# The MisAnthropic Statement

While Anthropic and its rivals are racing to build superintelligence through recursive self-improvement, Anthropic uniquely portrays itself as an AI safety company and leverages that reputation to advance its commercial and political ambitions. In light of its public offering, a broad coalition of signatories has endorsed the following statement:

> "Anthropic warns that recursively self-improving AI could concentrate power and escape human control, yet races to build it first. The company uses safety branding to court investors, employees, and the public—even as it has quietly weakened its safety pledges and lobbied for self-serving regulation. Anthropic is not a safe AI company, and we must reject its reckless attempt to amass power in the name of the greater good."

## Signatories

Steven Pinker
Johnstone Family Professor of Psychology, Harvard University

Thomas Wolf
Co-founder and Chief Science Officer, Hugging Face

Laurence Tribe
Carl M. Loeb University Professor Emeritus, Harvard Law School

Rumman Chowdhury
Founder and CEO, Humane Intelligence PBC

Igor Babuschkin
Co-founder and CEO, River AI; Co-founder, xAI

Oliver Hart
Nobel Laureate in Economics; University Professor, Harvard

Sheldon Lee Glashow
Nobel Laureate in Physics; Higgins Professor Emeritus, Harvard

Judea Pearl
Turing Award Laureate; Chancellor's Professor of Computer Science, UCLA

Tristan Harris
Co-founder and President, Center for Humane Technology

Guillaume Verdon (Beff Jezos)
Founder and CEO, Extropic

Max Tegmark
Professor of Physics, MIT; Founder and Chair, Future of Life Institute

Douglas Hofstadter
Distinguished Professor of Cognitive Science, Indiana University

Lawrence Lessig
Roy L. Furman Professor of Law and Leadership, Harvard Law School

Nate Soares
President, Machine Intelligence Research Institute

George Hotz
Founder and CEO, the tiny corp; Founder, comma.ai

Tim Estes
Chairman, Alliance for a Better Future; Founder and CEO, AngelQ

Thomas Dietterich
Former President, AAAI; Distinguished Professor Emeritus, Oregon State University

Bill McKibben
Founder, Third Act; Co-founder, 350.org

Geoffrey Irving
Co-founder and Chief Scientist, Resolution

Gary Marcus
Professor Emeritus of Psychology and Neural Science, NYU

Peter Scholze
Fields Medalist, Director, Max Planck Institute for Mathematics

Geoff Ralston
Founder, Safe AI Fund; Former President, Y Combinator

George Church
Professor of Genetics, Harvard Medical School; Founding Core Faculty, Wyss Institute

Rob Pike
Co-creator of Go and UTF-8; Formerly Google and Bell Labs

Baburam Bhattarai
Former Prime Minister of Nepal, Society of Nepalese Architects

James Maynard
Fields Medalist, Regius Professor of Mathematics, Oxford

David Mumford
Fields Medalist, National Medal of Science Laureate; Professor Emeritus, Brown and Harvard Universities

Mary Rosenberger
CEO, The Authors Guild

Katharina Pistor
Edwin B. Parker Professor of Comparative Law, Columbia Law School

David Haussler
Professor of Biomolecular Engineering, UCSC; Member, NAS and NAE

Ed Newton-Rex
Founder and CEO, Fairly Trained; Former VP of Audio, Stability AI

Robert Wright
Author and Journalist; President, Nonzero Foundation

Sarah Gardner
Co-founder and Former CEO, Heat Initiative

Rutger Bregman
Historian and Author, Humankind; Co-founder, The School for Moral Ambition

Joanne Harris
Author, Chocolat; Former Chair, Society of Authors

Chuck Wendig
New York Times Bestselling Author

Jeff McMahan
Emeritus Saxyro and White's Professor of Moral Philosophy, Oxford

Peter Railton
Distinguished University Professor Emeritus, University of Michigan

Gretchen Krueger
Affiliate, Berkman Klein Center, Harvard; Former Researcher, OpenAI

