

# **Evidence Commission Approach to Data Sharing & Sensitivity**

*and what the Evidence Act now requires*

**Nick Hart, Ph.D.**

President, Data Foundation

*Presentation for COPAFS Tiered Access Workshop. 9/15/2019*

# U.S. Commission on Evidence-Based Policymaking



- Created in 2016 by law to study how to make government data more accessible and useful
- 15 politically-appointed members with expertise in program administration, data, and research
  - 5 appointed for specifically knowledge about privacy and confidentiality
  - Included former congressional staff, former statistical agency leaders, and former senior executives
- Final report included 22 unanimous recommendations

# CEP Findings

- Many government agencies don't adequately assess resources to determine access controls
  - Viewed Tiered Access as a privacy mechanism and one way to support the data minimization FIPPs
  - Acknowledged PSAs are leaders in tiered access
  - "A well-designed and properly implemented data minimization strategy like tiered access can reduce the risk of unauthorized use and unintended harm to individuals." – CEP 2017, p. 38
  - Suggests agencies should review sensitivity of data to better determine tiering approaches – though a consistent system does not exist today

# Tiered Access in Commission Process

- Commission was presented information to suggest tiered access was a “major strategy for addressing re-identification risk” (see CEP staff memos)
- Viewed the NSDS as an opportunity to implement a consistent tiered access approach, with the organization using OMB’s standard setting authority

# CEP Findings

- Existing systems for considering sensitivity levels offer potential for more widespread adoption in Federal government
  - Suggest such systems benefit from feedback from stakeholders, researchers, privacy advocates, etc.
  - Sensitivity is not static and can change over time

Figure 5. Model of Sensitivity Levels for Federal Data

Level	Sensitivity	Description
5	Crimson	<b>Maximally restricted.</b> Highly sensitive. Identifiable records from data collected with a promise of confidentiality.
4	Red	<b>Restricted.</b> Sensitive. Identifiable records from data collected with a promise of confidentiality.
3	Yellow	<b>Restricted.</b> Crimson or Red datasets modified by technologies that mask individual records (e.g. data query tools, differential privacy).
2	Green	<b>Minimally restricted.</b> Not sensitive. Data files made available to the public but subject to procedures designed to raise accountability by users, such as registration before accessing.
1	Blue	<b>Public data.</b> Most safe. Open data.

# CEP Recommendations

- 2-1: Establish the NSDS
  - Make information publicly available about sensitivity, risk or re-id, etc.
- 2-8: develop single standard for access to data that “are not publicly available” – including with acknowledgements to data sensitivity with appropriate access restrictions (implemented with NSDS playing a key role)
  - NSDS to help agencies review sensitivity levels over time, including periodic reviews
  - Suggest could be based on a standard classification scheme
- 3-1: Data Officials envisioned with key role in establishing sensitivity levels with access controls
- 4-5: Establish data inventories with sufficient metadata.
  - Metadata for the NSDS should be sufficient to allow for determination of sensitivity and appropriate protection levels.

# Translation to the Evidence Act

- Foundations for Evidence-Based Policymaking Act of 2018 (PL 115-435) addresses half the CEP recommendations
- Does not establish NSDS (but does set up a new FACA)
- Does:
  - Require risk assessment in the statistical system
  - Establish CDOs government-wide
  - Statistical officials in all 24 CFO Act Agencies
  - Require data inventories with sufficient meta-data
  - Mandate open data, when practicable and consistent with the FOIA standard – taking into account “risks and restrictions related to the disclosure of [PII]...”
  - Direct OMB to establish standards for the statistical system to categorize sensitivity levels of “data assets” – with the info being made available via data inventories (see CIPSEA Sec. 3582)

## About the Data Foundation

- Non-profit research think-tank based in DC
- Support efforts to provide research, training and education about data policy
  - E.g., roundtables, white papers, technical assistance, trainings, and task forces

[www.datafoundation.org](http://www.datafoundation.org)  
[Nick.Hart@DataFoundation.org](mailto:Nick.Hart@DataFoundation.org)