AMENDED IN ASSEMBLY JULY 3, 2024

AMENDED IN ASSEMBLY JUNE 20, 2024

AMENDED IN ASSEMBLY JUNE 5, 2024

AMENDED IN SENATE MAY 16, 2024

AMENDED IN SENATE APRIL 30, 2024

AMENDED IN SENATE APRIL 16, 2024

AMENDED IN SENATE APRIL 8, 2024

AMENDED IN SENATE MARCH 20, 2024

SENATE BILL No. 1047

Introduced by Senator Wiener
(Coauthors: Senators Roth, Rubio, and Stern)

February 7, 2024

An act to add Chapter 22.6 (commencing with Section 22602) to Division 8 of the Business and Professions Code, and to add Sections 11547.6 and 11547.7 to the Government Code, relating to artificial intelligence.

LEGISLATIVE COUNSEL'S DIGEST

SB 1047, as amended, Wiener. Safe and Secure Innovation for Frontier Artificial Intelligence Models Act.

Existing law requires the Secretary of Government Operations to develop a coordinated plan to, among other things, investigate the feasibility of, and obstacles to, developing standards and technologies for state departments to determine digital content provenance. For the purpose of informing that coordinated plan, existing law requires the

secretary to evaluate, among other things, the impact of the proliferation of deepfakes, defined to mean audio or visual content that has been generated or manipulated by artificial intelligence that would falsely appear to be authentic or truthful and that features depictions of people appearing to say or do things they did not say or do without their consent, on state government, California-based businesses, and residents of the state.

Existing law creates the Department of Technology within the Government Operations Agency and requires the department to, among other things, identify, assess, and prioritize high-risk, critical information technology services and systems across state government for modernization, stabilization, or remediation.

This bill would enact the Safe and Secure Innovation for Frontier Artificial Intelligence Models Act to, among other things, require that a developer, before initially training a covered model, as defined, comply with various requirements, including implementing the capability to promptly enact a full shutdown, as defined, and implement a written and separate safety and security protocol, as specified. The bill would prohibit a developer from using a covered model commercially or publicly, or making a covered model or a covered model derivative available for commercial or public use, if there is an unreasonable risk that the covered model or covered model derivative can cause or enable a critical harm, as defined. *The bill would require a developer, beginning January 1, 2028, to annually retain a third-party auditor to perform an independent audit of compliance with the requirements of the bill, as provided.*

This bill would require a developer of a covered model to submit to the Frontier Model Division, which the bill would create within the Government Operations Agency, a certification under penalty of perjury of compliance with these provisions, as specified. By expanding the scope of the crime of perjury, this bill would impose a state-mandated local program. The bill would also require a developer of a covered model to report each artificial intelligence safety incident affecting the covered model or any covered model derivative controlled by the developer, as specified, to the Frontier Model Division.

This bill would require a person that operates a computing cluster, as defined, to implement written policies and procedures to do certain things when a customer utilizes compute resources that would be sufficient to train a covered model, including assess whether a

prospective customer intends to utilize the computing cluster to train a covered model.

~~This bill would punish a violation of these provisions with a civil penalty, as prescribed, to be recovered by the Attorney General.~~

This bill would specify unlawful acts under these provisions and authorize the Attorney General or the Labor Commissioner to bring a civil action, as provided. The bill would also provide for whistleblower protections, including prohibiting a developer of a covered model or a contractor or subcontractor of the developer from preventing an employee from disclosing information, or retaliating against an employee for disclosing information, to the Attorney General or Labor Commissioner if the employee has reasonable cause to believe the developer is out of compliance with certain requirements or that the covered model poses an unreasonable risk of critical harm.

This bill would create the Board of Frontier Models within the Government Operations Agency, independent of the Department of Technology, and provide for the board's membership. The bill would also create the Frontier Model Division within the Government Operations Agency and under the direct supervision of the board, and would require the division to, among other things, review annual certification reports from developers received pursuant to these provisions and publicly release summarized findings based on those reports. The bill would require the division to, on or before January 1, 2027, and annually thereafter, issue regulations to update the definition of a "covered model," as provided. The bill would authorize the division to assess related fees and would require deposit of the fees into the Frontier Model Division Programs Fund, which the bill would create. The bill would make moneys in the fund available for the purpose of these provisions only upon appropriation by the Legislature.

This bill would also require the Department of Technology to commission consultants, as prescribed, to create a public cloud computing cluster, to be known as CalCompute, with the primary focus of conducting research into the safe and secure deployment of large-scale artificial intelligence models and fostering equitable innovation that includes, among other things, a fully owned and hosted cloud platform.

The California Constitution requires the state to reimburse local agencies and school districts for certain costs mandated by the state. Statutory provisions establish procedures for making that reimbursement.

This bill would provide that no reimbursement is required by this act for a specified reason.

Vote: majority. Appropriation: no. Fiscal committee: yes. State-mandated local program: yes.

The people of the State of California do enact as follows:

SECTION 1. This act shall be known, and may be cited, as the Safe and Secure Innovation for Frontier Artificial Intelligence Models Act.

SEC. 2. The Legislature finds and declares all of the following:

(a) California is leading the world in artificial intelligence innovation and research, through companies large and small, as well as through our remarkable public and private universities.

