Spring 2018

Predictive Analytics in the Criminal Justice System: Media Depictions and Framing

Kar Mun Cheng
Rollins College, kcheng@rollins.edu

Predictive Analytics in the Criminal Justice System: Media Depictions and Framing

Kar Mun Cheng

Rollins College

Departments of Computer Science and Critical Media and Cultural Studies

1. Introduction

Artificial intelligence and algorithms are increasingly becoming commonplace in crime-fighting efforts. For instance, predictive policing uses software to predetermine criminals and areas where crime is most likely to happen. Risk assessment software are employed in sentence determination and other courtroom decisions, and they are also being applied towards prison overpopulation by assessing which inmates can be released. Public opinion on the use of predictive software is divided: many police and state officials support it, crediting it with lowering crime rates and improving public safety. Others, however, have questioned its effectiveness, citing civil liberties concerns as well as the possibility of perpetuating systemic discrimination.

According to the Prison Policy Initiative, over 2.3 million Americans were incarcerated in 2017 [1]. Of this population, 60 per cent were made up of people of color. African-American men are disproportionately targeted by the U.S. judicial system; they are more likely to be stopped and frisked by police, as well as receive stiffer sentences than white men for the same crimes [2].

In light of these facts, using algorithms and predictive methods to make decisions—especially ones that may affect the freedom of individuals—requires further study. Investigating the increasingly intertwined relationship between technology and human liberties can help develop a better understanding of how artificial intelligence can help make lives more efficient and the judicial system more transparent.

The news media plays a significant role in shaping opinions on controversial issues. Articles and reports on predictive policing not only inform the public, but they also influence how people perceive the use of artificial intelligence in law enforcement, and ultimately how we, as citizens, want to be policed.

This study evaluates the role of news media in shaping public opinion on two fronts: (a) the use of predictive analytics in the justice system, and (b) the integration of artificial intelligence in everyday life. Working with a corpus of articles from major journalistic outlets, we apply a qualitative methodology based on grounded theory to identify the key frames that govern media representation of predictive policing.

This study makes the following contributions:

- A survey of current predictive policing techniques, including hot spot analysis, regression methods, near-repeat, and spatiotemporal analysis
- Application of grounded theory methods to a qualitative analysis of a corpus of 51 online articles on the U.S. criminal justice system's use of predictive software and algorithms
- Identification of two frames most commonly adopted by elite journalists writing for national news outlets

Two dominant frames were identified from a corpus of 51 articles: *fear of the future* and *fear of the past*. The first frame elaborates on the potential consequences of implementing predictive algorithms in policing efforts, using specific examples to emphasize the difficulty of removing bias from software systems and the likelihood of perpetuating racial discrimination. The second frame argues that using data effectively can help combat rising crime rates, especially in metropolitan areas like Chicago and New York City. It bolsters its claim by attributing the ability of using predictive analytics to forecast crime as well as national threats before they happen — it focuses on preventing crime as opposed to combating it.

2. Literature Review and Background

Predictive policing algorithms take historic crime data as their input, identify trends of interest, and make predictions on likely outcomes. Police have used predictive software to determine, for example, subjects that may have been involved in gun violence as either the perpetrator or the victim [3]. Given a number of factors such as an individual's age, employment, and criminal histories of family members, an algorithm calculates a numerical score and classifies individuals as low, medium, or high risk of recommitting in the future. Similar models have also been used to determine the risk of a defendant in committing crimes in the future and the likelihood of recidivism [4].

Other uses of predictive methods include identifying areas where crime is likely to occur. Using police data and real-time surveillance information, the software distinguishes patterns in the dataset. Clusters of crime that previously took place in the area are marked on a map and the algorithm helps law enforcement officers make informed decisions on places to patrol. Given the overrepresentation of minorities in the criminal justice system, many news reports have voiced concern that the data for these systems, and hence their predictions, are inherently biased [5].

2.1 Classes of Predictive Techniques

Artificial neural networks provide models to solve problems through pattern recognition, forecasting, and classification. There are three forms of learning paradigms: supervised, unsupervised, and reinforcement learning. To goal of the supervised learning model is to learn a mapping of inputs to outputs; the decision function that determines the mapping is learned from a training set. Unsupervised learning models identify patterns of interest in the input data. Reinforcement learning involves learning a model by rewarding and punishing an agent through

environmental interactions. Techniques that have been applied to predictive policing typically fall under the supervised and unsupervised learning categories. The learning involved in building these models depends upon the space between connected neurons. In supervised learning models, it focuses on learning a mapping between inputs and outputs, whereas unsupervised learning learns using information associated with a group of neurons.

Supervised Learning

Supervised learning assumes that each training example has a known class value and builds a model that can accurately predict class labels for new data points. Classification methods assign labels to events in order to establish rules; they are used to make predictions on the type of crime that will occur in a given space. Supervised learning involves training the model on a dataset to identify patterns determining an observed class, such that it can be used to recognize the class in future observations.

As Figure 1 illustrates, the classifier takes in *n* inputs with varying weights. Given a known class label for the output, supervised learning involves understanding a mapping between the inputs and the output. This approach is commonly utilized in risk assessment tools. Given the profile of a defendant, the algorithm takes in descriptors such as their gender, education, and employment status as inputs. The developer determines the weight of each input and threshold of the activation function to calculate a risk score for the defendant. Evidently, supervised learning is highly subjective to the programmer's decisions and perspectives.

Prior to building a supervised learning model, the data must be separated into two sets: (a) the training set, which will be used to construct the model, and (b) the testing set, which will be used to evaluate how well the model generalizes to previously unseen data. The classifier is then

