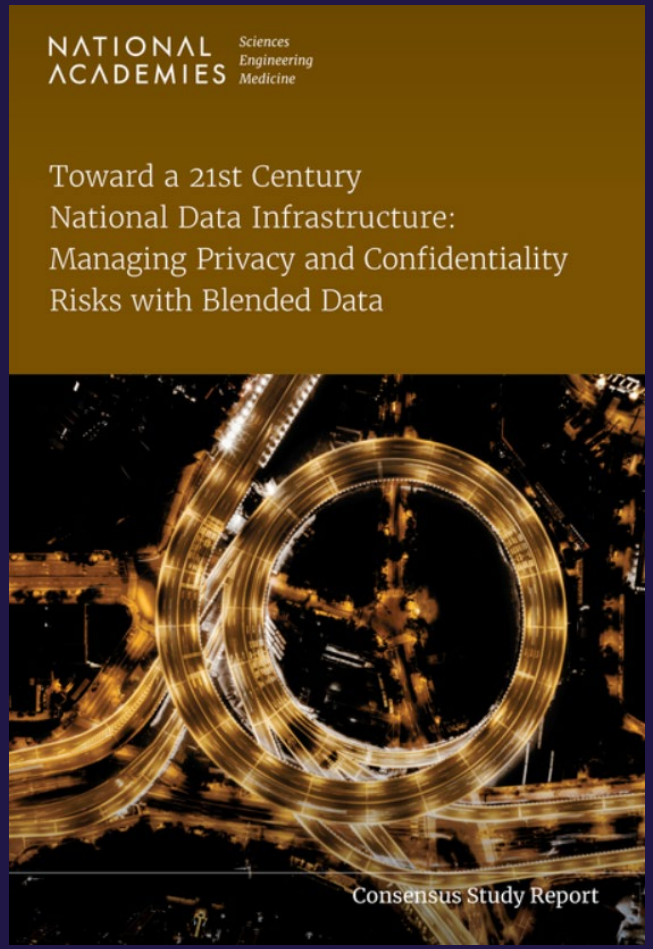
AI and Blended Data

H. V. Jagadish

University of Michigan



Toward a 21st Century National Data Infrastructure: Managing Privacy and Confidentiality Risks with Blended Data

https://nap.nationalacademies.org/catalog/27335/toward-a-21st-century-national-data-infrastructure-managing-privacy-and-confidentiality-risks-with-blended-data

Panel Members

- **Jerry Reiter,** Duke University
- **Claire Bowen**, Urban Institute
- **Aloni Cohen**, University of Chicago
- **Diana Farrell**, National Bureau of Economic Research
- **Robert Goerge**, University of Chicago
- **Nick Hart**, Data Foundation
- **Hosagrahar Jagadish**, University of Michigan
- **Dan Kifer**, The Pennsylvania State University
- **Karen Levy**, Cornell University
- **Salomé Viljoen**, Michigan Law School
- **Mark Watson**, Federal Reserve Bank of Kansas City (formerly)

Issues In Managing Blended Data Risk:
Risk Spans the Blended Data Life Cycle

- The blended data life cycle spans:
  - initial conceptualization of blended data;
  - identifying and accessing ingredient data sources;
  - blending the data from those sources; and
  - sharing the resulting data products.
- Each of these stages presents potential risks to privacy and confidentiality, and subsequent harms to data subjects and data holders.
- Disclosure risks and harms can be magnified in blended data.
- AI can greatly increase these risks

Mini Case Study in Blended Data Risk:
Loan qualification in FinTech

- The Equal Credit Credit Opportunity Act prohibits lenders from considering race (and several other protected attributes) in making their decisions about creditworthiness.
- But information not explicitly including protected attributes is fair game.
- Permissible attributes can be used to predict race with high probability.
- E.g. where do you worship?

Issues In Managing Blended Data Risk:

Risks In Blended Data Can Be Managed

- No non-trivial data release method guarantees zero risks to privacy. Generally, providing greater access enhances usefulness, but also increases disclosure risks for data subjects.
- As a general rule, enhancing the usefulness of blended data requires accepting greater disclosure risks.

  Trade-offs in disclosure risks, disclosure harms, and data usefulness are unavoidable and are central considerations when planning data-release strategies, particularly for blended data. Effective technical approaches to manage disclosure risks prioritize the usefulness of some analyses over others. (Conclusion 2-1)

  AI just makes this harder

Issues In Managing Blended Data Risk:
Knowing vs. Guessing: Does it matter??

- Traditional privacy concerns have focused on hard facts.
- With AI we often get likely facts.
- At what level of certainty will a user act on the data?
- What does the data subject feel about “probabilistic disclosure”?
- How do we interpret regulations in light of probabilities?
- Compare with traditional obfuscation methods such as releasing data aggregated to a minimum cell size, or k-anonymity.

Key Attributes of a Framework for Managing Risk:
Responds to Stakeholder Interests

- Engagement with stakeholders, including data holders, data users, and decision makers, is important for effective management of trade-offs.
- Communication plans may differ depending on the needs of relevant groups:
  - For the public, use plain language to describe context-specific protections.
  - For data users, include methods for demonstrating data quality after privacy protections are applied.
- Explanation/communication much harder with complex imputations.

