

Governance, Social and Organisational Perspective for Al

A theme report based on the 5th meeting of the All-Party Parliamentary Group on Artificial Intelligence [APPG Al]. **Governance, Social and Organisational Perspective for AI** is a theme report based on the fifth meeting of the **AII-Party Parliamentary Group on Artificial Intelligence (APPG AI)** - held on 11 September 2017 at the House of Commons.

This meeting was chaired by 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
- Lord Holmes of Richmond Vice Chair
- Lord David Willetts Vice Chair
- 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 I Ergon House, Horseferry Road, London SW1P 2AL T <u>+44 (0)20 3713 4036</u> Email: appg@biginnovationcentre.com

Any enquiries regarding this publication, contact: CEO and CoCreator: Professor Birgitte Andersen APPG AI Rapporteur: Niki Iliadis

Governance, Social and Organisational Perspective for Al

A theme report based on the fifth All-Party Parliamentary Group on Artificial Intelligence [APPG Al] Evidence Giving meetingEvidence Meeting.

11 September 2017 - House of Lords, Committee Room 4A



Overview

The aim of the fifth APPG AI Evidence Meeting centred on understanding the social and organisational implications of AI technologies - and how they should be governed.

The meeting was chaired by Lord Tim Clement-Jones (also Chair of the Lords Select Committee on AI). Eight experts were invited to provide evidence reflecting their views on (1) AI impacts on social, cultural, and organisational structures and (2) AI governance, in both the private and public sector.

139 TOTAL PARTICIPANTS

8 Pieces of Oral Evidence

8 Pieces of Written Evidence

The panel included: **Miles Brundage** (AI Policy Research Fellow at Oxford Future for Humanity Institute); **Dr Joanna Bryson** (Reader at University of Bath's Department of Computer Science and Affiliate at Princeton University's Center for Information Technology Policy); **Dr Stephen Cave** (Executive Director at the Leverhulme Centre for the Future of Intelligence); **Dr Kate Devlin** (Senior Lecturer at Goldsmiths' Department of Computing); **Dr Julian Huppert** (Director of the Intellectual Forum at Jesus College, University of Cambridge); **Rodolfo Rosini** (Co-Founder and CEO of Weave.AI); **Krishna Sood** (Technology Lawyer at Microsoft); and **Dr Sandra Wachter** (Postdoctoral Researcher in Data Ethics and Algorithms at Oxford Internet Institute).

Al's opportunities for growth and productivity, and the numerous social benefits, have repeatedly been highlighted in past Evidence Meetings as well as national and global reports.

- Dame Wendy Hall and Jerome Pesenti's AI Review recognized AI's potential to add an additional £630 billion to the UK economy by 2035, increasing the annual growth of GVA from 2.5 to 3.9%.¹
- PwC, in June 2017, concluded that AI could bring an additional £232 billion to UK GDP, making it 10.3% higher in 2030.²
- An independent report led by Professor Juergen Maier, CEO of Siemens UK, for BEIS - revealed the positive impact of faster innovation and adoption of industrial digital technologies to reach the amount of £455 billion for UK manufacturing over the next decade.³

Certainly, the windows of opportunity are mass. However, the same evidence we have

¹ Hall, W., Pesenti, J., (October 2017), 'Growing the Artificial Intelligence Industry in the UK.' Department for Digital, Culture, Media, and Sport; Department for Business, Energy, & Industrial Strategy. https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk

² PwC, (June 2017), 'The Economic Impact of Artificial Intelligence on the UK Economy.' PwC. https://www.pwc.co.uk/services/economics-policy/insights/the-impact-of-artificial-intelligence-on-the-uk-economy.html

³ Maier, J., (November 2017), 'MADE Smarter 2017,' Department for Business, Energy, and Industrial Strategy. https://www.gov.uk/government/publications/made-smarter-review

gathered has also addressed a set of challenges with AI, in regards to data privacy, unemployment, inequality, automated decision-making, and other ethical and philosophical matters. These issues are causing public concern; and, consequently, building low confidence in the emergence and application of AI technologies.

Governance is necessary moving forward, providing UK with the means to capitalise on AI benefits and, simultaneously, mitigate its risks. The APPG AI's fifth Evidence Meeting extracted five main themes in the discussion exploring what the impacts of AI on social and organisational structures are and how AI should be governed in both the private and public realms looking ahead.

| Theme | Description |
|--|---|
| Al is resurfacing existing issues in our social and organisational structures. | Many of the challenges now linked to AI are far from new. For instance, concerns about increasing inequality gaps, stereotypes and biases, shortages of skills, and abuse of power have existed in our society for centuries now. AI is not the creator of these problems. Rather, in many ways, AI is simply resurfacing prevailing problems and urging society to acknowledge their existence and provide solutions. |
| Al is creating a new set of issues for society to address. | On the other hand, AI technologies are of such high impact and progress at such rapid speeds that some issues developing are authentically new. Some of these include increasingly automated decision-making, potentially catastrophic security threats, technological unemployment, and transformations in current notions of privacy, agency, consent, and accountability. |
| Governance is necessary to build public confidence in Al. | The combination of existing problems put back in the spotlight and the emergence of a new set of challenges is creating social turbulences. Most people are not well informed of AI implications and, in consequence, there is a growing sense of uncertainty, fear, and mistrust building around AI. Governance, by both the private and public sectors, is necessary for individuals to build trust and faith in these AI technologies. |
| Al governance depends on corporate responsible innovation. | The private sector shares responsibility for creating AI that is made and used for the benefit of humanity. Therefore, the private sector must commit in taking corporate responsibility for the AI they produce and apply in their functions. Companies must be encouraged to create AI with social value and to consider the ethical and social consequences of the technologies they deploy. |

| Al governance relies on forward- looking policies and regulation. | Government must set the boundaries for what AI can and cannot do. Forward-looking policies, prioritising the good of the wider public, should be adapted. Taking into account the remarkable speed technology is progressing, Government must act quick in guiding the norms and standards for AI and set the appropriate regulation where need be. Furthermore, Government must evolve with technology to harness the opportunities and protect society from the risks. |
|--|--|
|--|--|

This theme report is not research-oriented but aims to summarize these key themes, using the evidence gathered at the fifth APPG AI Evidence Meeting (details above). It is not exhaustive but reflects what was discussed at the session, 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.

