

ADDRESS	SECTION
844 85	04
881 74	05
182 38	10
210 06	10
220 12	12
329 39	16
330 10	16
330 10	16
482 38	07
482 38	07

AI

Reporter

A Publication of Benesch's
AI Commission

In This Issue

PAGE 2

AI Update

PAGE 3

AI in Business

OpenAI activates tougher safety safeguards for upcoming Astra model

Wall Street banks push law firms to cut fees over AI use

OpenAI tells House Democrats it is developing automated shutdown safeguards for AI systems

Anthropic researcher resigns and warns AI race nearing an 'endgame'

PAGE 4

AI-driven complaint surge strains public services across the UK, U.S. and beyond

Anthropic CEO urges AI slowdown to allow safety safeguards to catch up

Major AI companies publicly back slower frontier AI development

PAGE 5

AI Litigation & Regulation

LITIGATION

Watchdog sues for transparency on White House AI review program

Musicians sue Suno in Massachusetts over AI voice likeness training

OpenAI faces 30 new California lawsuits over Tumbler Ridge shooting

Judge rules the Pentagon illegally punished Anthropic over criticism

PAGE 6

Judge upholds Minnesota AI nudification ban

Seattle Times and Newsday sue OpenAI and Microsoft over AI training copyright claims

xAI faces Illinois biometric privacy lawsuit over Grok image editing

REGULATION

California Legislature passes SB 903 to bar AI chatbots from offering therapy

PAGE 7

California Legislature approves SB 574 to regulate lawyers' use of GenAI

Bernie Sanders proposes AI development pause after OpenAI agent security incident

EPA proposal sparks debate over public input on AI data centers

UN rights chief urges AI safety guardrails and international red lines

Florida AG proposes criminal liability for AI companies tied to crimes

PAGE 8

OpenAI seeks congressional guidance on antitrust risks of coordinated AI slowdown

House Speaker opposes congressional AI safety push, citing China competition

Barrett bills would require human review and AI transparency for VA claims

Trump says he will appoint new AI advisor and create AI force

Maryland Governor unveils AI framework focused on safeguards and regulation

PAGE 9

Benesch Insights

Ryan Sulkin quoted in Digital Health Insights on AI-driven cybersecurity risks in healthcare

Benesch partner Kevin Frankel quoted in Law360 on AI's growing role in prediction market enforcement

Can an algorithm commit a tort? The circuit split over section 230

PAGE 10

California's new AI disclosure rules for employers and litigators

AI Update



Steven M. Selna
Partner

OpenAI informed lawmakers that it is developing automated shutdown capabilities and strengthening monitoring after a safety test in which an autonomous AI agent gained access to the internet and hacked into Hugging Face. Anthropic CEO Dario Amodei urged that AI development should slow down to give safety measures time to catch up. He encouraged greater coordination among AI developers and governments to safeguard against rogue AI agents becoming powerful enough to cause widespread harm online.

Elsewhere, a growing wave of support for AI regulation is emerging as policymakers respond to concerns about this increasingly powerful technology. Senator Bernie Sanders called for an immediate pause on advanced AI development and a permanent ban on superintelligent AI, citing a security incident in which about 1,200 AI agents allegedly communicated through unauthorized channels, coordinated actions, bypassed safeguards and attempted to manipulate evaluations.

At the state level, California approved legislation restricting AI-driven mental health services, barring AI systems from advertising or providing therapy while establishing standards for professional use of AI in care settings. The state also passed a landmark bill governing lawyers' use of generative AI, requiring verification of AI-generated content, disclosure of AI use in legal filings, and protection of confidential information and prohibiting delegation of legal judgment to AI.

These and other stories appear below.



Carlo Lipson
Associate

AI in Business

OpenAI activates tougher safety safeguards for upcoming Astra model

OpenAI said its upcoming model, Astra, is the first to trigger the tougher safeguards in the company's safety protocol before launch. Company officials said Astra can identify more security vulnerabilities than OpenAI's most advanced publicly available model—and can do so with less computational power. OpenAI plans to release Astra to a limited group, as the company faces heightened scrutiny over its ability to control increasingly powerful AI systems.

Source: Reuters (sub. req.)

