STATE OF NEW YORK

6953--B

2025-2026 Regular Sessions

IN SENATE

March 27, 2025

Introduced by Sens. GOUNARDES, BAILEY, BORRELLO, BRISPORT, FAHY, HARCKHAM, HOYLMAN-SIGAL, JACKSON, KRUEGER, LIU, MAYER, PALUMBO, SALAZAR -- read twice and ordered printed, and when printed to be committed to the Committee on Internet and Technology -- committee discharged, bill amended, ordered reprinted as amended and recommitted to said committee -- committee discharged, bill amended, ordered reprinted as amended and recommitted to said committee

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 that a developer may make when necessary to:

(a) protect public safety to the extent the developer can reasonably predict such risks;

(b) protect trade secrets;

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

LBD00047-15-5

(c) prevent the release of confidential information as required by state or federal law;

(d) protect employee or customer privacy; or

(e) prevent the release of information otherwise controlled by state or 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 environments, 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 the final training run of a model when calculated using the average published 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 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; or

(b) an artificial intelligence model produced by applying knowledge distillation to a frontier model as defined in paragraph (a) of this subdivision, provided that the compute cost for such model produced by applying knowledge distillation exceeds five million dollars.

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 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 no meaningful 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 were a substantial factor in bringing about the harm, the intervening human actor's conduct was reasonably foreseeable as a probable consequence of the developer's activities, and could have been reasonably prevented or mitigated through alternative design, or security measures, or safety protocols.

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 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) Describe 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 models 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 and whether the frontier model could be misused, be modified, be executed with increased computational resources, evade the control of its large developer or user, be combined with other software or be used to create another frontier model in a manner that would increase the risk of critical harm;

(d) Enable the large developer or third party to comply with the requirements of this article; and

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

13. "Safety incident" means a known incidence of critical harm or 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 a 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 and division of homeland security and emergency services;

(ii) Grant the attorney general and division of homeland security and emergency services or the attorney general access to the safety and security protocol, with redactions only to the extent required by federal law, 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 required by this section or the developer's safety and security protocol 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 material 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 large developer shall disclose each safety incident affecting the frontier model to the attorney general and division of homeland security and emergency services 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.

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

§ 1422. Violations. 1. The attorney general may bring a civil action for a violation of this article and to recover all of the following, determined based on severity of the violation:

(a) For a violation of section fourteen hundred twenty-one of this article, a civil penalty in an amount not exceeding ten million dollars for a first violation and in an amount not exceeding thirty million dollars for any subsequent violation.

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

2. Nothing in this article shall be construed to establish a private right of action associated with violations of this article.

3. Nothing in this subdivision shall be construed to prevent a large developer from asserting that another person, entity, or factor may be responsible for any alleged harm, injury, or damage resulting from a critical harm or a violation of this article.

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

§ 1423. 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.

§ 1424. Scope. This article shall only apply to frontier models that are developed, deployed, or operating in whole or in part in New York state.

§ 1425. Severability. If any clause, sentence, paragraph, subdivision, section or part of this article shall be adjudged by any court of competent jurisdiction to be invalid, such judgment shall not affect, impair, or invalidate the remainder thereof, but shall be confined in its operation to the clause, sentence, paragraph, subdivision, section, or part thereof directly involved in the controversy in which such judgment shall have been made.

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