

AI in the Criminal Justice System

epic epic.org/issues/ai/ai-in-the-criminal-justice-system



Background

Automated decision-making tools are used widely and opaquely both directly in the criminal justice system and in ways that directly feed the criminal justice cycle in the U.S.

The flowchart below is a necessarily incomplete representation of these tools. One of the largest reasons is a lack of transparency by design. Affected people are often unable to know what tools the jurisdictions they live in use because of trade secret carveouts in Open Government laws, as well as similar roadblocks in evidentiary and discovery rules. The explanations, documents, and chart showing what tools are used where below are a representation of the patchwork of jurisdictions that are occasionally forthright about the tools they use, the outputs of Open Government requests, and news items often exposing a problematic use of one of the tools.

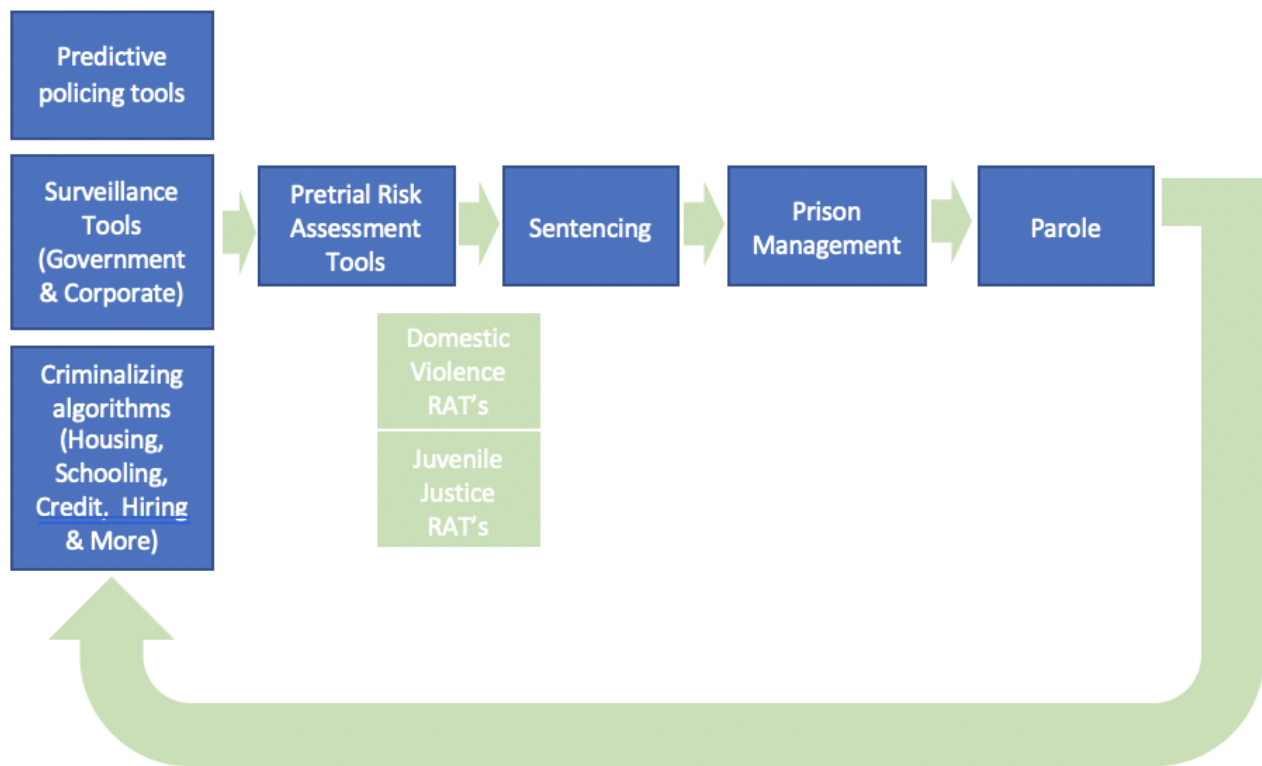


Figure 1: A rough cycle of the different algorithms/tech used in the criminal justice cycle

Figure 1: A rough cycle of the different algorithms used in the criminal justice cycle

THE TOOLS: PREDICTIVE POLICING

Predictive Policing Tools include “any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention,” according to the National Institute of Justice. Predictive policing comes in two main forms: location-based and person-based. Location-based predictive policing works by identifying places of repeated property crime and trying to predict where they would occur next, while person-based predictive policing aims to pinpoint who might be committing a crime – trying to measure the risk that a given individual will commit crimes. Both are used in different jurisdictions, and use past policing data as the main driver of these predictions, necessarily creating a self-fulfilling prophecy of arresting resources. The Bureau of Justice Assistance has and continues to give grants to Police departments around the country to create and pilot these programs. However, two high profile systems in Chicago and Los Angeles have been shut down due to limited effectiveness and significant demonstrated bias.

THE TOOLS: SURVEILLANCE

Surveillance tools encompass a large swath of technologies and functions that can be used to watch, track, and store information about a person. This ranges from Ring doorbells, whose servicer has direct partnerships with law enforcement, to facial recognition systems at the border and in U.S. cities. Learn more about this topic and EPIC’s work here.

