

119TH CONGRESS
2D SESSION

S. _____

To direct the Secretary of Energy to establish an extremophile whole genome sequencing initiative, and for other purposes.

IN THE SENATE OF THE UNITED STATES

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

A BILL

To direct the Secretary of Energy to establish an extremophile whole genome sequencing initiative, 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 “Homestake AI Act of
5 2026”.

6 **SEC. 2. EXTREMOPHILE WHOLE GENOME SEQUENCING INI-**
7 **TIATIVE.**

8 (a) DEFINITIONS.—In this section:

1 (1) COVERED FACILITY.—The term “covered
2 facility” means the long-term storage facility
3 planned for under subsection (c)(1)(E).

4 (2) INITIATIVE.—The term “initiative” means
5 the extremophile whole genome sequencing initiative
6 established under subsection (b).

7 (3) SECRETARY.—The term “Secretary” means
8 the Secretary of Energy, acting through the Under-
9 secretary of Science.

10 (b) ESTABLISHMENT.—Not later than 90 days after
11 the date of enactment of this Act, the Secretary shall es-
12 tablish an initiative, to be known as the “extremophile
13 whole genome sequencing initiative”, that shall—

14 (1) be consistent with the goals of the Genesis
15 Mission announced in Executive Order 14363 (90
16 Fed. Reg. 55035; relating to launching the Genesis
17 Mission); and

18 (2) facilitate the collection of new genomic data
19 that will contribute to advances in scientific dis-
20 covery and may support the development of new
21 healthcare treatments.

22 (c) INITIATIVE.—

23 (1) IN GENERAL.—Under the initiative, the
24 Secretary shall—

1 (A) collect, catalogue, and sequence the
2 whole genomes and metabolites of samples from
3 animals, plants, fungi, and microbes found in
4 locations selected under paragraph (3);

5 (B) establish a database for the data col-
6 lected under subparagraph (A), or use an exist-
7 ing database, such as the database of the Na-
8 tional Center for Biotechnology Information, in-
9 cluding sequence data, genome assemblies,
10 peptides, metabolite data, annotations, and as-
11 sociated metadata, including genome and
12 transcriptome data, metabolome data, and
13 proteome data;

14 (C) collaborate with industry, academia,
15 and other government agencies to collate exist-
16 ing relevant data for inclusion in the database
17 established or used under subparagraph (B);

18 (D) make that database accessible to per-
19 sons in the United States, subject to the condi-
20 tion that the Secretary shall, to the maximum
21 extent practicable, prevent the sharing of such
22 information with adversary nations or foreign
23 persons; and

24 (E) not later than 1 year after the date of
25 enactment of this Act, develop a plan for the

1 construction of a new long-term storage facil-
2 ity—

3 (i) for, at a minimum, the samples
4 collected under subparagraph (A); and

5 (ii) in such a manner that the facility
6 may continue to operate in the case of a
7 catastrophic incident (as defined in section
8 602 of the Post-Katrina Emergency Man-
9 agement Reform Act of 2006 (6 U.S.C.
10 701)).

11 (2) THIRD PARTY COLLECTION.—The Secretary
12 may enter into contracts with third party entities,
13 including institutions of higher education or private
14 sector entities, to collect samples under paragraph
15 (1)(A).

16 (3) LOCATIONS FOR COLLECTION.—

17 (A) IN GENERAL.—Not later than 180
18 days after the date of enactment of this Act,
19 the Secretary shall select not fewer than 5 loca-
20 tions from which to collect samples under para-
21 graph (1)(A), at least 1 of which shall be a
22 mine or former mine that—

23 (i) is deeper than 4,500 feet; and

24 (ii) houses an underground science fa-
25 cility.

1 (B) CONSIDERATIONS.—

2 (i) IN GENERAL.—The Secretary shall
3 select locations under subparagraph (A)—

4 (I) that are unique or are ex-
5 posed to unique properties, such as
6 extreme heat or pressure;

7 (II) that are under risk of de-
8 struction before 2050 from natural or
9 unnatural forces;

10 (III) based on the potential for
11 finding useful organisms that will con-
12 tribute to advances in scientific re-
13 search and discovery; and

14 (IV) that may support the devel-
15 opment of new healthcare treatments
16 relating to cancer, including cancers
17 affecting children and individuals with
18 high mortality rates, diabetes, and de-
19 mentia, including Alzheimer's disease.

