



Department for Digital, Culture, Media and Sport

Establishing a pro-innovation approach to regulating AI
An overview of the UK's emerging approach

LONDON, 18 July 2022

*Presented to Parliament by the Secretary of State for Digital, Culture,
Media and Sport by Command of Her Majesty*

July 2022

CP 728



Department for Digital, Culture, Media and Sport

Establishing a pro-innovation approach to regulating AI
An overview of the UK's emerging approach

LONDON, 18 July 2022

*Presented to Parliament by the Secretary of State for Digital, Culture,
Media and Sport by Command of Her Majesty*

July 2022

CP 728

National AI Strategy



© Crown copyright 2022

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at www.gov.uk/official-documents

Any enquiries regarding this publication should be sent to: enquiries@dcms.gov.uk

Department for Digital, Culture, Media & Sport, 100 Parliament Street, London, SW1A 2BQ, United Kingdom

ISBN 978-1-5286-3639-1

E02773700 07/22

Printed on paper containing 40% recycled fibre content minimum.

Printed in the UK by HH Associates Ltd. on behalf of the Controller of Her Majesty's Stationery Office.

Establishing a pro-innovation approach to regulating AI

An overview of the UK's emerging approach

Across the world, AI is unlocking enormous opportunities, and the UK is at the forefront of these developments. As Secretary of State for the Department of Digital, Culture, Media and Sport, I want the UK to be the best place in the world to found and grow an AI business and to strengthen the UK's position so we translate AI's tremendous potential into growth and societal benefits across the UK.

Our regulatory approach will be a key tool in reaching this ambition. A regulatory framework that is proportionate, light-touch and forward-looking is essential to keep pace with the speed of developments in these technologies. Such an approach will drive innovation by offering businesses the clarity and confidence they need to grow while making sure we boost public trust.

Getting this right is necessary for a thriving AI ecosystem and will be a source of international competitive advantage. We will continue to advocate internationally for our vision for a pro-innovation approach to AI regulation recognising that both the opportunities and challenges presented by AI are fundamentally global in nature.

I am therefore pleased to publish this paper which sets out our emerging thinking on our approach to regulating AI. We welcome views on our proposals from across business, civil society, academia and beyond ahead of publishing a White Paper later in the year.



The Rt Hon Nadine Dorries MP
Secretary of State for Digital, Culture, Media and Sport
Department for Digital, Culture, Media & Sport

Executive Summary

In the National AI Strategy, the Government set out an ambitious ten-year plan for the UK to remain a global AI superpower. The UK is already a leader in many aspects of AI, with a thriving ecosystem and a strong track record of innovation. But there is more to do to harness the enormous economic and societal benefits of AI while also addressing the complex challenges it presents.

Establishing clear, innovation-friendly and flexible approaches to regulating AI will be core to achieving our ambition to unleash growth and innovation while safeguarding our fundamental values and keeping people safe and secure. Our approach will drive business confidence, promote investment, boost public trust and ultimately drive productivity across the economy.

The UK has a world leading regulatory regime - known for its effective rule of law and support for innovation.¹ We need to make sure that our regulatory regime is able to keep pace with and respond to the new and distinct challenges and opportunities posed by AI. This is key to remaining internationally competitive.

We are therefore proposing to establish a pro-innovation framework for regulating AI which is underpinned by a set of cross-sectoral principles tailored to the specific characteristics of AI, and is:

- **Context-specific.** We propose to regulate AI based on its use and the impact it has on individuals, groups and businesses within a particular context, and to delegate responsibility for designing and implementing proportionate regulatory responses to regulators. This will ensure that our approach is targeted and supports innovation.
- **Pro-innovation and risk-based.** We propose to focus on addressing issues where there is clear evidence of real risk or missed opportunities. We will ask that regulators focus on high risk concerns rather than hypothetical or low risks associated with AI. We want to encourage innovation and avoid placing unnecessary barriers in its way.
- **Coherent.** We propose to establish a set of cross-sectoral principles tailored to the distinct characteristics of AI, with regulators asked to interpret, prioritise and implement these principles within their sectors and domains. In order to achieve coherence and support innovation by making the framework as easy as possible to navigate, we will look for ways to support and encourage regulatory coordination - for example, by working closely with the Digital Regulation Cooperation Forum (DRCF) and other regulators and stakeholders.
- **Proportionate and adaptable.** We propose to set out the cross-sectoral principles on a non-statutory basis in the first instance so our approach remains adaptable - although we will keep this under review. We will ask that regulators consider lighter touch options, such as guidance or voluntary measures, in the first instance. As far as possible, we will also seek to work with existing processes rather than create new ones.

¹ See [World Bank Ease of doing Business](#), [OECD International Regulatory Cooperation](#) (May 2020), and [Global Innovation Index](#)

The approach outlined above is aligned with the regulatory principles set out in the Better Regulation Framework, which emphasise proportionate regulation.² It is also aligned with the government's vision set out through the Plan for Digital Regulation. It describes how we will take a pro-innovation approach to regulating digital technologies, which will deliver on the UK's desire to establish a more nimble regulatory framework now that we have left the EU.³

We recognise the cross-border nature of the digital ecosystem and the importance of the international AI market, and will continue to work closely with key partners on the global stage to shape global approaches to AI regulation. We will support cooperation on key issues, including through the Council of Europe, OECD working groups and the Global Partnership on AI and through global standards bodies such as ISO and IEC.

