

SENATE JUDICIARY COMMITTEE
Senator Thomas Umberg, Chair
2023-2024 Regular Session

SB 721 (Becker)
Version: March 22, 2023
Hearing Date: April 18, 2023
Fiscal: Yes
Urgency: No
CK

SUBJECT

California Interagency AI Working Group

DIGEST

This bill creates the California Interagency AI Working Group tasked with delivering a report to the Legislature regarding artificial intelligence (AI). The bill includes a sunset of January 1, 2030.

EXECUTIVE SUMMARY

Owing to recent advances in processing power and the rise of big data, artificial intelligence's (AI) capacity and the scope of its applications have expanded rapidly, impacting how we communicate, interact, entertain ourselves, travel, transact, and consume media. 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.

California does not regulate AI as such, but several laws address some of its most prominent applications, including autonomous vehicles, social media, bots, fake news, drones, privacy, and job displacement. These efforts offer creative approaches to novel problems, but the rapid proliferation and breadth of AI's manifestations underscores the challenge of regulation in this field and the need for principled governance of AI.

This bill establishes the California Interagency AI Working Group tasked with delivering a report to the Legislature regarding artificial intelligence (AI). Group members are to be of varied expertise and shall take input from a broad range of stakeholders to, among other things, recommend a definition of AI and determine the

relevant agencies to develop and oversee artificial intelligence policy and implementation of that policy. The bill includes a sunset of January 1, 2030. The bill is author-sponsored. There is no known support or opposition. It passed through the Senate Governmental Organization Committee on a vote of 14 to 1.

PROPOSED CHANGES TO THE LAW

Existing law:

- 1) Establishes the California Department of Technology (CDT) within the Government Operations Agency, 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).)

This bill:

- 1) Creates the California Interagency AI Working Group. The group is required to deliver a report to the Legislature regarding AI. Working group members must be Californians with expertise in at least two of the following areas:
 - a) computer science;
 - b) AI;
 - c) the technology industry;
 - d) workforce development; or
 - e) data privacy.
- 2) Provides that the Director shall be the chair of the working group and the working group shall select a vice chair from among its members. The groups shall consist of the following 10 members:

- a) two appointees by the Governor;
 - b) two appointees by the President Pro Tempore of the Senate;
 - c) two appointees by the Speaker of the Assembly;
 - d) two appointees by the Attorney General;
 - e) one appointee by the California Privacy Protection Agency; and
 - f) one appointee by CDT.
- 3) Requires the working group to take input from a broad range of stakeholders, including from academia, consumer advocacy groups, and businesses of all sizes affected by AI policies.
- 4) Requires the working group to do all of the following:
- a) recommend a definition of AI as it pertains to its use in technology for use in legislation;
 - b) study the implications of the usage of AI for data collection to inform testing and evaluation, verification and validation of AI to ensure that AI will perform as intended, including when interacting with humans and other systems, develop common metrics to assess trustworthiness that AI systems will perform as intended, and minimize performance problems and unanticipated outcomes;
 - c) determine proactive steps to prevent artificial intelligence-assisted misinformation campaigns and unnecessary exposure for children to the potentially harmful effects of artificial intelligence;
 - d) determine the relevant agencies to develop and oversee artificial intelligence policy and implementation of that policy; and
 - e) determine how the working group and the Department of Justice (DOJ) can leverage the substantial and growing expertise of the California Privacy Protection Agency (PPA) in the long-term development of data privacy policies that affect the privacy, rights, and the use of AI online.
- 5) Requires the working group, on or before January 1, 2025, and every two years thereafter, to submit a report to the Legislature regarding the items described above.
- 6) Provides that it shall remain in effect until January 1, 2030, and as of that date is repealed.

