Arch Sci (2017) 17:305–330
DOI 10.1007/s10502-017-9281-4

ORIGINAL PAPER

A balance of trust in the use of government administrative data

Anna Sexton[1,2] · Elizabeth Shepherd[1] · Oliver Duke-Williams[1] · Alexandra Eveleigh[1,3]

Published online: 10 October 2017

Abstract Government departments and agencies around the world routinely collect administrative data produced by citizen interaction with the state. The UK government increasingly frames data as an 'asset'. The potential in administrative data can be exploited by sharing and linking across datasets, but when the rhetoric of the benefits of data sharing is bound up in commercial exploitation, trustworthy motivations for sharing data come into question. Such questions are framed around two apparently conflicting public goods. The public good in re-using data to increase government efficiency and to enhance research is set against the public good in protecting privacy. Privacy is a collective as well as an individual benefit, enabling the public to participate confidently in citizen-state interactions. Balancing these public goods is challenging given rapidly evolving technology and data science. The analysis presented here draws on research undertaken by the authors as part of the Administrative Data Research Centre in England. Between 2014 and 2017, four case studies were conducted on government administrative data across education, transport, energy and health. The purpose of the research was to examine stakeholder perspectives in relation to administrative data sharing and re-use. The themes of trust, risk and consent were chosen to articulate the research questions and analysis: this article focuses on the findings related to trust. It explores the notion of trust in the collection, analysis, linkage and re-use of routinely collected government administrative data in England. It seeks to demonstrate that securing public trust in data initiatives is dependent on a broader balance of trust between a network of actors involved in data sharing and use.

✉ Elizabeth Shepherd
e.shepherd@ucl.ac.uk

[1] Department of Information Studies, UCL, Gower Street, London WC1E 6BT, UK

[2] The National Archives, London, UK

[3] The Wellcome Library, London, UK

Keywords Government administrative data · Trust · Records and data · Data sharing · Privacy · Health data · Education data

Introduction

Government departments and agencies around the world routinely collect administrative data produced by citizen interaction with the state. The UK government increasingly frames data as an 'asset' (Davies and Bawa 2012; Harrison et al. 2015). The potential in administrative data can be exploited by sharing and linking across datasets, but when the rhetoric of the benefits of data sharing is bound up in commercial exploitation, trustworthy motivations for sharing data come into question (Bates 2012). Such questions are framed around two apparently conflicting public goods. The public good in re-using data to increase government efficiency and to enhance research is set against the public good in protecting privacy. Privacy is a collective as well as an individual benefit, enabling the public to participate confidently in citizen-state interactions. Balancing these public goods is challenging given rapidly evolving technology (volume of storage, faster information retrieval and processing) and data science (more powerful statistical techniques and algorithms) which have led to increasingly powerful means to 'collect, manage, combine, analyse and derive insight' from data (Nuffield 2014, p. xvii, 4). In sophisticated technological environments, these public goods are entangled in complex ways (Grace and Taylor 2013). Questions of trust begin to emerge, the most central of which is whether the public can put their trust in the government's ability to strike the right balance between public goods in extending sharing of administrative data across and beyond government. For those with responsibility for advising the government on policy in this area, the importance of securing public trust is well understood, and gaining trust from data subjects and the broader public is an essential ingredient in enabling government data-sharing initiatives to develop (Caldicott 2016).

This article explores the notion of trust in the collection, analysis, linkage and re-use of routinely collected government administrative data in England. It seeks to demonstrate that securing public trust in data initiatives is dependent on a broader *balance of trust* between a network of actors involved in data sharing and use. In examining this balance of trust, the article will explore how processes and systems intended to build and monitor trustworthiness across these interrelations can sometimes have an unintended detrimental impact on the balance of trust between stakeholders. Although the research setting is England, we believe that the issues it illustrates about trust in data have wider resonance (Yoon 2017).

Research methods

The analysis presented here draws on research undertaken by the authors as part of the Administrative Data Research Centre in England (ADRC-E). Between 2014 and 2017, four case studies were conducted on government administrative data across education, transport, energy and health. The purpose of the research was to examine stakeholder perspectives in relation to administrative data sharing and re-use. Each case study constitutes one bounded or instrumental case (Stake 2005), whereby the case 'plays a supportive role and … facilitates our understanding of something else': in this study, the secondary use of administrative data. The qualitative study used semi-structured interviews as the main data collection method, supported by documentary analysis and a systematic literature review. In the interviews, the re-use (or secondary use) of government administrative data by academic researchers was the central lens through which data issues and stakeholder perspectives were examined. The themes of trust, risk and consent were chosen at the outset to articulate the research questions and structure dialogue with stakeholders. Following Stake's constructivist methodology, data collection proceeded from a flexible, relatively unstructured conceptual framework (around the three themes), began with stakeholder interviews and triangulated with documentary evidence, in order to obtain a holistic understanding of the issues. Gathering perspectives from academic researchers at various levels of seniority who use government administrative data in their research formed the core of each case study. However, their views were contextualised through interviews with other stakeholders including government bodies acting as data providers, policy makers, advisors, regulatory bodies, research funders and lobby groups. In the education case study, data subjects were also included.

This article draws on the education and health data case studies, comprising 30 individual interviews and four respondents in a focus group, and focuses on the findings related to trust. The extensive range of data providers and datasets in the fields of health and education, the higher levels of public awareness and interest in data use relating to health and to education, and the larger numbers of well-established researchers using quantitative data allowed for more extensive studies in terms of the breadth of stakeholders and number of participants included. In each field there is a core dataset which is commonly used by researchers (described below) so that we could interrogate researchers' experiences of using the same data. We were not seeking to produce a single subject-based analysis or to make direct comparisons between the two subject fields, but rather to extrapolate the higher level issues around trust and government administrative data. Details of interviewees in these two case studies are given below. Interviewees were anonymised and extracts are referred to by the anonymisation code:

Education case study

Categorisation of interviewees	Numbers interviewed
University data manager	1
Academic researchers	7
Research board member	1
Department for education	2
Undergraduate student	1
Postgraduate students	4 in focus group
Total/reference codes	16/FG1, A6-A17

Health case study

Categorisation of interviewees	Numbers interviewed
Data provider (NHS digital and others)	4
Academic researchers	6
NHS England	4
Privacy lobby group (Medconfidential)	1
Research funding body (Wellcome Trust)	1
Health research authority	1
Senior policy advisor on health data	1
Total/reference codes	18/A36–A53

Interview protocols were developed for each group of interviewees. For example, the interview questions for academic researchers were structured according to the five phases of the Data Documentation Initiative's research data lifecycle model (DDI 2014) chosen because of its likely greater familiarity to the respondents than, for example, the records continuum: discovery and planning, data collection, data preparation and analysis, publication and sharing, and long-term management. Initial questions were such as, can you give me a brief outline of the type of research that you do, including any relevant recent projects which have made use of government administrative data? Does your research involve purely administrative data, or do you link administrative data to other data sources (e.g. survey data, longitudinal cohort studies)? Does your research use administrative data from a single government source, or from two or more departments? Do you use this data with a single research purpose in mind, or do you (intend to) re-use the same dataset to investigate other research questions? Followed by detailed questions about the respondent's experience of designing the data research project, finding and discovering existing data sources, consent for sharing, checking, validating, cleaning and anonymising data, data security, data interpretation and so on. To provide a contextual backdrop to our findings we describe here the types and range of government administrative datasets that the researchers we interviewed used.

In the education case study, the main dataset used by researchers was the National Pupil Database (NPD), a person-level database which matches pupil and school characteristic data to pupil attainment for the primary purpose of tracking student attainment. Raw data from the school census form part of the NPD (https://data.gov.uk/dataset/school-census). Access to the NPD including sharing for wider purposes was subject to a government consultation in 2012 (Department for Education 2013). Researchers can apply to the Department for Education for linked extracts of the NPD to Higher Education Statistics (HESA) data, and to records of students in Further Education (Individual Learning Record or ILR).

