To establish a Secure Research Data Network.

IN THE SENATE OF THE UNITED STATES

Mr. Wyden (for himself, Mr. Portman, Mr. Boozman, and Mrs. Gillibrand) introduced the following bill; which was read twice and referred to the Committee on

A BILL

To establish a Secure Research Data Network.

1 Be it enacted by the Senate and House of Representa-
2 tives of the United States of America in Congress assembled,
3
4 SECTION 1. SHORT TITLE.
5 This Act may be cited as the “Secure Research Data
6 Network Act”.
7
8 SECTION 2. DEFINITIONS.
9 In this Act:
10 (1) Section 3561 of title 44 definitions.—
11 The terms “evidence”, “identifiable form”, “statis-
12 tical activities”, and “statistical purpose” have the
meanings given the terms in section 3561 of title 44, United States Code.

(2) ANALYST.—The term “analyst” means a person either employed by, or working on behalf of, a Federal or State agency in empirical programmatic or data analysis.

(3) DATA ASSET.—The term “data asset” has the meaning given the term in section 3502 of title 44, United States Code.

(4) DATA STEWARD.—The term “data steward” means an individual employed by a Federal or State agency who is familiar with the agency’s data, and has a statutory responsibility to protect the confidentiality of such data and ensure its integrity and quality.

(5) DIRECTOR.—Except as otherwise provided, the term “Director” means the Director of the National Science Foundation.

(6) REPORTING ENTITY.—The term “reporting entity” means a Federal or State agency with data relevant to governmentwide evidence-building activities.

(7) STATE.—The term “State” has the meaning given the term in section 502 of the National Science Foundation Authorization Act of 2010 (42
U.S.C. 1862p note), except that the definition shall be applied by striking “, or any other territory or possession of the United States”.

**SEC. 3. SECURE RESEARCH DATA NETWORK.**

(a) **Establishment.**—

(1) **In general.**—The Director, in consultation with the Statistical Official of the National Science Foundation, the Chief Statistician of the United States, and the Director of the National Artificial Intelligence Initiative Office, shall, subject to the availability of appropriations, enter into an agreement for the establishment of the “Secure Research Data Network” or “SRDN”, which shall be operated as a pilot program.

(2) **Term of pilot program.**—The pilot program operated under paragraph (1) shall run for 3 years, with the possibility of not more than two 1-year extensions, upon consideration of the Director, in consultation with the SRDN Advisory Board.

(3) **Additional employees.**—The Director may hire additional employees as necessary to support the operation of the SRDN, including full-time equivalent Federal employees.

(4) **Avoid duplication.**—The Director shall coordinate with the Chief Statistician of the United
States to identify potential areas of overlap between the SRDN and efforts carried out at, or financially assisted by, the National Science Foundation, such as the America’s DataHub Consortium, on the date of enactment of this Act. The Chief Statistician of the United States shall seek to ensure that the activities of the SRDN enhance and complement those efforts existing on the date of enactment of this Act in order to avoid duplication and maximize the use of Federal resources.

(b) Responsibilities.—The Director shall direct the SRDN to carry out the following:

(1) Support governmentwide evidence-building activities as required under section 312 of title 5, United States Code, including implementation of agency multiyear evidence-building plans.

(2) Develop, deploy, maintain, and operate a SRDN platform for authorized analysts to calculate statistics on data for evidence-building activity purposes using data assets made available by reporting entities for approved projects.

(3) Execute a number of approved projects on the SRDN platform described in paragraph (2) and make the results publicly available.
(4) Ensure an appropriate number of approved projects will re-examine and attempt to fully or partially replicate the results of linked data studies in existence on the date of enactment of this Act as proof of concept for the SRDN platform described in paragraph (2).

(5) Maintain, in consultation with the Chief Statistician of the United States and other relevant Federal data strategy stakeholders, a public SRDN website with up-to-date information on all approved projects, including their results and documentation of the evidence-building value of each project for policymakers.

(6) Consult with the National Artificial Intelligence Research Resource Task Force established under section 5106 of the National Artificial Intelligence Initiative Act of 2020 (15 U.S.C. 9415) and consider how to integrate the Task Force’s recommendations and road map for expanding access to critical artificial intelligence resources and educational tools into the SRDN.