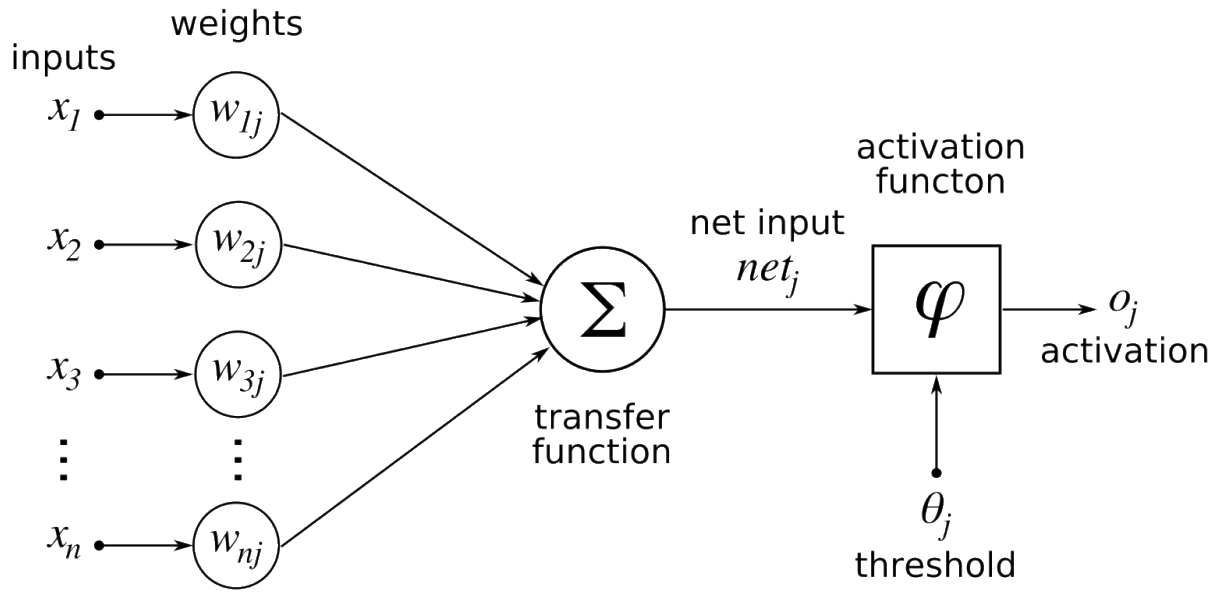

Figure 1: A single-layer neuron model. (Adapted from Wikimedia Commons)

repetitively re-trained and recalibrated based on accuracy figures obtained during validation. The process of building a supervised learning model, especially based on a limited training dataset, runs into the issue of *overfitting*, wherein the model becomes too specific to one particular dataset and does not generalize well to unseen data. In predictive modeling, "signal" refers to the pattern of interest desired to be learned, and "noise" refers to irrelevant information within the dataset. A model that is overfit leads to making predictions that are interfered by noise, leading to poor performance when applied to new data outside of its training data. Thus, a model is considered overfit when it is learning algorithmically from the noise as opposed to the signal.

Some common supervised learning algorithms include support vector machines (SVM), decision trees, and perceptrons. SVMs are classifiers used to linearly discriminate data; an optimal hyperplane that is determined from labeled training data is used as a separator in two dimensional space and categorizes new, unseen data. In contrast, the decision tree classifier uses repetition to divide the data into subsets of data by identifying vertical and horizontal lines.

```
initialize network weights to a small random value
do:
  calculate average error for all inputs
    if error < tolerance: exit
  for each input:
    calculate gradients for all weights, including bias
    weights, and average the gradients
      if length of gradient < small tolerance value: exit
      else: modify all weights by adding negative multiple of
```

Figure 2: Pseudocode for the backpropagation algorithm.

Currently popular in deep learning research is the multi-layer perceptron (MLP) model, which consists of three distinct features:

1. With enough neurons, the MLP is a universal approximator
2. Hidden layers that consist of neurons not in the inputs or outputs enable the network to solve complex problems, and so the more hidden layers the more representational power the MLP has
3. Changes hard thresholds to a softer differentiable activation function

Training a supervised neural network model is also referred to as the backpropagation algorithm outlined in Figure 2. When an error signal is found in the output, it is distributed back through the hidden layer(s). So, each training input t is compared to the corresponding output o_k. If the comparison is unfavorable, then all weights are updated to generate a more accurate mapping. Gradient descent, the strategy employed to update the weights of neurons, involves using derivatives to recursively pass from an unfavorable neuron to a new neuron where accuracy is improved. Eventually, a point of minimum error will be achieved.

Supervised learning is efficient when applied to classify linearly separable data, making it an appropriate algorithm to use in predictive policing to tackle problems such as classification (of criminals), forecasting (future crimes), and predictions.

Regression methods

Regressions involve fitting a mathematical relationship between independent variables and a different variable that the analyst desires to predict. A supervised learning method, regression models can be used to incorporate other factors including crime data to make future projections. Three common regression methods are:

- **Linear regression** fits an equation to model the linear relationship between two or more input variables and the output. Least-squares is most commonly used to calculate the best-fitting equation.
- **Nonlinear regression** methods fit complex mathematical equations between input and regression variables.
- **Regression splines** are used to model data of the same dependent variable over different regions with different regression methods.

Unsupervised learning

With no known class labels and vectors of observed features given, unsupervised learning algorithms are used to identify interesting structures or patterns in the input data without providing error signals. As suggested by the name, the lack of direction in how the algorithm learns provides certain advantages as it can backtrack to patterns that were not previously considered. The main characteristics of unsupervised learning include:

1. Clustering can be performed on both linearly and non-linearly patterned data
2. Neurons are arranged into rows and columns to form a single-layer feedforward structure

```
initialize k, the desired number of clusters

repeat:
  assign every point to the nearest centroid
  move centroids to the average of their assigned points
```

Figure 3: Pseudocode for the K-means algorithm.

3. Neurons that contain related pieces of information are grouped in close proximity to each other

A common application of unsupervised learning, especially within predictive policing methods, is clustering. Clustering methods divide the dataset and group data with similar attributes into clusters such that observations within clusters have maximum similarity to each other. Once trends have been identified models are built to make predictions by comparing the likelihood of a future situation being identical to previous clusters.

The K-means is one of the oldest and most popular clustering techniques. As Figure 3 delineates, the algorithm involves picking *k* points to be the cluster centroids, then recursively assigning every other data point to its nearest centroid and moving each centroid to the average of their assigned points.

Hot spot analysis

Hot spot analysis uses historical crime data to predict areas, called “hot spots”, where the risk of crime is higher. It assumes that crime typically occurs in the same pattern as the past. A research study found that applying hot spot analysis on crimes over short periods of time yields inaccurate predictions as crime rarely occurs in the same locations within temporal proximity to each other [6].

Methods used in hot spot analysis are related to clustering, an unsupervised learning algorithm. Geographic information systems (GIS) are employed to identify hot spots; spatial statistical software is used to generate sets of graduated circles that represent the cluster with the highest concentration of points on a map. However, overlapping ellipses on a map makes it difficult to identify concentration patterns of interest, and choosing the "best" and favorable result is highly subjective to the researcher.

In addition to covering ellipses, which serves as the simplest approach for identifying hot spots, other methods include kernel density estimation and heuristic methods. Kernel density estimation determines hot spots based on how close they are located to actual crime occurrences; predictions are made using crime data from nearby locations. A kernel function is used to smooth crime incident data to produce a contour, heat, or surface view map, such that the concentration of crime in an area is represented by its weight. It is important to note that making changes to the thresholds used to define hot spots in the function, regardless of how small it is, ultimately affects the results yielded. Much like the covering ellipses approach, kernel density estimation is highly subjective.

