119TH CONGRESS
2D SESSION

S. _____

To amend title 10, United States Code, to establish policy for the Department of Defense on maximizing autonomy and artificial intelligence systems, to establish requirements relating to Department review and verification of autonomous weapon systems and artificial intelligence capabilities, and for other purposes.

IN THE SENATE OF THE UNITED STATES

Mr. COONS (for himself and Mr. REED) introduced the following bill; which was read twice and referred to the Committee on _____

A BILL

To amend title 10, United States Code, to establish policy for the Department of Defense on maximizing autonomy and artificial intelligence systems, to establish requirements relating to Department review and verification of autonomous weapon systems and artificial intelligence capabilities, 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 ''Responsible Artificial Intelligence Defense Act of 2026''.

SEC. 2. POLICY AND GUIDANCE RELATED TO AUTONOMOUS WEAPON SYSTEMS AND ARTIFICIAL INTELLIGENCE CAPABILITIES ACQUISITION.

(a) IN GENERAL.—Chapter 345 of title 10, United States Code, is amended by adding at the end the following new section:

"§ 4577. Autonomous weapon systems and artificial intelligence capabilities acquisition; planning and oversight processes

''(a) POLICY.—It is the policy of the Department of Defense to maximize uses of autonomy and artificial intelligence capabilities to the extent practicable, while ensuring and maintaining that implementation of such autonomy and artificial intelligence capabilities provides continuous and rigorous human oversight to ensure that operations are conducted in accordance with the law of war, applicable treaties, weapon system safety rules, applicable rules of engagement, and long-standing frameworks protecting the privacy and civil liberties of United States persons.

''(b) REQUIREMENT.—In accordance with the policy set forth in subsection (a), the Secretary of Defense shall—

''(1) ensure personnel exercise appropriate levels of human judgment and consistently monitor any deployed artificial intelligence-enabled autonomous

weapons systems, while remaining responsible for the development, deployment, and use of autonomous weapon systems and artificial intelligence capabilities;

''(2) take deliberate steps to ensure accuracy in autonomous weapon systems and artificial intelligence capabilities;

''(3) develop and deploy autonomous weapon systems and artificial intelligence capabilities in a manner that promotes an appropriate understanding of the technology, their development processes, and operational methods applicable to autonomous weapon systems and artificial intelligence capabilities;

''(4) subject prototype and deployed autonomy and artificial intelligence capabilities to routine testing and information assurance across their entire life cycles to ensure that capabilities meet defined safety, security, and effectiveness parameters; and

''(5) design and engineer autonomy and artificial intelligence capabilities to fulfill their intended functions, and deploy these capabilities such that human operators—

''(A) retain the ability to detect and avoid unintended consequences or behaviors; and

"(B) retain a means for human intervention to disengage or deactivate deployed systems that demonstrate unintended or illegal behavior.

"(c) REVIEW FOR AUTONOMOUS WEAPONS SYSTEMS AND SUPPORTING ARTIFICIAL INTELLIGENCE CAPABILITIES.—

"(1) LEVELS OF HUMAN JUDGMENT.—(A) The Secretary shall ensure that any autonomous weapon system or artificial intelligence capability to be utilized by the Department, except as exempted by subsection (c), is categorized by the appropriate level of human judgment required to mitigate risks to life, safety and health of Department personnel or noncombatant civilian harm.

"(B) For purposes of subparagraph (A), appropriate levels of human judgment for an autonomous weapon system shall consist of two levels as follows:

"(i) Level 1, which means that the weapon system poses little to no risk to human life or safety should the system fail to act as designed.

"(ii) Level 2, which means that the weapon system poses moderate to high risk to human life or safety should the system fail to act as designed.

"(C) Appropriate levels of human judgment for an artificial intelligence capability shall consist of two levels as follows:

"(i) Level 1, which means that the artificial intelligence capability poses little to no risk as a result of compromise of data integrity or operational support should the system fail to act as designed.

"(ii) Level 2, which means that the artificial intelligence capability—

"(I) uses data protected by section 552a of title 5 (commonly known as the 'Privacy Act of 1974') or regulations promulgated under the Health Insurance Portability and Accountability Act of 1996 (Public Law 104–191); or

"(II) poses a risk such that the failure of the artificial intelligence system to act as designed would severely affect the ability of the Department to perform the designated mission of the artificial intelligence capability.

