



all	
but	
names	

- [Introduction](#)
- [The Missing Ethics framework](#)
- [A better “better use of data”](#)
- [Individual Level Data is the defining criteria](#)
- [“Digital services so good people choose to use them”](#)
- [Lessons from the Department of Health](#)
- [What is identifiable data?](#)
- [Bulk Personal Datasets in the non-secret parts of Government](#)
- [Consultation questions](#)
 - [Improving Public Service Delivery](#)
 - [Providing assistance to citizens living in fuel poverty](#)
 - [Access to civil registration to improve public service delivery](#)
 - [Combating fraud against the public sector through faster and simpler access to data](#)
 - [Improving access to data to enable better management of debt owed to the public sector](#)
 - [Access to data which must be linked and de-identified using defined processes for research purposes](#)
 - [Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research](#)
- [Research](#)
- [Statistics](#)
- [Citizens](#)

medConfidential & All But Names
coordinator@medConfidential.org
coordinator@allbutnames.com
April 2016

New “data legislation” is necessary. The illustrative clauses and principles underlying them in this consultation are unfit for purpose.

Irrespective of the choices made after this consultation, every citizen should be able to know how data about them is used, or Parliament has decided a public interest otherwise.

“Making the civil service work for modern Britain”¹ was the title of a post-election speech by the Minister for the Cabinet Office. The details and approach of this consultation suggests that was misheard as making every citizen in modern Britain be required to have their data work for each civil service silo. That is probably not what the Minister intended.

The questions asked in this consultation are systematically and fundamentally flawed. The important issues that should be addressed are treated with a bureaucratic contempt for which Whitehall is infamous.

The civil service needs to demonstrate public transparency on how it uses data. The recording of deaths questions in this consultation are the same figleaf for broad data sharing as in the Coroners and Justice Bill in 2009 - when the Government accepted the problem and withdrew the clause. It therefore is entirely understandable that the data bit of an old Coroner’s Bill raises its zombified ugly head again now as the death bit of the forthcoming Digital Economy bill. Reemergence under a different Government strongly suggests that this legislation is not a native Ministerial priority, but is a civil service priority.

The high quality data work of the Cabinet Office seems to be no more. Under past leadership it has been subject to well deserved public praise and resulted in the appointment of a Chief Data Officer. From the details of this consultation, and the process that led to it, the past vision has degraded to the state of a minor project of the Department of Administrative Affairs, while simultaneously lacking any of the leadership qualities that made Sir Humphrey Appleby the idol of the authors² of this consultation...

About medConfidential & AllButNames

medConfidential is an independent non-partisan organisation campaigning for confidentiality and consent in health and social care, which seeks to ensure that every flow of data into, across and out of the NHS and care system is *consensual, safe and transparent*.

AllButNames is a response to toxic data initiatives within the NHS, such as care.data, seeping into the rest of the HM Government. Despite huge public outcry at the misuse and sale of our medical records the Government, under lobbying from vested interests, is trying to do the same thing is always has, just more so. Alternative approaches based on this consultation will be published over the summer.

The Missing Ethics framework

¹ <https://www.gov.uk/government/speeches/making-the-civil-service-work-for-modern-britain>

² We would say the consultation designer, but there is scant evidence that there was one.

As care.data was for the NHS bureaucracy, this consultation is about doing more of what Government been doing already: Not better sharing, just more copying.

If this consultation wasn't about databases, the same questions could be asked about buying more filing cabinets, ink, and scribes.

The approach and consultation strengthens the groupthink of the last decade, where parts of this legislation originate. It may look as if there were no lessons from recent years, but there is a reason this consultation looks that way: the hard lessons have been entirely ignored.

"These legislative proposals are part of a broader programme to modernise the UK data landscape. Our goal is to transform and improve the relationship between the citizen and the state"³ says the Ministerial introduction to the consultation, however, it looks a lot like doing more of the same. This is the approach that led to the care.data fiasco.

The published "[data science ethics](#)" framework could have been used to justify care.data, in any of its disastrous forms. No one involved in that programme in 2013 would have had any problem ticking the current boxes.

The [original draft](#) framework⁴ was much more challenging. Even when it had been turned into the same 10 point civil service language before the election, the challenge to the status quo remained. The language is now exceptionally bland and unchallenging. It is not about doing better, it is now about doing more.

