

The Global Catastrophic Risks (GCR) Fund

Interventions to mitigate the biggest risks facing humanity.

Keeping Humanity Safe

We live in a time of perils. Humanity faces large and growing risks, including war between great powers, natural and engineered pandemics, nuclear war, threats from advanced artificial intelligence (AI), and emerging risks from frontier military technologies. These global catastrophic risks have the potential to kill billions of humans alive today.

Scientists, policymakers, and the public can come together to mitigate these risks. We've done it before. During the Cold War, leaders cooperated to reduce stockpiles of weapons of mass destruction. At the turn of the millennium, scientists spearheaded an unprecedented effort to compute and track the threat of large asteroids and comets in Earth's vicinity to protect humanity from catastrophic impacts. A growing global movement now seeks to increase pandemic preparedness in the wake of COVID-19.



The <u>Global Catastrophic Risks (GCR) Fund</u> seeks to identify and fund interventions to reduce the largest known risks to humanity today. The Fund is designed both to tackle threats to our long-term future, and to take action now and help protect every human being alive today.

Our Perspective



The Global Catastrophic Risks (GCR) Fund will find opportunities to support highly impactful and neglected initiatives to reduce the probability and mitigate the impact of worldwide catastrophes. These include but are not limited to: great power conflict; nuclear war; devastating natural pandemics; engineered bioweapons; catastrophic laboratory accidents; emerging risks from frontier military and civilian technologies, including artificial intelligence (AI); and more. This is a complex mission, with an ever-changing threat landscape.

The Fund's grantmaking will take a balanced approach to existential and catastrophic risks. Those who take a <u>longtermist</u> perspective in principle put special weight on <u>existential risks</u>—those that threaten to extinguish or permanently curtail humanity's potential—even where interventions appear less tractable. Not everyone shares this view, however, and people who care mostly about current



generations of humanity may prioritize highly tractable interventions on global catastrophic risks that are not directly "existential".

In practice, however, the two approaches often converge, both on problems and on solutions. A common-sense approach based on simple cost-benefit analysis points us in this direction even in the near-term. To illustrate this, both longtermists and those who prioritize current generations care about risks from catastrophic pandemics (whether natural or engineered) and both may want to fund work on developing advanced personal protective equipment (PPE). Complex cascading risks and risk factors further support a broad approach — for example, conflict between the Great Powers may not pose an extinction risk, but may cause a breakdown in international cooperation and thereby prevent future governance of high-consequence emerging technologies. It is therefore valuable for some funders to bridge philosophical divides and incorporate both perspectives outlined above, as this Fund seeks to do. We give special consideration to threats that could curtail humanity's future, but lean towards tractable solutions today.

Whether or not you share our <u>commitment to longtermism</u>, or believe that reducing existential risks is particularly important, you may still be interested in the Fund for the simple reason that you want to help prevent the deaths and suffering of billions of people. These risks are with us today, and more loom on the horizon, with advances in synthetic biology, AI, and modern weapons applications of emerging technologies. There is much to be done.



Our decision-making will be guided by three core values: impact, innovation, and flexibility.



To **maximize impact**, our grant portfolio will include both direct interventions (for instance, funding the development of new crisis communications technology or PPE) as well as research and hits-based bets (where success is less certain, but with the potential to improve many more lives if successful). Further, by seeking opportunities that are neglected by other funders, we aim to ensure that our Fund is as high-leverage as possible. For example, in our recent <u>research report on military AI and autonomous weapons</u>, we found that traditional philanthropists focused on eye-catching ideas of "killer robots", while neglecting the more mundane but higher-impact strategic-stability risks of AI applications.

We are committed to **innovation**, working to develop new and better approaches to grantmaking, and helping to provide seed funding for novel projects that we believe have high expected value. Through these grants, we hope to build a deeper understanding of how we can best evaluate and mitigate global catastrophic risks. We will iterate on our sourcing and evaluation processes to minimize the burden imposed on potential grantees. We will seek to experiment with the benefits of probabilistic forecasting and prediction markets to inform decision-making processes.

