

119TH CONGRESS
2D SESSION

S. 4069

To direct the Director of the National Institute of Standards and Technology to establish definitions, standards, resources, and frameworks to ensure certain biological datasets are ready for use in artificial intelligence models, and for other purposes.

IN THE SENATE OF THE UNITED STATES

MARCH 12, 2026

Mr. YOUNG (for himself and Mr. LUJÁN) introduced the following bill; which was read twice and referred to the Committee on Commerce, Science, and Transportation

A BILL

To direct the Director of the National Institute of Standards and Technology to establish definitions, standards, resources, and frameworks to ensure certain biological datasets are ready for use in artificial intelligence models, 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 “AI-Ready Bio-Data
5 Standards Act”.

1 **SEC. 2. DEFINITIONS, STANDARDS, RESOURCES, AND**
2 **FRAMEWORKS BY THE NATIONAL INSTITUTE**
3 **OF STANDARDS AND TECHNOLOGY FOR CER-**
4 **TAIN BIOLOGICAL DATASETS.**

5 (a) ESTABLISHMENT.—

6 (1) IN GENERAL.—Not later than 2 years after
7 the date of the enactment of this Act, the Director
8 of the National Institute of Standards and Tech-
9 nology (in this section referred to as the “Direc-
10 tor”), pursuant to recommendations from the advi-
11 sory group under subsection (f) and taking into ac-
12 count any feedback received under subsection (e),
13 shall establish definitions, standards, resources, and
14 frameworks to ensure each biological dataset gen-
15 erated as a result of qualified federally funded re-
16 search is artificial intelligence-ready.

17 (2) REQUIREMENTS FOR DEFINITIONS, STAND-
18 ARDS, RESOURCES, AND FRAMEWORKS.—

19 (A) DEFINITIONS.—

20 (i) IN GENERAL.—In carrying out
21 paragraph (1), the Director shall establish
22 definitions for the following terms:

23 (I) Artificial intelligence-ready.

24 (II) Biomanufacturing.

25 (III) Biotechnology.

(IV) Qualified federally funded research.

(ii) REQUIREMENTS FOR DEFINITION OF ARTIFICIAL INTELLIGENCE-READY.—

(I) IN GENERAL.—In defining “artificial intelligence-ready” under clause (i)(I), the Director shall develop a definition that, when applied to a biological dataset, requires that the dataset is generated and formatted in a manner that—

(aa) enables the effective use of the dataset for training artificial intelligence models; and

(bb) supports advancements in research relating to artificial intelligence and biotechnology.

(II) DISCRETION.—With respect to a biological dataset that otherwise meets the definition of “artificial intelligence-ready” established under clause (i)(I), the Director may, in consultation with the Chief Data Officer of the Federal agency that is responsible for such biological dataset, deter-

mine that such dataset is not artificial intelligence-ready.

(iii) REQUIREMENTS FOR DEFINITION OF QUALIFIED FEDERALLY FUNDED RESEARCH.—In defining “qualified federally funded research” under clause (i)(IV), the Director shall include certain conditions that, if satisfied, will result in certain federally funded research being qualified federally funded research. Such conditions shall include the following:

(I) The amount of Federal funding awarded to a recipient.

(II) The capability of the recipient to generate a biological dataset, which may include negative data, that is artificial intelligence-ready, regardless of the ability of the recipient to publish such dataset.

(III) The expertise of the recipient in generating a biological dataset that is artificial intelligence-ready.

(IV) The size of the biological dataset generated by the recipient.

1 (V) Any other condition the Di-
2 rector considers appropriate.

3 (B) STANDARDS.—In carrying out para-
4 graph (1), the Director shall establish stand-
5 ards relating to making biological datasets arti-
6 ficial intelligence-ready, in accordance with the
7 definition of artificial intelligence-ready estab-
8 lished under subparagraph (A)(i)(I) of this
9 paragraph.

10 (C) RESOURCES AND FRAMEWORKS.—In
11 carrying out paragraph (1), the Director shall
12 establish data management resources and cy-
13 bersecurity frameworks for the following:

14 (i) Federal departments and agencies
15 that provide either full or partial Federal
16 funding for research that generates biologi-
17 cal datasets.

18 (ii) Federally funded researchers who
19 are collecting, cleaning, curating, or gener-
20 ating biological datasets to make the
21 datasets artificial intelligence-ready.

