

AI SAFETY SUMMIT 2023

WHAT YOU NEED TO KNOW

November 2023

The AI Safety Summit saw the international community come together for the first time to grapple with what some view as the foremost challenge facing humanity in the coming years: generative AI. Lawmakers across the globe are desperately seeking to understand and effectively manage the potential risks of this technology, whilst not stifling innovation and the potentially revolutionary opportunities it can offer to human advancement.

UK Prime Minister Rishi Sunak aims to position the UK as the convener, pioneer and thought leader on frontier AI, building consensus about the serious risks if AI advancement is left unchecked. The Bletchley Declaration – signed by countries at the outset of the Summit – is seen as the first step towards that, arguing that any threats are “best addressed through international cooperation”. Underpinning this will be a new global hub – the AI Safety Institute – based in the UK and tasked with testing the safety of emerging types of AI. It’s notable that this has been backed by those leading AI companies and nations present, at least publicly. Sunak said that the declaration, the action on testing and a pledge to set up an international panel on risk would “tip the balance in favour of humanity.”

Against a backdrop of increasing geopolitical tensions, there was a risk that the Summit could pale into insignificance. As the dust settles, many may eventually decide that it did. But Sunak’s success in bringing together the United States, China and the EU as signatories of the world’s first AI declaration on Wednesday shouldn’t be underestimated. Though perhaps symbolic more than anything else, the outcome – preceded by them sharing a stage for the opening speeches – marked an important milestone and the first time the three real key players in this space have held such formal talks.

Though many other world leaders chose not to attend, including President Macron and Chancellor Scholz, (President Trudeau arrived via video link on Thursday), the presence of US Vice President Kamala Harris, including her bilateral meeting at Downing Street, was also a big win for the UK Government – the Biden Administration’s decision to announce its own Executive Order on AI ahead of the Summit aside.

That said, it is no secret that the United States remains dubious about China’s presence at the Summit given the ongoing tensions around the tech sector and the sanctions on the chips that will power the AI revolution. However, the UK Prime Minister has been clear that engagement with China on an issue of this magnitude is paramount to building consensus in the global industry. Whether Sunak can simultaneously view Beijing as a ‘systemic challenge’ whilst engaging on the risks of AI remains to be seen though; it will require an enormous level of political skill to maintain such an uneasy alliance, and his attention will soon have to turn back to domestic matters with a general election on the horizon.

The star of the show, however, was not a world leader, but Elon Musk. His 50-minute ‘in conversation’ session with Prime Minister Sunak on social media platform X concluded the two-day Summit. Musk’s assertion that AI will eventually render the workforce redundant may have been said in jest, but went to the heart of public and some political concern surrounding AI.

Overall, the real success of the Summit will be judged in its long-term outcomes. It has the potential, as Rishi Sunak hopes, to underpin a global drive towards the effective management and harnessing of a revolutionary technology. The fear will be that, as attention turns to other matters, nations will soon revert to going it alone. If the most pessimistic of analysts are to be believed, such an outcome could be catastrophic.

Geopolitical Context

When historians look back on this week, they are likely to describe it as the week where AI firmly entered the world of geopolitics.

This week alone we have seen the President of the United States issue an Executive Order on the Safe, Secure and Trustworthy Development and Use of Artificial Intelligence and the leaders of the G7 agree on International Guiding Principles on Artificial Intelligence. So how does the UK AI Summit fit within this geopolitical race to “lead” the conversation on AI?

The Washington Post’s commentary of the AI Summit described how the UK had “struggled for relevance in the post-Brexit world” and that the Summit is viewed by the wider world as an attempt by an unpopular Prime Minister to enhance his image and secure his legacy. The Canadian Innovation Minister, François-Philippe Champagne, sought to dampen expectation of the UK’s idea to launch a global network of AI researchers in the image of the Intergovernmental Panel on Climate Change by saying “I don’t think

anything could really replace Global Partnership on Artificial Intelligence.”

This view was rejected by Bruno Le Maire, the French Finance Minister, who was clear in saying that the Summit was a “key milestone in the definition of fair and effective regulation of artificial intelligence.” This upbeat tone was echoed by a senior UK Conservative MP who told EGA that “The UK’s leadership role in securing the signatures of 27 countries to the Bletchley Declaration, including China and the United States, should not be underestimated.”

Whether you believe that the AI Summit was an irrelevant distraction away from other international efforts on AI or instead proved that the importance of the UK’s soft power in the drive to establish political oversight on the future of AI, it is clear the geopolitical dial has shifted this week.

China’s Presence

China’s very presence at the AI Summit was ultimately a huge diplomatic win for the UK. Rishi Sunak has been under significant pressure from some in his own party to rescind China’s invite. But at the same time, many Western policymakers acknowledge, albeit not always publicly, that there can be no international agreement on frontier AI that excludes one of the world’s leading technological states.

Despite being shut out of day two’s talks on national security, the mood music from the Chinese delegation at the Summit was positive. In his opening speech, China’s Vice Minister of Science and Technology Wu Zhaohui told attendees “China is willing to engage on AI governance for the promotion of all mankind... that’s our objective.” Furthermore, Yi Zeng from the Chinese Academy of Sciences, said “So for that [convening the Summit] we should thank all together for the UK.”

Even the UK Prime Minister’s fiercest critics will be forced to admit this has been nothing short of a diplomatic coup for Sunak, getting 27 countries, including China and the US, to sign the Bletchley Declaration listing the opportunities, risks and need for global action on the most advanced forms of the technology.

Business Reaction

This Summit's focus has been on "Frontier AI" – nascent, cutting-edge systems far beyond the tools of today – but in the short-term, these require huge investment and research beyond all but the very deepest pockets.

