**Humanistic Values to Consider in Criminal Justice**

Numerous risk assessment tools have been trained and used in courts to intelligently predict the risk of recidivism, failure to appear in court for a charged criminal, and assisting in pretrial detention decisions. COMPAS is one such system.

The criminal justice system in the U.S. is overwhelmed for two reasons: a growing population and an increasingly complex law, which, while keeping criminal elements out of society, tends to exhaust itself of resources. The advent of AI powered tools like COMPAS offer a solution. In an overwhelmed system, incorporating every available efficiency is arguably a sensible approach. Examples such as predictive policing, facial recognition and license plate reading are other aspects of the justice system that are seeing increased use of AI (“AI and Predictive Policing: Transforming Criminal Justice in 2024 | Richard P. Davies, Esq.,” 2024; *The Implications of AI for Criminal Justice - Council on Criminal Justice*, 2024).  
 The question is if AI tools are the means to an end goal of efficiency. Are autonomous vehicles and other similar technologies better than (or for) humans? Justice systems worldwide pride themselves on ‘values’ and ‘humanistic’ elements. The dissenting judge’s minority opinion in major cases is the stuff of legacy. The role of the judge is to make unbiased decisions after factual evidence and human argument. Judges sit on benches for specific reasons: no judge can decide and preside over a court the next day they pass their bar exam. They practice as attorneys in courts before going through rigorous systems to be appointed judges. They move up these systems based on knowledge and human experience (Sanger, R., 2024). It is important to concede here that judges have an inherent bias as humans. In our analysis, judge decisions across counties in the given datasets are similar. While they treat males and females equally, they treat black populations more strictly on bail decisions. Across all counties, judges decide to deny bail to any black individual having a risk score greater than 4, while that threshold is 2 points higher for other races. It is also important to note that AI doesn’t fare any better: AI algorithms have marked black individuals with risk scores higher than other individuals in the counties by around 1 point. Both human and algorithm exhibit bias.

| ***Race*** | **Black** | **Other** | **White** |
| --- | --- | --- | --- |
| **Average Risk Score** | 5.815 | 4.625 | 4.881 |
| Warren County Data Analysis | | | |

| *Judge Decision* | Black | Other | White |
| --- | --- | --- | --- |
| Denied | 6.644 | 7.556 | 7.583 |
| Granted | 3.239 | 3.774 | 4.233 |

One argument AI enthusiasts present is informational advantage. A machine can process, store, and refer to a larger knowledge base than any judge can. Can that machine also decide in binary terms about a pretrial detention the same way a judge can? Yes–binary decisions. However, they cannot process the multitude of values, ethics and emotions as a human, and neither can they understand the context or consequence of their decision like a human (Winters, n.d.).

**Ethical and Value-Based Implications of AI in the Criminal Justice System**

For the non-immediate stakeholders, AI tools used in governance systems present several ethical implications where discrimination is exacerbated, and isn’t entirely traceable. This breaks down any negligible trust those on the margins may still have.

***Racial Bias, Discrimination***

When trained on biased data, AI will present biased outcomes. Historical crime data and legacy policing practices carry over the bias into the systems that are increasingly being empowered to make or break cases (Srinivas, 2023 | NAACP, 2024 | Bernstein, 2024 | Olsson, 2024).

In these cases, those who are marginalized are negatively impacted–the legacy of social prejudice affects their outcomes by an algorithm as judge. Although human judges will hold biases, their experience, context of socio-cultural norms, philosophy, and knowledge of ethics allows the most stubborn judges to have in them a humanity. No machine (today) can possibly evolve as a judge has over their study of jurisprudence.

In the examples of the counties, the data used to train the AI is biased. First, the algorithm has marked black individuals with an average higher risk score across all counties. Second, the threshold for bail denial is 4 for black people, whereas for white people it is 6. This double jeopardy strengthens existing marginalization of black people.

There is also the case of high false positive and false negative rates. False negative rates across all counties were much higher for black individuals, meaning that the AI tool incorrectly predicted an individual’s risk when they did not actually re-offend. The high false positive rates for white individuals implies they are often incorrectly granted bail when they will actually re-offend. These biases pose serious concerns to society about AI adoption. The question of efficiency seems insignificant in face of this observation from the counties because these tools deepen a divide the society is largely trying to bridge. The data below is from Warren county, which had the most even split between the black and white populations.

| **Statistical parity** | | | **Equal Opportunity** | | |
| --- | --- | --- | --- | --- | --- |
| **Black** | **White** | FALSE | **Black** | **White** | FALSE |
| **Black** | **Other** | FALSE | **Black** | **Other** | FALSE |
| **White** | **Other** | FALSE | **White** | **Other** | FALSE |

|  | **FNR** | **FPR** |
| --- | --- | --- |
| **Black** | 72.62% | 17.27% |
| **White** | 15.19% | 69.40% |
| Warren County Data Analysis | | |

***Data Transparency and Accountability***

Private corporations argue for proprietary needs considering the advanced technologies they develop. This is a fair argument until the point where accountability is not voluntarily accepted by these corporations for the same technologies. Additionally, it is simultaneously unlikely that judges be held wholly accountable for decisions in the criminal system made using inputs from AI tools (*AI, Data Analysis and Algorithms | Campaign*, n.d. | NAACP, 2024b | Bernstein, 2024 | Szilagyi, 2024).