The All-Party Parliamentary Group on Artificial Intelligence [APPG Al] was created in January 2017 to explore the impact and implications of Artificial Intelligence, including Machine Learning. We aim 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 Al entrepreneurs have come together in an effort to share evidence and beliefs, and assist in setting an agenda for how the UK should address Al 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 Meeting approached Artificial Intelligence through a general lens, identifying the key issues within the umbrella term that stakeholders should focus on. The second and third APPG AI Evidence Meeting deep dived into ethical and legal issues in AI, regarding decision-making and the data-driven economy. While the fourth APPG AI Evidence Meeting focused on changes in the economy, market structures, and business models.

Event Summary

The fifth APPG AI Evidence Meeting was chaired by **Lord Tim Clement-Jones** and focused on the social and organisational perspectives in the sphere of AI, exploring the social purpose of AI technologies. The conversation engaged with the debate on how AI should be governed.

Krishna Sood, from Microsoft and the first of eight panellists, discussed specific principles stakeholders in AI should follow. These included: AI products should be designed to assist humanity, they should be transparent, they should aim to maximize efficiency without sacrificing human dignity, they should respect the notion of privacy, they should address the issue of algorithmic accountability, and they should guard against biases and stereotypes. Companies should hold responsibility to uphold these principles and form collaborations such as Partnership on AI to encourage discussion on AI impacts. Krishna recommended the UK Government to promote the free movement of data, implement General Data Protection Regulation (GDPR), continue to invest in research and development, and promote relevant skills across all generations.

Miles Brundage, from the Future for Humanity Institute, took the microphone and started his speech reminding the group that technological progression is exponential. He argued we cannot really anticipate how fast technology will develop and in what way. In consequence, the UK Government should not base their policies on a specific time frame. He also noted some likely near-term and long-term implications. Near-term impacts will be seen in the job market and in the security sector. Long-term impacts include safety concerns (i.e. countries racing to compete in an era of Advanced AI might not pay as much attention to safety standards and the consequences could be catastrophic).

Joanna Bryson, from the University of Bath, highlighted the fact that AI and humans are different. She says: 'What sets AI apart from Natural Intelligence is that Artefacts are made deliberately, by humans.' We shouldn't over-personify machines and remember the importance of the human-factor when making decisions. Furthermore, she noted that many of the ethical implications we now link to AI have existed in society for centuries. For instances, biases already exist within our cultures. AI only surfaces the issues and urges stakeholders to react. She advises the Government to enable the on-demand and routine auditing of AI and algorithmic systems.

Stephen Cave, from the Leverhulme Centre for the Future of Intelligence, discussed the overlapping commonalities amongst the issues addressed in the meeting. Thinking of ethical and governance issues in three big categories, he explained how autonomy, data, and intelligence are linked together. For example, the issue of transparency is directly linked to autonomy and data and, also, indirectly linked to intelligence. Hence, when considering holistic solutions to these problems it is important to recognize their interplay. He called for the UK Government to create an advisory body that will be responsible for AI governance and suggested for this body to be closely tied to Royal Society's proposal for a data governance agency.

Kate Devlin, from Goldsmiths' Department of Computing, spoke about the increasing trend we are seeing in which technology is being humanized. She shed light on the fact that AI is impacting all sectors and industries, even those we might not feel so comfortable discussing (i.e. sex robotics). The sex tech industry is a \$30 billion market, she stated, and the impact of AI has been tremendous. There is much ethical debate in technologies like sex robots that stakeholders must consider. Government must encourage more research to understand impact across every sector, even one that might be taboo. It is important to regulate based on evidence, she emphasized.

Julian Huppert from the Intellectual Forum at University of Cambridge, agreed with the others: AI is a hugely exciting field with potential for much good but also potential harm. The challenge is that regulating too tightly might mean losing out on some benefits but regulating too loosely might mean social harms. He focused his talk on three major implications of AI: power, governance, and work. The UK should be concerned about the overconcentration of power in the hands of one entity and the Government should promote open standards, open interactions, and a competitive landscape. Stakeholders should adapt governance models – like that of DeepMind Health - in which they are audited by external affiliates and held accountable to these results. Furthermore, stakeholders need to recognize the impact AI will have on jobs but consider the positive outcomes of this disruption. Perhaps, automation will lead us to rethink one's purpose and the typical 40 hour working week? He urged the community to not think in simple trade-offs.

Rodolfo Rosini from Weave.AI was next to speak and highlighted the technological acceleration in the global context. Applied AI is changing the world and is a force multiplier for other technologies, he stated. UK has the opportunity to lead not only technologically but also in regard to governance. The national strategy on AI needs to address lack of entrepreneurship in the country, job destruction, migration policies, and educational challenges. Throughout history, UK has performed well in developing new technologies but not so well in exploiting their commercial usage. We have the chance to change that.

Sandra Wachter from Oxford Internet Institute was the last to speak and focused her talk on how the GDPR – expected to be put into force in March 2018 – will affect society. She argued that GDPR has some flaws because it only gives citizens the right to contest if the decision is made through a fully automatic process. Furthermore, it does not give a right to explanation. She urged the UK to lead in closing this accountability gap by (1) ensuring GDPR applies for decisions solely or predominantly based on automated making and (2) ensuring the right to explanation is legally binding. She encouraged stakeholders to work together for a future in which AI and humans can work side by side.

Lord Clement-Jones thanked the panel and opened the discussion to the wider audience. There were several questions zooming in on specific ethical implications such as algorithmic biases, privacy, the use of autonomous weapons, etc. A key question surfaced: even if we can technically make something, should we? The Evidence Meeting concluded with a positive message encouraging stakeholders to make AI with social purpose that can be used to make us better humans.

