

Assessing the Health of the Federal Statistical Agencies: Vision, Purpose, and Challenges

FCSM Conference

Nancy Potok

October 2023

US Federal Data Ecosystem

- Federal Data Strategy
- Evidence Based Policy Act/CHIPS Act
- Administrative Policies and Guidance to Agencies
- OMB Trust Regulation
- Infrastructure
- Tools
- People (workforce, stakeholders)

Statutes - Evidence Act of 2018/CHIPS

Foundations for Evidence-Based Policy Making Act of 2018 includes recommendations from the U.S. Commission on Evidence Based Policy Making

- Contains three titles:
 - I. Federal Evidence Building Activities
 - II. Open Government Data Act
 - III. Confidential Information Protection and Statistical Efficiency Act

CHIPS and Science Act of 2022 (CHIPS Act) authorizes NSF-NCSES to establish a pilot project for a National Secure Data Service

PUBLIC LAW 115-435—JAN. 14, 2019	132 STAT. 5529
Public Law 115-435 115th Congress	
An Act	
To amend titles 5 and 44, United States Code, to require Federal evaluation activities, improve Federal data management, and for other purposes.	
<i>Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,</i>	
SECTION 1. SHORT TITLE; TABLE OF CONTENTS.	
SHORT TITLE.—This Act may be cited as the “Foundations for Evidence-Based Policymaking Act of 2018”.	
TABLE OF CONTENTS.—The table of contents for this Act is as follows:	
Sec. 1. Short title; table of contents.	
TITLE I—FEDERAL EVIDENCE-BUILDING ACTIVITIES	
Sec. 101. Federal evidence-building activities.	
TITLE II—OPEN GOVERNMENT DATA ACT	
Sec. 201. Short title.	
Sec. 202. OPEN Government data.	
TITLE III—CONFIDENTIAL INFORMATION PROTECTION AND STATISTICAL EFFICIENCY	
Sec. 301. Short title.	
Sec. 302. Confidential information protection and statistical efficiency. Sec. 303. Increasing access to data for evidence.	
TITLE IV—GENERAL PROVISIONS	
Sec. 401. Rule of construction.	
Sec. 402. Use of existing resources. Sec. 403. Effective date.	
5 USC 301 prec.	
Foundations for Evidence-Based Policymaking Act of 2018.	
5 USC 101 note.	
Jan. 14, 2019	
[H.R. 4174]	

Assessing U.S. Data Infrastructure: 2024 onwards

Reactive Monitoring

- Challenging to reverse threats underway
- Damage may be done
- Reaction may be perceived as partisan

Proactive Monitoring

- Identify/shine light on needs, weaknesses, and vulnerabilities
- Address before problems arise
- Help agencies make their cases

Proactive Monitoring

Success hinges on engagement of and participation of federal statistical community of stakeholders

Publish report in 2024 Q1 and annually thereafter, expanding coverage to more agencies

Promote additional conversation and study of federal statistical agencies

Defining the “Health” of Any Agency



High Level Stat Agency Health



Can Assessment Dive Deeper Quantitatively?

Autonomy Protections

Innovation

Workforce

Resources

Host agency support

Data Usage

Deeper Dive Potential indicators



Resources (e.g., budget levels, staffing, contracting)



Professional autonomy (e.g., safeguards against political meddling)



Innovation/modernization capabilities (e.g., new products and data sources, pilot projects)



Workforce (e.g., employee satisfaction, professional development, recruitment and retention, diversity)



Host agency support (e.g., budget requests, data governance processes)



Data use and user engagement (e.g., downloads, citations, data linkages)

Gathering Indicator Data for Analysis



Collection sources

- Publicly available data: budgets, publication schedules, OPM employee satisfaction, etc.
- Agency questionnaire, possible data user survey



Analysis plans

- Conduct qualitative and quantitative analyses to ensure data and expert accounts agree
- Present findings in public report

Challenges We Have Encountered With a Deeper Dive

What does “Health” Mean?

- No standard definition or vision that is specific enough
- Most principles are high level but don't define such that a metric is obvious (e.g. how much autonomy is enough? What does that look like for each agency?)

Sensitive Areas

Relationship between statistical agency and parent agency

- Not a lot of transparency – discussions are not public
- Many interactions (budgets, IT plans, staffing levels, Secretarial priorities)
- Need cooperation to be successful
- Varying viewpoints on what is “helpful”
- OMB Trust Regulation not yet in place, thus effectiveness not yet evaluated

Major Differences Between Agencies

- Example: Cannot directly compare Statistics of Income at IRS, Bureau of Transportation Statistics at DOT, and Census Bureau at Commerce (size, control of resources)
- Widely varying statutory authorities that affect autonomy, agenda setting, and placement in agency
- Different appointment authorities for Directors (Presidential, political, career)

Budgets Can Be Difficult to Decipher

- A flat budget may not be a healthy budget if inflation is high
- No insight available on what an agency requested before its request went through the budget process of parent agency and OMB review
- No insight into parent agency instructions to agencies even before requests go forward (e.g., all increases must have equal offsets)
- Congress may add work without adding funds during the appropriations process
- Larger agencies may be better able to absorb cuts or level funding

Collaboration is Not Centrally Collected Information

- Cross-agency collaborations may not be noted anywhere in publicly available information
- Important to measure collaborations that lead to innovations and evidence-building
- Can that information be found without surveying agencies directly?

Agencies Track Activities Differently

- Some information requires use of proxies, such as training and travel
- Some training is in the travel budget but some training is virtual or local
- Agencies track these activities differently
- Travel restrictions during the pandemic created anomalies in trend data

A Majority of the Agencies Have Limited Usage Data

- Most of the agencies can only track website hits or downloads
- Only four agencies currently can track systematically how data are actually being used by researchers in scientific journal publications
- Need usage data beyond journal publications (e.g., conference papers, policy documents, reports)
- Need to capture agency reports usage but no common nomenclature being used for report citation (e.g., DOIs) and users don't have incentives for including citations

Attempts to Survey Agencies Directly had Varying Results

- Some agencies did not have enough staff available to answer questions
- Questions were interpreted differently by each agency
- May be difficult to collect information annually
- Different agencies have different clearance processes through their parent agencies, so information was missing in several instances if not already public

So What Were We Able to Collect?

- Public budget data
- Timeliness of data releases
- Some resources data
- Inconsistent and low quality FEVS data
- Inconsistent data on professional development and innovation
- Inconsistent data on usage

Here are some examples to date... (Connie and Jonathan)