COMMENTS

1. The need for principled governance of AI

- a. *What is AI?*

Intelligence is an amorphous concept. Broadly, a thing can be considered intelligent if it processes information and solves problems in uncertain conditions. Artificial intelligence, then, is intelligence exhibited by the things we build. The term also refers to a field of study pioneered by Dartmouth professor John McCarthy in the 1950s, who described the endeavor of developing AI as “that of making a machine behave in ways that would be called intelligent if a human were so behaving.”¹

In certain narrow fields, computers have far surpassed human capabilities. Some of the most dramatic milestones in AI’s advancement have come when computer programs have bested humans at games, prompting varied reactions in their human counterparts. When IBM’s Deep Blue defeated World Chess Champion Gary Kasparov in 1997 in their decisive match, Kasparov insisted that a human grand master must have been controlling the machine. Better sportsmanship was displayed by *Jeopardy* legend Ken Jennings upon realizing he was certain to lose to IBM’s Watson in 2011. On his *Final Jeopardy* question video screen, Jennings wrote: “I, for one, welcome our new computer overlords.” By contrast, when the world’s best Go player, a Chinese national, succumbed to Google DeepMind’s AlphaGo, the Chinese government blocked the live-stream.

This most recent illustration of AI’s ascendance is especially notable because until recently, the prospect of a computer program beating a human professional player in Go was viewed as at least a decade away. This is because Go, an ancient Chinese strategy game, is more complex than chess by orders of magnitude. But whereas AlphaGo was programmed with input about the best moves in given situations and learned to improve upon existing strategies, a successor program, AlphaGo Zero, was programed to train itself based on the rules of the game alone. As a result, it developed startlingly creative strategies, defeating AlphaGo in 100 consecutive games. A generic spin-off of this program, AlphaZero, which was not specifically designed to play any particular game, mastered chess, Go, and shogi in just a few hours of self-play.²

Such developments reflect the recent advances in a subset of AI known as “machine learning,” a technique that “aims to help computers discover such fuzzy rules by themselves, without having to be explicitly instructed every step of the way by human programmers.”³ The most prominent type of machine learning is “deep learning,” which “uses artificial neural networks – simplified computer simulations of how biological neurons behave – to extract rules and patterns from sets of data.”⁴ The application of systems that simulate billions of neurons to massive datasets has driven

¹ Kaplan, *Artificial Intelligence: What Everyone Needs to Know* (2016) p. 1 (internal citations omitted).

² *DeepMind’s AI became a superhuman chess player in a few hours, just for fun* (Dec. 6, 2017) The Verge, <https://www.theverge.com/2017/12/6/16741106/deepmind-ai-chess-alphazero-shogi-go>. All internet citations are current as of April 7, 2023.

³ *How machine learning works* (May 14, 2015) The Economist, <https://www.economist.com/the-economist-explains/2015/05/13/how-machine-learning-works>.

⁴ *Id.*

recent advancements in areas previously thought difficult for AI to penetrate, including image recognition, voice recognition, translation, and autonomous vehicles.

Developments in just the last year have shown how quickly the technology can advance. Most are likely familiar with the recent explosion of AI tools known as “generative AI,” such as ChatGPT, Midjourney, Stable Diffusion, and Bing Chat. The advantages these new tools have over previous iterations include scale and ease of use. They have integrated unfathomably large amounts of data from which to learn and function and have sophisticated natural language interfaces that are able to process user requests in plain language.⁵

b. Hidden bias in algorithms

Machine learning “enables computer systems to learn and make predictions based on historical data. The machine learning process is powered by a machine learning algorithm, a function that is able to improve its performance over time by training itself using methods of data analysis and analytical modelling.”⁶ Algorithms process enormous datasets and make decisions with speed and reliability that vastly exceed human capabilities. “They determine everything from what ads we see on the Internet, to whether we are flagged for increased security screening at the airport, to our medical diagnoses and credit scores. They lie behind two of the most powerful products of the digital information age: Google Search and Facebook’s Newsfeed.”⁷ The most sophisticated algorithms need no supervision and use deep neural networks to “discover hidden patterns in data, typically those unrecognizable to, or difficult to discern by, humans.”⁸ In addition to organizing vast troves of data, algorithms offer the possibility of eliminating human biases in areas such as hiring decisions, credit scores, and criminal sentencing.

