GOV•UK

Department for Digital, Culture, Media & Sport

Policy paper

National Data Strategy

Updated 9 December 2020

This was published under the 2019 to 2022 Johnson Conservative government

Contents

This publication is available at https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

Ministerial foreword

When I became Digital Secretary, I vowed to be unashamedly pro-tech. This has to begin with data. Data is now the driving force of the world's modern economies. It fuels innovation in businesses large and small, and has been a lifeline during the global coronavirus pandemic. The fact that governments, businesses, organisations and public services were able to share vital information quickly, efficiently and ethically during the pandemic has not only saved countless lives, but has enabled us to work from home, keep the economy running and stay connected with loved ones during a period of unprecedented disruption. As we enter into recovery, it is vital that we make the most of what we have learnt.

This National Data Strategy aims to do exactly that, building on our manifesto pledge to improve data use in government, and going further. It seeks to maintain the high watermark of data use set during the pandemic, and to free up businesses and organisations to keep using data to innovate, experiment and drive a new era of growth. It seeks to harness the power of data to boost productivity, create new businesses and jobs, improve public services and position the UK as the forerunner of the next wave of innovation.

Under this strategy, data and data use are seen as opportunities to be embraced, rather than threats against which to be guarded.

This means asking fundamental questions about what data should and should not be made available across the UK. It means maintaining a regulatory regime that is not overly burdensome for smaller businesses and that supports responsible innovation. It means driving a radical transformation of how the government understands and unlocks the value of its own data to improve a range of public services and inform decisions at scale, through a whole-government approach driven from the centre. It means taking the risks of increased data use seriously. And it means positioning the UK as a global champion of data use, and encouraging the international flow of information across borders.

The strategy is a central part of the government's wider ambition for a thriving, fast-growing digital sector in the UK, underpinned by public trust. We want the UK to be a nation of digital entrepreneurs, innovators and investors, the best place in the world to start and grow a digital business, as well as the safest place in the world to go online. We will set out more on how we propose to support a digital drive for growth in our Digital Strategy, which we will be publishing in the Autumn.

This document is a framework for the action this government will take on data. It is not the final answer, but part of a conversation about the way that we support the use of data in the UK. We lay out the opportunities that we want to realise, the pillars that we have identified as core to unlocking the power of data for the UK, and the missions that we must prioritise now.

Rt Hon. Oliver Dowden CBE MP
Secretary of State for Digital, Culture, Media and Sport

Executive summary

The opportunity

Better use of data can help organisations of every kind succeed – across the public, private and third sectors. It can support the delivery of existing services, from manufacturing to logistics, and it can be used to create entirely new products. It is a driver of scientific and technological innovation, and central to the delivery of a whole range of vital public services and societal goals, from tackling climate change to supporting the National Health Service. As businesses embrace technology, data creates jobs,

opens up whole new markets and drives demand for a highly skilled workforce.

On an individual level, the use of data benefits us every day – from the lives saved due to data-driven medical discoveries, to personal budgeting, understanding how much we have exercised and identifying better transport routes.

The UK is already a leading digital nation. The data market in the UK (i.e. money made from products or services derived from digitised data) is the largest in Europe (http://datalandscape.eu/study-reports/final-study-report-european-data-market-monitoring-tool-key-facts-figures-first-policy). UK tech grew dramatically in 2019 (https://technation.io/news/2019-a-record-year-for-uk-tech/), with the UK securing 33% of European tech investment.[footnote 1] Globally, the UK now sits behind only the US and China (https://technation.io/report2020/#forewords) in terms of venture capital investment.

But the last five years have seen huge technological changes, and national governments need to respond accordingly. We need a data strategy that reflects the opportunities and challenges of our new hyper-digital world, one that ensures we weigh the priorities and potential trade-offs of data in a deliberate and evidence-based way, and, above all, one that drives growth in the UK economy and powers our recovery from the coronavirus pandemic.

This strategy looks at how we can leverage existing UK strengths to boost the better use of data across businesses, government, civil society and individuals. Having left the European Union, we will take advantage of being an independent, sovereign nation to maximise those strengths domestically, and position ourselves internationally to influence the global approach to data sharing and use. We will act ambitiously at home and on the international stage, aligning our history of problem-solving in science and technology with progressive values and the competence and pragmatism of our regulatory institutions.

The UK response to the global coronavirus pandemic has powerfully illustrated the potential benefits of data. Our understanding of this disease, our ability to support people and our cooperation across borders have all relied on the responsible and effective use and sharing of data.

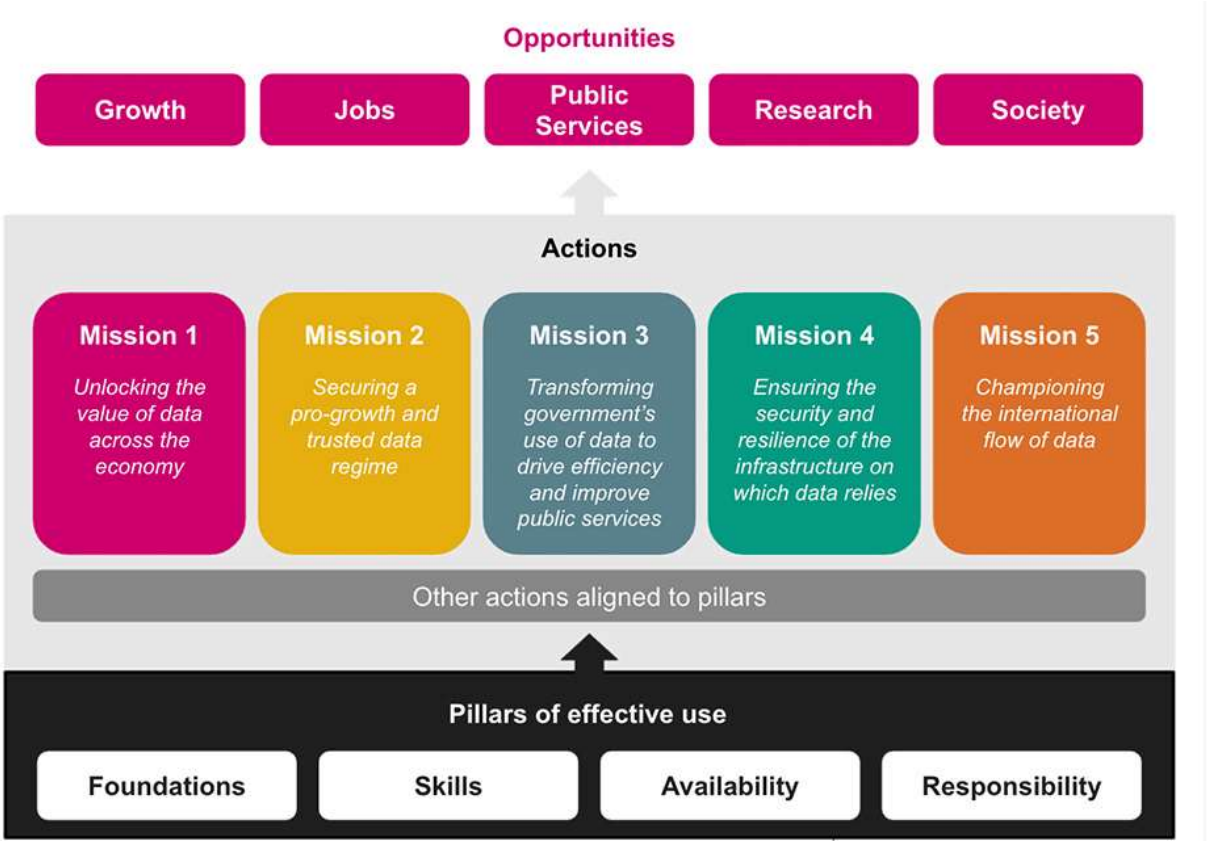
But we have a duty to do more – especially with the data that the government itself holds, which can be used and shared for the benefit of society. Individual transactions, from applying for social security benefits to buying a house, are more resilient when personal information can be shared from trusted sources – for example, in the form of a digital identity.

Data is a non-depletable resource in theory, but its use is limited by barriers to its access – such as when data is hoarded, when access rights are

unclear or when organisations do not make good use of the data they already have. These barriers undermine the performance of public services and our economy, risking poorer outcomes for citizens. We will ensure that data can be leveraged to deliver new and innovative services, promote stronger competition, and better prices and choice for consumers and small businesses. We will drive an approach to data that holds that all can benefit when data is used responsibly, and that withholding data can negatively impact society.

This strategy

As part of this document, we are asking for your views. Consultation questions are included throughout the text and in an accompanying document (https://www.gov.uk/government/consultations/uk-national-data-strategy-nds-consultation). In future updates, we will lay out the steps that we will take to implement the strategy, and the way that your responses and the evidence you have provided have shaped our approach.



The pillars

A number of interconnected issues currently prevent the best use of data in the UK. These are reflected in the core pillars of this strategy:

1. **Data foundations**: The true value of data can only be fully realised when it is fit for purpose, recorded in standardised formats on modern, future-proof systems and held in a condition that means it is findable, accessible, interoperable and reusable. By improving the quality of the data, we can use it more effectively, and drive better insights and outcomes from its use.

2. **Data skills**: To make the best use of data, we must have a wealth of data skills to draw on. That means delivering the right skills through our education system, but also ensuring that people can continue to develop the data skills they need throughout their lives.

