SB 892 (Padilla)
Version: April 10, 2024
Hearing Date: April 16, 2024
Fiscal: Yes
Urgency: No
CK

SUBJECT

Public contracts: automated decision systems: AI risk management standards

DIGEST

This bill requires the Department of Technology (CDT) to establish an artificial intelligence (AI) risk management standard regarding procurement and use of automated decionmaking systems (ADS) that is informed by leading established standards. The standard must detail specified procedures for assessing and controlling risks, prohibited use cases, and an assessment for impact on vulnerable communities. State agencies are prohibited from entering into contracts for ADS unless the contract contains specified provisions.

EXECUTIVE SUMMARY

The transformative power of AI and specifically generative AI (GenAI) is unquestionable, and it offers numerous benefits for society, including state government. It can enhance efficiency and effectiveness in various sectors freeing up human resources. Specifically, the deployment of ADS, algorithm-driven applications that assist or replace human discretionary decisionmaking, can improve decision-making processes by analyzing vast amounts of data to identify patterns, trends, and potential insights. However, as with most technologies, there are also inherent risks and challenges. One major concern is the potential for bias in ADS applications, especially when high-risk ADS are determining individuals access to housing, credit, employment, or other benefits. Ensuring transparency, accountability, and fairness in ADS is crucial to mitigate these risks and maintain public trust.

This bill directs CDT to establish an AI risk management standard, modeled after mainstream, widely-accepted publications, to guide the procurement, use, and oversight of ADS in government agencies. Agencies are required to include specific clauses in contracts for the procurement of ADS after CDT promulgates attendant regulations. This bill is author-sponsored. It is supported by a number of civil rights

and advocacy organizations. No timely opposition has been received. The bill passed
out of the Senate Governmental Organization Committee on a vote of 14 to 0.

PROPOSED CHANGES TO THE LAW

Existing law:

1) Establishes CDT within the Government Operations Agency (GovOps), under
   the supervision of the Director of Technology (Director), also known as the State
   Chief Information Officer. (Gov. Code Sec. 11545(a).)

2) Provides that the duties of the Director include:
   a. advising the Governor on the strategic management and direction of the
      state's information technology (IT) resources;
   b. establishing and enforcing state IT strategic plans, policies, standards, and
      enterprise architecture, as specified;
   c. minimizing overlap, redundancy, and cost in state IT operations by
      promoting the efficient and effective use of information technology;
   d. providing technology direction to agency and department chief information
      officers to ensure the integration of statewide technology initiatives,
      compliance with IT policies and standards, and the promotion of the
      alignment and effective management of IT services;
   e. working to improve organizational maturity and capacity in the effective
      management of IT; and establishing performance management and
      improvement processes to ensure state IT systems and services are efficient
      and effective. (Gov. Code § 11545(b).)

3) Expresses the intent of the Legislature that policies and procedures developed by
   CDT and Department of General Services (DGS) pertaining to the acquisition of
   IT goods and services provide for all of the following: the expeditious and value-
   effective acquisition of IT goods and services to satisfy state requirements; the
   acquisition of IT goods and services within a competitive framework; the
   delegation of authority by DGS to each state agency that has demonstrated to
   DGS's satisfaction the ability to conduct value-effective IT goods and services
   acquisitions; and the review and resolution of protests submitted by any bidders
   with respect to any IT goods and services acquisitions. (Pub. Con. Code § 12101.)

4) Requires CDT, on or before September 1, 2024, to conduct, in coordination with
   other interagency bodies as it deems appropriate, 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. (Gov't Code
   § 11546.45.5(b).)

5) Requires the comprehensive inventory to include a description of all of the following:
   a) any decision the ADS can make or support, the intended benefits of that use, and the alternatives to that use;
   b) the results of any research assessing the efficacy and relative benefits of the uses and alternatives of the ADS described above;
   c) the categories of data and personal information the ADS uses to make its decisions;
   d) the measures in place, if any, to mitigate the risks, including cybersecurity risk and the risk of inaccurate, unfairly discriminatory, or biased decisions, of the ADS, including performance metrics, cybersecurity controls, privacy controls, risk assessments or audits for potential risks, and measures or processes in place to contest an automated decision. (Gov't Code § 11546.45.5(c).)

