Mapping data in the UK government

Open Data Institute

Contents

About this project

In August 2021, at the Open Data Institute (ODI), we launched a project to crowdsource data-related organisations and initiatives inside the UK government, as part of our work in support of the National Data Strategy. Contributors added organisations and initiatives to an open document and provided a series of comments, critiques and questions. The ODI published a series of blogposts, looking at the exercise through the lens of the six themes in our manifesto: infrastructure, capability, innovation, equity, ethics and engagement. We also held a discussion as part of DataConnect21, and the project has a theme song, of sorts.

The lead author is Gavin Freeguard (Special Adviser) and the wider project team included Dr Jeni Tennison (Vice President and Chief Strategy Adviser), Dr Mahlet "Milly" Zimeta (Head of Public Policy), Lisa Allen (Head of Consultancy for Data Programmes, and Julie King (Delivery Manager).

> ### Mapping data in the UK government
>
> 'Mapping data in the UK government' explores some of the ways that the UK government is transforming its use of data can unlock the value of data across the wider economy

Better data for better government

On 27 October 2021, the UK Chancellor of the Exchequer will set out the government's spending plans in the Autumn Budget and Spending Review 2021. The former will outline the state of the economy and proposed changes to taxation; the latter is expected to set the spending for government departments for the next three financial years, after a few years of unusual one-year Spending Review settlements because of the disruption of Brexit and the Covid-19 pandemic.

The Treasury has taken more of an interest in data in its recent fiscal events. In March 2020, there was investment in better data sharing across government departments as part of the Budget 2020. The 2020 Spending Review brought several digital- and data-related projects as part of the Shared Outcomes Fund 'to fund pilot projects to test innovative ways of working across the public sector', and the most recent iteration is understood to have an even stronger focus on data-related projects. This year's Budget and Spending Review come with work on the National Data Strategy and a host of other initiatives (the National AI Strategy, Ten Tech Priorities and the forthcoming National Digital Strategy) already underway.

But if the UK government wants to maximise the impact of its investments in data, and ensure that such investments can support its wider priorities (including 'levelling up', 'building back better', net zero, and making 'Global Britain' a key hub for digital and data), it needs to understand what data-related work is currently underway and which parts of the government are responsible for what.

The government risks unnecessary duplication between different data projects it has invested in, inside and outside of government. While tailored approaches in different teams can lead to benefits, there are also risks of different parts of government duplicating work or implementing initiatives without learning from the experiences of others.

In addition to aligning good practice across government organisations, the government will need to ensure it also aligns with other sectors because of key private sector and third sector roles and relationships involved in delivering public services and ensuring data works for the wider economy and society. Given the government's global data and digital ambitions in the 'Integrated Review of Security, Defence, Development and Foreign Policy' and elsewhere, such an alignment could give the UK a comparative international advantage and ensure its place in conversations about global standards and best practice. And the possible magnitude of forthcoming reforms to the UK's data protection regime, currently out for consultation, means it is particularly important for the government to understand its own data landscape.

Quantity is not always quality

In recent years, the rapid development of new technologies and techniques, the increasing interest in data, the proliferation of data initiatives across the public sector, and the transfer of policy responsibilities for data between different government bodies have all led to confusion – inside and outside government – about who is responsible for what and what work is underway (let alone how successful such projects are proving to be). The National Audit Office has highlighted the challenges of using data across government that existed even before the pandemic; the House of Commons Public Administration and Constitutional Affairs Select Committee is among those bodies to have highlighted the particular problems during the pandemic, including the difficulty in understanding who is accountable and responsible for data in government.

Our project to map the data-related organisations and initiatives in government is designed to give everyone – inside and outside government – a better understanding of the different responsibilities held by different government organisations, and the different projects underway. Although the government itself should be maintaining similar lists and maps, we believe that an initiative like this from the ODI may be able to convene information from - and bridge across different sectors - in a way that the government by itself can't.

The continuing proliferation of data organisations and initiatives in government, the possible reforms to data protection, and the government's own ambitions and its intention to plan for the next few years makes such understanding of the data landscape particularly urgent in light of this Spending Review.

How better data in government can support better data across the economy

In September 2020, the UK government published the draft framework National Data Strategy (NDS) for public consultation, updating it in May 2021 in response to the views and evidence received. The National Data Strategy is built around five missions.

Mission 3 of the National Data Strategy is concerned with 'Transforming government's use of data to drive efficiency and improve public services'. It notes the 'massive untapped potential in the way the government uses data' and the 'duty' to maintain the 'high watermark' of data use during the pandemic, when the public sector was able to share information and solve problems quickly. The benefits of our mapping exercise to achieving Mission 3 are obvious: the government has a duty to get its own use of data right, to improve how it makes policies, delivers services and otherwise effectively and efficiently runs its operations; understanding who is responsible for what and what can be learned from whom allows this to happen.

