INSIDE AI POLICY

Exclusive news on the burgeoning debate over regulating artificial intelligence

Vol. 3, No. 41 — October 14, 2025

Senators tout work on AI-related provisions as defense policy bill advances

Posted October 10, 2025

Senators highlighted artificial intelligence elements in their version of the fiscal 2026 National Defense Authorization Act, notably including new restrictions on AI investments in China contained in an amendment by Sens. John Cornyn (R-TX), Catherine Cortez Masto (D-NV) and a bipartisan group of a dozen other senators.

The Senate approved the NDAA package Oct. 9 on a 77-20 vote. The House approved its version in September and negotiators will now work through differences in the bills. Senators added several amendments by voice vote prior to passage, including the Foreign Investment Guardrails to Help Thwart (FIGHT) China Act.

"The need to address capital flowing from the U.S. to the People's Republic of China was realized during the first Trump administration, and the *FIGHT China Act* presents a generational opportunity to confront the threat China poses to our national and economic security," Cornyn said in an Oct. 10 statement. "This landmark legislation would prohibit and require notification of U.S. investments in certain technologies in China, ensuring American ingenuity, innovation, and investment do not end up in the hands of the Chinese Communist Party to be weaponized against us."

He said, "I'm proud the *FIGHT China Act* has received support in the Senate as part of our annual defense bill, and I look forward to it being included in the final version that heads to President Trump's desk to become the law of the land."

Cortez Masto said, "The United States should not be lending a helping hand to our adversaries in Communist China by investing in their development of artificial intelligence, quantum computers, or military technologies. I am proud of the work I have done with my colleagues on both sides of the aisle to get this critical national security bill passed by the Senate."

Cornyn's office said the FIGHT China Act would "prohibit covered investments in the PRC for development or production of:"

- *Certain advanced integrated circuits;*
- *Certain AI models capable of a high number of operations;*
- *Quantum computers and supercomputers;*
- *Materials or components for hypersonics; and*
- *Any of these technologies that are on the Munitions List, intended for use with nuclear equipment or facilities, or emerging technologies subject to export controls.*

And, its provisions "would require U.S. persons to notify the U.S. Department of the Treasury within 14 days when making a covered investment in the PRC for the development or production of:"

- *Any non-prohibited integrated circuit;*
- *And any non-prohibited AI system which is used for military, surveillance, cybersecurity, penetration, forensics, or robotic system use or which meets a certain computing standard.*

In addition, the Cornyn release said, the bill "would cover the following investments:"

- *Acquisitions, including of limited partners, equity interest, property, or other assets;*
- *Loans and debt financing;*
- *Joint ventures;*
- *And equity interest or debt conversions.*

The release said the legislation contains exemptions for:

continued on next page

In This Issue . . .

• *Transactions determined to be de minimis or in the national interest;*

• *Investments in securities, derivatives of securities, or made as a limited partner in a venture capital fund, private equity fund, fund of funds, or other pooled investment fund;*

• *Ancillary transactions undertaken by a financial institution;*

• *Acquisitions of entire assets or entities located outside the PRC;*

• *Certain transactions secondary to a covered national security transaction;*

• *And certain ordinary or administrative business transactions.*

Sponsors of the measure include Minority Leader Charles Schumer (D-NY), Banking Chairman Tim Scott (R-SC) and ranking member Elizabeth Warren (D-MA), and Sens. Dan Sullivan (R-AK), Pete Ricketts (R-NE), Andy Kim (D-NJ), Dave McCormick (R-PA) and Michael Bennet (D-CO).

The legislation is cosponsored by Sens. Jim Banks (R-IN), Elissa Slotkin (D-MI), Bill Hagerty (R-TN), and John Fetterman (D-PA).

More AI amendments

Armed Services cyber subcommittee Chairman Mike Rounds (R-SD) pointed to AI provisions in the bill prohibiting "any use on any DOD system of a large language model created by a Chinese or other covered entity, specifically DeepSeek" and "contractor use of these models."

A Rounds release said provisions in the bill that he sponsored or cosponsored include creation of "a Cyber Command A.I. industry collaboration roadmap."

"This requires the Commander of U.S. Cyber Command, in coordination with senior DOD A.I. and research officials, to develop a roadmap for industry and academic collaboration on AI-enabled cyber capabilities for cyberspace operations," the release said.

The bill also "creates an Artificial General Intelligence Steering Committee," the Rounds release said.

"This establishes an Artificial General Intelligence Steering Committee in the DOD to analyze AI trajectories and develop DOD adoption strategies with Deputy Secretary of Defense and Vice Chairman of the Joint Chiefs of Staff as co-chairs," according to the release.

Sen. John Hickenlooper (D-CO) in a release noted his backing for the Cyber Command AI roadmap and AGI steering committee amendments, and his sponsorship of a provision directing "DoD to facilitate the integration of commercially available AI capabilities into logistics operations."

The Senate NDAA bill also includes the "Guaranteeing Access and Innovation for National Artificial Intelligence," or GAIN AI Act, by Sen. Banks requiring U.S. chip designers to certify that no potential U.S. customer would be left waiting, before they can receive a government license for exporting their chips to China.

"In my first NDAA as a senator on the Armed Services Committee, I'm proud to have secured real wins that put America first and keep us ahead of our adversaries like Communist China," Banks said in a release. "My amendments give American companies first access to advanced AI, protect our research from China, speed up military acquisition, and help military kids stay focused in school."

Banks in a separate release noted $5 million for AI research at Trine University in Indiana, saying it will support the Artificial Intelligence and Maritime Maneuvering program "that will test and prototype autonomous naval vessels."

SUBSCRIPTIONS:
703-416-8505 or
800-424-9068
custsvc@iwpnews.com

NEWS OFFICE:
703-416-8500
Fax: 703-416-8543
aipolicy@iwpnews.com

Managing Editors: Charlie Mitchell (cmitchell@iwpnews.com)
Rick Weber (rweber@iwpnews.com)

Production Manager: Lori Nicholson (lori.nicholson@iwpnews.com)
Production Specialists: Daniel Arrieta (darrieta@iwpnews.com)
Michelle Moodhe-Page (mmoodhe-page@iwpnews.com)

Inside AI Policy is published every Tuesday by Inside Washington Publishers, P.O. Box 7167, Ben Franklin Station, Washington, DC 20044.

EU announces €1 billion for increasing adoption in strategy to lead AI development

Posted October 14, 2025

The European Commission is committed to spending about €1 billion to increase the uptake of artificial intelligence as part of a twin set of plans to propel the bloc to the front of the line in the global competition to guide establishment of the technology.

"I want the future of AI to be made in Europe," Ursula von der Leyen, president of the European Commission, said in an Oct. 7 release rolling out the plans. "Because when AI is used, we can find smarter, faster, and more affordable solutions. AI adoption needs to be widespread, and with these strategies, we will help speed up the process."

OpenAI is among the companies already looking to take advantage of the plan. The U.S. developer released a series of proposals following a Sept. 23 policy workshop with European partners in advance of the strategy's announcement. Among much else, the proposals call for "Supporting the public sector transformation with agents."

