APPG on ARTIFICIAL INTELLIGENCE

EVIDENCE MEETING 1
DATA GOVERNANCE
COLLECTION & USE

PARLIAMENTARY BRIEF

28 January 2019 | Committee Room 1, House of Lords

Evidence Meeting 1 | Data Governance: Collection & Use is a Parliamentary Brief based on the **All-Party Parliamentary Group on Artificial Intelligence (APPG AI)** evidence meeting held on 28 January 2019 at the House of Lords.

This meeting was chaired by **Stephen Metcalfe MP** and **Lord Clement-Jones.**

We would like to express our appreciation to the following people for their oral evidence: **Dame Wendy Hall (University of Southampton), Roger Taylor (Centre for Data Ethics and Innovation), Paul Copping (Sightline Innovation), Professor Birgitte Andersen (Big Innovation Centre), Jeni Tennison (Open Data Institute),** and **Lauren Sager Weinstein (Transport for London).**

We would also like to acknowledge the APPG AI Data Governance Task Force for the input and feedback: **Rumman Chowdhury (Accenture), Andy Forrester (HypeAccelerator Solutions), Rachel Free (CMS), Max Dillon (PwC), Christine Chow (HERMES Investment Management),** and **Paul Copping (Sightline Innovations).**

Rapporteur: **Niki Iliadis | n.iliadis@biginnovationcentre.com**

Contents

EVIDENCE MEETING 1 OVERVIEW

Details

- Date: 28 January 2019
- Time: 5:30 – 7:00 pm
- Location: Committee Room 1, House of Lords
- Participants: **146 registered attendees**

Speakers

- Dame Wendy Hall, Regius Professor of Computer Science, University of Southampton
- Roger Taylor, Chair, Centre for Data Ethics and Innovation
- Jeni Tennison, CEO, Open Data Institute
- Prof. Birgitte Andersen, CEO, Big Innovation Centre
- Lauren Sager Weinstein, Chief Data Officer, Transport for London
- Paul Copping, International Advisor, Sightline Innovation

Questions

- **How is data currently being used by AI technologies?**
- **What is the proper way of seeking consent for using an individual's data?**
- **Who should ultimately oversee the use of personal data for AI systems?**

INTRODUCTION

Context

It is impossible to think of AI and data in silos. Access to accurate and quality data is critical to make AI work and, in many ways, data is useless or insurmountable without AI-driven analytics and insight.

Given this symbiotic relationship, for society to reap the socio-economic benefits AI promises, we need to first ensure we have the right data governance models in place.

Reoccurring news of data misuse, violation of individual privacy rights, ingrained dataset biases, and huge data-driven corporate profits show how far we are from data governance models fit for our twenty-first century.

In this context, in 2019, APPG AI launched the Data Governance Pillar to further explore the need for improved data governance. Specifically, the pillar was created:

- to raise awareness on how data is collected and used by AI systems,
- to shed light on the challenges of the current data governance structures, and
- to showcase new data governance models.

Although the primary audience for APPG AI are members from the House of Commons and the House of Lords, the group also aims to inform the general public and key stakeholders across business, academia, and civil society.

Skeleton

This Parliamentary brief is broken into three parts, highlighting the main takeaways from APPG AI's first evidence meeting under the Data Governance Pillar. The meeting, held 28 January 2019, was chaired by Stephen Metcalfe and Lord Clement Jones, and orchestrated by the Secretariat of APPG AI, Big Innovation Centre. It brought together key stakeholders to gather and share evidence on how AI systems use data and to apply that evidence to help empower decision-makers to make policies in the sphere of data governance.

The first part of the brief builds on the oral evidence provided by Dame Wendy Hall, Roger Taylor, Jeni Tennison, Professor Birgitte Andersen, Lauren Sage Weinstein, and Paul Copping – examining the multidimensional and interdependent relationship of AI and data.

Having highlighted the interdependence of the former and the latter, the second part explores the current challenges of our current data governance models. This section builds on a recent paper co-authored by Dame Wendy Hall [Four Internets: The Geopolitics of Digital Governance](#).

The third part focuses on what a new type of data governance model could look like. It looks explicitly at the evidence provided by Jeni Tennison (ODI) and Paul Copping (Sightline Innovation) to demonstrate how data trusts could potentially serve as the solution for preventing data misuses while also maximising public and private value derived from AI systems. It also considers the role of new institutions such as the Centre for Data Ethics and Innovation in this context.

