Criminal Law Forum (2022) 33:121–148
https://doi.org/10.1007/s10609-022-09437-5

JIAHUI SHI*

ARTIFICIAL INTELLIGENCE, ALGORITHMS AND SENTENCING IN CHINESE CRIMINAL JUSTICE: PROBLEMS AND SOLUTIONS

Published online: 27 April 2022

ABSTRACT. The State Council of the People's Republic of China has declared its intention to introduce AI into the Chinese criminal justice system including the imposition of criminal sentences. This plan has, however, raised a range of troubling questions and concerns which include the misinterpretation of court decisions by AI, the incapability of AI to make value judgements, possible biases of algorithms, selectivity of data used by AI, the "black box" character of sentencing by AI, diminished acceptance of AI-supported sentencing systems by the public, uncertain quality of algorithms, etc. The positive e ¤ ect of AI on the goal of "same case, same sentence" therefore should not be overstated, and an unlimited application of AI must be avoided. Chinese policy makers should therefore use great caution when integrating AI into sentencing. AI should be employed not as a decision-maker but only as an "assistant", providing information for judges and aiding them in making sentencing decisions. The final determination should in any event remain in the hands of the judges. Moreover, algorithms should be made transparent so that judges can review their operation. A Committee supervised by the Chinese Supreme Court should be established to guarantee the quality of judicial data used by AI and to operate a centralized AI system on sentencing. These measures would contribute to making the best use of judicial data and to introducing a fair, accurate, and e ¤ cient sentencing system.

* Shi, Jiahui is a research assistant at the Law Faculty, Sichuan University, China. She received her doctorate in law (Dr. iur.) from the University of Cologne, Germany, and Bachelor and Master of Law Degrees from the China University of Political Science and Law, Beijing. Her research areas cover criminal law, criminal procedure law, comparative law, technological law, and law on privacy.Dr. iur., University of Cologne (Germany); research assistant, Law Faculty, Sichuan University, China. E-mail: shijiahui2006zxl@gmail.com

Driven by the progress of information technology, Artificial Intelligence (AI)[1] has made great strides, especially regarding its accuracy[2] and efficiency, in many areas, such as business, industry, medical treatment, and elderly care. Now AI is making its way into criminal justice. For example, in the U.S., AI systems are being applied in the pre-trial phase, e.g., for predicting when and where crimes might occur, and for assessing the risk of recidivism in bail and sentencing decisions.[3] Programs such as the Public Safety Assessment (PSA) and Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) systems are widely used in the United States. In some European countries, electronic case management systems have been adopted to assist courts in managing their cases.[4]

In the wake of a general trend toward digitalization, AI has also attracted the interest of Chinese policy makers. The Chinese State Council declared its intention to introduce "Intelligent Courts", based on digitalizing judicial administration and on using AI for judicial decision-making, including sentencing.[5] In the context of Chinese criminal justice, AI refers to any type of algorithm with the capacity of analysing information provided by judges, enforcement agents, and defense lawyers and of relating it to existing criminal justice data in order to assist legal professionals in making decisions, such as reviewing evidence and sentencing. The digitalization of

[1] A standard definition of AI does not yet exist. See Benedikt Kohn, *KÜNSTLICHE INTELLIGENZ UND STRAFZUMESSUNG* (Nomos 2020), p. 26. The core feature of AI is the ability to analyse input provided by humans and to produce outcomes for certain purposes. See Johannes Kaspar, Katrin Höffler, Stefan Harrendorf, "Datenbanken, Online-Votings und künstliche Intelligenz" (2020) 32 *NEUE KRIMINALPOLITIK* 35, 40; see also Stuart Russell and Peter Norvig, *ARTIFICIAL INTELLIGENCE: A MODERN APPROACH* (Pearson Global Edition 4th Edition 2021), pp. 19–20.

[2] According to empirical studies, algorithms based on big data are, on average, 10% more accurate than clinical predictions on human health and behaviour. William Grove, David Zald, Boyd Lebow, Beth Snitz, Chad Nelson, "Clinical Versus Mechanical Prediction: A Meta-Analysis" (2000) 12 PSYCHOL. ASSESSMENT 19, 19.

