Policy Paper

Advancing the Global AI Discourse

BritishAmerican Business Recommendations for the AI Safety Summit

Holly Harris, Policy Coordinator, BritishAmerican Business
Emanuel Adam, Chief Trade and Policy Officer, BritishAmerican Business

Introduction

On June 7th, 2023, the UK announced that it would host the first global summit on artificial intelligence (AI) on November 1st and 2nd at Bletchley Park. Relevant government and business leaders plan to attend, with US and UK stakeholders likely to make up a significant share of attendees.

The Summit arrives at a time when AI has become one of the top issues for business and government alike. In fact, AI has already become a transformative and revolutionary technology that promises further significant economic and societal benefits. According to McKinsey's 2023 annual Global Survey, one third of businesses regularly incorporate AI into at least one business function.[1] This ratio is expected to continue to grow as AI technology becomes more accessible.

However, the increased awareness of the growing power of AI technology has triggered a widespread debate around AI use, risk, and regulation. Notable efforts include the OECD AI Principles, the draft AI Act in the European Union (EU), the White House voluntary AI commitments and subsequent Executive Order on AI, and the G7 Hiroshima AI Process.

According to information released by the UK Government, the discussions at the AI Summit will focus on "frontier" AI, which the UK government has defined as "cutting-edge AI which holds the most opportunities while also presenting new and dangerous risks."[2] Discussions are meant to complement existing efforts in the wider international discourse on AI governance.

The event has been framed as an opportunity for the UK to signal its leadership in global AI discourse, offering a forum for governments alongside private businesses, including tech firms, and civil society to have meaningful conversations on steps needed for safety in the face of rapidly developing frontier AI technology.

Looking at the Summit through a transatlantic lens, this BAB Policy Paper builds on the 2023/24 BAB Policy Agenda and seeks to establish how the discussions held at the AI Safety Summit in Bletchley Park can and should help advance the global AI discourse.

State of Play

The concept of AI has been discussed in the abstract by scientists and mathematicians since the 1950s. Broadly speaking, the term AI refers to the ability of machines or software to simulate human intelligence. The application of AI is relevant to all industries and can contribute to tasks ranging from customer service chatbots to increasing the accuracy of cancer diagnoses.

Since the 1990s, AI has found an increasing application in products and services. Common use-AI like virtual assistants and search engines saw advancements in the 2000s and 2010s, though it is only in the 2020s that the view on AI has shifted from being a helpful and relatively benign tool to a powerful technology with massive economic and societal potential, but which also comes with significant risks that need to be considered and managed carefully.

The breakthrough of AI as a wider discussion topic came with the public release of OpenAI's ChatGPT in November 2022. The tool was not only a milestone in technological advancement, but also a watershed moment in public awareness of AI. Increased public knowledge and understanding of ChatGPT and similar generative AI models have sparked further public interest in AI and served as a catalyst for

                                 Advancing the Global AI Discourse - Recommendations for the AI Safety Summit

governments across the world to turn their attention and resources to AI use and regulation.
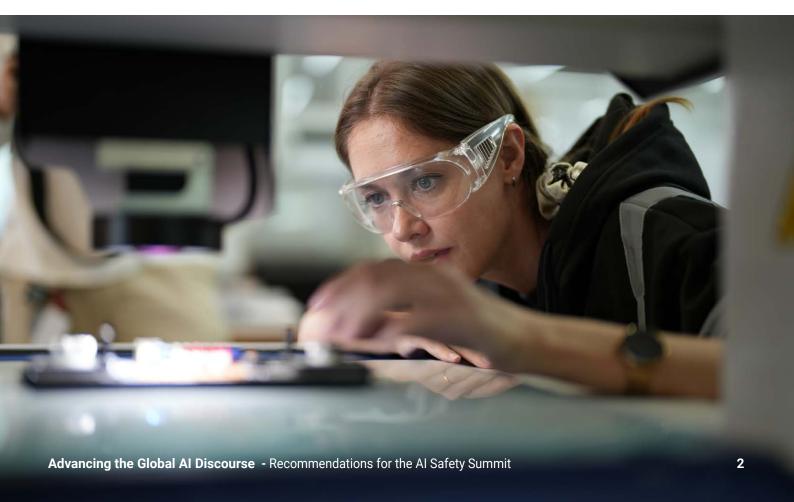
The prominence and influence of AI technology is augmented by its economic value. Both the UK and the US rank among the most successful markets for AI development globally.[3] For example, the UK's AI market is currently valued at over $21 billion with the potential to add $1 trillion to the UK economy by 2035.[4] The US market is valued even higher at $100 billion with expectation to grow to nearly $2 trillion by 2030.[5]

In terms of the UK Government's focus on AI, the transformative potential of the emerging technology is championed through the UK National AI Strategy.[6] Published in 2021, the National AI Strategy sets out a 10-year plan for securing Britian as a global AI superpower. The strategy delineates the plan under three pillars: 1) investing in the long term needs of the AI ecosystem, 2) ensuring AI benefits all sectors and regions, and 3) governing AI effectively.

