

EXTREME RISKS AND THE UK NATIONAL RESILIENCE STRATEGY

The threat of extreme risks

As Covid-19 has demonstrated all too clearly, we are currently living with an unsustainable level of vulnerability to extreme risks.

Extreme risks are high-impact threats with a global reach. Climate change is the one we all know about, but pandemics and risks from emerging technologies pose a similar and perhaps even graver threat.

- With regard to pandemics, Covid-19 has had devastating global effects, and considerably more severe naturally occurring pandemics are entirely plausible. We are also highly vulnerable to threats from bioweapons and accidental laboratory leaks, which risk even worse consequences.
- As for artificial intelligence ('AI'), the widespread deployment of emerging capabilities is likely to contribute to a range of extreme risks. Looking further forward, the potential future development of human-level AI could have globally catastrophic consequences, if it is not adequately aligned with the objectives and values of humans.

In his book *The Precipice*, Dr Toby Ord estimates the likelihood of the world experiencing an existential catastrophe over the next one hundred years as being as high as one in six. AI and pandemics are listed as the most likely threats to cause such a catastrophe.

Boosting resilience to extreme risks in the UK

The UK has made encouraging progress in this area in the last twelve months:

- Last year's **Integrated Review** committed to building the UK's resilience to "low probability, catastrophic-impact events".
- The Cabinet Office's Call for Evidence on the **National Resilience Strategy** in July 2021 also set a promising direction:
 - It recognised the need for a "more effective system for handling these complex risks".
 - The then Paymaster General, Penny Mordaunt, said: "We need to follow and understand the work of organisations like the Oxford Institute on the Future of Humanity, to develop our understanding of questions about the prospects of humanity against AI, hostile acts and natural disasters at a scale not seen before. Without a fundamental step change we simply won't be as safe in future as we have been in the past."
 - The **Select Committee on Risk Assessment and Risk Planning's report** set out what this "more effective system" might look like, calling for an "Office for Preparedness and Resilience" and a "newly created post of Government Chief Risk Officer". It also drew attention to shortcomings in the National Security Risk Assessment, which, it notes, neglects "risks which emerge over a longer timeframe or risks which have a low likelihood of occurring within the next two years."
- The Government's first **National AI Strategy** affirmed for the first time that it "takes the long term risk of non-aligned Artificial General Intelligence [...] seriously", as well as "risk, safety and national security concerns that must be considered here and now."

We welcome this significant progress, and look forward to the Government putting in place a risk management process that gives extreme risks the attention they deserve. A more robust risk management system will not only make the UK safer: it will also help achieve the Government's broader strategic objectives, including those set out in the Integrated Review, by ensuring that we can effectively identify, assess and mitigate the most serious risks we face.

The opportunity to transform the UK's resilience to extreme risks — and for the UK to play a global leadership role

There is a great deal that the UK can do domestically to boost its resilience to extreme risks. It is a natural area for the UK to be a global leader — it has world class extreme risks policy experts, including at [Cambridge University's Centre for the Study of Existential Risk](#) and [Oxford University's Future of Humanity Institute](#). Further, [The Centre for Long-Term Resilience's Future Proof](#) report contains costed, actionable recommendations for UK risk management, AI, biosecurity and research and development policy.

By putting extreme risks at the heart of the National Resilience Strategy, the UK will have the credibility to play a leading role in safeguarding the wider world against the most serious risks we face. Post-Covid, there may never be a better opportunity to build global resilience in this area.

