Center for AI and Digital Policy

CAIDP Update 7.25

June 30, 2025

Welcome to CAIDP Update!

NATO establishes 5% defense spending target with AI integration timeline while federal judges deliver narrow wins to Anthropic and Meta but warn copyright training remains likely illegal.

China launches AI Safety Institute as DeepSeek rivals US models, prompting congressional "Cold War" rhetoric and bipartisan legislation banning adversarial AI systems.

The EU faces high-stakes AI Code battle as a coalition including Nobel laureates, academics and civil society organizations warns against industry pressure to weaken safety rules before August deadline. India releases comprehensive AI policy templates for organizational deployment.

CAIDP supports Dutch authority rejecting scraped web data under GDPR and calls for Israel to ban rights-violating AI in public services. The Center urges federal privacy laws for healthcare AI and stronger oversight of authoritarian systems while opposing Senate provisions blocking state AI safeguards.

CAIDP President Hickok delivers train-the-trainer sessions on AI literacy for civil servants at UNESCO's 3rd Global Forum on AI Ethics in Bangkok.

AI POLICY NEWS

NATO Sets 5% Defense Spending Target, Establishes AI Integration Timeline

NATO members agreed to increase defense spending to 5% of GDP by 2035, with up to 1.5% allocated for technology and infrastructure development, according to The Hague summit declaration.

The commitment divides spending between core defense requirements (3.5% of GDP) and areas including innovation, critical infrastructure protection, civil preparedness, and defense industrial base strengthening. Members will submit annual plans showing incremental progress toward these targets, with a review

scheduled for 2029.

The alliance endorsed a Rapid Adoption Action Plan aimed at integrating new technologies into armed forces within 24 months. The plan seeks to accelerate national procurement processes and create NATO approval standards.

AI represents one of nine priority technology areas, alongside autonomous systems, quantum technologies, biotechnology, space, hypersonic systems, novel materials, energy systems, and communications networks. NATO's Data and Artificial Intelligence Review Board is tasked with developing responsible use principles and certification standards.

The alliance operates a €1 billion Innovation Fund for investments in defense technology startups. The Defence Innovation Accelerator for the North Atlantic (DIANA) comprises over 180 test centers and 23 accelerator sites across member countries.

The new spending targets represent a significant increase from the current 2% of GDP benchmark. Direct military support to Ukraine will count toward members' defense spending calculations under the new framework.

Judges Hand AI Companies Narrow Wins While Warning Practice Likely Illegal

Two federal judges delivered procedural victories to Anthropic and Meta last week in copyright lawsuits over AI training, but both emphasized their rulings were narrow and warned that using copyrighted books without permission will likely be illegal in most cases.

Judge William Alsup ruled that Anthropic's training of its Claude AI constituted fair use, writing that "the copies used to train specific LLMs were justified." However, Alsup ruled that Anthropic still faces trial for using pirated books, stating "Anthropic had no entitlement to use pirated copies for its central library."

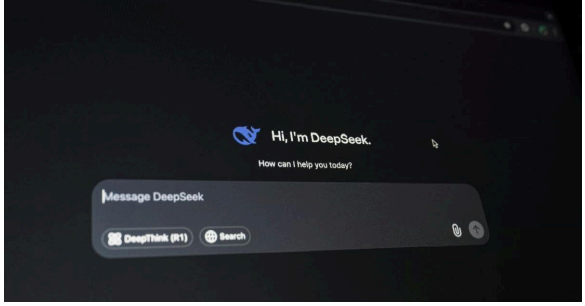
Judge Vince Chhabria dismissed a lawsuit by authors including Sarah Silverman and Junot Díaz against Meta over its Llama models. But Chhabria emphasized his ruling "does not stand for the proposition that Meta's use of copyrighted materials to train its language models is lawful." Meta won only because "these plaintiffs made the wrong arguments and failed to develop a record in support of the right one."

Chhabria warned that "in many circumstances it will be illegal to copy copyright-protected works to train generative AI models without permission" and that AI "has the potential to flood the market with endless amounts of images, songs, articles, books, and more."

The legal rulings coincided with a direct challenge to the publishing industry itself, as an open letter from authors demanded a pledge to reject books created with AI tools built on what they call "stolen voices."

China Builds AI Safety Institute While

DeepSeek Rivals US Tech Giants



China launched its AI Safety and Development Association in February 2025, positioning itself in the global AI governance conversation as its models increasingly rival American counterparts, according to a Carnegie Endowment for International Peace analysis.

