

APPG AI EVIDENCE MEETING 7 – NEXT STEPS

I. DETAILS

- Date: 5 November 2018
- Time: 5:30 – 7:00 pm
- Location: Committee Room 4A, House of Lords
- Participants: **135 registered attendees**

II. PURPOSE

The All-Party Parliamentary Group on Artificial Intelligence (APPG AI) was set up by co-chairs Stephen Metcalfe MP and Lord Clement-Jones CBE to explore the impact and implications of Artificial Intelligence.

In 2018, the APPG AI has decided to focus on building a roadmap to understand the practical steps for addressing key AI implications. The group has prioritised six policy areas: data, skills, accountability, innovation & entrepreneurship, infrastructure, and trade. Each meeting will explore one of the six policy areas' respective economic, social, and ethical implications.

Evidence Meeting 7 concentrated on: **Next Steps**

III. SPEAKERS

- **Scott Steedman**, Director of Standards, BSI
- **Prof. John McDermid OBE**, Director, Assuring Autonomy International Programme, University of York
- **Christina Blacklaws**, President, The Law Society
- **Adrian Joseph**, Partner: AI & Advanced Analytics, EY
- **Dr Spiros Denaxas**, Associate Professor, UCL

IV. QUESTIONS FOR INSPIRATION

- What are the practical steps of setting international rules, norms and standards?
- What is our vision of the new AI and data driven world?
- What does it mean to be human in 2025 and 2052?
- Is the roadmap national or international?

V. BACKGROUND: SETTING THE SCENE

Artificial intelligence (AI) has quickly become the most powerful narrative of our century. Its impact has been seen across regions, industries, and sectors.

Although AI offers many opportunities for both our economies and our societies, it simultaneously raises many concerns - related to matters including security, inequality, privacy, employment, and education.

Recently, countries worldwide have been launching national AI strategies to reap the benefits of AI technologies and protect their nations from potential harms. However, it has become very clear that many issues related to AI cross national borders. No nation alone can address these issues without working together. While each country must consider its specific needs, we need a global framework to help us solve complex and pressing global issues.

Global coordination is essential to truly address the heart of these concerns. Policymakers and other stakeholders need to coordinate in order to champion AI and its implications on society.

[Speaking at the World Economic Forum in Davos](#), the UK's Prime Minister said: "When technology platforms work across geographical boundaries, no one country and no one government alone can deliver the international norms, rules and standards for a global digital world."

Organisations like the IEEE Standards Association and the British Standards Institution are bringing together stakeholders across industries and sectors to build standards guiding the development and deployment of these technologies. International organisations including the OECD and the United Nations are forming high-level groups to further explore AI's impact on an international scale. Governments are forming partnerships, committing to collaboration when addressing some of these issues related to cybersecurity, taxation, and data regulation.

Together, stakeholders worldwide are trying to agree on a vision our policies and strategies should aim to move towards. As AI transforms nearly everything around us, we need to convene individuals from different backgrounds to discuss our vision for a new AI and data driven world. We are in a pivotal time in history in which we can decide what our future looks like – what good looks like and what good doesn't look like.

The very essence of what it means to be human is changing as AI becomes an increasingly bigger part of our daily lives. Children are now being brought up in a society in which technologies affect them from the very day they are born. AI technologies are part of the homes they grow in as well as the teaching environments they learn in. Data is being collected from the toys children play with, the learning material they engage with, the social platforms their parents are subscribed to, the health providers they visit, and much more. As this generation enters adult

live, it follows that their lives will be dramatically different from those of today. What it means to be human will be completely different.

VI. MEETING OVERVIEW

On the 5th of November, the APPG AI brought together policymakers, industry representatives, academics, philanthropists, and members of the public to discuss next steps in our journey towards an AI and data-driven world. There were 135 individuals who registered to attend the evidence meeting. Five speakers were invited to provide their insights on questions around international norms and standards, what it will mean to be human in the future, and the debate between national and international roadmaps.

Stephen Metcalfe MP welcomed the attendees. After providing a short overview of the APPG AI programme for 2019, he asked the panel to provide the Parliamentarians and the wider audience with their views of what our next steps should be.

