

Centre for the Study of Existential Risk Six Month Report: April 2019 – September 2019

We send short monthly updates in our newsletter – subscribe [here](#).

Contents

1. Overview.....	1
2. Policy Engagement – Impact:.....	2
3. Industry Engagement – Impact:.....	4
4. Academic Engagement – Field-building:.....	4
5. Public Engagement – Field-building:.....	7
6. Recruitment and research team	8
6.1 Changes to CSER Management Team	8
6.2 New Postdoctoral Research Associates and Research Assistants	8
6.3 Visiting Researchers	8
7. Upcoming events.....	9
8. Publications	10

1. Overview

The Centre for the Study of Existential Risk (CSER) is an interdisciplinary research centre within the University of Cambridge dedicated to the study and mitigation of risks that could lead to civilizational collapse or human extinction. Our research focuses on Global Catastrophic Biological Risks, Extreme Risks and the Global Environment, Risks from Artificial Intelligence, and Managing Extreme Technological Risks. Our work is shaped around three main goals:

- **Understanding:** we study existential risk.
- **Impact:** we develop collaborative strategies to reduce existential risk.
- **Field-building:** we foster a global community of academics, technologists and policy-makers working to tackle existential risk.

Our last Six Month Report was in [April 2019](#). Since then, we have continued to advance existential risk research and grow the field. Highlights include:

- Publication of **two academic books** by Sir Partha Dasgupta on the sixth mass extinction and on population ethics.

- Publication of **seven academic papers** on risks associated with AI and nuclear weapons, release norms for machine learning, mediation in climate diplomacy, avoiding polarisation around negative emission technologies, and new technologies for producing animal feed more sustainably.
- Publication of a **policy report** on how governments can better understand extreme risks, influencing the **EU AI ethics guidelines**, and engaging with leading UK and international policy-makers.
- Ellen Quigley's appointment to **advise Cambridge University** on responsible investment.
- **Ten expert workshops** on AI risks, climate change, managing global catastrophic risks, and biosecurity.
- **Raising awareness** of existential risk through articles, podcasts, TV appearances, lectures and the Ground Zero Earth' exhibition.
- Hosting our **first internship scheme**, with four Summer Visitors.
- Hiring **five new team members**, with recruitment about to begin for four more.

2. Policy Engagement – Impact:

We have had the opportunity to speak directly with policymakers and institutions across the world who are grappling with the difficult and novel challenge of how to unlock the socially beneficial aspects of new technologies while mitigating their risks. Through advice and discussions, we have the opportunity to reframe the policy debate and to hopefully shape the trajectory of these technologies themselves.

- We published a [Policy Report](#) **Managing global catastrophic risks: Part 1 Understand**. Its purpose is to inform policy-makers at a national level how they can better understand global catastrophic risks. We put out a [press release](#) and it received media coverage in [Vox](#), the [Irish Times](#), and elsewhere. We had several meetings in London, New York, Washington DC, San Francisco and Canberra to discuss the report with policy-makers.
- The EU's [AI ethics guidelines](#) were published. It reflected advice that Haydn Belfield and Dr Shahar Avin had submitted to the High-Level Expert Group, drawing attention to the recommendations in the [The Malicious Use of Artificial Intelligence](#) report. The EU's Ethics Guidelines are likely to affect policy and corporate behaviour across Europe.
- Dr Catherine Rhodes was involved in **three submissions to the UK Parliament's Joint Committee on the National Security Strategy** Inquiry on *Biosecurity and Human Health: Preparing for Emerging Infectious Diseases and Bioweapons*:
 - Co-author of [BHH0010](#) – Written evidence submitted by Dr Cassidy Nelson et al. This developed from an FHI roundtable.

