Proyecto de ley del Senado No. 53
CAPÍTULO 138

[Aprobado por el Gobernador el 29 de septiembre de 2025. Presentado ante el Secretario de Estado el 29 de septiembre de 2025.]

COMPENDIO DEL ASESOR LEGISLATIVO

SB 53, Wiener. Modelos de inteligencia artificial: grandes desarrolladores.

(1) La ley vigente regula en general la inteligencia artificial, incluso exigiendo, el 1 de enero de 2026 o antes, y antes de cada momento posterior, que un sistema o servicio de inteligencia artificial generativa, o una modificación sustancial de un sistema o servicio de inteligencia artificial generativa, publicado el 1 de enero de 2022 o después, se ponga a disposición del público para su uso por parte de los californianos, y que el desarrollador del sistema o servicio publique en su sitio web documentación sobre los datos utilizados por el desarrollador para entrenar el sistema o servicio de inteligencia artificial generativa, según lo prescrito.

Este proyecto de ley promulgaría la Ley de Transparencia en la Inteligencia Artificial Fronteriza (TFAIA), que, entre otras cosas relacionadas con la garantía de la seguridad de un modelo base, según se define, desarrollado por un desarrollador fronterizo, exigiría a un gran desarrollador fronterizo redactar, implementar y publicar de forma clara y visible en su sitio web un marco de IA fronteriza que se aplique a sus modelos fronterizos y describa cómo este aborda, entre otras cosas, la incorporación de normas nacionales, normas internacionales y las mejores prácticas consensuadas por la industria en su marco de IA fronteriza. La TFAIA también exigiría a un gran desarrollador fronterizo transmitir a la Oficina de Servicios de Emergencia un resumen de cualquier evaluación del riesgo catastrófico, según se define, resultante del uso interno de sus modelos fronterizos, según se especifique. La TFAIA requeriría que la Oficina de Servicios de Emergencia establezca un mecanismo que pueda ser utilizado por un desarrollador fronterizo o un miembro del público para informar, según lo prescrito, un incidente de seguridad crítico, según lo definido, y también requeriría que la Oficina de Servicios de Emergencia establezca un mecanismo que pueda ser utilizado por un gran desarrollador fronterizo para presentar de manera confidencial resúmenes de cualquier evaluación del potencial de riesgo catastrófico resultante del uso interno de sus modelos fronterizos, según lo prescrito.

La TFAIA eximiría de la Ley de Registros Públicos de California un informe de un incidente de seguridad crítico presentado a la Oficina de

Servicios de Emergencia, un informe de evaluaciones de riesgo catastrófico por el uso de Internet y un informe de un empleado cubierto realizado de conformidad con las protecciones para denunciantes que se describen a continuación.

La TFAIA impondría una sanción civil por el incumplimiento de la TFAIA, que sería aplicada por el Fiscal General, según lo prescrito.

(2) La legislación vigente establece el Departamento de Tecnología dentro de la Agencia de Operaciones Gubernamentales. Esta legislación exige que el departamento realice, en coordinación con otros organismos interinstitucionales según lo considere oportuno, un inventario exhaustivo de todos los sistemas de toma de decisiones automatizadas de alto riesgo que se hayan propuesto para su uso, desarrollo o adquisición por cualquier agencia estatal, o que estén siendo utilizados, desarrollados o adquiridos por ella.

Este proyecto de ley establecería, dentro de la Agencia de Operaciones Gubernamentales, un consorcio encargado de desarrollar un marco para la creación de un clúster público de computación en la nube, denominado "CalCompute", que impulse el desarrollo y la implementación de inteligencia artificial segura, ética, equitativa y sostenible, fomentando, entre otras cosas, la investigación y la innovación en beneficio del público, según lo prescrito. El proyecto de ley exigiría a la Agencia de Operaciones Gubernamentales que, a más tardar el 1 de enero de 2027, presentara un informe del consorcio a la Legislatura con dicho marco, y lo disolvería tras la presentación de dicho informe. El proyecto de ley haría efectivas estas disposiciones únicamente tras la asignación de fondos en una ley presupuestaria u otra medida para sus fines.

(3) La ley vigente prohíbe a los empleadores y sus agentes crear, adoptar o hacer cumplir una norma, reglamento o política que impida a un empleado revelar información a ciertas entidades o proporcionar información o testificar ante cualquier organismo público que realice una investigación, audiencia o indagación si el empleado tiene causa razonable para creer que la información revela una violación de una ley, según se especifica, y prohíbe las represalias contra un empleado por, entre otras cosas, ejercer estos derechos.

Este proyecto de ley, entre otras cosas relacionadas con la protección de los denunciantes que trabajan con modelos de fundación, prohibiría a un desarrollador fronterizo crear, adoptar, hacer cumplir o celebrar una norma, reglamento, política o contrato que impida a un empleado cubierto, según se define, revelar, o tomar represalias contra un empleado cubierto por revelar, información al Fiscal General, una autoridad federal, una persona con autoridad sobre el empleado cubierto u otro empleado cubierto que tenga autoridad para investigar, descubrir o corregir el problema denunciado, si el

empleado cubierto tiene causa razonable para creer que la información revela que las actividades del desarrollador fronterizo plantean un peligro específico y sustancial para la salud o seguridad pública resultante de un riesgo catastrófico o que el desarrollador fronterizo ha violado la TFAIA.

Este proyecto de ley exigiría a las grandes empresas promotoras de la frontera establecer un proceso interno mediante el cual un empleado sujeto a la ley pueda divulgar información anónimamente a dichas empresas si considera de buena fe que dicha información indica que sus actividades representan un peligro específico y sustancial para la salud o la seguridad pública, como resultado de un riesgo catastrófico, o que han infringido la TFAIA. El proyecto de ley especificaría disposiciones específicas para la aplicación de dichas protecciones a los denunciantes y autorizaría el pago de honorarios de abogados a quienes presenten una demanda exitosa por una infracción.

