

AN ETHICAL AI FUTURE

GUARDRAILS & CATALYSTS TO MAKE ARTIFICIAL INTELLIGENCE A FORCE FOR GOOD



***We should
regulate AI before
it regulates us.***

Yuval Noah Harari,
Essay In The Economist,
20 May 2023

19th June 2023

This report was written by Oona Muirhead CBE, Fellow, Policy Connect.

Policy Connect
7-14 Great Dover Street
London
SE1 4YR

www.policyconnect.org.uk

The views in this report are those of the author and Policy Connect.
Whilst these were informed by the contributors to our inquiry, they do
not necessarily reflect the opinions of either individuals or organisations.

Contents

Foreword	5
Recommendations	6
Context and Key Findings	8
Looking Internationally	13
Governance for culture-change	16
A National AI Centre	19
Laws and Tools	23
Conclusion and Next Steps	25
Methodology	26
Acknowledgements	29

***How we can retain
power for ever
over entities far
more powerful
than ourselves.***

Professor Stuart Russell,
Reith Lectures 2021,
AI: A future for Humans

Foreword

The All-Party Parliamentary Group on Data Analytics (APGDA), in its first report in 2019, described artificial intelligence (AI) and data as our next major societal opportunity and challenge. Four years on, the APGDA continues to believe very strongly that the UK has a real opportunity to become a world leader in the ethical use of data and AI as a means of innovation in our lives and work. The purpose of this report is to set out how this can be achieved.

The extent of the opportunity and challenge is now part of our daily diet of news. The capabilities of generative AI, such as ChatGPT, have generated both excitement and deep concern. Excitement about AI's ability to support step changes in the way we use technology for beneficial purposes, such as life-changing medical treatments, but concern about the potential misuse of our personal data and the creation of fake facts that can damage individuals and society, such as through the undermining of democracy by 'bad actors'. We are told by some that we are on the cusp of major AI-based technological developments that have the scope to help humankind solve existential problems such as climate change, while others fear the next generation of generative AI could destroy humankind.

In this report we do not make predictions about where AI might lead. Instead, we aim to identify what kind of governance and regulatory ecosystem could help us now to achieve maximum benefits from AI and manage current risks such as deep fakery, as well as deal with big future risks as they become known.

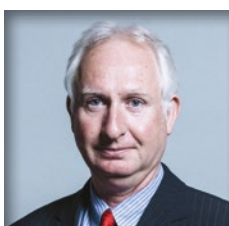
During the course of our inquiry the Government published its White Paper on the regulation of AI. We are pleased that the community now has proposals on which to comment. We agree with much of the White Paper's analysis of the problem and that it is important to introduce no-regret measures given the breakneck speed of technological developments. However, the solutions proposed in the White Paper neither remove the regulatory uncertainty that so many in the business and other sectors have stressed leads to loss of confidence and investment, nor manage the risks.

This report sets out a more comprehensive and practical set of proposals for regulation and governance, which will also turn into actionable mechanisms for the White Paper's principles. We have looked more broadly than its focus on regulators, and propose building on existing duties on companies and their Boards. Better strategic governance will be adaptable to any future direction in technological innovation, requiring both private and public sector bodies to make judgements about benefits and harms before introducing the results of research into society. And we continue to believe there needs to be a re-balancing of power and accountability between government, Parliament and regulation; Parliament has a critical role to play.

All participants emphasised the need for compatibility across jurisdictions through standards and tools, and want to see the Government re-engage internationally with pace and profile. Finally, many spoke about the importance of skills, and of the energy and climate impact of generative AI development and use; these themes may be the focus of the APGDA's follow-on work.

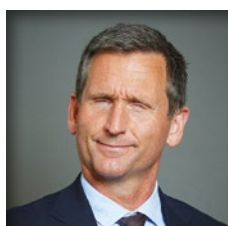
We would like to thank all those who participated in four tremendous evidence roundtables and sent in thoughtful written evidence. Huge thanks go to Zurich Insurance, EY, Bright Data and Jisc, whose sponsorship made this work possible.

Inquiry co-chairs



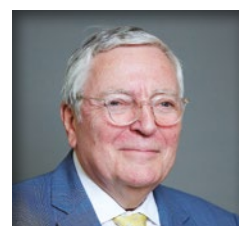
Daniel Zeichner MP

Chair of APGDA



Lord Chris Holmes

Co-Chair of APGDA



The Lord Clement-Jones

Vice-Chair of APGDA

Recommendations

LOOKING INTERNATIONALLY:

1. The Government must step up engagement with the EU, US, and other bilateral partners. The UK needs to regain its position in discussions about regulatory systems, to influence how these develop and ensure the UK is not out on a limb that will affect UK companies' commercial competitiveness.
2. In this engagement, the UK should seek to lead in its areas of strength, such as standards. For example, getting the UK Algorithmic Transparency Standard launched, tested, and then mandated by influential governments.
3. The UK should be at the heart of work towards an international AI Ethics Convention and supporting watchdog. At the practical level, the Government should advocate for common elements of governance and regulation across jurisdictions.

GOVERNANCE AND CULTURE CHANGE IN THE UK:

4. The Government should introduce statutory duties worded such as to require organisations to achieve the objective of 'doing no harm' through 'cultural embedding' governance; the definition of these duties should be subjected to meaningful consultation and testing.
 - a. A statutory Duty on an organisation's leadership to take account of the requirement to achieve 'doing no harm'. This could be tied into the design and bringing to market of 'high-risk' AI systems.
 - b. A statutory Duty, in cases where the AI system is judged to be 'high-risk' in terms of its impact on lives, for the leadership to have a 'go/no go' sign-off process, so that a deliberative decision to proceed with a significant new AI system is taken prior to release/use.
 - c. A requirement that the leadership team, at Board level, should include a person accountable for ensuring due diligence on AI and ethics. This might be a Chief Information Officer or Senior Responsible Individual (currently Data Protection Officer), or their equivalent in a small company.
 - d. A requirement, in organisations of over 250 employees, for an Ethics (Advisory) Committee (along the lines of Audit Boards or indeed as part of an Audit, Risk, and Ethics Board) to bring together technical, customer-facing, and ethics experts to support the balance of judgements in Board decision-making.
 - e. The option for pre-market, controlled real-conditions testing to be carried out in regulatory sandboxes to support Board 'go/no go' decisions on 'high-risk' systems. This would allow testing to eliminate bias and could help to identify the organisation's legal liabilities.
 - f. Approved/kite-marked standardised algorithmic and privacy risk assessment/impact tools (which have the potential to become global norms).
 - g. The option of engaging People's Panels (potentially shared within a locality such as Greater Manchester).

A NATIONAL AI CENTRE:

5. A single independent, central body with strong regulatory authority, properly resourced, should be established in statute. Its functions should be to:
 - a. Convene existing sector regulators (e.g., lead the Digital Regulation Cooperation Forum (DRCF)) and ensure AI principles are prioritised in their business plans and that their regulatory functions are properly resourced, proportionate, and agile. To ensure the right powers and resources are provided, the DRCF could also be placed on a statutory corporate basis.¹
 - b. Identify and fill regulatory gaps, and address overlaps between regulators.
 - c. Ensure regulators have the right roles, skills, and enforcement powers to deliver smart regulation.
 - d. Commission a registration scheme for professionals in the AI industry.
 - e. Provide guidance to the public and private sectors to encourage responsible uptake of AI.
 - f. Horizon-scan to anticipate AI developments/functionality/national ethical issues.
 - g. Initiate investigations into new AI deployments if it has reasonable cause to suspect 'harm'. It should have powers to demand information and/or the power to require another regulator to do this.
 - h. Initiate/oversee investigations into incidents, especially where regulatory roles are confused. This could include having powers, pending the investigation, to request an immediate injunction on a system from the Secretary of State if there is evidence that serious harm is already occurring.
 - i. Provide guidance to the public to help them guard against AI harms; alongside a single Ombudsman also to be established (potentially the Information Commissioner's Office).
 - j. Make an annual report to government and Parliament, ideally to a new Joint Parliamentary Committee including members from the relevant Select Committees.

