Understanding Patient Data

involve

CarnegieUK TRUST

CHANGING MINDS • CHANGING LIVES

Better use of Data: Balancing Privacy and Public Benefit
Background briefing paper for workshop participants

Introduction

Government and other agencies providing public services increasingly collect, store and use personal data about citizens as part of the business of delivering services. Effective use of this data would support agencies to deliver better targeted and more efficient services in ways that stand to benefit the public. However, the sharing of, what can sometimes be, highly personal data also raises legitimate concerns about privacy and prompts questions about what the public see as acceptable uses of this type of data.

The upcoming Data Sharing and Public Benefit workshop is part of a wider project being undertaken by Involve, the Carnegie UK Trust and Understanding Patient Data, which seeks to build a better understanding of how different groups (government, civil society and advocacy groups) make sense of, and balance, the tensions and trade-offs inherent in data sharing.

The attraction of delivering public benefit through more effective public sector data sharing is often considered sufficient motivation to pursue a course of wider data sharing, notwithstanding the potential risks this may pose to individual privacy. The concept of 'public benefit' itself however is rarely interrogated and appears to mean different things to different stakeholders, and in different contexts.

Understanding how different groups define and value the public benefits that may be delivered by the better use of data, and making sense of where an acceptable balance between risks and benefits may lie, is therefore central to the goal of this project.

Additionally, these workshops are designed to explore whether, in trying to balance these risks, the tipping point of acceptability shifts depending on, for example:

- the type of information being shared e.g. financial, health or criminal records;
- the type of agencies it is shared with (e.g. public sector, private sector, voluntary sector);
- the context in which the data was originally collected?

And if it does shift, how?

The issues and debates surrounding data sharing cut across all aspects of public service delivery, however for these workshops we are focusing on data sharing across and between the housing, criminal justice, health and social care and welfare sectors. Across these areas of public service there is evidence of increasing demand for the sharing of personal data to support more effective multi-agency working at a local level. Further, in all of these fields, decisions about what data to share, when to share it, and who to share it with can not only create ethical dilemmas for professionals but also have potentially significant impacts for individuals. This makes data sharing across these sectors an ideal focus for the workshops.

Data Sharing between Public Service Providers

Increased data sharing across and between public service providers is a cornerstone of ambitions to modernise government and transform the way public services across the UK are delivered.

The ability to make easy data driven decisions is becoming vital to the way that we all live and work. This should be the way that government provides services.
(UK Government's Transformation Strategy 2017-2020)

The overarching rationale behind the drive to extend the way data is used in public service provision is that the better use of data has the potential to:

- produce direct benefits to individual service users by allowing more personalised and targeted services to be provided; and
- create wider public benefit by increasing the efficiency and effectiveness of public services overall.

Understanding Data Sharing

What is Data Sharing?

The process of data sharing is, fundamentally, the 'disclosure of data from one or more organisations to a third party organisation or organisations, or the sharing of data between different parts of an organisation.' (Information Commissioner's Office, 2011) This includes both systematic, routine forms of data sharing i.e. where the same data sets are shared between the same organisations for an established purpose, as well as exceptional, one-off processes of data sharing for different purposes.

Advances in information technology, alongside the increasing number of organisations involved in delivering public services, have led to a significant increases in the amount of data being produced across governments and other agencies in recent years. Moves towards e-government have also resulted in the 'automatic capture' of vast quantities of data by public authorities as part of their routine business, for example for the purposes of registration, financial or service transactions or record keeping. (Wilson et al. 2011) For the most part, however, data generated across government and public service providers, is held and used solely within the organisation or department that collected it. This creates a situation where there can be both duplication and, potentially, contradiction in the datasets agencies hold, and effective data management is an ongoing challenge.

These same developments in information management and communications technology have also given organisations the ability to link and process large amounts of additional data in ways that, if used effectually, can provide insight into how services can deliver better outcomes. It is, however, not always easy, for government and service providers to gain access to data held by other departments, particularly in a timely manner.

The legal context for data sharing between public agencies (and even between government departments) is highly complex and is continually evolving. It has tended to involve bespoke, bilateral legal gateways established for a specific purpose. The recently passed Digital Economy Act aims to streamline these processes for a limited number of areas of public service delivery.

The Act will create a more permissive environment for data sharing across and between agencies providing public services and support the better use of the information to inform policy, planning and service delivery.

Legal Context for Data Capture and Sharing

A public body may only share data if it has legal authority to do so. The first question that agencies wishing to share data need to ask therefore is whether they have the expressed or implied legal powers to perform a function necessitating data sharing. The power may be set out expressly in statute, or it may be implied from the body's other statutory powers and functions.

Until very recently, the legal power to share data has come from a variety of specific legislative 'gateways' by which information can be disclosed or received for particular purposes. Examples of such permissive statutory gateways include:

- section 115 of the *Crime and Disorder Act 1998*, allowing anyone to pass information to certain authorities if it is necessary or expedient for the purposes of any provision of the Act;
- section 14 of the *Offender Management Act 2007*, allowing data sharing between specified bodies for various purposes relating to offenders;
- section 111(1) of the *Local Government Act 1972*, providing that they "shall have power to do anything...which is calculated to facilitate, or is conducive or incidental to, the discharge of any of their statutory functions.";
- section 6 of the *Crime and Disorder Act 1998* which gave police and local authorities the implied power to share data to formulate and implement strategies for reduction of crime in their area; and
- section 25 A, B and C of the *Health and Social Care (Safety and Quality) Act 2015* which places a legal duty on health and adult social care organisations to share information when it will facilitate care for an individual. (UK Government 2015)

In May 2017, the Digital Economy Act was passed by Parliament to enable greater data access for defined public interest purposes by public authorities. Broadly defined, clause 30 of the Digital Economy Act contains provisions for a 'single gateway to enable public authorities, specified by regulation, to share personal information for tightly constrained reasons agreed by parliament, where its purpose is to improve the welfare of the individual in question. To use the gateway, the proposed sharing of information must be for the purpose of one of the specified objectives, which will be set out in regulations.' (UK Government 2017)

The Digital Economy Act therefore provides new legal mechanisms to allow data sharing between specified public sector bodies to support the better use of data for targeted interventions; improving the welfare of citizens; reducing debt owed to the public sector; fraud prevention; the sharing of civil registration information; and producing better statistics and research. These powers are to be regulated by codes of practice that have yet to be published.