Este proyecto de ley invalidaría cualquier norma, reglamento, código, ordenanza u otra ley adoptada por una ciudad, condado, ciudad y condado, municipio o agencia local a partir del 1 de enero de 2025, específicamente relacionada con la regulación de los desarrolladores fronterizos con respecto a su gestión del riesgo catastrófico.

Las disposiciones constitucionales existentes requieren que una ley que limite el derecho de acceso a las reuniones de organismos públicos o a los escritos de funcionarios y agencias públicas se adopte con conclusiones que demuestren el interés protegido por la limitación y la necesidad de proteger ese interés.

Este proyecto de ley introduciría conclusiones legislativas a tal efecto.

Clave de resumen

Votación: MAYORÍA Asignación: NO Comité Fiscal: SÍ Programa Local: NO

Texto del proyecto de ley

EL PUEBLO DEL ESTADO DE CALIFORNIA DECRETA LO SIGUIENTE:

SECCIÓN 1.

La Legislatura encuentra y declara todo lo siguiente:

(a) California es líder mundial en innovación e investigación en inteligencia artificial a través de empresas grandes y pequeñas y de las notables universidades públicas y privadas del estado.

(b) La inteligencia artificial, incluidos los nuevos avances en los modelos fundamentales, tiene el potencial de catalizar la innovación y el rápido desarrollo de una amplia gama de beneficios para los californianos y la economía de California, incluidos los avances en medicina, prevención y pronóstico de incendios forestales y ciencia del clima, y de ampliar los límites de la creatividad y la capacidad humanas.

(c) El Grupo de Trabajo Conjunto de Políticas de California sobre Modelos de Frontera de IA ha recomendado principios sólidos para las políticas en inteligencia artificial.

(d) Las intervenciones específicas para apoyar una gobernanza eficaz de la inteligencia artificial deben equilibrar los beneficios de la tecnología y el potencial de riesgos materiales.

(e) Al crear un entorno de evidencia sólido y transparente, los encargados de formular políticas pueden alinear los incentivos para proteger simultáneamente a los consumidores, aprovechar la experiencia de la industria y reconocer las prácticas de seguridad líderes.

(f) A medida que los actores de la industria realizan investigaciones internas sobre los impactos de sus tecnologías, la confianza pública en estas tecnologías se beneficiaría significativamente del acceso a la información sobre las capacidades de la IA de frontera y de un mayor conocimiento de las mismas.

(g) Una mayor transparencia también puede mejorar la rendición de cuentas, la competencia y la confianza pública.

(h) La protección de los denunciantes y el intercambio de información pública son instrumentos clave para aumentar la transparencia.

(i) Los sistemas de notificación de incidentes permiten monitorear los impactos de la inteligencia artificial posteriores a su implementación.

(j) A menos que se desarrollen con diligencia cuidadosa y precauciones razonables, existe la preocupación de que los sistemas avanzados de inteligencia artificial podrían tener capacidades que planteen riesgos catastróficos tanto por usos maliciosos como por fallos de funcionamiento, incluidos ataques informáticos basados en inteligencia artificial, ataques biológicos y pérdida de control.

(k) Dado que la frontera de la inteligencia artificial evoluciona rápidamente, es necesario que haya una legislación que siga la frontera de la investigación en inteligencia artificial y alerte a los responsables de las políticas y al público sobre los graves riesgos y daños que plantean los sistemas de inteligencia artificial más avanzados, evitando al mismo tiempo sobrecargar a las empresas más pequeñas que se encuentran detrás de la frontera.

(l) Si bien los principales desarrolladores de inteligencia artificial ya han establecido voluntariamente la creación, el uso y la publicación de marcos de IA de vanguardia como una buena práctica del sector, no todos proporcionan informes consistentes y suficientes para garantizar la transparencia y la protección necesarias del público. Es necesario que los desarrolladores de vanguardia presenten informes obligatorios, estandarizados y objetivos para proporcionar al gobierno y al público información oportuna y precisa.

(m) La notificación oportuna de incidentes críticos de seguridad al gobierno es esencial para garantizar que las autoridades públicas estén informadas con prontitud de los riesgos actuales y emergentes para la seguridad pública. Esta notificación permite al gobierno monitorear, evaluar y responder eficazmente en caso de que surjan capacidades avanzadas en modelos de inteligencia artificial de vanguardia que puedan representar una amenaza para la ciudadanía.

(n) En el futuro, los modelos de base desarrollados por empresas más pequeñas o que se encuentran detrás de la frontera pueden plantear un riesgo catastrófico significativo, y tal vez se necesite legislación adicional en ese momento.

(o) La reciente publicación del Informe del Gobernador de California sobre la Política de IA de Frontera y el testimonio de las audiencias legislativas sobre inteligencia artificial ante la Legislatura reflejan los avances en las capacidades del modelo de IA que podrían representar un riesgo catastrófico potencial en la inteligencia artificial de frontera, que esta ley pretende abordar.

(p) La intención de la Legislatura es crear más transparencia, pero la seguridad colectiva dependerá en parte de que los desarrolladores fronterizos tengan el debido cuidado en el desarrollo y despliegue de modelos fronterizos proporcionales a la escala de los riesgos previsibles.

SEC. 2.

Se agrega el Capítulo 25.1 (que comienza con la Sección 22757.10) a la División 8 del Código de Negocios y Profesiones, para que se lea:

CAPÍTULO 25.1. Ley de Transparencia en la Inteligencia Artificial Fronteriza

22757. 10.

Este capítulo se conocerá como la Ley de Transparencia en la Inteligencia Artificial Fronteriza.

22757. 11.

Para los efectos de este capítulo:

(a) “Afiliado” significa una persona que controla, es controlada por o está bajo control común con una persona específica, directa o indirectamente, a través de uno o más intermediarios.