Finally, we will seek to maintain **flexibility** in order to rapidly respond to emerging crises and windows of opportunity. A main value proposition of this Fund is that it can step in quickly in response to new opportunities. We will work with networks of domain experts, trusted partners, and government decision-makers to identify new opportunities, and deploy funds in the most effective ways. We will also collaborate with the <u>Patient Philanthropy Fund</u>; we view the time horizons of these two philanthropic tools as complementary, such that philanthropists can allocate between the two funds depending on their view about the rate of emergence of new risks and the tractability of existing interventions.

Our Approach to Grantmaking

When evaluating potential grants, we consider several key points:

- **Focusing on counterfactual impact**. The Fund is designed to make grants that otherwise wouldn't have been made, and grants that individuals would be otherwise unable to make.
- Collaborating with trusted partners. The Fund will seek to collaborate with other funders who share our views and approach to effective giving and whose counterfactual impact we consider similar to that of the Fund. Trusted partners include Effective Giving, Giving What We Can and Longview Philanthropy's new Longtermism Fund, the Long-Term Future Fund, Open Philanthropy, and more. We plan to keep these and other philanthropic actors abreast of grantmaking and possibly to work together on co-funding promising projects.





- Avoiding harm and information hazards. Many global catastrophic risks revolve around dangerous technologies and issues of national and international security. In our grantmaking, we will develop rigorous processes to avoid doing harm or spreading harmful information ("information hazards"). For example, if we believe that a promising funding opportunity has a chance of developing dual-use technology that can be weaponized, we will not fund it.
- Filling funding gaps. Not all organizations are set up to scale their programming and many
 can't productively absorb large amounts of unsolicited funding. We evaluate all potential
 grantees' capacity for growth and consider how much money they could put to good use.
 This means there will be many great organizations we do not recommend because they are
 already sufficiently funded.
- Evaluating organizational strength. Particularly with unrestricted funding and hits-based giving on highly uncertain issues, it's important to have trust in the leadership and capacity of the organizations we support. We look at their track record, their calibration, and their transparency about failures as well as successes.
- Seizing time-sensitive opportunities and policy windows. International crises—as Russia's invasion of Ukraine and increased tension between the U.S. and China over Taiwan have illustrated—can move fast. The flexibility of the Fund will seek to enable time-sensitive grantmaking, find policy windows, and step in quickly where traditional funders can't.

Why Contribute to the Global Catastrophic Risks (GCR) Fund?



- A focus on impact and high-leverage grantmaking. The Fund Manager will conduct detailed, comparative analyses of potential funding opportunities to find those where we conclude that financial support will reduce the probability or mitigate consequences of catastrophe.
- Unique access to funding opportunities. By pooling contributions from multiple donors, the Fund makes accessible opportunities that are otherwise unavailable to individuals (e.g. organizations may only

accept grants above a certain threshold).

• **Zero administrative burden**. Contributing to the Fund is easy for both Founders Pledge members—who can contact their Community Manager or Advisor—and non-members, who can donate through Every.org or (soon) Giving What We Can. Founders Pledge does not take a cut of any donations—our operational expenses are covered by generous donors—and your donations are fully tax deductible. Every donation contributed to the Fund will be granted to a highly impactful organization.



• Learning opportunities. We will share regular updates and maintain a high degree of transparency. Fund contributors will have unique opportunities to learn alongside the Fund Manager about effective giving within the space of global catastrophic risks.

Past and Future Sample Grants

The following list illustrates the kinds of grants that the Global Catastrophic Risks (GCR) Fund may make. The first is a real grant to seed a new organization that the Fund Manager recently evaluated, and which a generous Founders Pledge member funded. The other section includes examples of the *kinds* of projects that the Fund could support, conditional on organizational strengths, cost-benefit analyses, due diligence, and an assessment of potential harms and information hazards.