The effectiveness of heuristic methods is attributed to the familiarity of law enforcement officers with the environments they police. Hot spots of concentrated activity are manually identified based upon an officer's judgement. Their input is then used by data analysts in quadrant thematic mapping, which involves mapping crime data to a grid, where shaded cells represent "hot" areas.

Near-repeat

Future crimes are assumed to take place very close to the time and place that current crimes have occurred in. In other words, near-repeat methods operate under the belief that "crime spreads through local environments (micro-time and micro-place) much like a contagious disease" [7]. So, areas surrounding a crime event are predicted to experience higher levels of crime in the immediate future. Near-repeat methods were developed based on models used to predict earthquakes.

A simple heuristic utilized in this "self-exciting process" consists of grid mapping, estimating the "background rate" at which crime appears in each grid square, and an "aftershock rate", which refers to the assumption that new crimes temporarily rise at a given rate [8]. Near-repeat methods work best in policing efforts against burglaries.

Spatiotemporal analysis

Other than basing predictive methods on the crime itself, spatiotemporal analysis focuses on the relationship between the environment and the crime that occurs in it over time. Thus, it allows the incorporation of environmental and temporal factors regarding the crime location to algorithmically calculate a prediction for the location and time of future crimes.

Techniques used on large data sets usually consider frequency distribution and modes. Heuristic methods applied in spatiotemporal analysis generally include features such as:

- The time or day of the crime event
- How closely in time it occurs around other events, such as payday and holidays
- The season and weather during the occurrence
- The amount of time between each crime offense
- The number of times a specific type of crime occurs in the same location

- How the same type of incident progresses geographically in a crime series
- Spatial patterns in crime occurrences in an area
- The type of location the crime event takes place in and what is around it, such as bus stops, liquor stores, and shopping malls
- Socioeconomic data regarding the crime area being observed

2.2 Framing

Frames in media organize and establish one perception of reality with the intention of increasing its salience; they provide audiences with schemas on how to interpret reported events. According to Entman, the salience of an event involves "making a piece of information more noticeable, meaningful, or memorable to audiences" [9]. Scholars Wilson, Ballman, and Buczek have also found that the salience of media messages transfers to society [10].

News organizations act as the main source of information for the public. Consequently, how reporters chose to frame a topic or event will define what aspects audiences are exposed to. As Kuypers notes, frames are powerful because they "subtly induce us to filter our perceptions of the world in particular ways; they make some aspects of our reality more noticeable than other aspects" [11]. Despite journalists' pursuit for objective storytelling and reporting, dominant framing can still occur. Gamson argues that facts "take on their meaning by being embedded in a frame or storyline that organizes them and gives them coherence, selecting certain ones to emphasize while ignoring others" [12]. It brings into light the role of journalists and editors as media gatekeepers, and the implications this entails—with the power to decide how to *tell* a story, they are able to determine the version of social reality to establish.

3. Methods

3.1 Corpus

The study focuses on frames used by national news media outlets in their reporting on predictive policing. The corpus consists of 51 online articles taken from the BBC, Bloomberg, *The Economist*, *The New York Times*, *The Wall Street Journal*, and *The Washington Post* (refer to Appendix I). In order to control the effects of ideological influences in news reporting, sources were selected based on how much the public trusted the publication. According to a study conducted by Pew Research Center, *The Economist*, BBC, *The Wall Street Journal*, CNN, and Bloomberg are generally trusted by readers regardless of ideological backgrounds [13]. Although *The New York Times* and *The Washington Post* are distrusted by conservatives, they are included in this study as they have a widespread online presence.

The topics of the articles included informative pieces, opinion articles, and news reports on the use of predictive software and algorithms in crime-fighting efforts as well as within the judicial system. Additional class labels have been attached to each article to bolster analysis:

- 38 informative articles and 13 opinion pieces;
- 9 short (< 600 words), 34 medium (< 1600 words), and 8 long (1600 > words) stories.

These articles were subjected to a qualitative analysis based on the principles of grounded theory, with the goal of identifying the key salient frames that summarize media depictions of predictive policing. These frames give insight into how journalists present the issues associated with predictive policing, and hence, how citizens receive information that influences their opinions on the role of predictive policing in their communities.

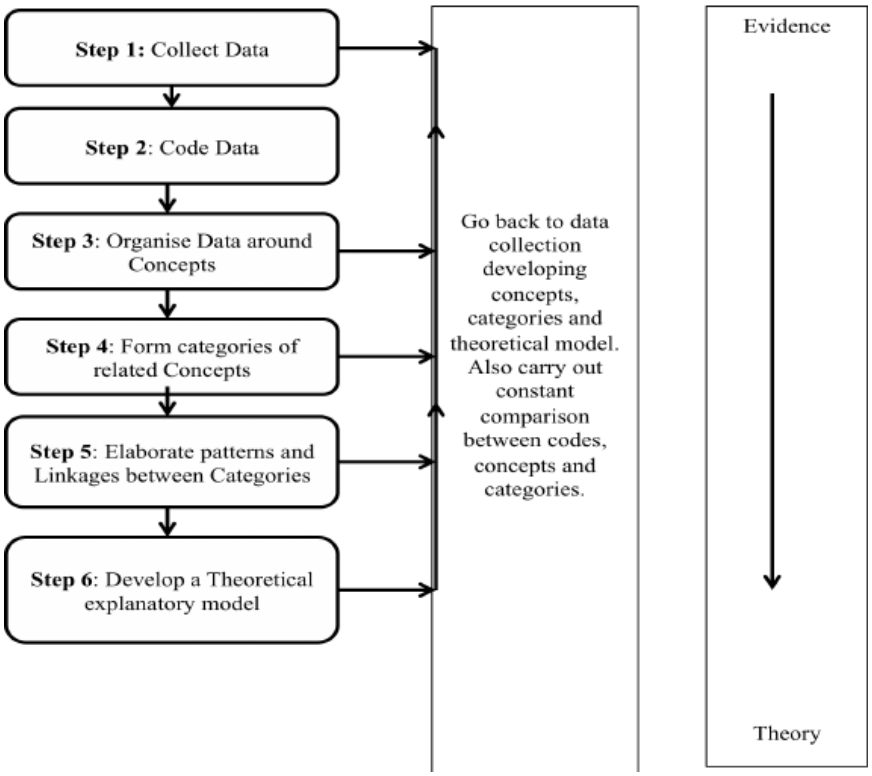

Figure 4: The grounded theory process. (Adapted from O'Hagan and O'Connor)

3.2 Grounded theory

Grounded theory methods consist of data collection and analysis to develop a theory. In a collaboration between Barney G. Glaser and Anselm L. Strauss, grounded theory methods developed out of a research study on death and hospital settings [14]. In particular, Glaser and Strauss observed how medical professionals and terminal patients acted when the latter were told about their deaths. The systematic methodological strategies involving *developing* theories from research grounded in data were then introduced in Glaser and Strauss's book, *The Discovery of Grounded Theory*, which was published in 1967.

