



California's AI Safety Bill: A Groundbreaking Move Amidst Industry Controversy and What AI Developers Can Do to Prepare

By: Tiana Garbett and Rachel Schuster

In a historic turn of events, California is poised to become the first state to enact comprehensive AI safety legislation with the introduction of [SB 1047](#). This bill, designed to address the potential risks associated with advanced AI technologies, has ignited intense debate within the tech community and among policymakers.

The bill passed the California State Assembly in a bipartisan 49-15 vote on [August 28, 2024](#), and now awaits an approval or veto from California Governor Gavin Newsom, which is anticipated to be received by September 30th. If approved, SB 1047 makes will significantly impact the future of AI development and regulation.

The Core Provisions of SB 1047

SB 1047, crafted by State Senator Scott Wiener, aims to place stringent requirements on developers of large-scale AI models. The bill targets any AI project with a development cost exceeding \$100 million, imposing a set of rigorous safety and accountability measures. Key provisions include:

- **Safety Testing and Safeguards:** Developers must conduct thorough safety evaluations and implement safeguards to mitigate risks associated with their models.
- **Third-Party Audits and Kill Switches:** The bill mandates independent third-party audits and requires that all AI systems incorporate a "kill switch" to disable the technology if necessary.
- **Whistleblower Protections:** The legislation proposes protections for individuals who report unsafe AI practices, as further discussed below.
- **Legal Accountability:** The bill makes AI model developers liable for their harms and the state attorney general is empowered to take action against developers if their AI models result in severe harm, such as mass casualties or damages exceeding \$500 million.

The Backlash and Support

The bill's ambitious scope has generated a spectrum of reactions from industry leaders and lawmakers. Supporters, including prominent figures in the AI field, argue that SB 1047 is a necessary step toward ensuring AI safety and responsibility. Leading figures in AI research have publicly backed the bill and supporters emphasize the importance of regulation to mitigate risks associated with advanced technologies.

However, the bill has also faced substantial opposition from major tech companies and influential figures, arguing that the bill could stifle innovation and potentially drive AI development out of California. Critics fear that the rigorous requirements could place smaller firms and startups at a disadvantage, thus harming the state's competitive edge in the tech sector.

Concerns have also been raised about unintended consequences of the bill, suggesting that its penalties and restrictions could negatively impact less-resourced sectors of the AI ecosystem, including academics and small tech companies.

Notably, opposition has not been limited to the tech industry. Some lawmakers have criticized the bill as being well-intentioned but flawed. Their concerns highlight a broader debate about the appropriate level of state versus federal involvement in regulating emerging technologies.

Expansion of Whistleblower Protections

Existing laws generally protect whistleblower activity related to the reporting of illegal activities by shielding employees from retaliation, e.g., disclosing hazardous or unsafe working conditions that violate safety regulations or reporting unethical financial practices within a company. SB 1047 expands these protections further to cover disclosures specially-related to AI risks by introducing comprehensive whistleblower protections specifically aimed at employees involved in the development of AI models.

Codified as new Section 22602 of the Business and Professions Code, these protections prohibit employers, as well as their contractors and subcontractors, from preventing employees from or retaliating against employees who report that (a) the developer employer is out of compliance with the bill's requirements, or that (b) an AI model poses an unreasonable risk of causing or materially enabling critical harm, even if the developer employer is *not* out of compliance with any law.

Among other provisions, Section 22602 also requires that developer employers must clearly notify employees of their rights and provide an anonymous internal reporting and investigations process. Section 22602 also provides that employees can report concerns to the California Attorney General or Labor Commissioner without fear of punitive actions, reinforcing the legal safeguards against retaliation under the existing Labor Code and ensuring that employees harmed by violations can seek legal relief. There is even a provision that allows the Attorney General or Labor Commissioner to publicly release complaints if it is in the public interest (so long as any confidential information is protected).

These protections reflect a growing demand from workers for more robust whistleblower laws in an ever-growing AI industry which evolves much more rapidly than existing regulations have the ability to. Although, developer employers should remember that these protections are geographically limited to California and thus not enforceable on a federal level. Section 22602 also does not apply to the extent that it is preempted by federal law.

The Broader Implications

SB 1047's potential passage represents a pioneering effort to establish a framework for AI safety at the state level. If enacted, it could set a precedent for other states and possibly influence federal legislation, particularly as Congress determines how to regulate rapidly advancing technologies.

While the bill's future is not yet certain, its passage underscores the urgency of addressing AI's risks while nurturing its potential. As the debate continues, key decisionmakers will need to navigate the balance between ensuring public safety and promoting technological innovation. The outcome of SB 1047 could very well shape the trajectory of AI regulation in the United States.

Preparing for SB 1047: Essential Tips for AI Developers and Companies

In preparation for the potential enactment of SB 1047, AI developers should proactively prepare for the changes that may come with this legislation to ensure a smooth transition and compliance:

- 1. Conduct a Thorough Compliance Review:** Evaluate current AI projects and development processes to identify any areas that might fall under SB 1047's requirements. This includes assessing the cost and scale of your AI models and ensuring that they meet the safety testing, safeguards, and kill switch mandates.
- 2. Implement Safety Protocols:** Develop and integrate comprehensive safety protocols and evaluation procedures for your AI models. Ensure that these measures are not only compliant with SB 1047 but also reflect best practices in AI safety and ethics.
- 3. Prepare for Third-Party Audits:** Arrange for independent third-party audits of your AI systems. Engaging with certified auditors who specialize in AI safety can help you identify potential risks and demonstrate compliance with the new regulations.
- 4. Establish Whistleblower Protections:** Develop internal policies and training programs to inform employees about their whistleblower rights and the process for reporting unsafe AI practices. Ensure that your organization has a clear and anonymous reporting mechanism in place, as well as policies against retaliation.
- 5. Stay Informed and Adapt:** Continuously monitor legislative updates and industry discussions related to SB 1047. Being proactive about changes in the regulatory landscape will help you stay ahead of compliance requirements and adjust your strategies as needed.

By taking these steps, AI developers and companies can better position themselves to navigate the complexities of SB 1047.

Given the legal complexities and compliance obligations, it is advisable to work with an attorney well-versed in the legal issues with AI and the requirements under this bill. More broadly, it is important for companies developing AI to ensure that there is sufficient AI legal training and that they develop comprehensive policies on AI legal compliance. For more information on the need for legal training and AI policies see [Why Companies Need AI Legal Training and Must Develop AI Policies](#) and [The Need For Generative AI Development Policies and the FTC's Investigative Demand to OpenAI](#).

We will continue to track updates to this legislation and how these developments might affect the future landscape of AI and technology policy.

For More Information, Please Contact:



Tiana Garbett

Associate | New York
212.896.0656
tgarbett@sheppardmullin.com



Rachel Schuster

Associate | San Diego (Del Mar)
858.876.3555
rschuster@sheppardmullin.com

This alert is provided for information purposes only and does not constitute legal advice and is not intended to form an attorney client relationship. Please contact your Sheppard Mullin attorney contact for additional information.