(b) Artificial intelligence, including new advances in generative artificial intelligence, has the potential to catalyze innovation and the rapid development of a wide range of benefits for Californians and the California economy, including advances in medicine, wildfire forecasting and prevention, and climate science, and to push the bounds of human creativity and capacity.

(c) If not properly subject to human controls, future development in artificial intelligence may also have the potential to be used to create novel threats to public safety and security, including by enabling the creation and the proliferation of weapons of mass destruction, such as biological, chemical, and nuclear weapons, as well as weapons with cyber-offensive capabilities.

(d) The state government has an essential role to play in ensuring that California recognizes the benefits of this technology while avoiding the most severe risks, as well as to ensure that artificial intelligence innovation and access to compute is accessible to academic researchers and startups, in addition to large companies.

SEC. 3. Chapter 22.6 (commencing with Section 22602) is added to Division 8 of the Business and Professions Code, to read:

Chapter 22.6. Safe and Secure Innovation for Frontier Artificial Intelligence Models

22602. As used in this chapter:

(a) "Advanced persistent threat" means an adversary with sophisticated levels of expertise and significant resources that allow it, through the use of multiple different attack vectors, including, but not limited to, cyber, physical, and deception, to

generate opportunities to achieve its objectives that are typically to establish and extend its presence within the information technology infrastructure of organizations for purposes of exfiltrating information or to undermine or impede critical aspects of a mission, program, or organization or place itself in a position to do so in the future.

(b) "Artificial intelligence" means 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.

(c) "Artificial intelligence safety incident" means an incident that demonstrably increases the risk of a critical harm occurring by means of any of the following:

(1) A covered model autonomously engaging in behavior other than at the request of a user.

(2) Theft, misappropriation, malicious use, inadvertent release, unauthorized access, or escape of the model weights of a covered model.

(3) The critical failure of technical or administrative controls, including controls limiting the ability to modify a covered model.

(4) Unauthorized use of a covered model to cause or enable critical harm.

(d) "Computing cluster" means a set of machines transitively connected by data center networking of over 100 gigabits per second that has a theoretical maximum computing capacity of at least 10^20 integer or floating-point operations per second and can be used for training artificial intelligence.

(e) (1) "Covered model" means either of the following:

(A) Before January 1, 2027, "covered model" means either of the following:

(i) An artificial intelligence model trained using a quantity of computing power greater than 10^26 integer or floating-point operations, the cost of which exceeds one hundred million dollars ($100,000,000) when calculated using the average market prices of cloud compute at the start of training as reasonably assessed by the developer.

(ii) An artificial intelligence model created by fine-tuning a covered model using a quantity of computing power equal to or greater than three times 10^25 integer or floating-point operations.

(B) (i) Except as provided in clause (ii), on and after January 1, 2027, "covered model" means any of the following:

(I) An artificial intelligence model trained using a quantity of computing power determined by the Frontier Model Division pursuant to Section 11547.6 of the Government Code, the cost of which exceeds one hundred million dollars ($100,000,000) when calculated using the average market price of cloud compute at the start of training as reasonably assessed by the developer.

(II) An artificial intelligence model created by fine-tuning a covered model using a quantity of computing power that exceeds a threshold determined by the Frontier Model Division.

(ii) If the Frontier Model Division does not adopt a regulation governing subclauses (I) and (II) of clause (i) by January 1, 2027, the definition of "covered model" in subparagraph (A) continues to be in effect until the regulation is adopted.

(2) On and after January 1, 2026, the dollar amount in this subdivision shall be adjusted annually for inflation to the nearest one hundred dollars ($100) based on the change in the annual California Consumer Price Index for All Urban Consumers published by the Department of Industrial Relations for the most recent annual period ending on December 31 preceding the adjustment.

(f) "Covered model derivative" means any of the following:

(1) An unmodified copy of a covered model.

(2) A copy of a covered model that has been subjected to post-training modifications unrelated to fine-tuning.

(3) (A) (i) Before January 1, 2027, a copy of a covered model that has been fine-tuned using a quantity of computing power not exceeding three times 10^25 integer or floating point operations.

(ii) On and after January 1, 2027, a copy of a covered model that has been fine-tuned using a quantity of computing power not exceeding a threshold determined by the Frontier Model Division.

(B) If the Frontier Model Division does not adopt a regulation governing clause (ii) of subparagraph (A) by January 1, 2027, the quantity of computing power specified in clause (i) of subparagraph (A) shall continue to apply until the regulation is adopted.

(4) A copy of a covered model that has been combined with other software.

(g) (1) "Critical harm" means any of the following harms caused or enabled by a covered model or covered model derivative:

(A) The creation or use of a chemical, biological, radiological, or nuclear weapon in a manner that results in mass casualties.

(B) Mass casualties or at least five hundred million dollars ($500,000,000) of damage resulting from cyberattacks on critical ~~infrastructure, occurring either in a single incident or over multiple related incidents.~~ *infrastructure by a model providing precise instructions for conducting a cyberattack or series of cyberattacks on critical infrastructure.*

(C) Mass casualties or at least five hundred million dollars ($500,000,000) of damage resulting from an artificial intelligence model ~~autonomously~~ engaging in conduct that ~~would constitute a serious or violent felony under the Penal Code if undertaken by a human with the requisite mental state.~~ *does both of the following:*

(i) Acts with limited human oversight, intervention, or supervision.