The organization primarily focuses on international engagement rather than domestic model testing, the Carnegie report found. Led by former vice foreign minister Fu Ying and Turing Award winner Andrew Yao, the organization represents institutions including Tsinghua University and the Beijing Academy of Artificial Intelligence.

Chinese startup DeepSeek shocked the tech world in January 2025 by creating an AI system that matched or exceeded the performance of leading American AI models from Meta and Anthropic. The model, which cost just $5.6 million to develop according to a World Economic Forum article, was a central point of concern for witnesses at the House Select Committee on China's June 25 hearing.

The achievement demonstrated China's ability to produce cutting-edge AI technology at a fraction of Western costs. According to the World Economic Forum, China now files four times as many AI patents as the United States. While the U.S. maintains a lead in producing top AI models, China leads in the overall number of AI publications.

The developments prompted bipartisan legislation (No Adversarial AI Act) to ban Chinese AI systems from federal agencies, the House Select Committee on China announced. "We are in a new Cold War, and AI is the strategic technology at the center," Committee Chairman John Moolenaar said during the hearing.

EU Faces High-Stakes Battle Over AI Code as Nobel Laureates Warn Against Industry Pressure

The European Commission will stage a last-ditch workshop this week to convince AI companies to sign its Code of Practice before the August 2 deadline, as Nobel laureates, academics and civil society organizations warn against bowing to industry pressure.

Major tech firms continue resisting the voluntary rules amid transatlantic tensions. The July 2-3 workshops will present the final code, including copyright provisions that haven't been circulated during the nine-month drafting process.

Meta has already refused to sign. US tech giants including Amazon, Google, Microsoft and OpenAI urged the EU to keep the code "as simple as possible," according to documents seen by Euronews.

The resistance comes as a coalition including Nobel laureates Daron Acemoglu and Geoffrey Hinton, UC Berkeley's Stuart Russell, a CAIDP board member, and 40+ academics and civil society organizations sent a June 25 letter to Commission President Ursula von der Leyen, warning that "major GPAI model providers have drastically scaled back transparency and the rigour of safety-testing around model releases."

The Center for AI and Digital Policy, a plenary participant, warned the March draft "moves further away from the spirit of the EU AI Act" by making fundamental rights assessments optional rather than mandatory.

The coalition argued that weakening rules "shifts the burden of AI Act compliance to startups and SMEs." They called for mandatory third-party testing and scaling the AI Safety unit to 100 staff.

India Releases AI Policy Template for Organizations

The Data Security Council of India published an AI Security, Privacy, and Governance Policy template, outlining requirements for organizations deploying AI systems.