- Lead in drafting [BHH0005](#) - Written evidence submitted by Biosecurity Research Initiative at St Catharine's College (BioRISC), Cambridge.
- Individual submission [BHH0007](#) - Written evidence submitted by Dr Catherine Rhodes, Executive Director, Centre for the Study of Existential Risk 'International Governance of Biosecurity and Human Health: Challenges and Opportunities for Coordination and Coherence'.
- The All-Party Parliamentary Group for Future Generations held two events in Parliament, continuing our engagement with MPs and Peers on long-termism:
 - 11 June: **Drones, Swarming and the Future of Warfare**, with David Hambling, Journalist and Author of 'Swarm Troopers', and Sebastian Brixey-Williams, Programme Director, BASIC. Report [here](#).
 - 22 May: **Negative emissions technologies: a necessary step or a false hope?** with Dr Naomi Vaughn, Charlotte Morgan and Dr David Reiner. Report [here](#).
- 21 May: CSER was a partner in the **Zero Carbon Futures Symposium** organised by Carbon Neutral Cambridge. This also involved participants from the Greater Cambridge Shared Planning Service, South Cambridgeshire District Council, Cambridge City Council, and local businesses. Its aim was to help accelerate the local transition to Net Zero Carbon by informing policy development for the new Greater Cambridge Local Plan. Report [here](#).
- 21 May: Dr Catherine Rhodes and Dr Sam Weiss Evans participated in a policy roundtable on emerging technologies where there is a need for global coordination and frameworks co-organised by the Centre for Science and Policy (CSaP) and the **Cabinet Office G7/G20 team**.
- 30 May: **Responding to Catastrophic Climate Change and Environmental Collapse** (led by Dr Simon Beard). With leading UK think-tanks IPPR and Demos.
- 19 June: Catherine Rhodes participated in **CSaP / Defence Science and Technology Laboratory Policy (DSTL) Workshop** on the risks of emerging technologies to national security.
- 26 June: CSER sponsored and organised a [panel](#) on 'Extreme Risks: Challenges for Evidence and Policy' at **CSaP's annual conference**. Audio available [here](#).
- Researchers from CSER and the Leverhulme Centre for the Future of Intelligence (CFI) (Dr Jess Whittlestone, Dr Sean O hEigeartaigh, Dr Shahar Avin, and Haydn Belfield) participated in a one day workshop with the **Centre for Data Ethics and Innovation** (the UK's national AI advisory body). Topics included horizon-scanning and foresight, targeting and misinformation.

- Dr Jess Whittlestone spoke at the **OECD 2019 Forum** on AI ethics principles.
- September: Dr Ellen Quigley participated in the **United Nations Principles for Responsible Investment** (PRI) conference. She has been [appointed](#) to work with Cambridge University's Chief Financial Officer to establish a **'responsible investment' research programme**. We have hired two research assistants to support her in this work.
- CSER researchers continued meetings with top UK civil servants as part of the **policy fellows program** organized by CSaP.

3. Industry Engagement – Impact:

Researchers continued their extensive and deep collaboration with industry. Extending our links improves our research by exposing us to the cutting edge of industrial R&D, and helps to nudge powerful companies towards more responsible practices.

- August / September: Haydn Belfield spent a month embedding at a **leading AI company** in San Francisco, deepening our links and collaborating on a multistakeholder report. He also contributed to two reports on publication/release norms. Previously in San Francisco, he chaired several sessions at Effective Altruism Global San Francisco and participated in an invite-only workshop on AI and international security.
- Dr Catherine Rhodes attended the **Australian Leadership Retreat**, and was a speaker in the sessions 'Climate Change and Ecological Breakdown', 'Leadership for Existential Threats', and 'Is Australia's Sovereignty at Risk?'. The ALC is Australia's premier forum for top-100 CEOs and national policymakers.
- Dr Sean O hEigeartaigh gave a keynote talk at the **Times Higher Education Innovation Summit in Korea**, and met with experts from KAIST. He also participated in several meetings and working groups of the **Partnership on AI**.
- Dr Shahar Avin continued running '**scenario exercises**' exploring different possible AI scenarios. He has run over thirty so far, with some participants from leading AI labs. He aims to explore the realm of possibilities, and educate participants on some of the challenges ahead.

