Privacy, ethics and data access

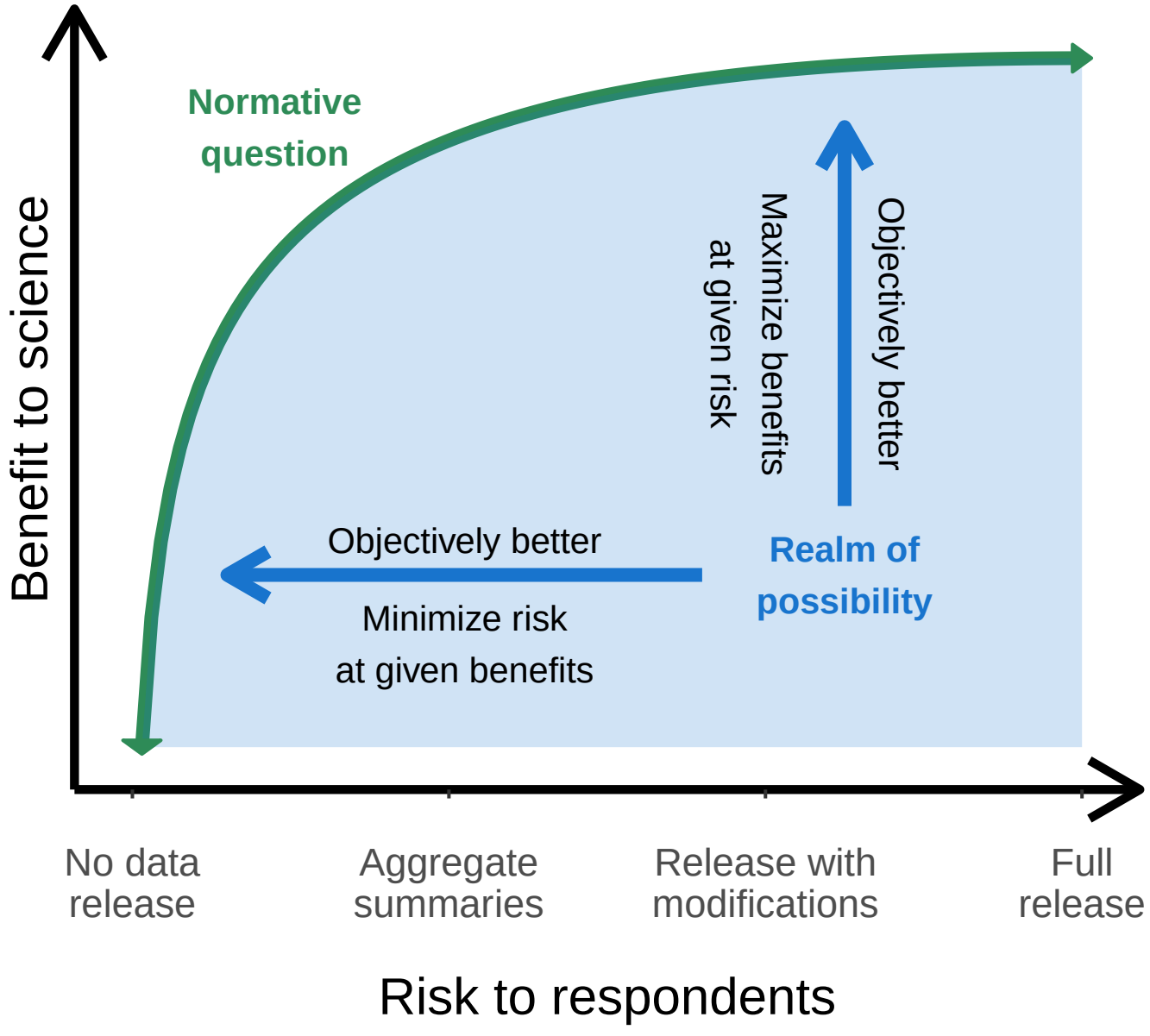
A case study of the Fragile Families Challenge

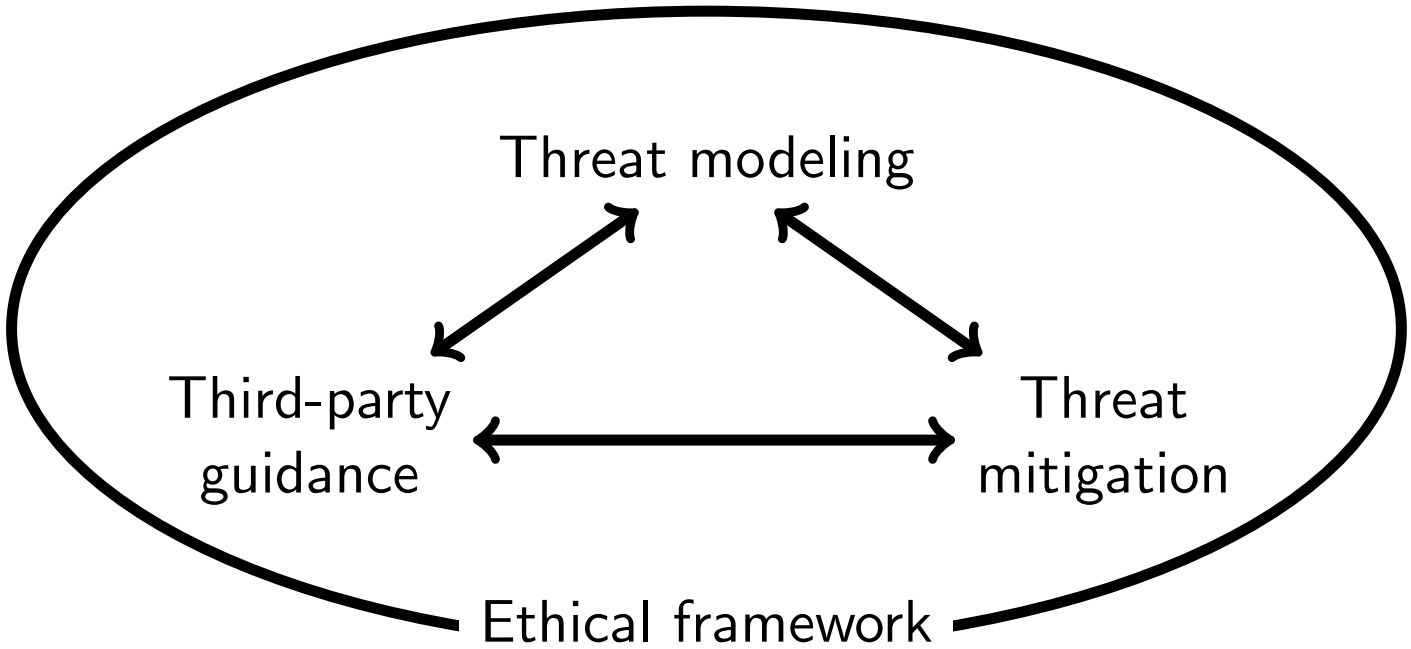
Ian Lundberg — Sociology, Princeton
Arvind Narayanan — Computer Science, Princeton
Karen Levy — Information Science, Cornell
Matthew J. Salganik — Sociology, Princeton

13 August 2018

Annual Meeting of the American Sociological Association

This research is supported by the Russell Sage Foundation. We are grateful to the members of the Board of Advisors of the Fragile Families Challenge. Source for these slides: www.github.com/fragilefamilieschallenge.