LAWS AND TOOLS:

6. A very short Bill should be developed to implement these recommendations, put forward as a draft for meaningful consultation with stakeholders and to ensure cross-party support, and tested against real-life scenarios including with citizens. Getting this right should ensure that companies only need to meet one set of obligations in order to satisfy compliance requirements in other jurisdictions.
7. The suite of standards currently under development by the international community (ISO etc) and nationally (BSI, the AI Standards Hub) should be trialled in a number of sectors and made part of a robust accreditation system.
8. The Government should lead the way in driving change through procurement policy.

¹ As recommended by the Joint Committee on the Draft Online Safety Bill and the Lords Communications and Digital Select Committee

Context and Key Findings

This third inquiry in our series on data, AI and ethics involved seeking expert views from parliamentarians, industry, academia and the third sector in a series of roundtables and bilateral meetings, alongside a written call for evidence. Our evidence gathering took place at a time of significant policy development and huge technological change impacting on the public perception of, and trust in AI.

The use of AI is not new. AI has been deployed to improve productivity for many years in well-regulated sectors such as finance, insurance, and medical diagnosis and products, supporting more efficient human decision-making, freeing up time for decision-making and customer engagement. However, the recent launch of ChatGPT, based on generative² AI able to produce original content, closely followed by Google's Bard internet search chatbot and other such tools, has contributed to AI being in the news on an almost daily basis.

ChatGPT has in particular sparked widespread discussion over the risks as well as benefits of AI able to hold human-like conversations (including for criminal purposes such as cyber grooming), and to be used to create deep fakes instantly through image generation tools. On 22 March 2023, the Future of Life Institute published a letter co-signed by Elon Musk, the Apple co-founder Steve Wozniak, and Professor Stuart Russell, and over 1000 other scientists, calling for a six month moratorium on developing models more powerful than GPT4, on the basis that AI systems with "human-competitive intelligence" pose profound risks to humanity.³

We do not pick up the proposal for a 'pause' in our report as we agree with those who say it would be unenforceable and may simply give a greater lead to those less concerned about ethics. This was, however, just the start of a series of warnings about AI. In May 2023, Geoffrey Hinton, one of the so-called 'godfathers of AI', said that "It is hard to see how you can prevent the bad actors from using [AI] for bad things" while the CEO of Google admitted that the rapid pace of AI development could be "very harmful if deployed wrongly".⁴ These warnings – albeit not accepted by everyone working in the AI field - are reminiscent of the warning about the atom bomb and nuclear power after the Second World War, which led to international treaties and rigorous domestic regulation.

There have also been very rapid policy developments over recent weeks. The EU had previously been at one end of the policy spectrum, being the first to seek to take an international lead on legislation through the publication by the Commission of the draft EU AI Act on 21 April 2021.⁵ This sets out legal requirements that kick in at the point when AI enters the EU market, thereby affecting all companies trading in the EU whether based there or not. The AI Act, the broad elements of which were agreed by the EU Parliament on 11 May 2023, will place obligations on member states and impose new transparency measures on generative AI applications like ChatGPT. The EU Parliament debate included some toughening of the Commission's original proposals: MEPs voted in a blanket ban on remote biometric identification (live facial recognition) in public venues. Fines worth potentially billions of euros are on the cards for businesses that unleash manipulative AI tools or advanced facial recognition surveillance. AI systems will be regulated according to their perceived level of risk, and member states will be required to implement regulatory sandboxes so that AI systems – ideally following self-assessment by companies – can be stress-tested for regulatory compliance before entering the market. The Bill is to be put to a plenary vote of the European Parliament in June 2023, final terms will then be agreed. Trilogue discussions to bring together the Commission, Council and Parliament versions will start at the end of June/beginning of July and are anticipated to conclude towards the end of 2023. After the Bill becomes law there is intended to be a grace period of around two years to allow affected parties to comply with the regulations.⁶

² In this report we use the term 'generative AI' rather than 'foundational AI'. We are conscious that definitions are an area of concern as the term AI covers a very broad range of functionality, right up to artificial general intelligence (AGI), defined as a system capable of any intellectual task that a human being can perform.

³ Pause Giant AI Experiments: An Open Letter, online. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

⁴ Widely reported in the media, e.g., the Daily Telegraph <https://www.telegraph.co.uk/world-news/2023/05/01/geoffrey-hinton-ai-google-artificial-intelligence-chatgpt/>

⁵ Adapted from: The AI Act, online <https://artificialintelligenceact.eu/the-act/>

⁶ Reuters, 11 May 2023, <https://www.reuters.com/technology/eu-lawmakers-committees-agree-tougher-draft-ai-rules-2023-05-11/>

The US Administration, in late 2022, published a blueprint for an AI Bill of Rights based on five citizen-centric principles that aim to guide the design, use and deployment of automated systems.⁷ The Bill of Rights did not envisage enforcement by the Government; the intention had been that companies would voluntarily implement these principles. In early May 2023, however, following a realisation amongst policy makers that generative AI may be moving at a breakneck speed, the White House started taking a more interventionist approach, telling companies that “the private sector has an ethical, moral, and legal responsibility to ensure the safety and security of their products”, and indicating that the US Government would not lag behind in setting up new laws and regulations.⁸

In April 2023, China published draft regulations under which Chinese tech companies will need to register generative AI products with China’s cyberspace agency and submit them to a security assessment before they can be released to the public.⁹ Among other things the regulations require tech companies to ensure that the training data used reflects the “core value of socialism”. This suggests a key aim of the regulation is political, to ensure AI plays its place in what President Xi calls “the great rejuvenation of the Chinese people”, through a combination of tightening his grip on the party and revitalising and revalidating it as the Leninist-Mandarin vanguard of the people.¹⁰ President Xi’s stated ambition is for China to be the global leader in AI by 2030.

In the UK, the Government followed up its national AI Strategy of 2021 with an AI Action Plan and AI Regulation policy paper in 2022, and the publication in March 2023 of its long anticipated White Paper *‘A Pro-Innovation Approach to AI Regulation’*¹¹, out for consultation until 21 June. The White Paper advocates for an incremental approach, using cooperation between existing regulators, with some (as yet unclear) form of ‘central function’ to carry out supporting functions including monitoring regulatory effectiveness, awareness raising, and promoting international interoperability. Alongside the White Paper, the Government launched the AI Foundation Model Taskforce with £100 million, saying that it intends the Taskforce to lay the foundations for the safe use of foundation models across the economy.

Our inquiry was initiated prior to these policy developments, reflecting the ongoing focus of the All-Party Parliamentary Group on Data Analytics (APGDA) on the importance of AI for the UK both economically and societally. Events over recent months have emphasised the unanimous call from participants to the inquiry that the Government take action at scale and speed if the UK is to regain its place at the heart of global developments and ensure that research, investment, and commercial exploitation is not hindered by an absence of clear, lasting policy, and regulatory direction. The APGDA’s inquiry and this report should therefore be seen through the lens of contributing to the debate following the publication of the White Paper, with more work needed on the specifics of implementation.

⁷ US White House Office for Science and Technology, October 2022

⁸ Meeting between the Vice President and AI leaders reported in the Times on 4 May 2023 and the Telegraph on 5 May 2023.

