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 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 ~~Systems~~ *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~~ *authorize* a developer of a covered model, as defined, to determine whether ~~it can make a positive safety determination with respect to~~ a covered model *qualifies for a limited duty exemption* before initiating training of that covered model, as specified. The bill would define ~~"positive safety determination"~~ *"limited duty exemption"* to mean a ~~determination~~ *determination, made as specified,* with respect to a covered model, that is not a derivative model, that a developer can reasonably exclude the possibility that the covered model has a hazardous capability, as defined, or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications.

This bill would require that a developer, before initiating training of a nonderivative covered model, comply with various requirements, including implementing the capability to promptly enact a full shutdown of the covered model until that covered model is the subject of a ~~positive safety determination.~~ *limited duty exemption.*

This bill would require a developer of a nonderivative covered model that is not the subject of a ~~positive safety determination~~ *limited duty exemption* to submit to the Frontier Model Division, which the bill would create within the Department of Technology, an annual certification *under penalty of perjury* of compliance with these provisions 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. 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 to report each artificial intelligence safety incident affecting a covered model to the Frontier Model Division in a manner prescribed by the Frontier Model Division.

This bill would require a person that operates a computing cluster, as defined, to implement appropriate 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 deploy 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 also create the Frontier Model Division within the Department of Technology 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 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 model" means an engineered or machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs that can influence physical or virtual environments and that may operate with varying levels of autonomy.

(c) "Artificial intelligence safety incident" means any of the following:

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

user that materially increases the risk of a hazardous capability being used.

(2) Theft, misappropriation, malicious use, inadvertent release, unauthorized access, or escape of the model weights of a covered model that is not the subject of a ~~positive safety determination.~~ *limited duty exemption.*

(3) The critical failure of technical or administrative controls, including controls limiting the ability to modify a covered model that is not the subject of a ~~positive safety determination.~~ *limited duty exemption.*

(4) Unauthorized use of the hazardous capability of a covered model.

(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) "Covered guidance" means ~~any~~ *either* of the following:

(1) ~~Applicable guidance~~ *Guidance* issued by the National Institute of Standards and Technology and by the Frontier Model ~~Division.~~ *Division that is relevant to the management of safety risks associated with artificial intelligence models that may possess hazardous capabilities.*

(2) Industry best practices, including ~~relevant~~ safety practices, precautions, or testing procedures undertaken by developers of comparable ~~models, and any safety standards or best practices commonly or generally recognized by relevant experts in academia or the nonprofit sector.~~ *models that are relevant to the management of safety risks associated with artificial intelligence models that may possess hazardous capabilities.*

~~(3) Applicable safety-enhancing standards set by standards setting organizations.~~

(f) "Covered model" means an artificial intelligence model that meets either of the following criteria:

(1) The artificial intelligence model was trained using a quantity of computing power greater than 10^26 integer or floating-point operations.

(2) The artificial intelligence model was trained using a quantity of computing power sufficiently large that it could reasonably be expected to have similar or greater performance as an artificial

intelligence model trained using a quantity of computing power greater than 10^26 integer or floating-point operations in 2024 as assessed using benchmarks commonly used to quantify the general performance of state-of-the-art foundation models.

(g) "Critical harm" means a harm listed in paragraph (1) of subdivision (n).

(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) (1) "Derivative model" means an artificial intelligence model that is a derivative of another artificial intelligence model, including either of the following:

(A) A modified or unmodified copy of an artificial intelligence model.

(B) A combination of an artificial intelligence model with other software.

(2) "Derivative model" does not include an entirely independently trained artificial intelligence model.

(j) (1) "Developer" means a person that creates, owns, or otherwise has responsibility for an artificial intelligence model.

(2) "Developer" does not include a third-party machine-learning operations platform, an artificial intelligence infrastructure platform, a computing cluster, an application developer using sourced models, or an end-user of an artificial intelligence model.

(k) "Fine tuning" means the adjustment of the model weights of an artificial intelligence model after it has finished its initial training by training the model with new data.

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

(m) "Full shutdown" means the cessation of operation of a covered model, including all copies and derivative models, on all computers and storage devices within custody, control, or possession of a person, including any computer or storage device remotely provided by agreement.

(n) (1) "Hazardous capability" means the capability of a covered model to be used to enable any of the following harms in a way that would be significantly more difficult to cause without access to a covered model:

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

(B) At least five hundred million dollars ($500,000,000) of damage through cyberattacks on critical infrastructure via a single incident or multiple related incidents.

(C) At least five hundred million dollars ($500,000,000) of damage by an artificial intelligence model that autonomously engages in conduct that would violate the Penal Code if undertaken by a human.