Health data however is considered particularly sensitive and there are additional restrictions and conditions on its sharing, including the common law duty of confidentiality. The Digital Economy Act, for example, explicitly excludes the use of health data from its permissions for research purposes and health services are not currently included in the list of specified public bodies. The 2013 Caldicott Review of Information Governance established four legal bases for processing personal confidential health and social care information which meet the common law duty of confidentiality. These are: with consent, through statute, through a court order and 'when the processing can be shown to meet the 'public interest test', meaning the benefit to the public of processing the information outweighs the public good of maintaining trust in the confidentiality of services and the rights to privacy for the individual concerned.' (National Data Guardian 2013)

Once it has been established that the parties have the necessary powers to share data, the next step is to consider whether the proposed sharing is compatible with other legal provisions regulating the use of personal data. For example, data sharing by public authorities must also comply with the European Convention of Human Rights (now part of the UK domestic law as a result of the Human Rights Act 1998), and in particular Article 8, which provides: Everyone has the right to respect for his private and family life, his home and his correspondence. It also must comply with the requirements of the Data Protection Act 1998 (DPA).

From May 2018 the General Data Protection Regulation (GDPR) will apply in the UK, and the government has confirmed that leaving the EU will not affect the commencement of the GDPR. (Information Commissioners Office 2017) The GDPR will replace the Data Protection Directive 95/46/ec as the primary law regulating how personal data is protected and is intended to harmonize data privacy laws across Europe, to protect and empower all EU citizens' data privacy and to reshape the way organizations across the region approach data privacy.

Understanding Identifiability

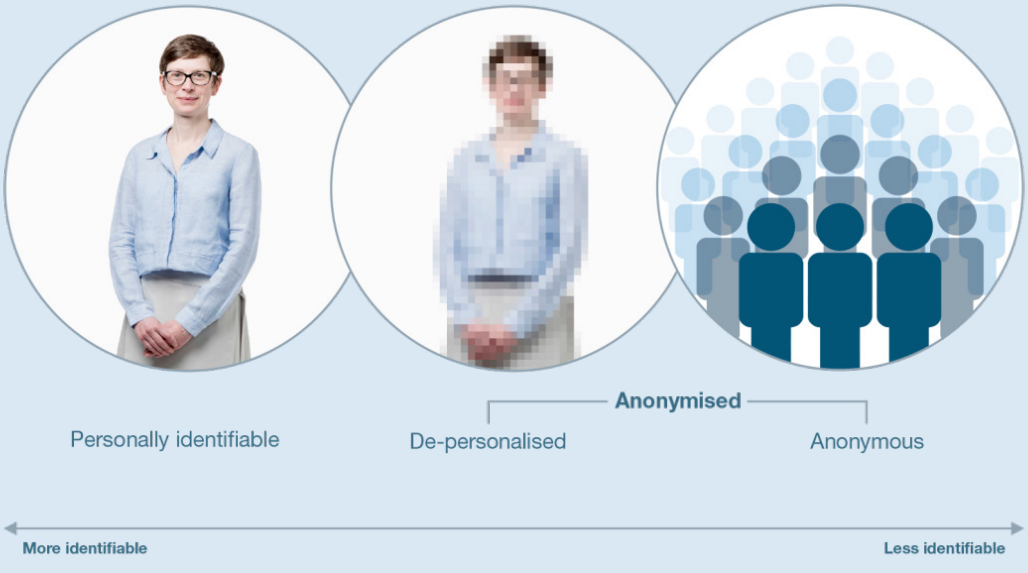
Generally people want to know whether they could be identified when data about them is used. The technical language of identifiability is complex. Many different words are used to describe the same thing, and many are unnecessarily technical (for example pseudonymised, key-coded, de-identified for limited disclosure). It is important to explain clearly what it means when information is 'anonymised' and what the likelihood of re-identification is when using different types of data

An important part of improving conversations is getting the words right, using words that are accurate but also accessible and meaningful. The words suggested here are the result of research commissioned by Understanding Patient Data, based on testing with the public and healthcare professionals, and will be used throughout the workshop.

(Understanding Patient Data 2017

Identifiability Demystified

Using pictures is the most helpful way to explain the concepts



Specturm of Identifiability

At one end of the spectrum, a person is fully identifiable. As you remove or encrypt information, you blur the image more and more, and it becomes more difficult to identify who that person is. At the other end of the spectrum, it is not possible to identify who someone is — they are effectively anonymous. Different controls are needed at different points along the spectrum depending on the risk of re-identification, which is why you may hear people talking about the 'environment' or 'context' in which data is used. The controls that are taken to protect the data are just as important as the data itself. It may also be possible to work out who someone is by joining together information from different sources — like joining together different pieces of a jigsaw puzzle.

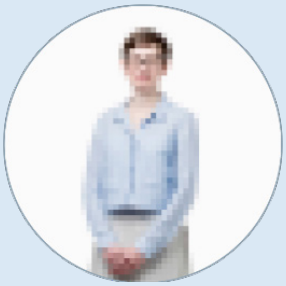
Personally Identifiable Data

This is information that identifies a specific person. Identifiers include: name, address, full postcode, date of birth or NHS number. Personally identifiable information will always be stored in a highly secure way. There are strict laws that safeguard how personally identifiable information can be used if you are not asked for consent. There are also sanctions under the Data Protection Act if personally identifiable data is misused.

Anoynmised Data

The Information Commissioner's Office gives guidance about what details must be removed or masked, and the safeguards that must be followed to anonymise data effectively. There are two different types of anonymised information: one individual-level, one grouped ('de-personalised' and 'anonymous'). It's important to distinguish between them, because the risks of re-identification are different, and therefore the data has to be protected in different ways.

