

THE PERILS AND PROMISES OF ARTIFICIAL INTELLIGENCE IN CRIMINAL SENTENCING

*Md Tasnimul Hassan**

ABSTRACT: *The proliferation of technology has resulted in an increased reliance of individuals on technological devices. The world is progressively being automated using Artificial Intelligence (AI) that are implemented to achieve specific objectives. Automation is now an important part of people’s lives, and it could heavily change the criminal justice system. This article looks at the use of AI in criminal sentencing in India, arguing that AI-based sentencing can reduce sentencing disparity, either if it is subject to a human element or it is fed with ample data that no case-fact is left out. The paper argues that since AI produces results based on the data it is fed, removing the “human” component from sentencing would be double-edged sword. The article delves into the potential challenges and concerns, such as algorithmic bias, opacity and accountability, proposing safeguards that can help in responsible and equitable implementation of AI in criminal sentencing. The article asserts that sentencing uniformity can be achieved by AI, which can bring about a suggestive sentencing score to be looked up to by the judge, instead of the judge enjoying blatant discretion, however, the risks associated with AI must be mitigated.*

<p>I. INTRODUCTION 2</p> <p>II. AI IN CRIMINAL SENTENCING: LIMITING JUDICIAL DISCRETION..... 8</p> <p style="padding-left: 20px;">A. Indian Sentencing Policy: Whether Judge is Omniscient..... 8</p> <p style="padding-left: 20px;">B. AI-based Sentencing: The Prospects and Challenges..... 12</p> <p>III. THE PROMISES OF AI IN CRIMINAL SENTENCING 16</p>	<p>IV. TESTING THE PERILS AGAINST THE PROMISES 20</p> <p style="padding-left: 20px;">A. Algorithmic Inefficiency and Bias... 21</p> <p style="padding-left: 20px;">B. Wrong Sentences and the Problem of Opacity..... 25</p> <p style="padding-left: 20px;">C. Diminishing the Scope for Individualised Justice..... 27</p> <p>V. TOWARDS AN EQUITABLE AI FRAMEWORK..... 31</p> <p>VI. CONCLUSION..... 35</p>
--	---

* The author is an Advocate based in New Delhi. He expresses his sincere gratitude towards Ameen Jauhar and Naveed Mehmood Ahmed for their guidance during the writing of this paper. The author is thankful to the anonymous peer-reviewers of the Journal for their thorough evaluation and comments.

I. INTRODUCTION

Artificial Intelligence (AI) has transformed various aspects of human life, from smartphones to automobiles to finances and health, and it is “as revolutionary as mobile phones and the internet.”¹ AI is the capacity of a system to autonomously perform errands that would commonly require human insight and dynamic cycles, however without direct human intervention.² AI uses algorithms to analyse and synthesise enormous amounts of data to respond to a question or solve a problem,³ and make predictions to facilitate or fully automate decision-making. The potential of AI is particularly promising in the field of criminal justice, where automation can be leveraged to improve investigation processes and enhance public safety, and decision making.⁴

The ability to learn through experience is an aspect of human intelligence, as humans can distinguish different circumstances through experience. This cognitive processing that humans are capable of can be replicated (not replaced) by algorithms, which process information quickly.⁵ Due to its algorithmic foundation, AI mimics human behaviour in many respects,⁶ including the ability to reason, make decisions and mistakes.⁷ It is to be noted that an “algorithm is a sequence of steps to move towards a goal,” while an AI-enabled algorithm learns from what

¹ Bill Gates, ‘The Age of AI Has Begun’ (*Gates Notes*, 21 March 2023) <<https://www.gatesnotes.com/The-Age-of-AI-Has-Begun>> accessed 25 March 2023.

² BJ Copeland, ‘Artificial intelligence’ (*Encyclopaedia Britannica*, 20 July 1998) <www.britannica.com/technology/artificial-intelligence> accessed 17 February 2023 (AI is defined as the “ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings...endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience”); Lorenzo Belenguer, ‘AI Bias: Exploring Discriminatory Algorithmic Decision-Making Models and the Application of Possible Machine-Centric Solutions Adapted from the Pharmaceutical Industry’ (2022) 2 *AI and Ethics* 771, 773 <<http://dx.doi.org/10.1007/s43681-022-00138-8>> accessed 22 February 2023 (specifying that Machine Learning (“ML”) is an advanced but tailored application of AI that simulates adaptation skills by allowing machines and their software to learn from their mistakes, more precisely, experiences).

³ Mirko Bagaric and others, ‘Erasing the Bias Against Using Artificial Intelligence to Predict Future Criminality: Algorithms Are Color Blind and Never Tire’ (2020) 88(4) *University of Cincinnati Law Review* 1037, 1042 <<https://scholarship.law.uc.edu/uclr/vol88/iss4/3>> accessed 9 February 2023.

⁴ For the purposes of this paper, AI-based decision making refers to automated computer programs that can augment or replace the human judges.

⁵ Camilo Miguel Signorelli, ‘Can Computers Become Conscious and Overcome Humans?’ (2018) 5 *Frontiers in Robotics and AI* 1 <<http://dx.doi.org/10.3389/frobt.2018.00121>> accessed 18 February 2023 (arguing that “trying to achieve conscious machines to beat humans implies that computers will never completely exceed human capabilities, or if the computer were to do it, the machine should not be considered a computer anymore”).

⁶ Christopher Rigano, ‘Using Artificial Intelligence to Address Criminal Justice Needs’ (2019) (280) *NIJ Journal* 1, 2 <www.ojp.gov/pdffiles1/nij/252038.pdf> accessed 8 February 2023.

⁷ S Mo Jones-Jang and Yong Jin Park, ‘How Do People React to AI Failure? Automation Bias, Algorithmic Aversion, and Perceived Controllability’ (2022) 28 *Journal of Computer-Mediated Communication* 1 <<http://dx.doi.org/10.1093/jcmc/zmac029>> accessed 22 February 2023.

they do, which causes the algorithms themselves to change.⁸ The latest AI developments have brought forward two groups: one supporting the AI growth and the other supporting a moratorium or at least strict regulations.⁹ Indeed, unscrupulous AI systems can have a distorting impact on the foundational tenets and structure of democracies.¹⁰ It has been contended that “governments deploy algorithms as social control mechanisms to contain and criminalize marginalized populations.”¹¹

AI integration in the Indian law enforcement has been on pace, as several tools have been deployed by agencies.¹² In the United States (US), AI-based predictive policing is a major part of law enforcement measures, and as a result, several companies like Geolitica¹³ have flourished ever since.¹⁴ The idea of using AI in sentencing has received heavy interest in recent years, mostly in the developed world; while in India, across judicial verticals and positions, sentiments have been echoing that AI should be deployed in deciding matters such as cheque bounce, motor vehicles and traffic law violations.¹⁵ In such petty matters, AI-led criminal

⁸ John Villasenor and Virginia Foggo, ‘Artificial Intelligence, Due Process, and Criminal Sentencing’ (2020) *Michigan State Law Review* 295, 301 <<https://hcommons.org/deposits/item/hc:36261/>> accessed 1 February 2023.

⁹ Pranav Dixit, “‘Why Stop?’ Bill Gates Questions Elon Musk’s Attempt to Pause ChatGPT Development” (*Business Today*, 5 April 2023) <<https://www.businesstoday.in/technology/news/story/why-stop-microsoft-cofounder-bill-gates-questions-elon-musks-attempt-to-pause-chatgpt-development-376143-2023-04-05>> accessed 5 April 2023.

¹⁰ Aleš Završnik, ‘Criminal justice, artificial intelligence systems, and human rights’ (2020) 20(4) *ERA Forum* 567, 581 <<http://dx.doi.org/10.1007/s12027-020-00602-0>> accessed 20 February 2023.

¹¹ Sarah Valentine, ‘Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control’ (2019) 46(2) *Fordham Urban Law Journal* 364 <<https://ir.lawnet.fordham.edu/ulj/vol46/iss2/4/>> accessed 14 February 2023.

¹² Parvez Hayat and Nidhi Singh, ‘How AI can play a catalytic role in managing criminal justice system for more equitable outcomes’ (*ET Government*, 11 February 2022) <<https://government.economictimes.indiatimes.com/news/digital-india/how-ai-can-play-a-catalytic-role-in-managing-criminal-justice-system-for-more-equitable-outcomes/89502089>> accessed 18 February 2023.

¹³ Geolitica, ‘Geolitica’s Advanced Predictive Policing Technology Now Available to Corporate Customers’ (*Geolitica*, 16 April 2018) <<https://geolitica.com/blog/geoliticas-advanced-predictive-policing-technology-now-available-to-corporate-customers/>> accessed 17 February 2023.

¹⁴ Herbert B. Dixon Jr., ‘Artificial Intelligence: Benefits and Unknown Risks’ (*American Bar Association*, 15 January 2021) <https://www.americanbar.org/groups/judicial/publications/judges_journal/2021/winter/artificial-intelligence-benefits-and-unknown-risks/> accessed 23 March 2023.

¹⁵ Awstika Das, “‘AI Is a Game-Changer in Legal Field’: Justice Hima Kohli on Why Artificial Intelligence Does Not Pose a Threat, But an Opportunity” (*LiveLaw*, 12 February 2023) <<https://www.livelaw.in/top-stories/artificial-intelligence-threat-opportunity-game-changer-supreme-court-judge-hima-kohli-221379>> accessed 15 February 2023; *The Indian Express*, ‘Use of AI Can Help Judiciary Dispose of Pending Cases: Gujarat HC CJ’ (*The Indian Express*, 20 July 2022) <<https://indianexpress.com/article/cities/gandhinagar/use-of-ai-can-help-judiciary-dispose-of-pending-cases-gujarat-hc-cj-8042351/>> accessed 3 February 2023; Swati Deshpande, ‘Can AI Speed up Disposal of Cases? Verdict Awaited’ (*The Times of India*, 2 August 2021) <<https://timesofindia.indiatimes.com/city/mumbai/mumbai-can-ai-speed-up-disposal-of-cases-verdict-awaited/articleshow/84959302.cms>> accessed 15 February 2023; News Desk, ‘AI Can Help Dispose Off Court Cases Quickly, Says India’s Former Chief Justice’ (*CXO Today*, 21 December 2020) <<https://www.cxotoday.com/news-analysis/ai-can-help-dispose-off-court-cases-quickly-says-indias-former-chief-justice/>> accessed 12 February 2023.

sentencing models can indeed help cut the clutter and help reduce the burden from the criminal justice system. Moreover, with the advancement of AI, India has already found itself on the road to judicial automation.¹⁶ The fact that AI carves out a result based on the type and format of the data fed to the system, the elimination of the “human” element would prove to be crucial.¹⁷

Once an accused is convicted of a crime, sentencing becomes the ultimate responsibility of a judge.¹⁸ While some theorists have focused on the obstacles and hazards associated with using AI in sentencing, for example, issues of algorithmic openness and trust, others have focused on some of the possible benefits of this process. Not only has it been claimed that AI may improve sentencing efficiency by saving time or resources, which is logical, but it has also been proposed that such technologies have the potential to address more fundamental issues such as sentencing discrimination.¹⁹ Thus, while most scholars agree that AI has made or can make

¹⁶ Rakhi Bose, ‘Dawn Of AI: As ChatGPT Enters Courts, A Look At The Question Of Jobs And Ethics’ (*Outlook India*, 2 April 2023) <<https://www.outlookindia.com/national/in-the-crosshairs-chat-gpt-the-job-market-and-the-dawn-of-ai--news-275505>> accessed 2 April 2023; Ben Cost, ‘Judge Asks ChatGPT to Decide Bail in Murder Trial’ (*New York Post*, 29 March 2023) <<https://nypost.com/2023/03/29/judge-asks-chatgpt-for-decision-in-murder-trial/>> accessed 30 March 2023; Jyoti Prakash Dutta, ‘Orissa High Court CJ Muralidhar Lauds Justice Anoop Chitkara Of Punjab & Haryana HC For Using ChatGPT While Writing Bail Order’ (*LiveLaw*, 31 March 2023) <<https://www.livelaw.in/news-updates/chief-justice-muralidhar-punjab-haryana-high-court-bail-order-chatgpt-technology-in-law-225253>> accessed 2 April 2023.

¹⁷ Lorenzo Belenguer (n2) 774 (noting that “Algorithms rely on data, and their outcomes tend to be as good as the data provided and labelled and the way the mathematical formulations are devised”).

¹⁸ Kate Stith and José A Cabranes, *Fear of Judging: Sentencing Guidelines in the Federal Courts* (University of Chicago Press 1998) 80-81, cited in Michael E Donohue, ‘A Replacement for Justitia’s Scales: Machine Learning’s Role in Sentencing’ (2019) 32(2) *Harvard Journal of Law & Technology* 657, 658 <<https://jolt.law.harvard.edu/assets/articlePDFs/v32/32HarvJLTech657.pdf>> accessed 8 February 2023.

¹⁹ Jesper Ryberg, ‘Sentencing Disparity and Artificial Intelligence’ (2021) *The Journal of Value Inquiry* <<http://dx.doi.org/10.1007/s10790-021-09835-9>> accessed 12 January 2023. See also, Alberto De Diego Carreras, ‘The Moral (Un)Intelligence Problem of Artificial Intelligence in Criminal Justice: A Comparative Analysis under Different Theories of Punishment’ (2020) 25(1) *UCLA Journal of Law & Technology* 1 <<https://uclajolt.com/the-moral-unintelligence-problem-of-artificial-intelligence-in-criminal-justice-a-comparative-analysis-under-different-theories-of-punishment/>> accessed 1 February 2023 (arguing that incompetence of AI tools in sentencing is less problematic under a utilitarian framework, where moral culpability is less important than long-term goals, using AI in sentencing may be desirable). See further, Christopher Rigano, (n6) 8 (noting that, everyday “holds the potential for new AI applications in criminal justice, paving the way for future possibilities to assist and ultimately improve public safety in the criminal justice system”). See also, Mirko Bagaric and others, ‘The Solution to the Pervasive Bias and Discrimination in the Criminal Justice System: Transparent and Fair Artificial Intelligence’ (2022) 59(1) *American Criminal Law Review* 95 <<https://www.law.georgetown.edu/american-criminal-law-review/in-print/volume-59-number-1-winter-2022/the-solution-to-the-pervasive-bias-and-discrimination-in-the-criminal-justice-system-transparent-and-fair-artificial-intelligence/>> accessed 2 February 2023 (arguing that when making decisions, algorithms can outperform humans in objectivity and predictability, hence they can overcome the prevalent bias and discrimination in the criminal justice system due to humans). See also, Brandon Garrett and John Monahan, ‘Assessing Risk: The Use of Risk Assessment in Sentencing’ (*Judicature*, 24 March 2020) <<https://judicature.duke.edu/articles/assessing-risk-the-use-of-risk-assessment-in-sentencing/>> accessed 12 January 2023 (arguing that the use of risk assessment tools in sentencing can be a valuable tool for judges, but they must be used carefully with an eye towards fairness and equity). See also, Sarah

crucial changes to the criminal justice system, they find differences of opinions on whether these technological innovations are a “panacea for the criminal justice system.”²⁰

India’s overall AI growth has been phenomenal in the last decade which is expected to grow further.²¹ The Supreme Court of India (“SC”) is relying on several AI-based tools like SUVAS,²² SUPACE²³ and TERES²⁴ in its administration of justice. Undoubtedly, AI has the potential to become a fundamental part of the criminal justice system.²⁵ However, AI in criminal justice may violate fundamental rights as well as the settled principles of criminal law, due to its intrinsic ambiguous nature.²⁶ The technical sophistication of the AI systems used in

L Desmarais, ‘The Role of Risk Assessment in the Criminal Justice System: Moving Beyond a Return to the Status Quo’ (*Harvard Data Science Review*, 31 March 2020) <<https://hdsr.mitpress.mit.edu/pub/60jfy7hm/release/6>> accessed 12 January 2023 (arguing that predictive risk analysis may increase the effectiveness and fairness of the criminal justice system, but in order to prevent unintended consequences, it must be used responsibly and with caution).

