Date of Hearing: July 2, 2024

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION
Rebecca Bauer-Kahan, Chair
SB 892 (Padilla) – As Amended June 21, 2024

AS PROPOSED TO BE AMENDED

SENATE VOTE: 37-0

SUBJECT: Public contracts: automated decision systems: AI risk management standards

SYNOPSIS

Over the past thirty years, automated decision tools (ADTs) have become ubiquitous in society. ADTs are enabled by artificial intelligence, and when they are embedded into a decisionmaking process they can increase efficiency and productivity many-fold compared to an equivalent human worker. But ADTs share many faults with the humans they replace. If they are not thoroughly tested, the decisions they produce can be incorrect or even biased. This becomes especially problematic when an ADT is used to make a consequential decision: for example, to determine whether to hire an employee, or prioritize a patient seeking healthcare, or evaluate a student's scholarship application.

California's state government regularly uses ADTs to engage in consequential decisionmaking. This bill, as proposed to be amended, would require the Department of Technology (CDT) to implement a procurement standard for ADTs in order to assess these tools for bias before they can affect the lives of Californians. Committee amendments would make this bill consistent with Proposition 24 (2020), which granted the California Privacy Protection Agency (Privacy Agency) regulatory authority over businesses' use of automated decisionmaking technologies. A second Committee amendment would align various definitions in this bill with those in a related bill, AB 2930 (Bauer-Kahan, 2024.) Finally, a Committee amendment would prohibit the procurement of ADTs by the state until CDT has developed and implemented necessary procurement standards. The amendments are set forth in full below.

This bill is author-sponsored and supported by a variety of cities and advocacy groups, including Oakland Privacy, Techequity Collaborative, and the Electronic Frontier Foundation. It has no opposition. If this bill passes out of this committee it will next be heard by the Assembly Appropriations Committee.

SUMMARY: Charges CDT with crafting a procurement standard for the acquisition of ADTs, and imposes a moratorium on the procurement of new ADTs until that standard is in place. Specifically, **this bill**:

1) Defines "artificial intelligence" to mean "an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments."

2) Defines “automated decision tool” (ADT) to mean “an artificial intelligence system or service that makes a consequential decision, or is a substantial factor in making consequential decisions.”

3) Defines “consequential decision” to mean a decision or judgment that has a legal, material, or similarly significant effect on an individual’s life relating to access to government benefits or services, assignments of penalties by government, or the impact of, access to, or the cost, terms, or availability of, any of the following:

    a) Employment with respect to all of the following:

        i. Pay or promotion.

        ii. Hiring or termination.

        iii. Automated task allocation that limits, segregates, or classifies employees for the purpose of assigning or determining material terms or conditions of employment.

    b) Education and vocational training as it relates to all of the following:

        i. Assessment or placement.

        ii. Detecting student cheating or plagiarism.

        iii. Accreditation.

        iv. Certification.

        v. Admissions or enrollment.

        vi. Discipline.

        vii. Evaluation.

        viii. Financial aid or scholarships.

    c) Housing or lodging, including rental or short-term housing or lodging.

    d) All of the following essential utilities:

        i. Electricity.

        ii. Heat.

        iii. Water.

        iv. Internet or telecommunications access.

        v. Transportation.

    e) Family planning.

f) Adoption services, reproductive services, or assessments related to child protective services.

g) Health care or health insurance, including mental health care, dental, or vision.

h) Financial services, including a financial service provided by a mortgage company, mortgage broker, or creditor.

i) All of the following aspects of the criminal justice system:

i. Risk assessments for pretrial hearings.

ii. Sentencing.

iii. Parole.

j) Legal services.

k) Private arbitration.

l) Mediation.

m) Voting.

4) Defines "substantial factor" to mean an element of a decisionmaking process that is capable of altering the outcome of the process.