Finally, an appendix with copies of the full written evidence provided by Jeni Tennison, Paul Copping, and Lauren Sager Weinstein is also available at the end of the brief.

1. AI AND DATA'S SYMBIOTIC RELATIONSHIP

AI Feeds on Data

Without data, AI technologies would lack the raw material essential for their development and deployment.

AI is a broad set of technologies with the ability to perform tasks that would otherwise require human intelligence and, also, the capacity to learn or adapt to new experiences or stimuli. Although AI technologies can live separate to data, they would be of little value without the right data inputs. AI works best when large amounts of data are available. The more facets that data covers, the faster the AI can learn. This is precisely why many refer to data as the lifeblood of AI.

For example, AI chatbots like Siri and Alexa entering the global markets would have little use if it weren't for the mass amounts of data fed into them, including millions of hours of verbal language that has trained these systems to learn various languages. Likewise, driverless cars would not be able to roam the streets if it wasn't for the data gathered from sensors on existing vehicles.

Data Thrives with AI

Data, of course, has been gathered and used by various actors in society for centuries now. However, the emergence of new technologies and digital platforms have caused a remarkable spike in data generation and collection in recent decades. Nowadays, most of the activities people engage in (during the day and the night) generate amounts of data that were previously unimaginable.

Petabytes of data are being produced at an exponential speed and collected in revolutionary ways, by entirely new actors and through completely different channels. Consider the data generated from contactless travel cards, smartphone applications, wearable health devices, TV streaming, credit card usage, email exchanges, and social media platforms.

Most of this data, however, would be of little use without AI. AI has the power to connect the points between these oceans of data, identify patterns, make accurate predictions and ultimately generate valuable insight able to bring about transformational economic and social benefits.

AI technologies can also create more and more data that can be fed back into the system hence further accelerating its growth.

AI Governance and Data Governance

As AI and data thrive and feed from each other, problems arise when there are critical issues in the governance of either the former or the latter. Simply put, if the data being collected and used has ingrained issues so does the AI. In fact, rather than solving these issues, AI makes them more pronounced.

Therefore, given the codependent and symbiotic relationship of AI and data, AI governance and data governance must adapt a coherent strategy, addressing key areas of growing concern, including ownership, control, misuse, bias, access, and monopolisation.

2. CURRENT DATA GOVERNANCE CHALLENGES

Data Ownership

When it comes to data ownership, it would appear that there aren't any easy answers.

First, unlike most other goods people can own, data is non-rivalrous. One player having access and using data does not prevent, in theory, another from doing so.

Second, the vast majority of data today is generated by different players through a handful of technologies hence it is becoming increasingly difficult to allocate who owns what data. Organisations, for example, no longer rely on just their own, independent processes of collecting data but rather use the data that is generated in interconnected open networks. Data can be used by different organisations in the same industry but also across sectors to increase the value further. Although this means more opportunity, it also means that the task of figuring out ownership and accountability for a specific data set becomes more complex. Furthermore, this creates a true challenge to differentiate between well-founded, clean data and the reverse.

Third, data can be about multiple people. With the aide of AI systems, data about 'me' can now also provide insight about 'others.' Personal data is defined as information that relates to an identified or identifiable person who could be identified, directly or indirectly based on information. However, as Jeni Tennison expressed at the evidence meeting: "Data about one person is about lots of others: their friends and family, their doctors and nurses, their teachers, or even people who are similar to them." Hence it becomes difficult to say that data can be identified to one individual. Rather, it is often connected to several people or organisations.

For these three key reasons, the concept of data ownership is quickly becoming impossible. In consequence, rather than thinking of who owns data, APPG AI evidence suggests we should shift to questions of who controls the data and who benefits from the data.

Data Control and Transparency

Lack of data control and agency are two of the key problems underlying the need for new data governance models. Increasingly so, people (and, also, organisations and communities) are feeling they are unable to control how their data is being collected and used by AI systems in both the private and public sectors. This issue is twofold.

First, organisations are often not as transparent as they could be about how their AI technologies rely on data, what data their AI technologies rely on, and how that data is

collected and used by those AI systems to generate specific output. Few organisations share with individuals or other organisations what they know about them, how that information is used, and in what way that usage impacts them or others. This lack of openness prevents individuals from being able to control their data from the very start.

