The UK foundation models opportunity

Briefing note (April 2023)

Cover note

In December 2022, the AI Council's memo on the policy implications of advances in foundation models highlighted the importance of rapid Government action to build national capability in this increasingly important technology area. The memo noted the importance of seeking input from across industry, research, and civil society to examine options to build national capability, and to create an action plan to position the UK for international leadership. Responding to this need, a subgroup of the AI Council has convened to deliver this briefing on the UK foundation model opportunity, based on engagement across its networks within the UK and internationally. This subgroup was led by Neil Lawrence, in collaboration with Tabitha Goldstaub, Wendy Hall, Nick Jennings, and Marc Warner.

This briefing note highlights two types of policy intervention required to position the UK for international leadership in foundation models:

- *Drive further progress delivering foundational AI policy framework*s: Progress in foundation models highlights the importance of establishing a policy infrastructure that enables the safe and effective development of AI, supports its adoption across sectors, and ensures all in society can have confidence in its use. The AI Roadmap and National AI Strategy provide a framework for progress.
- *Establish national capabilities to deliver core policy functions in relation to foundation models*: These capabilities should provide a proving ground to assess the functionality and limitations of foundation models before deployment; develop protocols for responsible research and innovation; build private sector capability; encourage adoption in areas of high potential benefit; provide an infrastructure for open research and innovation; increase collaboration with allies internationally; and ensure the UK has access to cutting edge capabilities needed for national security.

The Council has already established networks of expertise that Government can use to move quickly in developing its response to the opportunities and challenges presented by foundation models. Current interest in foundation models, coupled with the creation of the Department for Science, Innovation and Technology, provides an opportunity to establish an adaptive approach to AI policy, drawing from independent, expert advice that moves quickly to identify emerging issues, facilitate the development of policy responses, and support effective implementation. The AI Council can support Government in galvanising the community to deliver this progress, through both responding to rapid shifts in technical capabilities that demand near-term responses and building the UK's AI ecosystem over the long-term.

Summary

Rapid developments in Large Language Models, and other foundation models, are driving a new wave of progress in AI. This gearshift in AI capabilities creates opportunities for the UK, as powerful AI tools become more accessible, unlocking applications of AI in domains including healthcare, education, and science. However, if not carefully designed and implemented these models can harm individuals and society and contribute to malicious applications. In December 2022, the UK's AI Council highlighted the need for rapid Government action to identify the policy implications of foundation models, both to position the UK for international leadership in this strategically significant area and to ensure advances in AI increase wellbeing and prosperity across the UK. We welcome Government's action to establish a taskforce to focus on these issues, alongside the continuing

commitment to building the UK's AI sector signalled through Budget 2023's announcement of funding for exascale computing, further support for AI R&D, and the AI regulation White Paper. In pursuing this agenda, Government needs to be able to understand and respond to rapidly changing technological capabilities. Connecting this policy context to changing capabilities in foundation models, and the opportunities and challenges they create, this briefing sets out two types of policy intervention that can help deliver the Government's ambitions to build a world-leading UK AI community and translate advances in AI to social and economic benefits. It highlights the need for further progress on the fundamentals of AI policy, alongside interventions to increase national capability in foundation models.

The UK has created a policy infrastructure for AI. Through the AI Roadmap, National AI Strategy, Science and Technology Framework, Integrated Review, and AI White Paper, Government has identified levers that can create the conditions for a flourishing UK AI sector. These include investment in research and innovation, skills and capability-building, public dialogue and engagement, support for business, and governance and regulation. Many of the debates surrounding foundation models highlight the importance of securing these fundamentals of AI policy. Capitalising on the opportunity to work strategically and collaboratively through the Department for Science, Innovation, and Technology, further work is needed to drive progress in these policy building blocks. The AI Council has already played an important role in identifying these building blocks and supporting Government to create a policy package to support the AI sector. Its links and networks can enable Government to move quickly in developing its policy response, providing independent, expert input to policy development and implementation.