(b) “Modelo de inteligencia artificial” significa un sistema diseñado o basado en máquinas que varía en su nivel de autonomía y que puede, para objetivos explícitos o implícitos, inferir a partir de la información que recibe cómo generar resultados que pueden influir en entornos físicos o virtuales.

(c) (1) “Riesgo catastrófico” significa un riesgo previsible y material de que el desarrollo, almacenamiento, uso o despliegue de un modelo fronterizo por parte de un desarrollador fronterizo contribuya materialmente a la muerte o lesiones graves a más de 50 personas o a más de mil millones de dólares ($1,000,000,000) en daños o pérdida de propiedad que surjan de un solo incidente que involucre a un modelo fronterizo que haga cualquiera de las siguientes cosas:

(A) Proporcionar asistencia a nivel de expertos en la creación o liberación de un arma química, biológica, radiológica o nuclear.

(B) Participar en una conducta sin supervisión, intervención o control humano significativo que sea un ciberataque o que, si la conducta hubiera sido cometida por un ser humano, constituiría un delito de asesinato, agresión, extorsión o robo, incluido el robo mediante falsas pretensiones.

(C) Eludir el control de su desarrollador o usuario fronterizo.

(2) “Riesgo catastrófico” no incluye un riesgo previsible y material derivado de cualquiera de los siguientes:

(A) Información que genera un modelo de frontera si dicha información es accesible públicamente en una forma sustancialmente similar desde una fuente distinta a un modelo base.

(B) Actividad lícita del gobierno federal.

(C) Daño causado por un modelo de frontera en combinación con otro software si el modelo de frontera no contribuyó materialmente al daño.

(d) “Incidente crítico de seguridad” significa cualquiera de los siguientes:

(1) Acceso no autorizado, modificación o exfiltración de los pesos del modelo de un modelo de frontera que resulte en muerte o lesiones corporales.

(2) El daño resultante de la materialización de un riesgo catastrófico.

(3) Pérdida de control de un modelo fronterizo que cause muerte o lesiones corporales.

(4) Un modelo de frontera que utiliza técnicas engañosas contra el desarrollador de la frontera para subvertir los controles o la supervisión de su desarrollador de frontera fuera del contexto de una evaluación diseñada para provocar este comportamiento y de una manera que demuestre un riesgo catastrófico materialmente aumentado.

(e) (1) “Implementar” significa poner un modelo de frontera a disposición de un tercero para su uso, modificación, copia o combinación con otro software.

(2) “Implementar” no incluye poner un modelo de frontera a disposición de un tercero con el propósito principal de desarrollar o evaluar dicho modelo.

(f) “Modelo de base” significa un modelo de inteligencia artificial que es todo lo siguiente:

(1) Entrenado en un amplio conjunto de datos.

(2) Diseñado para generalidad de salida.

(3) Adaptable a una amplia gama de tareas distintivas.

(g) “Marco de IA de frontera” significa protocolos técnicos y organizativos documentados para gestionar, evaluar y mitigar riesgos catastróficos.

(h) “Desarrollador de frontera” significa una persona que ha entrenado, o iniciado el entrenamiento de, un modelo de frontera, con respecto al cual la persona ha utilizado, o pretende utilizar, al menos tanta potencia de cómputo para entrenar el modelo de frontera como para cumplir con las especificaciones técnicas que se encuentran en la subdivisión (i).

(i) (1) “Modelo de frontera” significa un modelo base que fue entrenado utilizando una cantidad de potencia computacional mayor a 10^26 operaciones de números enteros o de punto flotante.

(2) La cantidad de potencia de cómputo descrita en el párrafo (1) incluirá el cómputo para la ejecución de entrenamiento original y para cualquier ajuste fino posterior, aprendizaje de refuerzo u otras modificaciones materiales que el desarrollador aplique a un modelo de base anterior.

(j) “Gran desarrollador fronterizo” significa un desarrollador fronterizo que junto con sus afiliados colectivamente tuvo ingresos brutos anuales superiores a quinientos millones de dólares ($500,000,000) en el año calendario anterior.

(k) “Peso del modelo” significa un parámetro numérico en un modelo de frontera que se ajusta a través del entrenamiento y que ayuda a determinar cómo las entradas se transforman en salidas.

(l) “Propiedad” significa propiedad tangible o intangible.

22757. 12.

(a) Un gran desarrollador de fronteras deberá redactar, implementar, cumplir y publicar de forma clara y visible en su sitio web de Internet un marco de IA de fronteras que se aplique a los modelos de fronteras del gran desarrollador de fronteras y describa cómo el gran desarrollador de fronteras aborda todo lo siguiente:

(1) Incorporar estándares nacionales, estándares internacionales y mejores prácticas de consenso de la industria en su marco de IA de frontera.

(2) Definir y evaluar los umbrales utilizados por el gran desarrollador de fronteras para identificar y evaluar si un modelo de frontera tiene capacidades que podrían representar un riesgo catastrófico, lo que puede incluir umbrales de múltiples niveles.

(3) Aplicar mitigaciones para abordar el potencial de riesgos catastróficos basándose en los resultados de las evaluaciones realizadas de conformidad con el párrafo (2).

(4) Revisar las evaluaciones y la idoneidad de las mitigaciones como parte de la decisión de implementar un modelo de frontera o utilizarlo ampliamente a nivel interno.

(5) Utilizar a terceros para evaluar el potencial de riesgos catastróficos y la eficacia de las mitigaciones de dichos riesgos.

(6) Revisar y actualizar el marco de IA de frontera, incluidos los criterios que desencadenan actualizaciones y cómo el gran desarrollador de frontera determina cuándo sus modelos de frontera se modifican lo suficiente como para requerir divulgaciones de conformidad con la subdivisión (c).

(7) Prácticas de ciberseguridad para proteger los pesos de modelos no publicados de modificaciones o transferencias no autorizadas por parte de partes internas o externas.

(8) Identificar y responder ante incidentes críticos de seguridad.