Achieving Mission 3 of the National Data Strategy is a public good in itself. But it is also essential to, and cannot be separated from, achieving other missions of the National Data Strategy, including Mission 1 - 'Unlocking the value of data across the economy - in the following ways:

- **Publishing:** The government is a major publisher of open data, which can be used by businesses and others for economic (and social) growth

- **Leading:** The government has a major influence on what others do through setting expectations and sending signals - it can inspire others and lead by example

- **Collaborating:** The government's own ambitions are likely to require support from private (and voluntary) sector partners, providing them with opportunities which could lead to wider economic benefit

- **Supporting:** The government can produce resources of wider interest and benefit, such as toolkits and research, which other actors in the data economy can learn from

- **Stewarding:** The government is a major steward of data that can be valuable to data economies and data flows

Making information – such as that in our mapping exercise – more widely available and understanding which data-related organisations and initiatives exist within the UK public sector should help encourage collaboration, learning and implementation across sectors, and lead to more of the value of data being unlocked across the economy.

What we found

A template for getting this right is [the ODI's manifesto](#), which identifies six key areas essential to trustworthy and trusted data ecosystems: [infrastructure](#), [capability](#), [innovation](#), [equity](#), [ethics](#) and [engagement](#). We propose this model should be at the heart of the National Data Strategy, and therefore also the UK's cross-sector data ecosystems – which includes the government. As the next section of this report shows, there is work underway in all of these areas within the government, but some are more developed than others: for example, while there are longstanding initiatives and institutions with clear responsibilities in some areas (such as infrastructure or ethics), this is not the case in others (such as equity and engagement).

Every organisation is a data organisation - but some play a greater role than others

Our crowdsourced document – [Mapping: data in the UK government](#) I – shows that there are well over 100 different bodies with significant responsibility for, or a strong focus on, data. This includes:

- more than 65 departments (ministerial and non-ministerial), units and other public bodies

- more than 12 cross-government professions and functions, and

- more than 30 advisory boards, networks and communities.

This is likely to be an underestimate for several reasons. First, in 2021, every organisation is a data organisation. Some contributors to our crowdsourced document wondered where to draw the line, for example between organisations with wider responsibilities and any organisation with a data controller. Any public sector body should be thinking about how it uses data and its role in different data ecosystems, and to fully map data initiatives across the government would (or should) require mapping every government organisation. Nonetheless, our exercise has tried to concentrate on those organisations with cross-government responsibilities or with a particular influence on particular sectors and bring some coherence to the most important and influential organisations; if everyone is responsible, nobody is responsible. Highlighting the major players is therefore crucial, and these include:

- The **Cabinet Office** and its various units and related bodies

- The **UK Statistics Authority (UKSA)** and its associated organisations (including the **Office for National Statistics (ONS)** and **Office for Statistics Regulation (OSR)**), for whom ministerial responsibility ultimately resides in the Cabinet Office

- The **Department for Digital, Media, Culture and Sport (DCMS)** and its public bodies

- The **Department for Business, Energy and Industrial Strategy (BEIS)** and its public bodies (including the research councils)

- The **National Health Service (NHS)** and other parts of the health system.

The Treasury is, as per the Budget and Spending Review and its associated power at the centre of government, also vital.

Gaps and blurred lines in the public sector

A second reason for our figures being an underestimate is particular gaps in our knowledge – for example, in the devolved administrations in Scotland, Wales and Northern Ireland and in English local government.

A third is the sometimes blurred line between what should count as 'government' and the public sector and what shouldn't: there are several government companies, research council investments, even a limited liability partnership between a government company and the Local Government Association (GeoPlace LLP) and a charity (HDR UK) on our list, but there are likely to be other organisations with similar structures and claims to inclusion that we have missed.

A final reason is that we may simply have missed organisations, particularly more informal networks and communities across government that are particularly important for convening across organisations and sharing knowledge and experiences. In some cases, important bodies lack any sort of public presence on GOV.UK – the Government Digital, Data and Technology Steering Group has been mentioned to us but has no such presence, while the data focus of the latest Shared Outcomes Fund round is not on the record anywhere obvious. Some entries have been partly based on job adverts or obscure minutes in lieu of more complete information. This lack of transparency obviously makes it more difficult to fully understand the government data landscape.

In other cases, we might be including some bodies which have long since ceased to operate: the Ministerial Group on Government Digital Technology is mentioned in a 2015 press release but not since, and the Data Steering Group has a membership list which is very out of date. Both are still cited on the 'governance' page for the Government Digital Service.

Our crowdsourced document is, despite its length (more than 100 pages), inevitably a snapshot and not definitive. Nonetheless, we hope it proves useful – especially with the forthcoming Spending Review and in the midst of various data-related initiatives. We are extremely grateful to those who have filled in the gaps in our knowledge and added further information on the organisations and initiatives related to data.

