bright data

Bright Data's (formerly Luminati Networks) response to DCMS consultation on UK National Data Strategy

Introduction

The National Data Strategy is a pioneering initiative run by the UK Government in an effort to build a data-driven economy and deliver positive social outcomes and good public services. The NDS looks at several pillars, among them transparency, ethics, and education, that are the focus of our response, which in our experience, will maximize the value of effective and ethical data use.

Read our comprehensive response below

Bright Data's response to DCMS consultation on UK National Data Strategy

About Bright Data

Bright Data Networks is a leading online data collection company that serves thousands of customers worldwide, including multiple Fortune 500 companies. Owned by a UK company, EMK Capital, Bright Data works across multiple countries to help businesses from all industries to view the internet with complete transparency, aiding ad verification, brand protection, price comparison, fraud prevention, data collection, and cyber security.

We enable our customers to gather openly available (public) online data in a responsible manner, whether it's competitive data on a mass scale, reviews, market research, social media sentiments and more. In an era of massive data growth, these actionable insights are vital to make critical decisions fast and effectively.

Ethical standards are at the heart of Bright Data's DNA, and the company is proud to work alongside public bodies, businesses, academic institutions, and others to collect data through an ethical-by-design approach.

We have a strong commitment to building capacity and capability for effective and ethical use of online data for the long term, so we have established The Bright Data Academy, which works with over 20 academic and research institutes around the world. These include Princeton University, King's College London, Oxford University, Royal Holloway University, The Technion, ETH Zurich Research University, Tel Aviv University, Hong Kong University, Northeastern University, among others. In partnership with these institutions, we run practical workshops and seminars, providing data collection tips as well as promoting ethical and responsible data collection practices. We also support academic research with evidence and online data collection services and products.

We are responding to this consultation as we share the UK Government's ambition to build a world-class data economy in the UK while maximising the value of effective and ethical data use by public bodies to deliver positive social outcomes and good public services.

As a global leader in collecting, analysing and applying online data, we are pleased to apply our expertise, insight and experience to help ensure that the UK's NDS delivers on these aims. We would welcome the opportunity to contribute further by working in partnership with DCMS to further develop and implement the principles set out in the NDS.

Please contact Keren Pakes via kerenp@brightdata.com if you would like to further discuss how we can support the development and implementation of the NDS.

Response overview

Bright Data welcomes the publication of the National Data Strategy (NDS) as a clear show of the UK Government's commitment to embracing all of the opportunities of ethical and effective use of large-scale online data. We recognise the NDS as a substantial document that outlines the ways in which the UK can develop as a thriving data economy with more effective public services and positive social outcomes. We also recognise that the Strategy gives a detailed account of the issues that need to be accommodated in developing an advanced national approach to utilising data.

In further developing and implementing the NDS, we recommend that the Government focus on three priorities:

1. The NDS needs to focus explicitly on using data to enhance transparency across society and the economy in order to secure positive outcomes for consumers/service users and to have a transformative social impact;

2. This needs to be underpinned by high ethical standards and clear, fit-for-purpose regulatory guidelines that build trust and confidence among businesses, service providers and individuals to allow innovation to flourish.

3. There needs to be an investment of time, effort and resource into data skills and education that goes beyond data practitioners to develop the long-term capability of all businesses and service providers while empowering individuals as well-informed data users.

Detailed responses

Bright Data Networks offers the following responses to the consultation questions about where our insight and expertise is most relevant.

Q1. **To what extent do you agree with the following statement: Taken as a whole, the missions and pillars of the NDS focus on the right priorities**

Bright Data broadly agrees that the missions and pillars of the NDS focus on the right priorities

but believes that greater emphasis should be placed on the international context that the UK's data economy sits within. It is noted that the international flow of data is factored into the strategy, and features as an action-set, but we believe that it is an issue of high enough importance to be included as one of the Strategy's underlying pillars. This would recognise that online data is agnostic of national borders and will be most effective when open for non-

UK stakeholders to collate, analyse and apply, and for UK stakeholders to do so with data generated overseas. It would also reflect Bright Data's own experience, where clients work with us to gain real-time insight of large-scale data that allows them to listen to customer reviews, test products, offers and strategic plans and even go as far as identifying future market trends. Doing so offers huge value, not just to the businesses themselves but also to the consumers they serve, with the same 'true of service' providers and their users. It is based on the principle that there should be free, trusted and transparent access to data irrespective of geographic restrictions, so it is vital the NDS prioritises international data flows accordingly.

We believe that adding an additional pillar related to the international flow of data will give the issue a greater strategic prominence and thus help the UK secure Data Adequacy status with the EU following the Brexit transition period.