⁹ Adapted from: China races to regulate AI after playing catchup to ChatGPT, Al Jazeera, online <https://www.aljazeera.com/economy/2023/4/13/china-spearheads-ai-regulation-after-playing-catchup-to-chatgpt>

¹⁰ The US-China Race for Artificial Intelligence, Harvard Kennedy School, Belfer Centre, August 2020, <https://www.belfercenter.org/publication/china-beating-us-ai-supremacy#footnote-015>

¹¹ Adapted from: A pro-innovation approach to AI regulation, online. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>

The key findings of our research are as follows:

- Clear unanimity on the importance of cross-jurisdictional compatibility from both UK based and multinational companies. Many believe that, given Brexit and the pandemic, as well as other preoccupations, the Government had dropped the ball on international engagement on AI in the last few years. On the plus side, there was strong agreement that the UK has much to offer and could yet be at the heart of international work. Responders pointed to global initiatives that could be built on such as the UNESCO recommendations on the Ethics of AI.
- In terms of the current situation on ethical AI, there was a general view that the use of ethics tools is very patchy; many judged the quality of ethical governance and practice not to be universally high. Many stressed the need for a holistic approach (ecosystem) focussed on the 'customer'.
- A strong concern was expressed that the benefits of regulation are underestimated by government. Many stressed that regulation is not the enemy of innovation and growth; far from stifling innovation, good regulation can enable and stimulate it. In that context it is interesting that a recent survey of 3,500 small and medium-sized businesses across six countries showed that British businesses are more cautious about using AI software in the workplace than their US or French peers.¹² A powerful example of the destabilising effect of the absence of clear rules was that of local authorities, who will be at the forefront of ethical AI questions given the range of services they provide. Many authorities are unwilling to use AI until it is clear how due diligence can be carried out and impacts audited.
- In relation to the principles set out in the White Paper, no-one could object to these. It was, however, pointed out that a gap analysis of what is currently based in law would show that while some principles have an existing legal basis (such as privacy), others do not. This puts the priority on tackling the gaps in the new regulatory framework: that is, on transparency and explainability.
- The regulatory framework was described as about much more than regulators, including: ethics debates with the public and citizen engagement more generally; involvement of employees in AI discussions and development within organisations; standards and audit; incident investigation; professionalisation of the sector; and accreditation (both against standards and certification of individuals in the profession).
- The perspective from industry about the adequacy of the current regulatory position in the UK varies. In general, those companies operating in well-regulated sectors with a mature governance infrastructure consider that sectoral regulation works well. This was the case for example for financial services, medical devices, and automated vehicles. However, concerns were expressed about the risks in relation to unregulated sectors or sectors where there is a confusion of regulators leading to regulatory overlaps and gaps.
- In addition, many - even those who judged regulation to be broadly fit for purpose in their own sector for the current state of AI - acknowledged that there are issues around transparency, explainability, and accountability in relation to third party/outsourced AI system development. For example, attention was drawn to the difficulty of testing for bias in third party systems. Regulation of third party-provided AI was considered a complex issue which will need careful working through. How to have confidence in third party algorithms is a particular concern for local authorities and public service provision. Councils are shying away from using AI if there is no transparency on the data and how algorithms are designed and trained.

¹² Reported in the Times newspaper 17 April 2023

- Concern was also expressed about the balance of power between big tech and small start-ups, and big tech and citizens. In both cases the citizen should be at the heart of governance and regulation frameworks, and measures to regulate big tech should not be used to overwhelm SMEs, or force them out of the marketplace. This includes in relation to the accessibility of big data sets.
- Overall, most respondents considered that some cross-sectoral (horizontal) measures are needed such as wider governance arrangements. Differences of view were more in the detail, such as whether horizontal measures should include an overarching cross-sectoral regulator.
- It was felt that horizon-scanning and assessment of new AI impacts and risks could not be left to individual bodies in the public and private sectors, each working on their own. It needed to be a core central function carried out at a strategic level, with a clear consensus about the definition of AI risk. This should include the capacity to debate big ethical issues for the future that can only be done on a cross-sectoral basis.
- In relation to the functions required within the new governance and regulation framework, there was a majority view that central functions should include the following: act as a socket for sectoral regulators; monitor and be ready to act proactively on new AI development and risks; incident investigation by AI specialists; robust enforcement action; accredited third-party certification against standards; a registration system for AI professionals; and trusted advice and guidance including simple guidance on the do's and don'ts of using AI in service delivery.
- There was also a strong call for a single point of contact in the form of a lead Ombudsman, and questions around whether this function could be fulfilled by the ICO. If not, another organisation would be needed.
- On legislation, a significant majority of respondents judged there is a need for a legal basis for the governance and regulatory framework. While many felt that the EU AI Act had a number of flaws, nonetheless there should be some legal requirements, based on the risk level to society and individuals. A strong argument was that legislators should learn from recent examples where legislation has been introduced only after significant harms have already occurred. For example, the harms to children from social media, gambling harms, and the forced installation of pre-paid meters. The Government should be in the business of preventative policy and regulatory work to avoid harm to individuals and society, not running to catch up belatedly. Legal obligations should be carefully designed so that if a company has delivered on them, it will also pass compliance requirements in other jurisdictions.
- The scope and length of legislation should learn lessons from the complexity of the Online Safety Bill; an AI regulation bill should be short and broad to allow for adjustment as the AI scene develops. The development of legislation should be in parallel with the rapid introduction and testing out of voluntary Codes of Practice, this would help ensure the legislation reflects effective governance measures. Contributors emphasised that the scope and timing of legislation would also need to dovetail with UK's international work, so that domestic and international approaches remained in step. It was suggested that there may be opportunity to use the Data Protection and Digital Information Bill as a legislative vehicle in the immediate term.

A number of the themes raised during the course of the inquiry will need further, detailed work following on from this report and the Government's White Paper. These include defining 'harm' and 'high-risk' in the development and consultation on legislation. There are also important aspects around how the education and skills system can support development of the skills that will be needed across the private and public sector alike, within regulators and – last but not least – the public, in order to fully exploit the opportunities of these technological advances and manage the risks.

While the tenor of discussion is often centred around 'how AI will take your job' the governance and regulation framework set out in this report will create many new advisory, consultancy and audit roles that will require highly skilled workers. Equally, the advantage that generative AI will give to cyber criminals to be able to unleash billions of hacking attempts individually tailored to their target means there will be a need for additional cyber professionals to give the public on-going and up to date information, advice and support on how to stay safe.¹³

Finally, another important strand raised in the evidence sessions was that of the very large energy and carbon footprint of generative AI. While as a tool AI might help mankind solve climate catastrophes, it contributes itself to the problem through its huge energy bill, and this needs to be assessed and tackled. In the time available and given the focus on the standards and regulation framework we did not feel we could do justice – in this inquiry – to these important issues around skills, capacity, and climate impact.

We need companies that develop powerful AI systems to be registered, governments need to be able to follow and audit them. That's just the minimum thing we do for any other sector like building aeroplanes, cars or pharma.

Professor Yoshua Benjio,
BBC Radio 4 Today Programme,
31st May 2023

¹³ On 3 May 2023 the Home Office published its Fraud Strategy, which includes a role for the National Cyber Security Centre (NCSC).

Looking Internationally

1. The Government must step up engagement with the EU, US, and other bilateral partners. The UK needs to regain its position in discussions about regulatory systems, to influence how these develop and ensure the UK is not out on a limb that will affect UK companies' commercial competitiveness.
2. In this engagement, the UK should seek to lead in its areas of strength, such as standards. For example, getting the UK Algorithmic Transparency Standard launched, tested, and then mandated by influential governments.
3. The UK should be at the heart of work towards an international AI Ethics Convention and supporting watchdog. At the practical level, the government should advocate for common elements of governance and regulation across jurisdictions.

Many jurisdictions are addressing policy and regulatory issues around the question of AI development and regulation. Many of the (largely voluntary) playbooks being developed have a lot in common in terms of the principles on which they are based, such as transparency over how the AI works and stronger data protection rights. However, the models vary significantly on whether they are government or industry-led, and therefore in terms of the level of independent regulation and audit envisaged. At one end of the spectrum of regulatory approaches, the EU AI Act outlaws specific 'harmful' uses for the technology (though it could be argued that 'harmful' is ill-defined), at the other end, some governments assume that companies will 'do the right thing' through self-regulation, without independent intervention.

Alongside governments, international organisations are doing work that has the potential to provide a global framework for action. The Organisation for Economic Co-operation and Development (OECD) set out five Principles on Artificial Intelligence (AI Principles) that are viewed as the most comprehensive Western playbook for how to approach the technology. The principles promote AI that is innovative and trustworthy and that respects human rights and democratic values. These principles were adopted by the 38 members of the OECD in May 2019 when approving the OECD Council Recommendations on Artificial Intelligence.¹⁴ Some in the private sector have expressed support for the OECD approach – for example Microsoft was one of the first companies to call for so-called Responsible AI, aligned with the Organization for Economic Cooperation and Development's AI principles.