3. **Data availability**: For data to have the most effective impact, it needs to be appropriately accessible, mobile and re-usable. That means encouraging better coordination, access to and sharing of data of appropriate quality between organisations in the public, private and third sectors, and ensuring appropriate protections for the flow of data internationally.

4. **Responsible data**: As we drive increased use of data, we must ensure that it is used responsibly, in a way that is lawful, secure, fair, ethical, sustainable and accountable, while also supporting innovation and research.

The missions

From these pillars, we have identified five priority areas of action. These missions address key challenges that can prevent us from taking advantage of the opportunities that data offers.

1 **Unlocking the value of data across the economy.** Data is an incredibly valuable resource for businesses and other organisations. However, there is increasing evidence to suggest its full value is not being realised because vital information is not getting to where it needs to be. To ensure the UK is a world leader in data, our first mission will be to set the correct conditions to make data usable, accessible and available across the economy, while protecting people's data rights and private enterprises' intellectual property. Using a considered and evidence-based approach, we will develop a clear policy framework to determine what government interventions are needed to do so.

2 **Securing a pro-growth and trusted data regime.** We want the data revolution to benefit businesses large and small. That means maintaining a data regime in the UK that is not too burdensome for the average company; one that helps innovators and entrepreneurs to use data responsibly and securely, without undue regulatory uncertainty or risk, to drive growth across the economy. But we also want the public to be active agents in the thriving digital economy and to have confidence and trust in how data, including personal data, is used. The UK's data regime will support vibrant competition and innovation, building trust and maintaining high data protection standards without creating unnecessary barriers to data use.

3 **Transforming government's use of data to drive efficiency and improve public services.** The coronavirus pandemic showed that there is massive untapped potential in the way government and public services use and share data to help and protect people. To sustain the high watermark set by the pandemic, the government will undertake an ambitious and radical transformation of its own approach, driving major improvements in the way information is efficiently managed, used and shared across government. To succeed, we need a whole-government approach that ensures alignment around the best practice and standards needed to drive value and insights from data; and the creation of an appropriately safeguarded, joined-up and interoperable data infrastructure to support this. We also need the right skills and leadership within the public sector to understand and unlock the potential of data.

4 **Ensuring the security and resilience of the infrastructure on which data relies.** The use of data is now a central part of modern life, so we need to make sure that the infrastructure underpinning it is safe and secure. The infrastructure on which data relies is a vital national asset that needs to be protected from security risks and other concerns, such as service disruption. Interruption to data-driven services and activities can cause disruption to businesses, organisations and public services. While these are also commercial risks to manage, the government has a responsibility to ensure that data and its supporting infrastructure is resilient in the face of established, new and emerging risks, protecting the economy as it grows.

5 **Championing the international flow of data.** The flow of information across borders fuels global business operations, supply chains and trade, powering growth across the world. It also plays a wider societal role. The transfer of personal data ensures people's salaries are paid, and helps them connect with loved ones from afar. And, as the coronavirus pandemic has demonstrated, sharing health data can aid vital scientific research into diseases while uniting countries in their response to global health emergencies. Having left the European Union, the UK will champion the benefits that data can deliver. We will promote domestic best practice and work with international partners to ensure data is not inappropriately constrained by national borders and fragmented regulatory regimes so that it can be used to its full potential.

Together, the steps identified in this strategy build on UK strengths to drive better use of data – data use that is more secure, more innovative and more widely recognised as a force for good. Better use of data will drive growth and productivity, improve our society and public services and position the UK as a leader of the next wave of data-driven innovation. We have an obligation to realise this ambition.

1. About the National Data Strategy

This strategy sets out how best to unlock the power of data for the UK. It builds upon initiatives such as the Industrial Strategy (https://www.gov.uk/government/topical-events/the-uks-industrial-strategy), the AI Review (https://www.gov.uk/government/publications/growing-the-artificial-

intelligence-industry-in-the-uk), the AI Sector Deal (https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal) and the Research and Development Roadmap (https://www.gov.uk/government/publications/uk-research-and-development-roadmap) – setting out a framework for how we approach and invest in data to strengthen our economy and create big opportunities for us in the future. The government believes that unlocking the value of data is key to driving growth both within the digital sector and across the economy. This will be part of our Digital Strategy, which will be published in the Autumn and will consider more broadly how we can support a digital drive for growth.

In this publication, we set out the framework for the approach this government will take, the improvements we seek to deliver and the priority missions we will focus on now to realise that change. The increasing importance of data raises novel and complex policy questions. Some of these need further consideration before the UK government confirms its direction. As such, we are also asking for your views in relation to our general framing, along with some of the actions we are considering. These questions are included throughout the document and also collated in an accompanying publication.

This consultation is on a UK-wide basis: we welcome responses from organisations and individuals across the UK. The strategy covers both reserved and devolved areas: where the strategy covers reserved areas (and, in respect of Northern Ireland, excepted areas), it does so for the whole of the UK, and where it covers devolved or transferred areas, it applies to England only. We will publish a response to this consultation in early 2021.

A National Data Strategy will require activity and focus beyond government. In this framework strategy we have focused on the government's role in harnessing data. Following this consultation period, and as we move to implementation, we will work with stakeholders to set out how we will work with business and actors across the wider data landscape to land a strategy for the whole of the UK.

What we mean by data

Data is notoriously hard to define – and it means different things to different people. For an application developer, data is what enables the creation of rich and complex digital services. For a scientist, it is what is collected as part of experiments or surveys. For a data protection practitioner, it is the names and addresses of staff organised in a spreadsheet. For a personal trainer, it is the information in an app recording our heart rate during a workout. And for every one of us, it is the tool that powers our online maps, helps us book supermarket delivery slots, and allows us to check tomorrow's weather forecast.

When we refer to data, we mean information about people, things and systems. While the legal definition of data covers paper and digital records, the focus of this strategy is on digital information. Data about people can include personal data, such as basic contact details, records generated through interaction with services or the web, or information about their physical characteristics (biometrics) – and it can also extend to population-level data, such as demographics. Data can also be about systems and infrastructure, such as administrative records about businesses and public services. Data is increasingly used to describe location, such as geospatial reference details, and the environment we live in, such as data about biodiversity or the weather. It can also refer to the information generated by the burgeoning web of sensors that make up the Internet of Things.

When thinking about the government's own data use, the strategy covers administrative, operational and transactional data – that is, data collected in the process of running services or businesses – as well as analytical and statistical data.

To ensure that data dependency risks are well managed, we are also interested in the infrastructure underpinning the storage of data, such as physical and virtualised data centres/the Cloud.

1.1 Evidence and working with stakeholders

This strategy, and the structure and substance of the pillars on which it is built, is drawn from a range of evidence sources, including desk research covering both case studies and published academic/sector research.

In June 2019, the government launched a call for evidence on the proposed framework of the strategy, receiving over 100 responses. Alongside this, a programme of stakeholder engagement was undertaken, including the hosting of 20 roundtables and workshops, with representatives from over 250 organisations across business, the third sector and local government. Through our call for evidence, roundtables and workshops held across the country in 2019, we consulted on our parameters and objectives, and gathered evidence that has underpinned this framework National Data Strategy.

We have analysed the call for evidence submissions and the discussions at the roundtables to make sure our evidence base is as wide and inclusive as possible. A summary of our call for evidence and of the stakeholder engagement can be found in the accompanying publication

(https://www.gov.uk/government/publications/uk-national-data-strategy/call-for-evidence-and-roundtable-engagement-summaries).

Data policy is a rapidly evolving area globally, and for many of the issues highlighted in this strategy a number of questions remain unanswered, with further research and analysis required. As the government takes this strategy forward, we will continue to draw upon available evidence to inform our actions. We will develop a monitoring and evaluation process for the strategy to help ensure it is achieving its intended outcomes, as well as building the evidence base on which we develop and evaluate future policy decisions. As we move to implementation, we will further continue to work with stakeholders to set out how we will engage with business and the wider data economy to land a strategy for the whole of the UK.

2. The data opportunity

We are currently in the middle of a fourth industrial revolution. Technological innovation has transformed our lives, changing the way we live, work and play. At the same time, this innovation has brought with it an exponential growth in data: in its generation and use, and in the world's increasing reliance upon it.[footnote 2] By embracing data and the benefits its use brings, the UK now faces tangible opportunities to improve our society and grow our economy. If we get this right, data and data-driven technologies like AI could provide a significant boost to the entire economy. Data can improve productivity and provide better-quality jobs. But it can also transform our public services and dramatically improve health outcomes nationally. It can keep us safe and assist the reduction of crime, speed the journey to decarbonisation, and, used well, drive efforts to create a more inclusive, less biased society.

Importantly, data can also be used to harness the potential of regions right across the country,[footnote 3] ensuring that people and organisations from the whole of the UK can benefit from the full value of the digital revolution.

Like many things, the use of data also presents risks; those risks need to be fully understood and taken into account. Used badly, data could harm people or communities, or have its overwhelming benefits overshadowed by public mistrust. Equally, misplaced government reluctance to securely share and use data undermines the performance of public services and risks causing harm by missing opportunities to help those most in need. In the same way, unnecessary barriers to technological innovation could drive inefficiencies and slow down growth. So it is vital that we take decisive and evidence-led steps to make the most of data's potential.

We have identified five concrete and significant opportunities for data to positively transform the UK:

1. Boosting productivity and trade
2. Supporting new businesses and jobs
3. Increasing the speed, efficiency and scope of scientific research
4. Driving better delivery of policy and public services
5. Creating a fairer society for all

Achieving these opportunities will not be easy. While they are already being realised in some contexts, the means to do so are missing in many others. There is also a great deal of competition in the data space internationally, combined with differing global views on data and its use. As a digital leader, the UK is well placed to overcome this challenge.