We also believe that the pillars could better reflect the social drivers of the Strategy. While 'Public Services' feature among the opportunities, the pillars and actions suggest a more economic emphasis. We believe that the potential of data to improve almost every part of life needs to be better represented, perhaps through an adaptation or addition to the existing pillars so that 'social outcomes' are highlighted.

Q2. **We are interested in 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.**

Bright Data has been involved in a number of international initiatives to support people and communities during the coronavirus pandemic, applying our expertise as well as our data and data collection capability. The following are two examples.

1. Using Artificial Intelligence and Machine Learning to enable early COVID-19 detection

At the onset of the pandemic, Bright Data offered all COVID-19 researchers worldwide the opportunity to use our Data Collector, an automated solution to accelerate their research. With a pandemic of this size, we knew that we needed to equip the people fighting it with the most up-to-date intelligence and information possible. As a company that usually works with large-scale organisations in the e-commerce and travel sector, we are used to collecting massive amounts of publicly available data in short time frames.

The international Sprint COVID-19 initiative approached us, and we subsequently participated in a project aimed at using smart devices to help enable early COVID-19 detection and prevent mass infection.

Their research in China and Italy indicated that 30%-50% of the COVID-19-infected population suffered from silent hypoxia. This means having a low level of oxygen in the blood

without any obvious shortness of breath or other visible symptoms, which translates into a higher chance of them infecting other people without even knowing it.

Sprint COVID-19 found that many commonly owned smartphones or smartwatch devices have a built-in sensor, known as a SpO2 sensor, that can enable testing your own blood oxygen levels independently and easily, which could be a key first step in identifying the virus early. Based on those results, the team developed a heuristic test that can be carried out on smartphones and smartwatches to identify asymptomatic, pre-symptomatic, and mild cases which would otherwise go undetected. The team turned to Bright Data to identify which of the possible 8,000 different device models were able to deploy this test.

We knew time was of the utmost essence in this sprint to conquer COVID-19, so we quickly used data collection automation to sift through the mass amount of publicly available online data to identify all suitable smart devices. We quickly found that smart devices carrying the SpO2 sensor included many older smartphone models.

Moving forward, we were able to assess 8,000 models of different smart devices. In just under 24 hours, we found out that over 110 smartphone models and 165 smartwatch models and smart bands have this sensor to test blood oxygen levels. In addition, Apple has now announced that their new watches will also carry the SpO2 sensor.

2. Using automated data collection to tackle growing unemployment rates in Israel

With the outbreak of COVID19, Israel, along with countries around the world, encountered a dramatic rise in unemployment levels as businesses were forced to close their doors or downsize significantly. In response, a number of public bodies and NGOs came together, including JDC Tevet, the Israel Labor Ministry, and Nova to provide programmes of support to help people into work.

Bright Data has been working closely with this coalition, applying our automated data collection solution to quickly and effectively collect data around labor market trends. While generally applied to map business and e-commerce trends in real time, this solution is providing the coalition with an automated and customized flow of data about employment trends and opportunities in one simple dashboard. We collect data from over 40,000 web pages each month, each giving details of around 15 employment opportunities. It would have taken the coalition many weeks to obtain the same level of data without the automated solution. As a result, data reports are able to be submitted twice-weekly, providing targeted and timely support to help job-seekers into work.

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

We believe that the development and full implementation of a NDS will, in itself, have a profound and positive impact on reducing inequalities through greater transparency. There are two dimensions to this:

- An effective data strategy will provide businesses and services with greater transparency over consumer/user behaviour and preferences, allowing for more personalised experiences that fully account for diversity at both a protected characteristic and an individualised level;
- An effective NDS will allow for more empowered consumers/ service users whohave greater transparency around how their data is being applied, placing an obligation on businesses/service providers to directly link data collection to improved, more personalised, inclusive user experiences.

Key to realising these benefits will be the issue of individual trust, so the strategy must precipitate a culture (and regulatory regime) where individual data is collected and applied based on 'opt-in' principles. Data on individuals with protected characteristics (and others) which is collated and applied without well-informed and active consent would have a detrimental impact. Achieving such a culture of data transparency must account for the varying needs of people with differing protected characteristics.

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?**

As with inequalities based on protected characteristics, we believe that the strategy has the potential to tackle regional socio-economic disparities through giving businesses/service providers greater transparency on how to build personalised experiences; and in empowering individuals with greater transparency about how their data is used and the expectations they should have of those using it.

To realise this, the Strategy must incorporate a strong educational element that builds the capability of individuals to make informed, proactive decisions about how their data can be most effectively applied; and equips businesses and service providers to most effectively utilise data.