De-Personalised Data



This is information that does not identify an individual, because identifiers have been removed or encrypted. However, it would in theory be possible to reverse that process and re-identify someone, so safeguards are still important. It is just like a blurred photo of someone. We can't immediately see who the person is, but we know it is a specific person. If we had the right computer power, and really needed to know who the person was, it might be possible to work it out.
There are strict safeguards on how de-personalised information can be used, because there is the potential that it might be possible to re-identify someone. The higher the possibility of re-identification, the greater the level of control needed.

Anoynmous Data

This is information from many people combined together, so that it would not be possible to identify an individual from the data. It may be presented as general trends or statistics. Because it would not be possible to identify someone, this information does not need special protection and can be published openly. Information about small groups or people with rare conditions could potentially allow someone to be identified and so would not be considered anonymous.

Benefits of Data Sharing

In a context where increased pressure is being placed on limited public sector resources data sharing across and between public agencies is increasingly being seen as a vital tool for enabling more joined up working.

'Data is a critical resource for enabling more efficient, effective government and public services that respond to citizen's needs. Data acts as the foundation upon which everything else rests.' (Cabinet Office 2017)

The sharing of personal data between public service agencies - in anonymised (whether de-personalised or aggregate) or personally identifiable forms - is not, however, a new concept. It is already commonplace for the personal information governments and other agencies hold about individuals to be used to compile statistics. Anonymised datasets like this are used and published in an aggregated form that makes it impossible to identify individuals. They are vital tools for identifying trends, informing evidence based policy and decision making and enabling a diverse array of social, clinical and economic research. For example:

Age UK wanted to better understand the prevalence of loneliness among people aged 65. They wanted to know what makes older people at risk of being lonely and which neighbourhoods had the highest risk order to help them plan their interventions. To achieve this they commissioned a study that used statistics generated from the English Longitudinal Study on Ageing (ELSA) survey, alongside data from the last census, to identify the particular characteristics that increase the prevalence of loneliness among older people. The 'loneliness heat maps' generated by this research have helped to identify hotspots for the risk to loneliness across England and enabled Age UK to develop tailor-made interventions to address loneliness in these areas. (Age UK 2015)

There are also many cases where personally identifiable information about individuals is routinely shared between public service providers as part of their day-to-day functions, for example between a local authority and the Department for Work and Pensions (DWP) to allow a pensioner's application for housing benefit to be processed, or between a hospital and a GP to ensure continuity of care after discharge.

Improved digital integration across governments already means that data held about individuals is progressively being connected across different sections of government in order to drive efficiency. For example, effective data sharing can reduce the need for citizens and government to update address data multiple times when, for example, applying for a passport, a driving licence and a Blue Badge. Therefore, as well as sharing, linking data is seen as an important means for achieving more efficient and responsive public services and, as the example below shows both personally identifiable and de-personalised data can be linked for a variety of purposes.

Health and social care services in East and North Hertfordshire (including hospitals and GPs) have linked patient data to:

- *better understand the needs of their local population;*
- *facilitate data-enabled decisions;*
- *improve the process of identifying individuals at risk or in need of a specific intervention.*

Having access to this linked data has, for example, enabled them to undertake an impact analysis of their re-ablement service. By using de-personalised data with, identifiers such as name, NHS number, and full postcode were coded, rather than removed altogether, when an individual is identified as being at risk or in need of a specific intervention, the relevant health and care professional involved in the care of the patient are able to re-identify the individual and make the necessary intervention. (National Data Guardian for Health and Care 2016)

Data Linking involves bringing together identifiable information from two or more data sets, from two or more sources. This enables information a specific individual or an event to be linked and used in ways that were not possible using any single set of records separately. The resulting data set is usually de-personalised for the end user (and potentially for the intermediaries if a trusted third party sharing system is used) because individual identifiers have been removed or encrypted.

For example, to assess the impact of living in damp housing conditions on rates of illness and long term health an independent researcher could link NHS health records (Dataset A) with information held by HMRC's Valuation Office about housing quality (Dataset B):

i. First, an NHS analyst would attach a unique serial number to each record. They would then separate the identifiable details (for example name, address, date of birth) from the rest of the information held by the NHS in Data set A but retain the serial number. This information (Dataset C) is then sent to HMRC.
ii. A HMRC analyst then uses the identifiable information in Dataset C to link the serial number to information they hold in Dataset B, before deleting the identifiable details. This leaves a file of unique serial numbers and housing condition data. This file is then sent to the researcher as Data Set D.
iii. The NHS analyst then provides the researcher with de-personalised health records coded using the same unique serial number (Dataset E).
iv. The researcher can now merge the housing information held in Dataset D with the health information held in Dataset E using the unique serial number. The unique serial number is then deleted and a new random one added.
v. This de-personalised data (Dataset F) can then be analysed, trends identified and conclusions drawn.

When a process like this is used to facilitate the linking process the risk of breaching individual privacy is reduced, as none of the three actors involved in the sharing arrangement – the NHS, HMRC or the researcher - has access to both housing and health records with personal identifiers attached at any time. However, because the information remains about an individual person, there is a risk that the data may lead for re-identification. This risk increases with the number of data sets containing personal information that are linked

In a context where digital technology is increasingly being positioned as a tool to transform public service provision, and data is being viewed as an untapped resource, there is now an increased focus on linking data between different sources to maximise its value. However, this means there is also increasing interest in linking and sharing personally identifiable data held by different departments, sometimes for purposes other than it was explicitly collected for. The example below demonstrates how one of the existing legal gateways allows a council to use data collected for different purposes, by different departments, for new purposes without contravening the Data Protection Act.

Camden Council created the Camden Residents Index to streamline processes and provide a more responsive service to citizens. This index brings together data from 16 council business systems, covering 123 fields of primarily demographic information, to create a complete picture of each resident. The Camden Residents Index has helped the Council reduce administrative costs by eliminating duplication, identify instances of fraud and error, and ultimately provide a more seamless and efficient service to residents. (Symons 2016)

High quality, accurate and timely data clearly has a role to play in helping to inform the decisions made by government and public service providers, at both national and local levels. However, there may be scope to better use the data available. Ambitions for the wider and more systematic use of data include:

- **Enabling more tailored public service delivery** by ensuring that those delivering frontline services have the information required to offer the right services, to the right users, at the right time. This is particularly relevant when users with multiple and complex needs may be interacting with many different parts of the system independently. Effective data sharing can support coordinated interventions between agencies and also enable efficiency savings by reducing duplication.

