***This document is open for comments**. Please share your feedback: What have we done well? What's missing? Where do you disagree with us? Please respond before midday on Tuesday 8 December (GMT).*

UK National Data Strategy: Open Data Institute response

9th December 2020

About the ODI

The Open Data Institute (ODI) is an independent, non-partisan, not-for-profit organisation founded by Sir Nigel Shadbolt and Sir Tim Berners-Lee in 2012.

The ODI wants data to work for everyone: for people, organisations and communities to use data to make better decisions and be protected from any harmful impacts. We work with companies and governments to build an open, trustworthy data ecosystem. Our work includes:

- **pilots and practice:** working as a critical friend with organisations in the public, private and third sectors, building capacity, supporting innovation and providing advice
- **research and development:** identifying good practices, building the evidence base and creating tools, products and guidance to support change
- **policy and advocacy:** supporting policymakers to create an environment that supports an open, trustworthy data ecosystem

We believe that these six points, outlined in our manifesto, will help achieve our vision of a world where data works for everyone:

- Infrastructure: sectors and societies must invest in and protect the data infrastructure they rely on. Open data is the foundation of this emerging vital infrastructure.
- Capability: everyone must have the opportunity to understand how data can be and is being used. We need data literacy for all, data science skills, and experience using data to help solve problems.
- Innovation: data must inspire and fuel innovation. It can enable businesses, startups, governments, individuals and communities to create products and services, fuelling economic growth and productivity.
- Equity: everyone must benefit fairly from data. Access to data and information promotes fair competition and informed markets, and empowers people as consumers, creators and citizens.
- Ethics: people and organisations must use data ethically. The choices made about what data is collected and how it is used should not be unjust, discriminatory or deceptive.
- Engagement: everyone must be able to take part in making data work for us all. Organisations and communities should collaborate on how data is used and accessed to help solve their problems.

We have a mixed funding model and have received funding from multiple commercial organisations, philanthropic organisations, governments and intergovernmental organisations to carry out our work since 2012.

Consultation response

This is the ODI's response to the National Data Strategy Consultation.

Questions on the framing of the strategy

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.

Neither agree nor disagree – we are, in general, supportive of the NDS and welcome the government's efforts to establish a coherent strategy to 'unlock the power of data for the UK'. We are, in general, also supportive of the overall structure of the strategy into Opportunities, Pillars and Missions and recognise the challenge of setting out the framework for such a complex strategy for the whole of the UK. However, we do think that there is unnecessary overlap between some of these elements – such as Pillar 1 on 'data foundations' and Pillar 3 on 'data availability' and stress the importance of aligning the framework that is used within the strategy to the areas of responsibility of different departments or groups within government.

We also felt that the current Opportunities could have been strengthened by being more ambitious, inclusive, or rounded in some of their aspects. For example, we believe that Opportunity 1 ("Boosting productivity and trade") could have also included consideration of improving environmental sustainability and/or economic inequalities. Opportunity 2 ("Supporting new businesses and jobs") could have also considered the role of data in and data capability in supporting existing businesses and jobs, as well as a vision for social diversity and inclusion in these opportunities. Opportunity 3 ("Increasing the efficiency and scope of scientific research") could have also considered the breadth of research outside scientific disciplines and academia and industry - for example research in the arts, humanities, and social sciences, and research by or with civil society organisations or public sector organisations. Opportunity 4 ("Driving better delivery of policy and public services") could have included consideration of the role of new data sources, alongside existing strategic data assets. And Opportunity 5 ("Creating a fairer society for all") could have included a more proactive vision for the potential of data and digital technologies to counter or help ameliorate current societal biases, exclusions, and inequalities.

In addition, we feel the strategy should:

1. Set out a clearer, more coherent and ambitious vision for data in the UK, that builds on our unique strengths. The NDS recognises the critical role of data in the economy and the fact that the UK is already a leading digital nation. But there are many different ways in which the UK could lead around data, and it is not clear how the UK's approach to data will be distinct from that of, say, the EC, US, Canada or India. We believe that the UK should take advantage of its historic leadership in open data, data ethics, and data institutions to build a vision oriented around recognition of data as a public good.

2. Be more balanced in recognising both the opportunities and the risks around data. Currently, the NDS has a welcome tone of optimism about data but this can downplay the harms that it can bring. The NDS must be clear-sighted about the risks of data, both to ensure that they are effectively mitigated and to build trust in the approach that the government will take around data. Data ethics and responsible use of data cannot be seen as a separable add on: it must be embedded into every aspect of data policy.

3. Think carefully about how to empower the devolved administrations, local government, regulators, arms-length bodies and government departments outside DCMS and the Cabinet Office, to improve data across the country and across sectors. Making best use of data requires coordinated but distributed effort. The NDS needs to be clearer about which duties, responsibilities and powers around data are held by which public bodies, and how DCMS and Cabinet Office will support and coordinate these efforts.

4. Take a more holistic approach in considering the needs and opportunities around data skills. The NDS currently has more detail on relatively advanced data skills in central and local government, and also in research and the private sector. But comparable capabilities in civil society will help with the development of a balanced economy. And a strong foundation of core data literacy across all society will be a key component of engaged citizenship - supporting the development of trust between citizens, government, businesses and civil society through meaningful accountability and transparency about the management, use, and governance of data and digital technologies.

5. Pay more attention to the role of civil society within the data ecosystem. The NDS currently focuses on the role of government and of the private sector. But charities and other third sector organisations are also collectors, maintainers, users and reusers of data, and are important intermediaries in getting data and information into the hands of those who need it. They also play a key role in facilitating innovation with data - as well as being key potential innovators themselves. The NDS should recognise their crucial role, particularly in a difficult economic climate, and factor in the support they need to grow their data capability and participate in data ecosystems.

6. Give greater weight to the issues of equality diversity and inclusion, particularly with reference to the range of opportunities offered for data related jobs and skills development, and other economic opportunities around data use - for example, for start-ups and SMEs. We would also urge more consideration around the ways in

which data use and digital technologies could be used to proactively counter or ameliorate existing societal biases, discrimination or exclusion.

7. Overall, we would urge that the strategy covers the ODI's manifesto points of infrastructure, capability, innovation, equity, ethics and engagement in order to deliver a data ecosystem that works for everyone.

Finally, we recognise that this National Data Strategy is currently a framework strategy. We look forward to subsequent work that will map the intention of this strategy to actions, including a commitment to provide greater funding and resources, and to set SMART targets that would drive progress and measure success. We hope that this will provide additional clarity on ownership and accountability for delivering each part of the strategy and a structure that aligns with any divisions of responsibility.

Q2. We are interested in examples of how data was or should have been used to deliver public benefits during the coronavirus (COVID-19) pandemic, 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.

The Covid-19 pandemic has demonstrated the importance of data in a time of national crisis. It plays a vital role both to inform the policies taken to tackle the spread of the virus and as a communication tool to inform the public. Overall, data has been used to good effect during the outbreak, although there are areas where improvements could have been made, particularly with how data was shared by the public and private sectors and to ensure that local authorities have greater powers over decision making with data that will benefit local areas.

While this question asks for examples outside health and social care data, it is worth noting that this is an artificial boundary, especially in the context of a public health emergency that necessarily binds together health, social care, and the wider activities of societies and economies. There is a need to consider how health data informs non-health-system actions (for example business planning) and social or economic data informs health-system actions (such as shipping delays feeding into planning around PPE availability). The lesson for government is to focus on making the boundaries between data silos more porous, for example by adopting more open ways of working to enable low friction, ad-hoc collaboration and data sharing.

During times of crisis, decision making needs to be informed by the highest quality, most up-to-date data we have. This is made much easier if it is made available in a reusable form and as open data. This will allow others to build on the data, localise it, combine and compare it with other datasets. As an example, the ODI supported the Trade Union Congress (TUC) with their work supporting the publication of workplace risk assessments

during the pandemic[1]. This made use of Octopub[2], an experimental tool developed by the ODI to facilitate open data publishing. To date, the TUC have published over 240 records on their covidsecurecheck.uk[3] service. This has provided a valuable source of information for both employers and individuals and shown the benefits of standardising data collection and publication. It is likely that the government could have played a more central role in delivering such a service and, in doing so, given it greater impact and reach. This illustrates the potential role for government as a data steward for data from the private sector that could have a beneficial impact on the wider economy.

