MINORITY REPORT:
Public Entity Analytics

AGRiP
2016 Governance & Leadership Conference
Nashville, TN

Chris Kogut, FCAS, MAAA
Allison Van Steensburg, ACI, CCRA
Areas of Focus

Background on how public entities are using predictive analytics:
- Predictive Policing
- Road Safety
- Child Welfare

Details on what data is being used and how the models are being developed

Insight on how pools might be impacted and ways pools might incorporate similar strategies
State of Affairs

Public entities seeking to be more efficient and more effective

Decreased Funding

Predictive analytics is translating insight into action

Proactive resource allocation

Data driven decision-making

Targeted intervention techniques
Caveats

1. Affiliation: descriptions not endorsements
2. Concerns: civil liberties, privacy, etc.
3. Sources: published research, vendor presentations, news articles
4. Recognition: schools and scalability
5. Clarification: forecast vs. prediction vs. prescription

Sources: published research, vendor presentations, news articles
Spectrum

Value

Difficulty

![Spectrum Diagram]

How can we change the outcome?
Predictive analytics is the use of data mining, statistics, and machine-learning techniques to identify the likelihood of future outcomes based on historical data.

- Field of study that gives computers the ability to learn without being explicitly programmed
- There are different ways ("algorithms") to accomplish this learning

- The step-by-step procedure to accomplish a given calculation
- Different algorithms have different qualities
PREDICTIVE POLICING
Connecting the dots between analysts and operations
Eras of Innovation
Law Enforcement

Professional
- Random Patrol
- Rapid Response
- Reactive Investigation

Community
- Partnership
- Problem Oriented
- Place Based Prevention

Intelligence-Led
- Information and Communications Technology
- Collaboration and Accountability
- Data-Driven and Offender Focused

Predictive
- Machine Learning Based Forecasting
- Enhanced Situational Awareness
- Intervention and Assessment

Technology + Prevention + Analysis = NOW
Milestones

- **1933**: First 2-way radio transmissions
- **Mid-1990s**: CompStat Hotspots Koper Curve
- **1997**: NIJ initiative to develop and implement data-driven interventions
- **2005**: Memphis PD launches Blue CRUSH
- **2007**: LAPD partners with UCLA to develop predictive policing program
- **2009**: NIJ First Symposium on Predictive Policing
- **2007**: LAPD partners with UCLA to develop predictive policing program
- **2005**: Memphis PD launches Blue CRUSH
- **2009**: NIJ First Symposium on Predictive Policing
- **2012**: PERF survey finds 38% currently using predictive policing 70% to implement next 2 to 5 years
- **2013**: RAND Report “Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations”
- **2015**: RAND Report focuses on technology needs for law enforcement
Defining Predictive Policing

“Predictive policing refers to any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention.”

National Institute of Justice
Predictive Policing Symposium
November 2009

“Predictive policing tries to harness the power of information, geospatial technologies, and evidence-based intervention models to reduce crime and improve public safety.”

National Institute of Justice
Overview of Predictive Policing
June 2014
## Predictive Policing Methods

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>Crimes</em>: forecast places and times with an increased risk of crime</td>
</tr>
<tr>
<td>2</td>
<td><em>Offenders</em>: identify individuals at risk of offending in the future</td>
</tr>
<tr>
<td>3</td>
<td><em>Perpetrator Identities</em>: create profiles that accurately match likely offenders with specific past crimes</td>
</tr>
<tr>
<td>4</td>
<td><em>Victims</em>: identify groups, or in some cases, individuals who are likely to become victims of crime</td>
</tr>
</tbody>
</table>