Scott Steedman, Director of Standards at BSI, spoke first on the work the British Standards Institute is doing to provide the infrastructure for the standards needed around AI technologies. He emphasised the need for standards in AI ethics before the technology becomes ubiquitous. Furthermore, he urged stakeholders to prioritise global standards first, regional standards second, and national standards third. Addressing the Parliamentarians, Scott urged the government to engage with standards bodies and the wider community to ensure the new standards developed reflect the interests of the society.

Professor John McDermid, Director of Assuring Autonomy International Programme at University of York, spoke next. Professor John argues that rules, norms and standards (especially at a technical level) tend to be domain specific. Government must collaborate with organisations like BSI and ISO to set these but should also be conscious that technology moves fast and standards take a longer time to be create. Therefore, he spoke about the importance of mechanisms such as Publicly Available Specifications (PAS) to develop material quickly and to evolve it as the technology changes.

The third speaker was President of The Law Society, **Christina Blacklaws**. Christina's evidence reflected the work of the Technology and the Law Policy Commission to examine the ethical implications of artificial intelligence (AI). She summarized most of the work they've done in three points: impact, fairness, and expert approval. She reminded the audience that moving forward we need a multidisciplinary approach that reflects a diverse range of voices.

Ernst and Young's **Adrian Joseph** spoke fourth on the panel, agreeing with the others on the need to approach these issues through an international lens. He focused on four myths around AI, five predictions for the future, three challenges for the economy and the society, and three recommendations for policymakers to take on. He suggests policymakers help build trust around AI, focused on values of: Performance, Bias, Resiliency, Explainability and Transparency. Furthermore, Adrian called for the creation of an ethics code of conduct, and investment into the creation of modern skills for younger and older generations.

Last to speak was **Dr. Spiros Denaxas**, Associate Professor of Biomedical Informatics at University College London. Spiros introduced two big challenges stakeholders must urgently address. He called for increased data access and the building of a talent pipeline of individuals empowered to safely develop and deploy AI. Focusing on healthcare, he asked the Parliamentarians to invest in a robust national infrastructure which empowers patients to see what data are recorded, when, by who and how are they being used for their care. He also asked stakeholders to promote capacity building in skills from relevant social sciences that are required to understand the ethical, social and political challenges of AI in healthcare.

VII. WRITTEN EVIDENCE – SCOTT STEEDMAN

Short Biography of Speaker

Scott Steedman, Director of Standards, BSI

Scott Steedman joined the BSI Group in January 2012 and was appointed to the Board in October 2012. An engineer by background, he started his career at the University of Cambridge before moving to industry where he spent over 20 years as a consultant working in the Built Environment sector. He was a Non-Executive Board Member of the Port of London Authority from 2009 to 2015 and served as Vice-President (Policy) for the European Committee for Standardization (CEN) from 2013 to 2017. He is currently Vice-President (Policy) of the International Organization for Standardization (ISO).

Written Evidence

BSI is appointed by HMG as the National Standards Body, providing the infrastructure for UK experts to participate in international, European and national standards development, managing the catalogue of industry standards needed for the UK economy, and representing industry, government and society, reflecting the public policy interest.

Comment is invited on the three interconnected areas of “international rules, norms and standards”. BSI’s work as the UK standards organisation supports all these tools by offering WTO compliant standards that can be used to underpin regulation, to support performance based co-regulatory approaches and to provide opportunity for organisations to deliver aspirational performance independently of regulation.

BSI’s policy is to approach standards in the same way as BSI’s stakeholders, prioritising international standards first, then regional and then national.

BSI standards are developed in accordance with international rules, meeting the objectives of the WTO’s “Technical Barriers to Trade” Agreement and WTO principles of open public consultation, stakeholder engagement and consensus.

Artificial Intelligence (AI) is a developing area of strategic importance to the UK where there is a widely recognised need for standards.

There has been widespread recognition of the need for standards in AI ethics before the technology becomes ubiquitous.

BSI and the US organisation IEEE are working together as the driving force behind the 'OCEANIS' initiative (Open Community for Ethics in Autonomous and Intelligent Systems). BSI is one of 12 founder members (which are all National Standards Bodies or Standards Development Organisations, such as IEEE). The members will use OCEANIS to provide a high level global forum for discussion, debate and collaboration for organizations interested in the development and use of standards to further the development of autonomous and intelligent systems, with a focus on ethics.

