

Policy Brief:
GENERATIVE AI

Dr Ann Kristin Glenster & Sam Gilbert

This report is authored by:

Dr Ann Kristin Glenster

Minderoo Centre for Technology and Democracy

Sam Gilbert

Bennett Institute for Public Policy

October 2023

CONTENTS

FOREWORD

Which path should the UK take to build national capability for generative AI?

The rapid rollout of generative AI models, and public attention to Open AI's ChatGPT, has raised concerns about AI's impact on the economy and society. In the UK, policy makers are looking to large language models and other so-called foundation models as ways to potentially improve economic productivity.

This policy brief from Dr Ann Kristin Glenster and Sam Gilbert outlines which policy levers could support those goals. They argue that the UK should pursue becoming a global leader in applying generative AI to the economy. Rather than use public support for building new foundation models, the UK could support the growing ecosystem of startups that develop new applications for these models, creating new products and services.

A UK approach to generative AI could leverage the existing national strengths in safe, responsible and ethical AI to put human safety and flourishing at the forefront of innovation. A national approach could achieve these goals by increasing understanding of and access to generative AI tools throughout the economy and society.

This policy brief answers three key questions:

1. What **policy infrastructure and social capacity** does the UK need to lead and manage deployment of responsible generative AI (over the long term)?
2. What **national capability** does the UK need for large-scale AI systems in the short- and medium-term?
3. What **governance capacity** does the UK need to deal with fast moving technologies, in which large uncertainties are a feature, not a bug?

Thanks to Ann Kristin and Sam's extensive research, this policy brief maps out an ethical framework for the governance of generative AI, through the creation of an AI Bill.

We hope that this policy brief will be useful to a wide range of stakeholders and address how we can use regulatory and legislative power today, to ensure that the British public can trust how this technology is used.

We are also excited that this policy brief brings together expertise from three groups at the University of Cambridge: the Bennett Institute for Public Policy, Minderoo Centre for Technology and Democracy and ai@cam.

Evidenced-based, science-informed research like this brief is what our three organisations do best, and we hope that our insights can help decision-makers navigate public debates and policy choices with more clarity.

Professor Dame Diane Coyle

Bennett Professor of Public Policy,
Bennett Institute for Public Policy,
University of Cambridge

Professor Gina Neff

Executive Director, Minderoo Centre
for Technology and Democracy,
University of Cambridge

Professor Neil Lawrence

DeepMind Professor of Machine Learning,
University of Cambridge



EXECUTIVE SUMMARY

This policy brief aims to give the policy community an overview of the generative artificial intelligence (AI) field and highlight the key policy issues raised by its rapid development and adoption.

Our main findings and recommendations are as follows:

- Generative AI represents a significant technological advance, of comparable importance to the web, and offers a material opportunity for the United Kingdom (UK) to improve economic productivity
- The aspiration for the UK to become a global leader in the development of the foundation models that support generative AI products and services is unrealistic given the capital investment and compute capacity required
- The UK should focus on being a leader in applying foundation models in the real world, to change *how* things are produced, responsibly, safely, and fairly
- Expanding understanding of and access to generative AI tools throughout the economy and society is the most important way that the UK can build capacity in responsible AI implementation
- Innovation and skills policy levers can be applied to this challenge, including lobbying major cloud computing infrastructure providers to establish GPU-clusters in the UK, and introducing tax incentives for businesses to apply generative AI technologies to their existing operations
- There are potential legal, regulatory, cultural and societal impediments to the adoption of generative AI which need to be addressed, including uncertainty over the applicability of data protection, intellectual property, and product safety laws
- The sectoral approach to regulation based on value-based principles rather than enforceable legislation means there is a risk that regulators will lack the capacity to enforce their regulatory frameworks, or that sectoral regulatory frameworks will develop with contradictory and incoherent rules
- Currently, the UK's approach to regulating generative AI combines value-based sectoral regulation with efforts to shape international agreements. As a result, businesses lack incentives to comply with Responsible AI principles, with negative consequences for public trust in organisations' use of generative AI
- We believe this can be addressed through an AI Bill and sectoral legislation designed to embed an ethical framework for the governance of generative AI in domestic law, along with investment in strengthening regulatory capacity



1. INTRODUCTION

This policy brief aims to give the policy community an overview of the generative AI field.

It highlights the key policy issues raised by its rapid development and adoption ([section 2](#)).

We focus in particular on the questions of what is needed for the UK to unlock the productivity improvements promised by generative AI ([section 3](#)), and what impediments will need to be addressed to reconcile generative AI with emerging legal and ethical frameworks ([section 4](#)).

Finally, we make a set of recommendations for building capabilities to augment productivity through generative AI ([section 5](#)). Explanatory infographics, case studies, and a glossary of generative AI terminology (denoted by *italics*) are interspersed throughout.

We note that AI is a contested term. For the purposes of this brief, we assume a narrow definition of AI, taking it to mean computer systems which can improve themselves without explicit instructions, by making inferences from patterns in data.

The 'AI' we are concerned with is the kind that (among other things) organises social media newsfeeds, determines the sentiment of online comments, decides which adverts should be displayed on a webpage, classifies medical images, or recommends music, films, or books people might enjoy based on what they have previously consumed.

We recognise the debate about the potential for developments in AI research to engender machine superintelligence that poses an existential risk to humanity, but do not enter into it here – not least because *AI Safety* is already addressed extensively elsewhere as the focus of the [UK's Frontier AI Taskforce](#).¹

We likewise acknowledge important critiques which have drawn attention to the ways AI systems can reproduce bias and injustice, taking as a given that all AI should be [responsible AI](#).²

Glossary

AI Safety – efforts to pre-empt AI causing serious harm to humanity

Frontier AI – foundation models that are so advanced they pose serious risks to public safety

Responsible AI – the ethical practice of developing and deploying AI systems in a way that is fair, transparent, trustworthy, and accountable to society

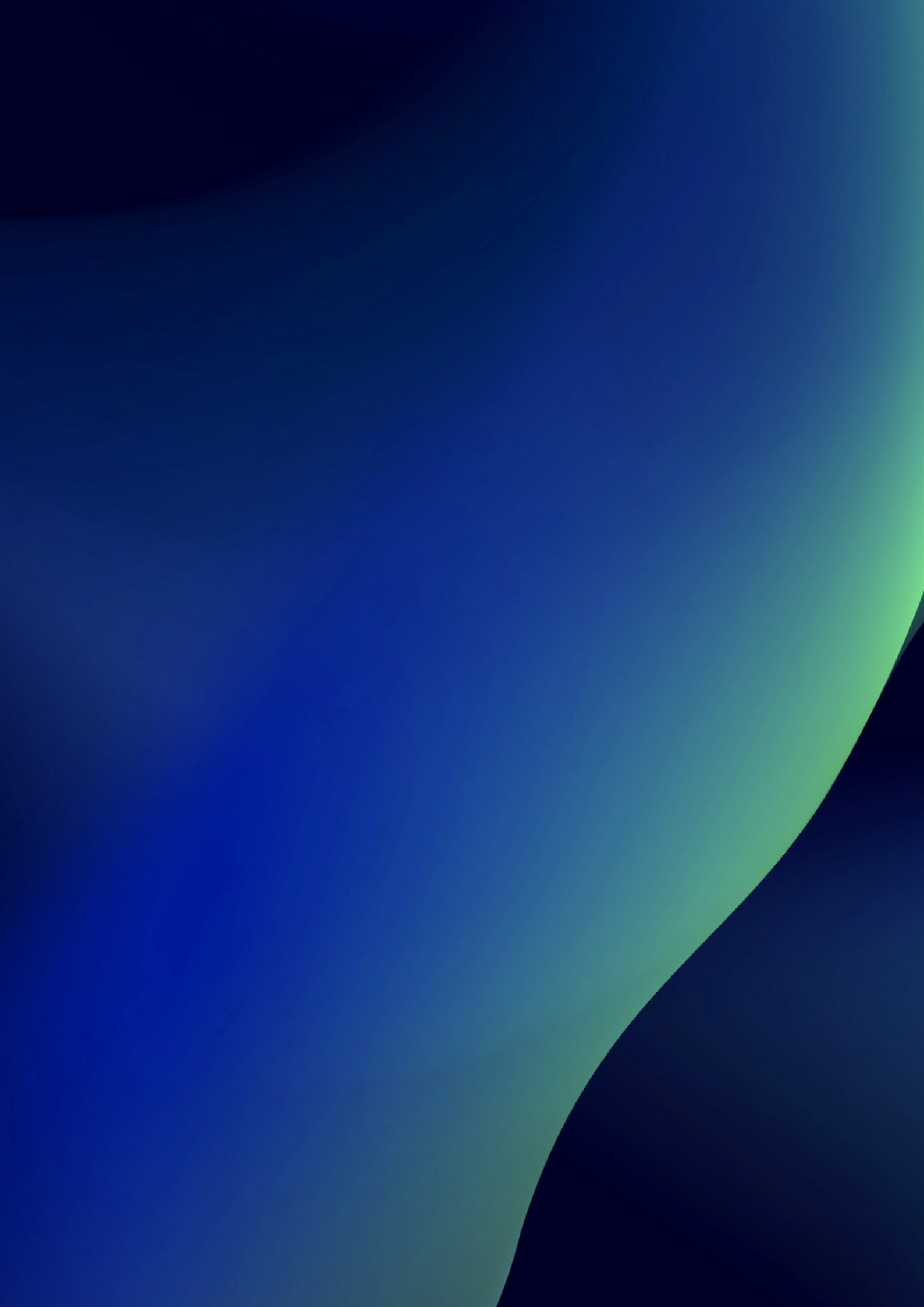
Large language model (LLM) – an AI model that can interpret, generate, and translate text

Prompt – the instructions a user gives to a generative AI model

Token – a unit of text or computer code, used by LLMs to interpret and generate text; can be a single character, part of a word, or a whole word

1. Department for Science, and Technology. 2023. "Tech Entrepreneur Ian Hogarth to Lead UK's AI Foundation Model Taskforce," Gov.uk <<https://www.gov.uk/government/news/tech-entrepreneur-ian-hogarth-to-lead-uks-ai-foundation-model-taskforce>> [accessed 21 September 2023]

2. "Responsible AI UK." [n.d.]. Responsible AI UK <<https://www.rai.ac.uk/>> [accessed 21 September 2023]



2. GENERATIVE AI

2.1 What is generative AI capable of and how does it work?

Generative AI involves running the kind of pattern-matching that machine learning systems do, only **in reverse**.³ Rather than looking at data and finding existing examples that fit a particular pattern, it draws on data to 'generate' new examples of that pattern. Generative AI systems can therefore output original high-quality text, images, audio, or video at mindboggling speed and scale.

