CENTER FOR
AI POLICY

What RAIA Does and Why

The Problem

Within the next few years, the largest artificial intelligence models will likely be smarter – and therefore more powerful – than their human controllers. This means we need to be *very* sure that these "frontier AIs" will remain firmly under the control of responsible humans at all times. Otherwise, frontier AIs might create bioweapons, hack into critical infrastructure, or take over the planet's resources.

Unfortunately, frontier AI is unsafe by default. Because frontier AI is trained through a process of trial-and-error, we do not know exactly what it will do. Unlike a bridge or an airplane or a power line, we cannot be sure that a very large AI model will behave as intended, nor can we assume that all of the model's users will have benign intentions.

Under current law, a private company can deploy any AI model regardless of the danger that creates for public safety. It is unreasonable to bet the world's future on the chance that every frontier AI developer will always be perfectly responsible. Instead, we need a commonsense system of third-party testing.

The Primary Solution – Permitting for the Largest AIs

- ***Who is Covered*** - Our model legislation, the Responsible AI Act of 2025 ("RAIA"), would apply only to the largest general-purpose systems. Developers who spent less than $1 billion on training their AI would be exempt, as would developers whose AI systems return only narrow types of data like a price, location, or translation.
- ***What is Needed*** – Developers who are covered by RAIA would need to have their largest AI systems evaluated by an independent auditor, who would certify that the system has enough safeguards to prevent it from causing catastrophic destruction. The government would then review the audit reports to confirm that they are adequate and complete. If they are not, the developer could be required to add more safeguards or run more tests before receiving a permit to deploy their AI.
- ***Enforcement*** – Developers who intentionally deploy frontier AI without first receiving a permit would be subject to civil and criminal penalties enforced by a small new federal office called the "Frontier AI Administration."

Other Solutions for Dangerous AI

- ***Hardware Security*** - Large data centers with billions of dollars in equipment could probably be hacked with a $20,000 effort. To fix this and stop terrorists and rival states from stealing or corrupting our AI, we need to require AI data centers to maintain a minimum level of physical security, cybersecurity, and KYC protocols.
- ***Monitoring & Reporting*** – Right now, the government is dangerously ignorant about what is happening in the world of AI. There is simply no organized process for tracking trends and surprises in the way AI is manufactured, developed, improved, and deployed, even though AI will soon be at least as important as other strategic resources like uranium and lithium. To fix this, we need to build a dedicated team of government experts who can investigate the most powerful AI systems and report on what they learn to the rest of the federal government.
- ***Liability Reform*** – It's unlikely that communities would be able to hold an AI developer accountable for reckless behavior under current law, because there are too many loopholes. If AI destroys a dam or a power plant, it could be impossible to prove that any particular developer was uniquely responsible. Worse, the field evolves so quickly that it is difficult to point to agreed-upon "best practices" whose violation would be clearly negligent. Our model legislation plugs these loopholes by specifying some best practices and explicitly providing for joint and several liability.
- ***Emergency Powers*** – If a rogue AI is in the process of escaping from its servers or destroying a city, the government might have a very narrow window of time in which to try to intervene. Rather than encourage the government to act outside the law or risk having the government act too slowly, it's better to take time now to specify what should happen in an emergency and how innocent bystanders should be compensated.

Using CAIP's Model Legislation

This model legislation, the Responsible AI Act of 2025 (RAIA), is free to use for anyone who is interested in promoting the security of America's AI. Please feel free to edit or adapt parts of it to suit your needs.

RAIA was developed by the Center for AI Policy (CAIP), a nonprofit, nonpartisan organization working to protect the American public. In addition to the full text of the bill, CAIP is also providing a section-by-section summary that walks people through the bill's contents in plain English, and a short policy paper that explains more about our reasons for supporting the bill.