In realising these opportunities, the UK can further enhance its world-class status in science and technology, and its reputation for finding pragmatic and innovative solutions to difficult problems. In the years to come, it can use these strengths as a springboard to become a global leader in data as well.

2.1 Boosting productivity and trade

Data is knowledge. By having access to more of it, combined with the ability to analyse it through modern techniques, we get greater insight into what works and what does not – both in terms of selling products and services, and in terms of making our own processes and practices more efficient. Data therefore has the potential to significantly enhance economic competitiveness and productivity across the UK economy, through new data-enabled business models, as well as through the adoption of data-driven processes by existing businesses.

There are various ways of defining and measuring the data market and the data economy.[footnote 4] A range of estimates suggest that the data economy grew about twice as quickly as the rest of the economy during the 2010s, making up about 4% of UK GDP in 2020 (http://datalandscape.eu/sites/default/files/report/D2.9_EDM_Final_study_report_16.

06.2020_IDC_pdf.pdf). Beyond the impacts of data-driven products and services (i.e. the direct 'data economy'), the use of data has a more general role in underpinning digitally delivered trade. ONS estimates show that in 2018 the UK exported £190 billion in digitally delivered services (67% of total UK services exports) and imported £90 billion digitally delivered services (52% of UK services imports) (https://www.ons.gov.uk/releases/modesofsupplyexploratoryestimatesfortheuk). Enabling and growing this data-driven trade will be a priority in our approach to free-trade negotiations.

While research into the business impact of increasing and improved data use is in its infancy and methodologically challenging, the existing evidence suggests wide-ranging economic benefits arising from better data use, in particular an association between efficiency, productivity and data-driven business practices.[footnote 5] There are also significant economic advantages from individual companies increasing data access and sharing. For example, Transport for London (TfL)'s opening up of its data sets to travellers and third-party providers contributed up to £130 million per year to the London economy through time saved by travellers (http://content.tfl.gov.uk/deloitte-report-tfl-open-data.pdf).

In recent years, the UK government has taken significant and unprecedented steps to position the UK as a world leader in data-driven innovation. This includes committing to raising investment in (often data-heavy) research and development by 2.4% of GDP by 2027, establishing institutions such as the Centre for Data Ethics and Innovation (CDEI) and the Alan Turing Institute, launching brand new conversion courses in data science and AI, and conducting pioneering work on 'data trusts' – a novel data-sharing framework.[footnote 6] The importance of continuing and furthering this work will be even more vital in ensuring the UK's recovery from the coronavirus pandemic, and its economic success in the years beyond.

Case study: Sharing data to drive open innovation

Open innovation allows companies to apply externally-developed data, ideas and technology to help address challenges. Sharing operational data provides the necessary insights into business challenges, allowing collaborators to analyse and use that data to deliver better insights and demonstrate the value of new technologies.

The Open Data Institute documented how an open innovation programme, Data Pitch, has allowed Greiner Packaging International GmbH, a company producing rigid plastic packaging, to share data with logistics intelligence company, Obuu, to help monitor the resilience and efficiency of its supply chain. To do this, Obuu used data to map out shipping, storage and manufacturing flows so they could investigate three key performance indicators for efficiency in the supply chain: whether spare parts were available when needed; the average time the system is down when a part is not available; and the overall investment in the accessible stock. From this, Obuu was able to identify reductions in fixed asset investment, resulting in a significant cost saving (https://datapitch.eu/wp-content/uploads/2020/01/London-Economics-Data-Pitch-evaluation-FINAL-PDF.pdf).

2.2 Supporting new businesses and jobs

Data skills – like many digital skills – are increasingly important for all aspects of life, but especially for the working environment. Increasing numbers of jobs require technical data skills. One estimate suggests there was over a 50% increase in data professionals between 2013 and 2020 – increasing from 1.1m to 1.7m employees. However, there are also over 100,000 unfilled data professional posts in 2020.[footnote 7]

But beyond these jobs requiring technical data skills, most jobs require some data and digital skills, and they are only set to increase in importance. Estimates from a DCMS-commissioned module of the February 2020 ONS 'Opinions and Lifestyle' survey found that nearly half (48%) of the working population use 'basic' data skills at work a lot, and just under a quarter (24%) use more advanced skills, such as data analysis and making graphs, a lot in their current job.[footnote 8]

We know that data is a basis for the creation of new jobs that use both general and more technical data skills. TfL's work on open data (below) is estimated to have created 500 jobs directly, and another 230 indirectly.[footnote 9] More widely, the European Data Portal predicts a baseline scenario of 1.12m open data employees and an optimistic scenario of 1.97m open data employees in the EU by 2025 (https://www.europeandataportal.eu/sites/default/files/the-economic-impact-of-open-data.pdf). And over 6% of new companies in the UK in 2019 were tech start-ups (https://centreforentrepreneurs.org/cfe-research/business-startup-index/#:~:text=Over%2045%2C000%20tech%20startups%20were,in%20Bournemouth%2C%20Poole%20and%20Christchurch.), with data and the aligned need for increasingly sophisticated data skills underpinning their business models.

By encouraging and supporting the use of data in the UK, we can ensure that the coming waves of technological innovation do not just drive new services but also foster the creation of new businesses and new jobs for the UK.

Case study: Using data in the transport sector

'Intelligent Mobility', the moving of people and goods in an easier and more efficient way, provides an intersection between traditional transport and exciting new products and services relating to mobile devices, open data, wireless communication or the Internet of Things. According to research commissioned by the Transport Systems Catapult, the international Intelligent Mobility market is estimated to grow to just over £900bn a year by 2025, with data comprising an estimated £32bn of this per year by 2025 (https://ts.catapult.org.uk/wp-content/uploads/2016/08/TSC-Intelligent-Mobility-Market-Update-2016-003.pdf).

TfL already uses technology and data to make journeys easier for their customers. By making live travel information available, developers can create software and services, such as online maps and journey planners. The data is provided via an API, which unifies the data for different modes of transport into a common format and structure, where historically they have differed and been hard to layer onto each other.

This data has been used to develop applications used by the public to plan journeys and check for disruptions and has helped create hundreds of jobs. Recent research by Tech City noted London's digital economy was worth £30bn in GVA supporting over 300,000 jobs (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/642813/15780_TSC_Market_Forecast_for_CAV_Report_FINAL.pdf), and the government's Future of Mobility strategy (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/846593/future-of-mobility-strategy.pdf) describes some of

the actions being taken to maximise the benefits of data use in the UK's transport sector.

2.3 Increasing the efficiency and scope of scientific research

The UK is a leader in science and research, and data is at the heart of it. New scientific developments driven by data have potentially game-changing applications across the economy, such as tracking public health risks and aiding decarbonisation through smarter energy grids, predictive maintenance of infrastructure or better traffic management.

While data is critical in all research, some of the clearest examples of the benefits to society are in the life sciences. For example, data has been crucial in recognising and understanding the side effects of drugs, identifying the benefits of surgery for patients with Inflammatory Bowel Disease and demonstrating the impact of anti-smoking laws on the number of babies born prematurely in Scotland (https://mrc.ukri.org/research/initiatives/health-and-biomedical-informatics/value-of-using-data/). The UK has also published five principles which will underpin the government's policy framework to govern fair, ethical and appropriate use of health data, while also supporting innovation. More advanced applications of data-driven technology have also provided responses to the coronavirus pandemic, with AI-driven systems being used to predict the virus's protein structure (https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery) and determine which drugs may be effective for treatment, helping prioritise promising candidates for real-world trials (https://www.benevolent.com/news/potential-treatment-for-covid-19-identified-by-benevolentai-using-artificial-intelligence-enters-randomised-clinical-trial).

However, barriers to accessing data represent a significant limitation on research; these range from legal barriers (real and perceived) through to cultural blockers and risk aversion. These barriers must be addressed if the UK is to remain at the forefront of science and research. For example, recent research into data use by the pharmaceutical and life sciences industry (https://s3-eu-west-1.amazonaws.com/media.newmd.catapult/wp-content/uploads/2019/10/22170649/health-data-report.pdf) identified a number of systemic barriers that limit access to data. Most companies surveyed noted having experienced delays and uncertainties. These include time taken to access data, access constraints for commercial users, the effort to identify and assess the quality of data sets and, most notably, the cost of the data itself.

When dealing with sensitive data, the way forward must be considered and appropriate. For example, NHSX (https://www.nhsx.nhs.uk/) is developing a Data Strategy for Health and Social Care in Autumn 2020, which will aim to build on the permissive approach to data sharing for care seen during the coronavirus response, while still protecting the absolute need for patient confidentiality. A similar balance is needed in other instances where risks must be managed while unlocking the significant opportunities from data use at the forefront of science, research and technological development.

Case study: Data-driven clinical trials to tackle coronavirus

Conducting clinical trials that are better, faster and more efficient is a UK priority.

Electronic Health Records (EHR) play a key role in identifying and enabling individuals from across the UK to take part in clinical trials. The UK is well-placed to take advantage of EHRs due to the unified nature of the NHS within the four nations and because a large portion of NHS records are electronic. Using securely stored data, we can find patients from across the UK that might benefit from a particular trial, rather than just a single NHS Trust or geographical location. All while protecting patients' data.

