

Date of Hearing: April 11, 2023

ASSEMBLY COMMITTEE ON PRIVACY AND CONSUMER PROTECTION

Jesse Gabriel, Chair

AB 331 (Bauer-Kahan) – As Amended March 30, 2023

As Proposed to be Amended

SUBJECT: Automated decision tools

SYNOPSIS

This author-sponsored measure is a significant, meaningful attempt to set state policy with respect to private and public use of automated decision tools (ADTs): systems and services that employ artificial intelligence, and make, or are a controlling factor in making, consequential decisions. Consequential decisions, in turn, are those that materially impact the necessities of a person's life, such as housing, education, employment, and health care (including reproductive health care), as well as the person's involvement with critical public functions such as voting and the criminal justice system.

The bill would establish requirements for informed, conscious deployment of ADTs, based on the White House's recently-released Blueprint for an AI Bill of Rights. Key aspects of the bill include the following:

- *Developers and deployers must prepare impact statements that address intended uses and potential effects of ADTs, as well as evaluation of their validity.*
- *Developers and deployers must also prepare governance programs to manage the risk of algorithmic discrimination from use of ADTs.*
- *Deployers must notify any natural person who will be the subject of a consequential decision if an ADT will be used to make, or will be a controlling factor in making, the decision.*
- *A deployer must accommodate, where technically feasible, a natural person's request to not be subject to an ADT and to instead be subject to an alternative selection process or accommodation, if a consequential decision will be made solely based on the ADT's output.*

Broadly speaking, the bill only prescribes one policy outcome: it forbids algorithmic discrimination against individuals on the basis of characteristics protected under state civil rights law, such as race, age, national origin, disability, limited English proficiency, and sex (including pregnancy, childbirth, and related conditions; gender identity; intersex status; and sexual orientation). Much of the bill is instead devoted to ensuring that ADT developers and deployers thoroughly consider the design, testing, and functioning of ADTs before releasing them into the world. This appears to be sound policy, as developers and deployers are the entities that possess the most knowledge about the systems they work with.

This bill is supported by the California-Hawaii State Conference of the NAACP, the Israeli-American Civic Action Network, Oakland Privacy, and the Santa Monica Democratic Club. It is

opposed by a coalition of nineteen business trade associations, led by the California Chamber of Commerce.

If passed by this Committee, this bill will next be heard by the Assembly Judiciary Committee.

SUMMARY: Establishes a statutory framework to further the safe and informed use of automated decision tools in California. Specifically, **this bill:**

1) Defines the following terms:

- a) “Automated decision tool” (ADT) means a system or service that (i) uses artificial intelligence and (ii) has been specifically developed and marketed to, or specifically modified to, (iii) make, or be a controlling factor in making, consequential decisions.
- b) “Artificial intelligence” means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing a real or virtual environment.
- c) “Consequential decision” means a decision or judgment that has a legal, material, or similarly significant effect on an individual’s life relating to the impact of, access to, or the cost, terms, or availability of, any of the following:
 - i) Employment, worker management, or self-employment, including, but not limited to, pay or promotion; hiring or termination; and automated task allocation.
 - ii) Education and vocational training, including, but not limited to, assessment (including, but not limited to, detecting student cheating or plagiarism); accreditation; certification; admissions; and financial aid or scholarships.
 - iii) Housing or lodging, including rental or short-term housing or lodging.
 - iv) Essential utilities, including electricity, heat, water, internet or telecommunications access, or transportation.
 - v) Family planning, including adoption services or reproductive services, as well as assessments related to child protective services.
 - vi) Health care or health insurance, including mental health care, dental, or vision.
 - vii) Financial services, including a financial service provided by a mortgage company, mortgage broker, or creditor.
 - viii) The criminal justice system, including, but not limited to, risk assessments for pretrial hearings; sentencing; and parole.
 - ix) Legal services, including private arbitration or mediation.
 - x) Voting.
 - xi) Access to benefits or services or assignment of penalties.