The research process is not linear and consists of repetitively going through the data, as shown in Figure 4. The initial phase of a grounded theory "cycle" consists of memo-writing,

IT WAS Eddie Johnson's first big test. Memorial Day weekend usually marks the start of the most violent period of the year, as the crime rate rises along with the temperature (see chart).	Scenario: it was an officer's first big test as he is policing the Memorial Day weekend, the start of rising crime rates of the year.
Thousands of officers patrolled the city's parks, beaches and neighbourhoods, including Mr Johnson, the boss of Chicago's besieged police force since April, who worked a night shift.	Many officers are deployed to patrol areas in the city.
Fixed-wing aeroplanes circled the area's expressways, which have recently seen a spike in shootings.	Planes fashioned with cameras are used to monitor areas prone to shootings.
In the run-up to the weekend Mr Johnson launched one of the biggest anti-gang raids in Chicago's history, resulting in the arrest of 140 gang members and the seizure of numerous guns, as well as drugs apparently worth tens of thousands of dollars.	Many gang members were arrested and guns seized, becoming the biggest anti-gang raids in Chicago history.
Considering the steep rise in gun violence this year, the sheer size of Chicago's territory, the complexity of its social problems, the large number of fractious gangs with ever-younger members and the recent breakdown in trust between residents and the Chicago Police Department (CPD), Mr Johnson has taken on perhaps the toughest job in law enforcement in the country.	Many circumstances regarding gang, violence, and guns in Chicago makes law enforcement the toughest job in the United States.
The results of the Memorial Day operations were mixed: killings were down this year, with six murders, including one of a 15-year-old girl, between Friday morning and Tuesday morning, compared with 12 last year. Shootings were higher: 63 compared with 56.	Memorial Day operations resulted in reduced number of killings but a rise in shootings.
From the start of the year until mid-May, the number of murders increased by 62% to 216. Shootings also rose by 60%.	The number of murders and shootings rose prior to Memorial Day.
Many theories compete to explain why.	There are many theories to the rise in statistics.
One is the low morale of CPD officers, many of whom feel they are unfairly vilified and "are all being grouped with Jason," says a former cop, referring to Jason Van Dyke, a white police officer who shot a black teenager 16 times as he lay in the road in 2014.	Officers have experienced low morale as the community is regarding the police department with suspicion after cases of police brutality in the previous years.
A task-force subsequently appointed by the mayor to look at race and policing concluded in April that the CPD has "no regard for the sanctity of life" when it comes to black Chicagoans.	A task-force appointed by the mayor found that the police department did not care about taking the lives of black residents.
Mr Johnson reckons that the problem is lack of confidence in the justice system.	The community is not confident with the justice system.
He argues that trust has broken down—between the police and the policed, between the police and an "overburdened and broken" judiciary, as well as between the officers and their leaders.	The relations between police and the community has broken down. So has the trust between officers and their leaders.
In some ways, though, the CPD has suffered from an excess of trust, among officers at least.	Officers trust too much.
On May 31st the city paid out $2m to settle a lawsuit with two police officers who say they were targeted by colleagues and even suffered death threats after they informed on corrupt cops who ran a criminal fief in a housing complex on the South Side.	Corrupted cops targeted colleagues and sent death threats.
Trust will take time to rebuild, but Mr Johnson hopes that technology will pay dividends sooner.	It is hoped that technology will help rebuild the trust.

The CPD confiscates 150-200 guns per week on average, more than New York and Los Angeles combined.	150-200 guns are confiscated every week by the CPD.
(Though Chicago and Illinois have strict gun laws, it is easy to buy a gun in Indiana or Wisconsin, a short drive from the city.)	Strict gun laws do not make it difficult to purchase firearms from neighboring states.
The department wants to make more use of the CPD's "Strategic Subject List" (SSL), which is based on a computer algorithm developed by the Illinois Institute of Technology that calculates the propensity of individuals to get shot or shoot.	Strategic Subject List is created using an algorithm that calculates the likelihood of an individual being involved in a shooting, either as the shooter or the victim.
The fourth iteration of the SSL, the one now in use, has become really good in its murderous predictions, according to Mr Johnson.	The latest iteration of the software has helped in predicting murders.
The software looks at ten variables, including a person's previous arrests and convictions, gang affiliations and involvement in shootings.	Ten variables are evaluated by the software, which includes an individual's criminal history and gang affiliations.
People are ranked according to their probability of becoming a "party to violence" (PVE), either as victim or assailant.	They are then ranked based on their calculated score.
According to the CPD, a mere 1,400 people are responsible for most of the gun violence in a city of 2.7m.	Only 1,400 people are responsible for the majority of gun violence that happens in Chicago.
Of the 140 arrested in the recent gang raid, 117 were already on the SSL.	Most of those arrest in a recent gang raid were on the SSL.
Three-quarters of shooting victims and more than 80% of those arrested for shootings so far this year were also found to be on the list.	Many shooting victims and perpetrators were present on the list.
The SSL is a work in progress: the police department is constantly updating the list and fine-tuning its technology.	The SSL is still under development and the police department is constantly updating it with improvements.
It is also trying to use its data to prevent crime.	The tool may help prevent crime from happening.
The SSL score ranges from one to 500, with higher scores representing greater risk.	Scores on the list range from 1 to 500; higher scores indicate higher risk.
Since 2013 officers, social workers and community leaders have visited the homes of more than 1,300 people with high scores; this year the CPD aims to reach 1,500 people likely to be involved in violence and to meet gang members every other week.	Social workers and community leaders have been contacting people on the list with high scores.
Mr Johnson's next big test will be the weekend of the 4th of July. Last year ten people were killed and 55 shot while everyone else was celebrating Independence Day.	The next big test will be July 4th; people are likely to be killed or shot during the Independence Day weekend.