Benefit to science
Normative question
Maximize benefits at given risk
Objectively better
Objectively better
Realm of possibility
Minimize risk at given benefits
No data release
Aggregate summaries
Release with modifications
Full release
Risk to respondents


Threat modeling
Third-party guidance
Threat mitigation
Ethical framework

Fragile Families
& Child Wellbeing Study
PRINCETON | COLUMBIA

- Birth cohort panel study
- $\approx$ 5,000 children born in 20 U.S. cities
- Followed from birth through age 15

Fragile Families
& Child Wellbeing Study
PRINCETON | COLUMBIA

Features relevant to privacy:

1. Informed consent

Fragile Families
& Child Wellbeing Study
PRINCETON | COLUMBIA

Features relevant to privacy:

1. Informed consent
2. Already available to researchers

Fragile Families
& Child Wellbeing Study
PRINCETON | COLUMBIA

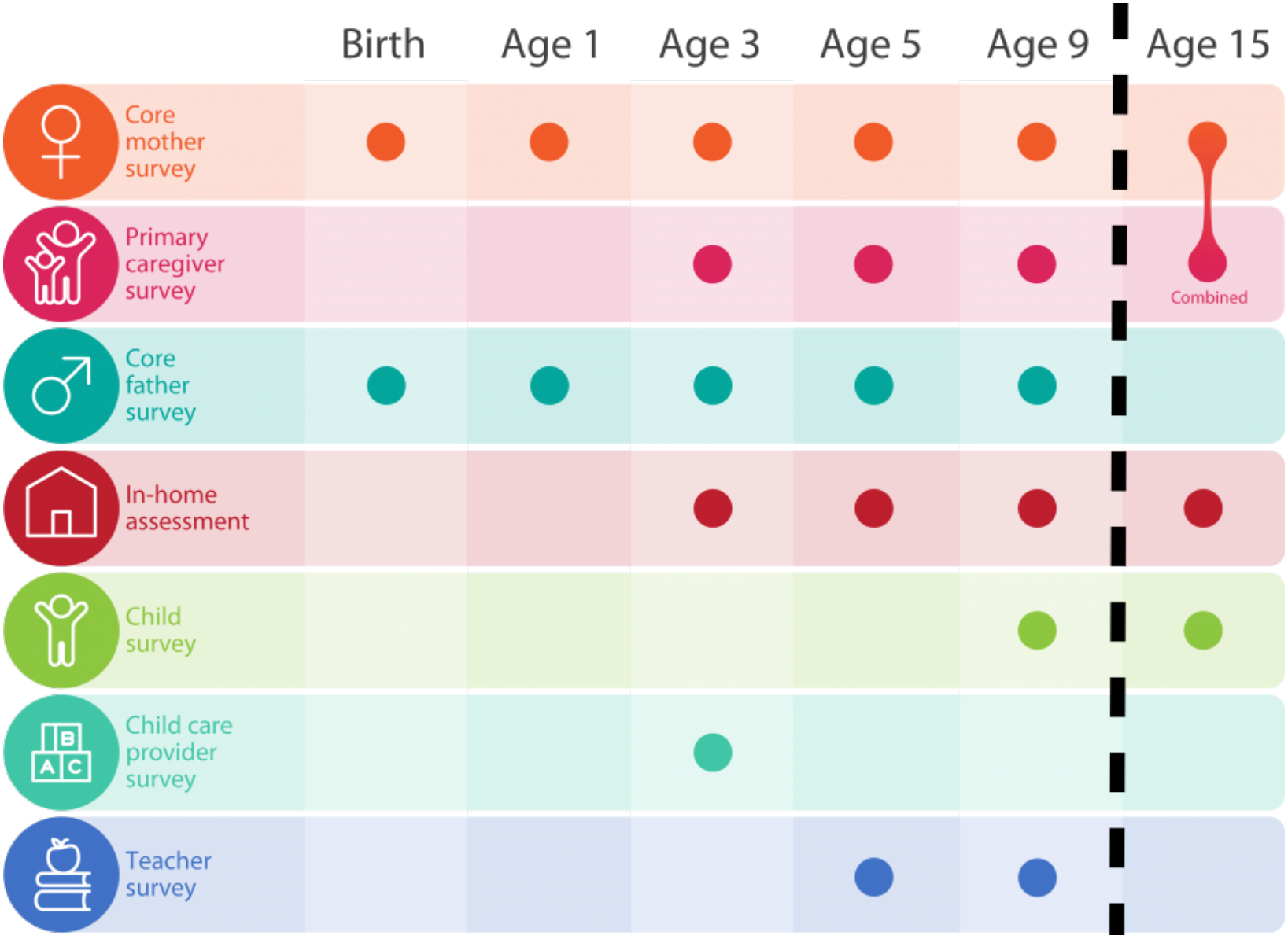
Features relevant to privacy:

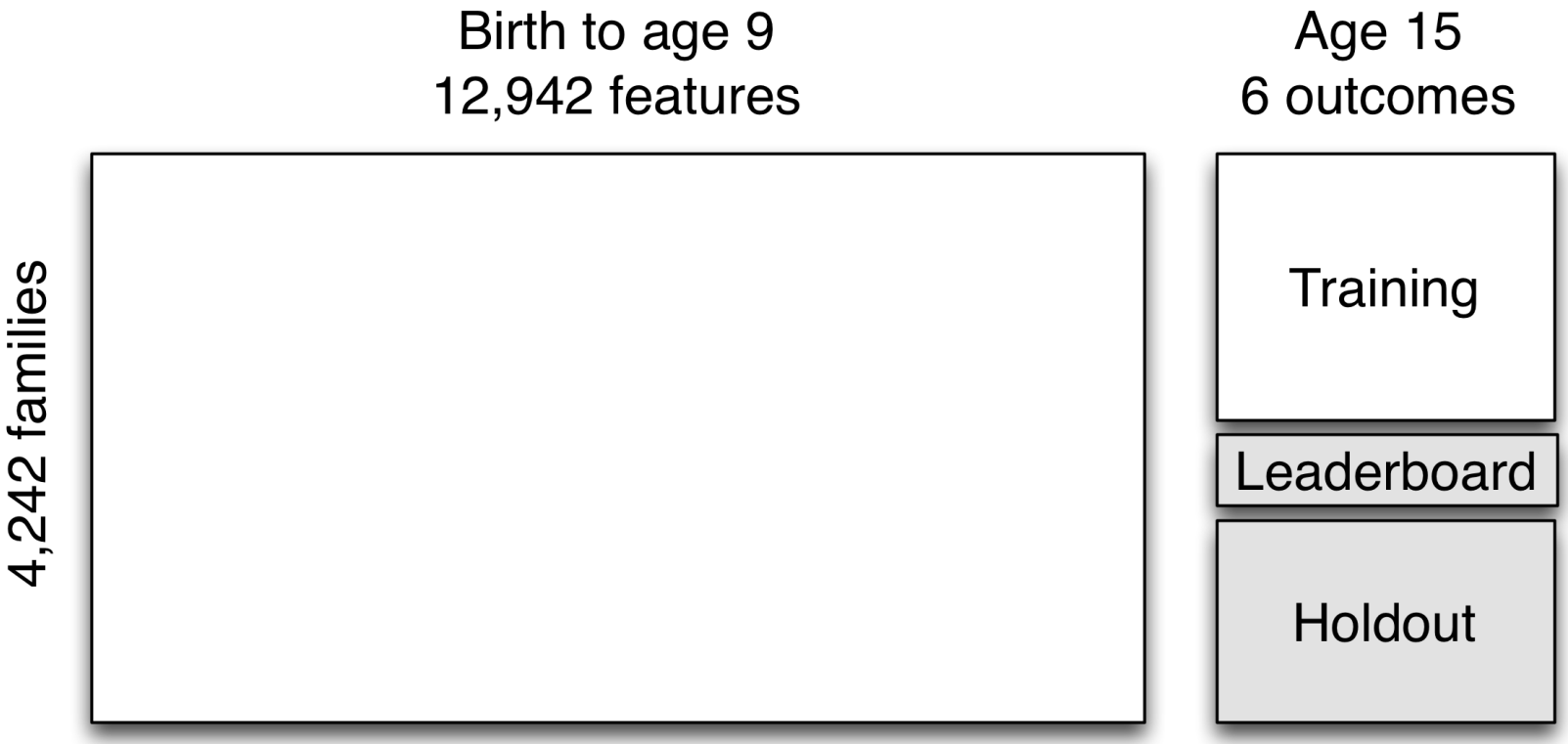
1. Informed consent
2. Already available to researchers
3. Already used in scientific and policy debates

Fragile Families
& Child Wellbeing Study
PRINCETON | COLUMBIA

Features relevant to privacy:

1. Informed consent
2. Already available to researchers
3. Already used in scientific and policy debates
4. Contain information from many respondents


Birth
Age 1
Age 3
Age 5
Age 9
Age 15
Core mother survey
Primary caregiver survey
Combined
Core father survey
In-home assessment
Child survey
Child care provider survey
Teacher survey


Birth to age 9
12,942 features
Age 15
6 outcomes
4,242 families
Training
Leaderboard
Holdout

The Princeton IRB approved the project.