We welcome stakeholders' views on our proposed approach to regulating AI. Ahead of setting out further detail on our framework and implementation plans through the forthcoming White Paper, we are keen to seek reflections from across the AI ecosystem, wider industry, civil society, academia and beyond on the approach set out here to inform how we best shape the rules that will form part of the wider approach to how we regulate AI.

² The [Better Regulation Framework](#) builds on an independent report by the Taskforce on Innovation, Growth, and Regulatory Reform ([TIGRR](#)) which made recommendations on how the UK can reshape its approach to regulation and seize new opportunities from Brexit with its newfound regulatory freedom. The TIGRR report frames a UK approach to regulation based on the following principles: a. Proportionate; b. Forward-looking; c. Outcome-focussed; 5 d. Collaborative; e. Experimental; and f. Responsive.

³ [Policy Paper: The Benefits of Brexit](#) (January 2022)

Context

The UK has a thriving AI ecosystem. In 2021, the UK was first in Europe and third in the world for private investment in AI companies (\$4.65 billion) and newly funded AI companies (49).⁴ The UK is also first in Europe for the number of AI publications in 2021, and only topped by China, the USA and India.⁵

AI is unlocking huge benefits across our economy and society. In Glasgow AI is being used to track asbestos cancer tumours,⁶ in the Southeast to help people facing fuel poverty,⁷ in Belfast to improve animal welfare on dairy farms,⁸ and across the country by HM Land Registry to compare property transfer deeds.⁹ AI is also being applied to fundamental challenges in biology that will revolutionise drug discovery,¹⁰ and is set to impact the future of mobility¹¹ and an accelerated reduction in emissions.¹²

Alongside the benefits that AI brings, it also creates a range of new and accelerated risks, such as those associated with the use of AI in critical infrastructure to algorithmic bias. It also presents new questions for governments and society; for example, how should we protect existing rights in the context of systems that use facial recognition, or from large language models trained on content harvested from the web? How do we ensure commercial customers can confidently buy ‘off the shelf’ systems that are evidenced, tested and robust?

The answer to these questions will ultimately rely on actions by governments, regulators, technical standards bodies¹³ and industry. Together these form an overall approach to AI regulation.¹⁴

Overview of the existing regulatory landscape

The success of our AI ecosystem is in part down to the UK’s reputation for the quality of its regulators and its rule of law. This includes the transparency of the UK’s regulatory regime, the detailed scrutiny that proposed regulation receives and comprehensive impact assessments. This certainty around how new regulation will evolve has promoted private investment in the UK for developing new technologies and has allowed AI innovation to thrive.^{15, 16} To maintain our leading regulatory approach, we must make sure that the rules that govern the development and use of AI keep pace with the evolving implications of the technologies.

⁴ [Artificial Intelligence Index Report](#), Stanford (2022)

⁵ [OECD AI Policy Observatory - Live data](#), OECD (2022)

⁶ [AI technology used to track asbestos tumours](#), BBC (April 2021)

⁷ [Artificial intelligence project to help people facing fuel poverty](#), Energy Systems Catapult (2022)

⁸ [Why cows may be hiding something but AI can spot it](#), BBC (February [2022])

⁹ [HM Land Registry: Using AI for intelligence document comparison](#), Kainos

¹⁰ [AlphaFold: a solution to a 50-year-old grand challenge in biology](#), DeepMind (2020)

¹¹ [A new approach to self-driving: AV2.0](#), Wayve (2021)

¹² [Reduce carbon costs with the power of AI](#), BCG (2021)

¹³ This includes the International Organisation for Standardization (ISO), International Electrotechnical Commission (IEC) and Institute of Electrical and Electronics Engineers Standards Association (IEEE SA)

¹⁴ The following is taken from the government’s [Plan for Digital Regulation](#), published in July 2021: “‘Digital regulation’ refers to the range of regulatory tools that the government, regulators, businesses, and other bodies use to manage the impact that digital technologies and activities can have on individuals, companies, the economy and society. These include norms, self-regulation, statutory codes of conduct, and rules in primary legislation. We use these tools to promote outcomes that the market alone cannot achieve efficiently. Non-regulatory tools can complement or provide alternatives to ‘traditional’ regulation. This includes industry-led technical standards, which benefit from global technical expertise and best practice.”

¹⁵ The World Bank in its [Global Indicators of Regulatory Governance](#) analysis gives the UK a score of 5/5.

¹⁶ The 2021 edition of the [Global Innovation Index](#) (GII) gives the UK a score of 92.4/100 for ‘Regulatory Environment’.

While there are no UK laws that were explicitly written to regulate AI, it is partially regulated through a patchwork of legal and regulatory requirements built for other purposes which now also capture uses of AI technologies. For example,¹⁷ UK data protection law includes specific requirements around ‘automated decision-making’ and the broader processing of personal data,¹⁸ which also covers processing for the purpose of developing and training AI technologies. The upcoming Online Safety Bill also has provisions specifically concerning the design and use of algorithms.