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

(2) "Hazardous capability" includes a capability described in paragraph (1) even if the hazardous capability would not manifest but for fine tuning and posttraining modifications performed by third-party experts intending to demonstrate those abilities.

(o) "Limited duty exemption" means an exemption, pursuant to subdivision (a) or (c) of Section 22603, with respect to a covered model that is not a derivative model that a developer can reasonably exclude the possibility that a covered model has a hazardous capability or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications.

~~(o)~~

(p) "Machine-learning operations platform" means a solution that includes a combined offering of necessary machine-learning development capabilities, including exploratory data analysis, data preparation, model training and tuning, model review and governance, model inference and serving, model deployment and monitoring, and automated model retraining.

~~(p)~~

(q) "Model weight" means a numerical parameter established through training in an artificial intelligence model that helps determine how input information impacts a model's output.

~~(q)~~

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

~~(r)~~

(s) "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.

~~(s) "Positive safety determination" means a determination, pursuant to subdivision (a) or (c) of Section 22603, with respect to a covered model that is not a derivative model that a developer can reasonably exclude the possibility that a covered model has a hazardous capability or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications.~~

(t) "Posttraining modification" means the modification of the capabilities of an artificial intelligence model after the completion of training by any means, including, but not limited to, initiating additional training, 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.

(u) "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 enabling or potentially enabling the creation of derivative models.

(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 initiating training of a covered model that is not a derivative model, a developer of that covered model ~~shall~~ *may* determine whether ~~it can make~~ *the covered model qualifies for* a ~~positive safety determination with respect to the covered model.~~ *limited duty exemption.*

(1) In making the determination ~~required~~ *authorized* by this subdivision, a developer shall incorporate all *applicable* covered guidance.

(2) A developer may ~~make a positive safety determination~~ *determine that a covered model qualifies for a limited duty exemption* if the covered model will have lower performance on

all benchmarks relevant under subdivision (f) of Section 22602 and does not have greater general capability than either of the following:

(A) A ~~non-covered~~ *noncovered* model that manifestly lacks hazardous capabilities.

(B) Another model that is the subject of a ~~positive safety determination.~~ *limited duty exemption.*

(3) Upon ~~making a positive safety determination,~~ *determining that a covered model qualifies for a limited duty exemption,* the developer of the covered model shall submit to the Frontier Model Division a certification under penalty of perjury that specifies the basis for that ~~conclusion.~~ *determination.*

(4) A developer that makes a good faith error regarding a ~~positive safety determination~~ *limited duty exemption* shall be deemed to be in compliance with this subdivision if the developer reports its error to the Frontier Model Division within 30 days of completing the training of the covered model and ceases operation of the artificial intelligence model until the developer is otherwise in compliance with subdivision (b).

(b) Before initiating training of a covered model that is not a derivative model ~~that~~ *and* is not the subject of a ~~positive safety determination,~~ *limited duty exemption,* and until that covered model is the subject of a ~~positive safety determination,~~ *limited duty exemption,* the developer of that covered model shall do all of the following:

(1) Implement administrative, technical, and physical cybersecurity protections to prevent unauthorized access to, or misuse or unsafe modification of, the covered model, including to prevent theft, misappropriation, malicious use, or inadvertent release or escape of the model weights from the developer's custody, 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 of the covered model.

(3) Implement all covered guidance.

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

(A) Provides reasonable assurance that if a developer complies with its safety and security protocol, either of the following will apply:

(i) The developer will not produce a covered model with a hazardous capability or enable the production of a derivative model with a hazardous capability.

(ii) The safeguards enumerated in the policy will be sufficient to prevent critical harms from the exercise of a hazardous capability in a covered model.

(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 reasonably exclude the possibility that a covered model has a hazardous capability or may come close to possessing a hazardous capability when accounting for a reasonable margin for safety and the possibility of posttraining modifications, and in addition does all of the following:

(i) Describes in detail how the testing procedure incorporates fine tuning and posttraining modifications performed by third-party experts intending to demonstrate those abilities.

(ii) Describes in detail how the testing procedure incorporates the possibility of posttraining modifications.

(iii) Describes in detail how the testing procedure incorporates the requirement for reasonable margin for safety.

(iv) Describes in detail how the testing procedure addresses the possibility that a covered model can be used to make posttraining modifications or create another covered model in a manner that may generate hazardous capabilities.

~~(iv)~~

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

(D) Describes in detail how the developer will meet requirements listed under paragraphs (1), (2), (3), and (5).