"(2) REVIEW AND VERIFICATION.—Subject to subsection (d), for any autonomous weapon system or artificial intelligence capability under development

or being fielded by the Department, the Secretary shall ensure that such autonomous weapon system or artificial intelligence capability requires—

"(A) before a decision to enter prototyping or formal development, review and verification by the Under Secretary of Defense for Research and Engineering that—

"(i) the system design incorporates the necessary capabilities to allow commanders, operators, and analysts to exercise appropriate levels of human judgment over the use of force in the envisioned planning and employment processes for the autonomous weapon system or artificial intelligence capability;

"(ii)(I) the autonomous weapon system or artificial intelligence capability is designed to complete engagements within a timeframe and geographic area, as well as other applicable environmental and operational parameters, consistent with commander and operator intentions; or

"(II) if not designed as described in subclause (I), the autonomous weapon system or artificial intelligence capability will

terminate engagements or obtain additional operator input before continuing the engagement;

''(iii) the combination of the design and concept of employment of the autonomous weapon system or artificial intelligence capability, such as its target selection and engagement logic and other relevant processes or measures, accounts for risks to nontargets, consistent with commander and operator intent and the laws of war;

''(iv) the design of the autonomous weapon system or artificial intelligence capability, including system safety, anti-tamper mechanisms, and the cybersecurity of the autonomous weapon system or artificial intelligence capability, in accordance with Department of Defense Instruction 8500.01 (relating to cybersecurity), or successor instruction, addresses and minimizes the probability and consequences of failures;

''(v) plans are in place for verification and validation and test and evaluation to

establish the reliability, effectiveness, and suitability of the autonomous weapon system or artificial intelligence capability under realistic conditions, including possible adversary actions, to a sufficient standard consistent with the potential consequences of an unintended engagement or unauthorized parties interfering with the operation of the autonomous weapon system or artificial intelligence capability prior to fielding; and

''(vi) a preliminary legal review of the autonomous weapon system or artificial intelligence capability has been completed—

''(I) in coordination with the General Counsel of the Department of Defense; and

''(II) in accordance with Department of Defense Directive 5000.01 (relating to Defense Acquisition System), or successor directive, Department of Defense Directive 2311.01 (relating to Department of Defense Law of War Program), or successor directive, and, where applicable, De-

partment of Defense Directive 3000.03E (relating to Department of Defense Agent for Non-Lethal Weapons and Non-Lethal Weapon Policy), or successor directive; and

''(B) before fielding, review and verification by the Under Secretary for Research and Engineering, in consultation with the Vice Chairman of the Joint Chiefs of Staff, that, with respect to the autonomous weapon system or artificial intelligence capability—

''(i) system capabilities, human-machine interfaces, doctrine, tactics, techniques and procedures, and training have been demonstrated to allow commanders and operators to exercise appropriate levels of human judgment over the use of force and to employ systems with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and definable rules of engagement that are applicable or reasonably expected to be applicable;

''(ii) system safety, anti-tamper mechanisms, cyber survivability, operational re-

silience, and cybersecurity capabilities have been implemented to minimize the probability and consequences of failures;

"(iii) for autonomous weapon systems or artificial intelligence capabilities that are being fielded that may have gone through the development pipeline and verified under subparagraph (A), an updated legal review of the weapon system or artificial intelligence capability has been completed—

"(I) in coordination with the General Counsel of the Department of Defense; and

"(II) in accordance with Department of Defense Directive 5000.01 (relating to Defense Acquisition System), or successor directive, Department of Defense Directive 2311.01 (relating to Department of Defense Law of War Program), or successor directive, and, where applicable, Department of Defense Directive 3000.03E (relating to Department of Defense Agent for Non-Lethal Weap-

ons and Non-Lethal Weapon Policy), or successor directive; and

"(iv) a monitoring regime is in place to identify and address changes in operational environment, data inputs, and use that could contribute to failure of the system or capability to act in a manner consistent with the intent for the system or capability.

"(3) VALIDITY OF VERIFICATION.—(A) The Secretary shall treat each verification under paragraph (2) or paragraph (4) as valid for a period of three years.

"(B) An autonomous weapon system or artificial intelligence capability that is a substantially similar variant of another autonomous weapon system or artificial intelligence capability that is verified under paragraph (2) or paragraph (4) shall also be treated as verified.

"(4) SUBSEQUENT REVIEW AND VERIFICATION.—(A) For any autonomous weapon system or artificial intelligence capability that was previously verified under paragraph (2)(A) or exempted under any predecessor review process under Department of Defense Directive 3000.09 (relating

to Autonomy in Weapon Systems), that does not currently have a valid verification pursuant to paragraph (3), the Secretary shall ensure that it undergoes subsequent review and verification under such paragraph.