In the health case study, across the National Health Service (NHS), data (including identifiable patient data) are collected at various levels of administration across primary (General Practice) and secondary (hospital) care settings. To ensure adequate safeguarding over patient data, the flow of data is tightly regulated and controlled, under legislation reinforced through data governance systems. Reviews by the National Data Guardian for Health and Care, Dame Fiona Caldicott, have been influential in highlighting areas of weakness in the system and advocating for stronger approaches to data governance (Caldicott 1997, 2013, 2016). Central control of administrative health data is mainly undertaken by NHS Digital. NHS Digital was established in 2013 as the Health and Social Care Information Centre (HSCIC), under the Health and Social Care Act 2012 with statutory powers to collect, process, and provide access to data. Its name changed in 2016. NHS Digital is the national provider for England of information, data and IT systems across health and social care.

The main dataset that researchers in the health case study used was Hospital Episodes Statistics (HES), a patient-level database containing over a billion records of patients attending Accident and Emergency units, admitted for treatment or attending outpatients clinics at NHS Hospitals including acute hospitals, primary care trusts and mental health trusts, in England. The primary function of HES is to allow hospitals to be paid for the care they deliver, but it is also a valuable resource for secondary use by NHS management and for health research. NHS Digital publishes annual HES data for 2009–2014 at provider level as open data (via data.gov.uk). Research access to patient-level data is via the Data Access Request Service (DARS) of NHS Digital.

What is trust?

Three of the authors of this paper have a background in archives and records management and therefore our own understandings of trust are bound up with those asserted within archival theory. Archival definitions of trust are rooted in the record as object and have traditionally revolved around the question of what makes records 'trustworthy'. Trust is synonymous with 'public faith' (MacNeil 2011, p. 176) which depends on showing that the record is what it purports to be (authenticity) and free from corruption and tampering (reliability). Public trust in reliability and authenticity depends on the record having adequate conditions of custody, secured by the twin notions of 'trusted custodian' and 'trusted repository'. The three-way relationship between trusted record, trusted custodian and trusted repository is what Heather MacNeil describes as a 'central trope' in archival science (2011, p. 175). However, the simplicity of this three-way relationship hides a more complex picture. The fluid nature of digital records, alongside their dynamism and multiplicity, forces a confrontation with traditional archival understandings. What archivists now see is that there is entanglement between the record, space–time and human agents. We are beginning to understand that public trust in the record is a fluid construct, dependent on an ever-changing dynamic between the record, the space–time in which it exists, and those that interact with it.

Although building public trust in the record, the repository, and the custodian has been central to archival science, there has been surprisingly little exploration of the nature of trust in them. A 'trusted custodian' and a 'trusted repository' are simply mechanisms for preserving the record's authenticity and reliability. The impacts on public trust made by broader human interactions with the record (by researchers, users and communities) are largely unexplored (Sundqvist 2011; Yoon 2014, 2017 are exceptions). Understandings of trust across sociology, philosophy and psychology do not take an object (i.e. a record) as their starting point for definition and exploration, but focus on how trust operates in human relationships. These other perspectives can enhance archival understandings of what trust is and helped the authors to conceptualise trust in relation to government data sharing and use.

Explorations of trust in human relationships have conceptualised trust as an attitude and an affect (Baier 1986; Holton 1994; Jones 1996). Jones (1996), for example, suggests that trust 'is an attitude of optimism that the goodwill and competence of another will extend to cover the domain of our interaction with her, together with the expectation that the one trusted will be directly and favourably moved by the thought that we are counting on her.' Nickel (2007) highlights how attitudes between truster and trusted are underpinned by a mutual sense of duty and obligation. Holton (1994) draws out the affective qualities attached to trust, through feelings of gratitude when trust in someone is rewarded and betrayal when trust is disappointed. Giddens (1991) draws on Simmel to suggest that trust is most accurately described as a form of faith because it expresses an affective commitment to something or someone (Giddens 1991, p. 25).

Others have conceptualised trust as a cognitive and rational judgment (Hardin 2002; O'Neill 2002b). Hardin (2002) describes trust as 'rational expectations of the self-interested behaviour of the trusted'. For Hardin, trust operates through a cognitive assessment by the truster on the motivations and 'encapsulated interest' of the trusted. Trust is linked to reciprocity and overlapping interests: 'I trust you because your interests encapsulate mine to some extent' (Hardin 2002, p. xix). Trust is contextual, a three-part relationship, 'grounded in the truster's assessment of the intentions of the trusted with respect to some action' (Hardin 2002, p. xix).

O'Neill sees trust as a rational judgment but 'displayed in making an overall judgment in the face of incomplete evidence' (Seemann 2007, p. 5). In other words, for O'Neill, trust is not a matter of 'blind deference' but is about placing trust with 'good judgment' in 'less than certainty' (Seemann 2007, p. 6). O'Neill makes the striking point that 'trust is needed not because everything is wholly predictable, let alone wholly guaranteed, but on the contrary because life has to be led without guarantees' (2002b, p. 24). Similar arguments have been put forward, for example, by Six et al. (2015) who describe trust as 'a lubricant' that enables the smooth running of actions, and view trust as an antidote to being stuck in inertia (p. 155). This links with the notion that trust has an important part to play in enabling cooperative social relations (Luhmann 1980; Putnam 1993; Fukuyama 1995).

In *Questions of Trust*, 'trust' and 'acting on trust' are somewhat conflated by O'Neill, who explores the tendency for individuals to say that they do not trust industries and institutions whilst choosing to buy their products or opt into their services (2002b, p. 11–14). For O'Neill, because they act as if they trust, people may trust even when they say they do not (2002b, p. 14). Hardin's formulation makes a distinction between 'trust' and 'acting on trust'. For Hardin, trust is cognitive not behavioural. Hardin separates the act of cooperation from trust: 'I can take the risk of cooperating with you even though I do not trust you' (Hardin 2002, p. 11). Therefore, an individual may 'not trust doctors' but will cooperate with a specific doctor in assessing options and risks in a given context. Luhmann argues that ‘trust is a solution for specific problems of risk’ (2000, p. 95). He separates passive and un-reflexive confidence from trust: we act confidently with little deliberation to avoid overwhelming uncertainty and paralysis. Trust, however, is a deeper reflexive process that ‘presupposes a situation of risk’ (2000, p. 97). ‘You can avoid taking the risk, but only if you are willing to waive the associated advantages’ of an action (2000, p. 97). For Luhmann, the decision to trust is an acceptance that the harm may outweigh the benefits (2000, p. 98), but trust ‘allows risk-taking decisions’ (2000, p. 103).

Rather than seeking to find a fixed and stable definition of trust, we would argue that each of the formulations of trust highlighted here offers an insight into what is a dynamic and multi-layered social construct. Trust is a relational attitude with both affective and cognitive attributes, akin to a form of faith. It results from reflexive deliberation that is always contingent on context. It supports the shared goodwill between parties in awareness of reciprocal duties and obligations. A precondition of trust is the absence of complete assurance, and it therefore relates to an assessment of risk.

As a multi-layered construct, trust in people and things can operate at various levels of aggregation. Trust can play a part in one-to-one relationships, in face-to-face as well as faceless interactions. It can figure at a collective level (such as ‘public trust’ in the government). In relation to trust in government data sharing and re-use, we have sought to understand trust at different levels of aggregation and between entities when people interact with government institutions, systems and processes.

