National Data Strategy

Consultation

November 2021

Station10

Contents

1.0 Foreword

We are living through what has been called the Fourth Industrial Revolution, in which data and automation come together across sectors, disrupting industries in ways not seen before. The Internet of Things, robotics, biotechnology, and artificial intelligence, amongst others, are at the forefront of this Fourth Industrial Revolution. Underpinning this is the collection, usage and processing of data in new ways to deliver value. For individuals, for businesses, and for society.

Clearly, therefore, data is a critical part of every organisation's future, whether that's making better decisions through insight and analysis, optimising existing business processes through automation or identifying new opportunities using AI and Machine Learning techniques. The most successful organisations in the 2020s will be those who understand, and embrace, this.

But such innovations don't happen in a vacuum. They need the right context and legal and cultural environment to thrive, and that's why national strategies are important. After many years where innovation has outstripped legislative progress, Governments recognise they have to create the right framework for businesses, individuals and society. GDPR has helped set the benchmark for legislative approaches around the world to ensure that individuals have legal rights around the information that relates to them, and this is a valuable platform on which to build a national data framework for the future. The UK Government has indicated its desire to be at the forefront of the Fourth Industrial Revolution, with leading approaches around AI Strategy, but this has to be founded upon a fair, balanced and forward-looking data collection, privacy and sharing framework that ensures individuals' personal details are protected, while enabling innovation.

At Station10, we have been pleased to host a series of roundtables, both virtual and physical, as part of the Department of Digital, Culture, Media and Sport's request for responses to the "Data: A New Direction" white paper. We have been delighted to act as a convenor of responses as part of this process. I would like to thank all those who have contributed opinions and thoughts as part of this process. We look forward to working with DDCMS and the data industry on the next steps to drive the Fourth Industrial Revolution forward in the UK.

David Ellis & Nick Willis
Station10

2.0 Consultation Overwiew

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

The opening paragraph from the Executive Summary is something we subscribe to on all levels.

Our strategic position as a business is to:

Decoding customers, Business and society

Station10 are experts in analysing omnichannel data to improve experiences; decoding society, businesses and people to unearth valuable insights into your customers. We translate your data into everyday language that drives positive change and accelerates strategic growth for your business, through defining, improving or transforming the digital experience.

Our commitment to the industry, our customers and our staff is to have conducted a full and formal consultation response to the Department of Digital Culture, Media and Sport to ensure the views of the businesses, individuals and key influencers of this strategy have been fully represented.

We analyse We decode We drive change

3.0 Consultation Outline

As part of the governments National Data Strategy, the key missions have been clearly outlined and are supported by a full and thorough chapter of proposed actions, implications and guidance.

Opportunities

Growth	Jobs	Public Services	Research	Society

Actions

Mission 1	Mission 2	Mission 3	Mission 4	Mission 5
Unlocking the value of data across the economy	Securing a pro-growth and trusted data regime	Transforming government's use of data to drive efficiency and improve public services	Ensuring the security and resilience of the infrastructure on which data relies	Championing the international flow of data

Other actions aligned to pillars

Pillars of effective use

Foundations	Skills	Availability	Responsibility

3.0 Consultation Outline

In line with the missions outlined and the supporting chapters and topics we performed a number of activities and events in the process of creating this consultation response, including;

Webinar events, Roundtables, Workshops, Individual interviews

These events and activities started at the beginning of November with an in person roundtable, and ended on the 16th of November with the final individual interviews with clients who wished to be involved in the response.

Though the consultation itself is structured around 5 chapters, to allow for a free flowing discourse around the topics our clients and partners would be most able to contribute to, we approached the roundtables and other sessions around key, and related themes. These were;

The Role of the ICO and Reducing Barriers to Data Flows

The role of the ICO

Should it regard economic growth, innovation and competition when discharging its functions?

Should complainants have to engage withi the data controller before raising a complaint?

Should the ICO have new powers to raise technical reports to inform investigations?

Reforming the Accountability Framework

Do you agree with the move to a risk based accountability framework?

Do you agree with the removal of legal requirements for a DPO?