Table of Contents

| Overview | 5 |
|---|---|
| Event Summary | 8 |
| 1. Al is resurfacing existing issues in our social and organisational structures1 | 1 |
| 2. Al is creating a new set of issues for society to address | 4 |
| 3. Governance is necessary to build public confidence in Al | 8 |
| 4. Al governance depends on corporate responsible innovation | 3 |
| 5. Al governance relies on forward-looking policies and regulation | 8 |
| Action Points | 2 |
| Acknowledgements | 3 |
| Contact details | 4 |

1. Al is resurfacing existing issues in our social and organisational structures

The rapid emerging and application of AI technologies has given rise to much discussion on the mass opportunities but also much debate regarding the potential risks and drawbacks these systems might cause in society. Media headlines publish worrisome articles warning individuals about unprecedented levels of unemployment, biased algorithms making decisions in sometimes life-critical situations, and even the potential of super intelligent agents destroying humanity. In many storylines, AI is posed as the biggest policy challenge for our generation.⁴

The panel in the fifth APPG AI Evidence Meeting was asked to contemplate on the overall impacts of AI on society, organisations, and their structures – paying special focus on some of the risks and drawbacks mentioned above. Ultimately, the questions the group explored were whether these implications required governance and, if so, what form should governance take.

One point was stressed from the beginning of the session: **we should be careful not to humanise AI**. AI itself is neither good nor bad. Many of the risks that are now credited to AI technologies have in fact existed in our societies for centuries. Many of these issues are wellembedded within our social, cultural, and organisational structures and AI only puts them back in focus.

For example, Dr Joanna Bryson (whose primary research focus is exploring how AI can be used to understand Natural Intelligence) discussed the issue of algorithmic biases with the group. In the second and third Evidence Meetings, APPG AI shed light on the problem of unreliable and inaccurate data being fed to algorithms and, precisely, its use to inform decisions related to important matters such as finance, health, education, security, and employment.

ProPublica, last year, published a well-circulated article, showing how a software used across the US to predict future criminals gave flawed results that were based on racial stereotypes.⁵ The software was likely to falsely flag black defendants, wrongly labelling them as future criminals at almost twice the rate as white defendants. Likewise, the software also mislabelled white defendants as low risk more often than black defendants.

Dr Kate Devlin, from Goldsmith's Department of Computing, also shared with the group a more

⁴ Yiu, C., (November 2017), 'Technology for the Many: A Public Policy Platform for a Better, Fairer Future,' Tony Blair Institute for Global Change. https://institute.global/insight/renewing-centre/technology-many-public-policy-platform-better-fairer-future

⁵ Angwin, J., Larson, J., Mattu, S., Kirchner, L., (May 2016), 'Machine Bias,' ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

recent example that has become popular in the media. A newspaper article argued that a new AI application can accurately guess whether people are homosexual or heterosexual based on photos of their faces.⁶ Various human rights groups have raised critiques against this software arguing that it is based on flawed data that does not represent the LGBT community.

These case studies are only two of hundreds of examples showing how bias embedded within algorithms can skew outputs and decisions, and raise ethical concerns.

However, as Dr Joanna Bryson informed the Parliamentarians and wider audience, the algorithms themselves are not the source of bias. Individuals are those who carry with them social biases and, consequently, embed these biases in AI. As she said, **'human biases tend to sneak into AI systems.'** We should be careful not to point the finger of blame on the algorithms themselves as AI is an artefact created by a group of humans. Humans should ultimately be the ones held to high standards of responsibility and accountability.



Dr Joanna Bryson

Reader, Department of Computer Science, University of Bath; and Affiliate, Center for Information Technology Policy at Princeton University

Firstly, AI is not mathematics but computation. Intelligence is not an abstraction, but rather a physical processes subject to natural laws and principles of computer science. Computation requires time, space, and energy. Even the number of possible chess games of 35 moves or less is greater than the number of atoms in the universe. Thus singularity-oriented concerns about AI are misplaced. Intelligent systems design is a long-term arms race for advantages in insight, comprehension, and planning.

The UK is well-positioned for AI expertise, but AI as ICT implies business potential to reach billions of customers. For this, the UK requires transnational cooperation. The EU presently leads globally in AI and data policy, and also embodies a leading model of transnational economic and legal cooperation -- a position gained with British ingenuity.

Al is definitionally an artefact, meaning it is built deliberately, and therefore is from inception a human responsibility. We should maintain human and corporate responsibility for all Al products, because our justice system rewards and dissuades humans, not machines. Auditing of Al does not

⁶ Levin, S., (September 2017), 'New AI can guess whether you're gay or straight from a photograph,' The Guardian. https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-orstraight-from-a-photograph

require knowing the 'meaning' of neural network weights any more than auditing accountants involves individual human neurons. By maintaining standard product liabilities for AI, we encourage not only responsible manufacturing and operation of intelligent systems, but clean, maintainable code benefiting also industry. We should not reward companies for poor systems engineering practices by reducing liability for systems they cannot predict or maintain.

Cybersecurity is essential to reliable AI, and AI to cybersecurity.

As decisions are increasingly being made based on the outputs of algorithms and computer software, individuals are voicing concern about what biased AI could mean for society. Individuals are looking for ways to address these injustices and, hence, tackle the issue of human stereotypes directly. **AI could potentially be part of the solution.** AI, Dr Joanna Bryson shared, has the ability to prevent biases and stereotypes through the development of explicit instructions that guide the algorithm to filter biases.

For instance, AI has the potential to help us avoid bias in recruitment, security, education, customer service, and many other important matters. Humans can programme AI with instructions we consider suitable and indicate that the best outcomes, ethically and economically, are those that are free from bias. In result, AI can help us create systems that are fair, more productive, and ultimately better for the wider society.

Hence, AI can be seen as a tool that is neither good nor bad. It is up to us to find ways to use AI to mitigate existing challenges including algorithmic biases but also inequality, low quality education, poverty, environmental concerns, and more.

2. Al is creating a new set of issues for society to address

On the other hand, AI is not like any other technology that has been introduced before. Its scale and speed make it truly unique. Consequently, some challenges posed by AI technologies are indeed authentically new.

