

26 June 2017 – 5:30pm GMT
House of Lords, Committee Room 4A



**BIG
INNOVATION
CENTRE**

Ethics and Legal in AI: Data Capitalism

A theme report based on the 3rd meeting of the **All-Party
Parliamentary Group on Artificial Intelligence [APPG AI]**.

Ethics and Legal in AI: Data Capitalism is a theme report based on the third meeting of the **All-Party Parliamentary Group on Artificial Intelligence (APPG AI)** - held on 26 June 2017 at the House of Lords.

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

The evidence presented in the report is not exhaustive but reflects what was discussed at the meeting, and the views and experiences put forward by the people giving evidence. Written submissions by individual expert advisors in relation to this meeting are also included.

The APPG AI was established in January 2017 and its officers include:

- **Stephen Metcalfe MP**- Co-Chair
- **Lord Tim Clement-Jones**- Co-Chair
- Chris Green MP- Secretary
- The Rt Rev Dr Steven Croft-Bishop of Oxford- Treasurer
- Baroness Susan Kramer- Vice Chair
- Lord Robin Janvrin- Vice Chair
- Lord Alec Broers- Vice Chair
- Mark Hendrick MP- Vice Chair
- Carol Monaghan MP- Vice Chair

For further information on the APPG AI, contact:

All Party Parliamentary Group on Artificial Intelligence (APPG AI) Secretariat
Big Innovation Centre | Ergon House, Horseferry Road, London SW1P 2AL
T [+44 \(0\)20 3713 4036](tel:+442037134036)
Email: appg@biginnovationcentre.com

Any enquiries regarding this publication, contact:

CEO and CoCreator: Professor Birgitte Andersen
APPG AI Rapporteur: Niki Iliadis (Business Intelligence Researcher)

Ethics and Legal in AI: Data Capitalism

A theme report based on the third **All-Party Parliamentary Group on Artificial Intelligence [APPG AI]** Evidence Giving meeting.

26 June 2017 – House of Lords, Committee Room 4A



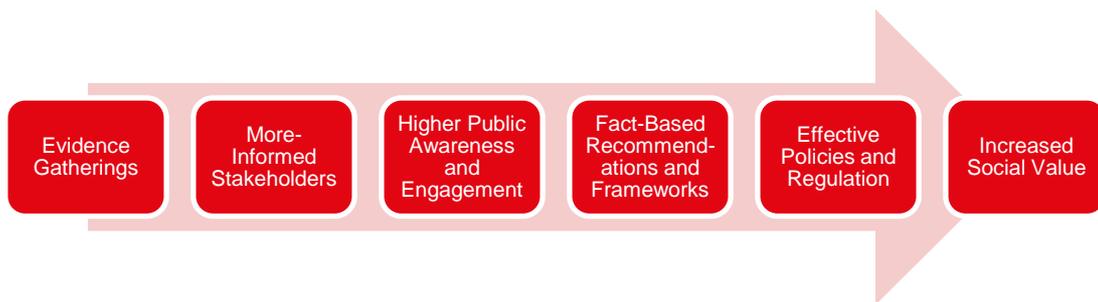
All Party Parliamentary Group on
Artificial Intelligence

Overview

The All-Party Parliamentary Group on Artificial Intelligence [APPG AI] was created in January 2017 to explore the impact and implications of Artificial Intelligence, including Machine Learning. We aim to: **to unpack the term, to gather evidence to better understand it, to assess its impact, and, ultimately, to empower decision-makers to make policies in the sphere.** Government, business leaders, academic thought leaders and AI entrepreneurs have come together in an effort to share evidence and beliefs, and assist in setting an agenda for how the UK should address AI moving forward.

Figure 1 illustrates the process of how APPG AI aims to contribute to increasing social value, through fact-based recommendations and well-informed stakeholders.

Figure 1. The Purpose of APPG AI



The first APPG AI Evidence Giving meeting approached Artificial Intelligence through a general lens, identifying the key issues within the umbrella term that stakeholders should focus on. The second APPG AI Evidence Giving meeting deep dived into issues regarding decision-making and AI - and addressed controversial questions such as: 'how can AI assist the decision-making process?' and 'where should we draw the line?'

The aim of the third APPG AI Evidence Giving meeting centred on the ethical and legal dimensions in AI, particularly in regards to data capitalism.

The meeting was co-chaired by Stephen Metcalfe MP (former Chair of the House of Commons Science and Technology Select Committee) and Lord Tim Clement-Jones (Chair of the House of Lords Select Committee on Artificial Intelligence). Six experts were invited to provide evidence reflecting their **views on data collection, use, management, and governance in a world of emerging AI technology.**

The third Evidence Giving meeting was held at 5:30pm on 26 June 2017, at Committee Room 4A in the House of Lords.

93 TOTAL PARTICIPANTS

6 Pieces of Oral Evidence

7 Pieces of Written Evidence

The panel included: **Sir Gordon Duff** (Principal at St.Hilda’s College – University of Oxford), **Professor Mark Skilton** (Professor of Practice, Information Systems Management and Innovation at the University of Warwick Business School), **Nicola Eschenburg** (Global Head of Analyst Relations at BAE Systems), **Stewart Room** (Partner and Global Data Protection Leader at PwC), **Clive Gringras** (Head of Technology, Media and Telecommunications at CMS Cameron McKenna Nabarro Olswang LLP), and **Dr. Maria Ioannidou** (Lecturer in Competition law at Queen Mary University London).

The group acknowledges the urgency to rethink data governance. There were two alternative approaches offered on how the UK should change the way data is collected, used, and managed. The first approach – the hard approach – called for a change in legislation. The second approach – the soft approach – called for the use of soft-structures (new and old) to address the issues of data capitalism.

This report is divided into five sections in which we discuss the transformation of data governance, consider different approaches moving forward, and, lastly, make some pragmatic and urgent recommendations.