Reducing the barriers to data flows

Thoughts and concerns around the approach to Adequacy and alternative transfer mechanisms

Views on certification schemes

The Role of the ICO and Reducing Barriers to Data Flows

The use of legitimate interests for data processing

Does an exhaustive list of legitimate interests help organisations alongside the LIA test?

What should be in the list of LI for processing?

Changes in approach to Cookies and online consent collection and management

Do the stated approaches for changes to the use of analytics go far enough?

Are there alternatives?

Are their technical issues created by these changes

the further processing of data

Are the proposed approaches to revisions going to aid your organisation to be innovative with the data it already has?

Anonymisation, Privacy, AI and Machine Learning

Is there an opportunity to derive more value through data anonymisation

Should there be more freedom to use personal data (either via pseudonymisation or anonymisation) to help train and detect unexpected biases in models?

What are the technical and legislative challenges that the moves towards an 'AI Super power' will create?

4.0 Consultation Participants

Though a small organisation when compared to some of the other designated conveners around this consultation, Station10 has a strong and varied list of clients, partners and general contacts in the field of data capture, analysis and insight. A view of a number of the organisations that we have worked with over the last 10 years can be seen in the figure below.

For the purpose of this consultation response, we engaged with our network and were fortunate to have input into the process, both directly and indirectly, from individuals from the following organisations:

- Adobe
- Another Peak
- Bee Herd
- Bunzl
- Capita
- The Children's Society
- Legal and General

- Observe Point
- Tealium
- Ted Baker
- Tesco
- Vue
- World Animal Protection

5.0 Consultation Response

5.1 The Role of the ICO and Reducing Barriers to Data Flows

The Information Commissioner's Office (ICO) is hugely important, and reforming it so it can deal with the future challenges is similarly significant.

Technical reviews and assessments are very beneficial for practitioners. However, the proposed solution may not go far enough, as this only relates to what happens when an issue is raised. There would be significant benefit in creating a certification scheme to show that the organisation is aligned with ICO standards. This is something that multiple contributors mentioned.

The assessments under this certification scheme could be (and practically would almost certainly have to be) outsourced to specialist companies to deliver within the certification framework. This is something that many organisations said they would pay for now. There are different possible existing models that could be used to deliver the "data kitemark" system. One is an ongoing "ISO" style assessment. This may be harder to achieve, as ISO certifications are often seen as something for larger companies, and may be difficult to implement, or to deliver, for SMEs or smaller organisations, like voluntary groups; the risk here is that you end up with a two-tier system.

Another model, however, could be an "MOT-style" certification, with a regular (annual, probably, but could be different) assessment by a certified practitioner, with "Major" and "Minor faults" noted, and with an action plan for those to be rectified before the next assessment. This has the advantage of being a well-known existing format, which could be applied nationwide for all types of organisation, and would mean a clear "UK standard kitemark"; indeed, there may even be an opportunity to export such assessments, thereby generating additional revenue for UK companies and the ICO, given how well respected the ICO is internationally.

It was agreed that the move to contacting an organisation first if there is a complaint, with the ICO acting as Ombudsman if required, was a good one.

The reuse of data is an important factor to be allowed.

One clear consistent statement was that Adequacy was hugely important. That's partly because many organisations have just spent years getting ready for GDPR, and the clear direction of travel from the ICO is that EU Adequacy would be achieved. So, the possibility of not being adequate would be a major blow to UK companies, especially when in competition for EU markets.

Adequacy is just as important for data flows in, as it is for data flows out. UK companies don't want to be operating in a country that is seen to be the "wild west of data".

The best thing is for clarity to be obtained. Each rule needs to have clear guidance on its application. This should be written both from a legislative but also a technical and commercial viewpoint; different roles in industry can have different nomenclature or wording, so it would be important to put these into the right language to ensure understanding.

One of the challenges for the ICO, and for legislators in general given the speed of innovation during the Fourth Industrial Revolution is that case law to create precedent works at a considerably slower pace to current technological development. This means that laws may need to be evolved or rewritten before they get properly tested in court. In a sense, this is already happening with UK GDPR – which is not even 3.5 years old - with this very consultation. This rate of change outstripping the law's ability to keep up is unlikely to stop any time soon, meaning a more consultative approach to data legislation and governance would be beneficial.