Figure 5: Line-by-line coding of 'Predictable Policing' by The Economist

diagramming concepts, and organizing the memos into thematic categories. The second phase involves conducting theoretical sampling to identify new data that follows the theme and furthering current memos with additional details and notes; this data collection process happens cyclically

IT WAS Eddie Johnson's first big test. Memorial Day weekend usually marks the start of the most violent period of the year, as the crime rate rises along with the temperature (see chart).	Scenario: it was an officer's first big test as he is policing the Memorial Day weekend, the start of rising crime rates of the year.	Narrative
Thousands of officers patrolled the city's parks, beaches and neighbourhoods, including Mr Johnson, the boss of Chicago's besieged police force since April, who worked a night shift.	Many officers are deployed to patrol areas in the city.	Policing
Fixed-wing aeroplanes circled the area's expressways, which have recently seen a spike in shootings.	Planes fashioned with cameras are used to monitor areas prone to shootings.	Surveillance
In the run-up to the weekend Mr Johnson launched one of the biggest anti-gang raids in Chicago's history, resulting in the arrest of 140 gang members and the seizure of numerous guns, as well as drugs apparently worth tens of thousands of dollars.	Many gang members were arrested and guns seized, becoming the biggest anti-gang raids in Chicago history.	Policing
Considering the steep rise in gun violence this year, the sheer size of Chicago's territory, the complexity of its social problems, the large number of fractious gangs with ever-younger members and the recent breakdown in trust between residents and the Chicago Police Department (CPD), Mr Johnson has taken on perhaps the toughest job in law enforcement in the country.	Many circumstances regarding gang, violence, and guns in Chicago makes law enforcement the toughest job in the United States.	Policing
The results of the Memorial Day operations were mixed: killings were down this year, with six murders, including one of a 15-year-old girl, between Friday morning and Tuesday morning, compared with 12 last year. Shootings were higher: 63 compared with 56.	Memorial Day operations resulted in reduced number of killings but a rise in shootings.	Crime
From the start of the year until mid-May, the number of murders increased by 62% to 216. Shootings also rose by 60%.	The number of murders and shootings rose prior to Memorial Day.	Crime
Many theories compete to explain why.	There are many theories to the rise in statistics.	Statistics
One is the low morale of CPD officers, many of whom feel they are unfairly vilified and "are all being grouped with Jason," says a former cop, referring to Jason Van Dyke, a white police officer who shot a black teenager 16 times as he lay in the road in 2014.	Officers have experienced low morale as the community is regarding the police department with suspicion after cases of police brutality in the previous years.	Police brutality
A task-force subsequently appointed by the mayor to look at race and policing concluded in April that the CPD has "no regard for the sanctity of life" when it comes to black Chicagoans.	A task-force appointed by the mayor found that the police department did not care about taking the lives of black residents.	Police brutality
Mr Johnson reckons that the problem is lack of confidence in the justice system.	The community is not confident with the justice system.	Relationships
He argues that trust has broken down—between the police and the policed, between the police and an "overburdened and broken" judiciary, as well as between the officers and their leaders.	The relations between police and the community has broken down. So has the trust between officers and their leaders.	Relationships
In some ways, though, the CPD has suffered from an excess of trust, among officers at least.	Officers trust too much.	Relationships
On May 31st the city paid out $2m to settle a lawsuit with two police officers who say they were targeted by colleagues and even suffered death threats after they informed on corrupt cops who ran a criminal fief in a housing complex on the South Side.	Corrupted cops targeted colleagues and sent death threats.	Corruption
Trust will take time to rebuild, but Mr Johnson hopes that technology will pay dividends sooner.	It is hoped that technology will help rebuild the trust.	Benefit
The CPD confiscates 150-200 guns per week on average, more than New York and Los Angeles combined.	150-200 guns are confiscated every week by the CPD.	Firearm
(Though Chicago and Illinois have strict gun laws, it is easy to buy a gun in Indiana or Wisconsin, a short drive from the city.)	Strict gun laws do not make it difficult to purchase firearms from neighboring states.	Legislation
The department wants to make more use of the CPD's "Strategic Subject List" (SSL), which is based on a computer algorithm developed by the Illinois Institute of Technology that calculates the propensity of individuals to get shot or shoot.	Strategic Subject List is created using an algorithm that calculates the likelihood of an individual being involved in a shooting, either as the shooter or the victim.	Risk assessment
The fourth iteration of the SSL, the one now in use, has become really good in its murderous predictions, according to Mr Johnson.	The latest iteration of the software has helped in predicting murders.	Predictive software
The software looks at ten variables, including a person's previous arrests and convictions, gang affiliations and involvement in shootings.	Ten variables are evaluated by the software, which includes an individual's criminal history and gang affiliations.	Variables
People are ranked according to their probability of becoming a "party to violence" (PVE), either as victim or assailant.	They are then ranked based on their calculated score.	Risk score
According to the CPD, a mere 1,400 people are responsible for most of the gun violence in a city of 2.7m.	Only 1,400 people are responsible for the majority of gun violence that happens in Chicago.	Criminals
Of the 140 arrested in the recent gang raid, 117 were already on the SSL.	Most of those arrest in a recent gang raid were on the SSL.	Results
Three-quarters of shooting victims and more than 80% of those arrested for shootings so far this year were also found to be on the list.	Many shooting victims and perpetrators were present on the list.	Results
The SSL is a work in progress: the police department is constantly updating the list and fine-tuning its technology.	The SSL is still under development and the police department is constantly updating it with improvements.	Development
It is also trying to use its data to prevent crime.	The tool may also be able to help prevent crime from happening.	Prevention
The SSL score ranges from one to 500, with higher scores representing greater risk.	Scores on the list range from 1 to 500; higher scores indicate higher risk.	Risk score
Since 2013 officers, social workers and community leaders have visited the homes of more than 1,300 people with high scores; this year the CPD aims to reach 1,500 people likely to be involved in violence and to meet gang members every other week.	Social workers and community leaders have been contacting people on the list with high scores.	Help
Mr Johnson's next big test will be the weekend of the 4th of July. Last year ten people were killed and 55 shot while everyone else was celebrating Independence Day.	The next big test will be July 4th; people are likely to be killed or shot during the Independence Day weekend.	Future

Figure 6: Focused coding of 'Predictable Policing' by The Economist

until the corpus has been exhausted and no new conclusions arise. Previous data and memos are then reexamined repetitively, allowing thematic ideas and a theory to organically develop.

3.3 Coding process

For the purpose of this study, the line-by-line coding strategy is adopted to conceptualize ideas in the initial phase. Coding allows the categorization of data segments using short names that

Fear of the future	Fear of the past
• Diamond narrative • 44 articles in total • 11 short articles • 26 medium articles • 7 long articles • 32 informative articles • 12 opinion pieces	• Inverted-triangle essay • 7 articles in total • 2 short articles • 4 medium articles • 1 long article • 6 informative articles • 1 opinion piece

Figure 7: Corpus breakdown between the two frames identified.

summarize and account for each piece of data. As seen in Figure 5, this involves making a memo for each line of every article in the corpus. This process is repeated three times for every article in the corpus. Following this, the focused coding strategy will be employed, which allows the separation, sorting, and synthesizing of the data collected (Figure 6). The coding process adopted in this study thus involves first summarizing each segment of data into a line, and finally into a single word.