²⁰ Aleš Završnik (2020) (n10) 568. See also, Danielle Kehl and others, ‘Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing’ (2017) Berkman Klein Centre for Internet & Society, Harvard Law School 1, 36 <<https://dash.harvard.edu/handle/1/33746041>> accessed 2 February 2023 (arguing that though algorithms can “have the potential to improve sentencing accuracy in the criminal justice system and reduce the risk of human error and bias,” they may also “reinforce or exacerbate existing biases and...undermine certain basic tenets of fairness that are central to...justice system”). See also, Willem Gravett, ‘Sentenced by an Algorithm — Bias and Lack of Accuracy in Risk-Assessment Software in the United States Criminal Justice System’ (2021) 34 South African Journal of Criminal Justice 31 <<http://dx.doi.org/10.47348/sacj/v34/i1a2>> accessed 11 January 2023 (arguing that “while AI tools could potentially increase sentencing accuracy and reduce the risk of human error and bias by providing evidence-based reasons in place of ‘ad-hoc’ decisions by human beings beset with cognitive and implicit biases, they also have the potential to reinforce and exacerbate existing biases, and to undermine certain of the basic constitutional guarantees embedded in the justice system”). See also, Cynthia Rudin and others, ‘The Age of Secrecy and Unfairness in Recidivism Prediction’ (*Harvard Data Science Review*, 31 March 2020) <<https://hdsr.mitpress.mit.edu/pub/7z10o269/release/7>> accessed 17 January 2023 (arguing that transparency and accountability are essential to ensure that the algorithms do not perpetuate or amplify existing biases and discrimination).

²¹ ET Tech, ‘AI Adoption to Add \$500 Billion to India’s GDP by 2025: Nasscom’ (*The Economic Times*, 23 June 2022) <<https://economictimes.indiatimes.com/tech/technology/integrated-adoption-of-ai-and-data-utilization-can-add-500-billion-to-indias-gdp/articleshow/92412781.cms>> accessed 12 January 2023.

²² SUVAS was launched in 2019, which has the capability to translate English language judicial documents into several vernacular Indian languages. See, Ajoy Karpuram, ‘The Supreme Court’s Translation Project is Slowing to a Halt’ (*Supreme Court Observer*, 12 November 2021) <www.scobserver.in/journal/the-supreme-courts-translation-project-is-slowing-to-a-halt/> accessed 4 February 2023.

²³ SUPACE was launched in 2021 to make legal research easier for judges. See, Express News Service, ‘CJI launches top court’s AI-driven research portal’ (*The Indian Express*, 7 April 2021) <<https://indianexpress.com/article/india/cji-launches-top-courts-ai-driven-research-portal-7261821/>> accessed 4 March 2023.

²⁴ TERES is a platform used to live transcribe the SC proceedings. See, Utkarsh Anand, ‘Live transcription of Supreme Court proceedings introduced’ (*Hindustan Times*, 22 February 2023) <www.hindustantimes.com/india-news/live-transcription-of-supreme-court-proceedings-introduced-101677004607162.html> accessed 25 February 2023.

²⁵ Christopher Rigano (n6) 2. See, Dan Hunter and others, ‘A Framework for the Efficient and Ethical Use of Artificial Intelligence in the Criminal Justice System’ (2020) 47 Florida State University Law Review 749 <<https://www.fsulawreview.com/wp-content/uploads/2022/08/ETHICAL-USE-OF-ARTIFICIAL-INTELLIGENCE.pdf>> accessed 11 February 2023 (proposing a framework for integrating AI with criminal justice to enhance its effectiveness while acknowledging human resistance to AI).

²⁶ Aleš Završnik (2020) (n10) 579.

criminal sentencing may often lead to a “black box” effect.²⁷ AI system must be transparent, as the explainability of outcomes and the clarity of reasoning are critical for the system to be called ‘just’,²⁸ since a process that is opaque and unintelligible cannot be legitimate.²⁹

The idea of criminal justice is absolute-abstract. Going with the current system, where judges view “fact and circumstances”, errors remain prevalent, indicating that even human judges may not be perfect. Particularly, in social settings where the State authority is biased towards specific social groups, the judiciary attracts their faith – which at times, calls out wrongful prosecutions.³⁰ Even in cases where State’s bias is not evident, it is substantially difficult to argue that other prosecutions, or sentencing for that matter, are fair.³¹ In any case, it remains impossible to “undo” what the prolonged trials and incarceration may do to such persons.³² Therefore, the need for a compensation regime appears like a starting point. Section 250 of the Code of Criminal Procedure, 1973 (“CrPC”) seeks to compensate for accusation without reasonable cause, yet this provision finds rare discussion or application, and is applicable only to summons and warrants cases.³³ In compliance of an order of the Delhi High Court,³⁴ the Law Commission of India published a report in 2018 titled “Wrongful Prosecution (Miscarriage of

²⁷ The black-box effect implies that the AI-tools (especially those used in sentencing) may be protected by trade-secrets, and can defy the control of its programmers, which can negate transparency and accountability – without people knowing how the decision was arrived at. For discussion on black box effect, see, Han-Wei Liu and others, ‘Beyond State v. Loomis: artificial intelligence, government algorithmization and accountability’ (2019) 27(2) International Journal of Law and Information Technology 122 <<http://dx.doi.org/10.1093/ijlit/eaz001>> accessed 18 January 2023. See also, Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2016).

²⁸ Aleš Završnik (2020) (n10) 568.

²⁹ Intentional opacity occurs when a system’s inner workings are hidden, either by a for-profit technology provider or by a public authority to prevent “gaming” the system. See, Gianna Seglias, ‘Bias and Discrimination in Opaque Automated Individual Risk Assessment Systems: Challenges for Judicial Review under the Equality Act 2010’ (2021) The Oxford University Undergraduate Law Journal 53, 62 <https://www.law.ox.ac.uk/sites/default/files/migrated/public_law_2.pdf> accessed 11 January 2023.

³⁰ Waqar Hasan, ‘Why The Percentage Of Muslim Prisoners In India’s Jails Is Disproportionate to Their Population In India’ (*Article 14*, 15 March 2022) <<https://article-14.com/post/why-the-percentage-of-muslim-prisoners-in-india-s-jails-is-disproportionate-to-their-population-in-india--6230001268058>> accessed 3 April 2024; Bilal Kuchay, ‘127 Indian Muslims charged with ‘terror’ acquitted after 19 years’ (*Al Jazeera*, 9 March 2021) <www.aljazeera.com/news/2021/3/9/127-indian-muslims-charged-with-terror-acquitted-after-19-years> accessed 3 April 2024; Rifat Fareed, ‘Kashmiris found innocent after 23 years in captivity’ (*Al Jazeera*, 28 July 2019) <www.aljazeera.com/news/2019/7/28/after-23-years-of-wrongful-imprisonment-kashmiri-men-return-home> accessed 3 April 2024.

³¹ Anand Mohan J, ‘Sentenced to death twice, Madhya Pradesh man cleared of rape-murder after 11 years in jail’ (*The Indian Express*, 1 April 2024) <<https://indianexpress.com/article/india/madhya-pradesh-court-acquits-man-rape-murder-case-9244198/>> accessed 3 April 2024.

³² Mahtab Alam, ‘Who Will Return Our Precious Years, Ask the Wrongfully Incarcerated’ (*The Wire*, 24 December 2016) <<https://thewire.in/rights/wongful-incarceration-priosners-tada-terrorism>> accessed 3 April 2024.

³³ *State of Rajasthan v. Jainudeen Shekh*, (2016) 1 SCC 514.

³⁴ *Babloo Chauhan @ Dabloo v. State Govt. of NCT of Delhi*, 2017 SCC OnLine Del 12045.

Justice): Legal Remedies”, recommending a framework to provide relief to the wrongfully prosecuted.³⁵ In 2019, a private member Bill was introduced in the Lok Sabha, titled as “The Protection of Rights of Wrongful Convicts Bill”, which provided a roadway to justice against wrongful convictions, however, this Bill could not occupy many hands.

Without any remedy against wrongful convictions, unlike in countries like the US and UK, and with prevalent sentencing disparity, there remains a wide void in the Indian criminal procedure.³⁶ This void can be filled – if AI is the decision maker, which can be held accountable for mistakes. However, the implementation of AI with raw information in prognostications and decision-making processes has the potential to minimize the degree of human subjectivity, and such technology may also incorporate prejudicial inclinations, thereby producing erroneous or discriminatory outcomes.³⁷ This paper argues that algorithmic sentencing serves the intended purpose, that is, reducing sentencing disparity, either if it is subject to a human element or it is fed with ample data that no case-fact is left out. The paper in Part II discusses criminal sentencing in India and elaborates the judicial precedents, to argue that there is no codified sentencing policy in India, because in matters of criminal sentencing, a judge is deemed omniscient. The paper proceeds with describing the application of AI in criminal sentencing. The paper then analyses the promises and perils of the application of AI to sentencing in Parts III and IV, respectively, and argues that the perils outweigh the promises, unless certain principles are adhered to, while designing the sentencing algorithm, failing which AI would reincarcerate the human biases in its decisions. Part V provides recommendations for equitable and responsible deployment of AI in sentencing and Part VI provides concluding remarks.

³⁵ Law Commission of India, *Wrongful Prosecution (Miscarriage of Justice): Legal Remedies* (Report No. 277, Ministry of Law and Justice, Government of India 2018).

³⁶ Pratiksha Basarkar and Sakshi Jain, ‘Supreme Court Acquittal Of 8 Men On Death Row Reveals Failures by Police, Prosecution & Lower Courts’ (*Article 14*, 16 January 2023) <<https://article-14.com/post/supreme-court-acquittal-of-8-men-on-death-row-reveals-failures-by-police-prosecution-lower-courts-63c45f79e9af3>> accessed 14 February 2023; Gloria Pazmino and others, ‘New York City agrees to pay \$26 million to 2 men wrongly convicted of Malcolm X murder’ (*CNN*, 30 October 2022) <<https://edition.cnn.com/2022/10/30/us/nyc-settlement-malcolm-x-murder/index.html>> accessed 16 May 2023; John Leonard, ‘Fujitsu Horizon scandal: Postmasters offered £600k compensation for wrongful conviction’ (*Computing - The UK's leading source for the analysis of business technology*, 19 September 2023) <www.computing.co.uk/news/4126501/fujitsu-horizon-scandal-postmasters-offered-gbp600k-compensation-wrongful-conviction> accessed 16 May 2023.

³⁷ Genevieve Smith and Ishita Rustagi, ‘Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook’ (Berkeley Haas Centre for Equity, Gender and Leadership 2020) 16 <https://haas.berkeley.edu/wp-content/uploads/UCB_Playbook_R10_V2_spreads2.pdf> accessed 20 January 2023.

II. AI IN CRIMINAL SENTENCING: LIMITING JUDICIAL DISCRETION

A. Indian Sentencing Policy: Whether Judge is Omniscient

The procedure is a punishment in India's criminal justice system.³⁸ Socially justifiable deprivation of life and liberty is primarily motivated by deterrence and reformation.³⁹ In most jurisdictions including India, judges have the discretion to award sentence to a convict, depending on the aggravating and mitigating circumstances, that, respectively, increase or lessen a sentence.⁴⁰ Without any uniform sentencing policy in India, it resultantly gets based on the judges' personal beliefs and wide discretion.⁴¹ In the absence of any statutory sentencing guidelines, decisions vary in its reasoning and logic across the country.⁴² Thus, where there is no sentencing policy in place and every sentencing decision comes out of discretion of the judge, it remains to be seen whether AI can be deployed in criminal sentencing, a stage hugely neglected in the justice system.⁴³

Since the early 2000s, the core of the sentencing policy in Indian criminal law is the notion of proportionality, or "*just deserts*."⁴⁴ Criminal law upholds this principle when determining

³⁸ ANI, 'Process is punishment in India's criminal justice system: CJI Ramana' (*The Print*, 16 July 2022) <<https://theprint.in/india/process-is-punishment-in-indias-criminal-justice-system-cji-ramana/1042074/>> accessed 18 February 2023.

³⁹ *Rajendra Prasad v. Uttar Pradesh* AIR 1979 SC 916 [88].

⁴⁰ In *Gopal Singh v. State of Uttarakhand* (2013) 7 SCC 545 [19], the SC pointed out:

"A Court, while imposing sentence, has to keep in view the various complex matters in mind...The legislature in its wisdom has conferred discretion on the Judge who is guided by certain rational parameters, regard been had to the factual scenario of the case. In certain spheres the legislature has not conferred that discretion and in such circumstances, the discretion is conditional. In respect of certain offences, sentence can be reduced by giving adequate special reasons. The special reasons have to rest on real special circumstances."

See also, *Mohd. Giasuddin v. State of Andhra Pradesh* AIR 1977 SC 1926 [14] (The SC summed up the components of a proper sentence as laid down in the 47th Report of the Law Commission of India: prior criminal record; age; professional and social record; education and social background; emotional and mental condition; prospect of rehabilitation, treatment, training and return to normal life; and the possibility of the sentence serving as a deterrent).

⁴¹ Anju Vali Tikoo, 'Individualisation of Punishment, Just Desert and Indian Supreme Court Decisions: Some Reflections' (2017) 2(Winter) ILI Law Review 20, 42 <<https://ili.ac.in/pdf/tikoo.pdf>> accessed 10 January 2023.

⁴² Samarth Grover, 'Death Penalty: How Trial Courts in India Are Violating Sentencing Guidelines' (*The Quint*, 29 October 2022) <www.thequint.com/news/law/death-penalty-how-trial-courts-in-india-are-violating-sentencing-guidelines> accessed 18 February 2023.

⁴³ Maulshree Pathak, 'Need for sentencing guidelines in India' (*P39A Blog*, 9 December 2020) <<https://p39ablog.com/2020/12/need-for-sentencing-guidelines-in-india/>> accessed 18 February 2023.

⁴⁴ According to the theory of *just deserts* punishment, the severity of the penalty should match the seriousness of the crime. See, John J Sloan and J Langly Miller, 'Just Deserts, The Severity Of Punishment And Judicial Sentencing Decisions' (1990) 4(1) Criminal Justice Policy Review 19

liability in accordance with each type of criminal action's culpability.⁴⁵ Every court is responsible for imposing the appropriate penalty based on the type of offence committed and how it was carried out.⁴⁶ The judge has a moral obligation to forgo formalities and consider the issue of sentencing from a broad sociological perspective.⁴⁷ The fundamental problem with criminal law in India is that it is impossible to set standards, leaving judges with a great deal of latitude in determining the severity of sentence.⁴⁸ For determining proper punishment, it is crucial to take the type and gravity of the offence into account, not the perpetrator.⁴⁹ Although there are no legislative or judicial guidelines for sentencing, there are several principles like proportionality, deterrence, and rehabilitation – that ought to be considered by courts while sentencing an accused,⁵⁰ however, the Courts usually do not strictly follow these principles.⁵¹

<<http://dx.doi.org/10.1177/088740349000400102>> accessed 19 February 2023. In *Ankush Maruti Shinde v. State of Maharashtra*, (2009) 6 SCC 667 [28], the SC acknowledged:

“The criminal law adheres in general to the principle of proportionality in prescribing liability according to the culpability of each kind of criminal conduct. It ordinarily allows some significant discretion to the Judge in arriving at a sentence in each case, presumably to permit sentences that reflect more subtle considerations of culpability that are raised by the special facts of each case. Judges in essence affirm that punishment ought always to fit the crime; yet in practice sentences are determined largely by other considerations...Inevitably these considerations cause a departure from just desert as the basis of punishment and create cases of apparent injustice that are serious and widespread.”

