Criminal Justice Administrative Records System: A Collaborative Approach to Building Next Generation Criminal Justice Data Infrastructure

Michael Mueller-Smith

University of Michigan

November 4, 2021
Disclaimer

- Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. While this work was not subject to formal Census Bureau content review, Census staff reviewed all statistical output to ensure that no confidential information was disclosed.

- This presentation meets all of the U.S. Census Bureau’s Disclosure Review Board (DRB) standards and has been assigned DRB approval numbers CBDRB-FY19-444, CBDRB-FY20-ERD002-021, CBDRB-FY21-ERD002-008.
Outline

1. Project summary
2. Project detail
3. Project operations
4. Selected contributions and findings
What is CJARS?

CJARS, founded in 2016, is a joint data infrastructure project between University of Michigan and the U.S. Census Bureau

Building a novel data platform to modernize research and statistical reporting on the U.S. criminal justice system, including:

- Event-level criminal justice data with nationwide scope
- Tracking across key milestones in the justice system
- Capacity to link with individual-level survey and administrative data at the U.S. Census Bureau

Financial support from the National Science Foundation, the Bill and Melinda Gates Foundation, the Laura and John Arnold Foundation, the Robert Wood Johnson Foundation, the University of Michigan, and the U.S. Census Bureau
Incomplete criminal justice data infrastructure

- Limits how performance is measured, which policy questions are asked
- Leads to varying definitions of recidivism
- Makes it difficult to identify effective policy levers
- Prevents comprehensive benefit-cost analyses
CJARS project goals

- Collect, harmonize, and link longitudinal, multi-jurisdictional CJ data to track individuals and cases across space and time
- Embed criminal justice contact and policy within broader socio-economic data
- Enable timely and innovative federal statistical products about CJ system and justice-involved population
- Provide national research platform for qualified researchers on approved projects
Outline

1. Project summary
2. Project detail
3. Project operations
4. Selected contributions and findings
Project scope and data collection

Arrest
3k sheriff’s offices/3k jails

Pre-trial detention
3k prosecutor's offices

Adjudication
50 state court admins/3k criminal courts

Probation
50 state departments of corrections

Incarceration
18k police departments

Parole
51 state criminal records repositories

Community corrections agencies
Project scope and data collection

- Arrest
- Pre-trial detention
- Adjudication
- Probation
- Incarceration
- Parole

- 18k police departments
- 3k sheriff’s offices/3k jails
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections
- 51 state criminal records repositories
- Community corrections agencies
Project scope and data collection

- Arrest
- Pre-trial detention
- Adjudication
- Probation
- Incarceration
- Parole

3k sheriff’s offices/3k jails
3k prosecutor’s offices
50 state court admins/3k criminal courts
50 state departments of corrections

18k police departments
51 state criminal records repositories
Community corrections agencies

Mueller-Smith (Univ. of Michigan)
CJARS: A Collaborative Approach
November 4, 2021
Project scope and data collection

- Arrest
- Pre-trial detention
- Adjudication
- Probation
- Incarceration
- Parole

- 3k sheriff’s offices/3k jails
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections
- 18k police departments
- 51 state criminal records repositories
- Community corrections agencies

Mueller-Smith (Univ. of Michigan)
CJARS: A Collaborative Approach
November 4, 2021
Project scope and data collection

18k police departments

3k sheriff’s offices/3k jails

3k prosecutor’s offices

50 state departments of corrections

50 state court admins/3k criminal courts

51 state criminal records repositories

Community corrections agencies
Project scope and data collection

- 3k sheriff’s offices/3k jails
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections
- 18k police departments
- 51 state criminal records repositories
- Community corrections agencies
Project scope and data collection

- Arrest
- Pre-trial detention
- Adjudication
- Probation
- Incarceration
- Parole

18k police departments
3k sheriff’s offices/3k jails
3k prosecutor’s offices
50 state court admins/3k criminal courts
50 state departments of corrections
51 state criminal records repositories
community corrections agencies
Project scope and data collection

- 18k police departments
- 3k sheriff’s offices/3k jails
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections

51 state criminal records repositories

Arrest → Pre-trial detention → Adjudication → Probation → Incarceration → Parole

Community corrections agencies
Project scope and data collection

- 3k sheriff’s offices/3k jails
- 3k prosecutor’s offices
- 50 state court admins/3k criminal courts
- 50 state departments of corrections
- 18k police departments
- 51 state criminal records repositories
- Community corrections agencies

Arrest, Pre-trial detention, Adjudication, Probation, Incarceration, Parole
Project scope and data collection

Out of scope: civil court proceedings, traffic offenses below misdemeanor level, crime reports not associated with arrest, juvenile records not transferred to criminal court, records without PII
CJARS stakeholders

- **University of Michigan** acquires and harmonizes data from *criminal justice agencies*, including law enforcement, judicial, and corrections agencies.

