Stanford Institute for
Human-Centered Artificial Intelligence

POWERING THE MOONSHOT FOR HUMAN-CENTERED AI

The Stanford Institute for Human-Centered Artificial Intelligence (Stanford HAI) is an interdisciplinary initiative dedicated to advancing AI research, education, policy, and practice to improve the human condition. Drawing on expertise from faculty across the university, including Stanford's renowned computer science department, and collaborating closely with industry and government, the institute aims to guide the development of AI systems that are groundbreaking, ethical, and equitable—an AI moonshot to serve and uplift humanity.

AN AI LEAP FOR HUMANKIND

“We told the president now is America’s moonshot moment for AI. . . . Harnessing AI will be one of the defining tasks of the 21st century. The technology has the power to achieve once unfathomable feats like curing cancer. Reaching that potential will require deep, hands-on investment and collaboration to catalyze advancement, balanced by responsible stewardship of AI’s integration into society.”

—Fei-Fei Li on meeting with President Biden in June 2023, together with Rob Reich. Li is the Denning Co-Director of Stanford HAI and Sequoia Capital Professor of Computer Science; Reich, who serves as a HAI senior fellow, is the Marc and Laura Andreessen Faculty Director of the Stanford Center on Philanthropy and Civil Society and the McGregor-Girand Professor of Social Ethics of Science and Technology.

When humanity landed on the moon, a seemingly impossible feat was made possible by uniting the brightest minds across various disciplines. This moment in AI holds similar promise to expand the horizons of human potential. But to realize this potential, we must weave human-centered principles into every facet of AI development, making ethical, equitable, and impactful AI the standard, not the exception.

Our vision of an integrated, human-centered approach to AI addresses many uncertainties and risks, such as threats to privacy from data-driven platforms, the reinforcement of inequalities by biased algorithms, and automation’s impact on the future of work. Only thoughtful collaboration among diverse experts and stakeholders (industry and civil society) will enable us to make the leap to an AI future safely and successfully.

Stanford HAI is uniquely positioned to leverage Stanford’s world-class faculty, collaborative culture, and depth of knowledge across fields as disparate as computer science, engineering, human health, and public policy. We tackle the most pressing challenges our world faces today by investing in people (faculty and students), ideas (research grants), and global engagement (policy work).

By fostering interdisciplinary partnerships and providing pivotal resources—from seed grants that kick-start promising initiatives to Hoffman-Yee Grants that nurture transformative ideas—we pursue outside-the-box solutions. And because we recognize AI’s potential for unintended consequences, our multidisciplinary Ethics and Society Review panel assesses all funding proposals for societal risks.

We see the emerging AI landscape not just through the lens of algorithms and data but also through the eyes of experts who understand the human mind, the intricacies of social systems, and the nuances of ethical responsibility. Through our regular convenings, we incorporate the voices and concerns of individuals, industry, and government leaders to inform AI development. Bringing these perspectives together is the best way to navigate this dynamic frontier, addressing challenges once considered insurmountable and ensuring that AI enriches and respects human life.

LAUNCHING REAL-WORLD SOLUTIONS

Public Policy

Policymakers are crucial for AI's success but often lag behind industry and academia when it comes to understanding this technology. At Stanford HAI, we actively engage with state, national, and global leaders to shape policies that mitigate AI's risks and maximize its benefits. From serving as the State of California's chief AI partner to advising the United States Congress and participating in United Nations panels, our faculty influence policymaking at every level. In addition to advising policymakers, we have trained more than 80 congressional staffers through our bipartisan congressional boot camps and more than 8,000 government employees through symposia and other educational programs.

UNITED WE INNOVATE

Fei-Fei Li, John Etchemendy, and James Landay—the Stanford HAI Denning Co-Directors—were among the first to push for equitable access to powerful computational resources and large-scale government datasets for universities and nonprofits. They recognized that democratizing access to cutting-edge tools and data fuels innovation, enhances workforce development, and strengthens cybersecurity, so they championed the creation of a National AI Research Resource in the United States to balance the dominance of for-profit entities in the technology sector. Stanford HAI leaders galvanized support from universities and technology companies, drafting a blueprint for this initiative in 2021. In 2023, a bipartisan group of senators introduced the CREATE AI Act, which proposed establishing the National AI Research Resource (NAIRR) and led to the National Science Foundation's pilot program launch in 2024.