- d) “Algorithmic discrimination” means the condition in which an ADT contributes to unjustified differential treatment or impacts disfavoring people based on their actual or perceived race, color, ethnicity, sex (including pregnancy, childbirth, and related conditions; gender identity; intersex status; and sexual orientation), religion, age, national origin, limited English proficiency, disability, veteran status, genetic information, reproductive health, or any other classification protected by state law.
 - e) “Developer” means a person, partnership, state or local government agency, or corporation that designs, codes, or produces an ADT, or substantially modifies an artificial intelligence system or service for the intended purpose of making, or being a controlling factor in making, consequential decisions, whether for the developer’s own use or for use by a third party.
 - f) “Deployer” means a person, partnership, state or local government agency, or corporation that uses an ADT to make a consequential decision.
 - g) “Significant update” means a new version, new release, or other update to an ADT that includes changes to its use case, key functionality, or expected outcomes.
 - h) “Impact assessment” means a documented risk-based evaluation of an ADT that meets the criteria set forth under 2) and 3) below.
- 2) Requires, beginning on or before January 1, 2025, and annually thereafter, a developer to perform an impact assessment for any ADT the developer designs, codes, or produces. The impact assessment must include all of the following:
- a) A statement of the ADT’s purpose and its intended benefits, uses, and deployment contexts.
 - b) A description of the ADT’s outputs and how they are used to either make, or be a controlling factor in making, a consequential decision.
 - c) A summary of the type of data collected from natural persons and processed by the ADT when it is used to make, or be a controlling factor in making, a consequential decision.
 - d) An analysis of potential adverse impacts on the basis of sex, race, color, ethnicity, religion, age, national origin, limited English proficiency, disability, veteran status, or genetic information from use of the ADT.
 - e) A description of the measures taken by the developer to mitigate the risk known to the developer of algorithmic discrimination arising from the use of the ADT.
 - f) A description of how the ADT can be used by a natural person, or monitored when it is used, to make, or be a controlling factor in making, a consequential decision.
- 3) Requires, beginning on or before January 1, 2025, and annually thereafter, a deployer to perform an impact assessment for any ADT the deployer uses. The impact assessment must include all of the following:
- a) A statement of the ADT’s purpose and its intended benefits, uses, and deployment contexts.

- b) A description of the ADT's outputs and how they are used to make, or be a controlling factor in making, a consequential decision.
 - c) A summary of the type of data collected from natural persons and processed by the ADT when it is used to make, or be a controlling factor in making, a consequential decision.
 - d) A statement of the extent to which the deployer's use of the ADT is consistent with or varies from the statement required of the developer under 9) below.
 - e) An analysis of potential adverse impacts on the basis of sex, race, color, ethnicity, religion, age, national origin, limited English proficiency, disability, veteran status, or genetic information from the deployer's use of the ADT.
 - f) A description of the safeguards implemented, or that will be implemented, by the deployer to address any reasonably foreseeable risks of algorithmic discrimination known to the deployer at the time of the impact assessment arising from the use of the ADT.
 - g) A description of how the ADT will be used by a natural person, or monitored when it is used, to make, or be a controlling factor in making, a consequential decision.
 - h) A description of how the ADT has been or will be evaluated for validity or relevance.
- 4) Requires developers and deployers to, in addition to the requirements of 2) and 3), perform as soon as possible an impact assessment with respect to any significant update, as defined.
- 5) Requires a deployer to notify any natural person that is the subject of a consequential decision that an ADT is being used to make, or be a controlling factor in making, the consequential decision. The notification must be made at or before the time the ADT is used.
- 6) Specifies that notification under 5) must include:
- a) A statement of the ADT's purpose.
 - b) The deployer's contact information.
 - c) A plain language description of the ADT, including a description of any human components and how any automated component is used to inform a consequential decision.
- 7) Requires the deployer, in cases in which a consequential decision is made solely based on the output of an ADT, to accommodate a natural person's request to not be subject to the ADT and to instead be subject to an alternative selection process or accommodation—but only if this alternative is technically feasible.
- 8) Permits a deployer, after a request under 7) is made, to reasonably request, collect, and process information from a natural person in order to identify the person and the associated consequential decision. Absent this information, the deployer is not obligated to provide an alternative selection process or accommodation.
- 9) Requires a developer to provide a deployer with a statement regarding the intended uses of an ADT, as well as documentation regarding the following:

- a) The known limitations of the ADT, including any reasonably foreseeable risks of algorithmic discrimination arising from the ADT's intended use.
 - b) A description of the type of data used to program or train the ADT.
 - c) A description of how the ADT was evaluated for validity and explainability before its sale or licensing.
- 10) Clarifies that 9) does not require the disclosure of trade secrets, as defined under the Civil Code.
- 11) Requires deployers and developers to establish, document, implement, and maintain a governance program that contains reasonable administrative and technical safeguards to map, measure, manage, and govern the reasonably foreseeable risks of algorithmic discrimination associated with the use or intended use of an ADT.
- 12) Explains that the governance program required under 11) must be designed to do all of the following:
- a) Designate at least one employee to be responsible for overseeing and maintaining the governance program and compliance with this bill. This employee is authorized to assert to their employer a good faith belief that the design, production, or use of an ADT fails to comply with this bill's requirements; if this assertion is made, the employer must conduct a prompt and complete assessment of any compliance issues that the employee raised.
 - b) Identify and implement safeguards to address reasonably foreseeable risks of algorithmic discrimination resulting from the use or intended use of an ADT.
 - c) If established by a deployer, provide for the performance of impact assessments under 3).
 - d) If established by a developer, provide for compliance with 2) and 9).
 - e) Conduct an annual and comprehensive review of policies, practices, and procedures to ensure compliance with this bill.
 - f) Maintain the results of impact assessments for two years after completion.
 - g) Evaluate and make reasonable adjustments to administrative and technical safeguards in light of material changes in technology, the risks associated with the ADT, the state of technical standards, and changes in business arrangements or operations of the deployer or developer.
- 13) Exempts from the requirements of 2) and 12) a deployer with fewer than 25 employees, unless, as of the previous calendar year, the deployer deployed an ADT that impacted more than 999 people per year.
- 14) Requires developers and deployers to make publicly available, in a readily accessible manner, a clear policy that summarizes both of the following:
- a) The types of ADTs that the developer or deployer currently uses or makes available to others.

- b) How the developer or deployer manages the reasonably foreseeable risks of algorithmic discrimination that may arise from the use of the ADTs that it currently uses or makes available to others.
- 15) Prohibits a deployer from using an ADT that contributes to algorithmic discrimination.
- 16) Requires a deployer or developer that is required by this bill to complete an impact assessment to provide the assessment to the Civil Rights Department within 60 days after it is completed.
- 17) Specifies that the administrative fine for non-compliance with the requirement under 16) is \$10,000 per violation. Further specifies that each day that an ADT is used is a distinct violation if an impact assessment has not yet been submitted.
- 18) Permits the Civil Rights Department to share impact assessments with other state entities as appropriate.
- 19) Provides, beginning on January 1, 2026, a private right of action to a person against a deployer or developer for a violation of a provision of this bill.
- 20) Provides the following remedies to a prevailing plaintiff: compensatory damages, injunctive relief, declaratory relief, reasonable attorney's fees, and litigation costs.
- 21) Requires a prospective plaintiff to provide 45 days' written notice to a deployer or developer of an alleged violation before commencing an action for injunctive relief, except that this requirement shall not apply if the violation is substantially the same as one previously the subject of a notice.
- 22) Provides that if a developer or deployer receives a notice of prospective injunctive relief under 21), the developer or deployer has 45 days to provide the prospective plaintiff with an express written statement, made under penalty of perjury, that the violation has been cured and that no further violations shall occur. If the developer or deployer does so, the prospective plaintiff will no longer have a claim for injunctive relief.
- 23) Preempts any local law, regulation, rule, requirement, or standard related to performing an impact assessment, governance program, or other equivalent of an ADT.

EXISTING LAW:

- 1) Establishes the Civil Rights Department, and sets forth its statutory functions, duties, and powers. (Gov. Code § 12930.)
- 2) Enacts the Fair Employment and Housing Act. (Gov. Code §§ 12900 *et seq.*)
- 3) Enacts the Unruh Civil Rights Act. (Civ. Code § 51.)
- 4) Defines "trade secret" under the Uniform Trade Secrets Act as information, including a formula, pattern, compilation, program, device, method, technique, or process, that both:

- a) Derives independent economic value, actual or potential, from not being generally known to the public or to other persons who can obtain economic value from its disclosure or use; and
- b) Is the subject of efforts that are reasonable under the circumstances to maintain its secrecy. (Civ. Code § 3426.1(d).)