The UK government data ecosystem through an ODI lens

As part of our crowdsourcing exercise, we looked at the organisations and initiatives through the prism of the six points of [the ODI manifesto](#) for open and trustworthy data ecosystems.

Infrastructure

What we mean
Sectors and societies must invest in and protect the data infrastructure they rely on. Open data is the foundation of this emerging vital infrastructure.

When many people think about infrastructure, they think about utilities like electricity and water supplies, or transport infrastructure like railway lines and roads. Data should be seen as important as these energy and transport networks, given how it enables more effective and efficient government and public services, and economic and social growth.

[Data infrastructure](#) can be thought of as:

- data assets (like datasets, identifiers and registers)
- the standards and technologies that allow those assets to be curated and accessed
- the guidance and policies that inform the use and management of those assets
- the organisations that govern the infrastructure
- the communities that contribute, maintain or are impacted by it.

Data infrastructure should be as open as possible for us to get the most value from it – in particular, key datasets that are stewarded by the public sector and can be seen as national public data infrastructure.

How the government getting data right can unlock the wider value of data across the economy

- **Leading:** The government needs to get its own data infrastructure in order to know what data it holds and provide the right infrastructure for other actors across the economy and society to unlock value

- **Leading:** The government treating data as infrastructure should push others to do so too

- **Leading and publishing:** In building the right foundations inside government, the government can provide examples, resources and skills for others to benefit from – as well as behaviour for others to model.

UK government strengths and weaknesses

Improving data infrastructure has become more of a focus for the government in recent years. One of the 'pillars' of the National Data Strategy is ['data foundations',](#) which covers some infrastructure issues such as standards, consistency and quality. The recent reforms to data in government have started to translate this strategic interest into practical activity, with the [Data Standards Authority](#) (DSA), part of the new [Central Digital and Data Office](#) (CDDO), 'lead[ing] the cross-government conversation around data standards', including how to develop and enforce them. The CDDO is responsible for setting and maintaining other standards across government, including open standards, for which the [Open Standards Board](#) is accountable and which the open standards community [helps select](#).

Other parts of the Cabinet Office have some responsibility for data infrastructure, including:

- the [Government Digital Service](#) ('creating the cross-government reference architecture and identifying, enabling and standardising the data registers across government most critical to service delivery' and running [data.gov.uk](#))

- the new [Information and Data Exchange](#) (INDEX, supporting the better of sharing of information across government)

- [Government Business Services](#) (for corporate information such as finance and HR data)

- the [Geospatial Commission](#) (for oversight of location data, which is key foundational data for the economy and society).

The [UK Statistics Authority](#) is responsible for [the UK's respected official statistics infrastructure](#). [UK Research and Innovation](#), and research council-funded projects and [data institutions](#) such as [Health Data Research UK (HDR UK)](#), [Our Future Health](#) and [Administrative Data Research UK (ADR UK)](#) are responsible for social science and research data infrastructure.

There are bodies responsible for securing our data infrastructure, ranging from the [Office for National Statisics's Secure Research Service](#) which allows safe and secure access to research data, to the [National Cyber Security Centre](#) which provides support and guidance for keeping data safe across government. The devolved administrations have their own bodies (the [Scottish Government's Digital Directorate](#) 'provid[es] a foundation for the digital delivery of public services', the Welsh Government's [Centre for Digital Public Services](#) provides standards, guidance, direction and training).

In particular sectors, the [Department of Health and Social Care](#), [NHS Digital](#) and [NHSX](#) are responsible for health data infrastructure; the Department for Transport for transport data infrastructure; the Department for Education for educational data infrastructure and so on.

The Ministry for Housing, Communities and Local Government is responsible for both data infrastructure about housing and that stewarded by local authorities, and is particularly notable for its focus on 'fixing the plumbing'.

There has also been more direct government interest and involvement in unlocking the value of data *about* infrastructure: the National Infrastructure Commission's report on 'Data for the public good' found that improved sharing of data about the UK's infrastructure could lead to lower bills for consumers, a reduced impact on the environment and improved transport, while the National Digital Twin involves a focus on capturing, using and sharing data about infrastructure effectively, to 'improve how infrastructure is built, managed, operated and eventually decommissioned'.

For all the strengths of these initiatives, there is a potential weakness. The National Audit Office's report on the challenges in using data across government highlights 'coherent infrastructure' as one of the key conditions for success, and cautions that a lack of strategic leadership has undermined the effectiveness of previous initiatives. Government attempts to get a grip on data infrastructure are not new, from 'e-government: A strategic framework for public services in the Information Age' (2000) which said departments would implement common standards to allow interoperability, to the National Information Infrastructure (2010–15) which sought to 'contain the data held by government which is likely to have the broadest and most significant economic and social impact'. The 2017 Government Transformation Strategy pledged to build 'a national data infrastructure of registers (authoritative lists that are held once across government)' by 2020; the Government Digital Service's work on registers wrapped up unsuccessfully in 2021. Some of the
new DSA, CDDO and Government Digital Service initiatives sound a lot like some of these earlier attempts. Has the government learned the right lessons from previous failures?