The UK Government's Troubled Families programme seeks establish a new, co-ordinated way of supporting families with multiple problems, and who may be dealing with multiple service providers across the health, housing, criminal justice, anti-social behaviour, welfare and education sectors individually. The programme therefore incentivises services to come together, sharing the information they held independently, to better understand the needs of the whole family (instead of constantly reacting to individual problems) and coordinate support accordingly.

The rationale is that there are both individual and service level benefits to this approach as:

- *those reliant on public services are not required to repeat their information multiple times and spend significant energy and time meeting bureaucratic requirements;*
- *clients will no longer receive reactive, disjointed and potentially overlapping services from the range of agencies they deal with;*
- *efficiency savings can be made by co-ordinating interventions designed to achieve long term change by addressing the root cause of the problem. (Cabinet Office 2016)*

- **Better targeting of public services** by identifying areas of unmet need or isolating the inter-related factors underpinning complex social problems. This can be used to indicate where (and how) resources can be allocated to have the greatest impact and enable agencies to better define and solve problems that no single organisation or jurisdiction is able to address.

In Hackney de-personalised information from A&E attendances following an assault (including time, date, location and means of assault) are shared with the Community Safety Partnership to contribute towards violence prevention measures in the area. The key purpose was to identify 'hot-spots' for violent crime and allow for preventative policing. (Ford et al 2014)

- **Outcome monitoring** in order to better understand the impact of policy interventions on individuals and/or different sections of society. This would enable decisions about which services to run, or which policies to implement, to be based on evidence of 'what works', improving the overall quality and effectiveness of public service delivery

Kent have created one of the largest integrated data sets in health and social care (known as the KID) which they use to run matched cohort analysis, assessing the impact of services by comparing outcomes of service recipients with statistically similar people who don't receive the service. Examples of use include:

- *Evaluating the impact of the c.10,000 home safety visits carried out by Kent Fire and Rescue Service on A&E attendances originating from the home;*
- *Evaluating a pilot of a reconfigured GP practice, with additional services for people with long-term health conditions, for its impact on acute care usage;*
- *An interrupted time-series analysis of a falls prevention service.*

This analysis has enabled Kent to identify services which are highly impactful and support a business case for expansion which allows more people to benefit from access. It is also allows them to identify the types of people a programme is most useful for, and for whom it has limited impact, helping prioritise which groups should receive certain interventions or services. (Symons 2016)

Challenges of Data Sharing

In order to achieve the types of benefits outlined in the previous section, the information that agencies would need to share is likely to include personal data held about service users and/or the wider public. This information goes way beyond basic contact information. It could include information supplied for the purposes of claiming benefits, personal medical or financial information, or records demonstrating eligibility for support services like counselling, or housing assistance and so on. In short, information that many people would consider, and want to keep, private. Data sharing of this kind therefore raises profound questions about the ability of the individual to manage the way data about them is used, who has access to it, and how their right to privacy is protected.

Defining Personal Data

The Data Protection Act (1998) defines personal data as being 'data which relate to a living individual who can be identified from those data, or from those data and other information which are in the possession of, or are likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.'

The Act further notes that there are some forms of personal data that are likely to be of a private nature are additionally sensitive because information about these matters could be used in a discriminatory way. Sensitive personal data is taken to include information related to racial or ethnic origin, political opinions, religious beliefs, trade union membership, physical or mental health, sexuality and criminal justice history, and needs to be treated with greater care than other personal data. There are also additional protections included within the Act to ensure that sensitive personal data is processed and stored securely and its use is controlled.

The General Data Protection Act, which comes into force in May 2018, provides a more detailed and expansive definition of personal data. and makes it clear that information such as online identifiers e.g. an IP address – can be personal data, reflecting changes in technology and the way organisations collect information about people.

There are also a range of practical, cultural and ethical barriers that hinder effective data sharing across and between public service providers. On an operational level these include:

- Legacy data systems that make sharing difficult;
- Differing organisational cultures and professional codes of ethics;
- Lack of confidence in the systems that are facilitating the sharing;
- Incomplete, or not updated, data sets;
- Uncertainty about whether the information sharing is necessary;
- Misconceptions around what can and cannot be shared;
- Concerns about the trustworthiness or professionalism of other parties involved in sharing arrangements; and
- Doubt about the quality and reliability of the data collected by others. (Bellamy et al 2005, Wilson et al 2011, Wilson and Grey 2015)

Although resolving some of these technical issues is out with the scope of this project, it is important to recognise that this range of both real and perceived barriers can shape attitudes towards data sharing amongst public sector organisations. They also give rise to legitimate concerns from the public about the capacity of agencies to safely store, process and share their personal data.

Data Sharing and Privacy

Individual privacy is defined by Privacy International as the conditions that enable individuals 'to create barriers and manage boundaries to protect ourselves from unwarranted interference in our lives, which allows us to negotiate who we are and how we want to interact with the world around us. Privacy helps us establish boundaries to limit who has access to our bodies, places and things, as well as our communications and our information.' (Privacy International 2017)

Data Sharing therefore has intrinsic implications for individuals' rights to privacy. A study into public attitudes to data sharing found that one of the top reasons (32%) for the public opposing the increased use of the data held about them was that: 'People have a right to privacy'. (Ipsos Mori 2014a) The use of personal data held by agencies can also be of particular concern to people when data is used for purposes aside from that for which it was initially provided and/or without the explicit consent of the data subject.

It follows then that debates around public sector data sharing and privacy are generally concerned with what is acceptable practice in relation to accessing and disclosing personal and/or sensitive information about a person, for what purposes should this be allowed, and the safeguards necessary to minimise risks to privacy.