A relevant data project that emerged during Covid-19 is the Emergent Alliance[4], a not-for-profit community of corporates, individuals, NGOs and Governments, that aims to contribute expertise, data, and resources to inform decision making on regional and global economic challenges to aid societal recovery post Covid-19. This is a good instance of data from the private sector being used to illustrate and inform issues such as compliance with lockdown rules (mobility data) and economic recovery. For example, the sharing of mobility data – data about how motor vehicles, bicycles and people are moving across an area over time – has generated useful insights such as the transport needs or options for key workers. In general, government needs to create an environment that enables data held by the private sector to be shared and opened for the benefit of society. And it is essential that where data has demonstrated proven utility every effort should be made to retain access - at the very least for research purposes.

Decentralisation of data should also be considered. There is still a lot of useful local data which could help people make better decisions about, for example, when to travel and shop that doesn't appear to be easily accessible or published. Covid has shown where lots of useful data (such as mobility data, or local services data) should be being published openly to enable innovation and better decision making. Data institutions could support this by acting as focal points for collaboration across a community. The pandemic should also be the starting point for local authorities to invest in skills, training and open data more enthusiastically.

Q3. If applicable, please provide any comments about the potential impact of the proposals outlined in this consultation may have on individuals with a protected characteristic under the Equality Act 2010?

The ODI agrees with the NDS assessment that data can be a powerful tool to understand and tackle bias and exclusion. Digital technologies based on large, rich datasets, such as AI, can be a force for good, but can also reflect and entrench existing biases and discrimination. While we support the aims of the NDS to ensure that government, civil

[1] How greater transparency can help people get back to work with confidence – The ODI e
[2] Octopub – The ODI
[3] COVID-Secure Check
[4] Emergent Alliance

society and private sector actions treat people equitably, there needs to be more concrete proposals on how to address these issues and we have focused our response on these.

The Equality Act 2010 is clear about the need to protect people from discrimination. GDPR and the Data Protection Act 2108 also apply with respect to data about protected characteristics and the requirement to protect individuals' rights to privacy is essential. This can lead some organisations to be concerned about collecting data on protected and non-protected sensitive characteristics, fearing it is intrusive or does not comply with the legislation.

However, monitoring who uses digital services is a vital step in ensuring that they are fair and inclusive and there is, at present, no consistent process for the collection of protected characteristics by digital public services. This needs to be implemented before any impacts, good or bad, can be accurately assessed. In doing so, it is essential that data on protected characteristics should be anonymous, optional and stored separately from any other element of a service.

A programme dedicated to creating guidance and reusable templates for the collection of data to monitor equality impacts, from ICO, the Equalities and Human Rights Commission, and by the Government Digital Service would go some way towards correcting this misconception and ensuring data is collected and governed in trustworthy ways. Mandating publication of the results of ethnicity monitoring of digital public services through the Race Disparity Unit would increase accountability and transparency on race and ethnic disparities.

What is needed is for government to develop clear processes for the collection of data on protected characteristics on a case-by-case basis. In doing so, government should follow the recommendations in the ODI's report on *Monitoring Equality in Digital Public Services*[5]:

- Collect data to understand services users - we recommend that service designers start to embed, with respect and care, equalities monitoring into the provision of digital public services. Any collection of protected characteristics data must be designed such that: people can choose to provide information – or not; privacy is respected; standards and guidance are followed; and the results are transparent and follow good data practice.
- Collaborate to develop standards, guidance and training - we recommend that regulators, and other bodies supporting and monitoring the adoption of digital public services, collaborate to produce robust guidance, standards and training on how to collect, use and publish data about the people using those services. We further recommend that the Gov.uk Design System is developed to include styles, components and patterns to collect data about who uses the services and that this should be based on rigorous and collaborative user research.
- We recommend that an open standard for data for monitoring equality requirements be developed. We recommend that training for service

[5] Monitoring Equality in Digital Public Services (report) – The ODI

designers includes how to: design coherent, multi-channel services; design for opting-out; consider inclusive user research.

- Conduct further research - we recommend that further research could be undertaken to investigate: examples of monitoring equal access to services; what characteristics could be monitored; the impact of monitoring on users of a service; whether monitoring can be trusted; and how citizens feel about the collection of monitoring data?

The above recommendations highlight an important distinction between data specifically collected to shed light on the experience of people with protected characteristics and how data about protected characteristics should be embedded in the collection and use of other data. As well as monitoring the use of digital and public services, attention is needed on performing disaggregated analyses based on protected characteristics, to capture the different experiences of different communities.

The act of data collection about people with protected characteristics by those without can itself be seen as a form of disenfranchisement, disempowerment and exploitation, which leads to some communities becoming unwilling to take part in data collection, or distrustful of data-driven applications. Thinking arising from the indigenous data governance community[6] emphasises the importance of communities to be able to tell their own stories with data, and lead, or at least be involved, in determining how data is modelled and what questions are asked of it. The ONS has modelled an approach to this in its consultations[7] around questions about gender identity within the 2021 census.

Within all data collection and use, particularly across the public sector but also within the private and third sectors, questions need to be asked about what is actually being measured, which categories are most appropriate and whether the data should be collected at all[8]. Each protected characteristic is, in itself, a complex, multidimensional concept that intersects with other features. Data collection must take account of this complexity and be specific about what characteristics are actually being measured and why. Moreover, sensitivities about protected characteristics could make certain groups feel profiled and targeted so those collecting such data need to think carefully about whether it is appropriate to do so.

It may be the case that the protected characteristics set out in the Equalities Act 2010 are not the only important characteristics to understand inequities. Others such as socio-economic or regional characteristics may also be relevant. Also, the intersection of multiple characteristics may be of significance to identify and understand more nuanced cases of discrimination.

It is not possible for a national data strategy to predict all potential forms of impact on people with protected characteristics. Therefore what is important is to enable and support those who are collecting, using and sharing data to properly examine such considerations as part of their data and service design processes. We recommend that organisations are encouraged to carry out an equalities impact assessment, which may be part of a wider

[6] The CARE Principles for Indigenous Data Governance – Data Science Journal
[7] Gender identity update - Office for National Statistics (ons.gov.uk)
[8] The dividing line: how we represent race in data – The ODI

data ethics assessment, and publish it openly, to help motivate and enable others to do the same and to build trust that such issues are being considered during the development of data-based analyses or services.

We also urge that equality, diversity and inclusion considerations be brought to bear on all aspects of the NDS including skills, Smart Data and the protection of vulnerable consumers, and investment. This should go beyond the narrow remit of the 2010 Equality Act and consider both the specific considerations around intersectionality of different protected characteristics, and also individuals or communities who might not be protected by equality legislation but are statistically more vulnerable to forms of prejudiced exclusion and structural disadvantage. Ultimately, the goal should be an equitable economy that has fair opportunities for all.

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

There is substantial diversity across the different regions and places within the UK. Significant inequalities have been identified across these different regions in areas such as wealth, productivity and physical infrastructure. Various government policies, such as the Industrial Strategy, have been set out to tackle regional inequalities, but this is a challenging issue, not least in terms of coordinating funding and policies between the various devolved, local and metropolitan authorities. The NDS clearly recognises the potential benefits of the use of data to local communities and states that government 'will work to better support local government in maximising the benefits of data'. This is vital but much more is needed to specify how this will be realised in practice, the potential harms from getting it wrong (or not doing anything), and the support mechanisms that will be needed.