EXTINGUISHING MODERN SLAVERY

Funded by an HAI seed grant, Stanford's Human-Trafficking Data Lab developed an AI system that spots forced labor at deforestation sites in Brazil's Amazon rainforest. Both legal and illegal deforestation efforts routinely rely on forced labor, but the sites are remote and transitory, often abandoned before inspectors can reach them. The lab's system addresses that challenge by recognizing one of the telltale signs of an active deforestation site: large ovens that convert the fallen timber into charcoal. The researchers are working with the Brazilian Federal Labor Prosecution Office to develop a system that automatically notifies investigators, enabling rescuers to mobilize more rapidly.

AUDITING THE IRS FOR RACIAL BIAS

Daniel Ho, a senior fellow at Stanford HAI and the Wm. Benjamin Scott and Luna M. Scott Professor of Law, uncovered that the IRS is three to five times more likely to audit Black taxpayers due to flaws in its algorithmic audit selection process. This revelation sparked a congressional outcry and a commitment from the IRS to remedy the discrimination. Ho and his team developed predictive models to identify potential tax evasion more accurately, reducing unnecessary audits and concentrating resources on high-risk cases. "Researchers did a great service spotting the racial bias in the algorithms that guide audit selection," said Senate Finance Committee Chairman Ron Wyden.

Daniel Ho

Human Health

As society grapples with an aging population and a shortage of health care labor, AI has the potential to revolutionize medical diagnostics, treatment, and patient care. However, it also raises critical ethical and privacy concerns. Drawing on expertise from medicine, computer science, ethics, and policy, Stanford HAI aims to develop AI that enhances health care quality for all patients. The RAISE Health initiative, a partnership between Stanford Medicine and HAI, promotes ethical AI use in health care by fostering collaboration, sharing resources, and establishing standards for safe, equitable practices.

COMING CLEAN ABOUT HOSPITAL HYGIENE

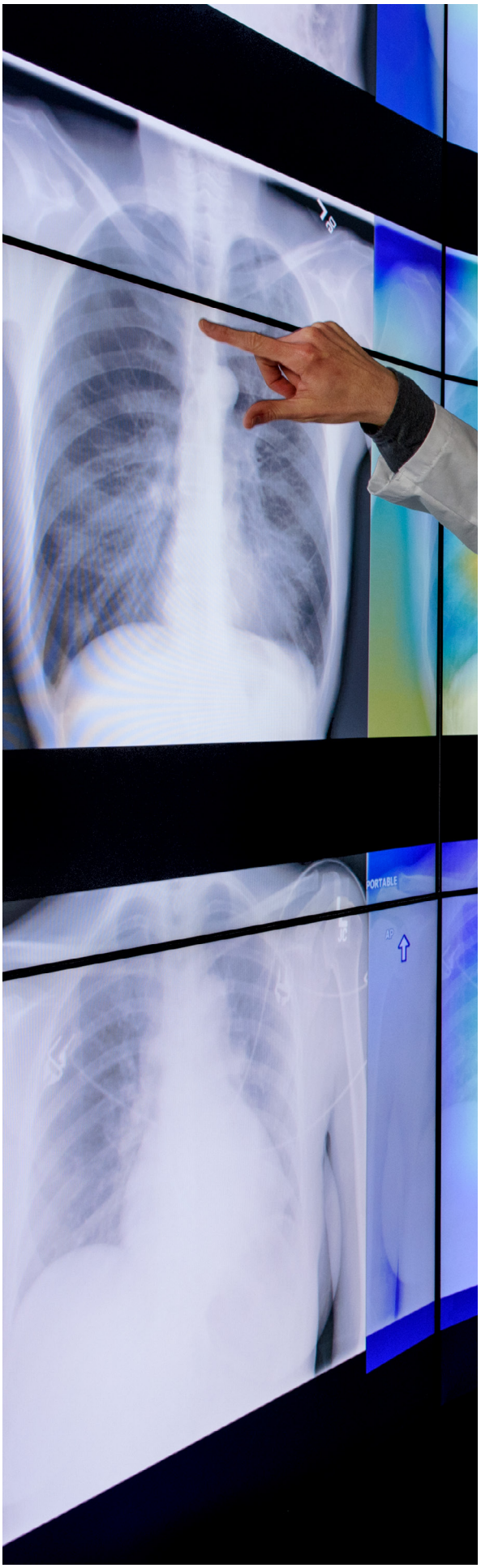
Hospital-acquired infections kill more than twice as many people as car accidents in the United States. In 2024 alone, more than 680,000 people in the United States developed infections during hospital stays, resulting in more than 72,000 deaths, compared to nearly 43,000 fatalities caused by car accidents. Can AI detect risks that humans often overlook without compromising patient and caregiver privacy? The Ambient Intelligence Lab strives to improve health and wellness by enabling smart physical spaces that use AI and privacy-preserving sensing to improve care and reduce costs while respecting human dignity. For instance, a simple network of sensors in hospitals can track staff hygiene compliance, sending real-time reminders to wash hands.