22 (D) ADDITIONAL REQUIREMENTS.—The
23 Director shall ensure that the definitions,
24 standards, resources, and frameworks estab-
25 lished under paragraph (1)—

1 (i) are not overly burdensome on re-
2 cipients of funding for qualified federally
3 funded research, such that the act of gen-
4 erating biological datasets that are artifi-
5 cial intelligence-ready requires resources
6 and expertise beyond those available to the
7 recipients; and

8 (ii) are tested and evaluated in ac-
9 cordance with subsection (c) and in con-
10 sultation with—

11 (I) the head of any Federal agen-
12 cy the Director considers appropriate;

13 (II) representatives from any pri-
14 vate sector entity the Director con-
15 siders appropriate; and

16 (III) any biotechnology re-
17 searcher the Director considers appro-
18 priate.

19 (3) ANNUAL UPDATES.—Not later than 1 year
20 after the establishment of the definitions, standards,
21 resources, and frameworks under paragraph (1), and
22 annually thereafter, the Director shall review the
23 definitions, standards, resources, and frameworks
24 and, if the Director considers it appropriate, shall
25 update the definitions, standards, resources, and

1 frameworks in accordance with the requirements of
2 this section.

3 (4) CONSULTATION.—To facilitate the estab-
4 lishment of the definitions, standards, resources, and
5 frameworks under paragraph (1), and any update
6 under paragraph (3), the Director shall consult with
7 the following:

8 (A) Private sector entities from the bio-
9 technology industry.

10 (B) Private sector entities from the fron-
11 tier artificial intelligence model industry.

12 (C) Members of academia.

13 (D) The following heads of Federal agen-
14 cies that provide funding for qualified federally
15 funded research relating to the generation of bi-
16 ological datasets:

17 (i) The Secretary of Agriculture.

18 (ii) The Secretary of Defense.

19 (iii) The Secretary of Energy.

20 (iv) The Director of the National Aer-
21 onautics and Space Administration.

22 (v) The Director of the National Insti-
23 tutes of Health.

24 (vi) The Administrator of the Na-
25 tional Science Foundation.

1 (vii) The head of any other Federal
2 agency the Director considers appropriate.

3 (5) PERSONNEL.—To facilitate the establish-
4 ment of the definitions, standards, resources, and
5 frameworks under paragraph (1), and any update
6 under paragraph (3), the Director shall hire staff as
7 the Director determines necessary.

8 (b) INFORMATION GATHERING.—

9 (1) IN GENERAL.—Not later than 1 year after
10 the date of the enactment of this Act, the Director
11 shall inventory the following:

12 (A) Existing biotechnology standards uti-
13 lized by recipients of Federal funding for bio-
14 technology research to generate biological
15 datasets.

16 (B) Existing biological datasets generated
17 by recipients of Federal funding for bio-
18 technology research.

19 (2) PUBLICATION.—Not later than 1 year after
20 the Director completes the inventory requirement
21 under paragraph (1), the Director shall make any
22 information inventoried under that paragraph avail-
23 able to the public through a website of the National
24 Institute of Standards and Technology.

1 (c) TEST AND EVALUATION.—Not later than 1 year
2 after the date of the enactment of this Act, and not less
3 frequently than every 2 years thereafter, the Director shall
4 coordinate with the Administrator of the National Science
5 Foundation to conduct a test and evaluation of the defini-
6 tions, standards, resources, and frameworks established
7 pursuant to subsection (a)(1) on a sample of biological
8 datasets generated as a result of qualified federally funded
9 research to determine the following:

10 (1) Whether the definitions, standards, re-
11 sources, and frameworks are clearly written, easy to
12 follow, and easily applicable for the generation of bi-
13 ological datasets that are artificial intelligence-ready.

14 (2) Whether compliance with the definitions,
15 standards, resources, and frameworks established
16 under subsection (a)(1) when generating or curating
17 biological datasets results in an undue burden on the
18 recipients of such qualified federally funded re-
19 search, and if so, how to modify the definitions,
20 standards, resources, and frameworks, as the case
21 may be, so as to reduce the burden.

22 (d) AGENCY-SPECIFIC DATA MANAGEMENT POLI-
23 CIES.—

24 (1) IN GENERAL.—Not later than 2 years after
25 the date of the enactment of this Act and in accord-

1 ance with requirements under subsection (e), the Di-
2 rector shall establish or, if already established, re-
3 view and revise agency-specific data management
4 policies for each Federal agency that provides fund-
5 ing for qualified federally funded research to ensure
6 implementation of policies that require that any bio-
7 logical dataset generated by a recipient of qualified
8 federally funded research is artificial intelligence-
9 ready.