In terms of benefits, cities and regions offer good opportunities for effective data sharing projects. They can often be set up to deliver specific local solutions that have more immediate outcomes for citizens. Such projects can demonstrate the benefits of data sharing more effectively than more complex national projects and help to build public trust. Projects may be run wholly by local authorities delivering public services. But, equally, the involvement of the private sector or local data collectives and institutions should be encouraged and supported. This could build on the so-called 'Triple Helix Approach' – collaboration between the public sector, private sector and academia – or collaborations between the public sector and academia such as the Urban Observatories[9].

Open innovation offers significant potential to drive economic growth at the regional level. As noted in the consultation, the Data Pich[10] programme was shown to enable startups from outside traditional innovation hubs in capital cities to participate in data innovation. Access to data outside that was supported by the programme improved startups' ability to attract funding.

[9] Urban Observatory
[10] Data Pitch

As with central government use of data, all the basic pillars set out in the NDS will need to be addressed at the local level. Perhaps most important will be to invest in local human capacity. Local authorities could lead on innovation and better public services using data but opportunities are being missed because the workforce lack data skills and data leadership. Here, the focus should be on foundational data skills across the workforce, and strategic data skills for leaders. Building these will provide the core understanding that can be expanded and pinpoint the practical skills that might be needed in the future. Many training courses and tools exist, such as those offered by the ODI, that could help build the skills needed at all levels of local government.

It is essential that local authorities understand their data before investing in new technology solutions and that they are able to be intelligent customers. Local government budgets are tight. Understanding the data they already have, or have access to, and investing in existing systems to improve its quality, accessibility and how it is put to use will create more value than investing funds in new tech solutions.

Combined authorities are playing an increasing role in terms of setting regional priorities. Greater Manchester Regional Authority, for example, is working to bring all its Local Authorities up to a similar level of capability on data and capturing impact from its use. This highlights the inequality in data capability among local authorities. Setting clear responsibilities and powers around data for combined authorities would help empower them to take on this coordination and capability building role.

Peer networks, such as Open Data Camp, LocalGovCamp and Local Digital Slack channels have proved to be useful tools for people working on data and digital services in local government. More investment in, and awareness of, peer networking opportunities could lead to more openness and collaboration in the local data communities.

Beyond investing in data capability within local authorities, there are a number of actions that should be taken within central government to provide the support they need.

1. Local governments, and other bodies such as Local Enterprise Partnerships (LEPs), need data that is collected centrally – including economic, demographic and health and social care data – to be disaggregated. National averages are seldom useful in understanding the unique needs encountered at a local level. Disaggregation to regional or local authority level is necessary to inform investments in different parts of the country. Disaggregation to postcode or MSOA level enables local decision makers to understand their context and the requirements of their citizens better.

2. Local governments need devolved powers that enable them to create the data infrastructure they need to carry out their duties efficiently and effectively. In particular, they need legal powers to access non-personal data from larger private sector organisations that operate in their areas, including from utilities, transport providers, and digital platforms such as AirBnB. Currently, local authorities and city regions have to engage in a time-consuming negotiation with companies that hold data they need; while some are amenable, others delay, provide out-of-date or low quality data, or restrict the potential uses of the data through licensing agreements. Legal powers would have to be designed carefully to avoid undue burden on business and to ensure there were proper ethical checks and citizen consultation in

place, but they would help abbreviate the negotiation process. The ODI's Data Ethics Canvas[11] would provide a useful guide here.

3. Local governments need targeted investment to improve their own data infrastructure, in particular – in common with other public bodies – to modernise and improve the quality of the data they hold. The restricted funds available to local authorities over the last decade have meant that the maintenance of digital and data systems have been neglected. The ODI's Data and Service Business Case Canvas[12], which supports people designing data-enabled public services to make a business case for a project or service that uses data, could help in this regard.

4. Central government has a role to play, either commissioning or building tools that can be used locally, or in enabling local authorities to learn from each other and to share and scale systems and solutions. Our work on scaling local data innovation[13] includes some case studies and recommendations for supporting this.

[11] The Data Ethics Canvas – The ODI
[12] Data and Public Services Business Case Canvas – The ODI
[13] R&D: Scaling data innovation – The ODI

Mission one: Unlocking the value of data across the economy

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) codes.

In today's economy, every business or organisation is a data business or organisation and, as such, every sector has a great deal to gain from better data availability. Identifying those sectors with the most to gain is a complex problem, not least given the challenge of quantifying the value of data[14]. However, preliminary work has been carried out by the European Data Portal in their report on the *Economic Value of Open Data*[15] which identifies: public administration; professional, scientific and technical activities; ICT; and transportation and storage as having proven and successful open data impact and: agriculture; financial services and insurance; health; education; wholesale retail and trade; and real estate as having high potential.

The ODI has worked with several sectors to assess the economic value of open data including the construction sector and the potential of digital twins[16], the energy sector through the Energy Data Taskforce[17] and the Modernising Energy Data Access programme, the transport sector with Deloitte and the (then) Transport Systems Catapult[18], and the engineering sector with Lloyd's Register[19].

Our experience is that drawing boundaries around data from different sectors is unhelpful: much of the potential value of data comes from sharing data across sectors and that a narrow focus on individual sectors is likely to miss opportunities for innovation, wealth creation and social value. This is equally true for sharing data between the public and private sectors. We find that it is more helpful to take a challenge-focused lens to prioritising activities around data, aligned with government priorities. For example, given the government's commitments towards the Green Industrial Revolution[20] and Sustainable Farming[21], we would recommend prioritising data that supports and enables those investments.

[14] Research projects at the Bennett Institute for Public Policy
[15] Open Data Impact
[16] Digital twins: needs, challenges and understanding – The ODI
[17] A strategy for a Modern Digitalised Energy System
[18] OpenData -Transport Systems Catapult
[19] Top engineering organisations sign data manifesto to improve safety – Lloyds Register Foundation
[20] PM outlines his Ten Point Plan for a Green Industrial Revolution for 250,000 jobs – GOV.UK
[21] Path to Sustainable Farming - GOV.UK

At the same time, there are a number of fundamental datasets that can add value in multiple sectors, such as geospatial data and demographics. All sectors will benefit if data about places and people are made available in sustainable and well-governed ways.

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

The more data is available, the more opportunities there will be for companies and social enterprises to make useful products and services using that data, creating jobs and driving economic growth in the digital economy. The government has a critical role to play in enabling better availability of data. As both a user and provider of data, it needs to be thinking about how data flows between all kinds of sectors, organisations, people and communities, and in all directions.

While doing so, government also has to recognise the human, environmental and market effects of the data economy. Excessive surveillance has knock-on effects on people's lives and wellbeing. Increasing access to sensitive data – whether about people, critical national infrastructure or our environment – creates a greater surface area for attacks, increasing the risk of data breaches and the knock-on harms they cause. Organisations that exclusively hold and get good at supplying certain types of data can become powerful, market-distorting, platforms for that data.

The data economy has particular characteristics that mean the market won't fix many of these problems[22]. In economic terms, there are various barriers that can't be addressed by the market. These include: the fact that those that collect data are unaware of who might be interested in using it, and those who might benefit are unaware of who is collecting it; asymmetries between organisations regarding what is known about datasets, such as their quality or how they could be used; misalignments that mean the incentives of those that collect data, those that could use data, and those who are affected by it, are mismatched; and contractual uncertainties about how other people will behave because of ambiguity in norms, regulations, licensing, or terms and conditions, or because there are few ways of assessing the trustworthiness of other parties.

Government must invest in improving access to data so that more people and organisations can use it, and to tackle the market-distorting effects of data monopolisation. This will mean making data as open as possible to enable us to innovate, create more efficient and effective services and products, and fuel economic growth and productivity. However, in doing so, people's privacy, commercial confidentiality and national security must be protected. Three forms of access will give particular benefits:

- Open data, which anyone can access, use and share without restriction, where the risks associated with making data available in this way are low.
- Secure data access, where access to data (usually bulk or large scale data) could be useful for analysis, research and innovation, but the data is sensitive and so cannot be made open to everyone.

[22] Research projects at the Bennett Institute for Public Policy

- Data portability, to enable people and organisations to access data that is about them and pass it to third party services to provide them with insights.