The UNESCO Recommendations on the Ethics of Artificial Intelligence were adopted by all 193 UNESCO Member States in November 2021. These put AI ethics in the context of human rights and fundamental freedoms, as well as the protection of the environment and ecosystems. It has been suggested in parliament that the UNESCO Recommendations could provide the basis for international alignment in practice; for example, for an international convention and enforcement body as we propose below.¹⁵

There is also work happening in industry-based international organisations such as the Institute of Electrical and Electronics Engineers (IEEE), a professional association for electronics engineering, electrical engineering, and other related disciplines whose core purpose is to foster technological innovation and excellence for the benefit of humanity. The IEEE new standard IEEE7001 on transparency, for example, includes the need for a watermark and – if adopted globally – could help to provide confidence and trust from the citizen perspective by demonstrating the difference between human and AI generated material.

In considering the UK's position, participants in the APGDA's inquiry welcomed the alignment of the principles in the UK White Paper with those of the OECD. There was unanimity about the importance for the UK to take action both internationally and domestically to build on this quickly to ensure cross-jurisdictional compatibility.

¹⁴ As at July 2021, 46 governments (38 OECD Members and 8 non-Members) had signed up to the AI Principles; the number has been static since then.

¹⁵ Darren Jones MP adjournment debate, House of Commons, 22 May 2023

Many stressed that, for the Government to regain its position internationally and be a central player within the debate on AI and regulation, considerable additional and sustained effort will be needed. While the Government is best placed to determine how to exploit all relevant diplomatic channels, contributors emphasised this should include both bilateral and multilateral channels including in the European Union and at the UN.¹⁶

The strategic issues around AI should be a regular item on the agenda of global gatherings such as the G7 and G20. Now is certainly an auspicious time to ramp up action amongst national leaders, off the back of the high profile scientific interventions about the global risks of AI and recent policy statements by politicians. In their Joint Statement on 20 May, the G7 political leaders called for a G7 working group to establish the “Hiroshima AI process” on generative AI and report back by the end of 2023.¹⁷ Options for coordinated global action include the establishment of a Conference of the Parties (COP) type forum, perhaps based on the Global Partnership on Artificial Intelligence initiated by the G7 in 2020. There is also relevant learning from the approach taken on nuclear weapons. It has been suggested that an international organisation along the lines of the International Atomic Energy Agency (IAEA) could be an enforcement body to ensure global compliance on AI risks, in support of a UN convention (see above in relation to UNESCO).¹⁸ It is relevant that the Prime Minister’s office has acknowledged – while this report was being written – that “there’s a recognition that AI isn’t a problem that can be solved by any one country acting unilaterally. It’s something that needs to be done collectively.”¹⁹

As well as re-engaging to be at the heart of strategic diplomatic discussions, the UK has much to offer at the research level. This could, for example, be an outcome from the UKRI to call at the end of 2022 for a leadership team to play a key role in driving the UK’s responsible and trustworthy (R&T) artificial intelligence (AI) agenda.²⁰

The UK also has considerable governance strengths. In later sections we will set out how a combination of regulation and governance can provide an ethics ‘ecosystem’ for future AI developments. This includes tools such as standards – on which the UK has taken a lead in the AI Standards Hub and BSI, for example with the Algorithmic Transparency standard. The work that is being carried out at the Responsible Technology Institute on an ethical black box standard for social robots – allowing interrogation of what went wrong and why – is another UK initiative that should be a model for international adoption, as was done decades ago for black box flight data recorders in the aviation sector. The inquiry heard from many that cross-jurisdictional compatibility can be achieved at the level of these sectoral and cross-sectoral standards, for example through these tools forming the conformity assessment required by the EU AI Act.

¹⁶ Some responders to this inquiry suggest an international approach based on UNESCO recommendations on the Ethics of AI.

¹⁷ adapted from: At G-7 summit, leaders call for international standards on AI, The Washington Post, Online <https://www.washingtonpost.com/world/2023/05/20/g7-summit-artificial-intelligence-ai/>

¹⁸ Sam Altman, CEO of OpenAI, in giving evidence to US Senators, 16 May 2023, reported in the Telegraph 17 May 2023.

¹⁹ The Times, Thursday 18 May 2023, 10:00 pm: We need ‘guardrails’ to regulate AI, Rishi Sunak says at G7 summit.

²⁰ Responsible and trustworthy artificial intelligence – UKRI – call issued 1 December 2022

Two final points arose.

First, it would be naïve to believe that all countries will sign up to and adopt international conventions and standards. It is worth recalling that in June 2017, President Putin said that “whoever reaches a breakthrough in developing artificial intelligence will come to dominate the world”.²¹ It would be safe to assume that under a Putin-type regime, Russia’s intentions towards the rest of the world in its use of AI will not be benign. And President Xi declared that this was a technology in which China had to lead, alongside setting specific targets for 2020 and 2025 aimed at putting China on a path to dominance over AI technology and related applications by 2030.

International conventions do, however, provide clarity over who are the ‘bad actors’ that need to be watched more carefully. The IAEA has flaws as a model, but it has provided powerful evidence for action by governments. At the commercial level, standards are used by responsible companies to differentiate themselves and their brand from the less responsible, and thereby to boost their own exports and trade.

The **second** point is on the sequencing of domestic and international efforts. Later in the report we recommend the adoption domestically of a short bill coupled with specific governance measures. Contributors to the inquiry rightly pointed out that the sequencing would need to be carefully handled to ensure that the UK’s domestic rules remain in step with its international aspirations, and the right level of equivalence is achieved to support business and trade. It is also worth bearing in mind that passing legislation is not a quick process; there will be a lead-in time during which the international position is likely to be clarified, especially through more active UK participation.

²¹ Putin: Leader in artificial intelligence will rule world. online. <https://www.cnn.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html>

Governance for culture-change

4. The Government should introduce statutory duties worded such as to require organisations to achieve the objective of ‘doing no harm’ through ‘cultural embedding’ governance; the definition of these duties should be subjected to meaningful consultation and testing.

The Government’s White Paper focuses mainly on AI regulation, and all contributors to our inquiry agreed that good regulation, and well-resourced regulators with clear responsibilities, have a critical role to play. In the next section we set out our findings to support our recommendations that there should be positive action by the Government on regulation and the current landscape of regulators.

There was a general consensus, however, that regulation is not the only mechanism that should be considered in a holistic framework or ecosystem to ensure that the development and bringing to market of AI is responsible and does not bring public harm.

Contributors stressed the need for a citizen-focussed approach in both the private and public sectors, through governance measures and ways of working that embed the principles of fairness, transparency, and explainability of AI into the culture of the organisation. It was pointed out that some of the principles set out in the White Paper have a statutory basis (such as safety), but that others do not (such as algorithmic transparency). Judgements about how AI systems deliver against such principles may be nuanced, and therefore need to be made at a senior level, involving a range of experts, both technical and ethical. Contributions from employee representatives will also be valuable in ensuring that the leadership has a good understanding of any cultural issues within its organisation.

Some organisations are already working on internal governance measures that provide ‘ethical grit’ in the technical development process, including through the adoption of codes of conduct, or with AI development teams that are multi-disciplinary, involving customer/citizen facing staff with the objective of achieving ‘ethical by design’.

On the other hand, we were told that there are many companies, especially SMEs, that have not even realised they need to think about ethical issues, governance, and regulation. Although they may appreciate the significant risk to company ‘Brand Reputation’ from loss of public or business-to-business confidence, knowing how to avoid or manage the risks and liabilities is a different matter especially in relation to technologies new to them like AI systems. Government can support sustainable growth by providing clear rules and a one-stop-shop source of guidance on how to tackle ethical considerations up front, including when company business models are being developed in start-ups, as well as through the product or service life cycle.

