SENATE JUDICIARY COMMITTEE
Senator Thomas Umberg, Chair
2025-2026 Regular Session

SB 468 (Becker)
Version: February 19, 2025
Hearing Date: April 22, 2025
Fiscal: Yes
Urgency: No
CK

SUBJECT

High-risk artificial intelligence systems: duty to protect personal information

DIGEST

This bill imposes a duty on a business that deploys a high-risk AI system that processes personal information to protect that information and requires such a deployer to maintain a comprehensive information security program that meets specified requirements.

EXECUTIVE SUMMARY

High-risk automated decision systems (ADS) powered by AI are being increasingly deployed in a multitude of contexts, including employment, housing, education, and health care. Major transparency and fairness concerns have been raised about the use of ADS to make consequential decisions, essentially determinations with significant legal or other material effect on people's lives. One particularly concerning aspect is the amount of consumers' personal information being handled by these systems. There is fear that the complexities of this new technology and the sheer volume of personal information being utilized is not adequately protected by existing cybersecurity laws and measures.

This bill responds to these concerns by imposing a duty on "covered deployers," businesses that deploy high-risk AI systems that process personal information, to protect the personal information they hold. Deployers are required to develop, implement, and maintain a comprehensive information security program that contains appropriate administrative, technical, and physical safeguards and that meets a series of specifications aimed at ensuring industry standards are met. To ensure these standards stay current, the California Privacy Protection Agency (PPA) is authorized to implement regulations to implement the law.

This bill is author-sponsored. It is supported by Transparency Coalition.AI and Oakland Privacy. It is opposed by the California Hospital Association.

PROPOSED CHANGES TO THE LAW

Existing law:

1) Requires a business that owns, licenses, or maintains personal information about a California resident to implement and maintain reasonable security procedures and practices appropriate to the nature of the information, to protect the personal information from unauthorized access, destruction, use, modification, or disclosure, and to contractually require nonaffiliated third parties to which it discloses such personal information to similarly protect that information. (Civ. Code § 1798.81.5.)

2) Establishes the California Consumer Privacy Act (CCPA), which grants consumers certain rights with regard to their personal information, including enhanced notice, access, and disclosure; the right to deletion; the right to restrict the sale of information; and protection from discrimination for exercising these rights. It places attendant obligations on businesses to respect those rights. (Civ. Code § 1798.100 et seq.)

3) Establishes the California Privacy Rights Act of 2020 (CPRA), which amends the CCPA and creates the PPA, which is charged with implementing these privacy laws, promulgating regulations, and carrying out enforcement actions. (Civ. Code § 798.100 et seq.; Proposition 24 (2020).)

4) Provides consumers the right to request that a business delete any personal information about the consumer which the business has collected from the consumer. (Civ. Code § 1798.105(a).)

5) Defines "personal information" as information that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household. The CCPA provides a nonexclusive series of categories of information deemed to be personal information, including biometric information, geolocation data, and "sensitive personal information." It does not include publicly available information or lawfully obtained, truthful information that is a matter of public concern. (Civ. Code § 1798.140(v).)

6) Requires the PPA to adopt regulations governing access and opt-out rights with respect to businesses' use of automated decisionmaking technology, including profiling and requiring businesses' response to access requests to include meaningful information about the logic involved in those decisionmaking

processes, as well as a description of the likely outcome of the process with respect to the consumer. (Civ. Code § 1798.185(a)(15).)

7) Authorizes a consumer whose nonencrypted and nonredacted personal information, or whose email address in combination with a password or security question and answer that would permit access to the account, is subject to an unauthorized access and exfiltration, theft, or disclosure as a result of a business' violation of the duty to implement and maintain reasonable security procedures and practices appropriate to the nature of the information to protect the personal information may institute a civil action for statutory or actual damages. (Civ. Code § 1798.150.)

8) Requires the California Department of Technology (CDT) to conduct a comprehensive inventory of all high-risk ADS that have been proposed for use, development, or procurement by, or are being used, developed, or procured by, any state agency. It defines the relevant terms:
    a) "ADS" means a computational process derived from machine learning, statistical modeling, data analytics, or AI that issues simplified output, including a score, classification, or recommendation, that is used to assist or replace human discretionary decisionmaking and materially impacts natural persons. "ADS" does not include a spam email filter, firewall, antivirus software, identity and access management tools, calculator, database, dataset, or other compilation of data.
    b) "High-risk ADS" means an ADS that is used to assist or replace human discretionary decisions that have a legal or similarly significant effect, including decisions that materially impact access to, or approval for, housing or accommodations, education, employment, credit, health care, and criminal justice. (Gov. Code § 11546.45.5.)

