Intimate Partner Violence (IPV) and Data Science: A case study of enhancing IPV data and implementing an IPV focused deterrence strategy

Kyle Ott, Data Analyst
Rachel Teicher, Director of IPVI

Applied Research in Crime and Justice Conference, Sydney
13 February 2019

The National Network for Safe Communities (NNSC)

- Well-known in the US for the group violence intervention (GVI)
    - "Operation Ceasefire" in Boston to over 60 cities implementing today
    - International adaptations underway

- New strategy being implemented: the Intimate Partner Violence Intervention (IPVI)

- Invited to Australia to present and hold roundtables on IPVI

- More in-depth presentation on IPVI on Wednesday February 20th at 11am at the University of Sydney Law School

Goals for this presentation

1) Brief overview of our unique way of working as an action research center

2) Quick introduction to the IPVI framework

3) Case study: example of new way of understanding IPV in a city implementing IPVI

Overview of the NNSC

NNSC is a partnership between action researchers at John Jay College of Criminal Justice and public safety stakeholders in cities around the United States and the world.

Together we focus on **implementing proven strategic interventions** to **reduce violence** and **improve public safety**, **minimize arrest** and incarceration, **strengthen communities**, and **improve relationships** between law enforcement and the communities it serves.

National Network for Safe Communities

Do no harm

Strengthen communities' capacity to prevent violence

Enhance legitimacy

Offer help to those who want it

Get deterrence right

Use enforcement strategically

NNSC Approach to Work

- Very applied, less on typical research: want to drive change on a daily/weekly/monthly basis

- Weekly advising calls with sites

- Embedded/frequent site visits

- Work directly with all levels of frontline practitioners: learn from the experts on the ground

- Build the coalition of the willing

- Peer learning: leverage (international) network of sites

- Map on strategies to fit local dynamics

NNSC: Theory of Change

1. Pick an important, intractable problem:

    → GVI / IPVI: most serious crime driven by small N

2. Assemble frontline coalition of the willing

3. Unpack the problem

4. Design and implement a solution

    → Create certainty; provide clear information about risk; mobilize moral voice of the community; offer support & outreach; face-to-face communication

5. Create new facts on the ground

6. Use new facts to drive change

    → Enhance legitimacy and procedural justice; follow-up and keep your promises; assess and evaluate

Intimate Partner Violence Prevalence

36% of women in the US have **experienced IPV** in their lifetime

Black et al 2011

40-55% of all murders of women are IPV homicides

Campbell et al 2003

Petrosky, 2017

15% of **all violent crime** is IPV

Truman & Morgan 2014

IPVI State of Play

- Piloted in High Point, North Carolina (2009-present)
- Actively working with 5 sites
  - Eventually 8 total
- Goal: address all IPV offenders known to CJ system
- In implementing the strategy, became clear that the real scope of IPV offending is unknown
- Knowledge of local offending dynamics is critical to adapting the strategy to the community it is intended to serve