The Princeton IRB approved the project.

The data met standards for de-identification.

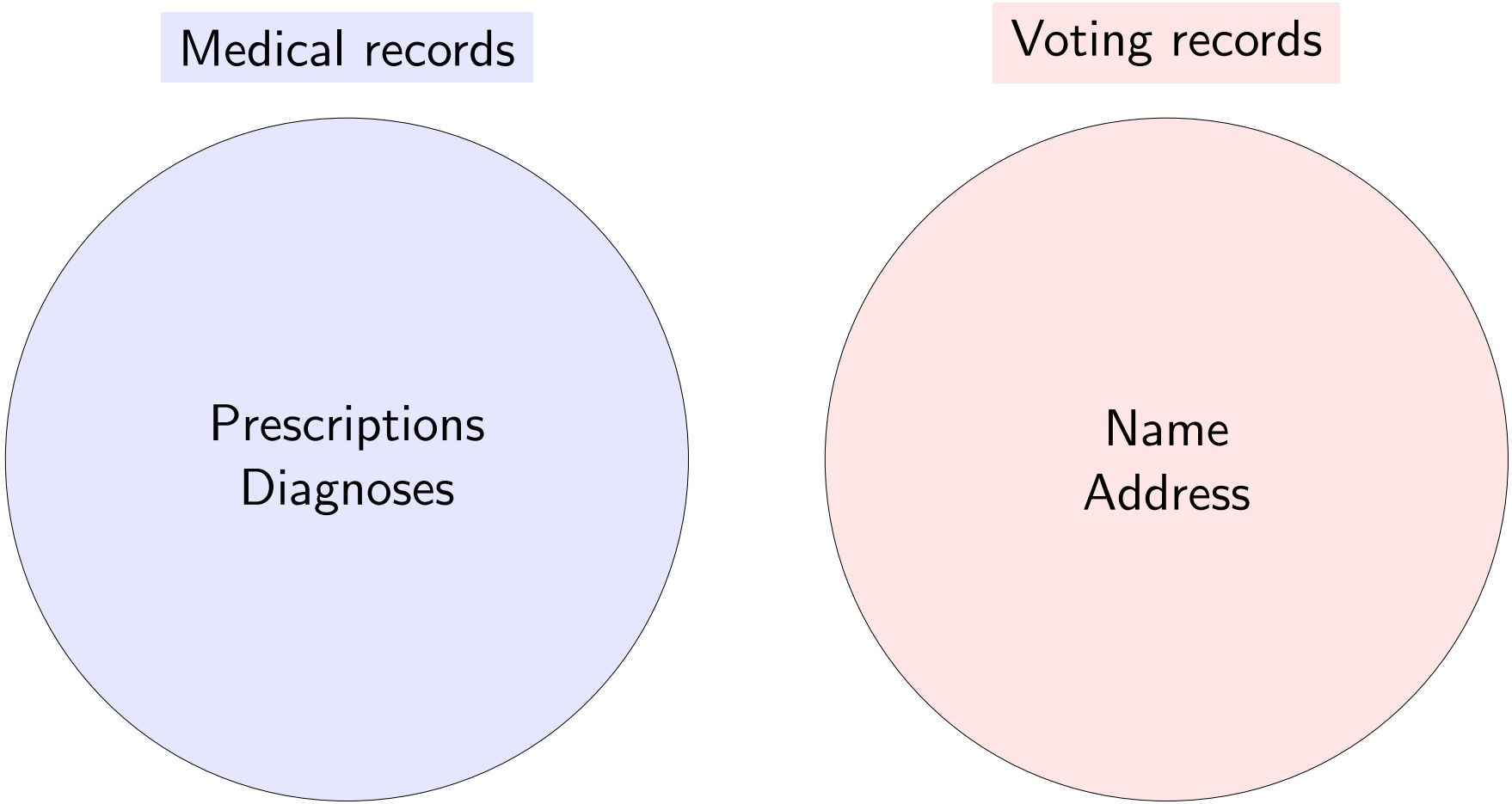
The Princeton IRB approved the project.

The data met standards for de-identification.

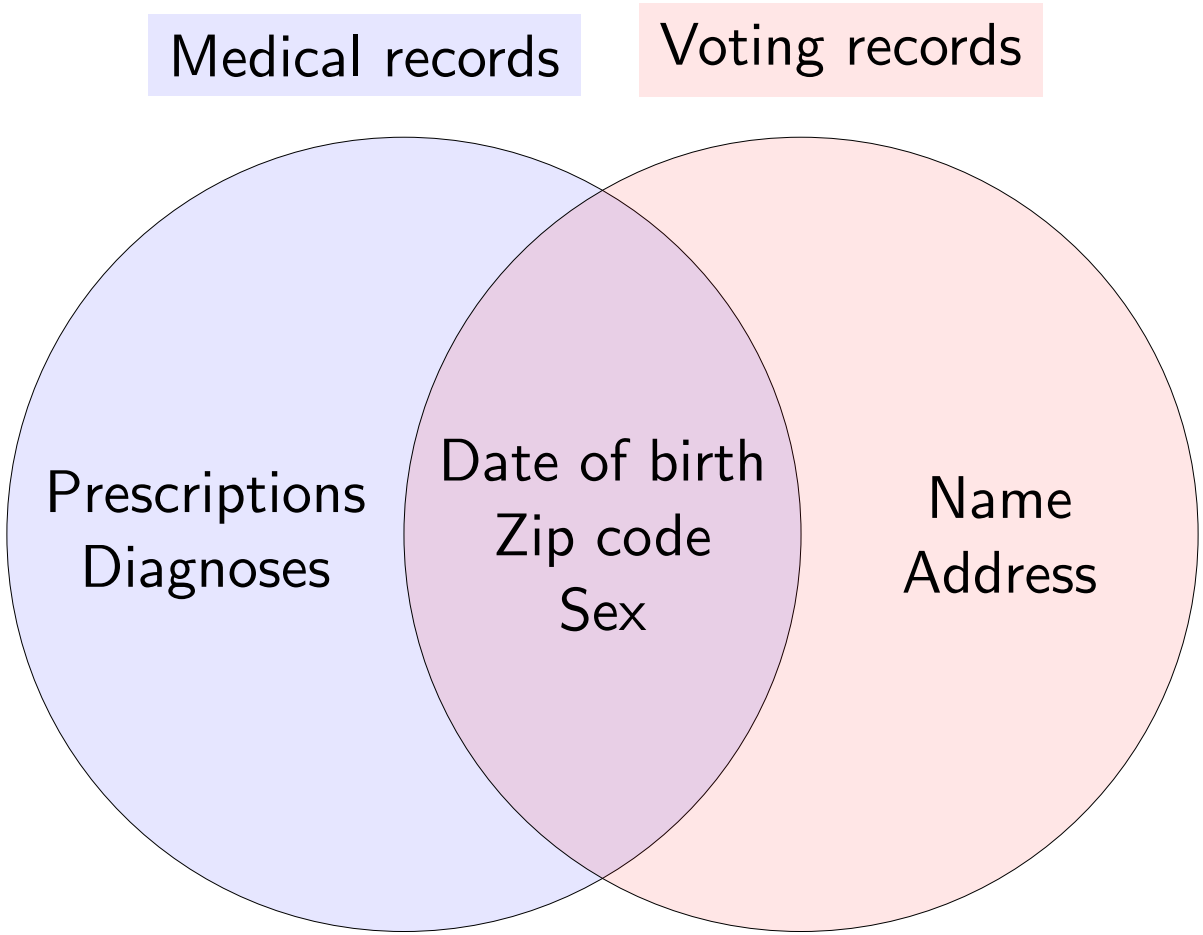
Why worry?

