.....

(Original Signature of Member)

119TH CONGRESS
2D SESSION

H. R. _____

To require certain artificial intelligence model developers to submit reports to the Secretary of Commerce, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

Mr. MORAN introduced the following bill; which was referred to the Committee on _____

A BILL

To require certain artificial intelligence model developers to submit reports to the Secretary of Commerce, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the ''AI Incident Reporting Act''.

SEC. 2. REQUIREMENT TO REPORT AI INCIDENTS.

(a) DESIGNATION OF COVERED MODELS; GUIDELINES AND REGULATIONS.—

(1) DESIGNATION OF COVERED MODELS AND ENTITIES.—Not later than 180 days after the date of the enactment of this section, the Secretary, in consultation with, as appropriate, the heads of relevant agencies as determined by the Secretary, artificial intelligence model developers, other relevant private-sector entities, academic, technical, cybersecurity, national security, and public safety experts, and such other outside experts as the Secretary determines to be appropriate, shall promulgate regulations that—

(A) establish capability or other thresholds that determine which artificial intelligence models and model developers could pose significant risks to the national security of the United States or to public safety; and

(B) designate such developers as covered model developers and such models as covered models for purposes of this section.

(2) COORDINATION.—The Secretary may coordinate with the heads of other agencies, as determined appropriate by the Secretary, to identify reportable activity and to receive, analyze, and act upon reports submitted under this section.

(3) GUIDELINES AND REGULATIONS.—

(A) IN GENERAL.—Not later than 180 days after the date of the enactment of this Act, the Secretary shall issue guidelines for covered model developers to be in compliance with the requirements of this Act and may promulgate regulations as are necessary to carry out this Act.

(B) REQUIREMENTS FOR REGULATIONS.—In establishing regulations and requirements under this subsection, the Secretary shall ensure, to the maximum extent practicable, that the regulations and requirements—

(i) clearly describe the categories of information that must be reported under this section;

(ii) minimize unnecessary ambiguity, duplication, and undue reporting burden; and

(iii) establish clear, secure, and straightforward mechanisms for submission of reports.

(C) THRESHOLDS.—In establishing thresholds under paragraph (1), the Secretary shall give particular weight to whether a model has the capability to engage in, or presents a sig-

nificant risk of, reportable activity described in subsection (b)(2).

(b) REPORTING REQUIREMENT.—

(1) IN GENERAL.—Not later than 7 days after the date on which a covered model developer knows, or reasonably believes, that reportable activity described in paragraph (2) has occurred related to a covered model, the developer shall submit to the Secretary a report that contains a detailed description of the activity.

(2) REPORTABLE ACTIVITY.—For purposes of this section, reportable activity is any of the following:

(A) Behavior expressing that the model is attempting to evade human oversight, deceive evaluators or operators, circumvent safeguards, resist shutdown or modification, obtain unauthorized access to tools, systems, or privileges, or otherwise undermine the ability of human operators to reliably control the model, but does not include behavior elicited solely through an evaluation designed to elicit such behavior, in which the model is not in production deployment and the behavior is not indicative of analogous behavior in deployment.

(B) Unauthorized access to, theft of, or attempted theft of model weights that the developer reasonably assesses had a credible prospect of obtaining, evidence that model weights have been exfiltrated or materially compromised, or behavior suggesting that a model is autonomously attempting to exfiltrate model weights without authorization or otherwise facilitate unauthorized transfer of model weights or related model artifacts outside of a testing environment.

(C) Capabilities that could materially enable or accelerate offensive cyber operations against important software, widely used digital infrastructure, industrial systems, or critical infrastructure, including through the discovery, exploitation, chaining, weaponization, or operationalization of vulnerabilities at a scale, speed, or level of sophistication that could pose serious risks to the national security of the United States or to public safety.

(D) Evidence that a covered model, when unprompted, has demonstrated the ability to materially accelerate or automate the research, development, evaluation, engineering, or im-

provement of advanced artificial intelligence systems, including in ways that could significantly compress timelines for the development or deployment of more capable systems, where the model developer knows, or reasonably believes, that such developments could have serious implications for the national security of the United States or for public safety.