Previous Relevant Grant: Launching a New Research Institute on Frontier Military Technologies

Recently, Founders Pledge supported the launch of the Berkeley Risk and Security Lab (BRSL), a new initiative at the University of California, Berkeley, that seeks to study emerging risks from frontier military technologies, like applications of AI to nuclear command, control, and communications (NC3) systems. With generous seed funding of \$1.86 million provided by a Founders Pledge member, we have enabled BRSL to operate for three years. We consulted several experts on military technologies and international relations to evaluate this grant.



Potential Future Grants

Other relevant opportunities we have recently recommended to Founders Pledge members, and which typify the kind of work the Fund would support, include a strategic dialogue between the U.S. and China on AI and basic research into probabilistic forecasting of low-probability high-consequence events. Additionally, we are working with experts and policymakers to explore the idea of an "International Autonomous Incidents Agreement" (see our report on military AI), to build confidence, reduce risks, and provide a focal point for longer-term collaboration on more powerful AI. We are also exploring the resilience of crisis communications links (like the U.S.-Russia and U.S.-China "hotlines"), ways of strengthening the <u>Biological Weapons Convention</u>, new data-generating methods on future risks, risks to critical space-based systems, and more.

Meet the Fund Manager



Christian Ruhl is an Applied Researcher based in Philadelphia.



Christian's research at Founders Pledge focuses on understanding, forecasting, and mitigating global catastrophic risks, including risks from great power conflict and weapons of mass destruction. Previously, Christian managed the research program on "The Future of the Global Order: Power, Technology, and Governance" at Perry World House, the University of Pennsylvania's global affairs think tank. After receiving his BA from Williams College, he studied on a Dr. Herchel Smith Fellowship at the University of Cambridge for two master's degrees, one in History and Philosophy of Science and one in International Relations and Politics, with dissertations on early modern state-sponsored science and Cold War nuclear strategy.

Christian was a member of the 2021 Project on Nuclear Issues (PONI) Nuclear Scholars Initiative, serves on the External Advisory Board of the Berkeley Risk and Security Lab (BRSL), and is a Mentor for summer fellows at the Cambridge Existential Risks Initiative (CERI). His writing has appeared in *The Atlantic, The Bulletin of the Atomic Scientists, Foreign Policy*, and more.

Learn More and Contribute

To learn more about Founders Pledge's <u>Global Catastrophic Risks (GCR) Fund</u>, or to contribute, please email your Community Manager or Advisor, or reach out to <u>funds@founderspledge.com</u>. Non-members can donate via Every.org. The Fund will also be listed on <u>Giving What We Can soon</u>.

Image Credits

Fig 1: <u>US Navy 081119-N-0000X-002 The Japan Maritime Self Defense Force destroyer JDS Chokai (DDG 176) launches a Standard Missile 3 during a Missile Defense Agency joint missile defense intercept test by U.S. Navy Photo, Wikimedia Commons. Public Domain.</u>



- Fig 2: <u>Army Team Equips Soldiers to Take on Chemical, Biological Warfare</u> by The U.S. Army, Wikimedia Commons, is licensed under <u>CC BY 2.0</u>.
- Fig 3: COVID-19 pandemic. Chernivtsi. Ukraine by Mstyslav Chernov, Wikimedia Commons, is licensed under CC BY 4.0.
- Fig 4: <u>USS John S McCain South China Sea 1</u> by Navy Petty Officer 3rd Class James Vazquez, Wikimedia Commons, Public Domain.
- Fig 5: <u>US Navy 061117-M-9827H-045 Marines and People's Liberation Army (Navy) (PLA(N)) 1st Marine Brigade Marines shake hands at the start of "Marine Day" 1st Marine Brigade's obstacle course by U.S. Marine Corps photo by Lance Cpl. Jeremy J. Harper, Wikimedia Commons. Public Domain.</u>
- Fig 6: Korea Civil Defense Exercise 03 (13142417285) by Republic of Korea, Wikimedia Commons, is licensed under \underline{CC} BY-SA 2.0.