The Use of Personal Information

Previous research exploring public attitudes towards public sector data sharing suggests that a key concern for many people about wider data sharing is whether the information will be personally identifiable. (Aitken 2011, Wellcome 2013, Davidson et al 2013). In most cases it seems that people intuitively understood this to mean whether their name, address or another unique identifier like NHS number or National Insurance number would be disclosed. Using the definition of personal data included in the Data Protection Act (page 13) however, it is clear that the identifiable personal data held by governments and public services can take many forms, not only the obviously identifiable fields of name, address, biometric information or national insurance number. It also encompasses information like date of birth, health records, ethnic background, disability, criminal record, income or credit history

which although it would not identify an individual in isolation, could render an individual identifiable if it was part of a collection of data held by an agency, or linked with data held by another agency.

Further, recent studies of public views on data use suggest that the public's understanding of what constitutes personal data is not fixed; it varies from person to person and is changing over time. For example, a Europe wide poll in 2011 found that over 70% of UK respondents considered data such as financial information, medical information, passport number, fingerprints and address to be personal, while less than half of respondents considered the websites you visit, your tastes, opinions and things you do to be personal information (Eurobarometer 2011). By 2013 however similar levels of opposition were expressed over the sharing of data held by internet providers and social media companies, with this data now being too regarded as too personal to be shared without appropriate considerations for sensitivity and privacy. (Davidson et al 2013) That said, the general view appears to be that data about *who you are* is generally considered by most people as more personal than *data about what you do*. (Sciencewise 2014)

It is also clear from previous research that there is no single *public view* about acceptability in relation to how personal data is shared. An individual's age and social class both appear to have some bearing on their views on data: with younger generations typically being happy to share more, but being less aware of the implications, and older generations sharing less but being more aware. More affluent or educated social groups are also typically more comfortable with sharing their personal data than those experiencing higher levels of deprivation. (Suherman-Bailey 2015, Wellcome Trust, 2013).

There also appears to be significant differences between the public's stated preferences and their actual behaviours. When asked, the public have generally expressed opposition to greater amounts of data about them being collected and used by government and other agencies. (Sciencewise 2014) However, numerous studies have revealed that most people are willing to make a number of privacy 'trade offs' to access particular benefites: for example, '71% of consumers would provide more information online if it helped them save money; 60% would be willing if the resultant service was better tailored to their needs; and 56% would be willing if it helped them to make better decisions.' (Coll 2015) How people view the use of personal information about them therefore must be considered within a context that questions not only what information is shared, but also for what purpose and what benefit it provides.

The Right to Privacy

Privacy is a qualified, fundamental human right. The right to privacy is articulated in all of the major international and regional human rights instruments, including:

United Nations Declaration of Human Rights (UDHR) 1948, Article 12: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

Individual privacy is, in the context of data sharing debates, fundamentally the ability of individuals to choose when they wish to disclose personal information about themselves, and who they want to disclose this information to. Individual privacy therefore can be threatened or breached through a number of practices associated with data sharing, each of which has the ability to produce a different form of harm. The list below, adapted from Solove's A Taxonomy of Privacy (2006), itemises the types of harm that different aspects of the process and outcomes of data sharing can have on individual privacy:

- **Information processing**
  - Aggregation - The combination of various pieces of data about a person
  - Identification - Linking information to particular individuals
  - Insecurity - Carelessness in protecting stored information from leaks and improper access
  - Secondary use - Use of information collected for one purpose for a different purpose without the data subject's consent
  - Exclusion - Failure to allow the data subject to know about the data that others have about her and participate in its handling and use, including being barred from being able to access and correct errors in that data
- **Information dissemination**
  - Breach of confidentiality - Breaking a promise to keep a person's information confidential
  - Disclosure - Revelation of information about a person that impacts the way others judge her character
  - Increased accessibility - Amplifying the accessibility of information
  - Distortion - Dissemination of false or misleading information about individuals
- **Invasion**
  - Intrusion - Invasive acts that disturb one's tranquillity or solitude
  - Decisional interference - Incursion into the data subject's decisions regarding her private affairs

Further, unlike most other rights and individual's privacy can be compromised without them necessarily being aware that is it taking place: 'With other rights, you are aware of the interference -- being detained, censored, or restrained. With other rights, you are also aware of the transgressor -- the detaining official, the censor, or the police.' (Privacy International 2017)

The Data Protection Act (DPA) gives individuals certain specific rights over their personal data. These include:

- the right to access personal data held about them;
- the right to know how their data is being used; and
- the right to object to the way their data is being used.

Individuals can object when the use of their personal data is causing them 'substantial, unwarranted damage or substantial, unwarranted distress'. The objection can be to a particular use of information or to the fact an organisation is holding their personal data at all. Organisations are required by law to respond to individuals who object in writing to the way their personal data is being used. 'However they do not need to comply with the request unless there is damage or distress and this is substantial and unwarranted.' (Information Commissioner's Office 2011)

Privacy Concerns

Whenever personal data is collected, accessed, analysed, shared or linked there is some risk to individual privacy. Privacy activists argue that any use of personal data 'should not lead to a widespread intrusion on people's privacy; should be proportionate, limited in scope and enshrine fundamental rights; and carry strong safeguards against wilful abuse and unintended consequences.' (Ruiz 2016) A key challenge therefore, in determining whether the sharing of personal data held by governments and public agencies is an appropriate use of this information, is to ensure that any risks to people's individual privacy from the sharing of personal data are proportionate to the benefits delivered.

The main threats to privacy from data sharing can be summarised as:

- the risk of data loss (through accident or malice) [1]
- statistical disclosure (the potential to identify an individual within a dataset by their unique or rare combination of characteristics)
- the potentially negative impacts of secondary usage of data (through the disclosure or linking of information about a person that they would prefer to have remained private in a given context).

Given the complex set of factors determining what is understood as personal data, information sharing across and between public service providers needs to be undertaken within a context that ensures there is greater public awareness about how data is being used, safeguards in place to protect individual privacy and clear routes for redress if individuals feel their privacy has been compromised.

