collecting personal health information' (Ipsos MORI, 2007)

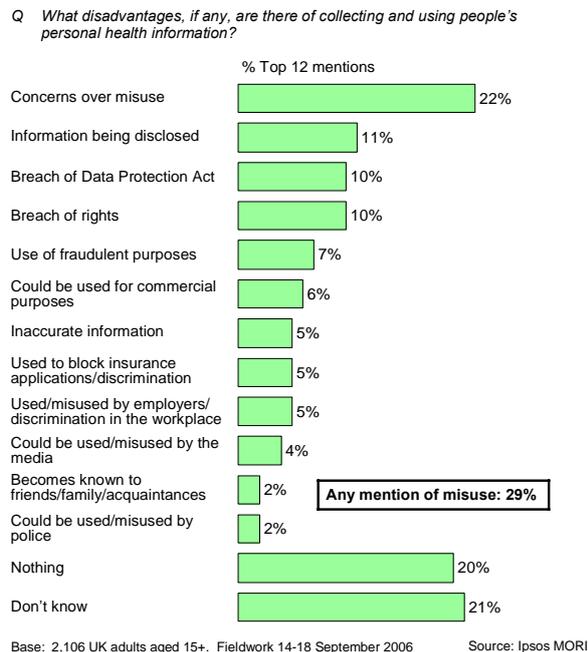
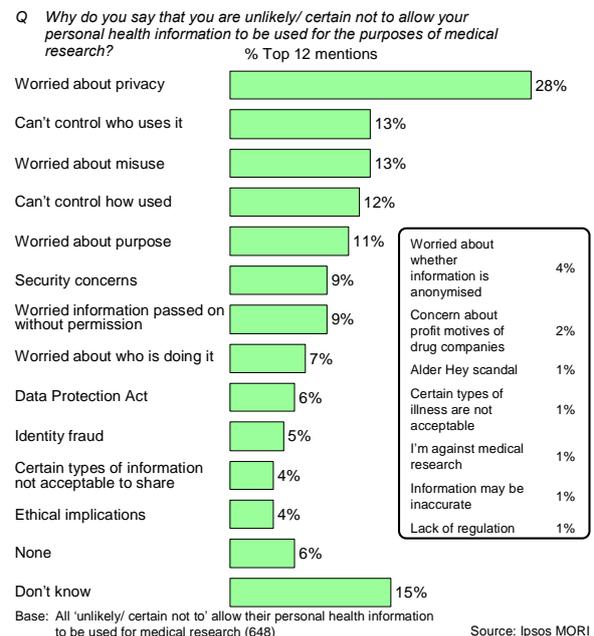


Figure 4: 'Reasons for not consenting to personal health information to be used' (Ipsos MORI, 2007)



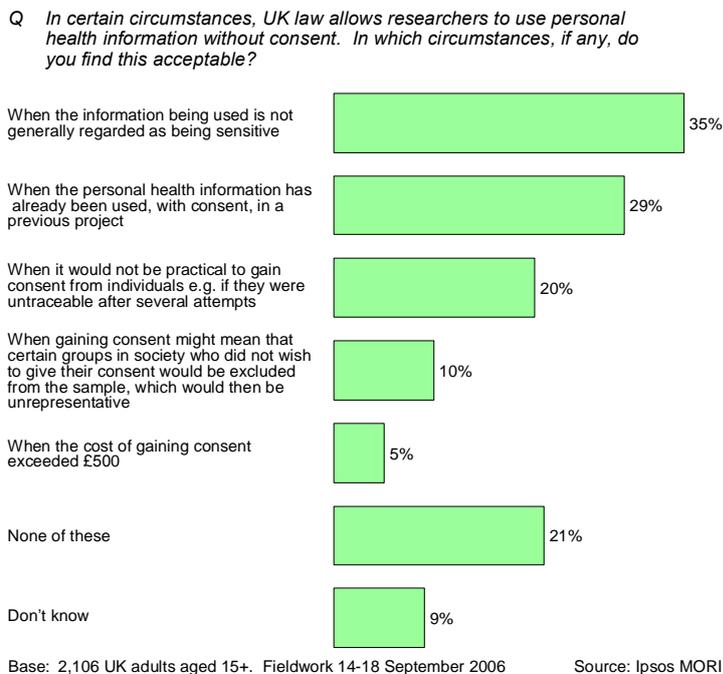
Qualitative studies (e.g. Wellcome Trust, 2013; Ipsos MORI, 2007) have shown that there is particular concern from the public about which organisations are given access to personal medical data and for what purposes, with particular opposition to it being used for commercial gain:

'The thought of personal health data getting into the 'wrong hands' outside the NHS did worry some people, especially if they had a particular history/condition. Examples of the wrong hands and inappropriate use were commercial companies for advertising purposes (e.g. targeting older/frail people with health products), and sharing personal data with employers, who might discriminate against individuals, e.g. because of mental health problems, or HIV status.' (Wellcome Trust, 2013)

This concern with commercial parties (e.g. insurance companies, schools, employers) accessing medical information is consistent across surveys and focus groups from the past decade (Wellcome Trust, 2013; Ipsos MORI, 2007; NPfIT, 2003; Scottish Consumer Council, 2005). That said, public views appear to be driven much more by the purpose of the data use, rather than a straightforward favouring of public organisations over private. The Wellcome Trust (2013), for example, report that ‘... the public does not seem particularly sensitive about who conducts research involving linking of health data, providing the objective is to increase knowledge around the causes and cures of ill health.’