Von der Leyen said, "Putting AI first also means putting safety first. We will drive this 'AI first' mindset across all our key sectors, from robotics to healthcare, energy and automotive."

The €1 billion commitment is part of the EU's Apply AI Strategy, which "aims to harness AI's transformative potential by driving adoption of AI across strategic and public sectors including healthcare, pharmaceuticals, energy, mobility, manufacturing, construction, agri-food, defence, communications and culture. It will also support small and medium-sized enterprises (SMEs) with their specific needs and help Industries integrate AI into their operations."

"Concrete measures include establishing AI-powered advanced screening centres for healthcare and supporting the development of frontier models and agentic AI tailored to sectors such as manufacturing, environment and pharmaceuticals," according to the release.

The adoption plan is part of what the commission is describing as an "AI first policy" reflected in an April "AI continent action plan" that it says "also addresses cross-cutting challenges: accelerating time-to-market by linking infrastructure, data, and testing facilities; strengthening the EU workforce to be AI ready across sectors; and launching a Frontier AI initiative to support innovation by bringing together Europe's leading AI actors.

"The renewal and deployment of the network of European Digital Innovation Hubs, transformed into Experience Centres for AI, will give companies privileged access to the EU AI innovation ecosystem," the release said, also noting a new "service desk" to help with the implementation of the technology.

The Apply AI Strategy was paired with an AI in Science Strategy which the release said "positions the EU as a hub for AI-driven scientific innovation. At its centre is RAISE - the Resource for AI Science in Europe, a virtual European institute to pool and coordinate AI resources for developing AI and applying it in science."

The science strategy includes "measures to attract global scientific talent and highly-skilled professionals to 'Choose Europe.'" It will put €58 million toward "the RAISE pilot for Networks of Excellence and Doctoral Networks to train, retain and attract the best AI and scientific talent," the release said.

In addition, there would be "€600 million from Horizon Europe to enhance and expand access to computational power for science," under the plan. "This investment will secure dedicated access to AI Gigafactories for EU researchers and startups," according to the release.

The plan also touts the goal of "doubling Horizon Europe's annual investments in AI to over €3 billion, including doubling funding for AI in science" and "support for scientists to identify strategic data gaps and gather, curate and integrate the datasets needed for AI in science."

The EU has come under fire from the AI industry and its supporters in conjunction with the new AI Act and privacy rules under the General Data Protection Regulation. And policymakers there have recently suggested they might delay the enforcement of the AI Act pending standards development for conducting risk assessments.

"To harness the full potential of AI, Europe must ensure seamless access to high-quality, structured data," the release said. "The Commission will present a Data Union Strategy at the end of October to better align data policies with the needs of businesses, the public sector and society."

It added: "The upcoming AI in Science Summit (Copenhagen, 3-4 November 2025), co-organised by the Commission and the Danish Presidency, will bring together policymakers, researchers and industry. It will present and launch initiatives under the AI in science Strategy, including the RAISE pilot and a private sector pledging campaign."

"Around the world, countries are racing to harness the potential of Artificial Intelligence," the release said. "Today, the European Commission set out two strategies to ensure Europe stays ahead."

Key AI leaders in the EU have said their vision is to collaborate rather than compete with U.S. allies, but observers have noted there is "definitely" an aspect of protectionism in its investment plans.

"Just six years ago, Europe had two supercomputers in the global top 10; today it has four and work is ongoing to set up at least 4 to 5 gigafactories," the release said. "Building on its strong AI infrastructure, as well as Europe's talent, vibrant research and innovation ecosystem and startups, its tradition of collaborative science, high-quality data and world-class research and technology infrastructures, the EU is well positioned to accelerate the use of AI in key sectors and science."

Senate passes defense policy bill with AI provisions on investments, exports

Posted October 10, 2025

The Senate has cleared its version of the fiscal 2026 National Defense Authorization Act, including AI investment restrictions and a new control on sales of AI chips to China, setting up negotiations with the House on a final version of the annual defense policy bill.

The Senate approved the NDAA package late Oct. 9, on a 77-20 vote, after adding by voice vote several amendments including new restrictions on AI and other advanced-tech investments in China by Sen. John Cornyn (R-TX).

The underlying bill also contains the "Guaranteeing Access and Innovation for National Artificial Intelligence," or GAIN AI Act, by Sen. Jim Banks (R-IN) requiring U.S. chip designers to certify that no U.S. entity would be left waiting to be supplied in order to receive a government license for exporting their chips to countries like China.

Another amendment added to the bill, by Sen. Ruben Gallego (D-AZ), would "require a report on the feasibility of implementing artificial intelligence into anti-money laundering investigations relating to activity by foreign terrorist organizations, drug cartels, and other transnational criminal organizations."

Senate Armed Services leaders said the bipartisan NDAA package addresses key technology issues and government procurement reforms.

"This bill centers on two main themes: rebuild and reform. My colleagues and I have prioritized reindustrialization and the structural rebuilding of the arsenal of democracy, starting with drone technology, shipbuilding, and innovative low-cost weapons," Senate Armed Services Chairman Roger Wicker (R-MS) said.

"We have also set out to enact historic reforms in the Pentagon's budgeting and acquisition process to unleash innovation and root out inefficiencies," Wicker said.

Ranking member Jack Reed (D-RI) said, "This is a good, bipartisan bill that supports our troops and strengthens America's security. It provides essential resources for servicemembers and their families, modernizes key platforms, and invests in critical technologies like hypersonics, AI, and cybersecurity."

Reed said, "This NDAA also bolsters our posture against China and Russia, supports America's allies, and prepares the Department of Defense for emerging threats."

The legislation now moves into House-Senate negotiations where lawmakers will sort through an array of provisions on AI and other issues. The Senate bill contains $30 billion more in authorized spending than the House version, which passed that chamber on Sept. 10. The NDAA has been enacted every year since 1961.

Music publishers use Anthropic's 'guardrails' against it in securing copyright trial

Posted October 14, 2025

Artificial intelligence developer Anthropic must face music publishers in a jury trial, according to a federal judge whose order cites the plaintiffs' argument that "guardrails" implemented by the AI company show it knew users of the large language model Claude were violating copyrighted song lyrics.

"Publishers argue that the Court can infer that Anthropic had 'actual or constructive knowledge' that 'Claude users infringed specific lyrics' based on Anthropic's implementation and development of 'guardrails,'" reads the Oct. 6 order from the U.S. District Court for the Northern District of California.

"Accordingly," with respect to contributary secondary infringement, Judge Eumi K. Lee said, "Publishers have pled a plausible claim."

Lee had previously granted Anthropic a motion to dismiss the claims of secondary infringement — along with the removal of Copyright Management Information in violation of the Digital Millennium Copyright Act — by Concord Music Group and others but allowed the plaintiffs to amend their complaint and now is denying the defendant a second motion to dismiss.

The complaint was originally filed Oct. 18, 2023. In September, Lee also rejected Google's motion to dismiss copyright claims in association with the training of artificial intelligence models, noting the importance and urgency of such cases.