The APGDA’s first report *Trust, Transparency and Tech – building ethical data policies for the public good*, made a number of recommendations to ensure that organisations delivering public services (whether public or private sector) have AI and data related governance mechanisms with the right internal checks and balances.²² We have revisited a number of our recommendations during this inquiry, to test their value in providing the ‘cultural embedding’ into an organisation of the principles in the White Paper (and OECD) of fairness, transparency, and explainability. Contributors pointed out that multiple surveys have demonstrated that individuals want to have trust in organisations and institutions, not just in the AI technology itself. And to provide that trust, surveys show that people want to see independent proxies such as independent oversight Boards and other means of third-party validation of how organisations approach the use of AI.

²² Policy Connect; *Trust, Transparency and Tech - building ethical data policies for the public good*, May 2019. <https://www.policyconnect.org.uk/research/trust-transparency-and-technology-building-data-policies-public-good>

The most effective and appropriate governance arrangements within an organisation will depend on many factors, and it would be unhelpful to be overly prescriptive in setting out rules for how a company delivers the objective of ensuring responsible AI (ie that 'does no harm'). Furthermore, a key message in our inquiry was that measures appropriate for big tech organisations might be too onerous – and unnecessary - for small enterprises and start-ups, and that anti-competitive issues should be taken into account in devising the governance and regulation framework.

It is generally accepted that good governance should start with the leadership of an organisation (for example by adopting best practice from management enterprise standards such as ISO9001-2015). In *Trust, Transparency and Tech – building ethical data policies for the public good* we proposed that – for organisations providing public services - there should be clear lines of accountability at the top of the organisation, by updating company director responsibilities.

Building on this recommendation, we propose that the leadership team should have, and be supported to deliver on, a new statutory duty, which would sit alongside its existing duty to 'promote the success of the company' or organisation. Under this duty a director would be required to ensure the company considers the ethical impact of AI on customers and society – and potentially on their own workers, ie to implement 'ethical by design'.

In *Trust, Transparency and Tech – building ethical data policies for the public good* we also recommended extending the Data Protection Officer (DPO) role. In the AI context, however, a focus on and requirement for a specific DPO role (or Senior Responsible Individual (SRI) as per the proposed changes in the Data Protection and Digital Information Bill) may be too prescriptive. While the new SRI role could in practice encompass the AI and ethics role, other arrangements might better suit the size and nature of each individual company or organisation.

This argues for a range of internal governance options supporting the leadership, who under this model would be accountable for responsible AI in the same way as they are for effective health and safety and anti-bribery. Contributors to this inquiry made a number of suggestions on governance, which could be included in a menu of obligations requiring an organisation to achieve the objective of 'doing no harm' and to introduce supporting 'cultural embedding' governance (the definition of these obligations should be subjected to meaningful consultation and testing):

- a. A statutory Duty on an organisation's leadership to take account of the requirement to achieve 'doing no harm'. This could be tied into the design and bringing to market of 'high-risk' AI systems.
- b. A statutory Duty, in cases where the AI system is judged to be 'high-risk' in terms of its impact on lives, for the leadership to have a 'go/no go' sign-off process, so that a deliberative decision to proceed with a significant new AI system is taken prior to release/use.
- c. A requirement that the leadership team, at Board level, should include a person accountable for ensuring due diligence on AI and ethics. This might be a Chief Information Officer or Senior Responsible Individual (currently DPO), or their equivalent in a small company.
- d. A requirement, in organisations of over 250 employees, for an Ethics (Advisory) Committee (along the lines of Audit Boards or indeed as part of an Audit, Risk and Ethics Board) to bring together technical, customer-facing, and ethics experts to support the balance of judgements in Board decision-making. Employee representative participation will also help the leadership to understand cultural issues within the organisation.

- e. The option for pre-market, controlled real-conditions testing to be carried out in regulatory sandboxes to support Board 'go/no go' decisions on 'high-risk' systems. This would allow testing to eliminate bias and could help to identify the organisation's legal liabilities.
- f. Approved/kite-marked standardised algorithmic and privacy risk assessment/impact tools (which have the potential to become global norms).
- g. The option of engaging People's Panels (potentially shared within a locality such as Greater Manchester, where local people have been trained to become volunteers on a People's Panel for AI).

Some measures are intended to be statutory, others not – at least in the first instance. Some measures could also depend on the size of organisation, applying to large tech companies and public service providers, but with SMEs and small bodies including in the third sector given freedoms to allow them a head-start while making clear that the direction of travel in governance terms has to be 'ethical by design'.

The effectiveness of an individual organisation's governance regime would be tested through a requirement – for example in government procurement processes – for companies to have third party accredited independent certification to relevant sector standards. Such standards are already under development and the UK could – as indicated earlier – make this a key element of its international work to drive global standards and cross-jurisdictional compatibility. It is likely that the EU regime will require conformity assessment of AI products and services prior to entering the EU marketplace. The UK could – through seeking international adoption of sector standards - make it easier for British companies to meet EU compliance.

In adding to Directors' Section 172 duties, we recognise that there are already many duties on Directors/the Board.²³ Experience has shown in many sectors (such as health and safety, fire safety in construction, modern slavery, anti-bribery), that external monitoring and enforcement is a necessary backstop. The onus could in the first instance be on the company to demonstrate the effectiveness of the internal governance tools they use as part of an annual report on AI governance and incidents as set out in the next section.

Alongside corporate company measures, submissions to the inquiry also addressed the question of training for and accreditation of individual professionals. This includes AI developers and AI ethics advisers and consultants. The introduction of a governance and regulatory regime will create demand for new roles to provide advice as well as to carry out audits, inspections, and investigations. It is important that such individuals are properly qualified and trusted.

We therefore recommend work to establish an accreditation scheme for AI developers, AI ethics advisers and consultants, and AI auditors, with individuals required to have professional registration. This recommendation aligns with the recent paper published by BCS: The Chartered Institute for IT.²⁴

²³ Section 172 of the Companies Act 2006 reflects the principle of "enlightened shareholder value" and provides for a specific duty to have regard to a number of factors in 'promoting the success of the company'.

²⁴ "Helping AI grow up without pressing pause" published on 4 May 2023 <https://www.bcs.org/articles-opinion-and-research/helping-ai-grow-up-without-pressing-pause/>

A National AI Centre

5. A single independent, central body with strong regulatory authority, properly resourced, should be established in statute. Its functions should be to:
 - a. Convene existing sector regulators (e.g., lead the DRCF) and ensure AI principles are prioritised in their business plans and that their regulatory functions are properly resourced, proportionate and agile. To ensure the right powers and resources are provided, the DRCF could also be placed on a statutory corporate basis.
 - b. Identify and fill regulatory gaps, and address overlaps between regulators.
 - c. Ensure regulators have the right roles, skills, and enforcement powers to deliver smart regulation.
 - d. Commission a registration scheme for professionals in the AI industry.
 - e. Provide guidance to the public and private sectors to encourage responsible uptake of AI.
 - f. Horizon-scan to anticipate AI developments/functionality/national ethical issues.
 - g. Initiate investigations into new AI deployments if it has reasonable cause to suspect 'harm'. It should have powers to demand information and/or the power to require another regulator to do this.
 - h. Initiate/oversee investigations into incidents, especially where regulatory roles are confused. This could include having powers, pending the investigation, to request an immediate injunction on a system from the Secretary of State if there is evidence that serious harm is already occurring.
 - i. Provide guidance to the public to help them guard against AI harms; alongside a single Ombudsman also to be established (potentially the ICO).
 - j. Make an annual report to government and parliament, ideally to a new Joint Parliamentary Committee including members from the relevant Select Committees.

Evidence submitted to this inquiry set out a mixed picture on current regulation of AI in the UK. Companies and representative bodies in sectors subject to statutory regulation enforced by regulators with a clear mandate and singular focus emphasised two points. First, that the existence of regulation in their sector helped not hindered growth as it provided a stable environment in which to innovate. Second, that – for their sector at least – no new regulator was needed for AI specifically as the principles set out in the White Paper could be addressed by existing regulators.

Similar views were expressed by cross-sectoral regulators such as the Equality and Human Rights Commission (EHRC) and the Information Commissioners Office (ICO). Existing regulators were concerned about the risk of overlap and muddle were there to be a single regulator for AI with, for example, issues around human rights protection and personal data privacy already covered.