5) Requires CDT to develop and adopt regulations to create an ADT procurement standard following the adoption of regulations by the Privacy Agency related to business' use of ADTs, or the passage of similarly comprehensive legislation by the Legislature. Requires CDT to consider principles and industry standards addressed in various relevant publications, as specified.

6) Requires the ADT procurement standard to include all of the following:

a) A detailed risk assessment procedure that analyzes:

i) Organizational and supply chain governance associated with the ADT.

ii) The purpose and use of the ADT.

iii) Any known potential misuses or abuses of the ADT.

iv) An assessment of the legality, traceability, and provenance of the data the ADT uses and the legality of the output of the ADT.

v) The robustness, accuracy, and reliability of the ADT.

vi) The interpretability and explainability of the ADT.

b) Methods for appropriate risk controls between the state agency and ADT vendor, including, but not limited to, reducing the risk through various mitigation strategies, eliminating the risk, or sharing the risk.

c) Adverse incident monitoring procedures.

d) Identification and classification of prohibited use cases and applications of ADT that the state shall not procure.

e) A detailed equity assessment, as specified.

f) An assessment that analyzes the level of human oversight associated with the use of the ADT.

g) Adherence to data minimization standards.

7) Requires CDT to do all of the following in developing the ADT procurement standard:

a) Collaborate with organizations that represent state and local government employees and industry experts, as specified.

b) Consult with the California Privacy Protection Agency.

c) Solicit public comment on the ADT procurement standard.

8) Requires any regulations adopted by CDT pursuant to this section be consistent with regulations adopted by the Privacy Agency pursuant to paragraph (16) of subdivision (a) of Section 1798.185 of the Civil Code, as well as any similarly comprehensive legislation passed by the Legislature.

9) Requires CDT to annually review and update the ADT procurement standard and any regulations adopted pursuant to this section.

10) Prohibits a state agency from procuring an ADT, entering into a contract for an ADT, or entering into a contract for any service that utilizes an ADT, until the department has adopted regulations pursuant to this section.

11) Prohibits a state agency from entering into a contract for an ADT or entering into a contract for any service that utilizes an ADT after CDT has adopted regulations pursuant to this section unless the contract includes a clause that does all of the following:

a) Provides a completed risk assessment, as specified.

b) Requires the state agency and/or ADT vendor, as appropriate, to adhere to appropriate procurement standards.

c) Provides procedures for adverse incident monitoring.

d) Requires authorization from the state agency before deployment of ADT upgrades and enhancements.

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

f) Provides a termination right in the event of a significant breach of responsibility or violation by the vendor.

EXISTING LAW:

1) Establishes the California Privacy Protection Agency and vests it with full administrative power, authority, and jurisdiction to implement and enforce the California Consumer Privacy Act of 2018. (Civ. Code § 1798.199.10.)

2) Requires the Attorney General to issue regulations governing access and opt-out rights with respect to businesses' use of automated decisionmaking technology. Passes along this authority to the California Privacy Protection Agency after a set period of time. (Civ. Code § 1798.185.)

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

4) 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.)

5) 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. Code § 11546.45.5(b).)

6) 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) "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).)

FISCAL EFFECT: As currently in print, this bill is keyed fiscal.

COMMENTS:

1) **Background.** The term "automated decisionmaking tool," or "ADT," refers to computer systems that replace or supplement human decisionmaking ability. ADTs are used to decide whom to hire, what treatments patients should receive, whether to grant parole, and many other life-altering decisions that used to be made entirely by human beings.