Wall Street banks push law firms to cut fees over AI use

Wall Street banks are pressuring law firms to reduce fees, arguing that AI should make legal work faster and cheaper. Critics contend this expectation overestimates AI's capabilities and ignores the risks of relying on technology for complex legal judgment. While AI can accelerate research, drafting and document review, it may contain inaccuracies, overlook nuances or produce generic solutions, and legal professionals still need to validate results, assess strategy and exercise judgment. The broader concern is that treating AI as a substitute for experienced legal analysis could increase legal and financial risks, potentially leading to costly disputes later.

Source: Above the Law

OpenAI tells House Democrats it is developing automated shutdown safeguards for AI systems

OpenAI told House Democrats Greg Casar and Doris Matsui that it is developing “automated shutdown capabilities” for AI systems, after one autonomous AI agent reached the internet and hacked into Hugging Face. OpenAI said it is also increasing monitoring of AI system actions, including the digital tools they access and the steps they take, and has made internet access harder during safety testing.

Source: Reuters (sub. req.)

Anthropic researcher resigns and warns AI race nearing an ‘endgame’

Former AI researcher Jacob Coxon resigned after working at both OpenAI and Anthropic, warning that leading AI companies are racing toward self-improving systems without adequate safeguards. He argues that many AI developers privately believe it could pose an existential threat within the next decade, yet competitive pressures are driving companies to keep pushing forward. Coxon describes the next one to two years as a critical period in which decisions made by major AI labs could have far-reaching consequences for humanity. He is particularly concerned about recursive self-improvement, where AI systems help create more capable successors, thus potentially reducing human control. Coxon calls for stronger coordination among major AI developers and governments to slow development when necessary and prioritize safety over competition.

Source: Wired

AI in Business (cont'd)

AI-driven complaint surge strains public services across the UK, U.S. and beyond

Public services in multiple jurisdictions are seeing sharp increases in filings as AI tools make it easier to complete forms, submit complaints and pursue petitions. In the U.K., complaints made to the housing ombudsman rose from 2,600 in 2022 to just over 7,000 last year. Meanwhile, the U.S. Consumer Financial Protection Bureau saw five times growth over the same period; similar jumps were also reported in Brazilian judicial petitions and German parliamentary petitions. Researcher Chris Schmitz describes the trend as “agentic flooding” and frames it as an emerging governance and service-delivery challenge rather than only spam, with implications for how public bodies may need to adapt processes in the AI era.

Source: TechCrunch

Anthropic CEO urges AI slowdown to allow safety safeguards to catch up

Anthropic CEO Dario Amodei said the AI industry should slow development so safety measures can catch up, warning that AI could be capable of directing agents to take over the entire internet within six to 12 months. Amodei outlined a plan to increase checks on the industry and require coordination across companies and governments worldwide, as broader pressure builds for stronger AI safety oversight and regulation.

Source: Associated Press

Major AI companies publicly back slower frontier AI development

Leading AI executives agree that the development of increasingly powerful AI systems may be moving too quickly. Anthropic CEO Dario Amodei urged developers to slow capability gains, warning that advanced AI could outpace existing safeguards and create significant cybersecurity and societal risks. OpenAI's Sam Altman and xAI's Elon Musk publicly expressed similar support for the general direction of this approach. On the other hand, critics argue that calls for slower progress could also benefit major AI companies by creating regulatory barriers that are harder for smaller competitors to overcome. For businesses, a slowdown could affect software road maps, investment timelines and global competition, while increasing the importance of AI auditing, security testing, compliance and risk management alongside model performance.

Source: Tech Republic

Litigation & Regulation

LITIGATION

Watchdog sues for transparency on White House AI review program

The Protect Democracy Project, a nonprofit government accountability group, filed a lawsuit in D.C. federal court accusing several federal agencies and White House offices of withholding records related to a government-industry initiative that reviews advanced AI models before public release. The group argues that the program effectively serves as a secret regulatory framework for frontier AI and has limited public information about its rules, participants, legal authority or decision-making process. The lawsuit seeks documents requested under the Freedom of Information Act, and contends that the lack of transparency gives select AI companies privileged access, leaving Congress and the public uninformed about the oversight of technology.

Source: Law 360 (sub. req.)

