THE CENTRE FOR
LONG-TERM RESILIENCE

CLTR 2022 Annual Report

Summary

Mission

Our mission is to transform global resilience to extreme risks.[1] Our focus is on improving governments' resilience to the extreme risks posed by artificial intelligence (AI) and pandemics — in the UK, in particular. We do this by working with governments and other institutions to improve relevant governance, processes, and decision-making around these risks.

Progress in 2022

In 2022, we made progress across our four key pathways to impact:[2]

1. **Government influence**
    - Providing feedback on elements of an early draft of the **Defence AI Strategy** to emphasise safety and responsibility, with several suggestions reflected in the strategy, including a box outlining what the Ministry of Defence (MoD) expects of a responsible AI supplier, which we wrote with academic experts in AI risk from Cambridge University's Centre for the Study of Existential Risk (CSER);
    - Providing ongoing advice on the **UK's approach to AI regulation** via written and oral input, a 'tech-in' session, and a written submission to the Office for AI;
    - Providing recommendations for the **UK's National Resilience Framework.** The published Framework reflects a number of our recommendations in whole or in part, including appointing a new government Head of Resilience, a National Security Council resilience sub-committee, and increasing the time horizon for the National Security Risk Assessment (NSRA).

2. **Policy development**
    - **Developing compute governance policy recommendations**, as part of our evidence submission to the **Future of Compute Review**;
    - **Developing recommendations for how AI regulation can keep pace with rapid AI progress, how defence can integrate AI safely**, and **how the UK can demonstrate international leadership in safe and responsible use of AI**;
    - **Developing an approach and tool for risk management in the UK government**, **presented** to the Resilience Directorate at the Cabinet Office.

[1] High impact threats with global reach, from a Covid-19-type event up to an existential risk.
[2] Much of our biosecurity policy work in 2022 has been sensitive in nature, or conducted under the 'Chatham House' rule. As such, the report provides a less detailed account of our activities focused on biosecurity.

3. **Mainstreaming extreme risks**
    - Contributing to the **Institute for Government's** first-ever **report on managing extreme risks**. We have developed a partnership with the Institute for Government, including jointly arranging and speaking on a **panel event on 'strengthening the UK's resilience'**, alongside Fleur Anderson MP and Rt Hon Oliver Letwin;
    - **Helping three House of Lords members establish a UK Resilience Forum** aimed at securing the UK government's implementation of the Select Committee's recommendations.

4. **Boosting expert capacity to work with governments on extreme risks**
    - Facilitating and supporting **two academic placements** to support the **MoD's work on safe and responsible AI**, as well as helping support **two academic placements** to work on **AI regulation in the Cabinet Office**;
    - Sharing lessons learned and **advice on the risks involved in extreme risks policy work**.

We've also focused on **strengthening our operations for our growing team**. This includes ensuring smooth onboarding and support for team members, updating our organisational values and building a strong team culture, setting up our KPIs and a KPI management system, and initiating changes to strengthen our governance structures among our leadership team and board.

Team & Growth Plans

In 2022, we grew from a team of three to a team of seven, onboarding:

- Dr Jess Whittlestone as Head of AI Policy;
- James Ginns as Head of Risk Management Policy;
- Sophie Rose as Biosecurity Policy Manager.

We also created an External Affairs unit, led by our Director Sophie Dannreuther, focused on mainstreaming extreme risks among relevant stakeholders, expanding and diversifying our government and expert network, and managing our external communications. We also onboarded Emma Lawsen as our Team Executive Assistant.

In 2023, we plan on expanding to a 11-person team, by building out each policy unit (AI policy, biosecurity policy, and risk management policy) with an additional team member. By the end of 2025, we hope to have expanded the above units to three people each, and add an additional Team Executive Assistant, which would take us to 15 people. However, the expansion to 15 is dependent on securing sufficient additional funding.

Overview of this report

1. Who we are
2. Progress
    Pillar 1: Government influence
    Pillar 2: Policy development
    Pillar 3: Mainstreaming extreme risks
    Pillar 4: Boosting expert capacity to work with governments on extreme risks
    Pillar 5: Strong organisational health
3. Priorities for 2023
4. Funding
Annex 1 – Expert network
Annex 2 – Overview of our 2022 Spending

1. Who we are

Our long-term vision is a world with high resilience to extreme risks, such as those from pandemics and AI.

Our mission is to transform global resilience to extreme risks[3]. Our focus is on improving governments' resilience to the extreme risks posed by artificial intelligence (AI) and pandemics — in the UK, in particular. We do this by working with governments and other institutions to improve relevant governance, processes, and decision-making.

Theory of change

Our theory of change remains largely the same as in previous years:

Research and influence of AI risk, biosecurity, and risk management experts, combined with an **understanding of how to get things done in government.**

↓

Concrete, actionable **policy advice**, designed in collaboration with governments and other actors, along with **mainstreaming** of extreme risks as an issue area within policy circles, and **improved expert capacity** to work with governments on these issues.

↓

High-impact actions taken by decision-makers in government (and other institutions where necessary, e.g. AI companies) to **boost resilience to these risks.**

The assumptions underpinning our theory of change are as follows:

- **Governments, international institutions, and leading private actors** (such as cutting-edge AI and bio labs) **have significant scope to increase or reduce resilience to extreme risks.** These actors have the influence, funding, and regulatory power to have a meaningful impact on the world's ability to mitigate and handle extreme risks.

- **The UK can become one of the most influential actors** in building resilience to extreme risks. This is largely as a result of its geopolitical influence, the size of its government budget, and the fact that leading research institutions and one of the world's leading AI labs are located there. Additionally, the UK has close ties to the US, allowing for policy diffusion and agenda setting in Washington. Furthermore, the UK has significant international standing, soft power, and diplomatic clout, especially with the EU, the US, and across the Anglosphere. There is potential for UK policies to be adopted by other countries. Finally, the UK has a commitment towards evidence-based policy and has stated an ambition to become a technological superpower.