Using an ADT is, in many cases, a good thing. Computers can generally draw upon far more digitally-encoded data, and process this data far more quickly, than any human being. A recent *New York Times* article described the development of a breast cancer screening system that uses artificial intelligence:

> From the millions of cases the system is fed, the technology creates a mathematical representation of normal mammograms and those with cancers. With the ability to look at each image in a more granular way than the human eye, it then compares that baseline to find abnormalities in each mammogram.
>
> Last year, after a test on more than 275,000 breast cancer cases, Kheiron [Medical Technologies, the system's developer] reported that its A.I. software matched the performance of human radiologists when acting as the second reader of mammography scans. It also cut down on radiologists' workloads by at least 30 percent because it reduced the number of X-rays they needed to read. In other results from a Hungarian clinic last year, the technology increased the cancer detection rate by 13 percent because more malignancies were identified.[1]

However, ADTs are far from flawless. Care should be taken to avoid what New York University Professor Meredith Broussard describes as "the notion that computers are more 'objective' or 'unbiased' because they distill questions and answers down to mathematical evaluation…an unwavering faith that if the world used more computers, and used them properly, social problems would disappear and we'd create a digitally enabled utopia."[2] Many aspects of the real world are not reducible to data, and many complex situations present competing demands that human beings remain best equipped to assess.

ADT usage continues to increase. Global consulting firm McKinsey & Company reports that the adoption of artificial intelligence among the organizations it surveyed more than doubled between 2017 and 2022, with 50 percent now reporting they use the technology.[3]

ADTs use artificial intelligence. ADTs use artificial intelligence (AI) to make consequential decisions. AI uses algorithms – sets of rules – to transform inputs into outputs. Inputs and outputs can be anything a computer can process: numbers, text, audio, video, or movement. This is because AI is not fundamentally different from other computer functions. Its novelty lies in its

---

[1] Santariano & Metz, "Using A.I. to Detect Breast Cancer That Doctors Miss," *New York Times,* Mar. 5, 2023, https://www.nytimes.com/2023/03/05/technology/artificial-intelligence-breast-cancer-detection.html.
[2] Broussard, "Artificial Unintelligence," *MIT Press,* 2019.
[3] McKinsey & Co., "The state of AI in 2022—and a half decade in review," Dec. 6, 2022, https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review.

application: unlike standard computer functions, AI is able to accomplish tasks that are normally performed by humans.

Bias and discrimination can result from poor training data. There is a famous saying in computer science: "garbage in, garbage out." The performance of an ADT is directly impacted by the quality, quantity, and relevance of the data used to train it.[4] If the data used to train the ADT is biased, the tool's outputs will be similarly biased. Over the past thirty years, several industries have been forced to contend with this fact as they have attempted to introduce ADTs into their workflows. Specific examples follow:

Hiring and recruitment tools. It is no secret that people of various races, genders, and cultures are not distributed equally throughout the workforce. An ADT that is trained on historical data to make hiring decisions will be predisposed to maintain the ratios it is trained on; as described by Aditya Malik, the Founder and CEO of Valuematrix.ai:

> Generative AI, for all its grandeur, has the potential to perpetuate latent biases inherited from human creators. A disconcerting echo of historical prejudices may inadvertently seep into the algorithms. Imagine a scenario where previous senior managers, driven by biases of gender, age, faith or race, rejected candidates for misguided reasons. The AI, if not vigilantly curated, might misconstrue these patterns as indicators of incompetence, thus exacerbating the exclusion of qualified candidates from underrepresented backgrounds.[5]

This was notoriously experienced by Amazon, who considered automating their hiring practices in the early 2010s. They opted against this approach in 2015 when they realized that their ADT-enabled system was not rating candidates in a gender-neutral way. In fact, their system was excluding women from the pool of acceptable candidates because it had been trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of inequities across the tech industry.[6]

Sentencing and bail decisions. ADTs are frequently used to inform sentencing and bail decisions. These tools are trained using historical data, and the predictions they make can therefore reflect historical bias. A 2016 Propublica study dove into the use of one such tool – COMPAS – in Broward County, Florida.[7] The study determined that Black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than Black defendants to be incorrectly flagged as low risk.

