STATE OF NEW YORK

6453

2025-2026 Regular Sessions

IN ASSEMBLY

March 5, 2025

Introduced by M. of A. BORES -- read once and referred to the Committee on Science and Technology

AN ACT to amend the general business law, in relation to the training and use of artificial intelligence frontier models

The People of the State of New York, represented in Senate and Assembly, do enact as follows:

Section 1. Short title. This act shall be known and may be cited as the "Responsible AI safety and education act" or "RAISE act".

§ 2. The general business law is amended by adding a new article 44-B to read as follows:

ARTICLE 44-B

RESPONSIBLE AI SAFETY AND EDUCATION (RAISE) ACT

§ 1420. Definitions. As used in this article, the following terms shall have the following meanings:

1. "Appropriate redactions" means redactions to a safety and security protocol or audit report that are reasonably necessary to protect any of the following:

(a) Public safety or United States national security;

(b) Trade secrets; or

(c) Confidential information pursuant to state and federal law.

2. "Artificial intelligence" means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments, and that uses machine- and human-based inputs to perceive real and virtual envi-

EXPLANATION--Matter in *italics* (underscored) is new; matter in brackets [-] is old law to be omitted.

LBD00047-07-5

ronments, abstract such perceptions into models through analysis in an automated manner, and use model inference to formulate options for information or action.

3. "Artificial intelligence model" means an information system or component of an information system that implements artificial intelligence technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs.

4. "Compute cost" means the cost incurred to pay for compute used in training a model when calculated using the average market prices of cloud compute in the United States at the start of training such model as reasonably assessed by the person doing the training.

5. "Deploy" means to use a frontier model or to make a frontier model foreseeably available to one or more third parties for use, modification, copying, or a combination thereof with other software, except as reasonably necessary for training or developing the frontier model, evaluating the frontier model or other frontier models, or complying with federal or state laws.

6. "Frontier model" means either of the following:

(a) an artificial intelligence model trained using greater than 10^26 computational operations (e.g., integer or floating-point operations), the compute cost of which exceeds one hundred million dollars.

(b) an artificial intelligence model produced by applying knowledge distillation to a frontier model as defined in paragraph (a) of this subdivision.

7. "Critical harm" means the death or serious injury of one hundred or more people or at least one billion dollars of damages to rights in money or property caused or materially enabled by a large developer's creation, use, storage, or release of a frontier model, through either of the following:

(a) The creation or use of a chemical, biological, radiological, or nuclear weapon; or

(b) An artificial intelligence model engaging in conduct that does both of the following:

(i) Acts with limited human intervention; and

(ii) Would, if committed by a human, constitute a crime specified in the penal law that requires intent, recklessness, or gross negligence, or the solicitation or aiding and abetting of such a crime.

A harm inflicted by an intervening human actor shall not be deemed to result from a developer's activities unless such activities made it substantially easier or more likely for the actor to inflict such harm.

8. "Knowledge distillation" means any supervised learning technique that uses a larger artificial intelligence model or the output of a larger artificial intelligence model to train a smaller artificial intelligence model with similar or equivalent capabilities as the larger artificial intelligence model.

9. "Large developer" means a person that has trained at least one frontier model, the compute cost of which exceeds five million dollars, and has spent over one hundred million dollars in compute costs in aggregate in training frontier models. Accredited colleges and universities shall not be considered large developers under this article to the extent that such colleges and universities are engaging in academic research. If a person subsequently transfers full intellectual property rights of the frontier model to another person (including the right to resell the model) and retains none of those rights for themself, then the receiving person shall be considered the large developer and shall

be subject to the responsibilities and requirements of this article after such transfer.

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

11. "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.