THE TOOLS: CRIMINALIZING ALGORITHMS

Criminalizing algorithms include algorithms used in housing, credit determinations, healthcare, hiring, schooling, and more. Many of these have been shown to make recommendations and decisions that negatively affect marginalized communities, encoding systemic racism, and contribute to entry of the Criminal Justice system. All of the other tools discussed here are affected by the results and data points produced by these criminalizing algorithms.

THE TOOLS: RISK ASSESSMENTS

Risk Assessment Tools are used in almost every state in the U.S. – and many use them pre-trial, although they exist at sentencing, in prison management, and for parole determinations. There are also specific risk assessment tools for different functions in the criminal justice system, such as domestic violence risk or juvenile justice risks, with the understanding that different factors are used in those contexts than in a general criminal risk or violent criminal risk of rearrest or re-offense.

Pretrial Risk Assessment tools are designed to attempt to predict future behavior by defendants and incarcerated persons, and quantify that risk. They use socioeconomic status, family background, neighborhood crime, employment status, and other factors to reach a supposed prediction of an individual's criminal risk, either on a scale from "low" to "high" or with specific percentages. Significant empirical research has shown disparate impacts of Risk Assessment Tools on criminal justice outcomes based on the race, ethnicity, and age of the accused. The concerns with the use of these tools don't stop there. The tools vary but estimate using "actuarial assessments" (1) the likelihood that the defendant will re-offend before trial ("recidivism risk") and (2) the likelihood the defendant will fail to appear at trial ("FTA"). These often proprietary techniques are used to set bail, determine sentences, and even contribute to determinations about guilt or innocence. Yet the inner workings of these tools are largely hidden from public view. As a result, two people accused of the same crime may receive sharply different bail or sentencing outcomes based on inputs that are beyond their control—but have no way of assessing or challenging the results.

As criminal justice algorithms have come into greater use at the federal and state levels, they have also come under greater scrutiny. Many criminal justice experts have denounced "risk assessment" tools as opaque, unreliable, and unconstitutional.

A 2016 investigation by ProPublica tested the COMPAS system adopted by the state of Florida using the same benchmark as COMPAS: a likelihood of re-offending in two years. ProPublica found that the formula was particularly likely to flag black defendants as future criminals, labeling them as such at almost twice the rate as white defendants. In addition, white defendants were labeled as low risk more often than black defendants. But the investigators also found that the scores were unreliable in forecasting violent crime: only 20

percent of the people predicted to commit violent crimes actually went on to do so. When considering a full range of crimes, including misdemeanors, the correlation was found to be higher but not exceedingly accurate. Sixty-one percent of the candidates deemed likely to reoffend were arrested for any subsequent crimes within two years. According to ProPublica, some miscalculations of risk stemmed from inaccurate inputs (for example, failing to include one's prison record from another state), while other results were attributed to the way factors are weighed (for example, someone who has molested a child may be categorized as low risk because he has a job, while someone who was convicted of public intoxication would be considered high risk because he is homeless).

COMPAS is one of the most widely used algorithms in the country. Northpointe published a validation study of the system in 2009, but it did not include an assessment of predictive accuracy by ethnicity. It referenced a study that had evaluated COMPAS' accuracy by ethnicity, which reported weaker accuracy for African-American men, but claimed the small sample size rendered it unreliable. Northpointe has not shared how its calculations are made but has stated that the basis of its future crime formula includes factors such as education levels and whether a defendant has a job. Many jurisdictions have adopted COMPAS, and other "risk assessment" methods generally, without first testing their validity.

Defense advocates are calling for more transparent methods because they are unable to challenge the validity of the results at sentencing hearings. Professor Danielle Citron argues that because the public has no opportunity to identify problems with troubled systems, it cannot present those complaints to government officials. In turn, government actors are unable to influence policy

Over the last several years, prominent groups such as Pretrial Justice Institute (PJI) strongly advocated for the introduction of these tools and the Public Safety Assessment among many other risk assessments has been adopted in nearly every state, up from only a handful in the beginning of the decade. However, in February 2020, PJI reversed this position, specifically stating that they "now see that pretrial risk assessment tools, designed to predict an individual's appearance in court without a new arrest, can no longer be a part of our solution for building equitable pretrial justice systems." One week later, Public Safety Assessment, a widely used risk assessment developed by the Laura and John Arnold Foundation, released a statement in which they clarify that "implementing an assessment cannot and will not result in the pretrial justice goals we seek to achieve."

Transparency is seldom required with pre-trial risk assessments. One of the primary criticisms of these risk assessment tools is that they are proprietary tools, developed by technology companies that refuse to disclose the inner workings of the "black box." Trade secret and other IP protection defenses have been given to demands of the underlying logic of the systems. In March 2019, Idaho became the first state to enact a law specifically promoting transparency, accountability, and explainability in pre-trial risk assessment tools. Pre-trial risk assessments are algorithms that help inform sentencing and bail decisions for

defendants. The law prevents a trade secrecy or IP defense, requires public availability of ‘all documents, data, records, and information used by the builder to build or validate the pretrial risk assessment tool,’ and empowers defendants to review all calculations and data that went into their risk score.

For a deeper dive into Pre-Trial Risk Assessments, including a state-by-state survey of Risk Assessment tools used in U.S. states up to September 2020, visit EPIC’s report: [“Liberty At Risk, Pre-trial Risk Assessment Tools in the U.S.”](#)

Ben Winters

EPIC Senior Counsel