Enabling access to data is not a cost free activity, so, whether the data is held by government or by organisations outside government, there needs to be both an incentive and a recognition of those costs to do so. Data held by the government, whose collection and maintenance is paid for by the public, should be as open as possible as a matter of course. Companies and other non-public sector organisations can be incentivised through a range of policy measures, from soft activities such as agenda setting through encouragement such as grant-giving all the way through to regulation. Government should not underestimate the power it can wield over access to data through its own procurement practices, or by attaching conditions to the provision of licences to operate (for example as the Oil and Gas Authority does on Offshore, Onshore, Gas, and Carbon Capture and Storage licences). Tax credits provided to organisations based on their creation of datasets should also be structured to incentivise organisations to make that data as open as possible.

Central government should ensure that the construction and strengthening of data infrastructure is built into existing programmes of work (which should already reflect government priorities), for example by:

- Providing funds for data infrastructure projects across the economy focused on providing additional investment to improve data infrastructure and availability within existing initiatives led by regulators, industry bodies, consortia and so on.
- Encouraging policymakers across the public sector to consider access to data as they design policy interventions. This should include a range of things aimed at helping those developing public policy interventions to understand how changing the availability of data can help contribute to (as well as monitor) the outcomes they desire. This could include training and changes to processes within the context of policy making (eg, through the Policy Lab), digital service delivery (eg, through the Government Digital Service Manual), public procurement, and activities like creating business cases for public investment. This should target policy makers within both central and local government, arms-length bodies and regulators.
- Give policymakers the powers they need to enact these interventions. Some activities simply require investment or other soft interventions; others require stronger levers over organisations, to require them to share data or adopt standards. Regulators and organisations like the Office for National Statistics already have (in the UK) some legal powers they can use to increase access to data from others – including companies. But many public bodies, particularly at a local level, have to make do with the carrot of persuasion as they lack any sticks of enforcement. Central government should legislate to give powers to public bodies, in particular regulators and local government, to enforce participation of organisations in data access initiatives (with appropriate checks and balances to ensure the access is well-governed and it doesn't become an excessive burden).

An emerging trend to facilitate the availability of data is the development of data institutions[23] – organisations whose purpose involves stewarding data on behalf of others, often towards

[23] Data institutions – The ODI

public, educational or charitable aims and often act as intermediaries between data holders and potential data users. These can take a variety of forms including data co-ops, data unions, data coalitions, and bottom-up data trusts. Data institutions play a number of vital roles, including:

- Holding data on behalf of an organisation or person, or group of them, and sharing it with others who want to use it for a particular purpose.
- Combining or linking data from different sources, and providing insights and other services back to those that have contributed data.
- Creating open datasets that anyone can access, use and share to further a particular mission or cause.
- Developing and maintaining common data infrastructure for a sector or field, such as by registering identifiers or publishing open standards.

Such data institutions support people and communities to take a more active role in stewarding data about themselves and thus provide a way to improve data governance and increase trust. Government needs to be aware of their utility, create an enabling environment to support them (for example by funding research that helps them tackle the common challenges they face), and in some cases fund their scoping, co-creation and/or operation.

Q6a. How should this role vary across sectors and applications?

The interventions that are necessary in different data ecosystems (in different sectors or to meet different challenges) are dependent on that ecosystem and its maturity. There are a range of different kinds of interventions that might be appropriate, including:

- Investing in creating and maintaining new datasets where they don't exist, such as those highlighted by Missing Numbers[24], to provide information that is needed to tackle a given issue. In many cases where data appears to be missing, there are likely to be existing datasets that partially meet the requirements; where possible, government should build on these existing efforts (including extending existing datasets, improving their quality, or committing to their maintenance, adopting existing standards, and using existing identifiers) rather than create new ones. Any investment in new datasets needs to consider potentially negative impacts of the data collection itself and should come with a sustainability plan – including a designated body to steward and maintain the data – that ensures it is not a one-off effort.

- Rationalising data stewardship where there are multiple bodies that are collecting and maintaining the same data. There are many places where several organisations maintain the same or very similar data, and where it may be possible to reduce overall costs to all those organisations, and reduce confusion about which is authoritative, by consolidating that activity into one organisation (a data institution[25]).

[24] Missing Numbers – the Centre for Public Data
[25] What do we mean by data institutions? – The ODI

There are equally cases where having multiple sources of data that provide different views and nuance on the same subject, and particularly counter-data[26] that provides a countervailing force against prevailing assumptions and power structures, is a good thing. Government has a role to play in bringing organisations together and identifying places where collaborative maintenance of datasets[27] could bring efficiencies and increase data quality, and supporting stewardship by the communities the data is about.

- Standardising and improving the quality of existing data, and the quality of its publication, so that it better meets wider needs. Not all data needs to be high quality and not all data needs to be standardised, but well structured data is easier to work with; higher quality supports more uses and more confidence; higher standardisation supports interoperability, aggregations and comparisons; disaggregation enables richer pictures of diverse experiences; APIs make it easier to get hold of data on demand. Government can invest in filling the gaps within datasets, creating open standards for data[28], driving their adoption, equipping people and organisations with the skills and technologies they need to improve the quality of the data they have, and incentivising those stewarding data to understand and consider the needs of others who could make use of that data as they structure it and publish it.

- Investing in improving access to data so that more (or fewer) people and organisations can use it, and to tackle the market-distorting effects of data monopolisation. This means moving data to being as open as possible, understanding both the benefits and risks of different types of access.

- Increasing capability to use existing data by making it easier to find, understand and use data, and training and supporting those who do so. This does not mean creating data portals and registers: we need more curated and service-oriented approaches to helping people and organisations locate and understand the data they need. These services, such as the support provided by the Covid-19 Mobility Data Network[29], are worth investment and should be designed to provide most support to those who have the least capability themselves. Government should lead the way in making the data it shares more findable and usable, through better documentation, machine-readable metadata that can be read by tools like Google Dataset Search[30], and potentially through creating a shared data agency with the responsibility of finding data for people across the public sector who need it, or finding public sector data for organisations in the wider economy.

- Encouraging the use of data – and identifying gaps – through open innovation. It is hard to understand the limits of what can and can't be done with the current data infrastructure until you test those limits. Challenge prizes are not only useful in terms

[26] Grassroots resource mobilization through counter-data action - Amanda Meng, Carl DiSalvo, 2018 (sagepub.com)
[27] R&D: Collaborative data maintenance – The ODI
[28] Welcome to the Open Standards for Data Handbook – The ODI
[29] COVID-19 Mobility Data Network
[30] Dataset Search – google.com

of the products and services that get built through them, but when they redress the advantages of existing market players, and for the insight that they provide about what innovators and technologists find they can't do. Innovation is often thought of as something that happens after data infrastructure is in place, but it can be just as important early on as an evidence-gathering activity for understanding gaps in data infrastructure and what should be built to fill them.

The type of interventions that are useful depend on the maturity and gaps in a given existing data ecosystem. Recent research at ODI has focused on how to identify the right kinds of interventions in a given context.[31]

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.

The ODI **strongly agrees** that the government should play a role in supporting data foundations in the wider economy.

In the first instance, government needs to be clearer in defining what is meant by 'data foundations' and, in particular, how this definition relates to data infrastructure. The definition of data foundations, as set out in the NDS, focuses on making data fit for purpose. As such, it incorporates some of the characteristics of strong data infrastructure, such as the use of standards, the infrastructure that's needed to support data, the availability of compute and discovery mechanisms, and aspects of data availability. As such, it shows considerable overlap with other Pillars, particularly 'availability' in terms of the data being accessible and reusable. Skills and responsibility also need to be woven into data foundations by recognising that people and processes are an integral part of data infrastructure.

As noted in the consultation, definitions can be difficult but one role for the NDS will be to take the lead in providing clear and consistent language relating to data. It is also inevitable that in designing the framework of a data strategy there will be overlaps and interconnections between the constituent components. But the government has the opportunity now to ensure that the basic framework is as clear and streamlined as possible. This will promote a commonality of language and understanding across all stakeholders.