Mapping relationships of trust in government administrative data

To understand how trust and trust relationships figure in the sharing and re-use of government administrative data, we needed to understand the interdependent relationships between the various stakeholders and actors, and the context surrounding government administrative data. We analysed all our interview data to identify the relationships between key actors as described by our interviewees. Descriptions from our interviewees were richest in describing the relational interdependencies around academic research re-use of data and form the focus of our model. As one interviewee said:

> …in the end it comes down to trust; developing trust between the data providers and those who are going to gain access to the data, [and] sometimes that trust requires intermediaries who can act to negotiate on behalf of researchers. (A15)

In the model given in Fig. 1, we categorise stakeholders into three broad groups: data providers, data users (of whom our focus is on researchers) and data subjects. These foundational stakeholder relationships are shaped by further independencies across oversight bodies, intermediary groups, the media and a wider public. Each stakeholder group represented in the model has a complex interdependent relationship with the other stakeholders and actors represented: the model attempts

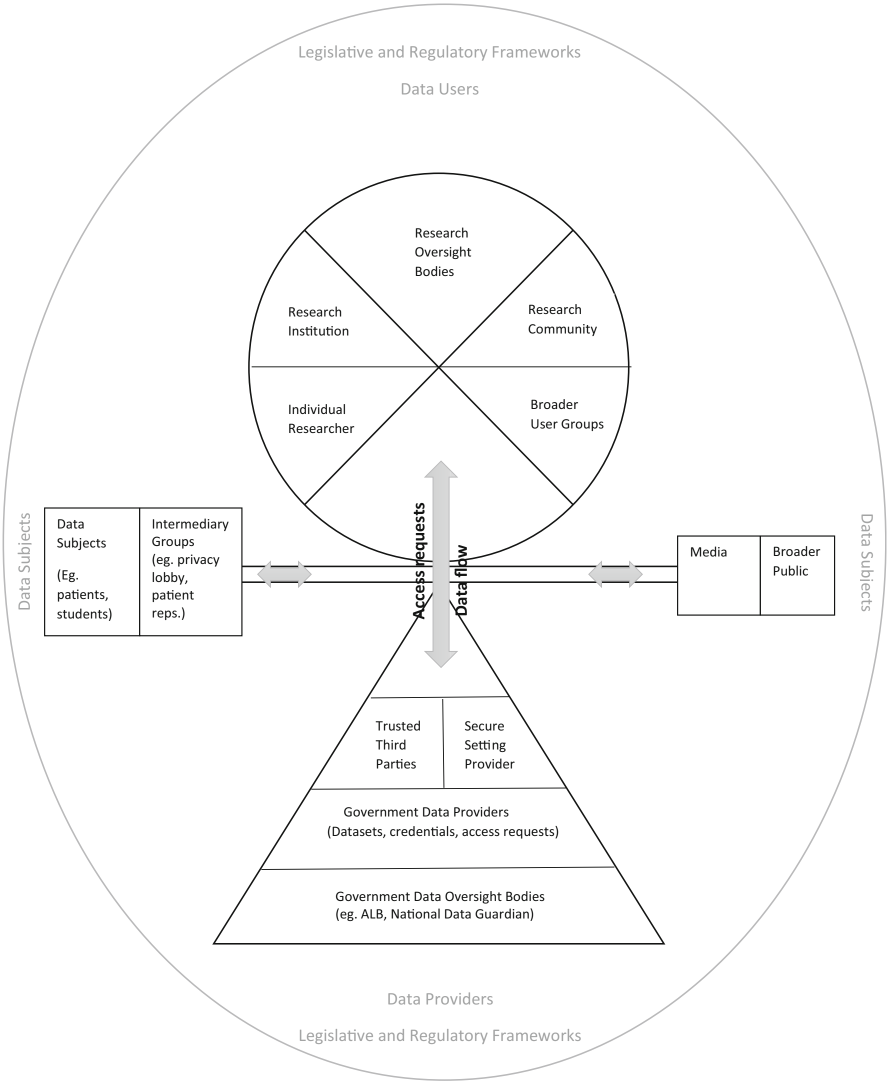
Legislative and Regulatory Frameworks
Data Users
Research Oversight Bodies
Research Institution
Research Community
Individual Researcher
Broader User Groups
Data Subjects
Data Subjects (Eg. patients, students)
Intermediary Groups (eg. privacy lobby, patient reps.)
Access requests
Data flow
Media
Broader Public
Data Subjects
Trusted Third Parties
Secure Setting Provider
Government Data Providers (Datasets, credentials, access requests)
Government Data Oversight Bodies (eg. ALB, National Data Guardian)
Data Providers
Legislative and Regulatory Frameworks

Fig. 1 Balance of trust

to visualise how these are suspended in a balance of co-dependency. We call this model the ‘trust balance’ because these complex interdependencies rely on building and maintaining a relational attitude of trust, bound by an array of duties and obligations.

Shifting relationships: faceless connections in an abstract system

Across our case studies, researchers who have used government administrative data in their research over a long period (i.e. for more than 5 years) spoke of a shift in the nature of the relationships underpinning their access to government administrative datasets (the research user—data provider relationship). Previously access depended on building up strong one-to-one relationships with key gatekeepers in government departments. Interviewees described a trend towards standardised, impersonal mechanisms for data access:

> Back then, there wasn't really an access route to getting the National Pupil Database. And it was all a lot more informal. And it typically was sort of given on a researcher by researcher basis, and additionally, only to those people who knew the correct people within the Department for Education to speak to. (A11)

Most researchers recognised that a standardised approach led to fairer access and that this was generally a positive step:

> So, you know, as that process formalised, which, you know, I'm greatly encouraging of, through the mechanism that you can apply for the data directly through DfE and there's no longer this sort of mechanism where you've got to know the right person, the assessment process is very fair in my experience, and also very uniform in terms of the things you can and can't do, whereas I think in the past that was probably not the case. (A11)

However, there was also a strong articulation amongst a subset of researcher interviewees on the challenges that occur when access becomes predicated on systemically driven interactions, rather than close interpersonal ties. For some, the loss of a more personal approach to custodianship and its replacement by generic access procedures reflected a higher turnover of gatekeeping staff. This was perceived to lead to a loss of data provider expertise which, in turn, has a detrimental impact on the researcher's ability to understand the data they receive from the provider:

> … if you wanted to know about that data you rang that nice man [redaction] and he would be very helpful… And then once it went into the Information Centre in Leeds it was a whole succession of people who were moved on very quickly and they no longer had any particular interest in the data…they tried to be helpful but they weren't very knowledgeable. (A44)

In *Consequences of Modernity* (1991), Anthony Giddens builds on his earlier work on structuration and the reproduction of social practices. A central theme is how trust operates in relation to modern abstract systems. Giddens argues that abstract systems comprise symbolic tokens which circulate in expert systems (i.e. systems of technical or professional expertise that organise the material and social environments in which people live). He suggests that the rise of modern abstract systems is a feature of an increasingly globalised world, where traditional patterns of local (face-to-face) interactions have given way to increasingly distant relationships with absent others. In abstract systems, our relational ties stretch, as we increasingly interact with the world through faceless processes. By 'disembedding' processes in abstract systems, face-to-face, local interactions are restructured across intervals in space and time. Abstract systems enable what Giddens refers to as time–space distantiation. Interactions in the abstract system are frequently ‘faceless’, as actors interact with the system itself (e.g. through automated information exchange) rather than through personal contact (‘facework’). The increasingly ‘faceless’ nature of relational ties in abstract systems resonates strongly with our interview data. Whilst ‘faceless’ interactions with a generic access process was seen as an increasing trend, opportunities within these processes for sustained ‘facework’ (i.e. personal one-to-one contact) with an experienced and knowledgeable individual acting on behalf of the data provider were not only highly valued and sought out, but were seen as essential in establishing a workable relationship:

> [In the end] we were lucky that someone in HSCIC turned out to be someone who [name of colleague] knew, and so we contacted him directly and [after weeks of hold ups] in a day he redrafted the agreements and got them though which just shows you what happens when you finally have contact with a competent person. (A44)

Giddens’ description of abstract systems also provides a frame for thinking about how routinely collected government administrative data circulates as a symbolic token within an increasingly distantiated expert system. Taking public sector health data as an example, patient information created as part of a face-to-face ‘doctor-patient’ relationship becomes disembodied from its originating context through data coding and extraction processes. It is further distantiated in its reconfiguration as part of large datasets, which in turn are expanded and stretched through time and space as datasets are linked across and beyond health and social care. Different data custodians are responsible for different datasets which are re-used by third parties, interactions which all take the data further from the creating context. As our researcher interviewees articulate, the relationship between data users and data providers is itself becoming more distanced, abstract and faceless as standardised protocols for access take over from personal gatekeeping as a means of governing the tie between provider and researcher. A complex legislative and regulatory environment, with many oversight institutions managing data quality, security and use, governs the relationship between provider and researcher. Faceless processes of governance are increasingly foregrounded over traditional relational bonds.

When the sharing and re-use of routinely collected administrative data is framed as an abstract system, we become aware of: (1) the stretching of the data away from its originating relational ties; (2) the complexity and opacity of the data flows within the system; (3) the number of bodies involved in managing or interacting with the data; and (4) the increasingly impersonal nature of interactions within the system. In the light of all of this, can we say that a balance of trusting relationships holds the system together?