4. Academic Engagement – Field-building:

As an interdisciplinary research centre within Cambridge University, we seek to grow the academic field of existential risk research, so that it receives the rigorous and detailed attention it deserves.

- 3-5 April: CSER supported the **EiM 2: The second meeting on Ethics in Mathematics** workshop. Drs Maurice Chiodo and Piers Bursill-Hall from the Faculty of Mathematics in Cambridge have been spearheading an effort to teach responsible behaviour and ethical awareness to mathematicians. [More](#).
- 5-6 April: **Tools for building trust in AI development** workshop (co-led by Dr Shahar Avin). This two-day workshop convened some of the world's top experts in AI, security, and policy to survey existing mechanisms for trust-building in AI and develop a research agenda for designing new ones.
- 6-7 June: **Evaluating Extreme Technological Risks** workshop (led by Dr Simon Beard). This brought together philosophers and economists to explore methodological problems in the evaluation of extreme risks.
- 21 June: Dr Catherine Rhodes, Prof Bill Sutherland, and Sam Weiss Evans visited the **Pirbright Institute**, the UK's leading research institute dedicated to the study of infectious diseases of farm animals, to discuss potential intersections with the biological risks work of CSER and BioRISC.
- 11-12 July: **Cross-Cultural Trust for Beneficial AI [workshop](#)**, led by Dr Yang Liu, Prof Huw Price and Dr Sean O hEigeartaigh. The aim was to address some potential obstacles for international cooperation for beneficial AI from a cross-cultural perspective, and, equally importantly, to connect a number of current initiatives to encourage trust-building dialogue between China and the West. Guests from China included Prof Yi Zeng (Chinese Academy of Sciences), Bing Song (Berggruen Institute China Center), Prof Zhe Liu (Peking University/CFI), Dr Chuang Liu (Fudan University) and Prof Victor OK LI (University of Hong Kong). It was followed by a public lecture on [Norms for Digital Technologies](#), by Prof. Onora O'Neill.
- 12-13 July: **Novel Practices of Biosecurity Governance [Workshop](#)** (led by Dr Sam Weiss Evans). Around 40 leading practitioners gathered to discuss the development of a system for sharing knowledge and learning from the practice of implementing biosecurity governance measures. It will result in a paper outlining the main findings and next steps to be taken.
- 16-17 July: the **Biosecurity Research Initiative at St Catharine's** (BioRISC) held its [launch event](#) at the House of Lords, to mark the first anniversary of the UK Biological Security Strategy. It was followed by a ['100 questions for UK Biosecurity' workshop](#) in Cambridge, that will result in a paper to inform the research agenda for biosecurity and provide a resource for communities engaged in national and global biosecurity efforts.

- 10-12 August: **AI Safety workshop** at leading machine learning conference *IJCAI 2019* (led by organising committee including Drs Sean O hEigeartaigh and Jose Hernandez-Orallo).
- 26-27 August: **Decision Theory & the Future of Artificial Intelligence** Workshop (led by Prof Huw Price and Dr Yang Liu) at the Australian National University (ANU). The [third workshop](#) in a series bringing together philosophers, decision theorists, and AI researchers to promote research at the nexus of [decision theory and AI](#). Co-organised with the Munich Center for Mathematical Philosophy.
- 23 September: **Black Sky Resilience Group (BSRG)** second roundtable. This was set up by Julius Weitzdörfer; this roundtable will be facilitated by Dr Luke Kemp.
- Dr Catherine Rhodes has been very active in academic engagement.
 - April: presented on ‘To what extent is international governance prepared for risks from new technologies?’ at the **Bennett Institute for Public Policy Conference**, Cambridge.
 - June: presented at the ‘Responsible Innovation, Risk, and Biotechnology’ workshop, **Warwick Integrative Synthetic Biology Centre**.
 - July: chaired Working Group 5: Foresight at a NATO sponsored workshop on **Security for Emerging Synthetic Biology Threats**, Lausanne. Dr Sam Weiss Evans also participated in the workshop.
 - July: participated in Royal Academy of Engineering Workshop on **Global Safer Complex Systems**.
 - August: led a **Schmidt Science Fellows roundtable** on ‘Interdisciplinarity and solving real-world problems’.
 - September: presented on ‘Mundane Crises and Failures of Governance’ at **MANCEPT workshop on Disasters and Crises**, Manchester University.
 - September: presented on the work of CSER and BioRISC at a meeting of the **UK Biosafety Strategic Leadership Group, PHE Colindale**. Dr Sam Weiss Evans also gave a remote presentation.
 - September: presented on ‘Power, Trust and Distrust in the Governance of (bio)Technologies’ at the **Cambridge Trust & Technology Initiative Symposium**.
- Haydn Belfield and Dr Luke Kemp presented at Princeton University’s [Workshop on "Historical Systemic Collapse"](#).
- Julius Weitzdörfer presented on ‘**Future Generations and Existential Risk**’ to Department of Engineering students and to the ICE International Summer Programme, and chaired a panel on ‘The mystery of risks - How can science help