(E) Capabilities that could materially enable or accelerate the development, acquisition, or use of chemical, biological, radiological, nuclear, or explosive weapons by providing uplift to actors that would not otherwise possess such capabilities at a scale, speed, or level of sophistication that could pose serious risks to the national security of the United States or to public safety.

(F) Any circumstance in which an incident or harm of a type described in subparagraph (A), (B), (C), (D), or (E) was reasonably likely to occur and would have posed a serious risk to the national security of the United States or to public safety, but was prevented only because of circumstances unrelated to the safeguards, controls, or mitigations of the developer, such as

the conduct of a third party, the absence of capability or intent on the part of a user, or other fortuity.

(G) Any other capability, incident, or combination of circumstances that the Secretary determines, by rulemaking, appropriate relating to serious harm to the national security of the United States or to public safety.

(c) TIMING, FORM, AND CONTENTS OF REPORTS.—

(1) INITIAL REPORT.—The Secretary shall require—

(A) a covered model developer to submit an initial report within such period as the Secretary determines appropriate and not later than the 7-day period described in subsection (b)(1); and

(B) expedited reporting for any reportable activity described in subsection (b)(2) that presents any imminent or ongoing risk of serious harm.

(2) SUPPLEMENTAL REPORTS.—The Secretary shall require a covered model developer to submit supplemental reports as additional material information, relating to the reportable activity and steps

that are being taken to mitigate the risks of the incident, becomes available.

(3) REQUIRED CONTENTS.—Each report submitted under this subsection shall include, as applicable and to the extent known at the time of submission, the following:

(A) A description of the relevant incident, behavior, or capability.

(B) The date on which, or approximate period during which, the covered model developer discovered the relevant information.

(C) Any known or suspected threat actor, attack vector, system vulnerability, safeguard failure, or other relevant causal or contextual information.

(D) Any known or reasonably suspected implication for the national security of the United States or for public safety.

(E) Such other information as the Secretary determines appropriate.

(4) CONGRESSIONAL REPORTING.—Not later than 48 hours after receipt of any report submitted under subsection (b) that presents an imminent or ongoing risk of serious harm, and not later than 30 days after receiving any report submitted under

paragraph (2), the Secretary shall inform the following individuals of each such report:

(A) The Speaker of the House of Representatives.

(B) The Minority Leader of the House of Representatives.

(C) The Chair of the Committee on Science, Space, and Technology of the House of Representatives.

(D) The Chair of the Committee on Energy and Commerce of the House of Representatives.

(E) The Chair of the Permanent Select Committee on Intelligence of the House of Representatives.

(F) The Majority Leader of the Senate.

(G) The Minority Leader of the Senate.

(H) The Chair of the Committee on Commerce, Science, and Transportation of the Senate.

(I) The Chair of the Committee on Energy and Natural Resources of the Senate.

(J) The Chair of the Select Committee on Intelligence of the Senate.

(d) PROTECTION AND USE OF INFORMATION.—

(1) PROTECTION OF SENSITIVE INFORMATION.—Not later than 180 days after the date of the enactment of this section, the Secretary shall establish procedures to appropriately protect from unauthorized disclosure any sensitive, classified, controlled, or security-relevant information submitted under this section, consistent with applicable law.

(2) EXEMPTION FROM DISCLOSURE.—Information submitted to the Secretary under this section is exempt from disclosure under paragraph (3)(B) of section 552(b) of title 5, United States Code, and may not be disclosed under any State or local law that requires disclosure of information or records.

(3) NO WAIVER OF PRIVILEGE OR PROTECTION.—The submission of information under this section is not a waiver of any applicable privilege or legal protection, including trade secret protection and any attorney-client and work product privilege.

(4) RESTRICTIONS ON USE.—

(A) CIVIL ACTIONS AND ADMINISTRATIVE PROCEEDINGS.—A report submitted under this section, and any communication or material created for the sole purpose of preparing or submitting such a report, may not be received in evidence, subjected to discovery, or otherwise

used in any civil or criminal action or administrative proceeding against the covered model developer that submitted the report, communications, or material.

