

# Data Science For Social Good

[Home](#) [Conference](#) [People](#) [Projects](#) [Blog](#) [Resources](#) [In the News](#)

[Get Involved](#) [FAQs](#)



Share  
This  
Post!



## Police Project Update: Expanding and Implementing the Early Intervention System

By Rob Mitchum, Jacqueline Genova & Lin Taylor



In recent years, the issue of **police violence and misconduct** has repeatedly made headlines and **inspired protests across the United States**. With each incident, **the trust** between police departments and the

### Recent Posts

[Connecting the Dots for Better Outcomes](#)  
[Data Science for Social Good](#)  
[Fellowship Update](#)  
[Preventive Care for the City's Circulatory System](#)  
[Police Project Update: Expanding and Implementing the Early Intervention System](#)  
[Scoping Data Science \(for Social](#)

communities they protect **takes damage**, and the work of police officers in the field becomes **more difficult and dangerous**. As a result, police departments seek new ways of identifying officers at risk of adverse interactions with the public, so that they may intervene and offer additional training or assistance before such incidents occur.

For the last two years, the [Data Science for Social Good Fellowship](#) (DSSG) and the [Center for Data Science and Public Policy](#) (DSaPP) have partnered with police departments to apply data toward this goal. In summer 2015, as part of the [White House Police Data Initiative](#), a DSSG team worked with the Charlotte-Mecklenburg Police Department (CMPD) on structuring the data and building a prototype predictive model. The following summer, that project expanded when [CMPD returned as a partner](#) to expand their scope and the [Metropolitan Nashville Police Department \(MNPd\)](#) [joined the effort](#) to see if the model was replicable in additional cities.

Many departments already use an Early Intervention System (EIS) to flag individual officers at high risk of an adverse event with the public, but most of these systems use simplistic, inaccurate rules such as thresholds (e.g. 3 complaints in 180 days) to flag officers.

## [Good\) Projects](#)

### [Recent Comments](#)

[Data for Good Exchange](#)

[\(Bloomberg\) –](#)

[claireduvallet on](#)

[The Real World:](#)

[DSSG](#)

[Using data science to confront policing challenges – Latest](#)

[Ethiopian and](#)

[World News on](#)

[Police Project:](#)

[Lessons From](#)

[Charlotte](#)

[Using Data Science](#)

[to Confront](#)

[Policing Challenges](#)

[\(Video\) - eNews](#)

[Park Forest on](#)

[Police Project:](#)

[Lessons From](#)

[Charlotte](#)

[You Say You Want](#)

[Transparency and](#)

[Interpretability? –](#)

Adverse events could include inappropriate use of force, a citizen complaint, or even an avoidable on-duty injury — basically any occurrence that would be filed to a department's internal affairs division. In all of these projects, the central mission was the same: **to build a data-driven Early Intervention System (EIS) that could be used to intervene with officers or dispatches before an adverse event occurs.**

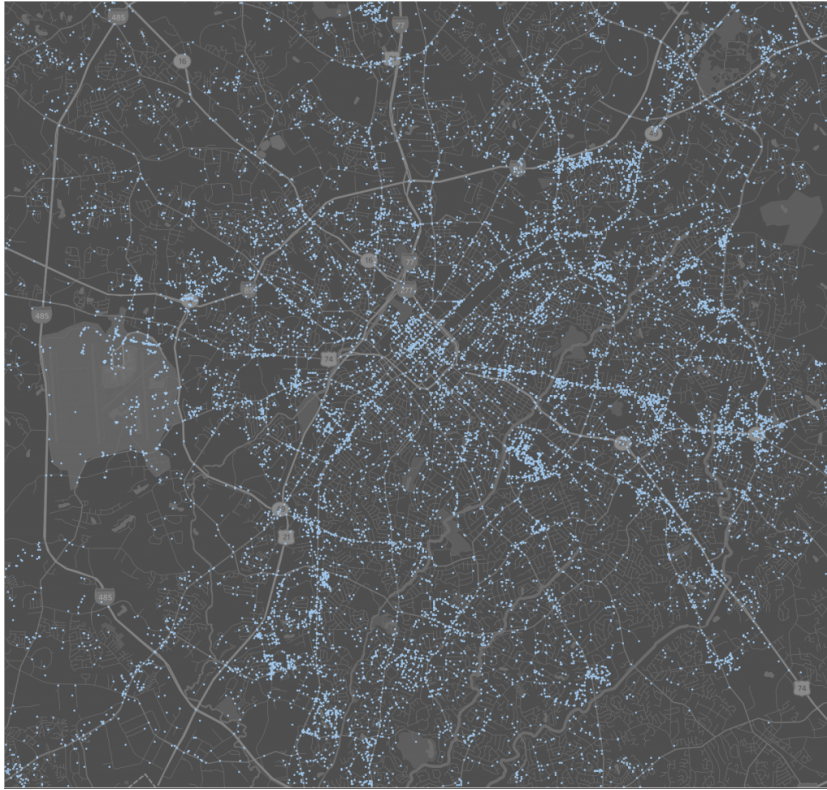
Having two partners and teams focused on the same challenge was a novel opportunity for DSSG. A second year working with CMPD also allowed us to further develop their EIS, exploring the risk of individual dispatches in addition to officers. Bringing in a second police department allowed us to test models on datasets from two municipalities, helping to build a more robust system. But most directly, it brought **twice the mental horsepower to bear on the problem**, with fellows Thomas Ralph Davidson, John Henry Hinnefeld, and Edward Yoxall (CMPD team) and Sumedh Joshi, Jonathan Keane, Joshua Mausolf, and Lin Taylor (MNPD team) united under mentors Joe Walsh and Jen Helsby and project manager Allison Weil.