reconcile perception and assessment?'. He also participated in the OECD-NEA International School of Nuclear Law.

- Drs Simon Beard, Catherine Rhodes and Lalitha Sundaram contributed lectures for the transferable skills module of the **MPhil Biotech**. Dr Rhodes has provided an interview for a Coventry University online **MSc Emergency Management and Resilience**.

5. Public Engagement – Field-building:

- Ariel Conn interviewed Dr Simon Beard and Haydn Belfield for an FLI podcast on [climate change as an existential risk](#). It has been listened over 4,750 times, and has prompted a lot of discussion in the existential risk community, a few local groups have had specific discussion groups about the podcast.
- Dr Luke Kemp was interviewed by [Australian radio](#) about why and how civilizations collapse, and wrote an [Aeon longread](#) on how collapse has surprisingly been often quite mild – but would not be nowadays. He also gave a long [video interview](#) about civilizational collapse.
- Haydn Belfield was interviewed by the [Naked Scientists](#) on a radio Q&A episode, answering questions like ‘what is the doomsday clock?’. He wrote a brief *Daily Mirror* [article](#) – read over 100,000 times – on a priority for the new UK Government: the malicious use of AI.
- Dr Simon Beard wrote about [Deep Ethics](#), and whether we could ever have shared universal principles, for BBC Future.

We are able to reach far more people with our research online:

- 11,243 website visitors over the last two months.
 - 6,942 newsletter subscribers.
 - 7,842 Twitter followers.
 - 2,537 Facebook followers.
-

- Dr Lauren Holt had a [long read](#) on ‘post-natural’ wildlife published in Aeon
- Drs Catherine Rhodes, Lauren Holt and Lalitha Sundaram had an [article](#) on artificial diseases published in the Metro Online, a widely read UK newspaper.
- Dr Asaf Tzachor had a Conversation article on [‘photo-bioreactors’ for feeding livestock](#), to accompany his new paper.
- Research Affiliate Dr Adrian Currie gave a 30-minute [TV interview](#) on existential risk.
- The month-long Ground Zero Earth exhibition exploring art and existential risk ended with a [screening](#) of ‘Ghost in the Machine’. We published an [overview](#) of the exhibition (with photos of the pieces).

- Lord Martin Rees lectured at [Hay Festival](#) and [Chatham House](#), and was [interviewed](#) by *the Observer* newspaper.

6. Recruitment and research team

6.1 Changes to CSER Management Team

From April, Dr Catherine Rhodes has taken on the role of Executive Director of CSER, with primary responsibility for leadership and management of the Centre's research and operations. Dr Seán O hÉigeartaigh's is now Co-Director of CSER, and Dr Simon Beard has started in the role of Academic Programme Manager. Haydn Belfield remains Academic Project Manager and Clare Arnstein remains Research Project Administrator.

6.2 New Postdoctoral Research Associates and Research Assistants

We have hired three Research Associates and two Research Assistants, and will be adding them to the [Team page](#) once they begin.

