

# Governor Newsom Vetoes SB 1047, Rejecting AI Whistleblower Protections

Article By:

Jesse Broad-Cavanagh

---

On Sunday, September 29, 2024, Governor Newsom (D-CA) [vetoed SB 1047](#), rejecting a framework designed to increase the safety and transparency of AI development. The bill, first introduced by State Senator Scott Wiener (D-San Francisco) and subsequently approved by both the state's Assembly and Senate, sought to address a number of safety concerns flowing from Big Tech's recent [AI spending spree](#). Crucially, the now-vetoed SB 1047 included anti-retaliation protections for employees.

California's attempt to include anti-retaliation protections as part of a larger scheme of regulating AI development reflects an impactful policy strategy. That strategy is rooted in the commonsense awareness that if certain activity is illegal and poses a direct risk of harm to the public, but such activity is easily hidden, the public benefits from empowering whistleblowers who are in a position to know about the illegal conduct.

Despite Governor Newsom's disappointing decision to veto SB 1047, the U.S. Congress has repeatedly embraced the strategy of empowering whistleblowers. The False Claims Act, which empowers individuals blowing the whistle on fraud in government spending, originated during the Civil War at a time when defense contractors fraudulently lowered their own costs at the expense of the public, including by cutting gunpowder with sawdust. More recently, in the wake of the Great Recession, Congress passed the [Dodd-Frank Act](#) to provide heightened protections and incentives to whistleblowers with information related to unlawful schemes by some of the world's largest financial institutions. Under Dodd-Frank, whistleblowers may be eligible for rewards if the information they provide leads to successful government recovery by the SEC or CFTC.

Given the successes of this policy strategy and amid Big Tech's race to develop AI, the federal government should pass protections specific to AI whistleblowers.

## Why should the government protect AI whistleblowers and how does SB 1047's framework accomplish that goal?

Stakeholders, including current and former employees in Big Tech, have been vocal regarding the dangers posed by the development of AI models. On June 4, 2024, a group of current and former employees at Open AI and Google Deepmind, responsible for two of the largest Big Tech AI models,

---

wrote a public [letter](#), warning of the risks posed by AI, including, “further entrenchment of existing inequalities,” “manipulation and misinformation,” and “the loss of control of autonomous AI systems potentially resulting in human extinction.” The signatories called on AI companies to commit to four core principles, all of which are designed to protect employees who seek to catch and stop dangerous decisions by AI developers: (1) ceasing the use of “anti-disparagement” provisions in employment contracts to chill whistleblowing; (2) creating an anonymous process for employees to report safety concerns; (3) supporting an internal culture of “open criticism” and disclosure; and (4) promising not to retaliate against employees who, should internal processes fail, blow the whistle externally.

The whistleblower framework in the now-vetoed SB 1047 took notable steps towards the June 4 signatories’ concerns, as it offered a measure of legal protection to employees who might be punished for raising concerns regarding their company’s failure to adhere to legal requirements for testing, design, and implementation of high-cost, large-scale AI models. Importantly, the bill prohibited retaliation against covered employees who disclosed information to the State Attorney General or the Labor Commissioner regarding the employee’s “reasonable belief” that their employer “is out of compliance with [the bill’s] requirements or that the covered model poses an unreasonable risk of critical harm.” Under the “reasonable belief” standard, an employee who has good reason to believe that their employer is violating AI requirements would be protected for reporting the conduct, regardless of whether the behavior about which the employee complained constituted an actual violation of the law.

The bill contained additional provisions designed to encourage an internal culture of openness and truthfulness, thereby increasing communication both within companies and with regulators. Internally, covered AI developers were required to provide an anonymous channel through which employees could report concerns related to violations of law, failures to disclose risk, or “false or materially misleading statements” made in a manner that violates California’s antitrust law. The bill’s anti-retaliation protections and reporting channel mandates extended to contractors and subcontractors working on behalf of covered developers, thus better ensuring that an appropriate pool of individuals with direct information about violations and safety concerns would be protected.

On September 9, 2024, a group of over 100 current and former employees at “frontier” AI companies – including OpenAI, Google DeepMind, Meta, and xAI – wrote to Governor Newsom in support of SB 1047. The signatories of the letter emphasized the “severe risks” posed by the most powerful AI models, and they pointed to SB 1047 as a step in the right direction for ensuring responsible AI development. The Governor’s veto has left this much needed legislation in limbo for now.

## **What legal protections are currently available for AI whistleblowers?**

The call for AI-specific whistleblower protections arises under a backdrop of piecemeal [federal and state whistleblower law](#). The anti-retaliation protections offered by these laws are not uniform across cybersecurity, data privacy, and technology companies, and, instead, their application depends on which company the employee works for and in which state they are employed. Under this framework, employees with concerns specific to AI development must navigate a patchwork of potentially relevant laws to find protection against retaliation.

## **Which anti-retaliation protections are relevant on the federal level?**

There are no federal protections uniformly available to AI whistleblowers. One option for some whistleblowers is the anti-retaliation provision contained within the Sarbanes-Oxley Act (“SOX”).