Members of the panel, including Miles Brundage from Oxford Future for Humanity Institute and Krishna Sood from Microsoft, identified one of these challenges to be labour disruption. The concept of technological unemployment is not new itself. In fact, the fear of automation leading to job losses has existed for decades. In every industrial revolution, there is much debate on how labour will be affected by the introduction of new, cheaper, and more efficient technologies. But, until now, although technology has created job losses, it has also created new jobs. History shows that labour replaced by machines brings down prices of products and, as a result, increases real incomes. Overtime, demand for new goods and industries to supply them increases and, consequently, more new jobs are created than lost.⁷

It is easy to assume that this time will be the same. Nonetheless, many at the fifth Evidence Meeting, argued there is a high chance this time is fundamentally different. The AI technologies of today have the power to do much more than the machines of the past. Most importantly, **they can replicate not just physical labour but also cognitive and problem-solving tasks.**

Therefore, although the problem of technological unemployment is not brand new, it has now taken different scale and scope.

Many reports have forecasted that AI technologies are being developed and deployed at such rapid speeds and are of such high impact that the disruption in the labour markets will be even more significant. Frey and Osborne estimated 47% of the total US employment to be in high risk for automation in the near-future⁸ and 35% of UK current jobs to be in risk over the next decade or two. ⁹ According to their findings, although technologies have created vast employment opportunities for workers throughout history, today's technology is distinct in that it does not provide the same opportunities, particularly for the less-skilled or less-educated workers. Deloitte's analysis in 2016 revealed UK jobs paying £30,000 to be five times more vulnerable to displacement than jobs paying £100,000 or more.¹⁰ Findings reveal occupations involving complex perception and manipulation tasks, creative intelligence tasks, and social

⁷ Brynjolfsson, E. and A. McAfee (2014), The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies, W.W. Norton & Company, New York.

⁸ Frey, C B, and Osborne, M A (2013), 'The future of employment: how susceptible are jobs to computerisation.', available at:

http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

⁹ Deloitte (2015), 'From brawn to brains: The impact of technology on jobs in the UK', available at http://www2.deloitte.com/uk/en/pages/growth/articles/from-brawn-to-brains--the-impact-of-technology-on-jobs-in-the-u.html

¹⁰ Deloitte (2016), London Futures. Agiletown: The relentless march of technology and London's response.

intelligence tasks are unlikely to be substituted by technology in the near future and, therefore, in less risk for automation. PwC's report in March 2017 also mirrored these findings. Their analysis suggests that up to 30% of UK jobs could be at high risk of automation by the early 2030s. These risks were found to be highest in sectors such as transportation and storage (56%), manufacturing (46%) and wholesale and retail (44%), but lower in sectors like health and social work (17%).¹¹

Technological unemployment was hence noted as a key issue Governance structures should address, but it was not the only. Other issues include: security risks involving the increased reliance on autonomous weapons by militaries across the world and new cybersecurity threats, autonomous decision-making, and concerns about transformations in current notions of privacy, agency, consent, and accountability

Dr. Stephen Cave, Executive Director and Senior Research Fellow at the Leverhulme Centre for the Future of Intelligence (CFI) at the University of Cambridge, stressed the connection between these various ethical and governance issues or challenges.

Breaking down the issues in three main categories **- autonomy, intelligence, and data**, - he shared with the MPs, Lords, and other key stakeholders how each of these categories raises key challenges and how they are all interlinked. Furthermore, he went on to prioritise some areas in which governance should focus on in the immediate, mid-term, and long-term horizons.

According to Dr Stephen Cave, the areas of immediate concern are:

• Accountability, robustness, security, interpretability, bias, privacy, and value alignment

The areas of mid-term concern are:

• Human dignity, jobs, public good, control, dependence, wealth, and manipulation

The areas of long-term concern are:

• Superintelligence, consciousness, and personhood

Figure 2 illustrates how all these issues are intertwined, and how a governance structure addressing one of the areas automatically impacts the entire set.

¹¹ PwC (2017), UK Economic Outlook, Available at: https://www.pwc.co.uk/economic-services/ukeo/pwc-uk-economic-outlook-full-report-march-2017-v2.pdf





Dr Stephen Cave urged Parliament to understand the connection of these various elements and to create governance structures that target the immediate concerns first and the others following. He also called for a stewardship body that would overlook AI-related issues and data governance issues to ensure new technologies' benefits are seen by the wider public and their risks are mitigated.



Although some of the ethical and governance issues posed by AI are related to those posed by other technologies, many are genuinely new. These are issues arising from AI systems' increasing autonomy (capacity to make decisions by itself) and intelligence (ability to make human-level or better-than-human decisions on issues of increasing complexity). They range from establishing accountability when decisions are made by machines, through preventing over-dependence and deskilling, to the impact of automated decision-making on a range of human rights.

The case for a stewardship body to assess and advise on AI-related issues is therefore at least as

strong as the need for such a body on data governance, as recently proposed by the Royal Society and British Academy. But at the same time, given that there are notable overlaps between data governance and Al governance, and given a desire to avoid duplication, there is a case for this to be one body.

However, as there are many issues distinctive to AI governance, it is crucial that the name and mandate of this body include AI from the outset. This will help to ensure it is properly established to give the best possible advice on these urgent and important issues into the future.

3. Governance is necessary to build public confidence in Al

Soon after the meeting began, it became clear that well-defined governance structures were essential to address both established challenges that AI has brought back in the spotlight and a new list of issues that have evolved and/or sprung up.

Al has to some extent impacted all industries and companies already - and will continue to do so exponentially. Therefore, it is important for the private and public sector to work closely to create and implement the appropriate governance structures that will drive the course of how Al technologies affect society in the future.

It is important to horizon scan across sectors and industries in order to gain a full understanding of how AI will transform society.

Dr Kate Devlin, investigating how people interact with and react to technology, urged the audience to consider all sectors in their analysis, not just the mainstream ones currently being discussed. Although healthcare, finance, and customer services are being disrupted by AI technologies, stakeholders must gather use cases from the rest of the sectors in parallel - to get a holistic understanding of AI impact.