AN X-RAY IS WORTH A THOUSAND WORDS

During the COVID-19 pandemic, the Stanford Center for Artificial Intelligence in Medicine and Imaging (AIMI), led by HAI Senior Fellow Curt Langlotz, developed AI models to detect COVID-19 from chest X-rays and CT scans. These models were critical in managing the pandemic, particularly in resource-limited areas, by enabling rapid and accurate diagnosis, predicting disease progression, and assessing patient outcomes. Langlotz's frequently quoted witticism drives home AI's game-changing potential in this space: "'Will AI replace radiologists?' is the wrong question. The right answer is: 'Radiologists who use AI will replace radiologists who don't.'"

ROBOTS THAT HELP AT-RISK ADULTS FIND THEIR FOOTING

Falls are the leading cause of injury among adults over age 65, with about 36 million incidents reported each year in the United States, often resulting in devastating consequences. To combat this, Stanford HAI researchers are developing an AI-powered wearable exoskeleton designed to predict and prevent falls before they occur. Led by C. Karen Liu, a computer science professor and HAI affiliate, a multidisciplinary team is creating a robotic system that helps users maintain their balance in real time—from adjusting a step to avoid a stumble to detecting hazards like uneven ground. Supported by a Hoffman-Yee Grant, this project aims to turn AI into a proactive guardian for at-risk adults, fostering safer mobility and greater independence.