6) Requires CDT, on or before January 1, 2025, and annually thereafter, to submit a report, as specified, of the comprehensive inventory to the Assembly Committee on Privacy and Consumer Protection and the Senate Committee on Governmental Organization. This requirement expires on January 1, 2029. (Gov't Code § 11546.45.5(d).)

7) Defines the following terms:
   a) "Automated decision system" (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 automated decision system" 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.
   c) "State agency" includes every state office, department, division, bureau, the California State University, the Board of Parole Hearings, and specified boards. It does not include the University of California, the Legislature, the judicial branch, or any board, except as provided. (Gov't Code § 11546.45.5(a).)

This bill:

1) Requires CDT to develop and adopt regulations to create an AI risk management standard consistent with President Biden's Blueprint for an AI Bill of Rights, the Artificial Intelligence Risk Management Framework (AI RMF 1.0), released by the National Institute of Standards and Technology (NIST), and the Risk Management Framework for the Procurement of Artificial Intelligence (RMF PAIS 1.0), authored by the AI Procurement Lab and the Center for Inclusive Change.

2) Requires the standard to include the following:
   a) Methods for appropriate risk controls between the state agency and ADS vendor, including, but not limited to, reducing the risk through various mitigation strategies, eliminating the risk, or sharing the risk.
   b) Adverse incident monitoring procedures.
   c) Identification and classification of prohibited use cases and applications of ADS that the state shall not procure.
   d) An analysis about how the use of high-risk ADS can impact vulnerable individuals and communities.

3) Requires the standard to also include a detailed risk assessment procedure for procuring ADS that analyzes all of the following:
   a) Organizational and supply chain governance associated with the ADS.
   b) The purpose and use of the ADS.
   c) Any known potential misuses or abuses of the ADS.
   d) An assessment of the legality, traceability, and provenance of the data the ADS uses and the legality of the output of the ADS.
   e) The robustness, accuracy, and reliability of the ADS.
   f) The interpretability and explainability of the ADS.

4) Provides, that in developing the standard, CDT shall collaborate with organizations that represent state and local government employees and industry experts, including public trust and safety experts, community-based organizations, civil society groups, academic researchers, and research institutions focused on responsible AI procurement, design, and deployment.

5) Requires CDT to adopt regulations. Commencing 6 months after the date on which those regulations are approved and final, a state agency is prohibited from entering into a contract for a high-risk ADS, or any service that utilizes a high-risk ADS, unless the contract includes a clause that does all of the following:
   a) Provides a completed risk assessment of the relevant high-risk ADS.
   b) Requires the state agency or the high-risk ADS vendor, or both, to adhere to appropriate risk controls.
   c) Provides procedures for adverse incident monitoring.

     d) Requires authorization from the state agency before deployment of high-risk ADS upgrades and enhancements.

     e) Requires the state agency or the high-risk ADS vendor, or both, to provide notice to individuals that would likely be affected by the decisions or outcomes of the high-risk ADS, and information about how to appeal or opt-out of high-risk ADS decisions or outcomes.

6) Defines the relevant terms, including:

     a) "Artificial intelligence" means an engineered or machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs that can influence physical or virtual environments and that may operate with varying levels of autonomy.

     b) "Automated decision system" or "ADS" means a computational process derived from machine learning, statistical modeling, data analytics, or artificial intelligence 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.

     c) "High-risk automated decision system" or "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.

COMMENTS

1. Frameworks for responsible development and accountability in AI

Owing to recent advances in processing power and the rise of big data, AI's capacity and the scope of its applications have expanded rapidly, impacting how we communicate, interact, entertain ourselves, travel, transact, and consume media. Since the widespread introduction of AI systems such as ChatGPT, the world has been in awe of the powers of GenAI, which is a type of artificial intelligence that can create new content, such as text, images, code, or music, by learning from existing data. GenAI models can produce realistic and novel artifacts that resemble the data they were trained on, but do not copy it. For example, generative AI can write a poem, draw a picture, or compose a song based on a given prompt or theme.

In ways we may not fully comprehend, AI empowers and encumbers us. It has been used to accelerate productivity, achieve efficiencies, liberate us from drudgery, write our college essay, help us understand and enjoy the world, connect with each other, and live longer, fuller lives. It has also been used to constrain personal autonomy, compromise privacy and security, foment social upheaval, exacerbate inequality, spread misinformation, and subvert democracy. For good or ill, its transformative potential seems boundless. With these recent dramatic advances in the capabilities of AI systems,

the need for frameworks for accountability and responsible development have become ever more urgent.