We would recommend the adoption and consistent use of the term 'data infrastructure'[32] as defined in the glossary, which can be broken down into:

1. Data assets, such as datasets, identifiers, and registers.
2. Standards and technologies used to curate and provide access to data assets.
3. Guidance and policies that inform the use and management of data assets and the data infrastructure itself.

[31] R&D: Data infrastructure for common challenges – The ODI
[32] 18. Data Infrastructure – The State of Open Data

4. Organisations that govern the data infrastructure.
5. The communities involved in contributing to or maintaining it, and those who are impacted by decisions that are made using it.

The strategy's definition of data foundations also includes the principles findable, accessible, interoperable and reusable - often cited as the acronym FAIR. In relation to making data fit for purpose these are a very reasonable set of principles. However, it should be recognised that they are just principles and do not necessarily provide guidance on how to answer specific questions about the data relating to issues such as metadata, standards or licensing. It should also be noted that the FAIR acronym does not incorporate any aspects of fairness which can lead to some confusion.

Part of government's role in supporting strong data infrastructure in the wider economy will be to provide practical guidance and tools that demonstrate best practice and support the development of the data ecosystem. In particular, government must lead the way by providing a strong data infrastructure for the wider economy with the publicly funded data that it stewards itself, including foundational datasets such as geospatial data, data about people, and data about organisations.

Government also has a role in setting and supporting the development and adoption of open data standards to strengthen the data infrastructure in the wider economy. Significant amounts of work have already been done to develop and promote the use of open standards for data, including by the Government Digital Service's Open Standards Board[33] and the ODI[34],[35].

Standards, in general, help to consistently publish and use data. They are a tool that can help to shape markets, create ecosystems and implement policy by making it easier for people and organisations to publish, access, share and use better quality data.

In economic terms[36], standards transform published data sets from opaque, bespoke items to commoditised products that are more reliable, safe, high quality, reproducible, interoperable, scalable and have economic utility. Data users would no longer need custom code and processes to use new datasets. This reduces the cost of both producing and using data, meaning an organisation can focus on providing value at any stage along the pipeline. The adoption of open standards for market-critical data can help to disaggregate authority: stakeholders (including market leaders and authorities) stop using bespoke and proprietary formats and instead use cooperatively produced and shared standards. This levels the playing field for data production and data use, allowing new uses of data and new entries to the market. An additional benefit can be certification of compliance from standards bodies, which can act as a badge of trust and build confidence.

However, ODI research shows that standards often fail to get adopted because of poor scoping and consultation resulting in them not being fit for purpose or the wrong things

[33] Open Standards Board - GOV.UK
[34] What are open standards for data? – The ODI
[35] Announcing the open standards for data guidebook – The ODI
[36] Economic impacts of open standards – The ODI

being standardised. Similarly, lack of support around adoption also leads to failed uptake. Government should provide support to convene stakeholders and enable and encourage the development and adoption of new standards rather than simply selecting preferred standards. There are particular gaps around standardising data available from cities and local government and in geospatial data where there is little alignment across data published through INSPIRE[37]. Much of this work should be done through regulators, to create sector-relevant standards; central government has a role in ensuring there is support for regulators who are taking these steps, such as that provided by the ODI.

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?

SMEs make up the majority of businesses in the UK and cover a very diverse range of companies in terms of size, sector and location. They often face greater challenges than large companies because of financial and resource constraints which, as noted in the consultation, can affect their ability to manage their data. However, this may not always be the case, particularly with more digitally native SMEs that work in the data rich sectors, and government should be aware that the diversity of SMEs means that support mechanisms should not be one-size-fits-all but should be tailored to the needs of particular SMEs. It is, however, important to remember that all businesses can benefit from data, and SMEs that are left behind are likely to find it hard to flourish.

All the basic pillars set out in the NDS to support good data infrastructure will be important for SMEs. Data foundations and availability will be essential to ensure that SMEs produce, can find and have access to high quality data to support existing operations and enable innovation on new products and services. This will require funding to develop and maintain strong data infrastructures that make data sharing, discovery and access easier. Data standards will be particularly important and government should fund the creation of data standards in industries that would benefit from higher digitalisation and competition, ensuring that SMEs have a voice at the table during their development.

Government should consider ways to directly incentivise large companies to make more, higher quality and diverse data openly available, so that SMEs can benefit from it. Possible options are to make it a requirement in all government procurement contracts, establishing it as a requirement for membership of trade bodies or providing tax relief for data that is made available. Better data portability can be enabled through mandating or encouraging open APIs for regulated industries such as has been done in the banking sector, as noted in the consultation in relation to Open Banking.

Government could provide funding to SMEs that are solving crucial societal challenges via organisations running support programmes for startups and SMEs, such as incubators, accelerators and innovation agencies. The Seed Enterprise Investment Scheme[38] and

[37] INSPIRE Geoportal – European Commission
[38] Use the Seed Enterprise Investment Scheme to raise money for your company – GOV.UK

Enterprise Investment Scheme are good examples of government support schemes, as is InnovateUK grant funding.

Skills can be a particular challenge for SMEs and we welcome the recognition given to this issue in the NDS. Addressing this will require access to suitably qualified and experienced staff (often in high demand) as well as upskilling and training of existing staff. Measures already announced or expected through the Office for Talent and the National Skills Fund are a start. But increased and continuing financial support, such as a voucher scheme, will be needed to provide access to education and training services for SME employees - particularly in sectors that are less digitally developed and need more digital innovation to meet the needs of consumers and people.

Capability in data protection and compliance with the law also require particular consideration with respect to SMEs. Failure to establish a coherent data protection framework will become a profound barrier for companies and businesses of all sizes to function. Similarly, a complex regulatory regime will have a disproportionate impact on SMEs and we welcome the proposals in the NDS to work with the ICO to address this issue.

Government should consider stimulating the creation of open source tools for common data tasks, such as data management, storage, publication, transformation, analysis, visualisation and so on, alongside non-software tools such as checklists, templates, canvases and so on. The availability of open source tools would lower the costs for SMEs to work with data.

Government should also extend the data portability right, which is currently focused on natural persons, to organisations, so that businesses are equally able to switch services and take advantage of insights and innovations made available by third parties. Open Banking was originally conceived as helping SMEs; Smart Data approaches should apply equally to businesses as to individuals.

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?

The ODI supports the work of the Smart Data Review to consider how to accelerate the development and use of new data-driven technologies and services to improve consumer outcomes. There is, however, much work to be done in this emerging field, as detailed in our work on open communications[39] and response to Ofcom's recent consultation on Open Communications[40].

The advantages of Smart Data are clear from the example of Open Banking as highlighted in the NDS and we welcome the proposals in the NDS to extend this to more sectors and to introduce legislation to mandate participation by market incumbents. Designed well,

[39] Open communications: an open trustworthy data ecosystem for the telecommunications sector [report] – The ODI
[40] Ofcom's call for input on Open Communications: the ODI's response

increasing the portability of consumers' data could enhance competitiveness, particularly with respect to SMEs competing against larger market incumbents. We recommend carefully monitoring the impact of Smart Data in these markets, recognising that they may also embed existing market inequalities if poorly implemented.

The potential of Smart Data should not be limited to simply enabling individual consumers to switch between service providers. It also opens up the opportunity to create new products and services that provide additional insights. We recommend prioritising the development of data portability standards towards facilitating the creation of these new services beyond switching.

Benefits may also be felt by the individual. Consumer data may be brought together to create public goods that will, in turn, benefit larger numbers of citizens. This is envisioned as the creation of 'data altruism organisations' within the European Commission's Data Governance Act[41] and in some conceptions of 'data trusts'. We recommend examining the proposals arising from the European Union and places like India[42] for ways to monitor and regulate these types of data institutions.

