

Adapting the UK's pro-innovation approach to AI regulation for foundation models

How the advent of generative AI and large language models will require policymakers to strengthen current approaches to regulation.

Contents

01

In brief04

02

The five challenges of
foundation models08

03

Adapting current
regulatory approaches12

04

Reshaping
the future16





Authors



Harvey Lewis
Partner, Client
Technology & Innovation,
Ernst & Young LLP EY UK



Ansgar Koene
EY Global AI Ethics
and Regulatory Leader
EY – Global

SECTION

01

In brief

- The UK faces five unique challenges in regulating foundation models, including their adaptability, accessibility, and pace of development.
- Building on rather than replacing current proposals can create a more proactive, user-centric and accountability led framework for foundation models.



The UK government is charting new territory with its proposal for a novel framework to regulate artificial intelligence (AI).¹ This principles-based approach seeks to foster responsible innovation, create a context-sensitive governance regime, and reinforce the UK's status as a global AI leader. At the heart of this initiative is the desire to build trust and confidence in the design, development, and use of AI systems so that the UK can unlock their full potential to bring societal and economic advantages.

The UK's proposals place a particular focus on foundation models, which are defined as "an emerging type of general purpose AI that are trained on vast quantities of data and can be adapted to a wide range of tasks."² In recent months, new foundation models, and especially generative AI systems, such as OpenAI's GPT-4 large language model (LLM) or Midjourney's photorealistic text-to-image application, have been making headlines.^{3,4} These innovations unlock a vast range of novel use cases, creating transformative opportunities for organisations. The UK government recognises that foundation models could be "paradigm-shifting" and have significant impacts on society and the economy.⁵

However, the expansive capabilities and fast-paced development of foundation models also bring novel challenges for governments seeking to regulate AI, as we have seen from recent amendments to the European Parliament's draft of the EU Artificial Intelligence Act, for example.⁶ Even the most forward-thinking regulatory proposals can struggle to address the slew of complexities that foundation models create. Compared with other emerging technologies and even conventional 'narrow AI' systems, foundation models are more:

- ▶ **Adaptable:** Foundation models are a general-purpose technology, which opens an incredibly diverse array of applications and brings greater unpredictability in usage and impact.
- ▶ **Adjustable:** Foundation models can be fine-tuned with new data, allowing innumerable model variants to be developed for more focused or domain-specific tasks.

- ▶ **Accelerable:** Foundation models are developing faster than other sub-fields of AI, making it difficult to predict and manage future impacts or keep regulations up to date.
- ▶ **Accessible:** Foundation models can be accessed by anyone, which makes it difficult to ensure that individual users, as well as businesses, adhere to ethical guidelines and commercial best practices.
- ▶ **Autonomous:** Foundation models can be used to automate a vast range of tasks and services, which could transform conventional organisational structures and make it more difficult to assign responsibility.

In light of these problems, policymakers face crucial questions: How well does the UK's proposed framework tackle the unique challenges posed by foundation models, such as LLMs? And how can these regulatory frameworks be adapted to remain agile in the face of ongoing rapid technological advancements?

To address these questions, we consider the emerging challenges of foundation models and explore how the UK's approach can be adapted to overcome them. We pay particular attention to opportunities that build on the UK's existing proposals for a collaborative, principles-based and pro-innovation framework. By meticulously monitoring the development and deployment of foundation models, raising awareness of ethical guidelines, building capacity within the regulatory community, and engaging with stakeholders across the AI ecosystem, the UK government can ensure that its regulatory framework remains responsive to the challenges and opportunities presented by novel AI technologies, fostering responsible innovation and upholding public trust.

1. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>, accessed 11 May 2023

2. Ibid

3. GPT-4 (openai.com), accessed 2 May 2023

4. "AI Deep Fake of the Pope's Puffy Coat Shows the Power of the Human Mind", Pope Francis' White Puffer Coat AI Image Sparks Deep Fake Concerns – Bloomberg, accessed 15 May 2023

5. <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>, accessed 11 May 2023

6. EU lawmakers pass draft of AI Act, includes copyright rules for generative AI | VentureBeat, accessed 2 May 2023

The UK's principles-based approach

The UK is proposing an innovative approach towards regulating AI, which aims to balance opportunity and risk while strengthening the country's global AI leadership. Unlike the more rigid and centralised approaches proposed in some other jurisdictions, the UK's framework is designed to be agile, aligning with the findings of Sir Patrick Vallance's Pro-innovation Regulation of Technologies Review.⁷ This approach aims to provide clarity to innovators and enable better trust and experimentation through collaboration among government, regulators, industry and civil society. It proposes a unique definition of AI and suggests a principles-based approach to respond to AI-related opportunities and risks.

