I

119TH CONGRESS
2D SESSION

H. R. 9341

To require the Director of the National Institute of Standards and Technology to develop guidelines to assist agencies with preparing open Government data assets to be used to train artificial intelligence models, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

JUNE 18, 2026

Mr. BABIN (for himself and Ms. LOFGREN) introduced the following bill; which was referred to the Committee on Science, Space, and Technology

---

A BILL

To require the Director of the National Institute of Standards and Technology to develop guidelines to assist agencies with preparing open Government data assets to be used to train artificial intelligence models, 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-Ready Federal Data Guidelines Act".

SEC. 2. AI-READY DATA GUIDELINES FOR FEDERAL AGENCIES.

(a) IN GENERAL.—The National Institute of Standards and Technology Act is amended by inserting after section 21 (15 U.S.C. 278g-4) the following new section:

"SEC. 21A. AI-READY DATA GUIDELINES FOR FEDERAL AGENCIES.

"(a) DEVELOPMENT OF GUIDELINES.—

"(1) IN GENERAL.—The Director, in consultation with the Director of the Office of Science and Technology Policy, the Secretary of Energy, the Director of the Office of Management and Budget, and the head of any other Federal agency the Director considers appropriate, shall develop voluntary guidelines to assist agencies with preparing datasets, including open Government data assets, to be used to train artificial intelligence models.

"(2) ELEMENTS.—In developing the guidelines under paragraph (1) the Director shall carry out the following:

"(A) Address, to the extent practicable, the following:

"(i) Data formatting and structure, including guidelines to ensure datasets are interpretable by artificial intelligence systems.

"(ii) Data labeling and annotation, including scalable methods such as programmatic, automated, and expert-guided approaches for preparing data for use in artificial intelligence development.

"(iii) Data quality evaluation, including guidelines to assess the suitability of datasets for use in artificial intelligence systems.

"(iv) Metadata and documentation, including information sufficient to enable appropriate interpretation and use of datasets for artificial intelligence development.

"(v) Data maintenance, including guidance for the ongoing management and updating of datasets to ensure continued suitability for use in artificial intelligence systems.

"(vi) Data availability, including guidelines for improving and expanding automated access to publicly available information for artificial intelligence model development and use.

"(B) Enable flexible implementation, to the extent practicable, for various use cases across sectors and scientific domains.

"(C) Ensure, to the extent practicable, consistency with Circular A–119 of the Office of Management and Budget.

"(D) Conformity assessment procedures, to the extent practicable.

"(b) PILOT PROGRAMS FOR AI-READY DATA GUIDELINES.—

"(1) IN GENERAL.—The Director, in coordination with the Director of the Office of Science and Technology Policy, the Secretary of Energy, and the head of any other Federal agency the Director determines appropriate, may carry out pilot programs to support the development of conformity assessment procedures for AI-ready datasets used in specific sectors and scientific domains.

"(2) REQUIREMENTS.—If pilot programs under this subsection are carried out, such programs shall—

"(A) not exceed one year in duration;

"(B) develop supplemental guidelines for AI-ready datasets used in specific sectors and

scientific domains in accordance with the guidelines published under subsection (a);

"(C) assess the impact of such guidelines on data usability, interoperability, and readiness for use in artificial intelligence systems in specific sectors and scientific or domains;

"(D) identify technical, operational, or resource challenges associated with future implementation and maintenance of such guidelines in specific sectors and scientific domains; and

"(E) develop, as practicable and appropriate, a process and materials for the transition to an appropriate non-Federal entity of such guidelines for future implementation of and updates to such guidelines.

"(3) SELECTION OF TOPICS.—If pilot programs under this subsection are carried out, such programs shall prioritize areas—

"(A) with significant national security and industrial competitiveness implications, such as biotechnology and biomanufacturing; and

"(B) with respect to which Federal agencies control and maintain AI-ready datasets.

"(3) PARTICIPATION.—If pilot programs under this subsection are carried out, the Director shall

carry out not more than two concurrent such programs through federally funded research programs, National Laboratories, institutions of higher education, or partnerships with the private sector.

''(c) CONGRESSIONAL BRIEFINGS.—Not later than one year after the publication of the guidelines under subsection (a) and annually thereafter for five years, the Director shall brief the Committee on Science, Space, and Technology of the House of Representatives and the Committee on Commerce, Science, and Transportation of the Senate on the implementation of this section.

''(d) PROHIBITION.—The Director may not transfer or reprogram any funds from any other program, project, office, or other entity or activity of the Institute to carry out this section.

''(e) DEFINITIONS.—In this section:

''(1) AGENCY.—The term 'agency' has the meaning given such term in section 3502 of title 44, United States Code.

''(2) ARTIFICIAL INTELLIGENCE.—The term 'artificial intelligence' has the meaning given such term in section 5002 of the National Artificial Intelligence Initiative Act of 2020 (15 U.S.C. 9401).

"(3) ARTIFICIAL INTELLIGENCE MODEL.—The term 'artificial intelligence model' means a component of an artificial intelligence system that is—

"(A) derived using mathematical, computational, statistical, or machine-learning techniques; and

"(B) used as part of an artificial intelligence system to produce outputs from a given set of inputs.

"(4) ARTIFICIAL INTELLIGENCE SYSTEM.—The term 'artificial intelligence system' means any data system, software, hardware, application, tool, service, or utility that operates in whole or in part using artificial intelligence.

"(5) BIOMANUFACTURING.—The term 'biomanufacturing' has the meaning given such term in section 10002 of the Research and Development, Competition, and Innovation Act (42 U.S.C. 18901; popularly referred to as the 'CHIPS and Science Act').

"(6) CONFORMITY ASSESSMENT PROCEDURE.—The term 'conformity assessment procedure' has the meaning given such term in section 451 of the Trade Agreements Act of 1979 (19 U.S.C. 2571).

''(7) DIRECTOR.—The term 'Director' means the Director of the National Institute of Standards and Technology.

''(8) INSTITUTION OF HIGHER EDUCATION.—The term 'institution of higher education' has the meaning given such term in section 101 of the Higher Education Act of 1965 (20 U.S.C. 1001).

''(9) NATIONAL LABORATORY.—The term 'National Laboratory' has the meaning given such term in section 2 of the Energy Policy Act of 2005 (42 U.S.C. 15801).

''(10) OPEN GOVERNMENT DATA ASSET.—The term 'open Government data asset' has the meaning given such term in section 3502 of title 44, United States Code.''.

(b) CONFORMING AMENDMENT.—Subsection (f) of section 22A of the National Institute of Standards and Technology Act (15 U.S.C. 278h-1) is repealed.

○