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

CC:
The Rt Hon Leader of the Opposition
The Rt Hon Secretary of State for Science, Innovation and Technology
The Rt Hon 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 the personal data and tax records of 195 million citizens. 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"*. More advanced models continue to be developed and are becoming 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 broadly felt: the 2017 WannaCry attack forced the cancellation of 19,000 NHS appointments. The Government's forthcoming Cyber Security and Resilience Bill is a welcome step, but it places the burden on the operators of essential services to keep pace with dangerous AI-fuelled offensive capabilities that are rapidly developing over the course of months, not years. That burden cannot rest on the defenders of our critical infrastructure alone.

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 progress, 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 race 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. The result is that AI developers face no meaningful financial or other punitive consequences for foreseeable harms to British citizens, and their incentives are not aligned with the public interest.

To ensure our safety, AI developers must be held responsible when their models cause serious harm to British people or property. Given the pace at which frontier AI capabilities are advancing, this work cannot wait. We urge you to introduce legislation which holds AI developers liable for critical harms caused by their models.

Yours sincerely,