12. "Safety and security protocol" means documented technical and organizational protocols that:

(a) Specify reasonable protections and procedures that, if successfully implemented would appropriately reduce the risk of critical harm;

(b) Describe reasonable administrative, technical, and physical cybersecurity protections for frontier models within the large developer's control that, if successfully implemented, appropriately reduce the risk of unauthorized access to, or misuse of, the frontier model leading to critical harm, including by sophisticated actors;

(c) Describe in detail the testing procedure to evaluate if the frontier model poses an unreasonable risk of critical harm;

(d) Describe in detail how the testing procedure assesses whether the frontier model could be misused, modified, combined with other software or be used to create another foundation model in a manner that would produce critical risk;

(e) State compliance requirements with sufficient detail and specificity to allow the large developer or a third party to readily ascertain whether the requirements of the safety and security protocol have been followed;

(f) Describe how the large developer will fulfill their obligations under this article, including with respect to any requirements, safeguards, or modifications; and

(g) Designate senior personnel to be responsible for ensuring compliance.

13. "Safety incident" means an incident of the following kinds that occurs in such a way that it provides demonstrable evidence of an increased risk of critical harm:

(a) A frontier model autonomously engaging in behavior other than at the request of a user;

(b) Theft, misappropriation, malicious use, inadvertent release, unauthorized access, or escape of the model weights of a frontier model;

(c) The critical failure of any technical or administrative controls, including controls limiting the ability to modify a frontier model; or

(d) Unauthorized use of a frontier model.

14. "Trade secret" means any form and type of financial, business, scientific, technical, economic, or engineering information, including a pattern, plan, compilation, program device, formula, design, prototype, method, technique, process, procedure, program, or code, whether tangible or intangible, and whether or how stored, compiled, or memorialized physically, electronically, graphically, photographically or in writing, that:

(a) Derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use; and

(b) Is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.

§ 1421. Transparency requirements regarding frontier model training and use. 1. Before deploying a frontier model, the large developer of such frontier model shall do all of the following:

(a) Implement a written safety and security protocol;

(b) Retain an unredacted copy of the safety and security protocol, including records and dates of any updates or revisions. Such unredacted copy of the safety and security protocol, including records and dates of any updates or revisions, shall be retained for as long as the frontier model is deployed plus five years;

(c) (i) Conspicuously publish a copy of the safety and security protocol with appropriate redactions and transmit a copy of such redacted safety and security protocol to the attorney general;

(ii) Grant the attorney general access to the unredacted safety and security protocol upon request;

(d) Record, as and when reasonably possible, and retain for as long as the frontier model is deployed plus five years information on the specific tests and test results used in any assessment of the frontier model that provides sufficient detail for third parties to replicate the testing procedure; and

(e) Implement appropriate safeguards to prevent unreasonable risk of critical harm.

2. A large developer shall not deploy a frontier model if doing so would create an unreasonable risk of critical harm.

3. A large developer shall conduct an annual review of any safety and security protocol required by this section to account for any changes to the capabilities of their frontier models and industry best practices and, if necessary, make modifications to such safety and security protocol. If any modifications are made, the large developer shall publish the safety and security protocol in the same manner as required pursuant to paragraph (c) of subdivision one of this section.

4. (a) Beginning on the effective date of this article, or ninety days after a developer first qualifies as a large developer, whichever is later, a large developer shall annually retain a third party to perform an independent audit of compliance with the requirements of this section. Such third party shall conduct audits consistent with best practices.

(b) The third party shall be granted access to unredacted materials as necessary to comply with the third party's obligations under this subdivision.

(c) The third party shall produce a report including all of the following:

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

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

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

(iv) The signature of the lead auditor certifying the results of the audit.

(d) The large developer shall retain an unredacted copy of the report for as long as the frontier model is deployed plus five years.

(e) (i) The large developer shall conspicuously publish a copy of the third party's report with appropriate redactions and transmit a copy of such redacted report to the attorney general.

(ii) The large developer shall grant the attorney general access to the unredacted third party's report upon request.

5. A large developer of a frontier model shall submit to the attorney general the updated total compute cost used to train their frontier models simultaneously with their annual third party's audit report.