Second, even if organisations are public about how AI systems collect and use data, the general public knows very little about it. Lack of public awareness means most people have no idea about how their data is being used even if that information is publicly accessible. Currently, most organisations rely on privacy controls which users can choose to 'opt in or opt out' to inform individuals about how their data will be used. Individuals have the opportunity to provide their consent at this stage and hence exercise their control; however, surveys show that the majority of citizens rarely take the time to thoroughly read these lengthy terms. In the rare case they do, the reality that there are very few alternatives providing the same service but with different models of data collection means most people choose to 'opt in' regardless of the conditions. The inability to negotiate the terms data is collected and used by AI systems means users must select over declining or accepting a service.

Professor Birgitte Andersen acknowledged the difficulty of the current models of consent and, hence, pushed policy makers to think of user rights rather than privacy rights. She highlighted: "Consent-based models of data governance put the responsibility on the individual to assess and determine whether an organisation's policies on data use are appropriate. Although, in theory, this respects individual freedom and privacy, in reality, we see there are several flaws engrained within these models. Policy makers must create a user rights framework to help consumers negotiate the fair and ethical way data should be used by AI systems across organisations."

Data Misuse

To add to the issues of data control, there have also been far too many real-world scandals of data being misused and overused, lost, and irresponsibly shared or sold.

The infamous story of Facebook and Cambridge Analytics is only one of many where data was exploited by the hands of people with too much power to influence and manipulate individuals on a global scale.

Data misuse is the inappropriate use of data as defined by global and national law, standards, and corporate laws. APPG AI argues data misuse can also be used to characterise scenarios in which data collection and use fail to take into account ethical implications and norms. Although data misuse has always existed, AI's speed and scale make the impact of this misuse much more alarming.

Data Bias

APPG AI's evidence has also shed light on the biases – deliberate and accidental and sometimes unavoidable – ingrained within the datasets fed into AI systems.

First, embedded within human beings are deep-rooted biases. As a consequence, data sets collected from observations of human behaviour, such as data from social media, is inherently biased. As human beings are responsible for constructing datasets, it follows that the same biases we have deep within us are also transferred through the data. And as this data is increasingly used to build AI, many of the outputs and insight these technologies generate are also, consequently, based on existing injustices and prejudices. Second, the way data is

collected and used may also be biased. There may be technical bias that has emerged through limitations of a program, computational power, design, or other system constraints.

Because humans often think of computers as being logical and rational, decisions made by AI are often considered, by lay users, to be neutral and unbiased. Hence, oftentimes, AI algorithms can inaccurately project greater authority than human expertise, and in some cases, reliance on AI can displace human responsibility for their outcomes. As AI-driven decisions are already playing a critical role in areas concerning our security, health, education, and wellbeing, the issue of biases and stereotypes ingrained within datasets is growing in priority and concern. This is particularly alarming as biases are often difficult to spot and diagnose, requiring a public scandal before being acknowledged and addressed

Data Access

On the other side of data misuse and overuse is the issue of data under-use. Access to data is essential to maximise the public and private socio-economic value promised by AI technologies. Restricting access to data can stop breakthrough innovation from solving some of the most severe problems our societies face. It can be difficult for new entrants to markets to succeed if they don't have access to the data needed to make their technology work.

Many stakeholders across industries and sectors are calling for open access to increased datasets to create a level playing field to develop AI technologies.

Some organisations, like Transport for London, are piloting how data access can be achieved while still protecting the individual from malicious use. As noted by Lauren Sager Weinstein, TfL has released their data through a single point of access, where the data is available subject to appropriate terms and conditions and with controls on privacy. Making data available through safe platforms such as this enables open innovation while restricting potential hazards.

Data Monopolisation

The last challenge highlighted by the APPG AI evidence is the risk of few, wealthy players dominating the entire arena. Few tech companies now own astronomical volumes of data and the power this data generates has created a fear for data monopolisation.

A 'winner-takes-all' model has been created. A small number of companies with huge quantities of data have massive advantage in developing and deploying AI technologies. These few companies have a unique competitive advantage over all others hence making market competition little to none.

It is now clear that while data has become one of the driving forces of the AI revolution, the value it generates is not fairly distributed. Organisations rely on the collection of personal data to build large datasets which they feed into their AI systems. Huge profits are produced from the output of these AI technologies, however that value rarely trails down to the original source: the individual.