(9) Instituir prácticas de gobernanza interna para garantizar la implementación de estos procesos.

(10) Evaluar y gestionar el riesgo catastrófico resultante del uso interno de sus modelos de frontera, incluidos los riesgos resultantes de que un modelo de frontera eluda los mecanismos de supervisión.

(b) (1) Un gran desarrollador de fronteras deberá revisar y, según corresponda, actualizar su marco de IA de fronteras al menos una vez al año.

(2) Si un gran desarrollador fronterizo realiza una modificación material a su marco de IA fronteriza, el gran desarrollador fronterizo deberá publicar de manera clara y visible el marco de IA fronterizo modificado y una justificación para esa modificación dentro de los 30 días.

(c) (1) Antes o simultáneamente con la implementación de un nuevo modelo de frontera o una versión sustancialmente modificada de un modelo de frontera existente, un desarrollador de frontera deberá publicar de manera clara y visible en su sitio web de Internet un informe de transparencia que contenga todo lo siguiente:

(A) El sitio web de Internet del desarrollador fronterizo.

(B) Un mecanismo que permite a una persona física comunicarse con el desarrollador fronterizo.

(C) La fecha de lanzamiento del modelo fronterizo.

(D) Los idiomas soportados por el modelo de frontera.

(E) Las modalidades de producción soportadas por el modelo de frontera.

(F) Los usos previstos del modelo de frontera.

(G) Cualquier restricción o condición de aplicación general sobre los usos del modelo de frontera.

(2) Antes o simultáneamente con la implementación de un nuevo modelo de frontera o una versión sustancialmente modificada de un modelo de frontera existente, un gran desarrollador de frontera deberá incluir en el informe de transparencia requerido por el párrafo (1) resúmenes de todo lo siguiente:

(A) Evaluaciones de riesgos catastróficos del modelo de frontera realizadas de conformidad con el marco de IA de frontera del gran desarrollador de fronteras.

(B) Los resultados de dichas evaluaciones.

(C) El grado en que estuvieron involucrados evaluadores externos.

(D) Otras medidas adoptadas para cumplir los requisitos del marco de IA de frontera con respecto al modelo de frontera.

(3) Un desarrollador fronterizo que publique la información descrita en el párrafo (1) o (2) como parte de un documento más grande, incluida una tarjeta de sistema o una tarjeta modelo, se considerará que cumple con el párrafo aplicable.

(4) Se alienta a un desarrollador fronterizo, pero no se le exige, a realizar divulgaciones descritas en esta subdivisión que sean consistentes con, o superiores a, las mejores prácticas de la industria.

(d) Un gran desarrollador fronterizo deberá transmitir a la Oficina de Servicios de Emergencia un resumen de cualquier evaluación de riesgo catastrófico resultante del uso interno de sus modelos fronterizos cada tres meses o de conformidad con otro cronograma razonable especificado por el gran desarrollador fronterizo y comunicado por escrito a la Oficina de Servicios de Emergencia con actualizaciones escritas, según corresponda.

(e) (1) (A) Un desarrollador fronterizo no deberá hacer una declaración materialmente falsa o engañosa acerca del riesgo catastrófico a partir de sus modelos fronterizos o su gestión del riesgo catastrófico.

(B) Un gran desarrollador fronterizo no deberá hacer una declaración materialmente falsa o engañosa sobre su implementación o cumplimiento de su marco de IA fronteriza.

(2) Esta subdivisión no se aplica a una declaración que se hizo de buena fe y que era razonable en las circunstancias.

(f) (1) Cuando un desarrollador fronterizo publique documentos para cumplir con esta sección, el desarrollador fronterizo puede hacer redacciones en dichos documentos que sean necesarias para proteger los secretos comerciales del desarrollador fronterizo, la ciberseguridad del desarrollador fronterizo, la seguridad pública o la seguridad nacional de los Estados Unidos o para cumplir con cualquier ley federal o estatal.

(2) Si un desarrollador fronterizo redacta información en un documento de conformidad con esta subdivisión, el desarrollador fronterizo deberá describir el carácter y la justificación de la redacción en cualquier versión publicada del documento en la medida permitida por las preocupaciones que justifican la redacción y deberá conservar la información no redactada durante cinco años.

(a) La Oficina de Servicios de Emergencia establecerá un mecanismo que será utilizado por un desarrollador fronterizo o un miembro del público para informar un incidente de seguridad crítico que incluya todo lo siguiente:

(1) La fecha del incidente crítico de seguridad.

(2) Las razones por las que el incidente califica como un incidente de seguridad crítico.

(3) Una declaración breve y sencilla que describa el incidente de seguridad crítico.

(4) Si el incidente estuvo relacionado con el uso interno de un modelo de frontera.

(b) (1) La Oficina de Servicios de Emergencia establecerá un mecanismo que será utilizado por un gran desarrollador fronterizo para presentar de manera confidencial resúmenes de cualquier evaluación del potencial de riesgo catastrófico resultante del uso interno de sus modelos fronterizos.

(2) La Oficina de Servicios de Emergencia tomará todas las precauciones necesarias para limitar el acceso a cualquier informe relacionado con el uso interno de los modelos fronterizos únicamente al personal con una necesidad específica de conocer la información y para proteger los informes del acceso no autorizado.

(c) (1) Sujeto al párrafo (2), un desarrollador fronterizo deberá informar cualquier incidente crítico de seguridad relacionado con uno o más de sus modelos fronterizos a la Oficina de Servicios de Emergencia dentro de los 15 días de descubrir el incidente crítico de seguridad.

(2) Si un desarrollador fronterizo descubre que un incidente de seguridad crítico representa un riesgo inminente de muerte o lesiones físicas graves, el desarrollador fronterizo deberá revelar ese incidente dentro de las 24 horas a una autoridad, incluida cualquier agencia de aplicación de la ley o agencia de seguridad pública con jurisdicción, que sea apropiada en función de la naturaleza de ese incidente y según lo requiera la ley.