In January of 2017, AI researchers, economists, legal scholars, ethicists, and philosophers met in Asilomar, California to discuss principles for managing the responsible development of AI. The collaboration resulted in the Asilomar Principles. Aspirational rather than prescriptive, these 23 principles were intended to initiate and frame a dialogue by providing direction and guidance for policymakers, researchers, and developers. Its endorsers include 1,200 leading experts in the field of AI, including DeepMind founder Demis Hassabis and the late Stephen Hawking.

The Legislature subsequently adopted ACR 215 (Kiley, Ch. 206, Stats. 2018), which added the State of California to that list by endorsing the Asilomar Principles as guiding values for the development of artificial intelligence and related public policy. In broad strokes, those principles aim to do the following:

- *Research issues:* create beneficial AI; direct funding toward beneficial innovation; maintain constructive and healthy exchanges between AI researchers and policymakers; promote a culture of trust, cooperation, and transparency among researchers and developers of AI; and avoid corner-cutting on safety standards.

- *Ethics and values:* promote safety, failure transparency, judicial transparency, and responsible innovation; align human values with innovation; protect privacy and liberty; ensure that the benefits and prosperity created by AI are broadly shared; maintain human control over AI; develop AI that supports rather than subverts social and civil processes; and avoid an AI arms race.

- *Longer-term issues:* avoid assumptions regarding the capabilities of AI; give AI its due attention; and **recognize that its risks are potentially catastrophic or existential**. [emphasis added]

As directed by the National AI Initiative Act of 2020, NIST developed the AI Risk Management Framework to assist entities designing, developing, deploying, and using AI systems to help manage the many risks of AI and promote trustworthy and responsible development and use of AI systems. That framework highlights the serious risks at play and the uniquely challenging nature of addressing them in this context:

> Artificial intelligence (AI) technologies have significant potential to transform society and people's lives – from commerce and health to transportation and cybersecurity to the environment and our planet. AI technologies can drive inclusive economic growth and support scientific advancements that improve the conditions of our world. AI technologies, however, also pose risks that can negatively impact individuals, groups, organizations, communities, society, the environment, and the planet. Like

risks for other types of technology, AI risks can emerge in a variety of ways and can be characterized as long- or short-term, high or low-probability, systemic or localized, and high- or low-impact.

While there are myriad standards and best practices to help organizations mitigate the risks of traditional software or information-based systems, the risks posed by AI systems are in many ways unique. AI systems, for example, may be trained on data that can change over time, sometimes significantly and unexpectedly, affecting system functionality and trustworthiness in ways that are hard to understand. AI systems and the contexts in which they are deployed are frequently complex, making it difficult to detect and respond to failures when they occur. AI systems are inherently socio-technical in nature, meaning they are influenced by societal dynamics and human behavior. AI risks – and benefits – can emerge from the interplay of technical aspects combined with societal factors related to how a system is used, its interactions with other AI systems, who operates it, and the social context in which it is deployed.

These risks make AI a uniquely challenging technology to deploy and utilize both for organizations and within society. [. . .]

AI risk management is a key component of responsible development and use of AI systems. Responsible AI practices can help align the decisions about AI system design, development, and uses with intended aim and values. Core concepts in responsible AI emphasize human centricity, social responsibility, and sustainability. AI risk management can drive responsible uses and practices by prompting organizations and their internal teams who design, develop, and deploy AI to think more critically about context and potential or unexpected negative and positive impacts. Understanding and managing the risks of AI systems will help to enhance trustworthiness, and in turn, cultivate public trust.

More recently the Biden Administration has published its Blueprint for an AI Bill of Rights, which is a set of five principles and associated practices to help guide the design, use, and deployment of AI to protect the rights of the American public:

- *Safe and Effective Systems*: You should be protected from unsafe or ineffective systems. Automated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system.

- *Algorithmic Discrimination Protections*: Designers, developers, and deployers of automated systems should take proactive and continuous measures to protect individuals and communities from algorithmic discrimination and to use and

design systems in an equitable way. This protection should include proactive equity assessments as part of the system design, use of representative data and protection against proxies for demographic features, ensuring accessibility for people with disabilities in design and development, pre-deployment and ongoing disparity testing and mitigation, and clear organizational oversight.