Matt White
Founding Executive Director, Open Intelligence Foundation

Joseph Jacks
Founder, OSS Capital

Joe Allen
Author, Dark Aeon

Dylan Hadfield-Menell
Associate Professor of EECS, MIT

Mark Nitzberg
Executive Director, Center for Human-Compatible AI, UC Berkeley

Anthony Aguirre
President and CEO, Future of Life Institute; Adjunct Professor, UCSC

Christopher Kutz
C. William Maxeiner Distinguished Professor of Law, UC Berkeley

Greg Conti
Associate Professor of Politics, Princeton University

Simone Browne
Associate Professor, University of Texas at Austin

Huw Price
Bertrand Russell Professor Emeritus of Philosophy, Cambridge

Zhiwei Yun
Professor of Mathematics, MIT

Andrea Miotti
Founder and CEO, ControlAI

Connor Leahy
US Director, ControlAI; Co-founder and Former CEO, Conjecture

Conner Brown
Executive Director, Finney Foundation for Open Technology

R. Jay Wallace
Distinguished Professor of Philosophy (Austahl Chair), UC Berkeley

Mark Gross
Professor of Pure Mathematics, University of Cambridge

Meetali Jain
Founder and Executive Director, Tech Justice Law

Luis Fernando Alday
Royal Society Research Professor of Pure Mathematics, Imperial College London

Richard Thomas
Royal Society Research Professor of Pure Mathematics, Imperial College London

David Krueger
Founder, Evitable; Assistant Professor, Université de Montréal / Mila

Vince Mastrangelo
Founder MML Builders AI Forum

Niko Kolodny
Professor of Philosophy, UC Berkeley

Thomas Metzinger
Professor Emeritus of Theoretical Philosophy, University of Mainz

Geoffrey Grimmett
Professor Emeritus of Mathematical Statistics, University of Cambridge

Kevin Costello
Professor of Mathematical Physics, Perimeter Institute

Émile P. Torres
Anti-TESCREAL Activist, Philosopher and Journalist; Affiliate, Data & Society

Lionel Levine
Professor of Mathematics, Cornell University

Stephen Cooper
Assistant Professor, Harvard Kennedy School

Paul S. Rosenbloom
Professor Emeritus of Computer Science, USC; Co-creator of Soar

Noah Giansiracusa
Associate Professor of Mathematics, Bentley University; Faculty Associate, Harvard

Wesley H. Holliday
Professor of Philosophy, UC Berkeley

Will Fithian
Associate Professor of Statistics, UC Berkeley

Emerson Spartz
Co-founder, Nonlinear; Founder, Dose

Tim Colbourn
Professor of Global Health Systems, Epidemiology and Evaluation, UCL

Roman Yampolskiy
Associate Professor, University of Louisville

Geoffrey Miller
Associate Professor of Psychology, University of New Mexico

Mounim A. El-Yacoubi
Professor, Institut Polytechnique de Paris

Jon Warnow
Executive Director, Humans in Control

Charbel-Raphaël Ségerie
Executive Director, Center for AI Security (CeSIA)

Tyler Johnston
Founder and Executive Director, The Midas Project

Jason Van Beek
Chief Government Affairs Officer, Future of Life Institute

Maxime Fournes
CEO, PauseAI Global

Holly Elmore
Founder and Executive Director, PauseAI US

Joep Meindertsma
Founder, PauseAI

Mikhail Samin
Executive Director, AI Governance and Safety Institute

John Sherman
Founder, The AI Risk Network; Host, For Humanity Podcast

Rohan Prasad
Executive Director, Sapiens First

Shon Pan
AI Safety Advocate

Adam Oberman
Professor of Mathematics, McGill; Canada CIFAR AI Chair, Mila

Phil Aroneanu
Founder and Executive Director, Irrepliceable; Co-founder, 350.org

Carah Ong Whaley
University of Virginia

Robert Creamer
Partner, Democracy Partners

Dan Hendrycks
Executive Director, Center for AI Safety

## Sign the statement

Full name *

Email *

Your title
Optional

Affiliation
Optional

Sign the statement

We'll email you to confirm your signature. Your email address will not be published. Privacy policy

Contact

Media

media@safe.ai

General contact

misanthropic@safe.ai

Privacy policy