Notably, in the above case, the SC found out that the accused persons were falsely implicated, and thus acquitted the six convicts (see, *Ankush Maruti Shinde v. State of Maharashtra*, 2019 SCC OnLine SC 317).

⁴⁵ *Shailesh Jasvantbhai v. State of Gujarat*, (2006) 2 SCC 359 [9]; In *Ram Naresh v. State of Chhattisgarh*, (2012) 4 SCC 257 [79], the SC observed:

“The principle of proportion between the crime and punishment is the principle of ‘just deserts’ that serves as the foundation of every criminal sentence that is justifiable. In other words, the ‘doctrine of proportionality’ has a valuable application to the sentencing policy under the Indian criminal jurisprudence. Thus, the court will not only have to examine what is just but also as to what the accused deserves keeping in view the impact on the society at large.”

See also, *Alister Anthony Pereira v. State of Maharashtra*, (2012) 2 SCC 648 [84-85], where the SC observed:

“One of the prime objectives of the criminal law is imposition of appropriate, adequate, just and proportionate sentence commensurate with the nature and gravity of crime and the manner in which the crime is done. There is no straitjacket formula for sentencing an accused on proof of crime. The courts have evolved certain principles: the twin objective of the sentencing policy is deterrence and correction. What sentence would meet the ends of justice depends on the facts and circumstances of each case and the court must keep in mind the gravity of the crime, motive for the crime, nature of the offence and all other attendant circumstances...The Court has to take into consideration all aspects including social interest and consciousness of the society for award of appropriate sentence.”

⁴⁶ *Bantu v. State of Uttar Pradesh*, (2008) 11 SCC 113 [21].

⁴⁷ *Muniappan v. State of Tamil Nadu*, AIR 1981 SC 1220 [2].

⁴⁸ *Jagmohan Singh v. State of Uttar Pradesh*, AIR 1973 SC 947 [26].

⁴⁹ *Ravji @ Ram Chandra v. State of Rajasthan*, (1996) 2 SCC 175 [24]. See also, Philip Petrov, ‘Proportionality in Criminal Sentencing: A Cognitive Hypothesis’ (2022) 43 Oxford Journal of Legal Studies 124 <<https://doi.org/10.1093/ojls/gqac020>> accessed 13 January 2023 (suggesting that people's judgments about the proportionality of sentences are based on their intuitive understanding of the severity of the crime and the offender's culpability).

⁵⁰ *Soman v. State of Kerala*, (2013) 11 SCC 382 [15-16]; *State of Madhya Pradesh v. Bablu Natt*, (2009) 2 SCC 272 [15].

⁵¹ *Rajendra Prahladrao Wasnik v. State of Maharashtra*, (2019) 12 SCC 460 [75].

The Indian Penal Code, 1860 (“IPC”), provides for a minimum and maximum punishment, between which, the judge must award the penalty,⁵² and if a particular offence would result in the minimum or maximum punishment depends entirely on the judge’s discretion. Under the IPC, more than 200 offences are punishable with imprisonment or fine or both, while about 350 offences give discretion to the judge to decide the type of imprisonment – simple or rigorous.⁵³ There are five different ways that criminals can be punished, depending on the nature of crime: death penalty, life sentence, incarceration (simple or rigorous), property forfeiture, and fine.⁵⁴ The CrPC discusses sentencing in Sections 235, 248, 325, 360 and 361, providing wide discretionary powers to the judge.

According to Section 235(2) of CrPC, the judge must first hear from the accused before passing the appropriate sentence, unless Section 360 of CrPC applies. The SC has held that Section 235(2) of CrPC contemplates the opportunity for both the prosecution and the accused to provide facts and materials pertaining to the punishment in addition to the right for oral submission.⁵⁵ Moreover, the purpose of Section 235(2) of CrPC is to allow the convicted individual a second chance to inform the court of any personal or social factors that would assist the court in imposing an appropriate sentence.⁵⁶ In accordance with Section 354(3) of CrPC, judges must give special justifications whenever a sentence authorises a life-or-death punishment, and Section 354(1)(B) of the CrPC directs judges to justify a sentence. Only the most serious of cases allow for the imposition of an extreme penalty, while sentencing must consider the [mitigating] circumstances of the convict.⁵⁷

The Court may not impose a punishment that is lesser than the minimum, although it shall be less severe than the maximum.⁵⁸ Section 325 of CrPC limits the power of a Magistrate to inflict a punishment greater than the prescribed maximum, by compelling the Magistrate to forward the matter to the Chief Judicial Magistrate, wherein the latter may pass a sentence as s/he thinks

⁵² Mandatory minimum sentencing refers to “a sentence which must be imposed without leaving any discretion to the court. It means a quantum of punishment which cannot be reduced below the period fixed.” See, *Mohd Hashim v. State of Uttar Pradesh*, (2017) 2 SCC 198 [19]; the SC is expected to review this judgment, though (see, Khadija Khan, ‘What are mandatory minimum sentences and what is the rationale behind them?’ (*The Indian Express*, 25 January 2023) <<https://indianexpress.com/article/explained/explained-law/mandatory-minimum-sentences-rationale-behind-them-8404166/>> accessed 18 February 2023).

⁵³ Maulshree Pathak (n43).

⁵⁴ IPC 1860, s 53.

⁵⁵ *Santa Singh v. State of Punjab*, AIR 1976 SC 2386 [4].

⁵⁶ *Tarlok Singh v. State of Punjab*, AIR 1977 SC 1747 [2].

⁵⁷ *Bachan Singh v. State of Punjab*, AIR 1980 SC 898 [206-207].

⁵⁸ *Vikas Yadav v. State of UP*, (2016) 9 SCC 541 [39].

fit. It remains the discretion of the judge, whether to have mercy or be harsh to the convict.⁵⁹ Not only must the sentence be fair to the accused, but also to the victim and society.⁶⁰ Sections 360 and 361 of the Criminal Code authorise the courts to release convicts and offenders based on their good behaviour while incarcerated; however, in the absence of a sentencing policy, it would depend upon the report of the jail authority.⁶¹ After hearing from the accused on sentence, the Magistrate must impose a sentence in accordance with the law if the accused is guilty, but he does not resort to Sections 325 or 360 of CrPC.⁶²

In 2003, the Malimath Committee proposed instituting sentencing guidelines to reduce sentencing uncertainty and inconsistency, and the suggestion was reasserted in the Madhava Menon Committee report.⁶³ The Malimath Committee pitted the adversarial system against the inquisitorial system and concluded that only the adversarial system allows for a fair trial.⁶⁴ However, India has failed to develop proper sentencing guidelines.⁶⁵ In such a situation, much

⁵⁹ In *Shailesh Jasyantbhai v. State of Gujarat*, (2006) 2 SCC 359 [7], the SC noted:

“Protection of society and stamping out criminal proclivity must be the object of law which must be achieved by imposing appropriate sentence. Therefore, law as a cornerstone of the edifice of ‘order’ should meet the challenges confronting the society...in operating the sentencing system, law should adopt the corrective machinery or the deterrence based on factual matrix. By deft modulation sentencing process be stern where it should be, and tempered with mercy where it warrants to be.”

⁶⁰ *State of MP v. Mehtaab*, (2015) 5 SCC 197 [7]. See also, *State of Andhra Pradesh v. Polamala Raju*, (2000) 7 SCC 75 [9], where the SC held:

“...it is an obligation of the sentencing court to consider all relevant facts and circumstances bearing on the question of sentence and impose a sentence commensurate with the gravity of the offence. The sentencing court must hear the loud cry for justice by the society and more particularly, in cases of heinous crime of rape of innocent helpless children, as in this case, of the victim of the crime and respond by imposing a proper sentence.”

⁶¹ In situations where the accused may have been handled by the court in accordance with Section 360 but chooses not to grant the benefit, it must specify those reasons in its judgment. See, *Chandreshwar Sharma v. State of Bihar*, (2000) 9 SCC 245 [3].

⁶² See, CrPC 1973, ss. 248(2) and 255(2). The SC, while referring the issue of same-day sentencing in matters of death penalty to Constitution Bench, has stated that “it is necessary to have clarity in the matter to ensure a uniform approach on the question of granting real and meaningful opportunity, as opposed to a formal hearing, to the accused/convict, on the issue of sentence” (*In Re: Framing Guidelines on Mitigating Circumstances to be Considered While Imposing Death Sentences*, 2022 LiveLaw SC 777).

⁶³ Report of the Committee on Reforms in Criminal Justice System (Chairperson: Justice V.S. Malimath) (Ministry of Home Affairs, Government of India, 2003). See also, Report of the Committee on Criminal Justice Reforms (Chairperson: Dr. Justice V.S. Malimath) (Ministry of Home Affairs, Government of India, 2007). Law Commission of India, Wrongful Prosecution (Miscarriage of Justice): Legal Remedies (Report No. 277, Ministry of Law & Justice, Government of India 2018).

⁶⁴ The adversarial system involves presenting the case before a neutral jury or judge – prevalent in the US, Canada, and the United Kingdom (“UK”). In contrast, in an inquisitorial system, the judge actively investigates the facts of the cases and gathers evidence – prevalent in France, Germany and Italy. For a comparison between the two systems, see, Francesco Parisi, ‘Rent-Seeking through Litigation: Adversarial and Inquisitorial Systems Compared’ (2002) 22 International Review of Law and Economics 193 <[http://dx.doi.org/10.1016/s0144-8188\(02\)00089-3](http://dx.doi.org/10.1016/s0144-8188(02)00089-3)> accessed 4 January 2023.

⁶⁵ The SC in *State of Punjab v. Prem Sagar*, (2008) 7 SCC 550 [2], observed:

“In our judicial system, we have not been able to develop legal principles as regards sentencing. The superior courts except making observations with regard to the purport and object for which punishment is imposed upon an offender, have not issued any guidelines...”

reliance is placed on the aggravating and mitigating factors in each case, which enables a judge to utilize their discretion and award the penalty. Mitigating factors, such as low age of the accused, previous criminal background, remorse⁶⁶ for the crime can lead to a more lenient sentence; while aggravating factors, such as lack of remorse⁶⁷ and hostility towards the victim, can lead to a more severe sentence.⁶⁸ The Court must justify the gravity of punishment depending on the facts and circumstances of the case.⁶⁹ The SC jurisprudence so far, on all fronts deciding the fate of an offender, seeks to analyse the mitigating and aggravating circumstances and then adjudicate on the quantum of sentence, putting a whole lot of discretion of fixing a sentence between the prescribed maximum and minimum on the judicial mind.⁷⁰

B. AI-based Sentencing: The Prospects and Challenges

In recent years, application of risk assessment in criminal sentencing has increased, with the aim of reducing recidivism and prison populations.⁷¹ Sentencing discretion can be replaced by automation, a crucial change that several jurisdictions globally are looking up to.⁷² Thus, it has

⁶⁶ Pieter du Toit, 'The Role of Remorse in Sentencing' (2021) 34(3) Nelson Mandela University Law Journal 558 <<http://dx.doi.org/10.17159/obiter.v34i3.12008>> accessed 27 February 2023.

⁶⁷ Ninawa Butrus, 'Judicial Sentencing Considerations in Cases of Violent Offenders versus Sexual Offenders' (2018) 25 Psychiatry, Psychology and Law 653 <<http://dx.doi.org/10.1080/13218719.2018.1473175>> accessed 27 February 2023.

⁶⁸ Matthew PJ Ashby, 'Comparing Methods for Measuring Crime Harm/Severity' (2018) 12(4) Policing: A Journal of Policy and Practice 439 <<http://dx.doi.org/10.1093/police/pax049>> accessed 12 February 2023.

⁶⁹ *State of Madhya Pradesh v. Bablu*, (2014) 9 SCC 281 [18]; *State of Madhya Pradesh v. Surendra Singh*, (2015) 1 SCC 222 [13]; *State of Karnataka v. Sharanappa Basanagouda Aregoudar*, (2002) 3 SCC 738 [6].

⁷⁰ In *Suresh Chandra Bahri v. State of Bihar*, AIR 1994 SC 2420 [101], the SC observed:

"...The criticism of judicial sentencing has raised its head in various forms — that it is inequitable as evidenced by desperate sentences; that it is ineffective; or that it is unfair being either inadequate or in some cases harsh. It has been often expressed that there is a considerable disparity in sentencing an accused found to be guilty for some offence. This sentencing variation is bound to occur because of the varying degrees of seriousness in the offence and/or varying characteristics of the offender himself. Moreover, since no two offences or offenders can be identical the charge or label of variation as disparity in sentencing necessarily involves a value-based judgment, i.e., disparity to one person may be a simply justified variation to another. It is only when such a variation takes the form of different sentences for similar offenders committing similar offences that it can be said to be disparate sentencing."

⁷¹ John Monahan and Jennifer L Skeem, 'Risk Assessment in Criminal Sentencing' (2016) 12 Annual Review of Clinical Psychology 489 <<http://dx.doi.org/10.1146/annurev-clinpsy-021815-092945>> accessed 23 January 2023 (discussing the different roles and meanings of risk assessment in sentencing, as well as four major problems with its implementation: conflating risk and blame, relying on group data, ignoring potential racial and economic disparities, and failing to differentiate risk assessment from risk reduction).

⁷² Danielle Kehl and others (n20) 10 (noting that US's Virginia first deployed risk assessment tool in sentencing in as early as 1994); Henry Zwartz, 'AI Is Creeping into the World's Courts. Should We Be Concerned?' (*UNSW Newsroom*, 27 September 2022) <<https://newsroom.unsw.edu.au/news/business-law/ai-creeping-world%E2%80%99s-courts-should-we-be-concerned>> accessed 4 January 2023; Andrew Lensen and Marcin Betkier, 'We Built an Algorithm That Predicts the Length of Court Sentences – Could AI Play a Role in the Justice System?' (*The Conversation*, 29 November 2022) <<http://theconversation.com/we-built-an-algorithm-that>>

been widely argued that sentencing is amenable to automated decision-making,⁷³ as such, the sentencing algorithms may be tried, and if the trials are successful, its deployment can be considered.⁷⁴ AI serves as the opportunity to replace the human mind (which is often biased) in sentencing, while offering speedy resolution and objective answers to complex questions. The first step is to create a dataset with cases in which judges have determined punishments for certain offences, based on which the system will predict a sentence, telling the judge how their peers handled similar cases in the past.⁷⁵ It has one simple problem though, the outcome that the AI would bring is entirely based on the type of data it would be fed.⁷⁶ The methodology of AI is to “eliminate humans from the equation,” as it can learn from its own experiences.⁷⁷ However, removing human discretion in criminal sentencing is a “double-edged sword” that can lessen bias and also intensify or create new ones.⁷⁸

The deployer purchases algorithms from private developers, meaning that the algorithm is a secret, and only the proprietors or the deployer can know how the tool functions and makes decisions. This lack of transparency is the Achilles heel of AI-enabled sentencing, which keeps these tools uncontrolled by keeping the algorithm hidden from the public.⁷⁹ One such apparent instance was witnessed in *Wisconsin v. Loomis*,⁸⁰ wherein the Wisconsin Supreme Court ruled that risk assessment⁸¹ tool did not violate due process even though neither the court nor the accused were informed of the assessment’s methodology.⁸² The court explained that

[predicts-the-length-of-court-sentences-could-ai-play-a-role-in-the-justice-system-193300](#)> accessed 5 January 2023; Rina Chandran, ‘As Malaysia Tests AI Court Sentencing, Lawyers Fear for Justice’ (*Thomson Reuters Foundation*, 12 April 2022) <<https://news.trust.org/item/20220411160005-k1a5o/>> accessed 22 January 2023.