(ii) Results in death, great bodily injury, property damage, or property loss, and would, if committed by a human, constitute a crime specified in the Penal Code that requires intent, recklessness, or gross negligence, or the solicitation or aiding and abetting of such a crime.

(D) Other grave harms to public safety and security that are of comparable severity to the harms described in subparagraphs (A) to (C), inclusive.

(2) "Critical harm" does not include ~~harms~~ *either of the following:*

(A) Harms caused or enabled by information that a covered model outputs if the information is otherwise publicly ~~accessible.~~ *accessible from sources other than a covered model.*

(B) Harms caused or materially enabled by a covered model combined with other software, including other models, if the covered model did not materially contribute to the other software's ability to cause or materially enable the harm.

(3) On and after January 1, 2026, the dollar amounts in this subdivision shall be adjusted annually for inflation to the nearest one hundred dollars ($100) based on the change in the annual California Consumer Price Index for All Urban Consumers published by the Department of Industrial Relations for the most recent annual period ending on December 31 preceding the adjustment.

(h) "Critical infrastructure" means assets, systems, and networks, whether physical or virtual, the incapacitation or destruction of which would have a debilitating effect on physical security, economic security, public health, or safety in the state.

(i) "Developer" means a person that performs the initial training of a covered model either by training a model using a sufficient quantity of computing power, or by fine-tuning an existing covered model using sufficient quantity of computing power pursuant to subdivision (e).

(j) "Fine-tuning" means adjusting the model weights of a trained covered model by exposing it to additional data.

(k) "Frontier Model Division" means the Frontier Model Division created pursuant to Section 11547.6 of the Government Code.

(*l*) "Full shutdown" means the cessation of operation of any of the following:

(1) The training of a covered model.

(2) A covered model.

(3) All covered model derivatives controlled by a developer.

(m) "Model weight" means a numerical parameter in an artificial intelligence model that is adjusted through training and that helps determine how inputs are transformed into outputs.

(n) "Open-source artificial intelligence model" means an artificial intelligence model that is made freely available and that may be freely modified and redistributed.

(o) "Person" means an individual, proprietorship, firm, partnership, joint venture, syndicate, business trust, company, corporation, limited liability company, association, committee, or any other nongovernmental organization or group of persons acting in concert.

(p) "Post-training modification" means modifying the capabilities of a covered model by any means, including, but not limited to, fine-tuning, providing the model with access to tools or data, removing safeguards against hazardous misuse or misbehavior of the model, or combining the model with, or integrating it into, other software.

(q) "Reasonable assurance" does not mean full certainty or practical certainty.

(r) "Safety and security protocol" means documented technical and organizational protocols that meet both of the following criteria:

(1) The protocols are used to manage the risks of developing and operating covered models across their life cycle, including risks posed by causing or enabling or potentially causing or enabling the creation of covered model derivatives.

(2) The protocols specify that compliance with the protocols is required in order to train, operate, possess, and provide external access to the developer's covered model.

22603. (a) Before a developer initially trains a covered model, the developer shall do all of the following:

(1) Implement administrative, technical, and physical cybersecurity protections to prevent unauthorized access to, misuse of, or unsafe post-training modifications of, the covered model and all covered model derivatives controlled by the developer that are appropriate in light of the risks associated with the covered model, including from advanced persistent threats or other sophisticated actors.

(2) Implement the capability to promptly enact a full shutdown.

(3) Implement a written and separate safety and security protocol that does all of the following:

(A) If a developer complies with the safety and security protocol, provides reasonable assurance that the developer will not produce a covered model or covered model derivative that poses an unreasonable risk of causing or enabling a critical harm.

(B) States compliance requirements in an objective manner and with sufficient detail and specificity to allow the developer or a third party to readily ascertain whether the requirements of the safety and security protocol have been followed.

(C) Identifies specific tests and test results that would be sufficient to provide reasonable assurance of both of the following:

(i) That a covered model does not pose an unreasonable risk of causing or enabling a critical harm.

(ii) That covered model derivatives do not pose an unreasonable risk of causing or enabling a critical harm.

(D) Describes in detail how the testing procedure assesses the risks associated with post-training modifications.

(E) Describes in detail how the testing procedure addresses the possibility that a covered model can be used to make post-training

modifications or create another covered model in a manner that may generate hazardous capabilities.

(F) Provides sufficient detail for third parties to replicate the testing procedure.

(G) Describes in detail how the developer will fulfill their obligations under this chapter.

(H) Describes in detail how the developer intends to implement the safeguards and requirements referenced in this section.

(I) Describes in detail the conditions under which a developer would enact a full shutdown.

(J) Describes in detail the procedure by which the safety and security protocol may be modified.

(4) Ensure that the safety and security protocol is implemented as written, including by designating senior personnel to be responsible for ensuring compliance by employees and contractors working on a covered model, monitoring and reporting on ~~implementation, and conducting audits, including through third-party auditors.~~ *implementation.*

(5) Provide a copy of the safety and security protocol to the Frontier Model Division.