20 (ii) PREFERENCE.—In selecting loca-
21 tions under subparagraph (A), the Sec-
22 retary may give preference to a location on
23 land owned or controlled by the Depart-
24 ment of Energy (including National Lab-
25 oratories).

1 (iii) RANGE OF EXTREME CONDI-
2 TIONS.—In selecting locations under sub-
3 paragraph (A), the Secretary shall ensure
4 that the locations reflect a broad range of
5 extreme conditions.

6 (4) LONG-TERM STORAGE FACILITY.—In select-
7 ing a location for the covered facility under the plan
8 developed under paragraph (1)(E), the Secretary
9 shall consider, among other factors determined rel-
10 evant by the Secretary—

11 (A) colocation in a Department of Energy
12 facility or a facility that receives significant
13 funding from the Department of Energy;

14 (B) the benefits of storing the applicable
15 materials underground in a remote area; and

16 (C) the ability of the location to withstand
17 natural disasters, such as earthquakes,
18 wildfires, or meteor strikes.

19 (5) DATA STANDARDS.—In carrying out the ini-
20 tiative, the Secretary shall—

21 (A) leverage existing and emerging data
22 standards to ensure—

23 (i) high-quality data are obtained
24 under the initiative; and

1 (ii) data collected under the initiative
2 are considered usable and interoperable,
3 with appropriate metadata to be included;
4 (B) ensure that data outputs under the
5 initiative—

6 (i)(I) meet applicable Federal data
7 standards for biological data; or

8 (II) if there are no applicable Federal
9 data standards described in subclause (I),
10 meet data standards for the initiative es-
11 tablished by the Secretary, in consultation
12 with the National Institute of Standards
13 and Technology and other relevant agen-
14 cies, based on findability, accessibility,
15 interoperability, and reusability data prin-
16 ciples and using industry best practices;
17 and

18 (ii) are updated to comply with any
19 applicable United States Geological Survey
20 standards for biological data; and

21 (C) ensure that appropriate cybersecurity
22 and access safeguards are implemented to pro-
23 tect the data generated or stored under the ini-
24 tiative, including by ensuring that data outputs
25 under the initiative are updated as necessary to

1 meet applicable Federal cybersecurity standards
2 for genomic data, including controls consistent
3 with the National Institute of Standards and
4 Technology special publications 800–53 (relat-
5 ing to Security and Privacy Controls for Infor-
6 mation Systems and Organizations) and 800–
7 111 (relating to Guide to Storage Encryption
8 Technologies for End User Devices) or any suc-
9 cessor framework.

10 (6) LONG-TERM STORAGE REQUIREMENTS.—

11 The Secretary may consult with the Secretary of Ag-
12 riculture and the Director of the National Museum
13 of Natural History of the Smithsonian Institution
14 with respect to storage capabilities for the manage-
15 ment and long-term storage of samples collected
16 under the initiative, subject to the following:

17 (A) Access to samples collected under the
18 initiative may be managed by the respective in-
19 stitutional collection managers, with—

20 (i) the natural history collections
21 managed by the Director of the National
22 Museum of Natural History of the Smith-
23 sonian Institution; and

1 (ii) the germplasm repository and the
2 covered facility managed by the Secretary
3 of Agriculture.

4 (B) Before any samples are stored at the
5 covered facility, the Secretary shall—

6 (i) evaluate the species to be sampled
7 to identify which species need to be col-
8 lected for long-term storage; and

9 (ii) from among the species identified
10 under clause (i), identify which species
11 should be considered to be high-priority
12 species.

13 (C) Any samples identified for long-term
14 storage under subparagraph (B)(i)—

15 (i) shall be collected and stored at the
16 covered facility after the covered facility is
17 operational; and

18 (ii) may be stored in an alternative lo-
19 cation until that date.