This must take into account the particular dynamics of each region in the UK, including location-specific skills and educational needs; major employer and industry clusters; and other demographic factors. With this in mind, we support the stated intention in the Strategy to take an evidenced-based approach to building data skills and organisational capability, encouraging the Government to adopt a geographically targeted approach that acknowledges and incorporates regional variation. Our own experience of working with universities shows the value of working in partnership with academic institutions embedded in their regions and localities. We encourage the Government to adopt a similar approach by ensuring that the value of UK universities as regional anchor institutions (working in partnership with other private and public sector organisations) is a core part of the NDS.

Q5. **Which sectors have the most to gain from better data availability?**

We believe that the sectors that should be prioritised through the NDS are those that stand most to benefit from a leap forward in terms of data capability rather than those seeking to optimize already strong positions. These sectors are:

- Agriculture
- Forestry and Fishing
- Charity or Non-Profit
- Construction, Education
- Human Health and Social Work Activities
- Manufacturing
- Mining and Quarrying
- Water Supply, Sewerage, Waste Management and Remediation Activities

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

The primary role of central government should be to facilitate the effective flow and utilisation of data by business, service providers and individuals. This will unleash the full power of online data - with business decisions and commercial transactions better informed through information and transparency; better design of consumer/service user experiences; and better-informed consumers/service users. Central Government can play this role through three primary functions:

1. Effective regulation of an open data market that allows for business/service providers and individuals to have easy, timely access to accurate data;

2. Investment in education at every level to drive the long-term data capability of businesses/service providers and the ability of individuals to make informed data decisions and utilise their own data most effectively;

3. Setting high ethical standards for the collection and utilisation of data that promote trust, positive intention and mitigate the misuse of data or action that run contrary to transparency principles.

Q6a. **What role do you think central Government should have in enabling better availability of data across the wider economy?**

The role for Government outlined above should be consistent across all sectors and applications, as it will allow an open and transparent data economy that all will be able to operate in without innovation being constrained.

Q7. **To what extent do you agree with the following statement: The Government has a role in supporting data foundations in the wider economy.**

Bright Data strongly agrees with this statement.

Closed, inaccessible data is the number-one risk to building the data economy that will deliver benefits for the public. It will stifle innovation and inhibit market competition, hampering efforts to deliver better outcomes for people as consumers and service users. It is therefore vital that the Government acts to establish strong data foundations, including proper regulations that set a framework for how data is collected and used.

If this role is not fulfilled, commercial and public institutions are likely to put a disproportionate balance of investment and effort into keeping data concealed and closed. Only through the Government working to develop data foundations and robust regulation will organisations have the confidence and motivation to make the data they hold available. This is essential to drive open competition forward as well as to keep developing products or services that address the public needs.

Sound data regulation and foundations will drive a free market of data innovation, with positive outcomes for consumers/service users and the wider economy - with new jobs created and greater commercial transparency.

However, it is important for the Government to maintain a supportive role through establishing data foundations and a framework of regulatory standards. It should not seek to inhibit free market innovation by taking an overtly direct role in seeking to shape the actual practices required to implement regulation or utilise data foundations.

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

While SMEs are more likely to have the resources to execute the technical aspects of data use, they are less likely to have the knowledge-based and in-house capabilities to fully get to grips with the ethical complexities involved. This is likely to be particularly relevant to data collection, with the potential for SMEs to miss out on the opportunities of effectively utilising online data due to a lack of confidence that they will not fall foul of ethical considerations. This poses a significant barrier, so it is here that the Government has the strongest role to play. In doing so, the Government should:

- Develop clear and unambiguous guidelines for the ethical collection and utilisation of data;
- Invest in programmes of education and skills development to ensure that SMEs are able to effectively implement ethical data processes;
- Create certification measures for ethical data collectors to provide SMEs with reassurance of safe and affordable access to business data.

The final point in particular is vital to ensure that SMEs are able to access and utilise data in the same way that bigger businesses do. Currently, larger companies have a disproportionate level of access to relevant cross-vertical business data that is necessary to make important business decisions. A certified hub of public data where any business, big or small, could easily search for datasets as easily as a Google search would do much to level the playing field and allow innovation and competition to flourish, driving jobs and growth as well as better results for consumers.

Q9. **Beyond the 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 most important thing that can be done is to empower consumers/service users with the knowledge and understanding to appreciate how their data is being used and the benefits they stand to enjoy as a result. This will ensure informed proactive choices about how data is shared and create expectations that drive innovation and outcomes. For example, we believe that consumers are right to expect that the data they share with businesses leads to better, more personalised experiences, beyond informing sales and marketing activity.