(6) Conduct an annual review of the safety and security protocol to account for any changes to the capabilities of the covered model and industry best practices and, if necessary, make modifications to the policy.

(7) If the safety and security protocol is modified, provide an updated copy to the Frontier Model Division within 10 business days.

(8) Implement other reasonable measures to prevent covered models and covered model derivatives from posing unreasonable risks of causing or enabling critical harms.

(b) Before using a covered model or covered model derivative, or making a covered model or covered model derivative available for commercial or public use, the developer of a covered model shall do all of the following:

(1) Assess whether the covered model is reasonably capable of causing or enabling a critical harm.

(2) Implement reasonable safeguards to prevent the covered model and covered model derivatives from causing or enabling a critical harm.

(3) Ensure, to the extent reasonably possible, that the covered model's actions and the actions of covered model derivatives, as well as critical harms resulting from their actions, can be accurately and reliably attributed to them.

~~(4) Beginning January 1, 2028, obtain a certificate of compliance from a third-party auditor who has been accredited pursuant to 11547.6 of the Government Code.~~

(c) A developer shall not use a covered model commercially or publicly, or make a covered model or a covered model derivative available for commercial or public use, if there is an unreasonable risk that the covered model or covered model derivative can cause or enable a critical harm.

(d) A developer of a covered model shall annually reevaluate the procedures, policies, protections, capabilities, and safeguards implemented pursuant to this section.

(e) (1) Beginning January 1, 2028, a developer of a covered model shall annually retain a third-party auditor that conducts audits consistent with best practices for auditors to perform an independent audit of compliance with the requirements of this section.

(2) The auditor shall produce an audit report including all of the following:

(A) A detailed assessment of the developer's steps to comply with the requirements of this section.

(B) If applicable, any identified instances of noncompliance with the requirements of this section, and any recommendations for how the developer can improve its policies and processes for ensuring compliance with the requirements of this section.

(C) A detailed assessment of the developer's internal controls, including its designation and empowerment of senior personnel responsible for ensuring compliance by the developer, its employees, and its contractors.

(D) The signature of the lead auditor certifying the results of the auditor.

~~(e)~~

(f) (1) A developer of a covered model shall annually submit to the Frontier Model Division a certification under penalty of perjury of compliance with the requirements of this section signed by the chief technology officer, or a more senior corporate officer, in a format and on a date as prescribed by the Frontier Model

Division. This paragraph applies as long as the covered model or any covered model derivatives controlled by the developer remain in commercial or public use, or remain available for commercial or public use.

(2) In a certification submitted pursuant to paragraph (1), a developer shall specify or provide, at a minimum, all of the following:

(A) The nature and magnitude of critical harms that the covered model or covered model derivatives may reasonably cause or enable, and the outcome of the assessment required by *paragraph (1) of* subdivision (b).

(B) An assessment of the risk that compliance with the safety and security protocol may be insufficient to prevent the covered model or covered model derivatives from causing or enabling critical harms.

(C) A description of the process used by the signing officer to verify compliance with the requirements of this section, including a description of the materials reviewed by the signing officer, a description of testing or other evaluation performed to support the certification, and the contact information of any third parties relied upon to validate compliance.

(D) Beginning January 1, 2028, ~~a certificate of compliance from an accredited third-party auditor.~~ *the most recent audit report pursuant to subdivision (e).*

~~(f)~~

(g) A developer of a covered model shall report each artificial intelligence safety incident affecting the covered model, or any covered model derivatives controlled by the developer, to the Frontier Model Division within 72 hours of the developer learning of the artificial intelligence safety incident, or within 72 hours of the developer learning facts sufficient to establish a reasonable belief that an artificial intelligence safety incident has occurred.

~~(g)~~

(h) A developer shall submit to the Frontier Model Division, under penalty of perjury, a certification of compliance with the requirements of this section no more than 30 days after making a covered model or covered model derivative available for commercial or public use for the first time. A developer need not submit a certification for a covered model derivative if the

developer has already submitted a certification for the applicable covered model.

~~(h)~~

(i) In fulfilling their obligations under this chapter, a developer shall consider applicable guidance from the Frontier Model Division, National Institute of Standards and Technology, and other reputable standard-setting organizations.

22604. (a) A person that operates a computing cluster shall implement written policies and procedures to do all of the following when a customer utilizes compute resources that would be sufficient to train a covered model:

(1) Obtain a prospective customer's basic identifying information and business purpose for utilizing the computing cluster, including all of the following:

(A) The identity of that prospective customer.

(B) The means and source of payment, including any associated financial institution, credit card number, account number, customer identifier, transaction identifiers, or virtual currency wallet or wallet address identifier.

(C) The email address and telephonic contact information used to verify a prospective customer's identity.

(2) Assess whether a prospective customer intends to utilize the computing cluster to train a covered model.

(3) If a customer repeatedly utilizes computer resources that would be sufficient to train a covered model, validate the information initially collected pursuant to paragraph (1) and conduct the assessment required pursuant to paragraph (2) prior to each utilization.

(4) Retain a customer's Internet Protocol addresses used for access or administration and the date and time of each access or administrative action.

