# Better Data in Government Consultation Response

### Submission from defenddigitalme - April 2016

## About defenddigitalme

defenddigitalme is a campaign group for children's privacy rights formed in mid 2015 in response to concerns from parents and privacy advocates about increasingly invasive uses of children's personal data in education. The campaign asks the Department for Education (DfE) to change their policies and practices to protect 20 million children's identifiable personal data in the National Pupil Database (NPD):

- stop handing out identifiable personal data to commercial third parties and press without consent
- start telling pupils, their guardians and schools what DfE does with personal data
- be transparent about policy and practice

We seek future-proofed and ethical, legal and regulatory frameworks in data policy and practice. More information: http://defenddigitalme.com/

In the interests of full disclosure we also mention that our Coordinator, Jen Persson is one of two lay members of the Administrative Data Research Network (ADRN) Approvals Panel.[1] She attended the consultation and code of practice Open Policy meetings in 2016.

## Summary

The Ministerial foreword[2] says: "we need to keep pace with both rising public expectations and the availability of new technology," and so we bring some of the policy making discussion on these themes into the written consultation. The structure of the consultation is such that themes cross into multiple questions, but the questions do not address all the issues. We respond instead with a structured approach that cuts across the six strands of the consultation.

We provide some specific examples of **positive public engagement** and hands-on research from **public attitudes to administrative data** sharing from 2013-6 to include wider voice.

While this legislation is designed to clarify the question 'can we' share individuals' personal data without consent, it offers little to support Public Service delivery users deciding 'should we.' We therefore suggest thorough attention is given to address the gap in an **ethics framework provision**.

Children, need particular attention given when sharing their data, and must trust that data given for one thing today, are not used for something different tomorrow, or that they or their future relations find themselves adversely affected as adults by decisions made today[3] or in **'Troubled Families'**. We outline **the National Pupil Database as a case study** of how-not-to-do datasharing legislation.

And we suggest some areas of consideration in the area of **'new technology'** that require greater thought in this legislation, which is written based on past and current practice, but is unfit for future data policy and lacks any ethical framework to support good practices, as technology changes.

---

[1] https://adrn.ac.uk/application-process/approvals-panel/

[2] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/100807/file47158.pdf

[3] http://blog.23andme.com/23andme-customer-stories/an-unexpected-discovery/

# Contents



# Introduction

We do not address the increased access to data for statistical purposes or in any depth the access to de-identified data for research purposes. These data uses are in safe settings, by safe people: data prepared by qualified data professionals, using trained data handling techniques, and raw data cannot leave the safe setting with the researcher. Those researchers use data, but don't hold onto it. We do however make reference to the public engagement work done by those data stakeholders.

Our concerns focus on increased access to and use of identifiable personal data in and across public services and with commercial energy companies, with a specific interest in children's privacy rights.

We agree that clear and consistent approaches will be beneficial across Departments and at local level if the public is to understand how government and citizens interact.

Consistent datasharing should seek to meet the best of current practice and most challenging of public expectation and engagement. Quick fix approaches are not future-proofed and at best in the past have only delayed public contempt when they went wrong, with damaging consequences to already fragile public trust, as proven in care.data[4], jeopardising the opportunities that good datasharing can offer. We suggest more forward thinking on technology and future uses of data.

It is positive that these proposals are 'not about selling data, collecting new data from citizens or weakening the Data Protection Act 1998' and we hope to see consistently good practices across government Departments develop over time. Right now this consultation could offer an opportunity for fixing where public bodies are selling public or personal data, collecting data from citizens into big databases, actions or principles that are out of step with the Data Protection Act 1998 today.

We suggest that in order to avoid arriving at the simplistic outcome from the thinking 'make it easier to share the public's confidential and personal data legally without consent' there needs to be a more detailed grasp of what these data plans (varying in scope) each change in practice, compared with what is done today. 'We don't share data very well' will not be improved by sharing more data.

Today, as a starting point we therefore propose thorough consideration is given to solutions looking back at some of the well known issues of data collection and releases in practice, and how they impact today, as was discussed in the policy discussions but not included in the consultation paper. Lessons need learned from applied practices which have been consistently in need of change over the last ten years;[5] investment of time and funding of training, understanding and applying existing DPA legislation, applying knowledge gained from data, from failures, and accountability.

---

[4] http://www.theguardian.com/commentisfree/2014/feb/28/care-data-is-in-chaos

[5] http://www.fipr.org/childrens_databases.pdf

# 1. The Current Landscape

At UK level, we should implement the best of what has already been proposed by the Select Committee. The consultation should support the call contained in the 2014 report 'Responsible use of data', *"the Government has a clear responsibility to explain how personal data is being used.*[6]

In line with current pro-European government talk, it is time now to act to support our citizens human rights. Government Departments should act on the CJEU in 2015[7] ruling that reiterated the existing need for all public bodies to ensure fair processing before sharing personal data with any other public body, as enshrined in Principle 1 of the Data Protection Act 1998.

The European Union is introducing a General Data Protection Regulation (GDPR)" [Consultation p8 item 25.][8] How this will safeguard children's rights should also be actively considered.

# 2. The United Nations Convention on Rights of the Child[9], especially the following articles should be given proper consideration

Article 3: the best interests of the child must be a top priority in all things that affect children
Article 12: a right to be heard and express views in decisions about them
Article 16: Every child has a right to privacy. The law should protect a child's family and home life
Article 29: Education must […] encourage the child's respect for human rights, as well as for their parents, their own and other cultures, and the environment

While the Consultation makes reference to The Human Rights Act 1998 and the European Convention on Human Rights and the DPA 1998, it does not mention children's rights legislation.