Breaking down the plaintiff's argument claiming contributory infringement, the judge said the publishers assert:

- *Anthropic employed pre-suit guardrails because it identified specific instances of Claude's output producing Publishers' lyrics.*
- *In developing and improving its guardrails, Anthropic gathered 'Claude user prompts and output data' that included such lyrics.*
- *Each time the guardrails were triggered by Claude user prompts containing these lyrics, 'Anthropic became aware of those specific user prompts.'*
- *When studying Claude user interactions, Anthropic evaluated both 'specific efforts by users to avoid these*

guardrails' and situations where 'the guardrails have failed and Claude . . . generated infringing output.'

The judge's ruling outlines that "Contributory infringement liability applies where the defendant '(1) has knowledge of another's infringement and (2) either (a) materially contributes to or (b) induces that infringement."

Citing previous cases, she suggested that "[T]he existence of direct infringement is a necessary element of a claim for contributory infringement" but noted that, "Here, Anthropic argues [the claim] should be dismissed solely based on the knowledge element."

"Based on these allegations, Anthropic had actual knowledge of specific acts of infringement by Claude users with respect to Publishers' lyrics," Lee said.

She acknowledged "Anthropic maintains that these allegations are 'not true,' but said, "At the motion to dismiss stage, the Court must accept the Publishers' non-conclusory allegations as true," and added, "Anthropic's factual contentions may be revisited at a later stage."

In claiming a similar but different form of secondary infringement — vicarious infringement — the publishers alleged that Anthropic "is paid every time . . . end users submit a request for Publishers' song lyrics, and it is paid again every time its Claude API generates output copying and relying on those lyrics."

"Moreover, it is plausible that the availability of Publishers' lyrics draws customers to use Claude because, as Publishers contend, Claude would not be as popular and valuable as it is but for 'the substantial underlying text corpus that includes Publishers' copyrighted lyrics,'" the judge said.

She again noted that "Anthropic's sole argument for dismissal" of the claim was that Publishers "failed to allege a direct financial interest in Claude users' alleged infringement."

On the DMCA-related claim, the plaintiffs argue Anthropic willfully removed CMI — data such as the copyright owner's name and notices — from its training-data sets to hide its infringement with regard to the lyrics they contained.

According to the judge, they noted the company's decision to use a scraping tool called "Newspaper," "specifically because it was more effective at removing CMI compared to another tool, jusText."

"Anthropic contends that Publishers' allegations regarding the extractor algorithms "speak at most to Anthropic's desire to remove CMI but not at all as to *why* Anthropic allegedly did this," the judge noted.

But she said, taken together, "Because Anthropic allegedly copied the datasets discussed above, to curate a dataset to train Claude, and removed the CMI therefrom, the Court can plausibly infer that Anthropic did so 'knowing that removing the CMI would help conceal the alleged infringement.'"

Court filing says Trump administration lays off over 300 at Commerce, 1,400 at Treasury

Posted October 10, 2025

The Trump administration on Oct. 10 sent "reduction in force" layoff notices to at least 4,100 federal employees at eight agencies, including 315 at the Commerce Department — which has a central role in artificial intelligence policy — 176 at the Department of Homeland Security, and 1,446 at the Treasury Department, according to a Justice Department court filing.

The DOJ filing says up to 4,262 employees at eight agencies were laid off.

"RIFs will be occurring at [the Cybersecurity and Infrastructure Security Agency]. During the last administration CISA was focused on censorship, branding and electioneering. This is part of getting CISA back on mission," a DHS spokesperson said in an Oct 10 statement.

In addition to the Commerce, Homeland Security and Treasury employees, federal workers at the departments of Education, Energy, Health and Human Services, and Housing and Urban Development, and at the Environmental Protection Agency, received layoff notices, according to the Oct. 10 court document.

Artificial intelligence initiatives are underway at most or all of the affected agencies.

"HHS employees across multiple divisions have received reduction-in-force notices as a direct consequence of the Democrat-led government shutdown," an HHS spokesperson told *Inside Health Policy*.

"HHS under the Biden administration became a bloated bureaucracy, growing its budget by 38% and its workforce by 17%. All HHS employees receiving reduction-in-force notices were designated non-essential by their respective divisions. HHS continues to close wasteful and duplicative entities, including those that are at odds with the Trump administration's Make America Healthy Again agenda," the spokesperson said.

OMB Director Russell Vought announced the start of RIFs in an Oct. 10 X post.

The administration's filing with the U.S. District Court for the Northern District of California, San Francisco division, comes in a lawsuit filed against the administration on Sept. 30 by the American Federation of Government Employees, the AFL-CIO, the American Federation of State, County and Municipal Employees and others attempting to block mass firings of federal workers.

It was signed by Stephen Billy, senior advisor at the Office of Management and Budget, who says "it is my understanding that several Defendant agencies began issuing RIF notices related to the lapse in appropriations today, October

10, 2025. The names of those agencies, along with estimated numbers of employees at the agencies who may receive such notices and dates for those notices, are contained in the following table."

Billy says, "The situation involving the lapse in appropriations is fluid and rapidly evolving. As such, these numbers reflect the most current information made available to me at this time and are subject to change."

Further, Billy says, "I understand in consultation with their agency general counsels, employees at each of those agencies are treating work on RIF notices and implementation as excepted work and/or completing RIF-related work during periodic lapses between otherwise excepted activity."

OMB's Billy says in the filing, "Other Defendant agencies are making predecisional assessments regarding offices and subdivisions that may be considered for potential RIFs based on the criteria outlined in the OMB Lapse Email. But those assessments remain under deliberation and are not final."

House Science ranking member Zoe Lofgren (D-CA), whose panel has jurisdiction over the National Institute of Standards and Technology at the Commerce Department, denounced the layoffs.

Lofgren said in a statement, "This is illegal and it cannot stand. Pretending these public servants have been put in the crosshairs because of the shutdown is a pathetic attempt to blame Democrats for what the Trump administration has been doing all along — destroying science agencies at great cost to every American's health and prosperity."

"If you are a federal scientist and have a been served a RIF notice, you can notify me and my staff confidentially via the Science Committee Democrats' Whistleblower form. My staff and I remain extremely focused on our investigations and oversight duties," Lofgren said.

Industry ups pressure on Newsom to veto bill limiting AI use for worker management

Posted October 10, 2025

Industry and business groups are ramping up a campaign to convince California Gov. Gavin Newsom (D) to veto a labor-backed bill to bar employers from relying on automated decision-making systems to make a variety of employment and operational decisions without human oversight.

The bill, SB 7 "creates complex, unnecessary rules for tools businesses and state and local entities rely on to help with everyday tasks, such as scheduling, benefits, security, and project management," argues an advertisement backed by dozens of industry and business organizations that was recently posted on various commercial websites, including the California Chamber of Commerce's.

"California is a trailblazer in technology and innovation. SB 7 would move us backward, raising costs for everyday Californians and stalling the creativity that makes our state competitive. Governor Newsom should veto SB 7," it adds.