''(B) For any autonomous weapon system or artificial intelligence capability that was previously verified under paragraph (2)(B) or exempted under any predecessor review process under Department of Defense Directive 3000.09 (relating to Autonomy in Weapon Systems) that does not currently have a valid verification pursuant to paragraph (3), the Secretary shall ensure that it undergoes subsequent review and verification under such paragraph as if it had not been deployed.

''(5) PRIVACY IMPACT ASSESSMENTS.—(A) For each artificial intelligence capability classified under subsection (c)(1)(C)(ii), the Secretary ensure that a privacy impact assessment is conducted by the Director for Privacy, Civil Liberties and Transparency.

''(B) In carrying out a privacy impact assessment under subparagraph (A), the Director may consult with such technical and policy experts in the Department of Defense or elsewhere in the Federal Government as the Director considers appropriate.

''(6) WAIVER OF UPDATED LEGAL REVIEW.—The Under Secretary of Defense for Research and Engineering may temporarily waive the requirement for an updated legal review under paragraph (2)(B)(iii) for longer than one year if the capability is being deployed in response to real-world conflict or for compelling national interest.

''(d) EXCEPTIONS.—The following categories of autonomous weapon systems are not subject to the policy set forth in subsection (a) or the requirements of subsection (b) and (c):

''(1) Operator-supervised autonomous weapon systems used to select and engage materiel targets to intercept attempted time-critical or saturation attacks.

''(2) Operator-supervised autonomous weapon systems used to select and engage materiel targets for defending operationally deployed remotely piloted or autonomous vehicles or vessels.

''(3) Autonomous or semi-autonomous cyberspace capabilities, reasonably judged to be non-lethal in nature.

''(4) Unarmed platforms, whether remotely operated or operated by onboard personnel, and whether autonomous or semi-autonomous.

"(5) Unguided munitions.

"(6) Munitions manually guided by the operator.

"(7) Mines.

"(8) Unexploded explosive ordnance.

"(9) Autonomous or semi-autonomous systems that are not weapon systems.

"(10) Any weapon system—

"(A) that is not continuously monitored by a human operator;

"(B) that the Secretary, after completing review of the system per subsection (c), has determined to be safer and more reliable for the intended use than alternative systems that are verified for such use per subsection (c) and involve continuous human supervision; and

"(C) for which—

"(i) the Secretary has notified the congressional defense committees of the intended use of the system; and

"(ii) no congressional defense committee has objected to during the 30-day period beginning on the date on which the notice was submitted under clause (i).

"(e) PROHIBITIONS ON CERTAIN USES OF AUTONOMY OR ARTIFICIAL INTELLIGENCE CAPABILITIES.—Except as may be provided in another statute, the Secretary may not use autonomy or an artificial intelligence capability for any of the following use cases:

"(1) For the decision to initiate the launch of a nuclear weapon.

"(2) For the monitoring, tracking, profiling, or targeting of an individual or group of individuals reasonably believed to be in the United States, without a warrant obtained based on probable cause of a crime with an individualized, articulable legal basis, or the collection, querying, or analysis of information about the same not otherwise permissible under the Constitution of the United States, regardless of the origin of the data used, except for activities conducted in accordance with applicable provisions of law.

"(3) In the employment of lethal force by autonomous weapon systems without incorporation of appropriate levels of human judgment.

"(f) VERIFICATION, VALIDATION, TESTING AND EVALUATION OF AUTONOMOUS WEAPON SYSTEMS LEVERAGING AUTONOMY OR ARTIFICIAL INTELLIGENCE SYSTEMS.—For each autonomy and artificial intelligence

system that is covered by the policy set forth in subsection (a), regardless of the acquisition pathway or test and evaluation oversight status for an autonomous weapon system or artificial intelligence capability, the Director for Operational Test and Evaluation shall ensure the autonomous weapon system or artificial intelligence capability functions as anticipated in realistic operational environments against adaptive adversaries and are sufficiently robust to minimize failures, including by ensuring—

''(1) such autonomous weapon system or artificial intelligence capability goes through rigorous hardware and software verification and validation and realistic system developmental and operational test and evaluation, including analysis of unanticipated emergent behavior to assess system performance, capability, reliability, effectiveness, and suitability under realistic conditions, including possible adversary actions, consistent with the potential consequences of unintended engagement or unauthorized parties interfering with the operation of the system or capability;