Issues of privacy and security will need to be taken extremely seriously when handling consumer data. Any breaches could have harmful immediate consequences and undermine trust more generally. Vulnerabilities will need to be well understood as these will change depending on the circumstances. Mitigating measures may need to be put in place such as anonymisation or synthetic data to protect consumer rights. It should also be noted that consumer data may be used to support non-digital services and that the benefits of all new products and services should be made available to all consumers - not just those that are digitally proficient. This may involve support for carers and intermediaries that provide support in the community. This raises the issue of data rights and ownership, and collective interest in data, which is an ongoing area of debate that the government should be aware of, participate in, and purposefully experiment around[43].

[41] Proposal for a Regulation on European data governance (Data Governance Act) | Shaping Europe's digital future – European Commission
[42] Practising data stewardship in India, early questions – Ada Lovelace Institute
[43] Data 2020: Data rights and ownership – The ODI

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

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

GDPR and the Data Protection Act 2018 have now had two years to become embedded in the economy. In that time businesses have had time to acclimatise to the changes they have brought. However, it was costly to implement GDPR so there is unlikely to be any desire to introduce alternative data protection approaches. Furthermore, as evidenced in the ODI, RSA and Luminate report *About Data About Us*[44], the public are familiar with many of the rights over personal data it provides and feel some sense of confidence about it.

In general, the key asks from the public from this work and around future of data protection related to:

- Honesty and transparency about what data about them is being used, how, when, and for what purpose by all organisations both public and private, and to be asked to opt in rather than opt out.
- Agency and control over their data with clear requests for permission to share data, choice of personalised services or adverts and clear signposting of data decisions.
- Rights and responsibilities relating to both individuals and organisations regarding the handling of data with appropriate legislation and independent oversight.
- Context and fairness to prevent prejudices and biases being replicated and clarity on any value-exchange relating to the sharing of data.
- Compliance and enforceability of data protection legislation.

In terms of the UK's data protection framework, we should build on GDPR rather than undermine it. In particular, rather than focus on improving individual data protection, we believe the government should develop and use a broader definition of responsible and trustworthy data governance based on the principles of ethics, engagement and equity, and democratic oversight that recognises the relational nature of data[45]. Additionally, the data portability right should be seen more broadly as a right of account holders including organisations (legal persons) as well as by people (natural persons).

There are five areas where we believe the government could act to increase the trustworthiness of the data ecosystem:

- Carry out a programme of work to improve the trustworthy collection of data about protected characteristics, including ethnicity, to monitor inequalities in public and private services.

[44] About Data About Us – The RSA
[45] Democratic Data: A Relational Theory For Data Governance by Salome Viljoen – SSRN

- Increase the capacity of regulators and professional bodies to support and enforce good data governance in their industries.
- Seek explicit alignment on a high standard of data protection in trade deals.
- Explore policies and regulation to extend data rights beyond individual rights over personal data to group or collective rights.
- Invest in continuing research and experimentation with new practices, technologies (including privacy enhancing technologies) and institutional forms to increase trustworthy access to data.

If changes are deemed to be necessary post-Brexit, the UK should create its own framework which uses GDPR as the foundation. This should include group rights, rights for business, more detail and scoping on data portability, consideration of how consent can be more dynamic, a guarantee that rights will be protected, recognition that data about us is not something that we 'own' but rather something we have rights over, and consideration as to how the data economy addresses the issue of online advertising and the consequent tracking and profiling of citizens.

Businesses do not want to have to spend more money changing their data protection approach having already incurred high costs implementing GDPR. This is especially the case for SMEs who have less capacity to cope with any challenges with the implementation of data protection and would need additional support. Divergence of the UK data protection framework from the EU's system in ways that undermine citizen's data rights is unlikely to be beneficial to the UK.

A data protection framework that is aligned with trading partners is important to ensure the free movement of data across national borders, which would help bolster our domestic digital services sector by enabling others to use their services. We agree with the emphasis in the NDS on achieving data adequacy with the EU. We need a data protection regime that provides trading partners with confidence that data about their citizens will be handled in trustworthy ways if it comes into the UK, and that ensures data about UK citizens is handled in trustworthy ways when it goes overseas. Data adequacy with global partners should not mean lowering our data protection standards as part of potential trade deals. Furthermore, agreeing to different data protection regimes as a consequence of future trade deals risks creating a confused patchwork of rules and regulations for UK business to adhere to and will undermine citizens' rights.

Overall, the UK Government should continue to pursue the vision of the UK as (in the Prime Minister's own words) a "global leader in ethical and responsible technology"[46]. This means more than retaining the rules laid out in GDPR; it means building and innovating on top of those rules to demonstrate leadership in data governance.

[46] PM speech to the UN General Assembly: 24 September 2019 – GOV.UK

Q11. To what extent do you agree with the functions set out for the Centre for Data Ethics and Innovation (CDEI) - AI monitoring, partnership working and piloting and testing potential interventions in the tech landscape? Please explain your answer.

The ODI **neither agrees or disagrees** with the role proposed for the CDEI.

We do strongly agree with the need to build principles of fairness, transparency and trust into government's use of data and AI. As noted in the consultation, public trust in central and local government to use personal data ethically is worryingly low at around 30%[47].

The government's Data Ethics Framework is a promising guide to help data stewards in government and the wider public sector understand and address ethical considerations in their use of data. It includes the essential principle of accountability, which along with transparency and fairness, is necessary to build trust. One way in which the government could increase trust in its own handling of data and use of AI would be for there to be an independent body actively monitoring and holding public sector bodies to account.

It is for the government to decide which organisations should hold the government to account for its use of data and algorithms, but there currently seems to be a gap here that none of CDEI, ICO or the National Audit Office are filling. Building trust in government's and regulators' use of data is what is important. And, as the A-level algorithm fiasco demonstrated, getting it wrong can have serious consequences.

We also see the importance of ongoing monitoring of the use of data and AI both within government and across the rest of the economy. And we see a role for supporting public sector organisations, including regulators, who are adopting data and AI, and helping them to adopt more innovative approaches in safe and considered ways (here we note there is some overlap with GDS's responsibilities around improving the use of data across government).

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

If CDEI were to take on a role that aimed to build trust by strengthening accountability around the ethical use of data and AI by the public sector, then it would need a change in statutory status to ensure its independence from government. In this case, CDEI could report to Parliament, in the same way as the NAO.

[47] Nearly 9 in 10 people think it's important that organisations use personal data ethically – The ODI

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

Q12. We have identified five broad areas of work as part of our mission for enabling better use of data across government: *We want to hear your views on any actions you think will have the biggest impact for transforming government's use of data.*

12. 1 General

All five areas identified in the NDS are critical and work will be needed to address them if the government is to improve its use of data. Furthermore, as none of the areas operates independently, it will be necessary to work on them in parallel, ensuring that they are integrated as part of a coherent whole.

In general, government will face a number of challenges, including but not limited to:

- Difficulties scaling up good data policies,
- A lack of best-practice examples,
- Lack of senior buy-in,
- Siloed working and a lack of data accountability,
- A fear of 'getting it wrong' with data.

Measures to address these challenges are given below, but it is critical that the government – and particularly GDS – leads by example at all levels.

12.2 Quality, availability and access

Improving the quality, availability and access of government data will require the application of all the measures addressed in earlier questions that relate to data foundations, data availability and data infrastructure for the wider economy.

We recommend that every government department and arms length body (ALB) develops a data strategy and action plan that takes into account the unique context of that organisation[48]. It is essential that each organisation feels ownership over – and has sufficient funding to see through – its own data transformation.

Government should ensure that these are coherent, and that they properly factor in the importance of facilitating transparent, accountable, well-governed access to data, by defining a set of data principles, reflecting the Government Design Principles[49], and a set of data standards, reflecting the Government Service Standard[50], to be adhered to. These should include recognition of the responsibility of public sector bodies to steward the data

[48] How to write a good open data policy – The ODI
[49] Government Design Principles – GOV.UK
[50] Service Standard - Service Manual – GOV.UK

they hold in a trustworthy way, and facilitate well-governed and ethical access, use and sharing of data across government, with local government, and outside government, including the publication of open data where data does not need to be protected.

