NCSES Data Management, Dissemination and User Perspective

May Aydin

Panel on Transparency and Reproducibility for NCSES

May 21, 2019

National Center for Science and Engineering Statistics
Directorate for Social, Behavioral and Economic Sciences
National Science Foundation
http://ncses.nsf.gov

What to expect

- Brief overview of ongoing activities, resources and needs
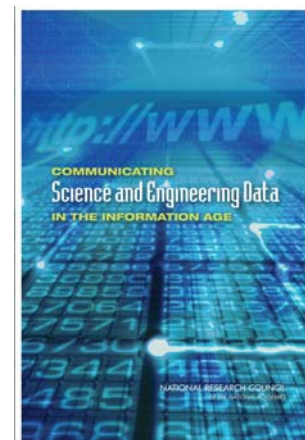- Success model and view
- Questions for Panel

Meeting the Mandate(s)

- Federal Data Strategy

- Evidence-Based Policy Act

- Improving Implementation of the Information Quality Act

Effectively manage, disseminate and communicate

- Recommendations from 2012 National Academy Panel Study,
  Communicating Science and Engineering Data in the Information Age
    - Modernize underlying and externally-facing technologies and systems
    - Centralize database management – automation, version control
    - Enhance access and usability of data
    - Engage with communities (data and non-data users, incl prof groups)
    - Improve metadata in content, reach and usability
    - Provide access and improve visualizations

Things to know

- Central, internal data repository houses most NCSES microdata and documentation, including metadata.

- Contractor-dependent for data collection, data/file preparation, and documentation (e.g. methods, user guides etc).

- Invest considerable effort to follow-up, store, and transform contractor delivered files for storage and publication.

- Redesign, development efforts ongoing (publication system, data system, website, intuitive tags/topics/labels, product design, usability of output files).
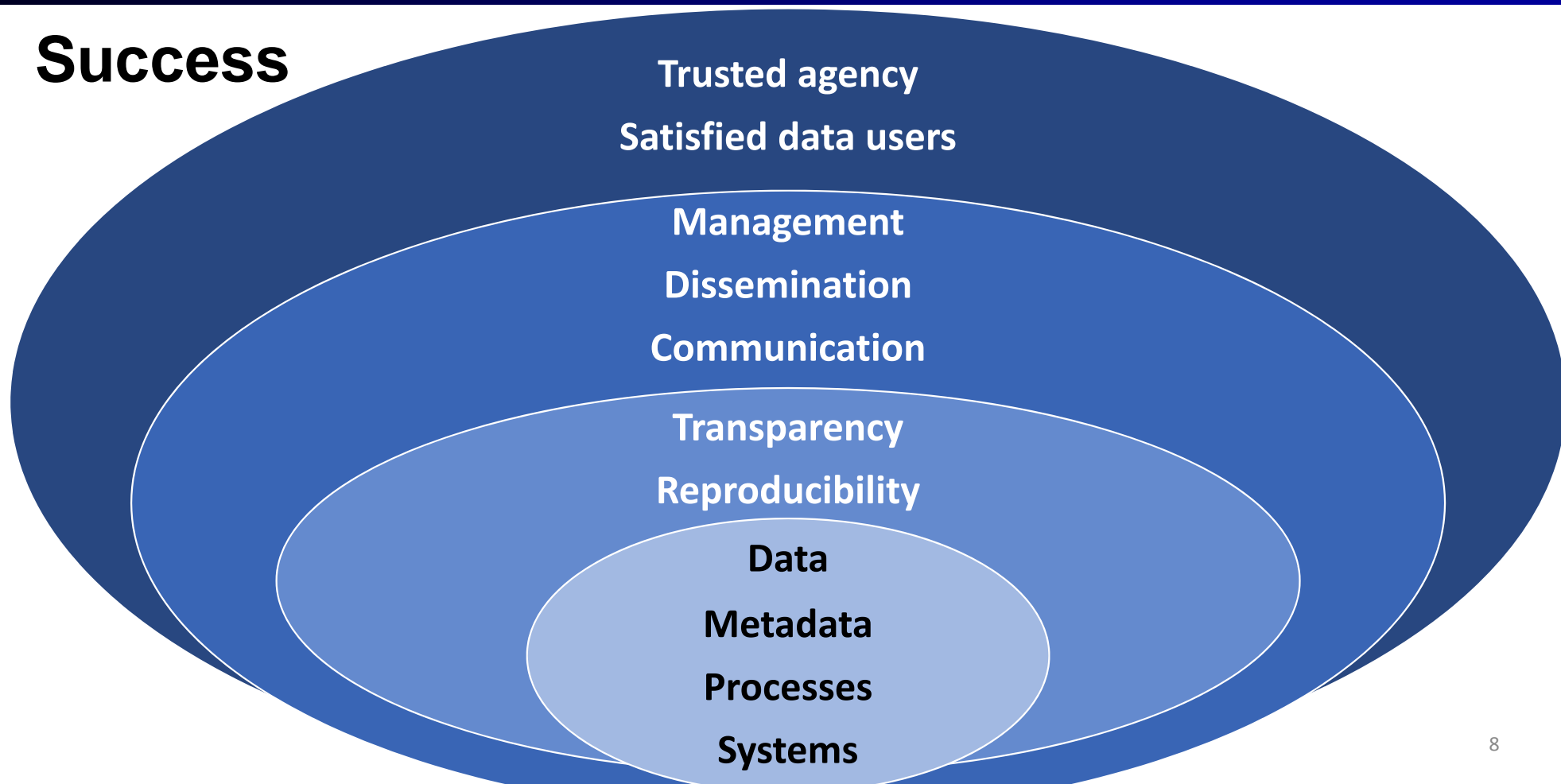
What does success look like?

- **Scenario 1:** A long-time NCSES data user quickly accesses the documentation online to understand changes in the survey's….

- **Scenario 2:** A journalist navigates the website to download the correct figure to support… (with appropriate NCSES citation).

- **Scenario 3:** A non-NCSES data user stumbles on NCSES data found in data.gov and eagerly visits NCSES's website to search across topics and subtopics to learn about…

- **Scenario 4:** An NCSES survey manager easily accesses an internal archived data set to reproduce a specific statistic for…

Agency success tied to transparency and reproducibility

Success

Trusted agency
Satisfied data users

Management
Dissemination
Communication

Transparency
Reproducibility

Data
Metadata
Processes
Systems

8

Questions for Panel

- A balancing act….

  - between public-facing metadata and internal metadata (content, presentation and breadth)?

  - between requiring minimum standards for metadata and additional metadata from subject matter experts and/or contractors?

  - between standardizing contractor-delivered data files (for increased efficiencies and usability) and allowing flexibility?

Questions for Panel

- Still balancing….

  - between automation in producing and populating metadata and the editorial process?

  - between traditional and non-traditional metadata (e.g. images) and assessing the value both bring to transparency and reproducibility?

Thank you.

maydin@nsf.gov