

RESPONSIBLE AI AND ROBOTICS

An ethical framework

THE RIGHT ENVIRONMENT

Laws and regulators can and should be enablers of the digital economy. Clarity and predictability combined with speed and flexibility are essential.

TRUST BY DESIGN

For the digital economy to flourish, trust needs to be baked into the system at every level. It is an essential ingredient of its licence to operate.

KNOW THE IMPACT

Impact: Essential. Key to digital success

Urgency: Critical. Short term

THE DETAILS

Artificial Intelligence (AI) is the science of using computers to do things that traditionally required the human mind. It is a technology that will accelerate the digital transformation of industry, and will prove essential to the success of our digital economy in what is an increasingly connected world.

For AI to deliver on its promise, however, it will require predictability and trust. These two are interrelated. Predictable treatment of the complex issues AI will throw up, such as accountability and permitted data uses, will encourage investment in and use of AI. Similarly, progress with AI requires consumers to trust the technology and the fairness of how they are affected by it and how their data is used; predictable and transparent treatment facilitates this trust.

A robust legal framework will be needed to deal with those issues too complex or fast changing to be addressed adequately by legislation. But the political and legal process alone will not be enough. For trust to flourish, an ethical code will be equally important.

The government should encourage discussion on the ethics of AI, and ensure all relevant parties are involved. Bringing together the private sector, consumer groups and academia would allow the development of ethical code that keeps up with technological, social and political developments.

Government efforts should be collaborative with existing efforts to research and discuss ethics in AI. There are many existing initiatives which could be encouraged, including the Alan Turing Institute, the Leverhulme Centre for the Future of Intelligence, the WEF Centre for the Fourth Industrial Revolution, work being done by the Royal Society, and the Partnership on Artificial Intelligence to Benefit People and Society.

There are industry bodies, such as the Institute of Electrical and Electronics Engineers, working on developing standards and certification systems for Ethically Aligned Design of AI and autonomous

systems. Where possible, making those efforts inclusive, rather than exclusive or isolated, is desirable.

OPPORTUNITIES AND CHALLENGES

As Artificial Intelligence becomes more sophisticated, it will start to make or assist decisions that have a greater impact on individual lives. This will raise ethical challenges as people adjust to the larger and more prominent role of automated decision making in society.

In the UK, AI has exciting implications in the consumer, healthcare, transport, professional services and defence industries. It promises to deliver learning and evolving systems that are more efficient, safer, smarter and better able to operate at scale. For example:

BIG DATA ANALYTICS

Applying AI across massive data sets enables functionality such as autonomous vehicles, remote healthcare and advanced network management.

LIQUID WORKFORCE

Use of AI in certain workforce processes will enable the separation of specific tasks and the exploration of new, disaggregated resourcing models.

AUTOMATION

The entrance of AI into industries including professional services will allow the automation of tasks previously reserved to humans, enabling them to be done faster and at greater scale.

But these opportunities come with associated ethical challenges:

Decision making and liability: As AI use increases, it will become more difficult to apportion responsibility for decisions. If mistakes are made which cause harm, who should bear the risk?

Transparency: When complex machine learning systems have been used to make significant decisions, it may be difficult to unpick the causes behind a specific course of action. The clear explanation of machine reasoning is necessary to determine accountability.

Bias: Machine learning systems can entrench existing bias in decision making systems. Care must be taken to ensure that AI evolves to be non-discriminatory.

Human values: Without programming, AI systems have no default values or "common sense". The British Standards Institute BS 8611 standard on the "ethical design and application of robots and robotic systems" provides some useful guidance: "Robots should not be designed solely or primarily to kill or harm humans; humans, not robots, are the responsible agents; it should be possible to find out who is responsible for any robot and its behaviour."

Data protection and IP: The potential of AI is rooted in access to large data sets. What happens when an AI system is trained on one data set, then applies learnings to a new data set? Could this infringe data protection and privacy rights? Who, if anyone, should own the output of AI thought processes?

Social dislocation: Extensive deployments of AI could lead to social dislocation.

Cybersecurity: As more reliance is placed on AI, the importance of cybersecurity will increase. Security must be a top priority for all actors if trust is to be maintained.

OUR VIEW

We welcome steps already being taken in this space, including the government's commitment to undertake a major review on the development of artificial intelligence. As such, the UK already enjoys strong foundations in AI, but more can always be done.

None of this is for government alone and it is important for the UK—government, business, academia and society more widely—to collaborate and build on its existing position of strength so that we can engage AI further to differentiate and boost growth. These opportunities include fostering a public discussion to build a set of fundamental ethical principles for AI development. Flexibility is necessary to allow for private bodies to develop and commit to more specific codes in certain sectors, building on these fundamental principles.

Such codes should focus on

Responsibility for mistakes - it must be clear where liability lies when systems make mistakes. General principles should guide accountability.

Transparency in decision making - it must be clear when AI systems need to explain their actions to humans to show why a decision was made, and when, if ever, such transparency is not necessary.

Avoiding bias - steps must be made to stop systemic bias. Core values such as equality, diversity and lack of discrimination must be promoted.

Core ethical values - what are the core ethical values of AI to be?

Data protection and IP - the importance of data protection, IP ownership and cyber security must be recognised and balanced against the need to use data to promote innovation.

Mitigating social dislocation - the codes should confront what obligations rest on actors who deploy AI to mitigate the social dislocation that results.

Cybersecurity - the need for strong protection against hacking will increase as AI systems take a heightened role in society.

Set up an AI Advisory Body

Set up an AI Advisory Body: To consider ethical issues, foster discussion forums and publish resulting guidance to the industry and regulators. Communicate developments to the public to show initiative.

Gather intelligence on and participate actively in the development of such codes internationally

Gather intelligence on and participate actively in the development of such codes internationally: The "Asilomar AI Principles" and the "Partnership on AI" codes should be considered, among others to pick up on the latest thinking.

Develop core ethical principles

Develop core ethical principles: Engage with stakeholders to put together and publish fundamental ethical principles.

Encourage the development of sector specific codes

Encourage the development of sector specific codes: Particularly in fast moving areas, as mentioned above.

CONNECT WITH US

Christina Demetirades

Christina Demetriades

Senior Managing Director,

Deputy General Counsel

Mail to Christina Demetriades. This opens a new window.

Tom McLaughlan

Managing Director,

European Government Relations