The scale of resources required to create and maintain foundation models distinguishes them from other current AI methods, and has resulted in an environment where technology development is led by a small number of large companies. Unless it takes action to establish national capability in this area, Government risks creating a technology gap that will become more difficult to bridge over time. This national capability should serve the UK in seven key ways. It should:

1. Provide a proving ground that connects research and practice, enabling rigorous testing of the functionality and limitations of foundation models.
2. Embed responsible research and innovation practices in foundation model development.
3. Enhance UK growth and productivity, building UK private sector capability.
4. Encourage responsible technology adoption in areas of societal need.
5. Provide an infrastructure that enables open research and innovation, bringing together researchers, businesses, and the public sector.
6. Increase international collaboration, boosting the UK's national profile in foundation models and working with partners to ensure maximum impact for UK investments.
7. Ensure the UK has access to cutting edge AI capabilities that are critical for maintaining national security.

Today's Large Language Models and foundation models represent the latest wave of progress in a rapidly changing technology area. Further developments in AI will yield more powerful systems, with new implications for science and society. Government needs to be agile in its response to such developments, working with the AI community to anticipate strategically significant changes and respond to rapid shifts in technical capabilities. The AI Council can support Government in galvanising the community to deliver these insights.

Background

In December 2022, a ministerial briefing from the AI Council recommended that Government takes action to build national capability in Large Language Models, and that it initiates this action by engaging across academia and industry to better understand what type of Government intervention would be most useful. This document updates that briefing, based on work by the Council, convening across research, industry, civil society, and policy communities, to identify a pathway towards building national capability.

Large Language Models are a type of AI model that can process, summarise, and generate text. The systems have typically been trained on large datasets to create models with billions of parameters. These represent one type of foundation model; a class of AI models that are developed using very large, unlabelled training datasets, and that can be adapted to perform different tasks, albeit of a similar nature. Both these model types are part of a field called generative AI, which analyses pre-existing data – such as images, video, audio, text – and uses its analysis to generate new outputs.

Three characteristics distinguish the recent progress in foundation models from previous waves of technological development in AI:

- Capabilities: Advances in AI over the last ten years have created systems that deliver highly accurate results when trained on tightly specified tasks in controlled environments. Foundation models provide a basis for broader-purpose AI tools. Once trained to generate text, for example, a Large Language Model can respond to user questions about a variety of topics, depending on the dataset on which it was trained. Continuing progress in this domain is expanding capabilities further by creating multi-modal systems, which can analyse and generate combinations of text, image, audio, and video content. These systems offer highly plausible results across a class of tasks, but these plausible results are not necessarily accurate or reliable. Foundation models continue to be limited by their training; they do not display the type of intelligence expected from humans.
- Accessibility: The ease with which users can interact with foundation models offers to make accessible AI capabilities that might otherwise have required specialist skills to deliver. This accessibility underpins much of the excitement surrounding the potential of AI across sectors. The integration of low-data learning strategies, such as one- or zero-shot learning, could further streamline their development, enabling rapid set-up and deployment of AI tools with broader purpose.
- Scale: Significant resources are required to develop, train, deploy, and maintain the Large Language Models in use today.[1] Some of the sophisticated capabilities demonstrated by these systems seem only to emerge with very large models and very large data. Resources to access compute, data, and skills are all necessary to operate at the scale required. As a result, this field has been dominated by big technology firms. Open-source approaches are emerging,[2] but still require resources for hosting and processing to enable their use.

The result is an environment where technical developments are rapidly changing the capabilities of AI systems; where widespread adoption of these new tools could accelerate the profound social and economic consequences of AI, creating both opportunities and risks; and where the scale of investment and pace of change pose challenges for national governments seeking to influence the direction of technological progress.

[1] While the exact cost of systems such as GPT3 is difficult to calculate, estimates suggest training costs of $5-12M; further costs can be expected to accrue for maintenance.
[2] See, for example: https://bigscience.huggingface.co/blog/bloom

Alongside the existing National AI Strategy,[3] policy announcements in Spring 2023 provide a framework for understanding potential Government action to build national capability in foundation models.