10 (2) ELEMENTS.—The data management poli-
11 cies described in paragraph (1) shall include the fol-
12 lowing:

13 (A) A mechanism to ensure sufficient Fed-
14 eral funding to a recipient to satisfy the re-
15 quirements of the definitions, standards, re-
16 sources, and frameworks established under sub-
17 section (a)(1).

18 (B) A process for the Chief Data Officer
19 of each Federal agency to designate an indi-
20 vidual of such agency to ensure compliance with
21 the policies established or revised under para-
22 graph (1).

23 (3) OVERSIGHT MECHANISMS.—As part of the
24 agency-specific data management policies under

1 paragraph (1), the Director shall establish the fol-
2 lowing:

3 (A) A regularly updated central repository
4 of the policies established at each Federal agen-
5 cy, made available to the public as the Director
6 determines appropriate, for the purpose of
7 tracking available policies.

8 (B) A publicly available database to serve
9 as a single point of access, updated as the Di-
10 rector determines necessary, on which the head
11 of any Federal agency may publish artificial in-
12 telligence-ready biological datasets.

13 (C) A reporting mechanism available to
14 each Federal agency to report to the Director
15 how the agency is complying with the policies.

16 (D) A mechanism available to each Federal
17 agency to request assistance from the Director
18 regarding compliance with the policies.

19 (e) PUBLIC INPUT AND FEEDBACK; CONSULTA-
20 TION.—In establishing the definitions, standards, re-
21 sources, and frameworks under subsection (a)(1) and the
22 agency-specific data management policies under sub-
23 section (d)(1), the Director shall carry out the following:

24 (1) Solicit input and feedback from the public
25 regarding the definitions, standards, resources, and

1 frameworks and the agency-specific data manage-
2 ment policies.

3 (2) Consult with the following heads of Federal
4 agencies that provide funding for qualified federally
5 funded research to ensure such agencies are able to
6 comply with the definitions, standards, resources,
7 and frameworks and the agency-specific data man-
8 agement policies:

9 (A) The Secretary of Agriculture.

10 (B) The Secretary of Defense.

11 (C) The Secretary of Energy.

12 (D) The Director of the National Aero-
13 nautics and Space Administration.

14 (E) The Director of the National Institutes
15 of Health.

16 (F) The Administrator of the National
17 Science Foundation.

18 (G) The head of any other Federal agency
19 as determined by the Director.

20 (f) ADVISORY GROUP.—

21 (1) IN GENERAL.—Not later than 180 days
22 after the date of the enactment of this Act, the Di-
23 rector shall establish an advisory group to carry out
24 the following:

1 (A) To provide recommendations relating
2 to the definitions, standards, resources, and
3 frameworks established under subsection (a)(1).

4 (B) To review and provide feedback on the
5 agency-specific data management policies estab-
6 lished or revised under subsection (d)(1).

7 (C) To provide recommendations to aca-
8 demic journals for guidelines relating to artifi-
9 cial intelligence-ready biological datasets.

10 (D) To solicit recommendations from the
11 academic community regarding implementation
12 of the definitions, standards, resources, and
13 frameworks.

14 (E) To provide any other guidance the Di-
15 rector may request.

16 (2) MEMBERSHIP.—

17 (A) IN GENERAL.—The advisory group es-
18 tablished under paragraph (1) shall be com-
19 posed of not fewer than 12 members to in-
20 clude—

21 (i) representatives of Federal agencies
22 that award funds to recipients to carry out
23 qualified federally funded research; and

1 (ii) representatives of academia, pri-
2 vate sector entities, and academic pub-
3 lishers.

4 (B) TERMS.—

5 (i) IN GENERAL.—Each member shall
6 serve for a term of 2 years.

7 (ii) RENEWAL.—Subject to the discre-
8 tion of the Director, the term of each
9 member may be renewed for an additional
10 2-year term.

11 (3) CHAIRPERSON.—

12 (A) APPOINTMENT.—The Chairperson of
13 the advisory group shall be designated by the
14 Director from among the members.

15 (B) TERM.—The term of the Chairperson
16 shall be 1 year.

17 (C) RENEWAL.—Subject to the discretion
18 of the Director, the term of a Chairperson may
19 be renewed for an additional 1-year term.