This could need an open forum for industry experts, the ICO and legislators to brief, improve and validate applications of the law to real-life examples on the cutting edge of technology.

If such a forum were to exist, it would need to be free from legal censure; if people feel they will go into it with the likelihood of being fined for anything they say, they are less likely to attend, and the forum method would break down. This would create an ongoing "wild west" environment that defined the first 20 years of this century, and wouldn't help Government manage the situation.

One other approach to foster an environment for the innovative use of data, is to have industry-specific certifications for particular data practices, with differing definitions and regulation activities to make managing compliance more straight-forward. For instance, there is existing regulation in sectors like Financial Services, so data used in FS will usually be different to that used in, say, Retail. FS will already have a set of definitions and regulations that they know they have to adhere to, but for data they now have to comply with definitions that apply to everyone. This makes it harder for compliance practitioners in that sector (and the same then goes for every other one) to keep up and understand whether new innovations in a totally different sector actually apply to them. In other areas of compliance, the regulation is clear, and in the case of FS fairly stringent, so it allows for greater focus. A similar approach in relation to data could be beneficial.

5.2 Reducing barriers to innovation and burdens to businesses

Lawful basis should still be needed for the reuse of data. Anonymised data is key to being able to use shared or reused data from elsewhere.

If we use that as a fundamental basis, there are some interesting ideas that could help innovation through data. Having a body that provides anonymised data sets to pass through AI or Machine Learning models would be beneficial, especially if this could avoid skew or bias in the data. It could be that these come from an official data source, whose role to ensure data quality is respected (the ONS, for example).

In addition, and in a similar way to some of the points raised in the previous section, a certification process for safely anonymised data to pass from business to business (2nd party data sharing, for instance) within GDPR content parameters would be hugely beneficial. Again, this could be a "kitemark" for data sharing, and systems and processes could be audited independently on a regular basis. This could be a valuable option, particularly for media companies looking to monetise their audiences, almost in a similar way to ABC and ABC Electronic for circulation data. Once again, this could be something that could be exported to organisations wanting to share data to "kitemark" standards internationally.

More broadly, however, there may be a larger opportunity. Borrowing from the Stock Exchange (an idea hatched in a coffee shop in London nearly 350 years ago), a stock is, in essence, a piece of information that says you own part of a company. It comes with effectively a guarantee that you own that, and have a set of established rights as part of that ownership (that you can sell it, that you can vote in relation to the company's governance, that it has a known, public value, and so forth). It might come with a share certificate, but other than that, there is nothing physical about the ownership, in the same way that data is not physical. Now, there are multiple Stock Exchanges around the world (FTSE, NYSE, NASDAQ, CAC, DAX, Hang Seng, etc); whilst they might have slightly different approaches, they all conform to a set of shared standards of what the basic definition of a share in a stock means.

However, there is currently no such idea of a set of rights in relation to data. There is no "Data Exchange" that guarantees certain standards that can be shared and can be consistently and universally applied. There is a very interesting idea of the "London Data Exchange", which defines, requires and guarantees any data that is shared through its exchange meets an established set of criteria, which are transparent and universally acknowledged. The organisation that underpins it is required, by charter or some other binding agreement that is backed in law, to manage and maintain these standards over the long-term. Given the UK, and London, has a historical link and expertise in such exchanges, this could differentiate the UK marketplace, at least in the short term.

To pick up and incorporate an earlier point, this could be managed centrally as one organisation, the "London Data Exchange" that manages all sectors. This is potentially a large job to establish from scratch, although the paradigm of the London Stock Exchange, which manages stocks across all sectors could show this could work. Alternatively, the requirements could be specified by different sectors, and so different organisations manage and "police" data relating to different sectors. For instance, the Turing Institute might become, by charter, responsible for the data sharing of certain "data commodities" for AI, the "Nightingale Institute" could be the organisation responsible for the Data Exchange in relation to Health, the "Attenborough Institute" for Data Exchange for Climate and the Environment, the "Portas Institute" for Retail, the "Rothschild Institute" for Financial Services, and so forth. Each would guarantee that a certain set of standards in which the data has been gathered, and processed, which means that they can be used in an agreed, and certifiably legal, way. It could effectively create a commodities market for data.