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

COMMENTS:

1) **Background.** The term “automated decision tool,” or “ADT,” refers to an ever-growing array of computer systems that are being relied upon to make, or to be a controlling factor in making, consequential decisions. 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. After all, computers can generally draw upon far more digitally-encoded data, and process this data far more quickly, than any human being. This can yield astonishing insights. 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. (Santariano & Metz, *Using A.I. to Detect Breast Cancer That Doctors Miss* (Mar. 5, 2023) N.Y. Times, available at

<https://www.nytimes.com/2023/03/05/technology/artificial-intelligence-breast-cancer-detection.html>.)

But we must be careful 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.” (Broussard, *Artificial Unintelligence* (MIT Press 2019) p. 8.) It remains the case that many aspects of the real world are not reducible to data, and that there are many complex situations presenting competing demands which human beings remain best equipped to assess.

Recent situations when ADTs substituted for human decision-making, to the detriment of human beings, include:

- A privately-built, error-prone software system operated with minimal employee oversight by the State of Michigan wrongly accused 40,000 people of unemployment fraud between 2013 and 2015, many of whom were forced to pay heavy fines, declare bankruptcy, or have their homes foreclosed upon. Upon appeal, less than 8% of those fraud charges were validated. (de la Garza, *States' Automated Systems Are Trapping Citizens in Bureaucratic Nightmares with Their Lives on the Line* (May 20, 2020) Time Magazine, available at <https://time.com/5840609/algorithm-unemployment/>.)
- A 54-year-old Black man was arrested in Maryland and spent nine days in jail after being misidentified by facial recognition software. (Johnson, *Face Recognition Software Led to his Arrest. It Was Dead Wrong*. (Feb 28, 2023) WIRED, available at <https://www.wired.com/story/face-recognition-software-led-to-his-arrest-it-was-dead-wrong/>.)
- A self-driving Uber car that killed a 49-year-old jaywalker in Arizona was programmed not to classify objects as pedestrians unless those objects were detected near crosswalks. (McCausland, *Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk* (Nov. 9, 2019) NBC News, available at <https://www.nbcnews.com/tech/tech-news/self-driving-uber-car-hit-killed-woman-did-not-recognize-n1079281>.)

As Broussard writes, “Because social decisions are about more than just calculations, problems will always ensue if we use data alone to make decisions that involve social and value judgments.” (*Artificial Unintelligence, supra*, p. 119.) Human judgment can mean the difference between being housed or unhoused, being free or imprisoned, and in extreme cases, between life and death.

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. (McKinsey & Co.. *The state of AI in 2022—and a half decade in review* (Dec. 6, 2022), available at <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review>.)

The premises of this bill are, first, that Californians ought to be much more informed about the ADTs that are being used to make consequential decisions that impact their lives, and second, that those who develop and deploy ADTs ought to be much more conscious and deliberate about what those impacts might be. In short, that we ought to get a handle on ADTs while we still can, before they become yet another phenomenon—like greenhouse gases and guns and urban sprawl—that we could have dealt with deliberately and intelligently in the past, but instead allowed to expand without any constraints, and now find ourselves struggling to contain.

2) Author’s statement.

AB 331 protects individuals from algorithmic discrimination by requiring developers and users to assess automated decision tools (ADTs) that make consequential decisions and mitigate any discovered biases. The use of ADTs has become very prominent within different sectors such as housing, employment, and even in criminal justice sentencing and probation decisions. The algorithms used within ADTs can be prone to unrepresentative datasets, faulty classifications, and flawed design, which can lead to biased, discriminatory, or unfair

outcomes. These tools can exacerbate the harms they are intended to address and ultimately hurt the people they are supposed to help. As the use of decision making via algorithm becomes more prevalent in our daily lives, it is crucial that we take the necessary steps to ensure that they are used ethically and responsibly.

3) **Key terms in this bill.** There are a few key terms that are first necessary to understand in order to understand what this bill would do.

The first term, of course, is “automated decision tool” (ADT), defined as a system or service that (i) uses artificial intelligence and (ii) has been specifically developed and marketed to, or specifically modified to, (iii) make, or be a controlling factor in making, consequential decisions.