Theme	Description
1. Data is AI’s key ingredient	Data is now one of the most valuable resources. It is now more than a knowledge asset and has become a key commodity worldwide. It is the key ingredient for AI. Data fuels AI technologies and, in turn, AI technologies curate more data.
2. Data collection, use, and management are vastly transforming	Throughout history, data has always been collected, used, and managed by different actors for different purposes. However, we have now entered a period of data capitalism – in which the volume and velocity of data has skyrocketed. The create use of data has completely transformed industries, routines at work and the way we live our lives. Social notions are with people, machines, and software.
3. Data governances needs a hard approach?	One view is that current legislation is outdated considering the transformations that have occurred. This report summarize the evidence of a call for a change in laws and policies to unlock the value that can be created with data in the AI era. Could a Data Charter setting the guidelines on how data are used create a trusted environment in which personal and business data are shared? Could this both empower people, entrepreneurship and set international standards for trade?
4. Data governance needs a soft approach?	Another view is that soft-structures are enough to unlock the value data can create in the AI era. On this front, the report explores several issues: Could new routines and more enforcement around privacy and consent unlock data and value creation in new

business models? Is the core issue lack of transparency or lack of skills or other?

5. Data governance needs an evidence-based, guidance framework

Two critical steps should be taken moving forward: **(1) An analysis of the current legislation landscape in place and (2) The creation of a set of standards on how data should be collected, used, and managed with social purpose.**

This theme report is not research-oriented but aims to summarize these key themes, using the evidence gathered at the third APPG AI evidence meeting (details above). It is not exhaustive but reflects what was discussed at the meeting, as well as the views and experiences put forward by the people giving evidence. Written excerpts by individual expert advisors in relation to the meeting are also included.

Event Summary

Following a short break leading to the General Election, the APPG on Artificial Intelligence held its 3th Evidence Meeting end of June. Co-chair Stephen Metcalfe MP welcomed back the group and focused the day's agenda on data capitalism, and the ethical and legal issues that are associated with it.

Sir Gordon Duff was first to speak amongst the six experts on the panel. Former Chairman of the Medicines and Healthcare Products, his particular interest and experience is in innovative technology in the health sector. He was responsible for implementing the largest-well curated clinical databases worldwide, holding records of over 20 million people. He argued that data has the potential of having huge positive impact in the health industry and most people are willing to donate their data for social purpose. The question government needs to address is how to raise public confidence about one's safety and security regarding data.

Mark Skilton, Professor at the University of Warwick Business School who has written 3 research books on the impact of AI, was next to speak. He talked about the key findings from his latest book, looking at how the 4th Industrial Revolution (4IR) is impacting business. He identified three main trends with AI: (1) large data sets creating neural networks that model human behaviour, (2) automating machine behaviour augmenting some types of work, and (3) AI being used to monitor and respond to cyber-attacks. He called for a data charter with sensible guidelines to help stakeholders manage data and figure out "where effective jurisprudence is going on."

The third panellist was **Nicola Eschenburg** from BAE Systems Applied Intelligence. Ethics and legal concepts are both self-enforced and politically-enforced, and are flexible and changeable depending on external variables. There is growing concern, she said, around who collects, owns, and uses data because we are shifting from a period of financial capitalism to data capitalism. According to her, regulation should be kept simple to reflect this fast-moving environment.

Dr. Maria Ioannidou from Queen Mary University agreed with Eschenburg, suggesting stakeholders should first analyse the law currently in place before creating new ones. Data is considered one of the most valuable resources and is changing business models worldwide. Yet, consumers lack understanding about how and when their data is being used. It is important to educate the public and empower them to understand how their data is being used, shared, and managed in exchange for a specific product or service. We should not forget, she reminds the group, that there is collective responsibility in every transaction.

Fifth to speak was **Stewart Room** from PwC. He has a background in law and now focuses mostly on what businesses are doing with their data. He argues that government should be cautious when changing existing legislature or creating new law. He believes that soft structures are the solution to the ethical issues related to data capitalism. Data needs to be put in the broader agenda to raise public awareness on the issue.

Lastly, **Clive Gringras** – Head of Technology, Media and Telecommunications at CMS - took the floor. He agreed with Room that "law should be incrementally slow when it comes to technology." He argued that change in legislature would only create definitional and jurisdictional arbitrage, as well as threaten the rights of human beings. Clive questioned whether GDPR would prevent society from reaping the full benefits of emerging technology.

He believes that data is the fuel that drives engines of AI and that current IP systems in the UK tighten the ability to access this data. Hence, he also called for the loosening of intellectual property rights.

Co-chair Lord Tim Clement-Jones thanked the panel and summarized the conversation, spotting two main outlooks: the need for more law to reflect AI and big data versus the need for soft structures to promote ethical conversation. Justin Madders MP questioned the panel about the role of privacy and how much of the public they believe is knowledgeable when giving out consent for using their data. The panel agreed that this is an awareness and educational issue and that government can play a key role in informing the citizens how their data is being collected, used, and managed.

Overall, most of the group concluded that the UK needs to be careful in the legislation we craft moving forward. AI can have huge benefits for society and laws might prevent society from reaping its full benefits. Hence, we need to collect more evidence and build use cases in order to act responsibly and strategically. New soft structures that educate the public and build trust in society are one solution for addressing the issue.

Table of Contents

Overview	4
Event Summary	7
1. Data is AI's key building block	10
2. Data collection, use, and management is vastly transforming.....	14
3. Data governance needs a hard approach?	19
4. Data governance needs a soft approach?	23
5. Data governance needs an evidence-based, guidance framework.....	26
Acknowledgements.....	30
Contact details	32

1. Data is AI's key building block

In the constantly increasing discussions on AI, there is much debate on how AI is impacting society economically, socially, and environmentally - and what this will mean for our futures. There are two general statements that all seem to agree on.

