



# DEMOCRATIZING DATA: THE VISION

Nancy A. Potok

# INFORMATION SATURATION

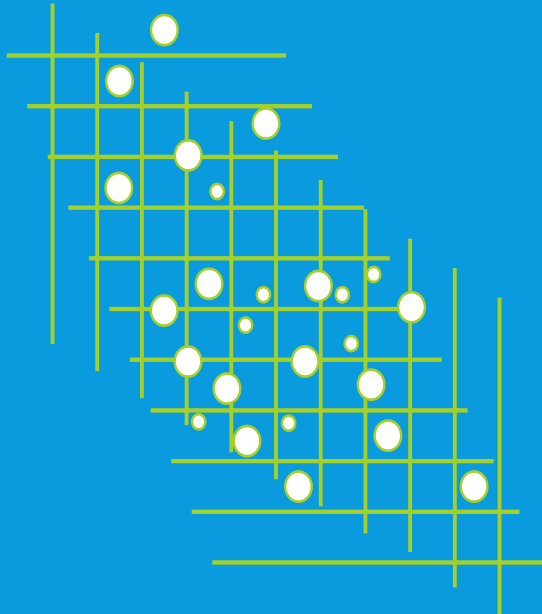


# URGENT NEED FOR SOCIETY'S FUTURE

Much data produced by  
researchers or  
government agencies  
could be transformed in  
valuable assets for the  
public good



# FEDERAL DATA ECOSYSTEM



---

Foundations of Evidence Based Policy Act

---

CHIPS and Science Act

---

Federal Data Strategy

---

OMB guidance

---

GAO/NIST Frameworks

---

Executive Orders

---

Data owners

---

Infrastructure Owners

---

Data users

---

Stakeholders

---

?

# EXECUTIVE GUIDANCE ON DATA

---

M-19-23 and M-21-27 (Evaluation Studies)

---

M-19-15 (Information Quality)

---

M-14-03 Administrative Data for Statistical Purposes

---

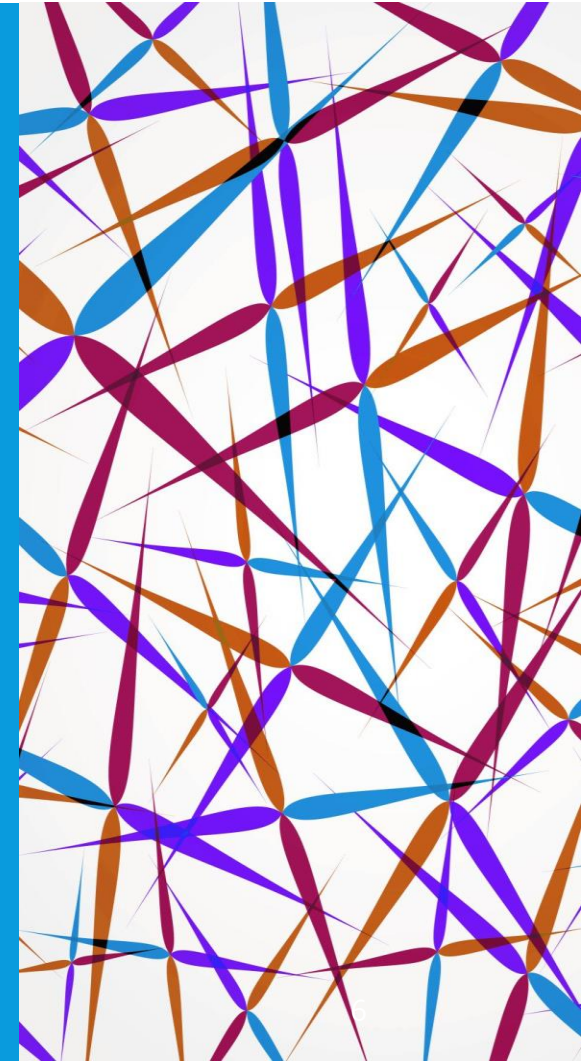
M-19-18 Federal Data Strategy

---

EO 13985 – Advancing Racial Equity for Underserved Communities

# AI-RELATED EXECUTIVE ORDERS

- **Executive Order 13859 of February 11, 2019: Maintaining American Leadership in Artificial Intelligence**
  - Instructed regulatory agencies to publish information about how they plan on regulating AI
- **Executive Order 13960 of December 3, 2020: Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government**
  - Ordered Federal Chief Information Officers Council to “identify, provide guidance on, and make publicly available the criteria, format, and mechanisms for agency inventories of...use cases of AI by agencies