Art by David Finch
Source: Wikipedia

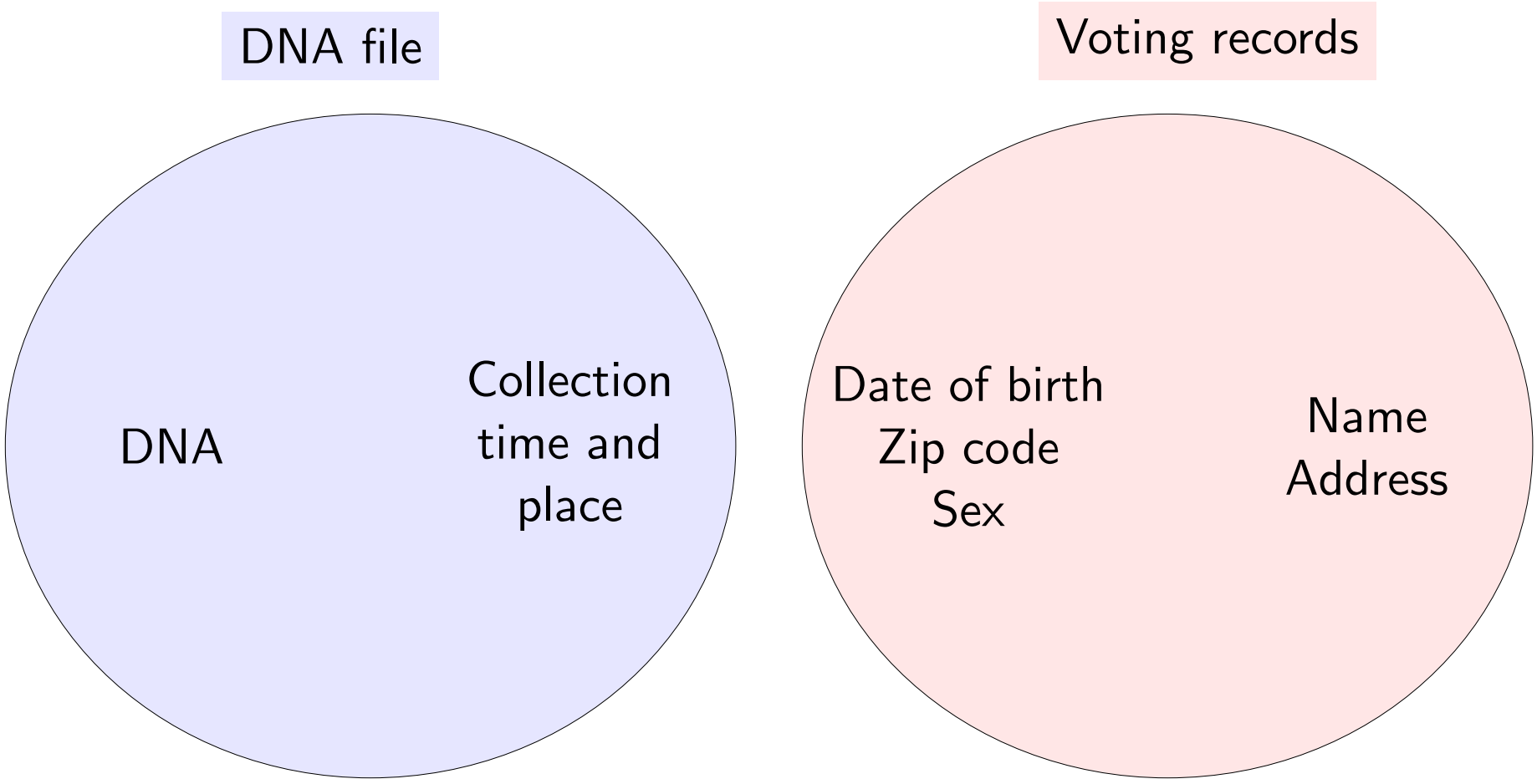
A. Sweeney (1997) re-identified Massachusetts medical records



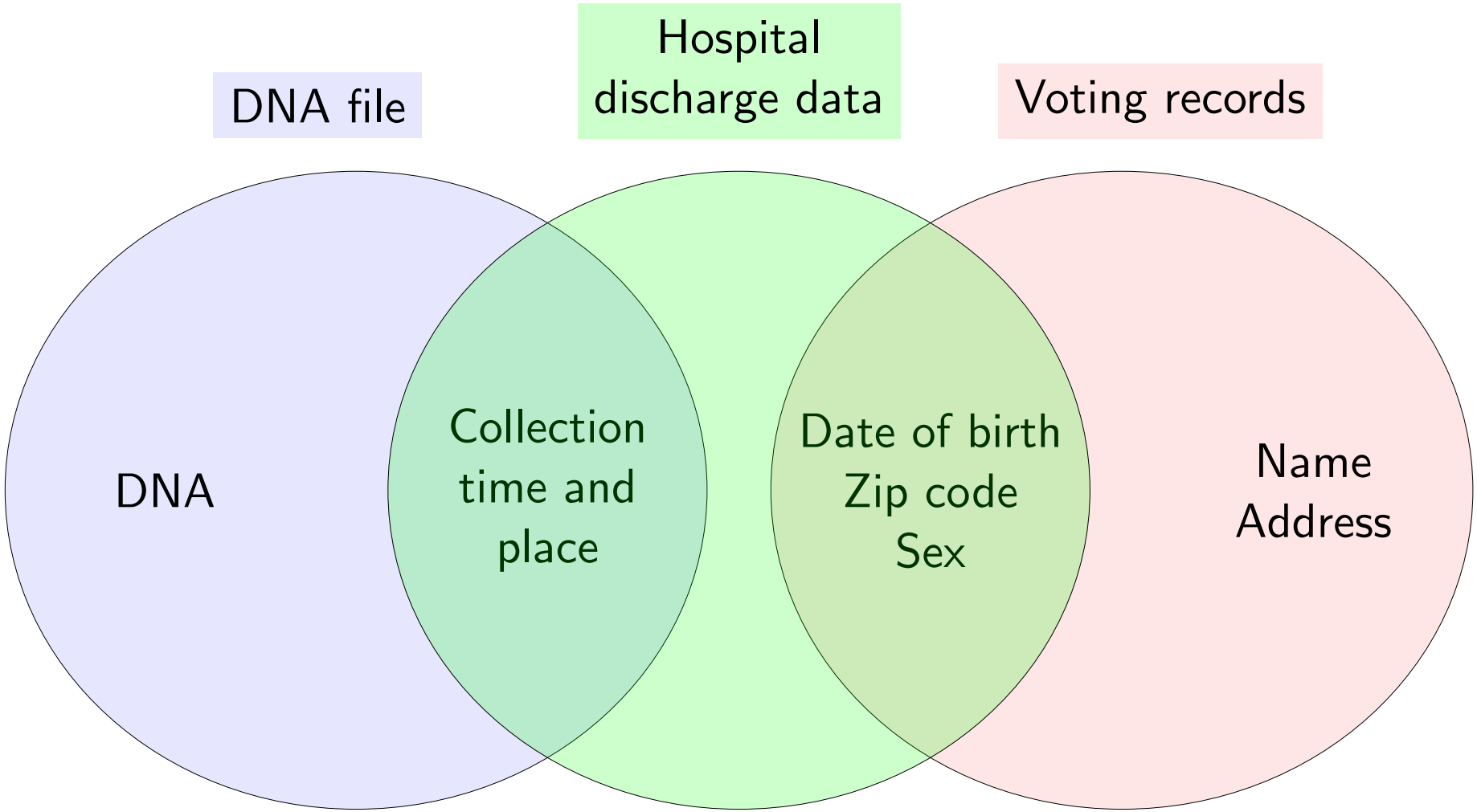
A. Sweeney (1997) re-identified Massachusetts medical records