Much of the excitement about generative AI has been driven by the runaway popularity of ChatGPT, a consumer facing app developed by OpenAI, which reached 100 million users **even faster than the app TikTok**.⁴

ChatGPT is underpinned by a type of generative AI system called a *large language model* (LLM). LLMs take instructions (or *prompts*) from users in natural language, and then output text in response—from stump speeches to Shakespearean sonnets and everything in between.

They work by predicting what word (or, strictly, *token*) ought to come next in a sequence, based on inferences from the vast corpus of data on which they have been *trained*, together with the user's instructions. While OpenAI's GPT-4 is the best-known LLM, there are many other examples (see **Figure 2**).

Although image-generation models like Midjourney use a different process called *diffusion*, from the perspective of the user they work in the same way as LLMs.

Natural language text prompts can yield **Van Gogh-inspired cover art for 'Stairway to Heaven', Pope Francis in a Balenciaga puffer jacket**, or more or less anything else than can be imagined and articulated.⁵

Glossary

Diffusion model – an image-generation model developed by corrupting a dataset of images with 'noise', then learning how to 'de-noise' the data and recover the images

Training – the process of teaching an AI system to interpret data

Prompt engineering – the practice of designing prompts with the objective of improving the quality of a generative AI model's output

Fine-tuning – a training technique used to customise a foundation model for a specific purpose

Plugin – a software add-on that enhances a system's capabilities. A number of **ChatGPT plugins** are available.

Foundation model – the generic name for LLMs, diffusion models and other general-purpose generative AI models which developers can use as the basis for more specialised apps

API – short for Application Programming Interface; a way of allowing different software applications to interact with each other

SaaS – short for Software-as-a-Service; software which is accessed over the web, rather than being installed locally

Compute – shorthand for the computational resources generative AI systems use to process data

3. Evans, Benedict. 2022. "ChatGPT and the Imagenet Moment," Benedict Evans <<https://www.ben-evans.com/benedictevans/2022/12/14/ChatGPT-imagenet>> [accessed 21 September 2023]

4. Tung, Liam. 2023. "ChatGPT Just Became the Fastest-Growing 'app' of All Time," ZDNET <<https://www.zdnet.com/article/chatgpt-just-became-the-fastest-growing-app-of-all-time/>> [accessed 21 September 2023]

5. [N.d.]. Prompthero.com <<https://prompthero.com/prompt/49b1d160343-midjourney-5-2-led-zeppelin-s-stairway-to-heaven-fine-detailed-pointillism-painting-low-angle-view-hyper-realistic-stairway-to-heaven-fantasy-vernacular>> [accessed 21 September 2023]; "Reddit - Dive into Anything." [n.d.]. Reddit.com <https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/> [accessed 21 September 2023]

2.2 Generative AI's limitations

At first sight these capabilities can seem miraculous, but it is important to be aware of their limitations. Diffusion models are not underpinned by an understanding of the physical world; **they don't 'know'** what text symbols mean, or that human hands usually have five fingers.⁶

The results can be comical, nightmarish, or simply wrong. Similarly, LLMs do not function like search engines, reliably retrieving information from a database.

Rather, LLMs generate new text probabilistically, meaning that they often invent facts and refer to seemingly plausible but non-existent academic studies and URLs (a phenomenon known as 'hallucination').

Overcoming these limitations requires a combination of *fine-tuning*, *prompt engineering*, and *plugins*.

2.3 Foundation models vs applications

Both LLMs and diffusion models are types of *foundation model*—a term describing models that others could 'build on top of' for many different purposes. This is enabled by giving third-party developers *API access*, allowing them to incorporate foundation model capabilities into their applications.

New startups have been able to develop software-as-a-service (*SaaS*) products that apply foundation models in specific contexts.

For example, **Harvey AI** uses OpenAI's GPT models in products designed to assist lawyers with research, contract drafting, and document review.⁷

Established tech companies have enhanced their products with generative AI features. For example, the graphic design platform Canva introduced a **text to image feature** powered by the DALL-E 2 model, and Microsoft added LLM-powered writing and editing features to its **Office 365 products**.⁸

2.4 The economics of generative AI

Providers of foundation models earn revenue by charging a small fee for each API request. As a result, their business model depends on the volume of API requests from applications being sufficient to offset the massive *compute costs* involved in developing and operating foundation models.

These costs are partly a function of the vast size of training datasets. For example, the text used to train OpenAI's GPT-3 model included a **45 terabyte archive of the web**, 11,000 books, and the entirety of Wikipedia.⁹

Processing such large quantities of data requires Graphics Processing Units (GPUs). A single GPU designed by market-leader Nvidia costs \$10,000; and thousands of GPUs are needed to train a single foundation model. Further compute costs accrue once models are released and begin processing prompts from users.

6. Mirjalili, Seyedali. 2023. "If AI Image Generators Are so Smart, Why Do They Struggle to Write and Count?," The Conversation <<http://theconversation.com/if-ai-image-generators-are-so-smart-why-do-they-struggle-to-write-and-count-208485>> [accessed 21 September 2023]

7. "Harvey." [n.d.]. Harvey.AI <<https://www.harvey.ai/>> [accessed 21 September 2023]

8. [N.d.-b]. Canva.com <<https://www.canva.com/apps/text-to-image/>> [accessed 21 September 2023]; Stallbaumer, Colette. 2023. "Introducing Microsoft 365 Copilot—A Whole New Way to Work," Microsoft 365 Blog <<https://www.microsoft.com/en-us/microsoft-365/blog/2023/03/16/introducing-microsoft-365-copilot-a-whole-new-way-to-work/>> [accessed 21 September 2023]

9. Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, and others. [n.d.]. "Language Models Are Few-Shot Learners," Arxiv.org <<http://arxiv.org/abs/2005.14165>> [accessed 21 September 2023]

Analysts estimate that ChatGPT costs \$40 million per month to run, and that Microsoft would need \$4 billion of compute if its GPT-powered Bing Chat product responded to all queries from Bing's users.¹⁰

A final nuance to note is that it is not necessary for foundation model developers to own GPUs themselves—they can rent GPU time from cloud providers as a service. Currently, most real-world end users of generative AI systems are paying nothing for the privilege, meaning foundation model providers' revenues are negligible – OpenAI projects just \$200 million for 2023.

Both development and usage of generative AI is therefore currently being funded by venture capital and the balance sheets of big tech companies – a situation which will clearly not last forever.

Glossary

GPUs – powerful chips originally developed to render 3D images in video games, now used for training foundation models

GPU Cluster – a group of computers containing GPUs

It seems likely that a small number of dominant foundation model providers will emerge and then increase prices to a level that produces attractive shareholder returns.

In the interim, the biggest beneficiaries are likely to be compute providers – Nvidia's share price, for example, is +150% year-on-year.¹¹ Economics are more benign for application developers, as their foundation model API costs rise and fall in proportion to usage of their products, and they can switch between different model providers easily.

	Examples	Value drivers	Cost drivers	Costs
Real-world uses	Business, public sector and individuals using generative AI-powered tools to fully or partly automate tasks and workflows	Productivity gains	Subscription fees paid to app & SaaS providers	£
Application layer	Consumer apps - e.g. ChatGPT SaaS tools - e.g. Harvey Product integrations - e.g. Microsoft 365 Copilot	Subscription fees from users; ad revenue	Transaction fees paid to model providers	££
Foundation model layer	Large language models (LLMs) - e.g. GPT-4 Diffusion models - e.g. Midjourney	Transaction fees from app developers	Compute costs; R&D	£££££
Infrastructure layer	"Compute" Cloud platforms - e.g. Azure AI Hardware - e.g. Nvidia GPUs	Hardware sales; service fees	Energy, materials; R&D	£££££

Figure 1: A simple schematic of the Generative AI 'stack'

10. Leswing, Kif. 2023. "Meet the \$10,000 Nvidia Chip Powering the Race for A.I.," CNBC <<https://www.cnbc.com/2023/02/23/nvidias-a100-is-the-10000-chip-powering-the-race-for-ai-.html>> [accessed 21 September 2023]

11. All market cap figures are correct as of 1 August 2023.

Company	Maturity	Generative AI Activities	Financial Position
US			
OpenAI	Late-stage growth	Develops foundation models for text, image, and audio-generation (e.g. GPT-4, DALL-E 2, Whisper); develops consumer apps (e.g. ChatGPT)	Valued at ~\$28bn in April 2023; has raised \$11.3bn in total ¹²
Meta	Public	Develops open-source LLMs (e.g. Llama2)	\$816bn market cap
Microsoft	Public	Provides compute as a service via the Azure AI platform; develops consumer applications (e.g. Bing Chat) and integrations (e.g. Microsoft 365 Copilot); major investor in OpenAI	\$2.50tn market cap
Nvidia	Public	Designs chips used in the training of foundation models; investor in Inflection AI and Synthesia	\$1.15tn market cap
Google	Public	Develops LLMs (e.g. PaLM2), consumer apps (e.g. Bard); integrates generative AI into existing products (e.g. Gmail); provides compute as a cloud service	\$1.68tn market cap
Anthropic	Growth-stage startup	Develops LLMs (e.g. Claude2)	Raised \$450m at ~\$4bn valuation in May 2023 ¹³
Inflection AI	Seed-stage startup	Develops LLMs and consumer apps (e.g. Pi)	Raised \$1.3bn at \$4bn valuation in June 2023 ¹⁴
Jasper	Growth-stage startup	Develops SaaS tools for copywriting, based on OpenAI LLMs	Raised \$125m at \$1.5bn valuation in October 2022 ¹⁵