Departmental and ALB data strategies should outline the types of data held by the organisation and refer to any relevant legislation or guidance that applies to the management of the data they steward. It should also define the processes for collecting and sharing or publishing data, particularly as they relate to sensitive and personal data; any IP or licensing issues; mechanisms for engaging with external stakeholders; transparency and accountability; and metrics to evaluate the effectiveness of the strategy. Crucially, these strategies should consider the ways in which different types of data are or should be shared, and their value outside organisational boundaries.

To ensure that proper attention and investment is given to the maintenance of foundational national data infrastructure, we recommend that government puts in place a legislative framework[51] that designates certain datasets as foundational national data infrastructure, alongside a set of consequences of that designation, which should include things like it needing to be available for free and as open data or (if it cannot be open data because it contains personal or otherwise sensitive data) through the ONS' secure research service, and designated roles and responsibilities for the stewards of those datasets. Datasets could be moved into this framework in three phases: important and easy-to-transition datasets such as those held by ONS; less important and easy-to-transition datasets such as the various public registers defined in existing legislation; and the important but hard-to-transition datasets such as address and other geospatial data.

Some public bodies and national data institutions (such as Ordnance Survey, Companies House and others) will need additional support to change their focus to maximising the use and social welfare value (which includes value to the economy) of the data they steward rather than the income they can receive for it. Substantial, long term investments and organisational changes will be necessary as they shift to providing free access to open public data. These are likely to involve investments in digitisation, purchase of data assets held by the private sector, and alterations to business models, for example to shift from paid data licences to paid data services.

12.3 Standards and assurance

The topic of open standards has been addressed in q.7 in relation to data in the wider economy but they are an equally important element of our local and national government data infrastructure. Standards help us to publish and use data in consistent and easy-to-access ways. They create a broad agreement on the processes and technical aspects of the data that means the data can be used more broadly, more effectively, and that shared tools and systems can be created. Ultimately, they create open ecosystems that can transform policy making and the delivery of public services.

[51] Legislation for data infrastructure and open data: what would it look like? – The ODI

As noted above, much work has been done to develop and implement open standards and government should draw on and expand this work going forward.

12.4 Capability, leadership and culture

The NDS describes a pronounced shift in the way that data will be managed throughout the whole of government. As such, it will require a change management programme to be implemented. Effective leadership will be essential to effect the required cultural changes successfully.

The proposal in the NDS that this should be led by a Government Chief Data Officer is welcomed. It is, however, critical that this post carries sufficient authority, levers and resources to effect the necessary changes, and is able to corral the efforts of Chief Data Officers across the public sector. Change must have leadership from the top but needs to have buy-in from across the public sector and at all levels.

Beyond leadership, the importance of building capability across the public sector cannot be understated. Clearly, there are many technical issues but human factors generally determine the effectiveness of data policies. Realising the full potential of data in government will be completely dependent on the abilities of the people involved. Attracting and retaining the right staff and upskilling those already in place will be essential.

In terms of delivering better public services, good data literacy in political and civil service leaders will be vital[52]. Data can be a powerhouse for transformation. But machines and algorithms aren't infallible and to maximise success and minimise risk requires the knowledge and understanding of how to use and apply data correctly. Data skills within the civil service need to be present outside the analysis and digital functions. For example, policymakers need to understand how to use data as a tool to achieve their goals. Procurement professionals need to understand how to build data access into procurement, as well as how to use open contracting practices, including data publication, to increase the fairness of procurement. The commercial function (and the newly announced BEIS unit focused on increasing the exploitation of government's intangible assets) need to understand the impacts of restrictive licences on the use, and therefore social welfare value, of data.

Skills is rightly chosen as one of the Pillars of the NDS and much work has been done to understand and promote data and digital skills at all stages of the educational pipeline. For example, work done by the ODI on The Data Skills Framework[53] breaks down the complex landscape of data skills into the sets of skills required by different people in an organisation. It illustrates how technical data skills must be balanced with 'softer' skills that promote leadership, good data governance and stakeholder engagement, along with business skills that enable data innovation.

At this stage of the NDS, when it is concerned with establishing the right framework, it is important that the correct building blocks are put in place to build this full range of

[52] Fiascos and the future – why data literacy is essential – The ODI
[53] Data Skills Framework – The ODI

capabilities. Government should publish a working definition of the various different aspects of data and digital skills for the public sector and the wider economy, and consider the benefits of providing information on pathways into data related careers.

12.5 Accountability and productivity

Accountability is a vital aspect of good data governance. It is closely linked to other important aspects such as transparency, ethics and trust. The government's intent in linking accountability to productivity is understandable, as it is important that the use of data does drive improvements in the effectiveness and efficient delivery of public services, but government must be held accountable across a range of outcomes for its collection, use and sharing of data, including fairness, access and value for money.

In the first place, a clear understanding of who is accountable for what will be needed. This should start with clarity about the division of responsibility between Cabinet Office and DCMS around government data. The Machinery of Government change on the government use of data[54] transferred this responsibility to the Cabinet Office. However, it is not clear who has responsibility for the use of government-held data by wider society and to support the economy, or data policy that might facilitate the use of private or third sector data by the public sector. It is also not clear how these responsibilities are devolved to nations, regions or local authorities, or to government departments and arms length bodies. Good data policies need to be context dependent, so some level of devolution of these responsibilities is necessary.

Responsibilities will also need to be defined for who has a role in carrying out independent scrutiny of the government's use of data. This may rest with regulators, the National Audit Office or other official bodies. Government's use of data should be designed such that it supports outside scrutiny by this body and by the public, civil society and the media.

We agree with the NDS's focus on using data for the monitoring and evaluation of the effects of government policy and delivery on our society, economy and environment. However, we would caution that care must be taken to balance the picture painted by things that can be measured quantitatively with qualitative analysis. Monitoring data used to evaluate the effectiveness of public policy or delivery should be published as open data, so that it can be analysed and visualised in different ways, combined with other data from different sources, and so that alternative interpretations and a richer picture can be built up in a collaborative process.

12.6 Ethics and public trust

Data ethics relates to good practice around how data is collected, used and shared; the importance of this was illustrated with the data and models used during the Covid-19 pandemic. Data ethics is a core principle of establishing good data infrastructure; it is especially relevant when the use of data has the potential to impact people and society, directly or indirectly, as is the case with public services and policy making.

[54] Written statements – UK Parliament

The complexity of evaluating the ethics of the use of data should not be underestimated, particularly where it involves contracts to share personal data between public and private actors. This was highlighted in a recent case where the Department for Health and Social Care entered into a partnership with Amazon[55] and has been a consistent theme throughout the pandemic. Government has to be proactively open about its use of data – including the details of its relationships with private sector organisations, and their access to data – to retain trust and to avoid misconceptions and conspiracy theories that undermine public trust in government.

Demonstrating that government collection, use and sharing of data is ethical is crucial to maintain public trust. A recent survey[56] conducted by the ODI and YouGov found that nearly 9 in 10 people (87%) feel it is important or very important that organisations they interact with use data about them ethically. Only around 30% of respondents trusted central or local government to do so. Government has to work proactively to build the trust of the public.

Building trust should be explicitly considered at all stages in the data lifecycle:

- When stewarding data, that is, the collection, maintenance and sharing of data.
- When creating information from that data in the form of digital services, analyses and insights, or stories and visualisations.
- When deciding what to do based on data, in automated and non-automated decision making, or during evidence-based policy-making.

Embedding ethical data practices is one part of building this trust. We welcome the Data Ethics Framework[57] developed by the government. However, it is not clear the extent to which it is being adopted and used, nor who is monitoring and holding to account government teams, to ensure data ethics assessments are completed properly. This needs to be remedied through clarity about which independent body (such as ICO, the NAO or CDEI if made independent) is responsible for holding the government to account over its use of data. As a matter of good practice, the results of these assessments also need to be made public so that they can be scrutinised and to provide assurance to the public that the ethics of data projects have been properly considered.