Key components of the proposed regulatory approach include:



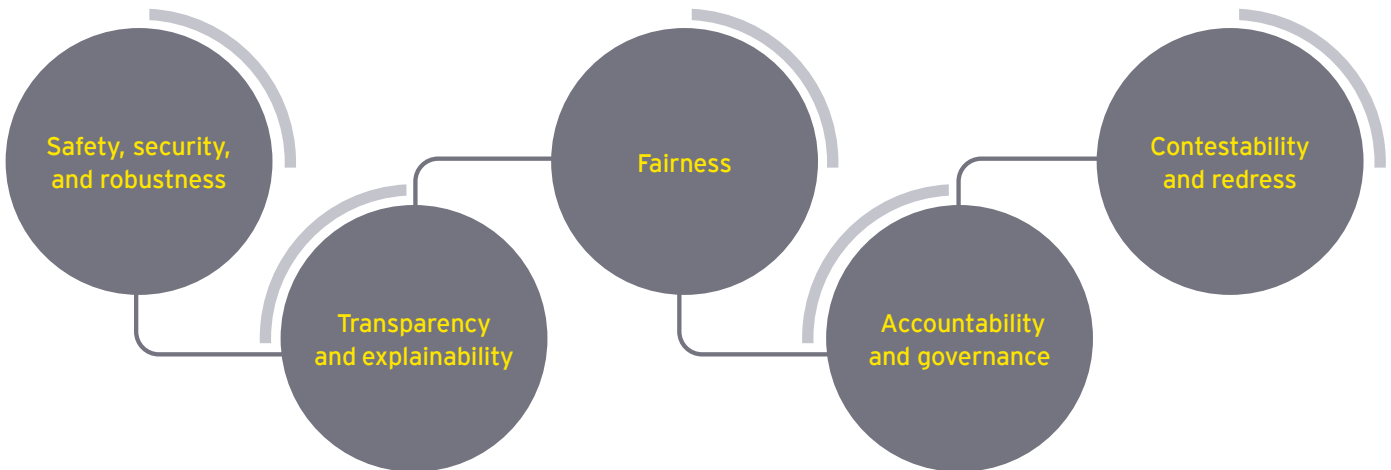
7. "Pro-innovation Regulation of Technologies Review: Digital Technologies", HM Treasury, Pro-innovation Regulation of Technologies Review: Digital Technologies – GOV.UK (www.gov.uk), accessed 15 May 2023

The framework is structured around four key elements:

- | | |
|--|--|
| <p>1 Defining AI based on unique characteristics, which supports regulator coordination and provides a clear understanding of AI.</p> | <p>3 Providing cross-sectoral principles, which guide regulatory responses to AI risks and opportunities and describe good governance at all stages of the AI life cycle.</p> |
| <p>2 Adopting a context-specific approach, which, instead of assigning rules to sectors or technologies, proposes to regulate based on the outcomes AI is likely to generate.</p> | <p>4 Delivering central functions to support regulators to ensure that the framework is coherent and maximises the benefits of the iterative approach.</p> |
-

The proposed AI definition is based on two characteristics that necessitate a bespoke regulatory response: the 'adaptivity' and the 'autonomy' of AI. The former refers to the capacity of AI systems to learn from data and perform new forms of inference, while the latter refers to the ability of AI systems to make decisions without human intervention.

The proposed framework's principles-based approach underpins five values-focused, cross-sectoral principles:



These principles reflect the OECD's values-based AI principles and are expected to guide businesses and regulators in implementing the framework.⁸

8. The OECD Artificial Intelligence (AI) Principles – OECD.AI, accessed 2 May 2023

SECTION

02

The five challenges of foundation models



Regulating the dynamic landscape of AI is an increasingly complex task, particularly when it comes to generative AI and foundation models, like LLMs.



Foundation models present unique challenges, from their broad range of applications to their rapid development and difficulties in establishing accountability. It is critical that regulators understand these nuances to strengthen proposals and address the associated gaps in the framework.