Some UK regulators are also starting to take action to support the responsible use of AI. For example:

- The Information Commissioner’s Office (ICO) has issued multiple pieces of guidance, such as [Guidance on AI and Data Protection](#),¹⁹ [Explaining decisions made with AI](#),²⁰ [AI and Data Protection Risk Toolkit](#),²¹ [AI Auditing Framework](#) and [AI blog resources](#);²²
- The Equality and Human Rights Commission has identified AI as a strategic priority in its [Strategic Plan 2022-2025](#), and has committed to providing guidance on how the Equality Act applies to the use of new technologies, such as AI, in automated decision-making;²³
- The Medicines and Healthcare products Regulatory Agency has launched a [Software and AI as a Medical Device Change Programme](#)²⁴ and consulted on possible changes to the regulatory framework²⁵ to ensure the requirements provide a high degree of assurance that these devices are acceptably safe and function as intended.²⁶
- The Health and Safety Executive committed to develop collaborative research with industry and academia in its [Science and Evidence Delivery Plan 2020-2023](#), to determine a clear understanding of the health and safety implications of AI in the workplace.

Regulators are also working together to understand the impact AI technologies could have on our economy and society. The Digital Regulation Cooperation Forum (DRCF)²⁷ has been exploring the impact of algorithms across their industries and regulatory remits. It recently published the outputs of its first two research projects looking at the harms and benefits posed by algorithmic processing (including the use of AI), and at the merits of algorithmic auditing.²⁸

¹⁷ Other examples include equality law, which would apply where the use of AI produces discriminatory outcomes. Sector specific regulation such as for financial services and medical research may also capture the use of AI in these sectors.

¹⁸ UK GDPR and Data Protection Act 2018

¹⁹ [Guidance on AI and data protection](#), ICO

²⁰ [Explaining decisions made with AI](#), ICO

²¹ [AI and data protection risk toolkit](#), ICO

²² [ICO consultation \(now closed\) on the AI auditing framework](#), ICO (February 2020)

²³ [Strategic Plan 2022-2025](#), Equality and Human Rights Commission (March 2022)

²⁴ [Software and AI as a medical device change programme](#), Medicines and Healthcare products Regulatory Agency (September 2021)

²⁵ [Consultation \(now closed\) on the future regulation of medical devices in the UK](#), Medicines and Healthcare products Regulatory Agency (October/November 2021)

²⁶ [Science and Evidence Delivery Plan 2020-2023](#), Health and Safety Executive

²⁷ The [Digital Regulation Cooperation Forum](#) (DRCF) comprises the Competition and Markets Authority (CMA), the Information Commissioner’s Office (ICO), the Office for Communications (Ofcom) and the Financial Conduct Authority (FCA). It was established to build on the strong working relationships between these organisations and to establish a greater level of cooperation, given the distinctive challenges posed by digital regulation.

²⁸ [Findings from the DRCF algorithmic processing workstream](#) - Spring 2022, DRCF (April 2022)

In addition, the Bank of England and the Financial Conduct Authority established the Artificial Intelligence Public-Private Forum (AIPPF) to further dialogue on AI innovation in financial services between the public and private sectors. It recently published its report of this work.²⁹

Standards can also play a key part in developing a coherent regulatory approach, and the government is also already taking steps to develop world leading AI standards in the UK. In January 2022, the Department for Digital, Culture, Media and Sport (DCMS) announced the pilot of an AI Standards Hub to increase UK engagement in the development of global technical standards for AI. Multiple global standards development organisations (SDOs) have already published AI-specific standards, and more are under development. The Hub will create practical tools and bring the UK’s multi-stakeholder AI community together to ensure that global AI standards are shaped by a wide range of experts, to deliver the tools needed for AI governance, in line with our values.³⁰ In November 2021, the Central Digital and Data Office³¹ (CDDO) also published one of the world’s first national algorithmic transparency standards to strengthen trust in government use of algorithms and AI.

Assurance also plays an important role in complementing our regulatory approach. The UK’s Centre for Data Ethics and Innovation (CDEI) highlighted the need for robust AI assurance tools and services to ensure that stakeholders can understand the performance, risk and compliance of AI. In December 2021, the CDEI published a roadmap towards building a world-leading AI assurance ecosystem in the UK, and is now delivering a programme to ensure that the UK capitalises on its strengths in professional and legal services to lead the growth of this nascent industry.

Key Challenges

The proliferation of activity; voluntary, regulatory and quasi-regulatory, introduces new challenges that we must take action to address. Examples include:

- **A lack of clarity.** Stakeholders³² often highlight the ambiguity of the UK’s legal frameworks and application of regulatory bodies to AI, given these have not been developed specifically with AI technologies and its applications in mind. The extent to which UK laws apply to AI is often a matter of interpretation, making them hard to navigate. This is particularly an issue for smaller businesses who may not have any legal support.
- **Overlaps.** Stakeholders also note the risk that laws and regulators’ remits may regulate the same issue for the same reason and this can exacerbate this lack of clarity. This could lead to unnecessary, contradictory or confusing layers of regulation when multiple regulators oversee an organisation’s use of the same AI for the same purpose.

²⁹ [The AI Public-Private Forum: Final Report](#), Bank of England and Financial Conduct Authority’s Artificial Intelligence Public-Private Forum (February 2022)

³⁰ Standards are often used as “soft law” in codes of conduct/practice and binding/non-binding guidance, but it can also be designated as voluntary tools to show legal compliance. See [Designated standards guidance](#) on GOV.UK.