(E) If applicable, describes in detail how the developer intends to implement the safeguards and requirements referenced in paragraph (1) of subdivision (d).

(F) Describes in detail the conditions that would require the execution of a full shutdown.

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

(H) Meets other criteria stated by the Frontier Model Division in guidance to achieve the purpose of maintaining the safety of a covered model with a hazardous capability.

(5) Ensure that the safety and security protocol is implemented as written, including, at a minimum, by designating senior personnel responsible for ensuring implementation by employees and contractors working on a covered model, monitoring and reporting on implementation, and conducting audits, including through third parties as appropriate.

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

(7) 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.

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

(9) Refrain from initiating training of a covered model if there remains an unreasonable risk that an individual, or the covered model itself, may be able to use the hazardous capabilities of the covered model, or a derivative model based on it, to cause a critical harm.

(10) Implement other measures that are reasonably necessary, including in light of applicable guidance from the Frontier Model Division, National Institute of Standards and Technology, and standard-setting organizations, to prevent the development or exercise of hazardous capabilities or to manage the risks arising from them.

(c) (1) Upon completion of the training of a covered model that is not the subject of a ~~positive safety determination~~ *limited duty exemption* and is not a derivative model, the developer shall perform capability testing sufficient to determine whether the developer can make a positive safety determination with respect to the covered model pursuant to its safety and security protocol.

(2) Upon making a ~~positive safety determination~~ *limited duty exemption* with respect to the covered model, a developer of the covered model shall submit to the Frontier Model Division a

certification *under penalty of perjury* of compliance with the requirements of this section within 90 days and no more than 30 days after initiating the commercial, public, or widespread use of the covered model that includes both of the following:

(A) The basis for the developer's ~~positive safety determination.~~ *determination that the covered model qualifies for a limited duty exemption.*

(B) The specific methodology and results of the capability testing undertaken pursuant to this subdivision.

(d) Before initiating the commercial, public, or widespread use of a covered model that is not subject to a ~~positive safety determination,~~ *limited duty exemption,* a developer of the nonderivative version of the covered model shall do all of the following:

(1) Implement reasonable safeguards and requirements to do all of the following:

(A) Prevent an individual from being able to use the hazardous capabilities of the model, or a derivative model, to cause a critical harm.

(B) Prevent an individual from being able to use the model to create a derivative model that was used to cause a critical harm.

(C) Ensure, to the extent reasonably possible, that the covered model's actions and any resulting critical harms can be accurately and reliably attributed to it and any user responsible for those actions.

(2) Provide reasonable requirements to developers of derivative models to prevent an individual from being able to use a derivative model to cause a critical harm.

(3) Refrain from initiating the commercial, public, or widespread use of a covered model if there remains an unreasonable risk that an individual may be able to use the hazardous capabilities of the model, or a derivative model based on it, to cause a critical harm.

(4) Implement other measures that are reasonably necessary, including in light of applicable guidance from the Frontier Model Division, National Institute of Standards and Technology, and standard-setting organizations, to prevent the development or exercise of hazardous capabilities or to manage the risks arising from them.

(e) A developer of a covered model shall periodically reevaluate the procedures, policies, protections, capabilities, and safeguards

implemented pursuant to this section in light of the growing capabilities of covered models and as is reasonably necessary to ensure that the covered model or its users cannot remove or bypass those procedures, policies, protections, capabilities, and safeguards.

(f) (1) A developer of a nonderivative covered model that is not the subject of a ~~positive safety determination~~ *limited duty exemption* shall submit to the Frontier Model Division an annual 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.

(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 hazardous capabilities that the covered model possesses or may reasonably possess and the outcome of capability testing required by subdivision (c).

(B) An assessment of the risk that compliance with the safety and security protocol may be insufficient to prevent harms from the exercise of the covered model's hazardous capabilities.

(C) Other information useful to accomplishing the purposes of this subdivision, as determined by the Frontier Model Division.

(g) A developer shall report each artificial intelligence safety incident affecting a covered model to the Frontier Model Division in a manner prescribed by the Frontier Model Division. The notification shall be made in the most expedient time possible and without unreasonable delay and in no event later than 72 hours after learning that an artificial intelligence safety incident has occurred or learning facts sufficient to establish a reasonable belief that an artificial intelligence safety incident has occurred.

(h) (1) *(A)* Reliance on an unreasonable ~~positive safety determination~~ *limited duty exemption* does not relieve a developer of its obligations under this section.

(B) A determination that a covered model qualifies for a limited duty exemption that results from a good faith error reported pursuant to paragraph (4) of subdivision (a) is not an unreasonable limited duty exemption.