When the post-election Director of Data took an ethical framework round the Departments, they hated it, and so it was watered down... "Start with clear user need and public benefit"

What the Departments wanted was to keep doing what they'd been doing all along. And so the Cabinet Office destroyed a credible ethical framework and became lap dogs to mass copying of bulk personal datasets... "*Use data and tools which have the minimum intrusion necessary*"

It isn't quite the same process that created care.data, but the outcomes will be the same... "*Be alert to public perceptions*"

The NHS has had 3 years of data pain as those who use data repeatedly fought to keep doing the same thing they've always done, or more of it. The Caldicott Review will say how that should change. A review whose publication has been delayed by Whitehall until the day before the Cabinet Office consultation closes – published before so they can say it was published before, but only hours before the closing deadline so it can't be digested... "*Be as open as possible*"

³ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/503905/29-02-16_Data_Legislation_Proposals_-_Con_Doc_-_final_3_.pdf

⁴ <http://blog.memespring.co.uk/2015/11/12/10-rules/>

As such, when we have read the Caldicott Review, we will publish a commentary in the context of this consultation at <https://www.medconfidential.org/news>

The Government response to this consultation should be simple:

- **“No, to secret or invasive copying of identifiers”.**
- Yes, to offering a citizen a choice whether their data can be copied, or shared, to make a transaction easier for them;
- Yes, to democratic decisions about defining exactly the circumstances when the above question isn't asked;
- Yes, to every citizen knowing how individual level data about them is used, and why.

A better “better use of data”

The Home Office part of the consultation creates arbitrary gateways, for the civil service, not for the citizen. The illustrative language is drafted as data copying done to the citizen, not for the citizen, and is not about the Home Office – they wrote the example, but the legislation came straight off the Cabinet Office shelf.

Local authorities see benefits to sharing medical data with landlords; and why wouldn't a local authority want to tell their department that deals with social housing to know what benefits a citizen is entitled to?

There are also many reasons a citizen may not want their landlord to know a fact that could be legitimately shared for others. This problem can not be addressed at a national scale. It can only be addressed by the citizen being able to give or withhold permission as part of each transaction. Insurance companies will take any data you offer them, which is why the GP profession uses a standard agreed form with just the information needed.

The same thought process needs to be applied to the rest of Government.

Individual Level Data is the defining criteria

Parents want better statistics on local schools, but giving parents access to the detailed school histories of 20 million children, including all their offspring's classmates, is likely unwise.

Even if the names and identifiers have been entirely removed from a dataset on your school days, you can still remember your classmates by name, and can reidentify them just on your knowledge, however long you have been out of school. We discuss later in this document Annex 7 of the first Caldicott Report, which is the list of criteria that make health data identifiable, even if the "direct" identifiers are seemingly removed.

Data on citizens is either aggregated statistics, or it is individual level data. Claims that data "without names" is somehow safe are fundamentally flawed, and a bureaucratic truism that led to the care.data fiasco.

When it comes to individual level data, much as it widely accepted when creating formal statistics, the creation process matters.

If the ethical framework and process are strong, project details matter less. If the ethical framework or process are weak, the project details don't matter at all.

"Digital services so good people choose to use them"

There are undoubted benefits for an individual and for government in data from one department being visible by another, it must only be done with the citizen's individual consent. If the Cabinet Office is now arguing that choice is not a necessity for improving public services, that is a discussion that will have wider ramifications.

Digital services should be so good that people choose to use them, not so creepy it doesn't matter whether citizens use them.

Legislation may be needed to create a gateway, but it must be up to the citizen whether they choose to walk through it, or choose to go a longer way round. The requirements for use of a service should be the same ("prove you were born"), but whether it is an API or a birth certificate should be the citizen's choice.

We understand that HMPO will only provide confirmations electronically if they have a statutory gateway to do so; we see no reason for that particular narrow gateway not to be possible, if it may only be used with the consent of an individual citizen.

Lessons from the Department of Health

Some citizens would be entirely happy with all their data being used; some, otherwise. But it is impossible for the central government data team in the Cabinet Office to tell the difference – they never deal with citizens in the course of those transactions. It is entirely possible to be well meaning, good intentioned, and utterly destructive.

There needs to be a good ethical framework, not one that is designed to be the lowest common denominator, acceptable to all projects that evolved without one. As the forthcoming Caldicott Review will show, some of the existing data projects have not met the standards that should be expected.

The Cabinet Office has lowered the standards and ignored the broken projects. The NHS tried the Cabinet Office approach at the start, which is why care.data is in it's third year of suspension. It's not that they don't know how to fix it, it's that the people who can prevent it being fixed like the current setup more than anything that is publicly acceptable. That is what the current independent Caldicott Review is needed to design a system those whose job it was refused to consider.