In various fora there have been numerous areas identified where there is a need for standards: aiding regulatory compliance in healthcare, safety, bias, privacy, etc. Ultimately, BSI as the UK standards organisation is here to take the areas where government, industry and consumers (as stakeholders together) need to develop consensus standards and ensure the mechanisms are in place for that to happen.

Given the global nature of technology development the goal for all participants is international standardisation which incorporates the views of the widest range of stakeholders and has international recognition. Practically, given the speed of change and the potential societal impact of AI, there is a need to develop and pilot standards at the European and national level that can meet short term needs and become base documents for international standards in due course. One such example of a national activity is the British Standard 'Robots and robotic devices, Guide to the ethical design and application of robots and robotic systems', published in 2016.

The UK government has been praised for the approach taken in the AI Sector Deal and for driving the role of bodies such as the Alan Turing Institute and the Centre for Data Ethics and Innovation. That said, more needs to be done if the UK is to assert leadership in AI standards.

In terms of the "practical steps" that government can take when it comes to standards HMG needs to get involved directly with the wider stakeholder community already engaged in this work, to bring their interests in regulatory policy for AI to the standards community. As it has already done in other sectors, if HMG sees global standards' leadership as being an important part of UK leadership in aspects of AI such as governance and public acceptance, government needs to invest to develop standards that will be trail-blazing through a strategy and roadmap that the community can endorse.

VIII. WRITTEN EVIDENCE – JOHN MCDERMID

Short Biography of Speaker

Professor John McDermid, Director of Assuring Autonomy International Programme at University of York

John is a Professor of Software Engineering and a member of the High Integrity Systems Engineering Group (HISE) within the Department of Computer Science at the University of York. He was Head of Department from 2006 to 2012.

His primary research interests are in high integrity computer systems, especially in safety and security. His work has influenced industrial practice both directly and via standards. He has taught extensively at postgraduate level, including on continuing professional development courses for industry.

Written Evidence

I have over thirty years' experience in research, education and consultancy on safety of complex computer-controlled systems. I currently run the Assuring Autonomy International Programme funded by the Lloyd's Register Foundation and the University of York. The Programme is addressing the global challenges in assuring the safety of robotics and autonomous systems (RAS). It is conducting research on the technical issues of safety of RAS using Artificial Intelligence (AI) and Machine Learning (ML), developing educational and training materials, supporting projects developing real-world demonstrators to address assurance and regulatory issues, and influencing the development of standards and regulations for safety of RAS.

What are the practical steps of setting international rules, norms and standards?

Rules, norms and standards tend to be domain specific, e.g. different in the maritime sector to aerospace or automotive.

At the technical level, e.g. assurance of ML, there is merit in developing domain independent norms and standards, and bodies such as the British Standards Institute (BSI) can form the bridge to international bodies, e.g. the International Organisation for Standardisation (ISO).

At the regulatory level, rules and norms must be developed in a domain specific way as, for example, the criteria for avoiding other vehicles (sense and avoid) are necessarily different to those at sea (although they are somewhat related). Further, the bodies who have regulatory authority are different, and often international, so co-ordinating across domains would be very challenging, and probably of limited benefit.

Technology moves fast, and standardisation often moves slowly. Thus it is important to use mechanisms such as Publicly Available Specifications (PAS) to develop material quickly and to evolve it as the technology changes. Pragmatically: work locally; develop PAS or similar; influence globally. To do this requires support, as it is necessary to establish collaborations between organisations with different skills. The Regulators Pioneer Fund and the Industrial Strategy Challenge Fund are potentially important mechanisms for enabling this.

What is our vision of the new AI and data-driven world?

It will be a net benefit to society, if enough attention is given to safety and acceptability of the technology. AI in the workplace, and embedded in systems such as factory robots or healthcare assistants, can relieve people of boring and repetitive tasks, remove people from harms way, support the elderly and infirm in independent living, and so on. However, to achieve these benefits requires work on safety and regulation (see above) and on public acceptance of the technology.

The forms of AI that will be widely used will be problem-specific, not “general AI”. Systems will be good at particular tasks, e.g. driving cars, monitoring patients’ health, or machining complex 3D shapes, but not able to engage with humans at a “peer level”, e.g. to take part in the full range of human activities. If such “general AI” is possible, it is a long-way off.