- **Research Associate, Global Justice and the Governance of Global Risks.** This post runs through to 2023 and is funded by the Isaac Newton Trust.
- **Research Associate in Population, Sustainability and Environmental Risk.** This post runs for 36 months and is funded by the Grantham Foundation.
- **Research Associate in Responsible Innovation and Extreme Technological Risk.** This post runs for 12 months and is primarily funded by the Templeton World Charity Foundation.
- **Part-time Research Assistants for Ellen Quigley.** As part of Ellen Quigley's work on responsible investment with the University Chief Financial Officer, funds have been provided for two part-time research assistant positions through to June 2020. They will support Ellen and other colleagues associated with the research project on sustainable finance, in investigating the ways in which the financial system can contribute to the transition to a zero-emissions economy.

6.3 Visiting Researchers

- **Dr Sam Weiss Evans** continued his year-long visit to CSER through to the end of July. He led a US/UK workshop on novel approaches to governance of dual use

research 11-13 July. Sam was joined by a student funded by Tufts University, **Stefan Lunte**, who supported his research and engagement activities over the next two months.

- **Rumtin Sepasspour**, Foreign Policy Adviser, Office of the Prime Minister of Australia, joined us for four months from the end of April.
- We had four **interns** this summer:
 - [Amritha Jayanti](#), worked with Shahar Avin on forecasting AI progress, and scenario mapping.
 - [Ross Gruetzemacher](#), worked with Shahar Avin on accountability gaps for military uses of AI.
 - [Siebe Rozendal](#), worked with Simon Beard on the relative importance of work on extinction-risks and collapse-risks.
 - [Nathaniel Cooke](#), worked with Luke Kemp on civilizational collapse.
- **Dr Megan Palmer**, Senior Research Scholar, Center for International Security and Cooperation, Stanford University, made two short visits in June and July, working - in particular - with Lalitha Sundaram and Sam Weiss Evans on their bio-risk related work.
- **Dr H. Orri Stefansson** will visit 13 - 17 October, and will give a work-in-progress on 14 October. Orri is currently a Pro Futura Fellow at the Institute for Future Studies in Stockholm, working mainly on decision theory and will be developing collaborations with Yang Liu, and other CSER researchers during his visit.
- **Dr Nick Evans**, University of Massachusetts, Lowell, will visit in September-December, working on a book project on scientific freedom as a factor in navigating dual-use risks in (but not limited to) biology. He will also work on a possible project with Lalitha Sundaram. He will give a talk while here.

7. Upcoming events

- 1 October: **Blavatnik Public Lecture - [Jason Matheny](#)**. Founding Director of the Center for Security and Emerging Technology at Georgetown University. Previously he was Assistant Director of National Intelligence and Director of IARPA. He is a member of the National Security Commission on Artificial Intelligence and was named one of Foreign Policy's "Top 50 Global Thinkers."
- October 9: Biological Engineering **Horizon-Scanning** workshop (led by Dr Luke Kemp). This follows our 2016 Biological Engineering Horizon-Scanning

[workshop](#) which produced an important [paper](#) on 20 emerging issues in biological engineering.

- 30 October: **Blavatnik Public Lecture** – [Zia Mian](#). Co-director of Princeton University’s Program on Science and Global Security. Received the 2014 Linus Pauling Legacy Award for “his accomplishments as a scientist and as a peace activist in contributing to the global effort for nuclear disarmament”.
- 18 November: **Blavatnik Public Lecture** – [Grethe Helene Evjen](#). Senior advisor at Norwegian Ministry of Agriculture and Food. Was key to the implementation and coordination of the Svalbard Global Seed Vault.
- 3 March: **Blavatnik Public Lecture** – [Rachel Bronson](#). President and CEO of the Bulletin of the Atomic Scientists. Will be visiting shortly after the 2020 Doomsday Clock announcement.
- 6-7 April 2020: CSER’s next international **Cambridge Conference on Catastrophic Risk**.