Figure 2: Selected Generative AI Companies, United States (US) and UK

12. Singh, Jagmeet, and Ingrid Lunden. 2023. "OpenAI Closes 300M Share Sale at 27B-29B Valuation," TechCrunch <<https://techcrunch.com/2023/04/28/openai-funding-valuation-chatgpt/>> [accessed 21 September 2023]; "OpenAI." [n.d]. Crunchbase <https://www.crunchbase.com/organisation/openai/company_financials> [accessed 21 September 2023]

13. Wiggers, Kyle. 2023. "Anthropic Raises \$450M to Build Next-Gen AI Assistants," TechCrunch <<https://techcrunch.com/2023/05/23/anthropic-raises-350m-to-build-next-gen-ai-assistants/>> [accessed 21 September 2023]

14. "Microsoft-Backed AI Startup Inflection Raises \$1.3 Billion from Nvidia and Others." 2023. Reuters <<https://www.reuters.com/technology/inflection-ai-raises-13-bln-funding-microsoft-others-2023-06-29/>> [accessed 21 September 2023]

15. Rogenmoser, Dave. [n.d.]. "Jasper Announces 125M Series A Funding Round, Bringing Total Valuation to 1.5B and Launches New Browser Extension," Jasper.AI <<https://www.jasper.ai/blog/jasper-announces-125m-series-a-funding>> [accessed 21 September 2023]

Company	Maturity	Generative AI Activities	Financial Position
UK			
DeepMind	Acquired	Developing a robot command language (RT-2) and LLM (Gemini) ¹⁶	Acquired by Google for ~\$500m (2014) ¹⁷
Stability AI	Seed-stage startup	Develops image-generation models (Stable Diffusion) and LLMs (StableLM)	Raised \$101m at \$1bn valuation (2022) ¹⁸
Synthesia	Growth-stage startup	Develops SaaS tools enabling users to create corporate training videos with realistic digital avatars, based on proprietary models	Raised \$90m at ~\$1bn valuation in June 2023 ¹⁹
Arm	Public	Designs chips used in the training of foundation models; developing a platform to power generative AI apps (TCS23) ²⁰	IPO-ed at ~\$55bn in September 2023
Graphcore	Growth-stage startup	Designs chips used in the training of foundation models	Has raised \$680m, but venture capital investor Sequoia has written off its stake ²¹

Figure 2: Selected Generative AI Companies, United States (US) and UK (continued)

16. Sharma, Shubham. 2023. "DeepMind Unveils RT-2, a New AI That Makes Robots Smarter," VentureBeat <<https://venturebeat.com/ai/deepmind-unveils-rt-2-a-new-ai-that-makes-robots-smarter/>> [accessed 21 September 2023]; Knight, Will. 2023. "Google DeepMind's CEO Says Its next Algorithm Will Eclipse ChatGPT," Wired <<https://www.wired.com/story/google-deepmind-demis-hassabis-chatgpt/>> [accessed 21 September 2023]

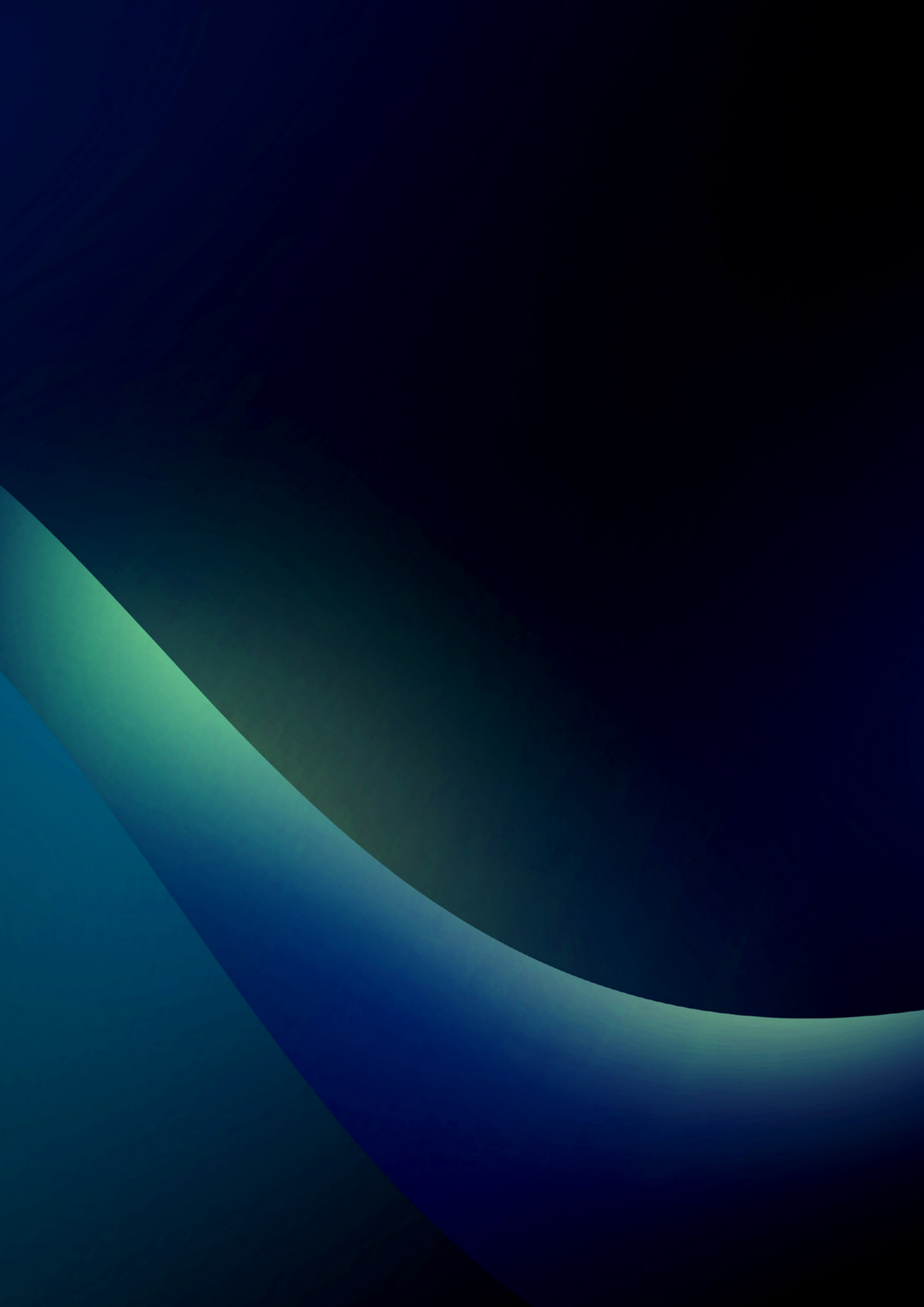
17. Shu, Catherine. 2014. "Google Acquires Artificial Intelligence Startup DeepMind for More than \$500M," TechCrunch <<https://techcrunch.com/2014/01/26/google-deepmind/>> [accessed 21 September 2023]

18. Wiggers, Kyle. 2022. "Stability AI, the Startup behind Stable Diffusion, Raises \$101M," TechCrunch <<https://techcrunch.com/2022/10/17/stability-ai-the-startup-behind-stable-diffusion-raises-101m/>> [accessed 21 September 2023]

19. Browne, Ryan. 2023. "Nvidia-Backed Platform That Turns Text into A.I.-Generated Avatars Boosts Valuation to \$1 Billion," CNBC <<https://www.cnbc.com/2023/06/13/ai-firm-synthesia-hits-1-billion-valuation-in-nvidia-backed-series-c.html>> [accessed 21 September 2023]

20. Arm Ltd. [n.d.]. "New Arm Total Compute Solutions Enable a Mobile Future Built on Arm," Arm | The Architecture for the Digital World <<https://www.arm.com/company/news/2023/05/new-arm-total-compute-solutions-enable-mobile-future-built-on-arm>> [accessed 21 September 2023]

21. "Blow for Tech Unicorn Graphcore as Sequoia Writes off Stake." [n.d.]. Times (London, England: 1788) (The Sunday Times) <<https://www.thetimes.co.uk/article/blow-for-tech-unicorn-graphcore-as-sequoia-writes-off-stake-jgrjrxsqw>> [accessed 21 September 2023]



3. PRODUCTIVITY AND GENERATIVE AI

We take the view that generative AI is a very significant technology, of comparable importance to the web.

However, it cannot be taken for granted that the adoption of generative AI will inevitably lead to whole-economy productivity growth—indeed, the digital innovations of the last 15 years have had no discernible impact on **measured UK productivity**.²²

It must also be acknowledged that there is still a lot of uncertainty about how generative AI will become economically useful. Google search data suggests the predominant use-cases for ChatGPT are currently **job applications and homework**, which have little relevance to the economy.²³

Meanwhile, most **capital investments in generative AI companies** to date have been at the foundation model and infrastructure layers; at the application layer, the majority of venture-backed companies are developing chatbots, virtual customer services assistants, writing tools, and features for video games.²⁴

While these may reduce operating costs in contact centres and increase copywriters' output and gamers'

play-time, they are unlikely to have a transformative economic impact.

If the UK is to benefit from generative AI, it needs to encourage direct application of the technologies to the productive economy, across multiple sectors.

STABILITY.AI: THE UK FOUNDATION MODEL LEADER?

Best known for the open-source image-generation model Stable Diffusion, Stability AI was founded in 2020 by former hedge fund manager Emad Mostaque. In 2022 the company raised \$101 million in a seed round led by Lightspeed Venture Partners and Coatue.

Stable Diffusion XL, released in July 2023, features the ability to generate words within images (see 'Generative AI's limitations'), and has been **favourably compared** by users to Midjourney and OpenAI's image models.

A number of controversies surrounding Stability AI should be noted. Recent months have brought a **lawsuit** from Getty Images, who claim that copyrighted material was included in Stable Diffusion's training data without permission, together with allegations from former partners and employees of **fraud**, financial irregularities and intellectual property **theft**.