*"Public authorities accessing and disclosing information under the proposed powers will need to ensure compliance with the Human Rights Act 1998, in particular Article 8 of the European Convention on Human Rights and other relevant measures relating to data protection set out in law." [Consultation p9 item 26]* [10]

We ask the consultation considers children's rights where their data will be included in changes.

We seek solutions to concerns with opening up vulnerable communities' and individuals' personal data in the name of helping, but without their consent, to wider audiences of police, debt collectors, energy companies and other public bodies, especially where children may be 'labelled for life'.

We include a case study from an NGO involved in education that was brought to our attention recently. Legislation designed to target individuals with the best of intents, but without safeguards, could easily become authoritarian if misused. Safeguards and transparency are needed to avoid this.

---

[6] http://www.publications.parliament.uk/pa/cm201415/cmselect/cmsctech/245/245.pdf

[7] http://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150110en.pdf

[8] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf

[9] http://www.unicef.org.uk/Documents/Publication-pdfs/UNCRC_PRESS200910web.pdf

[10] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf

# 3. Public Engagement on Administrative Data in Research

What does the wider public know or want to happen on the use of their personal data by others?

In 2013, the ESRC collaborated with the Office for National Statistics (ONS) to run public dialogues across the UK to understand how people view using and linking our data for research.

'The dialogues, run by Ipsos MORI, comprised of events in Manchester, London, Stirling, Cardiff, Wrexham, King's Lynn and Belfast. During the two day-long sessions, participants - recruited from a cross-section of people - worked with trained facilitators and experts to discuss the challenges of linking administrative data for research purposes.

The workshops explored attitudes around the re-use of sensitive data, mandatory and voluntary data collection and long-term data storage and data linking, and, specifically examined the re-use of public data for research purposes in de-identified formats in safe settings.

Later, in 2014 the Royal Statistical Society carried out research and found nearly all institutions suffer a "data-trust-deficit". Trust in them to use data appropriately is lower than trust generally.

These two sample research projects in 2013 and 2014 identified similar and consistent public opinions and concerns around administrative data sharing and public benefit, and was echoed in the care.data public debates of 2014. Generally there is a consistent lack of trust in government uses of data for individuals' well being, and trust in private companies' motivations is low. Strong public support exists for public benefit research and equally strong, lack of support in 'for-profit' uses. [11]

**Similar concerns included:**

- Lack of consent to data used about them
- Risks of profiling or pigeonholing individuals or areas
- Data accuracy and poor quality data, resulting in poor quality decisions and policy making
- Data Protection Act: widespread cynicism about others' respect for this and its enforcement

**Red lines included:**

- Allowing administrative data to be linked with business data
- Linking of passively collected administrative data, in particular geo-location data
- Allowing researchers for private companies to access data, either to deliver a public service or in order to make profit.
- Remote access to de-identified data safe settings was a no-no
- Creating large databases containing many variables/data from a large number of public sector sources

The consultation foreword[12] says: "we need to keep pace with […] rising public expectations " - public expectations are clear: to understand how data about them are used, for what and by whom. Linking of passively collected data, allowing administrative data to be linked with business data, private companies access for profit, profiling, and large databases collections are all unpopular.

How will the legislation and code of practice cater for this?

---

[11] http://www.esrc.ac.uk/public-engagement/public-dialogues/public-dialogues-on-using-administrative-data/

[12] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/100807/file47158.pdf

## 3a. 2013 Dialogue on Data[13] : ESRC & ONS workshops across the UK

*Sample comments:*

*"Concerns that administrative data could be inaccurate, especially where it is self-reported, and that this could have negative consequences for individuals or groups. For example: Using linked administrative data to justify and implement controversial policies, such as the bedroom tax."*

*"Further public dialogue or research would be needed for any expansion of the ADRNs remit. Specifically, further research should be done to understand the public's views on allowing businesses to access linked administrative data." [page 6]*

*"Profiling or pigeonholing individuals or areas. This was raised with reference to the case study examples of the National Pupil Database and the Index of Multiple Deprivation. While participants could see the value of both sets of linked data for improving services and allocating resources, they also thought that they could lead to unintended negative outcomes for specific types of people or those living in particular areas." [page 19]*

*"A few participants who had multiple interactions with government services repeatedly returned to the issue of lack of consent for further uses. They thought the lack of explicit consent had a bearing on whether data should be shared and linked or not–for example a participant who had been in care strongly disagreed with his data being linked for research purposes without his explicit permission."*

*"Concerns driven by experience (either personal or through friends and family) of their data being shared or sold by private companies without their express permission."*

*"Participants were also worried about personal data being leaked, lost, shared or sold by government departments to third parties, particularly commercial companies. Several participants had experiences that made them think that hospitals pass on data about those who have been in accidents to insurance companies."*

*"Low trust in government more generally seemed to be driving these views."*

Accountability is important but undefined in the consultation or code of practice. Challenging new agreements as a whole, and the outcomes of specific use of data both need named persons/positions as are assigned in the applications process for academic research uses of data and the accountable person is named in the ethics application for review at the ethics committee stage.

Named accountable data owners in every use of public data as done in research, may be helpful.

[13] http://www.esrc.ac.uk/files/public-engagement/public-dialogues/dialogue-on-data-exploring-the-public-s-views-on-using-linked-administrative-data-for-research-purposes/

## 3b. 2014 Ipsos MORI for the Royal Statistical Society: the Data Trust Deficit[14]

This research identified similar and consistent public opinions and concerns around administrative data sharing and public benefit.

A whopping 50% disagreed government and public services have the citizen's best interest at heart when they use their personal data, while only 11% tended to agree that they did.