Examples from other sectors and other countries
Europe's Gaia-X project brings together business, science and politics to develop 'an open, transparent digital ecosystem, where data and services can be made available, collated and shared in an environment of trust'.

France's view of data as essential infrastructure includes efforts to build a public data service and become a digital republic.

Further resources from the ODI

- Data infrastructure
- Demonstrating and assessing trustworthiness when sharing data
- Infrastructure for common challenges
- State of Open Data: Data infrastructure
- Understanding the common technical infrastructure of shared and open data
- R&D: Data infrastructure for common challenges
- The UK National Data Strategy 2020: consultation response

- [The UK National Data Strategy 2020: data as infrastructure, data for infrastructure](#)

Capability

What we mean
Everyone must have the opportunity to understand how data can be and is being used. We need data literacy for all, data science skills, and experience using data to help solve problems.

We need people with the skills to get the most out of data – to build the tools, to write the code, to manipulate and analyse it. But we also need people to think critically about the use of data in different contexts – to compare and contrast how data is presented, evaluate bias in the way data is collected, consider the limits and constraints of data and understand its impact on society. In other words, [data literacy](#).

How the government getting data right can unlock the wider value of data across the economy

- **Leading:** We need government plans for improving data skills (including data literacy) to build capability across sectors

- **Leading and supporting:** The government building its own capability provides examples, materials and role models to follow

- **Collaborating:** Government work will also develop experts who can move to other sectors; career paths which may include a stint in government; and opportunities for those in the private sector to sell their skills and services into government

UK government strengths and weaknesses
'[Skills](#)' is one of the pillars of the National Data Strategy. The government has longstanding expertise in developing skills around statistics and analysis – for example, through the work of the [Government Statistical Service](#). More recently, attempts to build the government's data capability have focused more on the use of data in digital services: the creation of the [Government Digital Service](#) in 2011 and the subsequent creation of digital teams in many departments helped bring data professionals into the public sector; and the 2017 [Government Transformation Strategy](#) also promised to embed digital skills across government.

Recent attempts to improve the capability of the civil service have focused on strengthening cross-cutting [professions](#) and [functions](#), each of which cover skills, expertise and activities common to many government organisations (including procurement, policy, finance and HR, as well as digital and data). Some of those particularly focused on data are relatively mature, including the [Digital, Data and Technology Profession](#), led by the Central Digital and Data Office and its work including developing capability frameworks and career paths for different data-related roles, and offering training; and the [Government Analysis Function](#), led by the National Statistician, which runs events like [Government Analysis](#)

Month and has a learning curriculum and career framework for civil servants working in the government's various analytical professions. There are other initiatives for data specialists to develop their skills, such as accelerator programmes led by the Government Data Science Partnership. And there are more informal cross-government communities and networks – on everything from data architecture to APIs and data exchange to civil servants working on understanding public attitudes to data.

In our response to the National Data Strategy, we noted the focus on relatively advanced data skills in central and local government, and called for more attention to be paid to building comparable capabilities in businesses and civil society. We also called for a strong foundation of core data literacy across all of society (and for other professionals in government, such as policymakers). There are some new schemes reaching beyond these specialists: Number 10 hosts innovation fellowships to bring people from business, academia and civil society into government, and 10DS runs data masterclasses for senior leaders while the Government Data Quality Hub has launched an 'introduction to data quality' course.

Various government strategy documents (such as the now superseded Industrial Strategy and the more recent Plan for Growth) mention the need to invest in the skills of the future, including data and artificial intelligence, across the economy and society. The National Data Strategy mentions several organisations with responsibility for raising digital and data skills, including the Data Skills Taskforce, The Alan Turing Institute, the National Innovation Centre for Data, the AI Council, the UK Cyber Security Council, The Data Lab and the ODI. DCMS published a report on Quantifying the UK Data Skill Gap back in May 2021 and recently launched an Online Media Literacy Strategy. While there is much promise in all this, there was a 2013 strategy for data capability which largely sank without trace.

Examples from other sectors and other countries
In the UK charity sector, DataKind UK helps 'social change organisations use data science' by connecting charities with data scientists, while the Catalyst network 'help[s] UK civil society grow in digital skills and confidence' (their recent projects include The Data Collective).

Ben Goldacre – currently undertaking a review into the use of health data for the government – has written (with others) about how to build data analysis capability in the NHS.

An international group of academics and experts have come together to develop skills to support the next generation of public servants, through Teaching Public Service in the Digital Age.

Further resources from the ODI

- [Data 2020: Skills, engagement and data literacy](#)
- [Data literacy: what is it and how do we address it at the ODI?](#)
- [Data Skills Framework](#)
- [ODI learning courses](#)
- [Skills and literacy](#)
- [The UK National Data Strategy 2020: skills that count](#)
- [We join forces with the IoD to help upskill a generation of company directors in data governance](#)

Innovation

What we mean
Data must inspire and fuel innovation. It can enable businesses, startups, governments, individuals and communities to create products and services, fuelling economic growth and productivity.