A further concern is that sharing and linking data for targeting public services can result in the production of generalisations that categorised individuals, social groups or geographic areas in ways that could result in discriminatory treatment and/or stigma. (ORG 2016) This concern was also highlighted in public engagement research that showed people were concerned that 'generalising about groups may lead to policies and/or interventions or treatments which did not adequately consider individual circumstances and needs.' (Aitken 2011) Cases where these types of concerns could be brought into play are, for example the UK Government's Troubled Families programme (outlined on page x), or the example below.

Crime reduction and community safety are held to be not only matters for the police, but depend on the policies, interventions and intelligence of other public services providers. Community Safety Partnerships, for example, pool and exchange data on a wide range of criminal and anti-social behaviours, and on people who engage, or who might engage, in it. Some data are required for analytic and crime auditing purposes and can be exchanged in de-personalised, aggregated form. This allows for intelligence-led analysis, using large datasets for crime mapping, and pinpointing specific areas and offender groups to be targeted for crime reducing initiatives.

Other data however may be sourced from individual case records kept by the police, the probation service, the courts, social services or health authorities and may be easily traceable to identified individuals or households. Further, to carry out their functions, these Partnerships increasingly rely on risk assessment tools drawing on data supplied by a range of agencies, especially those dealing with mental health, social care and social housing. (Bellamy et al 2005a)

1 The question of whether data can, in practise, be securely stored or not has major ramifications for whether agencies and individuals are supportive of data sharing. While overcoming technical challenges which limit people's confidence in the security of information sharing technologies is out with the scope of our project, the impact this has on attitudes to data sharing is real and has to be addressed if the ambitions for wider data sharing are to be realised.

Protecting Individual Privacy

The goal of ensuring that data sharing practices protect an individual's right to privacy is not starting from a blank canvas. The UK is a signatory to the European Convention of Human Rights, so have incorporated the right to privacy into national laws. However, as already noted in the outline of the legal context for data sharing, the right to privacy is not an absolute right, and public authorities are permitted to share information without consent if there are lawful gateways and clear and proportionate reasons for sharing. The Data Protection Principles set out in Schedule 1 to the Data Protection Act establish that the sharing or linking of personal data needs to be regulated and managed in ways that ensure that the powers are necessary (i.e. that there needs to be a specific purposes for which data is disclosed) and that the use of data is proportionate (i.e. that the minimum amount and type of data necessary is used). These the Digital Economy Act complies with these principles.

Data Protection Principles

The principal legislative provision relating to data protection is the Data Protection Act 1998 (DPA), which implements the Data Protection Directive 95/46/EC. The DPA gives individuals a number of rights to ensure that personal information covered by the Act is processed lawfully. It regulates the manner in which such information can be collected, used and stored, and so is of prime importance in the context of data sharing.

The eight data protection principles established in the DPA set out the obligations that organisation who hold personal data (data controllers) have to the person the data relates to (data subjects) and form the core of data protection regulation. These principles stipulate that data relating to living individuals must:

1. be obtained and processed fairly and lawfully, and subject to general conditions, either under consent, or under necessity for a limited number of general reasons;
2. be used only for the specified and lawful purpose(s) for which the data are collected and used;
3. be adequate, relevant and not excessive for the specified purpose(s);
4. be accurate and, where necessary, kept up to date;
5. be retained no longer than necessary for the specified purpose(s);
6. be processed in ways that respect the data subject's rights, include the right of subject access (the right of the individual to see information held about him or her);
7. be subject to appropriate technical and organizational measures to prevent unauthorized and unlawful processing, accidental loss of, destruction of, or damage to the information; and
8. not be transferred outside the European Economic Area, except to countries where levels of data protection are deemed adequate.

The General Data Protection Regulation (GDPR), which will apply in the UK from May 2018, sets out more stringent principles for government regarding the control and use of data for the public interest. A new accountability requirement will also be added that compels data controllers to demonstrate how they are complying with the principles.

In practice, data protection regulators and privacy experts tend to interpret the notion of 'proportionality' to imply that 'the risk must be proximal and specific: that is, the presumption of privacy and confidentiality can only be overruled if there is a real and present risk of determinate harm.' (Bellamy et al 2005). On this basis it has therefore been assessed as legitimate to: share personally identifiable information to, for example, control outbreaks of contagious diseases in the interests of public health; justify instances of telecommunications surveillance in the interests of national security; and allow public access to information held on the child sex offenders' register. This sharing of what would

otherwise be considered private information in these cases has, in these cases, been broadly accepted by the public as justified in the interest of what we might consider the 'greater good'.

But assessments of justified purpose and proportionality are not always clear cut and similar value based ethical assessments could also be made for a wide range of data sharing practices. The Data Protection Act can allow government departments to use data for purposes other than for which they were originally collected where this is determined as 'fair'. Fairness, however, will depend on the circumstances and should take into account, amongst other things, 'what the individual could reasonably have expected when s/he provided the information and whether the proposed use would adversely affect him/her.' (Suherman-Bailey 2015)

For example, in the interest of preventing a real and present risk of harm, a social housing provider could decide to inform a public utility companies that a vulnerable elderly client with dementia is in financial difficulties, in order to avoid the negative consequences that would result from the withdrawal of the power supply to that person's home. Similarly, assessments have been made that the incursion into an individual's privacy is justified in order to minimise the risks associated with essential services being unnecessarily diverted.

In order to reduce the number of unnecessary calls made to blue light services in Braunstone emergency service providers, the local authority and health services plan to data share information about repeat callers. The rationale is that those who repeatedly use emergency services are likely to be facing similar underlying difficulties which may be better served with a social care package or mental health support, freeing up the blue light services to respond to emergency calls instead. (LGA 2016)

It is grey areas like this, which rely on subjective assessments of whether the benefits outweigh the risks, or the interests of the many justify an intrusion into individual privacy, which mean that privacy concerns remain at the heart of debates about data sharing.

