NCES Data Confidentiality Considerations and Processes

IEA General Assembly

October 2019

Stephen Provasnik, Ph.D.
National Center for Education Statistics

U.S. Laws and Guidelines

Roots

- Belmont Report (1979) - Ethical principals and guidelines for protecting human subjects
- Common Rule (1991) - Provisions for IRBs, informed consent, assurances of compliance

Selected Laws

- Privacy Act of 1974 (Section 552a)
- Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)
- Education Sciences Reform Act of 2002 (PL 107-279)

U.S. Laws and Guidelines

Selected Agency Standards and Practices

- Census Bureau Standards
  - http://www.census.gov/srd/sdc/
  - http://www.census.gov/srd/sdc/FR_23693-94.pdf
- National Center for Education Statistics (NCES) Standards
  - http://nces.ed.gov/statprog/2002/std4_2.asp

- Disclosure risk cannot be eliminated
- The challenge is to reduce risk and keep distortion in data to a minimum
- NCES has a set of statistical confidentiality measures to minimize risk

General Process

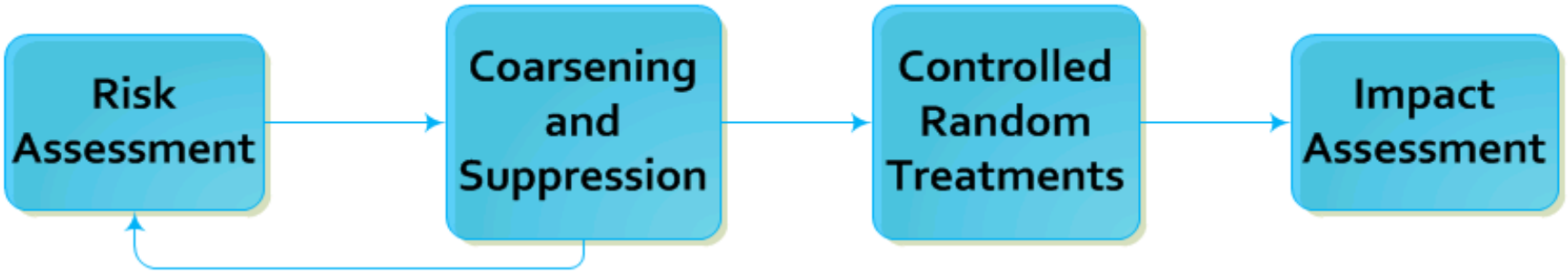
- In the United States, data cannot be delivered, reported, or shared before they are put through a process of statistical confidentiality measures.

Statistical confidentiality measures

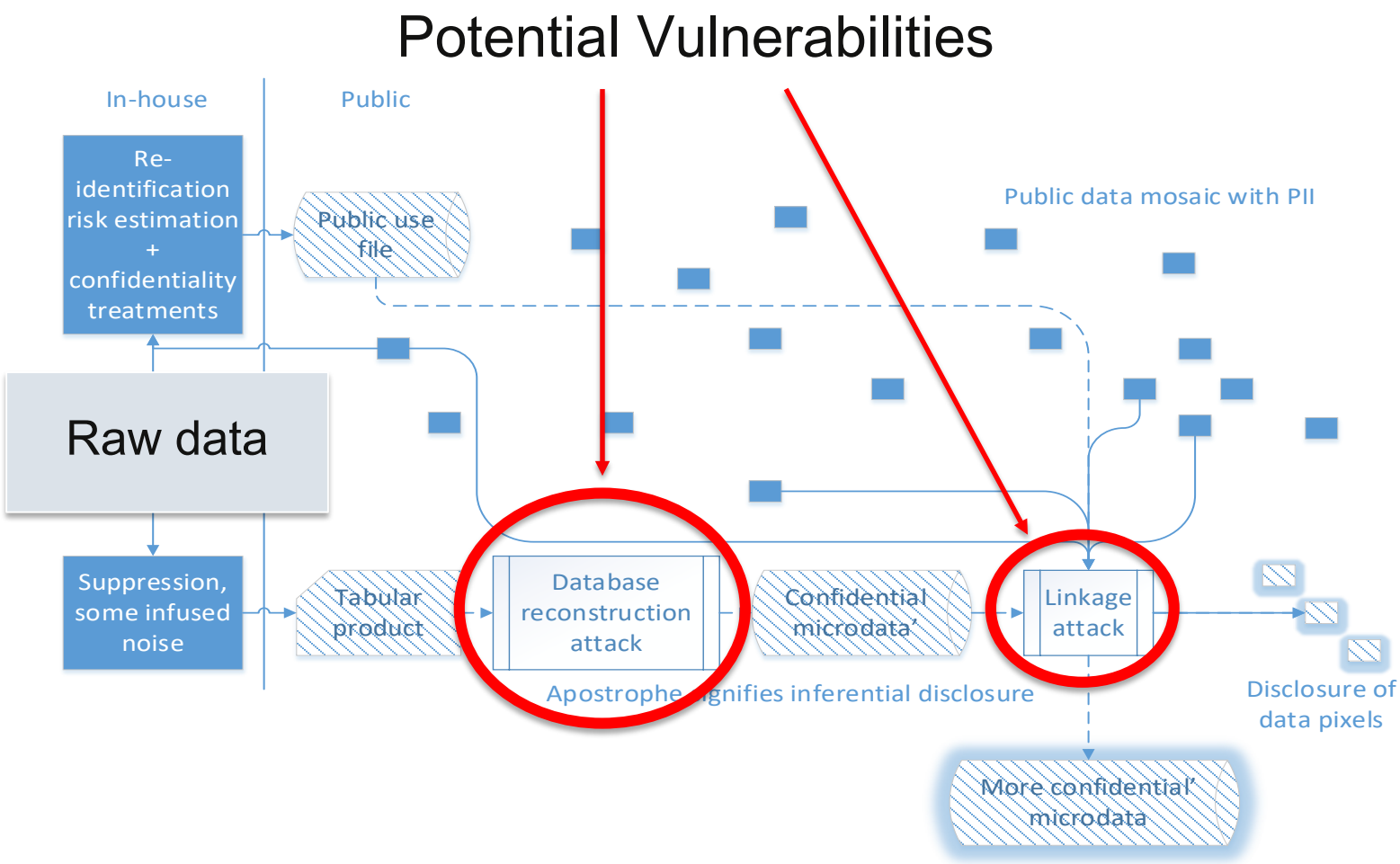
Step 1 is to submit a **Disclosure Analysis Plan** to the Disclosure Review Board (DRB) for review and approval.