(5) Maintain for seven years and provide to the Frontier Model Division or the Attorney General, upon request, appropriate records of actions taken under this section, including policies and procedures put into effect.

(6) Implement the capability to promptly enact a full shutdown of any resources being used to train or operate ~~such customer's administered models.~~ *models under the customer's control.*

(b) A person that operates a computing cluster shall consider applicable guidance from the Frontier Model Division, National

Institute of Standards and Technology, and other reputable standard-setting organizations.

22605. (a) A developer of a covered model that provides commercial access to that covered model shall provide a transparent, uniform, publicly available price schedule for the purchase of access to that covered model at a given level of quality and quantity subject to the developer's terms of service and shall not engage in unlawful discrimination or noncompetitive activity in determining price or access.

(b) (1) A person that operates a computing cluster shall provide a transparent, uniform, publicly available price schedule for the purchase of access to the computing cluster at a given level of quality and quantity subject to the developer's terms of service and shall not engage in unlawful discrimination or noncompetitive activity in determining price or access.

(2) A person that operates a computing cluster may provide free, discounted, or preferential access to public entities, academic institutions, or for noncommercial research purposes.

~~22606. (a) If the Attorney General finds that a person is violating this chapter, the Attorney General may bring a civil action pursuant to this section.~~

~~(b) Subject to subdivision (d), in a civil action under this section, the court may award any of the following:~~

~~(1) (A) Preventive relief, including a permanent or temporary injunction, restraining order, or other order against the person responsible for a violation of this chapter, including deletion of the covered model and the weights utilized in that model.~~

~~(B) Relief pursuant to this paragraph shall be granted only in response to death or bodily harm to another human, harm to property, theft of property, or an imminent risk or threat to public safety.~~

~~(2) Other relief as the court deems appropriate, such as monetary damages, including punitive damages, to persons aggrieved and an order for the full shutdown of a covered model.~~

~~(3) A civil penalty in an amount not exceeding 10 percent of the cost of the quantity of computing power used to train the covered model to be calculated using average market prices of cloud compute at the time of training for a first violation and in an amount not exceeding 30 percent of that value for any subsequent violation.~~

~~(c) A court shall disregard corporate formalities and impose joint and several liability on affiliated entities for purposes of effectuating the intent of this section if the court concludes that both of the following are true:~~

~~(1) Steps were taken in the development of the corporate structure among affiliated entities to purposely and unreasonably limit or avoid liability.~~

~~(2) The corporate structure of the developer or affiliated entities would frustrate recovery of penalties or injunctive relief under this section.~~

~~(d) (1) For a violation that occurs before January 1, 2026, a court shall not do any of the following:~~

~~(A) Order the deletion of a covered model and the weights utilized in that model.~~

~~(B) Order the full shutdown of a covered model that does not present an imminent threat to public safety.~~

~~(C) Award a civil penalty under paragraph (3) of subdivision (b).~~

~~(2) For a violation that occurs before July 1, 2025, a court shall not award monetary damages to persons aggrieved.~~

~~22607. (a) Pursuant to subdivision (a) of Section 1102.5 of the Labor Code, a developer shall not prevent an employee from disclosing information to the Attorney General if the employee has reasonable cause to believe that the information indicates that the developer is out of compliance with the requirements of Section 22603.~~

~~(b) Pursuant to subdivision (b) of Section 1102.5 of the Labor Code, a developer shall not retaliate against an employee for disclosing information to the Attorney General if the employee has reasonable cause to believe that the information indicates that the developer is out of compliance with the requirements of Section 22603.~~

~~(c) The Attorney General may publicly release any complaint, or a summary of that complaint, pursuant to this section if the Attorney General concludes that doing so will serve the public interest.~~

~~(d) Employees shall seek relief for violations of subdivisions (a) and (b) pursuant to Sections 1102.61 and 1102.62 of the Labor Code.~~

~~(e) Pursuant to subdivision (a) of Section 1102.8 of the Labor Code, a developer shall provide clear notice to all employees working on covered models of their rights and responsibilities under this section.~~

~~(f) (1) Developers shall provide a reasonable internal process through which an employee may anonymously disclose information to the developer if the employee believes in good faith that the information indicates that the developer is out of compliance with the requirements of Section 22603 or has made false or materially misleading statements related to its safety and security protocol that includes, at a minimum, a monthly update to the disclosing employee regarding the status of the employee's disclosure and the actions taken by the developer in response to the disclosure.~~

~~(2) The disclosures and responses of the process required by this subdivision shall be maintained and shared with nonconflicted officers and directors of the company on a regular basis and not less than once per quarter.~~

~~(g) As used in this section, "employee" has the same meaning as defined in Section 1102.5 of the Labor Code and includes both of the following:~~

~~(1) Contractors or unpaid advisors involved with assessing, managing, or addressing hazardous capabilities of covered models.~~

~~(2) Corporate officers.~~

22606. (a) The following are unlawful acts:

(1) For a developer to fail to comply with any of the requirements of Section 22603 or subdivision (a) of Section 22605.

(2) For a person who operates a computing cluster to fail to comply with the requirements of Section 22604 or subdivision (b) of Section 22605.

(3) For a developer to fail to comply with any of the requirements of Section 22607.