4. Results

Overall, the 51 articles studied follow one of two frames: fear of what predictive policing could bring in the future, and fear of the past and how algorithms can help prevent history from repeating itself. Basic organizational structures have also been identified; the former frame tends to follow the *diamond narrative*, whereas the latter adheres to the *inverted-triangle essay*. They have also been coded with the type of reporting and its length (short, medium, long). Figure 7 depicts the corpus breakdown between the two dominant frames identified.

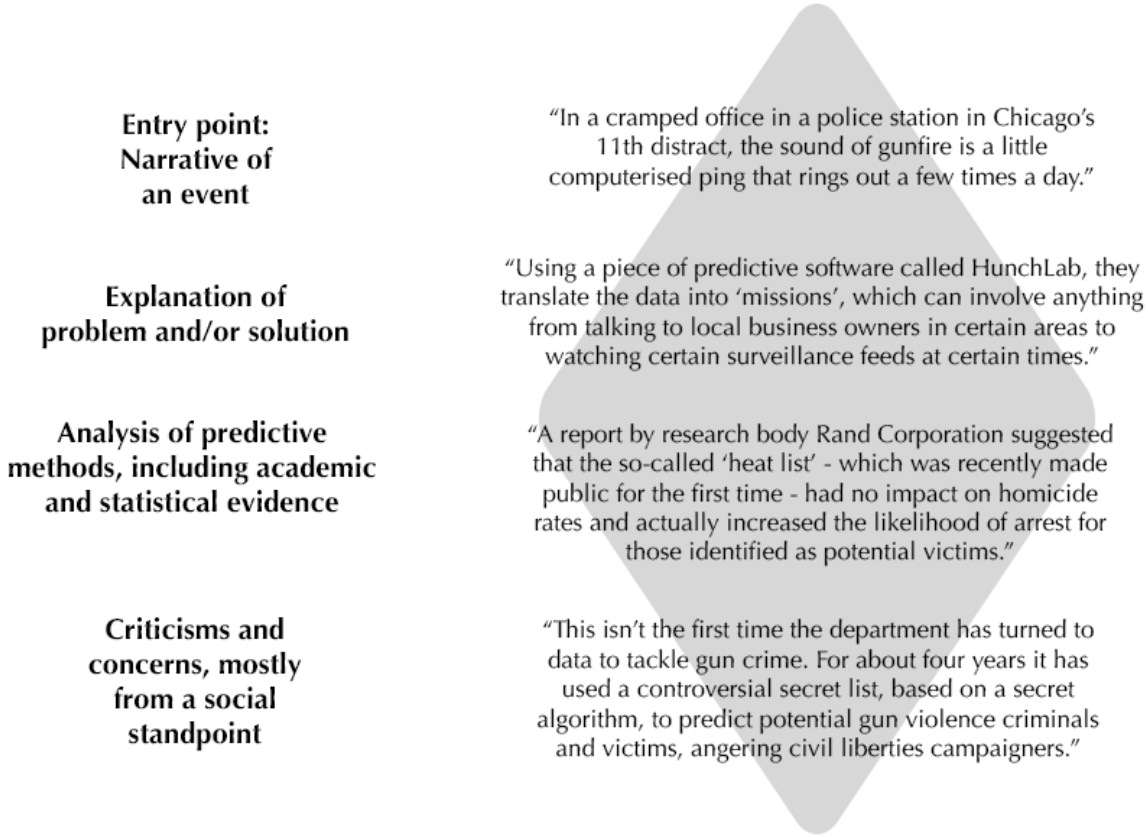

Figure 8: An example of the diamond narrative frame in the BBC article "Chicago goes high-tech in search of answers to gun crime surge" by Joel Gunter

4.1 Fear of the future and the diamond narrative

The fear of the future frame discusses the harm predictive policing pose to society over time. Authors tend argue that historic crime data are biased because they reflect social inequities of society. Resultantly, algorithms that take these statistics as inputs will yield inaccurate calculations. Authors bolster their claims and critique by quoting professionals from academia and civil liberties groups, as well as citing findings by research centers like ProPublica. A common approach of the frame involves comparing current events to Minority Report, a 2002 film depicting a future where police are able to predict and catch murderers before they commit the crime. In

essence, the argument is rooted in fear of perpetuating social inequities into the future with technology.

Figure 8 illustrates how the fear of the future frame organizes its story. Starting with a chronological description of an event, such as one outlining the interaction an individual has with the judicial system or the unfolding of a crime in city streets, the diamond narrative uses a real-life story as a hook. These articles then explain the problem as pointed out in the narrative. Diamond narratives then argue against the perceived effectiveness of predictive policing. Authors usually state that predictive methods have not been proven to lower crime rates or improve public safety. Finally, such narratives conclude with a specific concern raised by an identified authority figure. Resultantly, articles that adhere to the diamond narrative structure typically view algorithmic policing unfavorably and consider the issue from a societal perspective. The "diamond" structure of starting with a specific event, to a general discussion, and ending with specific criticism gives the framework its name.

4.2 Fear of the past and inverted-triangle essay

Given the rise in civic concern of mass shootings and terror attacks, the fear of the past frame states that predictive technologies can help prevent such disasters from happening. In other words, these articles argue for predictive analytics to be used to prevent crime as opposed to fighting crime. Authors use statistical results of falling crime rates and attribute them to predictive software; law enforcement administrators and state officials claim that it solves many issues including limited police resources and rising prison populations. Generally, the fear of the past frame views predictive policing methods favorably.

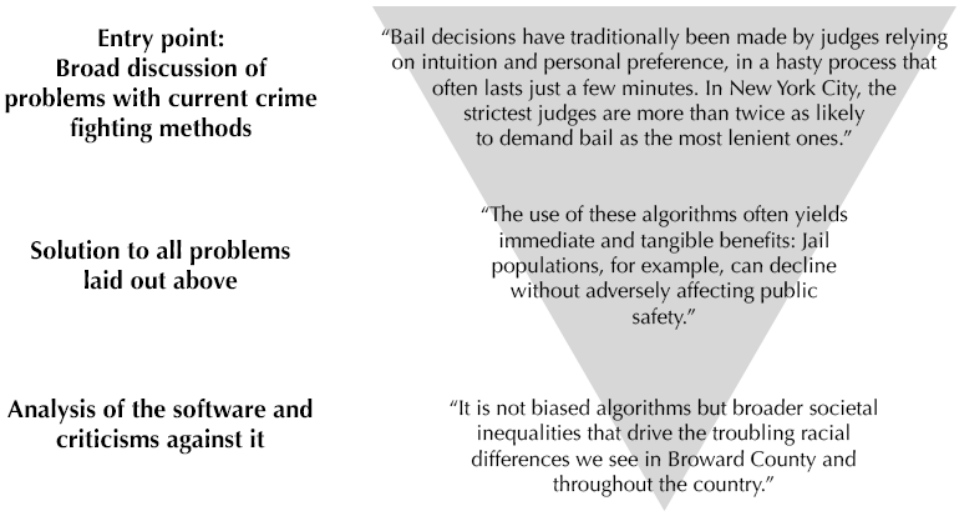

Figure 9: An example of the inverted-triangle essay frame in The New York Times article "Even Imperfect Algorithms Can Improve the Criminal Justice System" by Corbett-Davies, Goel, and González-Bailón

Figure 10: Word cloud generated from one-word summaries in focused coding.