Transparency is a core principle of the government's Data Ethics Framework, which also highlights the importance of engagement with external stakeholders, including experts, civil society organisations, target users, and the general public. To build trust with the public, we would strongly encourage the government to be proactively open and transparent about its collection, use and sharing of data, and to listen and respond to concerns. Many of the headlines that undermine trust in the government's handling of data are the result of a lack of openness and transparency, which means the public is surprised, that the government misses early opportunities to spot and fix problems, and that space is left for misinformation

[55] Jeni Tennison on Amazon's access to the NHS website – how has it come to this? – The ODI
[56] Nearly 9 in 10 people think it's important that organisations use personal data ethically – The ODI
[57] Data Ethics Framework – GOV.UK

and conspiracy theories to thrive. Government should invest in good communications support around data projects.

Q13. The Data Standards Authority is working with a range of public sector and external organisations to create a pipeline of data standards and standard practices that should be adopted. We welcome your views on standards that should be prioritised, building on the standards which have already been recommended.

We welcome the continuing work of the UK government to increase the use of open standards across the public sector, through the Data Standards Authority. We would encourage the proactive investigation, development and adoption of standards in areas that are government priorities (such as the personalisation of digital public services and the data exchanges that are necessary to support them); continuing adoption and support of transparency standards such as the Open Contracting Data Standard; the definition of profiles of adoption of schema.org, including in the publication of metadata about datasets; and the availability of registers that provide standard identifiers and other information about things that are referenced across multiple datasets.

Further, defining standards and blessing them for use in the public sector does not guarantee that they are adopted. We would therefore encourage the Data Standards Authority to pay at least as much attention, if not more, to activities that will get existing government standards adopted consistently. This should include documentation for standards, the creation of tools that support their use, mechanisms for discovering standards, and embedding references to standards within processes and practices, such as in training, procurement and service assessments.

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

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

No response

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?

No response

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?

No response

Q16. What are the most important risk factors in managing the security and resilience of the infrastructure on which data use 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.

No response

Q17. Do you agree that the government should play a greater role in ensuring that data does not negatively contribute to carbon usage? Please explain your answer. If applicable, please indicate how the government can effectively ensure that data does not negatively contribute to carbon usage.

The ODI **strongly agrees** that the government should play a greater role in ensuring that data does not negatively contribute to carbon usage. The use of data and IT services contribute a significant and increasing proportion of global carbon emissions - estimated currently at 1.4%. Digital services are powered exclusively by electricity. Any growth in emissions from the sector needs to be offset by a general reduction in the carbon intensity of the grid in line with the government's climate change targets. This could include, for example, locating data centres near to sources of renewable energy.

The Government Digital Service is to be commended for carrying out an assessment of their own climate impact[58]. However, it is instructive that they encountered problems in accessing the relevant data on their electricity usage.

Government should incentivise the use of renewable energy by data centres, and develop data infrastructure to publicly track energy usage by the infrastructure that supports data storage and processing, so that those developing digital services can make more informed decisions about their use of that infrastructure, and consumers using digital services can make informed decisions about the services they want to use.

[58] Measuring the climate impact of our digital services at GDS - Government Digital Service

Mission five: Championing the international flow of data

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

Trade, productivity and international innovation is a key area that the government is right to focus on[59]. Data flows in international trade support considerable economic activity across the world, but increasing trust in them could boost innovation and growth even further while ensuring that citizens' data is kept safe. To take part in the complex data sharing that supports trade in services and the development of frontier technology, countries will be competing on the quality of their national data infrastructure – the data that they have access to, how they share it, and their ability to enforce privacy regulations.

Access to data and the sharing of it across borders has become central to trade and economic competitiveness discussions across the world[60]. Chile, New Zealand and Singapore recently signed the Digital Economy Partnership Agreement, which puts data flows at the heart of their trade relationships. The European Union is encouraging international adoption of GDPR, and Japan has started the Osaka Track at the G20 for multilateral discussion of data standards in the digital economy.

To support international data transfer to the UK, and support domestic businesses, it is vital that the UK maintains a strong UK data protection framework that provides reassurance to other countries about how their citizens' data will be treated. The government must provide clarity and confidence about the direction of any future changes to the UK's data protection regime, and provide for the fact that such changes may require renegotiation of relevant trade agreements. The UK should also press other countries to be clear on their data protection frameworks, including controls over the transfer of data to third countries, as this will affect the ability of UK citizens to feel confident using services hosted in other countries. Given the level of adoption of GDPR worldwide, it is likely that continuing alignment with European data protection regulations will remain important; as has been highlighted repeatedly, achieving data adequacy with the EU is essential.

In addition, government should:

- Seek provisions with trade partners – including current negotiations with the EU, US, Japan, Australia and New Zealand – that remove unnecessary barriers to cross border data flows, with specific commitments to prevent the use of unjustified data localisation measures.

[59] Data 2020: Trade, productivity and international innovation – The ODI
[60] What are the links between data infrastructure and trade competitiveness? – The ODI

- Advocate for the importance of global data flows in the World Trade Organisation (WTO), G7, G20 and Organisation for Economic Co-operation and Development (OECD)

- Draw upon the expertise of the UK co-chaired Data Governance Working Group under the Global Partnership on AI to work with international partners and explore approaches to international data access and sharing.

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?

International data flows support considerable economic activity across the world[61]. Citizens need assurance that digital services provided by international companies will be safe and secure, and that data hosted outside their countries will be handled responsibly. These needs are reflected in trade deals and international regulatory cooperation worldwide. Alignment of data protection is not only critical in the UK's future relationship with the EU but to the other trade deals being negotiated, to ensure that the data rights of citizens continue to be protected and that wherever possible data localisation (keeping data within national borders) doesn't become embedded or a hindrance to the flow of data internationally. The UK Government should seek explicit alignment on a high standard of data protection in trade deals and be transparent about its approach to data protection within trade deals.

It is important that the UK achieves data adequacy with the EU and that its own data adequacy regime does not undermine that data adequacy (for example by permitting onward transfer of data about EU citizens to third countries). This is necessary to enable data to be transferred from organisations in the EU to the UK, which is important for many digital service businesses across the UK.

There are several challenges to the UK gaining data adequacy from the EU, that are also likely to arise when other countries assess the UK's data protection regime: data collection and retention under the Investigatory Powers Act 2016; rights removed in the Data Protection Act 2018 when data is used for immigration control; and the onward transfer of data to other countries. These challenges have been highlighted by the recent Schrems II ruling that rested on the fact that data about EU citizens transferred to the US is unprotected from US surveillance. Overall, this requirement for data adequacy has to keep the UK's data protection law from straying too far from GDPR and should cause the UK to re-examine its own regime around data use for intelligence and surveillance and for immigration control.

Data adequacy should be a two-way arrangement between countries, so while the UK seeks data adequacy from others, it must also assess their data adequacy and the protections granted to UK citizens and businesses. Government should establish an independent HMG capability to conduct the UK's own data adequacy assessments for

[61] Evolution of data requirements in international trade agreements – The ODI

transfers of personal data from the UK. This would include a review of the transitional arrangements for international data transfers and the use of alternative transfer mechanisms which ensure that transfers of personal data outside the UK are appropriately protected. We would encourage the UK's data adequacy assessment criteria and process to be clear and transparent, both to provide predictability about the process itself for other countries, and to provide assurance to UK citizens and businesses that the process is robust.

We suggest prioritising those countries with whom we have strong and/or rapidly growing digital services imports and exports. These should naturally include the EU and the US but also countries like Canada and India. We would also suggest exploring participation in the Digital Economy Partnership Agreement alongside Chile, New Zealand and Singapore and more generally aligning with countries that take a similar approach to the UK around both digital innovation – including open data – and data protection. And we would suggest identifying areas for international collaboration around particular topics – such as data to address the climate crisis or global health surveillance – and shaping narrower data agreements around these topics.