(c) PRIVACY REQUIREMENTS.—In developing the SRDN platform under subsection (b)(2), the SRDN—

(1) shall ensure the SRDN platform facilitates statistical activities for evidence-building activity
purposes while reducing the privacy and security
risks by developing, procuring, or adapting tech-
tology that, at a minimum, uses the latest cutting-
edge technical protection measures that reasonably
ensure that—

(A) the SRDN platform permits only au-
thorized analysts to perform statistical queries
necessary to answer approved project questions
using the data assets made available by the re-
porting entities;

(B) no information about the data assets
used in the SRDN platform is revealed to any
other party, except as incorporated into the
final result, which shall be used exclusively for
statistical evidence-building purposes and shall
not be released in an identifiable form;

(C) no individual entity’s data or informa-
tion is revealed by the SRDN platform to any
other party in an identifiable form;

(D) the SRDN platform prohibits any
other queries by the SRDN or any other party
through the SRDN platform; and

(E) the SRDN platform minimizes the pri-
vacy risks to individual entities whose data has
been made available by a reporting entity, in-
cluding those that could result from data breaches of any system operated by the reporting entity; and
(2) may—
   (A) use secure multiparty computation technologies; or
   (B) utilize technology other than secure multiparty computation technologies if the other technology—
      (i) fully complies with subparagraphs (A) through (E) of paragraph (1); and
      (ii) delivers greater or equivalent privacy and security than secure multiparty computation.

(d) SOFTWARE REQUIREMENTS.—
   (1) IN GENERAL.—The Director shall ensure the SRDN develops, deploys, operates, and maintains the SRDN platform described in subsection (b)(2), along with corresponding Application Programming Interfaces (APIs), to be used by reporting entities and authorized analysts who will interact with the SRDN platform to conduct the approved projects. The Director—
      (A) shall direct the SRDN to consult, design, and conduct usability testing of the
SRDN platform with relevant Federal and
State agencies, Federal coordinating councils,
subject matter experts, academia, and others
with expertise in technology development, main-
tenance, and governance, statistics, privacy, and
user-centered design, as the Director deter-
mines appropriate;

(B) in consultation with the SRDN Advi-
sory Board, shall engage in an open public re-
view and comment process on the development
of the SRDN platform and its governance poli-
cies; and

(C) shall enter into an agreement for the
establishment of the SRDN only with entities
based in the United States or in its allied coun-
tries.

(2) Public domain and open source soft-
ware.—The Director shall ensure the SRDN makes
all software developed for the SRDN platform de-
scribed in subsection (b)(2) available as public do-
main and open source software (as defined in section
1552.239–71 of title 48, Code of Federal Regu-
lations, or a successor regulation) both during develop-
ment and after completion, and endeavor to design
the architecture to ensure that appropriate compo-
ments can be reused independently. The SRDN shall publicly document the construction, operation, and functionality of the software technologies it develops on the SRDN public-facing website.

(3) PROVISION OF SOFTWARE.—The Director shall ensure the SRDN provides software to reporting entities, at no cost, that the reporting entities can use to connect their own systems to the SRDN platform described in subsection (b)(2). A reporting entity may use the public domain software the SRDN makes available to build their own software that interfaces using the publicly documented API, or use the services of another agency or organization with greater or equivalent privacy and security to help them connect their own systems to the SRDN.

(e) DATA QUALITY SERVICE TEAM.—

(1) IN GENERAL.—The Director shall direct the SRDN to develop a plan for and operate a data quality service team that is composed of data governance, information systems, statistics, cybersecurity, and disclosure avoidance experts, who will, at no cost to the reporting entity, help reporting entities evaluate their data and prepare it for use with the SRDN platform described in subsection (b)(2) to achieve approved project goals.
(2) AGENCY ASSISTANCE.—A reporting entity that is a Federal agency, and a reporting entity that is a State agency that chooses to receive assistance as described in paragraph (1), shall work with the SRDN to develop a plan for preparing its data for use with the SRDN platform described in subsection (b)(2), including adopting all necessary standards. The SRDN shall approve the cost estimates prepared by the reporting entity prior to the reporting entity and the SRDN undertaking work that is eligible for reimbursement, according to guidelines established by the Director. The Director shall approve the plans and enter into reimbursable agreements with reporting entities for expenses included in the approved cost plan. In addition, the reporting entity shall—

(A) make its employees responsible for the relevant data available to work with and assist the data quality service team for the extent of the project with reimbursement from the National Science Foundation for the employees’ worked hours;

(B) adopt the recommendations of the data quality service team necessary to prepare the
reporting entity’s data for use with the SRDN platform; and

(C) notwithstanding subparagraphs (B), (C), and (D) of subsection (c)(1), provide the data quality service team with access to the relevant data.