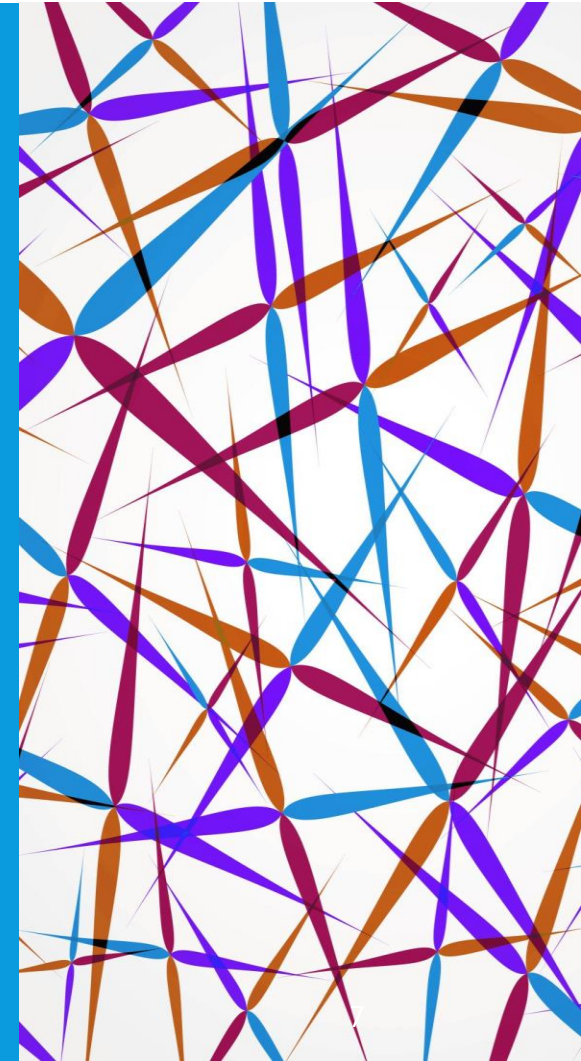


# AI-RELATED EXECUTIVE ORDERS

- Executive Order 14110 of October 30, 2023: *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*

Establishes a government-wide effort to guide responsible artificial intelligence (AI) development and deployment

Mandated the National Artificial Intelligence Research Resource (NAIRR), a national research infrastructure for responsible discovery and innovation in AI



# EVIDENCE BASED POLICY

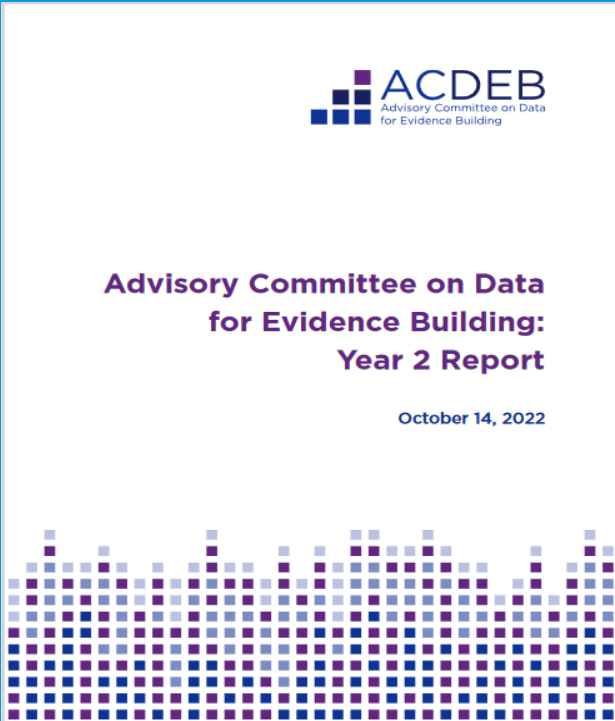
**Need for  
infrastructure  
and tools that  
are community  
driven and  
produce  
information  
that is valued  
and used.**



- Evidence-Based Policy Making Commission and Report (2016-2017)
- Foundations For Evidence-Based Policy Making Act (2018)



