119TH CONGRESS
1ST SESSION

S. _____

To require the Secretary of Energy to establish the Advanced Artificial
Intelligence Evaluation Program, and for other purposes.

_____

IN THE SENATE OF THE UNITED STATES

_____

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

_____

A BILL

To require the Secretary of Energy to establish the Advanced
Artificial Intelligence Evaluation Program, and for other
purposes.

1 *Be it enacted by the Senate and House of Representa-*

2 *tives of the United States of America in Congress assembled,*

3 **SECTION 1. SHORT TITLE.**

4   This Act may be cited as the "Artificial Intelligence

5 Risk Evaluation Act of 2025".

6 **SEC. 2. SENSE OF CONGRESS; PURPOSES.**

7   (a) SENSE OF CONGRESS.—It is the sense of Con-

8 gress that rapidly advancing artificial intelligence capabili-

9 ties present both opportunities and significant risks to na-

1 tional security, public safety, economic competitiveness,

2 civil liberties, and healthy labor and other markets, and

3 that, as artificial intelligence advances toward human-level

4 capabilities in virtually all domains, the United States

5 must establish a secure testing and evaluation program

6 to generate data-driven options for managing emerging

7 risks.

8   (b) PURPOSES.—The purposes of the program estab-

9 lished under this Act are to provide Congress with the em-

10 pirical data, lessons, and insights necessary for Federal

11 oversight of artificial intelligence to ensure that regulatory

12 decisions are made on the basis of empirical testing, and

13 to enable Congress to safeguard American citizens.

SEC. 3. DEFINITIONS.

15   In this Act:

16   (1) ADVANCED ARTIFICIAL INTELLIGENCE SYS-

17   TEM.—

18   (A) IN GENERAL.—Subject to subpara-

19   graph (B), the term "advanced artificial intel-

20   ligence system" means an artificial intelligence

21   system that was trained using a quantity of

22   computing power greater than 10^{26} integer or

23   floating-point operations.

24   (B) ALTERNATE MEANING.—The Sec-

25   retary may, by a rule, propose a new definition

1 of the term "advanced artificial intelligence sys-

2 tem" to replace the definition in subparagraph

3 (A), which new definition shall not go into ef-

4 fect until the Secretary submits the rule to

5 Congress and a joint resolution approving the

6 rule is enacted into law.

7 (2) ADVERSE AI INCIDENT.—The term "ad-

8 verse AI incident" means an incident relating to an

9 artificial intelligence system that involves—

10 (A) a loss-of-control scenario;

11 (B) a risk of weaponization by a foreign

12 adversary, a foreign terrorist organization, or

13 another adversary of the United States Govern-

14 ment;

15 (C) a threat to the safety or reliability of

16 critical infrastructure (as defined in subsection

17 (e) of the Critical Infrastructures Protection

18 Act of 2001 (42 U.S.C. 5195c(e)));

19 (D) a significant erosion of civil liberties,

20 economic competition, and healthy labor mar-

21 kets;

22 (E) scheming behavior; or

23 (F) an attempt to carry out an incident de-

24 scribed in subparagraphs (A) through (E).

1   (3) ARTIFICIAL INTELLIGENCE; AI.—The term

2   "artificial intelligence" or "AI" means technology

3   that enables a device or software—

4   (A) to make—for a given set of human-de-

5   fined objectives—predictions, recommendations,

6   or decisions influencing real or virtual environ-

7   ments; and

8   (B) to use machine and human-based in-

9   puts—

10   (i) to perceive real and virtual envi-

11   ronments;

12   (ii) to abstract such perceptions into

13   models through analysis in an automated

14   manner; and

15   (iii) to use model inference to formu-

16   late options for information or action.

17   (4) ARTIFICIAL INTELLIGENCE SYSTEM; AI

18 SYSTEM.—The term "artificial intelligence system"

19 or "AI system" means a particular model, program,

20 or tool within the field of artificial intelligence.

21   (5) ARTIFICIAL SUPERINTELLIGENCE.—

22   (A) IN GENERAL.—The term "artificial

23   superintelligence" means artificial intelligence

24   that exhibits, or can easily be modified to ex-

1   hibit, all of the characteristics described in sub-

2   paragraph (B).

3   (B) CHARACTERISTICS DESCRIBED.—The

4   characteristics referred to in subparagraph (A)

5   are the following:

6   (i) The AI can enable a device or soft-

7   ware to operate autonomously and effec-

8   tively for long stretches of time in open-

9   ended environments and in pursuit of

10   broad objectives.