However, an algorithm is only as good as the information it is analyzing. Flawed inputs will produce flawed outputs. And an algorithm may key on factors other than those intended by its designer. In one example, a software student was dismayed to learn his program that could reliably distinguish dogs from wolves had, in actuality, learned to recognize snow in the background of the pictures rather than the canine’s features.⁹

⁵ *Written Statement of Aleksander Madry for the hearing: Advances in AI: Are We Ready For a Tech Revolution* (Mar. 8, 2023) Cybersecurity, Information Technology, and Government Innovation Subcommittee, https://oversight.house.gov/wp-content/uploads/2023/03/madry_written_statement100.pdf.

⁶ *Intro to AI for Policymakers: Understanding the Shift* (Mar. 2018) Brookfield Institute, p. 4, <http://brookfieldinstitute.ca/research-analysis/intro-to-ai-for-policymakers/>.

⁷ Jacob Weisberg, *The Digital Poorhouse* (June 7, 2018) *The New York Review of Books*, <https://www.nybooks.com/articles/2018/06/07/algorithms-digital-poorhouse/>.

⁸ *AI for Policymakers, supra*, at p. 5.

⁹ Wolfson, *Husky or Wolf? Using a Black Box Learning Model to Avoid Adoption Errors* (Aug. 24, 2017) UCI Applied Innovation, Currents, <http://innovation.uci.edu/2017/08/husky-or-wolf-using-a-black-box-learning-model-to-avoid-adoption-errors/>.

When we do not fully understand how an algorithm works, we are unable to determine which aspects of data it is focusing on.

And in many cases algorithms may inadvertently pick up human biases. In *Weapons of Math Destruction: How Big Data Increases Inequality and Threaten Democracy*, Cathy O’Neill shows that ultimately, a person shapes an algorithm, mediating the datasets gathered and deciding how to weigh them. These decisions are colored by our inherent biases and cultural predilections. Jacob Weisberg writes that “[c]orrelations reflected in historical data become invisibly entrenched in policy without programmers having ill intentions. Quantified information naturally points backward.”¹⁰ Rather than eliminating bias, some algorithms reinforce it, cloaking discrimination with mathematical neutrality.

ProPublica explored this phenomenon in the field of criminal justice.¹¹ Some jurisdictions factor algorithm-driven risk assessments into criminal bail, sentencing, and parole decisions. In 2014, U.S. Attorney General Eric Holder warned that the risk scores might be injecting bias into the courts. “Although these measures were crafted with the best of intentions, I am concerned that they inadvertently undermine our efforts to ensure individualized and equal justice,” he stated, adding, “they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society.” After obtaining the risk scores assigned to more than 7,000 people arrested in Broward County, Florida, ProPublica found that the scores “proved remarkably unreliable in forecasting violent crime.”¹² ProPublica’s study validated Holder’s fears: “Black defendants were still 77 percent more likely to be pegged as at higher risk of committing a future violent crime and 45 percent more likely to be predicted to commit a future crime of any kind.”¹³

Wiesberg likens biased algorithms to a dog that barks at black people because it has absorbed its racist owner’s prejudices. He writes:

This kind of feedback loop helps to explain the “racist dog” phenomenon of ostensibly race-neutral criminal justice algorithms. If a correlation of dark skin and criminality is reflected in data based on patterns of racial profiling, then processing historical data will predict that blacks will commit more crimes, even if neither race nor a proxy for race is encoded as an input variable. The prediction brings more supervision, which supports the prediction.¹⁴

¹⁰ *The Digital Poorhouse*, *supra*.

¹¹ Julia Angwin, et al., *Machine Bias* (May 23, 2016) ProPublica, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

¹² *Id.*

¹³ *Id.*

¹⁴ *The Digital Poorhouse*, *supra*.