Rayid Ghani –  
Machine Learning,  
Analytics, Data  
Mining, Obama for  
America, Edgify,  
University of  
Chicago, social  
media, Data  
Science, Big Data  
on Increasing  
Graduation Rates  
and Improving  
College Readiness  
for High School  
Students  
Why does my  
house alarm go off  
in a power cut? |  
Eurotech Security  
Blog on CDOT:  
Using Data to Shine  
a Light on the  
Impact of City  
Services

Archives

March 2017

February 2017



*One week of dispatches in the Charlotte-Mecklenburg Police District, mapped by the DSSG Charlotte team.*

## Two Teams Are Better Than One

The unique circumstances produced both friendly competition and cross-team collaboration. In one marathon session, the teams united for a day of brainstorming and whiteboarding, dreaming up the ideal database schema for organizing the data that powers their models. More broadly, they worked together on creating a common data pipeline that both projects could use to ingest data and produce officer and dispatch risk scores. By designing a data schema and EIS for two departments instead of one, the project created a more generalizable product that could be

January 2017  
October 2016  
August 2016  
July 2016  
June 2016  
April 2016  
February 2016  
January 2016  
December 2015  
November 2015  
October 2015  
September 2015  
August 2015  
July 2015  
June 2015  
May 2015  
April 2015  
March 2015  
February 2015  
January 2015  
December 2014  
November 2014  
September 2014  
August 2014  
July 2014  
June 2014  
May 2014  
February 2014  
January 2014

better adapted to any future police departments interested in using the code.

These models used data on **officer attributes** (e.g. demographics, join date), **officer activities** (e.g. arrests, dispatches, training), **and internal affairs investigations** (the case and outcome), provided by the partner police departments. The teams **supplemented that data with publicly available sources**, such as American Community Survey data, weather data, shapefiles, and quality-of-life surveys. They then simulated history to test how well our model (and current models used by CMPD and MNPd) would have predicted officers involved in adverse incidents if each department had used it in the past.

Compared to CMPD's existing EIS, our system would have correctly flagged 10-20% more officers who were later involved in adverse incidents and reduced the number of incorrect flags by 50% or more. In Nashville, the top-performing model (a variation on a Random Forest) was able to correctly flag 80% of officers who would go on to have an adverse interaction, while only requiring intervention on 30% of officers in order to do so.

[December 2013](#)  
[November 2013](#)  
[October 2013](#)  
[August 2013](#)  
[July 2013](#)  
[June 2013](#)  
[May 2013](#)

#### Categories

[Uncategorized](#)  
[News](#)  
[Project](#)  
[Blog Post](#)  
[Technical](#)  
[Curriculum](#)

#### Meta

[Log in](#)  
[Entries RSS](#)  
[Comments RSS](#)  
[WordPress.org](#)

## 2016 DSSG Data Fest: Metro Nashville Police Department



Both results are important — by flagging fewer officers, police departments can better direct limited resources for retraining or counseling. Because our EIS offers more flexibility and provides continuous risk scores rather than binary flags, departments can rank all officers by risk in order to allocate resources to those most in need first.

The Charlotte team also explored dispatch-level predictions, assigning risk scores to different types of calls based on knowledge gathered while talking to officers during site visits. Responding to a domestic violence call, for example, will be more stressful to an officer than a routine auto accident, and could be more likely to result in an adverse interaction. While the work on this model is still preliminary, the officer-level model allows the department to determine how far ahead the system should predict, down to the next dispatch. This

gives the department the ability to decide if it should send lower-risk officers to a call even if they're farther away.

### 2016 DSSG Data Fest: Charlotte-Mecklenburg Police Department



### From Development to Implementation

In the months since the summer ended, both departments continue to work with the Center for Data Science and Public Policy on implementing the new EIS. The Nashville team gave their department a list of the highest-risk officers according to our model, which MNPd subsequently used to send letters to the officers and their supervisors informing them of the results and specific risk factors that led to their score. We're now helping them **integrate the EIS into their existing IT system**, so that it will continuously update with new data.

Similarly, CMPD awarded us a contract to help implement our EIS on their system. We're building a web interface to help them and other partners evaluate and understand the performance of the models — trying to avoid the “black box” mystery of some machine learning predictions. The interface will also allow for feedback from supervisors in the department on the quality of the predictions, providing valuable new data to further refine the model. CMPD hopes to bring the system live in the coming months. In addition, the Pittsburgh Bureau of Police will be involved in the expansion of this EIS in early 2017.

While working on a currently-controversial issue created some additional challenges for the teams, the fellows ultimately found it rewarding to engage with such a challenging and socially relevant problem.

“With recent events in the states, this is obviously a highly emotionally-charged topic,” wrote Lin Taylor, a fellow on the MNPD team. “It was challenging to maintain a balanced perspective for our work, while there was (and still continues to be) what felt like real social upheaval happening around us. At times this was difficult to navigate, but it was also an immensely valuable learning experience.”

---