This approach was tested during the response to coronavirus and the RECOVERY trial (https://www.recoverytrial.net/) for potential treatments. These trials normally take months to set up but, designated as an Urgent Public Health Research study, RECOVERY was set up in a

matter of days. 12,000 patients from 176 NHS hospital organisations were enrolled in the trial with data being tracked and analysed.

Within 100 days, the trial identified the world's first coronavirus treatment proven to save lives – Dexamethasone. The results were announced on 16 June 2020, adopted into UK practice later the same day and included in new US guidelines within 2 weeks.

RECOVERY was supported by NHS DigiTrials (https://digital.nhs.uk/services/nhs-digitrials), the Health Data Research UK (https://www.hdruk.ac.uk/) hub for clinical trials, which provided centrally collected and curated data on a weekly basis so that progress and outcomes of trial participants could be closely tracked. The data was drawn from extracts from routine EHR data, and supported rapid access to new data assets collected in response to the coronavirus pandemic, such as test data, intensive care data, and GP data. The approach not only removed the burden on the NHS 'front line' to field additional data requests, but enabled the RECOVERY trial team to make rapid decisions.

2.4 Driving better delivery of policy and public services

Data can revolutionise the public sector, creating better, cheaper and more responsive services. Public services are complex to deliver, with services such as the pensions system, the benefits system, the NHS, tax and the courts each engaging with millions of people across the UK every year. Likewise, keeping people safe requires access to the right information. These services and capabilities rely heavily on data, but the systems that handle this data have grown iteratively and independently, increasing in complexity over time. Many legacy systems are out of date, costly to operate and incapable of exchanging data with one another, presenting challenges in a world where public services are increasingly interconnected – be that between health and social care provision, tax and benefits or across policing, courts, prisons and probation.

Our experience responding to the coronavirus pandemic has demonstrated that when we treat data as a strategic asset and improve coordination between organisations, the delivery of services can be more agile, more innovative, more effective and more cost-effective. Indeed, it has underlined the need for the public sector to move away from a culture of risk aversion towards a joined-up approach, where the presumption is that, with appropriate safeguards, data should be shared to drive better outcomes. The rollout of the Coronavirus (COVID-19) Shielded Patients List

(https://digital.nhs.uk/coronavirus/shielded-patient-list) showed how much can be achieved through appropriate data sharing across central and local government and the private sector, with over four million support packages distributed to some of the most vulnerable people in society.

For central government, better data also means better decision-making. It means policies that can be tailored and delivered more efficiently and significant savings for the public purse. Better evidence on whether policies are delivering their intended effects in different areas and for different groups means interventions can be far more effectively designed. This aligns with the public's new expectations in our increasingly digital context. As highlighted by the CDEI's public attitudes research (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/864167/CDEJ7836-Review-of-Online-Targeting-05022020.pdf), for example, 'there is an expectation that the public sector should use online targeting to ensure that advice and services are delivered as effectively as possible.

The problem and the opportunity are not limited to central government. Some of the biggest benefits can be realised by better, more coordinated use of data across the wider public sector – in education, the justice system, health and within local government. As we move to implementation, we will work with partners to better understand the needs and barriers faced by local government in utilising data to its fullest potential. We will cut down on bureaucratic burdens, tackle risk aversion and strengthen the incentives to share data across the public sector. Non-standardisation and a lack of coordination on data mean that data collected by one organisation cannot easily be used by another. This results in duplication of effort and wasted resources. Treating data in the public sector as a strategic asset, with appropriate governance, will save time and money and drive better outcomes for us all.

Case study: Data First

Data First is an ambitious, pioneering data-linking programme led by the Ministry of Justice (MoJ) and funded by UK Research and Innovation through Administrative Data Research UK. The projects it supports will enable researchers in government, universities and other institutions to securely access anonymised extracts of linked administrative datasets held by MoJ and its executive agencies. This can then be linked with data held by other government departments, such as the Department for Education.

Data First will allow researchers to understand how people interact with courts over time and analyse which characteristics influence patterns of frequent use, all to build a much better understanding of what MoJ policies and services are most effective. Researchers will be able to explore how users of the justice system interact with other government services. This will enable a deeper understanding of how the economic, social and educational backgrounds of people who use the justice system influence their needs, the pathways they follow through the system (for example, between the civil and criminal courts) and the outcomes they experience. Such understanding will enable more evidence-informed, targeted support and lead to lower cost, higher-quality public services for everyone in the UK.

Data access will be facilitated by the controlled circumstances of the ONS Secure Research Server, an accredited processor under the Digital Economy Act 2017, which complies with the highest standards of data security and protection outlined by the principles of the Five Safes

(https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes).

2.5 Creating a fairer society for all

Data holds great potential to empower people and civil society, delivering benefits that reach beyond the economy. Powered by better data, civil society organisations can be better equipped to reach the people most in need, at the time they most need it. Better data use could also significantly decrease operating costs, allowing charities to focus resources on protecting the most vulnerable parts of our society. Charities and other non-profits, and particularly smaller organisations, rarely have access to large enough datasets to be able to prove, to very high levels of certainty, the effectiveness of different interventions. Better coordination, re-use and sharing of data between civil society organisations can also lead to better understanding of societal issues, and of what interventions are effective in supporting those most at need. Data can drive applications that make our digital lives better. Artificial intelligence is increasingly being used to drive automated content moderation online, particularly in social media contexts, where it can help tackle misinformation (https://www.ofcom.org.uk/__data/assets/pdf_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf). Data-driven online profiling technologies can help identify potentially vulnerable web users (such as people suffering from gambling addiction), and target support or prevent them from seeing potentially harmful content.

We can harness data as part of struggles to tackle bias and exclusion. Data can be used to hold a mirror up to society – to understand how different groups are faring, and to ensure that government and private sector actions treat people fairly, and are not unintentionally discriminating against protected groups.[footnote 10] We must further ensure that data-driven technologies and AI are a force for good. Biases arising from data or algorithm use will need to be addressed to ensure that data's potential is harnessed to drive a better, more inclusive and less biased society rather than entrenching existing problems. This is part of using data in an ethical and responsible way.

Case study: Domestic abuse statistics tool

The Office for National Statistics (ONS) domestic abuse publication brings together data from a range of sources, including the Crime Survey for England and Wales, police recorded crime, other government organisations and domestic abuse services. When taken in isolation, these data sources may not provide the context required to understand the national and local picture of domestic abuse. However, bringing this data together enables appropriate action to be taken to improve victims' experiences and to help provide a clearer understanding of the criminal justice system's response to perpetrators of abuse.

Alongside the publication, there is also an interactive data tool (https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/domesticabuseinenglandandwalesdatatool) for domestic abuse statistics. This allows users to explore data for their police force area in more detail and compare this with data from other areas. The tool is intended to help shape the questions that need to be answered by police forces and other agencies working with victims and responding to perpetrators of domestic abuse.

The statistics are used to monitor the UK's progress toward the 17 Sustainable Development Goals adopted by all UN Member States in 2015, as part of the 2030 Agenda for Sustainable Development.

2.6 Realising the Data Opportunity

While the preceding examples show the significant promise of better data use, there are considerable challenges preventing us from realising this more broadly and consistently across our economy and society. Organisations do not always make the best use of the data they hold, whether due to a lack of skills, a lack of leadership or a lack of resources – government and the wider public sector provide examples of this. Many organisations are limited in their access to data, much of which is controlled by a small number of key players. When data is available, it may be in formats that are unhelpful or of undetermined accuracy. And while the UK does have a wealth of data skills, these are concentrated in areas of UK expertise like science and technology; we have identified an overall lack of data skills across the workforce as a whole.

To harness the opportunities and realise our vision, we need to drive improvement across the entire data landscape. Through thematic analysis of the responses to our call for evidence, stakeholder engagement and reviewing the wider evidence base, we have organised this into four highly interconnected pillars that describe the basis for better data use. These are:

These are:

> **Data foundations**: The true value of data can only be fully realised when it is fit for purpose, recorded in standardised formats on modern, future-proof systems and held in a condition that means it is findable, accessible, interoperable and reusable. By improving the quality of the data, we can use it more effectively and drive better insights and outcomes from its use.

> **Data skills**: To make the best use of data, we must have a wealth of data skills to draw on. That means delivering the right skills through our education system, but also ensuring that people can continue to develop the data skills they need throughout their lives.

> **Data availability**: For data to have the most effective impact, it needs to be appropriately accessible, mobile and re-usable. That means encouraging better coordination, access to and sharing of data of appropriate quality between organisations in the public sector, private sector and third sector, and ensuring appropriate protections for the flow of data internationally.

> **Responsible data**: As we drive increased use of data, we must ensure it is used responsibly, in a way that is lawful, secure, fair and ethical, sustainable and accountable, while supporting innovation and research.

To ensure that we drive focused change, we have identified five priority missions (outlined in the next section) where the government will emphasise activity across these pillars to begin realising the data opportunity set out above.

Questions on the framing of the strategy

We want to ensure that we produce a forward-looking strategy that takes into account public opinion and delivers real change. These questions will help to inform future work that the government will take in this space. It will provide evidence for the government to target areas for intervention in future policy.

Q1. To what extent do you agree with the following statement: Taken as a whole, the missions and pillars of the National Data Strategy focus on the right priorities. Please explain your answer here, including any areas you think the government should explore in further depth.

NB: For question 2, we are only looking for examples outside health and social care data. Health and social care data will be covered in the upcoming Data Strategy for Health and Social Care.