## Predictive Policing Myths

1. The computer actually knows the future.

2. The computer will do everything for you.

3. You need a high-powered (and expensive) model.

4. Accurate predictions automatically lead to major crime reduction.

The Prediction-Led Policing Business Process

### Predictive Tools of Today

<table>
<thead>
<tr>
<th>PredPol®</th>
<th>Beware®</th>
<th>Lumen®</th>
<th>HunchLab®</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses historical criminal data focused on location, time of day, and type of crime to generate 500 feet by 500 feet boxes around locations with highest likelihood of crime occurring during each 10 hour shift.</td>
<td>Searches, sorts, and scores publically available commercial records about people, locations, vehicles, and properties to inform public safety personnel about potential risks.</td>
<td>Collects, identifies, and indexes unstructured data from disparate data sources and combines with traditional structured fields for purposes of rapid search and analysis.</td>
<td>Advanced statistical model that uses historical criminal data, external data, and crime theory concepts to create shift by shift mission areas reflecting department specific priorities.</td>
</tr>
</tbody>
</table>

- **Value**...is in the tool’s ability to provide situational awareness of the crime risks.
- **Move beyond predictions to offer explicit support for resource allocation**.
- **Design intervention programs that include freedom, accountability, and synchronized support within both the department and the community**.

PredPol®
Hot-Spotting on Steroids

- Geospatial crime data analysis indicates hotspots don’t remain the same over time
- Understand criminal behaviors underlying complex mathematical models
  - Fixed (Environment)
  - Dynamic (Repeat Victimization)
- Assigns probabilities of future crime events to regions of space and time using only three data points
  - Past time, place, and type of crime
- Property crime and gun violence specific
- More efficient and more accurate resource deployment

Source: http://newrisingmedia.com
Intrado Beware®
Situational Awareness App

- Data mines publically available commercial records to provide predicted threat level (red, yellow, green) to responding officers in 60 seconds
- Drill-down capabilities provide basis of threat
- Does not include response or intervention recommendations
- Location, property, vehicle, and people specific search
- Enhances situational awareness to promote optimal outcomes and increase officer safety

Source: https://blog.pilpul.me/
Police records contain an enormous amount of critical information in unstructured data fields:

- Motive operandi (MO) details
- Gang names
- Known affiliations

Hybrid search engine which allows for the real-time retrieval of virtually any kind of electronic data

Indexing of unstructured data allows for its use in advanced analytical modeling

Built-in data sharing

Source: http://www.numerica.us/i2s/lumen/
HunchLab®
Integrated Operational Analytics

- Incorporates criminal records, criminal theory, and external data variables such as:
  - Seasonality, time of day/month/year
  - Sporting events, holidays, festivals
  - Terrain, points of interest

- Identifies shift by shift mission areas by building models that automatically determine what data sets best predict each type of crime

- Provides transparent and customizable crime type prioritization based on societal impact and patrol efficacy

Source: Various HunchLab YouTube videos
Severity acts as weights against modeled frequency by crime type in order to predict mission areas.

Default severity based on RAND Corp societal cost of crime index.

Department priorities can be incorporated into the model by varying either the default severity or the patrol efficacy.

Integrates analytics into all aspects of the asset allocation decision making process.

Source: Various HunchLab YouTube videos.
“It’s a meta-model in the sense that we are representing concepts from several approaches within one model. The model itself is most succinctly described as: a stacked model comprised of a gradient boosted set of decision trees trained with the AdaBoost loss function which is then wrapped in a generalized additive model regression.”
Predictive Analytics models (particularly in insurance) are often discussed in terms of segmentation – simple, accurate, explainable – but only one version of the story.

Minority Report: the information from the one precog whose prediction differed from the other two
Best models balance **bias** (how far average prediction is from actual) and **variance** (how different predictions are from each other). Ensemble methods are one way to do this.