To achieve this level of empowered, well-informed consumer behaviour, we again emphasise the importance of investing time, effort and resource into data-education.

Bright Data's ongoing work with universities around the world offers examples that can be adapted and applied to this. In particular, our recent collaboration with King's College London as well as Royal Holloway University and University of Oxford showed the value of bringing industry expertise together with learners in established educational settings. We were able to help build students' understanding of things like the international data revolution and the shift in online data consumption amongst large enterprises in the UK. We were also able to work with students to explore real-life use cases from the health sector, where data has played a key role in saving lives during the global pandemic.

In addition to this work with King's College London, Bright Data is actively involved with leading education programmes at the Technion - Israel Institute of Technology, Amsterdam University of Applied Sciences, ETH Zurich Research University, Northeastern University, Hong Kong University, Hong Kong's data-science community and has recently held a learning session with Royal Holloway University and University of Oxford. We recommend that the NDS expands upon and applies the principles of education/industry collaborations like these.

Building on our recommendation (in Q8) of establishing certified data hubs for businesses of all sizes to access, we would also recommend that a second stage should be to develop such hubs as consumer propositions. Such hubs of readily available, trusted and curated data would provide consumers with a resource that fairly compares product prices and service levels and provides that information to consumers, even at a local geographic level.

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

Effective regulatory standards and clear guidance are essential to driving technology advancement forward, so we welcome the commitment in the NDS to a regime that is 'not too burdensome to the average company' nor 'unnecessarily complex or vague.' We also welcome the commitment to achieve data adequacy status, providing certainty and transparency to international stakeholders in particular.

In addition, we would encourage the Government to develop a 'central learning hub' where the experience, expertise and insight of subject matter experts (such as Bright Data and other data-driven businesses) can be captured and applied to drive forward innovation and leading edge practice.

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?**

We somewhat agree with the functions set out for the CDEI.

In particular, while AI monitoring is an obvious ethical issue for the CDEI to focus on, it is some way down the 'ethics funnel'. That is, AI considerations are only relevant once data collection and preparation processes and practices have been accounted for. We would therefore encourage the CDEI to place greater emphasis on these preliminary data processes as an area of prime focus.

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

Bright Data does not have a strong view on the statutory status of the CDEI.

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

Accountability and productivity

Capability, leadership and culture

Ethics and public trust

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

Bright Data believes that these are the right areas of focus to enhance the use of data across Government - and the wider UK public sector. In particular, we believe that 'capability, leadership and culture' should be prioritised as the underlying facilitator of all other areas. There are two dimensions to this:

1. Establishing a culture of understanding across Government of the value of effective data use and the aims should be geared towards - better outcomes and experiences for consumers/service users. This focus on end-user outcomes - whether through using data to make technological advances, drive system/process enhancements or create more personalised user experiences - should be recognised across Government as the sole purpose of data use. We recommend that establishing such a culture be a core element of the Chief Data Officer's remit.

2. An investment of time, effort and resource into education and knowledge exchange, allowing academic institutions and the private sector to help build understanding and capability across Government.

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 been recommended**

Bright Data recommends that the following standards are prioritised:

1. Standards around transparency in data collection practices that give confidence to businesses and public sector organisations in opening up their data and reassures consumer/service users that their data is being put to use for their benefit.

2. Standards around ethical data use that build trust. In particular, we suggest that ethical standards are prioritised to ensure that the way that data is collected and used is 'ethical by design' (technology and process).

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

Virtual and physical data infrastructure survive providers should be held to the following requirements:

- They should meet baseline data security standards - SOC2 certification for private data handlers and ISO27001 for service providers handling Government data;

- They should have verifiable business continuity plans in place;

- They should build data protection policies that align with GDPR (or any other regulatory regime that may replace it in the future) and suit data processing and storing functions, with verifiable proof of implementation.

We would also make the case for compliance with further regulation to be developed as part of the NDS, as recommended elsewhere in this 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?**

Bright Data's own experience is that clients place huge value on formal certifications, demonstrable compliance regimes and ethical considerations being woven into the core fibre of the business. Our clients and partners are reassured by a system that is based on deterrence, prevention and enforcement. This three-tiered process is the executive arm of Bright Data's core value – ensuring our network remains safe and ethical. Putting the process into practice involves a wide range of measures, including:


Ensuring that full and informed consent is obtained from all of the people and organisations we work with to access publicly available data;


Maintaining infrastructure that ensures data is shared only in accordance with strict conditions;


Investing in user-experiences that ensure all who we work with are able to exercise tight control over how they work with us and that they will enjoy benefits in doing so;


Implementing thorough vetting and on-boarding processes to ensure that all who we work with meet our own strict standards;


Restricting access to our network and data sources only for verified uses, and monitoring usage closely through automated and manual checking processes.