The for-profit company that developed this tool, Northpointe, does not publicly disclose the calculations used to arrive at defendants' risk scores, so it is not possible for either defendants or

---

[4] Rohit Sehgal, "AI Needs Data More Than Data Needs AI", *Forbes,* Oct. 5, 2023, https://www.forbes.com/sites/forbestechcouncil/2023/10/05/ai-needs-data-more-than-data-needs-ai/.
[5] Aditya Malik, "AI Bias In Recruitment: Ethical Implications And Transparency," *Forbes,* Sep. 25, 2023, https://www.forbes.com/sites/forbestechcouncil/2023/09/25/ai-bias-in-recruitment-ethical-implications-and-transparency/.
[6] Jeffrey Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters,* Oct. 9, 2018, https://www.reuters.com/article/amazoncom-jobs-automation/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSL2N1VB1FQ/.
[7] Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, "Machine Bias", *Propublica,* May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

the public to see what might be driving the disparity. These discrepancies mirror historical injustices perpetuated against Black Americans by California's criminal justice system.[8]

Healthcare. When ADTs are deployed in healthcare, biased historical data can lead to patients being recommended substandard care on the basis of their race or ethnicity. In 2007, an ADT was developed to help doctors estimate whether it was safe for people who had delivered previous children through cesarean section to deliver subsequent children vaginally – a risky procedure. The ADT considered various health relevant factors as it made its decision, such as the woman's age, her reason for the previous cesarean, and how long ago the cesarean had been performed. However, a 2017 study found that the ADT was biased; it predicted Black and Latino people were less likely to have a successful vaginal birth after a cesarean than non-Hispanic white women. As a result, doctors performed more cesareans on Black and Latino people than on white people.[9] These discrepancies perpetuate historical biases – Black Americans, for example, have historically received a lower standard of healthcare than their white counterparts.[10]

Credit and loan approval. Financial tools that utilize ADTs are similarly susceptible to bias and discrimination. An investigation by The Markup (and co-published by the Associated Press) revealed that in 2019, lenders were more likely to deny home loans to people of color; in particular, lenders were 40 percent more likely to turn down Latino applicants for loans, 50 percent more likely to deny Asian/Pacific Islander applicants, and 70 percent more likely to deny Native American applicants than similar white applicants. Lenders were 80 percent more likely to reject Black applicants than similar white applicants. In every case, the prospective borrowers of color looked almost exactly the same on paper as the white applicants, except for their race.[11]

2) **What this bill would do.** This bill would require CDT to develop a set of procurement standard for ADTs, and would impose a moratorium on the procurement of ADTs until these standards are in place. The bill requires CDT to draw from a variety of sources when creating the standards, including the California Privacy Protection Agency.

The premises of this bill are, first, that the State should be much more informed about the ADTs it procures, and second, that vendors of ADTs ought to be much more conscious and deliberate about what the impacts of their tools might be. In short, California ought to get a handle on ADTs while it still can before they become yet another phenomenon—like social media, greenhouse gases, guns, and urban sprawl—that could have been dealt with deliberately and intelligently in the past, but was instead allowed to expand without any constraints past the point of being manageable.

3) **Author's statement:**

---

[8] California Task Force to Study and Develop Reparation Proposals for African American, "Final Report," California.
[9] Caleb J Colón-Rodríguez, "Shedding Light on Healthcare Algorithmic and Artificial Intelligence Bias," *U.S. Department of Health & Human Services Office of Minority Health,* Jul. 12, 2023, minorityhealth.hhs.gov/news/shedding-light-healthcare-algorithmic-and-artificial-intelligence-bias.
[10] California Task Force to Study and Develop Reparation Proposals for African American, "Final Report," California.
[11] Emmanuel Martinez and Lauren Kirchner, "The Secret Bias Hidden in Mortgage-Approval Algorithms," *Markup,* Aug. 25, 2021, themarkup.org/denied/2021/08/25/the-secret-bias-hidden-in-mortgage-approval-algorithms.