³¹ [Central Digital and Data Office \(CDDO\)](#)

³² [AI Barometer Part 1 - Summary of Findings](#), Centre for Data Ethics and Innovation (December 2021)

- **Inconsistency.** There are differences between the powers of regulators to address the use of AI within their remit³³ as well as the extent to which they have started to do so. AI technologies used in different sectors are therefore subject to different controls. While in some instances there will be a clear rationale for this, it can further compound an overall lack of clarity.
- **Gaps in our approach.** As current UK legislation has not been developed with AI in mind, there may be current risks that are already inadequately addressed, and future risks associated with widespread use of AI that we need to prepare for. For example, around the need for improved transparency and explainability in relation to decisions made by AI, incentivising developers to prioritise safety and robustness of AI systems, and clarifying actors' responsibilities. There is also concern that AI will amplify wider systemic and societal risks, for instance AI's impact on public debate and democracy, with its ability to create synthetic media such as deepfakes.

These issues across the regulatory landscape risk undermining consumer trust, harming business confidence and ultimately limiting growth and innovation across the AI ecosystem, including in the public sector.^{34, 35, 36} By taking action to improve clarity and coherence, we have an opportunity to establish an internationally competitive regulatory approach that drives innovation and cements the UK's position as an AI leader.

³³ For example, while the Information Commissioner's Office has the power to issue fines on parties that breach data protection law, the Equality and Human Rights Commission cannot issue fines on parties that breach equality law. The Equality and Human Rights Commission can, however, pursue damages in judicial review proceedings.

³⁴ 70% of surveyed businesses said they "desired more information to help them navigate the often complex legal requirements around data collection, use and sharing". [AI Barometer 2021](#), Centre for Data Ethics and Innovation (2021)

³⁵ 31% of respondents feel concerned that the benefits of data and AI use will not be felt equally across society. Public attitudes to data and [AI: Tracker Survey](#), Centre for Data Ethics and Innovation (2022)

³⁶ In November 2021, Meta (previously known as Facebook) announced it was "shutting down the Facial Recognition system on Facebook" citing unclear rules from regulators. Similarly, IBM is to stop offering its own facial recognition software for certain activities including mass surveillance.

The scope

To develop a clear framework for regulating AI, it will be critical to clarify its scope. However, there is currently little consensus on a general definition of AI, either within the scientific community or across national or international organisations.

AI is a general purpose technology like electricity, the internet and the combustion engine. As AI evolves, it will touch on many areas of life with transformative implications - although the precise impact of this technology will vary greatly according to its context and application.

The EU has grounded its approach in the product safety regulation of the Single Market, and as such has set out a relatively fixed definition in its legislative proposals.³⁷ Whilst such an approach can support efforts to harmonise rules across multiple countries, we do not believe this approach is right for the UK. We do not think that it captures the full application of AI and its regulatory implications. Our concern is that this lack of granularity could hinder innovation.

An alternative approach would be to put no boundaries on what constitutes AI, and leave regulators or relevant bodies to decide what technology and systems are in scope as they see fit. While such an approach would offer maximum flexibility, it raises the risk that businesses and the public would not have a consistent view of what is and is not the subject of regulation. A further risk is that any scope becomes defined via case law in the absence of a definition, which may vary by sector - and could add to further confusion.

Our preferred approach therefore is to set out the core characteristics of AI to inform the scope of the AI regulatory framework but allow regulators to set out and evolve more detailed definitions of AI according to their specific domains or sectors. This is in line with the government's view that we should regulate the use of AI rather than the technology itself - and a detailed universally applicable definition is therefore not needed. Rather, by setting out these core characteristics, developers and users can have greater certainty about scope and the nature of UK regulatory concerns while still enabling flexibility - recognising that AI may take forms we cannot easily define today - while still supporting coordination and coherence.

³⁷ [Proposal for a Regulation laying down harmonised rules on artificial intelligence](#), European Commission (April 2021)

Defining the core characteristics of AI

AI can have a wider number of characteristics and capabilities, depending on the techniques used and specifics of the use case. However, in terms of regulation, there are two key characteristics which underlie distinct regulatory issues which existing regulation may not be fully suited to address, and form the basis of the scope of this work:

The ‘adaptiveness’ of the technology - explaining intent or logic

AI systems often partially operate on the basis of instructions which have not been expressly programmed with human intent, having instead been ‘learnt’ on the basis of a variety of techniques;

AI systems are often ‘trained’ - once or continually - on data, and execute according to patterns and connections which are not easily discernible to humans. This ability underscores the power of modern AI, enabling it to produce incredibly intricate artwork based on a paragraph of text input,³⁸ diagnose illness in medical scans which are imperceptible to a human,³⁹ or complete missing elements of ancient texts.⁴⁰

For regulatory purposes this means that the logic or intent behind the output of systems can often be extremely hard to explain, or errors and undesirable issues within the training data are replicated. This has potentially serious implications, such as when decisions are being made relating to an individual’s health, wealth or longer term prospects, or when there is an expectation that a decision should be justifiable in easily understood terms - such as legal dispute.

The ‘autonomy’ of the technology - assigning responsibility for action

AI often demonstrates a high degree of autonomy, operating in dynamic and fast-moving environments by automating complex cognitive tasks. Whether that is playing a video game or navigating on public roads, this ability to strategise and react is what fundamentally makes a system ‘intelligent’, but it also means that decisions can be made without express intent or the ongoing control of a human.