The inverse was similar. A majority 63% felt government and public services *"used my personal data, for their benefit, not mine."*

*"The trust in companies using personal data for the best interests of the individual not the company was even lower, at 6%."*

*"However, even within the public sector, public support for data-sharing is very dependent on the exact situation, and need reassurances to allay their concerns."*

*"When asked to give their overall views towards data-sharing within government, people are more worried about the risks to their privacy and security than the benefits that data-sharing might bring, by 44% to 33%. But when told that data will be anonymised, this turns around to a majority in favour of data-sharing (55%, to 28% who think the risks outweigh the benefits)."*

# 4. Ethics - can we vs should we? A decision making support tool

Without a strong ethical framework in either legislation or code-of-practice, it is hard to see how decisions will be applied consistently or always in the spirit that legislation was intended.

The de-identified academic research strand has ethical review built-by-design into every data use in the form of RECs as part of the application process. These consider whether the research will have contact with individuals (rare) or whether the outputs from the research could be discriminatory.

What it does not do is consider individual ethics which are not taken into account.

Some individuals would, given the choice, object to their data being used at all in research. For others, some object based on conscientious objection;  for example related to contraceptives, or pregnancy terminations, or having their children's data included in MOD research. [15]

Their rights of objection should be addressed if research intends to aim for high standards across the board. The right to opt out of some identifiable data sharing in health research will take effect shortly. If ethical, data uses should not be exploitative, risk harm, prejudice, identification and more.

---

[14] https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-lessons-for-policymakers.aspx

[15] http://schoolsweek.co.uk/mod-makes-inappropriate-request-by-mistake/

The 2013 Dialogue on Data[16] : ESRC and ONS surveys and fourteen workshops across the UK found that there were boundaries for the most sensitive data, even for use in research in de-identified form, rightly realising that data are either aggregated statistics or are at individual level and identifiable or potentially identifying.

Hence the reason why academic research using linked data well, is carried out in safe settings, by trained accredited researchers, after projects go through an ethical review process.

*"Within this overall view though, some particularly sensitive types of information were seen as too personal to be shared outside of the agency that collected it, for example records of domestic violence, or HIV status, because of the potential consequences of the data getting into the wrong hands." [ESRC Dialogue on data, 2013]*

Sensitive, fully identifying data will not be used in safe settings for public services targeting, but at ordinary work desks, and not by managed researchers[17] after a consistent accreditation process to ensure a standard approach to data analysis, protection, or practical handling, but staff who have been shown 'what they need to know'. Using identifiable data, without any ethical review.

What contrast between the safe-setting best practices for only potentially identifying data and the keep-your-fingers-crossed potential risks of increasing identifiable sensitive data used 'in the wild'.

From the public body side, sharing identifiable data for any indirect uses, code of practice might consider what safeguards are in place to prevent unethical requests being rejected by one body but accepted by another, or decisions based on who gets the application on which day of the week.

If identifiable data are to be used only for direct interventions then it should be more clearly defined what does not already exist today that is required. What don't we do today in our public services that will be done to individuals in future as a result of these changes?

Ethical considerations in research[18] may be a better basis for citizens, than 'user needs' to prioritise in codes of conduct and should include considerations of autonomy, anonymity and risk analysis.

Using an individual's data should not be punitive or harmful to that individual, but in the case of families, what may be good for one, could be seen as detrimental to other family members. What ethical considerations are to be made for data sharing for the tailored public services sharing?

Academic research creating heat list mapping using predictive technologies are thorny ethical issues that are already in play in the UK, and applied with police in practice and in social media projects[19]. What solutions are there to known problems[20] and will these changes address or ignore them? Personal data is increasingly commodified, and this is widely felt unacceptable [21]. For many children their lack of ability to give informed consent exploits trust, so additional protections must be assured by providers/users. It is unethical our children's personal data are exploited today.

---

[16] http://www.esrc.ac.uk/files/public-engagement/public-dialogues/dialogue-on-data-exploring-the-public-s-views-on-using-linked-administrative-data-for-research-purposes/

[17] http://eprints.uwe.ac.uk/22329/2/_nsta-uwe12_users$_fj-ritchie_Windows_Downloads_wp.15.e.pdf Desai, T, Ritchie, F (2009) Effective researcher management

[18] https://adrn.ac.uk/media/1172/ethics-and-administrative-data-guidance_00_08_pub.pdf

[19] http://www.cosmosproject.net/

[20] http://research.gold.ac.uk/11079/1/mcquillan-algorithmic-states-of-exception.pdf

[21] http://www.esrc.ac.uk/public-engagement/public-dialogues/

## 5. Prescribed Persons: A Current Case Study - The National Pupil Database

The National Pupil Database, is a case study in data handling not fit for the volume and the sensitivity of administrative data that it contains that can be shared at the press of a return key.

The National Pupil Database is now "one of the richest education datasets in the world" [22] holding a wide range of confidential and sensitive personal data from almost every child in England since 2000, through their education from age 2 to 19, and now includes nearly 20 million individuals.[23]

Today's practices were found lacking by The Government Internal Audit Agency (GIAA) who have rated assurance as 'Limited'. Improvements should be made over vetting and validation of applications to access the National Pupil Database, information retention procedures, and data handling guidance, according to The DfE Consolidated Annual Report and Accounts 2014-15 published on April 20, 2016.[24] [p40]

The UK Statistics Authority[25] also support our position in April 2016, that increased transparency to parents and pupils is required, as well as improved arrangements for ensuring the secure handling and end of project functions of NPD data (for example, ensuring that data is destroyed by third parties post research).

Confidential personal data, and sensitive or highly sensitive data are shared with a wide range of third parties. These data include candidate numbers, names, ethnicity and disability, special needs, detailed breakdowns of special educational needs, whether the child has local authority looked after status, are children of military service personnel, their reasons for absence or exclusion and whether they are recipients of free school meals, as well as a lifetime of educational attainment and notes.