22. [N.d.-c]. Parliament.uk <<https://commonslibrary.parliament.uk/research-briefings/sn02791/>> [accessed 21 September 2023]

23. Gilbert, Sam. 2023. "I Find That Homework Is Actually the #2 Application of ChatGPT (as Measured by US Google Search). There Is a Higher Volume of Searches Relating to Job Applications (Eg 'Chatgpt Resume', 'Chatgpt Cover Letter'), & a Comparable Volume for Code (per <https://t.co/VNSb0q30wG>) <https://t.co/Xu0dUnxWtQ> Pic.twitter.com/fcxqckaci3," Twitter <<https://twitter.com/samgilb/status/1673800626269048835>> [accessed 21 September 2023]

24. "No Title." [n.d.]. Dealroom.Co <<https://app.dealroom.co/lists/33530>> [accessed 21 September 2023]

3.1 Foundation model leadership

Policy discussion has focused on how the UK could become a **world leader in the development of novel commercial foundation models**.²⁵

We doubt that this is realistic, despite the UK benefitting from a world-leading research base in underpinning technologies. Training foundation models requires vast amounts of compute, and little compute capacity is available in the UK. The £900m supercomputer announced by the chancellor in March 2023 will not be online until 2026, and neither Amazon Web Services, Microsoft Azure, nor Google Cloud have **UK-located GPU clusters**.²⁶

Stability AI trains its foundation models on clusters in the US. However, the idea of sending sensitive data offshore is very unpalatable for all organisations concerned with privacy (including, say, the NHS), and such data transfers are not reconcilable with UK law.

A related barrier is the limited availability of investment capital to fund compute. Modest government **support for the UK chip industry**—which has strategic importance well beyond generative AI—speaks to constraints on state spending relative to China and the US.²⁷

Unlike the US, the UK has no big tech companies with balance sheets large enough to invest meaningfully in foundation model developers, and the UK venture capital market is far smaller (**\$31bn** vs **\$235bn** in 2022).²⁸

Traditional startup funding models where companies raise seed capital (~£1m) to develop a minimum viable product, followed by larger and larger amounts of investment once they have gained traction with customers will not work at the scale needed for foundation model development.

OTHER UK GENERATIVE AI STARTUPS

Criteria: Raised >\$10 million; HQ in the UK

PolyAI – develops voice assistants for enterprise clients that can handle tasks like hotel room bookings, food orders and insurance claims

Papercup – develops software that dubs existing video content into different languages

Lifescore – platform generating endlessly varying music based on original compositions and recordings

UnlikelyAI – in stealth mode; founder previously contributed to development of Amazon’s virtual assistant Alexa

Instadeep – machine-learning platform provider, acquired by BioNTech for £562 million in July 2023. Does not describe itself as a generative AI company

25. Department for Science, and Technology. 2023a. “Initial £100 Million for Expert Taskforce to Help UK Build and Adopt next Generation of Safe AI,” Gov.uk <<https://www.gov.uk/government/news/initial-100-million-for-expert-taskforce-to-help-uk-build-and-adopt-next-generation-of-safe-ai>> [accessed 21 September 2023]

26. The Economist. 2023. “How to Make Britain’s AI Dreams Reality,” Economist (London, England: 1843) (The Economist) <<https://www.economist.com/britain/2023/06/14/how-to-make-britains-ai-dreams-reality>> [accessed 21 September 2023]

27. Bradshaw, Tim, and Anna Gross. 2023. “UK Government Unveils Long-Awaited £1bn Semiconductor Strategy,” Financial Times <<https://www.ft.com/content/757cfa86-adeb-4d8e-ad71-034c9a4d2f7d>> [accessed 21 September 2023]

28. [N.d.-d]. Dealroom.Co <<https://dealroom.co/guides/united-kingdom>> [accessed 21 September 2023]; [N.d.-d]. Dealroom.Co <<https://dealroom.co/guides/usa>> [accessed 21 September 2023]

The upfront capital requirements to develop foundation models are of a different order of magnitude, making them unsuitable for UK-style startup investing.

The foundation model layer is also not the most economically attractive part of the generative AI 'stack'. Most models have been trained on the same openly-available data, rather than proprietary sources, meaning there is limited scope for competitive differentiation and defensible market leadership.

It is at least plausible that competition between the likes of OpenAI, Google, Anthropic, and Inflection will drive down prices, leading to foundation models becoming increasingly commoditised. Meta's **open-sourcing of Llama 2** means that a **powerful** LLM is now available for commercial use, without the upfront capital costs associated with building these models, undermining the business model of the closed-source foundation model developers.²⁹

There remain, however, significant compute costs associated with their use. There are also indications that the **performance of open source models** is progressing at pace.³⁰

Given these market conditions, it is unclear how foundation model leadership would contribute to economic productivity, even if it could be attained.

GLOSSARY

No-code – an approach to software development which uses intuitive drag-and-drop interfaces to allow people without programming skills to build applications

Web framework – a set of tools and resources designed to make it easier to build web applications

3.2 Real-world applications of foundation models

Rather than building publicly or privately funded competitors to the likes of OpenAI and Google, ***we see greater opportunity for the UK in becoming a leader in how foundation model are applied in the real world.***

With smaller funding requirements, application layer products which customise foundation model capabilities to specific use-cases are a better fit for the UK venture capital market, and can build on existing strengths in sectors like fintech, healthtech and cybersecurity.

A further opportunity could be leveraging the UK's research capabilities to drive progress in underpinning technologies and to develop products which address specific major challenges at the foundation model and infrastructure layers of generative AI, such as the **detection of AI-generated content** and **cooling of data centres**, as well as AI safety solutions.³¹

29. Facebook company. 2023. "Meta and Microsoft Introduce the next Generation of Llama," Meta <<https://about.fb.com/news/2023/07/llama-2/>> [accessed 22 September 2023]; "Alpaca Eval Leaderboard." [n.d.]. Github.io <https://tatsu-lab.github.io/alpaca_eval/> [accessed 22 September 2023]

30. "AI Foundation Models: Initial Report." 2023. Gov.uk <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1185508/Full_report_.pdf> [accessed 22 September 2023]

31. Cooke, Elizabeth. 2023. "AI Model Collapse Could Spell Disaster for AI Development, Say New Studies," Verdict <<https://www.verdict.co.uk/ai-model-collapse-could-spell-disaster-for-ai-development-say-new-studies/>> [accessed 22 September 2023]; [N.d.-e]. Datacentremagazine.com <<https://datacentremagazine.com/articles/the-liquid-cooled-future-of-high-performance-compute>> [accessed 22 September 2023]

CASE STUDY: SOFTWARE DEVELOPMENT WITH GENERATIVE AI

Ankur Shah is a London-based technology entrepreneur, whose previous exits include footwear brand Mahabis and adtech platform Techlightenment.

"I trained as a barrister and my coding skills are adequate for simple proofs of concept, but I've always relied on outsourced developers when building new projects, which is time-consuming and costly. But in the last 12 months generative AI has changed everything. For people like me who want to build websites, apps, and workflow automations it's akin to a superpower.

"One simple example I really like is **Meoweler** – a light-hearted travel site, ostensibly for cats. It's beautifully executed and provides a nice snapshot of thousands of cities around the world. But what's significant is that it **cost only \$140** to build, and the guy who made it is a designer with no formal training in software development.

"He found a freely available database of cities, then wrote GPT and Midjourney prompts to generate the content and images for each city in a consistent format and style. Then he used the Svelte web framework to create URLs, page components, and site search. It's a similar approach to the one we've

taken to programmatically **reviewing insurance products***, albeit we use a different web framework and are more focused on data quality.

"Features that used to take months and cost tens of thousands of pounds, I can now build myself in an afternoon with ChatGPT. It's insane.

"But sites and apps only scratch the surface. What's exciting me at the moment is systems that use LLM capabilities recursively. I love the idea of 'teams' of AI agents that can take a request like 'get me some quotes to have a heat-pump installed' and then automate the whole series of linked tasks needed to fulfil it – background web research, shortlisting and prioritising suppliers, contacting them for quotes, and so on. My intuition is, it will be scrappy, bedroom-hacker types – not computer science graduates or corporate IT departments – who get there first."

** Author disclosure: Sam Gilbert is involved with this project.*

Generative AI's primary contribution to productivity will be in *changing how things are produced*.³² The biggest benefits to productivity will not come from a small number of technologically-sophisticated companies using generative AI to invent new products, or cut their costs.

Rather, generative AI's promise lies with changing production itself, just as occurred with interchangeable parts (19th century), assembly lines (1910s), just-in-time production (1980s), and globalised supply chains (2000s). The best example is software.

32. Coyle, Diane. 2023. "The Promise and Peril of Generative AI," Social Europe (SE) <<https://www.socialeurope.eu/the-promise-and-peril-of-generative-ai>> [accessed 22 September 2023]

The **code interpreter plugin** for ChatGPT and LLM-powered tools like GitHub Copilot are already enabling developers to **write code up to 55% faster** than before, presenting a potential solution to the UK's chronic **developer labour shortage**.³³

Even more significant is how generative AI expands the scope of no-code, enabling people without programming knowledge to build increasingly sophisticated software applications. In the past, which systems and automations could be developed was constrained by the availability of workers with skills in programming languages. Generative AI tools effectively remove this constraint for some types of development, meaning that the capability to imagine what a system might do and to articulate how it ought to function becomes more valuable than formal computer science training—a paradigm brought to life by the case study on the previous page.

When it comes to productivity, in our view ***the most important national capability is a means of widely disseminating understanding of and access to generative AI tools through the economy and society.***

There is good evidence that only a minority of firms are adopting existing digital tools in ways that enhance their

productivity and commercial success, pulling ever further ahead of the pack. The gap could grow with the powerful new capabilities afforded by generative AI. The national economic challenge is to spread know-how among businesses and employees. There is a role for government and AI experts to encourage learning about the potential of generative AI, not only through sharing techniques and examples but also through the range of business support tools available.

