July 2026

Holding AI Companies Accountable for Catastrophic Harm

Model legislation from the Future of Life Institute to hold frontier AI developers strictly liable when they cause catastrophic harm.

Contact: **Alex Tsalidis**, alexandra@futureoflife.org

View 1-pager and complete bill online

The Problem

Today's most powerful AI models are being developed and released at a scale where a single failure can kill, injure, or disrupt on a magnitude no victim, deployer, or insurer can absorb.

- **Frontier AI can cause catastrophic harm.** When frontier AI goes wrong at scale, such as by helping develop a bioweapon, it can cause mass casualties or cripple critical infrastructure. This is exactly the kind of risk the law has always treated as abnormally dangerous.
- **We can't trust the most powerful AI companies to take reasonable care.** Like blasting, storing explosives, or handling toxic chemicals, frontier AI development carries a high degree of risk, even with safeguards in place.
- **Victims can't prove what they can't see.** Requiring proof of negligence or defect is an impossible burden when the design, training, and safety choices are locked inside a few frontier AI developers' labs. Catastrophic harm should not go unremedied because of an evidentiary maze.
- **Frontier AI companies push liability downstream to SMEs and third parties.** Developers use B2B contracts to disclaim responsibility and push it onto smaller providers and startups who have little ability to test or fix the underlying model.

The Solution

The same strict liability rule that governs every other ultra-hazardous activity should apply to those who develop and release the most powerful AI systems IF they cause catastrophic harm.

What counts as catastrophic harm?

- Death of **100+** people, serious injury to **500+**, significant disruption of **critical infrastructure**, or **$1B+** in property damage. Measured per event or over any 12-month period.

Who would be covered?

- Only frontier AI models: general-purpose systems capable across multiple cognitive domains, trained with more than **10^{26} FLOPs**, and powerful enough to cause catastrophic harm. Narrow, single-task tools are excluded.
- Liability falls only on frontier AI developers, the small number of companies that train or control these models (e.g., OpenAI, Meta). **NOT** ordinary software developers, university researchers, startups, deployers, or end users.

Key Provisions

- **AI is not a legal person.** A model cannot bear legal responsibility. It is never a defense that it "acted on its own," acted unexpectedly, or appeared to have intent. Responsibility rests with the companies behind it, not the machine.
- **Frontier AI development is an ultra-hazardous activity**, but only when a frontier AI model causes catastrophic harm. A frontier AI developer is strictly liable for all catastrophic harm caused by its model. Victims get a remedy without having to prove a defect, which is a significant hurdle for potential plaintiffs attempting to claim under product liability.
- **Accountability across the chain of development.** All frontier AI developers in the chain are jointly and severally liable, with contribution apportioned by each party's proximate causal contribution.
- **Open-weight releases carry responsibility.** A developer that releases model weights is strictly liable for reasonably foreseeable catastrophic harm from any downstream deployment or derivative model for five years after release.
- **Protecting downstream parties from predatory contracts.** Deployers, integrators, distributors, infrastructure providers, end users, tooling providers, and universities are exempt. Contracts that force others to indemnify a frontier developer are void as against public policy.
- **Real enforcement.** A private right of action (including class actions), Attorney General enforcement with civil penalties, a bar on forced arbitration and liability waivers, and a six-year statute of limitations from discovery of the injury.
- **Builds on existing law.** The Act adds a new basis for liability without displacing negligence, product liability, or other existing claims and agency authority.

Why It Matters

Frontier AI is one of the only industries capable of causing catastrophe on a mass scale while facing almost none of the accountability we demand of far less dangerous activities. Strict liability gives frontier model developers a powerful incentive to make their models safe before release, not after a major incident has occurred.