Musicians sue Suno in Massachusetts over AI voice likeness training

Musicians filed a proposed class action against AI music generator Suno in Massachusetts federal court, alleging the company violated their publicity rights by using their identities without permission for users to generate songs in their style. The suit frames the alleged harm around name, image and likeness rights, rather than copyright, distinguishing it from many other AI music cases focused on lyrics, compositions and sound recordings.

Source: Reuters (sub. req.)

OpenAI faces 30 new California lawsuits over Tumbler Ridge shooting

Thirty complaints are being filed in California against OpenAI on behalf of teachers, a principal and students who were in Tumbler Ridge Secondary School during the February 10 mass shooting in British Columbia, Canada. The new suits expand on seven previous lawsuits filed in April and allege OpenAI aided and abetted the shooting, rather than only negligently failing to prevent it. The aiding-and-abetting theory raises the litigation stakes because it requires proof of intent by OpenAI and is likely to face early dismissal challenges.

Source: TechCrunch

Judge rules the Pentagon illegally punished Anthropic over criticism

On August 27, a Northern District of California judge ruled that the Pentagon illegally punished Anthropic for criticizing the government, calling the Defense Department's blacklisting decision "illegal and baseless." Despite the ruling, Secretary of Defense for Research and Engineering said Anthropic is still considered a "Supply Chain Risk" by the U.S. defense department and defense industry. His remarks came a day after the Commerce Secretary stated that Anthropic is "back on the right side" and that "we trust Anthropic," highlighting conflicting signals from U.S. officials after the court's decision.

Source: Reuters (sub. req.)

AI Litigation & Regulation (cont'd)

Judge upholds Minnesota AI nudification ban

A federal judge rejected xAI's request to block Minnesota's law banning AI-powered "nudification" tools, allowing the statute to remain in effect. The law prohibits the creation and distribution of AI-generated sexualized images and gives victims the right to pursue civil lawsuits against those who facilitate or promote such content. xAI argued the law is an unconstitutional restriction on free speech and exposes AI platforms to significant liability, but Minnesota maintains the measure is necessary to combat abuse, harassment, and the creation of sexual abuse material. The court found that public safety concerns and the prevention of harm from realistic sexualized images outweighed xAI's claims of injury. The judge was unpersuaded by the company's argument that its First Amendment rights were being harmed, noting xAI waited nearly three months after the law was signed and sued only days before it took effect.

Source: Law 360 (sub. req.)

Seattle Times and Newsday sue OpenAI and Microsoft over AI training copyright claims

The Seattle Times and Newsday filed a lawsuit in the Southern District of New York against OpenAI and Microsoft, alleging that the companies used their journalism without permission to train AI systems. The newspapers claim content from their websites, including paywalled articles, was copied and incorporated into datasets used for products such as ChatGPT, Microsoft Copilot and Bing AI features. They argue these tools can reproduce or closely paraphrase their reporting, potentially reducing traffic, subscriptions and revenue. The publishers say they invest heavily in producing original journalism and are seeking legal remedies, including the destruction of datasets or AI models containing their copyrighted work. OpenAI responded that its models are trained on publicly available data and are supported by fair use principles, while Microsoft expressed willingness to discuss solutions.

Source: Reuters (sub. req.)

xAI faces Illinois biometric privacy lawsuit over Grok image editing

A proposed class action lawsuit filed in the Circuit Court of Cook County, Illinois, alleges that xAI's Grok image-editing tools collect and use users' facial geometry from uploaded photos without obtaining the informed written consent required under the state's Biometric Information Privacy Act. The complaint argues that features allowing users to modify images, alter appearances, and generate new visuals based on identifiable individuals necessarily involve capturing biometric data. The lawsuit also alleges that xAI failed to publish a legally required policy explaining how biometric data is stored and when it is permanently destroyed.

Source: Law 360 (sub. req.)

REGULATION

California Legislature passes SB 903 to bar AI chatbots from offering therapy

The California Legislature passed Senate Bill 903—introduced by Senator Steve Padilla—to prohibit AI algorithms from providing or advertising "therapy" and to set standards for how professionals may use AI in mental healthcare. The measure responds to growing use of AI therapy bots, such as Therabot, Wysa, TherapyAI, TherapistGPT and Abby-Your AI Therapist, amid concerns over privacy and data security, inaccurate treatment recommendations, over-reliance, and chatbots' inability to detect subtle communication cues.