Like previous technological revolutions, there will be a substantial change in the nature of work, and there may be some resistance to change, but it seems unlikely that the level of employment will drop dramatically. However, there is likely to be a need for “up-skilling” as RAS take on relatively mundane and repetitive tasks and also re-skilling as AI and RAS bring about changes to the role of humans, e.g. to monitoring autonomous vessels rather than piloting them.

What does it mean to be human in 2025 and 2052?

The same as now! The environment in which humans live and work will be richer due to the development of AI, including RAS. It is not realistic to believe that such systems will supplant humans, or make humans subservient. Well-defined (specific) AI and RAS, solving particular problems, will be supportive and assistive, if designed to be so. AI and RAS will be a force for good in helping humans to have fulfilling lives, but not alter what it means to be human.

Is the roadmap national or international?

International. Having international rules and norms is important economically, to avoid unnecessary costs in assurance, and obstacles to the uptake of technology. There are international bodies in many domains, e.g. the International Maritime Organisation (IMO), the overarching body for the marine sector. The UK can (and must) take a lead, but it is essential to influence the international community to avoid UK suppliers being disadvantaged by working to standards that are not harmonised across the globe.

IX. WRITTEN EVIDENCE – CHRISTINA BLACKLAWS

Short Biography of Speaker

Christina Blacklaws, President, The Law Society

Christina is the 174th president of the Law Society of England and Wales and the fifth woman to hold this office. During the presidential year 2018/19, Christina's main priorities are: (a) Diversity and inclusion - focusing on social mobility and women in

leadership in the law, (b) The future of the profession - including legal technology and leading the LawTech Policy Commission, (c) Promoting access to justice for all, and (d) Mental health and wellbeing in the legal profession.

Written Evidence

The Law Society of England and Wales have launched our Technology and the Law Policy Commission to examine the ethical implications of artificial intelligence (AI).

The Commission is an ongoing exploration of the impact of technology and data on human rights and the justice system, examining the future use of algorithms in the justice system of England and Wales, and the implications of the use of algorithms on human rights, fundamental freedoms, and the rule of law.

The issues raised by witnesses at this stage can be crystallised into three points:

1. Impact: The use of algorithms in the justice system can have enormous benefits or devastating consequences, depending on how they are coded, the data they use, the situations and circumstances they are used in, the knowledgeability of the users, and the perceived fairness and legitimacy of the algorithms.
2. Fairness: The key question is how can we ensure that algorithms are used in the right areas/spheres, using the right datasets, in a way that is fair, non discriminatory, and that there is proper human oversight and remedies for the users and those on the receiving end of decisions.
3. Expert approval: There are various models that can be used to achieve those aims, but so far we have heard a lot of experts suggesting we need a multi-disciplinary panel of experts from various sectors that would approve, or do an accreditation, of specific algorithms for use in specific circumstances, according to specific set of criteria, both ethical and practical.

We will publish our final report next year on the findings of the Commission.

X. WRITTEN EVIDENCE – ADRIAN JOSEPH

Short Biography of Speaker

Adrian Joseph, Partner: AI and Advanced Analytics, EY

Adrian is currently a partner at Ernst and Young in the UK Financial Services team, leading the Data and Analytics group. Previously, Adrian was the head of Client Services, Google Analytics, EMEA. Prior to this he was head of Search Advertising for Northern and Central Europe and Managing Director of Google Enterprise EMEA. He joined Google over 10 years ago with 15 years of sales, marketing and general management experience gained within industry and management consultancy. Prior to

Google, he spent 6 years at Ford Motor Company in sales and brand management roles in Europe, was a principal consultant at AT Kearney and was on the main board of Trafficmaster Plc.

In January 2012, Adrian was appointed chair of Race for Opportunity, the race campaign of Business in the Community, a Prince of Wales charity, which stands for responsible business and strives to ensure that all ethnic groups are appropriately reflected in the UK workforce. In 2014 he was appointed to the board of Business in the Community, whose president is the Prince of Wales.

