RESEARCH NOTES[1]

THE FUTURE OF AI IN CRIMINAL JUSTICE

DAVID E.G. CONWAY

A recent edition of the *New Statesman* magazine led with a front-page questioning whether Artificial intelligence was an existential threat or humanity's salvation.[2] The substantive article interviewed three "Godfathers of AI," who whilst they disagreed as to whether AI would surpass humanity in consciousness or intelligence, agreed that this technology would have radical impacts on humanity's reliance on AI. The solution to such a crisis, accordingly to one of these "Godfathers," Geoffrey Hinton, was to retain, within AI, dependence on humanity indefinity.[3] My current research is focused upon the relationship between Christian concepts of justice and mercy and the increasing use of Artificial Intelligence in the criminal justice system. Central to such considerations is whether the AI tail, could in fact wag the dog of justice and mercy, by ceding dependence in the criminal justice system from humanity to AI. Inherent within the deployment of such technologies are deeply encoded assumptions about risk management and economic efficiency prevalent within Western Democracies, and the perfectibility of algorithmic justice. My argument is that much greater scrutiny is required of the underlying power and ethical assumptions of what might be lost from human interactions with justice, with uncritical reliance on the apparent perfectibility of the criminal justice system by these economically efficient tools. A critical aspect of such research is to refocus us upon the relational aspects of what has been encoded into our criminal justice system, by Christian concepts of justice and mercy and an understanding of the temporal imperfectability of humanity. Fundamental to such ethical considerations is the exemplified solidarity of Christ with those who are considered deviant and at the excluded margins of society.

The House of Lords Justice and Home Affairs Committee has recently reported on the use of AI technologies within the criminal justice system.[4] This report focused upon the use of algorithmic technologies which can be deployed in criminal investigation, deterrence of criminality and in punitive and rehabilitative offender management. The central aspects of their deliberations were upon 'predictive policing,' 'visa streaming' and facial recognition tools. Whilst they recognised the resource saving benefits of utilising such tools, the

---

[2] *The New Statesman*, 23-29 June 2023.
[3] Ibid, 23.
[4] House of Lords Justice and Home Affairs Committee, *AI technology and the justice system: Lords Committee Report,* House of Lords, London: 23 November 2022; https://lordslibrary.parliament.uk/ai-technology-and-the-justice-system-lords-committee-report/

Committee expressed considerable concern regarding the frontier type approach to the rigorous monitoring, regulation and transparent evaluation of such technologies within the criminal justice system. The focus of their work was to ensure that a framework of monitoring and evaluation is put in place to ensure that the deployment of AI in criminal justice is subject to pre-emptive legal controls.

Whilst such endeavours are meritorious a more fundamental question arises as to what image of justice and mercy, if any, is being presented by seeking to have UK as technological world leaders in AI in criminal justice. Clearly this Justice and Home Affairs Committee's advocacy is focused upon wresting greater human control over how this technology is to be used. However, the greater challenge lies in providing regulation which moves beyond impact assessments and ethics committees to the content of the data sets and justice assumptions which are being drawn into AI systems when utilised in criminal justice systems. In short how can the relational and restorative ethics of our jurisprudential heritage inform and critique the AI tools used in the selection, detection, administration and restoration of offenders?

We must understand that the use of AI in criminal justice systems, is an amplification of an ongoing process of algorithmic justice, whereby criminality and potential recidivism may be risk-assessed, quantified and addressed with management performance targets. This response forms part of a wider "risk society" in Western industrial society whereby technology and processes seek to manage and control risk and contingency. For Ulrich Beck these management tools are inherently entwined with the social and economic systems of modernity.[5] Specifically, from a legal perspective, Richard Posner, the former American Federal Judge and legal scholar at the University of Chicago, in his Economic *Analysis of Law*[6] promoted a jurisprudence of economic efficiency in the administration of justice generally. In essence he argued that the common law and legal systems ought to render justice by the most economically efficient, and societally acceptable, disposal of each case.

