

Stanford



Max Lamparth

Postdoctoral Scholar, Infectious Diseases

Bio

BIO

Max is a postdoctoral fellow at the Center for International Security and Cooperation, the Stanford Center for AI Safety, the Stanford Existential Risks Initiative, and the Brainstorm Lab at the Department of Psychiatry and Behavioral Sciences at Stanford University. He is advised by Prof. Clark Barrett, Prof. Steve Luby, and Prof. Paul Edwards.

With his research, he wants to make AI systems more secure and safe to use. Specifically, he is focussing on improving the ethical behavior of language models, making their inner workings more interpretable, and increasing their robustness against misuse.

Max received his Ph.D. in August 2023 from the School of Natural Sciences at the Technical University of Munich and previously a B.Sc. and M.Sc. in Physics from the Ruprecht Karl University of Heidelberg.

STANFORD ADVISORS

- Stephen Luby, Postdoctoral Faculty Sponsor

LINKS

- Homepage: <http://www.maxlamparth.com>
- Twitter/X: <https://x.com/MLamparth>
- LinkedIn: <https://www.linkedin.com/in/maxlamparth/>

Teaching

COURSES

2024-25

- Introduction to AI Safety: CS 120 (Aut)
- Research in AI Alignment: CS 362 (Aut)

2023-24

- Introduction to AI Safety: CS 120, STS 10 (Spr)