⁷³ Nigel Stobbs and others, ‘Can sentencing be enhanced by the use of Artificial Intelligence?’, (2017) 41(5) *Criminal Law Journal* (accepted pre-print version) 1, 3 <<https://eprints.qut.edu.au/115410/>> accessed 20 January 2023.

⁷⁴ *ibid* 1.

⁷⁵ Vincent Chiao, ‘Predicting Proportionality: The Case for Algorithmic Sentencing’ (2018) 37(3) *Criminal Justice Ethics* 238, 240 <<http://dx.doi.org/10.1080/0731129x.2018.1552359>> accessed 20 January 2023.

⁷⁶ Endre Begby, ‘Automated Risk Assessment in the Criminal Justice Process: A Case of “Algorithmic Bias”?’ *Prejudice: A Study in Non-Ideal Epistemology* (Oxford University Press 2021) <<https://doi.org/10.1093/oso/9780198852834.003.0009>> accessed 17 January 2023 (arguing that the bias is embedded in the data itself, because of the historically institutionalized discrimination).

⁷⁷ Aleš Završnik (2020) (n10) 568.

⁷⁸ Aleš Završnik (2020) (n10) 581.

⁷⁹ Christopher Thomas and Antonio Pontón-Núñez, ‘Automating Judicial Discretion: How Algorithmic Risk Assessments in Pretrial Adjudications Violate Equal Protection Rights on the Basis of Race’ (2022) 40 *Minnesota Journal of Law & Inequality* 371, 371 <<http://dx.doi.org/10.24926/25730037.649>> accessed 6 April 2023 (arguing that “the state cannot meet its burden of proving that the algorithms are narrowly tailored, due to their opacity”).

⁸⁰ *State of Wisconsin v. Eric L. Loomis*, 881 N.W. 2d 749 (2016).

⁸¹ AI-based decision making, of which risk assessment tools are a subset, make decisions with little or no human intervention. See, Gianna Seglias (n29) 55.

⁸² “*State v. Loomis*” (*Harvard Law Review*, 10 March 2017) <<https://harvardlawreview.org/2017/03/state-v-loomis/>> accessed 12 January 2023.

COMPAS⁸³ may not be used “to determine whether an offender is incarcerated” or “to determine the severity of the sentence,”⁸⁴ meaning that the system could not determine if a criminal would be convicted or imprisoned, nor could it be used to calculate the length of a sentence.⁸⁵ However, the court acknowledged that to guard against any potential errors or a lack of individualised decision-making, accused have the right to due process, which calls for warnings to be given prior to sentencing using risk assessment algorithms.⁸⁶ Still, the logic fades away, since the algorithm is being used to “foresee” an accused’s future, based on the recidivism risk score presented by COMPAS.⁸⁷ The US Supreme Court denied the appeal for a writ of certiorari, and thus, COMPAS was validated.⁸⁸

Sentencing algorithms operate on a dataset compiled from actual judicial decisions, or the record is created on potential estimates of what might have been suitable penalties in particular circumstances.⁸⁹ The method would, in theory, give an ethically sound framework to judges, but it would be difficult to enforce in real-life situations when there exists a substantial split between punitive theory and punitive practice.⁹⁰ The right of offenders to contest the veracity and applicability of the data used in sentencing may give rise to due process issues when risk assessment tools are used.⁹¹ When introducing an AI algorithm into the criminal justice system, due process must be the primary consideration.⁹² As a result, the judge is faced with a difficult

⁸³ COMPAS assigns ratings from 1 to 10 based on several factors, and accused with scores of 5-10 (high risk) are more likely to be detained than those with scores of 1-4 (low risk). See, Sam Corbett-Davies and others, ‘A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased against Blacks. It’s Actually Not That Clear’ (*Washington Post*, 17 October 2016) <<https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>> accessed 8 January 2023.

⁸⁴ *State v. Loomis* (n 82).

⁸⁵ Ellora Israni, ‘Algorithmic Due Process: Mistaken Accountability and Attribution in *State v. Loomis*’ (*Harvard Journal of Law & Technology*, 31 August 2017) <<https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1>> accessed 17 February 2023.

⁸⁶ *State v. Loomis* (n82). See also, Katherine Freeman, ‘Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in *State v. Loomis*’ (2016) 18(5) *North Carolina Journal of Law & Technology* 75, 106 <<https://scholarship.law.unc.edu/ncjolt/vol18/iss5/3/>> accessed 11 January 2023 (arguing that the Wisconsin Supreme Court erred in ruling against *Loomis*, and that the courts should not be allowed to use COMPAS in sentencing until an open-source algorithm is developed).

⁸⁷ Mitch Smith, ‘In Wisconsin, a Backlash Against Using Data to Foretell Defendants’ Futures’ (*New York Times*, 22 June 2016) <www.nytimes.com/2016/06/23/us/backlash-in-wisconsin-against-using-data-to-foretell-defendants-futures.html> accessed 8 January 2023. See also, Sandra G Mayson, ‘Bias In, Bias Out’ (2019) 128(8) *Yale Law Journal* 2218, 2251 <www.yalelawjournal.org/article/bias-in-bias-out> accessed 9 February 2023.

⁸⁸ Danielle Kehl and others (n20) 20. See also, Ellora Israni (n85).

⁸⁹ Jesper Ryberg (n19).

⁹⁰ Jesper Ryberg (n19).

⁹¹ John Villasenor and Virginia Foggo, ‘Algorithms and sentencing: What does due process require?’ (*Brookings*, 21 March 2019) <www.brookings.edu/blog/techtank/2019/03/21/algorithms-and-sentencing-what-does-due-process-require/> accessed 17 February 2023.

⁹² John Villasenor and Virginia Foggo (n8) 353.

decision regarding how to use the risk assessment score, and their reaction may greatly be influenced by the theory of punishment that they follow or believe in.⁹³ However, in order to adhere to the specifications of due process, an opportunity must be provided to challenge the classification that an algorithm assigns to an individual, which must include challenging both the data and the design of the algorithm itself.⁹⁴ Moreover, data such as race, or on information that is materially inaccurate,⁹⁵ or in the Indian context, caste or religion, must not be used to sentence accused.

The AI assessment would only be able to provide the judge with an idea of what the sentence range has been in similar cases in the past; the judge would still need to provide sentencing reasons to avoid doing “an exercise in rubber stamping.”⁹⁶ Concerns about the fairness of AI-based decision-making, especially in comparison to conventionally made human decisions, have arisen and are growing.⁹⁷ Algorithms’ self-learning abilities are misunderstood, leading to mistrust. There is a breakdown in accountability between those who govern and those who are governed when decisions are made via incomprehensible processes.⁹⁸ Algorithms can learn on their own, but they can only do this by gathering information about the area in which the system works and putting it into a formula where it can be used.⁹⁹ Importantly, the algorithms are designed by humans who set the decision-making parameters, so they can be altered to fit a narrative.¹⁰⁰ AI-enabled sentencing might be opaque and constantly evolving, posing a threat to a judge’s capacity to make an educated decision.¹⁰¹ Conversely, unlike an algorithm, the

⁹³ Danielle Kehl and others (n20) 14.

⁹⁴ Vincent Chiao, ‘Fairness, Accountability and Transparency: Notes on Algorithmic Decision-Making in Criminal Justice’ (2019) 15 International Journal of Law in Context 126, 134 <<http://dx.doi.org/10.1017/s1744552319000077>> accessed 3 February 2023.

⁹⁵ John Villasenor and Virginia Foggo (n8) 339.

⁹⁶ Vincent Chiao (2018) (n75) 246.

⁹⁷ Md Abdul Malek, ‘Criminal courts’ artificial intelligence: the way it reinforces bias and discrimination’ (2022) 2(1) AI and Ethics 233 <<http://dx.doi.org/10.1007/s43681-022-00137-9>> accessed 6 May 2023.

⁹⁸ Gianna Seglias (n29) 63.

⁹⁹ Iqbal H Sarker, ‘Machine Learning: Algorithms, Real-World Applications and Research Directions’ (2021) 2(160) SN Computer Science 1 <<http://dx.doi.org/10.1007/s42979-021-00592-x>> accessed 4 February 2023 (noting that the efficacy of the AI-made decision depends on the accuracy of the data).

¹⁰⁰ Rebecca Heilweil, ‘Why Algorithms Can Be Racist and Sexist’ (*Vox*, 18 February 2020) <<https://www.vox.com/recode/2020/2/18/21121286/algorithms-bias-discrimination-facial-recognition-transparency>> accessed 7 March 2023; See also, Rishika Pardikar, ‘Social Media Companies Like Instagram Are Censoring Dissent’ (*Jacobin*, 1 June 2021) <<https://jacobin.com/2021/06/censorship-facebook-instagram-twitter-india-palestine-colombia>> accessed 18 February 2023.

¹⁰¹ Jason Tashea, ‘Courts Are Using AI to Sentence Criminals. That Must Stop Now’ (*Wired*, 17 April 2017) <www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/> accessed 8 January 2023; Adam Liptak, ‘Sent to Prison by a Software Program’s Secret Algorithms’ (*The New York Times*, 1 May 2017) <<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-program-secret-algorithms.html>> accessed 11 February 2023; Ed Yong, ‘A Popular Algorithm Is No Better at Predicting Crimes

human mind forms connections on its own rather than just being explicitly programmed. Therefore, the accused should be given the opportunity to challenge the sentencing data or the algorithm itself, on due process grounds.

III. THE PROMISES OF AI IN CRIMINAL SENTENCING

Deploying AI in India's criminal justice system presents an ocean of opportunities, which is plagued with whimsical arrests,¹⁰² staggering pendency,¹⁰³ bizarre sentences,¹⁰⁴ and peculiar bail orders.¹⁰⁵ However, to an extent only, the shift towards application of AI in courts, allowing it to supplement human judgment, could be beneficial.¹⁰⁶ Recidivism is the most crucial factor in sentencing, and "properly designed" algorithms are better than judges at determining it because computers do not have real or unconscious biases.¹⁰⁷ Computers that have been properly programmed will offer predictable and consistent responses to many similar questions.¹⁰⁸ Humans make many evaluative judgments based on unconscious biases and intuitive emotions, unlike machines. However, because algorithms lack complete "cognition," they are only as good as the data and questions they are given.¹⁰⁹ The present approach takes up a lot of time, and reduced expenditure, of both money and time, in the process is the essence

Than Random People' (*The Atlantic*, 17 January 2018) <<https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/>> accessed 9 March 2023.

¹⁰² Shahrukh Alam, 'A solution to the problem of judicial pendency, posed by the law minister – don't mechanically arrest a suspect' (*The Indian Express*, 23 November 2022) <<https://indianexpress.com/article/opinion/columns/judiciary-pending-cases-supreme-court-cji-chandrachud-law-ministry-shahrukh-alam-8282849/>> accessed 16 February 2023.

¹⁰³ For pendency in the Indian criminal justice system, see, Abhinav Sekhri, 'Pendency in the Indian Criminal Process: A Creature of Crisis of Flawed Design?' (2019) 15 *Socio Legal Review* 1 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256609> accessed 6 February 2023.

¹⁰⁴ Megha Satyanarayan Prasad, 'Rahul Gandhi Disqualified As MP After 'Bizarre' Sentence, Says Kapil Sibal' (*NDTV*, 24 March 2023) <<https://www.ndtv.com/india-news/rahul-gandhi-disqualified-as-mp-after-bizarre-sentence-says-kapil-sibal-3887812>> accessed 27 March 2023.

¹⁰⁵ Umang Poddar, 'Serve Cows, Wash Women's Clothes, Tie Rakhis: How Bizarre Bail Conditions Punish without Trial' (*Scroll*, 9 June 2022) <<https://scroll.in/article/1025621/serve-cows-wash-women-s-clothes-tie-rakhis-how-bizarre-bail-conditions-punish-without-trial>> accessed 11 February 2023. See also, Faizan Mustafa, 'Strange and Arbitrary Bail Orders: Are Indian Judges Going Too Far?' (*The Wire*, 28 April 2020) <<https://thewire.in/law/judges-bail-orders>> accessed 3 March 2023.

¹⁰⁶ Matt O'brien And Dake Kang, 'AI in the court: When algorithms rule on jail time' (*Phys.org*, 31 January 2018) <<https://phys.org/news/2018-01-ai-court-algorithms.html>> accessed 8 January 2023. See also, Vyacheslav Polonski, 'AI is convicting criminals and determining jail time, but is it fair?' (*World Economic Forum*, 19 November 2018) <www.weforum.org/agenda/2018/11/algorithms-court-criminals-jail-time-fair/> accessed 8 January 2023.

¹⁰⁷ Mirko Bagaric and others (2020) (n3) 1079-1080.

¹⁰⁸ Neha Saini, 'Explained: What is ChatGPT, how it works and can it replace humans?' (*Live Mint*, 7 February 2023) <www.livemint.com/technology/tech-news/explained-what-is-chatgpt-how-it-works-and-can-it-replace-humans-11670244352183.html> accessed 16 February 2023.

¹⁰⁹ George Siemens and others, 'Human and artificial cognition' (2022) 3 *Computers and Education: Artificial Intelligence* 1 <<http://dx.doi.org/10.1016/j.caeai.2022.100107>> accessed 6 February 2023. See also, Daniel N Cassenti and others, 'Editor's Review and Introduction: Cognition-Inspired Artificial Intelligence' (2022) *Topics in Cognitive Science* 652, 661 <<http://dx.doi.org/10.1111/tops.12622>> accessed 6 February 2023.

of an effective justice system. AI would help make sentencing decisions faster, more consistent, and easier to predict.¹¹⁰ The transition from traditional risk assessment methods to algorithmic solutions may result in a more standardised approach.¹¹¹ In the future, risk evaluations employed in criminal punishment will most likely increasingly rely on AI.¹¹²

AI is used in many industries to streamline time-consuming tasks like reviewing applications, determining creditworthiness, and analysing images.¹¹³ The fact that humans are not fully rational beings but have a potential to be rational is well known,¹¹⁴ but this tendency of irrationality may make one a better decision maker.¹¹⁵ Sometimes human being's lack of rationality is due to conscious or unconscious biases, which is when the problem starts.¹¹⁶ Human beings might be biased towards one group before awarding sentences because of their belief or morality compass, however, AI does not have that tendency, and thus, AI might bring lesser arbitrary decisions than human beings. Aside from quantitatively uneven penalties, there are also drawbacks to having extensive sentencing discretion. However, AI based sentencing may mitigate such problems, as it may bring fairer sentences,¹¹⁷ instead of putting this all on judicial discretion.¹¹⁸

¹¹⁰ Dan Hunter and others (n25) 783.

