# OpenUK's response to the UK Government's 2020 Consultation on the National Data Strategy

9th December 2020

Overall

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

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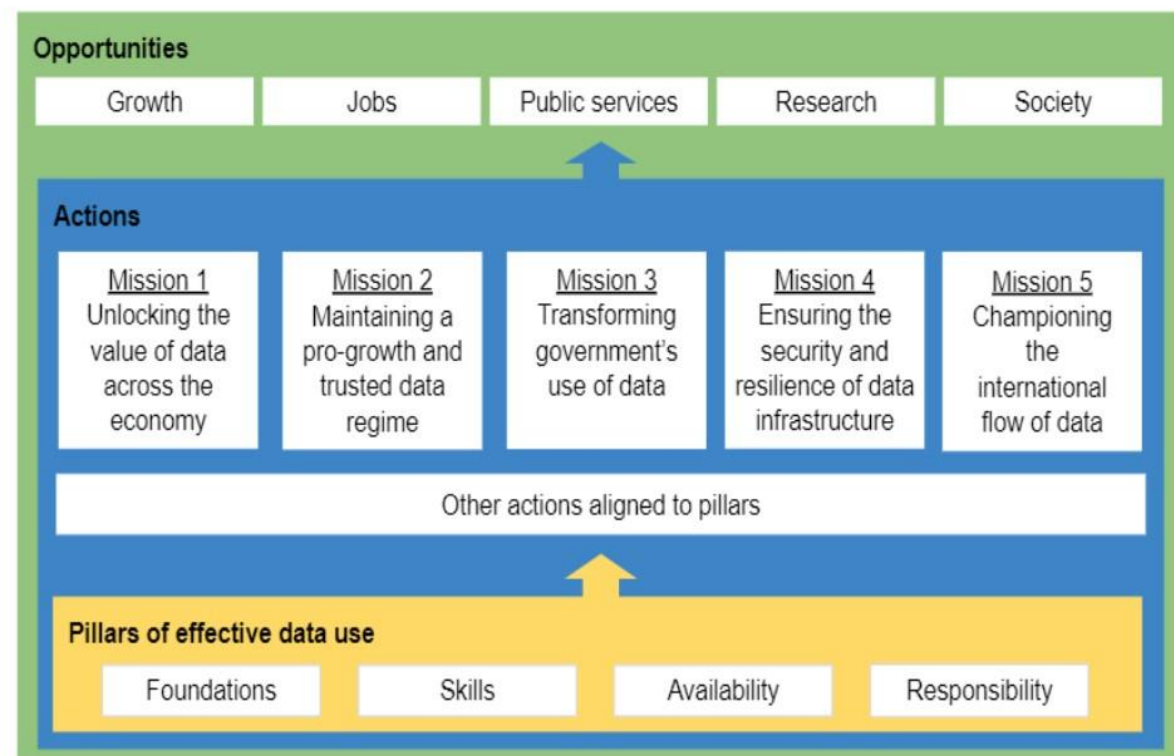
Q1. To what extent do you agree with the following statement: Taken as a whole, the missions and pillars of the National Data Strategy focus on the right priorities. Please explain your answer here, including any areas you think the government should explore in further depth.

Agreed.

We recognise this strategy as being a central part of the government's wider ambition for a thriving, fast-growing digital sector in the UK, underpinned by public trust. The UK has an opportunity to take on a world-leading role, sitting between the major tech powers, the US, China, and potentially Europe as a point of connection post-Brexit. This opportunity to become the hub or central 'data laundry' is one that is there for the taking.

Care must be taken to ensure trust (as espoused in Mission 2) as this underpins success in the other Missions. Championing international data transfer mustn't undermine the potential for UK to be trusted as a data destination, and our ability to process data in the UK which is subject to the GDPR.

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.

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

Citizens and their ability to control what data is captured and used should have been at the centre of data and its management. We also consider the management of data should have been localised as much as possible to empower local government to respond in a more agile way.

It has been correct not to use data for law enforcement purposes, as this would undermine fundamental freedoms necessary for public trust in government. In an ideal world citizens would trust Government before corporates, and the transparency that is created from using both: (i) open source software in public infrastructure; and (ii) open data, is paramount in building trust in Government.

We have seen examples of data being used in transport infrastructure in Germany to enable train operators to maintain social distancing, by analysing where individuals are seated on trains and pre-loading platforms accordingly.

We also saw Google, Apple, and others release data about mobility in various areas, and where this was lower than usual (for example). We believe this will have been useful for some business owners (and individuals) for planning purposes. We note that organisations such as Expedia collated this information, along with data on government responses, to create public data sets around how travel is impacted in different countries—these in turn could be used by those in the travel industry to plan based on which markets were recovering and which weren't.

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?