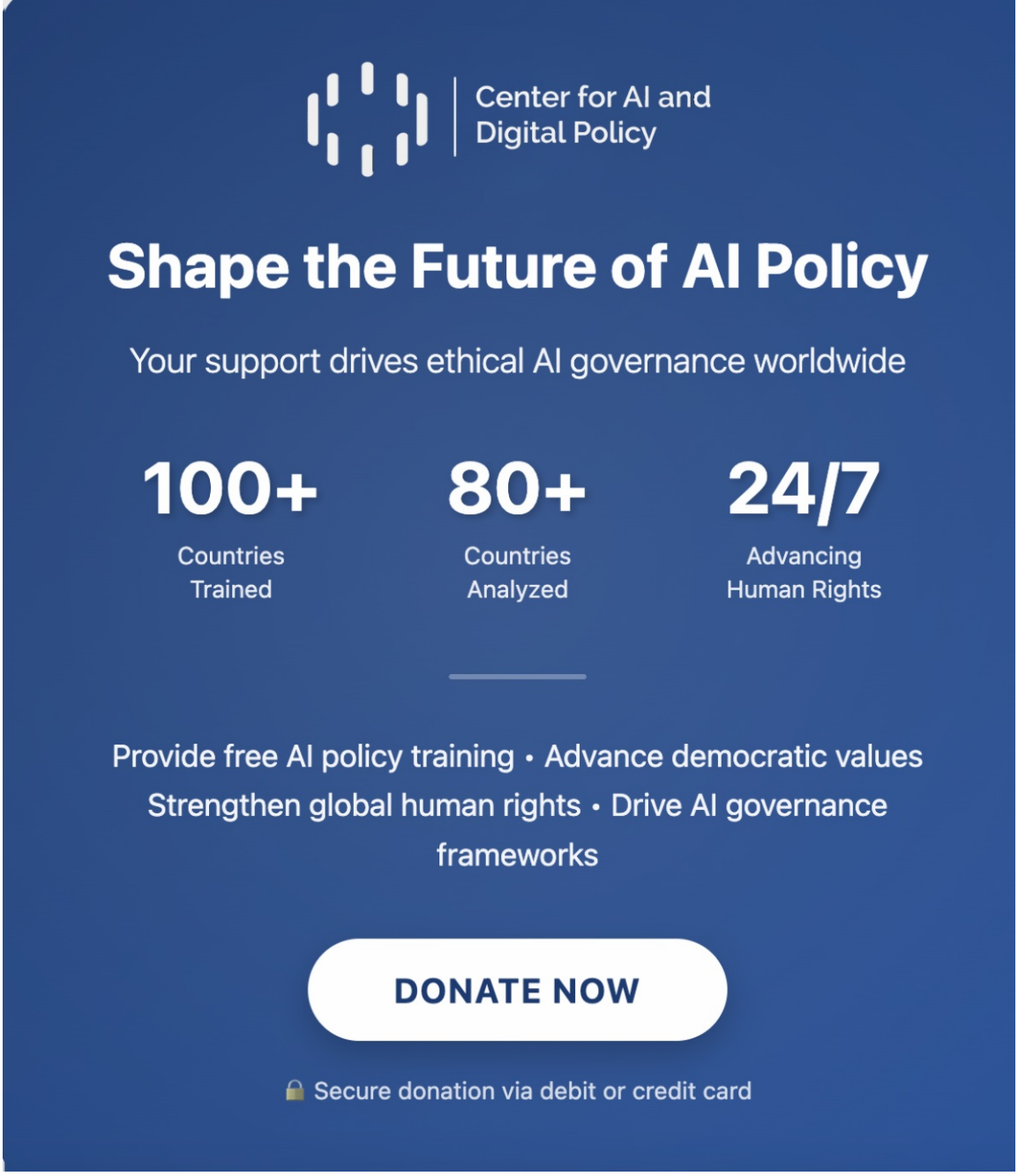
The template covers machine learning, generative AI, and agentic AI, autonomous systems capable of independent decision-making. Organizations must form AI Governance Committees with representatives from IT, risk, compliance, legal, and business units to oversee implementation.

Key requirements include documenting data sources, model selection criteria, and validation results. High-risk applications such as loan approvals and fraud detection require human oversight with defined intervention protocols. Organizations must implement safeguards against prompt injection, data poisoning, and unauthorized model extraction.

The framework treats policies as "living documents" requiring regular updates. It differentiates between enterprise and consumer-facing AI, with specific transparency obligations for each. Consumer applications must disclose AI interaction and provide feedback mechanisms for reporting errors or bias.

For autonomous AI systems, the template mandates boundary constraints, fail-safe mechanisms, and accountability structures for emergent behaviors. It requires monitoring for cascading failures and unintended agent interactions.

Organizations can select applicable provisions based on their contexts. The template includes guidance on data governance, secure deployment, input/output controls, and privacy protections throughout the AI lifecycle.



CAIDP ACTIONS

CAIDP Urges Dutch Authority to Close Generative AI Governance Gaps

The Center for AI and Digital Policy submitted comments to the Dutch Data Protection Authority June 27, praising its "proactive stance on regulating generative AI technologies to uphold human rights and democratic values."

CAIDP "fully supports the AP's assessment that foundation models trained on scraped web data, especially containing special categories of personal data, do not meet the requirements of GDPR lawfulness, particularly under Articles 6 and 9." The Center backed the AP's critique of the "legitimate interest loophole."

The submission identified accountability gaps, noting that "while the AP's framework

clearly distinguishes between developers and deployers as data controllers, it remains unclear how the role of joint controllers and data processors is defined in generative AI contexts."

Key recommendations included requiring "transparent data provenance tracking and opt-out mechanisms," establishing "verifiable anonymization" standards, and ensuring "data subjects must have effective redress mechanisms, including the ability to challenge automated decisions and seek removal or compensation where data is used unlawfully."