Exploring interdependencies

The notion of balancing trusting relationships emerges through a deeper analysis of our interviewees’ descriptions of the relational interdependencies that exist between stakeholders. Our interviewees talked in depth about how behaviour, actions and processes around data sharing are shaped, bounded and modified by attempts to build and maintain the good will of other stakeholders.

The researchers we interviewed frequently spoke of a determination to act responsibly with data, and how this determination is driven partly from their own vested interest in doing so. Acting irresponsibly would result in reputational damage for the individual and their institution, and likely exclusion from future research participation. Researchers are well aware of their reliance on building and maintaining the favour and goodwill of other stakeholders (data providers, data subjects, broader public, research institutions, and the research community). They are also aware that their behaviour as an individual has an impact on this wider network:

> Most researchers work in a research institute or a university, if they're getting access to administrative data. And they, when they're given access, their employer is bound into a set of conditions about how they must behave with respect to those data, and the penalties for disobeying that or abusing that set of rules and guidelines is fairly extreme. … a lifetime ban on their receipt of further ESRC [Research Council] funding. It can also be a ban on funding to their organisation …. So the penalties are quite draconian, so it means then that the individual and the institution has a very strong incentive to ensure that their employees do not do anything inappropriate. (A15)

This interdependence is in line with Nickel's (2007) framing of trust as driven by a mutual sense of duty and obligation and Hardin's (2002) idea of trust driven by 'encapsulated interest'. Researchers and research institutions have a stake in complying and this relational interdependence enables individuals and institutions to form a bond of trust.

If the formation of these trusting bonds relies on the balance of the interdependencies between stakeholders, then the behaviour of each stakeholder affects all the surrounding bonds of trust. One researcher in the health case study spoke of the impact a researcher can have on the reputation of both the data provider and the broader research community:

> You will always have this tension that comes from liability, we only need one researcher to sell the data, or lose the data for 100,000 patients and that is a big issue for the HSCIC, they will not release any datasets for about three years because why would you? You are going to be the one up in front of Parliament, why did you release it to them? Did you go and check their server? Did you do a security check? No of course they didn't, they are relying on someone signing a piece of paper and doing what they are said they would. (A41)

Researchers in the education case study also described this effect on the interdependent bonds between stakeholders:

> Nobody else has the kind of access that the research community has, and any abuse of the trust that people put in the research community can impact very strongly on that privilege. (A15)

> … about ten years ago, civil servants, you know, disks containing missile deployments were left on trains and buses, and we then all suffered for that. … education researchers and education data. … it just struck me that we're all now suffering because of some mistakes made by other people. Which is always the way. Instead of just saying, right, you made a mistake, we're going to make the system harder for everybody. (A9)

The interdependencies between stakeholders make the balance of trust across the abstract system a precarious one. One stakeholder's commitment to meeting their duties and obligations may not be enough to establish a balance of trust. Unfulfilled duties and obligations by others initiate a loss of goodwill which spreads across the bonds of trust, placing a stress on all relational interactions.

Tensions in securing public trust in abstract systems

All types of interviewees (researchers, data providers, policy makers) in both the education and health case studies spoke of establishing and maintaining public trust as a prerequisite for data-sharing initiatives, yet there was also a consensus that securing public trust is a difficult task:

> Public trust is very very easy to lose, and the moment you have headlines like Tesco owns your medical records, or insurance companies have got your medical records, this instils fear into people—fear, uncertainty, the unknown—who is accessing my data and why? And losing public trust has wide ramifications. It can potentially impact on the relationship between the doctor and patients. If you went to your GP would you be reluctant to tell them something important because you were worried about where the data could end up? That would be absolutely catastrophic…So public perception does matter and losing it has an impact on the potential to use data for research. (A41)

When doubt is cast over how the system operates, the public do not have enough knowledge to make a rational calculation of risk. Giddens highlights how the processes of disembedding and time–space distantiation inherent in abstract systems have a disorienting effect, we experience life as a 'careering juggernaut' (1991, p. 53), and this leads to increasing levels of uncertainty and anxiety and an increased perception of risk. Similarly, O'Neill highlights how modern life is framed by sociologists as a 'risk society' where 'living amongst highly complex institutions and practices whose effects we cannot control or understand' means that we 'see ourselves as subject to hidden and incomprehensible sources of risk' (2002b, p. 15). O'Neill suggests that our heightened perception of risk, rather than 'the seriousness of the hazards to which people are exposed, or the likelihood that those hazards will harm them' (2002b, p. 16), has led to an apparent 'crisis of trust'. She explores whether the loss of trust in our governments, institutions, professionals, and experts is symptomatic of the 'public mood of suspicion' inherent in a risk society or a 'justified response to growing untrustworthiness' (2002b, p. 16).

We examined this phenomenon in the health setting, which has suffered adverse media coverage of data release in the growing culture of suspicion over the government's care.data programme (discussed later), reports that the NHS has one of the worst records for data protection breaches reported to the Information Commissioner's Office (The Guardian 2016a), and that patient data have been passed onto commercial companies without consent, for example, the data-sharing agreement between the Royal Free London NHS Foundation Trust and Google's Artificial Intelligence company DeepMind (The Guardian 2016b; Financial Times 2016; ICO 2017). The bad news of untrustworthiness grabs public attention more easily than the good news of trustworthiness (O'Neill 2002b, p. 17). As O'Neill (2002a, p. 11) sums up, 'A culture of blame and accusation is widespread, both in the media and in the literature of campaigning organisations where fingers are pointed variously at government, scientists and at business.'

Our interviewees understood this:

> So [there is now] a much greater awareness from the public about how much data is connected, and the various uses to which data is put. An awareness that has tended, over the last decade, I would say, to have presented itself to most of the public through the media, which has not been helpful because much of the media has a particular, and I would say biased, slant on this. (A51)

The difference between trust and trustworthiness

One response to a loss of public trust is the 'audit explosion' (O'Neill 2002b, p. 21), the aim of which is to reduce risk and increase trustworthiness by expanding measures to demonstrate the competence, security and reliability of the abstract system. The rise of information governance in institutional contexts including the NHS as a system of monitoring to control risk and increase accountability is symptomatic of the audit culture. The question raised by O'Neill is a vital one: do these measures have any impact on securing or maintaining trust?

Following scrutiny of HSCIC by the UK Parliament Health Select Committee, Sir Nick Partridge (2014) led a review of data releases made by one of its predecessor organisations, the NHS Information Centre (NHS IC). The review found that in only nine of the 3059 releases under review (0.3%) were there any grounds for concern over data sharing. Yet, Partridge recommended that HSCIC should further standardise and tighten mechanisms for compliance and accountability in an attempt to eliminate personal data breaches entirely. The Partridge review illustrates a risk society where cycles of 'prevention and sanction' symptomatic of the audit explosion are the prescribed remedy to the question of trust (O'Neill 2002b, p. 44).

Measures to improve trustworthiness, such as those which demonstrate effective and proportionate governance rather than disproportionate levels of audit, may increase the trustworthiness of a system. Sundqvist (2011, p. 289) showed that control documentation designed to ensure measurable trustworthiness in record-keeping procedures (for instance through compliance with trusted standards and guidance) may have a positive relationship with building trust, by boosting a truster's positive opinion of a trustee's reliability, but she warns 'not necessarily so.' One data provider said:

> Do we audit compliance, um, no. We can't really… a lot of it is down to trust. We're trusting what the customer is telling us as being true. We can't audit everything that they're telling us. (A13)

Trustworthiness and trust are not equivalents. Some interviewees recognised tightening data governance as symptomatic not of gaining public trust but of the diminution of trust between data providers and researchers:

> There's been less trust of researchers, more, yeah, more bureaucracy, a lot more form filling and a lot more hassle to get access to any of the stuff that you could use, even when, it appears to me, you're not asking for anything particularly sensitive. (A9)

O'Neill's argument is in line with this interviewee's comments. The audit explosion with its focus on increasing the trustworthiness of systems as a means of raising public trust in fact seeks to generate trust by eradicating the need for it. 'It is an agenda of replacing traditional relations of trust, now grown problematic, with stronger systems for securing trustworthiness, an agenda, as John Thompson puts it, for economising on trust' (O'Neill 2002a, p. 130). A focus on trustworthiness undermines the need for trust.