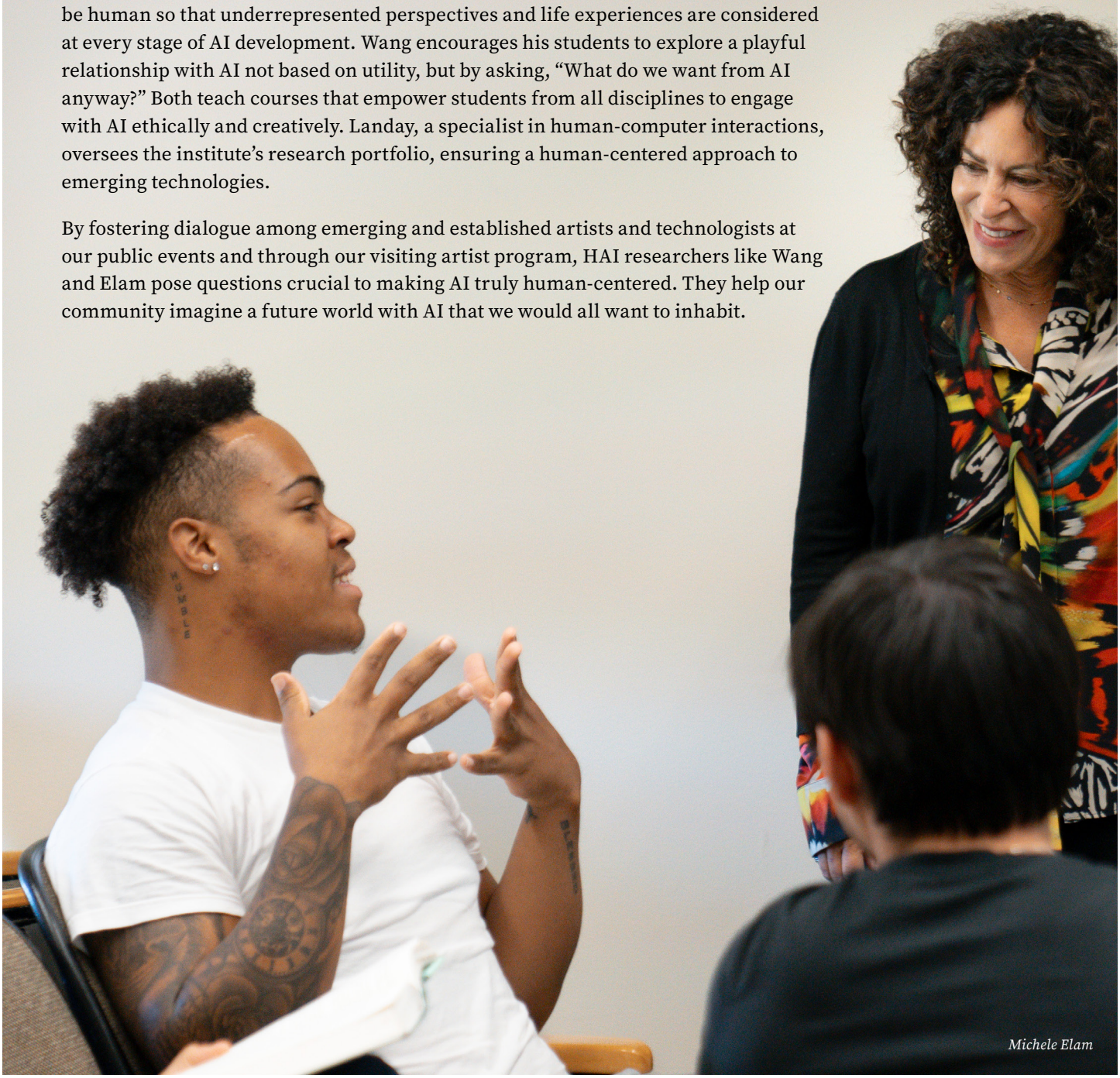
Humanity in the Loop

What is human-centered AI, and how can we design it? How can we advance AI while respecting the unique facets of human thought and creativity? What does it mean to be principled and pioneering?

James Landay, Denning Co-Director of Stanford HAI and the Anand Rajaraman and Venky Harinarayan Professor in the School of Engineering; HAI Associate Director Ge Wang, an associate professor of music; and HAI Senior Fellow Michele Elam, the William Robertson Coe Professor in the School of Humanities and Sciences, all address such concerns by bringing humanistic perspectives to AI research.

Elam urges us to draw on literature and art to expand our notion of what it means to be human so that underrepresented perspectives and life experiences are considered at every stage of AI development. Wang encourages his students to explore a playful relationship with AI not based on utility, but by asking, "What do we want from AI anyway?" Both teach courses that empower students from all disciplines to engage with AI ethically and creatively. Landay, a specialist in human-computer interactions, oversees the institute's research portfolio, ensuring a human-centered approach to emerging technologies.

By fostering dialogue among emerging and established artists and technologists at our public events and through our visiting artist program, HAI researchers like Wang and Elam pose questions crucial to making AI truly human-centered. They help our community imagine a future world with AI that we would all want to inhabit.



Michele Elam

Education

Stanford HAI equips students and professionals with the knowledge to develop and apply AI responsibly through interdisciplinary courses, workshops, and seminars integrating computer science with the humanities, social sciences, and ethics. In partnership with the Stanford Graduate School of Education and the Stanford Accelerator for Learning, HAI hosts the annual AI+Education Summit, uniting educators, industry executives, and researchers to explore and advance the role of AI in transforming education. HAI also prepares undergraduates to become future leaders in AI through a wide range of courses, such as AI Awakening *with Erik Brynjolfsson and* Human-Centered AI *with Peter Norvig.*

Teacher training

MATH MADE PERSONAL

Research shows that tailored instruction significantly improves students' academic and life outcomes. However, 70 percent of educators struggle to address diverse student needs due to time constraints and inadequate tools. Dora Demszky, an assistant professor in education data science, is addressing this by combining machine learning, natural language processing, linguistics, and practitioner input to create scalable, equitable, and student-centered educational tools. Supported by a Stanford HAI seed grant, Demszky partnered with a network of school districts to develop AI-powered resources that help middle school math teachers customize their curriculum to meet students' specific needs.