(3) Un desarrollador fronterizo que descubra información sobre un incidente de seguridad crítico después de presentar el informe inicial requerido por esta subdivisión puede presentar un informe enmendado.

(4) Se alienta a los desarrolladores de fronteras, pero no se les exige, a informar sobre incidentes de seguridad críticos relacionados con modelos de cimentación que no sean modelos de fronteras.

(d) La Oficina de Servicios de Emergencia revisará los informes de incidentes de seguridad críticos presentados por los desarrolladores

fronterizos y podrá revisar los informes presentados por miembros del público.

(e) (1) El Procurador General o la Oficina de Servicios de Emergencia pueden transmitir informes de incidentes críticos de seguridad e informes de empleados cubiertos realizados de conformidad con el Capítulo 5.1 (comenzando con la Sección 1107) de la Parte 3 de la División 2 del Código Laboral a la Legislatura, el Gobernador, el gobierno federal o las agencias estatales correspondientes.

(2) El Procurador General o la Oficina de Servicios de Emergencia deberán considerar seriamente cualquier riesgo relacionado con secretos comerciales, seguridad pública, ciberseguridad de un desarrollador fronterizo o seguridad nacional al transmitir informes.

(f) Un informe de un incidente crítico de seguridad presentado a la Oficina de Servicios de Emergencia de conformidad con esta sección, un informe de evaluaciones de riesgo catastrófico de uso interno de conformidad con la Sección 22757.12 y un informe de empleado cubierto realizado de conformidad con el Capítulo 5.1 (que comienza con la Sección 1107) de la Parte 3 de la División 2 del Código Laboral están exentos de la Ley de Registros Públicos de California (División 10 (que comienza con la Sección 7920.000) del Título 1 del Código de Gobierno).

(g) (1) A partir del 1 de enero de 2027, y anualmente a partir de entonces, la Oficina de Servicios de Emergencia producirá un informe con información anónima y agregada sobre los incidentes de seguridad críticos que haya revisado la Oficina de Servicios de Emergencia desde el informe anterior.

(2) La Oficina de Servicios de Emergencia no incluirá información en un informe conforme a esta subdivisión que pueda comprometer los secretos comerciales o la ciberseguridad de un desarrollador fronterizo, la seguridad pública o la seguridad nacional de los Estados Unidos o que esté prohibida por cualquier ley federal o estatal.

(3) La Oficina de Servicios de Emergencia transmitirá un informe de conformidad con esta subdivisión a la Legislatura, de conformidad con la Sección 9795, y al Gobernador.

(h) La Oficina de Servicios de Emergencia puede adoptar reglamentos que designen una o más leyes, reglamentos o documentos de orientación federales que cumplan con todas las siguientes condiciones para los fines de la subdivisión (i):

(1) (A) La ley, el reglamento o el documento de orientación impone o establece normas o requisitos para los informes de incidentes de seguridad

críticos que son sustancialmente equivalentes o más estrictos que los requeridos por esta sección.

(B) La ley, reglamento o documento de orientación descrito en el subpárrafo (A) no necesita exigir que se informe al Estado de California sobre incidentes de seguridad críticos.

(2) La ley, el reglamento o el documento de orientación tiene por objeto evaluar, detectar o mitigar el riesgo catastrófico.

(i) (1) Un desarrollador fronterizo que tenga la intención de cumplir con esta sección cumpliendo con los requisitos o cumpliendo con los estándares establecidos por una ley, reglamento o documento de orientación federal designado de conformidad con la subdivisión (h) deberá declarar su intención de hacerlo a la Oficina de Servicios de Emergencia.

(2) Después de que un desarrollador fronterizo haya declarado su intención de conformidad con el párrafo (1), se aplican ambas de las siguientes condiciones:

(A) Se considerará que el desarrollador fronterizo cumple con esta sección en la medida en que cumpla con los estándares o los requisitos impuestos o establecidos por la ley, regulación o documento de orientación federal designado hasta que el desarrollador fronterizo declare la revocación de esa intención a la Oficina de Servicios de Emergencia o la Oficina de Servicios de Emergencia revoque una regulación relevante de conformidad con la subdivisión (j).

(B) El incumplimiento por parte de un desarrollador fronterizo de los estándares o de los requisitos establecidos por la ley federal, la reglamentación o el documento de orientación designado de conformidad con la subdivisión (h) constituirá una violación de este capítulo.

(j) La Oficina de Servicios de Emergencia revocará un reglamento adoptado conforme a la subdivisión (h) si ya no se cumplen los requisitos de la subdivisión (h).

22757. 14.

(a) El 1 de enero de 2027 o antes, y anualmente a partir de entonces, el Departamento de Tecnología evaluará la evidencia y los desarrollos recientes relevantes para los propósitos de este capítulo y hará recomendaciones sobre si se debe actualizar, y cómo hacerlo, cualquiera de las siguientes definiciones para los propósitos de este capítulo para asegurar que reflejen con precisión los desarrollos tecnológicos, la literatura científica y los estándares nacionales e internacionales ampliamente aceptados:

(1) “Modelo de frontera” para que se aplique a los modelos fundamentales en la frontera del desarrollo de la inteligencia artificial.

(2) “Desarrollador de frontera”, de modo que se aplica a los desarrolladores de modelos de frontera que están en la frontera del desarrollo de la inteligencia artificial.

(3) “Gran desarrollador fronterizo”, de modo que se aplique a los desarrolladores fronterizos con buenos recursos.

(b) Al hacer recomendaciones de conformidad con esta sección, el Departamento de Tecnología tendrá en cuenta todo lo siguiente:

(1) Se utilizan umbrales similares en normas internacionales o leyes, directrices o reglamentos federales para la gestión del riesgo catastrófico y deberán alinearse con una definición adoptada en una ley o reglamento federal en la medida en que sea consistente con los propósitos de este capítulo.