Whilst the Chicago school approach has been influential, rather than prescriptive for the administration of justice on both sides of the Atlantic, clearly in economically straightened times, such economic drivers will increasingly hold sway in governmental approaches. In an interesting exchange between the aforementioned House of Lords committee and the UK Home Office, whilst the Home Office acknowledged the potential risks of Artificial Intelligence within the Criminal Justice system, it argued that the cost and efficiency benefits of AI autonomation ought to trump undue safeguarding and transparency concerns.[7] Unsurprisingly

---

[5] Ulrich Beck, *Risk Society: Towards a New Modernity* (London: Sage, 1992), 19.
[6] Posner, Richard, *Economic Analysis of Law, (*Aspen Publishing, Frederick, MD 9th edition, 2014).
[7] AI technology and criminal justice, 3.

this approach is consistent with the joint report on AI in government, published by Sir Tony Blair and Lord Hague, which proposed a light touch and agile regulatory approach to ensure the United Kingdom could become a world leading AI provider.[8] Whilst inherent economic and societal benefits are commended, scant regard is paid to the underlying power imbalances perpetuated by the untrammelled use of AI technologies.

In computing terms the risks lie for these technologies in terms of their inputs and outputs. It is clearly economically advantageous to law enforcement to be able to identify a suspect by way of AI systems, analysing massive DNA sets or decompressed images, whether the analysis be based on human directed or deep-learning parameters. Certainly there are serious evidentiary questions that may arise regarding the probity of such findings, however more fundamental questions arise as to the objectivity of the algorithm or data sets used by AI. Will AI be directed, or direct itself, to search for certain types of characteristics, in certain types of areas? Whatever the efficiency justification for the use of AI, critical consideration is required to the biases of its creators and contributors as to how deviance is defined and who should be subject to investigation, within society. Some predict that the heralded benefits of AI technologies may soon not be confined merely to investigation, but expanded to the trial process and indeed to the determination and implementation of sentencing. Already, risk management tools and mandatory sentences are already utilised by Probation and Judiciary alike, in an effort to provide objective and uniformity in decision making. As AI develops, might perceived mistakes in data inputs and outcomes by fallible human increasingly be militated out of such decision-making processes by AI systems, in an effort to remove risky outcomes? Placing uncritical confidence in the assumed objectivity of these systems can occlude clear thinking about extant ethical and moral coding in all that is, and has been, done to definite and administer criminal justice.

Whilst considerable jurisprudential questions arise due to the increasing rise of algorithmic and AI technologies in the criminal justice system, there remain many resources of resistance in the Judeo-Christian traditions of justice and mercy encoded into Western legal traditions. Oliver O'Donovan in *The Ways of Judgment*[9] reminds us that each act of judgement falls under the delegated authority of God. There is no theological basis for excluding AI judgments from that authority, and indeed, to do otherwise would be a blasphemous elevation of AI to a perfection rivalling God's own. No judicial decisions are perfect, but, as O'Donovan argues, we are nevertheless obligated to strive towards that perfection. We do so, even in the

---

[8] A new national purpose: AI promises a world-leading future of Britain, 13th June 2023; https://www.institute. global/insights/politics-and-governance/new-national-purpose-ai-promises-world-leading-future-of-britain
[9] Oliver O'Donovan, *The Ways of Judgment,* (Grand Rapids: Eerdmans, 2005).

imperfection of the judgement giver, by focusing upon the common good of the community for whom that judicial function is administered. Conversely, actuarial justice , of the type propounded by the Chicago school, is not likely to weigh the long-term consequences of a decision, beyond the most economically efficient disposal of the present complaint.

My research will argue that AI represents the apotheosis of risk management impulses within the criminal justice system. Government, and symbiotic societal demands increasingly call for justice solutions informed by economic efficiency and statistical analysis. However, whilst risk analysis is a valid tool, it provides evaluation over a large cohort of individuals, in which statistical probability assumes the binary inevitability of further deviance. A more cruciform form of justice might well place more analytical sway upon the depth of relational knowledge of the individuals involved within the criminal justice system, whether as victim or offender and the possibility of redemption. Such costly identification will not necessarily be more economically or resource efficient, but it is more likely to identify underlying causes, whether structural or individual, for offending behaviour and to produce a just and merciful restoration of all parties within their community. It will draw upon my previous considerations of deviance within youth justice, in which fear and risk management analysis are critiqued by the solidaristic and self-sacrificial lens of the radical, and costly, identification of Christ, to provide the moral interpretive tools. Such an approach will intentionally sacrifice economic efficiency in the service of higher goals: justice mercy, and redemption. In so doing a radical and iconoclastic data set of sacrifice and identification, may be introduced into the self-validating risk analysis of Artificial Intelligence, and its algorithmic justice assumptions.