- *Data Privacy*: You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used. You should be protected from violations of privacy through design choices that ensure such protections are included by default, including ensuring that data collection conforms to reasonable expectations and that only data strictly necessary for the specific context is collected. Designers, developers, and deployers of automated systems should seek your permission and respect your decisions regarding collection, use, access, transfer, and deletion of your data in appropriate ways and to the greatest extent possible; where not possible, alternative privacy by design safeguards should be used. Systems should not employ user experience and design decisions that obfuscate user choice or burden users with defaults that are privacy invasive. Consent should only be used to justify collection of data in cases where it can be appropriately and meaningfully given. Any consent requests should be brief, be understandable in plain language, and give you agency over data collection and the specific context of use; current hard-to-understand notice-and-choice practices for broad uses of data should be changed. Enhanced protections and restrictions for data and inferences related to sensitive domains, including health, work, education, criminal justice, and finance, and for data pertaining to youth should put you first. In sensitive domains, your data and related inferences should only be used for necessary functions, and you should be protected by ethical review and use prohibitions. You and your communities should be free from unchecked surveillance; surveillance technologies should be subject to heightened oversight that includes at least pre-deployment assessment of their potential harms and scope limits to protect privacy and civil liberties. Continuous surveillance and monitoring should not be used in education, work, housing, or in other contexts where the use of such surveillance technologies is likely to limit rights, opportunities, or access. Whenever possible, you should have access to reporting that confirms your data decisions have been respected and provides an assessment of the potential impact of surveillance technologies on your rights, opportunities, or access.

- *Notice and Explanation*: You should know that an automated system is being used and understand how and why it contributes to outcomes that impact you. Designers, developers, and deployers of automated systems should provide generally accessible plain language documentation including clear descriptions of the overall system functioning and the role automation plays, notice that such systems are in use, the individual or organization responsible for the system, and

explanations of outcomes that are clear, timely, and accessible. Such notice should be kept up-to-date and people impacted by the system should be notified of significant use case or key functionality changes. You should know how and why an outcome impacting you was determined by an automated system, including when the automated system is not the sole input determining the outcome.

- *Human Alternatives, Consideration, and Fallback*: You should be able to opt out from automated systems in favor of a human alternative, where appropriate. Appropriateness should be determined based on reasonable expectations in a given context and with a focus on ensuring broad accessibility and protecting the public from especially harmful impacts.[1]

TechEquity, an organization committed to ensuring technology's evolution benefits everyone equitably, has also laid out their straightforward AI Policy Principles:

- People who are impacted by AI must have agency to shape the technology that dictates their access to critical needs like employment, housing, and healthcare.
- The burden of proof must lie with developers, vendors, and deployers to demonstrate that their tools do not create harm—and regulators, as well as private [individuals], should be empowered to hold them accountable.
- Concentrated power and information asymmetries must be addressed in order to effectively regulate the technology.

Earlier this year, the AI Procurement Lab and the Center for Inclusive Change (woman-owned non-profit and business, respectively) released their own framework, the Risk Management Framework for the Procurement of AI Systems (RMF PAIS 1.0) "to provide organizations and procurement teams with an essential tool that classifies the risks embedded within each procurement opportunity for the purposes of risk awareness, assessment, measurement, mitigation, treatment, control, monitoring, and management." They explain the framework and the need for it:

Unfortunately, certain AI use cases can present new and novel risks to organizations (beyond reputational and legal damages). Such use cases include systems designed to deliver critical, sometimes life-altering, decisions in the context of employment, health, education, housing, finance, public assistance, critical infrastructure, essential utilities, law enforcement, immigration, justice, legal services, biometric identification, safety components and other consequential decision systems. Systems developed for these types of use cases are commonly referred to as "high-

---

[1] *Blueprint For An AI Bill Of Rights* (October 2022) Office of Science and Technology Policy, https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

risk" systems. They pose a high-risk to those impacted either directly or indirectly by the AI system in question in significant or critical ways.