She focused her evidence on Al's impact on the markets for sex robots. The market of these technologies is on the rise and, in consequence, there is also a growing debate on the legal and ethical implications of these robots.

She called for **governance structures that look at evidence coming from all sectors**, including more 'taboo' industries. For example, when one looks at privacy concerns in the case of sex robots, it is clear that data privacy on matters dealing with sexuality will take a different form than in other sectors like healthcare due to its extreme sensitivity. UK should invest and encourage further research to understand Al's impact in all of these categories. With exhaustive evidence, the governance structures can apply a truly customised approach depending on the use case in scrutiny.

The below excerpt is taken from one of Dr Kate Devlin's articles in July 2017, exploring the impact of sex robots for the society and the individual. It is an example of how a 'taboo' industry can start being analysed, in an effort to address important questions relating to gender prejudices, inequality, violence implications, and human relationships.



Dr Kate Devlin Senior Lecturer, Department of Computing, Goldsmiths, University of London

(excerpt taken from Dr Devlin's article in iNews.co.uk¹²)

Long a staple of sci-fi films, recent years have seen a rise in interest in real-life sex robots – a machine, often a life-like doll, to take the place of a human partner in sexual activity. Companion robots (of the non-sexual variety) are already in use today and the pros and cons of these are the focus of academic research. Extending this into sexual companion robots is clearly a next step, and one that is becoming a commercial reality with companies now offering mechanised sex dolls that come with artificial intelligence. A report released this week by the Foundation for Responsible Robotics is an overview of the current issues around sex robot development.

I have looked at some of the key questions it raises below.

What would a relationship with a sex robot look like?

The report contains a somewhat rambling account of what might constitute a relationship followed by the rather obvious 'we have noted that robots cannot feel love and tenderness or form emotional bonds'. Yes, we've all noticed that too. The ethical concern here, apparently, is that some people might be deceived into thinking that these emotional bonds are possible. Indeed – it's a concern in human relationships as well. But in all seriousness, if people find happiness in this self-deception (for self-deception it will clearly be), who are we to judge? If someone has no one, wouldn't something be better than nothing?

Will robot sex workers in brothels be acceptable?

On this question, the report concludes with 'we have no evidence'. Of course we don't – for a number of reasons, not least the difficulty in researching an area considered sensitive and taboo. There are already love doll brothels in existence: Japan has them, for example, and there's a new one in Spain. Tenuously, the report's states that 'although we found no evidence for the notion that sex robots would stop sex trafficking, we found some evidence to the contrary.' That 'evidence' is a statement saying 'sex trafficking does not diminish where customers have an appetite for abuse or child sexual abuse. In fact, an increase is seen because sex is known to be available in these areas'. To make the jump from a sex doll brothel to abuse and sex trafficking is disingenuous.

Could sex robots change our understanding of gender?

It's one of the most commonly debated topics in the sex robot debate. A better phrasing would be 'will the current form of sex robots be detrimental to women?' 'Up until now sex robots have been designed

¹² Devlin, K., (July 2017), 'Sex robots tend to be feminine and hyper-sexualised – but let's change that,' iNews: The Essential Daily Briefing. https://inews.co.uk/opinion/sex-robots-tend-feminine-hyper-sexualised-lets-change/

by men, for men' Silicon Valley is not just hostile to women employees. In fact, the technology itself is gendered. Tech is still very much a pale, male and stale world, with the artificial intelligence sector specifically criticised as being 'sea of dudes'. The products that are created are designed and marketed in the image of their makers. Think of the objects we use everyday: seatbelts, initially designed for a man's body, or the pacemakers only suitable for 20 per cent of women. Like much of the technology we use, up until now sex robots have been designed by men, for men.

Do sexbots pose a threat to women?

Does the hyper-feminised, hyper-sexualised sex robot in development today pose a threat to women? The report says 'we just don't know' – and that's fair, because we don't. The arguments on this subject mirror those of video game violence. Does playing violent computer games make the user more likely to act this out in real life? Despite years of research and meta-analyses, the evidence is both contradictory and inconclusive. So, too, is the worry of men acting out violent fantasies with sex robots an as yet unsubstantiated fear. Many current sex robot owners, like the customers of RealDoll, give their dolls a personality and treat them quite reverently. That said, the gendered, stereotyped and clichéd options we have today don't have to dictate that this what a sex robot must look like. Yes, the longstanding narrative is the female sexbot – but let's change that. 'It is clear overall that men are keener on sex with robots than women' states the report, with very little to go on and and without any attempt to discuss sexism, sex toys, or marketing bias.

Could relationships with sex robots isolate us?

Concern about social isolation as a result of intimacy with robots is a longstanding narrative. It's a common worry that humans will be displaced by robots and it plays into our deep fears about loss of agency – that we might be made redundant by robots, and not just when it comes to jobs. The report doesn't mention the reasons behind the concern but it does say 'we have no direct evidence' to say if human intimacy will be damaged. Related to this is whether robots could help with sexual healing and therapy, but it's poorly discussed. Sex technology is already being used in this area to help people who want to have sex but who are unable to, for physiological or psychological reasons, but the report doesn't mention this at all. It's that taboo again: yes, you can have social robots that provide a companionship function, but heaven forfend that it be of a sexual nature.

Would sex robots help to reduce sex crimes?

This is a subject that is already being debated at length in the research community, with strong viewpoints on both sides, particularly around the area of child sex robots, where on one hand they could potentially be used as a therapeutic proxy, but on the other they could potentially lead to an escalation of abuse. That said, UK law already criminalises a broad range of child-related sexual imagery and a failure to include robots is an unintended gap: another instance of law not keeping space with technology. Overall, sex with a robot (or a doll) may be distasteful to some but it is vitally important to remember that we need to regulate based on evidenced harms and threats to society, rather than on the basis of subjective feelings of discomfort or concern. This technology is happening now and we need to shape it in a responsible and fair manner. No, we don't have the evidence yet. Academics working in this field have been pushing for this debate for years, so although this week's report lacks depth – and is at times rambling and hard to follow – it is welcomed for the fact that it states what we've been saying all along: we need more research.