- The UK Science and Technology Framework sets out an overarching vision to delivering prosperity and security for the UK, based on achieving strategic advantage in priority areas of science and technology, creating a world-leading foundational science base, and establishing the agility to rapidly respond to emerging discoveries.[4]
- Budget 2023's commitment to invest £900M in increasing the UK's compute capacity,[5] alongside wider actions to implement the recommendations of the Ghahramani Future of Compute Review,[6] outlines an approach to building compute infrastructure.
- The Integrated Review Refresh sets out Government's approach to securing strategic advantage through science and technology, including AI, through both investment nationally and cooperation internationally.[7]
- The White Paper on AI Regulation refreshes Government's approach to Pillar 3 of the National AI Strategy – *Governing AI effectively* – by increasing collaboration across Government and with regulators in support of a risk-based regulatory regime.

This policy and regulatory landscape offers opportunities to bring together researchers, businesses, civil society, policymakers, regulators, and international collaborators to help achieve the UK's ambitions of being an AI superpower.

Foundation models and the UK's AI policy infrastructure

The Council's December 2022 Large Language Models memo noted how this step-change in AI capabilities could unlock opportunities across sectors: *In education, [Large Language Models] could act as personal tutors to enhance teaching and learning by providing accurate answers to student questions; in healthcare, they could synthesise complex patient information from diverse sources of text and data to support doctors in identifying treatment pathways; across the public sector, they could help citizens access the information and services they need, responding to user questions by synthesising large volumes of information from different websites and presenting a user-friendly answer. As the ability of these models to generate convincing, novel text increases, they may also be deployed to analyse genomic data (identifying the effects of genetic changes from nucleotide letter sequences) or environmental datasets (supporting sustainable living).*

If successfully developed in the UK, such systems could boost economic growth, alleviate pressures on public services, and enhance the UK's position in AI internationally. To deliver these benefits, careful stewardship is needed to support AI development and deployment, while tackling the associated risks or harms. The Large Language Models in use today have already demonstrated their potential to reproduce harmful biases, present false information, or mislead users, while also highlighting wider concerns about data use, privacy, monopolisation, job losses, and marginalisation of vulnerable communities.[8] Both the opportunities and challenges associated with rapid technological change need to be considered by Government as part of its taskforce on foundation models.

[3] https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version
[4] https://www.gov.uk/government/publications/uk-science-and-technology-framework/the-uk-science-and-technology-framework
[5] https://www.gov.uk/government/publications/spring-budget-2023/spring-budget-2023-html
[6] https://www.gov.uk/government/publications/future-of-compute-review#:~:text=The%20Future%20of%20Compute%20Review,a%20science%20and%20technology%20superpower.
[7] https://www.gov.uk/government/publications/integrated-review-refresh-2023-responding-to-a-more-contested-and-volatile-world
[8] https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Understanding the implications of rapid technical progress and widespread adoption requires insights from across disciplines and sectors. Leveraging its networks across the AI community within the UK and internationally, the AI Council provides a forum for generating these insights.

Community engagement carried out by the AI Council in March 2023 highlighted four areas where policy action is needed to help steward the development of these technologies:

- *Research, development, and innovation*: Continued progress in the power and deployability of foundation models will rely on advances in the core underlying technologies that support their development. Delivering their benefits through new applications in healthcare, transport, education, science, and other domains will require overcoming the implementation issues that affect the use of these models, to deliver robust, reliable, human-centric techniques. Alongside these AI capabilities there is a need for continued progress in the software and hardware infrastructures that support the development of large models. Responsible research and innovation is vital across the development pipeline, requiring new research practices and standards.
- *Education and skills*: User interactions with Large Language Models have already signalled how these systems will shape the online information environment. Convincingly 'human' responses to user queries can encourage false confidence in their results, which is compounded by a lack of clarity surrounding their capabilities and limitations, or false assumptions about their maturity. In response, users need tools to navigate this new information environment.[9] Such tools might include signifiers of AI-generated material, educational interventions to increase understanding of the uses and limitations of these systems, and wider public engagement to build a well-founded conversation about the capabilities of AI. In the longer term, as can automate more complex tasks, there is also a need to future-proof the labour market, through interventions to upskill the workforce and enable flexible retraining.
- *Regulation*: Changing technical capabilities highlight the importance of policymakers having access to insights and expertise from the frontiers of AI development. While some regulatory bodies have invested in building capacity in data and AI, there is a need to support regulators across sectors to steward the development of these technologies across all stages of the development pathway from data collection to AI deployment, for example through mechanisms for market surveillance, sandboxes, or equivalent agile regulatory interventions.
- *National security:* Progress in AI has already shaped the cybersecurity environment, creating new defensive and attack capabilities. Better understanding of the implications of future progress in this area is needed, alongside infrastructures for the safe and secure development of foundation models in sensitive domains, and spaces to test the capabilities of the models available on the market.

Many of these opportunities, issues, and concerns are shared across AI techniques, and across digital technologies more broadly. Progress in foundation models therefore increases the importance of establishing a policy infrastructure that enables the safe and effective development of AI, supports its adoption across sectors, and ensures all in society can have confidence in its use. The AI Council's Roadmap[10] and National AI Strategy[11] provide a basis for this infrastructure. Extending this work, current interest in Large Language Models, coupled with the creation of the Department for Science, Innovation and Technology, provides a moment to establish an adaptive approach to AI policy. Such an approach would

[9] Guidance from the DfE on the use of LLMs in education is an example of the action being taken to develop institutional responses to this risk: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1146540/Generative_artificial_intelligence_in_education_.pdf

[10] https://www.gov.uk/government/publications/ai-roadmap/executive-summary

[11] https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version#summary-of-key-actions

build on existing frameworks, driving deeper progress in areas of policy need. Annex 1 maps the areas for action emerging from current conversations about Large Language Models to the four pillars of the AI roadmap, indicating how renewal of these policy frameworks can support long-term capability building for AI in the UK.

The need for national capability in foundation models

Building the policy foundations that enable the safe and effective use of foundation models is necessary to ensure the UK remains a leader in AI technologies over the long-term. In the near-term, action is also needed to position the UK for strategic influence in this domain. The precise form of this action is the current focus of a Government taskforce on foundation models, announced in March 2023. This preparatory work by the AI Council can be used to refine the scope and focus of that taskforce.[12] The Council's expertise and networks offer a resource to enable the taskforce to move quickly in developing its work.

To achieve strategic advantage in this area of AI, and translate that advantage into prosperity and security for the UK, the policy landscape needs to deliver seven foundation-model-functions:

1. *Scrutinise*: Provide a proving ground that connects research and practice, enabling rigorous testing of the functionality and limitations of foundation models.
2. *Steward*: Embed responsible research and innovation practices in foundation model development, creating and ensuring implement of standards, protocols, and best practices for responsible AI.
3. *Grow*: Enhance UK growth and productivity, building UK private sector capability through smart procurement that incentivising private investment and developing public-private partnerships for long-term strategic advantage.
4. *Promote*: Encourage responsible technology adoption in areas of societal need, positioning the public sector as an early customer of current technologies while implementing guardrails to ensure safe use.
5. *Innovate*: Provide an infrastructure that enables open research and innovation, bringing together researchers, businesses, and the public sector to ensure the UK is at the forefront of the next generation of AI technologies.
6. *Collaborate*: Increase international collaboration, boosting the UK's national profile in foundation models and working with partners to ensure maximum impact for UK investments.
7. *Secure*: Ensure the UK has access to cutting edge AI capabilities that are critical for maintaining national security, and can respond to the security threats that emerge from advances in AI technologies.