The inverted-triangle essay frame takes a top-down approach when it comes to reporting. As Figure 9 shows, articles within this category build a broad discussion of problems with current

crime fighting strategies and propose a recommendation of predictive software as the up-and-coming solution. In some cases, authors may also provide an analysis of the use of predictive policing methods, particularly providing statistics of crime rates falling and proving that concerns regarding civil liberties are overstated.

4.3 Word frequency analysis

During the coding process, each sentence of every article in the corpus is summarized into a single word and inserted into a word cloud, as shown in figure 10. In total, the top ten words to occur are:

- Predictive (policing, model, software) (99 times)
- Discrimination (98 times)
- Risk (91 times)
- Data (83 times)
- Software (75 times)
- Concern (67 times)
- Algorithm (58 times)
- Crime (56 times)
- Secrecy (53 times)
- Bias (51 times)

5. Discussion

The recurrent occurrence of "discrimination" in the word frequency analysis suggests that it is a common topic for journalists to bring up when writing about predictive policing. Given the

heavy reliance on the accuracy and cleanliness of crime data being fed to algorithms—which are often proprietary software of for-profit companies—many authors have expressed concern over the reliability of calculated results. Moreover, it is also important to note that historic crime data are more than likely to contain social biases. Many concerns regarding discriminative policing is rooted in the fact that people from communities of color and poor backgrounds are more likely to interact with law enforcement and the judicial system as compared to white individuals [15, 16]. As a result, crime data that disproportionally represents people—particularly men—of color will cause such patterns to resurface in predictive algorithm since it is used as training data.

Another word of significance that shows up frequently in the corpus is "secrecy," which refers to the proprietary nature of algorithms. Many authors voiced their concerns over the lack of transparency in the use of predictive software, especially in making judicial decisions. Among the articles that fall under the fear of the future frame, one-fifth of them refer to the Loomis v. Wisconsin case when discussing the violation of civil rights in being judged by an algorithm without knowledge of how it makes its conclusion. This demonstrates that the fear of the future frame not only consists of concerns over biases in input data and the algorithm, but it also discusses the accountability of software developers in perpetuating discrimination in their code.

The fear of the future frame makes up the majority of the corpus, suggesting that it may be a popular perspective utilized by journalists. This may be attributed by the fact that writers of national media outlets are more comfortable with framing stories that critique law enforcement and technology. Speculatively, news journalists may be reluctant in adopting the fear of the past narrative because as it stands, there are more voices speaking against it than for it. As such, they may be more likely to err on the cautious side before publishing a grandiose claim that predictive analytics will evolve the way law enforcement police.

For the public, news media is the first point of contact on current events. However, when it comes to nuanced topics like artificial intelligence and algorithmic modeling, news media move from being an informant to becoming an educator. Considering that journalists generally do not possess the same level of expertise as someone in the computing profession, it increases the likelihood for miscommunication of information. The public thus receives second-hand information, consuming the reports by journalists after they have interpreted the findings regarding predictive analytics. Undoubtedly, this hierarchy of how information flows between computer scientists, journalists, and the public indicates that there is high likelihood for communication breakdown to occur.

Journalists write in layman terms in order to keep the language of reports accessible to the public, regardless of their education backgrounds. As a consequence, news writers may not be able to convey the full depth of complexity of artificial intelligence. Especially in articles under the fear of the past frame, the majority of authors use sweeping terms such as "artificial intelligence" and "algorithms." This creates a different type of communication barrier between programmers and society—generalizing an entire field of computer science not only makes it difficult to distinguish between concepts and terms, but it also increases the inaccuracy of these articles.

Framing artificial intelligence as a threat to humanity impedes society from reaping the full benefits of an automated world. According to Pew Research, the more Americans express concern over the concept of automating everyday life, with 72% fearing that in the future, robots and advanced computers will take jobs away from the economy as well as exacerbate existent economic inequalities [17]. The fear of the future frame depicts artificial intelligence as a terrifying unknown, citing events such as the case of a self-driving Uber running over a pedestrian in Arizona [18]. The frame not only instills fear into its audience, but it further drives society away from the

attainment of an automated life, and the potential benefits of efficiency it poses. Moreover, it may also prevent new innovations and technological progress, resulting in the stagnancy of a potentially promising computing field.

6. Conclusion

As the main source of information for the public, the news media has a considerable amount of influence over society. A salient frame that is consistently adopted in reports and articles will render that perspective the only reality, especially if the masses are overexposed to it. In our study, we put forth two frames that journalists use regarding predictive policing. The fear of the future frame focuses on using real-life events to establish problems with algorithms used in predictive methods, outlining the high likelihood that current social inequities will be reflected in analytics and perpetuated into the future by machines. On the other hand, the fear of the past frame discusses the benefits of using data and machines in criminal justice efforts, arguing that it serves as a method of crime prevention and not crime fighting. Moreover, the frame also accentuates the possibility of using predictive technologies to identify and address national threats or terror attacks before they occur. Results show that elite journalists at national media outlets are more likely to adopt the fear of the future frame as opposed to the the fear of the past frame, and this may be attributed to their reluctance of adopting a stance on artificial intelligence due to their inexperience. This brings to light a burgeoning problem that both the journalism and computer science industries will have to face—considering the rapid speed of technological progress occurring today, it becomes difficult for individuals outside of specialized fields (such as artificial intelligence) to comprehend these technologies. If a news agency fails to describe and report topics like artificial intelligence correctly and accurately, how would the common citizen get information? Moreover, a

programmer must be aware of the benefits and implications of technological implementations. In order to bridge this widening gap, computer scientists must be cognizant of the consequences technology may bring, as well as the importance of arming the public with accurate facts.

References

[1] P. Wagner and B. Rabuy, "Mass incarceration: The whole pie 2017." https://www.prisonpolicy.org/reports/pie2017.html, March 2017.

[2] M. Alexander, *The New Jim Crow* (1st. ed.). The New Press, 2012.

[3] M. Davey, "Chicago police try to predict who may shoot or be shot." https://www.nytimes.com/2016/05/24/us/armed-with-data-chicago-police-try-to-predict-who-may-shoot-or-be-shot.html, May 2016.