6. A large developer shall disclose each safety incident affecting the frontier model to the attorney general within seventy-two hours of the large developer learning of the safety incident or within seventy-two hours of the large developer learning facts sufficient to establish a reasonable belief that a safety incident has occurred. Such disclosure shall include: (a) the date of the safety incident; (b) the reasons the incident qualifies as a safety incident as defined in subdivision thirteen of section fourteen hundred twenty of this article; and (c) a short and plain statement describing the safety incident.

7. (a) This section shall not apply to products or services to the extent that the requirements would strictly conflict with the terms of a contract with a federal government entity and a large developer.

(b) This section applies to the development or deployment of a frontier model for any use that is not the subject of a contract with a federal government entity, even if such frontier model has already been developed, trained, or used by a federal government entity.

8. A large developer shall not knowingly make false or materially misleading statements or omissions in or regarding documents produced pursuant to this section.

9. Any person who is not a large developer, but who sets out to train a frontier model that if completed as planned would qualify such person as a large developer (i.e. at the end of the training, such person will have spent five million dollars in compute costs on one frontier model and one hundred million dollars in compute costs in aggregate in training frontier models, excluding accredited colleges and universities to the extent such colleges and universities are engaging in academic research) shall, before training such model:

(a) Implement a written safety and security protocol, excluding the requirements described in paragraphs (c) and (d) of subdivision twelve of section fourteen hundred twenty of this article; and

(b) Conspicuously publish a copy of the safety and security protocol, with appropriate redactions, and transmit a copy of such redacted safety and security protocol to the attorney general.

§ 1422. Protections, rights and obligations of employees. 1. A large developer or a contractor or subcontractor of a large developer shall not prevent an employee from disclosing, or threatening to disclose, or retaliate against an employee for disclosing or threatening to disclose, information to the large developer or the attorney general, if the employee has reasonable cause to believe that the large developer's activities pose an unreasonable or substantial risk of critical harm, regardless of the employer's compliance with applicable law.

2. An employee harmed by a violation of this section may petition a court for appropriate temporary or preliminary injunctive relief.

3. A large developer shall inform employees of their protections, rights and obligations under this article by posting a notice thereof. Such notice shall be posted conspicuously in easily accessible and well-lighted places customarily frequented by employees.

4. Nothing in this section shall be deemed to diminish the rights, privileges, or remedies of any employee under any other law or regulation or under any collective bargaining agreement or employment contract.

5. As used in this section, the following terms shall have the following meanings:

(a) "Employee" has the same meaning as defined in subdivision five of section two of the labor law and includes both of the following:

(i) Contractors or subcontractors and unpaid advisors involved with assessing, managing, or addressing the risk of critical harm from frontier models; and

(ii) Corporate officers.

(b) "Contractor or subcontractor" means any person, sole proprietor, partnership, firm, corporation, limited liability company, association or other legal entity who by oneself or through others offers to undertake, or holds oneself out as being able to undertake, or does undertake work assessing, managing, or addressing the risk of critical harm from frontier models on behalf of the large developer.

§ 1423. Violations. 1. The attorney general may bring a civil action for a violation of this article and to recover all of the following:

(a) For a violation of section fourteen hundred twenty-one of this article, a civil penalty in an amount not exceeding five percent of the total compute cost used to train the large developer's frontier models as reported by the large developer pursuant to this article for a first violation and in an amount not exceeding fifteen percent of that value for any subsequent violation.

(b) For a violation of section fourteen hundred twenty-two of this article, a civil penalty in an amount not exceeding ten thousand dollars per employee for each violation of such section to be awarded to the employee who was retaliated against.

(c) For a violation of section fourteen hundred twenty-one or fourteen hundred twenty-two of this article, injunctive or declaratory relief.

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

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

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

(ii) As the result of the steps described in subparagraph (i) of this paragraph, the corporate structure of the large developer or affiliated entities would frustrate recovery of penalties, damages, or injunctive relief under this section.

3. This section does not limit the application of other laws.

§ 1424. Duties and obligations. The duties and obligations imposed by this article 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.

§ 1425. Preemption. This article shall not apply to the extent that it is preempted by federal law.

§ 3. This act shall take effect on the ninetieth day after it shall have become a law.