- **First, AI is transformative.** Governments worldwide have identified AI as one of the most disruptive forces of the decade.
 - Some national/international initiatives include: (a) The UK has put forward a major review this year into what AI means for the country and has established various committees including the Lord's Select Committee on Artificial Intelligence and the APPG AI to gather evidence on AI impacts; (b) China has recently announced their 3-phase strategy of how to become the AI leader by 2030¹; (c) the US has prepared three in-depth reports on how to prepare for the Future of AI²; (d) Canada has announced a \$125 million funding for a Pan-Canadian AI Strategy³; and (e) in December 2016, the UN has voted to explore issues linked to automated weapons.
- **Second, AI fuels on data.** Without data, AI technologies would lack the raw material essential for their designing, development, and sustainability. Furthermore, AI creates more and more data that can feed back into the system and further accelerate its growth. Hence, the relationship becomes symbiotic and self-reinforcing (Figure 2).

The community built around the APPG AI recognises AI's dependency on data and, in consequence, organised the third Evidence Giving meeting to focus on today's data-driven environment, referred to by many as "data capitalism."

As Sir Gordon Duff (current Principal at St. Hilda's College of Oxford and former Chairman of the Medicines and Healthcare Products Regulatory Agency) noted in the meeting, "**it comes to no surprise that almost all conversations on AI quickly go back to data.**" The two concepts are closely interlinked, and it would be fair to say there wouldn't be the former without the later.

Siri, Alexa, or other chat-bots entering the global markets would have little use if it weren't for

¹ Kharpal, Arjun, *China wants to be a \$150 billion world leader in AI in less than 15 years*, CNBC, (21 July 2017). <https://www.cnbc.com/2017/07/21/china-ai-world-leader-by-2030.html>

² The U.S. Executive Office of the President prepared the "Artificial Intelligence, Automation, and the Economy" report in December 2016, the National Science and Technology Council Committee on Technology prepared "The National AI Research and Development Strategic Plan" in October 2016, and together they drafted the "Preparing for the Future of Artificial Intelligence" Report in October 2016.

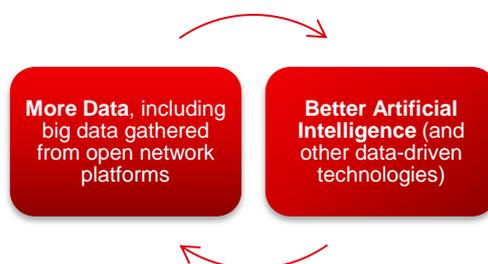
³ Pan-Canadian Artificial Intelligence Strategy Overview, CIFAR, (30 March 2017). <https://www.cifar.ca/assets/pan-canadian-artificial-intelligence-strategy-overview/>

the mass amounts of data fed into them, including billions of hours of verbal language that has trained these systems to learn various languages.

The Facebook AI Research (FAIR) would not be able to deploy the revolutionising new technologies and research it has launched if it didn't have access to the data provided by over 2.01 billion monthly Facebook users.⁴

Driverless cars would not be able to roam the streets if it wasn't for the data gathered from sensors on existing vehicles. According to a study at MIT, a modern car generates 25 gigabytes of data every hour.⁵ This data is exactly what companies such as Google and Tesla used to build and program automated cars. It is important to note, that automated vehicles – and other emerging technologies for that matter – will only cultivate more and more data. For example, it has been reported that Google's automated car can generate 750Mb of data per second.⁶

Figure 2. Symbiotic Relationship of Data and AI



Altogether, the panellists at the APPG AI third **Evidence Meeting** emphasised the fact that we are living in an increasingly data-driven world. Data has given rise to a plethora of AI technologies able to perform tasks that, till recently, most of us only believed only existed in science-fiction movies.

According to the report published by the House of Common's Science and Technology Committee in April 2016⁷:

- **The value of BIG DATA to the UK economy from 2012 to 2015 is £216 billion.**
- **£520 million has been spent by government in big data infrastructure since 2011.**

⁴ "The Top 20 Valuable Facebook Statistics – Updated August 2017" Zephoria Digital Marketing (1 August 2017). <https://zephoria.com/top-15-valuable-facebook-statistics/>

⁵ Barr, Julie, "Cars, data, and the Internet of Things," *Slice of MIT* (20 October 2016). <https://slice.mit.edu/2016/10/20/cars-data-and-internet-of-things/>

⁶ Wells, Peter, "Automated cars and data," Hackernoon (17 November 2016). <https://hackernoon.com/automated-cars-and-data-786dfb1e3eb4>

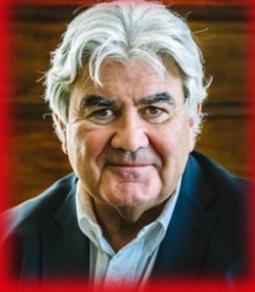
⁷ "The Big Data Dilemma" Science and Technology Science Committee (April 2016). <https://publications.parliament.uk/pa/cm201516/cmselect/cmsctech/992/99204>

- There are over **20,000 datasets** published on data.gov.uk.
- **90%** of all data available on the Internet was created less than two years ago.
- The NHS could save **£16 to £66 billion** if they employed data analytics better.

The numbers above highlight the growing importance of data for society – in respects to the economy and other ethical, environmental, and social considerations.

The benefits that society can reap from these data-driven technologies are enormous. We are already seeing improvements in health care treatments, financial services, cost-saving business practices, and much more.

Sir Gordon Duff noted that **‘AI is one of the areas the entire UK economy will soon be dependent on.’** In his post as Chairman of the Medicines and Healthcare Products Regulatory Agency, he was responsible for implementing the largest well-curated clinical database (holding records on over 20 million people) and shared with the group the life-saving opportunities that have been introduced because of the availability of this valuable data.