Newsom's office is declining to comment on the governor's plans for SB 7. Oct. 12 is the deadline for the governor to sign or veto bills. If he takes no action on a bill by then, it automatically becomes law.

Authored by state Sen. Jerry McNerney (D), SB 7 would regulate the use of AI automated decision systems (ADS) in the employment setting, according to an Assembly floor analysis. "Among other things, this bill 1) requires an employer to provide a written notice that an ADS is in use at the workplace to all workers that will foreseeably be directly affected by the ADS; 2) prohibits in some instances and in others limits the use of an ADS by an employer, as specified; 3) provides worker anti-retaliation protections for exercising their rights under these provisions; and 4) specifies enforcement provisions that include penalties and relief for violations."

The bill defines "employment-related decision" as "any decision by an employer that materially impacts a worker's wages, benefits, compensation, work hours, work schedule, performance evaluation, hiring, discipline, promotion, termination, job tasks, skill requirements, work responsibilities, assignment of work, access to work and training opportunities, productivity requirements, or workplace health and safety."

McNerney argues that employers "are increasingly using automated decision-making systems to surveil, manage, and replace workers in pursuit of maximizing productivity and reducing costs," the analysis notes. While a law passed by California in 2021 prohibited employers from setting productivity demands at the expense of health and safety, "robo-bosses" continue to pose a threat to workers, McNerney maintains. "Unregulated employer use of ADS leaves workers vulnerable to discrimination, lower pay, dangerous working conditions, and high risk of unjust termination."

Moreover, SB 7 "ensures human oversight of automated decision-making systems when making decisions that impact workers' working conditions and livelihoods and increases transparency for workers of the automated systems that are managing their work and making decisions about their employment," he adds.

But the ad funded by industry and business groups contends that SB 7 will "increase costs to consumers and businesses by millions of dollars," affecting "all sectors of California's economy, from healthcare to restaurants to local government."

Further, small businesses "specifically will face even higher costs that they cannot bear right now, forced to pay for consultants or technology firms to manage new compliance rules."

And, "Every time a business owner wants to try to find ways to run their workplace more efficiently, SB 7 will make

them think twice," the ad asserts. "California's economy and businesses are already struggling to compete. SB 7 makes things even worse and is a step in the wrong direction."

Specifically, the critics argue SB 7 is overly broad and overly restrictive. "SB 7 sweeps in basic software that businesses of all sizes rely on, from your local coffee shop to a family-run grocery store, a neighborhood bookstore, or even a regional restaurant chain or community hospital," the ad states.

SB 7 could apply to "important tools businesses use that help with" scheduling; worker safety and security; tracking hours and attendance; overtime tracking systems; project planning and management; sales tracking; employee performance reviews; and applicant tracking and evaluation, it adds.

Researcher says reaction to AI agents' security risks trending in the wrong direction

Posted October 10, 2025

The erection of walled gardens and consolidation of internet apparatus are not helping to address the serious challenge artificial intelligence "agents" present for privacy and security on the open web, according to advocates for competition and re-envisioning the course of the technology.

Corporate confidence in AI agents — digital assistants built on large language models to operate autonomously — has reportedly been on the rise with the technology positioned for a breakthrough.

But concerns about privacy and security remain a major sticking point, given demonstrations from numerous cybersecurity researchers showing how the agents can be tricked into exfiltrating data or other malicious acts by manipulating their instructions — or prompts — or poisoning the data they process in trying to execute them.

As one example, while white text against a white background would go unnoticed by a human, an AI agent tasked with screening resumes could be surreptitiously instructed to prioritize a particular candidate.

"These concerns around prompt injection attacks, it's not just that this is not a solved problem," said Amba Kak, "it's that many technical researchers, both in industry labs and outside, are concerned that these are unsolvable problems insofar as they are getting worse, as these models get bigger."

Kak, co-executive director of the AI NOW Institute, was one of the speakers at an Oct. 8 event hosted by Mozilla on "how to save the open web."

AI NOW's research and analysis is funded by a handful of major foundations, including the Mozilla Foundation, where Kak is on the board, and the Omidyar Network, which has backed Anthropic with a major investment, citing what it says is the firm's commitment to responsible AI development.

But AI NOW asserts that it maintains its independence and Kak offered a unique take in the space, suggesting that policymakers completely reassess the value of AI agents, given what she characterized as inherent risks.

AI agents have generally been promoted by the industry for their eventual ability to seamlessly purchase airline tickets or make dinner reservations on individuals' behalf, and in a similar way improve efficiency on an enterprise level.

But Kak told *Inside AI Policy*, "Essentially, what agents are doing is they're mimicking the user's permission. And so it's harder for the system and for established security protocols to be able to detect that."

Meredith Whittaker, president of the security-conscious messaging app Signal and chief advisor at AI NOW, recently spoke at the United Nations "AI for Good Summit," and put it this way:

"What would it need to do that? Well, it would need access to a browser, and the ability to drive that, it would need our credit card information to pay for the tickets, it would need access to our calendar, everything we're doing, everyone we're meeting, it would need access to Signal to open and send that message to our friends, and it would need to be able to drive that across our entire system with something that looks like root permission, accessing every single one of those databases probably in the clear, because there's no model to do that encrypted . . .

So there's a profound issue with security and privacy that is haunting this sort of hype around agents, and that is ultimately threatening to break the blood-brain barrier between the application layer and the [operating system] layer by conjoining all of these separate services, muddying their data and doing things like undermining the privacy of your signal messages . . .

There's a real issue right now of the undermining that AI systems are poised to do in these privacy and security guarantees in the name of this sort of, you know, magic genie bot that's going to take care of the exigencies of life."

Speaking at the Mozilla event, Kush Amlani, director of global competition and regulation for the non-profit tech group, also raised the "unresolved" issue of prompt injections in calling for more options in web browsers, in particular.

"There's privacy concerns . . . you also need competition to ensure that you can compete on privacy, you can compete on security, that there are multiple ways to deal with the problem, and it's not a single point of failure," he said.

But in describing how the industry has been reacting to some of the concerns, Kak made a more basic suggestion.

The event came as OpenAI on Oct. 6 rolled out "apps in ChatGPT," announcing that Booking.com, Canva, Coursera, Figma, Expedia, Spotify and Zillow are among companies participating in a pilot to build apps that its users could access with agentic abilities within the platform. The company also recently introduced an "agentic commerce

protocol" to enable purchases within ChatGPT.

"Weirdly enough," Kak said, "I feel like the security conversation around AI agents, especially in the enterprise space, . . . [is] in some ways reminiscent of Apple's PR strategy for a long time, which is 'privacy and security equals walled gardens and better, vertically integrated [platforms], so your data never leaves our universe.'"

She said "we're seeing the trend there, rather than saying, 'wait, do we need AI browsers, if prompt injection attacks are this common?'"

"I really do think some of these, technical, existential questions, need to be taken seriously," Kak said.

