United States Senate

WASHINGTON, DC 20510

December 20, 2022

The Honorable Sethuraman Panchanathan, Ph.D.
Director
National Science Foundation
2415 Eisenhower Avenue
Alexandria, Virginia 22314

Dear Director Panchanathan:

We write to you as the lead authors of the privacy requirements for the recently authorized National Secure Data Service (NSDS), to make clear that the National Science Foundation should secure Americans' data within the NSDS platform using advanced encryption technology.

The NSDS platform will enable government agencies to collaborate by using data for research projects. This research will help policy makers improve government programs, and will shed light on the effectiveness of federal policies. However, the NSDS program will only live up to its promise if it facilitates research while protecting Americans' data from hackers, foreign spies and misuse by government agencies. For that reason, Congress made sure that the CHIPS And Science Act, which created the program, requires NSF to engineer strong privacy and security protections into the NSDS platform.

The NSDS platform is required to use technologies that ensure no individual entity's data or information is revealed to any other party in an identifiable form. This requirement ensures that it can't become a warehouse of sensitive, identifying data about Americans. The 2014 breach of tens of millions of records held by the federal Office of Personnel Management demonstrated that digital warehouses of government data become targets for hacking and theft by cyber adversaries, and Congress was determined to avoid repeating this mistake. To ensure identifiable data within the platform is inaccessible to any agency other than the one who originally provided it — including NSF itself — NSF should require agencies to encrypt the information using an encryption key only they control. If sensitive data is encrypted when it enters the platform and throughout its storage there, individuals who appear in that data are protected in the event of a hack or breach of the NSDS system. And, by avoiding holding a "master key" that can access all of the data, NSF will remove a massive cyber-target from its back.

Requiring agencies to encrypt their identifiable data within the NSDS platform doesn't preclude the NSDS project from supporting important, cross-agency research that uses individual-level data records. Privacy-enhancing technologies such as multi-party computation, already in use in the commercial sector, make it possible for organizations

to collaborate on research without sharing unencrypted data. With multi-party computation, the NSDS program can support vital research that relies on sensitive data, such as studying the efficacy of programs to help our nation's veterans, without requiring agencies to share individuals' sensitive data. Multi-party computation provides stronger privacy and security guarantees than methods such as data de-identification: numerous studies have shown that many de-identification techniques are reversible, and the FTC recently warned businesses that, "claims that data is 'anonymous' or 'has been anonymized' are often deceptive." Multi-party computation will also help NSF fulfill other security requirements in statute, such as the requirement that the platform prohibit any queries that aren't run by authorized analysts and designed to answer approved project questions. It was important to Congress that Americans be able to trust that their data won't be used for any purpose that hasn't been reviewed by NSF and publicized on the program's website. NSF must use encryption technology to back up that trust with hard technical guarantees.

The use of privacy-enhancing encryption technology isn't just the best way for NSF to comply with the requirements in statute — it's also an opportunity for the agency to support the federal government's ongoing development of a national strategy for privacy-preserving data analytics. Multi-party computation was developed with heavy investment from the NSF, the Defense Advanced Research Projects Agency, and the Intelligence Advanced Research Projects Activity. It has already been proven for government research use cases: Allegheny County, Pennsylvania, used multi-party computation to perform research on highly-sensitive services data held by different state agencies, without ever combining or sharing the data. The NSDS program is an opportunity for NSF to support the federal government in using best-in-class privacy and technology practices when it comes to Americans' data.

As NSF begins to implement the NSDS demonstration project, we want to ensure that you are aware of the privacy and security requirements that Congress enshrined in statute, and of our intention that you implement the NSDS using technology — not just policies and promises — to protect Americans' data. We will continue to work with NSF to ensure that these standards are upheld. We request that you provide answers to the following questions by January 31, 2023:

1. Will NSF commit to using multi-party computation, or another privacy-enhancing technology that prevents unencrypted data from being available within the NSDS system, for all data pertaining to individual Americans?
2. Will NSF commit to having agencies encrypt their own data within the NSDS platform, preventing NSF from holding a "master key" that would be a target for hacking and theft?

We also request that you provide our offices with updates as NSF determines:

3. What guidance or processes it will use to determine whether data used by an NSDS project is in an "identifiable form" or is otherwise sensitive, and thus must be encrypted or otherwise inaccessible to other parties in the NSDS system.
4. How it plans to implement the NSDS platform to enforce the requirement that it allow only authorized analysts to run approved queries.

Sincerely,

Ron Wyden
United States Senator

Rob Portman
United States Senator