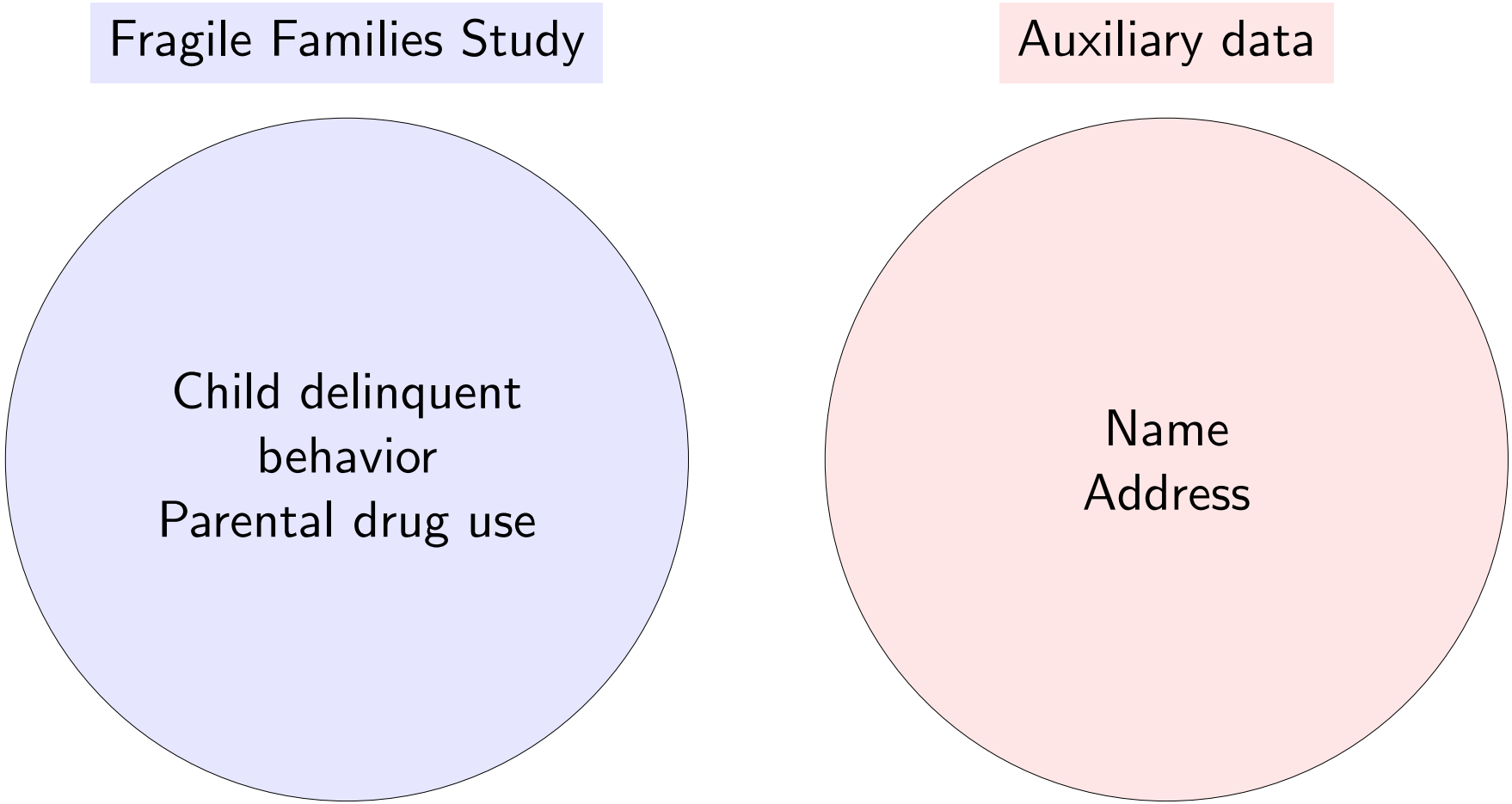
B. Malin and Sweeney (2004) re-identified genomics data



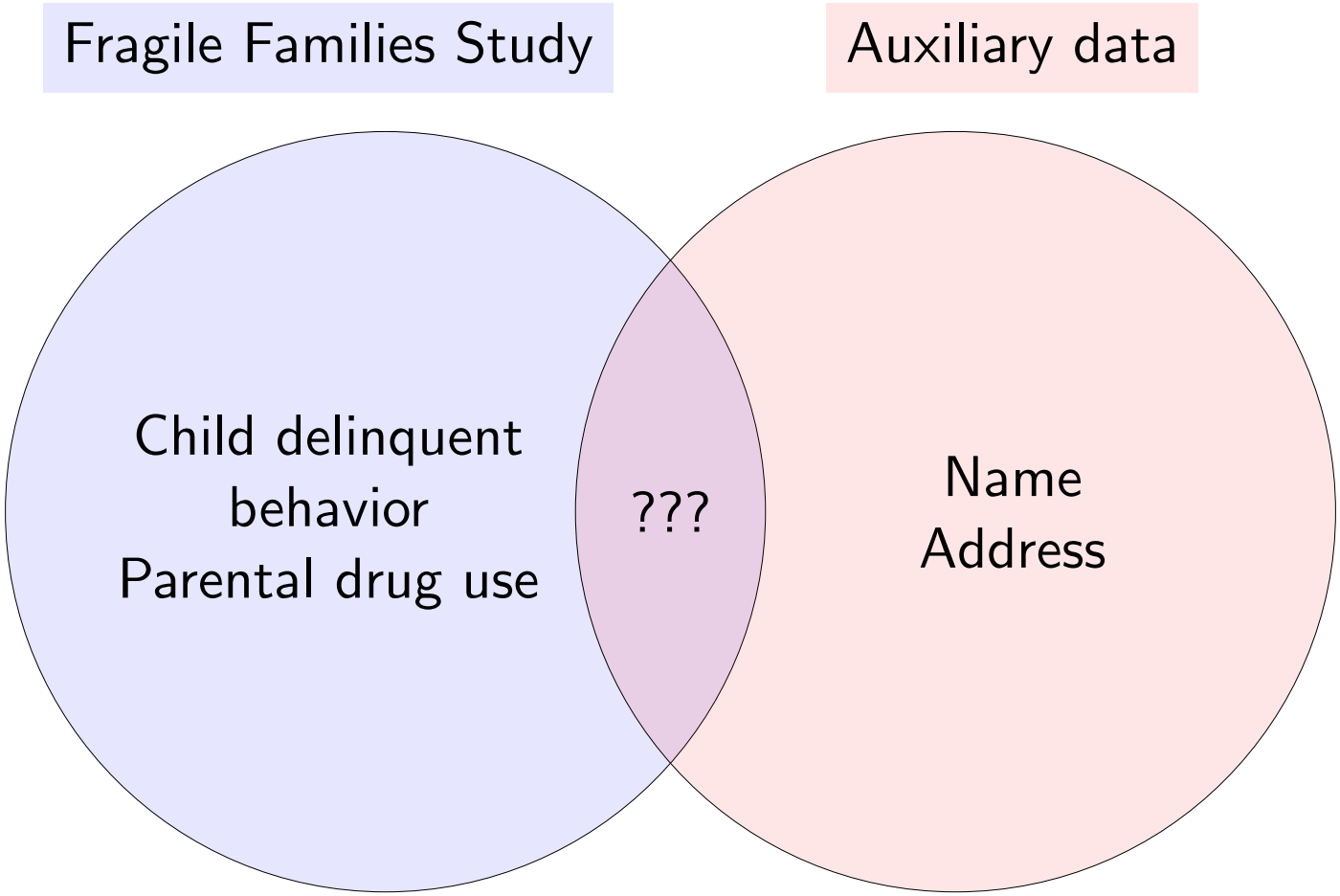
B. Malin and Sweeney (2004) re-identified genomics data



C. We don't know the auxiliary data that may exist.



C. We don't know the auxiliary data that may exist.



Provable privacy protections

Promising areas of active research:

- Differential privacy
- Cryptography

Provable privacy protections

Promising areas of active research:

- Differential privacy
- Cryptography

At their present development,
neither applied in our setting.

