

Go read this data analysis that uncovers predictive policing's flawed algorithm

A co-reported investigation with *Gizmodo* and *The Markup* into PredPol

By [Jasmine Hicks](#) Dec 6, 2021, 4:40pm EST

<https://www.theverge.com/2021/12/6/22814409/go-read-this-gizmodo-analysis-predpol-software-disproportionate-algorithm>

Illustration by Alex Castro / The Verge

Gizmodo [released a deep-dive look](#) into the data collection process behind [its co-reported investigation with *The Markup*](#) into PredPol, a software company specializing in predictive policing (hence the name, which it has since changed to Geolitica) through machine learning. PredPol's algorithm is supposed to make predictions based on existing crime reports. However, since crimes aren't equally reported everywhere, the readings it provides to law enforcement could simply copy the biases in reporting over each area. If police use this to decide where to patrol, they could end up over-policing areas that don't need a larger presence.

When *Gizmodo* and *The Markup* evaluated the areas, they found that the places PredPol's software targeted for increased patrols "were more likely to be home to Blacks, Latinos, and families that would qualify for the federal free and reduced lunch program."

23 OF THE 38 LAW ENFORCEMENT AGENCIES TRACKED ARE NO LONGER PREDPOL CUSTOMERS
Even as police tactics evolved to include crime and arrest data, there has been a historical disparity in how these tactics affect communities of color. As *Gizmodo* points out in its analysis, in New York in the 1990s, researchers at the time found that the methods reduced crime without simply displacing it to other areas. However, the approach included tactics like stop-and-frisk, which have been criticized as violations of civil rights.

PredPol's algorithm [has already been looked into](#) and criticized by academics more than once. As *Vice* quoted Suresh Venkatasubramanian, a member of the board of directors for ACLU Utah, in 2019:

"Because this data is collected as a by-product of police activity, predictions made on the basis of patterns learned from this data do not pertain to future instances of crime on the whole," Venkatasubramanian's study notes. "In this sense, predictive policing is aptly named: it is predicting future policing, not future crime."

Still, there hasn't been an investigation as thorough as this one. This investigation used figures retrieved from public data available via the web. According to *Gizmodo* and *The Markup*, they found an unsecured cloud database linked from the Los Angeles Police Department's website. That data contained millions of predictions stretching back several years.

Besides supposedly predicting individual crimes, a 2018 report by *The Verge* [looked into Pentagon-funded research by PredPol's founder Jeff Brantingham about using the software to predict gang-related crime](#). The former University of California-Los Angeles anthropology professor adapted earlier research in forecasting battlefield casualties in Iraq to create the platform, and the paper — "Partially Generative Neural Networks for Gang Crime Classification with Partial Information" — raised concern over the ethical implications.

Critics said this approach could do more harm than good. "You're making algorithms off a false narrative that's been created for people — the gang documentation thing is the state defining

people according to what they believe ... When you plug this into the computer, every crime is gonna be gang-related,” activist Aaron Harvey told *The Verge*.

Relying on some algorithms can work magic for some industries, but their impact can come at a real human cost. With bad data or the wrong parameters, things can go wrong quickly, even in circumstances that are less fraught than policing. Look no further than Zillow recently [having to shut down](#) its house-flipping operation after losing hundreds of millions of dollars despite the “[pricing models and automation](#)” it thought would provide an edge.

Overall, [Gizmodo and The Markup’s reporting](#) is a good consideration of how significantly predictive algorithms can affect the people unknowingly targeted by them. The [accompanying analysis](#) by *Gizmodo* provides relevant data insight while giving readers a behind-the-scenes look into these measures. The report indicates that 23 of the 38 law enforcement agencies tracked are no longer PredPol customers, despite initially signing up for it to help distribute crime-stopping resources. Perhaps by using methods that build transparency and trust on both sides, law enforcement could spend less time on tech that leads to pieces like this, which highlights the exact opposite approach.