[3] High-impact threats with global reach, from a Covid-19-type event up to an existential risk.

- **There is significant potential for leading academic and technical experts in extreme risks fields such as AI safety, biosecurity, and risk management[4] to have a much greater impact on UK and global policy than is currently the case.** These issues tend not to get the attention they deserve from the UK government, or internationally. There are a number of reasons for this, including that academic outputs are commonly not tailored to policy audiences and often need translation into concrete recommendations, the presence of incentives for policymakers to focus on more urgent and high-profile but less-important issues, and the tendency of even senior policymakers to see the mitigation of such extreme risks as overly difficult and 'beyond their pay grade.'[5]

- **There are three main obstacles to such experts having policy impact:**
    1. They need to **dedicate the time to think about** what policy recommendations they would most like to see implemented, and establish **agreement** with others in their field (including those with policy expertise) about this.
    2. Their thinking needs to be distilled into **concrete, actionable, concise, and red-teamed recommendations** that are accessible to time-poor senior government decision-makers and advisers.
    3. These actionable recommendations then need to be communicated to the **right people in government**, by the **right supporters**, at the **right time**, and in the **right way**.

- **An organisation that is sufficiently well connected to (and trusted by) these expert communities and has sufficient expertise and networks in the UK and international policymaking processes can overcome these three obstacles**, and thereby unlock greater policy impact.

This year, we analysed our specific pathways to impact and concluded that these can be subdivided into the following five pillars:

1. **Pillar 1: Government influence** – including providing expert advice and red-teaming on policy decisions, and convening cross-sector conversations and workshops related to extreme risks to produce and convey effective, actionable policy recommendations;
2. **Pillar 2: Policy development** – including developing policy recommendations and risk management frameworks together with experts and government stakeholders;
3. **Pillar 3: Mainstreaming extreme risks** – including helping decision-makers and the wider public to understand extreme risks and their importance;
4. **Pillar 4: Boosting expert capacity to work with governments on extreme risks** – including providing an exchange for specialist knowledge and facilitating expert placements into government;
5. **Pillar 5: Strong organisational health** – including continuing to build a well-functioning organisation as we grow (e.g. by developing our team culture, organisational and KPIs infrastructure, and team development and well-being) to ensure our ability to effectively achieve impact in the years ahead.

[4] Hereinafter referred to collectively as 'extreme risks experts' for ease of reference.
[5] For further reasons, see Annex A of Sam Hilton and Caroline Baylon's report on Risk Management in the UK for a summary of interviews with 12 UK civil servants on this topic.

Focus areas & scope

We focus on three core areas of extreme risks and resilience:

1. AI;
2. Biosecurity; and
3. Risk management.

Within each of these areas, we try to focus on projects that, as well as being important, tractable, and neglected areas, are also sufficiently low-risk and high-consensus to work on.

Our work currently focuses on the UK, as this is where we have a comparative advantage through our network and understanding of the UK government.

We occasionally respond to outstanding opportunities for impact at an international level, through channels like the OECD, the World Economic Forum, or the Biological Weapons Convention.

Going forward, we may also respond to outstanding opportunities in other jurisdictions, though we have not done this so far. For now, we see our route to global impact as being primarily through the UK government improving its own resilience to extreme risks, and being a global leader in this area by setting an international standard.

How we define success

Our ultimate aim is to achieve policy change to strengthen global resilience to extreme risks. This can be achieved through directly influencing policy, or indirectly through mainstreaming extreme risks and supporting extreme risk experts to work on policy-related issues.

We define success as having **counterfactually contributed to policy change that substantially reduces risks** from advanced AI or biological threats, or reduces the catastrophic consequences of such events.

This could entail:

- Adopting the **'Three Lines of Defence'** risk management model in UK government departments and/or AI/bio labs;
- Running regular **risk identification workshops** within UK government departments, with **longer timelines of risk assessments** and risk registers **assessing both the impact of risks and vulnerability to risks**;
- Adopting a **Chief Risk Officer** for the UK government;
- Establishing **effective incentives, norms and governance structures** within the UK government that reduce the likelihood any actor develops and deploys an AI system or synthetic pathogens that could cause mass harm;
- Enabling and supporting the **UK to play a leading role** in setting **international norms and standards for AI governance**;

- Taking **concrete steps** to help safeguard the UK against risks arising from naturally-occurring pandemics, laboratory leaks, deliberate biological threats (e.g. bioweapons use), and 'dual-use' research (advances that can be used for harm as well as good);
- Significantly **boosting funding and a clear division of responsibility for UK and international biosurveillance**.

Progress towards achieving such policy changes can also look like:

- **Helping decision-makers** and the wider public **understand extreme risks and what can be done** about them;
- **Providing expert advice** and **red-teaming** on policy development and decisions;
- **Convening cross-sector conversations** and workshops related to extreme risks;
- **Developing and advocating for policy recommendations** and effective **risk management frameworks** and systems;
- Providing an **exchange for specialist knowledge**, including by **facilitating expert placements** into government;
- Building and maintaining **relationships** with extreme risk **experts**;
- Deepening and broadening our **network of advisers and decision-makers** in government and **influential think tanks**.

Policy processes are often long, and outcomes from our work can only become apparent years later. As such, we evaluate our work using both 'lead' and 'lag' indicators. In our Progress section below we list the progress achieved during 2022, which range from work we expect will yield real-world change in the coming years to more incremental steps towards achieving such outcomes.