[4] M. Smith, "In Wisconsin, a backlash against using data to foretell defendants' futures." https://www.nytimes.com/2016/06/23/us/backlash-in-wisconsin-against-using-data-to-foretell-defendants-futures.html, June 2016.

[5] S. P. Gangadharan, "Predictive algorithms are not inherently unbiased." https://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/predictive-algorithms-are-not-inherently-unbiased, November 2015.

[6] L. W. Sherman, "Hot spots of crime and criminal careers of places," *Crime and Place* (4th. ed.). Criminal Justice Press, 1995.

[7] W. L. Perry, B. McInnis, C. C. Price, S. C. Smith, and J. S. Hollywood. "Making predictions about potential crimes," *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations.* RAND Corporation, 2013.

[8] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. "Self-exciting point process modeling of crime," *Journal of American Statistical Association,* vol. 106, pp. 100-108, 2011.

[9] R. M. Entman, "Framing: Toward clarification of a fractured paradigm," *Journal of Communication,* vol. 43, pp. 51-58, 1993.

[10] L. C. Wilson, A. D. Ballman, and T. J. Buczek. “News content about mass shootings and attitudes toward mental illness,” *Journalism & Mass Communication Quarterly,* vol. 99, pp.644-658, 2016.

[11] J. A. Kuypers and P. D’Angelo, *Doing News Framing Analysis: Empirical and Theoretical Perspectives*. Routledge, 2010.

[12] W. A. Gamson, “News as framing: Comments on Garber,” *American Behavioral Scientist,* vol. 33, pp. 157-161, 1989.

[13] Pew Research Center, *Political Polarization & Media Habits.* Pew Research, 2014.

[14] K. Charmaz, *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. SAGE, 2011.

[15] A. J. Ritchies and J. L. Mogul, “In the shadows of the war on terror: Persistent police brutality and abuse of people of color in the United States,” *DePaul Journal for Social Justice,* vol. 1, pp. 175-250, 2008.

[16] C. Solis, E. L. Portillos, and R. K. Brunson, “Latino youths’ experiences with and perceptions of involuntary police encounters,” *The ANNALS of the American Academy,* vol. 623, pp. 39-51, 2009.

[17] A. Smith and M. Anderson, “Automation of everyday life.” http://www.pewinternet.org/2017/10/04/automation-in-everyday-life/, October 2017.

[18] D. Wakabayashi, “Self-driving Uber car kills pedestrian in Arizona, where robots roam.” https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html, March 2018.

Appendix I: Corpus of 51 articles

BBC

- Can technology solve Chicago's gun crime problem? (Jane Wakefield)
- Chicago goes high-tech in search of answers to gun crime surge (Joel Gunter)
- Convict-spotting algorithm criticized (BBC)
- Crime fighting with big data weapons (Mark Ward)
- Gun crime tech 'failed to save lives' in Chicago (BBC)
- How maths can get you locked up (Simon Maybin)
- London police trial gang violence 'predicting' software (Leo Kelion)
- Police surveillance: The US city that beat Big Brother (Brian Wheeler)
- Police warned about using algorithms to decide who's locked up (BBC)
- Predicting crime - a step towards a safer world? (Katia Moskvitch)
- Predicting crime 'cut Medway street violence by 6%' (BBC)
- West Yorkshire Police predict crime spots using data (BBC)

Bloomberg

- How Big Data Could Help Identify the Next Felon -- Or Blame the Wrong Guy (Jordan Robertson)
- Serial Killers Should Fear This Algorithm (Robert Kolker)
- The Future of Policing Is Being Hashed Out in Secret (Noah Feldman)

The Economist

- Are programs better than people at predicting reoffending? (The Economist)
- Don’t even think about it (The Economist)
- Predictable Policing (The Economist)

- The power of learning (The Economist)

The New York Times

- ‘Intelligent’ Policing and My Innocent Children (Bärí A. Williams)
- Artificial Intelligence’s White Guy Problem (Kate Crawford)
- Be Cautious About Data-Driven Policing (Faiza Patel)
- Can Software Predict Crime? Maybe So, but No Better Than a Human (Niraj Chokshi)
- Chicago Police Try to Predict Who May Shoot or Be Shot (Monica Davey)
- Crime, Bias and Statistics (Charles M. Blow)
- Even Imperfect Algorithms Can Improve the Criminal Justice System (Sam Corbett-Davies, Sharad Goel and Sandra González-Bailón)
- In Hot Pursuit of Numbers to Ward Off Crime (Somini Sengupta)
- In Wisconsin, a Backlash Against Using Data to Foretell Defendants’ Futures (Mitch Smith)
- Inside the Algorithm That Tries to Predict Gun Violence in Chicago (Jeff Asher & Rob Arthur)
- Police Program Aims to Pinpoint Those Most Likely to Commit Crimes (John Eligon & Timothy Williams)
- Police Technology Shouldn’t Replace Community Resources (Kami Chavis Simmons)
- Predictive Algorithms Are Not Inherently Unbiased (Seeta Peña Gangadharan)
- Sending the Police Before There’s a Crime (Erica Goode)
- Sent to Prison by a Software Program’s Secret Algorithms (Adam Liptak)
- Sentencing by the Numbers (Emily Bazelon)
- Sentencing, by the Numbers (Sonja B. Starr)

- The Ivory Tower Can't Keep Ignoring Tech (Cathy O'Neil)
- The Risk to Civil Liberties of Fighting Crime With Big Data (Quentin Hardy)
- When a Computer Program Keeps You in Jail (Rebecca Wexler)
- When an Algorithm Helps Send You to Prison (Ellora Thadaney Israni)
- When it Comes to Policing, Data Is Not Benign (Anderson B. Francois)

The Wall Street Journal

- Chicago Designing Predictive Software Platform to Identify Crime Patterns (Joel Schectman)
- Chicago Police Take a Page From 'Minority Report' (Shibani Mahtani)
- Court: Judges Can Consider Predictive Algorithms in Sentencing (Joe Palazzolo)
- State Parole Boards Use Software to Decide Which Inmates to Release (Joseph Walker)
- A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. (Avi Feller, Emma Pierson, Sam Corbett-Davies, & Sharad Goel)
- Big data may be reinforcing racial bias in the criminal justice system (Laurel Eckhouse)
- Police are using software to predict crime. Is it a 'holy grail' or biased against minorities? (Justin Jouvenal)
- Relax, the futuristic pre-crime system of 'Minority Report' is still a long way from becoming reality (Dominic Basulto)
- The machines that could rid courtrooms of racism (Max Ehrenfreund)
- The new way police are surveilling you: Calculating your threat 'score' (Justin Jouvenal)