Support for personal data use in medical research is not without its conditions. **Consent and anonymity are often considered by the public to be two prerequisites for the use of their personal data.** Ipsos MORI (2007), for example, found that the public thinks that consent should always be sought to use their personal data for medical research. *Figure 5* shows that the public is relatively unimpressed by possible reasons that consent would not be sought, including consent being given in the past (29%), it being not practical to gain consent (20%) and it being expensive to gain consent (5%). 21 per cent of the public think there is no scenario in which it would be acceptable for personal data to be used without consent.

Figure 5: ‘When is consent not needed?’ (Ipsos MORI, 2007)

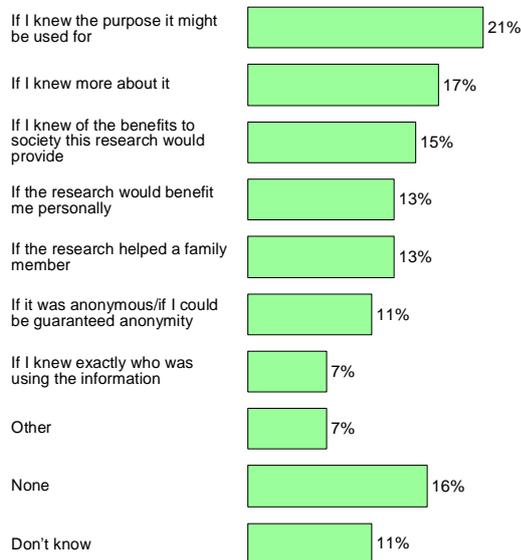


The public appears to take quite a principled approach to the question of consent, with the vast majority (79%) agreeing with the statement: ‘I have the right to be consulted about any use of my personal health information for research, even if that made research impractical’ (Ipsos MORI, 2007).

However, as shown in *figure 6*, there are a number of things that can be done to increase the public’s willingness to give their consent, including providing information about the purpose, the benefits and who will be using data.

Figure 6: 'What would encourage consent?' (Ipsos MORI, 2007)

Q What, if anything might make you more inclined to allow your personal health information to be used for the purposes of medical research?



Base: All not 'certain to' allow their personal health information to be used (1,807)

Source: Ipsos MORI

As well as information about the purpose of medical research and data use, information on how data use will be controlled also appears to have a positive effect on consent; Ipsos MORI (2007) found that:

- 'Assurance beforehand that the information they provide would probably be kept confidential prompts just over six in ten (62%) to say they would be certain or more likely to provide their information.
- However, when asked about the impact of knowing that their consent would not be sought for *further* research using their information, significantly fewer say they would be more likely or certain to participate (42%).
- Leaflets giving information about the project in advance would inspire half the general public to consider allowing their personal health information to be used, whilst websites would have a lesser effect (36% would be more likely).
- Just over half (56%) say that information about the risks and benefits of a research project would make them more likely or indeed certain to allow their information to be used.
- Six in ten would be more predisposed to allowing their personal health information to be used if they knew that the research it was intended for has the approval of an independent ethics committee.' (Ipsos MORI, 2007)

Conclusion

Despite considering personal medical information to be sensitive and private, a large majority of the public is in favour of its use for the purpose of medical research as long as some conditions are met. Where this is the case, the public is not particularly concerned about who conducts the research as long as the primary purpose is to benefit them, their family or the public. They are, however, very concerned about their health data being used for commercial purposes, and being accessed by private companies for private benefit.

While the public is generally happy for their data to be used for medical research, they are keen to maintain some ownership and control over its use, placing a particular importance on anonymity and consent. Providing citizens with information on both the purpose and governance of personal data use is important for gaining their trust and consent.

7. Appendix B: Summary of sources of information

Title	Type*	Produced/delivered by	Year	Outline
Dialogue on data: Exploring the public's views on using administrative data for research purposes	Public dialogue	Ipsos MORI (for ESRC and ONS)	2014b	Six sets of reconvened six-hour public dialogue workshops, each involving 16 to 20 participants, were conducted in London, Manchester, King's Lynn, Cardiff, Wrexham, Stirling and Belfast. A total of 129 participants attended the seven reconvened workshops.
Public attitudes to science 2014. 'Attitudes to Big Data' section.	Mixed methodology	Ipsos MORI	2014a	<p>Main face-to-face survey of 1,749 UK adults aged 16+, which was carried out from 15 July to 18 November 2013 using a probability sampling approach</p> <p>Booster face-to-face survey of 315 16-24 year-olds carried out over the same period using a quota sampling approach, so that the attitudes and behaviour of young adults could be compared and contrasted with those of the overall population</p> <p>Social listening, tracking how various science topics were discussed online.</p> <p>Four waves of online qualitative research with members of the Ipsos MORI Connects online community to explore in more depth the attitudes and behaviours of those who are already online</p> <p>Eight follow-up face-to-face observational interviews with members of the online community, observing how they sought out science-related information online</p> <p>Day of Discovery workshop with 106 members of the general public in London on 11 January 2014 to further explore issues raised by the</p>