With a good understanding of how AI impacts society as a whole, as well as different industries and sectors, the right governance structures can start being created.

Robust governance structures can have two direct results and, simultaneously, can build trust in these emerging technologies. (Figure 3).



Figure 3. Al Governance structures

As mentioned, the opportunities these AI applications promise are mass; however, uncertainty and fear from the drawbacks the fifth APPG AI Evidence Meeting discussed prevents the unlocking of their full potential.

Until people feel there are systems and platforms in place to protect them from the potential challenges (old and new), they will not feel comfortable to accept AI. Without transparent and robust governance frameworks society will not feel at ease.

Al Policy Research Fellow, Miles Brundage, discussed the importance of setting the right governance structures. Highlighting job displacement and security-related concerns as two key areas of concern, he noted that the UK can lead in setting the right governance for Al. He emphasized the need for international collaboration as most of these issues cross geographical borders.



Miles Brundage Al Policy Research Fellow, Oxford Future for Humanity Institute, University of Oxford

The UK has a unique opportunity to lead on both the technical and governance aspects of AI.

Al presents many challenges and opportunities in the near term and the long term. Already, we are beginning to see myriad business applications of Al, and the possibility of substantial job displacement associated with Al in the coming decades should not be discounted by the UK Government, though the pace of development more than a few years out is very hard to predict.

Additionally, there are significant security-related applications of AI, both for good (e.g. better crime prediction and detection) and for ill (e.g. automated cyber-attacks from criminals) which the UK Government should explore more fully.

Over the longer term, there is substantial uncertainty about how quickly AI will progress to humanlevel capabilities and beyond, but many experts see this as likely this century. Such a transition would present many deep safety and governance challenges, about which the UK Government should support more research and lead the conversation.

Finally, the **UK Government should commit to ensuring that AI developed within its borders is** of wide benefit to humanity, and explore cooperation with countries around the world to avoid reckless racing in the development of advanced AI systems. Doing all of this will require the UK to improve its recruitment and retention of top AI talent within the Government so that it remains on the cutting edge as the underlying technology develops.

4. Al governance depends on corporate responsible innovation

In a technology reception for British entrepreneurs and business, Prime Minister Theresa May emphasized the need for businesses, educators, innovators and Government to work as partners to find the ways new technologies can be most beneficial. Specifically, she said: 'We all have our part to play – let's work together to achieve it.'¹³

The role of the private sector in ensuring AI opportunities are reaped and the risks are avoided is huge. The private sector is often the one creating, programming, managing, and monitoring AI. The **UK must encourage businesses to create and adapt governance structures that ensure AI benefits all off humanity.**

Krishna Sood shared some of Microsoft's vision when creating AI of social purpose with a 'human-centred approach.' Specifically, she pinpointed six areas corporates should focus on when creating and applying AI.



The power of Artificial Intelligence (AI) will bring great benefits to society. Current AI technologies already enable important advances in education, healthcare, and many other areas. For example, the biological computation group at our Microsoft Research Lab in Cambridge is working at the intersection of machine learning, computer-aided design, and biology to pioneer new approaches to challenges such as treating cancer by using advanced computational methods to understand the behaviour of cells and their interaction, which will help to 'debug' an individual's cancer and provide personalised treatment. Moreover, it is estimated that AI will boost economic growth in the UK by adding £140 billion to the UK economy by 2034, and boost labour productivity by 25% across all sectors, including in Britain's strong pharmaceutical and aerospace industries.

Despite its potential benefits, AI has raised several concerns, including questions of governance, social change and job displacement. At Microsoft, we believe a 'human-centered' approach to AI is

¹³ May, T., (November 2017), 'PM tech reception: 'We all have our plart to play – let's work together to achieve it.' Prime Minister's Office. https://www.gov.uk/government/speeches/pm-tech-reception-we-all-have-our-part-to-play-lets-work-together-to-achieve-it

important to fully realise the power of AI. Microsoft's CEO, Satya Nadella articulated that AI must:

- 1. Be designed to assist humanity;
- 2. Be transparent;
- 3. Maximise efficiencies without destroying the dignity of people;
- 4. Be designed for privacy;
- 5. Have algorithmic accountability so that humans can undo unintended harm; and
- 6. Guard against bias.

Recommendations for Government and policy-makers:

- Despite advancements, AI is at a nascent stage of development. Continuing collaboration between Government, business, civil society and academic researchers, for example, through the Partnership on AI (PAI), is essential to shaping the technology and realising its benefits. Engagement should be viewed through the lens of diversity and inclusion to ensure that policies take into consideration all individuals that may be impacted by proposed technologies.
- The UK Government should implement the GDPR through the Data Protection Bill and agree to a successor to the Privacy Shield post Brexit to foster digital trust and facilitate the free movement of data.
- Continued investment in research and development of AI technologies in the UK as well as a lifelong commitment to STEM education and skills training initiatives for people at all stages of the job continuum.

Microsoft is enthusiastic about the promise of AI to transform and improve every aspect of our lives. We look forward to contributing to the UK Government's ongoing efforts to develop an enabling policy framework to realise this vision.

Many companies are already adapting policies that share Microsoft's vision. DeepMind, for instance, announced the creation of 'DeepMind Ethics & Society' in October 2017.¹⁴ The group aims to help technologists put ethics into practice by exploring the real-world impacts of Al. It brings together a group of experts, including Professor Jeffrey Sachs and Professor Nick Bostrom, to serve as independent advisors who help provide oversight, critical feedback and guidance for the company's research strategy and work program.

Dr Julian Huppert also discussed with the APPG AI another one of DeepMind's initiatives, as part of their DeepMind Health division. They have set up a committee of external affiliates to overlook the governing practices of the company. At the end of the evaluation, the affiliates will create a report with their findings which will be made public. This is a completely new approach to governance and encourages companies to be held accountable to the decisions they have made.