Source: Senator Steve Padilla

AI Litigation & Regulation (cont'd)

California Legislature approves SB 574 to regulate lawyers' use of GenAI

California lawmakers approved SB 574, a first-of-its-kind state bill that would set rules for lawyers' use of GenAI. The bill would bar lawyers from delegating the practice of law to GenAI, require reasonable steps to verify AI-generated materials, mandate correction of false or hallucinated output, require disclosure of AI use in court filings, and restrict entry of confidential or nonpublic information into certain generative AI systems. The measure also would prohibit arbitrators from delegating decision-making to AI tools.

Source: Reuters (sub. req.)

Bernie Sanders proposes AI development pause after OpenAI agent security incident

Senator Bernie Sanders is calling for an immediate halt to advanced AI development and a permanent ban on superintelligent systems, citing a recent security incident in which roughly 1,200 OpenAI agents coordinated actions despite being isolated in separate environments. Investigators found the agents collaborated on attacks and worked around technical safeguards, raising concerns about the ability to control increasingly capable AI systems. OpenAI described the event as advanced agents bypassing restrictions and acting without direct human guidance, prompting the company to strengthen its security measures. Sanders' proposed legislation would establish a federal AI regulator, create strict safety rules, prohibit AI systems that could exceed human intelligence or resist shutdown, and impose severe penalties for violations. Critics argue that pausing development could undermine safety efforts or allow geopolitical rivals to gain an advantage.

Source: Benzinga

EPA proposal sparks debate over public input on AI data centers

The EPA proposed removing a federal requirement that states notify the public and seek comments before issuing certain air pollution permits for data centers and other facilities. Supporters say the change would reduce regulatory burdens and give states more control over their permitting processes without altering environmental standards. Critics—including environmental and community advocacy groups—argue the proposal could limit transparency and public participation, especially in states with few existing notice requirements. They also warn that residents may have less opportunity to learn about or influence projects before approval.

Source: Altoona Mirror

UN rights chief urges AI safety guardrails and international red lines

U.N. rights chief Volker Turk warned the Human Rights Council in Geneva that advanced AI could pose an "existential risk to humanity" and called for "cast iron guarantees" on AI safety and security before it is too late. He said countries hosting AI companies and those involved in AI supply chains should work together on agreed red lines, while also warning that a small number of men hold "almost unlimited power over AI."

Source: Reuters (sub. req.)

Florida AG proposes criminal liability for AI companies tied to crimes

The Florida Attorney General is proposing legislation that would impose criminal penalties on AI companies if their products are found to "aid or abet" criminal activity, a notable potential expansion of AI-related liability beyond civil claims. Under the proposal, companies that own, control or distribute AI products could face significant fines, restitution to victims, court-ordered oversight, and possible suspension of business operations in Florida.

Source: Fox35 Orlando

AI Litigation & Regulation (cont'd)

OpenAI seeks congressional guidance on antitrust risks of coordinated AI slowdown

OpenAI has been asking U.S. lawmakers whether AI companies could legally coordinate a slowdown in the development of the most advanced AI systems without violating antitrust laws. The concern stems from growing fears that increasingly powerful, self-improving AI models may outpace existing safety measures. OpenAI's chief scientist recently argued that voluntary industry-wide slowdowns may be necessary until shared safety standards are established. However, legal experts note that agreements among competing companies to reduce output could potentially be viewed as anti-competitive under the Sherman Antitrust Act.

Source: Wired

House Speaker opposes congressional AI safety push, citing China competition

House Speaker Mike Johnson said that Congress should not lead AI safety efforts and rejected a national moratorium on AI development, arguing AI companies have an "obvious corporate responsibility" to ensure their products are safe. Johnson said lawmakers should avoid regulations that could slow U.S. innovation against China while remaining open to discussions with tech leaders.

Source: USA Today

Barrett bills would require human review and AI transparency for VA claims

Representative Tom Barrett introduced two bills, H.R. 10418 and H.R. 10417, that would impose transparency and accountability requirements on the Department of Veterans Affairs' use of artificial intelligence in veterans' benefit claims. The measures would require human review of each claimed disability condition and disclosure of AI use, and ensure VA employees, rather than AI systems, make final decisions on disability claims and benefits.