9) Establishes the Unfair Competition Law (UCL), which provides a statutory cause of action for any unlawful, unfair, or fraudulent business act or practice and unfair, deceptive, untrue, or misleading advertising, including over the internet. (Bus. & Prof. Code § 17200 et seq.)

10) Defines "unfair competition" to mean and include any unlawful, unfair, or fraudulent business act or practice and any unfair, deceptive, untrue, or misleading advertising, and any act prohibited by the False Advertising Law, Business and Professions Code section 17500 et seq. (Bus. & Prof. Code § 17200.)

11) Provides that any person who engages, has engaged, or proposes to engage in unfair competition may be enjoined in any court of competent jurisdiction. (Bus. & Prof. Code § 17203.)

12) Requires actions for relief pursuant to the UCL be prosecuted exclusively in a court of competent jurisdiction and only by the following:
    a) the Attorney General;
    b) a district attorney;
    c) a county counsel authorized by agreement with the district attorney in actions involving violation of a county ordinance;
    d) a city attorney of a city having a population in excess of 750,000;
    e) a county counsel of any county within which a city has a population in excess of 750,000;
    f) a city attorney in a city and county;
    g) a city prosecutor in a city having a full-time city prosecutor in the name of the people of the State of California upon their own complaint or upon the complaint of a board, officer, person, corporation, or association with the consent of the district attorney; or
    h) a person who has suffered injury in fact and has lost money or property as a result of the unfair competition. (Bus. & Prof. Code § 17204.)

13) Provides that any person who engages, has engaged, or proposes to engage in unfair competition is liable for a civil penalty not to exceed $2,500 for each violation. The court shall impose a civil penalty for each violation. (Bus. & Prof. Code § 17206.)

This bill:

1) Defines the relevant terms:
    a) "Covered deployer" means a business that deploys a high-risk AI system that processes personal information.
    b) "High-risk AI system" has the same meaning as "high-risk ADS," as that term is defined in Section 11546.45.5 of the Government Code.
    c) "Business," "consumer," "personal information," and "processes" have the same meaning as defined in the CCPA.

2) Imposes a duty on a covered deployer conducting business in this state to protect personal information held by the covered deployer, as provided.

3) Requires a covered deployer whose high-risk AI systems process personal information to develop, implement, and maintain a comprehensive information security program that is written in one or more readily accessible parts and contains administrative, technical, and physical safeguards that are appropriate for all of the following:
    a) The covered deployer's size, scope, and type of business.
    b) The amount of resources available to the covered deployer.
    c) The amount of data stored by the covered deployer.

d) The need for security and confidentiality of personal information stored by the covered deployer.

4) Provides that the required information security program must meet specified requirements, including:
   a) Incorporating safeguards that are consistent with the safeguards for the protection of personal information and information of a similar character under state or federal laws and regulations applicable to the covered deployer.
   b) Designating one or more employees of the covered deployer to maintain the program.
   c) Requiring the identification and assessment of reasonably foreseeable internal and external risks to the security, confidentiality, and integrity of any records containing personal information, and the establishment of a process for evaluating and improving, as necessary, the effectiveness of the current safeguards for limiting those risks, as provided.
   d) Including security policies for the covered deployer's employees relating to the storage, access, and transportation of records containing personal information outside of the covered deployer's physical business premises.
   e) Providing disciplinary measures for violations of a policy or procedure established under the program.
   f) Including measures for preventing a terminated employee from accessing records containing personal information.
   g) Providing policies for the supervision of third-party service providers, as provided.
   h) Providing reasonable restrictions on physical access to records containing personal information, including by requiring the records containing the data to be stored in a locked facility, storage area, or container.
   i) Including regular monitoring to ensure that the program is operating in a manner reasonably calculated to prevent unauthorized access to or unauthorized use of personal information and, as necessary, upgrading information safeguards to limit the risk of unauthorized access to or unauthorized use of personal information.
   j) Requiring the regular review of the scope of the program's security measures that must occur at least annually and whenever there is a material change in the covered deployer's business practices that may reasonably affect the security or integrity of records containing personal information.
   k) Requiring the documentation of responsive actions taken in connection with any incident involving a breach of security, including a mandatory postincident review of each event and the actions taken, if any, in response to that event to make changes in business practices relating to protection of personal information.