¹⁴ DeepMind, (October 2017), 'DeepMind Ethics & Society,' https://deepmind.com/applied/deepmind-ethics-society/



Dr Julian Huppert Director, Intellectual Forum, Jesus College, University of Cambridge

(excerpt taken from Dr Huppert's article in City AM)¹⁵

Al can liberate humans to lead happier lives, if we get it right

Discussions around artificial intelligence (AI) provoke both excitement and fear in strong measure.

Just last week, Elon Musk predicted that AI will be the most likely cause of World War III (although perhaps he just wasn't following President Donald Trump closely enough).

As happens so often with new advances, policymaking lags significantly behind. We must think now about the consequences that AI will have very shortly in our lives – changing work, society, economics and more.

There's some encouragement that at last governments around the world are waking up to the size of some of the societal questions that AI innovation will ask of us. Today I am back in Westminster to give evidence to a cross-party group of MPs and peers, to discuss and debate these seismic shifts that are coming our way.

There are doomsayers and evangelists, hype merchants and luddites, and as usual the truth lies somewhere in the middle. There are incredible opportunities in AI, and there are also incredible uncertainties and risks.

How do we get the former, and avoid the latter?

One particular challenge is how society should have control of companies and organisations doing Al. Too lax, and we will find that unwanted harms increase. Too tight, and we throw away the benefits from innovation. There are some great organisations already grappling with the potential of Al for this country. The Big Innovation Centre (a London-based think tank), is leading the debate on how Al may shape a more egalitarian and efficient society, in terms of business governance, effective policy making, and social interaction.

Another is DeepMind Health (DMH), owned by Google's parent company Alphabet, which is taking an innovative approach. The application of AI to healthcare (and, indeed, a Google-linked company having access to medical records) is deeply controversial, though also has large potential gains.

¹⁵ Huppert, J., (September 2017), 'AI can liberate humans to lead happier lives, if we get it right,' City AM. http://www.cityam.com/271681/ai-can-liberate-humans-lead-happier-lives-if-we-get-right

Boldly, DMH has set up an Independent Review Panel, which I chair. We have been given huge independence, a generous budget to commission independent investigations into anything DMH do, very broad access to everything that they are working on, and the freedom to publicly review and critique them.

I know of no other private organisation that is even trying to be so open and transparent. I hope others will try to follow this example.

I'm a liberal, and as a starting philosophy I am worried about any over-concentration of power. There is a problem whenever too much power rests with any one person, group or organisation. We should therefore guard against companies which seek to lock users of their AI products into using their systems and no others. We must encourage companies to provide open source interfaces, and be replaceable, so that others can come in and innovate.

This will stimulate competition and ultimately drive better products and services. That is not a comfortable position for many companies, but that is how society will get the greatest benefit from AI. Then there are worries about the impact of AI on work. The Rustat Conferences, which I direct, has looked at the consequences of this. A lot of attention has been paid to the jobs that will be lost or replaced by AI-based systems, and this is obviously important. However, it doesn't have to be a bad thing – if we address it properly.

There is nothing to say that our economic system requires a constant amount of human labour; we should treat the reduction in need for labour as a huge positive, rather than seeking to generate jobs simply to keep people occupied.

The Romans seem to have worked nine days in 10. This then changed to six in seven. We now work five. There is no reason that the status quo should remain as such. Shorter working weeks would give more people a stake in employment, and also give them more time to do other things with their lives.

Al has the potential to not only create new jobs and solve problems in different ways, but also free us up from drudgery, so we can focus our lives on things we actually enjoy. As such, Al can hugely boost our wellbeing.

If the industry, politicians, and society enter this new world together, we can ensure that it liberates us to lead better and more fulfilled lives.

Dr Huppert advised the others not to think of AI issues in simple trade-offs. When creating governance structures, it is often argued that there is a trade-off between one 'good thing' such privacy and another 'good thing' such as healthcare benefits. If regulation is drafted smartly, individuals should still be able to gain healthcare benefits without compromising their individual privacy (Figure 4). There is a limiting curve in which you can get more of the good thing and less of the bad thing.

Figure 4. AI Trade-Offs is Often a Misconception



There have also been a growing number of partnerships and associations amongst technology giants and corporates to explore the social implications of AI and help shape the agenda. The **Partnership on AI** is one of the most impactful, 'established to study and formulate best practices on AI technologies, to advance the public's understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influences on people and society.'¹⁶ Founded by Amazon, Apple, DeepMind, Google, Facebook, IBM, and Microsoft, the Partnership has now grown to 53 companies including non-profits (OpenAI, AINow, Future of Life Institute, the Future of Humanity, etc.) and for-profits (NVIDIA, Intel, ebay, Accenture, etc.).

Private sector initiatives are well positioned to create governance structures because they have a strong understand of the technology that they are working on and have unique perspectives about the specific challenges.

Such efforts by companies, including internal systems looking at the social implications of AI and external collaborations, should be encouraged and supported.

¹⁶ Partnership on AI. https://www.partnershiponai.org/#

5. Al governance relies on forward-looking policies and regulation

Governments, of course, have a big responsibility in regulating AI and helping set the norms and standards for how new technologies affect society economically and socially. **Without international standards for AI, it is difficult for these technologies to be widely implemented.**

The panel agreed that the Government should begin regulating now. Lack of consensus on the progress of AI technologies and their implications, in both the short-term and long-term horizons, is raising concern and uncertainty around critical issues - including those related to job losses, privacy, skill shortages, and national security. In many ways, this high degree of concern and uncertainty also inhibits the unlocking of AI's mass opportunities to raise UK's growth rate and productivity.

Given the scope and scale of AI impact, **Government must commit to taking imminent action**. Starting by drafting regulation adapted for specific use cases of AI, policymakers can work to create an ecosystem of responsible and widely beneficial innovation and, consequently, build public confidence in these technologies.

One of the ways Government can help is to foster healthy and beneficial innovation. Dr Julian Huppert recognized that regulation should not restrict innovation but rather guide it to move forward in a socially beneficial path. Investment is key in this approach, but so is creating an ecosystem that attracts responsible innovation companies and top talent.