The settlement that is being developed for medical records has a strong ethical framework underpinning it, which is widely known and absolutely concretely accepted: medical ethics.

Attempting any form of bulk personal dataset copying without a strong ethical framework is doomed to fail.

What is identifiable data?

The definition of what constitutes “identifiable” data in the health arena was clearly defined in 1997, in Appendix 7 of the first of the Caldicott Reports.⁵ It is entirely clear:

“The Working Groups identified a number of items by which a person’s identity may be established. These include:-

- *Surname*
- *Forename*
- *Initials*
- *Address*
- *Postcode*
- *Date of Birth*
- *Other Dates (i.e death, diagnosis)*
- *Sex*
- *NHS Number*
- *N.I. Number*
- *Local Identifier (i.e. hospital or GP Practice Number)*
- *Ethnic Group*
- *Soundex Code*
- *Occupation*

The groups determined that an individual item from this list, taken with another item from a particular flow, may in certain circumstances enable identity to be inferred, e.g.:

- *Age linked to a diagnosis;*
- *Postcode and the medicine prescribed;*
- *Address and the item of service provided”*

While not all of those specific criteria will apply to other parts of Government, the fundamental and underlying ideas do. The work of Professor Sweeney, Professor of Government and Technology in Residence at Harvard University and Director of the Data Privacy Lab at Harvard, is the canonical practical example here.⁶ It would be deeply unwise for the rest of Government to ignore the lessons learnt so painfully regarding the identifiability of health data.

⁵ http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068404.pdf

⁶ Sweeney L. Only You, Your Doctor, and Many Others May Know. *Technology Science*. 2015092903. September 29, 2015. <http://techscience.org/a/2015092903>

Bulk Personal Datasets in the non-secret parts of Government

There are reasons for bulk personal dataset use powers, sometimes with opt-outs, sometimes not; but they must all be based in statute and subject to a discussion and vote in Parliament. Where opt-outs are unavailable or chosen to be overridden, that should only be by the express will of Parliament.

Some aspects of research are only possible at large scales, and whatever the reason, citizens should know how their data is used, and the results of those researches.

Civil Servants should serve the citizenry. It is data about a citizen, which is deeply personal, in a similar way to it being “taxpayers’ money”.

The Cabinet Office data legislation needs to reflect that.

The coverage of this legislation should be all individual level data - which it appears not to be. We note with deep concern a FOI response⁷ from the Department of Health about their bulk personal datasets. That list is surprisingly short - is that really all they use?

It would be perverse for Government to argue that only individual level data with names attached is data of interest to citizens. That was the argument suggested by former Cabinet Office Data supremo Tim Kelsey when he was subsequently designing care.data, and is no more likely to gain public acceptance in central government than it did in the NHS.

There was at least some form of expectation that NHS uses would have some connection to “the promotion of health”, even if there is no agreement of what that means in practice. We do not expect that DWP will be given the same benefit of the doubt.

It is a perverse effect of the lack of understanding of data, that a dataset without names on can include the most intrusive data on citizens, yet be treated more carelessly than an email containing a lunch menu. The classification of individual level datasets at rest in the non-secret parts of Government should be reviewed and enhanced.

The increasing use of data in Government is important, but it is necessary that it be consensual, safe and transparent. Citizens should know how data about them is used, all uses should be done safely, and either a citizen or Parliament should give informed consent to that that usage.

Anything else will lead to care.data style debacles in many departments. That is unlikely to be the intent of this consultation, although it may well be the outcome.

⁷ <http://whatdotheyknow.com/request/320194/response/791962/attach/2/FOI%201022297%20reply.pdf>

Consultation questions

Given the structural flaws in this consultation, we respond to the sections rather than limiting ourselves to the specific narrow and often irrelevant questions.

Improving Public Service Delivery

As we cover in greater detail above, it is more important that citizens know every way that data about them has been used, and have a choice over that topic, than it is who uses it.

Every organisation that delivers public services should expect that the citizens whose data they use, will know when it was used, and why.⁸ Whether that organisation is public sector or private sector, in some ways, matters less.

There are obvious cases where who the user is matters more for particular areas.

In particular, since the consultation attempts to use the “Troubled Families Programme” as a justification, it is clear to note that families in the Troubled Families Programme are there with some form of consent.

