The Rt Hon the Prime Minister
10 Downing Street, London SW1A 2AA

CC:
The Rt Hon the Leader of the Opposition
The Rt Hon Secretary of State for Science, Innovation and Technology
The Shadow Secretary of State for Science, Innovation and Technology

Dear Prime Minister,

AI now poses substantial risks to our national security.

In February of this year, criminals used commercially available AI tools in cyber attacks on nine Mexican government agencies, exfiltrating 195 million taxpayer records. The UK's own AI Security Institute has found that Anthropic's latest model, Claude Mythos, can *"discover and exploit vulnerabilities autonomously – tasks that would take human professionals days of work"*. The offensive cyber capabilities of AI are rapidly progressing and are readily available to malicious actors.

Britain depends on the functioning of computer systems in our critical infrastructure. Disruptions to those systems can threaten our drinking water supply, electrical grid, hospitals and financial services. The consequences of disruption can be severe and widely felt: the 2017 WannaCry attack forced the cancellation of 19,000 NHS appointments.

The Government's Cyber Security and Resilience Bill is a welcome step, but it places the burden of security primarily on the operators of essential services. AI developers whose models are used in cyber attacks must share some of the responsibility for minimising the risks that they create.

Moreover, cyber attacks are just the beginning. The International AI Safety Report 2026 has found that current AI systems can already provide *"expert-level laboratory instructions"* for the creation of biological and chemical weapons. As AI capabilities advance, the range of potential catastrophic harms will expand. Many experts are warning of the risk of catastrophic loss of control if AI developers succeed in their efforts to create models that exceed human capabilities across all cognitive domains.

Under the existing legal framework, there is substantial doubt as to whether AI model developers can be held liable for even the most severe and foreseeable harms, including cases in which their models were clearly decisive in enabling the harm. As a result, AI developers face no meaningful financial penalties or other sanctions, even for serious and predictable harms to our national infrastructure.

To ensure our safety, the incentives of AI developers must be aligned with the public interest. Given the pace at which frontier AI capabilities are advancing, this work cannot wait. We urge you to introduce legislation that holds AI developers liable for critical harms caused by their models.

Yours sincerely,