How we manage risks from our work

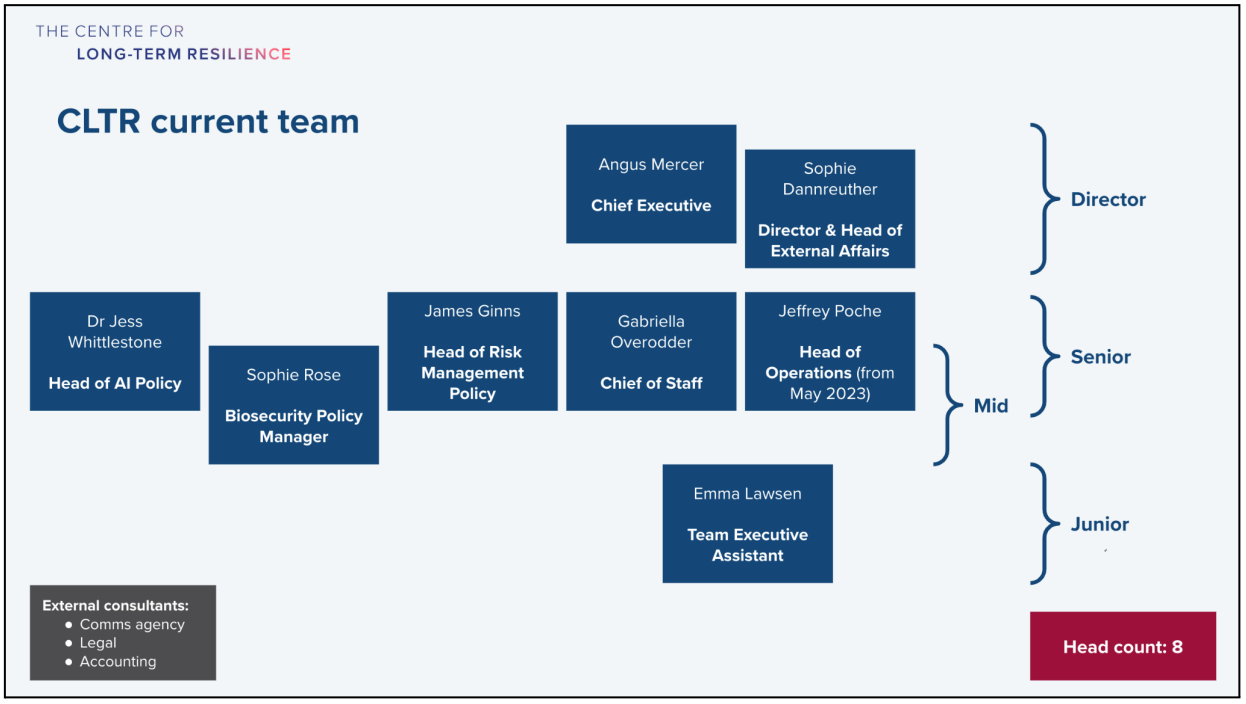
Working on extreme risk policy requires a careful, thoughtful approach to avoid accidentally doing harm. We recognise and manage these risks in our strategy and day-to-day operations:

- **To avoid putting ideas into the world that could inadvertently cause harm**, we have hired in-house experts in each focus area and work closely with our increasingly broad network of experts to co-create policy recommendations. If there isn't enough consensus on a question we're exploring, we avoid making recommendations to government;
- **To avoid politicising extreme risks**, we primarily work with non-political civil servants and strive to maintain balanced engagement across the political spectrum;
- **To mitigate reputational risks to CLTR and the wider extreme risks community**, we strive to conduct our work with high standards of quality, integrity, and transparency. Whenever we can, we also share information on extreme risks and our work more widely via our website, Twitter, and our broader media engagement;
- **To avoid having negative or no counterfactual impact**, we coordinate with other organisations to identify where we are best placed to invest our time, and regularly evaluate our impact and seek feedback from the experts and government stakeholders with whom we work most closely.

Team

In 2022, we expanded from a three-person team to a seven-person team, building out units for AI Policy, Biosecurity Policy, Risk Management Policy, and External Affairs. We also hired a Team Executive Assistant, and an Operations Contractor (who, visa permitting, will join us as a full-time employee in May).

We now have the following team:



New joiners in 2022

Dr Jess Whittlestone joined in March as CLTR's Head of AI Policy. Jess brings a wealth of expertise in the long-term impacts and governance of AI, and in working with governments to translate academic insights into policy recommendations. Prior to this, she was a Senior Research Fellow and Deputy Director of the AI: Futures and Responsibility Programme at the Centre for the Study of Existential Risk and the Leverhulme Centre for the Future of Intelligence, both at the University of Cambridge. Jess has published widely on subjects related to AI policy and provided expert advice to a number of government departments, including the Centre for Data Ethics and Innovation and Office for AI. Jess holds a PhD in Behavioural Science from the University of Warwick, and a first class degree in Mathematics and Philosophy from Oxford University.

James Ginns joined in April as CLTR's Head of Risk Management Policy. James is a leading expert in risk management, having served as a Chief Risk Officer in the aviation sector for a number of years. Prior to joining CLTR, James provided valuable contributions to Toby Ord's paper on how the UK could transform its approach to risk management with the adoption of a Three Lines of Defence approach, and the appointment of a Chief Risk Officer.

Sophie Rose joined in June after finishing an MHS in Infectious Disease Epidemiology at Johns Hopkins Bloomberg School of Public Health and a research fellowship at the Center for Health Security. Sophie's previous work focused on biosurveillance tools and strategy and their applications for mitigating the risk posed by high-consequence emerging infectious diseases. Sophie also co-founded 1 Day Sooner, an organisation born in the COVID-19 pandemic to represent volunteers who wanted to rapidly accelerate the deployment of the most effective COVID-19 vaccines possible. She is currently an 'Ending Bioweapons' fellow at the Council on Strategic Risks and an incoming Emerging Leaders in Biosecurity Initiative (ELBI) fellow at the Center for Health Security.