A lack of transparency reduces accountability, again underscoring the importance of oversight. “[A]lgorithms simply grind out their results, and it is up to humans to review and address how that data is presented to users, to ensure the proper context and application of that data.”¹⁵

As directed by the National Artificial Intelligence Initiative Act of 2020, the National Institute of Standards and Technology (NIST) developed the AI Risk Management Framework. Speaking to these issues, the framework asserts:

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. Without proper controls, AI systems can amplify, perpetuate, or exacerbate inequitable or undesirable outcomes for individuals and communities. With proper controls, AI systems can mitigate and manage inequitable outcomes.¹⁶

New York University School of Law Professor Sarah Valentine puts a finer point on it: “Helpful as algorithms may be, they inevitably target marginalized populations and exacerbate the social stratification and vast inequality that already exists in our society.”¹⁷

c. Automation of jobs

The rapid progress in AI has transformed how we experience everyday life, including how we communicate, interact, entertain ourselves, travel, transact, and consume media. These developments constitute significant gains for society and promise more benefits in the near future.

But even the most beneficial forms of AI entail trade-offs, the most immediate and obvious of which is job reductions. For instance, a transition to autonomous vehicles in

¹⁵ Keith Kirkpatrick, *Battling Algorithmic Bias* (Oct. 2016) Communications of the ACM Vol. 59, No. 10, pp. 16-17, <https://cacm.acm.org/magazines/2016/10/207759-battlingalgorithmic-bias/abstract>.

¹⁶ *Artificial Intelligence Risk Management Framework* (Jan. 2023) NIST, <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>.

¹⁷ *Artificial Intelligence and Predictive Algorithms: Why Big Data Can Lead to Big Problems* (2019) 46 Fordham Urb. L.J. 364, 365.

transportation industries, particularly trucking, promises to displace millions of jobs.¹⁸ And automation's decimation of the workforce is no longer limited to blue collar jobs. Algorithms that use machine learning, including generative AI, and image technology are disrupting white collar jobs, including medicine, finance, and law.¹⁹ AI expert Kai Fu Lee has stated that 40 percent of the world's jobs could be automated within the next 15 years.²⁰ How long before robots begin analyzing legislation?

The social, economic, and political consequences of job reductions present a major challenge for policymakers.²¹ The 2016 election was partly determined by a swelling population of disgruntled, economically dislocated individuals. And, as more and more humans do less and less, power and resources will continue to accumulate in the hands of fewer people, potentially exacerbating inequality and political instability. The future may be ruled not by AI, but by the governments or trillionaire oligarchs who control AI.²² Russian President Vladimir Putin has even warned that "the one who becomes the leader in this sphere will be the ruler of the world."²³

d. Security threats

AI poses profound security challenges. Some forms of AI are designed to be deployed maliciously, jeopardizing safety and privacy. A report by several AI research organizations describes these threats in chilling detail.²⁴ One form of malevolent AI are cyberattacks, which include computer viruses capable of disguising themselves, automated hacking programs, the malicious publication or distortion of private information, massive surveillance programs both private and governmental, and the use of speech synthesis and image manipulation for impersonation. A prominent

¹⁸ Scott Santens, *Self-Driving Trucks Are Going to Hit Us Like a Human-Driven Truck* (May, 14, 2015) Medium, <https://medium.com/basic-income/self-driving-trucks-are-going-to-hit-us-like-a-human-driven-truck-b8507d9c5961>.

¹⁹ Robert Gloy, *How AI threatens white-collar jobs* (Oct. 8, 2018) Technologist, <https://www.technologist.eu/the-threat-to-white-collar-jobs/>.

²⁰ Don Reisinger, *AI Expert Says Automation Could Replace 40% of Jobs in 15 Years* (Jan. 10, 2019) Fortune, <http://fortune.com/2019/01/10/automation-replace-jobs/>.

²¹ AI also raises novel ethical questions. For instance, autonomous vehicles may calculate life-or-death risks differently than humans. In a variant of the classic thought experiment known as the "trolley problem," we can expect scenarios in which an automated vehicle, in the split second before an unavoidable collision occurs, must decide among human lives. "Of course, it isn't the car making the decision. The software engineers are making it, cosseted in their dim engineering warrens. They will play God." *Ethicists Will Solve the Robocar 'Trolley Problem'* (May 28, 2017) Wired, <https://www.wired.com/2017/05/autonomous-vehicles-trolley-problem/>.