Core elements of IPVI

→ Conduct qualitative and quantitative data analysis of local dynamics

→ Identify levels of offenders

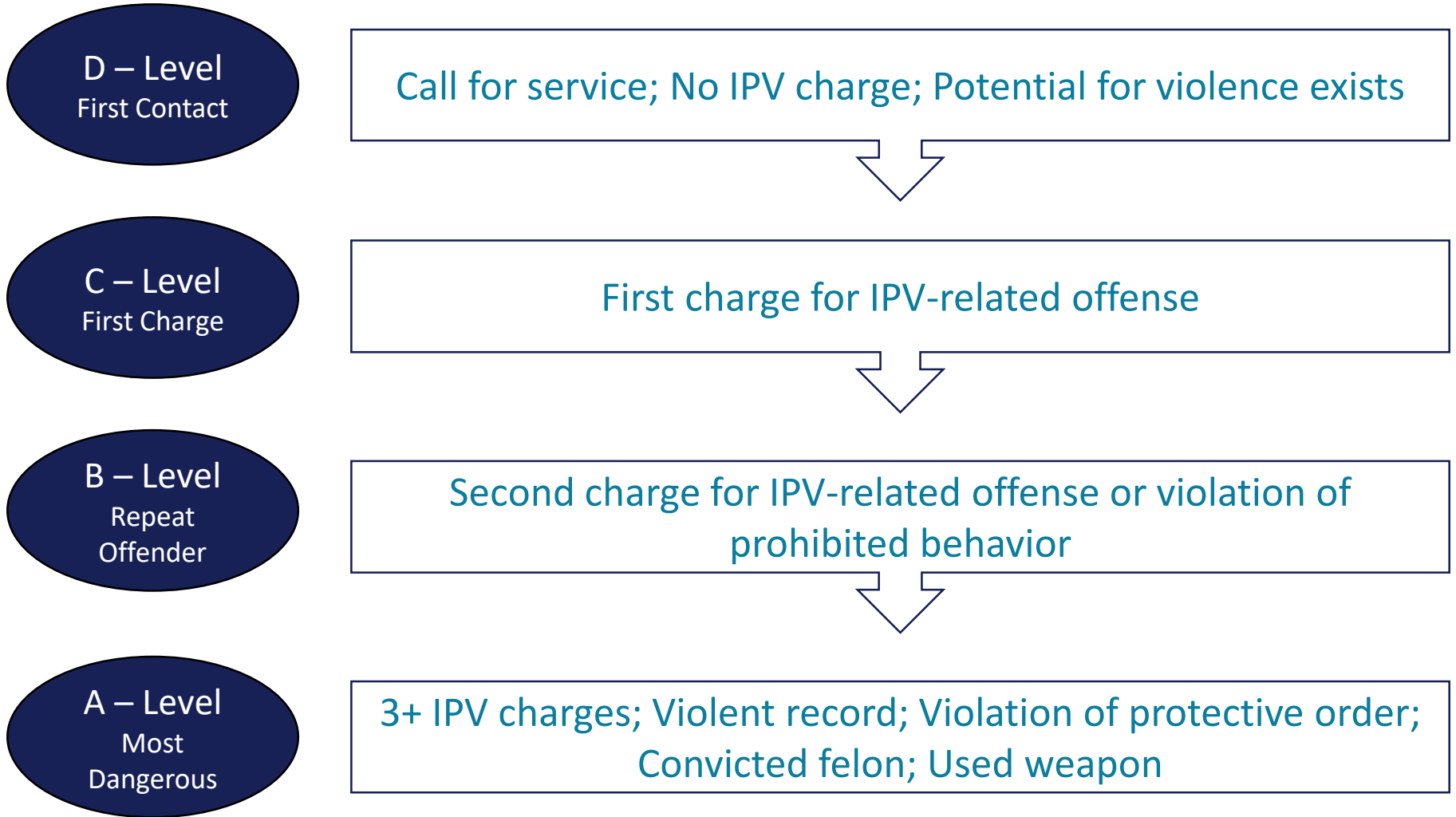→ Engage each level of offender with a specific approach

→ Promote offenders to the appropriate higher level if continued offending occurs

→ Provide affirmative outreach to victims at each level of offending

Sample Categories of Offending

D – Level
First Contact

Call for service; No IPV charge; Potential for violence exists

C – Level
First Charge

First charge for IPV-related offense

B – Level
Repeat Offender

Second charge for IPV-related offense or violation of prohibited behavior

A – Level
Most Dangerous

3+ IPV charges; Violent record; Violation of protective order; Convicted felon; Used weapon

We Use Data (Differently)

- Data is leveraged to move the needle on how partners think about a given type of crime
    - Important: often the data can validate our partners (e.g. many manually flag IPV)
- For IPV: just about no one codes it in the way we'd like
    - Australia may be different than the US ☺
- Our work: ___not___ about predicting the next IPV offender/incident, but rather understanding IPV offending in a local context
    - Think: not predictive but preventative policing

What Do We Need to Know?