(b) The following parties may bring a civil action pursuant to subdivision (a):

(1) The Attorney General to enforce any provision of this chapter.

(2) The Labor Commissioner to enforce any provision of Section 22607 that would constitute a violation of the Labor Code.

(c) The parties listed in subdivision (b) are entitled to recover all of the following in addition to any civil penalties specified in this chapter:

(1) A civil penalty for a violation that occurs on or after January 1, 2026, in an amount not exceeding 10 percent of the cost of the quantity of computing power used to train the covered model to be calculated using average market prices of cloud compute at the time of training for a first violation and in an amount not exceeding 30 percent of that value for any subsequent violation.

(2) (A) Injunctive or declaratory relief, including, but not limited to, orders to modify, implement a full shutdown, or delete the covered model and any covered model derivatives controlled by the developer.

(B) The court may only order relief under this paragraph for a covered model that has caused death or bodily harm to another human, harm to property, theft or misappropriation of property, or constitutes an imminent risk or threat to public safety.

(3) (A) Monetary damages.

(B) Punitive damages pursuant to subdivision (a) of Section 3294 of the Civil Code.

(4) Attorney's fees and costs.

(5) Any other relief that the court deems appropriate.

(d) (1) A provision within a contract or agreement that seeks to waive, preclude, or burden the enforcement of a liability arising from a violation of this chapter, or to shift that liability to any person or entity in exchange for their use or access of, or right to use or access, a developer's products or services, including by means of a contract of adhesion, is void as a matter of public policy.

(2) A court shall disregard corporate formalities and impose joint and several liability on affiliated entities for purposes of effectuating the intent of this section to the maximum extent allowed by law if the court concludes that both of the following are true:

(A) The affiliated entities, in the development of the corporate structure among the affiliated entities, took steps to purposely and unreasonably limit or avoid liability.

(B) As the result of the steps described in subparagraph (A), the corporate structure of the developer or affiliated entities would frustrate recovery of penalties or injunctive relief under this section.

(e) This section does not limit the application of other laws.

22607. (a) A developer of a covered model or a contractor or subcontractor of the developer shall not do either of the following:

(1) Prevent an employee from disclosing information to the Attorney General or the Labor Commissioner, including through terms and conditions of employment or seeking to enforce terms and conditions of employment, if the employee has reasonable cause to believe either of the following:

(A) The developer is out of compliance with the requirements of Section 22603.

(B) An artificial intelligence model, including a model that is not a covered model, poses an unreasonable risk of causing or materially enabling critical harm, even if the employer is not out of compliance with any law.

(2) Retaliate against an employee for disclosing information to the Attorney General or Labor Commissioner, if the employee has reasonable cause to believe either subparagraph (A) or (B) of paragraph (1).

(3) Make false or materially misleading statements related to its safety and security protocol in a manner that violates Part 2 (commencing with Section 16600) of Division 7 or any other provision of state law.

(b) (1) An employee harmed by a violation of this subdivision may petition a court for appropriate temporary or preliminary injunctive relief as provided in Sections 1102.61 and 1102.62 of the Labor Code.

(2) An employee of the Frontier Model Division may report any violation of this chapter by the Frontier Model Division to the State Auditor pursuant to the provisions of the California Whistleblower Protection Act (Article 3 (commencing with Section 8547) of Chapter 6.5 of Division 1 of Title 2 of the Government Code) which shall govern any such report.

(c) The Attorney General or Labor Commissioner may publicly release or provide, to the Frontier Model Division or the Governor, any complaint, or a summary of that complaint, pursuant to this section if they conclude that doing so will serve the public interest.

(d) A developer and any contractor or subcontractor of the developer shall provide a clear notice to all employees working on covered models of their rights and responsibilities under this section. A developer is presumed to be in compliance with the requirements of this subdivision if the developer does one of the following:

(1) At all times post and display within all workplaces maintained by the developer a notice to all employees of their rights and responsibilities under this section, ensure that all new employees receive equivalent notice, and ensure that employees who work remotely periodically receive an equivalent notice.

(2) No less frequently than once every six months, provide written notice to all employees of their rights and responsibilities under this chapter and ensure that such notice is received and acknowledged by all those employees.

(e) (1) A developer and any contractor or subcontractor of the developer shall provide a reasonable internal process through which an employee may anonymously disclose information to the developer if the employee believes in good faith that the information indicates that the developer has violated any provision of Section 22603 or any other law, or has made false or materially misleading statements related to its safety and security protocol, or failed to disclose known risks to employees, including, at a minimum, a monthly update to the disclosing employee regarding the status of the employee's disclosure and the actions taken by the developer, contractor, or subcontractor in response to the disclosure.

(2) The disclosures and responses of the process required by this subdivision shall be maintained for a minimum of seven years from the date when the disclosure or response is created. Each disclosure and response shall be shared with officers and directors of the developer and any contractor or subcontractor of the developer whose acts or omissions are not implicated by the disclosure or response no less frequently than once per quarter.

(f) Nothing in this section shall be construed to limit protections provided to employees by Section 1102.5 of the Labor Code, Section 12964.5 of the Government Code, or other provisions of California law.