Delivering these functions should accelerate the UK's overall capability in foundation models, while ensuring in areas of critical need or infrastructure there is sufficient national technology infrastructure to enable progress. The delivery mechanisms available to support this functionality span the own, access, collaborate framework set out in 2021's Integrated Review.[13] Government can invest in areas of critical need, leverage private sector partnerships to increase the UK's national capabilities, and collaborate with international allies to drive progress.

[12] https://www.gov.uk/government/publications/integrated-review-refresh-2023-responding-to-a-more-contested-and-volatile-world/integrated-review-refresh-2023-responding-to-a-more-contested-and-volatile-world
[13] https://www.gov.uk/government/publications/global-britain-in-a-competitive-age-the-integrated-review-of-security-defence-development-and-foreign-policy

Looking ahead

As with any technology, progress in AI is not linear. The last ten years have seen exciting waves of technical progress from innovations in different AI techniques. Current developments in Large Language Models increase the power and potential of AI technologies. In the coming years and decades there will be further advances that yield more powerful AI techniques. The challenge for policymakers is to act today amidst uncertainty about the pathway of technology development, ensuring technological progress yields AI techniques that are both technically sophisticated and aligned with societal interests in their development and use. Responding to this challenge will require agility. Government needs to be able to identify and respond to strategically significant developments, which might include rapid discontinuities in AI capabilities. Working across Government, the Council can galvanise action from the AI community, anticipating emerging technology issues, facilitating rapid policy response, and supporting policy implementation.

Annex 1: Areas for policy action emerging from current conversations about Large Language Models

AI Roadmap pillar	Emerging areas for action
Research, development, and innovation	● Invest in the AI research and innovation infrastructure, encouraging development of model, software, and hardware tools that can create next generation AI systems. ● Drive progress in core underlying technologies, ensuring the building blocks are in place to deliver human-centric, deployable AI, and to advance the capabilities of foundation models. ● Develop standards, best practices and protocols for responsible research and development, including co-design of research challenges with affected communities, consideration of duel use risks or unintended consequences, and release protocols. ● Build an infrastructure to open use of foundation models. ● Horizon-scan to anticipate for emerging developments in AI and enable rapid policy responses. ● Support international collaborations to drive progress with aligned partners.
Skills and diversity	● Integrate AI and data literacy across education, adapting curricula across schools, further, and higher education to increase understanding of AI, encouraging critical thinking about its use. ● Work with industry to better understand the potential impact of AI on workers, acting to future-proof the workforce, through upskilling in AI and flexible retraining programmes. ● Bring underrepresented voices and viewpoints into the design and development of AI technologies.
Data, infrastructure, and public trust	● Develop tools for detecting AI-generated content, such as watermarking. ● Create resources to demystify AI, clarifying definitions and terms in debate. ● Increase public engagement and dialogue efforts to build understanding of AI capabilities and limitations. ● Create use cases that can increase understanding of how to use foundation models, their opportunities and their risks. ● Work with regulators to build capacity in AI monitoring and evaluation, empowering them to act while enabling experimentation. ● Create sandboxes and living labs to allow promising use cases to develop in ways that are guided by regulatory expertise. ● Support the development of AI tools that embed societal interests in AI by-design, including privacy, security, and ethics. ● Collaborate with international partners to align with emerging standards, for example the US NIST risk management framework.
National, cross-sector adoption	● Create proving grounds or test beds that support the development of new applications, and build mutual understandings about the capabilities of foundation models across regulators, companies, and researchers.

	● Create challenge-led, or problem-based, approaches that deploy AI in areas of social need, leveraging these programmes to address public interest or safety concerns. ● Use public sector procurement to drive the development of foundational capabilities in industry. ● Work with international allies to enhance information sharing across industry and governments. ● Review national security frameworks through the lens of 'AI readiness'. ● Review IP and copyright legislation to help users understand the contours of this regulations in the context of foundation models.