[RMF PAIS 1.0] focuses primarily on risk management for high-risk systems. . . . More specifically, given that high-risk systems can produce great advantages in terms of efficiency gains and consistency in decision-making output, they also have the potential to impact a person's safety, civil rights, and/or fundamental human rights and dignity. AI researchers have repeatedly identified a variety of harms that require our attention when deploying such systems in general and especially when deploying high-risk AI systems. The known harms that have been perpetrated by these systems are particularly troubling because they stem from historic systemic bias that can be widely scaled across vulnerable populations through powerful and non-transparent algorithmic computations, generating unfair, unequal, disproportionate, and potentially life-altering outcomes.

Without proper governance and risk mitigation practices, the computational power of high-risk systems can pose unwanted threats to our fellow humans--causing more than just reputation or legal damages for organizations. Unmitigated risks and negative outcomes can, in some cases, mean life or death to the end uses. Hence, the stakes are high, and organizations must find ways to control the risks in order to establish trust for all stakeholders. The RMF PAIS 1.0 provides a guide to identifying and controlling these risks through the use of standard procurement lifecycle processes in a practical and responsible way.[2]

Seeking to establish a framework for California, Governor Gavin Newsom issued Executive Order N-12-23 "to study the development, use, and risks of artificial intelligence (AI) technology throughout the state and to develop a deliberate and responsible process for evaluation and deployment of AI within state government."[3]

The executive order includes the following provisions:
- **Risk-Analysis Report**: Directs state agencies and departments to perform a joint risk-analysis of potential threats to and vulnerabilities of California's critical energy infrastructure by the use of GenAI.

---

[2] Dr. Cari L. Miller & Gisele Waters, Ph.D., Risk Management Framework for the Procurement of AI Systems (2024) AI Procurement Lab & The Center for Inclusive Change, https://20596328.fs1.hubspotusercontent-na1.net/hubfs/20596328/Risk%20Management%20Framework%20for%20AI%20Procurement-4.pdf.

[3] Press Release, *Governor Newsom Signs Executive Order to Prepare California for the Progress of Artificial Intelligence* (September 6, 2023) Office of Governor Gavin Newsom, https://www.gov.ca.gov/2023/09/06/governor-newsom-signs-executive-order-to-prepare-california-for-the-progress-of-artificial-intelligence/.

- **Procurement Blueprint**: To support a safe, ethical, and responsible innovation ecosystem inside state government, agencies will issue general guidelines for public sector procurement, uses, and required training for application of GenAI – building on the White House's Blueprint for an AI Bill of Rights and the National Institute for Science and Technology's AI Risk Management Framework. State agencies and departments will consider procurement and enterprise use opportunities where GenAI can improve the efficiency, effectiveness, accessibility, and equity of government operations.

- **Beneficial Uses of GenAI Report**: Direct state agencies and departments to develop a report examining the most significant and beneficial uses of GenAI in the state. The report will also explain the potential harms and risks for communities, government, and state government workers.

- **Deployment and Analysis Framework**: Develop guidelines for agencies and departments to analyze the impact that adopting GenAI tools may have on vulnerable communities. The state will establish the infrastructure needed to conduct pilots of GenAI projects, including CDT-approved environments or "sandboxes" to test such projects.

- **State Employee Training**: To support California's state government workforce and prepare for the next generation of skills needed to thrive in the GenAI economy, agencies will provide trainings for state government workers to use state-approved GenAI to achieve equitable outcomes, and will establish criteria to evaluate the impact of GenAI to the state government workforce.

- **Legislative Engagement**: Engage with Legislative partners and key stakeholders, including academic institutions, in a formal process to develop policy recommendations for responsible use of AI, including any guidelines, criteria, reports, and/or training.

- **Evaluate Impacts of AI on an Ongoing Basis**: Periodically evaluate for potential impact of GenAI on regulatory issues under the respective agency, department, or board's authority and recommend necessary updates as a result of this evolving technology.

2. <u>A framework for California: guiding procurement and use of ADS</u>

ADS are algorithm-driven applications that can assist or supplant human decisionmaking processes in areas such as credit decisions, employment screening, insurance eligibility, and the delivery of government services. ADS process enormous datasets and make decisions with speed and reliability that vastly exceed human capabilities. However, poorly designed or poorly understood systems can create unfair, biased, and inaccurate results. When deployed by government agencies, flawed ADS

may disproportionately harm low-income families and communities of color and undermine trust in the public sector. Moreover, norms of participatory governance and due process may be jeopardized when ADS affect agency policymaking, adjudications, or enforcement.