CAIDP Calls for Red Lines on AI Systems in Israeli Public Sector

The Center for AI and Digital Policy filed comments to Israel's National Digital Directorate June 30, urging the government to prohibit AI systems that violate human rights in its Guide to Responsible Use of AI in the Public Sector.

CAIDP recommended Israel "designate systems that lack scientific validity and undermine human dignity and rights as 'unacceptable' AI systems," specifically naming social scoring, emotion recognition, biometric categorization, and predictive policing.

The Center called for establishing "red lines against AI used which invade privacy, are used in surveillance, and/or for social scoring" as a baseline for protecting fundamental rights.

CAIDP urged mandatory "independent impact assessments prior to deployment of AI systems and throughout the AI lifecycle," proposing a "go/no go" approach where systems presenting unacceptable risks would not be deployed.

CAIDP called on Israel to ratify the Council of Europe AI Treaty it signed, implementing "robust public documentation practices, contestability mechanisms, and redress pathways for affected individuals."

CAIDP Calls for Federal Privacy Law to Address AI Healthcare Risks

The Center for AI and Digital Policy submitted a statement to the House Ways and Means Committee June 25, urging federal data privacy legislation and stronger oversight of AI systems using digital health data.

CAIDP noted that while AI could reduce healthcare spending by $200-360 billion annually, public trust is eroding. Though 66% of physicians now use AI tools, double from 2023, a 2025 study found 65.8% of respondents don't trust healthcare providers to use AI responsibly, and 60% would be uncomfortable with providers relying on AI for patient care.

The statement highlighted significant privacy gaps. Period-tracking app Flo secretly shared 100 million users' health data with Facebook, while AI lab DeepMind accessed patient data without consent. Neither faced penalties under current health privacy laws because they aren't direct healthcare providers.

CAIDP recommended enacting comprehensive federal privacy legislation to close regulatory loopholes and requiring agencies to follow White House guidance mandating "pre-deployment testing, impact assessments, ongoing monitoring, and human oversight" for high-impact AI.

The Center urged the FDA to finalize its AI-enabled medical device guidance and commended federal civil rights officials for clarifying that anti-discrimination laws prohibit AI-based discrimination in healthcare settings.

CAIDP Urges Congress to Counter Authoritarian AI and Lead Global Governance

The Center for AI and Digital Policy sent a statement to the House Select Committee on Strategic Competition with China June 26, urging stronger oversight of AI systems that threaten civil liberties and American leadership in global AI governance.

CAIDP commended the committee's bipartisan No Adversarial AI Act, which would bar federal agencies from using AI developed by adversaries like China.

CAIDP's 2025 AI Index found China uses AI primarily for control and surveillance rather than protecting human rights. The Center cited DeepSeek's collection of American users' data, including chat history and typing patterns, which flows directly to China.

The statement highlighted regulatory gaps allowing both Chinese and American companies to deploy AI without meaningful oversight. CAIDP noted that U.S. firms offshore data work to Chinese companies and that American open-source models power Chinese AI development.

CAIDP's recommendations included ratifying the Council of Europe AI Treaty, maintaining U.S. leadership in international AI standards organizations, prohibiting AI systems that violate civil liberties, and enforcing the TikTok ban.

The Center warned that without proper governance, AI systems can manipulate public opinion and compromise Americans' personal data.

Opposition Mounts Against Federal AI Safeguards Moratorium

Growing bipartisan opposition has emerged against a provision in Senate budget legislation that would restrict states from establishing AI safeguards,

with Senator Marsha Blackburn announcing plans to introduce an amendment removing the measure reported by Axios.

Blackburn will seek to strike what she called a "misguided provision" from the final bill. Republican Senators Blackburn, Josh Hawley, and Rand Paul signed a letter urging the Senate to remove the language, while all Democratic senators are expected to vote against the moratorium.

The provision would limit states' authority to implement AI regulations and safeguards at a time when federal AI legislation remains stalled. Critics argue this would create a dangerous regulatory vacuum as AI systems rapidly advance.

CAIDP leadership has been vocal in opposing the measure. In May, CAIDP Executive Director Marc Rotenberg, President Merve Hickok, and Associate Director Christabel Randolph wrote in Tech Policy Press that the proposed moratorium was "short-sighted and ill-conceived."