1. Adaptability

One of the most pressing challenges arises from the general-purpose nature of foundation models. Unlike their narrow AI counterparts, which are designed for specific tasks, foundation models are built to be versatile and adaptable. They are like Swiss Army Knives, capable of being used in a multitude of ways that are difficult to foresee, which greatly complicates the creation of comprehensive regulations and the assignment of responsibility across a model's full value chain.

The UK's sector-based regulatory approach has worked well for many industries, like healthcare or banking, where the application of technology is specific, and the context is well-defined, which is the case for more traditional narrow AI systems. For example, new amendments to the UK Highway Code make it clear that drivers travelling in autonomous vehicles on the road must be alert and ready to take back control in a timely manner.⁹

However, when it comes to regulating foundation models, the sector-based approach could falter. Given the broad range of applications, from writing news articles to creating music to predicting stock prices, it becomes nearly impossible to devise sector-specific regulations that can effectively govern everything a single technology touches. What was already difficult for narrow AI – ascribing responsibility for intent or knowledge of wrongdoing, a concept known as *mens rea* – becomes much more difficult when a single AI model can be used to make decisions across a swathe of different contexts, including for some applications where the model's inherent complexity generates unpredictable results.

Let's take the example of an LLM being used in the financial sector to predict stock market trends. While it could be regulated under the existing financial regulatory framework, the same model could be used in healthcare to predict disease outbreaks or in journalism to write articles. Each of these uses falls under a different regulatory body, each with its own set

of rules and regulations. This poses the question: under whose purview does the regulation of such a versatile model fall?

Furthermore, context-sensitive regulations, which are designed around specific use cases, may also struggle to keep up with the unpredictable and wide-ranging impacts of foundation models. The same LLM can generate both beneficial and harmful outputs, depending on the specific use case.

For example, an LLM could be fine-tuned to generate medical advice. In a beneficial context, this might provide health information to individuals who otherwise lack access to medical professionals. However, in a harmful context, the same model could disseminate misleading or incorrect health information, with potentially dangerous consequences. Context-sensitive regulations may struggle to anticipate these varied outcomes and, thus, fail to adequately protect against potential misuse.

2. Adjustability

The very adaptability and versatility that make foundation models such a powerful tool also pose significant challenges for accountability. These challenges become even more pronounced when we factor in the ability to fine-tune these models, particularly those that are open-sourced. Fine-tuning allows users to adjust the base model to perform specific tasks, or to exhibit certain behaviours, based on new data. This could range from tuning a model to generate poetry in a particular style, to programming it to write convincing phishing emails. While the former is a creative and potentially beneficial use, the latter can be used to facilitate harmful activities.

This fine-tuning process complicates the assignment of accountability. For instance, if a fine-tuned model is used to generate harmful content, who is responsible? Is it the creators of the original foundation model, the individuals or organisations that fine-tuned the model, or the person who used the fine-tuned model to generate the content? The lines of

9. "UK paves way for public use of autonomous vehicles", Autovista24, UK paves way for public use of autonomous vehicles | Autovista24 (autovistagroup.com), accessed 22 May 2023

responsibility become blurred, leading to legal grey areas and difficulties for users in contesting decisions or seeking redress.

In addition, although historically open-source has been seen as a safe, trusted, and transparent mechanism for technology release, particularly in academia, the power of today's AI models means that open-source availability can also facilitate misuse. Consider a scenario where an open-source LLM is fine-tuned to generate disinformation. The individual who fine-tuned the model might be in one country, the person who used the model to spread disinformation in another, and the victims of the disinformation in yet another. Given the cross-border nature of this issue, determining legal jurisdiction and responsibility can be a regulatory headache.

3. Acceleration

The mismatch between the speed of technological development and the typically slower pace of policy response, can lead to a regulatory gap, leaving new technologies and their impacts unaddressed, at least temporarily. Consider the debates around privacy, 'hallucinations and potential plagiarism affecting ChatGPT soon after its launch.'¹⁰

One of the significant challenges in this period of technological turbulence is monitoring and evaluating the risks (or benefits) associated with advancements. Tracking the evolution of these models and assessing their capabilities becomes an onerous task. It's not just about monitoring the creation of new models, but also keeping tabs on fine-tuned models and their uses across various domains.

For instance, let's consider a situation where a new LLM is developed and released into the open-source domain. Regulators would need to understand the model's capabilities and potential uses. However, once the model is released, it can be fine-tuned by various users, each modification leading to a new variant with potentially different capabilities, applications, and risks. This proliferation of variants makes the task of monitoring and evaluating the models akin to herding cats.