In some ways this runs counter to prevailing trends: many organisations have **banned their employees** from using generative AI applications, reasonably fearing it could lead to data leaks and/or the loss of intellectual property.³⁴

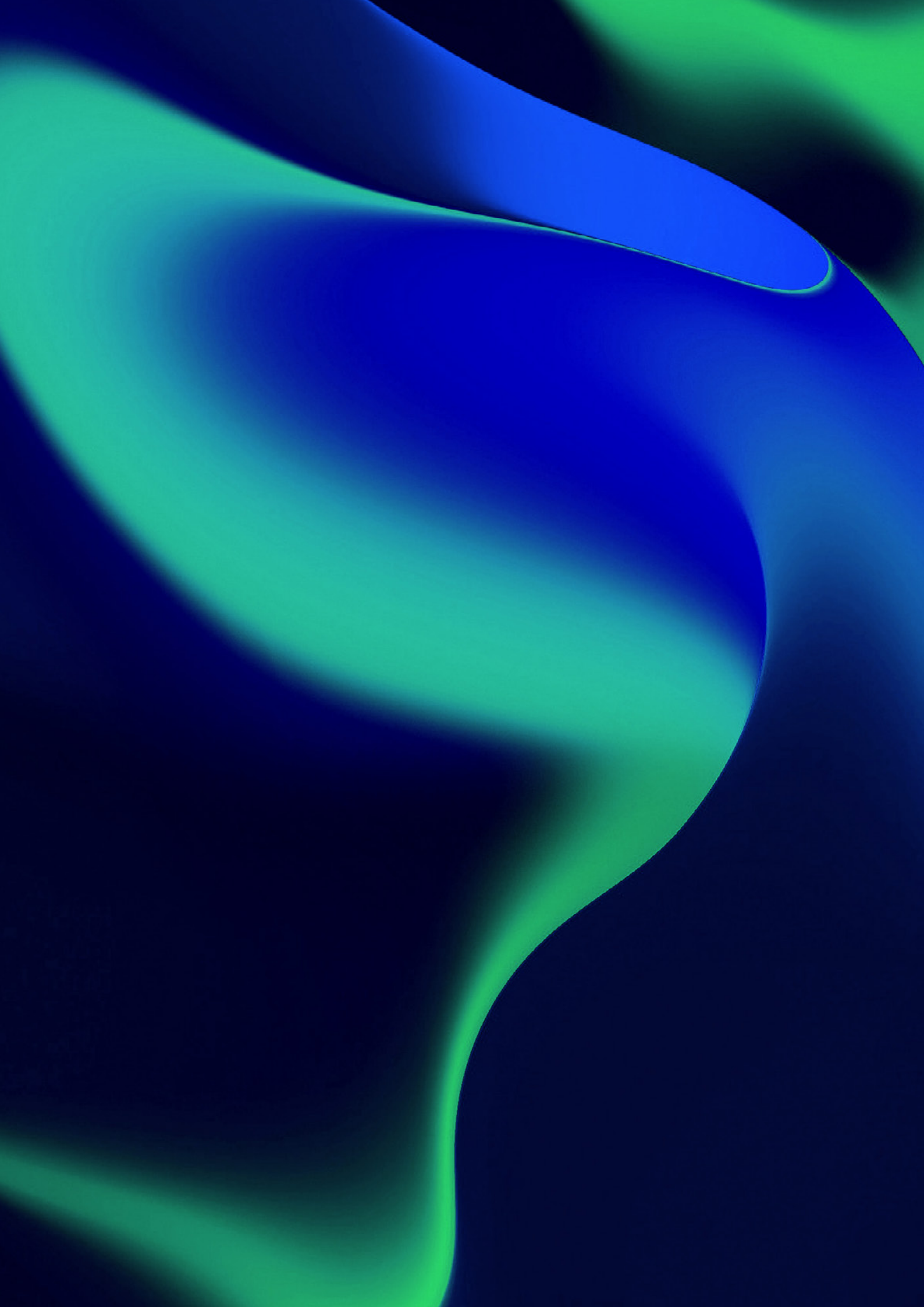
While understandable, such practices inhibit the bottom-up emergence of productivity opportunities inside organisations. There is also some anecdotal evidence that productivity gains from generative AI are already being realised, but lost to forms of arbitrage.

Remote workers **secretly use ChatGPT** to get more free time or impress their superiors; marketing agencies outsource content writing to LLMs while leaving their client fees and service-level agreements unchanged.³⁵ ***Incentives must be created for expert users of generative AI tools to share their techniques.***

33. "ChatGPT Plugins." [n.d.]. Openai.com <<https://openai.com/blog/chatgpt-plugins>> [accessed 21 September 2023]; Dohmke, Thomas. 2023. "GitHub Copilot for Business Is Now Available," The GitHub Blog <<https://github.blog/2023-02-14-github-copilot-for-business-is-now-available/>> [accessed 22 September 2023]; McDonald, Clare. 2022. "Around 750 New Software Developer Jobs Advertised Every Day," Computerweekly.com <<https://www.computerweekly.com/news/252523586/Around-750-new-software-developer-jobs-advertised-every-day>> [accessed 22 September 2023]

34. Milmo, Cahal. 2023. "ChatGPT Limited by Amazon and Other Companies as Workers Paste Confidential Data into AI Chatbot," INews <<https://inews.co.uk/news/technology/chatgpt-limited-amazon-companies-workers-paste-confidential-data-ai-chatbot-2254091>> [accessed 22 September 2023]

35. Ito, Aki. 2023. "Employees Are Secretly Using ChatGPT to Get Ahead at Work," Business Insider <<https://www.businessinsider.com/chatgpt-secret-productivity-work-ai-technology-ban-employees-coworkers-job-2023-8>> [accessed 22 September 2023]



4. IMPEDIMENTS TO DEVELOPING THE UK'S NATIONAL CAPABILITIES IN GENERATIVE AI

Several impediments may hamper efforts to unlock the full potential of the UK's capabilities for generative AI.

There are economic impediments in terms of lack of investment, and impediments from the challenges of scale of the technical infrastructure, as explained in previous sections. While the UK has not adopted specific legislation to regulate generative AI, there are some restrictions in existing laws, notably concerning personal data protection and intellectual property.

Further, there is an impediment to the uptake of national capabilities in generative AI in that these technologies are considered unethical and untrustworthy by some.³⁶ Thus, national capability will depend on generative AI tools which are reliable, safe, responsible, and trustworthy.

We have identified how generative AI can unlock the UK's potential for

augmented productivity by changing the ways things are produced. However, there are several impediments to UK businesses' access to and use of generative AI.

Figure 3 sets out some of the chief legal, regulatory, economic, cultural, and societal impediments to the adoption of generative AI in the UK. This section gives an overview of impediments to the uptake of generative AI in the UK. It specifically addresses risks associated with generative AI and what is meant by ethical and responsible AI.

The section also addresses the concerns regarding, personal data, privacy, and data governance, particularly in relation to copyright, that arise from the development and use of generative AI tools.

36. 'Trustworthy AI' is a contested term. The European Commission's Independent High-Level Expert Group on Artificial Intelligence identifies three components of Trustworthy AI: (1) it should be lawful, complying with all applicable laws and regulations; (2) it should be ethical, ensuring adherence to ethical principles and values; and (3) it should be robust, both from a technical and social perspective.

Impediment	Explanation
<p>Legal & regulatory (Law & regulation can both be a facilitator and an impediment if too restrictive; the absence of law & regulation can also be an impediment)</p>	<p>There is currently no omnibus Bill in Parliament dedicated solely to regulating AI in the UK. This is in contrast to other jurisdictions, especially the European Union (EU), where stringent legislation is being adopted. While the UK is considering regulating aspects of AI, notably through the Online Safety Bill and particularly through regulators, no legislative initiative specifically addresses generative AI or foundation models—as illustrated by the examples below.</p> <p>The UK Government has taken steps to regulate AI through a 'pro-innovation approach' by which the Government wants to use regulators to encourage business to adopt five ethical principles when using generative AI. The five principles are modelled on the Organisation for Economic Co-operation and Development (OECD)'s principles for the regulation of AI. The UK has also signed up through its membership of the UNESCO recommendations on the Ethics of AI.</p> <p>While there is no specific legislation for generative AI in the UK, the use of these technologies must still conform to existing law, such as the Data Protection Act 2018 or intellectual property laws. The UK's Intellectual Property Office is currently working on a draft code for copyright and AI that will address inter alia the contentious issue of text and database mining exceptions, which the Government had earlier proposed for the development of AI models and tools. <i>There is a chance that generative AI will be regulated through the regulatory framework being developed by the Competition & Markets Authority based on the regulator's new statutory powers in the Digital Markets, Competition and Consumers Bill (DMCC), which is expected to enter into force in the second half of 2024.</i></p> <p>A recent report from the Competition & Markets Authority suggested a collection of principles to guide regulatory intervention in support of competitive generative AI markets, built on ready access to the materials to create foundation models, diversity of business models, choice for businesses in how to use foundation models and flexibility for consumers in which provides to engage, the prevention of anti-competitive practices, and transparency about the risks and limitations of the foundation model products they are using.</p> <p>A cluster of policy initiatives also seek to set guardrails for AI development, with different levels of implementation. For example, while the UK has set out a data sharing governance framework as part of its national data strategy, it has not adopted specific legislation to give effect to the framework in the private sector. In contrast, the EU is adopting the Data Act and the Data Governance Act.</p> <p>In contrast to the UK, the EU is adopting the AI Act, expected to come into force at the end of 2025. The AI Act will regulate AI according to perceived risks: unacceptable risk (banned), high risk (transparency, oversight, and accountability requirements), and low-to-minimal risk (safety and user protection requirements). Canada has taken a similar approach with its Artificial Intelligence and Data Act.</p> <p>The EU is also considering specific AI product safety liability rules for how products are manufactured and how they should be used. See for example the European Commission's proposal for an AI Liability Directive, or the work of the European Centre for Algorithmic Transparency (ECAT).</p> <p>While there are numerous initiatives to introduce legislation to regulate AI in the US, the US so far has encouraged voluntary self-regulation based on the White House's Blueprint for an AI Bill of Rights, setting out five principles: (1) safe and effective systems, (2) algorithmic discrimination protection, (3) data privacy, (4) notice and explanation, and (5) human alternatives, consideration and fallback. The White House has also published a set of eight voluntary commitments pledged by leading companies in the AI industry. <i>In addition, in August 2023 it was announced that the Biden Administration is fast-tracking an Executive Order</i> to address risks associated with AI.</p>