Written Evidence

AI Fact vs. Fiction

Myth 1: General AI – Machines will be capable of experiencing consciousness and become self-aware

Fact: Use Case Driven AI – Narrow AI is the reality when it comes to AI in most business contexts.

Myth 2: AI will take all the jobs – Humans will be fully replaced by robots in the workplace. As robots will do all of the work, unemployment will increase rapidly.

Fact: The workforce will be impacted in some way – Some jobs will go; some will be partially automated; some will be barely affected.

Myth 3: AI will only be for the technological elite – AI will only be accessible by elite companies such as Google, Amazon, and Facebook, and will have no impact in other industries or companies.

Fact: AI can be accessible and utilised by many people and many types of companies – AI will only be developed by a (relatively) small number of tech firms, even though it can be widely accessed by individual consumers and businesses.

Predictions for the Future

Prediction 1: AI will be collar-blind

- AI will have an impact on many different levels, it's not limited to blue-collar work.

Prediction 2: Person + machine > person or machine

- AI will have the most impact when it used by people, rather than when it replaces them.

Prediction 3: There will be uneven adoption of AI

- AI will not be utilised by all industries or countries at the same time or speed.

Prediction 4: Voice assistants may be one of the greatest AI disruptions

- The greatest disruption (opportunity and challenge) is going to occur through mass adoption of relatively 'simple' virtual assistant technologies.

Prediction 5: Convergence of technologies

- Future developments in AI will increasingly hang off the platforms/applications that have already been or are likely to be very widely adopted

Challenges

Challenge 1: Data

- Generally, the state of data in business is that much of the data currently in use is lives in legacy systems, and the data quality can often be unreliable.

Challenge 2: Trust

- It is important to create a culture of trust around AI. People must be able to trust both the data and the algorithms.

Challenge 3: Skills

- Need for workers with higher cognitive skills, social and emotional skills, advanced IT and programming skills.

Recommendations

Recommendation 1: Trusted AI framework

- There are five key attributes necessary to sustain trust in AI: Performance, Bias, Resiliency, Explainability and Transparency.

Recommendation 2: Governance model

- Need to build universal but culturally adaptive ethical guidelines, codes of conduct and codes of practice.

Recommendation 3: Education & Skills Development

- Adjust the curriculum to embed STEM and creative skills at an early age to start creating the foundations for the jobs of the future.
- Additionally, there needs to be programs to re-educate and reskill the workforce

XI. WRITTEN EVIDENCE – SPIROS DENAXAS

Short Biography of Speaker

Dr Spiros Denaxas, Senior Lecturer at Institute of Health Informatics, UCL

Spiros is a leading biomedical data scientist operating at the intersection of computer science, clinical research and healthcare delivery. He leads the UCL Institute of Health Informatics Data Science Lab and his research focuses on applying and evaluating machine learning and data-driven methods for analyzing big biomedical data and answering clinically-significant questions. He has a proven track-record of understanding and translating real-world questions and challenges into statistical learning problems across industries. He has demonstrated success in developing cross- functional collaborations by bringing together stakeholders across engineering and healthcare and succinctly conveying complex technical concepts to non-technical audiences. His research has informed international clinical practice in the prevention and treatment of millions of patients with, or at risk of, cardiovascular disease.

Written Evidence

Good afternoon and thank you for the opportunity to address this All Parliamentary Party Group meeting on Artificial Intelligence on the vision of a data-enabled, AI-driven future.

My name is **Spiros Denaxas** and I am a computer scientist by background. I am currently an Associate Professor of Biomedical Informatics at University College London. I am also an investigator in the newly-created national Health Data Research UK Institute (<https://www.hdruk.ac.uk/>). HDR UK is a joint investment led by the Medical Research Council, together with the National Institute for Health Research (England), the Chief Scientist Office (Scotland), Health and Care Research Wales, Health and Social Care Research and Development Division (Public Health Agency, Northern Ireland), the Engineering and Physical Sciences Research Council, the Economic and Social Research Council, the British Heart Foundation and Wellcome.

My **lab** (<http://denaxaslab.org/>) focuses on analyzing electronic health records which are data that are generated during our interactions as patients with the NHS such as for example when you visit your General Practitioner or undergo surgery in hospital or end up in A&E. We analyze these data using, amongst other, machine learning and AI approaches, to do everything from uncovering hidden patterns of diseases, to building better tools to predicting when somebody is more likely to develop diabetes or to be admitted in hospital with a heart attack.