Annex 2: AI Proving Grounds: A Potential Model for UK Investment in Foundation Models

This document proposes an approach to building UK capability in Large Language Models and generative AI. The proposal is based on the idea of an *AI proving ground*, which in turn is inspired by the structure and development of NACA, the US National Advisory Committee on Aeronautics. NACA played a key role in establishing the US lead in aeronautics and space flight in the 20th Century. Over time it evolved from a research committee, to a testing ground for aircraft, to a foundational component of NASA.

Scientific Context: Transformative Advances in Generative AI

With the rapid adoption of Large Language Models and other forms of generative AI, significant and highly accessible forms of artificial intelligence are becoming available that will disrupt UK society at all levels across the private and public sectors. The opportunities and challenges arising from these new technologies make it imperative that the UK puts "technology steerage" in place to ensure outcomes that align with our national priorities and values.

Historical Context: Focussed Innovation in Aeronautics for US National Advantage

The United States formed NACA in 1915 when the US government became concerned it was losing its early lead in aeronautics to European powers engaged in the First World War. The objective was to undertake, promote, and institutionalize aeronautical research. By the outbreak of the Second World War NACA operated a proving ground in Langley Field, Virginia. This tested each new aircraft before deployment as well as working with industry to develop aircraft when challenges revealed themselves in use. The testing was done through a combination of wind tunnel and flight tests, by a team of test engineers and pilots. NACA played a key role in developing the technologies that allow high-altitude flight and supersonic flight, which were fundamental to US air supremacy in the post-war years.

Lessons to Learn

From the AI proving ground perspective, three important characteristics of the laboratory are as follows:

1) The laboratory had national profile.[1]
2) The laboratory did not build aircraft, it tested aircraft that had been built and/or commissioned.
3) Research and practice were tightly coupled.[2]

These three characteristics ensured that the proving ground became not only a centre of excellence but a collaborative meeting point between government, research, and industry and a trusted authority that set the standards of flight. [3]

NACA had access to world-leading testing facilities, and made fundamental research advances to improve these facilities. For example, by 1931 they had the most powerful wind tunnel in the world. Importantly, the structure of NACA also proved adaptable over time: when needs changed, focus changed.

Proposal

An AI proving ground, with national profile and substantial resource behind it, dedicated to testing advances in generative AI and foundation models. This would test AI technologies that have been built by or commissioned from private industry: researchers there would collaborate closely with engineers and data scientists in industry and government to solve the practical and foundational challenges that arise when deploying AI models to address

UK national priorities. The key here is *practice-led research*, with world-class scientists in tightly focussed collaborations underpinned by flexible, agile support. The proving ground would be led by data scientists and engineers with substantial international profile, and would provide plentiful, flexible support and world-class facilities to attract and sustain the best young researchers and practitioners.

Appendix: More on NACA

Alongside revisions and refinement to the aircraft produced for the US during the Second World War, a short list of research achievements of NACA includes:

1) The research behind the “laminar flow wing”, a collaboration between North American Aviation and NACA. This wing was deployed on the P51 Mustang, making it the longest-range fighter aircraft the Allies possessed in the war.
2) The research behind the X-1 experimental plane, a collaboration with Bell Aircraft. This was the first aircraft to break the sound barrier in level flight.
3) Setting the first national standards for handling characteristics of aircraft.

Once it was integrated into NASA, Bob Gilruth (the NACA lead) led the Space Task Group that commissioned the Mercury capsule, the Gemini capsule and designed and commissioned the Apollo Command Module and Lunar Module that Armstrong piloted to the Moon. Armstrong himself was a former NACA test pilot who flew from the Muroc Flight Test Unit, set up in 1946 for flying and testing rocket planes such as the X-1 and the X-15.

---

[1] For example, the committee engaged aviation luminaries such as Orville Wright and Charles Lindbergh (although the work was done by bright young engineers and test pilots).

[2] Test pilots worked closely with engineers on programmes of research that recreated and solved problems that had been discovered in operation. This was fundamental, for example, in solving the design challenges caused by transonic flow and enabling supersonic flight.

[3] See Bob Gilruth’s 1941 NACA Report R755, “Requirements for Satisfactory Flying Qualities of an Airplane”.