The UK's proposals suggest that one way of evaluating the potential impact of LLMs could be by monitoring the amount of compute used to train them. But computational efficiency is also improving rapidly. So much so that some large tech labs and the open-source community have already released approaches for pre-training LLMs on widely available devices like consumer GPUs, smartphones and tablets.¹¹ Unfortunately, the use of compute as the core metric in evaluating foundation models also fails to consider the much smaller computational effort required for fine-tuning.¹² This is already leading to a 'Cambrian explosion' of powerful, domain-specific LLMs, which consume exponentially fewer processing resources during their training.

Moreover, the rapid pace of development also raises the question of how to evaluate the effectiveness of regulations. Drawing a parallel with the regulation of the internet, we saw how initial attempts to regulate this rapidly developing technology often fell short due to the inability to anticipate its evolution. Policymakers had to grapple with emerging issues like data privacy, cybercrime, and online misinformation, often retrospectively, leading to a reactive rather than a proactive regulatory approach. With the technological landscape around foundation models in constant flux, gauging whether regulations are achieving their intended effect becomes a moving target. By the time the impact of a regulation is assessed, the technology it was designed to regulate may have evolved or been superseded.

4. Accessibility

Foundation models, like LLMs, are inherently more accessible than other, narrower types of AI because they can interact with people in ways that mimic humans. In addition, several companies developing these models have chosen to make them available to the public on a free-of-charge basis. This accessibility poses a significant regulatory challenge. On the one hand, we see the democratisation of AI capabilities, a world where anyone with a smartphone or laptop and an internet connection can harness the power of advanced AI for innovation and creativity. Yet, on the other hand, this

10. "Hallucinations, plagiarism and ChatGPT", AI News, Hallucinations, plagiarism, and ChatGPT (artificialintelligence-news.com), accessed 15 May 2023

11. "LLM on Android with Keras and TensorFlow Lite", Machine Learning – Toolkit LMM – Google AI, accessed 12 May 2023

12. "Fine-tuning 20B LLMs with RLHF on a 24GB consumer GPU", Hugging Face, Fine-tuning 20B LLMs with RLHF on a 24GB consumer GPU (huggingface.co), accessed 15 May 2023

accessibility potentially opens Pandora's box, offering up the same capabilities for misuse and malicious intent.

Traditionally, regulatory frameworks have been designed to govern the actions of businesses and organisations, rather than individuals. This is because businesses typically have a more substantial impact on society, operate at a larger scale, and are easier to monitor and hold accountable. But in the world of AI and foundation models, this approach may not be as effective.

Consider the example of an individual using a foundation model to create deepfakes – realistic but fake videos or images, often used to spread disinformation or perpetrate fraud. Under the current regulatory frameworks, it would be challenging to hold this individual accountable, particularly if they are located in a different jurisdiction or are using anonymising technology to hide their identity. This poses a significant challenge for regulators who are used to dealing with identifiable, regulated entities.

Furthermore, the sheer number of individual users makes monitoring and enforcement a daunting task. Regulating businesses and organisations often involves a manageable number of entities, allowing for in-depth inspections and enforcement actions. However, with potentially millions of individual users, regulators risk being overwhelmed.

This issue is further compounded by the lack of specialised training or equipment needed to use these models. This low barrier to entry means that even users with good intentions may lack a full understanding of the potential impacts and ethical implications of their actions, making misuse – either intentional or unintentional – more likely.

5. Autonomy

With the emergence of open-source frameworks like 'AgentGPT' and 'BabyAGI', the potential for LLMs to disrupt conventional organisations cannot be overlooked.¹³ This

shift towards radical automation, where a handful of people or even a single individual can create and manage an entire organisation of AI bots, reshapes traditional organisational structures and accountability lines, thus posing a significant regulatory challenge.

In a traditional organisation, accountability is distributed among various roles and responsibilities. If something goes wrong, there are human decision-makers, reviewers and auditors who can be held accountable. However, in the context of an organisation dominated by AI bots, the notion of accountability becomes murky. How can liability be assigned when decisions are made by algorithms? Who is responsible when an AI bot makes a decision that leads to harm or loss?

This is a particularly pertinent issue in the context of the General Data Protection Regulation (GDPR) in the EU. Under GDPR, individuals have a right to an explanation when subjected to automated decision-making that has legal or significant effects on them. But in an organisation run by AI bots, who would provide this explanation? If the AI's decision-making process is too complex to be easily understood, how can this right be upheld?