3. NEW DATA GOVERNANCE MODELS

Erosion of Trust in Data-Driven Society

The data governance challenges mentioned in the previous section have created public backlash and erosion of trust and confidence in the data-driven economy.

Individuals and organisations are starting to question how data is collected and used; and, ultimately, whether our current data governance models are fit for purpose. Data and AI have already changed our lives tremendously, helping generate great economic and social benefits across the UK. Simultaneously, however, reoccurring data breaches, lack of self-regulation by data-driven organisations, and little ethical judgment in the way day is collected and used are making the need for data governance one of high policy concern.

Governments and regulators are under pressure to clamp down.

In terms of legislative responses, regulations like GDPR and the Data Protection Act are already in place aiming to give individuals more control and protect them from violations of their privacy. Standards are also starting to be formed, pushing for data to be collected and used safely and ethically.

However, given the speed and scale of these transformations, more needs to be done to truly address the challenges unfolding. Ultimately, completely new data governance models - bringing together elements from law, regulation, ethics, and standards - must be built. New institutions like the Centre for Data Ethics and Innovation will be vital in this shift.

Data Trust

One new data governance model that has earned significant attention in the UK and around the world is that of the data trust. The concept of 'data trust' as an institution has existed for decades and is multi-dimensional.

In the UK, the concept first attracted policy attention after the launch of the Dame Wendy Hall's independent review into AI Growing the Artificial Intelligence Industry in the UK. The report recommended data trusts as a way to share data in a fair, safe and equitable way.

For the purpose of the brief, we've adapted the [Open Data Institute's definition](#), explaining the term "data trust" as "a legal structure that provides independent stewardship of data."

Simply put, a data trust takes the legal concept of holding and making decisions about assets such as property or investment and applies it to data. In a data trust, there are three key actors: the trustors, the beneficiaries, and the trustees.

The trustors are the individuals, organisations or other nodes that hold data, and grant some control of the data to a set of trustees. The beneficiaries are those who are provided access to the data and are usually those with direct benefit from what is created using that data. The trustees are legally bonded to make decisions about the data in the best interests of the beneficiaries. In a situation questioning how data was collected and used, trustees must demonstrate they acted in the best interests outlined in a pre-determined charter.

There are different examples of data trusts and various ways they can be structured, from 'bottom-up' to 'top-down.'

The Open Data Institute is currently working with the UK Office of AI to pilot [three data trust pilots](#) across different contexts. The aim of these pilots is to offer flexible and inclusive models that protect individual data control and privacy while also promoting the public value data-driven insight can generate. In one of the pilots, the ODI is working with The Royal Borough of Greenwich and the Greater London Authority to look at whether data collected through the Sharing Cities Programme could be made available in a data trust to bring about benefits for citizens. The second pilot aims to help reduce illegal wildlife trade by making wildlife data from across the world more accessible. The third pilot will look at how food data – particularly data on the nature of food waste and where it ends up – to help people track and measure food waste in supply chains.

Paul Copping, from Sightline Innovation, provided evidence to the group on how the company he now works for has taken the concept of data trusts and established a [technology framework](#) that enables control and sovereignty of data assets between trusted data partners. Arguing that data should belong to and be governed by those who create it, he shared with parliamentarians a new model of a secure data-exchange network and a smart contract system.

Ultimately, for all of these data trust models to work as new forms of data governance we need to ensure their creation and implementation is inclusive and collaborative.

For more information on the concept of 'data trusts,' please refer to NESTA's [The New Ecosystem of Trust](#) or ElementAI's [Data Trusts: A New Tool for Data Governance](#).

The Role of Government in Data Governance

Governments around the world have a key role to play in the data governance ecosystem.

Dame Wendy Hall presented to the group a recent paper she co-authored with Kieron O'Hara – looking at four different versions of governing the internet, each playing a geopolitical role and championed at the national level. The paper is titled: [Four Internets: The Geopolitics of Digital Governance](#).

In the paper, the two authors identify four versions of the internet that have evolved: (1) the Silicon Valley version based on openness and transparency, (2) the 'bourgeois' version, popular in Europe, where privacy is protected, (3) the authoritarian internet with surveillance is prioritised for social cohesion and security, and (4) the 'Washington DC' version thinking of online resources as private property, whose owners can monetise them and seek market rates

for their use. Most governments around the world, according to Dame Wendy Hall, have adopted a form of either of these four models.