Opening and sharing data outside an organisation means that a wider range of organisations have the opportunity to innovate with data or participate in data innovation, rather than just the larger organisation which already holds a significant amount of data. Greater access to data is necessary for innovation, but it is not sufficient on its own: practical support (including incubators, challenge funds and competitions, capital investment, open source tools and innovation-friendly policy changes) is needed to create incentives, incentivise risk-taking, and overcome barriers to participation from diverse backgrounds.

How the government getting data right can unlock the wider value of data across the economy

- **Stewarding and publishing:** As a major national steward of data, the extent to which government innovates with data shapes the level of innovation possible in the wider economy, including the data it makes open or otherwise available to businesses, civil society organisations and others to build new products and services, and its support for the creation of ecosystems where data is more widely shared.

- **Leading and supporting:** The government's own innovation in developing policy and delivering services provides a signal and an example to others, and potentially useful resources

- **Collaborating:** Provides opportunities for suppliers to work together on innovative projects

UK government strengths and weaknesses
The government has thought more about itself as an innovator in recent years, as shown by the 2018 ['Technology innovation in government survey'](#), which focused on the existing landscape of emerging technology use in government (including 'data driven concepts' and technologies like artificial intelligence and the internet of things), and the subsequent [Government Technology Innovation Strategy](#), which set out how government should use such technologies to improve public services.

The [Future Policy Network](#) brings together various teams from across government who bring innovation into policymaking (and includes [Analysis and Insight](#), a Cabinet Office unit, tasked with a watching brief on public sector innovation globally). The [CDEI](#) publishes much of its research and advice to the government on how to maximise the benefits of new technology in ethical ways.

The government has been at the forefront of innovation in particular areas, including its work on Smart Data (a new cross-government smart data working group aims to support initiatives across different sectors), while the Declaration on Government Reform and Mackintosh Report are just the latest efforts to make more data available, the latter also thinking seriously about how to value knowledge assets (work being continued by the new Government Office for Technology Transfer).

It has also made funding available for public sector innovation in recent years, including the Treasury's Shared Outcomes Fund 'to address cross-cutting issues in a way that improves outcomes and ensures value for money', which included a focus on innovation with data, and the Regulators' Pioneer Fund from BEIS.

There are also parts of government supporting innovation in the wider economy, such as:

- the Office for Artificial Intelligence (AI), tasked with helping to drive the 'responsible and innovative' uptake of AI technologies in the UK

- the Digital Markets Unit, part of the Competition and Markets Authority, which aims to promote competition and innovation in digital markets while protecting the consumer

- the Geospatial Commission, one of whose roles is to fuel innovation in the use of geospatial data

- Innovate UK, the 'UK's innovation agency' and part of UK Research and Innovation, which supports businesses 'to develop and realise the potential of new ideas' and has AI and the data economy as a priority area in its delivery plan.

In our response to the National Data Strategy consultation, we recommended that the government should pay more attention to civil society as data innovators and noted that publication of more open and shared data would allow innovation from a wider range of innovators (as well as the need to develop the right skills for people from all backgrounds.

Previous government reports on opening up data, such as the Power of Information review (2007) and taskforce report (2009), and 'Shakespeare review of public sector information' (2013), may have emphasised the innovation possible based only on public sector information being made more widely available at the expense of thinking about data from different sectors being brought together.

Examples from other sectors and other countries

The European Commission has innovation as one of the pillars of [the European Data Strategy](#), and says 'businesses will have more data available to innovate as a result' of the strategy.

[The Organisation for Economic Co-operation and Development's Observatory of Public Sector Innovation](#) surveys the landscape and aims to 'turn the new into the normal and provide trusted advice'.

Further resources from the ODI

- [Data innovation for the UK: research and development](#)
- [Innovation: ODI topic](#)
- [R&D: Scaling data innovation](#)
- [Seven reasons why businesses should be sharing data](#)
- [The UK National Data Strategy 2020: innovation with impact](#)
- [Using data in the public sector](#)

Equity

What we mean
Everyone must benefit fairly from data. Access to data and information promotes fair competition and informed markets, and empowers people as consumers, creators and citizens.

One aspect of equity is around diversity and inclusion. The importance of data to decision making, data-driven innovation through technologies like artificial intelligence, and data's role as important national [infrastructure](#) means it is vital that data is as representative as possible. Some groups can be under-represented in data and others over-represented, which may lead to discrimination and reinforce existing inequalities.

Another aspect is around fair competition and informed markets, and empowering consumers. Regulation and legislation like the General Data Protection Regulation (GDPR) guarantees people's rights as citizens and consumers, protects their privacy and gives them more control over rights about data about themselves. The [right to data portability](#) in GDPR – which allows individuals to obtain and reuse personal information – enables consumers to move between different services, which may be good for competition and allow the consumer to get a better deal.