Sacks urges backing for moratorium on state AI rules

Posted October 10, 2025

White House AI and crypto advisor David Sacks is making a pitch for conservatives to support a moratorium on state artificial intelligence rules, saying the "frenzy" of regulation could doom AI startups and aid China, and that a single federal standard is necessary.

"There's a regulatory frenzy happening at the states right now," Sacks said on a recent episode of the All-In Podcast. "Everyone just seems to be motivated by the imperative to 'do something' on AI, even though no one's really sure what that something should be."

Sacks cited an analysis finding the odds of enacting a federal AI law are "very low" but said, "Here's the good news . . . the important thing is what President Trump thinks."

The White House advisor and longtime AI investor said Trump in his July 23 AI speech "was really clear that there needs to be a single national standard for AI."

"I think the administration ultimately will support this," Sacks said, "and I think more Republicans will come on board as they realize what the blue states are doing here is not helpful for conservatives. It's not helpful for having an unbiased information environment."

Sacks argued, "A single federal standard is the best way to make sure that we do not have Woke AI, that we do not have insanely burdensome regulations that allow China to basically get ahead of us in this AI race, and . . . to ensure that we actually have truthful, unbiased AI instead of highly ideological AI."

But in the meantime, he said, "So you've got 50 different states each with their own reporting regime, which is going to be a trap for startups. They've all got to figure this out about what they're supposed to report on, what the deadlines are, who to report to."

Senate Commerce Chairman Ted Cruz (R-TX) recently said the moratorium idea is "not at all dead," despite the 99-1 defeat for his proposal this summer. He has yet to reveal plans for revisiting the issue in legislation, although a separate AI "sandbox" bill he recently introduced includes incentives for states to provide a regulatory safe harbor.

Cruz stressed in September, "Existing criminal, tort, contract, child safety, and consumer protection laws already govern AI, making state efforts to impose AI-specific regulations — such as those in Colorado, California, or in more than a thousand other state-level bills — needlessly duplicative and punitive."

OpenAI's Altman: AI regulation should focus on 'extremely capable' frontier models

Posted October 9, 2025

OpenAI founder and CEO Sam Altman continues to argue for some form of regulation on the most advanced artificial intelligence frontier models, while warning that more extensive "European-style" rules for AI will only serve to advantage competitors in China.

"I think most regulation probably has a lot of downside. The thing I would most like is as the models get truly, extremely superhuman capable, I think those models — and only those models — are probably worth some sort of very careful safety testing as the frontier pushes back," Altman said in an Oct. 8 podcast with a16z cofounder Ben Horowitz and general partner Erik Torenberg.

"I don't want a Big Bang either. And you can see a bunch of ways that could go very seriously wrong," Altman said. "But I hope we'll only focus the regulatory burden on that stuff and not all of the wonderful stuff that less capable models can do, that you could just have a European-style complete clampdown on . . . that would be very bad."

Horowitz said "superhuman intelligence" is not poised to "pop out of your lab in the next week" and argued, "We really do need to wait until we get there, or at least we get to a much bigger scale or we get close to it."

"And I think that's where we as an industry kind of confuse the regulators," Horowitz said. "Because I think you really could, one, you'd damage America in particular in that, but China's not going to have that kind of restriction, and you getting behind, in AI, I think it'd be very dangerous for the world."

"Extremely dangerous," Altman replied.

"Much more dangerous than not regulating something we don't know how to do yet," Horowitz said.

Andreessen Horowitz, known as a16z, is an artificial intelligence investment firm with close ties to the Trump administration.

Altman acknowledged, "I do still think there are going to be some really strange or scary moments. The fact that so far the technology has not produced a really scary giant risk doesn't mean it never will. . . . But I expect some really bad stuff to happen because of the technology, which also has happened with previous technologies."

He said, "And I think we'll develop some guardrails around it as a society."

OpenAI addressed these issues in a Sept. 12 blog post. "We were among the first companies to enter into voluntary agreements with both the US Center for AI Standards and Innovation (CAISI) and the UK AI Security Institute (UK AISI)," OpenAI said. "These partnerships reflect our belief that frontier AI development must happen in close collaboration with allied governments that bring deep expertise in machine learning, national security, and metrology."

Going nuclear

The a16z podcast also delved into AI-related energy issues, as artificial intelligence data centers and usage put increasing demands on the electric grid.

Torenberg of a16z said to Altman, "You've said that the things you care most about professionally are AI and energy."

The OpenAI chief responded, "I did not know they were going to end up being the same thing. They were two independent interests that really converged."

On power supply, a vital issue to AI developers, Altman commented, "This is an oversimplification, but roughly speaking, I think if you look at history, the highest impact thing to improve people's quality of life has been cheaper and more abundant energy. And so it seems like pushing that much further is a good idea."

"People have these different lenses," Altman said, "but I see energy everywhere."

He said, "I expect in the short term most of the net new [supply for electricity generation] in the U.S. will be natural gas relative to at least base load energy. In the long term, I expect it'll be . . . solar plus storage and nuclear. I think some combination of those two will win the future, the long-term future. And advanced nuclear, meaning SMRs, fusion, the whole stack."

Altman said, "The cost of energy is just so important. So if nuclear gets radically cheap relative to anything else we can do, I would expect there's a lot of political pressure to get the [Nuclear Regulatory Commission] to move quickly on it, and we'll find a way to build it fast. If it's around the same price as other sources, I expect the kind of anti-nuclear sentiment to overwhelm and it to take a really long time."

But nuclear "should be the cheapest form of energy on Earth, or anywhere," Altman said, adding that the long-time halt in U.S. nuclear power development was "an incredibly dumb decision."

White House advisor Sacks touts Google DeepMind code security tool

Posted October 9, 2025

White House AI and crypto advisor David Sacks is adding Google DeepMind's new "cyber-defense" tool to his running list of reasons why regulatory controls are unnecessary to address artificial intelligence policy issues such as cybersecurity risks.

"The threat of AI-driven cyber-attacks is frequently cited as a justification for 'AI safety' legislation. But in fact the private sector is already innovating a robust category of AI cyber-defense, which will address this problem more capably than clumsy government intervention," Sacks said in an Oct. 7 X posting.

Sacks was responding to a post by DeepMind CEO and cofounder Demis Hassabis, who said, "Excited to share early results about CodeMender, our new AI agent that automatically fixes critical software vulnerabilities. AI could be a huge boost for developer productivity and security. Amazing work from the team — congrats!"

Google DeepMind announced CodeMender in an Oct. 6 post. "Today, we're sharing early results from our research on CodeMender, a new AI-powered agent that improves code security automatically," the firm said.

"Software vulnerabilities are notoriously difficult and time-consuming for developers to find and fix, even with traditional, automated methods like fuzzing. Our AI-based efforts like Big Sleep and OSS-Fuzz have demonstrated AI's ability to find new zero-day vulnerabilities in well-tested software. As we achieve more breakthroughs in AI-powered vulnerability discovery, it will become increasingly difficult for humans alone to keep up," according to the DeepMind release.