(2) Aportes de las partes interesadas, incluidos académicos, la industria, la comunidad de código abierto y entidades gubernamentales.

(3) El grado en que una persona podrá determinar, antes de comenzar a entrenar o implementar un modelo de base, si esa persona estará sujeta a la definición de desarrollador fronterizo o de gran desarrollador fronterizo con el objetivo de permitir determinaciones más tempranas, si es posible.

(4) La complejidad de determinar si un modelo de persona o fundación está cubierto, con el objetivo de permitir determinaciones más simples si es posible.

(5) La verificabilidad externa de determinar si un modelo de persona o fundación está cubierto, con el objetivo de lograr definiciones que sean verificables por partes distintas del desarrollador fronterizo.

(c) Al desarrollar recomendaciones de conformidad con esta sección, el Departamento de Tecnología presentará un informe a la Legislatura, de conformidad con la Sección 9795 del Código de Gobierno, con dichas recomendaciones.

(d) (1) A partir del 1 de enero de 2027, y anualmente a partir de entonces, el Procurador General elaborará un informe con información anónima y agregada sobre los informes de los empleados cubiertos realizados de conformidad con el Capítulo 5.1 (comenzando con la Sección 1107) de la Parte 3 de la División 2 del Código Laboral que hayan sido revisados por el Procurador General desde el informe anterior.

(2) El Fiscal General no incluirá información en un informe conforme a esta subdivisión que comprometa los secretos comerciales o la ciberseguridad de un desarrollador fronterizo, la confidencialidad de un empleado cubierto, la seguridad pública o la seguridad nacional de los Estados Unidos o que esté prohibida por cualquier ley federal o estatal.

(3) El Procurador General transmitirá un informe de conformidad con esta subdivisión a la Legislatura, de conformidad con la Sección 9795 del Código de Gobierno, y al Gobernador.

22757. 15.

(a) Un gran desarrollador fronterizo que no publique o transmita un documento conforme que debe publicarse o transmitirse bajo este capítulo, haga una declaración en violación de la subdivisión (e) de la Sección 22757.12, no informe un incidente como lo requiere la Sección 22757.13, o no cumpla con su propio marco de IA fronterizo estará sujeto a una multa civil por un monto que dependerá de la gravedad de la violación y que no exceda un millón de dólares ($1,000,000) por violación.

(b) Una sanción civil descrita en esta sección se recuperará en una acción civil interpuesta únicamente por el Fiscal General.

22757. 16.

La pérdida de valor del patrimonio no se considera daño o pérdida de propiedad para los fines de este capítulo.

SEC. 3.

Se agrega la Sección 11546.8 al Código de Gobierno, para que se lea como sigue:

11546. 8.

(a) Por la presente se establece dentro de la Agencia de Operaciones Gubernamentales un consorcio que desarrollará, de conformidad con esta sección, un marco para la creación de un clúster de computación en la nube pública que se conocerá como “CalCompute”.

(b) El consorcio desarrollará un marco para la creación de CalCompute que promueva el desarrollo y la implementación de inteligencia artificial que sea segura, ética, equitativa y sostenible al hacer, como mínimo, ambas de las siguientes cosas:

(1) Fomentar la investigación y la innovación que beneficien al público.

(2) Permitir la innovación equitativa ampliando el acceso a los recursos computacionales.

(c) El consorcio realizará esfuerzos razonables para garantizar que CalCompute se establezca dentro de la Universidad de California en la medida de lo posible.

(d) CalCompute incluirá, pero no se limitará a, todo lo siguiente:

(1) Una plataforma en la nube totalmente propia y alojada.

(2) Experiencia humana necesaria para operar y mantener la plataforma.

(3) Experiencia humana necesaria para apoyar, capacitar y facilitar el uso de CalCompute.

(e) El consorcio deberá operar de conformidad con todas las leyes y normas laborales y de fuerza laboral pertinentes.

(f) (1) El 1 de enero de 2027 o antes, la Agencia de Operaciones del Gobierno deberá presentar, de conformidad con la Sección 9795, un informe del consorcio a la Legislatura con el marco desarrollado de conformidad con la subdivisión (b) para la creación y operación de CalCompute.

(2) El informe requerido por esta subdivisión deberá incluir todos los elementos siguientes:

(A) Un análisis del panorama de la infraestructura actual de la plataforma de computación en la nube pública, privada y sin fines de lucro de California.

(B) Un análisis del costo para el estado de construir y mantener CalCompute y recomendaciones para posibles fuentes de financiamiento.

(C) Recomendaciones para la estructura de gobernanza y el funcionamiento continuo de CalCompute.

(D) Recomendaciones para los parámetros de uso de CalCompute, incluyendo, entre otros, un proceso para determinar qué usuarios y proyectos recibirán soporte de CalCompute.

(E) Un análisis de la fuerza laboral tecnológica del estado y recomendaciones de vías equitativas para fortalecer la fuerza laboral, incluido el papel de CalCompute.

(F) Una descripción detallada de cualquier asociación, contrato o acuerdo de licencia propuesto con entidades no gubernamentales, incluidas, entre otras, empresas basadas en tecnología, que demuestre el cumplimiento de los requisitos de las subdivisiones (c) y (d).

(G) Recomendaciones sobre cómo la creación y gestión continua de CalCompute pueden priorizar el uso de la fuerza laboral actual del sector público.

(g) El consorcio, de conformidad con la ley constitucional estatal, estará integrado por 14 miembros de la siguiente manera:

(1) Cuatro representantes de la Universidad de California y otras instituciones públicas y privadas de investigación académica y laboratorios nacionales designados por el Secretario de Operaciones del Gobierno.

(2) Tres representantes de las organizaciones laborales afectadas, designados por el Presidente de la Asamblea.

(3) Tres representantes de grupos de partes interesadas con experiencia y conocimientos relevantes, incluidos, entre otros, especialistas en ética, defensores de los derechos del consumidor y otros defensores del interés público designados por el Comité de Reglas del Senado.