We have heard from our expert colleagues on the ethical considerations, legal challenges and technical standards so my evidence today comes as someone who is on the ground, building and applying some of the approaches for analyzing large amounts of health data for improving human health and healthcare. Importantly, I also provide evidence today as somebody who is educating the next generation of data-researchers in my capacity as Programme Director for the MSc in Health Data Science at UCL. This next generation of researchers will eventually be the enactors and consumers of these approaches.

What will an AI-driven healthcare system look like in 2052? Unsurprisingly, this is a very hard question to answer...! There are broadly two schools of thought around what an AI future holds, the utopian one and the dystopian one. If I were to paint an image of a utopian NHS, it would be a **rapidly learning** system where every single patient interaction is a learning point. The NHS will be using AI for converting huge amounts of heterogeneous health data into actionable knowledge, enacting this knowledge and rapidly evaluating its impact in terms of treatment, prevention and measurable benefits to patients. This process, which at the moment takes years, should in the future take minutes and enable the delivery of personalized, efficient and safe care to everybody.

Today, I would like to raise just two of the short-term challenges that we face today on our path to a rapidly learning NHS:

1. Data access

At the heart of AI, or any other data-driven approach for that matter, lie data. To achieve a rapidly learning healthcare system, it's critical to have streamlined access to the data itself at a national level. The UK has the unique potential of having high-quality, longitudinal health and social care data from birth to death on a national level, for 65 million citizens. At the moment, these data are siloed which make access to both academia and industry researchers very challenging. Our definition of what these data are is also likely to change with whole genome sequencing becoming a commodity, where all patients get sequenced at birth or with patient's bringing their own data from mobile phones, smartwatches or home sensors and integrating them with the healthcare system.

2. Strengthening the talent pipeline

The skills required to build and evaluate reliable and responsible AI approaches, especially in healthcare, are much broader and diverse than STEM subjects that the current focus is on. The focus is often on maximizing algorithmic accuracy without considering the implications of deploying the algorithm in routine care early on. Future AI researchers should not only have the solid technical background but also have a solid understanding the underlying healthcare processes that generate the data which in turn would enable them to understand recognize the biases that exist how to address them. Similarly, clinicians should be aware how the data they record and generate during care effectively influence the care they will deliver in the future. We should be encouraging lifelong learning, tackling new emerging challenges using new emerging technologies through a flexible and adaptable curriculum or we risk, as we say in machine learning, to overfit i.e. end up with a solution that is good now but will not adapt to the future needs and challenges.

To overcome these challenges, my recommendations are:

- The UK should Invest in a robust **national infrastructure which empowers patients** to see what data are recorded, when, by who and how are they being used for their care.

Patients should be aware when a decision with regards to their care is based on an algorithm or not. Furthermore, patients should be able to decide who uses what data for research and what are the benefits of such research to their health or the health of others.

- Comparably to genomics, the branch of molecular biology concerned with the structure and evolution of our genetic material, we need to establish a complementary **phenomics discipline** to systematically investigate, refine and redefine disease classification from information about health states (things that are observable, e.g. symptoms), health traits (things that can be measured, e.g. blood pressure) and their variation over time.

- We should **proactively promote capacity building in skills from relevant social sciences** that are required to understand the ethical, social and political challenges of AI in healthcare. In parallel, we should advocate and **reward reproducible research**, enable the sharing of algorithms and the use of open standards in research.

- Finally, we should encourage and reward the next generation of health researchers to go out there and co-design, **co-create AI tools and approaches with clinicians, patients and the public**.

To conclude, the potential of an AI-driven healthcare system is simply fascinating. The NHS will be able in the future to rapidly learn and act using large amounts of heterogeneous data and deliver better, personalized safer care.

XII. AWKNOWLEDGMENTS

Our supporters - Accenture, Barclays, BP, British Standards Institution, CMS Cameron McKenna Nabarro Olswang, Deloitte, EDF Energy, Ernst and Young, KPMG, Microsoft, Oxford University Computer Science, and PwC – enable us to raise the ambition of what we can achieve. The APPG AI Secretariat is Big Innovation Centre.