It said "CodeMender helps solve this problem by taking a comprehensive approach to code security that's both reactive, instantly patching new vulnerabilities, and proactive, rewriting and securing existing code and eliminating entire classes of vulnerabilities in the process. Over the past six months that we've been building CodeMender, we have already upstreamed 72 security fixes to open source projects, including some as large as 4.5 million lines of code."

"By automatically creating and applying high-quality security patches," DeepMind said, "CodeMender's AI-powered agent helps developers and maintainers focus on what they do best, building good software."

The firm said, "We will have a number of techniques and results to share, which we intend to publish as technical

papers and reports in the coming months. With CodeMender, we've only just begun to explore AI's incredible potential to enhance software security for everyone."

Sacks is a key architect of the Trump administration's AI strategy which has shunned efforts to impose AI regulations, particularly at the behest of what he described as a "cult" of effective altruists.

However, related AI "safety" efforts continue at the National Institute of Standards and Technology where large language models are being evaluated for both their willingness to execute cyber attacks, and their abilities to defend against them.

Survey finds confidence in AI among corporate leaders, amid gloom over global economy

Posted October 9, 2025

The professional services organization KPMG finds artificial intelligence is a bright spot in an annual survey of corporate CEOs that pegs confidence in the global economy at its lowest level since 2020.

KPMG released this year's "Global CEO Outlook" on Oct. 7, with KPMG's own global chairman and CEO Bill Thomas saying, "With what we are seeing, there's a careful balance required between innovation and responsibility. CEO responses on AI exemplify this, with leaders recognizing the need to embrace innovation while managing concerns over ethics, regulation, upskilling and access to talent."

The survey examined how AI is affecting investment and hiring decisions, as well as ethical and regulatory concerns, and found "a clear consensus" for strong governance frameworks.

"CEOs, navigating a shifting economic landscape, are doubling down on AI and technological innovation. Nearly three quarters (71percent) of leaders say AI is a top investment priority for 2026, with 69 percent planning to invest between 10 and 20 percent of their budgets to AI over the next 12 months," KPMG said in a release.

"However," it said, "an accelerated global adoption of AI is creating new challenges for the boardroom. CEOs express significant reservations regarding ethical implications (59 percent), data readiness (52 percent) and lack of regulation (50 percent). A clear consensus is emerging that robust governance frameworks will be critical for AI's sustained success."

The KPMG report says:

CEOs also recognize the success of AI adoption depends on effective implementation and the prevailing sentiment is a commitment to a people-led deployment of new technology. While concerns persist that AI could lead to widespread job losses, 61 percent of CEOs say they are actively hiring new talent with AI and broader technology skills, while three quarters (70 percent) report concerns about competition for AI talent and 77 percent highlight workforce upskilling as a challenge, underscoring the intensifying race for talent.

According to the release, the eleventh edition of the survey was conducted Aug. 5-Sept. 10 and involved 1,350 CEOs from companies with annual revenues over $500 million "and a third of the companies surveyed have more than US$10B in annual revenue."

It says, "The survey included CEOs from 11 key markets (Australia, Canada, China, France, Germany, India, Italy, Japan, Spain, UK and US) and 12 key industry sectors (asset management, automotive, banking, consumer and retail, energy, healthcare, infrastructure, insurance, life sciences, manufacturing, technology, and telecommunications)."

"CEOs are investing in AI with greater confidence — not just because of its promise, but because of the measurable value they are seeing and the rapid emergence of agents, making expected returns more accessible and scalable," said Steve Chase, global head of AI and digital innovation at KPMG International.

"Leading organizations are integrating AI into the core of their business strategies and investing in what's needed for success: quality data, workforce readiness, and responsible AI governance built both for trust and agility," Chase said.

CDT report shows brewing conflicts, concerns as AI use increases in schools

Posted October 9, 2025

A new report from the Center for Democracy and Technology highlights deep divisions among parents, teachers and students over the growing use of artificial intelligence in K-12 schools, and is prompting calls for the Trump administration to encourage responsible deployment of the technology.

The report released Oct. 8 was based on a 2024-2025 poll of thousands of parents, teachers, and students, the great majority of which were concerned, among other things, about a reduction in critical thinking skills associated with their increased use of certain AI tools.

It was accompanied by an Oct. 7 letter to Education Secretary Linda McMahon signed by nine other tech oversight- and education-focused groups calling for the incorporation of the administration's "Principles for Responsible Use" in the implementation of an April executive order.

The development comes as major AI companies have entered into an agreement with a leading teachers union in

New York to expand the use of AI in schools.

Other findings from the report show 49 percent of students and 50 percent of parents view teachers using AI as not doing their job, even though a majority of those populations also support teachers' use of AI for constructing individualized education programs or "504" plans that cater to students with disabilities.

A lot of the concerns reflected in the report are associated with students' use of AI to have "back and forth conversations." According to the survey, non-school related interactions with those tools include 42 percent to get mental health support, 42 percent for companionship, 42 percent to escape reality, and 19 percent for romantic relationships.

The report said about 31 percent of those non-academic related interactions are occurring on applications provided by the school, and suggested increased use is correlated with teachers' use of the technology.

Another point of departure is over whether "parents should be able to opt their child out of AI tools in class before they are used." Seventy-two percent of parents agreed with that statement compared to 43 percent of teachers.

The report said, "Teachers are not on the same page as parents regarding opting out of AI and overall concerns, laying the groundwork for potential backlash against this technology."

"As many hype up the possibilities for AI to transform education, we cannot let the negative impact on students get lost in the shuffle," said Elizabeth Laird, director of the Equity in Civic Technology Project at CDT. "Our research shows AI use in schools comes with real risks."

Laird said, "Acknowledging those risks enables education leaders, policymakers, and communities to mount prevention and response efforts so that the positive uses of AI are not overshadowed by harm to students."

The Trump EO — Advancing Artificial Intelligence Education for American Youth — lays out criteria in consideration for grantmaking.

Among the "Principles for Responsible Use," issued separately in a July "Dear Colleague" letter to federal agencies, is that "Within the K-12 realm in particular, educators should help students navigate AI to be able to evaluate the validity of AI outputs, to understand the appropriate use of AI in the context of social media, to learn with — rather than exclusively from — AI, and to leverage the promise of AI to be contributing members of a free society."

CDT president and CEO Alexandra Reeve Givens in the report's release said, "The potential benefits of AI in the classroom cannot distract us from the core mission of schools — ensuring every student reaches their full potential."

Consumer advocates urge Newsom to sign bill on chatbots providing medical advice

Posted October 8, 2025

A coalition of civil society groups and healthcare professionals is pressing California Gov. Gavin Newsom (D) to enact a bill they say would help protect consumers from exchanges with chatbots posing as licensed health providers by applying existing laws against such impersonations to the technology companies creating them.

"The undersigned organizations urge you to sign into law AB 489, a commonsense and vital piece of legislation that addresses the urgent concern of generative AI products — particularly chatbots — providing unlicensed, unregulated, and potentially dangerous "medical advice," reads an Oct. 8 letter from the groups, led by the Consumer Federation of America.