The second term is “consequential decision,” which has a lengthy definition in the bill, as set forth in 1) c) in the **SUMMARY** above. Briefly, a consequential decision can be understood as a decision that materially impacts the necessities of a person’s life, such as education, housing, employment, or health care (including reproductive health care), as well as the person’s involvement with critical public functions, such as voting and the criminal justice system.

It is important to distinguish ADTs, as defined under this bill, from the larger category of automated decision tools that have replaced jobs that humans used to do, such as elevator or telephone switchboard operator. These tools use artificial intelligence and have been specifically developed and marketed to make decisions—but not consequential decisions.

An ADT can be a system or service that uses artificial intelligence and has been specifically developed and marketed **to make** consequential decisions. For example, an automated system that schedules fast food workers’ shifts and directs them as to when to appear and leave would likely qualify as an ADT under this bill, since it makes decisions as to automated task allocation.

An ADT can also be a system or service that uses artificial intelligence and has been **specifically modified** to make consequential decisions. The author provides the example of an ADT that is meant to evaluate whether a person convicted of a crime is likely to reoffend then being re-deployed to determine whether the person should receive access to reentry services.

An ADT can also be a system or service that uses artificial intelligence and has been specifically developed and marketed to be a **controlling factor** in making consequential decisions. The term “controlling factor” is not defined in the bill, which gives it the necessary flexibility to encompass a wide variety of situations. But the terms suggests that, for an ADT to be a controlling factor in a consequential decision, the output provided by the ADT must ultimately determine the decision; if the ADT had provided output that was significantly different, the consequential decision would have also turned out differently. E.g., if a loan applicant’s credit score is a factor that could outweigh all other factors (age, marital status, income) in determining whether to give the applicant a home loan, then the system that produces the credit score would be an ADT for purposes of this bill.

Two other terms bear mention: “developer” and “deployer.” As the terms suggest, a developer designs, codes, or produces ADTs, while a deployer uses ADTs. Developers and deployers may be individuals, private entities (including corporations), or state and local agencies.

4) **What this bill would do.** The provisions of this bill are based upon the principles outlined in the *Blueprint for an AI Bill of Rights*, issued by the White House Office of Science and

Technology Policy on October 4, 2022. The *Blueprint* is available to read at <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>.

The *Blueprint* identifies five principles that, as it explains, should “guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence.” These principles are Safe and Effective Systems; Algorithmic Discrimination Protections; Data Privacy; Notice and Explanation; and Human Alternatives, Consideration, and Fallback. Each of these principles, other than Data Privacy (which has been the subject of extensive prior legislation), is implemented in this measure.

- Principle: Safe and Effective Systems

The *Blueprint* summarizes this principle as meaning that “you should be protected from unsafe or ineffective systems.”

The bill would implement this principle by requiring developers and deployers to complete impact assessments for any ADT that they develop or use, respectively. Specific criteria for these impact assessments are set forth in the **SUMMARY** above, in paragraphs 2) (for developers) and 3) (for deployers). As can be seen, the impact assessments require thoughtful consideration of what ADTs are intended to do, what their effects might be, and how they are validated. Once completed, these impact assessments must be shared with the Civil Rights Department, which can then share them with other state agencies.

Both types of impact assessments would have to be completed by January 1, 2025, and annually thereafter. Bill opponents argue that “the annual nature of these assessments will be burdensome, particularly for smaller businesses.” However, the author has anticipated this objection, and exempted deployers with fewer than 25 employees from this requirement, unless such a deployer previously used an ADT that impacted more than 999 people per year.

Developers and deployers would have to complete new impact assessments any time a significant update, as defined, is made to an ADT.

Bill opponents argue that, in preparing impact statements, deployers may lack the “knowledge and expertise of the underlying technology created by [a developer] to understand all the varying uses aside from their own intended use.” The bill already requires developers to share documentation about many aspects of the ADT with deployers that should help the latter prepare impact statements. The author may wish to consider whether developers should share additional documentation beyond that already specified in the bill.

- Principle: Algorithmic Discrimination Protections

The *Blueprint* summarizes this principle as meaning that “you should not face discrimination by algorithms[,] and systems should be used and designed in an equitable way.”