In 2023, the UK Government set up the Department for Science, Innovation and Technology (DSIT) and appointed a Minister for AI and Intellectual Property (IP). The Minister oversees DSIT's new Office for AI with the purpose of monitoring and ensuring the implementation of the UK National AI Strategy.

Further, in September 2023 the UK renamed and reconfigured its Foundation Model Taskforce, initially established in April 2023, to become the Frontier AI Taskforce. The Taskforce once aimed to build sovereign foundation model capability and boost UK AI adoption, but has pivoted its focus towards AI safety in advance of the Summit.[7]

On the regulatory side, in March 2023, the UK government published a white paper entitled "AI Regulation: A Pro-Innovation Approach,"[8] which includes proposals on how to regulate AI in a pro-innovation manner.[9] While the general approach to future AI regulation has been perceived as positive, questions around, for example, the pros and cons of a flexible definition of "AI systems" or the "sector-specific approach to regulating AI" remain.[10] The government is expected to set out next steps before the end of 2023, though there remains a degree of uncertainty on how closely future action will reflect the proposals of the white paper.

In the US, initial attempts to embrace AI in public policy began in 2016 with the US' first AI report under President Obama. Under the Trump Administration, there were multiple Executive Orders, as well as two pieces of legislation, the AI in Government Act and the National AI Initiative Act, which were signed into law in 2020. These initiatives focused on guiding principles in the US AI strategy, including an emphasis on facilitating research and development, removing barriers to AI innovation, and improving AI training in the workforce.[11] The initiatives also led to the establishment of the interdepartmental National AI Initiative Office (NAIIO) in early 2021 to oversee the implementation of the AI strategy.

Since the change in administration in 2021, the US approach towards AI and the accompanying governance structure has shifted. The National AI Advisory Committee (NAIAC) was established in 2022 as an interdisciplinary committee of AI-related experts to advise the NAIIO and the White House, though the NAIAC itself has limited powers.

The Biden Administration further developed their approach towards AI through a July 2023 meeting with seven leading AI development companies, Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI, which established a set of voluntary commitments to manage risks posed by AI.[12] These commitments, which broadly emphasize safety, security, and trust, include more specific measures such as internal and external security testing of AI systems before release, investing in cybersecurity and insider threat safeguards, and developing and deploying mechanisms that enable users to understand if audio or visual content is AI-generated. In September 2023, eight additional companies, Adobe, Cohere, IBM, Nvidia, Palantir, Salesforce, Scale AI, and Stability, joined the initial seven companies who signed the commitment in July.[13]

As anticipated, on October 30th 2023, President Biden signed an Executive Order on Safe, Secure, and Trustworthy AI which seeks to advance American leadership in responsible AI governance.[14] Building on principles from the voluntary commitments made in July, the Order establishes a series of standards for AI use, safety, and security, while also setting up frameworks for the innovation and development of AI. Further, the Order seeks to expand bilateral, multilateral, and multistakeholder engagements with the aim to establish robust international frameworks for AI. The Order also reaffirms the importance of the voluntary commitments and gives further momentum for future legislative action from US Congress.

At the transatlantic level, the UK and the US included AI in the Atlantic Declaration, a broad economic and political pact between the US and UK.[15] In the document, released in June 2023, the UK and the US committed to cooperation on AI with a focus on ensuring the safe and responsible development of the technology, which includes their engagement in international fora, such as the OECD, the UN, the Global Partnership for AI, the Council of Europe, and the International Organization for Standardization, as well as the G7 Hiroshima AI Process.

The G7 Hiroshima AI Process merits particular attention as an important platform that can help set interoperable rules for governing AI, especially as policymakers across G7 countries, and elsewhere, are considering legislation or regulation affecting the use of AI.[16] On October 30th, in a statement released leading up to the UK AI Safety Summit, leaders stressed that the Hiroshima AI Process 'will foster an open and enabling environment where safe, secure, and trustworthy AI systems are designed, developed, deployed, and used to maximize the benefits of the technology while mitigating its risks, for the common good worldwide, including in developing and emerging economies with a view to closing digital divides and achieving digital inclusion'.[17] A comprehensive policy framework featuring the results of the discussions is expected to be released by the end of 2023.

The latter is even more relevant as the EU has already advanced its efforts to establish the world's first comprehensive AI legislation, the draft EU AI Act.[18] The Act, itself the result of various EU strategy plans and groupings, is designed to lay out a horizontal, risk-based regulatory framework which establishes different rules for AI systems according to perceived risk levels. While some endorse the draft Act as a clear and comprehensive legal framework that protects users and raises the standard for AI safety,

other commentators view it as too restrictive in that it could stifle the innovation needed to advance the technology.

For the AI Safety Summit, the UK government has established five objectives, which will be subject to a series of roundtable discussions and meetings throughout the event:

1. Create a shared understanding of the risks posed by frontier AI and the need for action;

2. Advance a forward-looking process for international collaboration on frontier AI safety, including how best to support national and international frameworks;

3. Establish appropriate measures which individual organizations should take to increase frontier AI safety;

4. Identify areas for potential collaboration on AI safety research, including evaluating model capabilities and the development of new standards to support governance;

5. Showcase how ensuring the safe development of AI will enable AI to be used for good globally.

With these objectives, the UK government aims to use the Summit to begin a global conversation on how to identify, evaluate, and ultimately mitigate the potential risks of frontier AI.