Answering these questions leaves pro-innovation policymakers in something of a dilemma: the seemingly simple approach of ensuring that blame can be assigned to a 'responsible human' may act to discourage innovation and the creation of such autonomous organisations.

Moreover, there are considerable economic and societal impacts. While automation can lead to increased efficiency and cost savings, it also risks widening the inequality gap. As we are seeing, jobs in the technology and telecommunications sectors are already being lost to AI bots.^{14,15} And the benefits of such automation could be concentrated in the hands of a few corporations or individuals who control the AI models. This could lead to social disruption and exacerbate existing economic disparities.

13. "Auto-GPT, BabyAGI, and AgentGPT: How to use AI agents", Auto-GPT, BabyAGI, and AgentGPT: How to use AI agents | Mashable, accessed 12 May 2023

14. "IBM to Pause Hiring for Jobs that AI Could Do", Bloomberg, <https://www.bloomberg.com/news/articles/2023-05-01/ibm-to-pause-hiring-for-back-office-jobs-that-ai-could-kill>, accessed 22 May 2023

15. "BT to cut 55,000 jobs with up to a fifth replaced by AI", BBC News, BT to cut 55,000 jobs with up to a fifth replaced by AI – BBC News, accessed 22 May 2023

SECTION

03

Adapting current regulatory approaches

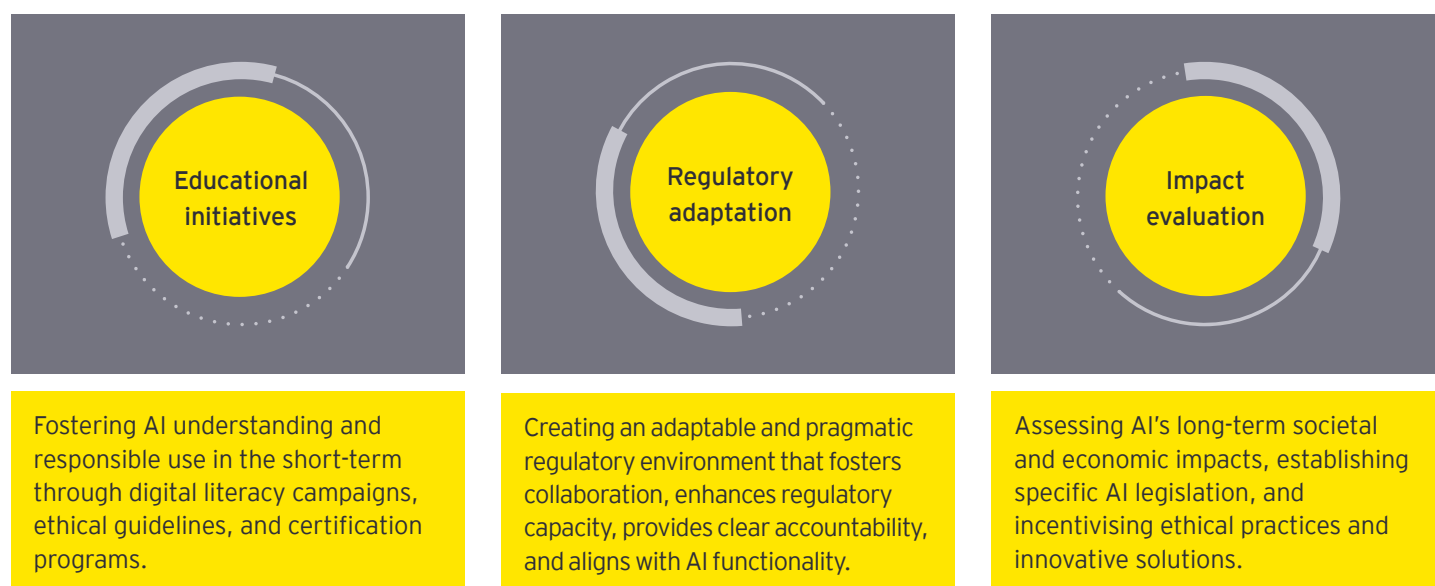


Clearly, regulating generative AI and foundation models is not a straightforward task.



Successful regulation of foundation models demands an understanding of the complex landscape these technologies inhabit, the ability to anticipate potential risks and opportunities, and the willingness to adapt as the technology evolves. Policymakers and regulators need to approach this with a mix of caution, innovation, and adaptability to ensure that the benefits of these technologies are maximised while the risks are mitigated.