The bill would implement this principle in at least three ways. First, developers and deployers would be required to prepare governance programs containing “reasonable administrative and technical safeguards to map, measure, manage, and govern the reasonably foreseeable risks of algorithmic discrimination associated with the use or intended use” of an ADT. Specific criteria for this governance program are set forth in the **SUMMARY** above, in paragraph 12).

Second, at least one employee would have to be designated to oversee and maintain the governance program and compliance with this bill's requirements. That employee would have "the authority to assert to [their] employer a good faith belief that the design, production, or use of an automated decision tool fails to comply" with one or more requirements of the bill; the employer would then be required to conduct a prompt and complete assessment of the compliance issues raised.

Third, the bill would flatly prohibit a deployer from using an ADT that contributes to algorithmic discrimination on the basis of any category protected under California's civil rights laws.

- Principle: Notice and Explanation

The *Blueprint* summarizes this principle as meaning that "you should know that an automated system is being used and understand how and why it contributes to outcomes that impact you."

The bill would implement this principle by requiring a deployer to notify any natural person that is the subject of a consequential decision that an ADT will be used to make, or will be a controlling factor in making, that decision. Requirements for the content of the notification are set forth under paragraph 6) in the **SUMMARY** above. The notification must be made at or before the time the ADT is used.

Bill opponents object that this notice requirement cannot be put into practice in some situations, such as if an ADT is being used to guide decision-making in an emergency room. The author may wish to clarify this requirement to address emergency situations where it is impossible to provide advance notice.

- Principle: Human Alternatives, Consideration, and Fallback

The *Blueprint* summarizes this principle as meaning that "[y]ou should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter."

The bill would implement this principle by requiring a deployer to accommodate a natural person's request to not be subject to an ADT and to instead be subject to an alternative selection process or accommodation. The bill limits this requirement in two significant ways, however. First, the alternative is only available in cases in which a consequential decision is to be made **solely** based on the output of an ADT; i.e., one over which there is no human discretion. Second, the alternative must be technically feasible.

5) **Other opposition arguments.** Several other arguments raised by bill opponents are as follows.

First, the opposition seeks greater clarity as to what constitutes a violation giving rise to liability under the bill. They write:

As currently drafted, it is possible that a violation could constitute not only a failure to complete an impact assessment altogether, but could also be interpreted as any single deficiency within that impact assessment. Even then, it is unknown whether the number of violations is based on that single error, or by any single error multiplied by the number of

individuals who were potentially impacted by the use of a particular type of ADT for which an assessment was required or who received an inadequate notice.

In response, the author may wish to further define what constitutes a violation.

Second, the opposition points out that at least two California agencies—the Civil Rights Department and the California Privacy Protection Agency—are currently engaged in rulemaking proceedings with respect to ADTs. The opposition argues:

[I]t is difficult to foresee how such laws and regulations will layer on top of one another and whether there will be conflicting public policy around the use of such tools and technologies. [...] Understandably, our members are alarmed by the likelihood of conflict and confusion at the conclusion of these efforts that are being run in parallel to each other, without any coordination or consideration of the other efforts underway.

The author may wish to clarify the boundaries between this bill and any regulations forthcoming from these state agencies, so as to better facilitate compliance.

Finally, opponents believe impact assessments should not be subject to California Public Records Act requests, and that greater restrictions need to be placed on sharing of impact assessments between state agencies. Yet imposing these restrictions may curtail informed future responses to ADTs by both the public and by state government. The author may wish to consider the appropriate balance regarding availability or confidentiality of impact assessments submitted to the Civil Rights Department.

6) **Author’s amendments—technical and clarifying changes.** Author’s amendments to the bill, reflected in the analysis above, make the following technical and clarifying changes:

22756.1 (a) [...]

(5) An analysis of potential adverse impacts on the basis of sex, race, color, ethnicity, religion, age, national origin, limited English proficiency, disability, veteran status, or genetic information *from the employer’s use of the automated decision tool.* [...]

(b) [...]

(4) An analysis of a potential adverse impact on the basis of sex, race, color, ethnicity, religion, age, national origin, limited English proficiency, disability, veteran status, or genetic information *from the use of the automated decision tool.*

The two preceding sets of amendments clarify that is the use of the ADT that may give rise to potential adverse impacts.