How the government getting data right can unlock the wider value of data across the economy

- **Publishing and leading:** Government collecting and publishing data that shines a light on inequalities within and beyond its own operations
- **Leading and supporting:** Improving the data the government holds and uses on social diversity characteristics
- **Leading:** More diverse social representation in government data efforts
- **Supporting:** Making available research and resources that can be used by others to support their own efforts
- **Collaborating:** Projects that improve the social diversity of the wider data science field
- **Collaborating:** Coordinating data projects to support consumers for fairer market competition

UK government strengths and weaknesses
At present, data equity is perhaps under-served by initiatives and organisations compared to the other ODI manifesto points. The [2020 A Level algorithm fiasco](#) is a prominent example of why it matters (with bias in the use of data and algorithms discriminating against certain schools and students) and why the government needs to take these issues more seriously. Nonetheless, there are some strengths in the government's current approach, and new initiatives are getting underway.

On the diversity and inclusion side of equity, this includes [Equality Hub](#), which works with the Cabinet Office, the Government Equalities Office, the Social Mobility Commission, Race Disparity Unit and Disability Unit to use data to

understand inequity and inequalities across society. Equality Hub's [equalities data audit](#) builds on [the Race Disparity Audit](#) and includes data on geographical inequalities – part of the government's 'levelling up' agenda – and the intersection between different characteristics, such as ethnicity, gender, disability status and socio-economic background. It is also responsible for [the Gender Pay Gap Service](#): organisations with more than 250 employees are required to publish gender pay gap data, to help understand and illustrate the problem.

The UK Statistics Authority's [Inclusive Data Taskforce](#) has recently published [its report](#) on how to improve the data held on some of these characteristics (and goes much further that [the (much-criticised) Report of the Sewell Commission on Race and Ethnic Disparities](#) which included a recommendation to 'use data in a responsible and informed way'). The The Centre for Data Ethics and Innovation's (CDEI's) review of how [algorithmic bias](#) might affect significant decisions about individuals looked at financial services and recruitment as well as the public sector, and its recommendations and resources will be of use beyond government. The National Data Strategy Forum – designed to help 'operationalise' the strategy and keep stakeholders involved – says it will '[prioritise diversity of representation](#), bringing together multidisciplinary expertise and fresh perspectives from up and down the country'.

On the markets and competition side, BEIS's work on [Smart Data](#) – 'putting consumers in control of their data and enabling innovation' – stands out. It was referenced multiple times in the National Data Strategy, including the creation of [a cross-government Smart Data working group](#) to support initiatives and benefit consumers across different sectors. The new [Digital Markets Unit](#) is tasked with approaching competition and harms in digital markets. The [Information Commissioner's Office](#) obviously has an important role, given its responsibility for enforcing legislation like GDPR.

Government actions in other areas will have a significant bearing on data equity. For example, government funding for degree conversion courses to data science and AI, 'including £10m for up to 1,000 scholarships for people from diverse backgrounds', should diversify data science intakes and lead to greater diversity in the collection, use and application of data in the future.

Examples from other sectors and other countries

The US is establishing an ['equitable data working group'](#), noting that a lack of disaggregated data on protected characteristics 'has cascading effects and impedes efforts to measure and advance equity'.

The European Commission has an [equality data handbook](#) and other initiatives underway.

There is also increasing action around the inequities created by tech monopolies on both sides of the Atlantic, with Europe's [Digital Markets Act](#) and the US government's appointment of Lina Khan as Federal Trade Commission chair [prompting complaints](#) from Facebook and Amazon.

Further resources from the ODI

- [About Data About Us](#)
- [Data enabled cities](#)
- [The dividing line: how we represent race in data](#)
- [Inclusive data: perspectives from a roundtable discussion](#)
- [Monitoring Equality in Digital Public Services (report)](#)
- [One year of Diversity, Equity and Inclusion at the ODI](#)
- [Objective data? Reflections on the Commission for Race and Ethnic Disparities report](#)
- [The ODI responds to the UK Statistics Authority Inclusive Data Consultation](#)
- [The UK National Data Strategy 2020: equity, fairness and rights](#)
- [The weird and the wonderful: reflections on the Commission for Race and Ethnic Disparities report](#)

Ethics

What we mean

People and organisations must use data ethically. The choices made about what data is collected and how it is used should not be unjust, discriminatory or deceptive.

Ethics is about using data well and avoiding harmful impacts. This means maximising the benefits of data use by people, organisations, and governments, as well as thinking through the [consequences of how data use might harm people](#), communities and infrastructure (for example, public goods such as the environment).