PMP – very few of our clients find the move to a PMP model useful, as they will likely need to maintain current DPO or DPIA processes for any SCC. PMP appears to be useful for smaller organisations with little or no exposure to international data, but for any organisation operating at an international level, the PMP approach is of limited value. Most organisations have a well-defined DPIA process, given the requirements under GDPR. Both those having to complete DPIAs, and those having to assess them, do not find this to be overly arduous as a process.

Legitimate Interests exhaustive lists will be valuable for organisations, but there are many complications.

Each defined purpose for use would require a clear definition, which can be understood by multiple different business functions and stakeholders with no ambiguity. Without this, the existing challenges for businesses will be maintained, or even exacerbated. That said, having the LIA test as a catch-all is an important point to ensure is understood within business and the industry. A better approach may be to establish a clearer process around LIAs – a clear flow diagram, for instance, or other assets for its application.

How the LI exhaustive list and the changes to cookie consent will be an important area to ensure correct legal application and compliance. It is clearly important to give explanations at the point at which individuals interact with anything that requires Legitimate Interest or is based on Strictly Necessary (for example, an AI recommender system on a page) to ensure people understand how this ad is being served, similar to the Google "Ad Purposes" link.

The technical implications of moving analytics cookies to Strictly Necessary will need to be made clear to businesses (for example, how can analytics data funnel into Personalisation engines?). That said, everyone agrees that the simplification of cookie consent will be beneficial, although greater education is needed on what is covered and how. Similarly, everyone agrees that making the individual the warden of their own data is the underlying principle of GDPR, and there may be better approaches to how this is managed; innovations in data sharing, such as Gener8, Brave and the BBC Databox, show ways in which this principle could be made more explicit, and to demonstrate value to the individual, rather than the perception that businesses are the main beneficiary.

5.3 Anonymisation, privacy, AI and Machine Learning

The complexity and breadth of current cutting-edge applications of AI and ML makes regulation and its assessment hard to even the most experienced individuals and organisations. In most areas of AI, the most experienced only have 20-25 years of practical experience, which again highlights how rapid change in the Fourth Industrial Revolution has been, and will continue to be.

The use of technical reviews by individuals who no "skin in the game" will become important, but would need legislation and confidentiality to enable this to happen. There is an opportunity to create an "AI Audit" industry, where ML systems' "black boxes" can be assessed and verified to be operating as intended, and in a fair way. There is already an emerging "Audit and Due Diligence" industry developing, particularly in relation to Private Equity investment in AI, but this is currently ad hoc, and there are no real standards associated with this. Once again, this could be an opportunity for a "kitemark" approach, but this is an area where the UK could lead this market, given its history in such areas. However, the UK would need to move fast, as this could emerge anywhere, and Silicon Valley, New York or Amsterdam could just as easily become the centre of such trade.

This could be a pillar on which the UK's "AI Superpower" strategy and positioning could be based (amongst others).

However, the main principle that will be key for the UK to become an AI Superpower is simplicity. As previously mentioned, sanitised datasets are the key for quickly developing AI technology, and vetted, "kitemarked" datasets enable this.

The rate of development of AI technology requires for regulation and governance to be easily navigated and clearly understood. Reference data will also be needed to be developed.

6.0 Recommendations

From the interviews, workshops and other sessions, we have the following recommendations to put forwards;

• Greater review should be put into the role of Technical Reviews, with many organisations wanting to have this as a service they can call upon, rather than be limited to breach reviews.

• Potentially a Kitemark-like or MOT style response from these reviews will help ensure transparency in how data is being used, and if used as part of the consent mechanisms, will help educate and build confidence in the overall data economy

• Anonymisation and the availability of anonymised datasets (either from a central body, or with certification available for production in the marketplace) are a key direction to enable the removal of bias and greater innovation in the AI and data space

• An exhaustive list of Legitimate interests will be of value, though having a genericised list to try and future proof may not be beneficial due to potential ambiguity caused. A regular review process, with input from the industry would be more beneficial.