Digital identity and biometrics are now literally our passport to the world. Our online and offline identity, and associated identifying data, online rights and responsibilities, should be equally respected and protected. How will all these sensitive data be future proofed to protect their futures?

It is not the stripping of identifiers that makes data secure. If that were the case then de-identified data would be handed out freely, not treated with the respect that serious data users give it. Inside a safe setting trained data users can take only a pencil, no phone. There are cameras, there is no internet. They can use the data, but not take it away. Safe settings, safe users and safe data.

In Department practice today, this is not the case. The Department for Education gives out copies of nearly 20 million children's confidential personal data from the National Pupil Database to commercial third parties and press own settings, without informing schools, parents or pupils. [26]

Releases [27] include giving out named data, and sensitive data to television and Fleet Street press.

---

[22] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/472700/NPD_user_guide.pdf

[23] https://www.whatdotheyknow.com/request/pupil_data_national_pupil_databa_2#incoming-764676

[24] http://defenddigitalme.com/2016/04/improvement-needed-over-access-to-childrens-national-database/

[25] https://www.statisticsauthority.gov.uk/wp-content/uploads/2016/04/Letter-from-Ed-Humpherson-to-Jen-Persson-220416.pdf

[26] http://defenddigitalme.com/ and https://www.whatdotheyknow.com/request/pupil_data_national_pupil_databa_2?nocache=incoming-764676#incoming-764676

[27] http://defenddigitalme.com/wp-content/uploads/2016/04/DDM_shared_examples_April2016.pdf

**Example 1: TV journalist, identifiable, highly sensitive personal data - population-wide data**

A BBC television journalist in August 2014 was given Tier 1 identifying and highly sensitive data. They describe in their application how they will take small number rules into account using the data, because they are identifying and *"School-level data is not helpful."*

A full copy of the data application request was obtained through a Freedom of Information Request and is publicly available via What Do They Know. [28]

**Example 2: Newspaper journalists, ca 10 million children, sensitive identifiable personal data**

Ten Telegraph journalists were given the personal data of ca. 10m children in February 2013.[29] The newspaper offered *"cast iron assurances that no pupil will be identified through our use of the data."* and received 5 years worth of identifying individual level and sensitive Tier 2 data. [see letter[30]] If the data were not identifying there would be no need to offer this assurance.

Processing sensitive data requires additional DPA conditions to be met.[31] The business cases in applications by journalists do not contain any indication how they meet schedule 3 conditions.

There is no proven 'need' nor 'benefit' to the individuals of releasing these identifying sensitive data like SEN or ethnicity. No audit had ever been done at the time we asked DfE this, to identify whether any benefits of the data release [see FOI 20 Aug 2015] [32] were achieved to meet the legal requirement for release: *"promoting the education or well-being of children in England."*[33]

## 5b. The wording of the Legislation matters - purposes and persons

Section 114 of the Education Act 2005, and section 537A of the Education Act 1996, together with the 2009 Prescribed Persons Act, updated in 2013, specifically allows the release of individual children's data to third parties which in practice has permitted data to get given to journalists, commercial third parties and charities.

*Education (Individual Pupil Information) (Prescribed Persons) (England) Regulations 2009 (Amended 2013).*
*persons who, for the purpose of promoting the education or well-being of children in England are—*

*(i) conducting research or analysis,*
*(ii) producing statistics, or*
*(iii) providing information, advice or guidance,*
*and who require individual pupil information for that purpose.*

How will prescribed purposes, persons and public bodies in the 'Public Services' legislation compare? Safeguards must prevent future scope creep as has happened to our pupil data.

---

[28] https://www.whatdotheyknow.com/request/293030/response/723407/attach/10/BBC%20Newsnight.pdf

[29] https://www.whatdotheyknow.com/request/293030/response/723407/attach/3/Daily%20Telegraph.pdf

[30] https://www.whatdotheyknow.com/request/293030/response/738135/attach/2/Annex.pdf

[31] https://ico.org.uk/for-organisations/guide-to-data-protection/conditions-for-processing/

[32] https://www.whatdotheyknow.com/request/pupil_data_application_approvals#outgoing-482241

[33] http://www.legislation.gov.uk/uksi/2013/1193/regulation/2/made

## 6. Considerations for particular regard to Children and Young People

We have three key concerns in regards the new powers specific to children and young people.

1. That the powers under Tailored Public Services are targeting, labelling and may last a lifetime
2. Decisions may be based on technology that children or their guardians cannot see or perhaps understand and are therefore disempowered by
3. That current powers may be sought to be misused and we have little safeguards in place to protect children from them today. Increasing data sharing may further increase this risk.

The case study we wish to present is of a non-governmental educational body which told us of an incident in which a sixteen year old was a victim of a crime. A police officer requested unrestricted access to the school pupil record, "to see if they could find a reason why the pupil would have had anything to do with what happened." The Data Compliance Officer was most concerned, knowing the pupil's record and knowing it contained nothing of any relevance and that there was no grounds for a search of the record. It was 'a fishing expedition', yet the police were insistent. On asking for further information and raising an objection, the school Data Compliance Officer was threatened with being taken to court.

To which the DCO told us, *"it was ironic, I was being threatened with the law, but if I did give the police officer what I felt was unnecessary, it could result in me breaking the law."*

The school Data Protection Office knew the student, the situation, and the school record, and the current law well enough to give robust defence of why it should not be shared. If a law is created for which the presumption is data would be shared between the prescribed bodies, how would a similar situation be in the best interests of the child and support the DCO to defend the pupil's privacy - better or worse than today?

## 6b. Geographical and Devolution Considerations for Children

Laws particular to children are not consistent across the UK. The geographical scope and issues specific to devolution will potentially affect what data are available to copy and transfer elsewhere.