²² Yuval Noah Harari, *Why Technology Favors Tyranny* (Oct. 2018) The Atlantic, <https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/>.

²³ *Putin: Leader in artificial intelligence will rule world* (Sep. 4, 2017) CNBC, <https://www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html>.

²⁴ See Miles Brundage, et al., *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (Feb. 2018) pp. 23-30, https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v_50335.pdf.

example of this last category is the recently viral images created using Midjourney's AI image generator, which included AI-generated deep fakes of former President Donald Trump being arrested and Pope Francis sporting stylish coats. Clearly these tools can be used to create widespread misinformation.

AI can also pose a threat to physical security. The Future of Life Institute, which argues that autonomous weapons – weapons capable of engaging targets without human intervention – may be “the third revolution in warfare, after gunpowder and nuclear arms.”²⁵ AI technology has reached the point where autonomous weapons, such as drones or even swarms of micro-drones, that hunt and kill people who meet predefined criteria, could feasibly be deployed in a matter of years. The development of such technology would almost inevitably precipitate a global arms race. Additionally, threats to physical safety include cyberattacks on physical systems, such as autonomous vehicles, airplanes, pacemakers, or the electrical grid. As AI grows more sophisticated, the magnitude of these threats grows.

Given these risks the Future of Life Institute recently published another open letter, this time calling for a pause on giant AI experiments:

Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: *Should* we let machines flood our information channels with propaganda and untruth? *Should* we automate away all the jobs, including the fulfilling ones? *Should* we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? *Should* we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. **Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.** This confidence must be well justified and increase with the magnitude of a system's potential effects. OpenAI's recent statement regarding artificial general intelligence, states that "At some point, it may be important to get independent review before starting to train future systems, and for the most advanced efforts to agree to limit the rate of growth of compute used for creating new models." We agree. That point is now.

Therefore, **we call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.** This pause should be public and verifiable, and include all key actors. If such a pause cannot be enacted quickly, governments should step in and institute a moratorium.²⁶

²⁵ Future of Life Institute, *Autonomous weapons: an open letter from AI & robotics researchers* (Feb. 9, 2016) <https://futureoflife.org/open-letter-autonomous-weapons/?cn-reloaded=1>.

²⁶ Future of Life Institute, *Pause Giant AI Experiments: An Open Letter* (2023) <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>. Signatories to the letter include Yoshua Bengio, Founder and

2. Developing a strategy for dealing with AI: roadmaps and legislative efforts

a. *Blueprints for effective regulation and oversight*

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.

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.²⁷

b. *Legislative attempts*

A host of bills in recent years have sought to take initial steps toward meaningful regulation of AI by creating advisory bodies to report to the Legislature.

AB 976 (Chau, 2019) would have established the Artificial Intelligence in State Government Services Commission and required the advisory commission to annually convene a public process to gather input on how artificial intelligence and data science could be used to improve state services. AB 459 (Kiley, 2019) would have required the Commission proposed by AB 976 to report to the Legislature on its recommended minimum standards for the use of AI in state government. AB 1576 (Calderon, 2019) would have provided for the creation of an advisory working group to report on specified issues relating to the use of AI by California-based businesses. SB 348 (Chang, 2019) would have required the Director to devise a strategic plan, as specified, to utilize AI technology to improve state services. Each of these bills died in the Senate or Assembly Appropriations Committees.

AB 594 (Salas, 2019) would have authorized the Director of CDT to designate a position with the department to evaluate the uses of AI in state government and to advise the Director on incorporating AI into state IT strategic plans, policies, standards and enterprise architecture, and required CDT to adopt guidelines by January 1, 2021 to govern the use and implementation of AI technologies in state government functions, as specified. However, the bill was vetoed by Governor Gavin Newsom, who asserted that the legislation was not necessary to accomplish the goal of examining the impact of AI on work and our economy.