8. Publications

- **Partha Dasgupta**, Peter Raven and **Anna Mclvor** (Eds.). (2019). [Biological Extinction](#). Cambridge University Press.
 - “The rapidly increasing human pressure on the biosphere is pushing biodiversity into the sixth mass extinction event in the history of life on Earth. The organisms being exterminated are integral working parts of our planet’s life support system, and their loss is permanent. Like climate change, this irreversible loss has potentially devastating consequences for humanity. As we come to recognise the many ways in which we depend on nature, this can pave the way for a new ethic that acknowledges the importance of co-existence between humans and other species. Biological Extinction features chapters contributed by leading thinkers in diverse fields of knowledge and practice, including biology, economics, geology, archaeology, demography, architecture and intermediate technology. Drawing on examples from various socio-ecological systems, the book offers new perspectives on the urgent issue of biological extinction, proposing novel solutions to the problems that we face.”
 - It draws upon the 2017 workshop with the Vatican’s Pontifical Academy of Sciences he co-organised.
- **Partha Dasgupta**. (2019). [Time and the Generations - population ethics for a diminishing planet](#). New York: Columbia University Press),

- “How should we evaluate the ethics of procreation, especially the environmental consequences of reproductive decisions on future generations, in a resource-constrained world? While demographers, moral philosophers, and environmental scientists have separately discussed the implications of population size for sustainability, no one has attempted to synthesize the concerns and values of these approaches. *Time and the Generations* blends economics, philosophy, and ecology to provide tentative answers to two fundamental questions: What level of economic activity can our planet support over the long run, and what does the answer say about optimum global population numbers? Dasgupta develops a population ethics that can be used to evaluate our choices and guide our sense of a sustainable global population and living standards. Structured around a central essay from Dasgupta, the book also features a foreword from Robert Solow; correspondence with Kenneth Arrow; incisive commentaries from Joseph Stiglitz, Eric Maskin, and Scott Barrett; an extended response by the author to them; and a joint paper with Aisha Dasgupta on inequalities in reproductive decisions and the idea of reproductive rights.”
- **Shahar Avin.** and Amadae, S. (2019). [Autonomy and machine learning at the interface of nuclear weapons, computers and people](#) in Boulanin, V. (Ed.) *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk: Euro-Atlantic Perspectives*. SIPRI.
 - “Increasing attention has been given in the literature to the impact of digital technologies, and in particular autonomy and machine learning, on nuclear risk. Most of this attention has focused on ‘first-order’ effects: the introduction of technologies into nuclear command-and-control and weapon-delivery systems. This essay focuses instead on higher-order effects: those that stem from the introduction of such technologies into more peripheral systems, with a more indirect (but no less real) effect on nuclear risk. It first describes and categorizes the new threats introduced by these technologies (in section I). It then considers policy responses to address these new threats (section II).”
- **Asaf Tzachor.** (2019). [The Future of Feed: Integrating Technologies to Decouple Feed Production from Environmental Impacts](#). *Industrial Biotechnology Vol. 15, No. 2*.
 - “Population growth, an expanding middle-class, and a global shift in dietary preferences have driven an enduring demand for animal products. Since animal products are playing a vital role in human diets, their consumption is predicted to increase further. However, the great dependency of animal husbandry on global staple feed crop soybean; the environmental consequences of soybean production; and barriers for soy cropland expansion cast doubt on food system sustainability. The need to mitigate

future demand for soy with other feed sources of similar nutritional profile, and thereby decouple food and feed production from ecological pressures, is compelling. Yet, the literature and science of sustainable agriculture is one of incremental improvements, featuring primarily crop production intensification. A different, more profound approach to the design of feed systems is required to ensure sustainable food security. The question arises if alternative technologies exist to support such a design. This paper explores a particular novel configuration of four advanced technologies recently deployed in the region of Hengill, Iceland: light-emitting diode systems, advanced indoor photobioreactors, atmospheric carbon capture technology, and geothermal energy technology. In situ system analysis and data triangulation with scientific literature and data from independent sources illustrate the potential of these integrated technologies to produce algal-based animal feed. The analysis suggests that a highly sustainable soybean equivalent is technically attainable for feed purposes. The integrated system requires less than 1% of arable land and fresh water compared with soybean cultivation and is carbon negative. In addition, it provides a pesticide- and herbicide-free cultivation platform. This new configuration provides one pathway for the future of feed.”