Some of these protections are already being strengthened, such as through the new consumer protection rules that will come into force at the end of July, legally requiring companies regulated by the Financial Conduct Authority (FCA) to put the interests of customers first. The FCA has said that it will act quickly to punish serious breaches, and harms to customers from AI systems should certainly fall into the category of not being in the primary interests of customers.

However, these tough new laws will only apply to financial services and financial markets, meaning that the lack of consistency with other industry sectors is likely to increase. Furthermore, many interlocutors pointed out that some sectors are currently unregulated, and even those in favour of sticking with the current regulatory landscape acknowledged that it is complex, with different regulators having varying degrees of regulatory ‘teeth’, some with enforcement powers (such as imposing fines) but others not.

Both regulators and third parties agreed that collaboration on AI regulation was helpful, such as through the Digital Regulation Cooperation Forum (DRCF) in relation to online regulation, but that the coverage of the Forum is limited. Furthermore, capacity and skills amongst and across regulators is acknowledged by regulators themselves to be a significant limiting factor both in working collaboratively and having the right access to technical AI skills.

From the industry perspective, a common problem is that of regulators being unable to respond quickly and proportionately. This is made worse where more than one regulator is involved, resulting in confusion and delay for a company wanting to move ahead quickly with innovative products and services. There is a risk that AI developers will be scrutinised more than is necessary, rather than less. The Competition and Market’s Authority (CMA), for example, has recently (4 May 2023) announced an investigation into the rise of artificial intelligence (AI) tools such as ChatGPT, to understand the competitive landscape for a new generation of AI algorithms and their potential threat to consumers. However, the CMA’s writ may not extend to all the aspects that need investigation. Finally, and importantly, the suggestion in the White Paper that a company will have to engage three regulators was described by contributors to our inquiry as unrealistic and slow.

The UK needs to be able to act decisively and quickly, and the relationship between industry and the regulator works fastest and best when there is a single sector regulator with the responsibility, authority, and right technical know-how to be proportionate, agile and responsive. It is interesting that Sam Altman, the Chief Executive of OpenAI (the creator of ChatGPT) told US Senators that a new government agency should licence large language models such as ChatGPT, and refuse licences if they do not comply with government standards and abide by safety guidelines or audit requirements.²⁵

Several contributors addressed the regulatory gap from the citizen perspective, arguing that the current levels of public awareness of AI and ethics is low, and that engagement and awareness-raising work needs to be done at scale. A major gap, looking from the citizen perspective, is in understanding who to go to when AI goes wrong, to complain or to seek redress. It was pointed out that the consequences of AI going wrong are felt not by the AI developer but by the person to whom the AI was applied. For example, the individual discriminated against in AI-based recruitment processes (which are not regulated); the worker (e.g., delivery person) who is unjustly penalised because AI-based performance tracking systems are not able to recognise the external causes of reduced work performance; or lives ruined by live facial recognition used by police forces into which bias has unconsciously been built.

Another failing of the current system is the lack of a strong horizon-scanning function to anticipate future AI-related developments and risks. This inquiry did not attempt to define ‘AI risk’ or ‘high-risk’ AI, but there was a common view that this should be done. The suggestion in the White Paper of putting this in the ‘central function’ was noted, but contributors commented that the scale of the task should not be underestimated.

²⁵ When testifying before a Senate Judiciary Privacy, Technology and the Law Subcommittee hearing on Capitol Hill in Washington on 16 May 2023

Against that background, there was unanimity that a central body is certainly required, but with clout. Although the White Paper does propose a central function or functions, this currently comes across as insufficiently high-profile and muscular, and putting it within Whitehall would rule out regulatory functions and independence. The scale and speed of developments in AI and regulation call for a strong national body able to drive change, anticipate events, and with sufficient organisational permanency to take the long view as well as operate in an agile way in the short term. When the Centre for Data Ethics and Innovation (CDEI) was created there was optimism that this might provide the kind of centrally driving force required. However, the CDEI was not put on the statutory basis which would have given it clout and longevity. Other AI-focussed bodies have since been created such as the Office for AI and the AI Standards Hub that – while positive in their own right – means that there is no single leadership agency.

We recommend that the Government remembers the creation of the National Cyber Security Centre (NCSC). The bringing together of a number of units previously dispersed across different parts of government into a single, high-profile NCSC has been a real success. Cyber activity, both reactive and proactive, has come on leaps and bounds in both the private and public sectors. The NCSC provides a single focal point of technical expertise and guidance to companies and regulators to help minimise the risk of incidents and respond to them.

There are two key differences between the NCSC model and the one we are recommending here. First, the NCSC is not a regulator, but has been involved in the setting of statutory requirements for Board and of technical standards. Second, it is not a body based in statute, but as part of the Government Communication Headquarters (GCHQ), one of the three UK Intelligence and Security Agencies, it has its own strong voice and is generally regarded as ‘independent’.

Our strong recommendation is for a statutory body given the societal and economic importance of AI issues, opportunities and risks. We made a similar proposal for a statutory body in *Trust, Transparency and Tech – building ethical data policies for the public good*; the national and international context now makes this a more urgent and critical requirement. A new and high-profile organisation could be built swiftly by bringing together existing bodies to create a core of skills and capacity in the National AI Centre. They could include the CDEI, the Office for AI, the AI Standards Hub, and the relevant elements of the Central Digital and Data Office in the Cabinet Office. Other skills would need to be added, in particular regulatory skills.

In terms of functions for this important new body, these need to be worked through in detail but should broadly fit with an ‘enabling regulatory’ approach: research, supportive guidance, regulatory coordination and oversight, and enforcement.

The inquiry also considered the question of an Ombudsman. Contributors held the strong view that the Government needs to look through the lens of providing one ‘go-to’ point for citizens, not expect the individual to find the right regulator or ombudsman. Adding this into the regulatory body seems to mix functions that should remain separate. There was some discussion as to whether the Information Commissioners Office (ICO) could provide the one-stop shop Ombudsman focal point (even if to signpost to the right organisation). However, the provisions of the Data Protection and Digital Information Bill seem to take the ICO in the opposite direction, allowing it greater latitude to decide not to investigate a complaint. We have not therefore made a recommendation as to where a focal-point Ombudsman function should sit.

The regulatory function would be focussed – certainly to start with - not on hands-on regulation, but a more strategic role of oversight and review of existing sectoral arrangements. In effect, a socket into which sectoral and cross-sectoral regulators (such as EHCR) can be plugged. The role would in the first instance include convening the existing sector regulators (through an expanded Digital Regulation Cooperation Forum), and using its overarching perspective to identify regulatory overlaps and gaps to deliver effective and agile regulation. This would allow objective decisions on how best to fill regulatory gaps, potentially through reassignment amongst existing regulators; possibly with a recommendation for a new sector regulator where none exists; or perhaps by taking on direct regulatory functions itself. As part of this it would keep under review the capacity and skills of regulators, and enforcement tools available to each. The Centre could provide technical and ethics expertise to other regulators where required.

The National AI Centre would oversee and potentially initiate accident investigations, including providing the necessary technical and ethics expertise to existing investigators. An example of this would be the provision of AI experts to the Health and Safety Executive to carry out forensic interrogation of the algorithms in the case of an accident involving a smart robot. It would provide or commission advice and guidance to the private, public, and third sectors. It could commission the development of professional skills schemes and standards for the creators, developers, and users of AI to ensure they have the appropriate professional and ethical skill sets.

The National AI Centre should have a strong anticipatory horizon-scanning function, to anticipate the direction of technical innovation and the benefits and risks involved, and help the Government stay ahead of the game in this fast-moving environment. For this purpose, it could connect in with the AI Watch work that is done by the European Commission's Joint Research Centre (JRC).