(3) **HARDWARE AND SOFTWARE SUPPORT.**—The Director shall direct the SRDN to provide hardware and software support technology the reporting entities need to stage and prepare data for use with the SRDN platform described in subsection (b)(2).

(4) **TRAINING MATERIALS AND TOOLS.**—The data quality service team, in coordination with the Secure Research Data Network training team described in subsection (f)(1), shall—

(A) produce training materials, documented runnable code, and other tools to help reporting entities prepare their data for use with the SRDN platform described in subsection (b)(2); and

(B) publish such resources on the SRDN public website.

(5) **DISCLOSURE AVOIDANCE.**—The data quality service team shall assist reporting entities as they conduct a disclosure avoidance review to ensure
that project results are not released in an identifiable form. No results shall be released until a disclosure avoidance review is conducted.

(f) TRAINING.—

(1) IN GENERAL.—The Director shall direct the SRDN to develop a plan for and operate a Secure Research Data Network training team that is composed of data science, social science, statistics, privacy, disclosure avoidance, and cybersecurity experts, which will, at no cost to the reporting entity, help reporting entities develop capacity to produce evidence using the SRDN platform described in subsection (b)(2), explain how the SRDN platform works and how it protects data assets, and evaluate the value of the evidence for policymakers and the public.

(2) TRAINING CURRICULA.—The Director shall—

(A) develop, in consultation with relevant Federal and State agencies, Federal and State coordinating councils, subject matter experts, academia, and others with expertise in user-centered design, privacy preserving technologies, data science, and statistics design, as the Director determines appropriate, training curricula
for agency staff and authorized analysts and
make it publicly available; and

(B) in consultation with the SRDN Advisory Board, engage in an open public review
and comment process on the development of the curricula.

(3) CURRICULA CONTENT.—The curricula developed under paragraph (2) shall build upon Federal data strategy and Office of Management and Budget evidence-building recommendations and include training in the use of the SRDN platform described in subsection (b)(2), preparation of data for use with the SRDN platform, testing and evaluation of the usefulness of the training materials and tools, and documentation of the evidence value for policymakers and the public.

(g) PROJECT PROPOSAL.—

(1) IN GENERAL.—The Director, in consultation with the SRDN Advisory Board and the Chief Statistician of the United States, shall develop criteria and guidelines for analysts to become authorized analysts and for project proposals to be submitted for consideration.

(2) FACILITATING PROPOSALS.—The Director shall facilitate project proposals from research com-
munities by soliciting questions and connecting re-
search communities with analysts from the appro-
priate reporting agencies through methods such as
workshops, conferences, or idea labs.

(3) PROPOSAL REQUIREMENTS.—The project
proposals shall be submitted by authorized analysts
and, at a minimum, include the following:

(A) Documentation of the relevant data as-
sets necessitated by the project, including de-
tails of their level of preparedness for analysis
with the SRDN platform under subsection
(b)(2).

(B) Identification of data stewards from
the relevant reporting entities who will work
with the data quality team to prepare data as-
sets for analysis with the SRDN platform under
subsection (b)(2).

(C) Attestation from the relevant reporting
entities and data stewards that they support
both the proposed project and the usage of
their data assets for the proposed project.

(D) Documentation of the evidence-build-
ing value the project would provide to policy-
makers.
(E) Feedback and comments on the proposed project collected from nonprofit organizations, Tribal communities and governments, relevant State and local governments, community leaders, and other members of the public, as appropriate.

(h) ADVISORY BOARD.—

(1) IN GENERAL.—The Director, in consultation with the Directorate for Social, Behavioral and Economic Sciences of the National Science Foundation, the Directorate for Computer and Information Science and Engineering of the National Science Foundation, and the Chief Statistician of the United States, shall establish a SRDN Advisory Board. The SRDN Advisory Board shall be responsible for receiving, evaluating, advising, and prioritizing a diverse set of project proposals based on National Science Foundation strategic priorities and established evidence-building plans and policy-relevant questions required of Federal agencies in accordance with section 312 of title 5, United States Code, from multiple different authorized analysts for the consideration of the Director.