> 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. By doing so, we can use California's purchasing power to influence the development of appropriate standards around automated-decision making systems.

4) **Committee amendments.** Committee amendments strengthen this bill by aligning it with other major ADT legislation and regulations, and by introducing a moratorium on the acquisition of new ADTs until safety measures have been enacted.

AB 2930 (Bauer-Kahan, 2024) would require developers and deployers of ADTs to perform risk assessments on their tools before using them to make consequential decisions. That bill defines a "state government deployer" to mean "a state government agency that uses an automated decision tool to make a consequential decision" – these are the same entities referred to in the current bill by the phrase "state agency." Furthermore, "developers" in AB 2930 are the likely source of any ADTs procured under the terms of this bill. A Committee amendment has been introduced to align the definitions of "ADT," "consequential decision," and "substantial factor" between these two bills. The author has committed to maintaining this alignment as the two bills progress through the legislative process.

California's Proposition 24 (2020) granted the California Privacy Protection Agency regulatory authority over access and opt-out rights with respect to businesses' use of automated decisionmaking technologies. In order to avoid conflicting with this mandate, an amendment has been introduced to delay the implementation of procurement standards by the Department of Technology until after the Privacy Agency has adopted regulations pursuant to Section 1798.185 of the Civil Code, or until the legislature has adopted a similarly comprehensive framework (which is necessarily consistent with Prop 24.) Another amendment requires that any regulations adopted by CDT pursuant to the terms of this bill be consistent with regulations adopted by the Privacy Agency pursuant to Section 1798.185 of the Civil Code.

Finally, a Committee amendment has been introduced which takes seriously this bill's premise: that ADTs are capable of great harm to individuals and society if they are not carefully assessed prior to deployment. According to the author:

> The procurement of technology without proper guardrails has already led to adverse impacts to communities in various states and California. Recently, the Electronic Privacy Information Center (EPIC) released a report titled "Outsourced and Automated: How AI Companies Have Taken Over Government Decision-Making", showcasing how many agencies across all fifty states have outsourced their services and decisions to private vendors. The technology

includes anything from automated facial recognition systems, applicant screening systems, and fraud detention systems.

- The Michigan Integrated Data System (MiDAS) was launched in October 2013 as a replacement for a 30 year old system that the Michigan Unemployment Insurance Agency used to administer unemployment benefits. Simultaneously with the system's rollout, the UI agency reduced its staff by one third. The system nearly immediately has significant issues and by the time the system was shut down in September 2015 under community and legal pressure, MiDAS had incorrectly accused more than 34,000 people of unemployment fraud, cutting critical receiving benefits, due to a poorly designed algorithm within a poorly designed system. Michigan state legislators have promised to seek at least $30 million in compensation for those falsely accused. Advocates are pushing for close to $100 million given the trauma faced by people who saw their credits and reputations ruined, filed for bankruptcy, had their houses foreclosed, or were made homeless.

- Use of the Allegheny Family Screening Tool in Pennsylvania which was meant to guide responses to calls to family service agency about potential child neglect. This is a predictive model that the ACLU analyzed for how its design functions as policy and its impact on families.

- In 2022, there was a Carnegie Mellon study conducted on the tool and a resulting AP investigation; included in that "CMU researchers found that from August 2016 to May 2018, the tool calculated scores that suggested 32.5% of Black children reported as being neglected should be subject to a "mandatory" investigation, compared with 20.8% of white children." As a result of that investigation, it was reported last January that the DOJ is investigating the use of this algorithm.

- Another example is COMPAS which is a risk assessment model used to predict recidivism, which was found to be racially biased.

- In 2022, an LAO report found that California's Employment Development Department (EDD), procured an automated review software that flagged 1.1 million claims as fraudulent, meaning worker's stopped receive much needed benefit. Ultimately, more than half of those claims (600,000) were confirmed as legitimate.