[3] Dan Hunter, Mirko Bagaric and Nigel Stobbs, "A Framework for the Efficient and Ethical Use of Artificial Intelligence in the Criminal Justice System" (2020) 47 FLA. ST. U. L. REV. 749, 752.

[4] The European Commission for the Efficiency of Justice, "Thematic Report: Use of Information Technology in European Courts" (2016), CEPEJ STUDIES No. 24, 22.

[5] Haibo Sun (孙海波), "Reflection on the Possibility and Limitation of Intelligent Judging (反思智能化裁判的可能及限度)" (2020) 5 *REVIEW OF THE NATIONAL COLLEGE OF PROSECUTORS (*国家检察官学院学报) 80, 81.

information on proceedings and court decisions is to improve judicial administration by providing more transparency; this has hardly met with any objections.[6] By contrast, the introduction of AI into sentencing, which is regarded as an important step toward an "Intelligent Courts" system, has been greeted by some applause but has also raised a range of troubling questions and concerns. This article will demonstrate the current development and application of AI in the sentencing phase in China and its main problems. In the end, some possible solutions will be suggested.

I SENTENCING AND THE APPLICATION OF AI IN CHINA

1.1 *Big Disparities in Chinese Sentencing Practice*

Big disparities in sentencing outcomes for similar cases between different judges and regions are a chronic and stubborn disease that impairs criminal justice and causes distrust in the Chinese judicial system.[7] Such disparities become more recognizable when more judgements are easily accessible online to the public.

The *Chinese Criminal Law* (hereafter referred to as CCL) provides for each offense a lowest and a highest term of imprisonment. The sentence range can be broad.[8] For example, a rapist can be imprisoned for a term between three years and ten years, and he can receive the death penalty or a prison term beyond ten years if there exist aggravating circumstances, which are not defined in the CCL (Art. 236 CCL).[9] In the past, Chinese judges did not discuss aggravating and mitigating circumstances but imposed sentences depending on their personal preference and experience within the framework provided by law for the crime of conviction.[10] This has led to two problems. First, sentences vary greatly among different judges; for

[6] *Ibid*.

[7] For two cases showing big disparities in Chinese sentencing see Julian V. Roberts and Wei Pei, "Structuring Judicial Discretion in China: Exploring the 2014 Sentencing Guidelines" (2016) 1 *CRIMINAL LAW FORUM* 27, 7.

[8] Chunyan Huang (黄春燕), "Sentencing Discretion of Judges and Realization of Sentencing Justice (法官量刑的自由裁量权与量刑公正的实现)" (2021) 296 *JOURNAL OF SHANDONG NORMAL UNIVERSITY (SOCIAL SCIENCES) (*山东师范大学学报(社会科学版))136, 136.

[9] Further examples can be found in: Roberts and Pei, *supra* note 7, 6.

[10] Tingguang Zhao (赵廷光), *EMPIRICAL STUDY ON FAIRNESS OF SENTENCING (*量刑公正实证研究) (Wuhan University Publishing 2005), p. 7.

example, female judges tend to show greater sympathy for defendants and impose more lenient sentences in rape cases than their male colleagues.[11] Second, due to the lack of guidelines on sentencing, judges' decisions can be arbitrary, also inviting corruption. The absence of written reasons in judgments further aggravates these problems.

1.2 *Development of Sentencing Guidelines*

It is widely believed that in an ideal sentencing system "any sentencer presented with the same case would reach the same decision as to the appropriate sentence," and "the sentence for any case would be predictable".[12] Such a system is free from judicial bias and inconsistent sentencing. With such a prospect in mind, the Chinese Supreme Court initiated a reform on reducing sentencing discretion in order to achieve the goal of "same case, same sentence". According to this reform, a quantitative sentencing method is to be introduced and Chinese judges are to be provided with detailed sentencing guidelines.[13] As a substantive outcome of this reform, the Chinese Supreme Court in 2013 issued the first *Sentencing Guidelines on Frequent Crimes*[14] for 15 offenses to restrict and guide judicial discretion at sentencing.[15] These guidelines provide details on the length of imprisonment for aggravating or mitigating circumstances. For example, if the defendant confessed to the crime, the judge can impose a prison sentence that is up to 20% shorter than without a confession.[16] Eight more crimes and details on imposing fines and probation were added in 2017.[17] In July 2021, the Supreme Court and the General Prosecution Office jointly issued the latest version of the

[11] Yiwei Xia, Tianji Cai and Hua Zhong, "Effect of Judges' Gender on Rape Sentencing" (2019) 19 *THE CHINA REVIEW* 125, 141.