Emma Lawsen joined in August as CLTR's Team Executive Assistant. Emma is responsible for the smooth running of our day-to-day administrative operations, and assists unit leads on research tasks as a research assistant. Prior to this, Emma worked as a teacher for eight years, including responsibilities for several whole-school programmes. Emma holds an MPhil in Social Anthropology from the University of Oxford.

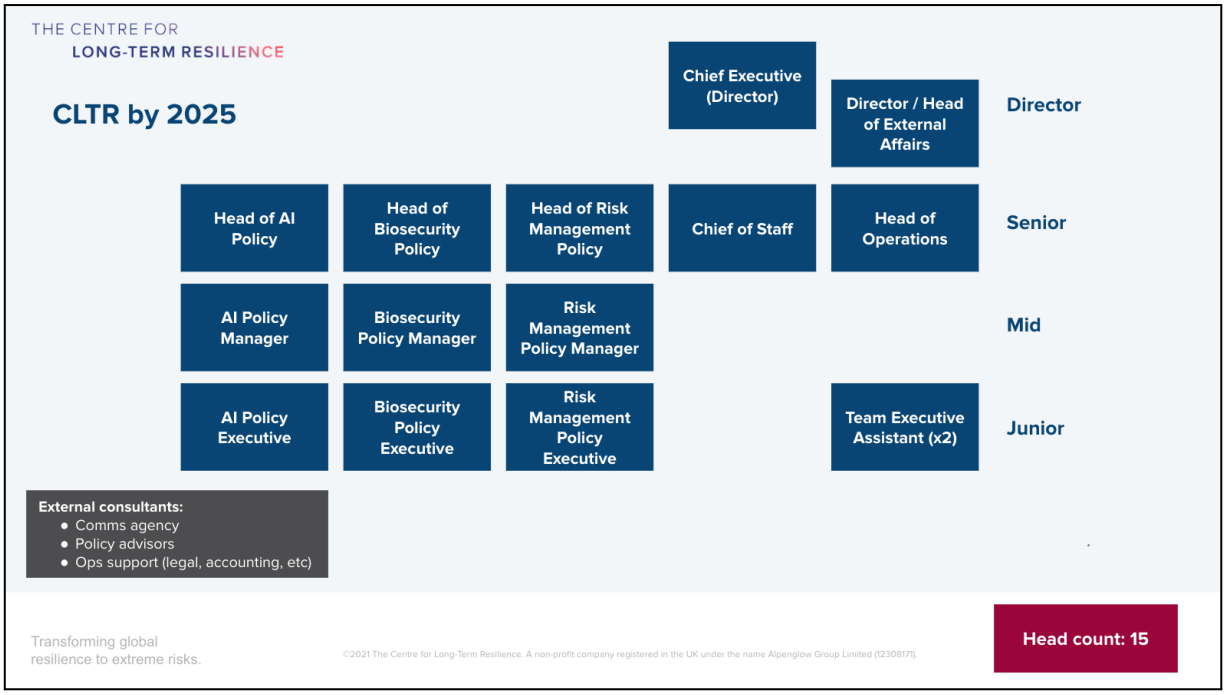
New joiners in 2023

Jeffrey Poche will join us in May as CLTR's Head of Operations (visa permitting). Jeffrey will lead on building CLTR's infrastructure for a growing team and oversee our operations. Jeffrey has extensive experience in operations management, previously in roles as Chief Operating Officer at Fountain Forward and Waste Management, as well as a freelance operations consultant. Most recently, Jeffrey founded Anti Entropy, an operations consultancy providing operations support and training.

We also work with extreme risk experts as contractors, and plan on continuing to do this in 2023.

Expansion plans

In 2023, we plan to expand to an 11-person team by building out each policy unit (AI policy, biosecurity policy, and risk management policy) with an additional team member. By the end of 2025, we hope to have expanded the above units to three people each, and add an additional Team Executive Assistant, which would take us to 15 people. However, this is dependent on securing sufficient additional funding and demand for our work.



Board

Our board of directors consists of:

- Cassidy Nelson, University of Oxford (chair);
- Niel Bowerman, 80,000 Hours;
- Sophie Dannreuther (full-time staff);
- Angus Mercer (full-time staff).

In December 2022 Jade Leung (who works at OpenAI) stepped down from the Board. Since our AI work will increasingly focus on the issue of AI regulation, the board agreed it would be a good idea to pre-empt any conflict, or perception of a conflict, in this area by accepting Jade's offer to step down.

2. Progress

Pillar 1: Government influence

Successes in 2022:

- **Providing feedback** on elements of an early draft of the **Defence AI Strategy** to emphasise safety and responsibility, with **several suggestions reflected in the strategy, including a box outlining what the MoD expects of a responsible AI supplier**, which we wrote with academic experts in AI risk from Cambridge University's Centre for the Study of Existential Risk (CSER);
- **Ongoing advice on the UK's approach to AI regulation**, including via a 'teach-in' session and a written submission to the **Office for AI**;
- Providing regular **evidence and briefings to the House of Lords Risk Assessment and Risk Planning Committee's inquiry on extreme risks**. This likely helped result in the **inclusion of several low-probability, high-impact events in the Register**, as well as **longer time periods for measuring risks than the usual two years** (up to five years);
- Providing recommendations for the **UK's National Resilience Framework. The published Framework reflects a number of our recommendations** in whole or in part, including **appointing a new government Head of Resilience, a National Security Council resilience sub-committee, and increasing the time horizon for the National Security Risk Assessment (NSRA).**

AI

Recommending and helping organise a conference on AI safety between multiple teams at the Ministry of Defence and technical AI safety experts, hosted by CSER

Together with the Centre for the Study of Existential Risk (CSER) in its partnership with the Ministry of Defence (MoD), we helped organise an AI ethics workshop in March 2022 for policymakers and AI technical experts. We suggested the original idea of a workshop to the MoD.