Q2. We are interested in examples of how data was or should have been used to deliver public benefits during the coronavirus (COVID-19) crisis, beyond its use directly in health and social care. Please give any examples that you can, including what, if anything, central government could do to build or develop them further.

Q3. If applicable, please provide any comments about the potential impact the proposals outlined in this consultation may have on individuals with a protected characteristic (https://www.equalityhumanrights.com/en/equality-act/protected-characteristics) under the Equality Act 2010?

Q4. We welcome any comments about the potential impact the proposals outlined in this consultation may have across the UK, and any steps the government should take to ensure that they take account of regional inequalities and support the whole of the UK.

3. Missions

This strategy sets out five priority areas of action for the government. By delivering against these missions, we will create the optimal environment for data to drive growth and productivity in the UK for the benefit of all, while helping to solve a number of societal and global issues.

Mission one: Unlocking the value of data across the economy

Data is an incredibly valuable resource for businesses and other organisations, helping them to deliver better services and operations for their users and beneficiaries. However, there is increasing evidence to suggest the full value of data is not being realised because vital information is not getting to where it needs to be.

For example, the Digital Competition Expert Panel's review of competition in Digital Markets and the Competition Market Authority (CMA)'s report into online platforms and digital advertising (https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf) highlighted that smaller companies often do not have the same access to data as tech giants, potentially limiting their participation and innovation in digital markets. Improved public sector access to data can also lead to better decision-making at scale. For example, if the government had better data about infrastructure, it could reduce the disruption caused when underground pipes and cables are struck by mistake, or drive more informed choices about where to build new housing.

Our first mission is to create an environment where data is appropriately usable, accessible and available across the economy – fuelling growth in organisations large and small.

Much of the transformative potential of data lies in the potential for linkage and re-use of datasets across organisations, domains and sectors. We must ensure that the right conditions and incentives are in place to encourage organisations to work together across the economy, ensuring appropriate and timely access to data that is of sufficient quality. This can aid innovation, ensure the benefits of data can be realised by the maximum possible people in society and aid scientific research.

This is not simply a case of opening up every dataset. We must take a considered, evidence-based approach: government interventions to

increase or decrease access to data are likely to have myriad consequences, intended and not. There is a balance to be struck between maintaining incentives to collect and curate data, and ensuring that data access is broad enough to maximise its value across the economy. For personal data, we must also take account of the balance between individual rights and public benefit.

This is a new and complex issue for all digital economies, one that has come to the fore as data has become a significant modern, economic asset. The first step must therefore be the development of a clearer policy framework to identify where greater data access and availability across and with the economy can and should support growth and innovation, in what form, and what government's role should be, in the UK and globally. We will move quickly to build that framework in the coming months by:

- undertaking research to further develop our evidence base on the timely availability of appropriate quality data, set out the economic case and understand the opportunities and rationale for intervention
- drawing on that work, and our existing evidence base, to scope out government's potential role, both with respect to short-term quick wins and longer-term projects
- piloting the most promising interventions, working closely with industry and expert groups
- working closely with the CDEI, the Open Data Institute (ODI) and others to leverage their expertise and capability to support delivery of this agenda

We are proposing the creation of a framework to identify where we can and should make data available in the wider economy. There are a number of ways the government can intervene to achieve this goal – including as a collaborator, steward, customer, provider, funder, regulator and legislator. Using Policy Lab's Style of Government Action (https://openpolicy.blog.gov.uk/2020/03/06/introducing-a-government-as-a-system-toolkit/) could be helpful in thinking about the next few questions.

These questions will provide an opportunity for the government to scope out areas of focus for the data availability framework.

Data availability: For data to have the most effective impact, it needs to be appropriately collected, accessible, mobile and re-usable. That means encouraging better coordination, access to and sharing of data of appropriate quality between organisations in the public sector, private sector and third sector, and ensuring appropriate protections for the flow of data internationally.

Q5. Which sectors have the most to gain from better data availability? Please select all relevant options listed below, which are drawn from the

Standardised Industry Classification (SIC) (https://onsdigital.github.io/dp-classification-tools/standard-industrial-classification/ONS_SIC_hierarchy_view.html) codes.

- Accommodation and Food Service Activities
- Administrative and Support Service Activities Agriculture, Forestry and Fishing
- Arts, Entertainment and Recreation
- Central/ Local Government inc. Defence
- Charity or Non Profit
- Construction
- Education
- Electricity, Gas, Steam and Air Conditioning Supply
- Financial and Insurance Activities
- Human Health and Social Work Activities
- Information and Communication
- Manufacturing
- Mining and Quarrying
- Transportation and Storage
- Water Supply; Sewerage, Waste Management and Remediation Activities
- Wholesale and Retail Trade; Repair Of Motor Vehicles and Motorcycles
- Professional, Scientific and Technical Activities
- Real Estate Activities
- Other

Q6. What role do you think central government should have in enabling better availability of data across the wider economy?

Q6a. If yes, what is it? If not, why not? How does this vary across sectors and applications?

Data foundations: The true value of data can only be fully realised when it is fit for purpose, recorded in standardised formats on modern, future-proof systems and held in a condition that means it is findable, accessible, interoperable and reusable. By improving the quality of the data we are using, we can use it more effectively, and drive better insights and outcomes from its use.

Q7. To what extent do you agree with the following statement: The government has a role in supporting data foundations in the wider economy. Please explain your answer. If applicable, please indicate what you think the government's enhanced role should be.

Q8. What could central government do beyond existing schemes to tackle the particular barriers that small and medium-sized enterprises (SMEs) face in using data effectively?

The Smart Data Review (https://www.gov.uk/government/consultations/smart-data-putting-consumers-in-control-of-their-data-and-enabling-innovation) in 2019 consulted on ways to make evolving schemes more coordinated across banking, finance, telecoms and energy. The focus of Smart Data is customers asking their providers to share information about them with third parties who then use this data to offer innovative services to consumers and SMEs.

Q9. Beyond existing Smart Data plans, what, if any, further work do you think should be done to ensure that consumers' data is put to work for them?

Mission two: Securing a pro-growth and trusted data regime

As the world becomes increasingly digitised, data has become a central driving force of the modern economy. So it is vital that the UK has a data regime that promotes growth and innovation for businesses of every size, while maintaining public trust.

The UK is already a world leader in technological innovation and robust data protection standards: two areas required to build and maintain privacy, security and public confidence.

We will build on these strengths to maintain a data regime that supports the future objectives of the UK outside of the EU. A pro-growth legal regime must include consideration of both regulation in the wider digital and technology landscape, which will be addressed in the government's forthcoming Digital Strategy, as well as our data protection laws.

As with all policy areas, the UK will control its own data protection laws and regulations in line with its interests after the end of the transition period. We want our data protection laws to remain fit for purpose amid rapid technological change. Far from being a barrier to innovation or trade, we know that regulatory certainty and high data protection standards allow

businesses and consumers to thrive. We will seek EU 'data adequacy' to maintain the free flow of personal data from the EEA, and we will pursue UK 'data adequacy' with global partners to promote the free flow of data to and from the UK and ensure that it will be properly protected.

But data is now a far more influential force in our economy than ever before – with the potential to affect the structure and competitiveness of entire markets. This has serious implications for innovators, not least in the way our approach to data affects the ease, costs and risks of developing new technologies and services. The government needs to create the conditions to support vibrant competition and innovation, which will in turn drive future growth.

To build a world-leading data economy, we must maintain and bolster a data regime that is not too burdensome for the average company – one that helps innovators and entrepreneurs to use data legitimately to build and expand their businesses, without undue regulatory uncertainty or risk in the UK and globally.

Given the rapid innovation of data-intensive technologies, we also need a data regime that is neither unnecessarily complex nor vague. Businesses need certainty to thrive, and the government will work with regulators to prioritise timely, simple and practical guidance, especially for emerging technologies, and create more opportunities to experiment safely.

We want to encourage the widespread uptake of digital technologies more broadly – both for the benefit of the economy and wider society. We will work with regulators to provide more support and advice to small and medium-sized businesses to help them expand online, lifting compliance burdens where possible. We will also prioritise the development of sector-specific guidance and co-regulatory tools to accelerate digitisation across the UK economy.

Amid all this technological change, we want people to be active agents in the digital revolution. This is a shared responsibility of both businesses and individuals.

- **Businesses** and other data-using organisations should be clear and transparent about how they collect and use data responsibly. Working with the CDEI, the government will partner with industry to identify and incentivise best-practice.
- **People** should be empowered to choose whether and how to share data in both the public and private sectors, including where the use of their data can help others. In turn, the government will remain committed to high data protection standards so that data processing is fair and does not result in discriminatory outcomes.

People, businesses and other data-using organisations need the right data skills to participate actively. Individuals should have the basic data skills to be able to engage with and understand what is happening to their data. UK organisations will need access to top talent in data science, data engineering and related fields, as well as data-literate workforce.

Q10. How can the UK's data protection framework remain fit for purpose in an increasingly digital and data driven age?

In section 7.1.2 we lay out the functions of the Centre for Data Ethics and Innovation (CDEI), set up in 2018 to advise the Government on the use of data-driven technologies and AI.

Q11. To what extent do you agree with the following statement: the functions for the Centre for Data Ethics and Innovation (CDEI) should be Artificial Intelligence (AI) monitoring, partnership working and piloting and testing potential interventions in the tech landscape?

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

Q11a. How would a change to statutory status support the CDEI to deliver its remit?