---

SOX prohibits publicly traded companies and certain subsidiaries from retaliating against employees for reporting concerns regarding specific categories of misconduct, including several types of fraud and violations of SEC rules and regulations. Importantly, a whistleblower covered under SOX has a federal right to bring their case to court and in front of a jury, even if the employee has signed an employment agreement purporting to restrict the employee's ability to bring claims outside of non-public arbitration.

Many employees with information concerning AI-related wrongdoings or development concerns are not covered by SOX's anti-retaliation protections because (1) SOX is available only to employees of publicly traded companies and their subsidiaries; and (2) many violations of rules governing AI development would not fall squarely within one of the categories of misconduct covered by SOX. Thus, if an employee works for a private company – for example, OpenAI or Inflection AI – they are not protected under SOX. Moreover, even if the employee does work for a publicly traded company, their concerns regarding their employer's development of AI may not qualify as “protected activity” unless their employer engages in fraud, such as by materially misrepresenting to investors how safe or advanced their AI model is.

## **Which anti-retaliation protections are relevant on the state level?**

AI whistleblowers may be covered by the statutes or common law provisions in the state in which they live or work. In some states, including California, New York, New Jersey, and Virginia, broad whistleblower statutes prohibit employers from retaliating against employees for reporting unlawful activity, such as fraud. Relatedly, many states, again including California, provide a common law claim for wrongful discharge in violation of public policy.

The specific requirements for establishing a common law claim for wrongful discharge in violation of public policy vary by state. However, a common thread is that an employee may be protected from retaliation if they object to their employer's violation of an existing statutory or constitutional provision, regardless of whether the provision explicitly provides for whistleblower protections. Thus, the viability of such a claim for an AI whistleblower will often depend on whether there is a state or federal statute which prohibits the employer's conduct in the first instance.

Over the past year, a number of state legislators have acted on a desire to regulate AI, thereby potentially opening new statutory hooks for employees attempting to establish a claim for wrongful discharge in violation of public policy. For instance, in 2023, state legislators across the country introduced [nearly two hundred](#) AI-related bills, many of which focused on deepfake technology and government use of AI. Most of these bills died prior to passage, but several states – Connecticut, Florida, Illinois, Louisiana, Minnesota, Montana, Texas, Virginia, and Washington – successfully passed AI legislation in 2023, though they did not include explicit whistleblower protections. Overall, these developments are encouraging for whistleblowers, because if a state (1) recognizes a claim for wrongful discharge in violation of public policy, and (2) passes legislation rendering certain activities in AI development unlawful, then an employee in that state may have a viable pathway to blowing the whistle in a manner that renders subsequent retaliation unlawful.

Amid continued public pressure to regulate AI, there should continue to be a noticeable increase in the number of states placing statutory obligations on AI developers, thus providing provisions with which AI whistleblowers can ground their wrongful discharge claims. AI whistleblowers and their advocates must remain vigilant for such developments.

## What changes are needed under the existing legal framework?

In the U.S. Senate, lawmakers on both sides of the aisle are actively grappling with how to structure future laws related to AI oversight. On September 17, 2024, the Senate Judiciary's Subcommittee on Privacy, Technology, and the Law heard [testimony](#) from experts with experience in Big Tech's AI industry. At the end of the hearing, committee chair Senator Richard Blumenthal (D-CT) pointedly asked what could be done to improve whistleblower protections. The experts responded with recommendations including whistleblower training and the development of ethics rules specific to the AI industry, much like the professional rules of conduct governing lawyers.

One expert, Helen Toner, Director of Strategy and Foundational Research Grants at Georgetown's Center for Security and Emerging Technology, stated simply, "I think the need for whistleblower protections goes hand in hand with other rules." That is, in the absence of action from states or the federal government regulating AI, it is difficult for a potential whistleblower to know whether voicing their objections qualifies as "protected" conduct. In such a climate, many will choose silence.

Given the current piecemeal nature of legal protections for AI whistleblowers, the federal government should use the groundwork laid by SB 1047 and pass AI-specific anti-retaliation protections applicable across state lines. SB 1047's framework is desirable in that it unequivocally provides anti-retaliation protections to a subset of employees working on AI development. Instead of working through a patchwork of law to determine if they are protected, employees would be able to determine (1) if their employer is covered under the law; (2) if the conduct regarding which they are blowing the whistle is illegal; and (3) how to blow the whistle. Moreover, SB 1047 extends anti-retaliation protections to contractors, further increasing the applicable pool of knowledgeable individuals who would have legal protection if they face retaliation.

Overall, employees working within or adjacent to AI development are stifled by a culture of secrecy. If lawmakers want to catch and stop dangerous practices in this rapidly developing field, stable whistleblower protections are desperately needed. If you are concerned about possible retaliation for speaking out against your employer's questionable practices in the AI arena, the knowledgeable whistleblower attorneys at KBK can help you navigate the complex legal landscape to determine whether you have any protection.

Katz Banks Kumin LLP Copyright ©

---

National Law Review, Volume XIV, Number 281

Source URL: <https://natlawreview.com/article/governor-newsom-vetoes-sb-1047-rejecting-ai-whistleblower-protections>