Last year, AB 302 (Ward, Ch. 800, Stats. 2023) required 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. This bill directs CDT to develop and adopt regulations to create an AI risk management standard that will govern the use of ADS in California. State agencies are to use the framework and attendant regulations to guide their use and procurement of ADS.

The standard must include methods for appropriate risk controls between the relevant state agency and its ADS vendor, adverse incident monitoring procedures, and an analysis about how the use of high-risk ADS can impact vulnerable individuals and communities. The standard is also required to draw a bright line to identify and classify specific use cases and applications for which ADS will not be allowed in state entities.

The bill requires the standard to include a detailed risk assessment procedure for procuring ADS that analyzes all of the following:
- Organizational and supply chain governance associated with the ADS.
- The purpose and use of the ADS.
- Any known potential misuses or abuses of the ADS.
- An assessment of the legality, traceability, and provenance of the data the ADS uses and the legality of the output of the ADS.
- The robustness, accuracy, and reliability of the ADS.
- The interpretability and explainability of the ADS.

The bill calls out a number of the frameworks discussed above to guide creation of California's framework, namely President Biden's Blueprint for an AI Bill of Rights, the AI RMF 1.0, and the RMF PAIS 1.0.

According to the author:

> Artificial intelligence stands to have the largest influence on society since the dawn of the Digital Age. It has the potential to provide incredible societal benefits if harnessed appropriately, but threatens to pose terrible consequences if safeguards are not put in place as it becomes integrated into everyday life. The research and guardrails around generative AI services will become the standard that guides the technology as it proliferates throughout every sector of our economy. The rapid growth of this technology's capability over even just the past year is clear warning, we must set these safety parameters now. The public has been left

vulnerable to the dangers AI poses because of congressional failure to act and the regulatory desert they've created. We cannot wait for Congress to overcome their dysfunction, so California must step in and step up to lead. We are the home of Silicon Valley, we are leading the way in AI development and innovation and we are also one of the largest purchasing entities of technology in the world. But, as companies develop the cutting edge of this technology, we must be sure the guardrails around its use are built with California values.

3. Stakeholder positions

The Electronic Frontier Foundation writes in support:

SB 892 rightly recognizes the potential risks of AI technology alongside its potential for fostering innovation. It sets the foundation for a framework that reflects principles that are necessary to create an environment that the public can trust, including requirements to:
- **Assess** the risks associated with procuring high-risk AI systems and ADS, particularly in areas that have a high impact on individual rights, including but not limited to health, education, employment, insurance, utilities, critical infrastructure, public services, and justice/legal.
- **Require** the development and adoption of risk management standards that include risk assessment, appropriate risk controls, and adverse incident monitoring when procuring such systems.
- **Require** engagement and consultation with a diversity of stakeholders, including those most affected by government use of these systems, such as public trust and safety experts, community-based organizations, civil society groups, workers, academic researchers and members of the public.
- **Ensure** people have access to robust notices and opportunities to appeal or opt out of high-risk AI system outcomes or ADS decisions.

Writing in support, the Greenlining Institute states:

This bill, focused on the safe procurement of artificial intelligence (AI) decision systems in the public sector, is necessary to ensure that government agencies do not deploy AI systems that are inaccurate, unfair, or otherwise harm Californians. This accountability is key to building trust and encouraging public adoption of AI technologies. The Greenlining Institute encourages further clarifying that risk assessments should be part of the safety, privacy and non discrimination standards developed by the Department of Technology.

SUPPORT

AI Procurement Lab
American Federation of Musicians
Electronic Frontier Foundation
Greenlining Institute
Secure Justice
Surveillance Resistance Lab
TechEquity Collaborative

OPPOSITION

None received

RELATED LEGISLATION

Pending Legislation:

SB 893 (Padilla, 2024) requires GovOps, the Governor's Office of Business and Economic Development, and CDT to collaborate to establish the California Artificial Intelligence Research Hub in GovOps, as prescribed. SB 893 requires the hub to serve as a centralized entity to facilitate collaboration between government agencies, academic institutions, and private sector partners to advance AI research and development that seeks to harness the technology's full potential for public benefit while safeguarding privacy, advancing security, and addressing risks and potential harms to society, as prescribed. SB 893 is currently in this Committee.