- How much of a given location's violence is IPV?
    - Specifically IPV and IPV-related (aka spillover)

- Are high-level and/or chronic IPV perpetrators generalists or specialists?

- Are these individuals known to law enforcement practitioners?

- Can we apply a tailored deterrence regime to engage with IPV offenders of all levels (w/ parallel victim engagement)?
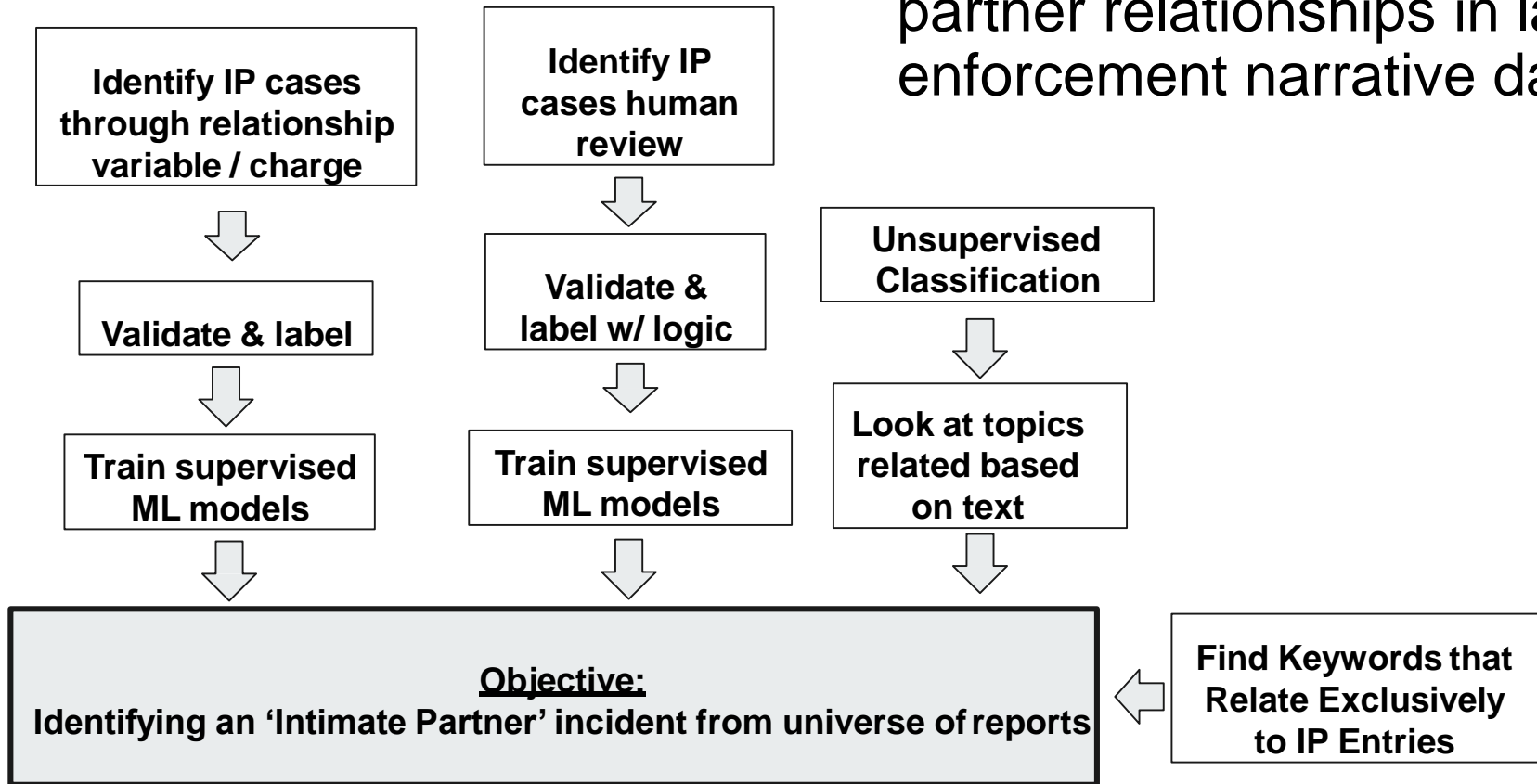
Case Study: Baton Rouge, LA