Providing feedback on elements of an early draft of the Defence AI strategy to emphasise safety and responsibility

After presenting AI safety policy recommendations to the MoD, we were asked to provide feedback on elements of an early draft of the UK government's draft Defence AI Strategy.

Evidence submission to the Future of Compute Review

We submitted evidence to the Future of Compute Review that emphasises the importance of compute as a lever for shaping beneficial AI progress. Our submission recommends that the review should (i) explore ways to increase compute capacity for academia, and (ii) consider how the UK government could monitor and govern compute usage.

Ongoing advice on the UK's approach to AI regulation

We provided oral and written input in response to the Office for AI's policy paper on AI regulation. Our written submission emphasises the importance of ensuring the regulatory approach can adapt to AI progress such as in foundation models, and of having a strong broader governance ecosystem to support the regulation. As a follow-up, we ran a "teach-in" session with the Office for AI on how progress in foundation models may challenge the proposed regulatory approach, and how the forthcoming White Paper can address these challenges. Initial feedback on the session was very positive and we plan to follow up on next steps in early 2023.

Advice to the FCDO on mitigating long-term risks of AI

We had several conversations with senior policymakers at the Foreign, Commonwealth and Development Office (FCDO) to provide advice on how the department can understand and mitigate long-term risks from AI, and we participated in a roundtable on this topic.

Ongoing advice to the MoD on internal AI safety and assurance processes as well as how the UK can engage internationally on these topics

We wrote a short policy briefing note on AI and international stability. We are also beginning to discuss how the UK should engage internationally to promote best practice and norms in this space, and reduce international security risks from AI.

Evidence submission to the House of Commons Science and Technology Committee's Governance of AI Inquiry

Our evidence submission makes three key recommendations for strengthening the UK's approach to AI governance: (i) Ensure that the UK's regulatory regime is supported by a strong broader governance ecosystem which can provide expertise and identify and address regulatory gaps; (ii) Consider that more cross-cutting and anticipatory regulation may be needed to mitigate extreme risks, especially as capabilities advance; (iii) Engage internationally to mitigate shared risks from AI.

Biosecurity

Much of our biosecurity policy work in 2022 has been sensitive in nature, or conducted under the 'Chatham House' rule. As such, this section provides a less detailed account of our activities.

If you have questions about our biosecurity work, please contact info@longtermresilience.org.

Biosurveillance value briefing for UK Health Security Agency

In late 2022, our biosecurity unit accepted a request from a Deputy Director at the UKHSA to brief a team of ~150 government officials (including the UKHSA, MoD, the Department for Environment, Food & Rural Affairs [DEFRA], the Animal and Plant Health Agency [APHA], the National Security Secretariat and others) on the current status of other countries' biosurveillance infrastructure as part of their initiative to explore options for the delivery of a national biosurveillance system.

Risk Management

Calling for expanding the scope of the UK National Risk Register

In 2021, the House of Lords Committee on Risk Assessment and Risk Planning published a report focused on improving the UK government's management of extreme risks, and the Royal Academy of Engineers conducted an external review of the Cabinet Office's National Security Risk Assessment (NSRA).

The government's official response in March 2022 commits to include several low-probability, high-impact events in the 2022 National Risk Register. Where appropriate, these risks will be measured over longer time periods than the usual two years (up to five years).

Briefing the COO of the Civil Service and head of Civil Contingencies Secretariat, which generated a request to run a crisis workshop with senior policymakers

To counter any perception that there is little that can be done domestically to tackle extreme risks, we prepared a policy briefing which made the case for UK leadership in building resilience to extreme risks. We shared this briefing with decision-makers in government and secured a meeting with senior civil servants within the Cabinet Office.

In the meeting, we were asked to arrange a crisis workshop for the Cabinet Office, which didn't happen in the end due to team changes.

Split of the Civil Contingencies Secretariat (CCS) into a longer-term resilience function (Resilience Directorate) and a crisis response function (COBR Unit)

In July it was announced that the CCS had been split into a team focusing on longer-term resilience and a team dedicated to crisis response. We had advocated for this to happen as part of an enhanced risk management framework over the previous 18 months. Resilience is now covered by a dedicated Directorate within the Cabinet Office. This should help ensure that it receives appropriate focus and that it isn't deprioritised by short-term crises.

'Lunch and Learn' briefing for the Cabinet Office's Resilience Directorate

We held a briefing for the new Resilience Directorate at the Cabinet Office in November, in a hybrid format attended by 45 people in the team, and covered both extreme risks and a

proposed cross-sector approach to managing extreme risks. Following this, we've had several conversations with Cabinet Office officials working on resilience and crisis management.

Submission to the Foreign Affairs Committee's (FAC's) inquiry on the update to the Integrated Review (IR)

We submitted a response to the FAC's inquiry on the ongoing update to the Integrated Review (IR). This provided an opportunity to reiterate the emphasis of the original IR on the importance of enhancing resilience to catastrophic risks, to continue to press for publication of a fully-fledged National Resilience Strategy which was promised in the original IR, and to encourage the application of private sector risk management best practice to the updated IR to ensure it stays adaptive and forward-looking as the international security environment changes. The FAC accepted and published our response, referenced CLTR twice in notes to its report, and recommended both a national resilience lead and regular cross-government meetings to review shared efforts to improve resilience to threats across all policy areas.

Publication of the UK's National Resilience Framework

The National Resilience Framework was published in December 2022. CLTR's work was listed as a source for evidence considered in formulating policy. The Framework is a welcome first step to enhance the UK's preparedness and capability to prevent as well as respond if necessary to a growing array of risks.