Provable privacy protections

Promising areas of active research:

- Differential privacy
- Cryptography

At their present development, neither applied in our setting.

Turned to approaches without provable guarantees.

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Main threats:

1. Privacy researcher*

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Main threats:

1. Privacy researcher*
2. Nosy neighbor

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Main threats:

1. Privacy researcher*
2. Nosy neighbor
3. Troll

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Main threats:

1. Privacy researcher*
2. Nosy neighbor
3. Troll
4. Journalist

Threat modeling

Criteria that represent a threat of a re-identification attack:

- skills
- auxiliary data
- incentives

Concrete experience: We attacked our own data for 1.5 months.

Main threats:

1. Privacy researcher*
2. Nosy neighbor
3. Troll
4. Journalist
5. Cheater

← **Mitigations** →

↓ **Threats** ↑	Low profile	Careful language	Challenge structure	Application process	Ethical appeal	Modifications to data
Privacy researcher	✓	✓	✓	✓	✓	✓
Nosy neighbor	✓			✓		
Troll	✓		✓	✓		✓
Journalist	✓	✓	✓	✓	✓	✓
Cheater		✓	✓		✓	✓

Response plan

We mitigated but did not eliminate risks.

We needed a team ready to respond in a crisis

- Computer scientist who had re-identified datasets previously
- Lawyer and sociologist who studies privacy and inequality
- Respected journalist

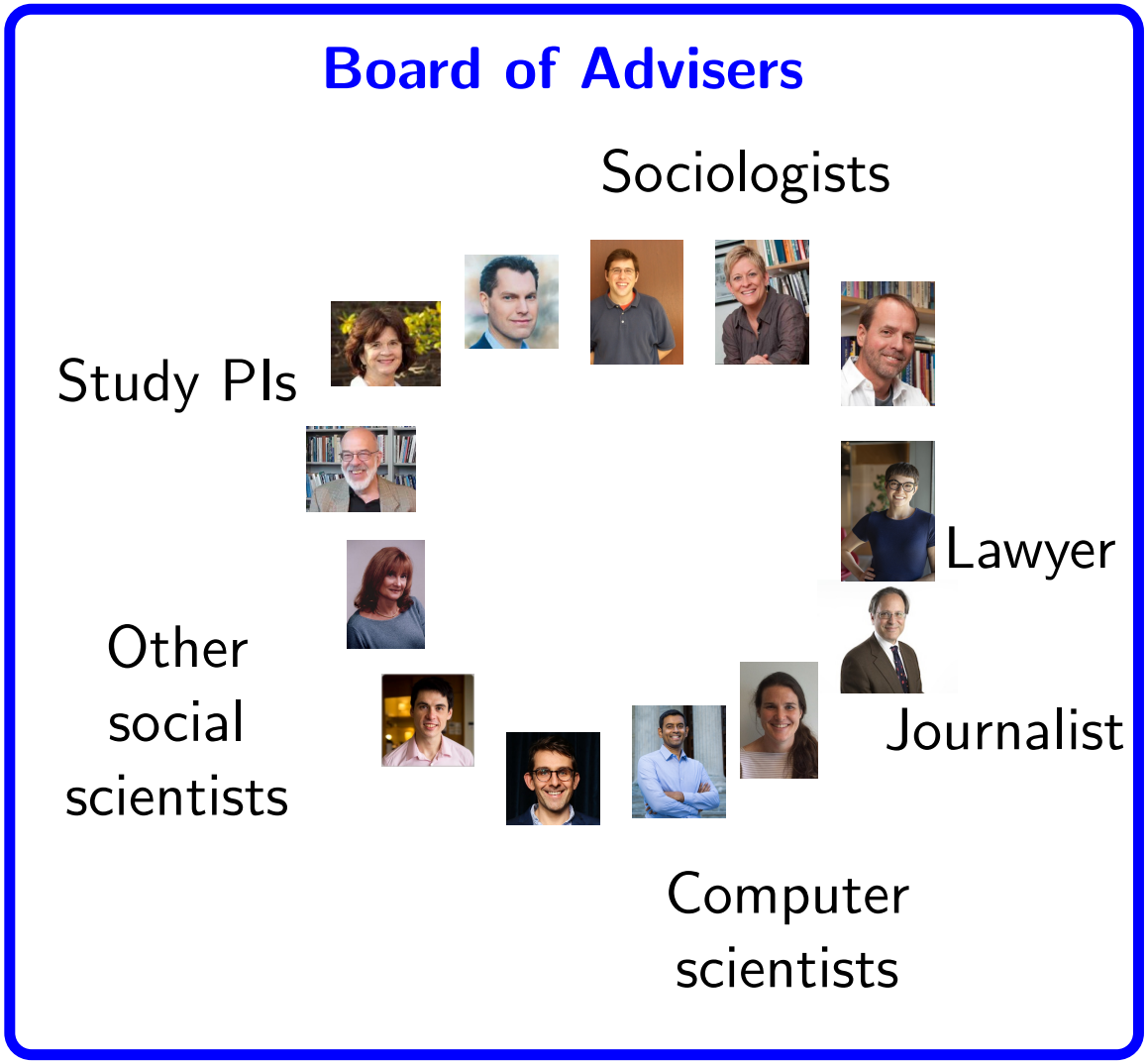
We were prepared to respond **quickly**.