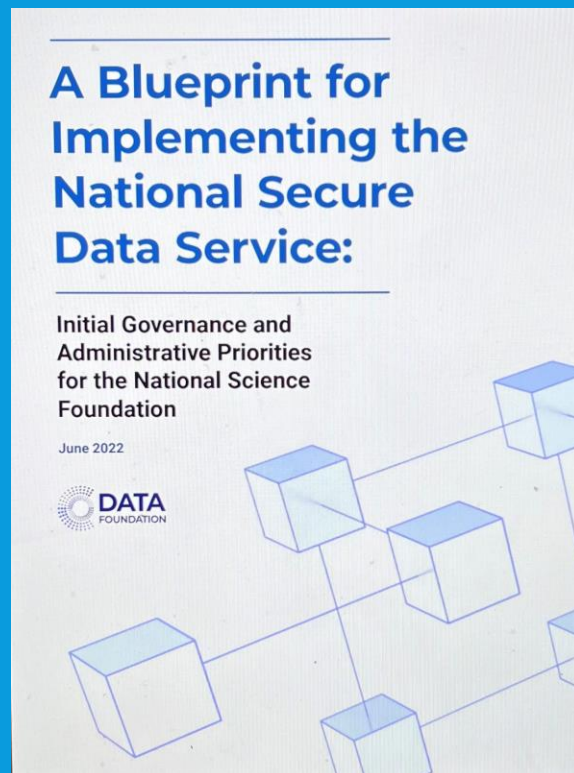
# NATIONAL ADVISORY COMMITTEE REPORT



Recommendations Part 3. NSDS Functions	
<i>Function 1. Coordination—Evidence Act Implementation for the Federal Government and Beyond</i>	
3.1	The NSDS should coordinate with the ICSP, communities of practice, and other key stakeholder groups to establish best practices for implementing the requirements of the Evidence Act and should serve as a model for testing and demonstrating those best practices for governments at all levels.
<i>Function 1. Coordination—Support a High-Quality User Experience, Including Providing Technical Assistance</i>	
3.2	The NSDS website should serve as a “front door” to the nation’s data assets, organized around a set of personas that reflect basic user needs.
3.3	The NSDS should identify opportunities for automation of its “intake process,” providing a high-quality user experience while focusing staff effort on complex user needs.
3.4	The NSDS should employ data concierges to help users refine their research projects, discover relevant data, and acquire access to that data.
3.5	The NSDS should employ technical assistance leads who develop educational resources for data providers and data users related to the methods and technologies used by the NSDS and in the broader evidence-building ecosystem.
3.6	The NSDS should actively encourage the development of communities to crowdsource support for users, complementing the work of the data concierges.
3.7	The NSDS should regularly sponsor projects that demonstrate the value of streamlining data sharing and increasing coordination, specifically projects that highlight cross-functional, cross-agency, and cross-governmental topics.
<i>Function 2. Communication</i>	
3.8	OMB should adopt a clear statement of purpose for the NSDS that is rooted in its core value.
3.9	Through the NSDS website, the National Center for Science and Engineering Statistics (NCSES) should create a public information hub that serves as a central repository for

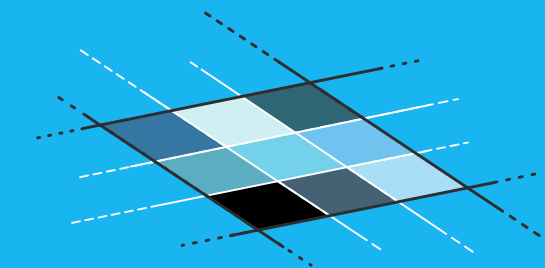
# Designing a National Secure Data Service