The data governance approach is strongly connected to which of the four models of internet governance has been adopted.

Data trusts, or other new data governance models, could potentially be the solution to the challenges of some of these approaches. Paul Copping argued: "The data trust protects the individual from the authoritarian and commercial, providing a healthy balance of open and bourgeois - so enabling innovation without compromising on privacy."

Roger Taylor explained to APPG AI the significant role institutions like the Centre for Data Ethics and Innovation can play in the creation and implementation of new data governance models such as data trusts.

Recently, the Centre for Data Ethics and Innovation launched their Work Programme 2019/20 explaining exactly how the institute will operate and what they will be focusing on. The main purpose of the institute will be to analyse and anticipate the opportunities and risks posed by data-driven technology and put forward practical and evidence-based advice to address them. The first two reports they will be focusing on will be on 'online targeting' and bias.

In conclusion, APPG AI's evidence highlighted the need for policy makers to build public awareness and to ensure as great number of stakeholders as possible engage with the process of building new data governance models. The wider society must be informed of the many implications and pilots of new data governance models must ensure they promote public good and respect human rights simultaneously.

WRITTEN EVIDENCE APPENDIX

Jeni Tennison, CEO, Open Data Institute

Biography: Jeni Tennison is the CEO of the Open Data Institute. She gained a PhD in Artificial Intelligence, then worked as an independent consultant specialising in open data publishing and consumption. She was the Technical Architect and Lead Developer for legislation.gov.uk before joining the ODI as Technical Director in 2012, becoming CEO in 2016. Jeni sits on the UK's Open Standards Board; the Advisory Board for the Open Contracting Partnership; the Board of Ada, the UK's National College for Digital Skills; the Co-operative's Digital Advisory Board; and the Board of the Global Partnership for Sustainable Development Data.

Written Evidence

How is data currently being collected and used by AI technologies, and is our data infrastructure fit for purpose?

- We wrote a report on AI business models last year. This emphasised the challenge of large tech companies determining what gets collected, and exclusive access to that data. It discussed the need to share data to increase the visibility of and reduce bias, and to support innovation, new entrants and new services.
- Data infrastructure consists of data assets, the organisations that operate and maintain them, and guides describing how to use and manage the data. Trustworthy data infrastructure is sustainably funded and is directed to maximise data use and value, meeting society's needs.
- We're all still building a data infrastructure that's fit for purpose. Right now there are issues with, among other things, quality, interoperability, accessibility, reliability, equity and sustainability.

What is the proper way of seeking consent for using an individual's data?

- We should be wary of talking about "an individual's data". Data about one person is about lots of others: their friends and family, their doctors and nurses, their teachers, or even people who are similar to them. Data is collected about people who aren't direct users of services as well as those who are.
- Individual consent is important but not the only mechanism for getting consent for the use of data. There are a number of cases where we have decided democratically that data should be collected and used whether individuals consent or not: the census and registration of company directors are examples.
- There are also challenges around existing approaches to consent. We can never give fully informed consent about future uses of data. Individual consent needs to be layered over fundamental regulatory protections. It is also worth looking at the roles of trusted intermediaries who may be granted delegated responsibility, and participative approaches such as citizen juries to help organisations understand public acceptability.

Who should ultimately oversee the use of personal data for AI systems?

- Many organisations are putting in place internal ethical principles and practices. It is good to see data ethics being recognised as important within businesses. However,

- these mean very little without some mechanism for accountability when ethical principles are violated.
- One layer of oversight has to come from regulators. These include sector specific regulators such as in finance or telecoms, as well as the ICO. They include local government when there are devolved powers such as around transport or accommodation as well as national and potentially multinational regulators.
- We should not forget the role of consumer groups, trade unions, professional bodies or the media. These are important in other places where there is a need for accountability, consumer protection, or citizen empowerment; they are no less important when data is involved.
- Both regulators and civil society require some level of transparency, monitoring frameworks, whistleblowers, auditing powers and so on to enable them to do their job.