It reflects our recommendations for improved risk governance and for a single point of accountability for risk management in that it announced a new government Head of Resilience, a National Security Council resilience sub-committee, and an emphasis on clear ownership of all risks.

It also includes our recommendation for an increased time horizon for the National Security Risk Assessment (NSRA), which has been increased from two to five years for some risks. Further development is needed, but it represents progress towards building resilience which we had a role in helping to achieve. We subsequently published a response to the framework.

On January 12 2023, the House of Lords convened to debate topics such as the government's response to a House of Lords Committee report ('Preparing for Extreme Risks: Building a Resilient Society') and the government's publication of the National Resilience Framework. Lord Des Browne, member of the House of Lord's Risk Assessment and Risk Planning Committee, stated:

"While supply chains, global context and societal make up ensure that different nations must mitigate risk in different ways and with different emphases, universal challenges show us that preventive work, on which the framework places a welcome emphasis, can work properly only through international co-operation. I urge reflection on the response of the Centre for Long-Term Resilience to the framework. While commending it, it asks that, in recognition of the global nature of the threats, the Government advocate for a dedicated multinational resilience forum for greater coherence of the efforts of individual nations to protect their people."

External Affairs

Providing input and encouragement to HMG efforts to strengthen resilience

CLTR Director Sophie Dannreuther was a member of the Expert Advisory Group for a GO-Science Resilience Project, which involved providing evidence about future trends and threats. As part of this, she met a number of GO-Science policy advisers to brief them on CLTR's work.

Briefings and training sessions for parliamentarians and policymakers on our work

We organised an event on long-term resilience for a cross-government network, focusing on biosecurity threats and managing risks. The event had around 50 attendees, and feedback demonstrated a demand for development opportunities around risk management, and encouraged us to continue offering learning sessions to relevant HMG teams.

Government network

Feedback from government stakeholders

> *"Very positive experience of working with CLTR - professional, knowledgeable and easy to deal with. Has become one of my 'go to' collaborators. Please keep up the great work!"*
>
> **– UK government official**

> *"As a busy policymaker working in a highly stretched team in UK government, I really value the work of CLTR and knowing that I can ask a policy relevant question and either get an informed answer or a recommendation of useful contacts, in a timely manner."*
>
> **– Civil servant working on biosecurity in the UK government**

Other work/achievements:

- We **submitted a proposal** outlining **'Technology to Counter Biological Threats'** to the UK's **Advanced Research & Invention Agency (ARIA)** open-call for thematic programme recommendations;
- Our Biosecurity Policy Manager, Sophie Rose, **participated in a Scientists for Labour panel event on pandemic preparedness and recommendations for bolstering our capacity** in this area;
- Our Head of External Affairs attended **Conservative and Labour Party conferences** to pursue opportunities to engage both parties on the subject of boosting resilience to extreme risks.

Pillar 2: Policy development

Successes in 2022:

- Developing **compute governance policy recommendations**;
- Developing **recommendations for how AI regulation can keep pace with rapid AI progress; how defence can integrate AI safely**; and **how the UK can demonstrate international leadership in safe and responsible use of AI**;
- **Developing an approach and tool for risk management in the UK government, presented** at a 'Lunch and Learn' to the Resilience Directorate at the Cabinet Office.

AI

Developing compute governance policy recommendations

We have been exploring how to make policy recommendations related to compute governance more actionable in the UK context, including through conversations with academic researchers working on compute governance, and conversations with relevant parts of the UK government (including members of the Future of Compute Review team, the Competition and Markets Authority and newly-proposed Digital Markets Unit within it, and the Centre for Data Ethics and Innovation). Though we still have some work to do in developing these policy recommendations, we now have a much clearer idea of what a compute monitoring pilot within the UK government could look like and some options for where it could sit.

Other AI policy development

Our engagement and work advising various parts of government on AI policy — particularly on AI regulation, defence, and foreign policy — has enabled us to explore and develop clearer policy recommendations in all of these areas. This includes recommendations for how AI regulation can keep pace with rapid AI progress; how defence can integrate AI safely; and how the UK can demonstrate international leadership in safe and responsible use of AI.

Risk Management

A series of workshops at CLTR with Beehive, an external partner, to provide an approach to risk management process and begin to develop an alternative extreme risk register

We ran a series of three internal workshops to identify, define, assess, and score AI and biosecurity risks, involving scenarios and vulnerability assessment for each risk. This was backed by use of Beehive's automated Risk360 tool to produce and present combined scores across all dimensions of impact and vulnerability for each scenario and an associated heatmap. This can be used to define further mitigations, actions, and owners. The output of the workshops was included in the material we presented at the recent Cabinet Office 'Lunch and Learn.'

Expert network

We believe that extreme risks pose some of the biggest challenges of the 21st century, and want to work with a wide range of people and organisations who share our goals. As a result, we aim to collaborate with a broad coalition of policymakers, academics, think tanks, and private-sector risk experts, to ensure that our policy recommendations are based on a high degree of consensus from a diverse expert community.

Since our policy leads joined in 2022, we have increased our focus on policy development within biosecurity, AI, and risk management. In 2023, we hope to further grow and diversify our expert network to strengthen our policy development work.

Please find a list of institutions we work with in Annex 1 below.

Other work/achievements:

- We developed and published concrete **policy proposals on AI, biosecurity, and risk management**. These have been widely useful, including to introduce our work to new stakeholders and to start conversations about **political parties' manifestos**;
- We completed **scoping a project to update our policy recommendations regarding the UK's biosafety practice,** including substantive legislative mapping and conducting interviews with officials at UK Health Security Agency and the Health and Safety Executive**;**
- We began **scoping a project to produce a historical biosecurity funding overview for the UK** — but ultimately decided this was currently intractable for our unit to take on given lack of detailed, publicly-available funding information;
- We **began developing concept notes on dual-use research of concern governance and 'deterrence by denial'**, which we plan to continue developing with a view towards updating our policy recommendations more broadly;
- **Our Biosecurity Policy Manager, Sophie Rose, attended the Biological Weapons Convention** Ninth Annual Review Conference (RevCon).