Recommendations for Policymakers

- We welcome the AI Safety Summit, the focus it provides on the risks of frontier technology, and its efforts to approach these risks in a coordinated and collaborative way. We welcome the leadership role the UK is playing as a driving force in international discussion around AI use, risk, and regulation.

- We agree that the Summit should complement and support existing national and international efforts, including advisory bodies, to manage risks associated with AI. We welcome the emphasis on international cooperation and encourage the Summit participants to work towards globally harmonized and interoperable regulation of high-risk AI use.

- We encourage conversations on risk mitigation to consider consequences of the forthcoming AI regulations on a sector-specific basis with a view to avoid unnecessarily inconsistent regulatory approaches between sectors. Further, we encourage efforts on risk mitigation to focus not only on the attributes of specific models, but also on the wider systems in which these models are embedded in order to better manage specific risks within specific contexts.

- We welcome the Summit to facilitate the development of concrete measures, including interoperable technical standards, which individual organizations should take to increase frontier AI safety. We recommend that these build on the principles found in the US Executive Order on AI and the White House voluntary AI commitments emphasizing safety, security, and trust. Further, we recommend that these measures be translated into risk-based, targeted regulation that is applicable to nations and organizations globally.

- While recognizing the necessity of data access for well-functioning AI models, we encourage policy makers to pursue international approaches that build public trust in data use and advance human rights. Building on the US-UK data bridge, we encourage continued discussion among international partners how to create interoperable international approaches to data transfer and protection. We are encouraged

to see a call to pass comprehensive federal privacy legislation to address AI-related challenges in the US.

- We welcome the allowance of open dialogue for researchers and developers to share evidence of rogue AI systems to avoid AI-related existential catastrophes.

- We acknowledge the potential implications of AI on national security, and we encourage Summit partners to collaborate on facing security challenges rather than adopting protectionist measures.

- We welcome the participation of technology companies in the AI Safety Summit and encourage the expansion of industry participation in subsequent discussions to include a wider cross-section of the AI technology ecosystem beyond solely model developers.

- Building on the commitments made in the Atlantic Declaration, we welcome and encourage the US and the UK to use the outcome of the Summit to further strengthen their partnership to shape global discussions on the safe and responsible development of AI technology in a joint manner.

Conclusion

On November 1st and 2nd, the UK will host the first global AI Safety Summit. This paper makes a case for the Summit to be ambitious in its ability to inform and complement relevant discussions on AI use, risk, and regulation at a national and international level, with the aim to ensure the safe and responsible development of AI technology globally.

References

[1] McKinsey & Company, The state of AI in 2023: Generative AI's breakout year

[2] UK Government - Department for Science, Innovation & Technology, AI Safety Summit: introduction

[3] Goldman Sachs, AI investment forecast to approach $200 billion globally by 2025

[4] US Department of Commerce - International Trade Administration, United Kingdom Artificial Intelligence Market

[5] Statista, Artificial intelligence (AI) market size worldwide in 2021 with a forecast until 2030

[6] UK Government, National AI Strategy

[7] UK Government, Industry and national security heavyweights to power UK's Frontier AI Taskforce

[8] UK Government - Department for Science Innovation & Technology, Department for Artificial Intelligence, A pro-innovation approach to AI regulation

[9] Goodwin Procter, Overview of the UK Government's AI White Paper | Insights & Resources

[10] Sidley, UK Sets Out It's "Pro-Innovation" Approach To AI Regulation - Data Matters Privacy Blog

[11] The White House, The White House Launches the National Artificial Intelligence Initiative Office

[12] The White House, FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI

[13] Fedscoop, Eight more tech companies join White House commitment to manage AI risks

[14] The White House, FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI

[15] THE ATLANTIC DECLARATION: A FRAMEWORK FOR A TWENTY-FIRST CENTURY U.S.-UK ECONOMIC PARTNERSHIP

[16] OECD, G7 Hiroshima Process on generative artificial intelligence (AI)

[17] G7, G7 Leaders' Statement on the Hiroshima AI Process

[18] European Commission, A European approach to artificial intelligence | Shaping Europe's digital future

For queries, contact Sub-Committee Lead:

Holly Harris
Policy Coordinator
BritishAmerican Business
hharris@babinc.org

Proudly part of

BritishAmerican Business is the leading transatlantic trade association incorporating the British-American Chamber of Commerce in the US and the American Chamber of Commerce in the UK.

We are committed to strengthening the economic corridor between the United States and the United Kingdom by supporting policies and actions that protect and enhance the environment for transatlantic trade and investment on behalf of our members.

We convene and serve a growing network of companies and business leaders through networking opportunities, bespoke programming and marketing platforms.

We actively promote trade and investment and support those who make the transatlantic corridor part of their business growth ambition.

For more information, please visit www.babinc.org

October 2023
BritishAmerican Business
www.babinc.org