l) Including, to the extent feasible, all of the following procedures and protocols with respect to computer system security requirements or procedures and protocols providing a higher degree of security, for the protection of personal information:
   i. The use of secure user authentication protocols, as specified.
   ii. The use of secure access control measures, as specified.
   iii. The encryption of transmitted records and files containing personal information that will travel across public networks and data containing personal information that is transmitted wirelessly.
   iv. The use of reasonable monitoring of systems for unauthorized use of or access to personal information.
   v. The encryption of all personal information stored on laptop computers or other portable devices.
   vi. For files containing personal information on a system that is connected to the internet, the use of reasonably current firewall protection and operating system security patches that are reasonably designed to maintain the integrity of the personal information.
   vii. The use of a reasonably current version of system security agent software that shall include malware protection and reasonably current patches and virus definitions and a version of a system security agent software that is supportable with current patches and virus definitions, and is set to receive the most current security updates on a regular basis.

5) Provides that a violation constitutes a deceptive trade act or practice pursuant to the UCL.

6) Authorizes the PPA to adopt regulations to implement and administer these provisions.

7) Finds and declares that it furthers the purposes and intent of the CPRA by ensuring consumers' rights, including the constitutional right to privacy, are protected by enabling and empowering Californians to request that covered deployers secure their high-risk AI systems that process personal information.

COMMENTS

1. Understanding the AI-related risks to personal information

With recent dramatic advances in the capabilities of AI systems, the need for regulatory frameworks for accountability and responsible development and deployment have become ever more urgent. This is especially true with respect to AI-powered ADS that are used to make, or assist in making, decisions that have a legal or other significant

effect. ADS introduce several concerning issues when deployed, including risks to consumers' personal information.

Therefore, comprehensive information security programs are essential for AI models handling sensitive personal information due to their unique vulnerabilities. Unlike traditional systems, AI models face specialized threats including model extraction attacks, training data exposure through carefully crafted prompts, and adversarial examples designed to manipulate model behavior. These systems can inadvertently memorize sensitive information during training, making them susceptible to extraction attacks that conventional security measures are not designed to prevent.

Internal risks compound these challenges, as technical staff often require extensive data access, while the "black box" nature of complex models makes security auditing difficult. System integration points create additional attack surfaces, such as APIs connecting AI systems to data sources.

External threats include supply chain vulnerabilities from pre-trained components, cloud infrastructure attacks, and increasingly sophisticated adversaries specifically targeting AI assets for their valuable data and capabilities.

IBM's recent exploration of AI-related privacy issues highlights these risks:

> **Data exfiltration**
> AI models contain a trove of sensitive data that can prove irresistible to attackers. "This [data] ends up with a big bullseye that somebody's going to try to hit," Jeff Crume, an IBM Security Distinguish Engineer, explained. . . . Bad actors can conduct such data exfiltration (data theft) from AI applications through various strategies. For instance, in prompt injection attacks, hackers disguise malicious inputs as legitimate prompts, manipulating generative AI systems into exposing sensitive data. Such as, a hacker using the right prompt might trick an LLM-powered virtual assistant into forwarding private documents.
>
> **Data leakage**
> Data leakage is the accidental exposure of sensitive data, and some AI models have proven vulnerable to such data breaches. In one headline-making instance, ChatGPT, the large language model (LLM) from OpenAI, showed some users the titles of other users' conversation histories. Risks exist for small, proprietary AI models as well. For example, consider a healthcare company that builds an in-house, AI-powered diagnostic app based on its customers' data. That app might unintentionally leak customers' private information to other customers

> who happen to use a particular prompt. Even such unintentional data sharing can result in serious privacy breaches.[1]

Warning calls are already being made, as indicated by recent guidance from a financial services regulator in New York:

> New York's financial regulator said firms need to address the specific cybersecurity risks arising from the use of artificial intelligence, as more regulators aim to ensure the safe use of this rapidly evolving technology.
>
> The New York State Department of Financial Services on Wednesday issued a new guidance document that advises the entities it regulates to monitor and assess risks from AI-enabled tools, as part of the agency's existing cybersecurity regulation. The department said financial-services firms need to better understand AI-related risks, including from social engineering, cyberattacks and the theft of nonpublic information.[2]

Experts have highlighted the need to focus not only on AI safety, but also AI security:

> Focusing on keeping AI models secure from those seeking to break in may seem more immediate and actionable than tackling the potential for all-powerful AI that could conceivably go off the rails. However, the world's best ethical hackers, or those who test systems in order to find and fix weaknesses before malicious hackers can exploit them, say AI security—like traditional cybersecurity—is far from easy.
>
> AI security risks are no joke: A user could trick an LLM into generating detailed instructions for conducting cyberattacks or harmful activities. An AI model could be manipulated to reveal sensitive or private data in its training set. Meanwhile, self-driving cars could be subtly modified; deepfake videos could spread misinformation; and chatbots could impersonate real people as part of scams.
>
> More than two years since OpenAI's ChatGPT burst onto the scene, hackers from the Def Con security conference, the largest annual gathering for ethical hackers, have warned that it is still far too easy to break into AI systems and tools. In a recent report called the Hackers'

---

[1] Alice Gomstyn & Alexandra Jonker, *Exploring privacy issues in the age of AI* (September 30, 2024) IBM, https://www.ibm.com/think/insights/ai-privacy#:~:text=Understanding%20the%20privacy%20risks%20of,Data%20leakage. All internet citations are current as of March 31, 2025.

[2] Mengqi Sun, *Financial Firms Need to Focus on Cyber Risks Posed by AI, New York Regulator Says* (October 16, 2024) The Wall Street Journal, https://www.wsj.com/articles/financial-firms-need-to-focus-on-cyber-risks-posed-by-ai-new-york-regulator-says-61c1203d.

Almanack published in partnership with the University of Chicago, they said that AI vulnerabilities would continue to pose serious risks without a fundamental overhaul of current security practices.[3]

2. Responding to the risks

Given these critical risks, organizations must implement AI-specific security measures including rigorous access controls, specialized security assessments, privacy-preserving techniques, and continuous monitoring for unusual model behavior to avoid potentially catastrophic breaches that could expose sensitive personal information at unprecedented scale. This bill attempts to provide a regulatory baseline for AI security measures.

Existing law provides some basic protections. A business that owns, licenses, or maintains personal information about a California resident is required to implement and maintain reasonable security procedures and practices appropriate to the nature of the information, to protect the personal information from unauthorized access, destruction, use, modification, or disclosure, and to contractually require nonaffiliated third parties to which it discloses such personal information to similarly protect that information.

The CCPA gives consumers some transparency into who is collecting their personal information and for what purposes it is being used, as well as some basic control over how it can be used. In addition, consumers are authorized to bring a civil action for breaches of their personal information that result from a business' violation of the duty to implement and maintain reasonable security procedures and practices appropriate to the nature of the information to protect the personal information and seek statutory or actual damages.

However, given the scope of the risks in the AI-context and the unique vulnerabilities detailed above, more should arguably be done.

The bill imposes a duty on "covered deployers," businesses that deploy a "high-risk artificial intelligence system" that processes personal information, to protect personal information held by the covered deployer. The bill equates "high-risk AI system" with "high-risk ADS," which is already defined under current law as an ADS that is used to assist or replace human discretionary decisions that have a legal or similarly significant effect, including decisions that materially impact access to, or approval for, housing or accommodations, education, employment, credit, health care, and criminal justice.

---

[3] Sharon Goldman, *AI security risks are in the spotlight – but hackers say models are still alarmingly easy to attack* (Feb. 18, 2025) Fortune, https://fortune.com/2025/02/18/ai-security-risks-are-in-the-spotlight-but-hackers-say-models-are-still-alarmingly-easy-to-attack/.

The bill requires a covered deployer whose high-risk AI systems process personal information to develop, implement, and maintain a comprehensive information security program that is written in one or more readily accessible parts and contains administrative, technical, and physical safeguards that are appropriate given various factors, including the size and scope of the business, the amount of data it stores, and the amount of resources it has.

The information security program must include a number of specified elements. This includes evaluation of reasonably foreseeable risks, both internal and external, to the security, confidentiality, and integrity of its records containing personal information. There is a strong focus on ensuring anyone coming into contact with the AI system and the personal information it processes and contains is vetted, trained, and held accountable for the security of the system and its data. The security programs must also integrate strong access controls and secure storage practices.

To ensure the required components stay current with the rapidly evolving technology, and the increasingly complex risks, the bill authorizes the PPA to adopt regulations to implement and administer the law.

Violations are deemed deceptive trade acts or practices pursuant to the UCL, ensuring some level of enforcement is available. Given the expertise of the PPA and its regulatory role, the author may wish to consider whether the PPA should also have administrative enforcement powers.