[12] Neil Hutton, "Sentencing, Rationality, and Computer Technology" (1995) 22 *J. L. & SOC'Y* 549, 552.

[13] Huang (黄春燕), *supra* note 8, 137.

[14] 最高人民法院关于常见犯罪的量刑指导意见.法发[2013]14号 https://www.nmql.com/flvfgui/zyfgui/6853_5.html accessed 15/09/2021.

[15] See also Roberts and Pei, *supra* note 7, 3.

[16] Part III, Section 7 of the *Guideline*.

[17] 关于常见犯罪的量刑指导意见(二)(试行).法发[2017]7号 https://www.waizi.org.cn/doc/62824.html accessed 15 September 2021.

Sentencing Guidelines[18] (hereafter referred to as *Guidelines*), which amended the former version by adding four more mitigating circumstances, such as plea bargaining, to keep pace with the changes of the CCL in 2018. Within the framework of the *Guidelines* and the former versions, provinces issued their own local sentencing guidelines with more details.[19]

1.3 *Current Application of AI in Chinese Sentencing*

The Chinese Supreme Court saw the development of AI as a new possibility for implementing the principle "same case, same sentence". In 2016, the Chinese State Council issued the *State Guidelines on the Development of Informatization*,[20] which aim at "introducing 'Intelligent Courts', improving the digitalization in all phases, including case registration, trial, execution, and supervision, and promoting the transparency of judicial information and justice". In 2017, the State Council issued a *Plan for a New Generation of Artificial Intelligence*,[21] and the Chinese Ministry of Industry and Information Technology issued a *Three-Year Plan to Promote the Development of the Next Generation of Artificial Intelligence (2018–2020),*[22] which invited all departments to explore the possible use of AI in their respective fields. The judiciary reacted positively to this strong push from the central government. The Supreme Court in 2018 introduced a computer system called "similar case" based on big data algorithms, which compares similarities among court decisions and produces a list of cases including factors (such as "theft" and "confession") considered by judges.[23] At the local level, some courts developed and introduced their own AI systems to assist in trial and sentencing, such

[18] 关于常见犯罪的量刑指导意见(试行).法发[2021]21号 https://www.chinalawtranslate.com/16861-2/ accessed 15/09/2021. These twenty-three crimes cover approximately 90% of all criminal cases. See https://www.sohu.com/a/476986647_114988 accessed 14/09/2021.

[19] For detailed rules issued by the Beijing High Court in 2014 see https://www.faxin.cn/lib/dffl/dfflcontent.aspx?gid=B1012559&nid=2650 accessed 14/09/2021.

[20] 国家信息化发展战略纲要 http://www.gov.cn/zhengce/2016-07/27/content_5095336.htm accessed 15/09/2021.

[21] 新一代人工智能发展规划, 国发[2017]35号 http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm accessed 15/09/2021.

[22] 促进新一代人工智能产业发展三年行动计划(2018–2020年), 工信部科 [2017]315号http://www.cac.gov.cn/2017-12/15/c_1122114520.htm accessed 15/09/2021.

[23] Weimin Zuo (左卫民), "How to Realize Similar Sentencing in Similar Cases with Artificial Intelligence (如何通过人工智能实现类案类判)" (2018) 20 *CHINA LAW REVIEW (*中国法律评论) 26, 27. This system is not applied nationwide.

as the "Rui Judge" (translated as "wise judge") system applied by Beijing courts and the "206 system" in the Shanghai court system[24] with multiple functions, e.g., to review arrest warrants and evidence, to assess defendants' danger to society, and to assist in sentencing.[25]