Sir Gordon Duff
Principal of St. Hilda's college in University of Oxford and
Former Chairman at the Medicines and Healthcare Products
Regulatory Agency

Data is incredibly important for the NHS, and the wider healthcare sector. One of the ambitions of NHS is to capture the large quantities recorded within various health services.

The first step is to use data to produce algorithms. The second step is to use data rights to seek patents for these algorithms. The third step is to apply these algorithms for large health-gains across the population.

Many questions need addressing in order for society to benefit from this data.

- How do we develop the security needed to bring public confidence about the safety of one's data?
- Can data be useful if it is anonymised?
- If a computer predicts something for me using my data, to whom does this information belong to?
- Who should have access to this information?

Although people are conscious about how data is transforming social notions such as privacy, there

is also an ethical negative about not using data for social purpose.

Still, like most new disruptive forces, data-driven technologies also have a set of risks. There are complex issues (i.e. privacy, consent, impact on democracy, accountability, transparency, etc.) embedded within this data-driven society that should be addressed forthwith.

Facing this challenge, the APPG AI, in the third **Evidence Giving meeting**, chose to explore the benefits that could arise from data-driven AI but also shed light on its potential risks. The discussion acknowledged the **transformation of data collection, use, and management in the last decade and concluded in an urgent need for change in data governance**. The attendees proposed two alternative approaches, as illustrated in Figure 3, for how data can be governed moving forward.

Figure 3. The Two Approaches to Rethinking Data Governance

The Hard Approach

- implying **a reform in current legislation** to reflect the social changes data capitalism and AI have brought about. (Section 3)

The Soft Approach

- implying the **use of soft-structures to better inform and educate the public** on data-related issues. (Section 4)

2. Data collection, use, and management is vastly transforming

Before unpacking the two different approaches, it is essential to have a basic understanding of how data collection, use, and management has transformed in the recent decades.

Data has been gathered and used by various actors in society for centuries now.

Fast-forwarding to now, the emergence of new technologies and digital platforms has caused a remarkable spike in data generation and collection.⁸ Nowadays, most of the activities people engage on in their daily routines (and also while sleeping) generate amounts of data that were previously unimaginable. Only consider the data generated from contactless travel cards, smartphone applications, wearable health devices, TV streaming, credit card usage, email exchanges, and social media platforms.

Based on an article featured on AnalyticsWeek in March 2017, some impressive numbers on data are⁹:

- More data has been created in the past two years than in the entire history of humanity.
- We perform 40,000 search queries on Google per second, totalling to 1.2 trillion searches per year.
- Every minute up to 300 hours of video are uploaded to YouTube.
- Data production will be 44 times greater in 2020 than it was in 2009.

Put simply, petabytes of data are being produced at an exponential speed and collected in revolutionary ways, by entirely new actors and through completely different channels.

Nicola Eschenburg (Global Head of Analyst Relations at BAE Systems Applied Intelligence), speaking on behalf of Kevin Bailey (Vice President – Applied Intelligence at BAE Systems Applied Intelligence), highlighted the **growing concern for who collects, owns, and uses data.**

⁸ Marr, Bernard. "A Brief History of Big Data Everyone Should Read." LinkedIn. (24 February 2015). <https://www.linkedin.com/pulse/brief-history-big-data-everyone-should-read-bernard-marr>

⁹ Kumar, Vishal. "Big Data Facts." AnalyticsWeek. (26 March 2017). <https://analyticsweek.com/content/big-data-facts/>



Kevin Bailey

Vice President – GTM Strategy at BAE Systems, responsible for proposing and overseeing its 5 year transformation strategy for Cyber Security and Financial Fraud & Compliance.

The meaning of "**ethics**" is hard to pin down, and the views many people have about ethics are diverse.

First, ethics refers to well-founded standards of right and wrong that prescribe what humans ought to do, usually in terms of rights, obligations, benefits to society, fairness, or specific virtues. **Second**, ethics refers to the study and development of one's changing ethical standards. Citizens will continue to find loopholes in ethics to align to their cause

Legal; once more we have ambiguity in a word! Legal can mean 'allowed by law', 'correct or acceptable according to a law or rule' or 'relating to the law'. However, we've had the ability to challenge the position of 'something's' status as 'Legal', in the law courts since the 18th century, using 'Lawful methods of dissent'.

The Irish DPC is challenging the use of 'Standard Contract Clauses' by Facebook et al., using the '2013 Snowden disclosures about U.S. government surveillance programs' as a 'Lawful method of dissent' to ensure critical safeguards, protecting EU citizens data' in the US are maintained

Ethics and Legal are both self-enforced and politically enforced, whilst also being flexible and challengeable that increases the opportunity of confrontation, dispute and change

Artificial Intelligence without 'data' is like a car with no petrol; Morecombe without a Wise; Fish without any Chips; or Stephen [Metcalf] without a Tim [Clement-Jones] and Brigitte, It just doesn't work!

Artificial Intelligence has caught the eye of the consumer technology market, but we need to pay increasing attention to the secondary consequences of these connected devices, specifically how it 'Concerns who **Collects**, **Owns** and **Uses** Data'.

We all know that Data are increasingly determining economic value, transforming from an era of financial capitalism into one of **data capitalism**.

Applying the simple 'six degrees of separation' logic we know have to think of Kevin Bacon as a Fridge (via its IP address), accessible to multi-Undecillion of devices on and above earth.

Regulations to oversee data capitalism associated with AI needs to be kept simple, not mis-interpretable by citizens who will change their ethical position, adopt periods of 'Social Cooling'¹⁰ whilst

¹⁰ **Social Cooling** describes the long-term negative side effects of living in a big-data-driven reputation economy.

protected by Human Rights Act. Compartmentalisation of B2B2C2C2B interactions will minimise heavy handedness and embrace collaboration and innovation.

If we fail to appreciate differentiation, how do we regulate AI in 2020 when - four billion citizens have access to the internet, each owning - 5.2Terabytes of data, accessible via - 340 Undecillion (34×10^{36}) IP addresses across diverse devices, capable of contributing towards the negativity of AI modernisation?