The consultation says that:"Discussions have taken place with officials in the devolved administrations about the proposals and how they might allow for UK-wide coverage should their Ministers and legislatures wish to adopt this legislation."[34]

Data collection and dissemination is different between the devolved nations. For example, the highly controversial role of a Named Person is a key part of the Children and Young People (Scotland) 2014 Act[35] will start to collect vast amounts of subjective opinion on a wide range of "risk indicators"[36] for every child in Scotland from August 2016 and has resulted in backlash[37]. Should public bodies be permitted to share across boundaries, one could see that the level of invasion of privacy for children in some regions could be potentially of vastly different degrees.

Would English police be able to access Scottish Named Persons data on children for example?

---

[34] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf Consultation paper p 10, item 30

[35] http://www.gov.scot/Topics/People/Young-People/gettingitright/about-named-person

[36] http://www.gov.scot/Publications/2012/11/7143/9

[37] http://no2np.org/named-person/

In general regards children and young people's data sharing we need "*a new framework for child protection, provision and participation online that results in clear and effective policy that is born of real needs, targets specific and evidence-based risks, includes measurable goals [...] policy implementation is independently evaluated.*"[38] *[Livingstone, Sonia and O'Neill, Brian (2014)]*

There are no clear legal obligations of confidentiality that apply to the deceased but they vary in different places and should be considered.

We support the work and aims of the Royal Statistical Society in their efforts for the timely registration of deaths and the production of statistics.

Nevertheless since the scope of legislation is cross border, and in Northern Ireland, for example, the DHSSPS, Department of Health and the General Medical Council previously agreed there is an ethical obligation requiring that confidentiality obligations continue to apply after death.[39]

The Common Law Duty of Confidentiality arguably applies to deceased patients' records, as per the Information Tribunal Appeal Number: EA/2006/0010 of 17 Sep 2007 between Pauline Bluck, the Information Commissioner and Epsom & St Helier University NHS Trust and Lewis v Secretary of State for Health [2008] EWHC 2196. [40]

In a world in which genomic data will live on longer the individual, to whom children are related and identifiable by their genomes, and for which the uses are not yet clear, we believe ethical considerations could be made for data use and sharing beyond death.

# 7. Plans for Protecting Privacy Today and Predictive Technology?

The outcomes from data used predictively based on algorithms suggesting a combination of factors and predicting likelihood, does not mean that a child will, for example, commit a crime, play truant, or become an underage parent. The reliance on risk factors could easily switch from seeing children 'at risk' who need welfare care and safeguarding, to 'at risk of becoming a burden to the State'.

This discussion is elegantly outlined by Anderson, R; Brown, I; Clayton, R; Dowty, T; Korff, D; Munro, E; in 'Children's Databases - Safety and Privacy a report by the The Foundation for Information Policy Research (FIPR) for the Information Commissioner's Office' in 2006.

*"One of the main issues identified … is the shift in meaning of the term "at risk" as used in work with children, from "at risk of significant harm or neglect" to "at risk from failing to achieve the government's five targets for children" and "at risk of social exclusion". If the purpose of data collection, processing and sharing is defined as "protecting children at risk" in these very broad senses, then clearly this shift in meaning leads to a major widening of the "specified purpose" for which the data are processed. The question then arises whether the wider purposes are still sufficiently specific in terms of the DPA1998."*

A case study example in the Database State, 2009, Professor Ian Brown, Ross Anderson, Terri Dowty et al in 2008 and in 'Stephen's case study'.[41] shows how a lifetime of labelling can have an adverse affect. Further, providers store schoolchildren's biometric data[42] which needs broad ethical consideration.

---

[38] http://eprints.lse.ac.uk/62276/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_Livingstone,%20S_Childrens%20rights%20online_Livingstone_Childrens%20rights%20online_2015.pdf Livingstone, Sonia and O'Neill, Brian (2014) Children's rights online: challenges, dilemmas and emerging directions

[39] http://www.publichealth.hscni.net/sites/default/files/good-mamagement-good-records_0.pdf

[40] https://www.dhsspsni.gov.uk/articles/common-law-duty-confidentiality

[41] A report commissioned by the Joseph Rowntree Reform Trust Ltd to map the main central UK government databases. http://www.jrrt.org.uk/sites/jrrt.org.uk/files/documents/database-state.pdf

[42] https://www.bigbrotherwatch.org.uk/files/reports/Biometrics_final.pdf

## 8. The Data Protection Act in practice: Fair Processing

*"A key guiding principle of the open policy making process was that the DPA 1998 should not be weakened." [Consultation paper.[43]]*

The principles of the Data Protection Act 1998 appear selectively included in the consultation. Principle 1: fairness is required across all six of these data sharing strands.

Public bodies have a legal duty which was reiterated by the CJEU in 2015[44], to ensure fair processing before sharing personal data with another public body.

Once it has been established that a data controller does have the "lawful" power to share personal data it would then need to satisfy a Schedule 2 condition for processing and where sensitive personal data is involved, a Schedule 3 condition. It should be remembered though that even where a condition or conditions for processing can be met this will not on its own ensure that the processing is fair or lawful. These issues need to be considered separately.

The organisation should make it clear to individuals about how their information is being used and where they can find out more about the processing and/or object to the processing (the latter point covering s10 of the DPA).

It is also important to ensure that the other Data Protection principles are complied with eg the information shared needs to be relevant and not excessive, it must be accurate and kept up to date, not kept for longer than necessary and kept secure. Existing government databases do not meet these criteria, and we should be pleased to see this consultation take the opportunity to move to better datasharing across the board as an investment in ethical and engaged public involvement.