Third-party guidance: Avoid groupthink

Third-party guidance: Avoid groupthink

Basic oversight
Princeton IRB

Third-party guidance: Avoid groupthink



Basic oversight
Princeton IRB

Third-party guidance: Avoid groupthink

Board of Advisers

Sociologists

Study PIs

Lawyer

Other social scientists

Journalist

Computer scientists













Basic oversight

Princeton IRB

Outside advice

Philosophy professor

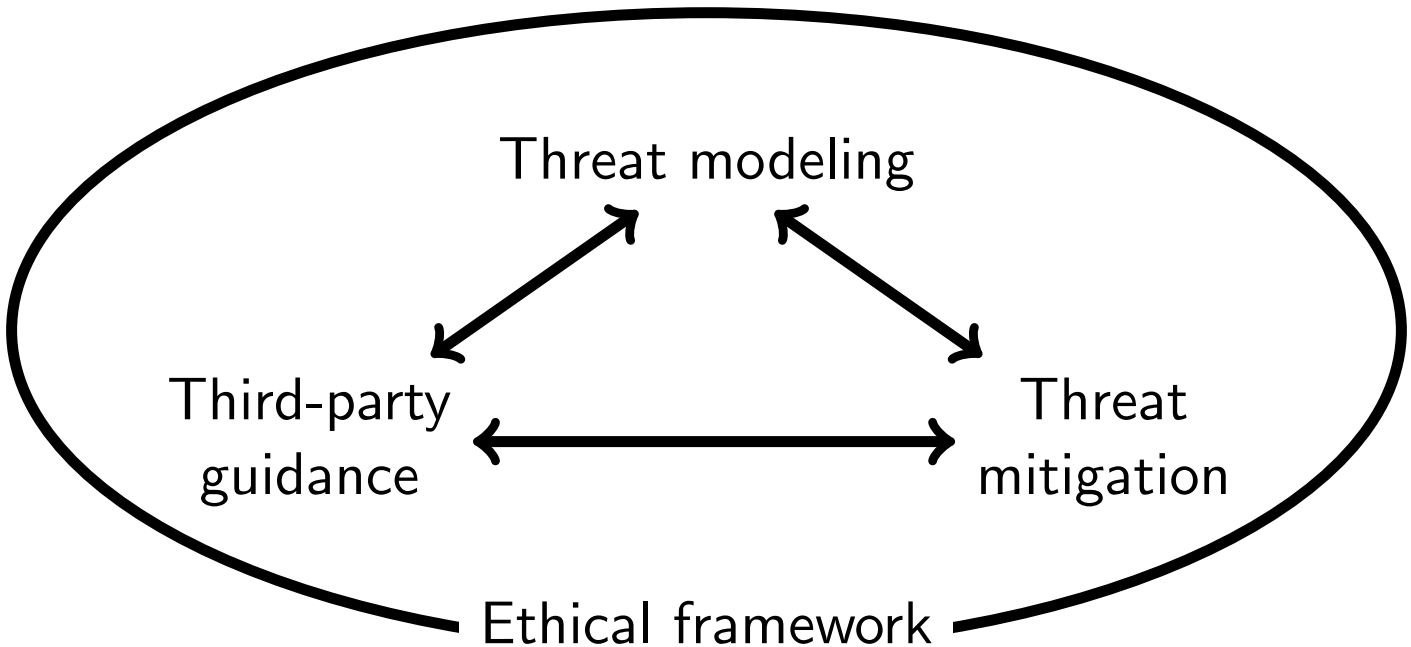
Health lawyer

Public interest lawyer

Member of military

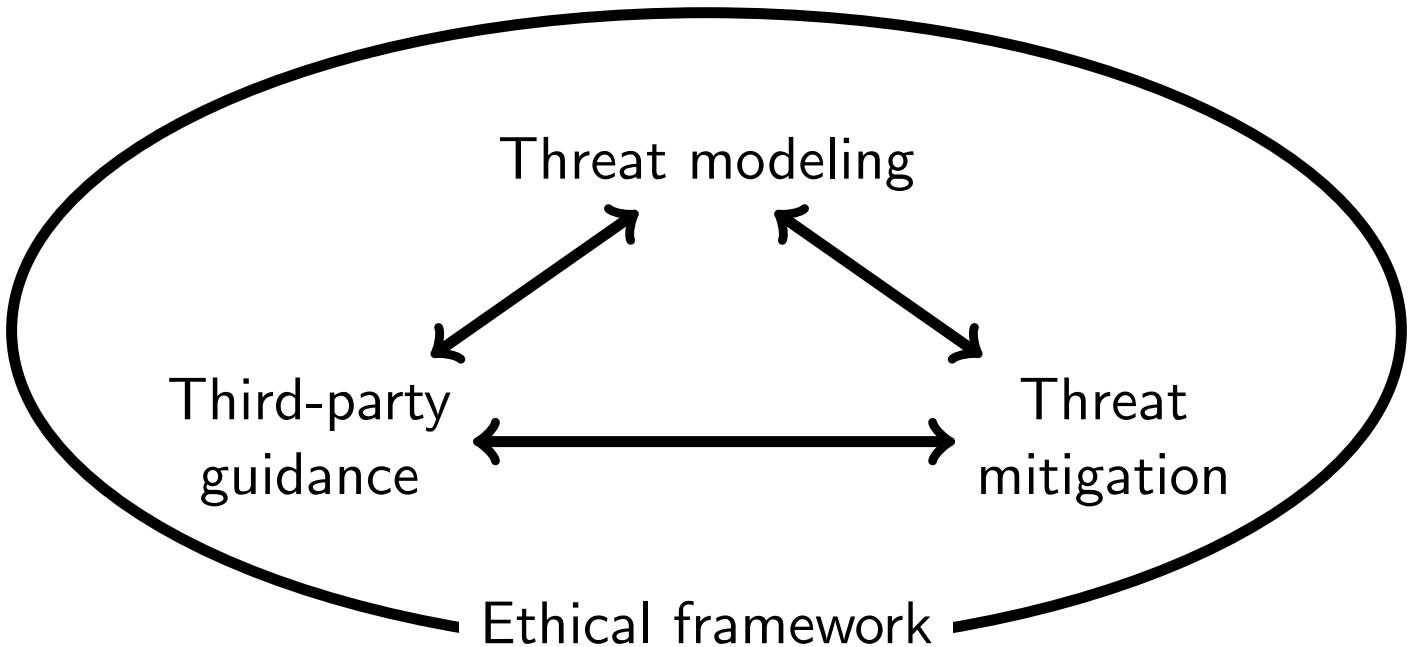
Ethics grounded in the Belmont Report

- Respect for persons
- Beneficence
- Justice



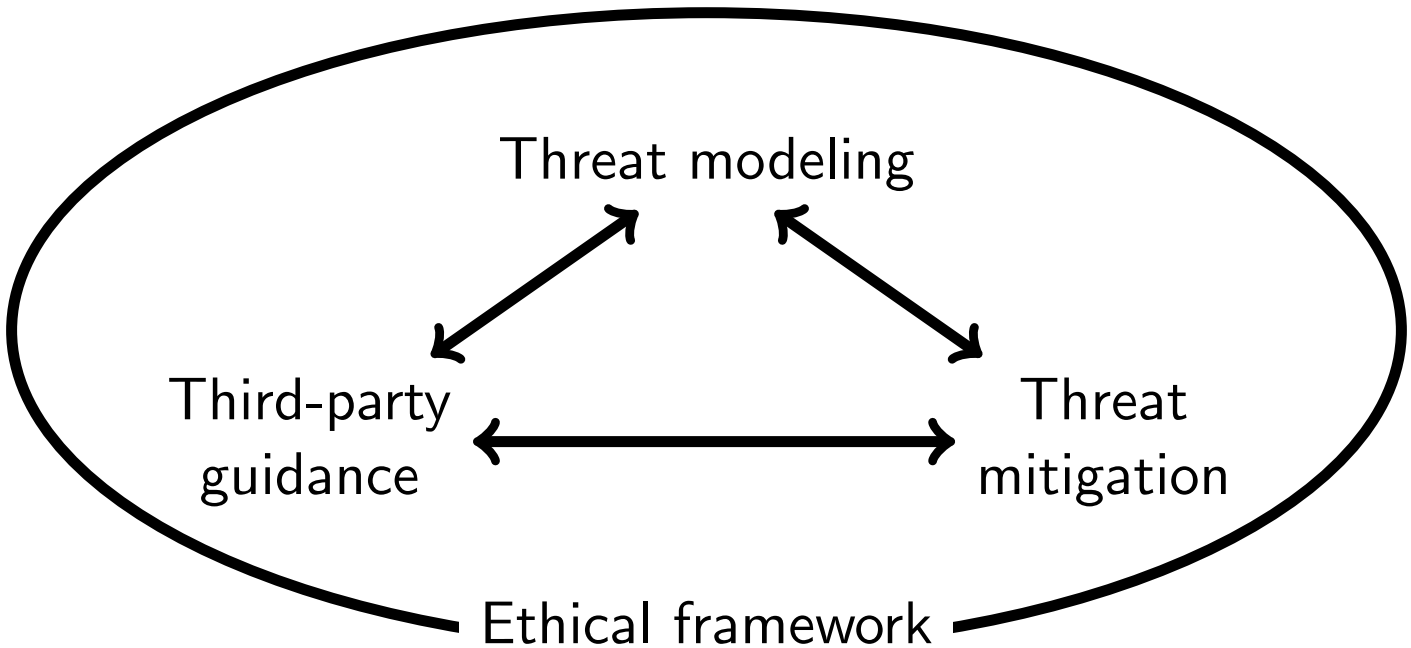
Ethics grounded in the Belmont Report

- Respect for persons $\rightarrow$ honor participants' agreement
- Beneficence
- Justice