Figure 3: Impediments to the uptake of generative AI in the UK

Impediment	Explanation
<p>Legal & regulatory (Law & regulation can both be a facilitator and an impediment if too restrictive; the absence of law & regulation can also be an impediment) <i>(continued)</i></p>	<p>There are also legal and regulatory initiatives on State-level, exemplified by the Governor of California's recent Executive Order N-12-23 on generative AI, and domain-specific guidance on the application of existing legislative, for example from blog posts by the Federal Trade Commission on consumer protections.</p> <p>The US National Institute of Standards and Technology (NIST) has also developed a voluntary AI Risk Management Framework (AIRMF) and Senator Chuck Schumer has developed a SAFE Innovation Framework for the regulation of AI. In addition, the US Consumer Product Safety Commission published a report on product safety and liability on AI in 2021.</p> <p>Numerous issues arise in relation to product liability and AI, including whether existing laws cover the systemic risk of harm or if a precautionary principle approach should be adopted, to whom liability should be assigned, and the resources and levers available to the regulators. The Department for Business and Trade and the Office for Product Safety and Standards are currently reviewing the UK's product safety regime post-Brexit, which offers an opportunity to also consider the need for national generative AI product safety standards.</p> <p>The Trade Union Congress (TUC) and the Minderoo Centre for Technology and Democracy at the University of Cambridge have set up a taskforce to draft a legislative proposal for the protection of workers and the use of AI. The taskforce will particularly examine risks associated with privacy, insecurity of work, and discrimination from the deployment of AI.</p>
<p>Technical</p>	<p>There are technical limitations to the capabilities generative AI can provide business, particularly when it comes to responsible, transparent, and trustworthy AI. The effectiveness of tools for auditing for bias or delivering required levels of explainability continue to be limited. AI 'hallucinations', where a generative AI tool makes up information is a weighty concern about the reliability of these technologies.</p>
<p>Economic</p>	<p>Lack of investment as outlined in earlier sections.</p>
<p>Cultural</p>	<p>Business may be reticent to make use of generative AI whilst employees may be using these technologies 'under the radar', without the quality or legal assurance, which poses a risk to competitiveness and regulation. There may also be reticence within the labour force to deploy generative AI either for fear that these technologies are not trustworthy, or that they will replace workers, thereby taking away the user's job.</p> <p>A variety of organisational processes or cultural factors will also influence patterns of AI adoption, from internal data management, and an executive understanding of the potential of AI, to employer-employee relations. Organisational AI readiness will be an important influence on overall patterns of adoption.</p>
<p>Societal</p>	<p>There are numerous concerns regarding the ethics of generative AI, which lead to questions of fairness, trustworthiness, transparency, and accountability. Without a robust and accountable ethics framework, the public will not trust the use of generative AI. There is also a risk that without a sound compulsory ethical framework, generative AI will perpetuate and advance biases and inequalities within the population, thereby contributing to greater systemic unfairness.</p>

Figure 3: Impediments to the uptake of generative AI in the UK (continued)

In its *interim report* published on July 2023, the House of Commons Science, Innovation and Technology Committee summarised the barriers to implementing safe and effective AI as 12 AI challenges:

1. The Bias challenge
2. The Privacy challenge
3. The Misrepresentation challenge
4. The Access to Data challenge
5. The Access to Compute challenge
6. The Black Box challenge
7. The Open Source challenge
8. The Intellectual Property and Copyright challenge
9. The Liability challenge
10. The Employment challenge
11. The International Coordination challenge
12. The Existential challenge

While not negating the importance of AI safety, this policy brief narrowly focuses on how to build UK's capabilities for productivity using generative AI. We therefore only consider risks that pose impediments to that goal. There are three chief impediments to building the UK's capabilities in this regard.

First, there is the risk that a lack of trust in generative AI becomes so pervasive that the deployment of these technologies is rejected by businesses and the public. Second, there is the risk that generative AI will be subjected to legal and ethical regimes which will be overly restrictive and thus hamper its full potential.

Third, the issue with AI hallucinations, whereby the generative AI tool makes up information, alongside other technical limitations, poses a challenge to their reliability which again is an impediment to their uptake nationally.

4.1 Risks with generative AI

This section briefly examines risks associated with generative AI and the legal and ethical frameworks that are emerging to address these. Fundamentally, the British public must be able to trust the use of generative AI. There are many conceptualisations of risks associated with generative AI.

The list below is not intended to be read as a complete overview, but rather a list of some of the most prominent concerns related to AI. Numerous risks are associated with generative AI, including risks to personal data, privacy, and intellectual property. There are risks that due to lack of transparency or accountability, generative AI may produce unreliable outcomes, or be used for hidden or unacceptable outcomes.

Key concerns with generative AI applications are the reliability or veracity of the outputs, especially as the capacity of non-technical users to produce deepfake images, audio, and video abounds.

Scholars have also identified risks of negative environmental consequences, the overrepresentation of hegemonic viewpoints and value-lock in training data, the risk of propagating toxic stereotypes and racist, sexist, and ableist ideologies, marginalising communities, violating personal data, and subjecting people to abusive language, hate speech, micro-aggressions, derogating language, dehumanising and denigrating content and framing, which could lead to psychological harm.³⁷

37. Emily M. Bender et al., On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?, *FaccT'* 21, March 3-10, ACM <https://dl.acm.org/doi/10.1145/3442188.3445922>

There are risks that data scraping for training foundation models violates copyright laws, or that the foundation models will reproduce bias, which may produce illegal outcomes, especially when generative AI is used in the context of social services, policing, and education. The cumulative effect of these risks is the erosion of trust in the technology, and of societal trust overall.

According to the Ada Lovelace Institute: "It is also unlikely that international agreements will be effective in making AI safer and preventing harm, unless they are underpinned by robust domestic regulatory frameworks that can shape corporate incentives and developer behaviour in particular." (Ada Lovelace Institute, [Regulating AI in the UK](#), p. 5).

4.2 Ethical and responsible generative AI

Responsible AI means demonstrating how the ethical principles are adhered to throughout all the stages of the generative AI lifecycle.³⁸ To do so, there must be appropriate accountability, risk mitigation, and liability.³⁹

In terms of building national capabilities for the workforce, there are particular concerns regarding automated decision-making and the role of humans in the loop. Numerous voices have expressed concern that generative AI is not responsible or ethical.

OECD's value-based principles for AI:

- *Inclusive growth, sustainable development, and well-being*
- *Human-based values and fairness*
- *Transparency and explainability*
- *Robustness, security, and safety*
- *Accountability*

To meet these concerns about the use of AI more broadly, the Government has proposed a guiding **principle-based framework**. The principles are drawn from the work of the **OECD** and as such build on the emerging international consensus for ethical and responsible AI. This principle-based approach is dependent on regulatory capacity to be effective.

The UK's government's value-based principles are:

- Safety, security and robustness
- Appropriate transparency and explainability
- Fairness
- Accountability and governance
- Contestability and redress

The principles are designed to be future-oriented and flexible, with the intention of promoting growth and innovation.

While not legally binding, the Government envisions that sector-specific regulators will adopt the principles as fit to their sectors and industries.

38. AI ethics is a growing academic field with numerous different interpretations of the term. Some relevant scholarly articles are Robert Ganna and Emre Kazim, Philosophical foundations for digital ethics and AI ethics: a dignitarian approach, *AI and Ethics* (2021) 1: 405-423; Samuele Io Piano, Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward, *Humanities and Social Sciences Communications* (2020) 7-9; Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Christopher Nagy, Madhulika Sri Kumar, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI, Berkman Klein Center For Internet & Society at Harvard University, 2020, Ibo van de Poel, Embedding Values in Artificial Intelligence (AI) Systems, *Minds and Machine* 2020 30:385-409.

39. Australian Government Department of Industry, Science and Resources, Safe and responsible AI in Australia (Discussion paper June 2023), pp. 8-9.

However, this approach may pose challenges in ensuring that regulators have the incentives, resources, or mandate to do so, especially as many regulators' remits are constrained by statutory language. Thus, the approach has been challenged by [leading academics](#), pointing to the need for more holistic thinking.

It is also problematic that the Government's principles are so vague as to be nearly vacuous.⁴⁰ It is, for example, difficult to discern with any certainty whether the principles are focused on outcomes or how those outcomes are to be achieved.

However, elsewhere, for example in data protection, the Government has suggested that regulation should be based on outcomes; an approach that could potentially be taken for the five value-based principles as well. (Department for Digital, Culture, Media & Sport, [Data: A new direction](#), 10 September 2021, updated 23 June 2023, p. 7.)

As the principles are not legally binding, it is unlikely that businesses will have an adequate incentive to adopt all the principles unless there are compelling competitive advantages to doing so. While the Government has provided tools such as the Algorithmic Transparency Recording Standard, which aims to support the implementation of ethical AI principles, the extent to which such tools are being implemented in practice is not clear.

Thus, the Ada Lovelace Institute has noted that: "The principles will not – initially – be placed on a statutory footing, and so regulators will have no legal obligation to take them into account, although the Government has said it will consider introducing a 'duty

to have regard' to the principles." (Ada Lovelace, *Regulating AI in the UK*, p. 16.)

The [House of Commons Science, Innovation and Technology Committee](#) has criticised the Government's unwillingness to consider AI-specific legislation, noting that: "[t]here is a growing imperative to ensure governance and regulatory frameworks are not left irretrievably behind the pace of technological innovation." ([The governance of artificial intelligence; interim report](#), Ninth Report of Session 2022-23, p. 3.).

Thus, rather than see legislation as an impediment to the development of the UK's competitiveness in generative AI, we echo the sentiment of the review of the digital technologies, led by Sir Patrick Vallance, that: "Well-designed regulation and standards can have a powerful effect on driving growth and shaping a thriving digital economy." (HM Government, [Pro-innovation Regulation of Technologies Review: Digital Technologies](#) (March 2023), p. 3.)