Society naturally categorises people, and this can be reflected in datasets. This needs to be recognised, and datasets interrogated more thoroughly, recognising that individuals within the data can (but not necessarily will) be adversely affected by the underlying bias derived from society. In particular we refer to the US Supreme Court case of Loomis v Wisconsin 881 N.W.2d 749. This describes a case in which inherent police bias led to a skewed dataset, causing the results of data-driven policing strategy to further reflect the underlying bias—i.e., a positive feedback loop arising out of confirmation bias.

Opening datasets in an appropriate way can assist in identifying and understanding bias, and therefore eliminating unwanted bias. We note that this isn't just a case of ignoring certain types of data (for example relating to a protected characteristic), it's about identifying where the data may lead to unfairness. Achieving fairness relies on recognising that bias doesn't only emerge when datasets don't reflect society, but also when datasets accurately reflect unfair aspects of society.

It is generally contrary to notions of equality to categorise the population into broad groups, and true equality comes from recognising and taking account of individuality. Where minorities appear in a dataset, it will typically not be possible to draw general conclusions about them that are accurate—certainly not to the degree necessary to form the basis for effective Government policy. Government needs to recognise this, and offer protections for those minorities that may be adversely affected by such issues created from data-driven decision making.

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?

Better numerical literacy is needed at all levels and in all parts of the country, suggesting that better teaching of maths both in schools and beyond will be critical for the better understanding of the results of data processing. It is important to ensure that those developing algorithms have a good understanding of data science—something which is currently not always (or even rarely) true.

Teaching about data, particularly open data, as part of the curriculum will be necessary to create a digitally skilled population equipped for the future. This should also include teaching about open technology more generally, of which open data is a part.
Datasets should be made open and available for education and training purposes.

Mission one: Unlocking the value of data across the economy

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

Our first mission is to create an environment where data is appropriately usable, accessible and available across the economy – fuelling growth in organisations large and small. We will create a clearer policy framework to identify where greater data access and availability across and with the economy can and should support growth and innovation, in what form, and what government's role should be, in the UK and globally.

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

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.

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

One of the key realisations of the transition to a digitised environment, the final steps of which we have seen as a consequence of the pandemic, is that all companies are tech companies and ultimately all rely on data for their future revenue streams.

Anti-competitive behaviour will be curtailed by making data open and putting all players on an even playing field.

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

We expect Government to lead by example: (i) opening Government data; (ii) using and promoting open standards to enable interoperability; and (iii) providing information as to the source of data and validation processes to enable trust in data.

Data sovereignty assurance will also be the responsibility of the Government. We have an opportunity to be a centre of data laundering and to be able to use our independence to promote UK business and services around data. With the data market in the UK already being the largest in Europe, the Government needs to ensure that we retain our position as a leading data economy and do not lose market share as a consequence of Brexit.

Companies will rely on Government to: (i) ensure that the UK has an adequacy ruling from Europe post-Brexit; (ii) ensure suitable rulings, treaties etc. with other nations to put processing of data in the UK at the centre of the global data economy; (iii) ensure future legislation promotes digital; and (iv) consider digital and data sovereignty.

Citizens will rely on Government to: (i) create transparency and trust; (ii) ensure that international data flows are handled in a suitable manner; and (iii) put citizen control at the heart of data.

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

No response

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

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

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

Strongly agree

In addition to the points we've made above, Government should help to establish best practices by: (i) educating organisations and supporting them to introduce good data foundations; and (ii) improving the educational curriculum to produce an adequately aware and skilled workforce and to ensure individuals are able to function effectively in a digital economy.

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

The Smart Data Review in 2019 consulted on ways to make evolving schemes more coordinated across banking, finance, telecoms and energy. The focus of Smart Data is citizens asking their providers to share information about them with third parties.

SMEs are more likely than large enterprises to have digital skills deficit within their workforce, with a smaller pool of talent and less resources to deliver training to staff. Further education and training should be made available to close the digital skills deficit.

SMEs have already invested heavily in complying with GDPR. Making international data transfers as simple as possible whilst ensuring those transfers continue to comply with European regulation would reduce the administrative burden on SMEs.

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

Ensuring that data is put to work by businesses and government in an open and transparent manner, in a way which is supportive of consumer and data subject rights. In particular, giving consumers more control over their data (and therefore better potential for benefits/rewards from allowing processing), and reducing data lock-in with particular service providers.

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

Building on our status as a world leader in technological innovation and our robust data protection standards, we will maintain a data regime that supports the future objectives of the UK outside of the EU and promotes growth and innovation while maintaining public trust. This regime will not be overly burdensome for the average company, nor will it be unnecessarily complex or vague; it will help innovators and entrepreneurs use data legitimately to build and expand their businesses, without undue regulatory uncertainty or risk at both the domestic and international levels.

To encourage the widespread uptake of digital technologies, we will also work with regulators to provide advice and support to small- and medium-sized businesses to help them expand online, and develop sector specific guidance and co-regulatory tools to accelerate digitisation across the UK economy.