There there should be a gateway for projects like the Troubled Families Programme is not controversial; however, every family, and every individual, whose data is used should only be used with consent.

As a voluntary programme, it would be inconsistent and bureaucratically perverse for a family or parent to be able to reject help (as it’s voluntary), but for those same families to have no choice on how data on that family was copied by the programme. The Programme is transparent to families about what help is provided, that transparency is entirely undermined if departments can copy data in secret.

While Ministers may advocate for “government using data like the private sector does”, that argument is fundamentally flawed. Those arguments have oft been addressed in part, and we do not propose to repeat them there, with the exception of querying whether Ministers and the creators of this consultation have spoken to, for example, the mobile phone operators, in detail about how they use data, and the proposals covered here. In conversations, the disregard for functional and meaningful safeguards in these proposals is deeply concerning. While departments sling data around whitehall like paper aeroplanes, the telcos ensure that, while they may use data in innovative ways, there are very very strong accountability and internal controls on what happens - because of the reaction. Those internal controls do not exist within Whitehall, where the driver can be political.

⁸ <https://medconfidential.org/2014/what-is-a-data-usage-report/>

Providing assistance to citizens living in fuel poverty

We limit our comments to *how* assistance should be facilitated. Whether assistance should be provided to particular groups is not an issue on which we take a view, and would like all views, even if contradictory, to be heard and balanced on question 5.

There will be views where the interests are contradictory - and where only the citizen can know which way they will choose to go, and that may be subtly different. The systems in place for data decisions involving individual level data must facilitate that.

The process currently followed, for the DECC Winter Fuel Payments, seems to be a privacy by design process that shares the minimal information, while minimising the costs to all involved, without any possibility of harm or distress to individuals given the nature of the data that is shared. Expanding that process does not, on the face of it, seem controversial or problematic, but the devil is in the details. We are happy to look further should they be provided.

Access to civil registration to improve public service delivery

We will answer the two, very narrow, questions posed, and then discuss the illustrative clauses which are far broader than the figleaf covered by the consultation questions on a single substantive issue.

The sharing of data is currently a substantially political decision, rather than a decision about citizens. Jeremy Hunt's database of women's genitalia, designed by the Home Office⁹ which had "prosecutions" as a justification for certain data items; DWP wanting every fit-note form filled out¹⁰ to measure GPs, or grab GP appointment books¹¹ which got walked back after a firestorm.

Those who wish data will always justify it as "necessary and proportionate". Without strong Departments, "necessary and proportionate" will become whatever is politically expedient, as it has in the past. We will publish further on this as part of our response to the Caldicott review.

8. Should a government department be able to access birth details electronically for the purpose of providing a public service, e.g. an application for child benefit?

⁹ <http://www.hscic.gov.uk/fgm> - see spreadsheet.

¹⁰ <https://www.gov.uk/government/publications/fit-notes-plans-for-collecting-anonymous-data-in-england>

¹¹ <http://www.dailymail.co.uk/news/article-3168803/Privacy-storm-GP-visits-No10-demands-details-millions-confidential-appointments.html>

As part of digital public service delivery, a citizen should be able to consent to a digital check being made for a particular purpose, in line with the same evidence being provided via other non-digital means (such as a paper birth certificate).

We understand that the Home Office, in particular the HM Passport Office, wish a particular, well defined, narrow legislative gateway to answer citizen consented electronic requests from other parts of Government.

That is not what the illustrative clauses discuss now, with the illustrative clauses being similar to the flawed s152 of the Coroners and Justice Bill 2009.

9. Do you think bulk registration information, such as details of all deaths, should be shared between civil registration officials and specified public authorities to ensure records are kept up to date (e.g. to prevent correspondence being sent to families of a deceased person)?

In line with the level of quality and attention to detail of much of this consultation, the case for this has not been made. That is not to say that it would not be made.

We support the work of the Royal Statistical Society in their work for the timely registration of deaths and the production of statistics.¹² This whole area of government handling is flawed due to the secondary interest of the Registrar General's work in the Home Office, where their primary job is that of Director General of the Passport Office.

We would support a machinery of Government change to move the registration of such life events either back to ONS, or to the HSCIC.

That is not what was proposed in the Coroners and Justice Bill in 2009, and it is not what the illustrative clauses discuss now. It is entirely understandable that the data bit of Coroner's Bill raises its zombified ugly head again now.

This reemergence strongly suggests that this legislation is not a Ministerial priority but is a civil service priority, and the civil service needs to demonstrate public transparency.