- **Census Bureau** links CJARS with other socioeconomic data to build data products for data providers, host CJARS access through FSRDC network, develop new data products, and improve operations.

- Qualified *researchers* on approved projects analyze CJARS in the FSRDCs.

- **Communities and policymakers** use CJARS-derived statistical products to understand the performance of the criminal justice system.
CJARS at a glance

Data Providers
- Federal Agencies
- State Agencies
- Local Agencies

U-M CJARS Repository
- U-M CJARS Secure Data Enclave Server
- Administrative Data
- Roster File
- Physical Archive

U.S. Census Bureau
- Person Identification Validation Servers
- Data Intake
- Project Approval
- Anonymized Research Servers
- Disclosure Review

External Researchers
- External research request
- Non-CJARS external publication

CJARS statistical reports
CJARS data coverage

- 100–300 data providers
- States with coverage of state court, state DOC, and/or state repository represent 48% of U.S. population
- 2b records
- 175m CJ events
- 36m unique individuals
Project documentation and resources

Data Documentation

Benchmarking Report

Proposal Guide

Fellowship Solicitation
CJARS can replicate BJS series

All results were produced without the use of Title 13 data.

CJARS: A Collaborative Approach

Mueller-Smith (Univ. of Michigan)
Outline

1. Project summary

2. Project detail

3. Project operations

4. Selected contributions and findings
Funding/resources

- UM has raised about $7m for project so far
- Census resources
- Increasingly difficult to fund general operations
- Growing research obligations could slow collection, harmonization, and Census production activities
- 2–3 year cycles increase risk of losing subject-matter expertise
Current staff

University of Michigan

- Lead: Mike Mueller-Smith
- 2 project managers
- 3 RAs
- 1 post-doc
- 2 PT grad students
- 2 PT undergrad programmers

US Census Bureau

- Lead: Keith Finlay
- Civic Digital Fellows (4 since 2020)
- 3 add’l researchers working on topical research projects
Governance

- Census Bureau acquiring UM data under legal agreement through 2026
  - Access to IRS tax data available through 2024
- Strategic data decisions consensus based
- Board of Directors/Scientific Advisory Board
- Census/FSRDC proposal review at UM is as light as possible
Project challenges

- Representative data system requires significantly more data acquisition
- Limited economies of scale in data acquisition and harmonization
  - Trade-offs in geographic coverage, procedural scope, schema detail, data quality
- Cannot depend on grant funds to UM for core Census Bureau outputs
Outline

1. Project summary
2. Project detail
3. Project operations
4. Selected contributions and findings
Innovations in data science/measurement

- Entity resolution at UM
  - Agency-validated identifiers across datasets/PII realizations used to algorithmically identify high-quality matches

- Automated offense string classification
  - Partnered with Measures for Justice
  - New offense categorization scheme based on FBI/BJS standard
  - ML model estimated from 400k hand-coded offense strings
  - Public classification service

- Family and residential history crosswalks at Census
  - Link residually unstable people to families and local labor markets

Mueller-Smith (Univ. of Michigan)
Cohort-level cumulative felony conviction risk

Note: Estimates based on CJARS data linked by Census-provided PIK to Census Numident (to identify Commuting Zone of birth, sex, race, and ethnicity). These results were disclosed by the U.S. Census Bureau’s Disclosure Review Board with approval number CBDRB-FY19-444.

Probability of any felony conviction by age 25

- Black men
- White men
- Hispanic men
- AIAN men
- Asian/PI men
- Black women
- White women
- Hispanic women
- AIAN women
- Asian/PI women

Cohort = year of birth
× commuting zone of birth
× sex
× race/ethnicity

Note: Estimates based on CJARS data linked by Census-provided PIK to Census Numident (to identify Commuting Zone of birth, sex, race, and ethnicity). These results were disclosed by the U.S. Census Bureau’s Disclosure Review Board with approval number CBDRB-FY19-444.
Children exposed to parental justice involvement

Note: Estimates based on CJARS data linked by Census-provided PIK to family relation crosswalk that was synthesized from a variety of survey and administrative data. These results were disclosed by the U.S. Census Bureau’s Disclosure Review Board with approval number CBDRB-FY20-ERD002-021.
Employment of the justice involved

All figures have been rounded according to Census Bureau DRB rules. All results were approved for release by the Census Bureau, authorization number CBDRB-FY21-ERD002-008.
Thank you

Michael Mueller-Smith
mgms@umich.edu

Learn more:
https://cjars.isr.umich.edu