CODE AND CONDUCT—AT STANFORD AND BEYOND

Stanford HAI continues Stanford's legacy of educating tech pioneers by ensuring that future computer scientists understand the ethical implications of their work. In partnership with the Stanford Computer Science Department and the McCoy Family Center for Ethics in Society, HAI developed Embedded Ethics—curricular materials that integrate ethics into core undergraduate computer science courses. This interdisciplinary approach provides students with both technical expertise and an understanding of the ethical, societal, and policy implications of AI technologies. Stanford HAI freely shares Embedded Ethics resources with schools at Stanford as well as other universities to achieve the greatest impact.

SENDING AI BACK TO HIGH SCHOOL

"Regardless of whether someone anticipates a technical career, being an informed person in society means understanding AI," says Victor Lee, an associate professor in the Graduate School of Education and HAI affiliate. To prepare young people for AI's impact, Lee initiated Project CRAFT (Classroom-Ready Resources About AI for Teaching), offering free, multidisciplinary curriculum resources for high school teachers through the Project CRAFT website. The program awards co-design fellowships to high school teachers to ensure these resources will make a difference in the classroom. In addition, HAI's AI4ALL summer camp, launched in 2022, immerses underrepresented students in hands-on AI research and mentoring at Stanford to increase diversity among the field's future leaders. The program saw a 90 percent increase in applicants from 2023 to 2024, and the 2024 cohort was notably diverse, with four continents represented.

HAI AI4ALL summer camp

Erik Brynjolfsson

Economy

Customer service chatbots, supply chain optimizing robots, and rideshare apps—AI is already transforming the economy, and its influence is set to grow exponentially. Led by HAI Senior Fellow and Jerry Yang and Akiko Yamazaki Professor Erik Brynjolfsson, Stanford HAI's Digital Economy Lab brings together researchers and experts to explore how digital technologies are reshaping work, organizations, and the economy. By analyzing the impacts of AI, automation, and digital platforms on labor markets, productivity, and economic growth, the lab provides critical insights to inform policymakers, businesses, and the public on harnessing digital innovations for inclusive and sustainable economic development.

A HELPING HAND, NOT A PINK SLIP

How can we ensure that AI tools will augment rather than replace human workers? To inform AI implementation strategies in the workplace, the Digital Economy Lab studied the introduction of a generative AI conversational assistant using data from 5,179 customer support agents. The findings revealed a 14 percent increase in productivity on average, with the greatest benefits seen among novice and low-skilled workers due to the dissemination of knowledge from experienced colleagues. In addition, AI enhanced customer sentiment, reduced the need for managerial intervention, and improved employee retention, highlighting AI's potential to transform the workplace positively.

THE COST OF MISINFORMATION

When AI is used to amplify misinformation, the consequences can be severe—from exacerbating public health crises to inciting political violence. The Digital Economy Lab investigated the financial mechanisms fueling the spread of online misinformation, focusing on how advertising revenue sustains misinformation outlets. The research uncovered that many companies inadvertently advertise on these outlets due to automated ad distribution. Furthermore, public awareness of a company's unintentional role in funding misinformation can decrease demand for its products. The lab's proposed low-cost, scalable interventions for digital platforms aim to curb the monetization of misinformation and reduce its spread online.

Stanford solar generating station

Sustainability

Can AI's power be harnessed to help mitigate a challenge as great as climate change? At Stanford HAI, we leverage AI to address critical sustainability issues such as energy efficiency and environmental protection with participants from HAI, the Stanford Doerr School of Sustainability, and researchers across Stanford and beyond.

LOOKING BEFORE WE LEAP INTO MINING PROJECTS

As clean energy transitions accelerate, the need for critical minerals for battery metals—sourced primarily from rainforests—raises environmental and human concerns. Initially funded by a Stanford HAI seed grant, the Mineral-X Initiative enables companies to assess water, energy, and land uses to identify unsuitable projects in the early exploration stages rather than at the later mining stages. The initiative also works to safely and efficiently capture and store carbon emissions.

