



Department for
Science, Innovation
& Technology

Rt Hon Michelle Donelan MP

Secretary of State for
Science, Innovation and
Technology

100 Parliament Street
London SW1A 2BQ

The Rt Hon the Baroness Stowell of
Beeston MBE,
Chair of the Select Committee on
Communications and Digital,
House of Lords,
London,
SW1A 0PW

E: secretary.state.donelan@dsit.gov.uk

W: www.gov.uk/dsit

2nd November 2023

Dear Baroness Stowell,

UK Frontier AI Taskforce Update

I wanted to write to update you on the transition of the UK's Frontier AI Taskforce into a more permanent structure as the AI Safety Institute.

AI Safety Institute

On 26th October, the Prime Minister announced the establishment of the AI Safety Institute. Today, the Government will publish the attached document outlining its functions and scope.

The Institute is the first state-backed organisation focused on advanced AI safety for the public interest. Its mission is to minimise surprise to the UK and humanity from rapid and unexpected advances in AI. It will work towards this by developing the sociotechnical infrastructure needed to understand the risks of advanced AI and enable its governance.

This mission stems from our conviction that Governments have a key role to play in providing publicly accountable evaluations of AI systems and supporting research. While developers of AI systems may undertake their own safety research, there is no common standard in quality or consistency. Additionally, only Governments can run evaluations on issues related to national security, as they require very sensitive knowledge and could have legal ramifications for private actors. Lastly, governments will only be able to develop effective policy and regulatory responses to AI if they understand the technology better than they do today. By building a body of evidence

on the risks from advanced AI, the Institute will lay the foundations for technically grounded international governance.

The Institute will focus on the most advanced current AI capabilities and any future developments - aiming to ensure that we are not caught off guard by progress at the frontier of AI in an area that is ripe with uncertainty. It will consider open-source systems as well as those deployed with various forms of access controls. Both AI safety and security are in scope.

The research of the AI Safety Institute will inform UK and international policymaking and provide technical tools for governance and regulation. Possible examples of such technical tools include secure methods to fine-tune systems with sensitive data, platforms to solicit collective input and participation in model training and risk assessment, or techniques to analyse training data for bias.

The Institute is not a regulator itself and will not determine government regulation. It will collaborate with existing organisations within government and the private sector to avoid duplication, ensuring that activity is both informing and complementing the UK's regulatory approach to AI as set out in the AI Regulation White Paper earlier this year. It will provide key foundational insights to our governance regime and be a key player in ensuring that we take an evidence-based, proportionate response to regulating the risks of AI.

The Institute will establish the UK as a global hub for safety research, deepening the UK's stake in this strategically important technology. By improving the safety of advanced AI, the Institute will pave the way for increased adoption of advanced AI in this country, so that the UK is well-placed to seize these benefits.

Yours sincerely,

A handwritten signature in black ink, reading "Michelle Donelan". The signature is fluid and cursive, with the first name "Michelle" written in a larger, more prominent script than the surname "Donelan".

Rt Hon Michelle Donelan MP

Secretary of State Science, Innovation and Technology