While calls for legislation are mounting, it does not mean that the content of AI legislation is self-evident. Legal rules that are too specific risk being quickly outdated while principles that are too broad or vague risks being meaningless.

The challenge is therefore how to find the regulatory approach that will be robust and future-proof. Legislation would also clarify the chain of liability throughout the value-chain and lifecycle of generative AI. For example, the All-Party Parliamentary Group on Data Analytics (APGDA) has noted that: "there are issues around transparency, explainability, and accountability in relation to third party/outsourced AI system development.

40. The Central Digital and Data Office expounded on these principles in its [Data Ethics Framework](#) for the use of digital technologies, including AI, in 2020. However, this guide is only for the public sector and does not have the force of law.

For example, attention was drawn to the difficulty of testing for bias in third party systems." (Policy Connect, [An Ethical AI Future: Guardrails & catalysts to make artificial intelligence a force for good](#), 19 June 2023, p. 10.)

Legislation could clarify the standards and responsibility of testing that would befall UK businesses using third-party generative AI systems. A key issue with applying the law or ethical principles to generative AI is that the outcome is personalised or bespoke, therefore making predictability or comparison difficult. "Generated content is probabilistically and randomly generated based on certain input (or 'prompts'), which are usually written by a human.

"Therefore, the output of any given generative AI model is likely to be different for each person prompting the model and may both resemble patterns in the training data or appear to be something completely new." (Forbrukerrådet, [Ghost in the Machine: Addressing the consumer harms of generative AI](#), June 2023, p. 8.) A [recent review](#) of 10 foundation models found that none met the compliance requirements set out in the EU's draft AI Act.

There is a question of whether generative AI should go through an approval or vetting process before being used, or if redress and contestability should be used as a deterrent for unacceptable practices. Accountable principles also means that there must be ways to audit the generative AI systems, which will require access to data for researchers and for regulators.

A right to access to data for researchers in relation to the processing of personal data has been proposed included in the [Online Safety Bill](#) in relation to the online

information environment, but this has yet to be adopted by Parliament. Further there are no legal stipulations for data access in the legislative pipeline with regard to generative AI in the UK.

4.3 Personal data and privacy concerns

Many of the ethical concerns regarding generative AI are linked to the use of personal data and privacy. These concerns span personal data that is being inputted into generative AI systems, personal data generated by these systems, and uses of generative AI systems for surveillance.

Some of these fears should be allayed with the [Data Protection Act 2018 \(DPA\)](#) and its forthcoming replacement, the [Data Protection and Digital Information \(No. 2\) Bill](#). The UK's data protection framework is based on the EU's [General Data Protection Regulation \(GDPR\)](#), which includes the stipulation that all processing of personal data must adhere to the data processing principles.

The data processing principles are: (1) lawfulness, fairness, and transparency; (2) purpose limitation; (3) data minimisation; (4) accuracy; (5) storage limitation; (6) integrity and confidentiality; and (7) accountability. That means that all uses of personal data by generative AI must respect these principles as a matter of law.

As the legislation covers all forms of personal data its remit is broader than processing that concerns privacy. The legal definition of personal data is technologically neutral and comprehensive to ensure that all forms of personal data falls under its scope.

The broadness of the definition has implications for the use of generative AI and may pose a considerable impediment from the uptake of these technologies by UK companies. The Court of Justice of the European Union (CJEU) has for example ruled that an **IP address can be personal data** when combined with other factors held by third parties.

Regardless of where in the process personal data is generated or the sources from which it is harvested, including public domain sources, or provided directly (and voluntarily) by an individual, all the data processing principles still apply in full. There are further restrictions on the use of sensitive data, which poses challenges for companies using generative AI as sensitivity may first become apparent once the system has generated output data.

The use of generative AI poses several challenges when it comes to compliance with data protection law. It may not be apparent whether data is personal or not, or the system may generate personal data unbeknown or unintended by the creator of the AI system. However, it must be noted that the DPA is not a privacy statute, and that the objective of the legislation is not to preclude the processing of personal data, but instead to ensure that the processing is legal. Thus, the DPA does not automatically prevent the generation and use of personal data in generative AI.⁴¹

Article 3 of the DPA defines personal data as:

"...any information relating to an identified or identifiable living individual... [meaning] a living individual who can be identified, directly or indirectly, in particular by reference to: (a) an identifiers such as a name, an identification number, location data or an online identifier, or (b) one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of the individual."

The UK Government clearly recognises the role personal data has in innovation and AI. In its **White Paper on data**, the Department for Digital, Culture, Media & Sport writes that: "Innovative uses of personal data are at the forefront of driving scientific discovery, and enabling cutting-edge technology, like artificial intelligence (AI)...This means maintaining a clear legal framework overseen by a regulator that takes account of the benefits of data use, while protecting against the harms that can come from using personal data irresponsibly." (p. 6).

The objective of the White Paper on data is to use "personal data responsibly" (p. 6.), which necessitates an ethical framework. One of the drawbacks in the context of generative AI of the current data protection regime is that it is focussed on the right to data protection of the individuals and as such does not address the potential for systemic risk of bias, discrimination, and inequality arising from the use of personal data at scale.

41. However, several European data protection authorities are examining whether generative AI tools comply with the GDPR. Notably, the Italian data protection authority has placed a **temporary ban** on an OpenAI generative chatbot for failing to provide information as required under the GDPR. The Spanish Data Protection Authority is also **investigating ChatGPT** for breaches of GDPR. Furthermore, the European Data Protection Board (EDPB) has set up a **taskforce** to examine whether generative AI is compatible with the GDPR. It must, however, also be noted that these concerns regard whether generative AI comport with the data-processing principles for the processing of personal data, not whether they should be banned outright as illegal.

4.4 Data governance

Generative AI needs data and there is therefore considerable concern and interest in the data that goes into the training of foundation models and the data that is input into generative AI systems.

There is growing concern that generative AI violates intellectual property rights. Legal challenges have been mounted in the US concerning the use of data scraping for training data which could violate **copyright**.

Generative AI has been front and centre of the recent labour dispute and **strike** by the SAG-AFRA trade unions representing actors in the US. Whether any legal dispute will be successful is highly uncertain, however, the broader point is that the labour force of the creative industries is under threat from generative AI, which will have a direct effect on the UK economy as these industries represent **5.6% of GDP**.⁴²

As mentioned above, the UK's Intellectual Property Office is currently drafting a **code for AI and copyright** in an attempt to answer some of these questions. In the meantime, there are signals of a **broader debate** about the societal value and risk associated with absorbing a large portion of human knowledge into large AI models, potentially impinging on fundamental human rights, such as access to culture.

As a countermeasure to copyright concerns, there is a chance that companies will hold data in so-called walled gardens.

That would give the public even less access to open data and would stifle innovation and productivity. There is still room for clarification of the legal framework in this regard.

4.5 Regulatory capacity

The government's approach to AI regulation would support individual regulators to develop sector-specific frameworks for the adoption of the value-based principles by UK industry.

In many ways, this is a more concrete and pragmatic approach than the approach taken by other jurisdictions, notably the EU, where centralised, overarching principles have been adopted in comprehensive legislation. As such, the UK is showing more willingness to operationalise the principles in ways that will have a direct impact on the development and uptake of generative AI.

For example, the Competition and Markets Authority has **proposed a set of principles** to regulate the development of AI models.⁴³ (Competition & Markets Authority, **AI Foundation Models: Initial Report**, 18 September 2023.)

42. See <https://lordslibrary.parliament.uk/arts-and-creative-industries-the-case-for-a-strategy/#:~:text=The%20creative%20industries%20sector%20contributed,the%20UK%20economy%20in%202021>.

43. The principles are: (1) ensuring that foundation model developers have access to data and computing power, and that early AI developers do not gain an entrenched advantage; (2) that both closed and open source models are allowed to develop; (3) that businesses have a range of options to access AI models – including developing their own; (4) that consumers should be able to use multiple AI providers; (5) that no anticompetitive conduct like 'bundling' AI models into other services take place; (6) that consumers and businesses are given clear information about use and limitations of AI models.

The effectiveness of the approach taken by the Government will depend on regulatory capacity and there is a risk that efforts will be unnecessarily duplicated, or that regulatory frameworks will promote contradictory rules.

The [Sir Patrick Vallance Review](#) of digital technologies, published in March 2023, found more than 10 different regulators of digital technologies. We concur with others who have observed a need for centralised regulatory oversight to coordinate the efforts of the many departments and regulators. This is necessary to ensure that the UK's value-based principled framework for the governance of generative AI is adopted in a consistent manner across the UK's industrial sectors.

These functions are today met by the [Office for Artificial Intelligence](#) and Centre for Data Ethics and Innovation under the Department of Science, Innovation and Technology (DSIT), and the [Digital Regulation Cooperation Forum](#), which was formed as a membership organisation consisting of four key regulators: the Competition & Markets Authority, OFCOM, the Information Commissioner's Office, and the Financial Conduct Authority.

Alongside recent internal changes to Government's policy delivery infrastructure, through the establishment of DSIT, further changes to the Government's interactions with the external expert community are also expected, which may influence the ability to rapidly identify and respond to emerging technological changes with regulatory implications.⁴⁴

4.6 International leadership

The UK Government has repeatedly set forth an ambition of international leadership in AI, both in terms of development and regulation. In March 2023, the [Sir Patrick Vallance Review](#) asserted that the UK had a window of no more than 24 months to realise that ambition. In relation to the development of regulatory frameworks, the UK is struggling to keep pace, as suggested by [Figure 3](#) in an earlier section.

While the UK's Government has resisted calls for legislation to allow for growth and innovation in the sector, the lack of AI-specific legal regulation opens the possibility that the safe and responsible deployment of AI solutions and products will depend on the enforcement of rules devised and overseen by other jurisdictions or the international community. The absence of robust legislation poses a serious risk to the safety and trustworthiness of generative AI solutions, especially when these are devised wholly or in part by foreign companies.