Increased trustworthiness may instead 'deepen the distrust it seeks to remedy' (Thompson 2000, p. 253–4 in O'Neill 2002a, p. 130). Systems that focus on trustworthiness at the expense of trust will only work if the system and those that interact with it are completely infallible. Yet no matter how tightly regulated, controlled, secure and audited the system is, human error or human deviance will emerge. As the Partridge review (2014) demonstrated the system will prove to be untrustworthy. A culture of distrust will pervade. A number of the interviewees acknowledged the inevitability of failure, summarised by one in the following terms:

> Those responsible for reviewing data [requests] they know that in the long, long, long term they will approve something they shouldn't because someone will lie to them and they won't catch it, and their job is to hope that it doesn't happen on their watch. …when people determined on misuse, read the rules and have time to decide how they will respond to them, they will work out a way round, and that is not a scenario where public trust in the long term will continue, simply because at some point somebody will do something stupid. (A49)

O'Neill draws on Thompson to suggest that measures to increase trustworthiness may 'increase inefficiency' by adding 'further layers of bureaucracy' and this may also 'exasperate rather than alleviate' a sense of distrust (2002a, p. 130). This is in keeping with the findings of our health case study in relation to levels of trust between researchers seeking to access data and data providers who act as data custodians. In response to the Partridge review recommendations, HSCIC halted processing data requests whilst they reviewed their procedures and introduced new measures. The researchers we interviewed described the far-reaching consequences for the research community, many of whom faced considerable delays in receiving data for research purposes. These delays affected the wider research governance framework as timeframes set by research funders became unachievable in the face of data provider delays. The reciprocal relationship of trust and goodwill between HSCIC and researchers was placed under considerable strain. Research approved by university ethics committees as being in the public interest, legal and ethical, was put on hold. Our interviews suggested that the delays were particularly problematic for PhD researchers working to very fixed timescales, forcing some of them to reduce the scope of their research in order to use more easily available datasets. One PhD supervisor we interviewed described it as ‘extraordinarily disturbing’ (A40) for the students involved.

The focus on measures to increase trustworthiness at the expense of the maintenance of relations of trust may damage the ‘encapsulated trust’ between researchers and data providers. Researchers trust data providers to enable timely access to data because doing so enables them to fulfil their remit in supporting beneficial research. The data provider trusts the researcher not to abuse the system (through incompetence or malfeasance) because it is in the researcher’s best interests to play by the rules. O’Neill suggests that institutions must aim, in the laudable pursuit of trustworthiness, for mechanisms of ‘intelligent accountability’ (2002b, p. 59). This entails recognising the separation between trustworthiness and trust, and looking for ways to build and maintain both.

Intelligent accountability recognises that detailed contractual agreements between data-sharing parties and complex auditing may replace trust. Detailed monitoring suggests to the public that data users cannot be trusted. Intelligent accountability seeks a balance between the bureaucracy necessary to ensure trustworthy systems and maintaining trust relationships. Considered in this way, Partridge’s recommendation that the HSCIC develop an audit function to monitor ‘other party’ compliance of its data-sharing agreements becomes questionable:

> I think [the Partridge review] contributed to the perception of system failure. I am not convinced there was necessarily a system failure. Is it the business of the HSCIC to track through contract after contract to see if people have done what they are supposed to do? Usually you would expect people to fulfil their contractual obligations, where they don’t you introduce sanctions, which is usually not doing business anymore and withdrawing. So I think they have gone out of their way to look for breaches…and they have created a climate of distrust through that. (A42)

Hardin (2002) argues that relational trust that does not involve direct interpersonal connection (such as in abstract systems): what stands in place of trust is citizen cooperation. Cooperation is based on undeliberated confidence which is itself dependent on our ‘inductive expectations’ from past behaviour and reputation (pp. 151–172). This works if the public’s confidence in the system remains relatively unchallenged. However, media reports and lobby groups casting doubt on the system force the public into a deliberative choice on whether to trust or distrust. If the public's confidence turns into distrust, then cooperation is likely to be withdrawn.

The 'debacle' (A36) surrounding the care.data programme illustrates this. NHS England announced the care.data initiative in 2013 to extend the data collection already undertaken by HSCIC across the NHS to include General Practice (GP) and other records in a national database, alongside the HES dataset of hospital episodes and other data, pseudonymised and centrally maintained. Datasets could be linked to enable better planning, managing and commissioning of healthcare across NHS England, more comprehensive whole population research and more effective care pathways and service models. However, the programme was subject to relentless public scrutiny and criticism. The loss of public confidence led to withdrawal of citizen cooperation, as over a million citizens opted-out. Deliberative public distrust had taken hold. In the face of this public mistrust and opposition from health professionals, the programme was put on (now indefinite) hold.

The failure of care.data has been explored from many different perspectives (Mann 2016; Torjesen 2014; Carter et al. 2015; Hays and Daker-White 2015; van Staa et al. 2016; Vezyridis and Timmons 2017). Concerns about trustworthiness of the proposals over privacy and data security, a lack of adequate communication around the safeguards for access and use of the data, and concerns over data commercialisation led to its demise. The nature of informed consent and the adequacy of the consent model as 'opt-out' rather than 'opt-in' undermined trust (Hays and Daker-White 2015 in van Staa et al. 2016). Did the care.data programme falter because the system underpinning it was not trustworthy? Or, did it falter because of inadequate communication with the public aimed at establishing trust in the programme and the benefits it would lead to? In contrast to successful analogous programmes in Wales and Scotland (Jones et al. 2014; SAIL 2017; SHIP 2017), care.data failed to generate 'social legitimacy' (Carter et al. 2015). A reliance on the legal basis of the programme as sufficient licence for it, combined with a belief that the one-to-one warrants of trust between a GP and patient over confidentiality and use of data could be unproblematically extended into a national linked data programme, reflected a failure to understand the necessity of securing adequate degrees of informed and deliberated public trust in a distantiated data initiative.

Generating public trust

Is transparency enough?

We invited our interviewees to share their perceptions of the best ways to secure public trust in government administrative data initiatives. Our analysis of interviewee responses indicates that transparency over the purposes of an initiative, the processes that underpin it, and the safeguards surrounding it, was the most often cited means through which to secure public trust:

> It is around making people aware of the purposes that their data will be used for. …this is why we are using your data, I think Joe Public are generally ok about it, it is when something happens to their data that they aren't aware of, that is when people start getting twitchy…Once you get on that back foot where it has been used without consent and without awareness, then … you are probably going to get distrust—I won't give my data again because you have used it in this way without my knowledge…As long as patients are informed, and they don't feel like it has been done behind their back, I think that is the key isn't it? (A38)

Interviewees agreed that levels of transparency over data flows, systems and programmes could and should be improved:

> There are two types of trust question. One the more you know about the problem the more concerned you get, bioweapons for example, and there is another type where the more you learn the less concerned you get, and health data should be in the latter category because there is actually quite a lot of governance. The problem is because the governance is secret, it is a case of are you sure that nobody has broken or abused the system? (A49)
> So there's a balancing act here between transparency, thinking about the ethics but with the default being if the data can be used for the public good it should be. (A12)

However, despite advocating for transparency, the majority of the interviewees also alluded to the complexities. Data flows, control systems, organisational roles, types of data, uses of data, variety of users, and the legislative and regulatory frameworks and oversight bodies that make up the monitoring and safeguarding system is complex, technical and opaque. The question is how to be transparent:

> For me, as a transparency advocate, I find this incredibly challenging because I want to be out there talking to the public about what we are doing with their data, and to have them give their views …but data sharing and data linkage are actually incredibly complicated ideas. … I think we have not been good enough about demonstrating to people what their information looks like, and my experience is that you get shunted into doing 20 min with a group of 15 patients trying to tell them what the programme is doing and getting them to say 'yeah it sounds like a good idea'. Actually it is very hard to do that … but if you want them to properly input into what we are doing you have to spend time educating them and building up that knowledge base before they can start to input properly. (A43)