Below, we outline ten recommendations for the UK’s policymakers to strengthen proposed regulatory approaches specifically for foundation models. We have grouped these into three thematic areas:



Educational initiatives

Build greater public and regulatory understanding

As foundation models become increasingly accessible, there's a pressing need for regulators to consider how the essential principles can permeate beyond traditional industry frontiers and encompass all users. This shift could be facilitated by implementing targeted digital literacy initiatives and awareness campaigns designed to educate individuals on foundation models and how to use or interact with AI responsibly.

For certain types of use of foundation models, a parallel can be drawn with obtaining a driver's licence: users should be able to demonstrate a rudimentary understanding of the potential risks and ethical dilemmas associated with the generation of content or fine-tuning before gaining access and using models. Encouraging model developers to publish ethical guidelines

alongside other terms and conditions of use, and asking users to acknowledge these as part of the sign-up process might help to improve levels of awareness and understanding of the risks associated with the technology. To ensure that as many users as possible are informed, model developers could also offer free-to-access tutorials and online demonstrations that showcase responsible use – an approach similar to that adopted by the gaming industry, which typically offers ‘tutorial modes’ to help users understand and learn the mechanics of gameplay before entering the real player arena.

For the most powerful generative AI and LLMs, a certification programme could be instated, which requires users to complete a basic form of training before being granted access to the models. This would not only deepen the understanding of ethics and potential risks but could also create a registry of certified users, enhancing the ability of regulators to track misuse.

Regulatory adaptation

Foster collaboration

Creating an inclusive platform for dialogue between government, regulators, AI developers, and foundation model users can foster collaboration. Just as the internet does not heed national borders, this platform needs to consider both UK interests and diverse international perspectives on the regulation of AI. It can serve as a real and virtual meeting point for sharing insights and co-developing new approaches to foundation models.

Regular roundtable discussions, public consultations, and co-creation workshops can bridge the gap between regulators, developers and users – both businesses and individuals. This ensures that regulations are practical, effective, and informed by a broad spectrum of stakeholders. For foundation models, it is more important than with other forms of AI to ensure that regulatory approaches are globally harmonised, acknowledging the inherently transnational nature of these technologies. Just as global trade agreements align economic policies across borders, international collaboration on foundation models can ensure a level playing field and prevent regulatory arbitrage.

Build additional regulatory capacity

Investing in capacity-building initiatives by equipping central functions and sector-based regulators with the requisite skills and knowledge ensures that the regulatory framework remains nimble and responsive to the rapid evolution of foundation models.

Incorporating an enhanced definition of AI to spotlight the unique attributes of foundation models could also help to concentrate regulatory attention on aspects that require additional scrutiny.

Expand regulatory sandboxes

Broadening the scope of regulatory ‘sandboxes’ to include dedicated capabilities for foundation models can provide a controlled environment for regulators, businesses, and potentially individuals, to experiment with model variants and use cases. These sandboxes would allow for safe testing and experimentation before full deployment. They would also provide opportunities for regulators to test technology-driven

approaches to improve the safety and security of foundation models. For example, regulators could explore the use of prompt engineering or prompt injection techniques to reduce bias, improve accuracy and embed stronger guardrails to prevent the generation of harmful content in a more robust or automated way.

Develop clearer accountability frameworks

Creating legal guidelines that delineate responsibilities across the AI lifecycle can establish clear accountability frameworks. For instance, a ‘human-in-the-loop’ requirement could ensure that significant decisions made by AI bots or agents once deployed are reviewed and approved by a human operator. This could provide a check on the AI’s decisions and ensure that there is a human who can be held accountable. Prior to deployment, an equivalent certification method could also be adopted for model design, development and testing, and for any downstream model enhancements, such as fine-tuning or open sourcing.

A further step could be to require the presence of a ‘responsible human’ for every AI bot or within every organisation. This individual would be legally accountable for the actions of the foundation model. Additionally, ‘bot registries’ could be created, which would track the ownership and usage of specific types of AI bots, such as those used for processing and decision-making involving personal data, which could enhance transparency and provide clear lines of responsibility.