(4) Cuatro expertos en tecnología e inteligencia artificial para brindar asistencia técnica designados por el Secretario de Operaciones de Gobierno.

(h) Los miembros del consorcio prestarán sus servicios sin remuneración, pero serán reembolsados por todos los gastos necesarios en que efectivamente incurran en el desempeño de sus funciones.

(i) El consorcio se disolverá al presentarse a la Legislatura el informe requerido por el párrafo (1) de la subdivisión (f).

(j) Si CalCompute se establece dentro de la Universidad de California, la Universidad de California puede recibir donaciones privadas con el fin de implementar CalCompute.

(k) Esta sección entrará en vigor únicamente mediante una asignación en una ley de presupuesto u otra medida para los fines de esta sección.

SEC. 4.

Se agrega el Capítulo 5.1 (que comienza con la Sección 1107) a la Parte 3 de la División 2 del Código del Trabajo, para que se lea como sigue:

CAPÍTULO 5.1. Protección de los denunciantes: Riesgos catastróficos en los modelos de la Fundación IA

1107.

Para los efectos de este capítulo:

(a) (1) “Riesgo catastrófico” significa un riesgo previsible y material de que el desarrollo, almacenamiento, uso o despliegue de un modelo de cimentación por parte de un desarrollador fronterizo contribuya materialmente a la muerte o lesiones graves a más de 50 personas o a más de mil millones de dólares ($1,000,000,000) en daños o pérdida de propiedad que surjan de un solo incidente que involucre a un modelo de cimentación que haga cualquiera de las siguientes cosas:

(A) Proporcionar asistencia a nivel de expertos en la creación o liberación de un arma química, biológica, radiológica o nuclear.

(B) Participar en una conducta sin supervisión, intervención o control humano significativo que sea un ciberataque o que, si fuera cometida por un ser humano, constituiría un delito de asesinato, agresión, extorsión o robo, incluido el robo mediante falsas pretensiones.

(C) Eludir el control de su desarrollador o usuario fronterizo.

(2) “Riesgo catastrófico” no incluye un riesgo previsible y material derivado de cualquiera de los siguientes:

(A) Información que genera un modelo de base si dicha información es accesible públicamente en una forma sustancialmente similar a partir de una fuente distinta a un modelo de base.

(B) Actividad lícita del gobierno federal.

(C) Daño causado por un modelo de base en combinación con otro software cuando el modelo de base no contribuyó materialmente al daño.

(b) “Empleado cubierto” significa un empleado responsable de evaluar, gestionar o abordar el riesgo de incidentes críticos de seguridad.

(c) “Incidente crítico de seguridad” significa cualquiera de los siguientes:

(1) Acceso no autorizado, modificación o exfiltración de los pesos del modelo de un modelo de cimentación que resulte en muerte, lesiones corporales o daños o pérdida de propiedad.

(2) El daño resultante de la materialización de un riesgo catastrófico.

(3) Pérdida de control de un modelo de base que cause muerte o lesiones corporales.

(4) Un modelo de fundación que utiliza técnicas engañosas contra el desarrollador fronterizo para subvertir los controles o la supervisión de su desarrollador fronterizo fuera del contexto de una evaluación diseñada para provocar este comportamiento y de una manera que demuestre un riesgo catastrófico materialmente aumentado.

(d) “Modelo de fundación” tiene el significado definido en la Sección 22757.11 del Código de Negocios y Profesiones.

(e) “Desarrollador fronterizo” tiene el significado definido en la Sección 22757.11 del Código de Negocios y Profesiones.

(f) “Gran desarrollador fronterizo” tiene el significado definido en la Sección 22757.11 del Código de Negocios y Profesiones.

1107. 1.

(a) Un desarrollador fronterizo no deberá crear, adoptar, hacer cumplir ni celebrar una norma, reglamento, política o contrato que impida a un empleado cubierto divulgar, o tomar represalias contra un empleado cubierto por divulgar, información al Fiscal General, una autoridad federal, una persona con autoridad sobre el empleado cubierto u otro empleado cubierto que tenga autoridad para investigar, descubrir o corregir el problema denunciado, si el empleado cubierto tiene causa razonable para creer que la información divulga cualquiera de los siguientes:

(1) Las actividades del desarrollador fronterizo plantean un peligro específico y sustancial para la salud o seguridad pública resultante de un riesgo catastrófico.

(2) El desarrollador fronterizo ha violado el Capítulo 25.1 (que comienza con la Sección 22757.10) de la División 8 del Código de Negocios y Profesiones.

(b) Un desarrollador fronterizo no podrá celebrar un contrato que impida a un empleado cubierto realizar una divulgación protegida bajo la Sección 1102.5.

(c) Un empleado cubierto puede utilizar la línea directa descrita en la Sección 1102.7 para realizar los informes descritos en la subdivisión (a).

(d) Un desarrollador fronterizo deberá proporcionar un aviso claro a todos los empleados cubiertos sobre sus derechos y responsabilidades bajo esta sección, incluso haciendo cualquiera de las siguientes cosas:

(1) Publicar y exhibir en todo momento dentro de cualquier lugar de trabajo mantenido por el desarrollador fronterizo un aviso a todos los empleados cubiertos sobre sus derechos bajo esta sección, asegurando que cualquier nuevo empleado cubierto reciba un aviso equivalente y asegurando que cualquier empleado cubierto que trabaje de forma remota reciba periódicamente un aviso equivalente.

(2) Al menos una vez al año, proporcionar aviso escrito a cada empleado cubierto sobre sus derechos bajo esta sección y garantizar que el aviso sea recibido y reconocido por todos esos empleados cubiertos.

