CLTC
Center for Long-Term Cybersecurity
UC Berkeley

AI Security Initiative

Analyzing Global Security Implications of Artificial Intelligence

As the capabilities of AI systems increase, we are experiencing a dramatic shift in the global security landscape. For all their benefits, AI systems introduce new vulnerabilities and can yield dangerous outcomes — from the automation of cyberattacks to disinformation campaigns and new forms of warfare. AI is expected to contribute more than $15 trillion to the global economy by 2030, but these gains are currently poised to widen inequalities, stoke social tensions, and motivate dangerous national competition. The AI Security Initiative works across technical, institutional, and policy domains to support trustworthy development of AI systems today and into the future.

AI SECURITY INITIATIVE

"Artificial intelligence may bring enormous benefits to the world, but not unless we develop training, standards, and policies that guide us toward safe and beneficial development and use."

– Jessica Newman, Director, AI Security Initiative

WHO WE ARE

Housed in the UC Berkeley Center for Long-Term Cybersecurity (CLTC), the AI Security Initiative is a growing hub for interdisciplinary research on the global security implications of artificial intelligence. We support research and dialogue to help AI practitioners and decision-makers prioritize the actions they can take today that will have an outsized impact on the future trajectory of AI safety and security around the world. Collaborating with a network of UC Berkeley researchers, we work across disciplines to expose and manage new threats and opportunities. Our long-term goal is to help communities around the world thrive with safe and responsible automation and machine intelligence.

WHAT WE DO

The AI Security Initiative conducts independent research and engages with technology leaders and policymakers at state, national, and international levels, leveraging UC Berkeley's premier reputation and our location in the San Francisco Bay Area near Silicon Valley. Our activities include conducting and funding technical and policy research, and translating research into practice. We convene international stakeholders, hold policy briefings, publish white papers and op-eds, and engage with world-class partner organizations in AI safety, governance, and ethics.

Our research agenda focuses on three key challenges: 1) Vulnerabilities. How does a lack of reliability, robustness, accuracy, and transparency affect AI implementation? 2) Misuse: How do AI technologies cause intentional and unintentional harm? and 3) Power. How are global power dynamics shifting as different entities vie for AI leadership?

AI POLICY HUB

In 2022, together with the CITRIS Policy Lab (part of the Center for Information Technology Research in the Interest of Society), AISI co-founded the AI Policy Hub to cultivate an interdisciplinary research community to anticipate and address policy opportunities for safe and beneficial AI.

The hub annually supports cohorts of UC Berkeley graduate student researchers across an academic year who make meaningful contributions to research and recommendations for the AI policy landscape. Their work focuses on helping to reduce the harmful impacts, and amplifying the benefits of AI and its rapidly evolving technological capacities.

With faculty and staff mentorship, access to world-renowned experts, training sessions, a public symposium, and opportunities to meet directly with policymakers and other decision-makers, AI Policy Hub participants help chart a future in which AI technologies do not exacerbate division, harm, violence, and inequity, but instead foster human connection and societal well-being.

QUICK FACTS

15
UC Berkeley-affiliated graduate students supported since 2019

20+
publications with CLTC and external media

17
convenings, including policy briefings, workshops, and events

WHY WE NEED YOU

We're expanding our research team of faculty, students, and staff to tackle the emerging security challenges stemming from artificial intelligence. Your generous support will enable us to empower more researchers through targeted grants awards, and to scale our impact by advancing more cutting-edge research toward practical application.

$5,000+ funds a summer stipend for a graduate student to initiate a research project within the AI Security Initiative.

$50,000+ provides an AI Security Initiative fellowship for one graduate student to conduct and publish original research on the global security implications of AI.

$250,000+ establishes an endowed fellowship within the AI Policy Hub, providing critical funding resources to students advancing AI policy research.

For more information, contact:

Shanti Corrigan
Senior Director of Philanthropy
School of Information
510-693-8062
shanti@berkeley.edu

Jessica Newman
Director, Artificial Intelligence Security Initiative
Center for Long Term Cybersecurity
510-918-0070
jessica.newman@berkeley.edu

Scan to see the AISI website, or visit cltc.berkeley.edu/ai-security-initiative.

Berkeley SCHOOL OF INFORMATION