As such, a Committee amendment has been introduced to prohibit the procurement of ADTs by the state until the procurement standard required by this bill has been developed and implemented. This moratorium does not prohibit the continued use of existing ADTs – only the procurement of new ADTs.

5) **Full text of bill as proposed to be amended:**

SECTION 1. Section 12100.1 is added to the Public Contract Code, to read:

12100.1. (a) For purposes of this section, the following definitions apply:

(1) "Artificial intelligence" or "AI" means an engineered or machine-based system that~~,~~ **varies in its level of autonomy and that can,** for explicit or implicit objectives, infer~~s,~~

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~~.

~~(2) (A) "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.~~

~~(B) "Automated decision system" does not mean a spam email filter, firewall, antivirus software, identity and access management tool, calculator, database, dataset, or other compilation of data.~~

(2) "Automated decision tool" or "ADT" means an artificial intelligence system or service that makes a consequential decision, or is a substantial factor in making consequential decisions.

(3) "Consequential decision" means a decision or judgment that has a legal, material, or similarly significant effect on an individual's life relating to access to government benefits or services, assignments of penalties by government, or the impact of, access to, or the cost, terms, or availability of, any of the following:

(A) Employment with respect to all of the following:

(i) Pay or promotion.

(ii) Hiring or termination.

(iii) Automated task allocation that limits, segregates, or classifies employees for the purpose of assigning or determining material terms or conditions of employment.

(B) Education and vocational training as it relates to all of the following:

(i) Assessment or placement.

(ii) Detecting student cheating or plagiarism.

(iii) Accreditation.

(iv) Certification.

(v) Admissions or enrollment.

(vi) Discipline.

(vii) Evaluation.

(viii) Financial aid or scholarships.

(C) Housing or lodging, including rental or short-term housing or lodging.

(D) All of the following essential utilities:

(i) Electricity.

(ii) Heat.

(iii) Water.

(iv) Internet or telecommunications access.

(v) Transportation.

(E) Family planning.

(F) Adoption services, reproductive services, or assessments related to child protective services.

(G) Health care or health insurance, including mental health care, dental, or vision.

(H) Financial services, including a financial service provided by a mortgage company, mortgage broker, or creditor.

(I) All of the following aspects of the criminal justice system:

(i) Risk assessments for pretrial hearings.

(ii) Sentencing.

(iii) Parole.

(J) Legal services.

(K) Private arbitration.

(L) Mediation.

(M) Voting.

(3) "Department" means the Department of Technology.

(4) "Privacy Agency" means the California Privacy Protection Agency.

~~(4) "High-risk automated decision system" or "high-risk ADS" means an automated decision system 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, free speech, housing or accommodations, education, employment, credit, health care, child welfare, immigration, and criminal justice.~~

(5) "Substantial factor" means an element of a decisionmaking process that is capable of altering the outcome of the process.

(b) **Following the adoption of regulations by the Privacy Agency pursuant to paragraph (16) of subdivision (a) of Section 1798.185 of the Civil Code, or following the enactment of similarly comprehensive statewide legislation that establishes a regulatory framework governing the development and deployment of ADTs, whichever occurs sooner,** ~~T~~**t**he department shall develop and adopt regulations to create an ~~AI risk management~~ **ADT procurement** standard.

(1) To develop regulations related to the ~~AI risk management~~ **ADT procurement** standard, the department ~~may apply~~ **shall consider** principles and industry standards addressed in relevant publications, including, but not limited to, ~~any~~ **all** of the following:

(A) The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People, published by the White House Office of Science and Technology Policy in October 2022.

(B) The Artificial Intelligence Risk Management Framework (AI RMF 1.0), released by the National Institute of Standards and Technology (NIST) in January 2023.

(C) 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 in 2024.

(D) The Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence Memorandum, published by the Executive Office of the President, Office of Management and Budget, dated March 28, 2024.