While AI systems vary greatly, we propose that it is this combination of core characteristics which demands a bespoke regulatory response and informs the scope of our approach to regulating AI.

To ensure our system can capture current and future applications of AI, in a way that remains clear, we propose that the government should not set out a universally applicable definition of AI. Instead, we will set out the core characteristics and capabilities of AI and guide regulators to set out more detailed definitions at the level of application.

³⁸ [DALL.E 2](#), Open AI

³⁹ [Artificial intelligence is improving the detection of lung cancer](#), Nature (November 2020)

⁴⁰ [Restoring and attributing ancient texts using deep neural networks](#), Nature (March 2022)

Table 1. Example case studies: The regulatory implications of AI’s adaptive & autonomous characteristics

Case study scenario	How it is Adaptive	How it is Autonomous	Potential AI related regulatory implications
Transformer Language Model used to output text-based, or image-based content	Transformer models have a large number of parameters, often derived from data from the public internet. This can harness the collective creativity and knowledge present online, and enable the creation of stories and rich, highly-specific images on the basis of a short textual prompt.	These models generate their output automatically, based on the text input, and produce impressive multimedia with next to no detailed instruction or ongoing oversight from the user.	<ul style="list-style-type: none"> - Security and privacy concerns from inferred training data - Inappropriate or harmful language or content output - Reproduction of biases or stereotyping in training data
Self-driving car control system	These systems use computer vision as well as iteratively learning from real time driving data to create a model which is capable of understanding the road environment, and what actions to take in given circumstances.	These models directly control the speed, motion and direction of a vehicle.	<ul style="list-style-type: none"> - Safety and control risks if presented with unfamiliar input - Assignment of liability for decisions in an accident/dispute - Opacity regarding decision-making and corresponding lack of public trust

A new pro-innovation approach

Often the transformative effects of AI will be rapid and - at times - unexpected. There is therefore an important need to establish a clear framework which sets out how the government will respond to these opportunities as well as new and accelerated risks. This will offer greater clarity regarding how we intend to drive growth while also protecting our safety, security and fundamental values. In order to promote innovation and to support our thriving AI ecosystem, our approach will be:

- **Context-specific** - we will acknowledge that AI is a dynamic, general purpose technology and that the risks arising from it depend principally on the context of its application.
- **Pro-innovation and risk-based** - we will ask regulators to focus on applications of AI that result in real, identifiable, unacceptable levels of risk, rather than seeking to impose controls on uses of AI that pose low or hypothetical risk so we avoid stifling innovation
- **Coherent** - we will ensure the system is simple, clear, predictable and stable.
- **Proportionate and adaptable** - we will ask that regulators consider lighter touch options, such as guidance or voluntary measures, in the first instance.

A context-based approach allows AI related risk to be identified and assessed at the application level. This will enable a targeted and nuanced response to risk because an assessment can be made by the appropriate regulator of the actual impact on individuals and groups in a particular context. This also allows domains that have existing and distinct approaches to AI regulation such as defence to continue to develop appropriate mechanisms according to context. Relying on our existing regulatory structures also provides the flexibility to identify and adapt according to emerging risks since it is unlikely that new risks will develop in a consistent way across the entire economy.

Our approach will also be risk-based and proportionate. We anticipate that regulators will establish risk-based criteria and thresholds at which additional requirements come into force. Through our engagement with regulators, we will seek to ensure that proportionality is at the heart of implementation and enforcement of our framework, eliminating burdensome or excessive administrative compliance obligations. We will also seek to ensure that regulators consider the need to support innovation and competition as part of their approach to implementation and enforcement of the framework.

We think this is preferable to a single framework with a fixed, central list of risks and mitigations. Such a framework applied across all sectors would limit the ability to respond in a proportionate manner by failing to allow for different levels of risk presented by seemingly similar applications of AI in different contexts.⁴¹ This could lead to unnecessary regulation and stifle innovation. A fixed list of risks also could quickly become outdated and does not offer flexibility. A centralised approach would also not benefit from the expertise of our experienced regulators who are best placed to identify and respond to the emerging risks through the increased use of AI technologies within their domains.

⁴¹ See section 1 of [Regulating AI in Europe](#), Ada Lovelace Institute (March 2022)

We do, however, acknowledge that a context-driven approach offers less uniformity than a centralised approach - by its nature, it varies according to circumstance. That is why we wish to complement our context-based approach with a set of overarching principles to make sure that we approach common cross-cutting challenges in a coherent and streamlined way.

We are taking an actively pro-innovation approach. AI is a rapidly evolving technology with scope of application and depth of capability expanding at pace. Therefore, we do not think the government should establish rigid, inflexible requirements right now. Instead, our framework will ensure that regulators are responsive in protecting the public, by focusing on the specific context in which AI is being used, and taking a proportionate, risk-based response. We will engage with regulators to ensure that they proactively embed considerations of innovation, competition and proportionality through their implementation and any subsequent enforcement of the framework.

Cross-sectoral principles

While context is critical, AI technologies feature a range of underlying issues and risks which require a coherent response, such as a perceived lack of explainability when high-impact decisions are made about people using AI. We propose to address this by developing a set of cross-sectoral principles tailored to the distinct characteristics of these technologies. Regulators would be tasked with interpreting and implementing these cross-sectoral principles within their regulatory remit in line with their existing roles and remit. Our expectation is that our cross-sectoral principles will also provide a basis for coordination with our global partners and will support our implementation of the global principles that the UK has already helped to develop.