Demonstrating that ethical processes are embedded in how data is collected, analysed and used can help build confidence: as [the National Data Strategy outlines](#), any 'transformation' in government's use of data 'will only be possible and sustainable if it is developed within a robust ethical framework of transparency, safeguards and assurance which builds and maintains public trust in the government's use of data'.

How the government getting data right can unlock the wider value of data across the economy

- **Leading:** Leading by example (including perhaps being seen to go beyond checklist exercises)

- **Supporting:** Providing resources and tools that can be used more widely outside government (such as the Data Ethics Framework and published research)

- **Collaborating:** Providing all the actors and agents in the economy with confidence about data flows through data assurance

- **Leading and collaborating:** Giving people and organisations confidence in the services it provides and strengthening their participation in the economy and society

- **Leading and collaborating:** Supporting an environment where businesses using data ethically can thrive through projecting ethical values globally

UK government strengths and weaknesses

'Ethics and public trust' is one of the five key areas of Mission 3 in the National Data Strategy: 'transforming government's use of data to drive efficiency and improve public services'. Both the National Data Strategy and the [Integrated Review](#) discuss shaping an open international order (including ethical values), which aims to support UK businesses. And the UK has attempted to establish and '[retain a global leadership position'](#) in the ethical and accountable use of data and AI (through establishing CDEI, work on standards and working on national strategies around data and AI) in recent years. All of this means a proliferation of ethics-related initiatives across government.

One of the most prominent was the Data Ethics Framework, 'a set of principles to guide the design of appropriate data use in the public sector', now overseen by the CDDO. As well as being used to embed ethical practice across the public sector, its principles can be useful to other sectors.

As well as the research and advice provided by CDEI, parts of the UK Statistics Authority (eg the National Statistician's Data Ethics Advisory Committee and the Centre for Applied Data Ethics) provide ethical advice to the National Statistician and the wider research community (their strategy, 'Statistics for the Public Good', says ethics will be ever more important); the NHSX AI Lab has an AI ethics initiative; and the Geospatial Commission has launched a project on the ethics of location data, which includes public engagement. We've previously argued reflective and deliberative approaches are better than prescriptive frameworks, and CDEI has established a Public Attitudes to Data and AI network across Whitehall to bring together various public engagement initiatives.

While all of these initiatives are encouraging and show some movement from theoretical considerations of ethics to more practical ones, there remain questions about how best to ensure individuals in the public sector, private sector and elsewhere pay attention to ethical concerns in their work on a practical, daily basis.

There may also be some work needed in the field of data assurance. Our data assurance work notes that there may be a need for a responsible body or named person to assure that data is being used and shared ethically. The National Data Strategy includes an example, not related to ethics, where the government can play such a role: the Food Standards Agency's Food Hygiene Ratings Scheme can help consumers and platforms make more informed decisions. The government's ongoing work around digital identity assurance – where providers of digital identity services are assured against a trust framework – is one example of where government work on ethics may help unlock wider economic value.

Examples from other sectors and other countries

Many civil society organisations and academic institutions have a strong focus on data ethics. As well as the ODI (for example, our Data Ethics Canvas), examples include the Ada Lovelace Institute and the Oxford Institute for Ethics in AI in the UK, and Data & Society in the US.

There are also private sector initiatives ranging from techUK's Digital Ethics Working Group to Microsoft's Responsible AI principles.

The Organisation for Economic Co-operation and Development keeps a watching brief on international initiatives and has produced its own Good Practice Principles for Data in the Public Sector.

Further resources from the ODI

- [Assurance, trust, confidence – what does it all mean for data?](#)
- [Covid-19: Identifying and managing ethical issues around data](#)
- [Data ethics and privacy](#)
- [Data ethics: how to be more trustworthy with data](#)
- [Expanding the global fight against misinformation with technology](#)
- [The next generation of data ethics tools](#)
- [The ODI Data Ethics Canvas](#)
- [ODI Summit 2020: Algorithmic transparency and the public sector roundtable](#)
- [Sharing models for Covid-19: guidance and tools](#)
- [The UK National Data Strategy 2020: doing data ethically](#)

Engagement

What we mean
Everyone must be able to take part in making data work for us all. Organisations and communities should collaborate on how data is used and accessed to help solve their problems.

The government's data initiatives will be more likely to succeed if they draw on the collective perspectives of the public, society and industry – they will highlight opportunities and challenges that the government itself may not have thought about, or have experience addressing. Engaging those outside government may also help build support for reforms and make them more likely to be implemented successfully.

But the ODI believes that 'engagement' goes beyond outreach activities. Collaborating on how data is used and accessed takes us into the realm of data governance, which should be interpreted as stewarding data for the best use, rather than a narrower definition focusing on data protection and compliance. Organisations that think seriously about their stewardship role and the positive economic and social outcomes they can achieve can unlock real value from their data. In our theory of change, we contrast this positive 'farmland' vision (where those who steward data act in ways that lead to the best social and economic outcomes for everyone) with more negative futures, such as the oilfield (where organisations decide to hoard data and little value is realised) and the wasteland (where organisations fear data and we end up with little data to build value upon).

