Int. No. 1109

By Council Members Hanks, Maloney, Louis, Narcisse, Epstein, Zhuang, Gutiérrez and Morano

A Local Law to amend the administrative code of the city of New York, in relation to requiring reporting and public disclosure of artificial intelligence safety incidents concerning city contracts

Be it enacted by the Council as follows:

Section 1. Chapter 1 of title 6 of the administrative code of the city of New York is amended by adding a new section 6-152 to read as follows:

§ 6-152 Artificial intelligence safety incidents concerning city contracts. a. Definitions. For purposes of this section, the following terms have the following meanings:

Artificial intelligence model. The term "artificial intelligence model" means an engineered or machine-based system that varies in its level of autonomy and that can, for explicit or implicit objectives, infer from the input it receives how to generate outputs that can influence physical or virtual environments.

Artificial intelligence safety incident. The term "artificial intelligence safety incident" means any of the following:

1. Unauthorized access to, modification of, or exfiltration of, the model weights of an artificial intelligence model that results in death, bodily injury, damage to, or loss of, property, or a breach of security;

2. Harm resulting from the materialization of a substantial risk;

3. Loss of control of an artificial intelligence model causing death, bodily injury, damage to, or loss of, property, or a breach of security; or

4. An artificial intelligence model that uses deceptive techniques against the developer to subvert the controls or monitoring of its developer outside of the context of an evaluation designed to elicit this behavior

and in a manner that demonstrates materially increased substantial risk.

Breach of security. The term "breach of security" shall mean the unauthorized access, acquisition, disclosure, or use of any data or system that compromises the security, confidentiality, or integrity of any information or technology maintained by an agency.

City chief procurement officer. The term "city chief procurement officer" has the same meaning as set forth in section 6-129.

Contract. The term "contract" means any written agreement, purchase order, or instrument by which the city is committed to expend or does expend funds in return for goods, professional services, standard services, or construction.

Contractor. The term "contractor" means a person who has been awarded a contract by an agency.

Covered contract. The term "covered contract" means a contract entered into on or after the effective date of the local law that added this section, the performance of which will foreseeably involve the development, storage, use, or deployment of an artificial intelligence model.

Model weight. The term "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.

Reportable artificial intelligence safety incident. The term "reportable artificial intelligence safety incident" means any artificial intelligence safety incident that: (i) involves an artificial intelligence model developed, stored, used, or deployed by a contractor for the purpose of performing under a contract; (ii) involves a breach of security; or (iii) otherwise relates to any contract or any information or technology maintained by an agency.

Substantial risk. The term "substantial risk" means a foreseeable and material risk that the development, storage, use, or deployment of an artificial intelligence model will materially contribute to (i) the death of, or serious injury to, a person, (ii) damage to, or loss of, property, or (iii) a breach of security, resulting from an artificial intelligence model doing any of the following:

powered by Legistar™

1. Providing expert-level assistance in the creation or release of a chemical, biological, radiological, or nuclear weapon;

2. Engaging in conduct with no meaningful human oversight, intervention, or supervision that is either a cyberattack or, if the conduct had been committed by a human, would constitute the crime of murder, assault, extortion, or theft, including theft by false pretense; or

3. Evading the control of its developer or user.

b. Establishing standards and procedures. 1. In consultation with the city chief procurement officer and the office of technology and innovation, the office of cyber command shall establish standards and procedures that a contractor that is a party to a covered contract shall use to identify the occurrence of any reportable artificial intelligence safety incident.

2. The office of cyber command shall periodically review the standards and procedures established pursuant to this subdivision and update such standards and procedures as the office of cyber command determines to be necessary.

3. Within 7 days after the establishment of, or any update to, the standards and procedures established pursuant to this subdivision, the office of cyber command shall submit copies of such standards and procedures to the mayor and the speaker of the council.

c. Contract provision. A covered contract shall include a provision requiring a contractor to:

1. Use the standards and procedures established pursuant to subdivision b of this section to identify the occurrence of any reportable artificial intelligence safety incident; and

2. Within 24 hours of becoming aware of the occurrence of any reportable artificial intelligence safety incident, provide to the office of cyber command, in writing: (i) a short and plain statement describing such safety incident; (ii) the date of such safety incident; and (iii) the reasons such incident qualifies as a reportable artificial intelligence safety incident.

d. Mandatory agency reporting. If an agency becomes aware of the occurrence of any reportable

artificial intelligence safety incident, such agency shall, within 24 hours of becoming aware of such occurrence, provide to the office of cyber command, in writing: (i) a short and plain statement describing such safety incident; (ii) the date of such safety incident; and (iii) the reasons such incident qualifies as a reportable artificial intelligence safety incident.

e. Public disclosure. Within 24 hours of the office of cyber command becoming aware of any reportable artificial intelligence safety incident, the office of cyber command shall post on the city's website: (i) a short and plain statement describing such safety incident; (ii) the date of such safety incident; and (iii) the reasons such incident qualifies as a reportable artificial intelligence safety incident. Such post shall not include information that would compromise the trade secrets or cybersecurity of the contractor, public safety, or the national security of the United States or that would be prohibited by any federal or state law.

f. Annual reporting. No later than one year after the effective date of the local law that added this section, and annually thereafter, the office of cyber command shall submit to the mayor and the speaker of the council, post on the city's website, and post on the single web portal described in section 23-502 a report that summarizes the artificial intelligence safety incidents about which a contractor or agency has notified the office of cyber command pursuant to subdivisions c and d of this section. Such report shall include the following information regarding each reported artificial intelligence safety incident during the immediately preceding 12-month period:

1. The contractor involved in the artificial intelligence safety incident;

2. The agency that awarded the covered contract to the contractor;

3. A short and plain statement describing such safety incident;

4. The date of such safety incident;

5. The reasons such incident qualifies as a reportable artificial intelligence safety incident; and

6. A short and plain statement describing the manner in which such safety incident was resolved.

§ 2. This local law takes effect 180 days after it becomes law.

powered by Legistar™

JPA
LS #26630
9/25/2026