¹¹¹ Alicia Solow-Niederman and others, 'The Institutional Life of Algorithmic Risk Assessment' (2019) 34(3) Berkeley Technology Law Journal 705, 713 <<https://lawcat.berkeley.edu/record/1137216>> accessed 9 February 2023.

¹¹² John Villasenor and Virginia Foggo (n8) 311.

¹¹³ Jonathan Barrett and Stephanie Convery, 'Robot Recruiters: Can Bias Be Banished from AI Hiring?' (*The Guardian*, 26 March 2023) <<https://www.theguardian.com/technology/2023/mar/27/robot-recruiters-can-bias-be-banished-from-ai-recruitment-hiring-artificial-intelligence>> accessed 28 March 2023. See also, Mahtab Fatima, '50+ Use Cases & Application Of Artificial Intelligence' (*Tezeract*, 1 February 2023) <<https://tezeract.ai/use-cases-and-applications-of-artificial-intelligence-in-our-daily-life/>> accessed 17 February 2023.

¹¹⁴ Andrea Kern, 'Human Life, Rationality and Education' (2020) 54(2) Journal of Philosophy of Education 268 <<http://dx.doi.org/10.1111/1467-9752.12412>> accessed 18 February 2023.

¹¹⁵ Olivia Goldhill, 'Humans are born irrational, and that has made us better decision-makers' (*Quartz*, 4 March 2017) <<https://qz.com/922924/humans-werent-designed-to-be-rational-and-we-are-better-thinkers-for-it/>> accessed 17 February 2023.

¹¹⁶ Michael E Donohue (n18) 676.

¹¹⁷ Mirko Bagaric and others (2022) (n19) 137-138; Research News, 'Using Algorithms to Determine Sentencing May Reduce Length of Prison Sentences, Increase Use of Evidence-Based Rehabilitative Programs' (*Vanderbilt University*, 28 July 2021) <<https://news.vanderbilt.edu/2021/07/28/using-algorithms-to-determine-sentencing-may-reduce-length-of-prison-sentences-increase-use-of-evidence-based-rehabilitative-programs/>> accessed 7 January 2023; John Monahan, 'Risk Assessment in Sentencing' in Erik Luna (ed.), *Reforming Criminal Justice: Punishment, Incarceration, and Release*, vol 4 (The Academy for Justice 2017) <https://law.asu.edu/sites/default/files/pdf/academy_for_justice/Reforming-Criminal-Justice_Vol_4.pdf> accessed 7 January 2023 (arguing that a form of risk assessment that is limited by morals could reduce the financial and human costs of mass incarceration).

¹¹⁸ John Monahan and others, 'Judicial Appraisals of Risk Assessment in Sentencing' (2018) 36 Behavioural Sciences & the Law 565 <<http://dx.doi.org/10.1002/bsl.2380>> accessed 22 February 2023 (most judges supported the principle of incorporating an offender's recidivism risk into sentencing decisions, however, they opposed a policy that compelled them to justify not intervening with "low-risk" offenders in writing).

Sentencing is a post-conviction subject, which brings a clear opportunity for AI to play a role in informing decision-making.¹¹⁹ AI is more precise than judges, which is slightly more successful in predicting recidivism than human beings who are presented with only the accused's age, sex, and criminal history, reach a verdict.¹²⁰ They also self-update as fresh judgments are fed into the system, and therefore, the use of AI in the sentencing process may have several significant benefits, including giving the judges instant access to the gravity of previous sentences for individual crimes. However, bias, flawed data, or false assumptions must not have influenced risk scores, if AI is to support sentencing equity.¹²¹ Crime data used to train predictive policing algorithms is often biased and unrepresentative of true crime patterns, as it reflects *inter alia* institutional biases in law enforcement practices, differential reporting rates across communities – which warrant transparency and accountability to be in place.¹²² Although, it cannot objectively be assessed what, if any, place does AI have in sentencing until its guiding principles are carefully defined.¹²³

The application of AI to risk assessment has been studied in various fields, including cryptocurrency investments,¹²⁴ occupational safety and health,¹²⁵ suicide prevention,¹²⁶ supply chain risk management,¹²⁷ and evidence management in risk assessment.¹²⁸ AI can improve the accuracy and efficiency of risk assessment by identifying the most crucial risk assessment

¹¹⁹ Mirko Bagaric and others (2020) (n3) 1038.

¹²⁰ Julia Dressel and Hany Farid, 'The accuracy, fairness, and limits of predicting recidivism' (2018) 4(1) Science Advances 5580 <<http://dx.doi.org/10.1126/sciadv.aao5580>> accessed 8 January 2023. See also, Mirko Bagaric and others (2022) (n19) 131 (opposing the criticism of algorithmic sentencing and noting that AI tools are more accurate in predicting recidivism).

¹²¹ John Villasenor and Virginia Foggo (n8) 354.

¹²² William S Isaac, 'Hope, Hype, and Fear: The Promise and Potential Pitfalls of Artificial Intelligence in Criminal Justice' (2018) 15 Ohio State Journal of Criminal Law 543 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3145308> accessed 18 March 2023.

¹²³ Alberto De Diego Carreras (n19) 38.

¹²⁴ Lin Li, 'Investigating Risk Assessment in Post-Pandemic Household Cryptocurrency Investments: An Explainable Machine Learning Approach' (2023) Journal of Asset Management <<http://dx.doi.org/10.1057/s41260-022-00302-z>> accessed 20 April 2023.

¹²⁵ André Steimers and Moritz Schneider, 'Sources of Risk of AI Systems' (2022) 19 International Journal of Environmental Research and Public Health 3641 <<http://dx.doi.org/10.3390/ijerph19063641>> accessed 17 March 2023.

¹²⁶ Alban Lejeune and others, 'Artificial Intelligence and Suicide Prevention: A Systematic Review' (2022) 65(1) European Psychiatry 1 <<http://dx.doi.org/10.1192/j.eurpsy.2022.8>> accessed 7 January 2023.

¹²⁷ Paul Souma Kanti and others, 'Artificial Intelligence Adoption in Supply Chain Risk Management: Scale Development and Validation' (2022) 12 Ho Chi Minh City Open University Journal of Science – Economics and Business Administration 15 <<http://dx.doi.org/10.46223/hcmcoujs.econ.en.12.2.2142.2022>> accessed 2 February 2023.

¹²⁸ Bersani C. and others, 'Roadmap for Actions on Artificial Intelligence for Evidence Management in Risk Assessment' (2022) 19(5) EFSA Supporting Publications 1 <<http://dx.doi.org/10.2903/sp.efsa.2022.en-7339>> accessed 2 February 2023.

determinants,¹²⁹ estimating a person's likelihood of suicide or spotting vulnerable people in a particular group,¹³⁰ and proactively enabling supply chain risk identification, evaluation, and mitigation.¹³¹ The inclusion of aggravating and mitigating circumstances can help provide a more accurate and comprehensive risk assessment. To incorporate aggravating and mitigating circumstances, the algorithm could be designed to consider many factors that could affect the level of risk associated with a particular situation. To do so, AI algorithms can use various methods, such as fuzzy set theory,¹³² certainty factor,¹³³ and explainable ML algorithms.¹³⁴ The ethical risks of AI decision-making include algorithmic discrimination, data bias, and unclear accountability, which can be mitigated by risk governance elements – like ethical AI norms and legal regulation.¹³⁵

If AI can determine what metrics human judges rely on, such as the sentencing objectives, penalty, and aggravating and mitigating factors, it would make sentence decisions more transparent, predictable, and uniform.¹³⁶ Risk assessment tools are less prone to unconscious or implicit biases, but they are not immune to racial/religious/ethnic biases, and AI can help tailor outcomes by screening out empirically or morally irrelevant factors.¹³⁷ The algorithm should be trained on large datasets to learn which factors are most predictive of risk and how to weight those factors in the final risk score. However, it is crucial to ensure that the algorithm is transparent, fair, and subject to ongoing evaluation to ensure that it is providing accurate and unbiased assessments. In India, stereotypical arguments, and personal traits such as caste, sex, and ethnicity, often creep into judicial decisions,¹³⁸ and by providing these algorithms with

¹²⁹ Lin Li (n124).

¹³⁰ Alban Lejeune and others (n126).

¹³¹ Paul Souma Kanti and others (n127).

¹³² Ekaterina Orlova, 'Fuzzy Model for Support Investment Decisions under Risk' (2018) Collection of selected papers of the IV International Conference on Information Technology and Nanotechnology <<http://dx.doi.org/10.18287/1613-0073-2018-2212-1-9>> accessed 20 January 2023.

¹³³ David Heckerman, 'The Certainty-Factor Model' (*Microsoft Research*, 5 August 1992) <<https://www.microsoft.com/en-us/research/publication/certainty-factor-model/>> accessed 12 March 2023 (describing the certainty factor model, a probabilistic framework for representing uncertainty in rule-based expert systems, assigning a certainty factor to each rule based on the available evidence and using these factors to calculate the overall degree of certainty or uncertainty in the system's output).

¹³⁴ Lin Li (n124).

¹³⁵ Hongjun Guan and others, 'Ethical Risk Factors and Mechanisms in Artificial Intelligence Decision Making' (2022) 12 *Behavioural Sciences* 343 <<http://dx.doi.org/10.3390/bs12090343>> accessed 7 March 2023.

¹³⁶ Nigel Stobbs and others (n73) 3.

¹³⁷ Vincent Chiao (2019) (n94) 131.

¹³⁸ Rintu Mariam Biju, 'Don't Mention Caste Of Parties In Cause Titles Of Judgements: Supreme Court To Trial Courts' (*LiveLaw*, 14 March 2023) <<https://www.livelaw.in/top-stories/caste-parties-shouldnt-mentioned-cause-titles-judgements-supreme-court-223743>> accessed 16 March 2023. See also, Sidharth Luthra and Ketaki Goswami, 'Courts Must Dispel Gender Stereotypes' (*Hindustan Times*, 16 April 2021)

skewed information, they will be more susceptible to biases.¹³⁹ It is not the case that human judges are always error free in their decisions, but in fact, India and even the US witnesses a massive wrongful convictions.¹⁴⁰ Due to its efficiency and ability to at least appear impartial, AI adjudication will probably contribute to the shift toward “codified justice,” an adjudicatory paradigm that favours standardisation over discretion.¹⁴¹

IV. TESTING THE PERILS AGAINST THE PROMISES

The Indian criminal justice system restricts individual freedom and errors can have severe negative effects, necessitating the decisions to be open, accurate, and efficient.¹⁴² The judge must take into account both the mitigating and aggravating circumstances and translate them into a specific punishment that is neither excessively harsh nor excessively lenient and is proportionate.¹⁴³ Therefore, in spite of several promises, the use of AI in sentencing raises several issues that must be addressed,¹⁴⁴ as it can lead to unpredictability.¹⁴⁵ However, to the extent consistency in sentencing is desired, the emphasis is on the consistent application of the appropriate sentencing principles rather than the lack of numerical uniformity in sentencing.¹⁴⁶ The single most crucial element in determining a sentence is the level of risk that the person will reoffend.¹⁴⁷ Apropos this prediction, Sandra Mayson wrote that “[t]he deep problem is the

<<https://www.hindustantimes.com/opinion/courts-must-dispel-gender-stereotypes-101618575195593.html>> accessed 12 February 2023.

¹³⁹ Aleš Završnik, ‘Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings’ (2019) 18(5) *European Journal of Criminology* 623, 633 <<http://dx.doi.org/10.1177/1477370819876762>> accessed 4 February 2023.

¹⁴⁰ Kent Roach, ‘Wrongful Prosecutions and Convictions’ (*Law School Policy Review*, 2 January 2021) <<https://lawschoolpolicyreview.com/2021/01/02/wrongful-prosecutions-and-convictions/>> accessed 27 January 2023; Hassan Kanu, ‘Rising Number of False Convictions Shows Stark Racial Patterns’ (*Reuters*, 27 September 2022) <<https://www.reuters.com/legal/government/rising-number-false-convictions-shows-stark-racial-patterns-2022-09-27/>> accessed 29 January 2023.

¹⁴¹ Richard M. Re and Alicia Solow-Niederman, ‘Developing Artificially Intelligent Justice’ (2019) 22 *Stanford Technology Law Review* 242 <https://law.stanford.edu/wp-content/uploads/2019/08/Re-Solow-Niederman_20190808.pdf> accessed 12 December 2022 (arguing that there are many potential problems with using AI for judging, including biasedness, opacity and unfairness, relating to its tendency to reduce discretionary moral judgment or equitable justice, and concluding by calling for a careful and thoughtful approach to the advancement of AI adjudication in order to ensure its use in a way that is consistent with the values of the legal system).

¹⁴² Mirko Bagaric and others (2022) (n 19) 147.

¹⁴³ Vincent Chiao (2018) (n75) 238.

¹⁴⁴ Christopher Markou, ‘Why using AI to sentence criminals is a dangerous idea’ (*The Conversation*, 16 May 2017) <<https://theconversation.com/why-using-ai-to-sentence-criminals-is-a-dangerous-idea-77734>> accessed 8 January 2023.

¹⁴⁵ Nigel Stobbs and others (n73) 17.

¹⁴⁶ Nigel Stobbs and others (n73) 17.

¹⁴⁷ Mirko Bagaric and others (2020) (n3) 1037 (arguing that, “many problems in the criminal justice system would be solved if we could accurately determine which offenders would commit offenses in the future”).

nature of prediction itself. All predictions look to the past to make guesses about future events. In a racially stratified world, any method of prediction will project the inequalities of the past into the future. This is as true of the subjective prediction that has long pervaded criminal justice as it is of the algorithmic tools now replacing it.”¹⁴⁸

Due process of law is essential to prevent innocent people from being wrongfully convicted of crimes or arbitrarily punished after being found guilty, given the imbalance of power between the State and the citizen. The possibility of arbitrary punishment increases in the absence of specific justifications for how and why any mitigating or aggravating factors affect the sentence that is given.¹⁴⁹ It has been noted that different verdicts in like cases are “likely to lead to an erosion of public confidence in the integrity of the administration of justice.”¹⁵⁰ If the procedure was unjust and unclear, punished individuals are less likely to wilfully obey sentencing order.¹⁵¹ Where there is no sentencing policy in place and everything rests on judicial discretion, deploying AI might raise more problems than solve existing ones, if certain risks are not mitigated.¹⁵² AI-based sentencing may reduce the likelihood of human error, but it may also introduce other types of errors;¹⁵³ and such tools may not be perfect, yet.¹⁵⁴

A. Algorithmic Inefficiency and Bias

Sentencing algorithm analyses previous crime statistics and recognizes patterns linked to crime.¹⁵⁵ The subconscious prejudice has a considerable influence on sentencing,¹⁵⁶ while personal traits like ethnicity and gender can impact sentence results.¹⁵⁷ Even the most

¹⁴⁸ Sandra G Mayson (87).

¹⁴⁹ Nigel Stobbs and others (n73) 16.

¹⁵⁰ *Rees v. The Queen* [2012] NSWCCA 47, [50], cited in Nigel Stobbs and others (n73) 19.

