00:04 / 02:01

f

y

in

✉

BY SEAN CAPTAIN
5 MINUTE READ

"No doubt the precogs have already seen this," says Chief John Anderton (played by Tom Cruise), head of Washington, D.C.'s experimental "Precrime" crime-prediction department in *Minority Report*, the 2002 Steven Spielberg movie based on Philip K. Dick's 1956 short story (which is also now a new Fox TV series).

Hitachi Says It Can Predict Crimes Before They Happen

Not quite Minority Report, but monitoring everything from weather to Twitter may be able to predict where and when crime will occur.

Of course, no one has found a trio of psychic mutant "precogs" who can unanimously foresee future crimes, but Hitachi today introduced a system that promises to predict where and when crime is likely to occur by ingesting a panoply of data, from historical crime statistics to public transit maps, from weather reports to social media chatter. Hitachi says that "about half a dozen" U.S. cities will join a proof of concept test of the technology beginning in October, and though Hitachi hasn't yet named them, Washington, D.C. could well be on the list. It's one of several dozen cities in the U.S. and Caribbean countries where the company already provides video surveillance and sensor systems to police departments with its Hitachi Visualization Suite. Hitachi execs provided several examples–even screenshots of the software–featuring D.C. in my conversations with them.
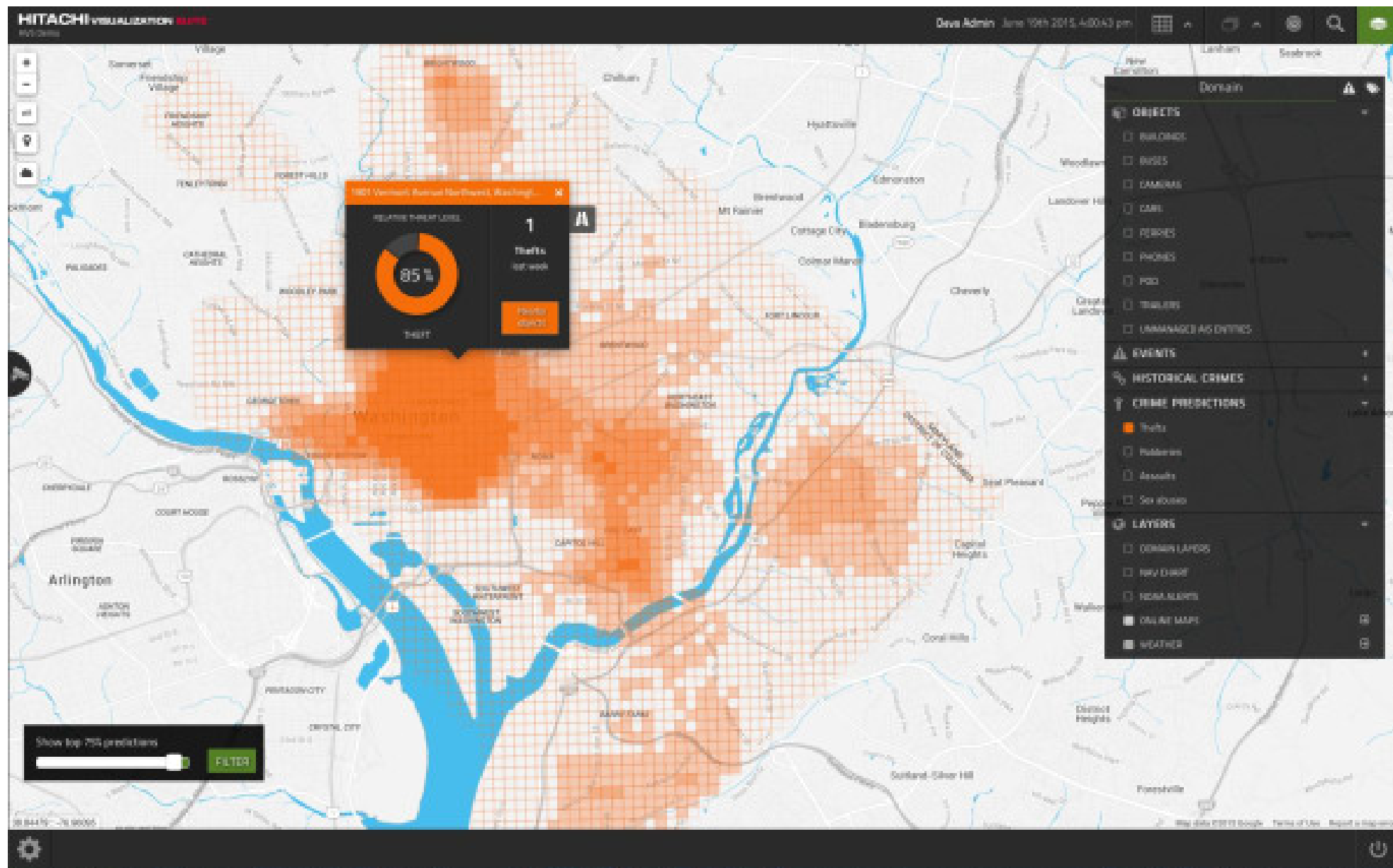
"We don't have any precogs as part of our system," says Darrin Lipscomb, cofounder of companies Avrio and Pantascene, which developed crime-monitoring tech that Hitachi later acquired. "If we determined that the precogs were actually somewhat accurate, we could certainly use their predictions to feed into our model," he says with perfect deadpan. What the new technology, called Hitachi Visualization Predictive Crime Analytics (PCA), does have is the ability to ingest streams of sensor and Internet data from a wide variety of sources.