Establish a regulatory monitoring body

Creating a regulatory monitoring body would enable surveillance of the web and the use of other tracking tools to provide immediate alerts about new foundation models or publicly available fine-tuned variants, applications and potential risks. This would help with ongoing systematic reviews of the regulations to ensure they remained fit for purpose as the technology continues to evolve. The role of such a monitoring body would be distinct from the crucial work of regulators, such as Ofcom or the Information Commissioner, to monitor compliance with the Online Safety Bill or to reduce the deliberate publication of mis- and disinformation that could be potentially harmful to consumers.¹⁶

16. Online Safety Bill, UK parliament, Online Safety Bill – Parliamentary Bills - UK Parliament, accessed 2 June 2023

This monitoring body could also conduct audits, assess risks, and provide enhanced guidance for the responsible use of open-source AI. In addition, a certification system for fine-tuned models could be established to ensure that they meet certain ethical and safety standards before they are deployed. If the monitoring body also created a public repository of such AI systems, this would provide assurance to users and developers that the models listed were accredited.

Regulate by functionality

Given the general-purpose nature of foundation models, regulating based on functionality rather than sector could prove more effective. For instance, a ‘text-generation’ regulation could apply across sectors and use-cases, from news generation to email marketing, providing a uniform approach that can accommodate the diversity of uses. This approach can help overcome the challenge of the versatility of foundation models, ensuring they are regulated based on what they do, rather than where they are used.

Impact evaluation

Promote a duty of care

Policymakers could play the role of societal influencers, for example by promoting a ‘duty of care’ approach to require providers of foundation models to prevent misuse and harm.

Such duty of care might include requiring organisations to establish a dedicated ethics board, including representation from technology teams, ethicists and independent parties, to encourage the development and regular update of corporate ethical guidelines. Transparency could also be enhanced by encouraging AI developers in companies and academia to share non-sensitive aspects of their models’ design, development and application.

Consider legislative measures

To tackle issues associated with AI-generated content, the UK government could consider specific legislation, much like rules governing copyright infringement. These laws could outline protocols for attributing AI-generated content, marking it as ‘AI-generated’, and preventing plagiarism.

Intensifying legal consequences for misuse of AI technologies could act as a deterrent, ensuring that individuals who misuse these technologies face appropriate consequences.

Furthermore, the government could work with tech companies to develop both protection and detection technologies to identify content as AI-generated and to detect and remove such content that violates the new rules. New legislation could require all AI-generated images, videos, or audio to carry an invisible watermark or metadata, acting as a digital signature indicating their AI origin for enhanced traceability.

Establish impact assessments and an impacts panel

Assessing the potential impact of foundation models is crucial for creating trust and embedding the concept of responsible AI. Before deploying a foundation model, an AI Impact Assessment could be conducted. This assessment would evaluate the potential risks and benefits of the model, consider the ethical implications, and plan for mitigation strategies. Much like Environmental Impact Assessments in construction projects or Privacy Impact Assessments to ensure the privacy-preserving nature of data handling processes, AI Impact Assessments could ensure responsible and ethical use of foundation models.

The assessment process could involve rigorous analysis of the model’s potential impact on various societal facets, including privacy, fairness, and employment, among others. The role of the assessment isn’t merely to identify risks, but also to propose solutions to mitigate them. Just as Health Impact Assessments in public health guide the formulation of healthier public policies, AI Impact Assessments can inform the deployment and use of powerful AI models.

Moreover, the Government’s recently proposed Foundation Model Taskforce could establish a panel to assess the societal and economic impacts of foundation models across the UK.¹⁷ This panel should include experts from various fields, such as AI, ethics, economics, and the social sciences. The panel would conduct regular reviews to assess the impact of foundation models on various sectors of the economy and society and propose recommendations to policymakers to optimise their use. By maintaining a bird’s eye view of the deployment of AI across the country, the panel can ensure that the benefits of AI are reaped while minimising potential harms.

17. “UK government creates AI taskforce to look at foundation models”, Tech Monitor, UK AI taskforce launch by government to look at foundation models (techmonitor.ai), accessed 15 May 2023

SECTION

04

Reshaping the future



Foundation models are at the forefront of the AI revolution, representing a seismic shift in the landscape of artificial intelligence. Their capabilities, versatility and adjustability, rapid development, and radical accessibility present both extraordinary opportunities and significant challenges for governments. Foundation models are like a double-edged sword; while the potential for societal advancement is immense, so too are the risks.