- 2nd largest city in Louisiana
- Population: 446,000
- Average ~10 IPV homicides annually
- Thousands of calls for service; multiple law enforcement agencies in the parish (county)
- Data issues:
  - Relationship field not reliable
  - Charges not helpful
  - Narrative not in easily accessible format (but! HTML format)

BR Data

- Police "DV" Data was far from complete:
    - 2% of incidents were IPV
    - Practitioners saying it was way higher
- Worked with IT department to extract narratives of 3+ years of incidents
    - 35 types of charges: ~13,000 unique incidents
    - Ave. report contains 642 words (max was 5,161)
- Qualitative: discussed homicide cases and top repeat DV offenders
- Analyzed criminal histories: generalists *not* specialists

Can We Extract IPV From Text Data?

- Engineering robust methods of identifying intimate partner relationships in law enforcement narrative data

Identify IP cases through relationship variable / charge

⬇

Validate & label

⬇

Train supervised ML models

⬇

Identify IP cases human review

⬇

Validate & label w/ logic

⬇

Train supervised ML models

⬇

Unsupervised Classification

⬇

Look at topics related based on text

⬇

Objective:
Identifying an 'Intimate Partner' incident from universe of reports

⬅

Find Keywords that Relate Exclusively to IP Entries

How Can We Detect IPV From Text Data?

- Note: Officers are *trained* to type out a sentence or two describing the relationship between a victim and suspect

- Supervised models – is this incident IPV?
  - Labeled random sample of data
  - Does the narrative have an IP(V) keyword?
  - Human reviewer: is this IPV?
  - Human reviewer: why is this IPV?
  - Apply insights from labeled sample to non-labeled incidents
  - Goal: can we get a better estimate of IPV in BR?

- Unsupervised models – future work

Keywords Indicative of IPV

Relationship: 'boyfriend', 'girlfriend', 'fiancé', 'beau', 'bae', 'baby', 'partner', 'couple', 'sex', 'sexual', 'romance', 'romantic', 'domestic', 'partnership', 'intimate', 'lover', 'love', 'dating', 'fling', 'hookup'

Marriage: 'marital', 'marriage', 'married', 'husband', 'wife', 'spouse'

Separation: 'ex', 'breakup', 'divorced', 'separated', 'cheated', 'affair', 'infidelity'

**Children/Custody:** 'custody', 'child', 'children', 'daughter', 'son', 'pregnant', 'mama', 'daddy'

Gendered Slurs: 'whore', 'bitch', 'mistress', 'slut'

Cohabitation: 'cohabit', 'roommate', 'bedroom',

IP Violence: 'abuser', 'batterer', 'consent', 'rape'

Human Label Feedback

- Individuals asked to review each report and respond whether they thought the incident constituted an act of intimate partner violence

- Individuals respond to one of the following:
    - 1: 'no'
    - 2: 'unknown'
    - 3: 'unclear'
    - 4: 'yes'

- Binary or scaled response variable?

- Human reviewers also asked to describe the logic for why IP(V) or not

- Issue of relationship undetermined ("unknown"), but charge seems to indicate IPV ("unclear")