To achieve this, we suggest a two-stage process:

STEP 1

First, the UK should put extreme risks management front and centre of the upcoming **National Resilience Strategy** by:

- **Expanding the timeline of the NSRA** well beyond two years to enable it to sufficiently capture extreme risks, and improving how it categorises, compares and communicates risks.
- **Implementing a 'three lines of defence' approach to Government risk management.** Each line of defence should be managed by a different Government body. It is promising that the 'three lines of defence' risk management model is already being explored across Whitehall. To implement it in a Government setting, there needs to be a single point of accountability for risk management in Government, akin to a 'Chief Risk Officer' (CRO). This is the second line of defence, which then holds departmental risk owners (the first line of defence) to account. A clear distinction between the second and third line (the independent review function) is also fundamental. See the Annex for details.
- **Boosting resilience to the full range of extreme risks we face.** For biological risks, this should include creating a National Centre for Biosecurity, as proposed by Oxford's [Future of Humanity Institute](#) and the [Joint Committee for National Security Strategy](#). For AI risks, the Government should build on the commitments made in its [National AI Strategy](#), including by improving foresight and progress tracking in AI development, and bringing in more AI technical expertise.

STEP 2:

Having boosted its domestic resilience to extreme risks, the UK should then become a global leader in this area by:

- **Calling on other countries to create their own 'three lines of defence' risk management structure,** and providing the blueprint to do so.
- **Encouraging annual meetings between Chief Risk Officers** from around the world, where countries share information and learn from each other's risk assessments. This would include exploring why some countries were better prepared for Covid-19 than others.
- **Developing a set of light structures and processes for UN member states to address extreme risks,** working initially with likeminded nations such as Australia, Japan and New Zealand. This could be a stepping stone on the way to creating a formal Treaty, with a series of UN Security Council resolutions to place the new framework on the strongest legal footing.

Dr Toby Ord

Senior Research Fellow
Future of Humanity Institute
University of Oxford

Dr Seán Ó hÉigearthaigh

Executive Director
Centre for the Study of Existential Risk
University of Cambridge

Mr Angus Mercer

Chief Executive
The Centre for Long-Term
Resilience

ANNEX

The Three Lines of Defence

The 'three lines of defence' risk management structure is a familiar method used in the private sector, including in finance. Applied to a Government setting, it aims to provide an overarching structure to drive risk management improvements through Government on an ongoing basis, and address three key issues:

- Extreme risks being deprioritised because of the tyranny of the urgent, the current two-year time horizon of the NSRA, and potential 'panic/neglect cycles' which follow crises;
- A lack of checks and balances for risk management, where risk owners are not held to account to mitigate or plan for their risks;
- Unclear accountabilities — including no single point of accountability for risk management — which might lead to inaction in the crucial early days of an emerging extreme risk event.

a) First line of defence: Government departments

Government departments are, and should be, responsible for the day-to-day management of risks relevant to their department, including extreme risks. In some cases, additional support and technical expertise is needed.

b) Second line of defence: Office of Risk Management, led by a Government Chief Risk Officer-type role

The CRO would provide oversight and a single point of accountability for ensuring the proper management of risks and vulnerabilities across Government. They would be empowered to assign responsibility for risks to ministers and hold them to account for their risk response strategy. They would also play a leadership role in driving risk management improvements across Government and in ensuring that risk management improves globally.

The CRO role is not designed to provide risk ownership or expertise in particular risk areas, just as a conductor is not expected to moonlight as a soloist. Where the 'three lines of defence' is deployed in the private sector, a CRO is almost invariably included in the second line.

c) Third line of defence: National Extreme Risks Institute

This plays an independent audit and advisory function, submitting its recommendations on extreme risks to the Office for Risk Management and its Chief Risk Officer. It provides independent advice and reviews the Government's approach to preparing for and mitigating extreme risks.

It also provides a role for independent experts to effectively provide a second opinion on the state of Government risk management. Such an audit role was recommended by the Select Committee on Risk Assessment and Risk Planning.

An Oversight Committee would bring the three lines of risk infrastructure together and avoid siloism. The Chief Risk Officer would report to its Chair.

Source: Ord, T. (2021): "[Proposal for a New 'Three Lines of Defence' Approach to UK Risk Management](#)", Extreme Risks Working Paper 2021, Future of Humanity Institute, University of Oxford.