To enhance parliamentary scrutiny and help ensure that the citizen remains at the heart of AI developments and the work of the National AI Centre, it should submit an annual report to Parliament to scrutinise. While it is for Parliament to decide how to organise its committee structures, we note there have been calls by parliamentarians for a new committee in relation to digital and online issues arising from the scrutiny of the Online Safety Bill.²⁶ We would recommend that Parliament consider the establishment of a new Joint Committee²⁷ on emerging technologies and ethics. This could be staffed by parliamentary members from expert committees across both Houses of Parliament, such as the Commons DCMS and Business and Trade Committees, and the Lords Communications and Digital Committee. Such a committee would chime with the suggestion that there needs to be a body “to oversee cross-cutting issues such as the regulation of emerging technologies that exist beyond a single sector. The present system simply does not take into account how regulators interact with one another, to the benefit or detriment of the public.”²⁸

²⁶ Reported in Politics Home online, 19 April 2023: Online Safety Bill “Could Pose Risk” To Balance Of Power Between Government And Tech Firms, Peer Warns.

²⁷ A Joint Committee is one in which Members from both Houses meet and work as one committee, and appoint a single chairman who can be an MP or a Member of the Lords.

²⁸ Article in The Times, 17 April 2023, by Lord Tyrie and Bim Afolami MP “MPs and peers need to regulate the regulators to get the best of Brexit”.

Laws and Tools – top down and bottom up

6. A very short Bill should be developed to implement these recommendations, put forward as a draft for meaningful consultation with stakeholders and to ensure cross-party support, and tested against real-life scenarios including with citizens. Getting this right should ensure that companies only need to meet one set of obligations in order to satisfy compliance requirements in other jurisdictions.
7. The suite of standards currently under development by the international community (ISO etc) and nationally (BSI, the AI Standards Hub) should be trialled in a number of sectors and made part of a robust accreditation system.
8. The Government should lead the way in driving change through procurement policy.

In the Regulation White Paper, the Government indicated it was reluctant to enact legislation and wanted a more hands-off approach with companies being asked to abide by five “principles” when developing AI. There was a suggestion that a statutory duty on regulators might be introduced in due course.

It is certainly the case that a complicated and inflexible bill could be counter-productive by failing to imagine the risks of tomorrow or by focussing on the technology rather than the objective of beneficial outcomes for people and society. This is not, however, an argument for eschewing legislation but for the nature of a parliamentary bill. We have proposed that it should focus on governance to safeguard individuals and communities, so that the measures will be effective however the technology develops.

Furthermore, if the UK is to achieve the “co-ordination with our allies” that the Prime Minister suggested during the May 2023 G7 meeting in Japan it will be important to have at least compatible legal approaches, as well as a clear lead for regulatory discussions with other countries.²⁹ Our proposals for a short and broad Act, and for an umbrella regulatory body, will put the UK in a better position to be heard internationally, and to achieve cross-jurisdictional compatibility.

We have indicated earlier that a concern has been expressed that bringing in a parliamentary bill could pre-empt what is happening globally on standards and regulation and might mean the UK moves at a different pace in different sectors, with unhappy consequences for trade and commerce. This tension can be managed in several ways. First, we have recommended a bill be published and consulted on in draft, including testing its provisions (in sandboxes and with citizens’ panels) to ensure it achieves the objective. This will take a little time. Second, it will be important to resist any calls to over-complicate the bill or make it too prescriptive. Third, the bill should focus on broad governance requirements via legal obligations on AI developer and provider organisations so that – should they act ‘recklessly and thoughtlessly’ – they may face criminal proceedings, and – while setting the National Centre for AI in statute – provide flexibility on its functions and tasks by concentrating the legislation on outcomes for citizens.

To minimise prescription and maximise regulatory agility, while many positive operational developments and tools were drawn to our attention, we do not suggest they should be included in primary legislation. These tools merit listing, as a start point for either guidance or regulatory requirements, and to form part of a robust standards and accreditation framework. The Centre for Data Ethics and Innovation has made a great start with the development of an AI Assurance roadmap, and this needs to be given further support.

²⁹ The Times, Thursday 18 May 2023, 10:00 pm: We need ‘guardrails’ to regulate AI, Rishi Sunak says at G7 summit.

Other tools that can be developed and brought into a standards and accreditation framework are as follows:

1. New relevant standards: IEEE7001, a new standard on transparency which includes the need for a watermark; the Ethical Black Box open standard published in 2022. Such standards, under development by the international community and nationally should be trialled in a number of sectors and made part of a robust accreditation system.
2. Algorithmic Impact Assessments (there is an NHS pilot) and Privacy Impact Assessments which can be used to support Boards to take the right decision on whether and how to introduce new ‘high-risk’ AI systems.
3. Regulatory Sandboxes used to trial new AI systems in ‘real world’ scenarios. Such Sandboxes should have access to robust databases to ensure responsible training of AI. This could help in terms of testing for ‘bias’, for example, which is judged to be a particular concern. It will be important for the training of AI to be underpinned by the responsible collection of public data made readily accessible. In *Our Place, Our Data – involving local people in data and AI-based recovery*, Policy Connect recommended that an important role for government is to publish – or enable the publishing of – open data sets of standardised, usable data.³⁰
4. User-friendly AI watermarks and ‘kitemarks’ on products/services so that people know if something has been created by AI. One such initiative is the Content Authenticity Initiative (CAI). The CAI represents a global, cross-industry, academia and third sector community (media and tech companies, NGOs and academics), seeking to promote the adoption of an open industry standard for content authenticity and provenance. The initiative also promotes tools that make it easy to indicate when AI was used to generate or alter content. This could be one way of uncovering mis- or dis-information and more generally demonstrating when content is created by AI rather than a human (i.e., a ‘watermark’ equivalent).
5. For public-facing AI systems, transparency and explainability in a ‘flight data recorder’ equivalent.
6. A standing Citizen’s Assembly on Data and AI, and People’s Panels for AI.

In addition, a registration scheme should be developed for specialist professionals working in the sector, as is the case in many other sectors whose activities may impact on human rights. Such a scheme should work towards internationally recognised qualifications.

Finally, procurement is a good place to drive change amongst suppliers. While the Government is in general concerned about stipulating measures in procurement requirements that could prove anti-competitive, we have been told that without procurement guidance public service providers such as local authorities are likely to be fearful of committing to AI-enabled service delivery solutions. The Government should lead the way in devising and testing procurement requirements that are citizen focussed – for example by requiring an Algorithmic Transparency Standard Report. This could be done in conjunction with other public procurement authorities in local authorities and mayoral teams. Work to develop procurement processes should address liability issues including in relation to third party-provided data and algorithms.

³⁰ Policy Connect, *Our Place, Our Data - involving local people in data and AI-based recovery*, March 2021: <https://www.policyconnect.org.uk/research/our-place-our-data-involving-local-people-data-and-ai-based-recovery>. See also: <https://fortune.com/2023/05/17/ai-revolution-threat-big-tech-web-data-public-domain-or-lenchner/>

Conclusion and Next Steps

This is the third APGDA report to date on data, AI and ethics, and the subject-matter has since our first report in 2019 become ever wider, deeper, and societally significant. There are a raft of issues around AI and its ramifications that could have been addressed in this report, and there will no doubt be some that its readers will wish had been covered. For example, we have not sought to define different aspects of AI, or to pull out any particular type of AI as the object of this report's recommendations.

It will undoubtedly also be the case that not everyone will agree with every part of this report and its recommendations. As the APGDA said in *Trust, Transparency and Tech – building ethical data policies for the public good*, these are complex and potentially controversial issues. What we have sought to do is to set out a governance and regulatory regime that has general application and focusses on outcomes not technology. This is to allow the UK to take the opportunity of AI alongside providing government with the agility and flexibility to deal with future surprises in generative AI.

In concluding, we should point out three messages that came through loud and clear as golden threads in this inquiry. First, **the Government needs to take action at scale and speed**, both internationally and domestically. The two are linked: on the international stage, the UK's credibility depends on strong domestic action; and what UK does domestically needs to fit with international developments if UK industry is to have a chance to succeed in external markets. Second, that **unambiguous smart regulation is an enabler to innovation**, not a blocker. And third, that it will require **enduring heavy lifting to bring together all necessary stakeholders** to create solutions for ethical dilemmas (which should involve business, academia, employee representatives, civil society bodies and the public as well as regulators). This in part underpins the rationale for a strong and independent body based in statute, sitting outside government.