- Goal: establish workflow with an easy to use GUI process
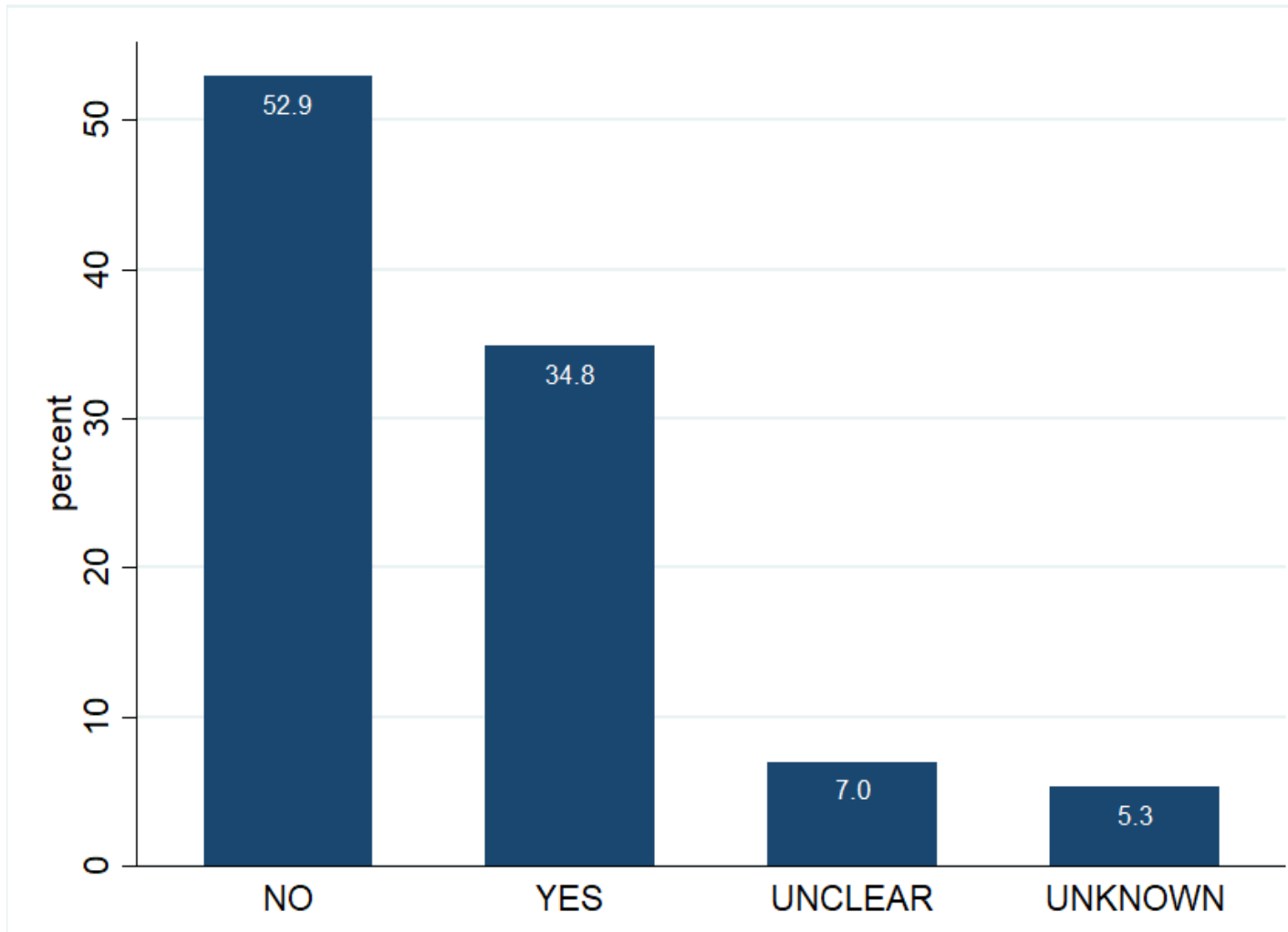
Results

- Machine learning results found that around 19-23%, depending on the model, of all the incidents in the text dataset were potentially IPV
    - This was meaningful and validating to our partners
    - Future analysis to fully validate results

- Caveats:
    - Testing the validity of any of these methods require that we have more training data
    - Insights gained from one city could be difficult to generalize
    - Plenty of other analyses to be done – but more interested in driving the intervention forward
- In the future: will have papers and package in Python/R

Thanks!
kott@jjay.cuny.edu
rteicher@jjay.cuny.edu