				survey data.
Data Nation 2013: Balancing growth and responsibility	Survey	Deloitte (with research conducted by Ipsos MORI)	2013	National survey, conducted face-to-face, with a sample of 2,006 people aged 15 and older, representative of the British population. http://www.deloitte.com/assets/Dcom-UnitedKingdom/Local%20Assets/Documents/Market%20insights/Deloitte%20Analytics/uk-da-data-nation-2013.pdf
Summary Report of Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data	Focus groups and telephone interviews	The Wellcome Trust	2013	Six 90-minute focus groups and six 45-minute telephone interviews were conducted with 50 members of the public, including men and women, aged 18-70, from socio-economic groups ABC1 and C2DE
Public Understanding of Statistics	Survey	Ipsos MORI	2013	National survey, conducted online, with a sample of 1,034 adults aged 16-75 across Britain, with data weighted to match profile of the population http://www.ipsos-mori.com/Assets/Docs/Polls/rs-s-kings-ipsos-mori-trust-in-statistics-topline.pdf
Data Nation 2012: Our lives in data	Survey	Deloitte (with research conducted by Ipsos MORI)	2012	National survey, with a sample of 1,036 people aged 15 and older, representative of the British population https://www.deloitte.com/assets/Dcom-UnitedKingdom/Local%20Assets/Documents/Market%20insights/Deloitte%20Analytics/uk-mi-da-data-nation-2012.pdf
The Data Dialogue	Survey	Demos (with research conducted by Populus)	2012	National survey, conducted online, with a sample of 5,010 people aged 18 and over from across Britain, weighted to the profile of all adults aged 18 and over. http://www.demos.co.uk/files/The_Data_Dialogue.pdf?1347544233
Public Perception of	Secondary analysis of	Hallinan, D and	2012	Secondary analysis of surveys and academic literature

the data environment and information transactions: A selected survey analysis of the European Public's views on the data environment and data transactions	surveys	Friedwald, M		relating to the views of European citizens http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2374358
Data Privacy: What the consumer really thinks	Survey	Direct Marketing Association (with research conducted by Future Foundation)	2012	National survey, conducted online, with a sample of 1,020 people, representative of the UK population http://dma.org.uk/sites/default/files/toolkit_files/data_privacy_-_what_the_consumer_really_thinks_2012.pdf
Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union	Survey	Eurobarometer, European Commission	2011	European survey, conducted face-to-face, with a sample of 26,574 people aged 15 and over, from European countries. Results reported here are just for the UK. http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf
Data and Privacy: How concerned are citizens about data sharing in the public services?	Survey	Institute for Insight in the Public Services	2008	Survey, conducted via telephone, with a sample of 1,002 people. http://www.tns-bmrb.co.uk/assets-uploaded/documents/iips-insight-data-and-privacy_1279573299.pdf
Public service delivery: What the citizen expects	Survey	Institute for Insight in the Public Services	2007	Survey with sample size of over 2,300 people. http://www.tns-bmrb.co.uk/assets-uploaded/documents/public-service-delivery-what-the-citizen-expects_1283439916.pdf
The Use of Personal Health Information in Medical	Survey and in depth interviews	Ipsos MORI on the behalf of the Medical Research Council (MRC)	2007	Survey, conducted face-to-face, with sample of 2,100 people 15 and over, representative of Great Britain.

Research				http://www.ipsos-mori.com/Assets/Docs/Archive/Polls/mrc.pdf
Public and Professional attitudes to privacy and healthcare data	A review of literature	Peter Singleton, Nathan Lea, Archana Tapuria, and Dipak Kalra (Cambridge Health Informatics and the General Medical Council)	2007	Review of 51 papers considered relevant to public attitudes to privacy undertaken between 1996 and 2007 http://www.gmc-uk.org/GMC_Privacy_Attitudes_Final_Report_with_Addendum.pdf_34090707.pdf
Health on-line: public attitudes to data sharing in the NHS	Focus Groups	Scottish Consumer Council	2005	Eight focus groups in Scotland http://www.statewatch.org/news/2005/nov/scot-nhs-database.pdf
The Public View on Electronic Health Records	Focus groups and survey	NHS National Programme for Information Technology (NPfIT) and Health Which?	2003	Focus groups Survey, with a sample of nearly 2000 people
Share with Care! People's views on Consent and Confidentiality of Patient Information	In-depth interviews, focus groups and survey	NHS Information Authority (NHSIA), Consumers Association and Health Which?	2002	In-depth interviews and focus groups. Survey, conducted face-to-face, with a sample of 2,087 adults aged 15 and over in Great Britain, representative of the population.