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

Ensure open data, where possible putting companies on an even playing field. Government should identify where companies have advantages by being a supplier in multiple places in the ecosystem giving them the ability to use data gathered in one area and using it to unfairly compete in another. An example of behaviour to be avoided is that alleged in the EU

Commission's recent Statement of Objections to Amazon for the use of non-public independent seller data..

Ensure we comply with best international practice in line with Canada and the EU, and in a way which enables data exchanges with other key jurisdictions, whilst being mindful of British data sovereignty.

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

Q11. To what extent do you agree with the 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.
- **Strongly disagree**
- **Somewhat disagree**
- **Neither agree nor disagree**
- **Somewhat agree**
- **Strongly agree**

Neither agree nor disagree.

The functions mentioned above appear to be along the right lines, but it is not currently possible to comment as there is insufficient detail. We suggest further information is published in the interest of transparency.

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

No response

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

There is massive untapped potential in the way the government uses data. We will implement major and radical changes in the way that the government uses data to drive innovation and productivity across the UK. In doing so, we will improve the delivery of public services, as well as our ability to measure the impact of policies and programmes, and to ensure resources are used effectively.

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

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

Q12. We have identified five broad areas of work as part of our mission for enabling better use of data across government:
- **Quality, availability and access**
- **Standards and assurance**
- **Capability, leadership and culture**
- **Accountability and productivity**
- **Ethics and public trust**

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

The structuring is less important than the breadth of delivery. The biggest impact the government could have is by delivering transparency in the use of data. Transparency is fundamental to this, as it underpins the critical element of trust.

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'd suggest prioritising the reuse of existing open standards wherever possible rather than creating new ones.

Open standards should be prioritised, and all standards should be opened. Schema definitions, standard formats for data and metadata, and standards for data interchange should be the priority. Wherever possible these should be defined and managed by an independent entity like the Open Data Institute, Apache Foundation, W3C etc.

Government should ensure all data formats used are 100% open (e.g., using open document over Excel, big data file formats and schemas like Avro, Parquet, ORC over

proprietary formats etc.). This would also apply to data containers like relational databases—again the priority should be on open source solutions like mySQL, Postgres, Cassandra, Mongo etc. Note that some of these are open source but are in effect controlled by commercial companies, so care needs to be taken to ensure there is no lock-in to any of them and that data can easily be exported from them and moved somewhere else. Hence the emphasis on schema definitions etc. above.

When working with external organisations this needs to be very transparent, particularly if they may have their own commercial interests in the outcomes, and the opportunity to participate should be opened to all.

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

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

The government will determine the scale and nature of risks and the appropriate response, accounting for emerging trends in the market landscape. We will also determine whether current arrangements for managing data security risks are sufficient to protect the UK from threats that counter our missions for data to be a force for good. And we will consider the sustainability of data use, exploring inefficiencies in stored and processed data, and other carbon-inefficient processes.

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

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

We have telecoms regulation that applies to providers of communications networks and services, implemented in the Communications Act 2013 and other instruments, which we believe is broadly fit for purpose. We do however believe that open technologies and standards can support security and resilience, for example the use of open server hardware in data centres, and openRAN in mobile networks. This may also have tangential benefits by enabling innovative developments (such as the integration of open server-based data centres in heat networks for environmental benefits).

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?

Collaboration with European colleagues in GaiaX would also be a sensible way forward.

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.

We believe a level playing field allowing new entrants to the market from the UK would be beneficial. This can be supported by clear regulation and action taken against bad actors, including under competition law. Opening up datasets and technology will assist in reducing barriers to entry.

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.
- **Strongly disagree**
- **Somewhat disagree**
- **Neither agree nor disagree**
- **Somewhat agree**
- **Strongly agree**

Strongly agree

The use of open hardware in data centres has enabled the implementation of data centres integrated into heat networks. The government should also ensure that data infrastructure is included in public planning.

We note the significant increase in demand for processing power in recent years, citing the use of GPUs to support blockchain as an example. It would be beneficial to look at other methodologies that have a reduced energy requirement (such as proof of space and proof of time, rather than proof of work). A reduction in energy and hardware requirements would have a significant beneficial impact going forwards.

OpenUK is actively working on a project for the COP26 in November 2021, looking at the future of the Data Centre, the circular economy and how carbon emissions can be reduced

by a combination of open source software, open hardware which if course lends itself to open data.

Mission five: Championing the international flow of data

In our hyper-connected world, the ability to exchange data securely across borders is essential.

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

Using our reputation as a world leader in digital, a champion of free trade and the rules-based international system, and an engaged, rule-abiding member of the global community, we will build trust in data's use, creating the regimes, approaches and tools to ensure personal data is appropriately safeguarded as it moves across borders. We will also facilitate cross-border data flows by removing unnecessary barriers to international data transfers that promote growth and innovation. And we will seek to promote data standards, data interoperability, and UK values internationally.

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

We believe the UK can achieve this by ensuring it establishes best practice, in line with Canada and the EU.

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

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

We agree EU data adequacy is a key priority.