We recommend that similarly robust measures be developed in the UK through the NDS, and that certification and verifiable evidence of implementation be used to provide assessment that data infrastructure service providers are compliant with regulatory regimes.

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?**

We believe that the respective roles and responsibilities should be as follows:

- Central Government should set the regulatory framework, set standards and implement/monitor them through a regulatory body or in partnership with an appropriate third partner;
- Data service providers should implement standards, abide by regulation and enforce the same through their supply chains;
- Data service provider clients should restrict work to those able to demonstrate certifiable compliance and verified implementation of regulations and standards.

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**

The most important risk factors and our recommended mitigation requirements are:

- The risk of unintended third-party access, requiring all access points to be monitored, authenticated and logged;
- The risk of unauthorised access to data by employees, requiring both physical and cloud access to be carefully managed and monitored;
- The risk of malicious external hacking, requiring stringent security measures to be in place;
- Supply chain risks, requiring authenticated meeting of standards and regulatory compliance;
- The risks of non-compliance with location-specific legal and privacy standards, or of those standards not meeting the UK's adequacy requirements (see answer to Q18);
- The risk of data being 'tainted' through non-compliance failure to meet security/privacy standards.

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.

Bright Data strongly agrees that the Government should play a greater role by defining clear carbon usage parameters and imposing clear fines where thresholds are exceeded.

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

Bright Data considers that any restrictions on international transfer of data should be well formulated and proportional, showing itself adequate to the aims being sought (i.e., protection of citizens' fundamental rights).

While current GDPR-related practices are very beneficial and show commitment from the regulators to protecting citizens, Bright Data believes that future regulations could benefit from a more risk-based, sector-specific approach: the higher the risk, or the more crucial the sector involved, the more stringent standards for flow of data should be applicable. This approach could replace the overall rigidity of international transfer rules and mechanisms, shifting to an accountability-based approach that considers actual potential for harm in particular circumstances.

Some of the factors to be considered in regard to risk could include elements such as:

- The sector in which a relevant company acts (different requirements could be created for specific sectors);
- The nature of the data being shared;
- The intention in sharing data;
- The volume and frequency of data being shared;
- The type of data subjects;
- Bespoke risk measurements for the specific organisation with which data is being shared, (i.e., did the organisation experience any breaches in the past three years? Did the organisation suffer any sanctions from relevant authorities in the past? Are they privacy certified? etc.);
- Mitigation measures through contractual, legal and technical provisions;
- The jurisdiction where data is being sent to.

According to these risk factors, different requirements and obligations could be imposed, including, for specific sensitive cases, the possibility of having to request special prior approval for the transfer from the competent data authority. This procedural approach of requesting notification from larger companies or for sensitive data could take inspiration from the antitrust law regime, which works in a combination of a flexible, principle-based law with strict requirements and enforcement both in a preventive and repressive manner. Other measures could include an obligation to conduct thorough due diligence on the organisation receiving data, on the same level as organisations in the financial arena.

From a technological perspective, Bright Data believes that, in addition to investing in more specific encryption requirements (such as homomorphic encryption and other novel approaches), the UK could set itself apart by establishing the parameters and architecture for a private blockchain solution, which would be the base for organisations in the UK to transfer data with higher levels of security and control.

As outlined above, Bright Data recommends that data transfer requirements be designed

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?**

with a risk-based approach, with different measures applicable to different situations. In such a context, instead of considering each jurisdiction or country as adequate or not, on a binary scale, the UK should classify countries according to a risk grade (as in international finance, for example), correlated with other group risk factors to determine objective criteria of adequacy.

Specific standards and requirements should also be applied to specifically sensitive sectors, such as surveillance, and taken into account alongside jurisdiction-based adequacy assessments. Any cases where clear adequacy assessments cannot be determined on this basis should require detailed submissions to appropriate authorities for approval, following the model of antitrust regulations.

We also recommend that the UK Government innovate by establishing a bilateral private blockchain foundation with key trade partners, which would provide a safe, private and control-enabling base for organisations to share data between each other.

Bright Data also believes that the UK could fulfil the role outlined in the NDS as a 'champion of good-quality, available data across the globe' by proposing that the topic of international data transfers be regulated in a more uniform way, either through multilateral agreements or through a dedicated panel for rule-based conflict resolution in an organisation such as the WTO.