¹⁵¹ Nigel Stobbs and others (n73) 16. See also, Dan Hunter and others (n25) 785.

¹⁵² Alicia Solow-Niederman and others (n111) 715-716 (noting that “concerns about the use of such tools include, among others, the risk of data inputs that reproduce and reinforce racial inequities in the criminal justice system overall; the failure to provide adequate procedural safeguards, including individualized, adversarial hearings for all accused; and the lack of transparency or access to the data or algorithms used by proprietary instruments”).

¹⁵³ Katherine Freeman (n86) 98.

¹⁵⁴ Seena Fazel and others, ‘The Predictive Performance of Criminal Risk Assessment Tools Used at Sentencing: Systematic Review of Validation Studies’ (2022) 81 *Journal of Criminal Justice* 1 <<http://dx.doi.org/10.1016/j.jcrimjus.2022.101902>> accessed 12 February 2023.

¹⁵⁵ Carlow University, ‘Artificial Intelligence in Criminal Justice: How AI Impacts Pretrial Risk Assessment’ (*Carlow*, 27 July 2021) <<https://blog.carlow.edu/2021/07/27/artificial-intelligence-in-criminal-justice/>> accessed 11 January 2023.

¹⁵⁶ For instance, it has been demonstrated that good looking criminals face less severe punishments through AI than other not-so-good appearing accused. See, Nigel Stobbs and others (n73) 22.

¹⁵⁷ Nigel Stobbs and others (n73) 22.

sophisticated AI algorithm can get biased towards races and genders.¹⁵⁸ Data bias can appear in many different ways and lead to discrimination, and AI bias occurs when an ML model's output results in bias against specific individuals or groups.¹⁵⁹ Bias in AI develops based on the data it is fed,¹⁶⁰ when the data does not contain proportional representation of different groups.¹⁶¹ Even though the law does not explicitly target or discriminate against anyone, research has shown that sentencing procedures take certain actions that are discriminatory.¹⁶² Users of AI in the justice system should be wary of self-reinforcing cycles that could lead to an algorithm becoming more biased with repeated use.¹⁶³ AI has the potential to amplify human biases and, through its judgments, contribute to the pervasive false beliefs and fears that promote incarceration.¹⁶⁴ It is possible that the people or communities who have historically been the focus of law enforcement's attention will receive less favourable algorithmic scores.¹⁶⁵

The outcomes may be biased if manifest prejudice affects the data which the algorithm is trained on.¹⁶⁶ For instance, based on 10 years of data, Amazon's resume-screening algorithm favoured men over women;¹⁶⁷ algorithms have been found to absorb obsolete gender norms, such as "doctors" identified as male and "receptionists" identified as female;¹⁶⁸ AI models fed

¹⁵⁸ Matthew Hutson, 'Even artificial intelligence can acquire biases against race and gender' (*Science*, 13 April 2017) <www.science.org/content/article/even-artificial-intelligence-can-acquire-biases-against-race-and-gender> accessed 8 January 2023; Rebecca Heilweil, (n100); See also, Varsha PS, 'How Can We Manage Biases in Artificial Intelligence Systems – A Systematic Literature Review' (2023) 3(1) International Journal of Information Management Data Insights 1 <<http://dx.doi.org/10.1016/j.ijime.2023.100165>> accessed 6 April 2023 (asserting that "AI biases and vulnerabilities experienced by people across industries lead to gender biases and racial discrimination").

¹⁵⁹ Lorenzo Belenguer (n2) 773. See also, European Union ("EU") Agency for Fundamental Rights, 'Bias in Algorithms – Artificial Intelligence and Discrimination' (Publications Office of the EU 2022) <<https://data.europa.eu/doi/10.2811/25847>> accessed 20 January 2023 (discussing the application of AI to offensive speech detection and predictive policing while highlighting how algorithmic bias can grow over time and potentially result in discrimination).

¹⁶⁰ Keith Kirkpatrick, 'It's Not the Algorithm, It's the Data' (2017) 60 Communications of the ACM 21 <<http://dx.doi.org/10.1145/3022181>> accessed 17 January 2023.

¹⁶¹ For instance, Google was forced to deliberately seek out more training examples of different shoe kinds, such as high heels and clogs, to compensate for representation gaps in an experiment. See, Reena Jana and Josh Lovejoy, 'Exploring and Visualizing an Open Global Dataset' (*Google AI Blog*, 25 August 2017) <<https://ai.googleblog.com/2017/08/exploring-and-visualizing-open-global.html>> accessed 8 January 2023.

¹⁶² Mirko Bagaric and others (2022) (n19) 100.

¹⁶³ Carlow University (n155).

¹⁶⁴ Carlow University (n155).

¹⁶⁵ Carlow University (n155).

¹⁶⁶ Sarah Valentine (n11) 392.

¹⁶⁷ Jeffrey Dastin, 'Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women' (*Reuters*, 11 October 2018) <<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>> accessed 22 January 2023. See also, Ignacio N Cofone, 'Algorithmic Discrimination Is an Information Problem' (2019) 70(6) *Hastings Law Journal* 1389, 1397 <https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/1/> accessed 2 February 2023.

¹⁶⁸ Vyacheslav Polonski (n106).

with images of previous elected public office holder have been shown to predict wholly male contenders as the likely winner of the elections.¹⁶⁹ Thus, if AI is used to make conclusions, training data, variable selection, or output interpretation may reveal unintentional biases;¹⁷⁰ for instance, facial recognition technology (“FRT”) has trouble recognising dark-skinned faces if it is trained on light-skinned faces, and is potentially discriminatory towards minority groups.¹⁷¹ The widely acclaimed inquiry by ProPublica revealed that AI algorithms amplify racial bias.¹⁷² Nearly two thirds of the time, COMPAS’s predictions of recidivism were accurate, the study found, although both false negatives and false positives tended to favour people of colour.¹⁷³ Nearly twice as many whites were found among low-risk offenders who reoffended, while over four times as many blacks were found among high-risk offenders who did not.¹⁷⁴ If the underlying data is prejudiced, structural inequities and unjust biases are at risk of being reinforced.¹⁷⁵ All predictions will reflect the algorithm’s inherent biases, to the extent that they are based on prior decisions that were discriminatory toward some members of society.¹⁷⁶ In addition, the cyclical nature of the system’s predictions influencing information used to make changes to the same system can lead to the phenomenon of a “self-fulfilling prophecy.”¹⁷⁷

¹⁶⁹ Vyacheslav Polonski (n106).

¹⁷⁰ Simon Chesterman, ‘Through a Glass, Darkly: Artificial Intelligence and the Problem of Opacity’ (2021) 69(2) *The American Journal of Comparative Law* 271, 281 <<http://dx.doi.org/10.1093/ajcl/avab012>> accessed 16 February 2023.

¹⁷¹ Ameen Jauhar, ‘Indian Law Enforcement’s Ongoing Usage of Automated Facial Recognition Technology – Ethical Risks and Legal Challenges’ (Vidhi Centre for Legal Policy 2021) <<https://vidhilegalpolicy.in/research/indian-law-enforcements-ongoing-usage-of-automated-facial-recognition-technology-ethical-risks-and-legal-challenges/>> accessed 20 February 2023 (highlighting the key risks associated with automated FRT in the criminal justice system); Jai Vipra, ‘The Use of Facial Recognition Technology for Policing in Delhi’ (Vidhi Centre for Legal Policy 2021) <<https://vidhilegalpolicy.in/research/the-use-of-facial-recognition-technology-for-policing-in-delhi/>> accessed 20 February 2023 (noting that the use of FRT in policing can discriminately target Muslims, women, and marginalised castes, among others).

¹⁷² Julia Angwin and others, ‘Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And It’s Biased Against Blacks’ (*ProPublica*, 23 May 2016) <www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 8 January 2023.

¹⁷³ *ibid.*

¹⁷⁴ *ibid.*

¹⁷⁵ Sidney Fussell, “‘Black Data’ Is the Reason Why Smart Policing Is Still Incredibly Biased’ (*Gizmodo*, 11 October 2017) <<https://gizmodo.com/black-data-is-the-reason-why-smart-policing-is-still-1819288923>> accessed 8 January 2023. See also, Genevieve Smith and Ishita Rustagi (n37) 37 (noting that “AI systems can infringe on civil liberties, while also reinforcing existing prejudices”).

¹⁷⁶ Jesper Ryberg (n19).

¹⁷⁷ EU Agency for Fundamental Rights (n159) 29-30 (pointing out that, “For example, if the AI system ‘detects’ more crime in district A and decides to send more patrols there, more crimes will be recorded in district A, and the corresponding data will be fed back into the system, reinforcing the system’s ‘belief’ that there is more crime in district A”).

COMPAS exacerbated the concerns about AI in sentencing, but the underlying problem is a logical fallacy.¹⁷⁸ COMPAS calculates the “violent recidivism risk score” using a formula that takes into account factors such as age, level of education, and criminal record.¹⁷⁹ The use of aggregate data for analysing individuals, which is the foundation of COMPAS-like models, creates stereotypes.¹⁸⁰ AI systems are being used by law enforcement agencies across India,¹⁸¹ however, from secretive procurement to non-disclosable deployment, the use is notoriously opaque.¹⁸² Predictive policing technology has been used by Delhi Police, based on the existing available data, that can potentially transfer the existing biases and errors into the technology.¹⁸³ Delhi Police’s data collection procedure disproportionately affect historically marginalised and vulnerable groups,¹⁸⁴ and the system’s discriminatory and arbitrary practices reflect problematic social norms.¹⁸⁵ Additionally, several forms of biases perpetuate in the data – like discrimination against minority groups and representational biases against underrepresented groups.¹⁸⁶ The systematic discrimination works against vulnerable communities, resulting in their over policing,¹⁸⁷ and causes a reiteration of discrimination within the institution grappling with discrimination based on race, gender, and class.¹⁸⁸ If AI tools are trained using such data, it will make the same mistakes that humans have made in the past.¹⁸⁹ The criminal justice

¹⁷⁸ This is a logical fallacy that has been around for a long time: *cum hoc ergo propter hoc* (“with this, therefore because of this”). See, Simon Chesterman (n170) 282.

¹⁷⁹ *ibid.*

¹⁸⁰ *ibid* 283.

¹⁸¹ Deepak Jha and Pradeep Kushwaha ‘AI Fortifying Indian Cities: Making a Haven for Citizens’ (*India AI*, 6 March 2023) <<https://indiaai.gov.in/article/ai-fortifying-indian-cities-making-a-haven-for-citizens>> accessed 10 March 2023.

¹⁸² See, Vidushi Marda and Shivangi Narayan, ‘Data in New Delhi’s Predictive Policing System’ (2020) Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency <<http://dx.doi.org/10.1145/3351095.3372865>> accessed 3 March 2023.

¹⁸³ For real-time mapping of spatial hotspots, analysis of criminal behaviour patterns, and profiling of potential suspects, the Crime Mapping, Analytics, and Predictive System (CMAPS) was introduced. To identify crime hotspots, data from the Dial 100 call centre would be used to plot call locations. See, Vidushi Marda and Shivangi Narayan, (n183) (noting that, “It is intended that CMAPS, with automated mapping will correct these inconsistencies, but having an inaccurate benchmark or a historical dataset, against which performance will be evaluated casts doubt over CMAPS being a panacea”).

¹⁸⁴ There are several major errors in the data collection measures, like geographical inconsistencies, database limitations, and knowledge and incentive problems. See, Vidushi Marda and Shivangi Narayan (2020) (n 182).

¹⁸⁵ Vidushi Marda and Shivangi Narayan (2020) (n182). See also, Vidushi Marda and Shivangi Narayan, ‘On the Importance of Ethnographic Methods in AI Research’ (2021) 3 *Nature Machine Intelligence* 187, 188 <<http://dx.doi.org/10.1038/s42256-021-00323-0>> accessed 3 March 2023 (noting that it could appear, based on numbers, that people in ghettos and slum settlements overreport crime, which led to the institutional assumption that people in those areas would exaggerate and lie about what happened when they called to complain).

¹⁸⁶ Vidushi Marda and Shivangi Narayan (2020) (n182).

¹⁸⁷ When police employ predictive policing algorithms to prioritise patrols, they are more likely to over-police high-crime regions. See, Aleš Završnik (2019) (n139) 635.

¹⁸⁸ Vidushi Marda and Shivangi Narayan (2020) (n182).

¹⁸⁹ Karen Hao, ‘AI Is Sending People to Jail—and Getting It Wrong’ (*MIT Technology Review*, 21 January 2019) <<https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/>> accessed 17 January 2023.

system has a history of horrifyingly mistreating marginalised populations because of human decision-making,¹⁹⁰ and if such data is used to generate sentencing scores, it will reproduce past injustices and biases. To create algorithms that are both effective and equitable, it is essential to check that they do not in any way favour or disadvantage any group or community.¹⁹¹ Moreover, merely removing bias from AI is not enough, instead the developers must constantly strive to make it more equitable and fairer.¹⁹²

B. Wrong Sentences and the Problem of Opacity

Judges have various contradicting directives from the legislature and society, making sentencing difficult.¹⁹³ The sentence must not only exact proportional retribution, but also dissuade the criminal from repeating the offence, and keep the society from witnessing that offence again.¹⁹⁴ The Indian criminal justice has been focused to ‘reform’ the criminal.¹⁹⁵ The penalty should be long enough to safeguard the society from the same person committing another offence, and the sentence should allow the convict to reintegrate into society after punishment.¹⁹⁶ In AI-based sentencing, however, there is a reliance on a single punishment philosophy.¹⁹⁷ Throughout human history, there has been a consistent pattern of compromised and frequently incorrect decisions made in the absence of a scientific methodology, thus, judges’ intuitive and unstructured recidivism assessments of the future criminal propensities are frequently inaccurate.¹⁹⁸ Usually, sentencing by a judge leads to opaque and inconsistent

¹⁹⁰ Mirko Bagaric and others (2022) (n19) 97-98.

¹⁹¹ This is possible, but it takes a deep understanding of how sentencing factors, like a person’s criminal history and level of education, can be used to stand in for things like race and gender. See, Mirko Bagaric and others (2020) (n3) 1040. See also, Saar Alon-Barkat and Madalina Busuioc, ‘Human–AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice’ (2023) 33(1) *Journal of Public Administration Research and Theory* 153 <<http://dx.doi.org/10.1093/jopart/muac007>> accessed 17 February 2023 (noting that in contrast to human decision making, which can be biased and discriminatory, algorithms offer the “promise of neutrality”).

¹⁹² Chelsea Barabas, ‘Beyond Bias: Re-Imagining the Terms of ‘Ethical AI’ in Criminal Law’ (2020) 12(2) *Georgetown Journal of Law & Modern Critical Race Perspectives* 83 <<https://www.law.georgetown.edu/mcrp-journal/in-print/volume-12-issue-2-fall-2020/beyond-bias-re-imagining-the-terms-of-ethical-ai-in-criminal-law/>> accessed 10 January 2023.

¹⁹³ Michael E Donohue (n18) 658.

¹⁹⁴ Michael E Donohue (n18) 658.

¹⁹⁵ *Bachan Singh v. State of Punjab*, (1982) 3 SCC 24.

¹⁹⁶ Michael E Donohue (n18) 658.

¹⁹⁷ Michael E Donohue (n18) 665-66. For example, COMPAS score is presented as bar charts, which must be accepted, without which the purpose or the working of the algorithm would itself be frustrated.