Giddens argues that public trust in abstract systems 'does not depend on full initiation into the systems processes or mastery of the knowledge they yield' (Giddens 1991, p. 27) but on pragmatic experience that such systems generally work as they are supposed to. Research by Ipsos MORI (2014, 2016) on the relationship between public understanding and public trust in the uses of data and of data linking and by Health e-Research Centre (2016) asking to what extent patients should control access to data, indicated that participants had very low awareness of these issues and tended to be sceptical initially. Two or three days of dialogue increased knowledge and anxiety about data sharing and social research but eventually led to more positive associations about its value, as ‘greater knowledge about the subject and exposure to the ideas tends to be related to acceptance’ (Ipsos MORI 2016). We can describe this as a tipping point where knowledge of data-sharing systems and processes becomes sufficient to enable trust. However, as one of our interviewees pointed out, such an intensive process cannot be replicated for the whole population. In the case of care.data, NHS England notified every household in England with brief information in a leaflet, which tended to increase public anxiety without increasing trust. Transparency should provide mechanisms for those that want to become more knowledgeable about initiatives, systems and processes, whilst keeping the programmes out of the headlines. The aim is to be quietly transparent so as to maintain undeliberated confidence. However, there is an inherent problem with too quiet an approach: since the legislative framework supports freedom of information, data initiatives do not stay quiet. When the public learn of data initiatives through the media rather than communication and dialogue initiated by public bodies, a culture of suspicion grows. Quiet transparency may therefore be counterproductive especially given the perspectives of the National Data Guardian, Dame Fiona Caldicott, who has consistently advocated for ‘no surprises’ in relation to data sharing for the patient and wider public (Caldicott 1997, 2013, 2016).

When handling transparency, the question is how to raise awareness and involve the public in discussion, without inducing hysteria and panic: how to get the balance between too much and too little information. We posed this question to the National Data Guardian for Health and Social Care, Dame Fiona Caldicott, interviewed as part of the study. Her response was:

> I will continue to think about it, but it is akin to giving a patient a difficult diagnosis in healthcare. When you are working with people, whether it is clinically or within a research setting, there is a rate at which an individual member of the public will want information, both in terms of rate and scale. I don’t think it is helpful to give people answers to questions they haven’t asked you on complex issues such as these, so my own approach would be; ‘I am here to explain these things to you, and I will give you an outline of this subject that you are interested in. You may have further questions, and there is more evidence and information available, but let’s start with the basics, and you are welcome to return and ask if you want to know more.’ I think it is about building layers of information and giving people opportunities to go on asking questions. In the end, a lot of this information is going to have to be online, isn’t it? One of the things many members of the public do now, as soon as there is a health issue, is to look it up on the web. We are going to need something available on data and its use, clearly documented in everyday language. People can look at the questions and answers and be offered a staged process, because for many people who we have spoken to, and there is no reason for me to think they are not typical, they don’t want to know all of the complexity, and certainly not at once, but they do want to know how they can

> get further information. So I think that is transparency, willingness to say more, but not all of it at the outset, is what is necessary to inspire confidence.

The concept of proactive public engagement through layered transparency is perhaps a useful means of approaching public conversations. In looking at the different perspectives given by our interviewees, whilst we found agreement on the need for transparency, interviewees recognise that getting that balance between too much and too little information is much easier said than done.

Several interviewees acknowledged that there are risks involved in being open with the public around data systems and data flows. A41 argued that if we let the public too far into the inner workings of the NHS, then uncomfortable questions inevitably unfold:

> [We need to] spark the realisation around something we should have been doing for the last three years but have studiously ignored, that we actually have to start to tell people what we do with data in the NHS and allowing the conversations and dialogue to happen. …the problem is that no politician wants the inner workings of the NHS exposed, because otherwise it is like you guys spend a lot of money on moving money around. That is not the story they have told about the NHS, so there are all sorts of vested interests. We can't quite tell the truth because the truth is a little bit unpalatable. (A41)

The difficulty for those involved in managing the sharing and re-use of health data is that whilst transparency over data systems in the NHS and beyond may contribute to public trust, there is no simple symbiotic relationship. Too much transparency is a risk, not just because a little bit of public knowledge can be dangerous, but also because it has the capacity to lay bare ethical questions and concerns. How can we use transparency to reach the tipping point where knowledge of data systems and processes becomes sufficient to enable public trust?

A question of ethics?

Routinely collected administrative data are increasingly shared between UK government departments and re-used in support of other government policies, such as immigration enforcement. For example, a Memorandum of Understanding (MoU) between the Department of Health, NHS Digital and the Home Office makes it clear that administrative data on individual patients held by NHS Digital can now be used directly for the purposes of immigration enforcement, to 'reduce the size of the illegal population and prevent harm caused by illegal migrants' (MOU 2017, p. 14). The public may be concerned when the government's re-use of the data is so far removed from the purpose of the original data collection (Mail Online 2017). Questions of trust are raised in relation to the balance of public goods bound up in the re-use of the data: whilst the government may deem immigration enforcement to be in the public interest, the fear of enforcement may deter individuals from accessing health services, such as vaccination, to the detriment of both the individual and the collective public. The apparent secrecy in which such data-sharing agreements have been established causes further distrust.

Another example is public reaction following revelations over the sharing of education data for immigration enforcement. Public controversy followed the UK government's decision to include country of birth and nationality as new categories on the school census, which is passed to the Home Office. Campaign groups such as Liberty have suggested that 'this isn't a data-sharing agreement—it is a secret government programme that turns the Department for Education into a border control force with an explicit aim to create a hostile environment in schools and assist with mass deportation of innocent children and their families' (The Guardian 2016c). Little effort was evident to build public trust through transparency. As a result, lobby groups, opposition MPs and the press urged parents and schools to boycott the school census. Here, as with care.data, the distrust felt by the public in relation to government data sharing led to a threat of withdrawal of cooperation. This distrust not only adversely affects government but radiates through the abstract system: research relying on the school census as a data source will be adversely affected if public cooperation is withdrawn.

In relation to the balance of trust across the system and the relationship between trust and trustworthiness, a complex picture emerges. In parts of the system, data sharing appears to be happening between government departments without adequate checks and balances, as described above; here measures to increase trustworthiness through greater degrees of scrutiny and transparency appear necessary. In other parts of the system, legitimate requests to re-use government administrative data for academic research are heavily scrutinised and controlled. For instance, a university researcher's application for access to health data held by NHS Digital is reviewed by the Independent Advisory Group on the Release of Data (IGARD), a panel of specialist and lay members. Proposals also go through institutional Research Ethics Committees, as well as monitoring by funders. In accordance with the Data Protection Act 1998 (UK Government 1998), the release must be proportionate to need, processed securely and not kept longer than necessary. The common law duty of confidentiality residing in medical information also governs health data access. Patient consent is required to enable re-use beyond the initial doctor-patient relationship. This part of the system is so complicated that the question is not whether there is enough scrutiny, but whether there are too many overlapping layers of governance. Is governance effective and proportionate to the risks, or is it excessive auditing? Is the prevailing culture of risk adversity inherent in these processes hampering potentially life-changing and life-saving research?

Carter et al. (2015) identify the need to build 'social legitimacy' in data-sharing initiatives. In keeping with their argument, we suggest that generating public trust in data sharing requires a strategy that goes beyond simple transparency and moves towards building collectively agreed boundaries in data-sharing practices. This depends on generating the political will to foster deep engagement around government data-sharing practices and a willingness to educate and involve experts and the broader public in exploring the underlying ethical questions and concerns. We argue that however challenging and difficult, robust transparency over the issues, coupled with a collective societal shaping of the parameters of legitimacy, is fundamental in moving towards reflexive and deliberated public trust.

Media reporting has shaped public knowledge of the ethics of sharing and re-using administrative data, which has become fixated on particular risks (such as re-identifying individuals from large datasets and commercial access and re-use). The public have been encouraged to worry over issues that are already tightly governed, whilst remaining ignorant of many deeper issues that have not yet been robustly or adequately addressed.

Trusting data integrity and reliability?