- **Luke Kemp.** (2019). [Mediation Without Measures: Conflict Resolution in Climate Diplomacy](#) in Wilkenfeld, J., Beardsley, K. and Quinn, D. (Eds). *Research Handbook on Mediating International Crises*. Edward Elgar.
 - “Current conceptions of mediation can often fail to capture the complexity and intricacy of modern conflicts. This Research Handbook addresses this problem by presenting the leading expert opinions on international mediation, examining how international mediation practices, mechanisms and institutions should adapt to the changing characteristics of contemporary international crises.”
- **Ovadya, A. and Jess Whittlestone.** (Working Paper). [Reducing malicious use of synthetic media research: considerations and potential release practices for machine learning](#). arXiv preprint.
 - “The aim of this paper is to facilitate nuanced discussion around research norms and practices to mitigate the harmful impacts of advances in machine learning (ML). We focus particularly on the use of ML to create “synthetic media” (e.g. to generate or manipulate audio, video, images, and text), and the question of what publication and release processes around such research might look like, though many of the considerations discussed will apply to ML research more broadly. We are not arguing for any specific approach on when or how research should be distributed, but instead try to lay out some useful tools, analogies, and options for thinking about these issues.”

- **Simon Beard.** (2019). [Book Review - Climate Justice: Integrating Economics and Philosophy](#). Economics and Philosophy.
 - “Debates about justice are increasingly seen as vital to policy-making and international dialogue on climate change and how we should respond to it. While many disciplines have participated in these debates, philosophers and economists are often the most vocal. However, given the many historical disagreements between these disciplines this raises the question of whether we are fighting on the same team. This important volume of essays, edited by a philosopher and an economist who have contributed to both academic debates and real-world policy forums on climate change, argues that we are.”

- **Haydn Belfield.** (2019). [How to respond to the potential malicious uses of artificial intelligence?](#) Journal of Unsolved Questions.
 - “Artificial intelligence (AI) is beginning to change our world – for better and for worse. Like any other powerful and useful technology, it can be used both to help and to harm. We explored this in a major February 2018 report *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. We co-authored this report with 26 international experts from academia and industry to assess how criminals, terrorists and rogue states could maliciously use AI over the next five years, and how these misuses might be prevented and mitigated. In this piece I will cover recent advances in artificial intelligence, some of the new threats these pose, and what can be done about it.”

- R.M. Colvin, **Luke Kemp**, Anita Talberg, Clare De Castella, C. Downie, S. Friel, Will J. Grant, Mark Howden, Frank Jotzo, Francis Markham, Michael J. Platow. (2019). [Learning from the Climate Change Debate to Avoid Polarisation on Negative Emissions](#). *Environmental Communication*.
 - “This paper identifies critical lessons from the climate change experience to guide how communications and engagement on negative emissions can be conducted to encourage functional public and policy discourse. Negative emissions technologies present a significant opportunity for limiting climate change, and are likely to be necessary to keep warming below 2°C. While the concept of negative emissions is still in its infancy, there is evidence of nascent polarization, and a lack of nuance in discussion of individual technologies. We argue that if negative emissions technologies are to be implemented effectively and sustainably, an effective governance regime is needed; built on functional societal discourse and avoiding the ideological baggage of the broader climate change debate or the controversies concerning

geoengineering. At its core, our argument is to avoid the ideological bundling of negative emissions; this can be pursued directly and via careful selection of communication frames and the use of non-partisan, trusted messengers. Whether these lessons are heeded may determine if negative emissions are governed proactively, or are distorted politically, misused and delayed.”