New American

Written by **Veronika Kyrylenko** on April 14, 2026

OpenAI Backs Bill Shielding AI Firms From Liability in Mass-harm Scenarios

AP Images

Artificial intelligence company OpenAI is backing an Illinois bill that seeks to limit when AI companies can be held responsible for catastrophic harm. That encompassess mass deaths, including those resulting from the use of weapons of mass destruction, as well as autonomous criminal behavior.

At the same time, AI lobbyists are pouring money into Congress and statehouses, aiming to shape the very system meant to protect the public through sound laws. In doing so, the AI industry appears to be following a well-worn path blazed by other major lobbies — defense contractors, Big Pharma, and the rest of corporate America's seasoned practitioners of influence.

The bill is already viewed by some observers as a possible model for future legislation in other states. That raises the stakes well beyond the Prairie State. In principle, the rule should be simple: If a consumer product causes harm, liability follows. That standard is now under pressure as AI moves into higher-risk domains.

The Bill

The proposal, known as SB 3444 in the Illinois General Assembly, introduces the "Artificial Intelligence Safety Act." On paper, it presents itself as a safety framework for "frontier" AI models. In practice, it offers major AI developers a path to avoid liability even in cases of extraordinary harm.

The bill defines a "frontier model" as one trained using more than "10^26 computational operations" or at a "compute cost that exceeds $100,000,000." In other words, this is not aimed at small developers or niche tools. It is written for the largest and most powerful players in the field.

At the heart of the bill is its liability shield. Section 10 states,

> A developer shall not be held liable for critical harms if the developer did not intentionally or recklessly cause the critical harms[.]

At the same time, the developer must publish certain documents on its website (more on that later).

The definition of "critical harm" is not trivial. It means "the death or serious injury of 100 or more people" or "at least $1,000,000,000 of damages to rights in property." The bill further specifies that such harm may occur through "the creation or use of a chemical, biological, radiological, or nuclear weapon." It also includes conduct by a model acting "with no meaningful human intervention" that, if committed by a human, "would constitute a criminal offense" involving "intent, recklessness, or negligence," or even the solicitation or aiding of such a crime.

That is what makes the bill so striking. It does not deal with minor errors, consumer annoyance, or routine software disputes. It addresses mass-casualty events, billion-dollar destruction, and criminal conduct carried out autonomously, then proceeds to outline the circumstances under which the developer may still not be liable.

Liability Shield

The mechanics matter because they show what the bill treats as accountability.

To qualify for protection, a developer must publish a "safety and security protocol." The bill says this document should describe the company's "technical and organizational protocols to manage, assess, and mitigate risk of critical harm." It lays out testing procedures, defines thresholds for "critical risk," and explains the mitigations in place. It also addresses whether third parties assess risk, how the company handles cybersecurity, and how it monitors for harm after deployment.

The developer must also publish a "transparency report." That report identifies the frontier model and summarizes "the results of assessments" and "the steps taken to address any identified risks."

On its face, the framework appears rigorous. In practice, it leaves room for discretion. The bill allows "appropriate redactions" to protect "model security and integrity, trade secrets, and proprietary information." Key details may remain hidden, even as the company claims compliance.

The bill offers additional paths to protection. A developer is deemed compliant if it agrees to follow requirements adopted under Article 56 of the European Union's AI Act. That provision relies on "codes of practice," which guide compliance but do not impose fixed safety standards.

A company may also qualify by entering into an agreement with a federal agency. That agreement should allow for government access to the model for "research and evaluation." Once in place, the developer files a certification with the Illinois attorney general.

In effect, the system centers on process. Publish the documents. Align with accepted frameworks. Establish the right relationships in Washington. Liability then depends less on outcomes and more on whether those steps were followed.

OpenAI's Position

OpenAI has publicly supported the bill.

According to a *WIRED* report from last week, policy experts see SB 3444 as more aggressive than earlier proposals OpenAI backed. It moves beyond defensive lobbying into shaping liability standards.

In a statement to the outlet, OpenAI spokesperson Jamie Radice framed the bill as pragmatic:

> We support approaches like this because they focus on what matters most: Reducing the risk of serious harm from the most advanced AI systems while still allowing this technology to get into the hands of the people and businesses — small and big — of Illinois.

The company also emphasized regulatory consistency. It warned against a fragmented system of state laws. That argument has become a common refrain across Silicon Valley.

In testimony, OpenAI's Caitlin Niedermeyer reinforced that view. She argued that policymakers should avoid "a patchwork of inconsistent state requirements that could create friction without meaningfully improving safety."

The position aligns with the federal push to override state laws. Niedermeyer also tied regulation to global competition, reflecting a worldview that prioritizes an "AI arms race" over normalization and cooperation.

Real-world Harm

The trajectory of AI deployment reinforces critical concerns. AI is embedded in surveillance platforms, targeting systems, and operational decision tools. Both federal and local law-enforcement agencies use drones and AI-assisted technologies in real-world environments.

The "national security" layer makes the picture darker. The proposed "Golden Dome," described as a missile-defense shield, is in practice a network of sensors, satellites, and interceptors coordinated by AI systems. That software will be built and refined by major AI labs and Deep State-aligned firms such as Palantir Technologies. Critics already warn that systems ostensibly designed for external threats would be turned inward. Adding to that concern is the Pentagon's legal dispute with Anthropic. The DOD blacklisted the company after it objected to having its systems used for autonomous killings and mass domestic surveillance, particularly in Golden Dome.

Beyond physical harm, the AI systems are already eroding civil liberties and weaken constitutional protections. They monitor speech and behavior, map relationships, and shape decisions through automated flagging, risk scoring, and predictive enforcement.

At the individual level, harm is already being tested in court. Several lawsuits have been filed against OpenAI, including claims involving psychological harm and alleged links to suicide. Notably, these cases fall outside the bill's "critical harm" threshold, which is focused on mass casualty and large-scale destruction. If passed, the bill would create an uneven, and in many respects illogical, legal landscape.

At the same time, if Donald Trump's administration and its allies in Congress succeed in federalizing AI regulation, state laws, whether lax or restrictive, would be effectively swept aside. The practical result would be a regime with fewer constraints on the industry, fewer remedies for the public, and more power concentrated in the hands of the very institutions building these systems.