A citizen driven world requires corporations and government to protect their needs, especially when the machines do not align to a citizen's definition of ethics. Obviously omitting the onus on the citizen to accept a level of data controls and responsibility themselves!

Would simplicity be aggressive enough to quench the citizen's thirst for advancement? And what should developers consider?

- The 'opt-in opt-out' challenges with the unpredictable ABC deviations that humans can influence in a nanosecond;
- An approach of innovation from the belief of a new age of civility and transparency as depicted by Dave Eggers in his 2013 novel 'The Circle'¹¹;
- Do developers take ethics and legality out of the conversation, targeting demographic groups only and not go down to the 'individual level'?

AI worries citizens, as they think they are losing control hence governments and regulators will be under pressure to clamp down, especially as law, regulation and ethics are the human constructions we put in place in order to restrict the power of wealth and privilege.

We can use data to create a far smarter world without sacrificing precious rights. If we truly believe in such a benign future, we had better hurry up and invent it. It's upon all of us here and our peers, to act rather than 'naval gaze'. Given the speed of technology evolution, there will never be such a thing 'as a right time' again; the only right time is the present time to effect positive change.

A recent report published in June 2017 by the British Academy and the Royal Society discusses how the data life cycle has changed in the last few years, identifying four key trends (Figure 4) contributing to the increasingly complex data environment.¹²

Figure 4. Four Key Trends of Data Capitalism (source: the British Academy and the Royal Society)



The six panellists seemed to agree with the findings from this report, also noting how data

¹¹ The Circle, Dave Eggers 2013 <https://www.theguardian.com/books/2013/oct/12/the-circle-dave-eggers-review>

¹² "Data Management and use: Governance in the 21st Century." A joint report by the British Academy and the Royal Society. (June 2017).

trends are completely different nowadays compared to a few decades ago. As mentioned, **data is generated at remarkably new volumes and velocities**. In fact, as noted by several attendees at the meeting, much of the time the public is not consciously aware of the data they are generating by simple tasks such as using a particular smartphone application or paying with a loyalty card. Justin Madders MP questioned, in the meeting, how informed people are when giving out consent for use of their data.

Furthermore, data is collected, used, and managed in revolutionary ways. Organisations no longer rely on their own, independent processes of collecting data but rather use the data that is generated in interconnected open networks. As a result, this means access to a lot more data and, hence, a lot more opportunity. Data can be used by different organisations in the same industry but also across sectors to increase the value further. Most companies, at the moment, take personal data and anonymise it in order to subsequently apply it for various value-generating processes.

However, this also means that the task of figuring out ownership and accountability for a specific data set becomes more complex. **It becomes a true challenge to differentiate between well-founded, clean data and the reverse**. Professor Skilton commented that we need to address the issue of scaling up. Data is collected for one purpose, used by an algorithm for a different purpose, and often reused by another algorithm for a completely alternate purpose.

Next, there is much conversation on the biases and stereotypes ingrained within datasets. We need to remember that humans construct data and, hence, the same biases we have deep within us are also transferred through the data. Consequently, if this data is increasingly being used to build AI that helps influential people make big decisions, there is a risk that these decisions might be based on existing injustices and prejudices in our cultures.

Lastly, debate has surfaced on whether current trends of data management are respecting the privacy of any individual. Andrew Machin and Fernanda Onions from the Central and Northwest London NHS Foundation Trust voiced their concerns for how personal data can be kept private in an era of data capitalism.



Andrew Machin

Associate Director of Charity Development and Project Manager
at the Finance Department of the Central and Northwest London
NHS Foundation Trust

The presenters seemed to have a common understanding of what constitutes AI. However, the wider audience seemed less clear – and at what point does AI become sentience? – are Auto pilot/ Self

driving cars examples of AI?

Although the presentations tried to move away from the subject of data, the conversation repeatedly came back to a debate of the identification and protection of personal data, and the use on anonymised data in order to improve AI. This raises the question of what personal data about us is private, and what is public, and what use can be made of either by commercial organisations (would it make any difference if it was used by Not for profit organisations – recent decisions by the Data Protection Regulator suggests not). Is it possible to donate data for the good of humanity (in the same way you donate blood or organs under the NHS).

A question arose as to the need to produce new law, or whether the existing law remains fit for purpose it is the application of it? Panellists were split.

As illustrated through several arguments and examples at the meeting, data capitalism has completely transformed and, in many ways, disrupted society.

Sir Gordon Duff suggested many of these issues can be described through the 3 V's: volume, velocity, and visualisation (Figure 6). Data is being generated in such remarkable speeds and has the capacity to do the unimaginable. The only way to reap its full benefits is to solve the visualisation issue. He argued that an engineering approach is the most appropriate solution.

We need to find how to visualise big data in such a way that it becomes usable and accessible for the public.

The current processes for collecting, using, and managing data seem to be outdated.

In consequence, the APPG AI panellists agreed that we are in a critical point in which we must rethink how data is governed. Throughout the discussion, the group was divided on the direction of this change. Some argued that government should adopt a hard approach (Section 3) and others pushed for a soft approach (Section 4). In the end, all seemed to conclude that we need to establish guidelines based on well-curated evidence - to set the standards for how data should be collected, used, and managed.

3. Data governance needs a hard approach?

The introduction of data-driven technologies has transformed most of what we knew as status quo only a couple of years ago. The world we now live in is very different from that of past generations; and, consequently, some at the **third APPG AI Evidence Giving meeting** argued that existing laws and policies are based on conditions no longer relevant.

Today, data and AI technologies have changed our daily routines tremendously and redefined concepts of accountability, privacy, consent and justice. We have already been benefiting from the opportunities these new technologies offer to make our lives easier and better. Still, simultaneously, we – as a society – are also at a critical point in deciding how to address the risks within these disruptive forces.