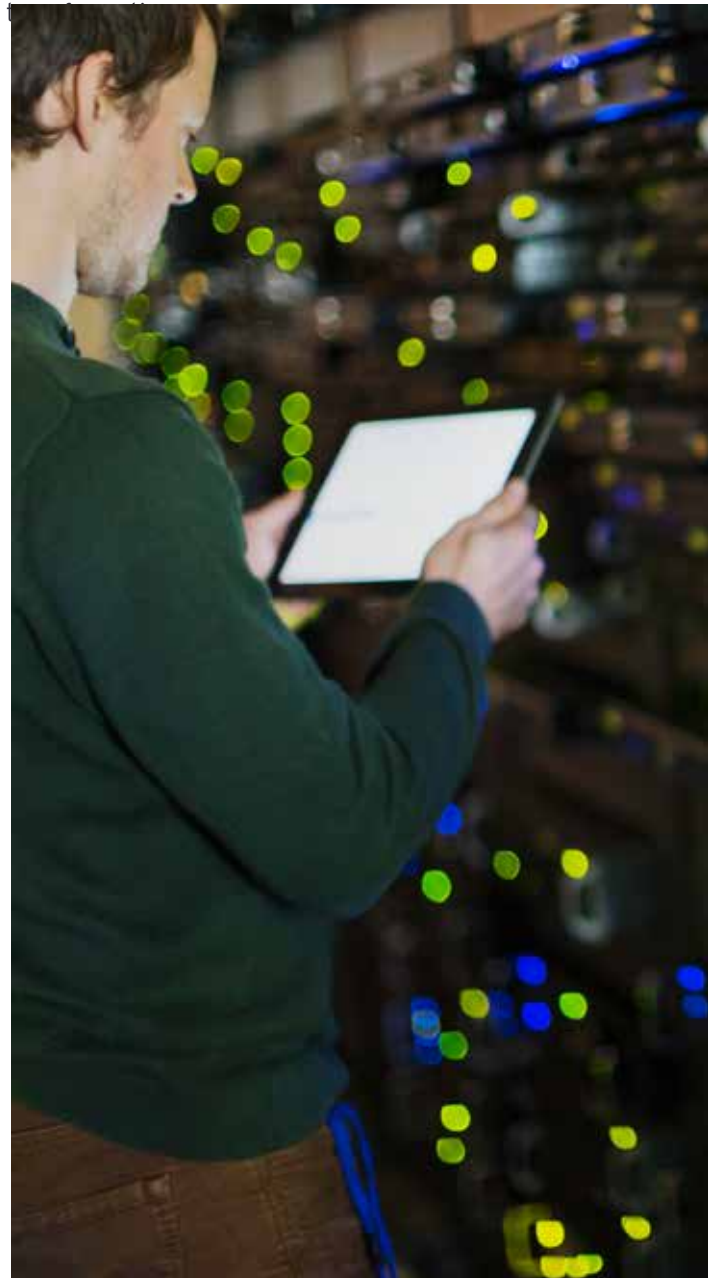
Imagine a world where the challenges of foundation models are not tackled. In their unregulated state, foundation models would come to permeate all corners of society, wielding their transformative power unchecked. The pace of their development, coupled with their adaptability and accessibility, means that anyone, regardless of their intentions, could wield this potent tool.

This lack of oversight could lead to a range of damaging consequences. Misuse of these models could undermine privacy, propagate misinformation, and – in the extreme – even destabilise economies. Without regulation, the potential for harm is significant.

Conversely, consider a world where these challenges are appropriately managed within the envelope of the UK's current approach. Foundation models would still transform society, but within a framework that promotes ethical use, fosters transparency, and ensures accountability. With a collaborative platform, rigorous impact assessments, and an established monitoring body, the AI revolution would be steered responsibly.

This proactive pro-innovation and principles-based approach to regulation ensures that the benefits of foundation models can be harnessed, while their potential harms are mitigated. Through this approach, The UK will shape an AI-driven future that is inclusive, ethical, and beneficial to all.

The road ahead is both exciting and fraught with challenges. The policy decisions made today will shape the UK's AI-driven future. Given the unique nature of foundation models, ignoring the challenges runs the risk of opening Pandora's box. However, an approach which adapts and builds on the current white paper, embracing comprehensive, proportionate and pro-innovation regulation, will ensure the responsible use of foundation models and create a future of opportunity and



EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

© 2023 EYGM Limited.
All Rights Reserved.

EYG no. 006095-23Gbl

BMC Agency
GA 174132617
ED None



In line with EY's commitment to minimize its impact on the environment, this document has been printed on paper with a high recycled content.

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

ey.com