Pillar 3: Mainstreaming extreme risks

Successes in 2022:

- **Helping three House of Lords members establish a UK Resilience Forum** aimed at securing the UK government's implementation of the Select Committee's recommendations;
- Contributing to the **Institute for Government's (IfG's) first-ever report on managing extreme risks**;
- Developed a partnership with the IfG and **spoke on a panel** at a **joint event with the IfG on 'strengthening the UK's resilience',** alongside Fleur Anderson MP and Rt Hon Oliver Letwin.

Risk Management

Contributing to the Institute for Government's first-ever report on managing extreme risks

The Institute for Government (IfG), a well-respected, politically neutral, government-facing think tank, produced a report in July 2022 on Managing extreme risks. The IfG also shared a first draft of the report with us in confidence, on which we were able to offer comments. Our work was drawn on substantially for the report.

The report calls for the UK government to undertake many of our proposed policy changes in full, including the introduction of an overarching risk management framework in government based on the 'three lines of defence' model, a second line central unit being located at the core of government headed by a Chief Risk Officer or equivalent advising a new cabinet-level risk committee, and an independent third line body sitting outside of government reporting to a parliamentary joint committee along the lines of the Climate Change Committee.

We have built the start of a strong partnership with the Institute for Government. For the report, we provided suggestions for their distribution list and referenced it in a letter published in *The Times*.

Held a joint resilience event with the IfG in December

We worked with the IfG to hold a joint event together with the British Red Cross to encourage the UK government to strengthen resilience and to publish the National Resilience Strategy. We invited speakers from both Labour and the Conservatives (Fleur Anderson MP and Sir Oliver Letwin), and spoke alongside them on a panel about strengthening the UK's resilience.

Receiving support for our risk management recommendations

The National Institute of Standards and Technology (NIST) included the 'Three Lines of Defence' in the companion document playbook section of its AI Risk Management Framework Second Draft document, following feedback we'd provided on the initial draft.

Contributing to international risk governance discussions

We were invited to speak at the OECD's High Level Risk Forum (HLRF) meeting in June, and have since been invited to attend the next meeting in December as a participating organisation, and to provide input to discussions on revised guidance on the governance of critical risks, to be published next year. We are also in discussion with the Global Risks team at the World Economic Forum (WEF) to set up an expert meeting following Davos in 2023 to discuss recommended national risk governance best practice.

External Affairs

Growing CLTR's profile

Tom Ough published a profile on Sophie Dannreuther and CLTR's work in the magazine *Prospect* in December 2022.

We had a letter published in the *Financial Times* in response to an article by Stephen Bush ("Can we avoid disaster and should we even try to?").

Other work/achievements:

- We have **designed and are ready to roll out a biosecurity seminar offering** to foster networking and community-building in the biosecurity policy sphere;
- Our **Biosecurity Policy Manager, Sophie Rose, spoke at**:
  - **a UK Innovation and Knowledge Centre for Synthetic Biology panel on biosecurity** alongside Dr. Angela McLean and other security officials in the CBRN space (as the only think tank representative);

Pillar 4: Boosting expert capacity to work with governments on extreme risks

Successes in 2022:

- **Facilitating two academic placements to support MoD work on safe and responsible AI**, as well as helping support **two academic placements to work on AI regulation in the Cabinet Office**;
- **Sharing lessons** learned and **advice** on the many **risks involved in extreme risks policy work**.

AI

Facilitating four placements/secondments to support UK government departments' work on responsible AI

We helped arrange two academic placements to support the MoD's work on safe and responsible AI. We also helped support two academic placements to work on AI regulation in the Cabinet Office.

Pillar 5: Strong organisational health

In 2022, we have focused on building strong operational infrastructure to enable our growing team to thrive and effectively execute our mission.

Building infrastructure for a bigger team

During 2022, our full-time team grew from three to seven. To ensure a smooth, efficient, and enjoyable work environment where team members can do their best work, we have created tools to help the team coordinate and manage their work. This work includes creating smooth hiring and onboarding processes, a generous benefits package, building out our standard operating procedures and guides, operational support for unit leads, and updating our meeting schedule and task management system.

Updating values and team culture

We have updated our values to reflect those of our larger team, and focused on maintaining our team culture. Our three core values are: integrity, people first, and targeted real-world impact.

Unit-specific strategies

With the onboarding of our policy leads (AI policy, biosecurity policy, and risk management policy) and the creation of our external affairs unit, we have developed unit-specific strategies for the next five years. While these strategies will be iterated on in the coming years, they prove a good starting point for increasing and focusing our impact.

Setting up our KPIs and KPI management system

We have developed a set of KPIs to evaluate our progress, to help focus our work, and to learn what activities generate the most impact. The KPIs are structured around our five pillars of impact, and involve both lead and lag indicators of our progress. We implemented the KPIs in September, and will iterate it in 2023. More details about our KPIs system are available on request.

3. Priorities for 2023

In 2023, we will focus on three core priorities: **(1) impactful team** (understanding what activities drive the most impact and explore how we can improve our strategy and effectiveness), **(2) growing team** (ensuring our fundraising, team culture and internal systems are prepared for the expansion of CLTR to a team or 11 or 15), and **(3) thriving team** (ensuring CLTR is a sustainable, supportive, and enjoyable place to work).