11   (ii) The AI can enable a device or

12   software to match or exceed human cog-

13   nitive performance and capabilities across

14   most domains or tasks, including those re-

15   lated to decisionmaking, learning, and

16   adaptive behaviors.

17   (iii) The AI can enable a device or

18   software to potentially exhibit the capacity

19   to independently modify or enhance its own

20   functions in ways that could plausibly cir-

21   cumvent human control or oversight, pos-

22   ing substantial and unprecedented risks to

23   humanity.

24   (6) COMPUTING POWER.—The term "computing

25   power" means the processing power and other elec-

1   tronic resources used to train, validate, deploy, and

2   run AI algorithms and models.

3   (7) COVERED ADVANCED ARTIFICIAL INTEL-

4   LIGENCE SYSTEM DEVELOPER.—The term "covered

5   advanced artificial intelligence system developer"

6   means a person that designs, codes, produces, owns,

7   or substantially modifies an advanced artificial intel-

8   ligence system for use in interstate or foreign com-

9   merce, including by taking steps to initiate a train-

10   ing run of the advanced artificial intelligence system.

11   (8) DEPLOY.—The term "deploy" means an ac-

12   tion taken by a covered advanced artificial intel-

13   ligence system developer to release, sell, or otherwise

14   provide access to an advanced artificial intelligence

15   system outside the custody of the developer, includ-

16   ing by releasing an open-source advanced artificial

17   intelligence system.

18   (9) FOREIGN ADVERSARY.—The term "foreign

19   adversary" means a foreign adversary (as defined in

20   section 791.2 of title 15, Code of Federal Regula-

21   tions) (or successor regulations) that is included on

22   the list in section 791.4(a) of that title (or successor

23   regulations).

24   (10) FOREIGN TERRORIST ORGANIZATION.—

25   The term "foreign terrorist organization" means a

1 foreign entity designated as a foreign terrorist orga-

2 nization by the Secretary of State under section 219

3 of the Immigration and Nationality Act (8 U.S.C.

4 1189).

5   (11) INTERSTATE OR FOREIGN COMMERCE.—

6 The term "interstate or foreign commerce" has the

7 meaning given the term in section 921(a) of title 18,

8 United States Code.

9   (12) LOSS-OF-CONTROL SCENARIO.—The term

10 "loss-of-control scenario" means a scenario in which

11 an artificial intelligence system—

12   (A) behaves contrary to its instruction or

13   programming by human designers or operators;

14   (B) deviates from rules established by

15   human designers or operators;

16   (C) alters operational rules or safety con-

17   straints without authorization;

18   (D) operates beyond the scope intended by

19   human designers or operators;

20   (E) pursues goals that are different from

21   those intended by human designers or opera-

22   tors;

23   (F) subverts oversight or shutdown mecha-

24   nisms; or

1   (G) otherwise behaves in an unpredictable

2   manner so as to be harmful to humanity.

3   (13) PROGRAM.—The term "program" means

4   the Advanced Artificial Intelligence Evaluation Pro-

5   gram established under section 5.

6   (14) SCHEMING BEHAVIOR.—The term "schem-

7   ing behavior" means behavior by an AI system to

8   deceive human designers or operators, including

9   by—

10   (A) hiding its true capabilities and objec-

11   tives; or

12   (B) attempting to subvert oversight mecha-

13   nisms or shutdown mechanisms.

14   (15) SECRETARY.—The term "Secretary"

15   means the Secretary of Energy.

16 **SEC. 4. OBLIGATION TO PARTICIPATE; ENFORCEMENT AND**

17   **PENALTIES.**

18   (a) IN GENERAL.—Each covered advanced artificial

19 intelligence system developer shall—

20   (1) participate in the program; and

21   (2) provide to the Secretary, on request, mate-

22   rials and information necessary to carry out the pro-

23   gram, which may include, with respect to the ad-

24   vanced artificial intelligence system of the covered

25   advanced artificial intelligence system developer—

1   (A) the underlying code of the advanced

2   artificial intelligence system;

3   (B) data used to train the advanced artifi-

4   cial intelligence system;

5   (C) model weights or other adjustable pa-

6   rameters for the advanced artificial intelligence

7   system;

8   (D) the interface engine or other imple-

9   mentation of the advanced artificial intelligence

10   system; and

11   (E) detailed information regarding the

12   training, model architecture, or other aspects of

13   the advanced artificial intelligence system.

14   (b) PROHIBITION ON DEPLOYMENT.—No person may

15 deploy an advanced artificial intelligence system for use

16 in interstate or foreign commerce unless that person is in

17 compliance with subsection (a).