According to the author:

> AI is advancing rapidly, becoming more a part of our daily lives and influencing many everyday decisions. Our security laws must keep up with these constantly evolving technologies.
>
> AI systems handle vast amounts of sensitive personal data, creating new vulnerabilities beyond traditional data security concerns. Cybercriminals can manipulate AI models through tactics like corrupting AI training data to create biased or incorrect decisions, extracting personal details by repeatedly querying AI models, and making it difficult to detect breaches or misuse. Without proper safeguards, AI systems that automate life-altering decisions could expose people's most sensitive information to data breaches, fraud, or manipulation.
>
> SB 468 ensures that businesses deploying high-risk AI systems processing personal information establish and maintain a comprehensive security program to protect consumers. This security program includes creating clear accountability by designating security managers and conducting risk assessments; employee training in AI security protocols; physical access

> restrictions for personal data; third-party oversight; and incident response plans to rapidly address security breaches when they occur.
>
> The public deserves to know their personal information is protected and that AI systems are operated responsibly and securely, and SB 468 will provide this crucial consumer protection.

3. Stakeholder positions

Transparency Coalition.AI writes in support:

> The leaking and inappropriate use of stored personal information continues to be a problem with wide-ranging and potentially lasting legal and ethical impact. AI systems, because of the size and scope of their underlying data sets, present a particularly attractive target for bad actors, and therefore must be protected using only the most effective technologies and practices. SB 468 effectively places this responsibility on deployers of these AI systems, while not demanding any actions that might be considered unduly unreasonable or impactful. It is for these reasons that TCAI is pleased to support SB 468.

The California Hospital Association writes in an oppose-unless-amended position and asks to be carved out of the bill:

> Senate Bill (SB) 468 (Becker) would add a new Title 1.81.28 (commencing with Section 1798.91.2) to Part 4 of Division 3 of the Civil Code. This new title uses many of the same definitions as in the CCPA and CPRA, further requiring "covered deployers" to develop, implement, and maintain a comprehensive information security program that contains administrative, technical, and physical safeguards. However, SB 468 does not include the same exemptions as do CCPA and CPRA, subjecting hospitals and other HIPAA-covered entities to duplicative and costly regulation, which would drive up health care costs without providing any benefit to patients.
>
> Given the existing comprehensive regulatory framework, and longstanding health information privacy and security laws, the California Hospital Association (CHA) requests that hospitals and other entities subject to HIPAA and/or CMIA be exempted from SB 468.

SUPPORT

Oakland Privacy
Transparency Coalition.AI

OPPOSITION

California Hospital Association

RELATED LEGISLATION

Pending Legislation:

SB 420 (Padilla, 2025) regulates the use of "high-risk ADS," defined the same as high-risk AI systems in this bill. SB 420 includes requirements on developers and deployers to perform impact assessments on their systems. SB 420 establishes the right of individuals to know when an ADS is being used, details about the systems, and an opportunity to appeal ADS decisions, where technically feasible. SB 420 is currently in the Senate Governmental Organization Committee.

AB 1018 (Bauer-Kahan, 2025) regulates the development and deployment of ADS used to make consequential decisions, as defined. It requires a developer of a covered ADS to take certain actions, including conduct performance evaluations of the ADS, submit to third-party audits, and provide deployers to whom the developer transfers the covered ADS with certain information, including the results of those performance evaluations. AB 1018 is currently in the Assembly Privacy and Consumer Protection Committee.

Prior Legislation:

SB 892 (Padilla, 2024) would have required CDT to develop and adopt regulations to create an ADS procurement standard, as specified, and prohibited a state agency from procuring ADS, entering into a contract for ADS, or any service that utilizes ADS, until CDT has adopted regulations creating an ADS procurement standard, as specified. SB 892 was vetoed by Governor Newsom, who stated in his veto message that aspects of the bill would disrupt ongoing work, "including existing information technology modernization efforts, which would lead to implementation delays and higher expenses for critical projects."

AB 2885 (Bauer-Kahan, 2024) established a uniform definition for "artificial intelligence" in California's code, which is used in this bill.

AB 302 (Ward, Ch. 800, Stats. 2023) requires CDT, on or before September 1, 2024, to conduct a comprehensive inventory of all high-risk ADS that have been proposed for use, development, or procurement by, or are being used, developed, or procured by, any state agency.

****