The question that arose at the meeting is: **Is our current data legislation landscape suitable for data capitalism?**

The most influential data protection law in the UK is **EU's General Data Protection Regulation (GDPR)** voted on in April 2016 and expected to take full force from May 2018.¹³ The law, in brief, is supposed to protect the personal data of individuals and also guarantee people with a 'right to explanation' for all decisions made by automated or artificially intelligence algorithmic systems. Nonetheless, since its passing, there has been much controversy on the legal existence and feasibility of such a right in the regulation.¹⁴

In the meeting, Alex Housley, CEO of Seldon, commented on the GDPR, shedding light on its integral trade-off between accuracy and efficiency. As a society, we need to consider and decide what trade-off we are willing to tolerate and what margin of error we are comfortable with.

In their June 2017 report, the British Academy and the Royal Society provide a table listing the main UK and EU regulations associated with data-related issues, including:¹⁵

- Data Protection Act (DPA) – 1998 - the current UK law on how personal data should be processed.
- The Freedom of Information Act – 2000 - UK law providing the public access to data held by public authorities.
- Digital Single Market Strategy – 2015 – EU objective to create a single digital market

¹³ REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

¹⁴ Wachter, S., Mittelstadt, B., & Floridi, L. (2016). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation, (Whi), 26–30. <https://doi.org/10.2139/ssrn.2903469>

¹⁵ "Data Governance: Landscape Review." A joint report by the British Academy and the Royal Society. (June 2017).

and smart, sustainable, and inclusive growth in Europe.

The stakeholders at the meeting called for further work to better understand the landscape before deciding which direction to respond in. Dr. Maria Ioannidou, Lecturer in Competition Law at Queen Mary University (University of London), has started an initiative on risks and challenges posed by the 4th Industrial Revolution. She advised the group: **“before looking at how we move forward, we need to understand what is already there.”**



Maria Ioannidou

Lecturer in Competition Law at Queen Mary University London

Data has been coined as **“the world’s most valuable resource”** (Economist 2017). **“Data capitalism”** encapsulates this reality of our data driven economy and implicitly adopts a market-oriented approach to data. Not any data – Big Data depicting unprecedented volume, variety, velocity and value. Inherent in this market driven approach is the tension between data as commodity and data as personal “identifiers”. The unprecedented attributes of **Big Data** has shifted the focus on treating data as a commodity and as an invaluable input processed by complex algorithms allowing digital decision-making. **Big Data** raise numerous legal and ethical questions triggering discussions about whether the instruments currently available suffice or whether they need to be re-evaluated. The complexity of different legal and ethical issues presents a spectrum corresponding to the nature of the data – anonymised, personal or sensitive – and is further magnified by the sheer volume of data and the growing analytical power of algorithms.

From a theoretical perspective, competition, consumer protection and data protection rules all regulate Big Data and they may at times lead to conflicting outcomes jeopardising people’s trust. In fact, in digital markets what appears to be lacking is people’s trust on the legitimate collection, aggregation, processing and sharing of our data. Consumers too easily give away their personal data in exchange for free services, often because they find themselves in a ‘take it or leave it’ situation, thereby conferring enormous powers to a handful of firms – acting as gatekeepers.

Recent changes in data protection rules with the General Data Protection Regulation aiming at improving privacy by design, accountability, enforcement and data portability attempt to improve this **‘trust gap’**. Equally, competition law enforcers have become increasingly aware and prone to address privacy and data protection considerations in competition law analysis and for good reasons. Competition rules present an adaptable tool to regulate Big Data addressing potential issues of market power, foreclosure, quality degradation, market transparency and collusion.

From a practical perspective on the regulatory intervention, it is crucial to get the balance right between a “fundamental rights” and a “market driven approach”. Crucial for technological development and AI improvements is the link between the data and the processing algorithm since that allows digital

decision-making, which indeed presents enormous potential to improve all aspects of our life.

Hence, before making bold pronouncements about a “**Data Charter**”, one needs to take a step back and reflect what the current frameworks are, what these laws deliver, what the gaps are and opt for an incremental approach through improving the current frameworks before introducing new rules. The current debates focus and explore a synergetic approach to consumer protection, competition and data protection rules and these synergies need to be further advanced.

In addition to this synergetic approach, it is important to understand and map the initiatives on the industry side and explore collective actions building a bottom up approach and conferring on consumers the power to better control their data in digital markets. In this vein, a wider array of stakeholders needs to be engaged and consulted in a bid to improve these tools and foster consumer education. The short/midterm solution in building consumer trust is a soft-law approach to boosting ethics rather than changing the current legal framework. The potential of developing best practices and guidelines for data sharing need also be explored possibly on a sector specific level, e.g. in the health sector.

Mark Skilton – Professor of Practice, Information Systems Management and Innovation at University of Warwick Business School - suggested that we need to adopt a more aggressive approach once having analysed the current landscape. Emphasizing the massive transformations and disruptions that have taken place, he suggested society will need new laws and policies to reflect the new conditions inherent to data capitalism.



Mark Skilton

Professor of Practice, Information Systems Management and Innovation at University of Warwick Business School

Digital data is gold for companies and cyber hackers alike; we need to drive innovation and competitiveness such as medical research and new supply chain productivity; but we also need new ways to monitor and govern A.I. misuse and propriety.

The lawyers say the law in the UK maybe perhaps sufficient to prosecute data theft and data miss use following GDPR for example, but the point is that Artificial Intelligence is a new form of technological manipulation different from previous industrial eras. It can modify data and create automated actions in software and through robotic action.

Today's governance problem is how do we separate the human from the machine? With increasing automation of process steps using A.I., how do you allocate human responsibility where necessary? I support we are ethically obliged to make it clear and transparent for the outcomes from use of A.I. Liability is and will become harder to track with complex algorithm use. Creating a new AI Data usage

charter will help, but we also need to create safe spaces for our digital data with encryption and separating rights to access from rights to appropriate use especially where automation use may blur these boundaries and violate liberties.