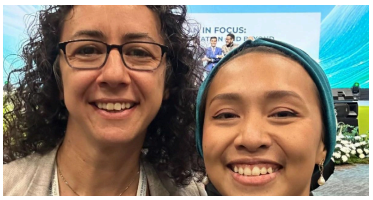
The opposition reflects growing concern that preempting state AI regulations without corresponding federal protections would leave Americans vulnerable to AI harms. As budget negotiations continue, the fate of the provision remains uncertain, with mounting pressure from both parties to remove it.

CAIDP at UNESCO AI Ethics Forum, Bangkok

The 3rd UNESCO Global Forum on the Ethics of AI gathered governments, civil-society groups, and industry to advance ethical, community-benefiting AI. Highlights led by CAIDP President Merve Hickok and our alumni:

• Train-the-Trainer sessions on AI Literacy for Civil Servants delivered by President Hickok
• RAM (Readiness Assessment Methodology) reports from the AI Ethics Experts Without Borders (AIEB) network and partner governments
• Launch of the Global Network of AI Supervising Authorities
• Launch of the Global CSO & Academic Network on AI Ethics
• AI and the Judiciary panel

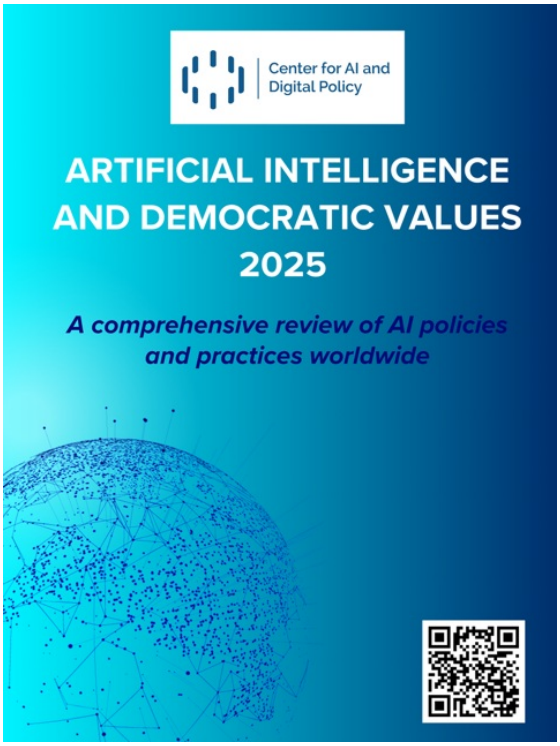
The accompanying photos capture President Hickok with CAIDP alumni.

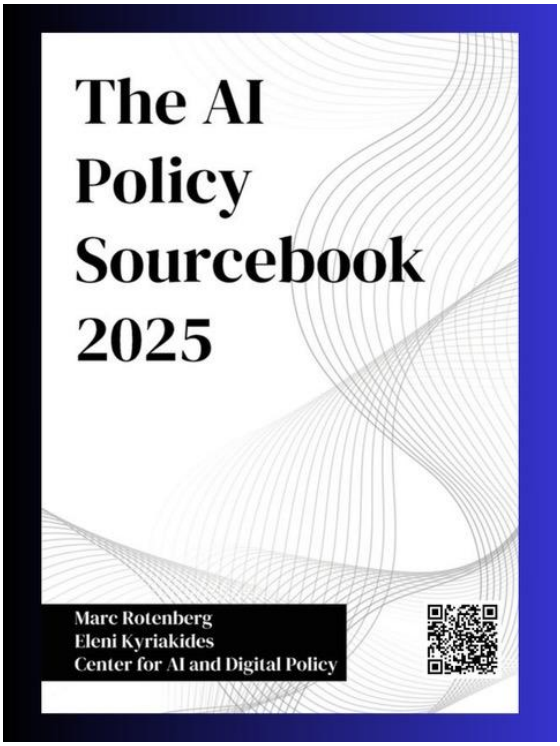


CAIDP AI POLICY EVENTS

- UNESCO, 3rd UNESCO Global Forum on the Ethics of AI with Merve Hickok (CAIDP), Thailand, Jun 24-27, 2025
- Conference of the European Society of Criminology with Marc Rotenberg

(CAIDP), Athens, Greece, Sept. 4, 2025
- Global Privacy Assembly, Seoul, Korea, September 16-19, 2025
- G20, Digital Economy Working Group Ministerial Meeting, Cape Town, Sept 25-26, 2025





Subscribe to CAIDP | Let us know your suggestions here.
| Privacy Policy | Contact | Donate |