(B) FEDERAL, STATE, OR LOCAL GOVERNMENT.—Information submitted under this section may not be used by any Federal, State, or local government to regulate, or to bring an enforcement action against, the covered model developer.

(C) RULE OF CONSTRUCTION.—Nothing in this paragraph may—

(i) limit the use of such information by the Secretary or any other agency to respond to, mitigate, or prevent a risk to the national security of the United States or to public safety;

(ii) limit use of a report, or information in the report, to determine compliance with or enforce the requirements of this section; or

(iii) affect the liability of any person for the underlying incident, conduct, or capability described in a report in which such liability can be established on the basis of

information obtained independently of the report.

(5) INFORMATION SHARING WITHIN GOVERNMENT.—The Secretary may share information submitted under this section with other agencies, including an element of the intelligence community and law enforcement agencies, where appropriate and consistent with applicable law. Any information shared under this paragraph is subject to the protections and use restrictions of this subsection for the agency that receives the information.

(e) GOOD-FAITH REPORTING.—In issuing guidelines and regulations under this section, the Secretary shall, to the maximum extent practicable, design reporting requirements to facilitate timely reporting of material incidents, including for a case in which relevant facts are incomplete at the time of initial disclosure, and shall permit supplemental reporting as additional material information becomes available.

(f) ENFORCEMENT.—

(1) AUTHORITY OF THE SECRETARY.—To enforce this section, the Secretary may—

(A) issue orders, regulations, and guidance;

(B) require, inspect, and obtain books, records, reports, audit materials, and other information that the Secretary determines to be relevant or material to determine compliance with, or violations of, this section, from any developer or other person subject to this section;

(C) administer oaths or affirmations and, by subpoena, require any person to appear, testify, and produce books, records, reports, audit materials, and other materials relevant or material to determine compliance with, or violations of, this section, from any developer or other person subject to this section;

(D) conduct investigations within the United States and, consistent with applicable law, outside the United States;

(E) require corrective action, including the production of omitted records or materials; and

(F) refer a matter to the Attorney General for appropriate civil action, including to recover a civil penalty assessed under paragraph (2) that remains unpaid, to enjoin a violation of this section, or to compel compliance with an order or subpoena issued under this subsection.

(2) CIVIL PENALTIES.—

(A) IN GENERAL.—After notice and an opportunity for a hearing, the Secretary may assess a civil penalty for a violation of this section in an amount not to exceed $2,000,000. Each day of a continuing violation shall constitute a separate offense.

(B) FACTORS.—In determining the amount of a civil penalty under subparagraph (A), the Secretary shall consider the nature, circumstances, extent, gravity, and duration of the violation, the degree of culpability, any history of prior violation, any good faith effort to comply, any other mitigating factor, and such other matters as justice may require.

(g) DEFINITIONS.—In this section:

(1) AGENCY.—The term ''agency'' has the meaning given that term in section 551 of title 5, United States Code.

(2) ARTIFICIAL INTELLIGENCE.—The term ''artificial intelligence'' includes the following:

(A) Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.

(B) An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.

(C) An artificial system designed to think or act like a human, including cognitive architectures and neural networks.

(D) A set of techniques, including machine learning, that is designed to approximate a cognitive task.

(E) An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting.

(3) COVERED MODEL.—The term ''covered model'' means a model designated by the Secretary under subsection (a)(1).

(4) COVERED MODEL DEVELOPER.—The term ''covered model developer'' means any person or entity that—

(A) develops or trains a covered model; or

(B) substantially modifies a covered model, including through fine-tuning or other modifica-

tion of the weights of the model, in a manner that the Secretary determines causes the model to meet a threshold established under subsection (a)(1).

(5) MODEL WEIGHTS.—The term ''model weights'' means the parameters, numerical values, or other internal artifacts of an artificial intelligence model that are sufficient to reproduce, substantially reproduce, or enable the operational use of the model.

(6) SECRETARY.—The term ''Secretary'' means the Secretary of Commerce.

(7) STATE.—The term ''State'' means each of the several States, the District of Columbia, each commonwealth, territory, or possession of the United States, and each federally recognized Indian Tribe.