It will – and indeed should - take a little time to develop and introduce effective, smart legislation. It follows that the government should commit now to legislation around the National AI Centre and the governance requirements on organisations. This will provide a clear direction of travel and the certainty that industry requires if it is to have the confidence to invest today and for the long term.

The development and consultation on legislation should not hold up short-term action – on the contrary such commitments should be used to increase the pace and profile of action on all fronts. This includes the establishment of the National AI Centre with the right level of resources to deliver good regulation; the further development by industry of responsible codes of practice; the establishment of regulatory sandboxes; the piloting of standards; and the development and piloting of accreditation schemes – both for standards and for the registration of professionals.

Finally, this report is intended as a contribution to the consultation on the White Paper. The APGDA stands ready to assist the Government with taking forward and operationalising these recommendations. We have in addition indicated aspects that require further investigation, in particular skills and education, and climate impacts.

Methodology

Policy Connect carried out this inquiry between September 2022 and May 2023. Evidence was gathered in a series of evidence sessions between January and April 2023, interviews with those working in and around AI, written submissions, and input from our steering Group.

We recognise that these are complex and potentially controversial issues and expect that not all of those listed as contributors will agree with every part of the report.

The views in this report are those of the author and Policy Connect. While they were informed by our contributors, they do not necessarily reflect the opinions of either individuals or organisations - whether Steering Group members, participants in evidence sessions, or contributors of written evidence.

Steering Group

Name	Organisation
Daniel Zeichner, MP	House of Commons
Lord Holmes of Richmond	House of Lords
Lord Clement-Jones	House of Lords
Ansgar Koene	EY
Dr. Adrian Weller	The Alan Turing Institute
Eve Lugg	Formerly Policy Connect
Jennifer Amphlett	Zurich Insurance
Marina Jirotko	The Responsible Technology Institute at the University of Oxford
Or Lenchner	Bright Data
Penny Jones	Zurich Insurance
Rashik Parmar	BCS
Roger Taylor	Independent Advisor
Sue Daley	TechUK
Tom Moule	Jisc

Evidence Session 1: Chaired by Lord Clement-Jones

Speakers included:

- Adrian Weller, The Alan Turing Institute
- Rachel Coldicutt, Careful Industries
- Rashik Parmar, BCS
- Sue Daley, TechUK

Evidence Session 2: Co-chaired by Lord Holmes of Richmond and Lord Clement-Jones

Speakers included:

- Ansgar Koene, EY
- Carly Kind, Ada Lovelace Institute
- Dr Aisha Naseer, Strategic Advisor on AI Ethics
- Marina Jirotko, The Responsible Technology Institute at the University of Oxford
- Or Lenchner, Bright Data
- Roger Taylor, formerly CDEI

Evidence Session 3: Chaired by Daniel Zeichner MP

Speakers included:

- Lord Clement-Jones, House of Lords
- Penny Jones, Zurich Insurance
- Tom Moule, Jisc

Evidence Session 4: Chaired by Lord Clement-Jones

Speakers included:

- Adam Leon Smith, BCS
- Carly Kind, Ada Lovelace Institute
- David Leslie, The Alan Turing Institute
- Stuart Holland, Equifax

Evidence Collected From:

ACCA
Access Partnership
Ada Lovelace Institute
AI Centre
Amazon
Anekanta Consulting
BCS
Bright Data
BS 30440 Development Panel
Careful Industries
Centre for Data Ethics and Innovation
Confederation of British Industry
Digital Catapult
DLA Piper
EKB Consulting
Equality and Human Rights Commission
Equifax
EY
Government Department of Science, Technology, and Innovation
Information Commissioner's Office
Institute of Customer Service
Jisc
Living with Data
London Office of Technology and Innovation
Manchester Metropolitan University
Massachusetts Institute of Technology
Microsoft
Responsible Technology Institute and RoboTIPS
TechUK
The Alan Turing Institute
The Market Research Society
The Open University
The Responsible Technology Institute at the University of Oxford
UK Finance
University of Bournemouth
University of Cambridge
University of Liverpool
University of Newcastle
University of Oxford Brookes
University of West England (UWE)
VISA
Wayve
Zurich Insurance

Individual Contributions

Ali Hessami, Vega Systems
Chris Reed
Eleanor 'Nell' Watson, European Responsible AI Office
Keri Grieman
Patricia Shaw, Beyond Reach Consulting Limited
Roger Taylor

Acknowledgements

This report, its recommendations, views and opinions expressed herein are the responsibility and the work of Policy Connect and the All-Party Parliamentary Group on Data Analytics.

We are grateful to Daniel Zeichner MP, Lord Holmes of Richmond and Lord Clement-Jones for their leadership and dedication to this project as co-chairs for the inquiry. Special thanks go to Bright Data, EY, Jisc, and Zurich Insurance for their kind sponsorship and expertise to inform our findings.

We also wish to thank our many partner organisations that were consulted and provided invaluable evidence and input over the course of this inquiry. Special thanks go to Alainah Amer, Claudia Jaksch, Victoria Zeybrandt, Robert McLaren, Rob Allen, and former Policy Connect members Eve Lugg and Floriane Fidegnon.

THE ALL-PARTY PARLIAMENTARY GROUP ON DATA ANALYTICS

The cross-party group's aims are to connect Parliament with business, academia and civil society to promote better policy making on big data and data analytics. This report follows the Group's previous two reports; entitled *Trust Transparency and Tech: Building Ethical Data Policies for the Public Good (May, 2019)* and *Our Place Our Data: Involving Local People in Data and AI-Based Recovery (March, 2021)*.



POLICY CONNECT

Policy Connect is a cross-party think tank. We specialise in supporting parliamentary groups, forums and commissions, delivering impactful policy research and event programmes and bringing together parliamentarians and government in collaboration with academia, business and civil society to help shape public policy in Westminster and Whitehall, so as to improve people's lives.



Our work focusses on five key policy areas which are: Education & Skills; Industry, Technology & Innovation; Sustainability; Health; and Assistive & Accessible Technology.

We are a social enterprise and are funded by a combination of regular annual membership subscriptions and time-limited sponsorships. We are proud to be a Disability Confident and London Living Wage employer, and a member of Social Enterprise UK.

This is not an official publication of the House of Commons or the House of Lords. It has not been approved by either Houses or its committees. All-Party Parliamentary Groups are informal groups of Members of both Houses with a common interest in particular issues. The views expressed in this report are those of the group.

BRIGHT DATA

Bright Data is the world's #1 web data platform. Fortune 500 companies, academic institutions, and small businesses all rely on Bright Data's solutions to retrieve crucial public web data in the most efficient, reliable, and flexible way. With web data, these companies can research, monitor, and analyse data to make better decisions. We believe that making public web data easily accessible is essential to keeping markets openly competitive, benefiting all of us.


 The logo for Bright Data, featuring the words "bright data" in a lowercase, sans-serif font. The word "bright" is in a light blue color, and "data" is in a darker blue color.

EY

At EY, our purpose is Building a better working world. The insights and quality services we provide help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.



JISC

Jisc is the UK digital, data and technology agency focused on tertiary education, research and innovation. We are a not-for-profit organisation and believe education and research improves lives and that technology improves education and research. We provide managed and brokered products and services, enhanced with expertise and intelligence to provide sector leadership and enable digital transformation.



ZURICH INSURANCE

Zurich Insurance provides a suite of general insurance and life insurance products to retail and corporate customers. We supply personal, commercial and local authority insurance through a number of distribution channels, and offer a range of protection policies available online and through financial intermediaries for the retail market and via employee benefit consultants for the corporate market. Based in a number of locations across the UK - with large sites in Birmingham, Farnborough, Glasgow, London, Swindon and Whiteley - Zurich employs approximately 4,800 people in the UK.





CONTACT

Policy Connect
CAN Mezzanine
7-14 Great Dover Street
London SE1 4YR

 @Policy_Connect
 policy-connect
 info@policyconnect.org.uk
 0207 202 8585