(g) As used in this section, the following definitions apply:

(1) "Employee" has the same meaning as defined in Section 1132.4 of the Labor Code and includes both of the following:

(A) Contractors or subcontractors, and unpaid advisors involved with assessing, managing, or addressing hazardous capabilities of covered models.

(B) Corporate officers.

(2) "Contractor or subcontractor" has the same meaning as in Section 1777.1 of the Labor Code.

22608. The duties and obligations imposed by this chapter are cumulative with any other duties or obligations imposed under other law and shall not be construed to relieve any party from any duties or obligations imposed under other law and do not limit any rights or remedies under existing law.

SEC. 4. Section 11547.6 is added to the Government Code, to read:

11547.6. (a) As used in this section, "critical harm" has the same meaning as defined in Section 22602 of the Business and Professions Code.

(b) There is hereby established the Board of Frontier Models. The board shall be housed in the Government Operations Agency and shall be independent of the Department of Technology. The Governor may appoint an executive officer of the board, subject to Senate confirmation, who shall hold the office at the pleasure of the Governor. The executive officer shall be the administrative head of the board and shall exercise all duties and functions necessary to ensure that the responsibilities of the board are successfully discharged.

(c) Commencing January 1, 2026, the Board of Frontier Models shall be composed of five members, as follows:

(1) A member of the open-source community, appointed by the Governor, subject to Senate confirmation.

(2) A member of the artificial intelligence industry, appointed by the Governor, subject to Senate confirmation.

(3) A member of academia, appointed by the Governor, subject to Senate confirmation.

(4) A member appointed by the Speaker of the Assembly.

(5) A member appointed by the Senate Rules Committee.

(d) The Frontier Model Division is hereby created within the Government Operations Agency under the direct supervision of the Board of Frontier Models.

(e) The Frontier Model Division shall do all of the following:

(1) Annually review certification reports received from developers pursuant to Section 22603 of the Business and Professions Code and publicly release summarized findings based on those reports.

(2) Advise the Attorney General on potential violations of this section or Chapter 22.6 (commencing with Section 22602) of Division 8 of the Business and Professions Code.

(3) (A) Issue guidance, standards, and best practices necessary to prevent unreasonable risks of covered models and covered model derivatives causing or enabling critical harms, including, but not limited to, more specific components of or requirements under the duties required under Section 22603 of the Business and Professions Code.

~~(B) Establish an accreditation process and relevant accreditation standards under which third-party auditors may be accredited for a three-year period, which may be extended through an appropriate process, to certify adherence by developers to their requirements under Section 22603 of the Business and Professions Code.~~

(B) Issue guidance regarding best practices for conducting an audit pursuant to subdivision (e) of Section 22603 of the Business and Professions Code.

(4) Publish anonymized artificial intelligence safety incident reports received from developers pursuant to Section 22603 of the Business and Professions Code.

(5) (A) Issue guidance describing the categories of artificial intelligence safety events that are likely to constitute a state of emergency within the meaning of subdivision (b) of Section 8558 and responsive actions that could be ordered by the Governor after a duly proclaimed state of emergency.

(B) The guidance issued pursuant to subparagraph (A) shall not limit, modify, or restrict the authority of the Governor in any way.

(6) Appoint and consult with an advisory committee that shall advise the Governor on when it may be necessary to proclaim a state of emergency relating to artificial intelligence and advise the Governor on what responses may be appropriate in that event.

(7) Appoint and consult with an advisory committee for open-source artificial intelligence that shall do all of the following:

(A) Issue guidelines for model evaluation for use by developers of open-source artificial intelligence models that lack the ability to cause or enable critical harms.

(B) Advise the Legislature on the creation and feasibility of incentives, including tax credits, that could be provided to developers of open-source artificial intelligence models that are not covered models.

(C) Advise the Frontier Model Division on future policies and legislation impacting open-source artificial intelligence development.

(8) Levy fees, including an assessed fee for the submission of a certification, in an amount sufficient to cover the reasonable costs of administering this section that do not exceed the reasonable costs of administering this section.

(9) (A) Develop and submit to the Judicial Council proposed model jury instructions for actions involving violations of Section 22603 of the Business and Professions Code that the Judicial Council may, at its discretion, ~~adopt.~~ *adopt consistent with its policies and procedures for the promulgation of jury instructions.*

(B) In developing the *proposed* model jury instructions required by subparagraph (A), the Frontier Model Division shall consider *and incorporate* all of the following ~~factors:~~ *factors into the proposal that it submits to the Judicial Council:*

~~(i) The level of rigor and detail of the safety and security protocol that the developer faithfully implemented while it trained, stored, and released a covered model.~~

~~(ii) Whether and to what extent the developer's safety and security protocol was inferior, comparable, or superior, in its level of rigor and detail, to the safety and security protocols of comparable developers.~~

~~(iii) The extent and quality of the developer's safety and security protocol's prescribed safeguards, capability testing, and other precautionary measures with respect to the relevant risk of causing or enabling a critical harm.~~

~~(iv) Whether and to what extent the developer and its agents complied with the developer's safety and security protocol, and to the full degree, that doing so might plausibly have avoided causing or enabling a particular harm.~~

~~(v) Whether and to what extent the developer carefully and rigorously investigated, documented, and accurately measured, insofar as reasonably possible given the state-of-the-art, relevant risks that its model might pose.~~

(i) All of the actions that a developer of a covered model must take pursuant to Section 22603 of the Business and Professions Code.