## 9a. Increased Datasharing for Debt

Creating a "single view of debtors", as discussed in the consultation discussions, requires a broader strategy on public debt management. Simply enabling the remaining 10% of debt owed to HMRC and DWP (the consultation impact assessment [45] itself states on page 6 that "there are already sufficient data sharing powers for data to be shared between HMRC and DWP which covers around 90% of the £24bn." On page 9 it appears to read that the 10% not covered by existing datasharing agreements is only £240 million.

Debt was brought back it appears hurriedly, into the 2016 discussions having been previously discussed and dropped, so there was little proper discussion as part of the process. However the business case is unclear and any framework of change management how it might be implemented, as well as safeguards for those who will be most impacted, defining "who cannot pay, and who will not pay", are missing. The ethical implications of automated targeting must also be clarified here.

A 'single view of debt' may indirectly disproportionately negatively affect children as is suggested did the single view of welfare payments, Universal Credit, and since it is potentially likely this

---

[43] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf page 6 item 17

[44] http://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150110en.pdf

[45] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/504023/Consultation_Stage_Impact_Assessment_for_Debt_power.pdf

'single view' would be outsourced to a central service[46] [Consultation, page 13, item 43], it is also potentially likely this 'single view' would not sufficiently take into account the personal factors and structural obstacles of individuals' lives[47] that face-to-face consensual discussion and agreement to data sharing can facilitate.

While the paper outlines the concept that people experiencing genuine difficulties can be offered the right support. and a managed payment plan can be tailored to take their personal circumstances into account, we are unable to see why existing arrangements (remembering they already cover an estimated 90% of datasharing requirements according to the Consultation Impact analysis) do not already enable this or how changes would positively impact what is done today in practice.

We would suggest it beneficial to seek separate inclusive discussion with third sector organisations most familiar with this specialist area return to the table rather than the proposal put back in again after they left when debt proposals had been taken out from discussions.

## 9b. Health data

Health data was explicitly excluded from the open policy making discussions and consultation. "Health and care data is particularly sensitive and rightly needs additional protections."

However not all health and social care data is held within healthcare organisations and some therefore will nonetheless potentially be more widely shared as a result of this proposed legislation. For example data types such as Special Needs and Types of Disability held on individuals in schools data at local and regional level and within the individual PLASC-type extractions of school data, and its linked master collection in the National Pupil Database.

It is a pity that the review by National Data Guardian, Dame Fiona Caldicott has been inexplicably delayed given that its expertise and recommendations would be helpful to understand and address consistency for similar data, stored in other silos, which may be shared as a result of this legislation.

It is also worth noting the health mention in the Civil Registration Information Impact Assessment:[48]

*"Information supplied could also benefit wider society in terms of providing data to deal with ad-hoc situations such as flu pandemics where there is currently no gateway in place to provide the information."*

Temporary population-wide needs for data sharing in scenarios such as pandemics have procedures in place already in Section 60 of the Health and Social Care Act 2001 as re-enacted by Section 251, NHS Act 2006.[49]

---

[46] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf

[47] http://www.welfareconditionality.ac.uk/wp-content/uploads/2014/09/Briefing_LoneParents_14.09.10_FINAL.pdf - S. Johnsen, (2014) Conditionality Briefing: Lone Parents

[48] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/504025/Consultation_Stage_Impact_Assessment_for_Civil_Registration_power.pdf

[49] http://www.hra.nhs.uk/documents/2014/02/cag-frequently-asked-questions-1.pdf

## 9c. Considerations specific to 'Troubled Families' datasharing

There is already extensive legislation[50] used in sharing data under the 'troubled families' banner and it reportedly comes with differing degrees of transparency to the people involved. The August 2015 privacy notice[51] for Troubled Families suggests that these data are already extracted together with control families' data for research purposes. Further, for example, Information can be shared with police forces under section 115 of the Crime and Disorder Act.[52] Therefore it is unclear why they need be included as Prescribed Persons / Public bodies on the new legislation.

There is also evidence that suggests every family included in the programme was turned around.

*"CLG told Manchester that it had precisely 2,385 troubled families, and that it was expected to find them and "turn them around"; in return, it would be paid £4,000 per family for doing so. Amazingly, Manchester did precisely that. Ditto Leeds. And Liverpool. And so on."* [53]

Given the evidence[54] that every family approached has been 'turned around', not one objected or slipped through the net, then it would appear the programme is entirely consensual. Why then datasharing legislation to enable non-consensual datasharing is required, is unclear.

Why data processing must sometimes 'necessarily be carried out without the explicit consent of the data subject being sought' so as not to prejudice the provision of that counselling, advice, support or other service, or for purposes of identifying fraud, is understandable. It is less clear how this applies to families with whom services intend to have direct interaction, which is consensual.

Where consent is not given[55], data should not be shared. This was the view given by the ICO in the past, "*we dislike is where "consent" is sought, it is refused but the data controller goes ahead anyway. That does not fit with how we view consent nor does it meet the Act's fair processing requirements*." And the duty of confidence and consent both have important considerations in the rights of the individuals involved to be respected, and transparency about data sharing.[56]

Although health data sharing is in itself outside of this consultation, one can look to its treatment of personal data and consent as an area with long established models of accepted practice.

The common law of confidentiality appears currently unfashionable but perhaps is the simplest and most effective principle to consider public expectations with a question: is information I give in confidence, kept in confidence? Children should be able to move into adulthood free from historical data labels. The 2009 Government document: Information Sharing: Further Guidance on Legal Issues[57] is very clear on the common law duty of confidentiality.