We need external causal accountability to humans in the use of A.I. to be much clearer to drive options for better A.I. use for benefits to society, but we also need stronger data and social contracts with improved awareness by companies and individuals to protect the liberties we enjoy in the physical world.

Many in the audience agreed with Professor Skilton, commenting on how data-driven technologies have changed business models and other social notions and, hence, the existing legislation no longer reflects the contemporary arena. Nicola Eschenburg noted that all legislation should provide citizens with increased value. Therefore, if the current landscape does not do this in the period of data capitalism, it is government's responsibility to ensure that the laws change.

4. Data governance needs a soft approach?

Others in the room argued that government should take a softer approach in respects to data governance in today's world.

Clive Gringras, Head of Technology, Media and Telecommunications at the law firm CMS Cameron McKenna Nabarro Olswang, argues: **“law should be incrementally slow when it comes to technology.”** He offered four reasons why the government should refrain from passing new legislation.



There are at least four reasons why our Parliament should resist the temptation to introduce new legislation to address the nascent and new innovations in Artificial Intelligence.

Equality of Rights. If citizens really need more protection – all citizens should be protected, not merely those unlucky enough to see their rights abused by the new, latest technology. We saw this when the 1984 Data Protection Act protected only digital, not offline filing systems. It was no consolation to victims that their privacy rights *were* breached but offline not online, so should they would have no remedy. So there should be no new rules for AI – specifically – that lift the standard expected of those who use AI in contrast to other technologies.

Definitional Arbitrage. Differentiating the regulation of one type of technology in contrast to another creates an incentive for definitional arbitrage. Businesses will spend time arguing that their innovation falls on the lenient side of the legislation; hurt citizens will, in contrast, spend their time arguing that the harm originates from the stricter side of the statute. All this will lead to lawyers and courts needing to figure out what the draftsman meant.

Jurisdictional Arbitrage. Fearing the new regulation, businesses in the UK might avoid innovation in the area. And so, inadvertently, the UK Parliament might tip the scales of investment away from the new and back to the old. Meanwhile, the technological developments from more indulgent overseas regimes will not wait just off the shores of the UK, frightened of lapping over the feet of King Canute. Those non-UK developments instead will simply end up being used by UK companies – just not sold by them.

Law Thicket. There is already a considerable thicket of data-obsessed laws. The UK already has two highly comprehensive legislative frameworks in place, or at least soon to be, given the Queen's

Speech: the GDPR is apparently on its way and the current Database Directive. If anything, current and new privacy laws should be constrained to take only a risk-based approach to de-identification and anonymization restrictions in the GDPR. And the legislator should permit databases and copyright works to be used for text and data mining – AI systems should be able to learn without requiring millions to be spent on licensing or litigation.

Clive Gringras gave two suggestions for policy-makers:

- GDPR has to be rethought, keeping in mind that the gains are far more than the drawbacks.
- Intellectual property rights should be loosened. IP Systems in Europe and the UK tighten the ability to assess the data.

Stewart Room, Partner and Global Data Protection Leader at PwC, shared the same view with Clive Gringras on data governance. He called for soft-structures as the appropriate channel to govern the trends of data capitalism. He argued that most stakeholders are already engaging with the ethical issues related to data and data-driven technologies. Companies are already proposing solutions to how to address ethical issues and government should encourage more of these discussions and facilitate collaboration amongst stakeholders across industries.



History shows that soft-structures have been really successful.

There are four recommendations for how the UK government should react:

- The UK needs to analyse the current law to see if it fits this new period of 'data capitalism' and emerging technology.
- Data needs to be put on the broader agenda, given its economic potential and the fact that it crosses several different policy areas.
- The government needs to encourage conversations on how we value data and how we deal with it.
- There might be no need to change the existing law. Simply, there might be a need to create different mechanisms.

Many of the panellists spoke about the user's responsibility in data transactions. When one signs up for a platform, for example, he/she often gives consent to give away his/her data in

exchange for a specific service. Dr Maria Ioannidou, reminded the group, that **there is collective responsibility in every transaction including that of data.** However, as noted by most of the attendees at the APPG AI meeting, a small percentage of the population actually understands what data they are giving away and how it will be used. We need to understand the problem in order to regulate it and intervene if need be.

The government should help empower citizens to understand the consequences of their actions better. Awareness and educational campaigns could help inform citizens how their data is being used and what rights they are giving away when giving data.

The chairs Stephen Metcalfe MP and Lord Tim Clement-Jones highlighted this issue of trust. They acknowledged the government's responsibility to help create trust and also educate the public on these critical issues. Stephen Metcalfe MP proposed a ranking system to illustrate how much data one is giving away when signing up for a service or a new platform.

5. Data governance needs an evidence-based, guidance framework

Robbie Stamp, Chief Executive of Bioss International, summed up the key takeaways from the meeting in the following way:

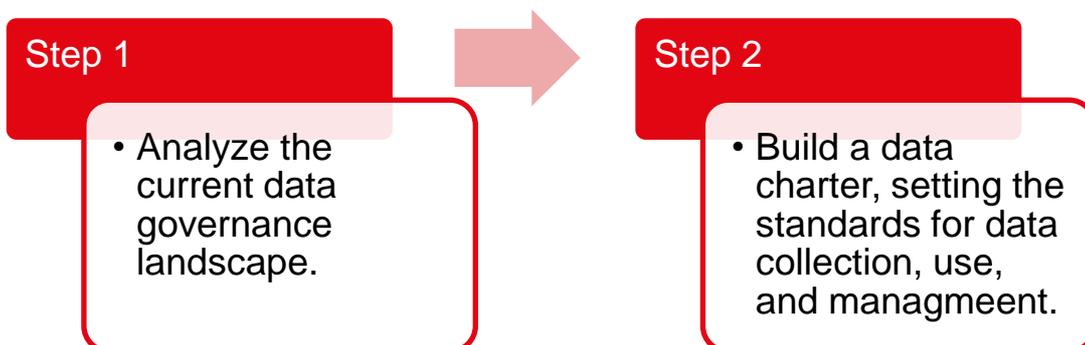


1. Look very hard at existing legislation before creating new.
2. It is very easy for a discussion like this to be overwhelmed by data issues - critical as they are they are only part of a wider set of ethics issues, which were scarcely touched on. So in a way only half the agenda was really covered.
3. Important work to be done on understanding what the various bodies around the world are already doing in 'ethics.'
4. The very notion of 'ethical' algorithms is problematic- both practically and philosophically.
5. Important to consider other immediate and practical frameworks for managing what I think of as a whole new area of 'working relationships' with new forms of intelligence.