When decisions are then made but no one stakeholder fully is held legally accountable, little can be done. The AI tools will output decisions based on prevailing information, and these decisions invoke recursion in the system, thereby reinforcing prevalent conditions. With respect to the dataset that we analyzed, no knowledge of the AI model used to train the system which gave these assessment scores. The technical barrier is a significant hurdle in understanding the number, but worse is the lack of context on what calculates the number. Increasing transparency thus allows each participant of the system–judges individuals–to have these questions answered.

***Erosion of Public Trust***

Rarely do people naturally trust law enforcement and justice systems. Many regions are culturally trained to beware of these systems to the effect of holding negative perceptions of the police, courts, and lawyers are normal amongst some marginalized communities who have historically been at the receiving end of related hostilities. Adding this context to the reinforcement of prevalent norms inherent within black box AI tools, the trust that exists will depreciate further in these communities (Bernstein, 2024 | *The Ethics of AI Decision-Making in the Criminal Justice System* |NAACP, 2024).

**AI is Still One Step Forward**

All of this, is not to say that using the efficiencies AI tools cannot ever benefit criminal justice systems. We would argue that marginalized stakeholders of the system can benefit the most when systems are implemented with certain caveats.

Natural Language Processing for example, is great technology that can be used to process information, reports, documents, and other police summaries much faster. Humans are creative beings and when they spend less time on the mundane, a space opens for them to pause and reflect on public questions they are trying to answer. Digital tools that standardize police reports when information is inputted, check for errors and provide dialogue which delivers informative summaries from inputted data can be really useful as it allows faster movement of cases throughout the system.

Likewise, other processes of the system, which are already regulated like chain of evidence maintenance, performance reviews of staff, etc., can incorporate digital intelligent systems that go through those tasks with rule-based precision. Humans do that too, when the guidelines exist–when decisions need to be binary. Machines, comparatively, can do that much faster and at great scale.

All of these steps and other similar ones, where processes require binary, probabilistically-solved decisions should be first places of consideration when deploying intelligent systems. They are administrative aspects - where human lives cannot be deeply affected due to machine-made errors, unlike bail decisions, which do have the potential to break or remake a potential criminal. These systems, if implemented in such administrative places wi

ll, in fact, best benefit the marginalized, considering that they are also the most affected when administrative systems are broken or are slow.

Technology, when incorporated into governance systems of any kind should ideally be benchmarked to transparency, accountability, and accessibility indices. The systems which score exceptionally on all fronts should be implemented before others and based on the experience of those deployments, should aid in eliminating weak and less reliable systems used in society.

**References**

AI and Predictive Policing: Transforming Criminal Justice in 2024 | Richard P. Davies, Esq. (2024). *Richardpdavieslaw.com*. https://doi.org/10907492753/A42LCNOYqMIDEJHDjNEo

*AI, data analysis and algorithms | Campaign*. (n.d.). Fair Trials. Retrieved November 10, 2024, from https://www.fairtrials.org/campaigns/ai-algorithms-data/

Bernstein, G. (2024, September 27). *The Ethics of AI Decision-Making in the Criminal Justice System*. CloudTweaks. https://cloudtweaks.com/2024/09/ethics-ai-criminal-justice-system/

NAACP. (2024, February 15). *Artificial Intelligence in Predictive Policing Issue Brief | NAACP*. Naacp.org. https://naacp.org/resources/artificial-intelligence-predictive-policing-issue-brief

Olsson, A. (2024, February 9). *Racial bias in AI: unpacking the consequences in criminal justice systems*. IRIS Sustainable Dev; IRIS Sustainable Dev. https://www.irissd.org/post/racial-bias-in-ai-unpacking-the-consequences-in-criminal-justice-systems

Sanger, R. (2024, January 12). *Artificial Intelligence and Criminal Law*. The Colleges of Law. https://www.collegesoflaw.edu/blog/2024/01/12/artificial-intelligence-and-criminal-law/

Srinivas, N. (2023, February 24). *The ethical debate of AI in criminal justice: Balancing efficiency and human rights*. ManageEngine Insights. https://insights.manageengine.com/artificial-intelligence/the-ethical-debate-of-ai-in-criminal-justice-balancing-efficiency-and-human-rights/

Szilagyi, K. (2024, October 9). *Maintaining the Rule of Law in the Age of AI*. Just Security. https://www.justsecurity.org/103777/maintaining-the-rule-of-law-in-the-age-of-ai/

*The Implications of AI for Criminal Justice - Council on Criminal Justice*. (2024, October 7). My WordPress. https://counciloncj.org/the-implications-of-ai-for-criminal-justice/

Winters, B. (n.d.). *AI in the Criminal Justice System – EPIC – Electronic Privacy Information Center*. Epic.org. Retrieved November 10, 2024, from https://epic.org/issues/ai/ai-in-the-criminal-justice-system/