22756.1 [...]

(d) ~~(1) Subject to paragraph (2),~~ This section does not apply to a employer with fewer than 25 employees *unless, as of the end of the prior calendar year,*

~~(2) This subdivision does not apply if the employer deployed~~ *eds* an automated decision tool that impacted *eds* more than 999 people per year ~~as of the end of the prior calendar year.~~

[...]

22756.4 [...]

(c) ~~(1) Subject to paragraph (2),~~ This section does not apply to a deployer with fewer than 25 employees *unless, as of the end of the prior calendar year,*

~~(2) This subdivision does not apply if the deployer deployed~~ *eds* an automated decision tool that impacted *eds* more than 999 people per year ~~as of the end of the prior calendar year.~~

The two preceding sets of amendments are intended to simplify the wording of these paragraphs.

22756.4. [...]

(b) [...]

(4) If established by a developer, provide for compliance with Sections 22756.4 2 and 22756.3.

This amendment corrects an erroneous cross-reference, as Section 22756.1 applies to deployers, not developers.

22756.5. A deployer or developer shall make publicly available, in a readily accessible manner, a clear ~~artificial intelligence~~ policy that provides a summary of both of the following: [...]

Use of the phrase “artificial intelligence policy” may have led some to believe that this section provided a statutory definition of an artificial intelligence policy, which was not the author’s intent.

7) **Related legislation.** AB 302 (Ward) would require the California Department of Technology (CDT) to conduct an inventory of all high-risk automated decision systems being used in state agencies. Status: Assembly Appropriations Committee.

SB 313 (Dodd) would establish 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.” Status: Senate Governmental Organization Committee.

SB 721 (Becker) would create the California Interagency AI Working Group, which would deliver a report to the Legislature regarding artificial intelligence, include a recommendation for a statutory definition of “artificial intelligence” for use in legislation. Status: Senate Governmental Organization Committee.

AB 13 (Chau, 2021) would have established the Automated Decision Systems (ADS) Accountability Act of 2021. The Act would have required state agencies seeking to procure ADS for high-risk applications to consider, among other things, steps taken by a prospective contractor to identify and mitigate potential disparate impacts that could result from use of that ADS; required a prospective contractor for an ADS for a high-risk application to submit an ADS

impact assessment containing specified information about the ADS; and required the contracting agency to submit to CDT a high-risk ADS accountability report containing specified information regarding their proposed use of the ADS. Status: Gut-and-amend for another purpose.

ARGUMENTS IN SUPPORT: The California-Hawaii State Conference of the NAACP emphasizes the importance of this bill for safeguarding civil rights:

ADTs have been found to exhibit biases and consequently have resulted in discriminatory impacts and harm to marginalized communities. A study published in *Science* showed that a clinical algorithm used across hospitals for determining patient care was racially biased against Black patients. The algorithm used healthcare spending as a proxy for health needs and falsely concluded that Black patients were healthier than equally sick white patients, depriving Black patients from needed high-risk care. These results were biased because Black patients face disproportionate poverty levels and spend less on health care.

ARGUMENTS IN OPPOSITION: The opposition coalition warns of unintended harmful consequences:

[O]verregulation in this space can easily undermine many beneficial uses of ADT—including the ability to develop and deploy these tools in a manner that can in fact reduce the instances and effects of human bias. To that end, we urge greater clarity, precision, and narrowing of the bill to avoid unintended consequences[.]

REGISTERED SUPPORT / OPPOSITION:

Support

California-Hawaii State Conference of the NAACP
Israeli-American Civic Action Network
Oakland Privacy
Santa Monica Democratic Club

Oppose unless Amended

California Association of Realtors

Opposition

American Financial Services Association
California Apartment Association
California Bankers Association
California Chamber of Commerce
California Credit Union League
California Financial Services Association
California Grocers Association
California League of Food Producers
California Manufacturers & Technology Association
California Retailers Association
Card Coalition
Civil Justice Association of California

Computer & Communications Industry Association
Insights Association
National Payroll Reporting Consortium
Netchoice
Software & Information Industry Association
State Privacy and Security Coalition, INC.
TechNet

Analysis Prepared by: Jith Meganathan / P. & C.P. / (916) 319-2200