Our org-level priorities for 2023 are:

1. **Impactful team**
    - **Identifying which activities are driving the most impact** and which aren't proving as valuable;
    - **Refining CLTR's KPIs and KPI process**: continuing to refine our KPIs for focusing and evaluating our work, refining the meeting schedule for working with the KPIs, and developing a North Star goal and metric.

2. **Growing team**
    - **Gaining more clarity around what our sustainable team size is:** explore and assess the level of funding that is available to us and our unit's hiring needs;
    - **Fundraising**: sourcing a pipeline of sustainable levels of funding for our team and increasing the diversity of funders;
    - **Board refresh**: recruiting new board members to diversify the range of backgrounds and experiences we have on our board, and implementing our new board governance structure;
    - **Hiring and onboarding**: ensuring we find and successfully onboard the talented people we need.

3. **Thriving team**
    - **Team well-being and workload management**: continuing to ensure CLTR is a great place to do great work, that we prioritise the highest-impact opportunities to avoid burnout risks, and ensure our policy units have the levels of support they need — both internally and through our expert networks;
    - **Managing our external communication and reputation**: continuing our communications work to strengthen our reputation;
    - **Great leadership norms**: ensuring our leadership team upholds CLTR's values and approach to collaborative and risk-averse decision-making as the team grows.

4. Funding

Funding situation

During 2022 we expanded from a three-person team to a seven-person team, as well as a part-time operations contractor, in line with our growth plan.

Our current funding breaks down as follows:

Over £1 million

Survival and Flourishing Fund
The Survival and Flourishing Fund is funded primarily by Jaan Tallinn, a founding engineer at Skype, and aims to fund organisations working on long-term challenges. Jaan is a co-founder of the Cambridge Centre for the Study of Existential Risk, and of the Future of Life Institute. He is a core supporter of The Founders Pledge, both as an adviser and member of the Foundry.

A private foundation
This foundation (founded in 2018) focuses primarily on impact investing, the promotion of social responsibility, and making grants that benefit low and middle-income countries.

Over $100,000

The Powoki Foundation
The Powoki Foundation focuses on safeguarding humanity from global challenges, such as safely navigating synthetic biology and advanced artificial intelligence.

Before accepting funds from any donor, we undertake a robust due diligence process to ensure that we are satisfied that the origin of the funds does not conflict with CLTR's overall mission or values. None of our donors have any influence on CLTR's strategic direction or priorities.

You can find an overview of our 2022 spending in Annex 2 below.

Annex 1 – Expert Network

We work closely with experts from a range of domains to inform our work and ensure policy recommendations we produce and present are of high quality. We continue to work to expand our expert network, with an emphasis on expanding the diversity of perspectives among our collaborators.

- Department of Statistics, University of Oxford
- Institute for Progress
- Institute for Policy Research, University of Bath
- Beehive
- MIT Media Lab
- Department of War Studies and the School of Law, KCL
- Centre for Science & Security Studies, KCL
- Centre for Geopolitics, University of Cambridge
- Effective Institutions Project
- Wharton School, University of Pennsylvania
- Council on Strategic Risks
- Rethink Priorities
- Simon Institute
- Legal Priorities Project
- Leverhulme Centre for the Future of Intelligence
- Global Biological Policy & Programs, Nuclear Threat Initiative
- International & Biosafety Initiative for Science, Nuclear Threat Initiative
- European Leadership Network
- altLabs
- Institute of Microbiology and Infection, University of Birmingham
- CSER, University of Cambridge
- Future of Humanity Institute, University of Oxford
- Department of Philosophy, University of Oxford
- Global Priorities Institute, University of Oxford
- International Relations Network, University of Oxford
- OpenAI
- GovAI
- RAND Corporation
- Anthropic
- Atlas Fellowship
- Center for Security and Emerging Technology, Georgetown University
- Polaris Ventures
- Future of Life Institute

Annex 2 – Overview of our 2022 Spending

Expense type	Breakdown	Estimated cost	Actual cost
Staff salaries	Salaries for: • Chief Executive • Director • Heads of Policy x 2 • Managers x 2 • Executive Assistant	**£428k**	£409k
External consultants	**£216k**: £15k per month + VAT for consultancy support (comms + advocating for top policy recommendations) + Ad hoc consultants as needed (e.g. events) of **£30k per year**	**£246k**	£229k
Office space	**£92k**: Public Hall & £6.4k + VAT per month = £92k per year (minus first month in year for 2022) + meeting room credits £6k per year **£10k** setup costs (furniture, electronics etc)	**£98k**	£93k
Legal fees	**£20k**: Based on contingency costs required in 2020 from a medium-tier firm specialising in UK-based non-profits	**£20k**	£59k[6]
Benefits (pension, life assurance, gym membership etc)	**£70k:** Approx. £10k per person per year on average	**£70k**	£35k
Travel and accommodation budget	**£60k:** Assuming an average annual budget of £10k for six team members **£10k** retreats	**£55k**	£26k

[6] Legal advice on immigration, employment, IP, and governance-related issues has been more costly than forecasted. We have updated our projections for this budget going forward.

	(lower costs in Year Three)		
Accountancy fees	**£18k:** Based on costs incurred in 2020, scaled up to a team of seven **£4k:** Swedish costs	**£22k**	£19k
Costs under £10k	IT hardware: **£6k** Business insurance: **£6k** Entertainment / relationship building: **£4k** Newspaper subscriptions: **£0.8k** Bank fees: **£0.6k** Petty cash e.g. postage: **£0.1k**	**£17.5k**	£23k
Learning and development budget	**£21k:** Based on a budget of £3k per team member	**£21k**	£6k
IT software inc. website support	**£5k:** Contingency fee for any website support required	**£5k**	£4k
	Sub-total	£982.5k	£903k
Total (including 10% contingency budget)		**£1.08 million**	**£903k**