A Disclosure Analysis Plan includes descriptions of

- The study
- Potential risk (or the "risk assessment")
- Proposed statistical confidentiality measures (including coarsening and suppression)



Risk Assessment



Coarsening and Suppression

NCES statistical confidentiality measures try to minimize risk in various ways, such as:

- removing the direct personal identifiers ("de-identify cases")
- Converting continuous data into categorical data (e.g., by releasing ages in age ranges -- aka "masking")
- anonymizing dataset by slightly changing values to prevent identification from indirect identifiers—such as changing the value of a birthday by one day ("perturbing")

Challenges

Constraints on statistical confidentiality measures

- Data integrity / minimal data distortion
- Reasonable costs
- Reasonable turnaround time

Overview of Disclosure Analysis Plan

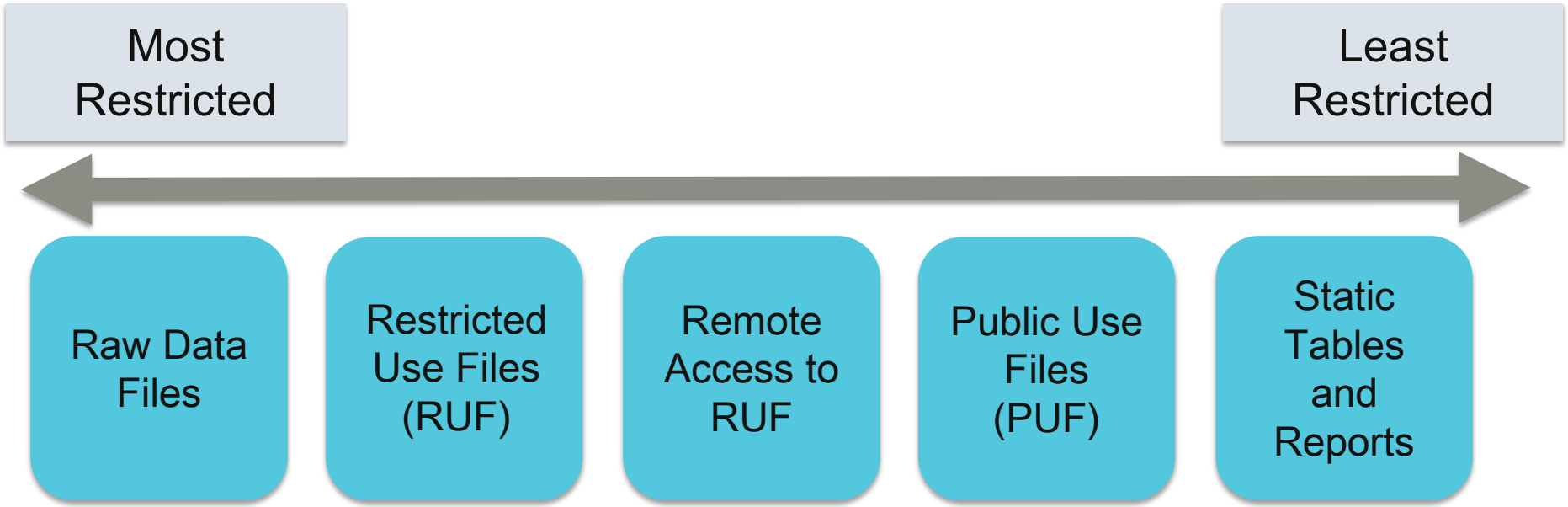
Once the Disclosure Analysis Plan is approved by DRB, step 2 is to prepare the data for dissemination.

Preparing data for dissemination includes

- Swap data using confidential swapping rates provided by DRB
- Prepare and review results of each procedure
- Check potential data distortion and/or any supplemental risk
- Revisit the Disclosure Analysis Plan to see if any changes would provide better results
- Provide a report to the DRB with the results of the plan implementation, assessing its overall impact and reporting any deviations from the original plan

Forms of released data

Once studies receive approval from DRB data can be released in any or all of these forms:



Thank You

For more information on NCES Data Confidentiality Measures:

- https://nces.ed.gov/pubs97/p97527/Sec1_txt.asp
- https://nces.ed.gov/statprog/conflaws.asp

Contact information:

- Stephen Provasnik Stephen.Provasnik@ed.gov