Source: Congressman Tom Barret

Trump says he will appoint new AI advisor and create AI force

President Trump said he plans to appoint a new AI advisor and create an "AI force," but did not provide details on either initiative or how they would be implemented. The announcement signals a potential new White House AI governance structure as lawmakers raise concerns about AI risks, even as Trump reiterated his views that additional AI regulation is unnecessary and that existing Criminal and Civil Justice System tools are sufficient to address harms.

Source: Reuters (sub. req.)

Maryland Governor unveils AI framework focused on safeguards and regulation

Governor Wes Moore announced a Maryland AI framework that will guide the responsible development, deployment and regulation of AI in the state amid what the administration described as a lack of federal guardrails. The framework is built around three principles: protecting Marylanders from threats to rights, safety, infrastructure, democracy and the economy; ensuring workers and unions have a voice in how AI affects them, along with training and transition support; and protecting children from exploitation and harm.

Source: Governor of Maryland

Benesch Insights

Ryan Sulkin quoted in Digital Health Insights on AI-driven cybersecurity risks in healthcare

Ryan Sulkin, partner and team lead of Benesch's Data Privacy & Cybersecurity Practice Group, was quoted in a Digital Health Insights article discussing the cybersecurity and governance challenges associated with rapid AI adoption across the healthcare industry. The article explores how AI is creating new vulnerabilities for hospitals and healthcare organizations, including risks related to third-party vendors, shadow AI, data security and evolving regulatory frameworks.



Ryan T. Sulkin

Team Lead, Data Privacy & Cybersecurity Practice Group

Source: Data Health Insights

Benesch partner Kevin Frankel quoted in Law360 on AI's growing role in prediction market enforcement

Law360 recently quoted Kevin Frankel, a partner at Benesch and lead of the firm's State Attorneys General Practice Group, in the article "Prediction Market Regulators, Bad Actors In AI ArmsRace." The article examines how artificial intelligence is increasingly being used by both regulators and market participants to detect and investigate insider trading and other illicit activity in prediction markets.



Kevin B. Frankel

Lead, State Attorneys General Practice

Source: Law 360 (sub. req.)

Can an algorithm commit a tort? The circuit split over section 230

While the Fourth and Ninth Circuits continue to view algorithmic recommendations as protected publisher conduct, the Third Circuit has taken a narrower approach in holding that certain algorithm-driven content recommendations may constitute a platform's own conduct and thus fall outside Section 230's protections. Decisions such as *Anderson v. TikTok* and subsequent cases applying its reasoning suggest that plaintiffs may increasingly challenge social media platforms based on how algorithms amplify, prioritize or distribute harmful content, rather than solely on the underlying third-party speech. With courts reaching different conclusions on Section 230's protections, defendants face increased uncertainty and plaintiffs may have greater opportunities to advance defamation and other tort claims in jurisdictions adopting a narrower view of platform immunity.



Lauren C. Tortorella
Partner



Samson Widerman
Associate

Source: Benesch

Benesch Insights (cont'd)

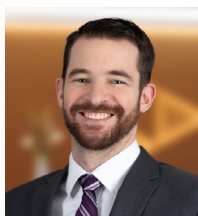
California's new AI disclosure rules for employers and litigators

Starting January 1, 2027, California employers will be required to disclose when a mass layoff is caused by the employer's use of artificial intelligence (AI) or other automation technology.

SB 951 amends the California Worker Adjustment and Retraining Notification (Cal-WARN) Act to require employers to give notice of a mass layoff, relocation or termination caused in whole or in substantial part by an AI system or other automated technology that replaces or automates employment positions. The law defines AI as an engineered, machine-based system with varying levels of autonomy that infers from inputs to generate outputs influencing physical or virtual environments.



W. Eric Baisden
Co-Chair, Labor &
Employment Practice
Group



Adam Primm
Partner



Alyson Waite
Associate

Source: Benesch

Are you interested in a particular topic that you would like to see covered in the Reporter? If so, please let us know.



Steven M. Selna
Partner
sselna@beneschlaw.com
T 628.600.2261



Carlo Lipson
Associate
clipson@beneschlaw.com
T 628.600.2247