We propose developing a set of cross-sectoral principles that regulators will develop into sector or domain-specific AI regulation measures.

Early proposals for our cross-sectoral principles

Our proposed cross-sectoral principles build on the OECD Principles on Artificial Intelligence⁴² and demonstrate the UK's commitment to them. Our principles will provide a clear foundation for our framework, tailored to the UK's values and ambitions, and will be delivered within existing regulatory regimes. The principles will complement existing regulation, with our vision being to increase clarity and reduce friction for businesses operating in the AI lifecycle.

Our principles are deliberately 'values' focused - we want to make sure AI-driven growth and innovation is aligned with the UK's broader values given the vital role AI plays in shaping outcomes that affect our society. They are not, however, intended to create an extensive new framework of rights for individuals. The principles describe what we think well governed AI use should look like on a cross-cutting basis, but taken as part of our broader context-based, pro-innovation approach. For example, we expect well governed AI to be used with due consideration to concepts of fairness and transparency. Similarly, we expect all actors in the AI lifecycle to appropriately manage risks to safety and to provide for strong accountability.

⁴² [OECD AI Principles Overview](#)

Our proposal is that the principles will be interpreted and implemented in practice by our existing regulators. We are examining how the government can offer a strong steer to regulators to adopt a proportionate and risk-based approach (for example through government-issued guidance to regulators). The principles will ultimately apply to any actor in the AI lifecycle whose activities create risk that the regulators consider should be managed through the context-based operationalisation of each of the principles. For example, regulators will be tasked with deciding what 'fairness' or 'transparency' means for AI development or use in the context of their sector or domain. Regulators will then decide if, when and how their regulated entities will need to implement measures to demonstrate that these principles have been considered or complied with depending on the relevant context. We are also exploring ways to ensure that regulators can coordinate effectively to ensure coherence between their respective approaches to the principles, including where possible by working together to interpret or implement the principles on a joint or cross-sectoral basis.

Below we set out our early proposals for these cross-sectoral principles.

<p>Cross-sectoral principles for AI regulation</p>
<p>Ensure that AI is used safely</p> <p>The breadth of uses for AI can include functions that have a significant impact on safety - and while this risk is more apparent in certain sectors such as healthcare or critical infrastructure, there is the potential for previously unforeseen safety implications to materialise in other areas.</p> <p>As such, whilst safety will be a core consideration for some regulators, it will be important for all regulators to take a context-based approach in assessing the likelihood that AI could pose a risk to safety in their sector or domain, and take a proportionate approach to managing this risk. Ensuring safety in AI will require new ways of thinking and new approaches, however we would expect the requirements to remain commensurate with actual risk - comparable with non-AI use cases.</p>
<p>Ensure that AI is technically secure and functions as designed</p> <p>AI is rapidly bringing new capabilities online and reducing the costs of existing business functions and processes. Ensuring that consumers and the public have confidence in the proper functioning of systems is vital to guaranteeing that the research and commercialisation of AI can continue apace.</p> <p>AI systems should be technically secure and under conditions of normal use they should reliably do what they intend and claim to do. Subject to considerations of context and proportionality, the functioning, resilience and security of a system should be tested and proven, and the data used in training and in deployment should be relevant, high quality, representative and contextualised.</p>

Table continues on page 14

Make sure that AI is appropriately transparent and explainable

Achieving explainability of AI systems at a technical level remains an important research and development challenge. Presently, the logic and decision making in AI systems cannot always be meaningfully explained in an intelligible way, although in most settings this poses no substantial risk. However, in some settings the public, consumers and businesses may expect and benefit from transparency requirements that improve understanding of AI decision-making. In some high risk circumstances, regulators may deem that decisions which cannot be explained should be prohibited entirely - for instance in a tribunal where you have a right to challenge the logic of an accusation.

Taking into account considerations of the need to protect confidential information and intellectual property rights, example transparency requirements could include requirements to proactively or retrospectively provide information relating to: (a) the nature and purpose of the AI in question including information relating to any specific outcome, (b) the data being used and information relating to training data, (c) the logic and process used and where relevant information to support explainability of decision making and outcomes, (d) accountability for the AI and any specific outcomes.

Embed considerations of fairness into AI

In many contexts, the outcomes of the use of AI can have a significant impact on people's lives - such as insurance, credit scoring or job applications. Such high-impact outcomes - and the data points used to reach them - should be justifiable and not arbitrary.

In order to ensure proportionate and pro-innovation regulation, it will be important to let regulators continue to define fairness. However, in any sector or domain we would expect regulators to:

- interpret and articulate 'fairness' as relevant to their sector or domain,
- decide in which contexts and specific instances fairness is important and relevant (which it may not always be), and
- design, implement and enforce appropriate governance requirements for 'fairness' as applicable to the entities that they regulate.

Define legal persons' responsibility for AI governance

AI systems can operate with a high level of autonomy, making decisions about how to achieve a certain goal or outcome in a way which has not been explicitly programmed or even foreseen - which can raise secondary issues and externalities. This is ultimately what makes them intelligent systems.

Therefore, accountability for the outcomes produced by AI and legal liability must always rest with an identified or identifiable legal person - whether corporate or natural.⁴³

⁴³ For example, the [Law Commission](#) recommends that self-driving vehicles will represent a shift in responsibility from driver to manufacturer and operators.