(2) A ~~positive safety determination~~ *limited duty exemption* is unreasonable if the developer does not take into account reasonably

foreseeable risks of harm or weaknesses in capability testing that lead to an inaccurate determination.

(3) A risk of harm or weakness in capability testing is reasonably foreseeable, if, by the time that a developer releases a model, an applicable risk of harm or weakness in capability testing has already been identified by either of the following:

(A) Any other developer of a comparable or comparably powerful model through risk assessment, capability testing, or other means.

(B) By the ~~United States Artificial Intelligence Safety Institute,~~ *National Institute of Standards and Technology,* the Frontier Model Division, or any independent standard-setting organization or capability-testing organization cited by either of those entities.

22604. A person that operates a computing cluster shall implement appropriate 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:

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

(1) The identity of that prospective customer.

(2) 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.

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

(4) The Internet Protocol addresses used for access or administration and the date and time of each access or administrative action.

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

(c) Annually validate the information collected pursuant to subdivision (a) and conduct the assessment required pursuant to subdivision (b).

(d) 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.

(e) Implement the capability to promptly enact a full shutdown in the event of an emergency.

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 ~~has reasonable cause to believe~~ *finds* that a person is violating this chapter, the Attorney General ~~shall commence~~ *may bring* a civil action ~~in a court of competent jurisdiction.~~ *pursuant to this section.*

(b) 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 harm or an imminent risk or threat to public safety.

(2) Other relief as the court deems appropriate, including monetary ~~damages~~ *damages, including punitive damages,* to persons ~~aggrieved~~ *aggrieved, punitive damages,* 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, excluding labor cost, to develop the covered model for a first violation and in an amount not exceeding 30 percent of the cost, excluding labor cost, to develop the covered model for any subsequent violation.

~~(c) In the apportionment of penalties assessed pursuant to this section, defendants shall be jointly and severally liable.~~

~~(d)~~

(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.

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 ~~this section~~ *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.

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:

(1) "Hazardous capability" has the same meaning as defined in Section 22602 of the Business and Professions Code.

(2) ~~"Positive safety determination"~~ *"Limited duty exemption"* has the same meaning as defined in Section 22602 of the Business and Professions Code.

(b) The Frontier Model Division is hereby created within the Department of Technology.

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

(1) Review annual 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 ~~sufficient~~ *necessary* to prevent unreasonable risks from covered models with hazardous capabilities including, but not limited to, more specific

components of or requirements ~~on~~ *under* the duties required under Section 22603 of the Business and Professions Code.

(B) Establish an *optional* accreditation process and relevant accreditation standards under which third parties may be accredited for a three-year period, which may be extended through an appropriate process, to certify adherence by developers to the best practices and standards adopted pursuant to subparagraph (A).

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

(5) Establish confidential fora that are structured and facilitated in a manner that allows developers to share best risk management practices for models with hazardous capabilities in a manner consistent with state and federal antitrust laws.

(6) (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.

(7) 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.

(8) 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 do not have hazardous capabilities.

(B) Advise the Frontier Model Division 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.

(9) Provide technical assistance and advice to the Legislature, upon request, with respect to artificial intelligence-related legislation.

(10) Monitor relevant developments relating to the safety risks associated with the development of artificial intelligence models and the functioning of markets for artificial intelligence models.

(11) 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.

(12) (A) Develop and submit to the Judicial Council proposed model jury instructions for actions ~~brought by individuals injured by a hazardous capability of a covered model.~~ *involving violations of Section 22603 of the Business and Professions Code that the Judicial Council may, at its discretion, adopt.*

(B) In developing the model jury instructions required by subparagraph (A), the Frontier Model Division shall consider all of the following factors:

(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 hazardous capability and related hazardous capabilities.

(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 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,~~ *state-of-the-art,* relevant risks that its model might pose.

(13) (A) On or before July 1, 2026, issue guidance regarding both of the following:

(i) Technical thresholds and benchmarks relevant to determining whether an artificial intelligence model is a covered model, as defined in Section 22602 of the Business and Professions Code.

(ii) Technical thresholds and benchmarks relevant to determining whether a covered model is subject to a limited duty exemption under paragraph (2) of subdivision (a) of Section 22603 of the Business and Professions Code.

(B) In developing guidance 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 having hazardous capabilities or not having hazardous capabilities when accounting for a reasonable margin for safety.

(ii) Similar thresholds used in federal law or regulations for the management of hazardous capabilities.

(14) At least every 24 months after initial publication of guidance under paragraphs (3) and (13), 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.

(d) 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