Mission three: Transforming government's use of data to drive efficiency and improve public services

There is massive untapped potential in the way the government uses data. The coronavirus pandemic showed how much can be achieved when government departments and the wider public sector share vital information to solve problems quickly. We have a duty to maintain that high watermark after the pandemic, and will implement major and radical changes in the way the government uses data to drive innovation and productivity across the UK. In doing so, we will improve the delivery of public services, as well as our ability to measure the impact of policies and programmes, and to ensure resources are used effectively.

There is already consensus amongst experts both inside and outside government – including academics, civil society and parliamentarians – on the need to address this challenge and to capitalise on the opportunities.

However, there are numerous obstacles to achieving our ambitions, many of which are long-term and systemic. These include: real and perceived legal and security risks of sharing data; a lack of incentives, skills or investment to drive effective governance and overhaul data infrastructure; and a lack of consistency in the standards and systems used across the government, making it hard to share data efficiently.

These obstacles are not insurmountable, and we have both the ambition and the commitment to tackle them.

To succeed, we need a whole-government approach led by a Government Chief Data Officer from the centre in strong partnership with organisations. We need to transform the way data is collected, managed, used and shared across government, including with the wider public sector, and create joined-up and interoperable data infrastructure. We need the right skills and leadership to understand and unlock the potential of data – and we need to do so in a way that both incentivises organisations to do the right thing, as well as build in the right controls to drive standardisation, consistency and appropriate data use.

To achieve this objective, we will need to drive change across five key areas:

1. **Quality, availability and access**: striving towards improved data quality that is consistent, a clear understanding of what data is held and where, better data collection, and efficient data-sharing between organisations. All should be the norm, rather than the exception.

2. **Standards and assurance**: setting and driving the adoption of standards for data, leading to greater consistency, integrity and interoperability, and enabling data to be used widely and effectively across government.

3. **Capability, leadership and culture**: developing world-leading capability in data and data science across central and local government, so that leaders understand its role, expert resource is widely available, staff at all levels have the skills they need, and a ‘data-sharing by default’ approach across government tackles the culture of risk aversion around data use and sharing.

4 **Accountability and productivity**: opening government up to greater scrutiny and increasing accountability, ensuring that this drives improvements in productivity, policy and services for people, while also ensuring data security; and using procurement to drive innovation and better outcomes.

5 **Ethics and public trust**: this transformation 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.

The government is going to set an ambitious package of work in this space and wants to understand where we can have the biggest impact.

Q12. We have identified five broad areas of work as part of our mission for enabling better use of data across government:

- Quality, availability and access
- Standards and assurance
- Capability, leadership and culture
- Accountability and productivity
- Ethics and public trust

We want to hear your views on which of these actions will have the biggest impact for transforming government's use of data.

Q13. The Data Standards Authority is working with a range of public sector and external organisations to coordinate or create data standards and standard practices.

We welcome your views on which if any should be prioritised.

Mission four: Ensuring the security and resilience of the infrastructure on which data relies

With data now a critical part of modern life, we need to ensure the infrastructure underpinning it is safe, secure and resilient. The infrastructure on which data relies is a vital national asset – one that supports our economy, delivers public services and drives growth – and we need to

protect it appropriately from security risks and other potential service disruption.

In the UK, the government already imposes safeguards and enforcement regimes to ensure that our data is handled responsibly. But we will also take a greater responsibility in ensuring that data is sufficiently protected when in transit, or when stored in external data centres.

The government will determine the scale and nature of risks and the appropriate response, accounting for emerging trends.

We will tackle the cyber threats that arise from those seeking to harm the UK head on. We will shape a more secure technology environment, and improve cyber risk management in the economy to make the UK resilient to cyber threats. The increasingly international nature of data collection, storage and transfer can present data security risks. We will determine whether current arrangements for managing data security risks are sufficient to protect the UK from threats that counter our missions for data to be a force for good.

Data use creates other risks. Better use of data has the potential to help solve wider climate change problems and help the UK meet its net zero 2050 target, but we will also consider government's and businesses' responsibility for the environmental impact of increased data use. We will look to understand inefficiencies in stored and processed data, and other carbon-inefficient processes.

The infrastructure on which data relies is the virtualised or physical data infrastructure, systems and services that store, process and transfer data. This includes data centres (that provide the physical space to store data), peering and transit infrastructure (that enable the exchange of data), and cloud computing that provides virtualised computing resources (for example servers, software, databases, data analytics) that are accessed remotely.

Q14. What responsibilities and requirements should be placed on virtualised or physical data infrastructure service providers to provide data security, continuity and resilience of service supply?

Q14a. How do clients assess the robustness of security protocols when choosing data infrastructure services? How do they ensure that providers are keeping up with those protocols during their contract?

Q15. Demand for external data storage and processing services is growing. In order to maintain high standards of security and resilience for the infrastructure on which data use relies, what should be the respective roles of government, data service providers, their supply chain and their clients of such services?

Q16. What are the most important risk factors in managing the security and resilience of the infrastructure on which data relies? For example, the physical security of sites, the geographic location where data is stored, the diversity and actors in the market and supply chains, or other factors.

Q17. To what extent do you agree with the following statement: The government should play a greater role in ensuring that data use does not negatively contribute to carbon usage?

- Strongly disagree
- Somewhat disagree
- Neither agree nor disagree
- Somewhat agree
- Strongly agree

Mission five: Championing the international flow of data

In our hyper-connected world, the ability to exchange data securely across borders is essential. Economically, it drives global business, supply chains, trade and development; it will also be critical in enabling the global recovery after coronavirus. On a personal level, people rely on the flow of personal data to ensure their salaries are paid and to connect with loved ones from afar; this data in particular needs to be suitably protected. Finally, it has a huge impact on international cooperation between countries, including for law enforcement and national security, keeping the public safe.

Having left the European Union, the UK now has a unique opportunity – as a world leader in digital and as a champion of free trade and the rules-based international system – to be a force for good in the world, shaping global thinking and promoting the benefits that data can deliver while managing malign influences.

Using our international engagement and influence, we will:

1. **Build trust in the use of data**: We will create the regimes, approaches and tools to ensure personal data is appropriately safeguarded as it moves across borders. This will include looking to secure positive adequacy decisions from the EU to allow personal data to continue to flow freely from the EU/EEA to the UK, implementing an independent UK Government capability to conduct data adequacy assessments for transfers of personal data from the UK, and working with the Information Commissioner's Office (ICO) to build cooperation between national data authorities. The importance of data to our daily lives has made it a geostrategic tool. We will establish clear expectations of accountability when processing data – to protect personal data when it moves across the globe. These criteria will align with the UK's stance on promoting its wider values, ethics and national interests.

2. **Facilitate cross-border data flows**: We will work globally to remove unnecessary barriers to international data flows. We will agree ambitious data provisions in our trade negotiations and use our newly independent seat in the World Trade Organisation to influence trade rules for data for the better. We will remove obstacles to international data transfers which support growth and innovation, including by developing a new UK capability that delivers new and innovative mechanisms for international data transfers. We will also work with partners in the G20 to create interoperability between national data regimes to minimise friction when transferring data between different countries.

3 **Drive data standards and interoperability internationally**: We will cooperate with nations to develop shared standards that align with the UK's national interests and objectives. In a global arena, technical standards are increasingly expressions of ethical and societal values, as well as industry best practice. Recognising this, the UK will support global work on interoperability, which will facilitate the combination and cross-referencing of different data sources. This will include support to the collaborative on interoperability, an outcome of the UN World Data Forum in January 2017. The UK Government will also work with like-minded states to seek to ensure our values are considered and incorporated into the standards for new technologies which substantially impact data and their data trail.

4 **Drive UK values internationally**: The UK will be a champion of good-quality, available data across the globe. We want to ensure that UK values of openness, transparency and innovation are adopted worldwide. Now the UK has left the EU, we have an opportunity to set the UK apart and take an independent, individual approach that extols UK values. National competitiveness and the balance of power internationally are increasingly based on technology and the data that drives it. We want to ensure that UK values of openness, transparency and innovation, as well as the protection of security and ethical values, are adopted and observed globally. The UK will continue to play a leadership role to meet the urgent need for open, inclusive data, and its commitments under the International Aid Transparency Initiative. And we will continue to support the work of the Open Government Partnership to open up governments across the globe.

As the UK leaves the EU, we have the opportunity to develop a new UK capability that delivers new and innovative mechanisms for international data transfers.

Q18. How can the UK improve on current international transfer mechanisms, while ensuring that the personal data of UK citizens is appropriately safeguarded?

We will seek EU 'data adequacy' to maintain free flow of personal data from the EEA and we will pursue UK 'data adequacy' with global partners to promote the free flow of data to and from the UK and ensure it will be properly protected.

Q19. What are your views on future UK data adequacy arrangements (e.g. which countries are priorities) and how can the UK work with stakeholders to ensure the best possible outcome for the UK?

4. Data foundations: ensuring data is fit for purpose

In this section:

If the UK is to fully realise the benefits of our ongoing technological transformation, we must start by getting the basics right. This means being more effective in how we collect, curate, store, manage and delete data. Left alone, data does not sort itself out. If it is to become a powerful tool that can transform organisations and society, data requires effective governance, management and stewardship. It also requires modern infrastructure, allowing data to be shared across systems that can interact with one another.

> **Data Foundations**
>
> In this strategy we are using the term ‘data foundations’ to mean data that is fit for purpose, recorded in standardised formats on modern, future-proof systems and held in a condition that means it is findable, accessible, interoperable and reusable.