Balancing the Benefits and Risks of Data Sharing

As noted in the earlier section of this paper, the increased sharing of personal data between public service providers is generally promoted and justified on the basis that it can improve outcomes for citizens through the provision of better services. For those keen to realise the ambitions of greater and wider data sharing, the key challenge seems to be finding a way to allow governments and public service providers the flexibility to better use the information they hold in order to plan and provide responsive services, whilst ensuring that processes are constrained to purposes that offer clear benefits to the public and respect individual's rights to privacy. However this approach also creates a risk that data sharing is simply seen as 'the solution', with the legal constraints and social and ethical concerns the practice raises something that needs to be 'managed', instead of constructive and necessary features of the process. (Ruiz, 2016)

It is clear that there are tensions between extending data sharing and enhancing privacy, and it cannot necessarily be assumed that these are compatible goals. While in an ideal scenario it would be possible to strike a balance that provides sufficient safeguards to allow both objectives to be pursued with limited compromises to either, in reality this may not be the case. Instead finding an acceptable public settlement may need to involve consenting to reductions in the quality of service delivery in some cases; or a greater than desirable intrusion into privacy in others. For the public to begin to grant agencies a greater 'social licence'[2] to use the data held about them

2 Social Licence is a term emerging from debates about data sharing that are taking place in New Zealand and rests on the assumption that 'When people trust that their data will be used as they have agreed, and accept that enough value will be created, they are likely to be more comfortable with its use.' (Data Futures Partnership 2015)

more widely, there first needs to be a clearer understanding, shared across all stakeholders, of what constitutes ' public benefit' and acceptable, 'beneficial uses'.

While research tends to suggest that the public are much more likely to accept or support data sharing if there is a public benefit, or at least the potential for public benefit (Ipsos Mori 2016, Aitken et al 2016, Davidson et al 2012, Sciencewise 2014, Ipsos Mori 2014), there has to date however been little examination of how either the public, or those involved in delivering public services, understand and evaluate the idea of the wider public benefits promised by systematic data sharing.

In reviewing a range of literature to prepare this report the term 'public benefit' was extensively used to describe the purpose of data sharing in a wide range of contexts, but it was rarely defined or discussed any further. The only formal definition found came from Statistics and Registration Services Act (s7(2)) which states that 'public benefit includes in particular (a) informing the public about social and economic matters, and (b) assisting in the development and evaluation of public policy'. (UK Government 2007) Elsewhere formulations of public benefit were variously presented as being 'in the public interest', 'the greater good', 'ethically sound', 'having a defined and measurable outcome', of 'real-world value or practical application' and as 'leading to the improvement of health, education or economic and social well-being'. (Cabinet Office 2016, Aitken et al 2016, Davidson et al 2012, Open Policy Making 2016)

It must also be acknowledged that 'what constitutes a benefit for a particular individual or community will depend on the circumstance, the needs, the values, and the cultural priorities and expectations of that individual or that community'. (Sheremata and Knoppers 2006 cited in Davidson et al 2013) It is also the case that in some instances the costs and benefits are not equally distributed across society. Therefore, without further investigation, the use of 'public benefit' as a blanket term to legitimise the wider use of data sharing in public service delivery, masks many of the real concerns the public and other stakeholders have about risks to privacy and the acceptability of different uses of data.

Moving Forward with the Debate

The examples provided throughout this paper highlight that the considerations involved in determining the acceptability of data sharing are necessarily complex, varied and context specific. While the overall rationale for data sharing may be arguably beneficial, the realities of practice open up a range of questions about appropriate purposes, the type of data shared and the relative privacy incursions that may be needed to achieve different types of benefits. For example:

- How do people's attitudes to data sharing vary depending on the type of personal information shared (e.g. financial, health or criminal records)?
- How important is anonymisation? And in what circumstances?
- How do people understand and define 'public benefit' in different circumstances?
- Do people balance the risks and benefits differently whether data is shared to deliver direct benefits to individuals or for wider, potentially less visible, public benefits?
- How important is the context in which the data was originally collected in determining what is seen as acceptable uses?
- Do attitudes change depending on the type of organisation or sector the information is shared with (including the voluntary sector)?

Determining what constitutes an acceptable settlement between the proportionate use of data to deliver services that benefit the public, and protecting people's privacy is a key challenge for policy makers, frontline staff, advocacy groups and the public at large if the ambitions for data sharing are to be realised. The key purpose of this wider engagement project, and the focus of the upcoming workshop, is to open up these issues for wider investigation and discussion

Bibliography

Age UK (2015) Age UK loneliness maps http://www.ageuk.org.uk/professional-resources-home/research/loneliness/loneliness-maps/ (accessed 28/5/2017)

Aitken, M. (2011) SHIP Public Engagement: Summary of Focus Group Findings. The Scottish Health Informatics Programme http://www.scot-ship.ac.uk/sites/default/files/Reports/Focus_Group_Findings_Briefing_Paper.pdf

Aitken, M., de St. Jorre, J., Pagliari, C., Jepson, R. and Cunningham-Burley, S. (2016) Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Medical Ethics 17:73 https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-016-0153-x

Bellamy, C., 6, P., Raab, C., Heeney, C. (2005) Data sharing and personal privacy in contemporary public services: the social dynamics of ethical decision making. Paper presented to Working Group 6 at: Annual Conference of the European Group of Public Administration (EGPA), University of Berne, Switzerland.

Bellamy, C., 6, P., Raab, C., Heeney, C. (2005a) Joined up government and privacy in the United Kingdom: Managing tensions between data protection and social policy: Part 2. Public Administration Vol 83 No 2

Cabinet Office (2016) Better Use of Data – Consultation Paper. UK Government https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf

Cabinet Office (2017) Government Transformation Strategy. UK Government https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-2020 (accessed 28/5/2017)

Coll, L. (2015) Personal Data Empowerment: Time for a Fairer Data Deal, Citizens Advice; https://www.citizensadvice.org.uk/about-us/policy/policy-research-topics/consumer-policy-research/consumer-policy-research/personal-data-empowerment-time-for-a-fairer-deal/

Data Futures Partnership (2015) What is Social Licence? http://datafutures.co.nz/our-work-2/talking-to-new-zealanders/social-licence/ (accessed 2/6/2017)