Lauren Sager Weinstein, Chief Data Officer, Transport for London

Biography: Lauren Sager Weinstein is chief data officer at Transport for London (TfL), where she is responsible for driving decision-making and improved customer services through provision of data products and services. Lauren oversees TfL's business intelligence / data science strategy, TfL's data technical platform development, as well as support for data products and services. Lauren has worked for TfL in various roles since 2002 and before this she was employed by the City of Los Angeles and the RAND Corporation, a non-profit institution that helps improve policy and decision making through research and analysis.

Written Evidence

We see that AI as an evolution in the way that we can use data to inform decision-making. We have a long history of making the best use of data and computer processing power available at the time to deliver analysis to drive our decision-making. We at TfL have used analysis and statistics for many years—starting with calculations that we could do by hand. For example, we were surveying people and manually counting them in order to plan our network as far back as 1939. Today we utilise the computing processing power available from cloud computing to collect and process our telemetry data to support the day to day running of our network to deliver essential services for our customers.

- We have become more sophisticated over time, and where we could we have looked to automation and using new computer technology to plan our network. For more than a decade our Oyster ticketing system has allowed us to construct patterns from depersonalised ticketing to understand our customers' journey flows to better plan and run our network.
- What used to take considerable computing even a decade ago, can now be calculated much faster. And we have employed new pattern analysis techniques now that the computing processing can support this. This includes taking advantage of new techniques such as machine learning and deep learning.
- As we as an industry move forward, we expect the technology to improve, allowing for even better data tools. However, underpinning this we need to make sure that our underlying data is fit for purpose. This means that we as an industry need to make sure our data foundations are strong, through understanding our data quality and the way in which it is collected and stored.
- Transparency is fundamental on this. Our approach has been to make sure that we are clear with our customers about how we use data and the benefits it can deliver.

Our broad aim is to understand patterns in how our customers move in aggregate. We give our customers the option to not be known to us when we are collecting personal data. And we treat our customers' privacy extremely seriously, when we do need to support them through delivering a personalised service (such as refunds, or when they ask to get service information from us).

- We rely heavily on the ICO's guidance, which underpins our approach. This assistance has been very helpful for us.

We recognise that the industry is evolving and changing, and as society, we need to all think about the ways that data is used, and potential and risks across society for this. There is a role for philosophical/ ethical discussions about this as a society, in terms of how we set guiding policies to protect our citizens, whilst delivering improvements for them and the wider community.

Paul Copping, International Advisor, Sightline Innovation

Biography: Paul is currently an international advisor to a Canadian AI company called Sightline Innovation. Sightline is a market leader in data trusts and participates in the work of the Centre for International Governance Innovation (www.cigionline.com) founded and chaired by Jim Balsillie, ex CEO Blackberry / RIM. Paul recently ended a four year assignment as Chief Innovation Officer at Digital Greenwich and during that period he chaired an ETSI Industry Specification Group on City Digital Profile. Paul also has experience as a senior executive in Nortel Networks, another Canadian flag carrier.

Written Evidence

On this Data Protection Day (28th January) I am calling for the UK Government to promote and adopt personal data trusts at scale and give citizens the means to have more control of their data.

We routinely share online data about ourselves. We click "YES" to accept cookies and (mostly unread) privacy policies. We share our card details, health and fitness data, financial records, location and other less intentional metadata. We all have the option under Article 17 of GDPR to take all this personal data back and be forgotten – known as the right to erasure. No one knows how much this right has so far been exercised, but it is by no means working at an industrial scale.

Valuable insights can be derived from fusing multiple data sources. For example, a trusted health advisor could bring together our clinical records, our exercise data and maybe even our DNA to provide deep insights into our well being and the life trajectory we can expect. NHS Digital would welcome our increased accountability for our own wellbeing. Similarly, a commercial aggregator could predict our life time value to a given retailer. More hostile aggregators could use their insights against us.

How does a personal data trust make this data model more efficient and secure?

The core element is that we provide our consent for the use of our data to the Trustee. The Trustee manages our data according to our agreed policies which govern how it interacts with all our data users.

We provide the data only once. Any subsequent changes (such as a change of address) can be updated once and accessed by all our data users. The Trustee can exercise the Right to Erasure on our behalf and control who has what visibility of our data.

The very helpful and timely concept of the "Four Internets" * – which I hope its co-author Dame Wendy Hall will expand today – sets out the various strengths and weaknesses of different geopolitical approaches in a shorthand of so called open, bourgeois, authoritarian and commercial regimes. The data trust protects the individual from the authoritarian and commercial, providing a healthy balance of open and bourgeois - so enabling innovation without compromising on privacy.