How the government getting data right can unlock the wider value of data across the economy

- **Collaborating:** For 'engagement' in the classic sense, engaging those outside government on government's plans for data is vital for support, challenge and ultimately success

- **Leading and supporting:** Government engagement can provide others outside government with a blueprint on how to engage and build trust

- **Stewarding and collaborating:** For 'engagement' in the wider sense, being a responsible steward of data and collaborating properly with others allows the best possible use of data across the economy

UK government strengths and weaknesses
There may be more that the government can do in engaging the public in a conventional sense. And while some government organisations are starting to think seriously about their role as data institutions and how they collaborate with others, this is another area where more work can be done.

The National Data Strategy itself has been built around engagement, in a public dialogue sense: the initial framework strategy was based on a series of roundtables and other interactions, that strategy was then put out for consultation, and DCMS has established a forum to support the implementation of the strategy

(co-chaired by techUK) and a subgroup, the International Data Transfers Expert Council, to 'enable the government to deliver on its mission to champion the international flow of data'.

There are also many other government advisory boards with members from outside government, including the Council for Science and Technology (CST), which advises the Prime Minister on science and technology policy issues; the Digital Economy Council, which aims 'to harness the expertise of industry and the wider tech community'; the AI Council, which provides advice on the artificial intelligence ecosystem; the Open Standards Board, which is 'accountable for transparent selection and implementation of open standards'; and the Privacy and Consumer Advisory Group, which 'advises the government on how to provide users with a simple, trusted and secure means of accessing public services'. And there are other cross-government networks and communities to try to ensure different parts of the public sector engage with one another.

There are fewer initiatives engaging directly with the public. The National Data Guardian is a notable exception, with several of its outputs based on public dialogue. The CDEI has set up a cross-government Public Attitudes to Data and AI (PADAI) network to try to keep abreast of work on public attitudes. And there are other cross-government networks and communities to try to ensure different parts of the public sector engage with one another.

We also need to think more broadly about 'engagement' in terms of data stewards and data institutions – and this includes organisations not traditionally (or primarily) seen as data institutions. In government, this might include bodies as diverse as the Charity Commission (which holds data on all the UK's charities), the Oil and Gas Authority (which has a national repository of data on petroleum-related information), Our Future Health (a health research programme) and the UK Data Service ('a 'one-stop shop' for researchers to access a wide range of high-quality social and economic data'). We have previously argued that the UK government needs to think about how to support and scrutinise these institutions – especially in support of ambitions such as global AI leadership. Institutions in all sectors need to start thinking about themselves as data institutions, and there may be a need for new forms of data institution, such as data trusts. An ecosystem of such institutions and data stewards may also help plug any gaps in data that exist through being able to curate and share data from different sources.

Examples from other sectors and other countries

A prominent recent example of public engagement was [OneLondon's work on the use of health and care data](#).

[Understanding Patient Data](#) also supports public conversations around health data.

The Ada Lovelace Institute has [a repository of public attitudes work](#) relating specifically to Covid-19.

The ODI's own work on [data institutions](#) provides a wealth of resources on thinking about data stewardship. [Our collaboration with the Institute of Directors](#) focuses on how to support private sector leaders to think more widely about data governance as part of their digital transformation.

Further resources from the ODI

- [Assurance, trust, confidence – what does it all mean for data?](#)
- [Data About Us: 'the people' know and care more than they are given credit for](#)
- [Data institutions](#)
- [The Data Institutions Register](#)
- [How leaders can develop a data strategy that addresses global challenges](#)
- [The IoD and the Open Data Institute (ODI) join forces to upskill a generation of company directors in data governance](#)
- [The ODI's written evidence for APPG on Artificial Intelligence](#)
- [Trustworthy data institutions](#)
- [The UK National Data Strategy 2020: engaging for resilience](#)
- [Why poor data governance could be a director's undoing (podcast)](#)

Mapping the future

This [crowdsourcing exercise](#) gives a sense of the different data-related organisations across government, and the data-related initiatives currently underway. We hope it proves useful to those in government confronting the Spending Review and Budget.

But the Spending Review is just one of many data-focused initiatives at the centre of government at the moment. The new [Downing Street Delivery Unit (10DU)](#) and [an Evaluation Task Force](#) are both designed to use data to hold departments to account for their performance on government priorities. The new [Information and Data Exchange (INDEX)](#) in the Cabinet Office and [Integrated Data Service and Platform at the Office for National Statistics](#) are both designed to better share data and insights across government. The [Declaration on Government Reform](#), the Treasury's own [Plan for Growth](#) and the [National Data Strategy](#) also make much of using data to track government progress.

This [report, and the full crowdsourcing document,](#) therefore provide a map to help further exploration for those inside and outside government, as the government seeks to improve its own use of data – and those of us outside government seek to learn from it.