It then applies what's called machine learning, using the popular statistical software known as "R," that crunches all this information in order to find patterns that humans would miss. "A human just can't handle when you get to the tens or hundreds of variables that could impact crime," says Lipscomb, "like weather, social media, proximity to schools, Metro [subway] stations, gunshot sensors, 911 calls."

LET THE DATA SPEAK FOR ITSELF

Machine learning is the hot new phase of artificial intelligence. Rather than trying to design a beautiful electronic mind, computer scientists are now building huge distributed computing systems that learn by sifting through fire hoses of data and ascertaining patterns or anomalies. This has become practical only recently with the development of big, cheap data storage and processing capabilities, like Amazon Web Services (AWS), Microsoft Azure, and Hitachi's own HDS cloud system–all of which Hitachi's PCA can run on.

Applying machine learning is a big switch from traditional police dispatching, say both Lipscomb and Mark Jules, his cofounder at Avrio and Panatscene. (Both are now execs in Hitachi's Public Safety and Visualization division.)
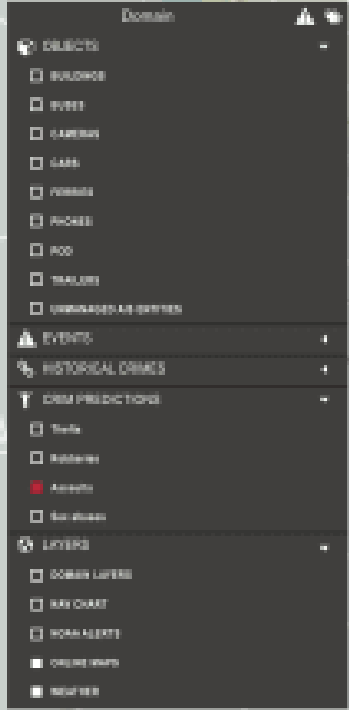
Hitachi isn't the only company to provide public safety monitoring and prediction services. Where it claims to be unique is in its use of machine learning, in allowing the data to drive the predictions rather than going in with any preconceived notions of what factors are important.

Traditionally, says Jules, police investigators build crime-prediction models based on their experience with certain variables, like the location of schools or slang words for drugs that pop up on Twitter. They assign a weight to each variable based on how important it seems to be. Hitachi's system, he says, doesn't require a human to figure out what variables matter and how much. "You just feed those data sets," says Jules. "And it decides, over a couple of weeks, is there a correlation."

Social media plays a big role in predicting crime, they say, improving accuracy by 15%. Hitachi uses natural language processing: the ability of a computer to ingest and understand colloquial text or speech.

Applying what's called a latent Dirichlet allocation, the system can sift through every tweet tagged to a specific geography to find significant words that indicate what's happening. "Gangs, for instance, use these different keywords to maybe meet up or perform some action," says Lipscomb. "I don't know what that keyword is…but with our approach we can actually pick out something that's abnormal, like someone's using an off-topic word, and using it in a very tight density or proximity, and that's going to get a bigger weight."

One thing social media indicates is tension between neighborhoods that could turn violent. "We were talking to [Washington] D.C., and they said, our biggest cause and effect is what neighborhood you're closest to," says Lipscomb. "There's these neighborhood rivalries going on in D.C." Normally, said his colleague Jules, police wouldn't realize the correlation between neighborhood tension flare-ups and crime until months later.

PCA provides a highly visual interface, with color-coded maps indicating the intensity of various crime indicators and even surprisingly cute icons for things like guns, cellphones, and surveillance cams. The system can pinpoint a location, down to a 200-meter square, and assign it a relative threat level from 0 to 100 percent. Jules calls this visual approach putting everything on a single pane of glass. That again brings up an image from *Minority Report*, with Chief Anderton standing in front of a massive screen displaying different data sources.

BIG BROTHER, OR ACTUALLY LESS DISCRIMINATION?

Anyone who has seen or read *Minority Report* knows (spoiler alert!) that things go awry. What if Hitachi's Visualization Predictive Crime Analytics makes mistakes and guesses wrong? No one is talking about preemptively arresting people, as in the story. But could this lead to a new kind of biased profiling of innocents as potential criminals?

Lipscomb claims it would be the opposite, at least better than New York City's controversial stop-and-frisk practice, in which police can search anyone in a targeted neighborhood. (Police aren't allowed to target people based on race, but 85% of those stopped have been Latino or African-American, according to the New York City Bar Association.) "We're trying to provide tools for public safety so that [law enforcement is] armed with more information on who's more likely to commit a crime," says Lipscomb. "I don't have to implement stop-and-frisk. I can use data and intelligence and software to really augment what police are doing." Lipscomb also says that Chicago has never used stop-and-frisk, perhaps hinting at another city that will implement Hitachi's new technology.

That still leaves open the question of accuracy: Will the technology really target the right places, where crime is likely to occur? Lipscomb acknowledges that he still has to prove the system will work. In the upcoming tech trials, some cities will be taking action based on the predictions, reallocating police to areas when the model predicts a higher likelihood of crime.

There will also be double-blind trials. Police departments will continue with business as usual, but the models will also be running in the background. Only after the test period will the police see what the model had predicted each day, so they can compare the predictions to what actually happened in the time frame. Hitachi has pledged to make all these results publicly available for scrutiny.

"We know that our approach is probably a little more innovative than some of the others, but we're not saying it's more accurate," Lipscomb says. "We want to prove it out with existing customers and then really go broad-based and say: Look, this works."

ABOUT THE AUTHOR

Sean Captain is a Bay Area technology, science, and policy journalist. Follow him on Twitter @seancaptain. More

---

You Might Also Like:

8 simple, inexpensive ways to be a more ethical consumer in 2019

The one dumb feature I wish Apple and Google would add to their phones