20 (D) Any high-priority species identified
21 under subparagraph (B)(ii) shall be collected
22 for storage in each of—

23 (i) the natural history collections to be
24 managed by the Director of the National

1 Museum of Natural History of the Smith-
2 sonian Institution;

3 (ii) the germplasm repository to be
4 managed by the Secretary of Agriculture;
5 and

6 (iii) the covered facility.

7 (E) Any physical samples collected under
8 the initiative shall not be transferred, exported,
9 or loaned outside the United States.

10 (F) Any physical storage, curation, and
11 laboratory processing of samples collected under
12 the initiative shall occur within facilities located
13 in the United States.

14 (G) The Secretary may accept donated
15 samples for storage from third parties or other
16 countries.

17 (7) SEQUENCING AND DATA PROCESSING.—In
18 carrying out the initiative, the Secretary shall—

19 (A) collaborate with the Secretary of Agri-
20 culture, the Director of the National Museum of
21 Natural History of the Smithsonian Institution,
22 and the heads of other applicable Federal agen-
23 cies to sequence genomic information from sam-
24 ples collected under the initiative;

1 (B) develop and use standardized sequenc-
2 ing protocols to ensure consistency and com-
3 parability of data; and

4 (C) incorporate automated data processing
5 and quality checks to ensure the reliability of
6 the data.

7 (8) TERMINATION OF COLLECTION.—No collec-
8 tion of samples under paragraph (1)(A) may occur
9 by the Secretary or a third party pursuant to a con-
10 tract entered into under paragraph (2) after the
11 date that is 5 years after the date of enactment of
12 this Act.

13 (d) COORDINATION WITH OTHER ENTITIES.—

14 (1) IN GENERAL.—In carrying out the initia-
15 tive, the Secretary may coordinate with—

16 (A) other offices within the Department of
17 Energy;

18 (B) the Department of Agriculture;

19 (C) any other relevant agency that may
20 benefit from utilizing the covered facility;

21 (D) the Smithsonian Institution, including
22 the National Museum of Natural History;

23 (E) the Department of Health and Human
24 Services, including the Centers for Disease Con-
25 trol and Prevention and the National Institutes

1 of Health, including the Childhood Cancer Data
2 Initiative;

3 (F) academia;

4 (G) nonprofit organizations, including or-
5 ganizations focused on the collection of data for
6 the discovery of treatments for cancer, diabetes,
7 and dementia, including Alzheimer's disease;

8 (H) nonprofit organizations that support
9 biodiversity conservation, including nonprofit
10 organizations in allied countries described in
11 subsection (e)(2); and

12 (I) private sector entities, including entities
13 in the pharmaceutical, biotechnology, and artifi-
14 cial intelligence industries.

15 (2) PRIVATE SECTOR CONTRIBUTIONS.—There
16 shall be established a foundation, which shall be an
17 organization described in section 501(c)(3) of the
18 Internal Revenue Code of 1986 and exempt from
19 taxation under section 501(a) of that Code, to re-
20 ceive in-kind contributions from biotechnology and
21 other private sector companies to assist in carrying
22 out the initiative and supporting the operation of the
23 covered facility.

1 (3) OTHER AGENCY RESOURCES.—The Sec-
2 retary may use the services, personnel, and facilities
3 of another agency if agreed to by that agency.

4 (e) WORKING GROUP.—

5 (1) IN GENERAL.—The Secretary shall establish
6 a working group to coordinate with private and pub-
7 lic sector initiatives in the United States and allied
8 countries described in paragraph (2) collecting rel-
9 evant genomic data to create alignment between sci-
10 entific groups—

11 (A) to maximize beneficial research output;

12 and

13 (B) to minimize duplicative spending of the
14 initiative.

15 (2) ALLIED COUNTRIES DESCRIBED.—The al-
16 lied countries referred to in paragraph (1) are—

17 (A) Australia;

18 (B) member countries of the European
19 Union;

20 (C) Japan;

21 (D) South Korea;

22 (E) New Zealand; and

23 (F) the United Kingdom.

1 (f) AUTHORIZATION OF APPROPRIATIONS.—There
2 are authorized to be appropriated to the Secretary such
3 sums as are necessary to carry out this section.