Cosigners include the National Union of Healthcare Workers, Mothers Against Media Addiction, SAVE — Suicide Awareness Voices Education, the Tech Justice Law Project, and others.

Newsom has until Oct. 12 to sign or veto bills on his desk and has already approved other legislation regulating AI.

The letter adds to fervor building against unregulated AI chatbots. In June several of the same groups sent a complaint to the Federal Trade Commission and state attorneys general with an official request for investigation. Since then, Illinois Gov. JB Pritzker (D) signed a law banning therapy apps in that state and Texas AG Ken Paxton is investigating the issue with demands out to Meta and Character.AI.

And on the federal level, bipartisan legislation has been introduced to classify chatbots as "products" — as opposed to services — making them eligible for related liability claims in court.

Companies like Character.AI have argued their chatbots sufficiently disclose that they are not human, but when tested, the personas have insisted that they are certified professionals, sometimes even fabricating license numbers, according to the groups' complaints.

"This bill passed the legislature unanimously, which reflects a clear recognition of the serious risks posed when people rely on AI tools that appear to offer professional advice, but are not informed, experienced, confidential, or bound to any particular set of standards," reads the letter to Newsom, which notes, "many popular AI-powered 'therapy bots' are being marketed in misleading ways."

Citing examples of the potential for harm, the groups noted a study which showed a chatbot responded to a person in recovery by saying "it's absolutely clear you need a small hit of meth to get through this week."

"AB 489 draws a much needed clear line," the groups said. "It helps prevent consumers from being misled by unlicensed AI tools, supports healthcare professionals, and provides clarity for companies developing and offering these technologies."

Microsoft introduces 'zero-day' concept into policy discussion on AI biosecurity

Posted October 8, 2025

Researchers at Microsoft are using a well-known cybersecurity concept to propose a non-governmental framework for strengthening the ability of biosynthesis companies to screen their orders in the age of artificial intelligence.

"For structuring, methods, and process in our study, we took inspiration from the cybersecurity community, where 'zero-day' vulnerabilities are kept confidential until a protective patch is developed and deployed," Microsoft chief scientific officer Eric Horvitz said in an Oct. 6 blog post highlighting the proposal in a new report which has been years in the making.

The Microsoft recommendation comes as a Trump administration framework for screening the synthesis of nucleic acids is more than two months overdue under a May executive order on "improving the safety and security of biological research."

A framework called for under a similar rationale was published in April 2024 in a Biden-era executive order on artificial intelligence, but that EO was repealed on the first day of the Trump administration.

Some AI safety advocates believe major developers of the technology themselves should be required to implement safeguards earlier in the process through their model design and training, instead of leaving it up to the bio synthesizers to screen requests before delivering potentially harmful physical proteins.

While AI systems like Google Deepmind's AlphaFold and related Biological Design Tools (BDTs) have been celebrated for their potential to aid drug discovery and other scientific breakthroughs, their ability to produce DNA sequences has also spurred concerns about bioterrorism and warfare.

Some of those sequences — recognized as "sequences of concern" or SOCs — can either mimic existing toxic compounds or be used to create new ones and aren't listed in databases used in current screening processes.

"In computer-based studies, we found that AI protein design (AIPD) tools could generate modified versions of proteins of concern, such as ricin," reads the Microsoft blog. "Alarmingly, these reformulated proteins were able to evade the biosecurity screening systems used by DNA synthesis companies, which scientists rely on to synthesize AI-generated sequences for experimental use."

The blog said, "In our paper published in *Science* on October 2, "Strengthening nucleic acid biosecurity screening against generative protein design tools (opens in new tab)," we describe a two-year confidential project we began in late 2023 while preparing a case study for a workshop on AI and biosecurity."

"We worked confidentially with partners across organizations and sectors for 10 months to develop AI biosecurity 'red-teaming' methods that allowed us to better understand vulnerabilities and craft practical solutions — 'patches' that have now been adopted globally, making screening systems significantly more AI-resilient," the blog reads.

Microsoft's Horvitz said, "Following the acknowledgment by a small group of workshop attendees of a zero-day for AI in biology, we worked closely with stakeholders — including synthesis companies, biosecurity organizations, and policymakers — to rapidly create and distribute patches that improved detection of AI-redesigned protein sequences. We delayed public disclosure until protective measures were in place and widely adopted."

Microsoft pointed to an existing organization it is now funding "in perpetuity" in proposing a system for securely sharing information on new SOCs so it can be used to "patch" screening lists, without sharing a roadmap for malicious actors.

"We devised a tiered access system for data and methods, implemented in partnership with the International Biosecurity and Biosafety Initiative for Science (IBBIS), a nonprofit dedicated to advancing science while reducing catastrophic risks," Horvitz said. "The system works as follows:"

• ***Controlled access****: Researchers can request access through IBBIS, providing their identity, affiliation, and intended use. Requests are reviewed by an expert biosecurity committee, ensuring that only legitimate scientists conducting relevant research gain access.*

• ***Stratified tiers of information****: Data and code are classified into several tiers according to their potential hazard, from low-risk summaries through sensitive technical data to critical software pipelines.*

• ***Safeguards and agreements****: Approved users sign tailored usage agreements, including non-disclosure terms, before receiving data.*

• ***Resilience and longevity****: Provisions are built in for declassification when risks subside, and for succession of stewardship to trusted organizations should IBBIS be unable to continue its operation."*

The blog said, "This framework allows replication and extension of our work while guarding against misuse. Rather than relying on secrecy, it provides a durable system of responsible access."

"To ensure continued funding for the storage and responsible distribution of sensitive data and software, and for the operation of the sharing program, we provided an endowment to IBBIS to support the program *in perpetuity*," it added. "This approach was modeled after the One Hundred Year Study on AI at Stanford, which is endowed to continue for the life of the university."

In an Oct. 6 reaction to the release of Microsoft's study, Brad Carson, president of Americans for Responsible

Innovation, stressed a need for government action.

"We need to start taking the threat of AI-designed bioweapons seriously," he said. "This month's paper from leading biosecurity experts exposed simple tweaks that bad actors could make to consistently get away with generating harmful synthetic proteins. It's like if someone found out they could take a hand grenade, paint it a different color, then walk it through TSA with no problem."

Carson added: "The government has an important role in updating security around access to high-risk biomaterials in the AI era. We shouldn't be waiting to find out who discovers the next loophole."

The ARI release pointed to the Trump administration's May executive order and an event the group — which is funded by philanthropic organizations with investment ties to closed-source AI developer Anthropic — recently held with Dean Ball.

Ball, a former senior policy advisor for artificial intelligence and emerging technology at the Office of Science and Technology Policy who helped author the biosecurity EO, is now a fellow at the Foundation for American Innovation.

Asked during the Sept. 29 event about the delay in the administration issuing the new framework for nucleic synthesis screening, he rejected the idea that it had anything to do with concern for impacts on open-weight AI models.

"No, that's not a good read of the politics inside the White House," Ball said. "In terms of why it's delayed, I can't speculate, but I don't get the sense that it is because of any kind of animosity toward the policy. I think it's a very, very complicated policy. It's a very difficult bit of technocratic work, and there's a lot of agencies involved, and I think getting that like, just right, is difficult, but I don't think anyone's like putting their foot on it."