¹⁹⁸ Mirko Bagaric and others (2020) (n3) 1038.

decisions, being based on a judge's personal predisposition, thus some groups are sentenced harsher than others,¹⁹⁹ which might be reproduced in the sentencing algorithm.

AI is prone to mistakes which may lead to wrong sentences being imposed, for instance, it can be manipulated in negative ways – to create deep fakes,²⁰⁰ fake videos,²⁰¹ or fake photographs²⁰² of events that never occurred in real life. In fact, instead of reducing human error, automated decision support systems frequently open the door for new types of errors.²⁰³ Any statistical model and its interpretation is susceptible to bias, particularly if the interpreter attaches their own preconceived notions and values to it.²⁰⁴ AI is typically used to gather data on various aspects of an offender's behaviour and correlate it with the likelihood of recidivism. As in the case of Loomis, COMPAS was used, which forecasted recidivism based on factors such as the accused's criminal past, education level, and Loomis was sentenced to 6 years.

Furthermore, the judges who relied on COMPAS were unaware of how the scores were calculated.²⁰⁵ This is troublesome since the data used to train AI is only as reliable as the data used to train them.²⁰⁶ While AI has the potential to outperform conventional methods, the lack of openness inherent in its black box approaches makes it difficult to evaluate its efficacy, which not only hides the decision-making process but also raises credibility concerns for users.²⁰⁷ Sentencing algorithms are difficult to evaluate and audit for accuracy and bias due to their opacity.²⁰⁸ Simon Chesterman argues that “judicial decisions are the clearest example of

¹⁹⁹ Mirko Bagaric and others (2020) (n3) 1064. See also, Mirko Bagaric and others (2022) (n19) 103 (arguing that harder punishments for some groups cannot be separated from the fact that judges' freedom of choice in sentencing always leads to sentences based on their own personal preferences).

²⁰⁰ Bloomberg Quicktake: Originals, 'It's Getting Harder to Spot a Deep Fake Video' (*YouTube*, 27 September 2018) <www.youtube.com/watch?v=gLoI9hAX9dw> accessed 8 January 2023.

²⁰¹ Aja Romano, 'Jordan Peele's simulated Obama PSA is a double-edged warning against fake news' (*Vox*, 18 April 2018) <www.vox.com/2018/4/18/17252410/jordan-peeel-obama-deepfake-buzzfeed> accessed 8 January 2023.

²⁰² Kimberly Powell, 'NVIDIA Researchers Showcase Major Advances in Deep Learning at NIPS' (*NVIDIA Blog*, 3 December 2017) <<https://blogs.nvidia.com/blog/2017/12/03/nvidia-research-nips/>> accessed 8 January 2023.

²⁰³ Linda J Skitka and others, 'Does Automation Bias Decision-Making?' (1999) 51 *International Journal of Human-Computer Studies* 991, 992 <<http://dx.doi.org/10.1006/ijhc.1999.0252>> accessed 5 January 2023.

²⁰⁴ Aleš Završnik (2019) (n139) 632.

²⁰⁵ Julia Angwin and others (n172).

²⁰⁶ Rajen Sheth, 'Steering the right course for AI' (*Google Cloud Blog*, 5 November 2018) <<https://cloud.google.com/blog/products/ai-machine-learning/steering-the-right-course-for-ai>> accessed 11 January 2023.

²⁰⁷ Simon Chesterman (n170) 283.

²⁰⁸ Danielle Kehl and others (n20) 28.

an area in which the use of opaque AI systems should be limited,” as AI has introduced efficiency and optimization to numerous decision-making procedures, but at a price.²⁰⁹

Technology has indeed brought many benefits, but that does not solidify technology as an improvement.²¹⁰ The problem with AI-based sentencing is that when it passes a wrong sentence, it possibly cannot be rectified without questioning the technology itself. Once the authenticity of the technology comes into question, the functions performed will naturally be considered inappropriate. It is important that the sentence passed is justifiable and authentic, lest it be on judicial discretion, as at least the right to appeal will rest with the convict. If a biased dataset is used in recidivism risk calculation, it could wrongly imply that a poor or less-represented accused is more likely to commit crimes in the future, leading to a higher risk score for that person,²¹¹ even though the prospect of further harm poses little relevance to the importance of meting out *just deserts*.²¹²

C. Diminishing the Scope for Individualised Justice

The human judge is more capable of dealing with the humans that pass through the court than the mechanical machine is, because the machine may not be trained to manage unique circumstances.²¹³ Different facts and circumstances require different sentences, which is the very basis of judging in common law countries, may not be the tendency of AI, which would diminish the scope for “individualised justice.”²¹⁴ Judges are tasked with balancing society’s demands one accused at a time, but AI technologies would force them to concentrate on a single assessment of the demands of justice in every case. They may observe aggravating and mitigating elements in front of them without being able to consider them, jeopardising their capacity to administer proportionate punishment.

²⁰⁹ Simon Chesterman (n170) 293.

²¹⁰ Christopher Markou, ‘Why using AI to sentence criminals is a dangerous idea’ (*The Conversation*, 16 May 2017) <<https://theconversation.com/why-using-ai-to-sentence-criminals-is-a-dangerous-idea-77734>> accessed 8 January 2023.

²¹¹ Danielle Kehl and others (n20) 29. See also, Genevieve Smith and Ishita Rustagi (n37) 37 (noting that biased AI “can pose a detriment to the safety of individuals”).

²¹² Danielle Kehl and others (n20) 14.

²¹³ Michael E Donohue (n18) 672.

²¹⁴ *Elias v The Queen*, [2013] HCA 31 [27], cited in Nigel Stobbs and others (n73) 4.

When an algorithm classifies a person based on data that may be inaccurate, the person is denied some procedural protections, such as the opportunity to challenge the evidence.²¹⁵ The concept of legal fairness incorporates the idea that everyone should be afforded a fair trial by the law.²¹⁶ Algorithms will always produce “cruder” recommendations than human judgment, because the latter is adaptable and well-suited to decision-making – being suited to the realities of different cases.²¹⁷ Research suggests that “machine-based decisions are too rigid in that they pre-specify a limited number of outcome variables to optimize.”²¹⁸ Human intervention can act as a safeguard against algorithmic failures.²¹⁹ Thus, the purpose of algorithmic outputs should not be to make decisions autonomously, but rather to aid human decision making, where the human decision-maker remains an integral part of the process.²²⁰

In the absence of any sentencing policy or guideline, the SC has designed a framework of sentencing, however, this individualisation can create severe disparity in the quantum of punishment in similar cases.²²¹ The idea of aggravating and mitigating factors allows a judge to administer punishments in a manner that is suited to the specific offence, which enables the process to be individualised.²²² The troubling concern is that, if judicial discretion is replaced by AI, convicts may receive harsher sentences because AI may not be unable to take into account the accused’s overall background or the victim’s actual suffering; even sentencing guidelines are not intended to replace judges, but rather to augment their knowledge and

²¹⁵ Vincent Chiao (2019) (n94) 134.

²¹⁶ Danielle Kehl and others (n20) 30.

²¹⁷ Vincent Chiao (2019) (n94) 130.

²¹⁸ Anna M Costello and others, ‘Machine + man: A field experiment on the role of discretion in augmenting AI-based lending models’ (2020) 70(2-3) *Journal of Accounting and Economics* 1, 37 <<http://dx.doi.org/10.1016/j.jacceco.2020.101360>> accessed 21 January 2023.

²¹⁹ Saar Alon-Barkat and Madalina Busuioc (n191) 166.

²²⁰ Saar Alon-Barkat and Madalina Busuioc (n191) 153-154.

²²¹ Anju Vali Tikoo (n41) 45.

²²² In *KP Singh v. State (NCT of Delhi)*, (2015) 15 SCC 497 [10], the SC observed:

“Determining the adequacy of sentence to be awarded in a given case is not an easy task, just as evolving a uniform sentencing policy is a tough call. That is because the quantum of sentence that may be awarded depends upon a variety of factors including mitigating circumstances peculiar to a given case. The Courts generally enjoy considerable amount of discretion in the matter of determining the quantum of sentence. In doing so, the courts are influenced in varying degrees by the reformatory, deterrent and punitive aspects of punishment, delay in the conclusion of the trial and legal proceedings, the age of the accused, his physical/health condition, the nature of the offence, the weapon used and in the cases of illegal gratification the amount of bribe, loss of job and family obligations of accused are also some of the considerations that weigh heavily with the Courts while determining the sentence to be awarded. The Courts have not attempted to exhaustively enumerate the considerations that go into determination of the quantum of sentence nor have the Courts attempted to lay down the weight that each one of these considerations carry. That is because any such exercise is neither easy nor advisable given the myriad situations in which the question may fall for determination.”

experience.²²³ It is also unclear if the private entities driving AI systems have any incentive to act in the public interest, which raises ethical questions about their application in criminal justice.²²⁴ AI has the potential to eliminate judicial discretion because it assigns a score to the accused, which is then approved by the court, and for a judge to “override” the score, their counterintuitions would have to be far more convincing than the score itself.²²⁵ This prospect gives rise to another major risk of AI-based sentencing – the automation bias.²²⁶

A. The Possibility of Automation Bias

In establishing one’s impressions of fairness – human behaviour, emotions, and societal standards play a crucial role.²²⁷ When dealing with a judge, if anything is wrong, one can debate it with them, but not against an algorithm, because it either scores one as high-risk or not, and a sense of finality permeates the score, even if it is incorrect.²²⁸ Depending on the social context, any human right may be compromised when computerised reasoning aids or replaces human decision-making.²²⁹ No technique for sentencing criminals may be flawless, but an option like computerised sentencing merely needs to be better than the *status quo*.²³⁰ Nonetheless, AI strengthens the highly perilous concept of “automation bias,” which exists within the system.²³¹

²²³ In the US, this happened when the Federal Sentencing Guidelines became mandatory, which eventually led to such mandates being made advisory. See, Michael E Donohue (n18) 670.

²²⁴ This argument becomes more significant given that the criminal justice system primarily engages with poorer and marginalised segments of society, who lack the resources to defend themselves against the system. See, Vincent Chiao (2019) (n94) 137. See also, Ameen Jauhar and Jai Vipra, ‘Procurement of Facial Recognition Technology for Law Enforcement in India: Legal and Social Implications of the Private Sector’s Involvement’ (Vidhi Centre for Legal Policy 2021) <<https://vidhilegalpolicy.in/research/procurement-of-facial-recognition-technology-for-law-enforcement-in-india-legal-and-social-implications-of-the-private-sectors-involvement/>> accessed 23 February 2023 (noting that questions of privacy, security, and public autonomy arise when private entities provide FRT to the state for law enforcement). See further, Sarah Valentine (n11) 404.

²²⁵ Michael E Donohue (n18) 665.

²²⁶ The author is grateful to Ameen Jauhar for an enriching discussion on this issue.

²²⁷ Mirko Bagaric and others (2020) (n3) 1049.

²²⁸ Vincent Chiao (2019) (n94) 133-134.

²²⁹ Aleš Završnik (2020) (n10) 580.

²³⁰ Vincent Chiao (2018) (n75) 240. See also, Frej Klem Thomsen, ‘Iudicium ex Machinae – The Ethical Challenges of Automated Decision-Making in Criminal Sentencing’ (2020) Danish Institute for Human Rights <<https://philarchive.org/archive/THOIE2>> accessed 17 January 2023 (arguing that using an automated sentencing system is possible, but it must be more secure, more transparent, and less biased than human sentencing).

²³¹ Sarah Valentine (n11) 396.

Automation bias is a cognitive phenomenon that describes people's propensity to over-rely on automated decision-making processes and accept their choices without proper scrutiny.²³² The assumption that algorithms are neutral and impartial is a significant source of automation bias. Unfortunately, this view ignores the fact that the objectivity of algorithms depends on the training data, and that this data can itself reflect and strengthen societal biases. This is especially significant in the setting of criminal sentencing, where automated methods are used to decide the severity of punishment, and human judges tend to dismiss contradicting information.²³³ People perceive computer systems as authority, which increases the likelihood that these systems will be blindly followed, even in the face of facts that indicates it would be preferable not to follow them.²³⁴ However, people also have an irrational preference for human decision-making over computers due to "algorithmic aversion."²³⁵

Those with financial incentives are particularly susceptible to the automation biases.²³⁶ When technology fails to signal an issue, humans become more susceptible to the effects of automation bias, which encourages a propensity to refuse or not look for contradicting facts regarding a computer-generated solution that is considered as correct.²³⁷ In the judicial system, automated decision-making technologies are used in some jurisdictions, yet even for the judges themselves, the systems are opaque.²³⁸ The power of the human mind to adapt the rules to new social contexts is undervalued by proponents of automated decision-making.²³⁹ Nonetheless, policymakers should let the reality of imprecision (both human and technical) guide their

²³² Automation bias is a well-known decision-making issue that arose from research in aviation and healthcare, both of which have traditionally relied heavily on computerised systems. See, Saar Alon-Barkat and Madalina Busuioc (n191) 155.

²³³ Kenneth A Bamberger, 'Technologies of Compliance: Risk and Regulation in a Digital Age' (2010) 88(4) Texas Law Review 669, 711-712 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1463727> accessed 14 February 2023.

²³⁴ Linda J Skitka and others, (n203) 993. See also, John Zerilli, 'Algorithmic Sentencing: Drawing Lessons from Human Factors Research' (2020) Preprint 1 <<http://philsci-archive.pitt.edu/id/eprint/19229>> accessed 27 January 2023 (arguing that "if a system becomes reliable enough, humans will become diffident to the point of adhering to the system's recommendations even when they have grounds to disbelieve them").

²³⁵ Mirko Bagaric and others (2022) (n19) 98.

²³⁶ Kenneth A Bamberger (n233) 676.

²³⁷ Kenneth A Bamberger (n233) 711-712.

²³⁸ Automated decision-making systems have been flagged with issues like, lack of consistent policies, ambiguity on transparency, absence of legal remedies in cases of disputes. See, Krzysztof Izdebski (ed.), *alGOvrithms: State of Play* (Wisegrad Grant No. 21820296, ePanstwo Foundation 2019) 41 <<https://crta.rs/wp-content/uploads/2019/05/alGOvrithms-State-of-Play-Report.pdf>> accessed 15 February 2023. See also, Aleš Završnik (2020) (n10) 572.

²³⁹ Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press, New York 2018) 80-81 cited in Aleš Završnik (2020) (n10) 580-581.

decisions, realising the need to safeguard against catastrophe in the face of uncertainty, anything otherwise would be to abandon regulatory responsibilities.²⁴⁰

Algorithms are often programmed to optimise for outcomes or to represent ideals, which might inject new biases into the decision-making process. Using automated systems helps lighten the mental load of decision making, which is especially helpful for humans given their finite cognitive capacities. But this might cause people to become complacent and overly reliant on the system, which is risky, if the system itself is defective. Understanding that algorithms are not inherently neutral and can perpetuate prejudices if not carefully monitored is crucial for preventing the unfair application of AI in sentencing. To address automation bias in algorithmic decision-making, it is necessary to recognize that algorithms are not inherently objective and that they can perpetuate biases if not carefully designed and monitored.²⁴¹ It is important to subject algorithms to rigorous testing to ensure that they are fair and unbiased, and to incorporate human oversight into the decision-making process. Although, lessons can be taken from the EU's and the UK's stance on automated decision making.²⁴²

V. TOWARDS AN EQUITABLE AI FRAMEWORK

Criminal sentencing is difficult, but it is not *that* difficult that it could not possibly be improved upon by an algorithm.²⁴³ The greatest potential of algorithmic decisions resides in enhancing the precision of decisions made in the setting of mundane, daily events.²⁴⁴ At the stage of sentencing, the majority of the facts have already been resolved, AI can determine the pertinent factors affecting sentencing by analysing the case history,²⁴⁵ to assess the quantum of punishment that the judge can refer to. Despite this, it is undeniable that opaque, biased, judicially influenced, and racially discriminatory algorithmic risk assessment in pretrial

²⁴⁰ Kenneth A Bamberger (n233) 739.