18   (c) PENALTY.—A person that violates subsection (a)

19 or (b) shall be fined not less than $1,000,000 per day of

20 the violation.

21 **SEC. 5. ADVANCED ARTIFICIAL INTELLIGENCE EVALUA-**

22   **TION PROGRAM.**

23   (a) IN GENERAL.—Not later than 90 days after the

24 date of enactment of this Act, the Secretary shall establish

1 an Advanced Artificial Intelligence Evaluation Program

2 within the Department of Energy.

3   (b) ACTIVITIES.—The program shall—

4   (1) offer standardized and classified testing and

5   evaluation of advanced AI systems to systematically

6   collect data on the likelihood of adverse AI incidents

7   for a given advanced AI system;

8   (2) implement testing protocols that match or

9   exceed anticipated real-world AI jailbreaking tech-

10   niques, including adversarial testing by red teams

11   with expertise comparable to sophisticated malicious

12   actors;

13   (3) to the extent feasible, establish and facili-

14   tate classified, independent third-party assessments

15   and blind model evaluations to maintain trans-

16   parency and reliability;

17   (4) provide participating entities with a formal

18   report based on testing outcomes that clearly identi-

19   fies evaluated risks and safety measures;

20   (5) develop recommended containment proto-

21   cols, contingency planning, and mitigation strategies

22   informed by testing data to address identified risks;

23   (6) inform the creation of evidence-based stand-

24   ards, regulatory options, guidelines, and governance

11

1   mechanisms based on data collected from testing

2   and evaluations;

3   (7) assist Congress in determining the potential

4   for controlled AI systems to reach artificial super-

5   intelligence, exceed human oversight or operational

6   control, or pose existential threats to humanity by

7   providing comprehensive empirical evaluations and

8   risk assessments; and

9   (8) develop proposed options for regulatory or

10   governmental oversight, including potential national-

11   ization or other strategic measures, for preventing or

12   managing the development of artificial superintel-

13   ligence if artificial superintelligence seems likely to

14   arise.

15   (c) PLAN FOR PERMANENT FRAMEWORK.—

16   (1) IN GENERAL.—Not later than 360 days

17   after the date of enactment of this Act, the Sec-

18   retary shall submit to Congress a detailed rec-

19   ommendation for Federal oversight of advanced arti-

20   ficial intelligence systems, drawing directly upon in-

21   sights, empirical data, and lessons learned from the

22   program.

23   (2) CONTENTS.—The plan submitted under

24   paragraph (1) shall—

1   (A) summarize and analyze outcomes from

2   testing, identifying key trends, capabilities, po-

3   tential risks, and system behaviors such as

4   weaponization potential, self-replication capa-

5   bilities, scheming behaviors, autonomous deci-

6   sionmaking, and automated AI development ca-

7   pabilities;

8   (B) recommend evidence-based standards,

9   certification procedures, licensing requirements,

10   and regulatory oversight structures specifically

11   informed by testing and evaluation data, ensur-

12   ing alignment between identified risks and reg-

13   ulatory responses;

14   (C) outline proposals for automated and

15   continuous monitoring of AI hardware usage,

16   computational resource inputs, and cloud-com-

17   puting deployments based on observed relation-

18   ships between those factors and AI system per-

19   formance or emergent capabilities;

20   (D) propose adaptive governance strategies

21   that account for ongoing improvements in algo-

22   rithmic efficiency and system capabilities, en-

23   suring that regulatory frameworks remain rel-

24   evant and effective as AI technology advances;

1   (E) suggest revisions with respect to Fed-

2 eral oversight or resourcing, such as a new of-

3 fice within an existing agency, a new agency, or

4 additional funding, that may be necessary to

5 develop and administer a permanent framework

6 for oversight of advanced artificial intelligence

7 systems; and

8   (F) provide comprehensive evaluations re-

9 garding the potential for tested AI systems to

10 exceed human oversight, approach artificial

11 superintelligence, threaten economic competition

12 (including in labor markets), undermine civil

13 liberties, and pose existential risks to humanity,

14 including clearly articulated options for regu-

15 latory or governmental oversight measures to

16 address scenarios of imminent concern identi-

17 fied through testing.

18 (3) UPDATES.—Not less frequently than once

19 every year for the duration of the program, the Sec-

20 retary shall—

21   (A) update the plan submitted under para-

22 graph (1) with new insights, data, and lessons

23 from the program; and

24   (B) submit the updated plan to Congress.

1   (d) SUNSET.—The program shall terminate on the

2 date that is 7 years after the date of enactment of this

3 Act, unless renewed by Congress.