Input from the public to inform the Advisory Committee



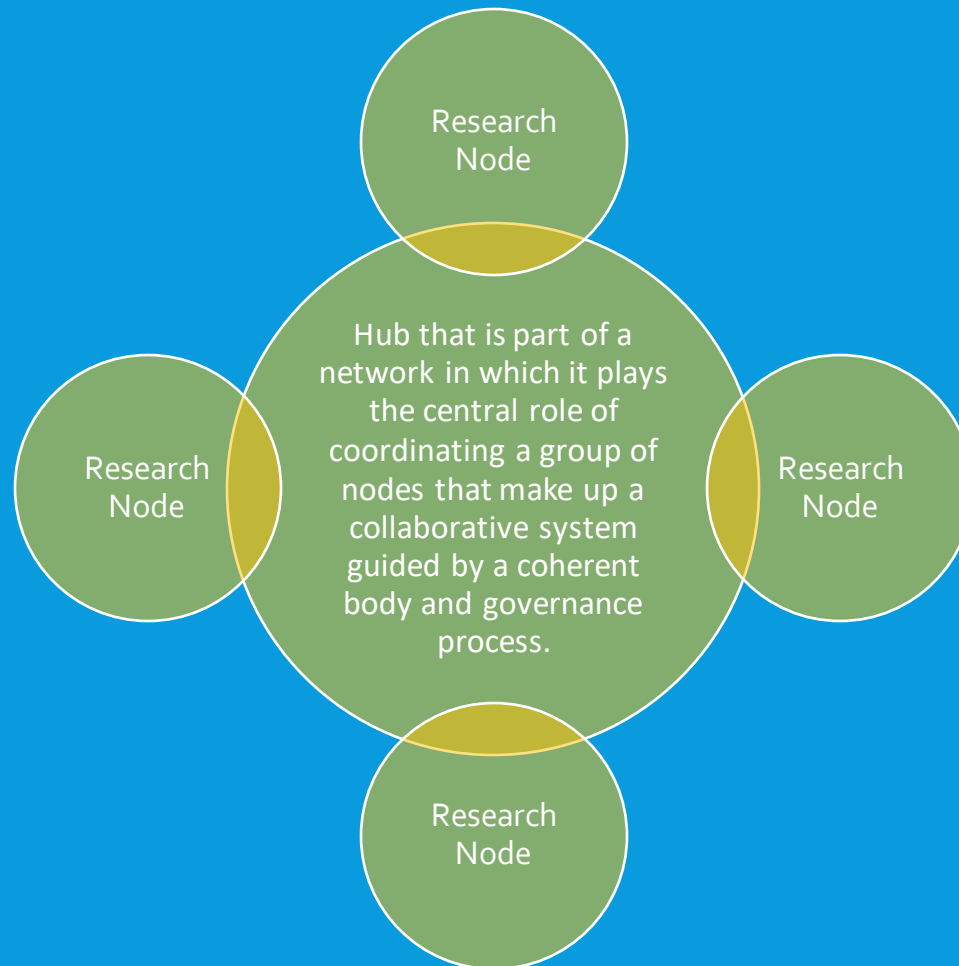
## MODERNIZING U.S. DATA INFRASTRUCTURE:

**Design Considerations for  
Implementing a National Secure  
Data Service to Improve Statistics  
and Evidence Building**



DATA FOUNDATION

# Vision for the NSDS: Coordinating Data Hub



# CHIPS AND SCIENCE ACT OF 2022: SEC. 10375 NATIONAL SECURE DATA SERVICE


Authorizes NSF to establish a National Secure Data Service pilot



Charged with developing, refining, and testing models to inform a full service NSDS



Authorizes USD \$9 million per year 2022-2026 (total \$45 million)



Requires pilot be set up within one year of enactment



OMB and the Interagency Council on Statistical Policy is governing the pilot

# NATIONAL AI RESEARCH RESOURCE (NAIRR)- EO14110

- Mandated that the National Science Foundation lead a consortium of agencies and non-governmental partners to create the National Artificial Intelligence Research Resource (NAIRR)
- National research infrastructure for responsible discovery and innovation in AI
- Building bridges with the NSDS to extent capabilities and include more communities

# VISION OF DEMOCRATIZING DATA MOVEMENT

**Goal:** Improve the practice of both government policy and research, by:

- providing evidence about how datasets are used; and
- building an ecosystem of agencies, researchers, and other stakeholders that are committed to revealing the value of data and evidence through its creation and use in collaborative communities.

*Data access, use and reuse are foundational to helping agencies and the data user community connect in meaningful ways.*

# DEMOCRATIZING DATA SEARCH AND DISCOVERY PLATFORM

- Identify **how** data are used
  - government agencies can better understand the utilization of their data investment portfolio and scientists can find other ways in which data have been used
- By **whom**
  - government agencies and their stakeholders as well as scientists (particularly junior faculty and graduate students) can find experts, and to study
- what **topics**
  - government agencies can identify how well their data investments are supporting their mission and scientists can find other work that is complementary to their own.

# HOW THE PLATFORM CAME TO BE

- Congressional imperatives
- Early philanthropic investment
- Tangible use case
- Engagement and investment by statistical agencies
- Dedicated research team led by NYU



# CRUCIAL TURNING POINT

- Open access to high quality, objective data to a wider community
- Work will continue to evolve –new topics for research on search and discovery will be picked up by the data ecosystem participants
- Galvanize the Democratizing Data community to continue the work as platforms such as the NSDS hubs and NAIRR National Data Platform become available

Our future depends on it.

# THE FUTURE IS BRIGHT

- Special issue articles show that providing evidence about how data sets are used spurs development of communities
  - Agencies
  - Researchers
  - Stakeholders
  - Policy Makers