The case for change in public sector data use is clear. Presently, data is not consistently managed, used or shared in a way that facilitates informed decision-making or joint working across government and the wider public sector (https://www.nao.org.uk/wp-content/uploads/2019/06/Challenges-in-using-data-across-government.pdf). Data remains undervalued and underexploited (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/752003/Getting_smart_about_intellectual_property_and_other_intangibles_in_the_public_sector_-_Budget_2018.pdf).

Modernising the way we manage and share data across government will generate significant efficiency savings and improve services. To succeed, we must expand work to treat data as a strategic asset, and create a whole-government, collectively responsible approach to investing in data foundations, so that everyone can benefit from the improved outcomes data can offer.

The picture is more varied across the private and third sectors. While the UK is home to many world-leaders in data use, driving innovation and better services for consumers, this is not universally the case. Responses to our call for evidence and the wider existing evidence base suggest that across all sectors of the economy – perhaps particularly for SMEs and the third sector – issues include a lack of understanding about how data can be used, and used well; these issues are felt at most levels of organisations. [footnote 11]

The lack of basic coordination and interoperability both within and between organisations can drive inefficiency, a lack of accountability and an inability to thoroughly evaluate or plan. Data that is not usable, linkable or comparable between organisations means that, nationally, we lose out on the 'positive externalities of data' (https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_summary_report_26_Feb.pdf) – on the ability to pool data from multiple sources and sectors to create new economic opportunities, or to save lives.

Indeed, even those working on advanced technologies report that poor data foundations can be a real blocker for driving the transformative power of data. For example, when the source data needed to power AI or machine learning is not fit for purpose, it leads to poor or inaccurate results, and to delays in realising the benefits of innovation (https://www.gov.uk/government/publications/cdei-ai-barometer).

With better data, we can unlock new opportunities for businesses to grow and innovate. We can vastly improve and streamline public service delivery and offer consumers greater power and choice in the market.

4.1 Data foundations in the wider economy and society

Poor data quality and, relatedly, a lack of agreed standards are clear barriers to the effective use of data, from basic record-keeping to cutting-edge applications of data-driven technology. Responses to our call for evidence and stakeholder engagements highlighted common issues around:

- a lack of (central) ownership of data standards/ metadata/ APIs
- a lack of skills in managing data
- the pace of change leading to a fragmentation in the systems used to manage data, with ongoing resourcing issues linked to set up and maintenance costs

Some respondents suggested these costs could be especially burdensome for smaller organisations, or for organisations who make data as a bi-product of their operations rather than as a discrete business product. There was anecdotal evidence from our call for evidence that in general SMEs find it more difficult than large companies to invest in and maintain high quality data. There are pockets of stronger evidence for particular types of business. For example, the construction industry has well recognised Building Information Management (BIM) standards.[footnote 12] However, a range of academic studies find SMEs in construction generally do not use BIM. The issues identified by SMEs include:

- perception that BIM is only of benefit for larger construction projects
- high set-up costs of software
- licensing of software
- lack of in-house skills and/or cost of training.
- information retention across platforms (interoperability) – despite the industrial strategy supporting BIM
- lack of demand from clients (so no push to adopt the greater functionality) [footnote 13]

Pulling together the pockets of evidence with the wider anecdotal points, it is likely these issues do act beyond just the construction industry. The wider evidence base on the impacts of improving data 'foundations', and conversely the effect of not investing in them, is not especially strong. There appears to be little robust and independent research into the case for – and means of – implementing data quality, standards and management-improvement measures. This could be for a number of reasons:

- definitions vary, with a lack of consensus on how to measure the constituent elements of 'data foundations'
- most organisations would not want independent analysts studying their proprietary data, because of the risks to reputation or intellectual property

This is an apparent evidence gap that the government intends to address. Taking the limited research that does exist, there appears to be widespread concerns for 'data foundations' in the private and third sectors, with some consensus that small businesses, charities and SMEs are particularly affected. These problems are seen to result in loss of time and risk lower quality business decisions and operation.

Any way forward must be carefully considered, and one size will not fit all. Some barriers to better data use, even when substantiated by evidence, will not necessarily warrant government intervention. Poor data quality within organisations, for instance, is unlikely to warrant government intervention, unless it stops them carrying out a legal or statutory requirement. Even then, it is more likely to be a question for enforcers or regulators (for

example, if poor records lead to poor or negligent care, or if an organisation's poor data management procedures lead to a data breach that warrants ICO involvement).

What could warrant intervention is the need to drive better quality, more standardised and interoperable data to help drive economic growth or enable a public good outcome, especially where the value of the data sits beyond its immediate use. The government has taken decisive action to unlock the power of location data and data about the built environment – as exemplified by the Geospatial Commission and through our world-leading National Digital Twin programme (https://www.cdbb.cam.ac.uk/what-we-do/national-digital-twin-programme). The government has acted to further ADR-UK's work to transform the way researchers are able to access public sector administrative data by using the Research Powers of the Digital Economy Act (2017). There may be a case for extending this approach to other areas of the economy. In the first instance, we will take the actions outlined in the section below, and are seeking your views where stakeholder input and debate might push the agenda forward.

4.1.1 Consolidating a clear framework for government action in the wider economy

The government is committed to tackling market failures that mean the foundations of data use in the wider economy are missing or misaligned. Interoperable and consistent data can bring wide economic benefits. However, data of sufficient quality comes with a cost, and businesses can lack the information they need to make resourcing choices. Organisations report being tied into contracts on legacy systems that make collecting and maintaining data in interoperable formats harder. We know that a lack of coordination can act as a barrier to interoperability. And where businesses require data in certain formats for their business practices, such data may be undersupplied by the current market.

This is a new and complex issue that has come to the fore as data has become a more significant modern, economic asset. We are committed to addressing these issues, as highlighted in Mission 1 (Unlocking the value of data across the economy). This is a complex area, with many actors and the potential for unintended consequences. Before moving to consolidate this framework, we are keen to work in partnership with stakeholders as highlighted by our consultation questions and plans for future engagement. We are further working to build on the evidence base by commissioning research that deepens how we understand the government's role in driving data availability in the economy.

The steps we will take to consolidate the framework and approach, working with regulators and the Better Regulation Executive where appropriate, will include the coordination and alignment of existing work to build on data foundations across the economy, such as:

- working across the physical environment, including the standardisation of data about location, the built and natural environment, and transport and other infrastructures. This will include further developing the Information Management Framework (https://www.cdbb.cam.ac.uk/news/pathway-towards-IMF), which will seek to establish a common language by which digital twins can communicate securely and effectively, part of the Centre for Digital Built Britain's work towards developing a National Digital Twin.
- maximising the use of trusted data in innovation at the national and local level through supporting the new UK R&D roadmap (https://www.gov.uk/government/publications/uk-research-and-development-roadmap). This will include taking decisive action to drive the standardisation and interoperability of data for research, science and innovation, and improving the access to trusted data resources at local and regional levels, with further actions to be confirmed as joint work on R&D progresses.
- bolstering efforts to ensure that consumers' data is put to work for them

Case study: the Food Hygiene Ratings Scheme

The Food Hygiene Rating Scheme (FHRS) is an established, government-led open data service. Diners across the UK will recognise the scheme from the green stickers on restaurant windows. The scheme was set up in 2010, with a web service and an associated API created in 2012. The scheme serves in the region of 120 million API calls a year and covers over 400 local authorities in the UK.

The Food Standards Agency makes these ratings available as open data to platforms such as Just Eat, Uber Eats and Deliveroo. This data can then be combined with other data sources to support consumers and platforms in making more informed choices about where to eat and which businesses to feature.

Behind the FHRS is a complete and rigorous method of food safety and hygiene assurance based on best practice developed over many years. Due to the open availability of its data, the scheme has also been used for purposes beyond its original intention. For example, the FHRS was used by the Department for the Economy in Northern Ireland to help close the business rates gap (https://www.odcamp.uk/open-data-case-study-how-belfast-found-350000-in-rates-revenues-using-open-fhrs-data/).

The usual method of targeting missing rates is to conduct manual inspection on high value properties, but the Department for Economy used FHRS data, along with other datasets, to work out which properties were likely to be occupied. This targeted approach increased the success rate of inspections in a two-week period and enabled the pilot to identify £350,000 in business rates that were not being collected. In another example, FHRS data was correlated with other datasets allowing it to be used to help identify risks as diverse as the location of fatbergs in the sewers. The power of the FHRS as a tool for consumers lies in its simplicity, which is achieved without compromising its underlying fidelity.

4.2 Data foundations across government and the wider public sector

In our call for evidence, issues related to ‘data foundations’ were particularly highlighted across government and the public sector. Key issues included:

- data quality issues, and different standards for data used at all stages of the data lifecycle from collection to publicly available datasets, and the (in)consistent use of metadata – where it was provided at all
- issues with legacy systems and different, often incompatible systems for inputting and recording data at different stages of the data journey
- a lack of resources for local authorities to deal with data issues
- a lack of senior buy-in and leadership on data
- a lack of alignment across government

The case for change for government and the wider public sector's use of data is well established. The Chancellor of the Duchy of Lancaster stressed the importance of using data more effectively (https://www.gov.uk/government/speeches/the-privilege-of-public-service-given-as-the-ditchley-annual-lecture) to measure the impact of policies and make the best value-for-money decisions, for the greatest benefit, while increasing accountability. As highlighted by the National Audit Office (NAO) report – Challenges in using data across government (https://www.nao.org.uk/wp-content/uploads/2019/06/Challenges-in-using-data-across-government.pdf) – too often data is not seen as a key priority, the quality of data is not well understood and there is a culture of tolerating and working around poor-quality data. Inefficiency and cost can arise not just from poor-quality data, but also from a lack of coordination in data systems.

