THE CENTRE FOR
LONG-TERM RESILIENCE

POLICY PROPOSALS
Artificial Intelligence

THE OPPORTUNITY TO TRANSFORM THE UK'S RESILIENCE TO EXTREME RISKS

September 2022

Artificial Intelligence (AI)

THE CENTRE FOR LONG-TERM RESILIENCE

Establish a world-leading, effective AI governance regime which enables the UK to harness the benefits of AI while mitigating its risks

Build high levels of AI expertise in the UK government to ensure AI capabilities, their impacts, risks and policy implications are well understood

- **Build up AI technical expertise within government departments and regulators:**
  - Create specific technical roles around AI in key departments, such as the Ministry of Defence and Information Commissioner's Office, and ensure these positions can be made attractive to experts.
  - Implement an AI and machine learning training programme for existing civil servants.

- **Make it easier for government to draw on outside expertise in AI, ethics and governance:**
  - Establish a secondment/fellowship programme (similar to TechCongress) placing experts in AI ethics and governance in relevant parts of government.

- **Support improved understanding across government of the potential risks and harms of evolving uses of AI:**
  - Develop capacity and infrastructure for monitoring progress in AI, including by collecting information on inputs such as data and compute, and tracking AI models affecting UK citizens and business.
  - Establish more coordinated AI foresight and horizon-scanning programmes across government which feed directly into policy and regulatory decisions.

Promote safe and trustworthy AI development and deployment via improved incentives, norms, processes, and governance structures

- **Invest in AI safety, security, and interpretability via academic and private sector research, and cultivate talent in these areas via student fellowships.**

- **Commit to building a thriving and effective AI assurance ecosystem in the UK**, in line with the recommendations of the CDEI's roadmap.

- **Commit to implementing the UK's regulatory regime for AI in line with the recommendations of the Office for AI's forthcoming whitepaper.**

- **Review and improve risk management processes used by government around the use of AI:**
  - Establish red-teaming exercises and throughout-lifetime stress-testing of AI systems used by government.
  - Explore use of industry best practices around risk management, including three lines of defence.

- **Commit to reviewing the use of AI in high-risk domains, including in weapons systems, nuclear command and control, and critical infrastructure**, with a view to implementing robust assurance mechanisms and potentially restricting use where risks outweigh benefit.

Facilitate cooperation across nations and sectors on shared AI ethics and governance challenges

- **Lead on establishing norms and standards for safe and responsible use of AI internationally:**
  - Facilitate international dialogue around standards for safe and responsible use of AI in warfare.
  - Promote technical standards for safe and responsible AI via bodies such as the OECD and UN.

- **Work with the private sector to boost cybersecurity resources and incentivise industry investment in safety.**

- **Collaborate with academia to more effectively verify and challenge the claims made by AI industry**, by supporting red-teaming efforts in academia and increasing computing power resources for researchers.

Dr Jess Whittlestone
Head of AI Policy,
The Centre for Long-Term Resilience
jess@longtermresilience.org