²⁴¹ Alicia Solow-Niederman and others (n111) 713-714 (noting that “As more complex ML methods are integrated into risk assessment instruments, it will become even more essential to resist ‘automation bias’ and ensure adequate oversight of the tool’s fairness and accuracy”)

²⁴² Article 22 of the EU's General Data Protection Regulation, 2016, provides that “data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.” While Article 49(1) of the UK's Data Protection Act, 2018, provides that “a controller may not take a significant decision based solely on automated processing unless that decision is required or authorised by law.”

²⁴³ Vincent Chiao (2018) (n75) 257.

²⁴⁴ Vincent Chiao (2019) (n94) 133.

²⁴⁵ Nigel Stobbs and others (n73) 44.

adjudication violates basic constitutional principles.²⁴⁶ Since it is unclear and difficult for laypeople to judge how the sentencing grounds translate into a specific punishment decision, algorithmic sentencing could improve sentencing accountability by making the foundation for sentencing easier to comprehend and analyse,²⁴⁷ yet it remains to be seen whether AI can explain aggravating or mitigating factors. Algorithm usage in criminal sentencing ought to be open and accountable, so that their decisions can be audited and challenged, if necessary.²⁴⁸ Considering the dynamic nature of algorithms, if algorithmic sentencing is to be deployed, certain foundational principles – audibility, transparency, and consistency – can help safeguard due process.²⁴⁹

Policies recommending the establishment of a regulatory government body, the development of a legal framework to strengthen transparency, the formation of a national advisory group of experts, and the development of national standards for procuring and auditing are all acceptable approaches to the government's (or the courts') use of algorithms.²⁵⁰ Moreover, the following AI fairness criteria for predictive justice can be relied upon while designing the sentencing algorithm:²⁵¹

- i. **Representation:** to avoid unjust prejudices, all subjects should have a proportional representation in the data to mitigate bias and discrimination.

²⁴⁶ Christopher Thomas and Antonio Pontón-Núñez (n79) 406.

²⁴⁷ Vincent Chiao (2018) (n75) 255.

²⁴⁸ Madalina Busuioc, 'Accountable Artificial Intelligence: Holding Algorithms to Account' (2020) 81 Public Administration Review 825, 833-834 <<http://dx.doi.org/10.1111/puar.13293>> accessed 22 January 2023. See also, Jule Pattison-Gordon, 'Criminal Justice Algorithm Predicts Risk of Biased Sentencing' (*GovTech*, 12 July 2022) <<https://www.govtech.com/computing/criminal-justice-algorithm-predicts-risk-of-biased-sentencing>> accessed 3 February 2023 (noting the creation of an "algorithm that predicts risks of biased, overly punitive sentencing," which can be used to audit the sentencing algorithm for bias).

²⁴⁹ John Villasenor and Virginia Foggo (n8) 354 (Auditability mandates the documentation of a static representation of the data that was used to produce the risk assessment; transparency ensures that the data is made available to the accused; consistency examines whether factors like ethnicity are used and whether similar cases are scored similarly); Varsha PS (n158) 6 (noting and discussing the responsible AI principles to be adhered to reduce bias); Cynthia Rudin and others (n20) (proposing a framework that incorporates three principles of transparency, auditability, and accountability to address the concerns around algorithmic decision-making in criminal justice); Danielle Kehl and others (n20) 32; Mirko Bagaric and others (2022) (n19) 140; Greg Satell and Josh Sutton, 'We Need AI That Is Explainable, Auditable, and Transparent' (*Harvard Business Review*, 28 October 2019) <<https://hbr.org/2019/10/we-need-ai-that-is-explainable-auditable-and-transparent>> accessed 22 March 2023.

²⁵⁰ Art Alishani and Krzysztof Izdebski (eds.), *alGOvrithms 2.0: State of Play* (Visegrad Grant No. 22020475, Open Data Kosovo 2021) 39-40 <https://metamorphosis.org.mk/wp-content/uploads/2021/04/algovrithms_2.0_report-2021.pdf> accessed 15 February 2023.

²⁵¹ Vyacheslav Polonski, 'AI is convicting criminals and determining jail time, but is it fair?' (*World Economic Forum*, 19 November 2018) <www.weforum.org/agenda/2018/11/algorithms-court-criminals-jail-time-fair/> accessed 8 January 2023.

- ii. Protection: individuals must be protected from unfair consequences of algorithms, and how an algorithm affects the most vulnerable individuals should be used to gauge its fairness.
- iii. Stewardship: algorithmic justice symbolises the active obligation to strive for fairness in AI systems and only a diverse team of developers that questions each other's underlying assumptions can bear the essence of stewardship.
- iv. Authenticity: it pertains to the legitimacy of AI predictions as they are utilised to guide human decision-making, as well as the validity of dataset used.

During the development of sentencing algorithms, it is crucial that aggravating and mitigating factors be utilised to decide the severity of the sentence, and that AI approaches be employed to establish the relative weight of the factors.²⁵² The sentencing algorithm must be designed while keeping the following rules in mind:

- i. Using anonymised data: all personal traits like religion, caste, ethnicity, gender must be deleted from the data and only the factual matrix must be fed in the system like the type of offence and the quantum sentence.
- ii. AI score non-binding on the judge: AI should be used to analyse the previous similar instances and analysing the previous sentencing, which must only augment the judges' decision making.
- iii. Making the algorithm public: the algorithm can be made publicly available which can be challenged, if necessary, and the public can flag potential pitfalls, if at all found.
- iv. Diverse team of AI developers: the developer of the sentencing algorithm must ensure that a diverse team constantly works to improve the algorithmic efficacy based on newest data and bugs that are identified.
- v. Compensation for wrongful sentencing: if the algorithm brings out a wrongful score based on which the judge awards sentence, the convict must be compensated commensurate with injury – which will ensure accountability towards both, the judge, and the algorithm.
- vi. Specifying the filters and inputs: the algorithm must be specified on the lowest possible level which will enable the input of specific aggravating and mitigating circumstances and it will enable the sentencing judge to contextualise their decision.

²⁵² Mirko Bagaric and others (2022) (n19) 119.

- vii. Appeal: in grave offences, once a sentence has been passed, the convict shall have the right to challenge the court's decision as well as the algorithm based on which the sentence was decided in the first place.

However, a dataset has proven to produce racially uneven predictions when race is excluded,²⁵³ and due to the variety of variables which could be linked to sensitive variables, even eliminating the characteristics like gender, ethnicity, and sexual, may not remove social disparities.²⁵⁴ Thus, AI developers must be particularly aware of their blind spots and underlying preconceptions in this regard; it is not the choice of algorithms that matters, but all small decisions²⁵⁵ about finding, classifying, and labelling data for AI models.²⁵⁶ AI programmers should be held to the same high moral and legal standards as judges, who have to explain not only their decisions but also why they made them.²⁵⁷ The European Parliament adopted the AI Act in 2024 – perhaps the world's first comprehensive AI regulation – which will hugely impact companies using AI, by providing *inter alia* for a compensation regime.²⁵⁸ Overall, we need to be more critical of the ways that the criminal justice system uses AI, and we need to develop new ways to use AI that are fair and equitable.²⁵⁹

In AI-based sentencing, too, the first principles of criminal justice system including presumption of innocence of the accused, the State's burden of proof beyond reasonable doubt, should never be done away with.²⁶⁰ Considering the current plight of justice system in India, it would not be wise to completely automate sentencing, rather, it can be piloted, and thereafter,

²⁵³ James E Johndrow and Kristian Lum, 'An algorithm for removing sensitive information: Application to race-independent recidivism prediction' (2019) 13(1) *The Annals of Applied Statistics* 189 <<http://dx.doi.org/10.1214/18-aos1201>> accessed 8 January 2023. See also, Vincent Chiao (2019) (n94) 129 (claiming that even a completely racially neutral process cannot ensure fairness in the criminal justice system).

²⁵⁴ *Carlow University* (n155); Aaron M Bornstein, 'Are Algorithms Building the New Infrastructure of Racism?' (*Nautilus*, 5 December 2017) <<http://nautil.us/issue/55/trust/are-algorithms-building-the-new-infrastructure-of-racism>> accessed 11 January 2023.

²⁵⁵ Rajen Sheth (n206).

²⁵⁶ Vyacheslav Polonski (n251).

²⁵⁷ Ellora Israni, 'Algorithmic Due Process: Mistaken Accountability and Attribution in *State v. Loomis*' (*Harvard Journal of Law & Technology*, 31 August 2017) <<https://jolt.law.harvard.edu/digest/algorithmic-due-process-mistaken-accountability-and-attribution-in-state-v-loomis-1>> accessed 17 February 2023.

²⁵⁸ Kirk J Nahra and others, 'The European Parliament Adopts the AI Act' (*WilmerHale*, 14 March 2024) <www.wilmerhale.com/en/insights/blogs/wilmerhale-privacy-and-cybersecurity-law/20240314-the-european-parliament-adopts-the-ai-act> accessed 2 April 2024.

²⁵⁹ Chelsea Barabas (n192).

²⁶⁰ Tatyana Sushina and Andrew Sobenin, 'Artificial Intelligence in the Criminal Justice System: Leading Trends and Possibilities', *Proceedings of the 6th International Conference on Social, economic, and academic leadership (ICSEAL-6-2019)* (Atlantis Press 2020) 432, 436 <<https://doi.org/10.2991/assehr.k.200526.062>> accessed 14 June 2023.

deployed in a phased manner. The broad categories of offences can be made, with the role of AI being defined and well-structured in each of such categories, which can be as follows:

- i. Offences punishable with imprisonment of up to 2 years: the judge may accept the algorithmic sentence without mandatorily having to justify the reasons thereof, having the discretion to modify the sentence ultimately.
- ii. Offences punishable with imprisonment of up to 7 years: if the algorithmic sentence is contested by the accused (on due process grounds) or the judge (who when considers the punishment to be illegitimate), the judge may modify the sentence, while explaining the reasons thereof.
- iii. Offences punishable with imprisonment above 7 years: the algorithmic sentence can be utilised by the human judge to augment their decision making, and the judge must give justification for the sentence they finally arrived at.

VI. CONCLUSION

The Indian criminal justice system is heavy, open ended, with heavy reliance on judicial discretion – which, algorithms present an opportunity to fix. Without adequate protections, algorithms risk undermining the rule of law and individual rights.²⁶¹ It is unlikely that an AI will ever become conscious, even if it does, it is unclear how this will happen or what kind of entity it will be.²⁶² This idea implies that AI should be assessed using a larger range of ethical and legal standards,²⁶³ rather than merely classification precision and mistake matrices. The authors of an algorithm might have good intentions, the algorithm may still take an unexpected path to its goals,²⁶⁴ thus, algorithmic decision-making may infringe fundamental rights.²⁶⁵ Nevertheless, it is acknowledged that if at all AI takes on judicial discretion in criminal sentencing without the risks being mitigated, the marginalised population will be at the receiving end of discriminatory and biased sentencing. AI presents multiple risks like automation bias, and thus, the algorithm's forecast should not be made binding on the judge

²⁶¹ Jason Tashea, 'Courts Are Using AI to Sentence Criminals. That Must Stop Now' (*Wired*, 17 April 2017) <www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/> accessed 8 January 2023.

²⁶² Camilo Miguel Signorelli (n5).

²⁶³ See generally, Eleanor Bird and others, 'The Ethics of Artificial Intelligence: Issues and Initiatives' (European Parliamentary Research Service 2020) <<https://data.europa.eu/doi/10.2861/6644>> accessed 7 February 2023 (exploring the ethical and moral issues raised by AI development and use).

²⁶⁴ Aleš Završnik (2019) (n139) 630.

²⁶⁵ Aleš Završnik (2019) (n139) 636.

and its sole purpose can be to inform the judge of what the punishment level has been in past similar situations, and the judge should be allowed to make their own ultimate decision.²⁶⁶ AI holds great promise if deployed to augment judicial discretion through partnership to help judges use their own judgment, instead of replacing them.²⁶⁷

The apparent flaws of computers can be corrected, but centuries of human bias, such as racial injustice or discrimination in the judicial system, cannot be undone overnight.²⁶⁸ Thus, outright opposition to AI-based sentencing cannot be advocated,²⁶⁹ and instead, a tailored and rights-based protectionist regime can be established albeit if risks like discrimination or bias are mitigated.²⁷⁰ Additionally, if private actors, who majorly develop the AI systems, are given the trade-secret benefits, they may, without any checks and balances, act as a political stooge,²⁷¹ which might hugely compromise the democracy. Therefore, it is important that appropriate AI regulation is in place, and more importantly, the regulations hold those accountable who make errors – a starting point is the EU’s AI Act 2024, which attempts to hold AI accountable while protecting the citizenry. It is imperative that when AI in sentencing should be deployed with the principles of responsible AI in mind,²⁷² that will not only boost market competition but also improve public trust in the justice system.²⁷³

²⁶⁶ Tim O’Brien, ‘Compounding Injustice: The Cascading Effect of Algorithmic Bias in Risk Assessments’ (2020) 13 *Georgetown Journal of Law & Modern Critical Race Perspectives* 39, 81 <<http://dx.doi.org/10.2139/ssrn.3694818>> accessed 17 January 2023 (arguing that “not all algorithms exist to simply make decisions on their own, but they serve the purpose of ‘decision support,’ providing humans with insights needed for more informed decision making”).

²⁶⁷ Jiahui Shi, ‘Artificial Intelligence, Algorithms and Sentencing in Chinese Criminal Justice: Problems and Solutions’ (2022) 33 *Criminal Law Forum* 121 <<http://dx.doi.org/10.1007/s10609-022-09437-5>> accessed 22 February 2023.

²⁶⁸ Mirko Bagaric and others (2022) (n19) 98.

²⁶⁹ Dan Hunter and others (n25) 800.

²⁷⁰ Bart Custers, ‘AI in Criminal Law: An Overview of AI Applications in Substantive and Procedural Criminal Law’, *Law and Artificial Intelligence* (T.M.C. Asser Press 2022) 205 <http://dx.doi.org/10.1007/978-94-6265-523-2_11> accessed 6 April 2023 (concluding that “the use of AI can offer many benefits in criminal investigation, but only if prejudice, discrimination, and other risks are avoided or mitigated”).

²⁷¹ Genevieve Smith and Ishita Rustagi (n37) 37.

²⁷² NITI Aayog, ‘Approach Document for India – Principles for Responsible AI (Part 1)’ (*India AI*, 24 February 2021) <<https://indiaai.gov.in/research-reports/responsible-ai-part-1-principles-for-responsible-ai>> accessed 10 January 2023.

²⁷³ Genevieve Smith and Ishita Rustagi (n37) 39. See also, Sarah L Desmarais (n19) (proposing that a responsible approach to predictive risk modelling should include transparency, fairness, and accountability, and should involve the input of stakeholders and experts in the design and use of the models).