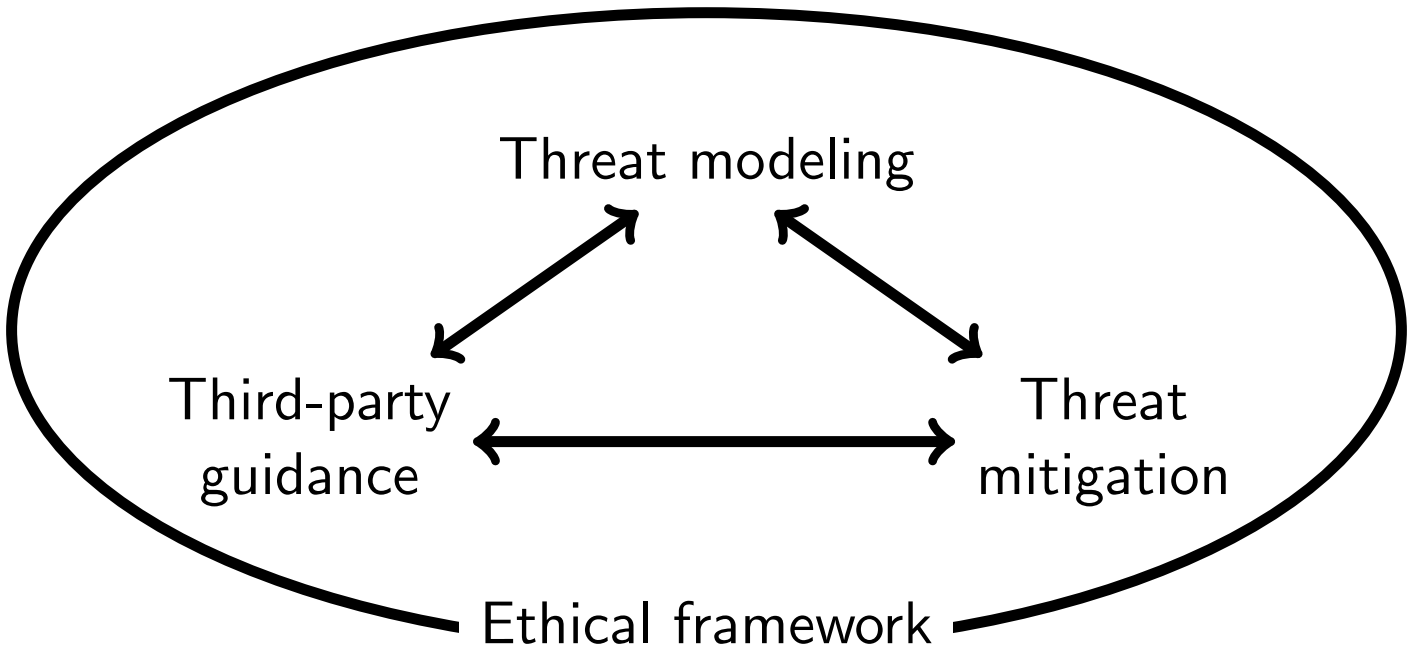
Ethics grounded in the Belmont Report

- Respect for persons $\rightarrow$ honor participants' agreement
- Beneficence $\rightarrow$ maximize benefits and minimize harms
- Justice



Ethics grounded in the Belmont Report

- Respect for persons $\rightarrow$ honor participants' agreement
- Beneficence $\rightarrow$ maximize benefits and minimize harms
- Justice $\rightarrow$ population to benefit is similar to study population



Decision time

Decision time

1. Already have IRB approval.

Decision time

1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?

Decision time

1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?
3. Central organizers proposed proceeding

Decision time

1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?
3. Central organizers proposed proceeding
4. Board of Advisers agreed

Decision time

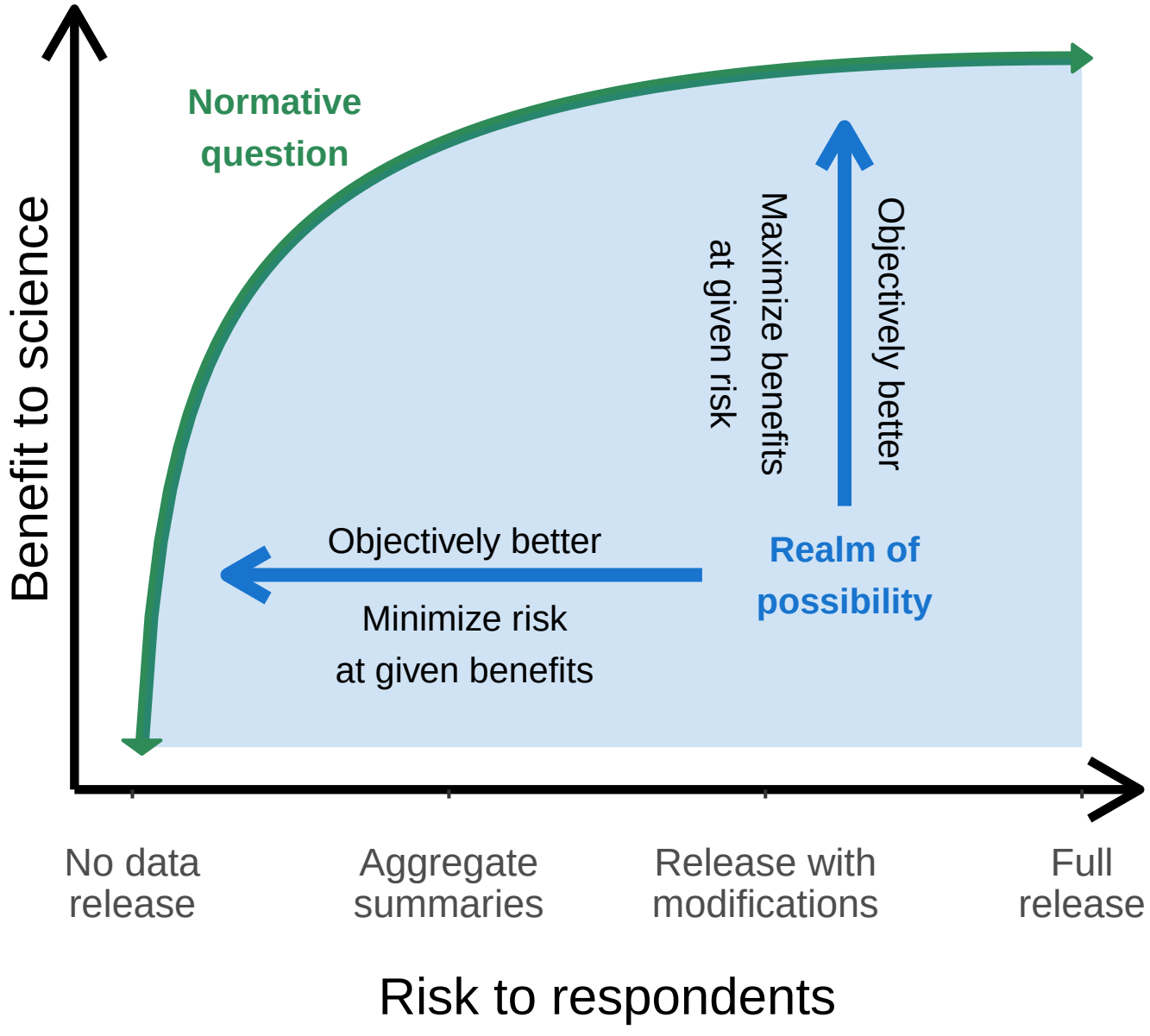
1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?
3. Central organizers proposed proceeding
4. Board of Advisers agreed
5. Pilot test in a machine learning class (early spring 2017)

Decision time

1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?
3. Central organizers proposed proceeding
4. Board of Advisers agreed
5. Pilot test in a machine learning class (early spring 2017)
6. Full-scale launch (mid-spring 2017)

Decision time

1. Already have IRB approval.
2. Stepped back. Are we comfortable proceeding?
3. Central organizers proposed proceeding
4. Board of Advisers agreed
5. Pilot test in a machine learning class (early spring 2017)
6. Full-scale launch (mid-spring 2017)
7. Continuous consideration overseen by Board of Advisers


Benefit to science
Normative question
Maximize benefits at given risk
Objectively better
Objectively better
Realm of possibility
Minimize risk at given benefits
No data release
Aggregate summaries
Release with modifications
Full release
Risk to respondents

Privacy, ethics, and data access: Generalizable principles

Key elements of our process may help promote the **ethical use of other data sources** by future researchers.