And so, these were discussions about (and for) Big Tech. Representation for the startups driving the next wave of innovation was limited and civil society groups have also noted the absence of a voice for affected workers.

In the middle of a global expansion of new and exciting tech companies defining this territory, this Summit focused on the existing big players in political and corporate spheres.

And while BCS CEO Rashik Parmar noted vibrant and positive discussions, one smaller AI developer told us they view the Summit as "a lot of noise" with little direct relevance. Many across the sector will feel the same.

Secondly, while questions remain on the regulation of the UK's AI ecosystem, it is arguable this week

has lured action from a US administration keen to demonstrate leadership.

Joe Biden's Executive Order is more substantive than the Bletchley Declaration – The Register noted the latter's "contents hardly amount to concrete, practical commitments" – but coming alongside this week's conversations on shared priorities, these could represent initial steps towards improving regulatory convergence – a prospect to be welcomed by many.

Lastly, it has been yet another demonstration of Sunak's enthusiasm for the domestic industry, underlining his view of AI as a legacy issue. And despite the safety talk, Downing Street remains emphatically positive on tech.

Where the UK sector has a good story to tell about their presence in the UK, it is likely the Government will continue to be excited to hear it, to share it and promote it.

Labour's Approach to AI

With strong and sustained poll leads as well a series of significant by-election wins, the consensus view is that the Labour Party will form the Government after a General Election at some point in 2024.

Labour thinks AI could transform how public services are delivered. In their Industrial Strategy, they see it as a vehicle to boost productivity and grow the UK economy. The party's relatively nascent views on the technology also stress the need to address risks with AI and build public trust.

Shadow Secretary of State for Science, Innovation and Technology Peter Kyle said the UK is "way behind" in bringing forward the right checks and balances. He says the Government is slow off the mark to deliver tangible results on domestic regulation. He also wants more clarity on the UK's role in coordinating global AI policy development.

As a centre-left party with close links to trade unions, Labour makes clear that AI's impact on the workplace and the transition to new jobs and industries needs to be managed carefully. They are clear that a "wild west" shift without support for those affected will not be acceptable in the UK.

Kyle has called for the biggest companies to be mandated to carry out and report safety tests – something that wasn't agreed at the Summit. The Prime Minister said the Government will not "rush to regulate" this technology – this concerns Labour who want Britain to play a coordinating and central role in this space and would take more of an activist role if they were in government.

What Comes Next?

In her opening speech at Bletchley Park, UK Science and Technology Secretary Michelle Donelan confirmed that the event will not be a one-off, but evolve into a recurring forum to discuss risks arising from advanced AI technology. South Korea will play host to a virtual AI Summit in six months' time, with France hosting the next in-person event later in 2024.

Despite the positive mood music about continued collaboration and the Bletchley Declaration itself, there is a significant risk that nations will continue to opt for a unilateral approach, attempting to dominate the development of frontier AI through fear of being left behind. The US and China will invariably continue to chart their own courses in that respect. But that

doesn't mean that such global talking shops won't have a significant role to play. There is a lot to be said for getting the right people around the table, and articulating common challenges, threats and opportunities.

Elon Musk's doomsday predictions about AI will also be dismissed by many, but the point remains a valid fear – and one which will likely dominate the domestic politics of AI for decades to come.

Reaction from the United States

When President Biden dropped his Executive Order on AI on October 30th, he sent Vice President Kamala Harris to the UK AI Summit as his most passionate ambassador. In attending Sunak's AI Summit and announcing more details of the States' approach to building guardrails around AI during the Summit, the Biden Administration has deliberately worked to lead the international discourse on AI governance.

By actively shaping these early-stage discussions, the US seeks to secure its seat at the head of the table. The reaction from the United States has been mixed. The Biden Administration feels they have checked all the right boxes and showed up with a plan for the world to follow. The private sector and the likes of Altman, Musk and others who have been very outspoken about the dangers of AI feel they have been taken seriously and encourage the dialogue. However, the long-term goals of the Biden AI Doctrine lie beyond Prime Minister Sunak and instead on idea diplomacy in other nations.

Reaction from China

China reacted positively to the UK's invitation to the Summit. At the same time, it viewed the Summit as a platform to demonstrate its independence and responsibility as a great power in AI governance. State-run media China Daily said the UK "cannot be faulted for inviting China", given the country's leading role in developing AI, adding: "No matter how Western the summit might be, its organisers are not adopting an ostrich policy."

In turn, Vice Minister Wu Zhaohui's speech at the Summit was highly aligned with the Initiative, promoting global cooperation and an international mechanism with broad consensus. However, many have noted that the Bletchley Declaration adopted many key points from The Global Artificial Intelligence (AI) Governance Initiative, which was launched at the third Belt and Road Forum for International Cooperation in Beijing last month, and presents China's approach to AI development and governance.

Reaction from the European Union

EU representatives made clear the bloc would stick to its plan to develop binding rules in the EU AI Act, because public authorities are ultimately responsible for the safety and security of citizens. Their power to intervene is seen as both complementing and acting as a backstop to self-regulation. European Commission President Ursula von der Leyen saw no contradiction between regulation and innovation: “guardrails are not barriers”. There is still an ambition to agree the Act before the European Parliament elections in June 2024. A new EU AI Office is set to enforce these rules, and President von der Leyen said it would collaborate with AI Safety Institutes, although Spain’s AI minister floated the idea of an EU Safety Institute with a broader focus on emerging technologies.

Edelman Global Advisory



Mohammed Hussein
President, UK

edelmanglobaladvisory.com
EGAcemms@edelmanega.com