“Ensemble Methods” generally refers to building a large number of somewhat independent predictive models (“learners”) and then combining them to improve stability and predictive power.
Models can be different for four key reasons (and combinations thereof):

- Population
- Hypothesis
- Modeling Technique
- Initial Seed

**Bagging** – learner trained on different subsets of data

**Boosting** – each learner attempts to correct the failures of the prior learner

**Stacking** – learner which combines outcomes of other learners

**Stratified Policing** integrates the effective elements of different policing approaches.

The challenge with both is how to weight the various approaches together.
ROAD SAFETY
Data driven decisions
CRASH: Crash Reduction Analyzing Statistical History

- Tennessee Highway Patrol (2014)
- Festivals, sporting events, weather patterns, time of day, historical accident data
- Predict where to deploy resources
  - Reduces accidents and violations
  - Reduces response times
- 5 x 6 mile squares; 4 hour periods
- Example: Between 7pm and 11pm, probability of serious crash is 63% at the corner of Broadway and Fifth
Types of Data

- Geospatial – location, terrain, school, bar or liquor store
- Live camera feeds (public and private) and city wide sensors
- Crime – historical crime records and real-time (active) crime events
- Weather data and other seasonality considerations
- Census and public (city records, property hazards) and private demographic data
- Criminal profiles – imagery, prior history, affiliations
- Research and expert analysis
- Community feedback and social media
- Census and public (city records, property hazards) and private demographic data
- Criminal profiles – imagery, prior history, affiliations
CHILD WELFARE
Targeting effective solutions
From Situational Awareness to Solution

Structured and unstructured data sets

Higher case loads

Budget cuts

Intra- and inter-departmental data sharing

Informed prioritization, targeted solutions, measurable impacts

Research on delinquency and recidivism

Social benefit by shifting from reactive to preventive

Closest to prescriptive analytics

Intra- and inter-departmental data sharing
Analytics Driven Intervention

- Probability of failure of reunification
- Information to assist case workers provided on consistent and timely basis
- Background information, information from school boards, geographic information
- Real-time information on each child in dashboards

Approach to Understanding Risk Assessment
- Predicting child abuse and likelihood of subsequent abuse
- Scoring, public records, crime data, research, social media, status changes, external factors

Delinquency Prevention Pilot program
- Predicting children in the child welfare system with likelihood of entering the juvenile justice system
- Prioritization of case management
- Create intervention approaches specific to the situation
- Importance of practical implementations

Permanency
- MindShare

Safety
- AURA

Well-Being
- DPP
Other Public Entity Analytics

- Property, fire marshal, and health inspections
- College crime incidents on game day
- Real-time location hazard assessments in route to 911 call
- Prison releases – predict least likely for repeat offenses
- Innovative community collaboration
- Spot unknown trends; confirm institutional folklore and intuition
- Performance management and financial accountability
- What Works Cities

Innovative community collaboration

Spot unknown trends; confirm institutional folklore and intuition

Performance management and financial accountability

What Works Cities

Real-time location hazard assessments in route to 911 call

College crime incidents on game day

Prison releases – predict least likely for repeat offenses

Property, fire marshal, and health inspections
Why Does it Matter to Pools?
Impact

- Law Enforcement Liability
- Unknown Exposures (technology outpacing laws and coverage exclusions)
- School Bus Safety
- Legal and Legislative Issues
- Cyber Risks
Extending the Data

**Traditional**
- Descriptive – historical trends
- Predictive – claims (case reserves) and underwriting (premiums)

**Prescriptive**
- Loss Control, Site Surveys, Risk Management Questionnaires, Targeted Programs Offerings, Performance / Accountability

**Trending**
Pools and Proactive Prevention

- Informed Resource Allocation
  - Claims staff
  - Member representatives

- Extending the Data
  - Provide support for underwriting assumptions based on a more granular view of your members
  - Move away from one size fits all assessment of risk
  - Use publically available data to avoid unwanted data calls to allow for analysis of potentially useful exposure information
  - Combine data across members for more robust information for pool and for membership
  - Capture data for future use; remove data silos

- Targeted Intervention Techniques
  - Combine claims data with exposure information to enhance loss control
  - Use analytics to create member specific risk assessments, exposure specific site surveys, and performance proven programs
Thank you