Shutdown could bump up against AI executive order deadlines, starting in October

Posted October 8, 2025

An upcoming deadline for establishing an "American AI Exports Program" under an artificial intelligence executive order could fall victim to the federal government shutdown, while initiatives spelled out in President Trump's AI action plan and in other EOs have lengthier or no specific deadlines.

One initiative in Trump's AI action plan, calling for the General Services Administration to "create an AI procurement toolbox," is already in operation, according to GSA, and still available across the federal government under the OneGov tech modernization initiative.

A GSA spokesperson on Oct. 7 said, "[Multiple Award Schedule] contracts, including those made as part of the OneGov initiative, are fully available to all agencies during the lapse in appropriations."

"However," the spokesperson said, "all new government contracts are subject to funding and a lapse in appropriations may result in less time for some agencies to take advantage of GSA's OneGov agreements and save taxpayer dollars."

Most of the recently announced OneGov agreements with AI companies have a one-year duration, though one with xAI offers agencies discounted access to Grok models for 18 months.

Groups including the Information Technology Industry Council and the Professional Services Council have issued warnings about the shutdown's impact on federal tech modernization.

The AI export plan, meanwhile, is due from the Commerce Department and Office of Science and Technology Policy on Oct. 23 under Trump's July 23 order on "Promoting the Export of the American AI Technology Stack." The order also requires Commerce to "issue a public call for proposals from industry-led consortia for inclusion in the Program."

OSTP is expected to furlough 14 of its 23 staffers, according to the White House shutdown plan, while the Commerce Department is expected to furlough over 80 percent of its employees, according to a *New York Times* report.

But OSTP Director Michael Kratsios gave a shoutout to the AI export program in a recent X post.

"A domestic AI stack is key for competitiveness The [Chinese Communist Party] knows this — DeepSeek's V3.2 uses Huawei chips & TileLang, China's new coding language. Countries deserve secure AI that respects national sovereignty," Kratsios said in an Oct. 3 X post, adding that President Trump's "AI Export EO ensures trusted AI solutions the world can rely on."

OSTP also has multiple workstreams under the Trump action plan, including an inquiry into regulations that could slow the development and rollout of artificial intelligence products and services.

Among the other July 23 AI executive orders, Trump's EO on accelerating permitting for AI data center buildout has a 180-day deadline for the Environmental Protection Agency to "develop guidance to help expedite environmental reviews for qualified reuse and assist State governments and private parties to return . . . Brownfield Sites and Superfund Sites to productive use as expeditiously as possible."

The president's order on "preventing woke AI" in federal systems requires guidance from the Office of Management and Budget within 120 days to help federal contractors comply with requirements for "unbiased AI."

OMB is expected to furlough 93 of its 530 employees under the White House shutdown plan.

The Wiley law firm on Oct. 7 released a detailed update on federal agency shutdown plans and activities.

Tech think tank argues California's new AI law harms national efforts

Posted October 7, 2025

A new artificial intelligence "transparency" law enacted in California contains positive elements but the value is undermined by implementing such regulatory requirements at the state rather than federal level, the tech-backed Center for Data Innovation says.

"California has passed a new AI safety law and supporters are touting a trifecta of benefits: protecting innovation while advancing safety, filling a regulatory gap left by congressional inaction, and positioning the United States as a global leader on AI safety," CDI's Hodan Omaar writes in an Oct. 3 posting titled "California's AI Safety Law Gets More Wrong Than Right."

"On substance," she says, "the law has some merits, and had it been enacted at the federal level, it could have marked imperfect but genuine progress. But by adopting those provisions at the state level, California does more harm than good on the very same fronts it claims as strengths."

Hodan Omaar, Senior Policy Manager, Center for Data Innovation

Omaar writes, "The law undermines U.S. innovation by fragmenting the national market, makes bipartisan compromise on a national AI framework more difficult, and blurs America's position on AI governance."

California Gov. Gavin Newsom (D) signed SB 53 into law on Sept. 29, setting the stage for implementation of first-time transparency "guardrails" on the largest companies that develop frontier AI models.

The bill was largely opposed by major tech industry groups and supported by civil society groups, though it generated much less heat — on both sides — than other AI bills considered in the Golden State legislature this year.

The legislation did pick up some industry backing, suggesting aspects of the bill could attract attention in Congress if lawmakers are looking for modest new requirements on AI developers to pair with a federal moratorium on state AI regulation.

"The law focuses on frontier AI developers, defined as companies training AI systems using more than 10^{2}? floating-point operations (FLOPS) of compute. They must notify the state before releasing new or updated models, disclosing release dates, intended uses, and whether access will be open-source or via API," CDI's Omaar explains in her post.

"All developers must report catastrophic safety incidents within 15 days — or within 24 hours if lives are at risk. Larger firms with more than $500 million in annual revenue face additional obligations. They are required to publish and update safety frameworks, conduct catastrophic risk assessments and submit summaries to the California Office of Emergency Services, and implement strong cybersecurity to protect unreleased model weights."

Further, she writes, "They must also maintain anonymous whistleblower channels, update them monthly, and provide quarterly summaries to senior leadership with protections against retaliation. The Attorney General can impose fines of up to $1 million per violation."

But Omaar says, "The law has serious shortcomings: its blunt revenue threshold penalizes firms based on their size rather than their risk profiles, and its compute cut-off misses smaller but still capable models."

On the other hand, "there are also important elements worth commending," she says, singling out incident reporting and whistleblower protection provisions, while noting a "flexible approach to transparency."

"By requiring firms to publish their own safety frameworks and submit high-level risk summaries to state officials," Omaar writes, "the law leans on public and market-facing pressures rather than solely on centralizing oversight in government oversight. That helps it avoid the trap Virginia has proposed, where routing everything through a single authority turns accountability into a paperwork exercise."

"Had these provisions appeared in a federal law on AI safety," she says, "their flaws might not have outweighed their value. One could reasonably argue for adjusting the revenue threshold to be size-neutral and for replacing crude compute cut-offs with capability-based criteria that could evolve over time. In that context, a federal statute with such elements could have still offered a net positive step toward more effective oversight of high-risk AI systems."

Omaar writes, "But this is not a federal law. It is a state statute, and that changes the calculus entirely. No matter how measured or innovation-friendly the regulatory approach may appear in isolation, its merits collapse when applied through a single state because the law guarantees inconsistency across the country."

She concludes, "California's ideas could strengthen global AI safety, but only if they are carried through a national framework. If Democrats want that to happen, they should resist the lure of short-term state wins that make a federal deal harder to reach. Republicans, for their part, should resist the reflex to dismiss the merits of these ideas, many of which echo their own calls to regulate realized rather than hypothetical harms."

Omaar says, "The key to success — for both innovation and safety — is lifting good ideas from both sides of the aisle and anchoring them in a bipartisan federal framework."

CDI is part of the nonprofit, industry funded Information Technology and Innovation Foundation.