A great deal of attention has been given to developing a consent model for the sharing of confidential patient data within and beyond the provision of direct patient care, most recently by Caldicott in her *Review of Data Security, Consent and Opt-Outs* (2016). Caldicott points to the public knowledge gap and the public engagement work necessary alongside a consent model:

> This has been a report about trust. It is hard for people to trust what they do not understand, and the Review found that people do not generally understand how their information is used by health and social care organisations…. public understanding of the use and benefits of information sharing is limited – in particular there is a knowledge gap about the crucial need to share information across organisations to integrate health and social care and to fully benefit the individual with its potential use (pp. 42–43).

The consent model proposed by Caldicott across health and social care exposes a dichotomy between uses of data for direct care (which are implicitly presented as unproblematic) and for secondary purposes (which are implicitly presented as risky by the presence of an opt-out). The reality shown in our study is more complex. A39 described NHS hospitals routinely linking patient-level secondary care data on hospital admissions to patient-level primary care (GP) data. Patients at risk are identified and notifications sent to their GP to review their care plans. However, the effectiveness of the extrapolations relies on the underlying data quality and match rates in the linkage, yet quality in data and in methods of linking have been inadequately explored.

An NHS England internal report (Brown 2016) showed that identifiers are used to link records, both within and across datasets. Incomplete or inaccurate recording of identifiers leads to data linkage errors, via 'missed matches' 'where records belonging to the same individual fail to be linked' and 'false matches where records belonging to different individuals are erroneously linked'. Of concern is the fact that 'these linkage errors are often not randomly distributed, leading to implications for clinical practice and bias in analyses' and 'that data quality is less robust for hard-to-reach populations… for example people from deprived areas, people from different ethnic groups, cross border patients and homeless people.' (Brown 2016, p. 5).

Data linkage is effectively being used as a form of 'diagnostic testing' (A39) yet its quality and reliability is not assured. The potential in linked data to enhance patient management is being exploited, before the robustness of the techniques underpinning it have been adequately evidenced and without the necessary governance, checks and balances and safeguards. A39 highlighted this as a broader ethical issue:

> We have regulation sitting over devices, breast implants, drugs, yet we have no research and development to investigate the provenance, the applicability, the effectiveness of using data in individual patient management. … In essence, [doctors] have no control over how information technology is being used. These decisions are being made by computer contractors, by HSCIC, by NHS England, outside the remit and the advocacy role of the Doctor-Patient relationship. … We have a national screening committee, we have clear approaches for evaluating new screening tests and deciding whether we should screen babies for cystic fibrosis, for example. We have no similar systems for evaluating this technology which is changing the way we work…. Can you imagine devices or drugs being purchased across the NHS when you had no idea how they were made? And whether they are really what they are supposed to be? And this is much worse than that, this issue of big data usage in individual patient management, it is much bigger money, and bigger consequences.

This gets to the heart of the arguments presented in this paper: the need for greater transparency over the complex issues around the use of data across and beyond government. Transparency is a starting point for enabling a common approach between experts and the public to shape the boundaries of what is socially acceptable in data use. The shaping of these boundaries should then enable the development of adequate safeguards, checks and balances from which the trustworthiness of the processes can be established and held to account. Such transparency, leading to engagement and collectively shaped boundaries of legitimacy, as the basis for appropriate checks and balances has the potential to build informed and deliberated public trust. Yet there are (perhaps insurmountable) risks in taking such an approach. How can the public become knowledgeable enough to enable the introduction of these complexities without inducing panic? In a 'risk society', a pervading culture of suspicion is made worse when public perceptions are shaped by media reports. O'Neill has argued that freedom of the press should not include the licence to deceive (2002b, pp. 92–99). Media reporting offers a 'robust and widely accessible' forum in which a 'complex and multi-faceted debate' plays out and may even 'contribute to restoring public trust' (2002a, p. 168). Regulation of media communication is therefore part of the broader question of how to enable the transparency and engagement that could ultimately build public trust without it leading to deeper levels of distrust.

Conclusions: risking trust to gain it?

We have elucidated here some of the tensions connected to health and education data that became evident to us through dialogue with interviewees in our case studies. We sought to map relationships of trust between data providers, data users and data subjects, exploring the shifting relationships and interdependencies. The trust issues raised are only a part of the picture of data-sharing initiatives involving government administrative data: other articles from this research will discuss consent and risk. Building public trust in data initiatives is complex as the data is stretched away from its originating context and held in abstract systems with multiple intersections of agents and processes. Data sharing and re-use in these environments is both increasingly difficult to grasp, and increasingly ethically challenging. There is no easy answer on how to break down this complexity and enable informed public debate.

We have drawn out differences between trustworthiness and trust and argued that it is necessary to engage in initiatives that build both. In trustworthy systems and processes, a balance must be struck between appropriate monitoring in the system whilst ensuring against excessive auditing that may counterproductively contribute to the erosion of trust. Intelligent accountability built on citizen cooperation is the aim. In forging public trust, we argue that transparency is needed to enable expert and public engagement with the issues in order to collectively shape the boundaries of legitimacy, which should lead to the checks and balances needed for trustworthiness.

Is it worth engaging the public in collectively shaping the boundaries of legitimacy in relation to data sharing? Is it even possible, given the level of knowledge that is required to understand the data-sharing landscape and reach a positive tipping point? And even if it is, will a collectively shaped sense of legitimacy ultimately lead to building and maintaining public trust?

Building public trust through public engagement where the ethical complexities of data use are opened out is a risky strategy: securing public trust cannot necessarily be guaranteed. Giddens warns that in a risk society, cultures of suspicion are hard to break down. Securing a consensus on the underpinning ethics of data sharing, if done well, with adequate investment and care, may well generate greater degrees of collective trust. Simple transparency may not be enough, opening up as many questions as it answers. Can we trust existing linkage and data quality as reliable diagnostic tools? The ethics of data sharing is often framed by media reports as risky: if gaining public trust cannot be guaranteed by enabling the public to shape the boundaries of legitimate data sharing, then should we take that risk? Perhaps, following O'Neill, the balance of trust needs to shift from securing public trust as the ultimate aim, to the more fundamental principle of avoiding deception. The question of how to successfully engage with the public demands further attention, but the argument on why it is necessary to seek public cooperation in shaping the boundaries of trust in data sharing, comes down to how much do we, and our governments, want to commit to the notion that there should be 'no surprises' for the public in relation to administrative data-sharing practices.

Acknowledgements This work was funded by the UK Economic and Social Research Council (ESRC) as part of the ADRC-E (Administrative Data Research Centre—England) project, 2013–2018, Grant Number ES/L007517/1. The authors are grateful to members of the sub-project Management Group (ADRCE-project 003.01 SHEPHERD) for their advice and guidance: Professor Ruth Gilbert (UCL Institute of Child Health), Professor Lorraine Dearden (UCL and Institute for Fiscal Studies) and Professor Paul Wilkinson (The London School of Hygiene & Tropical Medicine).

References

Baier A (1986) Trust and antitrust. Ethics 96(2):231–260. doi:10.1086/292745

Bates J (2012) 'This is what modern deregulation looks like': co-optation and contestation in the shaping of the UK's open government data initiative. J Community Inform 8(2). http://cijournal.net/index.php/ciej/article/view/845. Accessed 26 Sept 2017

Brown H (2016) Understanding the impact of data quality on data linkage. NHS England, Data Services for Commissioners (internal report)

Caldicott F (1997) Report on the review of patient identifiable information. Department of Health, London

Caldicott F (2013) To share or not to share? The information governance review. Department of Health, London

Caldicott F (2016) Review of data security, consent and opt-outs. Department of Health, London

Carter P, Laurie GT, Dixon-Woods M (2015) The social license for research: why care.data ran into trouble. J Med Ethics 41:404–409. doi:10.1136/medethics-2014-102374

Data Documentation Initiative (2014) Research data lifecycle model. http://www.ddialliance.org. Accessed 18 Sept 2017

Davies TG, Bawa ZA (2012) The promises and perils of open government data (OGD). J Commun Inform 8(2). http://ci-journal.net/index.php/ciej/article/view/929/926. Accessed 26 Sept 2017

Department for Education (2013) Prescribed persons consultation response final. https://www.education.gov.uk/consultations/downloadableDocs/Prescribed%20Persons%20Consultation%20ResponseFinal.pdf. Accessed 18 Sept 2017