---

[50] such as existing legislation SI 417 (2000)

[51] http://defenddigitalme.com/wp-content/uploads/2016/04/Privacy_notice_for_the_evaluation_of_the_Troubled_Families_programme.pdf

[52] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/11469/2117840.pdf

[53] http://www.niesr.ac.uk/blog/troubling-attitude-statistics - cross-reference Stephen Crossley 'A kind of Trouble' (2015) https://akindoftrouble.wordpress.com/2015/03/13/the-troubled-families-programme-the-perfect-social-policy/

[54] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/410715/Final_The_Benefits_of_the_Troubled_Families_Programme_to_the_Taxpayer.pdf

[55] https://khub.net/c/document_library/get_file?uuid=adbd557e-5082-48a3-b915-8ed16273ba72&groupId=5404774 a past view of the ICO on consent in the Troubled Families programme

[56] https://ico.org.uk/media/1424185/webinar-questions-with-ico-and-dclg-responses-20150423.pdf

[57] http://webarchive.nationalarchives.gov.uk/20130401151715/https://www.education.gov.uk/publications/eOrderingDownload/Info-Sharing_legal-issues.pdf

As advancing technology permits wider data sharing of more data with less effort, and collections of data are getting ever larger and 'too big to ask' it is often easy to forget that for the individual it is still about them, their data, and bulk data sharing has impacts on individual lives.

Technologies that have become able to assess massive amounts of data in short time frames permit ever greater predictive uses of data and the consultation and legislation has not addressed data for the future but rather datasharing as participants are familiar with today.
*"Predictive algorithms increasingly manifest as a force-of which cannot be restrained by invoking privacy or data protection."* [58]

There are several considerations which I can only touch on here. More expert contributors in cyber security and children's data, should be asked to inform the legislation on this:

- Predictive uses of data are indirect secondary use of data
- Punitive uses of algorithm attributed decisions can be harder to challenge than a person attributable decision 'because the system says so' when both the data user and data subject cannot see inside, nor review all the data on which a decision may have been based.
- Punitive outcomes may occur through exclusion too, rather than directly from authoritarian use.

For children all of these mean a shift in power which is already imbalanced due to the nature of adult child relationships. Decisions made about them without them are contrary to the UN Convention on the Rights of the Child.

How predictive technology is likely to warp future face-to-face interactions i.e. the frontline interactions of services and people is of concern while it is in use but poorly understood by both users and the people whose data are being used by them.

*"Under the emerging regime of big data there is little enforceable in the idea of consent. Moreover, the notion of fair processing is completely obfuscated by the scale and nature of the algorithms being used."* [D.McQuillan, 2016]

However automated decision making has particular data protection legal requirements[59] that this consultation and its code of practice should take into consideration.

**Steps must be taken to safeguard where decisions are automated:**

- the legitimate interests of the individual, such as allowing them to appeal the decision.
- against risks as result of decision making driven by financial targets [60] [61]
- for users of the data, ensuring when decisions are based on an algorithm, we still trust experience

What might be considered as using data for an individual's benefit, is often determined through using massive amounts of data on speculation 'to see if someone qualifies'. These uses of data in health are however not considered direct, but secondary, indirect uses of data: Risk stratification.

**"Examples of activities would be risk prediction and stratification, service evaluation, needs assessment, financial audit."[62]**

---

[58] http://research.gold.ac.uk/11079/1/mcquillan-algorithmic-states-of-exception.pdf McQuillan, D (2015, forthcoming) 'Algorithmic States of Exception', European Journal of Cultural Studies, Vol 18, No 4-5, ISSN 1367-5494

[59] https://ico.org.uk/for-organisations/guide-to-data-protection/principle-6-rights/automated-decision-taking/

[60] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/11469/2117840.pdf

[61] http://www.niesr.ac.uk/blog/troubling-attitude-statistics

[62] http://www.hscic.gov.uk/article/3638/Personal-data-access-FAQs

And as in health for example, a lawful basis for the processing of data for these reasons where the data is for the purposes of direct care is where consent has been gained. The workaround relied on increasingly in the last five years, section 251 support for the relevant purposes, is problematic. It is a workaround, to avoid consent, designed for temporary emergencies like pandemics, it has become increasingly relied on for standard bulk data transfers seen as 'too big' to ask for consent.

Therein lies much of today's challenge to the rights of the individual. "We can't ask everyone," for consent in a massive study. At this point we again draw on the fact that these bulk datasets are made up of individuals data and on the public engagement work referenced earlier from the ESRC and ONS, the Royal Statistical Society on broad administrative data, and also most recently by Wellcome on health administrative data. Not everyone wants to opt out but most want asked. The research suggests that there is some increased sensitivity to health data sharing, there is a significant population of 17% who does not want their personal data shared without consent in any secondary non-direct care circumstances. Data shared at individual level, consent is possible to manage.[63]

The Wellcome report 2016[64], A One-Way Mirror: Public attitudes to commercial access to health data found:

*"that in order to have a trusted system for patient data use it is absolutely crucial that there is honest and open communication and engagement with the public about how their health data could be used for purposes beyond their care, and what safeguards are in place."*[65]

What efforts are being made for these communications to happen across public data?

## 10. Direct service does not imply indirect uses of data by default or forever

Across the indirect uses of data, an important principle is missing from the legislation and code of practice consultation discussion and written statements:

Direct service must be possible without implied secondary uses being a requirement. How will this be communicated to individuals, that they can use a public service without being required to consent to other uses of data?

Coercive consent is invalid as by its nature, consent is an agreement. Users also have a right to be informed that even if agreeing to a direct service they are not compelled to consent to data sharing with service B in order to be eligible for the provision of service A.

Simply by agreeing that data should be shared for the purpose of obtaining a driver's licence should not by default mean research uses of my personal data are a requirement. This is a longstanding principle of health data management until recently, when coercive use was attempted in 2013 in the care.data programme, based on the thinking 'no one who uses a service should be able to opt out'.

Communications will be important if the provision of fair processing information to the individuals is involved, with more information being required where the data sharing is more extensive and taking into account the added requirements for automated processing.

Consent must also be revocable. There must also be practicable ways to revoke consent once given, for example in research today, at any future date.