SB 896 (Dodd, 2024) largely codifies Governor Newsom's executive order on the use of GenAI. The bill requires assessments of the beneficial uses, potential harms, and risks to critical infrastructure of GenAI. The bill calls for the development of guidelines for public sector procurement, uses, and required trainings for the use of GenAI. The bill places obligations on state entities with respect to the use of GenAI and ADS. SB 896 is currently in this Committee.

SB 942 (Becker, 2024) establishes the California AI Transparency Act, which, among other things, requires a covered provider, as defined, to create an AI detection tool by which a person can query the covered provider as to the extent to which text, image, video, audio, or multimedia content was created, in whole or in part, by a generative AI system, as defined, provided by the covered provider that meets certain criteria. Covered providers are required to include in AI-generated content a visible disclosure that, among other things, includes a clear and conspicuous notice, that identifies the content as generated by AI. SB 942 requires a covered provider to register with CDT and provide them a URL to any AI detection tool it has created. SB 942 is currently in this Committee.

SCR 17 (Dodd, 2023) affirms the California Legislature's commitment to President Biden's vision for a safe AI and the principles outlined in the "Blueprint for an AI Bill of Rights" and expresses the Legislature's commitment to examining and implementing those principles in its legislation and policies related to the use and deployment of automated systems. SCR 17 is currently in the Assembly Privacy and Consumer Protection Committee.

AB 331 (Bauer-Kahan, 2023) prohibits "algorithmic discrimination," that is, use of an automated decision tool to contribute to unjustified differential treatment or outcomes that may have a significant effect on a person's life. It requires any deployer of an automated decision tool to perform an impact assessment for those tools and to notify any natural person that is the subject of the consequential decision that an automated decision tool is being used to make, or be a controlling factor in making, the consequential decision. AB 331 was held in the Senate Appropriations Committee.

AB 2013 (Irwin, 2024) requires, on or before January 1, 2026, a developer, as defined, of an AI system or service to post on the developer's website documentation regarding the data used to train the AI system or service, as specified. AB 2013 is currently in the Assembly Privacy and Consumer Protection Committee.

AB 2930 (Bauer-Kahan, 2024) requires, among other things, a deployer and a developer of an automated decision tool to, on or before January 1, 2026, and annually thereafter, perform an impact assessment for any automated decision tool the deployer uses that includes, among other things, a statement of the purpose of the automated decision tool and its intended benefits, uses, and deployment contexts. The assessments must be provided to the Civil Rights Department within 7 days of a request. AB 2930 requires a deployer to, at or before the time an automated decision tool is used to make a consequential decision, notify any natural person that is the subject of the consequential decision that an automated decision tool is being used to make, or be a controlling factor in making, the consequential decision and to provide that person with, among other things, a statement of the purpose of the automated decision tool.
AB 2930 is currently in the Assembly Privacy and Consumer Protection Committee.

Prior Legislation:

AB 302 (Ward, Ch. 800, Stats. 2023) *See* Comment 2.

AB 13 (Chau, 2021) would have established the Automated Decision Systems Accountability Act, which, in the context of the State's procurement policies, promotes oversight over ADS that pose a high risk of adverse impacts on individual rights. The bill was eventually gutted and amended to address a different topic.

SB 444 (Umberg, 2019) would have requested the Regents of the University of California (UC) to enact a resolution authorizing the law schools at UC Berkeley and UC Irvine to

participate in a pilot project to develop AI or machine-learning solutions to address access to justice issues faced by self-representing litigants in their respective courts. The bill died in the Assembly Higher Education Committee.

AB 1576 (Calderon, 2019) would have required the Secretary of GovOps to appoint participants to an AI working group to evaluate the uses, risks, benefits, and legal implications associated with the development and deployment of AI by California-based businesses. The bill was held on the Senate Appropriations Committee suspense file.

SJR 6 (Chang, Res. Ch. 112, Stats. 2019) urged the President and the Congress of the United States to develop a comprehensive AI Advisory Committee and to adopt a comprehensive AI policy.

ACR 215 (Kiley, Resolution Ch. 206, Stats. 2018) *See* Comment 1.

PRIOR VOTES:

Senate Governmental Organization Committee (Ayes 14, Noes 0)

****