¹²

<http://www.rss.org.uk/Images/PDF/influencing-change/RSS-statement-late-registration-deaths-England-Wales.pdf>

Combating fraud against the public sector through faster and simpler access to data

It has been a repeated mantra of successive Governments that increased data sharing can combat fraud. Throughout the open policy making process, there was repeated mention of past attempts to prevent fraud through data sharing and a desire to do more, and no substantive evidence provided that any of it had made any difference at all.

While data sharing programmes may provide some assistance, there should be a new framework designed accounting for their cost, their benefit, and the outcomes. That is likely to involve the Parliamentary scrutiny process of programmes, rather than internal data sharing metrics that the public never see.

Combating fraud should be subject to safe and transparent reporting, and also subject to the democratic consent of Parliament, both in practice as well as in principle. In that framing, it is possible to consider error in the same context, looking at mechanisms to reduce institutional and individual error as part of the same process.

While piloting and testing programmes is important, given the past litany of failed programmes in this area that are limping along, there should be an expectation that each programme ends unless it is shown to meet or exceed predefined criteria, and that it continues to do so.

Improving access to data to enable better management of debt owed to the public sector

Whatever practice is developed, it should be the case that this offer is better to the citizen than the existing status quo of dealing with debt independently.

As such, each citizen who is in debt with Government should be able to choose to be part of this programme if they wish, for the aspects of their debt that they wish.

If a citizen wishes Government to merge the debts and act as a single creditor, then Government should be capable of doing so, but where the citizen does not explicitly wish that to happen, then the status quo should continue.

At a future point, the “default” choice may change, but Government has clearly not demonstrated that this new approach is better. Should it do so, to the satisfaction of the organisations that work with vulnerable individuals in these circumstances, an opt out model may be considered, if felt beneficial by those organisations.

For any legislation on this topic, there will have to be a statutory bar to privatisation, with any consent for a public body doing debt management being reconcented for a private organisation.

Similarly, should any part of this be outsourced, there must be clear statutory bars on the data from Government being reused for other purposes.

Access to data which must be linked and de-identified using defined processes for research purposes

It is unfortunate that the attempt to define trust third party matching in primary legislation is not the most flawed part of this consultation. Defining such a technical process in primary legislation is fundamentally and entirely unwise.

The legislation should define the outcome, and the requirements, and the restrictions, but should not define the mechanisms for matching. It should also require that any data for research purposes must be subject to a dissent mechanism (ie, opt-out; or be an opt-in process - as with surveys), and a full reporting mechanism to citizens of what research has been approved, and the new knowledge that came from their participation.

Access by UK Statistics Authority to identified data for the purpose of producing official statistics and research

There should be a separation between statistics and research. Official (or National) Statistics require a population dataset, which can not be consented.

For a worked example in the health arena, please see the existing medConfidential publications¹³ on this topic, which looks at risk stratification for A&E, which is said to be the most complex area and previously argued as the reason that individual level data was absolutely required. It isn't.

Research

Modelling and the discussion around the design of statistics are both actions and features of Research (which also use other Official/National Statistics to inform the population pyramids of non-response or dissent).

¹³ <https://medconfidential.org/2015/towards-protecting-data-in-secondary-uses/>

It is in this way, that research done by the UK Statistics Authority is little different in data access to that done by academic researchers, or others. There is no expectation of “privileged” access that some researchers get and others can not - other than the differentiation of organisation, purpose, and capabilities that vary between institutions and process already.

All research projects should be published in advance,¹⁴ and all publications which generate new knowledge should be available to the public.

Statistics

Any statistics produced (whether official, national, or experimental) are then created off the minimal dataset that is required for their production, as evidenced by the research process.

All statistics should be published.

Citizens

For individual level data, whether identified or de-identified data, citizens should know how data about them was used, and the outcomes of the research that came from the use of their data.

ONS has historically attempted to ensure all researchers reported back on their publications and “impact” to ONS or intermediate data providers (who passed that information on). It has previously been an important measure for understanding how non-ONS researchers use data, and a strong justification for some data products continuing to exist.

Given the large strides that the ONS website has made in recent months, and the renewed focus on accountability, it is necessary for UKSA to begin to think about how it can report back to citizens on the knowledge that is generated based on statistical knowledge.

MedConfidential

coordinator@medconfidential.org

coordinator@allbutnames.com

April 2016

¹⁴ Following the model of <http://opentrials.net>