This model also gives individuals the right to be forgotten if they wish. There remains an open issue, deep in the domain of the Ada Lovelace Institute. Can it be ethical to use AI from social media or surveillance systems to support people at risk of drug addiction, homelessness or suicide? Is it reasonable, as happened only last week, to request that depressing or suicidal materials on social media are hidden from people judged to be depressed?   When should our right to privacy be trumped by society's duty of care?

 * "Four internets – The Geopolitics of Digital Governance" by Kieron O'Hara and Wendy Hall, Centre for International Governance Innovation (CIGI), December 2018

Birgitte Andersen, CEO, Big Innovation Centre

Biography: As CEO and CoCreator of the Big Innovation Centre, Professor Birgitte Andersen (PhD Economics) is leading a London based open innovation and investment hub of a dozen global companies, world class universities and public organizations. Birgitte brings a wealth of experience across the innovation ecosystem. During the dot.com boom she set up and ran for 10 years (2000 to 2010) probably London's biggest interdisciplinary postgraduate programmes (across the Departments of Management, Economics and Computer Science) on E-Commerce and Business Innovation at Birkbeck College, University of London, where she is also Professor of the Economics and Management of Innovation.

Written Evidence

Ethics in an AI world is righty about fairness, transparency, trust, privacy and protocols. But data capitalism is much broader, depending on where you stand:

A. **Law** - about moral rights to own your data record. The core focus is on mainly exclusive individual rights.
B. **Economics** – about growing the value-pie in the economic system through the new growth engine. Think of big data as the new oil. Here ethics is about economic based rights to the value pie and to get a fair share. It is also about not blocking the growth of the pie by refusing taking part. The focus is on community rights and often user rights.
C. **Corporate governance** – about the Corporate Social Responsibility (CSR) around privacy and use. The focus is on business ethics.

Now there is a lot of focus on Data Ethics regarding the legal perspective, (i.e. (A) above), but to grow the value pie, from which users access economic and social value such as better health and more sustainable smarter homes, we have to enforce some kind of user-rights.

Also, while there rightly is a lot of focus on the Tech Giants' tax practices (in particular Base Erosion and Profit Shifting) and the Chancellor of the Exchequer increase in 2% tax on major platforms which transact more revenue than many entire countries, I strongly feel that the solution is data governance and creative commons.

We now opt-in (provide user-rights) into all the Tech giants one by one while we receive their services, but if we instead shared data into a creative commons pool which entrepreneurs and innovators can access, we will have competition on the best product or service offering through their data analytics. Taxing them higher may change the allocation of resources and profit in the short run, but not competition. Data governance can impact this and long industry and social effects.

This will also mean that we do not just have one platform in each sector – one for friends (Facebook), one for auctions (Ebay), one for photos (Instagram), one taxi firm (Uber) and one office (Microsoft Office), one bookshop (Amazon), one professional network (Linked-in) and so forth.

These are my 3 key points I'd like Government to address:

1. Community rights to ethical and secure data sharing (Creative Commons as a model) to the growth the value pie
2. User rights to data to access value
3. Competition is about data governance

ABOUT APPG AI

The [All-Party Parliamentary Group on Artificial Intelligence (APPG AI)](#) was set up in January 2017 with the aim to explore the impact and implications of Artificial Intelligence.

The APPG AI is co-chaired by Stephen Metcalfe MP and Lord Clement-Jones CBE.

The Group Officers are Chris Green MP, The Right Reverend Doctor Steven Croft, Baroness Kramer, Lord Janvrin, Lord Broers, Lord Holmes of Richmond, Lord Willetts, Baroness McGregor-Smith, Justin Madders MP, Mark Hendrick MP and Carol Monaghan MP.

The Group supporters – Accenture, Barclays, British Standards Institution (BSI), CMS Cameron McKenna Nabarro Olswang (CMS), Creative England, Deloitte, Ernst and Young (EY), KPMG, Microsoft, Osborne Clarke LLP, PwC, and Rialto – enable us to raise the ambition of what we can achieve.

Big Innovation Centre is the APPG AI Secretariat.

All Party Parliamentary Group on AI (APPG AI)

www.appg-ai.org