---

[63] http://www.hscic.gov.uk/catalogue/PUB20527/exp-care-info-choi-eng-ccg-apr-2016.pdf

[64] http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_grants/documents/web_document/wtp060244.pdf

[65] http://www.wellcome.ac.uk/News/Media-office/Press-releases/2016/WTP060240.htm

# 11. Specified consultation questions

**Question 2. Are there any public authorities that you consider would not fit under this definition?**
and
**Question three: Should non public sector bodies (such as private companies and charities) that fulfil a public service function to a public authority be included in the scope of the public service delivery power?**

The definition of public authority determines the criteria for which bodies can be added to the Schedule. It is intended to be broad enough to encompass all relevant public authorities. To achieve this the legislation defines a public authority as 'a person who exercises functions of a public nature'.

Should Father Christmas be added to the nice or the naughty list, is not a question we want this consultation to ask, yet the clause in subsection 3 proposes "a person who exercises functions of a public nature" - a definition so broad, that any person working for a local authority, indeed temporary Christmas staff, might be considered suitable to add to the list of Prescribed Persons by this standard.

It is this open definition which enables the legislation to stay flexible as needs change over time. It is the same openness that since 2012 allowed the Department for Education to hand out children's identifiable personal data from the National Pupil Database to data recipients considered suitable 'prescribed persons' that may meet legal requirements, but far exceed fair and reasonable public expectations today; namely journalists, charities and commercial companies.

The questions asked numerous times in the consultation discussion which did not get satisfactory answers to, included how private companies providing State services will be entitled to personal data they would not otherwise come into possession of, and what Chinese Walls will be sufficient safeguards within a firm in which one arm performs a public service but is also a for-profit player or offers punitive services. Key areas include Education, Prisons and Youth Offender Services.

A guiding principle of the policy making is that the uses of data are not to be punitive. How will the legislation ensure this stays the case when every police force is included on the list of specified public authorise for the public services data, Companies such as G4S, Serco, in punitive service delivery, also provide the services which will be seen as 'tailoring public service delivery'?

The current Code of Practice on principles for use of the power does not appear to address these risks or clearly exclude use with punitive or for-profit measures.

**Question 4. Are these the correct principles that should be set out in the Code of Practice for this power?**

A Department of Energy and Climate Change (DECC) proposal for data sharing to provide direct assistance to citizens living in fuel poverty is included in the package of measures, which was not part of the open policy process. [66] [p5, item 14a]

[66] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final__3_.pdf

We were given to understand that the process would be similar as to now, an annual date on which the Warm Home Discount (WHD) scheme assessment would be made and the information shared with the energy companies.

We would welcome that this measure may benefit "low income citizens of working age and families with children."

Two concerns are however unaddressed in the consultation.

The first is the assumption that the energy provider of the individual will be unknown by the state. Item 49 in the consultation suggests that the new power would "allow these data sets to be matched together within Government to identify the priority customers." We assume however that this does not identify which energy company, which customer is with. So it is likely that the identifying information, and flag of eligibility will be shared with 37 companies [67] or an unknown number in future across the UK, all but one of whom have no need to know, but would receive the individual's name, contact details and eligibility flag anyway. We are not sure how this disclosure will be justified given that it is neither necessary or proportionate.

Safeguards that prevent the energy companies then targeting a known vulnerable group, with marketing for example, would need to be specific as the consultation only defines purposes as "citizens being offered a service which aims to improve their wellbeing" and "the provision of assistance to citizens living in fuel poverty." The question asked includes the WHD "alongside information about energy efficiency support". We wonder how broad in time and content this is.

The second is to understand what safeguard is in place for the change to be in the best interest of the recipient, rather than the company, or indeed fuel poverty assistance paid for by the State. For example, today on a given annual date (we were told in July) those pensioners who qualify for the WHD rebate are automatically flagged to providers. Once you reach pensionable age you keep the status of pensioner.

The key difference that contrasts with those in a position of qualifying for a rebate based on welfare grounds, is that this status, unlike becoming a pensioner, can change.

If the means/qualifying test is intended to be carried out more frequently running up to winter, and automatically flagged by the system if no longer qualifying, it needs more safeguards than if annual.

Safeguards for when identification impacted the start and stoppage of the qualification for the energy rebate would be important to have in place, if the intent is to improve the welfare of the individual and not to simply identify the minimum number of individuals who qualify in real-time as opposed to over time, given the latter is a better measure of need.

**Question 15. Fees should not be charged by public authorities for providing data for research purposes.**

Public opinion is clear that any data transfers for which payments are made, is seen as selling. There is strong opposition to selling our personal data which make up public administrative data.

While we understand that costs are incurred in aggregating and cleaning work in order to prepare data for using in a range of formats, including as Open Data, the public benefit should outweigh the

[67] http://switch.which.co.uk/energy-suppliers/suppliers-atoz.html

cost and the cost will act as a barrier to access. Any fees charged already today as cost recouping, should ensure public interest research projects are not disadvantaged by commercial companies.

However this should not mean that the data could be repacked and sold, or the taxpayer is subsidising private gain. Referring again to the public engagement work since 2013 included in this submission, public interest research has wide support. For-profit uses do not.

Indirect or hidden costs, such as DBA checks recently imposed at the Department for Education for academics using data, should be taken into account as well.

**Question 16. Rejected applications with the reasons as well as accepted applications for research purposes should be published.**

Public transparency supports trust and professional accountability for the approvals panel and for the applicants. It may also foster discussion for any rejections in other data users or applicants. In the past year only three applications have been rejected to the best of my knowledge for ADRN access. It is our belief that transparency is key to trust and rejections should be as transparent as acceptance. The organisation name, rather than individual applicant could be used if felt necessary.