Financial Times (2016) Fears raised over Google's DeepMind deal to use NHS medical data. https://www.ft.com/content/f6bcce6e-b099-11e6-9c37-5787335499a0. Accessed 19 Sept 2017

Fukuyama F (1995) Trust: the social virtues and the creation of prosperity. Free Press, New York

Giddens A (1991) The consequences of modernity. Polity Press, Cambridge

Grace J, Taylor MJ (2013) Disclosure of confidential patient information and the duty to consult: the role of the Health and Social Care Information Centre. Med Law Rev 21(3):415–447. doi:10.1093/medlaw/fwt013

Hardin R (2002) Trust and trustworthiness. Sage, London

Harrison E, Shepherd E, Flinn A (2015) InterPARES Report Team Europe EU19 project 2014–15: a research report into open government data in NHS England (unpublished)

Hays R, Daker-White G (2015) The care.data consensus? A qualitative analysis of opinions expressed on Twitter. BMC Public Health 15:838. doi:10.1186/s12889-015-2180-9

Health e-Research Centre (2016) Citizens' jury: health data on trial. https://www.herc.ac.uk/get-involved/citizens-jury/. Accessed 18 Sept 2017

Holton B (1994) Deciding to trust, coming to believe. Aust J Philos 72:63–76

Information Commissioner's Office (ICO) (2017) Royal Free London NHS Foundation Trust ruling. https://ico.org.uk/action-weve-taken/enforcement/royal-free-london-nhs-foundation-trust/. Accessed 18 Sept 2017

Ipsos MORI (2014) Dialogue on data: exploring the public's views on using administrative data for research purposes. https://www.ipsos.com/ipsos-mori/en-uk/dialogue-data. Accessed 18 Sept 2017

Ipsos MORI (2016) The one-way mirror: public attitudes to commercial access to health data. https://www.ipsos.com/ipsos-mori/en-uk/commercial-access-health-data. Accessed 18 Sept 2017

Jones K (1996) Trust as an affective attitude. Ethics 107(1):4–25

Jones KH, Ford DV, Jones C et al (2014) A case study of the Secure Anonymous Information Linkage (SAIL) gateway: a privacy-protecting remote access system for health-related research and evaluation. J Biomed Inform 50:196–204. doi:10.1016/j.jbi.2014.01.003

Luhmann N (1980) Trust: mechanism for the reduction of social complexity in trust and power. Wiley, New York

Luhmann N (2000) Familiarity, confidence, trust: problems and alternatives. In: Gambetta D (ed) Trust: making and breaking cooperative relations. Department of Sociology, University of Oxford, pp 94–107. https://philpapers.org/rec/LUHFCT. Accessed 18 Sept 2017

MacNeil H (2011) Trust and professional identity: narratives, counter-narratives and lingering ambiguities. Arch Sci 11(3/4):175–192

Mail Online (2017) NHS data chief tried to block officials from receiving patient records to help them trace illegal immigrants. http://www.dailymail.co.uk/news/article-4182344/NHS-data-chief-refused-illegal-migrants-data.html. Accessed 18 Sept 2017

Mann N (2016) Learn from the mistakes of care.data. BMJ 354:i4289

MoU (2017) NHS digital, home office, department of health, memorandum of understanding between health and social care information centre and the home office and the department of health. http://kingsfund.blogs.com/health_management/2017/01/memorandum-of-understanding-between-the-home-office-nhs-digital-and-the-department-of-health.html. Accessed 26 Sept 2017

Nickel P (2007) Trust and obligation–ascription. Ethical Theory Moral Pract 10(3):309–319. doi:10.1007/s10677-007-9069-3

Nuffield Council on Bioethics (2014) The collection, linking and use of data in biomedical research and health care: ethical issues. Nuffield Council on Bioethics, London

O'Neill O (2002a) Autonomy and trust in bioethics. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511606250

O'Neill O (2002b) A question of trust. Cambridge University Press, Cambridge

Partridge N (2014) Review of data releases by the NHS Information Centre. Health and Social Care Information Centre, Leeds

Putnam R (1993) Making democracy work: civic traditions in modern Italy. Princeton University Press, New Jersey

Scottish Health Informatics Programme (SHIP) (2017) Core programme 4: public engagement. http://www.scot-ship.ac.uk/c4.html. Accessed 19 Sept 2017

Secure Anonymised Information Linkage (SAIL) Databank (2017) Public engagement. https://saildatabank.com/about-us/public-engagement/. Accessed 19 Sept 2017

Seemann A (2007) An interview with Onora O'Neill. Philos Manag 6(2):3–8. doi:10.5840/pom20086230

Six B, Zimmeren E, van Popa F, Frison C (2015) Trust and social capital in the design and evolution of institutions for collective action. Int J Commons. doi:10.18352/ijc.435

Stake RE (2005) Qualitative case studies. In: Denzin NK, Lincoln YS (eds) The Sage handbook of qualitative research. Sage, Thousand Oaks, pp 443–466

Sundqvist A (2011) Documentation practices and recordkeeping: a matter of trust or distrust? Arch Sci 11(3/4):277–291. doi:10.1007/s10502-011-9160-3

The Guardian (2016a) NHS seeks cure for its costly digital headache. https://www.theguardian.com/healthcare-network/2016/jul/01/nhs-seeks-cure-costly-digital-headache. Accessed 19 Sept 2017

The Guardian (2016b) Google given access to healthcare data of up to 1.6 million patients. https://www.theguardian.com/technology/2016/may/04/google-deepmind-access-healthcare-data-patients. Accessed 19 Sept 2017

The Guardian (2016c) Pupil data shared with Home Office to 'create hostile environment' for illegal migrants. https://www.theguardian.com/uk-news/2016/dec/15/pupil-data-shared-with-home-office-to-identify-illegal-migrants. Accessed 19 Sept 2017

Thompson J (2000) Political scandal: power and visibility in the media age. Polity, Cambridge

Torjesen I (2014) New patient database could undermine trust in NHS, risk analysis concludes. BMJ 348:g1624. doi:10.1136/bmj.g1624

UK Government (1998) Data Protection Act 1998 CHAPTER 29. http://www.legislation.gov.uk/ukpga/1998/29/contents. Accessed 02 Feb 2017

van Staa TP, Goldacre B, Buchan I, Smeeth L (2016) Big health data: the need to earn public trust. BMJ. doi:10.1136/bmj.i3636

Vezyridis P, Timmons S (2017) Understanding the care.data conundrum: new information flows for economic growth. Big Data Soc. doi:10.1177/2053951716688490

Yoon A (2014) End users' trust in data repositories: definition and influences on trust development. Arch Sci 14(1):17–34. doi:10.1007/s10502-013-9207-8

Yoon A (2017) Data reusers' trust development. J Assoc Inf Sci Technol: JASIST 64(8):946–956. doi:10.1002/asi.23730

Anna Sexton is Head of Research at the UK National Archives. She was previously Research Associate on the ADRC-E research project at UCL, having completed her doctoral thesis which used action research to explore participatory approaches to building life history archives in the context of mental health recovery, funded by an AHRC Collaborative Doctoral award with Wellcome Library. She is a qualified archivist and was previously Archives Manager at Peterborough Archives Service.

Elizabeth Shepherd is Professor of Archives and Records Management at UCL in the Department of Information Studies, where she is also Director of Research. Her research interests are in the management of public sector records and data and in the history of archives in twentieth century England. She teaches on the MA in Archives and Records Management at UCL.

Oliver Duke-Williams is a Senior Lecturer in Digital Information Studies in the Department of Information Studies at UCL and was previously a Senior Research Fellow at the School of Geography, University of Leeds. His research interests include web-based access to geographic data; disclosure control issues, especially those associated with migration and commuting data, and the past, present and future of census taking and other demographic information capture in the UK.

Alexandra Eveleigh is Collections Information Manager at the Wellcome Library, UK. She was previously Research Associate on the ADRC-E research project at UCL; having completed her doctoral research at UCL entitled 'Crowding out the Archivist? Implications of online user participation for archival theory and practice' (funded by an AHRC Collaborative Doctoral Award with The National Archives). She is also a qualified archivist and worked for a number of years in local government archive services.