Being first-to-the-post in adopting legislation to regulate AI is not necessarily a desirable objective if that legislation is not robust, balanced, and feasible. However, the UK's lack of binding regulation means that despite any ambition of the Government, the UK is failing to reach its ambition of international leadership in this regard.

44. The AI Council and Centre for Data Ethics and Innovation have recently come to the end of their term or been disbanded, with plans for an alternative approach to external engagement in development.



5. RECOMMENDATIONS TO BUILD CAPABILITY IN GENERATIVE AI

Although the [National AI Strategy](#) is concerned with the broader AI field and pre-dates the latest developments in generative AI, many of the key actions it sets out retain their relevance and do not need to be repeated here.⁴⁵

We focus instead on innovation and skills policy levers that both support the goal of making the UK a global leader in applying generative AI to the economy, and are not discussed in detail in the National AI Strategy. It is worth noting that exactly how these policy levers are used depends on whether the UK pursues *AI Nationalism* or a more open approach.

As noted by the National AI Strategy, increased compute capacity is a dependency for the development of most generative AI capabilities. An efficient way of mitigating the UK's compute deficit would be lobbying *hyperscalers* to establish GPU-clusters in the UK. This would allow organisations like the NHS to run fine-tuned foundation models with fewer concerns about data security and privacy.

In parallel, subsidies could be increased for companies developing capital-intensive proprietary and/or strategically important generative AI capabilities (e.g. chips; cybersecurity and defence applications). Tax incentives like the Seed Enterprise Investment Scheme (SEIS) could be enhanced to increase the supply of early-stage capital to generative AI startups at the application layer.

Tax credits could be introduced for all businesses to incentivise them to apply generative AI technologies to their existing operations and/or to develop

GLOSSARY

AI Nationalism – coined by Ian Hogarth to describe an approach to national AI policy which prioritises a country's strategic interests and/or the economic interests of its citizens

Hyperscaler – a company operating massive cloud computing infrastructure (e.g. Amazon Web Services)

new generative AI-powered products and services. Challenge prizes could be launched to identify and disseminate effective bottom-up uses of generative AI by teams and individuals inside organisations operating in industries where productivity gaps have been identified. They can also be used to motivate innovation in industries identified as potential growth areas for the UK economy.

An AI Nationalist approach would imply government acting assertively to steer market outcomes. Public sector procurement of generative AI capabilities could positively favour UK suppliers—for example, public funding for supercomputers could be made contingent on the use of chips designed by UK companies like Graphcore. Acquisitions of major UK generative AI companies by foreign rivals—comparable to the past acquisitions of Deepmind by Google, Arm by Softbank, or Instadeep by BioNTech—could be challenged.

45. "National AI Strategy - HTML Version." [n.d.]. Gov.uk <<https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version>> [accessed 22 September 2023]

By contrast, an open approach might involve designing a regulatory regime encouraging foreign generative AI entrepreneurs to set up in, or relocate their companies to, the UK. In addition to the National AI Strategy's plans to make visas easier to obtain, this might include corporation tax and entrepreneurs' relief incentives. It would not, however, be compatible with the kind of controls on mergers and acquisitions described above.

For generative AI to pervade the economy, school and higher education curriculums would need to be developed to increase both understanding of the technologies and critical thinking about how they are used in practice. Computer science education may need to be reformed, or a new discipline established, to teach software development using no-code and LLMs. These could also be the subject for new **Skills Bootcamps**, and/or upskilling programmes co-designed with employers and workers.⁴⁶

Regardless of whether an AI Nationalist or open strategy is pursued, our view is that legislation and regulation will be needed to remove impediments to the adoption of generative AI and ensure that the British public can trust organisations' use of the technology. We favour government adopting a principled approach to introducing legislation that would embed an ethical framework for the governance of generative AI in domestic law in multiple sectors.

It should forbid high-risk uses of generative AI, for example in the operation of critical infrastructure, where it could pose a significant threat to human safety or violate fundamental ethical rules.

Legislation takes a long time to pass. In the interim we recommend the adoption

of soft governance models, such as the **IEEE7001 Standard on Transparency**, together with moves to strengthen regulatory capacity. International standards may also be used as frameworks for legislative proposals.

We therefore support the All-Parliamentary Group on Data Analytics' (APGDA) recommendation for a centralised AI office with a renewed strategic focus to not only oversee and coordinate AI regulation across regulators, as set out in the Government's White Paper, but also to ensure that regulators enforce regulation. This could be achieved by, for example, bolstering the remit to the Office of Artificial Intelligence with a strategic focus on work programmes that identify regulatory gaps and empower existing regulators to deliver responsive regulatory interventions in their domains.

There continues to be a need for capacity building among regulators. Although this is well under way in some domains, as seen from the framework being developed by the Competition & Markets Authority, others will need further support to deliver the Government's current AI White Paper proposals.

In addition, as is already recognised, regulators need to enhance their existing co-operation to ensure clarity about responsibilities, as the technology will cut across all sectors. This coordinating function may need additional or more active guidance and support than is currently proposed. It is crucial that the regulatory oversight mechanism has sufficient resources and expertise to test and oversee the use of generative AI to build national capabilities for productivity, and be transparent about the oversight in order to inspire public confidence.

46. "Find a Skills Bootcamp." 2022. Gov.uk <<https://www.gov.uk/guidance/find-a-skills-bootcamp>> [accessed 22 September 2023]



ABOUT THE AUTHORS

Dr Ann Kristin Glenster is Senior Policy Advisor on Technology Governance and Law at the Minderoo Centre for Technology and Democracy. The Executive Director of the Glenlead Centre, she is a legal expert on information technology law and regulation in the UK, US, and EU. She holds a UK qualifying law degree, has been a doctoral visiting scholar at the Harvard Law School, and holds a PhD in Law from the University of Cambridge.

Sam Gilbert is an affiliated researcher at the Bennett Institute for Public Policy. He is the author of [Good Data: An Optimist's Guide to Our Digital Future](#) (Welbeck Publishing, 2021) as well as influential reports on data ethics, crypto, web3, the metaverse, and online safety. Previously, he was Employee No. 1 and Chief Marketing Officer at the fintech unicorn ManyPets, and held senior roles at Experian and Santander.

APPENDIX

The table below briefly summarises some of the other policy areas generative AI bears on—all of which deserve more thorough exploration than is possible here. The table is included to demonstrate how generative AI will have an impact across society, and also to show that while we are aware of that impact, this brief has too narrow a focus to allow a full investigation of these implications.

GLOSSARY

Jailbreak – to modify a model or device with the objective of removing restrictions put in place by its developer

Policy area	Issue
Competition	At the foundation model layer, access to compute and the concomitant capital requirements represent significant entry barriers. The market may tend towards monopoly, further entrenching incumbents (not least Google, Amazon, Microsoft, and Meta) and/or producing a new generation of big tech, with gatekeeping power over models enabling the extraction of economic rents. The release of open source models (e.g. Llama 2, Stable Diffusion) somewhat mitigates the threat to competition, but increases exposure to online harms (see above).
Labour market	Concerns over AI-driven job displacement are not new. While few jobs are at risk of full automation by generative AI, its aptitude for writing, classification, and summarisation seems likely to lead to job losses in customer service operations, administration, and creative industries.
Online harms	Open source foundation models can be run on users' own infrastructure, giving them the opportunity to circumvent controls on dangerous, illegal, or otherwise harmful uses. Stable Diffusion has been used to create child sexual abuse material, while jailbreaking LLMs allows them to be used to automate, optimise and scale harmful practices ranging from fraud (e.g. phishing, romance scams) to online radicalisation.
Information environment	Generative AI tools make webspam, misinformation, and disinformation easier and cheaper to produce at scale. Predictable consequences include a general dilution of the quality and factual accuracy of content available online; the proliferation of fake consumer reviews and inauthentic social media accounts; and increased volumes of 'fake news', political propaganda, and extremist material .
Education	ChatGPT has already had a disruptive impact on secondary and higher education institutions, thanks to its ability to produce plausible-sounding original essays and coursework with minimal input on the part of students. Written assignments and methods of assessment will obviously need to evolve—but this may present an opportunity to incorporate teaching of generative AI skills like prompt engineering into school and university curriculums (see below).
Social justice	Generative AI systems are prone to the same forms of embedded gender, class and racial bias as AI systems used for classification and decisioning tasks.
Climate	Training foundation models is computationally intensive and therefore energy hungry. The carbon footprint of generative AI development is likely to be exacerbated by arms-race dynamics, and could be in tension with Net Zero goals. Generative AI which raises questions about whether the benefit from the use of generative AI will outweigh its negative impact on the climate, and/or if that impact can be offset by other climate change action. For example, Australia's Chief Scientist observes that: "Managing the energy and water consumption of training and retraining (including data collection and cleaning) and operating LLMs and MFMs is a challenge. While techniques have improved the energy efficiency of algorithms, hardware upgrades and increasing levels of e-waste from computer components will heighten demand for critical minerals with resultant environmental and human rights impacts"—Australian Government Department of Industry, Science and Resources, Safe and responsible AI in Australia (Discussion Paper, June 2023) p. 13.
Geopolitics	Generative AI is completely dependent on the availability of GPUs, meaning it is subject to the dynamics of the chip market. GPUs contain rare earth metals, and as with other advanced chips, the majority are manufactured by Taiwan Semiconductor Manufacturing Company.

SELECTED BIBLIOGRAPHY

Australian Government Department of Industry, Science and Resources, Safe and responsible AI in Australia (Discussion paper June 2023)

Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Christopher Nagy, Madhulika Srikumar, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI, Berkman Klein Center For Internet & Society at Harvard University, 2020

Robert Ganna and Emre Kazim, Philosophical foundations for digital ethics and AI ethics: a dignitarian approach, *AI and Ethics* (2021) 1: 405-423

Samuele lo Piano, Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward, *Humanities and Social Sciences Communications* (2020) 7-9



UNIVERSITY OF
CAMBRIDGE

➤ www.cam.ac.uk

ai@cam

➤ ai.cam.ac.uk



Bennett Institute
for Public Policy
Cambridge

➤ www.bennettinstitute.cam.ac.uk