Effective communication with data holders and data users can help agencies understand and better manage disclosure risk/usefulness trade-offs. (Conclusion 2–2)

Key Attributes of a Framework for Managing Risk:
Adapts to Policy and Technology Changes

- As policy priorities change, data availability can change. As more data are made available, the potential for privacy risk also increase. Technical approaches to limit privacy risk are advancing.
- Even when regulatory guidance and procedures for managing privacy risks are established, social acceptance of sharing and use of blended data will change.
- More powerful imputation methods lead to increased privacy risks.

  The effectiveness of a framework for making decisions about acceptable disclosure risks given expected usefulness of data depends on whether that framework is dynamic. A dynamic framework allows for changing policy needs and data availability over time, in a way that accounts for the interests of data subjects, data holders, and data users. (Conclusion 3–1)

A Model Framework for Managing Risk:
Six Steps Supported with Guiding Questions

- Drawing from the panel's review of technical and policy approaches, the panel provides a framework that accounts for the attributes of blended data for making decisions about data-protection methods.
- Framework encourages agencies to answer a set of questions at each stage of the data-blending lifecycle to aid decision-making. Rather than attempting to cover all data-blending scenarios or stipulate precise approaches, the framework provides a lens to promote careful consideration of key questions.

  Technical and policy approaches in combination are necessary for effective management of disclosure risks. (Conclusion 4–1)

A Model Framework for Managing Risk:

1. Determine auspice and purpose of the project

a. What are the anticipated final products of data blending?

b. What are potential downstream uses of blended data?

c. What are potential considerations for disclosure risks and harms, and data usefulness?

A Model Framework for Managing Risk:

2. Determine ingredient data files

a. What data sources are available to accomplish blending, and what are the interests of data holders?

b. What steps can be taken to reduce disclosure risks and enhance usefulness when compiling ingredient files?

A Model Framework for Managing Risk:

3. Obtain access to ingredient data files

a. What are the disclosure risks associated with procuring ingredient data?

b. What are the disclosure risk/usefulness trade-offs in the plan for accessing ingredient files?

A Model Framework for Managing Risk:

4. Blend ingredient data files

a. When blending requires linking records from ingredient files, what linkage strategies can be used?

b. Are resultant blended data sufficiently useful to meet the blending objective?

A Model Framework for Managing Risk:

5. Select approaches that meet the end objective of blending

a. What are the best-available scientific methods for disclosure limitation to accomplish the blended data objective, and are sufficient resources available to implement those methods?

b. How can stakeholders be engaged in the decision-making process?

c. What is the mitigation plan for confidentiality breaches?

A Model Framework for Managing Risk:

6. Develop and execute a maintenance plan

a. How will agencies track data provenance and update files when beneficial?

b. What is the decision-making process for continuing access to or sunsetting the blended data product, and how do participating agencies contribute to those decisions?

c. How will agencies communicate decisions about disclosure management policies with stakeholders?

Figure 4-1. Model decision matrix of disclosure-protection strategies given potential harms and usefulness.

Potential Usefulness from Access *(potential usefulness decreases from top to bottom)*	**Potential Harms Resulting from Disclosure** *(potential harm increases from left to right)*				
	Negligible	Minor and fleeting	Significant and lasting	Life altering	Life threatening
Assess policy—major impact	• Access: RD; tiers • Privacy: minimal	• Access: RD; tiers • Privacy: light	• Access: RD; tiers • Privacy: rigorous	• Access: RD; tiers • Privacy: rigorous	• Access: RD • Privacy: very rigorous
Assess policy—modest impact	• Access: RD; tiers; DL • Privacy: minimal	• Access: RD; tiers • Privacy: light	• Access: RD; tiers • Privacy: rigorous	• Access: RD; tiers • Privacy: rigorous	• Access: RD; tiers • Privacy: very rigorous
General knowledge	• Access: tiers; DL • Privacy: minimal	• Access: tiers; DL • Privacy: light	• Access: tiers; DL • Privacy: rigorous	• Access: tiers; DL • Privacy: very rigorous	• Access: DL • Privacy: very rigorous
Educational	• Access: DL • Privacy: minimal	• Access: DL • Privacy: rigorous	• Access: DL • Privacy: rigorous	*Do not blend data*	*Do not blend data*

Summary of Conclusions (1)

- Agencies, policymakers, data users, and data subjects need to recognize that any blended (or nonblended) data release that offers nontrivial usefulness introduces disclosure risks; it is not productive or correct to think of disclosure risks as a “yes or no” feature.
- Data-release strategies need to balance disclosure risks with data usefulness. When usefulness is high, stakeholders may be willing to accept greater risks to realize the benefits. Agencies can use various disclosure-protection methods for differing data-analysis objectives, such as tiered access approaches.

Summary of Conclusions (2)

- Successful risk-management strategies are likely to involve both technical and policy approaches. Some existing approaches can be gainfully applied with blended data, but others are less effective given the magnified disclosure risks in blended data.
- Disclosure risk management approaches need to be dynamic, involve stakeholder input, and rely on best practices. These characteristics can help determine desirable disclosure risk/usefulness trade-offs.
- Agencies can be (and should be, in the panel's opinion) intentional in examinations of risks at all stages of the blended data lifecycle.

Summary of Conclusions (3)

- AI can greatly increase the value of blended data:
  - Through better record linkage
  - Through extrapolating more and better from the available data
- But AI also increases privacy risks.
- So, same tradeoffs, but with greater weights on both sides.