20 (4) REPORTS.—

21 (A) INTERIM REPORT.—Not later than 1
22 year after the date of the enactment of this Act,
23 the advisory group shall submit to the Director
24 a interim report that contains advice and guid-

1 ance with respect to the matters described in
2 paragraph (1).

3 (B) SUBSEQUENT REPORTS.—If the advi-
4 sory group or the Director determines appro-
5 priate, the advisory group shall submit to the
6 Director subsequent reports relating to the mat-
7 ters described in paragraph (1).

8 (g) FEDERAL ACQUISITION REGULATION REVI-
9 SIONS.—The Federal Acquisition Regulatory Council shall
10 revised the Federal Acquisition Regulation as necessary
11 to implement the definitions, standards, resources, and
12 frameworks established under subsection (a)(1).

13 (h) ANNUAL REPORT.—

14 (1) INTERIM REPORT.—Not later than 1 year
15 after the date of the enactment of this Act, the Di-
16 rector shall submit to Congress and the Comptroller
17 General of the United States an interim report that
18 includes information relating to the progress of es-
19 tablishing the definitions, standards, resources, and
20 frameworks under subsection (a)(1) and the agency-
21 specific data management policies under subsection
22 (d)(1).

23 (2) SUBSEQUENT REPORTS.—Not later than 2
24 years after the date of the enactment of this Act,
25 and annually thereafter, the Director shall submit to

1 Congress and the Comptroller General of the United
2 States a report that includes information relating to
3 the following:

4 (A) The establishment, implementation,
5 and, if applicable, revision, of the definitions,
6 standards, resources, and frameworks estab-
7 lished under subsection (a)(1).

8 (B) The establishment, implementation,
9 and, if applicable, revision of the agency-specific
10 data management policies established or revised
11 under subsection (d)(1).

12 (3) ADDITIONAL REQUIREMENT FOR FIRST
13 SUBSEQUENT REPORT.—With respect to the first
14 subsequent report under paragraph (2), the Director
15 shall include a summary of the testing and evalua-
16 tion under subsection (c) that includes information
17 relating to the following:

18 (A) The findings of the testing and evalua-
19 tion.

20 (B) The manner by which the Director ad-
21 dressed any concern identified as a result of the
22 testing and evaluation.

23 (C) An assessment of any burden on re-
24 cipients of funding for qualified federally fund-

1 ed research regarding ensuring that biological
2 datasets are artificial intelligence-ready.

3 (D) A cost-benefit analysis of the value of
4 ensuring that biological datasets are artificial
5 intelligence-ready in relation to any such bur-
6 den.

7 (i) GOVERNMENT ACCOUNTABILITY OFFICE RE-
8 PORT.—Not later than 5 years after the date of the enact-
9 ment of this Act, the Comptroller General of the United
10 States shall submit to Congress a report on the impact
11 of the definitions, standards, resources, and frameworks
12 established under subsection (a)(1), including the fol-
13 lowing:

14 (1) An assessment of the effectiveness of the
15 definitions, standards, resources, and frameworks in
16 ensuring each biological dataset generated by a re-
17 cipient of funding for qualified federally funded re-
18 search is artificial intelligence-ready.

19 (2) An assessment of whether the implementa-
20 tion of the definitions, standards, resources, and
21 frameworks, as the case may be, has resulted in an
22 undue burden on any recipient of Federal funding.

23 (3) Any recommendations with respect to the
24 implementation of the definitions, standards, re-
25 sources, and frameworks.

1 (j) SUNSET.—This section shall terminate on the
 2 date that is 10 years after the date of the enactment of
 3 this Act.

4 (k) DEFINITIONS.—In this section:

5 (1) BIOLOGICAL DATA.—The term “biological
 6 data” means information that is measured, collected,
 7 or aggregated for analysis, including associated
 8 descriptors, derived from the structure, function, or
 9 process of a biological system.

10 (2) BIOLOGICAL DATASET.—The term “biologi-
 11 cal dataset” means a discreet collection of biological
 12 data.

13 (3) BIOLOGICAL DATA REPOSITORY.—The term
 14 “biological data repository” means any centralized
 15 data storage capacity meant for managing or storing
 16 biological data.

17 (4) NEGATIVE DATA.—The term “negative
 18 data” means data that disproves, fails to support, or
 19 explains through previously unknown or contradic-
 20 tory means a research hypothesis but that otherwise
 21 still advances scientific knowledge or understanding.

○