(2) MEMBERS.—
(A) IN GENERAL.—The SRDN Advisory Board shall consist of 15 members from a broad range of specialties and institutions, including individuals with expertise in producing high-value evidence, data stewardship, cybersecurity, privacy, data governance, State and Federal program evaluation, State and Federal data infrastructure, State and Federal statistics infrastructure, and social science research.

(B) FEDERAL EMPLOYEES AND NON-FEDERAL EMPLOYEES.—The SRDN Advisory Board shall consist of members who are employed by a Federal agency and members who are not employed by a Federal agency.

(C) DIVERSITY.—The SRDN Advisory Board shall have a diverse membership based on gender, race, ethnicity, and geography. One-third of the members of the SRDN Advisory Board shall consist of individuals located in jurisdictions that participate in the program under section 113 of the National Science Foundation Authorization Act of 1988 (42 U.S.C. 1862g).

(3) DUTIES.—The SRDN Advisory Board shall consider, at a minimum, the following:
(A) The feasibility of the proposed project, including preparedness of necessary data components, formal support from relevant data stewards and reporting entities, plans and resources to implement an adequate data governance plan, and reporting entity technological capabilities.

(B) How the proposed project supports established evidence-building plans and answers policy-relevant questions required of Federal agencies in accordance with section 312 of title 5, United States Code, and to what extent the project represents a diverse group of data sources and statistical work in the Federal Government or in Federal-State partnerships.

(C) Feedback and comments included in the project proposal.

(i) Reporting.—

(1) In general.—Not later than 3 years after the date of enactment of this Act, the Director, in consultation with the SRDN Advisory Board, shall produce a final report, to be published on the SRDN website and shared with relevant committees of Congress, which includes—
(A) the technological considerations of the SRDN platform described in subsection (b)(2);

(B) recommendations for future SRDN projects;

(C) a summary of all the SRDN projects undertaken, including their results, details of the data sources used, and analysis of their disparate impacts on subgroups;

(D) a description of how project results are relevant to Federal agency evidence-building plans and policy-relevant questions, as required under section 312 of title 5, United States Code;

(E) lessons learned by the data quality service team, operated pursuant to subsection (e)(1), from working with reporting entity stakeholders to prepare their data for the SRDN platform described in subsection (b)(2);

(F) recommendations for a permanent program that would be called the Secure Research Data Network, including needs for additional capacity, functionality, and funding related to providing a secure privacy preserving statistical platform; and
(G) consideration of relevant recommendations from the Office of Management and Budget Advisory Committee on Data for Evidence-Building.

(2) **GAO EVALUATION.**—

(A) **IN GENERAL.**—Upon publication of the report under paragraph (1), the Comptroller General of the United States shall conduct an evaluation of the SRDN pilot program, which shall include—

(i) a recommendation for a potential permanent SRDN program;

(ii) a technical review of the SRDN pilot program’s ability to protect individual identity from disclosure and recommendations for how a future permanent SRDN program should be constructed in order to provide a secure privacy preserving statistical platform;

(iii) an analysis of the adequacy of allotted resources, issues with the solicitation of projects and public comment, and any issues faced in facilitating the collaboration of reporting entities involved in approved projects; and
(iv) a review of, and recommendations for, how a permanent SRDN program will comply with relevant privacy statutes.

(j) Requirements.—The agreement entered into under subsection (a) shall—

(1) be competitively awarded;
(2) last not more than 5 years;
(3) ensure any entity designated to operate the SRDN shall—

(A) coordinate with the Statistical Official of the National Science Foundation, as part of the Statistical Official’s role as an agent of the National Center for Science and Engineering Statistics and a member of the Interagency Council on Statistical Policy, in accordance with section 314(b) of title 5, United States Code, and section 3504(e)(8) of title 44, United States Code; and

(B) comply with applicable requirements provided in section 552a of title 5, United States Code (commonly known as the “Privacy Act of 1974”) and subchapter III of chapter 35 of title 44, United States Code (commonly known as the “Confidential Information Protec-
and

(4) if practicable, establish the pilot program as a Federally funded research and development center.

(k) AUTHORIZATION OF APPROPRIATIONS.—There is authorized to be appropriated to carry out this section $100,000,000.