Davidson, S. McLean, C., Cunningham-Burley, S., and Pagliari, C. (2012) Public Acceptance of Cross-Sectoral Data Linkages, Scottish Government http://www.gov.scot/Publications/2012/08/9455/0

Davidson, S., McLean, C., Treanor, S., Cunningham-Burley, S., Laurie, G., and Pagliari, C. and Sethi, N. (2013) Public Acceptability of Data Sharing Between Public, Private and Third Sectors for Research Purposes, Scottish Government http://www.gov.scot/Publications/2013/10/1304/0

Eurobarometer (2011) Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union. European Commission http://ec.europa.eu/public_opi nion/archives/ebs/ebs_359_en .pdf

Ford, K., Wood, S. Hughes, K. and Quigg, Z. (2014) Violence profile: Hackney Use of NHS data in local violence prevention Centre for Public Health http://www.cph.org.uk/wp-content/uploads/2014/12/Hackney-violence-profile.pdf

Information Commissioner's Office (2011) Data Sharing Code of Practice. https://ico.org.uk/media/for-organisations/documents/1068/data_sharing_code_of_practice.pdf

Information Commissioner's Office (2017) Overview of the General Data Protection Regulation (GDPR) https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/

Ipsos Mori (2016) The one-way mirror: public attitudes to commercial access to health data. Wellcome Trust https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf

Ipsos MORI (2016a) Public dialogue on the ethics of data science in government. Government Data Science Partnership, and Sciencewise http://www.sciencewise-erc.org.uk/cms/assets/Uploads/data-science-ethics-in-government.pdf , accessed 25 May 2017

Ipsos Mori (2014) Dialogue on data: Exploring the public's views on using administrative data for research purposes. Economic and Social Research Council (ESRC).

Ipsos Mori (2014a) Public attitudes to science. Department for Business, Innovation and Skills (BIS) and the Economic and Social Research Council (ESRC). https://www.ipsos.com/ipsos-mori/en-uk/public-attitudes-science-2014

Local Government Association (2016) Data experts grants support better use of local data mini projects http://about.esd.org.uk/news/lga-data-experts-grants-support-better-use-local-data-mini-projects

National Data Guardian (2013) Information: To Share Or Not To Share? The Information Governance Review https://www.gov.uk/government/publications/the-information-governance-review

National Data Guardian for Health and Care (2016) Review of Data Security, Consent and Opt-Outs https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF

NHS (2014) Information Governance Toolkit https://www.igt.hscic.gov.uk/Caldicott2Principles.aspx (accessed 28/5/2017)

Open Policy Making Process (2016) Conclusions of civil society and public sector policy discussions on data use in government http://www.datasharing.org.uk/conclusions/index.html

Open Rights Group (2016) Consultation Response Data Sharing https://www.openrightsgroup.org/ourwork/reports/orgs-response-to-data-sharing-consultation (accessed 28/5/2017)

Privacy International (2017) What is Privacy? https://www.privacyinternational.org/node/54 (accessed 28/5/2017)

Ruiz, J (2016) Government announces new data sharing legislation in Queen's Speech, Open Rights Group https://www.openrightsgroup.org/blog/2016/government-announces-new-data-sharing-legislation-in-queens-speech (accessed 28/5/2017)

Sciencewise (2014) Big Data Public views on the collection, sharing and use of personal data by government and companies Sciencewise http://www.sciencewise-erc.org.uk/cms/assets/Uploads/SocialIntelligenceBigData.pdf

Solove, D.J. (2006) A Taxonomy of Privacy. University of Pennsylvania Law Review 154(3): 477-560.

Suherman-Bailey, J. (2015) Data policy and the public: shaping a deeper conversation. Sciencewise http://www.sciencewise-erc.org.uk/cms/assets/Uploads/Data-policy-and-the-publicJan-2015.pdf

Symons, T (2016) Wise Council: insights from the cutting edge of data-driven local government Nesta and LGA http://www.nesta.org.uk/sites/default/files/wise_council.pdf

UK Government (2007) Statistics and Registration Service Act http://www.legislation.gov.uk/ukpga/2007/18/contents (accessed 28/5/2017)

UK Government (2015) Public Sector Data Sharing: Guidance on the Law http://webarchive.nationalarchives.gov.uk/20150730125042/http://www.justice.gov.uk/downloads/information-access-rights/data-sharing/annex-h-data-sharing.pdf

UK Government (2017) Digital Economy Act http://www.legislation.gov.uk/ukpga/2017/30/part/5/enacted (accessed 28/5/2017)

Understanding Patient Data (2017) What are the best words to use when talking about data https://understandingpatientdata.org.uk/what-are-best-words-use-when-talking-about-data

Wellcome Trust (2013) Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data. Wellcome Trust

Wilson, R., Cornford, J., Baines, S., Mawson, J. (2011) New Development: Information for Localism? Policy Sensemaking for Local Governance. Public Money & Management 2011, 31(4), 295-300. http://eprint.ncl.ac.uk/file_store/production/168469/7C270E73-CDF2-4AE2-8AC7-6298B6A4A073.pdf

Wilson, R. and Grey, A. (2015) Information Sharing: Easy to say harder to do. Centre of Excellence for Information Sharing http://informationsharing.org.uk/our-work/academic-thinking/

Involve is the UK's leading authority on public participation. Involve develops opportunities for engagement, collaboration and accountability that work for both citizens and organisations.

The Carnegie UK Trust works to improve the lives of people throughout the UK and Ireland, by changing minds through influencing policy, and by changing lives through innovative practice and partnership work. The Carnegie UK Trust was established by Scots-American philanthropist Andrew Carnegie in 1913

Understanding Patient Data has been set up to support better conversations about uses of health information. It aims to develop tools and resources to help inform discussions and support responsible uses of data. Understanding Patient Data is led by a small core team, based at the Wellcome offices in London, UK and has been set up to run for two years.

June 2017

Understanding Patient Data

involve

CarnegieUK TRUST

CHANGING MINDS • CHANGING LIVES

Carnegie United Kingdom Trust
Scottish charity SC 012799 operating in the UK and Ireland
Incorporated by Royal Charter 1917