(2) The ~~AI risk management~~**ADT procurement** standard shall include all of the following:

(A) A detailed risk assessment procedure ~~for procuring ADS~~ that analyzes all of the following:

(i) Organizational and supply chain governance associated with the ~~ADS~~**ADT**.

(ii) The purpose and use of the ~~ADS~~**ADT**.

(iii) Any known potential misuses or abuses of the ~~ADS~~**ADT**.

(iv) An assessment of the legality, traceability, and provenance of the data the ~~ADS~~**ADT** uses and the legality of the output of the ~~ADS~~**ADT**.

(v) The robustness, accuracy, and reliability of the ~~ADS~~**ADT**.

(vi) The interpretability and explainability of the ~~ADS~~**ADT**.

(B) Methods for appropriate risk controls between the state agency and ~~ADS~~**ADT** vendor, including, but not limited to, reducing the risk through various mitigation strategies, eliminating the risk, or sharing the risk.

(C) Adverse incident monitoring procedures.

(D) Identification and classification of prohibited use cases and applications of ~~ADS~~**ADT** that the state shall not procure.

(E) A detailed equity assessment that analyzes, at a minimum, all of the following:

(i) The individuals and communities that will interact with the ~~high-risk~~ ~~ADS~~**ADT**.

(ii) How the information or decisions generated by the ~~ADS~~**ADT** will impact an individual's rights, freedoms, economic status, health, health care, or well-being.

(iii) Any issues that may arise if the ~~ADS~~**ADT** is inaccurate.

(iv) How users ~~of~~ **with** diverse abilities will interact with the user interface of the ~~ADS~~**ADT** and whether the ~~ADS~~**ADT** integrates and interacts with commonly used assistive technologies.

(F) An assessment that analyzes the level of human oversight associated with the use of ~~ADS~~**ADT**.

(G) Adherence to data minimization standards, including that an ~~AI or ADS~~**ADT** vendor shall only use information provided by or obtained from an agency to provide the specific service authorized by the agency. Further, the data collected may not be used for training of proprietary vendor or third-party systems.

(3) ~~To~~ **In** develop**ing** the ~~AI risk management~~**ADT procurement** standard, the department shall ~~comply with~~**do** all of the following:

(A) Collaborate with organizations that represent state and local government employees and industry experts, including, but not limited to, public trust and safety experts, community-based organizations, civil society groups, academic researchers, and research institutions focused on responsible ~~AI~~ **ADT** procurement, design, and deployment.

(B) Consult with the California Privacy Protection Agency.

(C) Solicit public comment on the ~~risk management~~**ADT procurement** standard.

(4) **(A) Subject to subparagraph (B),** ~~The~~ **the** department shall adopt regulations pursuant to this subdivision in accordance with the provisions of Chapter 3.5 (commencing with Section 11340) of Part 1 of Division 3 of Title 2 of the Government Code.

(B) Any regulations adopted by the department pursuant to subparagraph (A) shall be consistent with both of the following:

(i) Any regulations adopted by the Privacy Agency pursuant to paragraph (16) of subdivision (a) of Section 1798.185 of the Civil Code.

(ii) Any statewide legislation that establishes a regulatory framework governing the development and deployment of ADTs.

(5) The department shall annually review and update both of the following:

(A) The ADT procurement standard.

(B) Any regulations adopted pursuant to this subdivision.

(c) A state agency shall not procure an ADT, enter into a contract for an ADT, or enter into a contract for any service that utilizes an ADT, until the department has adopted regulations pursuant to subdivision (b).

(**d**~~c~~) ~~Commencing six months after the date on which the regulations described in subdivision (b) are approved and final, A~~**a** state agency ~~shall not~~**may** enter into a contract for an ~~automated decision system~~**ADT**~~,~~ or ~~any~~ **a** service that utilizes an ~~automated decision system~~**ADT**~~,~~ ~~unless~~ **after the department has adopted regulations pursuant to subdivision (b) only if** the contract includes a clause that does all of the following:

(1) Provides a completed risk assessment of the relevant ~~ADS~~**ADT** that analyzes the items included in subparagraph (A) of paragraph (2) of subdivision (b).