Thinking this way opens up new avenues to consider ethics, accountability and authority in the 'everyday' work of governments and organisations.

Although two different approaches were offered in the way data should be governed, most agreed on the following two-step process moving forward.

Figure 5. Two Necessary Steps to changing Data-Governance



The **first step** to changing data governance is to analyse the landscape currently in place and gather evidence on the existing regulations, policies, and institutions existing. For example, given that the UK has already committed to enforcing the European General Data Protection Regulation in 2018, further research should be pushed forward to recognize the law's impacts now and in the future.

Using the data gathered from the landscape analysis, the **second step** is to provide society with a framework, guiding them on how data should be collected, used, and managed. **We need to understand how to manage data on a social level.**

Professor Birgitte Andersen, CEO of the Big Innovation Centre (BIC), shared BIC's vision of what a **data charter** should look like.



We need to be much more open to share our data if we want to unlock the benefits from the AI revolution. We need to understand that data sharing is a social benefit and the way value is created in the AI era. In fact, data sharing should be part of the 'social contract' between each person and society/government.

For example, for a driverless car to go on the road the users would need to share their data – just as we require driving license and a car insurance. This means that we adopt an '**opt-in unless you opt-out**' approach to personal data disclosure (opting out means that you can't enter the roads).

For this to happen, the Big Innovation Centre argues that government needs to establish a **Data Charter** with stakeholders on what can be done with personal and business data, so everyone will know how their data is used and not used, which in turn increases trust and creates incentives to allow data to be shared.

This means a change in policy and regulation – namely a **shift from policies around controlling the data itself to how the data is governed**. As a first for Europe, UK could lead. USA is already moving ahead with 'user rights'. Of course – we need to introduce a 'fair use' principle.

The **Data Charter** should guide **ethics boards** in companies to set transparent principles on how data will be governed. It should also trigger an official **AI watch dog** around which consumers can unite, enforcing trading standards.

There should also be an **equal access to data platform** or **shared information system** on which AI data be retrieved in a user-friendly way by the public, so people can know their public record and benefit from knowing information about them-selves in a structured way.

There was much discussion during the meeting on what the standards of these guidelines could potentially look like. Professor Skilton suggested a data charter with sensible guidelines on how to manage data. This charter should be built based on the values we have as a society. A common theme that was discussed was the use of data for purposes that have explicit social value, and the group agreed that **there is an ethical negative if data exists yet is not used to generate social purpose.**

A second point made by several of the panellists is offering the user the option to **“opt in”** or **“opt out.”** Professor Skilton suggested we create a middle option, in which one can “opt in” in some cases and “opt out” in others.

Furthermore, the panel agreed that the system of data governance must be transparent and easily accessible to communities. Transparency is the ingredient for building trust in the system. Stakeholders and the general public need to understand the process of data collection, use, and management. Also, the society needs to have a common understanding of who is accountable or liable for a certain outcome.

The government can play a significant role by creating confidence and ensuring society that their rights are being protected.

Education is a tool to move forward successfully. The society must be informed on what the data conversations are, what the gains of sharing data are, what the drawbacks might be, how their data is likely to be used, etc. However, the educational issue, transcends the agenda of simply informing the public on how data is being collected and managed. The government needs to help shrink the skills gap. Stewart Room noted the current lack of skills in the area on a global level, including in the UK. According to Room, the key is bringing in STEM skills in education from primary levels.

The following **action points** were taken away from the third Evidence Giving meeting:

Theme	Action Points
Data governance needs an evidence-based, guidance framework	<ol style="list-style-type: none">1. Invest in research to further understand and analyse the current legislative regime, specifically in regards to data and AI technologies.<ol style="list-style-type: none">a. Focus on the General Data and Protection Regulation (GDPR) and evaluate its anticipated impact in the UK environment. Questions to consider include: (A) Will GDPR help the UK economy? (B) Does the GDPR adequately address ethical implications of data usage? (C) Are UK

businesses ready to meet the requirements of the GDPR?

2. Using evidence-based research, build a data charter to:
 - a. Educate society on how their data is being collected, used, and managed.
 - b. Establish a standard of visibility and transparency.
 - c. Guide relevant stakeholders in how to use data when building AI technologies and / or training AI technologies.
 - d. Encourage AI technologies with a social purpose.

Acknowledgements

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, including Machine Learning.

Our supporters - Barclays, BP plc, Deloitte, EDF Energy, KPMG, Olswang, Oxford University Computer Science, PwC - enable us to raise the ambition of what we can achieve.

The APPG AI Secretariat is Big Innovation Centre.



Big Innovation Centre

Ergon House, Horseferry Road
Westminster, London SW1P 2AL

info@biginnovationcentre.com
www.biginnovationcentre.com

All rights reserved © Big Innovation Centre. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form without prior written permission of the publishers. For more information contact b.andersen@biginnovationcentre.com. Big Innovation Centre Ltd registered as a company limited by shares No. 8613849. Registered address: Ergon House, Horseferry Road, London SW1P 2AL, UK.