The NAO report referenced above also found that a lack of standards across government has led to inconsistencies in the way data is recorded between departments, including identifying numerous methods of capturing and storing data on individuals and businesses. This makes it extremely challenging for the government to get a holistic view of problems and limits the ability of departments to benefit from new technologies and tools.

These problems extend into the wider public sector. For example, effective electronic health recording systems play an important role in direct care, service delivery and research. While many of these systems are likely working to the same data standards and many neighbouring Trusts have interoperable systems, a recent study (https://spiral.imperial.ac.uk/bitstream/10044/1/75302/9/ImprovingDataSharingBetweenAcuteHospitals.pdf) indicated that, of the 117 NHS trusts using electronic health recording systems, 92 of them were using at least 21 different medical records systems, making it harder to coordinate and effectively share information. Ensuring these systems are fully functional and interoperable is vital if we want to continue to realise their benefits. The government took decisive steps at the March 2020 Budget, announcing the establishment of the Data Standards Authority, the Government Data Quality Hub and the development of an integrated platform for data across government. In July 2020, the Cabinet Office also assumed responsibility for government use of data to drive coordinated improvements in the use of data in policy making and service delivery. More remains to be done.

4.2.1 Data quality and technical barriers to data use and re-use across government

We will improve data quality across the public sector, ensuring that this data is not fragmented, siloed or duplicated across different organisations, and is deleted appropriately. Even the best-quality data cannot be maximised if it is placed on ageing, non-interoperable systems. The government is committed to removing the barriers to data interoperability presented by variations in the hardware and software used across government,[footnote 14] including by using processing techniques that make data 'independent' of the infrastructure that contains it.

We will:

- launch a programme of work to tackle the cultural and coordination barriers to good quality data, including:
  - creating a central team of experts able to ensure a consistent interpretation of the legal regime around data sharing
  - launching the Data Quality Framework
  - creating a Data Maturity Model for government
  - building a data management community of good practice
  - learning and setting best practice and guidance through a series of flagship demonstration – or 'lighthouse' – projects
- implement the recommendations of the 'Joined-up data in government: the future of data linking methods' (https://www.gov.uk/government/publications/joined-up-data-in-government-the-future-of-data-linking-methods) report to improve data linkage methods, application and skill sets across government
- commit to resolving the long-running problems of legacy IT and broader data infrastructure
- drive data discoverability across government through:
  - developing an Integrated Data Platform for government, which will be a safe, secure and trusted infrastructure for government's own data. It will be a digital collaborative environment that will support government in unlocking the potential of linked data, building up data standards, tools and approaches that enable policymakers to draw on the most up-to-date evidence and analysis to support policy development, improving public services and improving people's lives.
  - creating an audit of data inventories
- work to better support local government in maximising the benefits of data

4.2.2. Standards and assurance

To ensure that data is reusable and interoperable across government, [footnote 15] we have established a Data Standards Authority, with ongoing work to identify and agree a prioritised list of data standards to adopt across government. In the past, standards around data have been seen as voluntary. The result has been inconsistent adoption and a failure to realise

the benefits – we will tackle this through a prioritised approach to mandating certain standards and using spend controls to drive others.

> **We will:**
>
> - develop and validate a set of data principles to be applied across government
> - set out a strategy for standards, to include:
>   - clarity on where the Data Standards Authority will mandate some standards
>   - use of the DDaT spend controls process
>   - a parallel controls process for APIs and Technology Code of Practice to ensure consistent adoption of data standards across government

4.2.3. Productivity and accountability

To ensure that these changes are effective, we will tackle data governance across government, challenging risk aversion and data-hoarding, driving consistent levels of data maturity and ensuring a joined-up approach to establishing appropriate safeguards. A whole-government approach on data requires oversight and accountability from the top and centre of government, and through each department; we will ensure aligned accountability mechanisms, as well as a set of senior data leaders with the relevant expertise and backing and support from across government.

> **We will:**
>
> - recruit senior cross-government data leadership, including a Chief Data Officer for government
> - establish a cross-departmental governance mechanism with the authority to enforce standards across government
> - drive aligned governance structures across government by:
>   - undertaking a review of governance structures for data within departments
>   - ensuring central government departments include data management plans in their Single Departmental Plans

Case Study: Harnessing the power of administrative data to transform statistical systems for the public good

Timely population, migration, social and economic statistics reflecting regional variations better enable the government to proactively respond to trends, and to react more effectively to crises.

ONS is using administrative data to transform its statistical systems, efficiently providing more timely, flexible statistics and analysis to decision-makers. These transformed systems will, in turn, provide greater insight across key aspects of our society and the economy, ultimately leading to better outcomes for the public.

4.3 Supporting data foundations internationally

Beyond our borders, and as brought to the fore most recently, a lack of basic data maturity, standardisation and interoperability on the international stage can mean that it is difficult to thoroughly understand issues that affect us globally, such as the difficulty in comparing transmission or mortality figures in the early stages of the coronavirus pandemic.

In the global arena, technical standards are increasingly expressions of ethical and societal values, as well as industry best practice. As we set in place our domestic data strategy, we must engage our global partners to

adopt complementary measures so they can fully embrace and harness the innovations that data can bring. We must also work with like-minded states to ensure our values are considered and incorporated into the standards of new technologies which substantially impact data and their data trail.

We want to be a data champion across the world.

> **We will:**
>
> - support the global effort on data interoperability, which will facilitate the combination and cross-referencing of different data sources
> - collaborate with our international partners to build strong national statistical systems to drive economic growth and help to deliver inclusive, effective services

Case study: Ordnance Survey and Singapore's exploration of a 3D geospatial data model

In 2016-19, following prior joint research programmes, Singapore turned to the UK's Ordnance Survey (OS) for expertise in geospatial data standards and interoperability to help pave the way and enable enhancements to their 3D 'digital twin' of Singapore.

OS carried out two projects involving a variety of stakeholders and discovered data requirements. An efficient data-capture process for a variety of city 'themes' (e.g. buildings, vegetation, street furniture and transport) was identified and set up. A key part of the challenge was to

select relevant information 'locked away' in building information management (BIM) models, and to make it accessible through an open standards data model to wider stakeholders, such as city planners or regulatory bodies, who would benefit from the data to test new developments or city planning initiatives.

Based on these findings, joint collaboration between OS and the National University of Singapore led to the development of an 'IFC2CityGML transformation engine' – a software tool capable of automating the transfer of detailed building model information for different geospatial use cases. The project also helped to improve engagement, understanding and collaboration across the BIM and geographic information systems communities in Singapore, creating a new building theme with enhanced semantic representations for mobility, energy and urban planning. In the UK, OS is continuing to innovate with spatial-visualisation and data-integration techniques, combining them with other technologies such as Artificial Intelligence and machine learning. By applying common standards to practical solutions, OS is ensuring that the right information can get to the right people at the right time – and, crucially, in the right format.

5. Skills: Data skills for a data-driven economy and data-rich lives

In this section:

Data skills deliver benefits across the board. Businesses are more likely to be competitive in today's digital-driven economy if they can use data effectively. Likewise, data-literate individuals are more likely to benefit from and contribute to the increasingly data-rich environments they live and work in, while data-driven companies can deliver significant productivity benefits to their own business and the wider economy.

The need for data skills continues to grow across the economy. The Royal Society reports that demand for specialist data skills has more than tripled since 2013, while DCMS-commissioned analysis of 9.4 million online job adverts predicts that data analysis skills will be the fastest growing digital

skills cluster over the next five years. This characterises the exponential growth in the demand for advanced applications of data science and machine learning across all sectors of the economy, from cyber to construction. The growth in AI and cyber specialisms also drives the demand for broader supply of data skills at foundational level, to feed the pipeline of advanced skills and to provide business with the foundational skills they need to work with data. Notably, these scarce skills have been critical in the deployment of research capabilities to the coronavirus response. The portion of UK R&D that they support is significant and growing rapidly.

Data skills

There is no widely agreed definition of data skills. In this document we use the term broadly to cover the full range of basic, technical, governance and other skills – including project management, governance and problem solving – needed by practitioners to maximise the usefulness of data.

The required technical skills range from programming, data visualisation, analysis and database management, to core skills such as problem solving, project management and communication.

Planning for and delivering data at the right quality requires a wide range of skills that are sometimes underappreciated. Assurance of data requires people familiar with data laws and ethical oversight. Data processing and analysis skills, used to turn data into useful information, span a wide range of capabilities comprising both technical and soft skills.

Basic data literacy requires some knowledge of data uses, some ability to assess the quality of data and its application, and the skills to conduct basic analysis.

Consultation with data experts, the responses to the call for evidence and a review of existing research identifies a number of challenges involved in helping both individuals and companies develop the data skills they need. These include:

- **Lack of coordinated vision and leadership across multiple industry interests**. Many of the issues with data skills lie parallel to the issues with AI and Cyber skills needs. For example, almost half of UK businesses are facing a cyber skills gap (https://www.gov.uk/government/publications/cyber-security-skills-in-the-uk-labour-market-2020/cyber-security-skills-in-the-uk-labour-market-2020). A coherent approach across all skills stakeholders and