• An open forum for leaders in the cutting edge of technology to engage with the ICO (and potentially the DDCMS) needs to be made readily available, to ensure that the pace of technological advances being greater than the speed of case law and precedents is overcome, with a strong understanding with the regulators of what is now possible

• The regulations need to be interpreted with use-cases in language that is used by the people in organisations who will be innovating and creating value in the Data Economy. Activity like this will provide greater clarity to those who are trying to work within the regulations, and ensure more 'bad behaviours' are avoided up front

• The change to PMP will only be of benefit for smaller organisations, with many large organisations seeing benefit from the DPO/DPIA model in place.

• More innovation needs to occur in the 'data locker' space, giving data subjects more ownership and control of the distribution of their data

7.0 Consultation Conclusions

The 'Data; A new direction' consultation has been a hugely important discussion, which has created a huge amount of opportunity in the 'Data Industry' to create a better environment for both data subjects and the data controllers/processors that will engage with them.

The content within the consultation has created a great deal of conversation in the industry, but has also given rise to a few ideas which were not originally included in the consultation which we hope will be discussed at greater length between the ICO and the DDCMS.
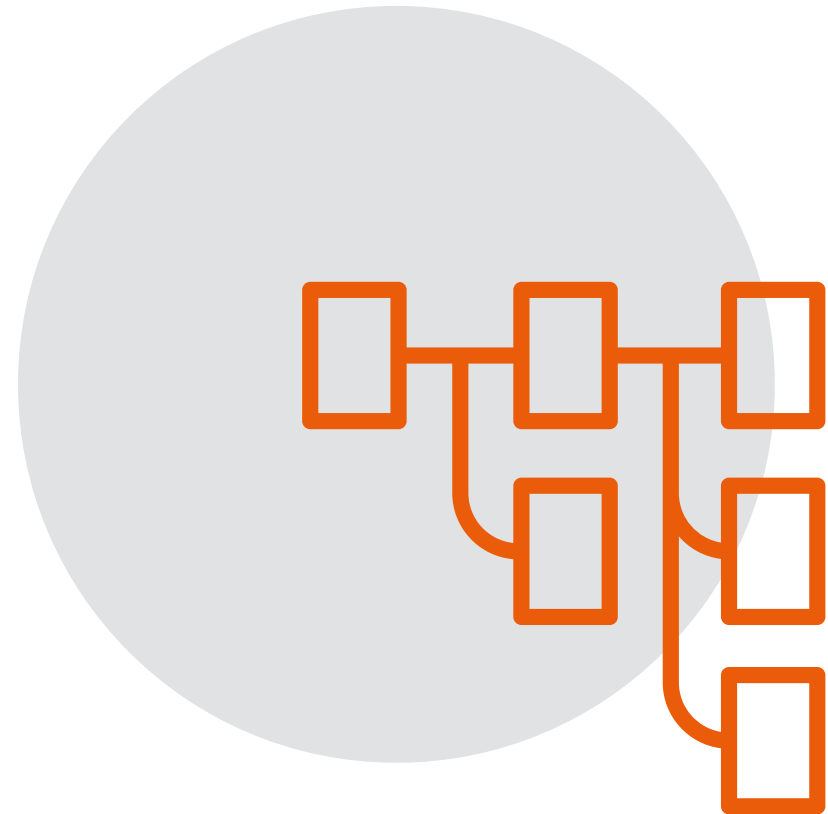
8.0 Acknowledgements

Thanks to the following for their attendance and involvement in this process:

Bayo Ade

Matthew Blake

Tim Coleman

Duncan Connal

Gavin Crouch

Matt Crow

Vincenzo De Simeone

Matthew Dixon

Mike Fong

Sarah Goulsbra

Bill Griggs

Tianxo Jia

Steve Long

Anthony Millest

Belinda Prescott-Collins

Chris Reid

Mike Stannard

Nathan Stockford