²⁷ *Blueprint For An AI Bill Of Rights* (Oct. 2022) Office of Science and Technology Policy, <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>.

This bill similarly seeks to take a modest approach. It establishes the California Interagency AI Working Group in order to study the relevant implications of AI, determine proper protocols and oversight infrastructure, and report it to the Legislature on or before January 1, 2025, and every two years thereafter. The Director chairs the group, which is made up of appointees by both houses of the Legislature, the Governor, the Attorney General, the PPA, and CDT. Members must have expertise in at least two specified areas, including computer science, AI, workforce development, and data privacy. The working group is disbanded by a sunset of January 1, 2030.

This represents another Legislative attempt to create an informed body that can help guide California forward in this era of AI. According to the author:

Artificial Intelligence (AI) is a rapidly developing area of technology that is already integrated into most parts of our everyday life, including voice assistants (e.g., Apple's Siri feature), personalized student learning, spam filters, facial recognition, and much more. In particular the use of neural networks – a mathematical system that uses vast amounts of data to find statistical patterns – has changed the field in the last decade. Neural networks have increased the personalization and accuracy of AI leading directly to some of its most notable uses, including self-driving cars and ChatGPT.

Existing law has created a framework to analyze the impact of AI technology within state agencies. However, this bill gives us an opportunity to look at the impact of AI holistically and suggest guardrails to protect against misuse and manipulation for both state agencies and the general public.

It is important that we continue to invest in innovation and new technology, but we need to ensure the path forward does not create undue harm. The California Interagency AI Working Group will bring together some of the best minds in a broad swath of fields to anticipate issues and brainstorm paths forward as we head into the future of AI.

SUPPORT

None known

OPPOSITION

None known

RELATED LEGISLATION

Pending Legislation:

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.

SB 313 (Dodd, 2023) establishes the Office of Artificial Intelligence. It requires state agencies to disclose when they are using generative artificial intelligence to communicate with a person and to provide them an option to speak with a natural person at the agency. SB 313 is currently in this Committee and is set to be heard the same day as this bill.

SB 398 (Wahab, 2023) establishes the Government Services Advanced Technology Act, which requires the Department of Justice to develop and implement a comprehensive research plan to study the feasibility of using advanced technology to improve state and local government services. The bill must include an analysis of the potential benefits and risks of using AI to assist disaster victims in finding and applying for disaster relief funds and to assist individuals in determining their eligibility for various public benefits programs. The bill require the department, on or before January 1, 2026, to provide a report to the Legislature on the findings of its research. SB 398 is currently in the Senate Public Safety Committee.

AB 331 (Bauer-Kahan, 2023) requires a deployer of an automated decision tool to perform an impact assessment for any automated decision tool the deployer uses and a developer of an automated decision tool to complete and document an assessment of any automated decision tool that it designs, codes, or produces. It requires deployers and developers to institute certain policies and provide certain disclosures. AB 331 is currently in the Assembly Appropriations Committee.

Prior Legislation:

SB 444 (Umberg, 2019) would have requested the Regents of the University of California to enact, no later than July 1, 2021, a resolution to authorize the university's Berkeley and Irvine law schools, in consultation with the superior courts of the Counties of Alameda and Orange, respectively, to participate in a research program to develop AI or machine-learning solutions to address issues of access to justice faced by pro se indigent litigants who are residents of those counties. This bill died in the Assembly Higher Education Committee.

SB 348 (Chang, 2019) *See* Comment 2.

AB 459 (Kiley, 2019) *See* Comment 2.

AB 594 (Salas, 2019) *See* Comment 2.

AB 976 (Chau, 2019) *See* Comment 2.

AB 1576 (Calderon, 2019) *See* Comment 2.

ACR 215 (Kiley, Ch. 206, Stats. 2018) *See* Comment 2.

PRIOR VOTES:

Senate Governmental Organization Committee (14 Ayes , 1 No)
