Fair, Safe, and Transparent: A vision of data strategy for a digital government

> "Citizens should know how data about them is used, in the same way taxpayers should know how taxpayers' money is spent."[1]

Data is fundamental to all aspects of service delivery, policy design, and digital transformation, whatever the service. Government needs a data policy that is fit for purpose, and which sustains public confidence into the long term.

The approach in this paper would have avoided many recent/current public sector data fiascos (care.data; NPD/Country of Birth; et al), and should also match the forthcoming data strategy of the NHS. If the NHS can now routinely publish a list of bulk data flows,[2] why can't Government?

The NHS was the canary in the coal mine - care.data showed the state of data handling in the NHS. It was not worse than Government data handling, it was simply noticed first. The problem is excessive secrecy and inadequate processes, both of which can be resolved.

Data policy should split into two stands - individual service delivery and bulk policy/service design. Data should inform policy, just as it is necessary to make decisions on service delivery, but the principles and purposes are different.

Nothing in this paper requires legislative change to implement by choice, although may require legislation to mandate. "Fair" data handling is defined by the Data Protection Act.

1. Individual data for Service Delivery

Service delivery is providing operational services, of any kind, to an identified or identifiable member of the public. They are the "user" in user needs. For decisions to be made for/about them, they are capable of consent (or at least notification) of what data has been used and why. Often, the consent will be a statutory mandate as necessary to acquire the service. Much of this can be done (at varying degrees of practicality) via attribute exchange in a Gov.UK Verify model. What is necessary for that citizen?

Accountability, in a digital by default world, comes from transparency via digital means back to the citizen - see Annex A.[3]

This should not be controversial, and service one person does not require access to data on other unconnected individuals. This is individual access; one citizen at a time as they use a service. The output of individual data for service delivery is better service to each citizen.

---

[1] Minister for the Cabinet Office, January 2016.
[2] http://hscic.gov.uk/dataregister
[3] For a longer description, see https://medconfidential.org/2015/implementing-data-usage-reports/

2. Bulk personal data for Policy Analysis and Service Design

Policy and Service Design requires a high level view across a service. It requires modelling of situations, likelihoods, etc, which are only possible using data. Some of these datasets will be subject to individual dissent, others individual consent, and many no form of consent at all. But there should be a public record of flows of bulk data, whether for policy analysis, delivery measurement, fraud protection, or other defined purposes.

Bulk projects are not about making decisions on individual cases today, but designing the services of tomorrow. The outputs are reports and analyses of a scientific (ie falsifiable) hypothesis, tested against data. The general knowledge gained may then be used in service delivery, or not, as the case may be. This work should be the evidence base upon which policy gets made. Some will be is “data science”, some of it is traditional analysis - the differences are irrelevant. It can be defined, and included in a register.

Increasingly, departments wishing to collaborate should utilise (virtual) “safe settings”[4] for their analysis, rather than sharing unrestricted datasets across boundaries and control. This provides all stakeholders confidence that “risky” data was used appropriately, without undue concerns. Safe settings and transparency over every flow of data means that a project that causes novel concerns can be seen as a single incident amongst many, rather than public concern being exacerbated by secrecy. The NHS does this already.[5]

Accountability comes from the public data release register of projects and data that that flows across organisational (departmental/ALB) boundaries. The register should cover what data flowed, where, and why. In doing so, that will demonstrate that data controllers are taking due care of their data, but also that data is being shared appropriately on a case by case basis, as Parliament or Departments intended.

Summary

The above proposals, together, provide for a fair, safe and transparent data system for Government. It starts from where Departments are, allowing them to move independently towards a data infrastructure which meets both departmental and cybersecurity needs.

A Ministerial commitment to a public register of flows requires each Department to main such a list. As such, all flows can be appropriately secure - transparency is the demonstration that safeguards are working, as policy development and innovation proceeds iteratively, project by project. By publicly listing flows, Departments can each show the public that their use of personal data on citizens is fair, safe, and transparent.

October 2016

---

[4] http://www.thinknpc.org/our-work/projects/data-labs/justice-data-lab/ or the ONS virtual microdata laboratory https://www.ons.gov.uk/aboutus/whatwedo/paidservices/virtualmicrodatalaboratoryvml

[5] http://hscic.gov.uk/dataregister

HM Government

Personalised Government Data Usage report

How the Government uses your data

Report on
Alison Smith

Period covered
October 2016

There have been no reported breaches or losses of your records this quarter.

For NHS records, please see your Health and Social Care Data Usage report.

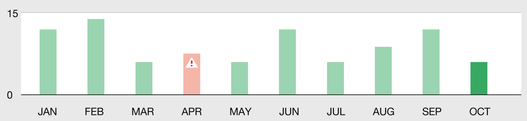
Operational uses

When you were present your data was accessed
6 times by **3 different services**

Service	When
JobCentre Plus Manchester, M22 5RA*	1 April, 14 April, 28 April
HM Passport Office **First Adult Passport Inteview** Manchester, M1 3hu*	27 April, 29 April
Online electoral roll registration GOV.UK Website	20 April

Report an issue

Monthly overview 2016
Amount of times your data was accessed, in your presence



When you weren't present your data was accessed
12 times by **4 different services**

Data relating to your use of items marked * will have been sent to ONS for the production of Official and National Statistics.

Research undertaken using your data

Your data is distributed as part of a variety of datasets. Academics and researchers use them to try and learn more about all aspects of our society. The research below may have been conducted in any of the last several years, on any event since the dataset was collected. You may have been included to examine differences against the general population from those at the focus of the study.

What we learnt

Your data was used in the creation of **8 academic research** reports

View these reports

There were **205 articles** published in the time period.

To read all publicly available scholarly papers that used data, and search by keywords, please see:
www.ons.gov.uk/whatwelearnt

New Data Releases

Data releases this month will lead to publications in the future. Most non-commercial users list all their outputs as part of the public benefit of receiving your data. Not all commercial recipients are required to report what they learnt.

- Requests for Statistics
- ONS Longitudinal Study
- National Pupil Database

If you have any other questions about the content of this report, please contact: enquiries@ons.gov.uk

Produced for you by the Office of National Statistics
Delivered by the Government Digital Service