(e) (1) Un gran desarrollador fronterizo deberá proporcionar un proceso interno razonable a través del cual un empleado cubierto pueda divulgar información de manera anónima al gran desarrollador fronterizo si el empleado cubierto cree de buena fe que la información indica que las actividades del gran desarrollador fronterizo presentan un peligro específico y sustancial para la salud o seguridad pública como resultado de un riesgo catastrófico o que el gran desarrollador fronterizo violó el Capítulo 25.1 (que comienza con la Sección 22757.10) de la División 8 del Código de Negocios y Profesiones, incluida una actualización mensual para la persona que hizo la divulgación sobre el estado de la investigación del gran desarrollador fronterizo sobre la divulgación y las acciones tomadas por el gran desarrollador fronterizo en respuesta a la divulgación.

(2) (A) Salvo lo dispuesto en el subpárrafo (B), las divulgaciones y respuestas del proceso requerido por esta subdivisión se compartirán con los funcionarios y directores del gran desarrollador fronterizo al menos una vez cada trimestre.

(B) Si un empleado cubierto ha alegado una irregularidad por parte de un funcionario o director del gran desarrollador fronterizo en una divulgación o respuesta, el subpárrafo (A) no se aplicará con respecto a ese funcionario o director.

(f) El tribunal está autorizado a otorgar honorarios de abogados razonables a un demandante que presente una acción exitosa por una violación de esta sección.

(g) En una acción civil interpuesta de conformidad con esta sección, una vez que se haya demostrado por una preponderancia de la evidencia que una actividad proscrita por esta sección fue un factor contribuyente en la supuesta acción prohibida contra el empleado cubierto, el desarrollador fronterizo tendrá la carga de la prueba para demostrar mediante evidencia clara y convincente que la supuesta acción habría ocurrido por razones legítimas e independientes incluso si el empleado cubierto no hubiera participado en actividades protegidas por esta sección.

(h) (1) En una acción civil o procedimiento administrativo interpuesto de conformidad con esta sección, un empleado cubierto puede solicitar al tribunal superior de cualquier condado en el que se alega que ocurrió la violación en cuestión, o en el que la persona reside o realiza negocios, una medida cautelar provisional o preliminar apropiada.

(2) Al presentarse la petición de medida cautelar, el peticionario deberá notificarla a la persona, y entonces el tribunal tendrá jurisdicción para conceder la medida cautelar temporal que considere justa y apropiada.

(3) Además de cualquier daño que resulte directamente de una violación de esta sección, el tribunal considerará el efecto paralizante sobre otros empleados cubiertos que hagan valer sus derechos bajo esta sección al determinar si la medida cautelar temporal es justa y apropiada.

(4) Se dictará la medida cautelar adecuada si se demuestra que existe causa razonable para creer que se ha producido una violación.

(5) Una orden que autorice una medida cautelar temporal permanecerá vigente hasta que se emita una determinación administrativa o judicial, o una citación, o hasta que se concluya una revisión conforme a la subdivisión (b) del Artículo 98.74, lo que sea más largo, o en un plazo determinado por el tribunal. Posteriormente, se podrá emitir una medida cautelar preliminar o permanente si se demuestra que es justa y adecuada. Ninguna medida cautelar temporal impedirá a una empresa desarrolladora fronteriza disciplinar o despedir a un empleado cubierto por conducta no relacionada con la reclamación de la represalia.

(i) No obstante lo dispuesto en la Sección 916 del Código de Procedimiento Civil, la medida cautelar concedida de conformidad con esta sección no se suspenderá en espera de la apelación.

(j) (1) Esta sección no afecta ni limita la aplicabilidad de la Sección 1102.5, incluso con respecto a los derechos de los empleados que no están cubiertos a denunciar violaciones de este capítulo o del Capítulo 25.1 (que comienza con la Sección 22757.10) de la División 8 del Código de Negocios y Profesiones.

(2) Los recursos previstos en esta sección son acumulativos entre sí y con los recursos o sanciones disponibles conforme a todas las demás leyes de este estado.

1107. 2.

La pérdida de valor del patrimonio no se considera daño o pérdida de propiedad para los fines de este capítulo.

SEC. 5.

a) Las disposiciones de esta ley son separables. Si alguna disposición de esta ley o su aplicación se declara inválida, dicha invalidez no afectará a otras disposiciones o aplicaciones que puedan surtir efecto sin la disposición o aplicación inválida.

(b) Esta ley se interpretará liberalmente para lograr sus fines.

(c) Los deberes y obligaciones impuestos por esta ley son acumulativos con cualesquiera otros deberes u obligaciones impuestos bajo otras leyes y no se interpretarán como que eximen a ninguna parte de cualesquiera deberes u obligaciones impuestos bajo otras leyes y no limitan ningún derecho o recurso bajo la ley existente.

(d) Esta ley no se aplicará en la medida en que entre estrictamente en conflicto con los términos de un contrato entre una entidad del gobierno federal y un desarrollador fronterizo.

(e) La presente ley no se aplicará en la medida en que esté sustituida por la ley federal.

(f) Esta ley prevalece sobre cualquier norma, reglamento, código, ordenanza u otra ley adoptada por una ciudad, condado, ciudad y condado, municipio o agencia local a partir del 1 de enero de 2025, específicamente relacionada con la regulación de los desarrolladores fronterizos con respecto a su gestión del riesgo catastrófico.

SEC. 6.

La Legislatura determina y declara que la Sección 2 de esta ley, que añade el Capítulo 25.1 (a partir de la Sección 22757.10) a la División 8 del Código de Negocios y Profesiones, impone una limitación al derecho del público de acceder a las reuniones de organismos públicos o a los escritos de funcionarios y agencias públicas, en el sentido de la Sección 3 del Artículo I de la Constitución de California. De conformidad con dicha disposición constitucional, la Legislatura realiza las siguientes conclusiones para demostrar el interés protegido por esta limitación y la necesidad de protegerlo:

La información contenida en informes de incidentes de seguridad críticos, evaluaciones de riesgos de uso interno e informes de empleados cubiertos puede contener información que podría amenazar la seguridad pública o comprometer la respuesta a un incidente si se divulga al público.