CARBON-CONSCIOUS CODING

The full suite of experiments needed to build and train an off-the-shelf AI language processing system can generate 78,000 pounds of carbon emissions. In collaboration with Facebook AI Research, Stanford HAI-affiliated faculty created an easy-to-use tool to measure AI projects' electricity use and carbon emissions, helping developers steer toward greener algorithms.

IT TAKES ALL OF US, INCLUDING YOU.

Stanford University
Human-Centered
Artificial Intelligence

Imagine a world where AI helps eradicate hospital-acquired infections, tailors education to individual student needs, and protects human dignity. At Stanford HAI, we are turning these possibilities into reality. We are reaching for the stars in AI innovation, harnessing the brilliance of interdisciplinary teams to revolutionize technology for the good of humanity.

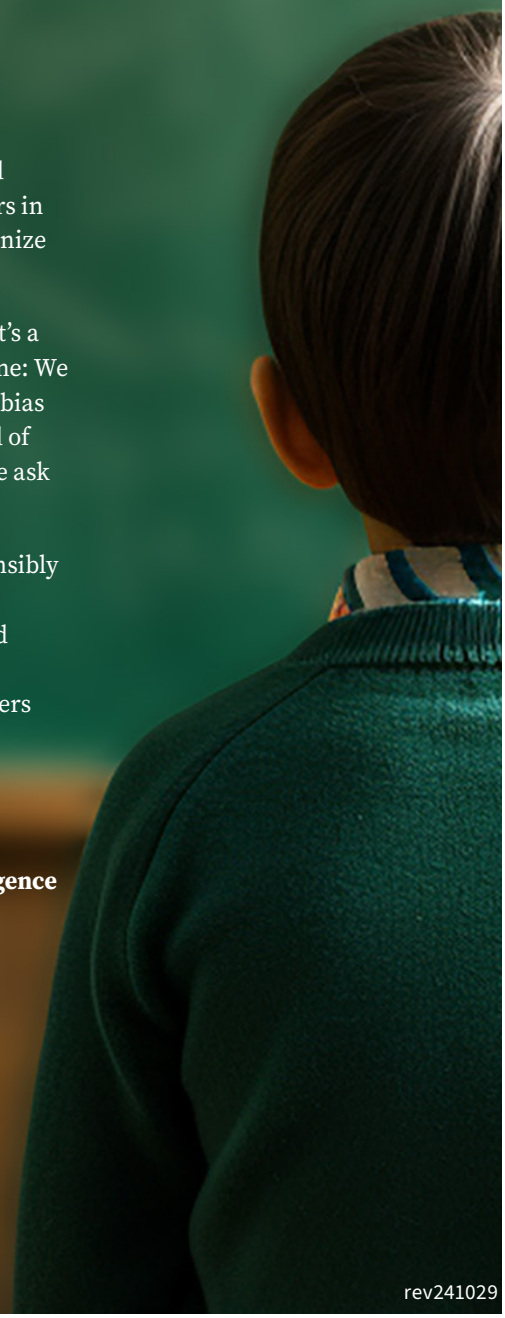
Aligning AI's remarkable advances with human values isn't just a goal for us—it's a moral imperative. Our success isn't measured by what happens at Stanford alone: We aim to ensure that people worldwide benefit from AI. Whether reducing racial bias in algorithmic systems, advancing sustainable practices, or curbing the spread of misinformation, Stanford HAI strives to create solutions felt by all. To do so, we ask for your support in this moonshot moment.

We invite you to join our community dedicated to harnessing AI's power responsibly and equitably. By partnering with Stanford HAI, you can help expand our groundbreaking research, foster innovative and urgently needed solutions, and cultivate the next generation of leaders who will drive ethical AI development. Together, we can shape an AI-enabled future that respects human dignity, fosters inclusivity, and drives positive change worldwide. Please join us.

For more information, please visit hai.stanford.edu

Stanford Institute for Human-Centered Artificial Intelligence
Gates Computer Science Building
353 Jane Stanford Way
Stanford, CA 94305

Contact: Stacey Sickels Boyce
Senior Associate Director for Development
Stanford Institute for Human-Centered Artificial Intelligence
sickels@stanford.edu
(650) 497-3772



rev241029