''(2) hardware and software verification and validation include iterative cyber test and evaluation in accordance with Department of Defense Instruction 5000.89 (relating to Test and Evaluation), or

successor instruction, to verify that autonomous weapon system or artificial intelligence capability is resilient and survivable in contested cyberspace, if relevant to the purpose and mission of the system or capability;

''(3) systems incorporating autonomy or artificial intelligence capabilities go through rigorous developmental and operational test and evaluation to verify and validate that autonomous weapon system or artificial intelligence capability is robust according to design requirements;

''(4) test and evaluation of systems incorporating autonomy or artificial intelligence capabilities include testing to confirm that their autonomy or artificial intelligence algorithms can be rapidly reprogrammed on new input data to enable timely correction of any unintended system behaviors that may be observed or discovered during future system operations;

''(5) adequate training, tactics, techniques, procedures, and doctrine are available, periodically reviewed, and used by system operators and commanders to understand the functioning, capabilities, and limitations of the system's autonomy or artificial intelligence in realistic operational conditions;

''(6) system design and human-machine interfaces are readily understandable to trained operators, with clear ability for trained operators to activate and deactivate system functions;

''(7) after initial operational testing and evaluation, as directed by the Director, system data is collected and any further changes to the system undergo appropriate verification, validation, test, and evaluation to ensure that critical safety features have not been degraded;

''(8) system software is tested using the best means and methods available to the Department to validate that critical safety features have not been degraded;

''(9) automated testing tools, such as modeling and simulation, are used whenever feasible;

''(10) testing identifies any new operating states and other relevant changes in the autonomous weapon system or artificial intelligence capability;

''(11) as directed by the Director—

''(A) each new or revised operating state undergoes appropriate and tailored additional test and evaluation to characterize the system behavior in that new operating state; and

''(B) whole system follow-on operational test and evaluation when required due to changes to the state transition matrix; and

''(12) in coordination with the Under Secretary for Research and Engineering and Director, the owning component of the Department provides for monitoring to identify and address when changes to the system design or operational environment require additional testing and evaluation to provide sufficient confidence that the system will continue to avoid unintended engagements and resist interference by unauthorized parties.

''(g) AUTONOMY AND ARTIFICIAL INTELLIGENCE SYSTEMS WORKING GROUP.—

''(1) ESTABLISHMENT.—(A) The Secretary shall establish and charter a working group for the purposes set forth in paragraph (2).

''(B) The working group established under subparagraph (A) shall be known as the 'Autonomy and Artificial Intelligence Working Group'.

''(2) PURPOSES.—The purposes set forth in this paragraph are as follows:

''(A) To support the Under Secretary of Defense for Research and Engineering, and the Vice Chairman of the Joint Chiefs of Staff in

considering the full range of relevant Department interests during the review of autonomous weapon systems and artificial intelligence capabilities before formal development.

''(B) To support the Under Secretary of Defense for Research and Engineering and the Vice Chairman of the Joint Chiefs of Staff, in considering the full range of relevant Department interests during the review of autonomous weapon systems before deployment.

''(C) When requested by appropriate representatives of the secretaries of the military departments, the Commander of United States Special Operations Command, or, when applicable, a director of a defense agency or a Department of Defense Field Activity—

''(i) to advise whether a given weapon system requires senior-level approval in accordance with this section; and

''(ii) to help identify and advise on addressing potential issues presented by a given weapon system during a potential senior-level review in accordance with this section.

"(D) To develop and issue safety standards for use of autonomy and artificial intelligence capabilities in evaluation of such capabilities.

"(h) ANNUAL REPORT.—Not later than January 31 of each year until January 31, 2037, the Secretary shall submit to the congressional defense committees an annual report on the administration of this section.

"(i) DEFINITIONS.—In this section:

"(1) The term 'artificial intelligence' has the meaning given the term section 5002 of the National Artificial Intelligence Initiative Act of 2020 (15 U.S.C. 9401).

"(2) The term 'autonomous weapon systems' means a weapon system that, once activated, can select and engage targets without further intervention by an operator. Such term includes operator-supervised autonomous weapon systems that are designed to allow operators to override operation of the weapon system but can select and engage targets without further operator input after activation.

"(3) The term 'autonomy' means a capability (or set of capabilities) that enables a particular action of a system to be automatic or, within specified

boundaries self-governing with minimal human oversight.''.

(b) CLERICAL AMENDMENT.—The table of sections for chapter 345 of title 10, United States Code is amended by inserting after the item related to section 4576 the following new item:

''4577. Autonomous weapon systems and artificial intelligence capabilities acquisition; planning and oversight processes.''.