(ii) How any regulations of the Frontier Model Division should be incorporated into the proposed model jury instructions.

(iii) The rigor and quality of the safety and security protocol that a developer is required to implement while training and releasing its artificial intelligence model, and how to determine whether this safety and security protocol was inferior, comparable, or superior to the safety and security protocols of similarly situated developers.

(iv) The rigor and quality of the developer's investigation, documentation, evaluation, and management of its model's potential hazardous capabilities, and associated risks.

(10) (A) On or before January 1, 2027, and annually thereafter, issue regulations to update the definition of a "covered model" to ensure that it accurately reflects technological developments, scientific literature, and widely accepted national and international standards and applies to artificial intelligence models that pose the greatest risk of causing or enabling critical harms. The updated definition shall contain both of the following:

(i) The initial compute threshold that an artificial intelligence model must exceed to be considered a covered model, as defined in Section 22602 of the Business and Professions Code.

(ii) The fine-tuning compute threshold that an artificial intelligence model must meet to be considered a covered model.

(B) In developing regulations pursuant to this paragraph, the Frontier Model Division shall take into account both of the following:

(i) The quantity of computing power used to train covered models that have been identified as being reasonably likely to cause or enable a critical harm.

(ii) Similar thresholds used in federal law, guidance, or regulations for the management of models with reasonable risks of causing or enabling critical harms.

(iii) Input from stakeholders, including academics, industry, and government entities, including from the open-source community.

(11) Every 24 months after initial publication of guidance under paragraphs (3), (5), and (10), review existing guidance in consideration of technological advancements, changes to industry best practices, and information received pursuant to paragraph (1) and update its guidance to the extent appropriate.

(12) On and after January 1, 2026, annually publish the inflation-adjusted dollar amounts described in paragraph (3) of

subdivision (g) and paragraph (2) of subdivision (e) of Section 22602 of the Business and Professions Code.

(f) There is hereby created in the General Fund the Frontier Model Division Programs Fund.

(1) All fees received by the Frontier Model Division pursuant to this section shall be deposited into the fund.

(2) All moneys in the account shall be available, only upon appropriation by the Legislature, for purposes of carrying out the provisions of this section.

SEC. 5. Section 11547.7 is added to the Government Code, to read:

11547.7. (a) The Department of Technology shall commission consultants, pursuant to subdivision (b), to create a public cloud computing cluster, to be known as CalCompute, with the primary focus of conducting research into the safe and secure deployment of large-scale artificial intelligence models and fostering equitable innovation that includes, but is not limited to, all of the following:

(1) A fully owned and hosted cloud platform.

(2) Necessary human expertise to operate and maintain the platform.

(3) Necessary human expertise to support, train, and facilitate use of CalCompute.

(b) The consultants shall include, but not be limited to, representatives of national laboratories, universities, and any relevant professional associations or private sector stakeholders.

(c) To meet the objective of establishing CalCompute, the Department of Technology shall require consultants commissioned to work on this process to evaluate and incorporate all of the following considerations into its plan:

(1) An analysis of the public, private, and nonprofit cloud platform infrastructure ecosystem, including, but not limited to, dominant cloud providers, the relative compute power of each provider, the estimated cost of supporting platforms as well as pricing models, and recommendations on the scope of CalCompute.

(2) The process to establish affiliate and other partnership relationships to establish and maintain an advanced computing infrastructure.

(3) A framework to determine the parameters for use of CalCompute, including, but not limited to, a process for deciding

which projects will be supported by CalCompute and what resources and services will be provided to projects.

(4) A process for evaluating appropriate uses of the public cloud resources and their potential downstream impact, including mitigating downstream harms in deployment.

(5) An evaluation of the landscape of existing computing capability, resources, data, and human expertise in California for the purposes of responding quickly to a security, health, or natural disaster emergency.

(6) An analysis of the state's investment in the training and development of the technology workforce, including through degree programs at the University of California, the California State University, and the California Community Colleges.

(7) A process for evaluating the potential impact of CalCompute on retaining technology professionals in the public workforce.

(d) The Department of Technology shall submit, pursuant to Section 9795, an annual report to the Legislature from the commissioned consultants to ensure progress in meeting the objectives listed above.

(e) The Department of Technology may receive private donations, grants, and local funds, in addition to allocated funding in the annual budget, to effectuate this section.

(f) This section shall become operative only upon an appropriation in a budget act for the purposes of this section.

SEC. 6. The provisions of this act are severable. If any provision of this act or its application is held invalid, that invalidity shall not affect other provisions or applications that can be given effect without the invalid provision or application.

SEC. 7. This act shall be liberally construed to effectuate its purposes.

SEC. 8. No reimbursement is required by this act pursuant to Section 6 of Article XIII B of the California Constitution because the only costs that may be incurred by a local agency or school district will be incurred because this act creates a new crime or infraction, eliminates a crime or infraction, or changes the penalty for a crime or infraction, within the meaning of Section 17556 of the Government Code, or changes the definition of a crime within

the meaning of Section 6 of Article XIII B of the California Constitution.

O