Rodolfo Rosini, CEO of weave.AI, asked the Parliamentarians to set the framework for such an ecosystem.



Exponential growth in technology is not a new concept. It was first experienced by humanity 2.5 million years ago with the invention of the Oldowan, the oldest stone tool. It was followed by the controlled use of fire 700k years Before Present, then language 200k Before Present, the wheel, continuing its

acceleration through the industrial age and into our current day.

Look the results: population growth, energy generation, GDP, GDP per capita, and information produced. All exponential growth enabled by technology.

This means that in the next few Parliament terms technology acceleration will bring fundamental changes to society, and you are going to have to debate legislation about the rights of a living being that is not based on DNA, or about the self of one person merging with the self of another to create a new organism, communicate with other mammalians or even cephalopods in their own language, discover that privacy as a concept might cease to exist when we read minds at a distance, or have computers that can perform calculations in parallel universes.

Everything I mentioned has existed in the lab for close to a decade, there are even teams building commercial ventures for some. This is the present.

And in this, AI is a rising tide lifting all boats.

It's not just that AI in itself is important. It's that applied AI is changing the world by acting as a force multiplier to other technologies.

This change is coming just from one branch of AI called Machine Learning.

It is very important that whatever legislation you are considering it cannot be based on the limitations of Machine Learning of today because they might not hold true in a few years.

As technologists, we are in control of the future as we are the ones inventing it. But are not in control of the timing of it.

Today AI is a land grab that will overwhelmingly benefit the winners in this race.

The UK is the best country in the world for global contributions to science and technology.

But it also has 100 years of history of failing to capitalize and generate wealth on those discoveries.

In conclusion:

We are at an inflection point and Britain is one of the leaders in AI.

But that could soon change.

China had its Sputnik moment and is making AI research an absolute priority after seeing the world's best players crushed at Go. The United States had their first congressional hearing on AI last week. Russia said that AI leaders will rule the world.

This calls for a national strategy on AI. Things like increasing funding, data sharing, like opening up the NHS data, forcing the private sector, the military, and crucially the intelligence community to adopt new technologies.

But having a national strategy on AI is not enough if we come from a place of weakness. We must

address lack of entrepreneurship, necessary immigration, jobs destruction, low productivity and poor education to align the whole country toward the goal of ensuring that our leadership in this field is not lost.

It's about securing our future wealth, about the defence of our borders, and our culture. It is about the survival of our society.

Rodolfo Rosini called for a coherent approach to tackling AI issues. He prioritised the UK Government to focus on:

- Current lack of entrepreneurship in the UK
- Immigration policies preventing talent from studying/working in the UK
- Potential jobs destruction
- Low productivity
- Poor education, not adequately preparing generations for the future

Also, as discussed in APPG AI Evidence Meeting 3, the Government must also scan the current legislative landscape and identify any gaps or areas that need to be updated to fit the current day demands.

Data, the fuel of AI technologies, is a policy area of high concern at the moment. Dr Sandra Wachter, Postdoctoral Researcher in Data Ethics and Algorithms at the Oxford Internet Institute, focused on the impact of the GDPR in the UK economy and society.



Dr Sandra Wachter

Postdoctoral Researcher in Data Ethics and Algorithms, Oxford Internet Institute

What is good about the GDPR: We will be able to contest automated decisions if they are solely automated and if they have legal or significant effects (Art 22). Being able to contest is a major step forward to increase transparency and accountability.

What is problematic about the GDPR: the right to contest will only apply if the process of making the decision is solely automated. meaning if there is no human in the decision making loop. Therefore, by putting a human in the decision-making process one can create a loophole which prevents the right to contest from applying. We had this problem in the past e.g. Germany. A person was denied credit based on an algorithmic generated credit score. He wanted to contest this decision. The judge denied

this because the algorithmic generated decision was communicated by a human and not by an algorithm. Therefore, the decision-making process was not seen as solely automated anymore, even though the bank fully relied on the algorithmic decision. But there is another hurdle to successfully challenge decisions. The framework does not mandate a legally binding right to explanation. That means I have no legally guaranteed right that the reasons for a decision are explained to me. This can affect the ability to contest decisions as I will have no grounds to assess if the decision was lawful.

How to improve it: The UK has the opportunity to lead the way and set the standard across Europe. One way to improve the framework is to return to the phrasing of a prior draft of the European Parliament (EP). The EP proposed that the right to contest should apply when decisions are solely or predominantly based on algorithmic processes. Further the EP suggested to create a legally binding right to explanation. By slightly changing the wording we can close these accountability gaps.

Action Points

The public sector must guide safe and beneficial innovation, and hence ensure policies are forward-looking, coherent, and robust. The fifth APPG AI Evidence Meeting came up with the following action points the UK should focus on in the construction of AI governance structures, in both the private sector and public sector realms.

| Theme | Action Points |
|--|--|
| | Set norms and assessment indicators for the 'social value' of an AI technology. |
| | Create financial and non-financial incentives for companies to create and use AI with social value. |
| Al governance depends on corporate responsible | Make companies accountable for the decisions made by the algorithms that they use. |
| innovation. | Encourage corporate internal governance structures that ensure responsible innovation (i.e. Ethic Boards). |
| | Encourage partnerships (i.e., Partnership on AI, OpenAI) amongst companies to explore the social implications of AI. |
| Al governance relies on | Adapt a coherent AI national strategy. |
| and regulation. | Gather evidence across industries and sectors, and pass regulation targeted for specific use cases of AI. |
| | Set up monitoring and auditing systems to ensure AI benefits the wider humanity. |

Acknowledgements

The All-Party Parliamentary Group on Artificial Intelligence (APPG Al) 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, Deloitte, EDF Energy, KPMG, CMS, Oxford University Computer Science, PwC - enable us to raise the ambition of what we can achieve.

The APPG AI Secretariat is Big Innovation Centre.



Contact details

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.

www.appg-ai.org www.biginnovationcentre.com