Clarify routes to redress or contestability

AI systems can be used in ways which may result in a material impact on people's lives, or in situations where people would normally expect the reasoning behind an outcome to be set out clearly in a way that they can understand and contest - for example, when their existing rights have been affected. Using AI can increase speed, capacity and access to services, as well as improve the quality of outcomes. However it can also introduce risks, for example that the relevant training data reproduces biases or other quality concerns into an outcome.

Subject to considerations of context and proportionality, the use of AI should not remove an affected individual or group's ability to contest an outcome. We would therefore expect regulators to implement proportionate measures to ensure the contestability of the outcome of the use of AI in relevant regulated situations.

Case study: How our principles could apply to an AI start-up

An AI-first startup has created a platform that can automate complex customer-facing processes such as providing advice, sales and customer services, built on top of a Large Language Model.⁴⁴ On their product roadmap, the company has plans to expand into multiple regulated domains, using their technology to offer legal advice, financial advice and potentially even medical advice.

As a business, they have strong technical expertise and want to develop products to enter and expand into these sectors, but are holding back from investing time and resource in market development activities because of the uncertainty that comes with regulatory compliance. The business leaders assume the costs of regulatory compliance to be high, making the justification for investment difficult.

Our proposed framework will bridge this gap and provide the clarity that this company needs:

- The specific measures introduced by the regulators to implement each of our cross-sectoral principles will communicate clearly and in a coherent way to businesses what expectations are around the technical, internal processes and likely regulatory requirements.
- Relevant regulators could issue guidance to highlight relevant regulatory requirements such as sector-specific licences, standards or the need for named individuals to assume particular responsibilities. They would do this either jointly or in a coordinated way to minimise the risk of confusion or excessive burdens.

As a result this company can integrate these requirements into their product roadmap, understand the rules more easily and spend more time and resource on product development or fundamental AI research, and less on legal costs. The UK benefits both from the increased investment but also from the disruptive power of new technology-led business models increasing access to financial or legal advice.

⁴⁴ A Large Language Model (LLM) is an AI model which is trained on vast amounts of data - often with billions of parameters - which can produce content, such as text or visual output, on the basis of a short prompt from a user. Some examples are Open AI's GPT-3, DeepMind's Chinchilla, and Google's LaMDA.

Putting our approach into practice

We are still at the early stages of considering how best to put our approach into practice, and will set out fuller details through the forthcoming White Paper.

We propose initially putting the cross-sectoral principles on a non-statutory footing. This is so that we can monitor, evaluate and if necessary update our approach and so that it remains agile enough to respond to the rapid pace of change in the way that AI impacts upon society. This position would be kept under review as part of an ongoing process of monitoring and evaluating the effectiveness of the framework, including the principles and existing regulatory structures.

We propose that regulators will lead the process of identifying, assessing, prioritising and contextualising the specific risks addressed by the principles. We anticipate that the government may issue supplementary or supporting guidance, for example focused on the interpretation of terms used within the principles, risk and proportionality, to support regulators in their application of the principles. These principles provide clear steers for regulators, but will not necessarily translate into mandatory obligations. Indeed we will encourage regulators to consider lighter touch options in the first instance - for example, through a voluntary or guidance-based approach for uses of AI that fall within their remit. This approach will also complement and support regulators' formal legal and enforcement obligations using the powers available to in order to enforce requirements set out in statute.

Many regulators will have the flexibility within their regulatory powers to translate and implement our proposed principles, but not all. There are also differences between the types of rules regulators can make when translating these principles, and the enforcement action regulators can take where the underlying legal rules are broken. We need to consider if there is a need to update the powers and remits of some individual regulators. However we do not consider that equal powers or uniformity of approach across all regulators to be necessary.

Regulatory coordination will be important for our approach to work and to avoid contradictory or very different approaches across regulators. It will also be important to maintain a clear overview of how coherently the regulatory landscape as a whole is operating and to be able to anticipate issues arising from the implementation of our framework. We will look for ways to support collaboration between regulators to ensure a streamlined approach. For example, we will seek to ensure that organisations do not have to navigate multiple sets of guidance from multiple regulators all addressing the same principle. To do this, we will need to ensure we have the right institutional architecture in place. The UK already benefits from close cooperation between some of its regulators at a statutory level, and - in the digital space - from the ground-breaking work of the Digital Regulation Cooperation Forum, whose members have already begun to think actively about their shared priorities and areas of interest in relation to AI regulation.⁴⁵ We need to identify what further mechanisms, if any, are needed to ensure that this existing infrastructure can successfully support our goals for a coherent, decentralised framework.

⁴⁵ [Findings from the DRCF algorithmic processing workstream - Spring 2022](#), Digital Regulation Cooperation Forum (April 2022)

We also need to ensure that UK regulators have access to the right skills and expertise to effectively regulate AI. While some have been able to make significant and rapid investment in their AI capabilities in recent years, not all regulators have access to the necessary skills and expertise required. We will need to consider how we can address these disparities in a proportionate and innovative way; this could include consideration of the role that ‘pooled capabilities’ can play, as well as the effectiveness of secondments from industry and academia.