(2) Requires the state agency or the ~~ADS~~**ADT** vendor, or both, to adhere to appropriate ~~risk controls~~**procurement standards**.

(3) Provides procedures for adverse incident monitoring.

(4) Requires authorization from the state agency before deployment of ~~ADS~~**ADT** upgrades and enhancements.

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

(6) Provides a termination right in the event of a significant breach of responsibility or violation by the vendor.

6) **Related legislation.** AB 2930 (Bauer-Kahan, 2024) would require developers and deployers of automated decision tools to perform impact assessments before using the tools, as well as disclose the tools use to affected consumers. This bill is currently pending in Senate Judiciary Committee.

AB 302 (Ward, Chap. 800, Stats. 2023) required CDT to conduct an inventory of all high-risk automated decision systems being used in state agencies. The bill was chaptered on October 8th, 2023.

SB 313 (Dodd, 2023) would have established an Office of Artificial Intelligence within CDT, with "the powers and authorities necessary to guide the design, use, or deployment of automated systems by a state agency to ensure that all AI systems are designed and deployed in a manner that is consistent with state and federal laws and regulations regarding privacy and civil liberties and that minimizes bias and promotes equitable outcomes for all Californians." The bill was held in Senate Appropriations.

AB 2269 (Chau, 2020) would have established the Automated Decision Systems (ADS) Accountability Act of 2020. The Act would have required businesses in California that provide a person, as defined, with a program or device that uses an automated decision system (ADS) to take affirmative steps to ensure that there are processes in place to continually test for biases during the development and usage of the ADS, conduct an ADS impact assessment on its program or device to determine whether the ADS has a disproportionate adverse impact on a protected class, as specified, examine if the ADS in question serves reasonable objectives and furthers a legitimate interest, and compare the ADS to alternatives or reasonable modifications that may be taken to limit adverse consequences on protected classes. The bill was held in Assembly Privacy and Consumer Protection.

ARGUMENTS IN SUPPORT:

Electronic Frontier Foundation writes:

> 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 standard that includes 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.
>
> - Ensure people have access to robust notices and opportunities to appeal or opt out of high-risk AI system outcomes or ADS decisions.

The City of San Jose writes:

> The City of San José has established itself as a leader in AI, having convened the nationwide GovAI Coalition, which is comprised of more than 600 officials at 250 local, county, and state governments to promote the responsible use of AI. While the City has been an early supporter and thought partner on the benefits of AI, we also recognize the potential risks and the need for legislation and regulation to provide guardrails throughout its development. This legislation strikes the appropriate balance of assessing and minimizing risks while not hindering growth. AI has the potential to provide incredible societal benefits if harnessed appropriately, and California has an immense opportunity to lead this effort in partnership with technology companies and Silicon Valley leaders.

Secure Justice writes:

As the fourth largest economy in the world, California must ensure that innovation does not come at the expense of our constituents. Given the state's purchasing power through our procurement of technology and services, California can drive the innovation of AI technology and services by requiring that any technology and services contracted by the state adhere to standards that protect Californians. While the Governor's Executive Order to Prepare California for the Progress of Artificial Intelligence required the issuance of general guidelines for public sector procurement, California must take further steps to safeguard the critical services we provide.

REGISTERED SUPPORT / OPPOSITION:

Support

AI Procurement Lab
American Federation of Musicians, Local 7
City of Long Beach
City of San Jose
Electronic Frontier Foundation
Greenlining Institute; the
Oakland Privacy
San Diego Regional Chamber of Commerce
Secure Justice
TechEquity Collaborative

Opposition

None on file.

Analysis Prepared by: Slater Sharp / P. & C.P. / (916) 319-2200