While we currently do not see a need for legislation at this stage, we cannot rule out that legislation may be required as part of making sure our regulators are able to implement the framework. For example, legislation may be necessary to ensure that regulators are able to take a coordinated and coherent approach. This could be relevant in the context of enabling and supporting regulatory coordination, or to make updates to regulatory powers. Alongside this, we may need to consider specific new powers or capabilities for regulators where risks associated with particular applications arise. However, we expect to pursue this approach by exception where it is the only viable option to address a high-impact risk

At this stage, we are considering implementing the principles on a non-statutory basis which could be supplemented by clear guidance from the government. This approach would be kept under review. We cannot, however, rule out the need for legislation as part of the delivery and implementation of the principles. For example, in order to enhance regulatory powers, ensure regulatory coordination, or to create new institutional architecture.

The international landscape

The inherent cross-border nature of the digital ecosystem and scientific collaboration as well as the importance of facilitating cross-border trade means it is imperative we work closely with partners. This is in order to prevent a fragmented global market, ensure interoperability and promote the responsible development of AI internationally.

We will continue to pursue an inclusive multi-stakeholder approach, to bring in relevant voices and expertise to help address these issues. We will also protect against efforts to adopt and apply these technologies in the service of authoritarianism and repression. To support this, the UK will continue to be an active player in organisations such as GPAI and the OECD, as well as acting as a pragmatic pro-innovation voice in ongoing Council of Europe negotiations. We will ensure that UK industry’s interests are well represented in international standardisation - both to encourage interoperability and to embed British values.

We will promote a pro-innovation international governance and regulatory environment for AI which fosters openness, liberty and democracy. We will work with partners around the world to ensure international agreements embed our values so that progress in AI is achieved responsibly, according to democratic norms and the rule of law. We will reject efforts to adopt and apply AI technologies to support authoritarianism or discrimination.

Next steps

This paper sets out our overall pro-innovation direction of travel on regulating AI. Over the coming months, we will be considering how best to implement and refine our approach to drive innovation, boost consumer and investor confidence and support the development and adoption of new AI systems. Specifically we will be considering:

- **The proposed framework** and whether it adequately addresses our prioritised AI-specific risks in a way tailored to the UK's values and ambitions while also enabling effective coordination with other international approaches. This includes considering whether any gaps exist in the existing regulator coverage that require a more targeted solution;
- **How we put our approach into practice.** This includes considering the roles, powers, remits and capabilities of regulators; the need for coordination and how this should be delivered across the range of regulators (statutory and non-statutory) involved in AI regulation; and whether new institutional architecture is needed to oversee the functioning of the landscape as a whole and anticipate future challenges. This includes the role of technical standards and assurance mechanisms as potential tools for implementing principles in practice, supporting industry, and enabling international trade. We will also consider if there are any areas of high risk that demand an agreed timeline for regulators to interpret the principles into sector or domain specific guidance. We will work with key regulators such as the Information Commissioner's Office, Competition and Markets Authority, Ofcom, Medicine and Healthcare Regulatory Authority and Equality and Human Rights Commission - as well as other stakeholders - to examine these questions⁴⁶;
- **Monitoring of the framework** to ensure it delivers our vision for regulating AI in the UK and that it is capable of foreseeing future developments, mitigating future risks and maximising the benefits of future opportunities. This includes designing a suitable monitoring and evaluation framework to monitor progress against our vision as well as criteria for future updates to the framework to ensure a robust approach to identifying and addressing evolving risks. This will be undertaken on two levels, both at the overall system level and at the individual regulator level. Our approach will also require consideration of how to ensure an effective and holistic horizon scanning function so we ensure our approach is suitable to address both immediate and long-term risks.

We will set out our position on these topics through the forthcoming White Paper and public consultation, which we plan to publish in late 2022.

⁴⁶ For example, the DCMS Secretary of State has already written to the Digital Regulation Cooperation Forum requesting their insights on AI governance.

Share your views

Through this paper, we want to invite stakeholder views about how the UK can best set the rules for regulating AI in a way that drives innovation and growth while also protecting our fundamental values. This will inform the development of the forthcoming White Paper.

We therefore welcome reflections on our proposed approach and we would like to specifically invite views and any supporting evidence that you can share with regard to the following questions:

1. What are the most important challenges with our existing approach to regulating AI? Do you have views on the most important gaps, overlaps or contradictions?
2. Do you agree with the context-driven approach delivered through the UK's established regulators set out in this paper? What do you see as the benefits of this approach? What are the disadvantages?
3. Do you agree that we should establish a set of cross-sectoral principles to guide our overall approach? Do the proposed cross-sectoral principles cover the common issues and risks posed by AI technologies? What, if anything, is missing?
4. Do you have any early views on how we best implement our approach? In your view, what are some of the key practical considerations? What will the regulatory system need to deliver on our approach? How can we best streamline and coordinate guidance on AI from regulators?
5. Do you anticipate any challenges for businesses operating across multiple jurisdictions? Do you have any early views on how our approach could help support cross-border trade and international cooperation in the most effective way?
6. Are you aware of any robust data sources to support monitoring the effectiveness of our approach, both at an individual regulator and system level?

The call for views and evidence will be open for 10 weeks, closing on 26 September 2022, to allow time for your consideration and response.

You can send your views on this to: evidence@officeforai.gov.uk. You can also write to us at:

Office for AI, DCMS, 100 Parliament Street, London, SW1A 2BQ

ISBN 978-1-5286-3639-1

E02773700 07/22