There are also AI systems developed by private companies as commercial products, such as "Little Judge Bao" for the prediction of sentencing.[26] The programmers designed the algorithms of "Little Judge Bao" based on factors named in the *Guidelines* and various local sentencing guidelines and imported judicial big data.[27] Users first select the jurisdiction where the case has been tried and then the offence of conviction. With that information, the system can determine the sentencing framework provided by the CCL. Users then select factual circumstances of the individual case, such as the number of victims, the degree of injury caused by the defendant, and whether the offence was committed in an especially cruel way. The next step is the selection of aggravating and mitigating circumstances from a list provided by the system, including the defendant's age if relevant to sentencing (below 12 or 16, or above 75 years), disabilities, confession, surrender to the police, plea bargaining, an agreement reached with victims, compensation for damages, attempt, recidivism, organized crime, etc. Some circumstances have subcategories, such as pleading guilty during the investigation, at charging, at the first instance, or at the appeal stage. Different circumstances may result in different ranges of increasing or decreasing a prison sentence. Algorithms further narrow down the range of sentences. Eventually,

[24] See Yadong Cui (崔亚东), *ARTIFICIAL INTELLIGENCE AND MODERNIZATION OF THE JUDICIARY (*人工智能与司法现代化) (Shanghai People Publishing 2019), pp. 111–6.

[25] Changshan Ma (马长山), "Reshaping Effect of Artificial Intelligence in Judiciary and its Limitation (司法人工智能的重塑效应及其限度)" (2020) 42 *RESEARCH ON LAW (*法学研究) 23, 28. Other AI systems in the judicial system are, for example, "Fawu Cloud" in Jiangsu Province, "Zhishen" in Hebei Province and "Fazhi Cloud" in Chongqing. Xi Zheng (郑曦), "Application and Regulation of Artificial Intelligence in Trials (人工智能技术在司法裁判中的运用及规制)" (2020) 32 *PEKING UNIVERSITY LAW JOURNAL (*中外法学) 674, 677.

[26] Homepage of "Little Judge Bao": https://www.xiaobaogong.com.html accessed 14/09/2021.

[27] "Little Judge Bao" has already adapted its prediction system to the latest version of the *Guidelines* issued in July, 2021. See 关于常见犯罪的量刑指导意见 (试行), 法发(2021)21号 https://mp.weixin.qq.com/s/I-htMj42zpNoauoo4BwutQ accessed 14/09/2021.

"Little Judge Bao" produces a table predicting the sentence for the case. The table[28] presents:

Basic Information	name of the defendant/crime	
Case Facts	key words, e.g., corruption, damage, confession	
Sentencing Framework in CCL	e.g., above three years and under 10 years	
The Base Range of Sentencing (C)	base line for sentencing (A)	e.g., 66 months to 69 months
	circumstances increasing the sentence (B)	e.g., 1 month because of large amount of money
	base range of sentence (C = A + B)	e.g., 67 to 70 months
Reduction of the Sentence (normally expressed in percentages) (D)	mitigating circumstances (e.g., plea of guilty, confession, compensation)[29]	recommended percentage for reduction, e.g.,0–60% by the Guidelines60% by the AI system
Recommended Sentence (E)E = C*(1−D)		27 to 28 months

Fines will be suggested if applicable. The option for probation can be ticked by judges. The system also allows users to adjust the degree of discretion (such as 30%) for the baseline for sentencing and to change the percentage of the reduction of sentence.[30] The system provides empirical data with graphs, for example, in how many theft cases Beijing courts imposed a sentence of imprisonment or probation. Users can also find similar cases suggested by the system.[31] This

[28] This table has been simplified by the author to only present the most important information. See https://mp.weixin.qq.com/s/I-htMj42zpNoauoo4BwutQ accessed 5/02/2022.

[29] There are two options for judges to choose: one is mitigation within the sentencing framework provided by CCL, the other is mitigation below the sentencing framework.

[30] See https://mp.weixin.qq.com/s/I-htMj42zpNoauoo4BwutQ accessed 14/09/2021.

[31] See https://mp.weixin.qq.com/s/UkEu_L4bedsHZArFGuDZpw accessed 14/09/2021. More information on this system can be found in an article published by its founder, Dr, Wang: Yanling Wang (王燕玲), "The Implementation and the Approach of Big Data Precise Sentencing (大数据精准量刑的实现方法与路径)" (2020) 38 *JOURNAL OF GUIZHOU UNIVERSITY: SOCIAL SCIENCE (*贵州大学学报(社会科学版)) 89, 97–100.

system is available online and accessible to anyone, including judges, prosecutors, practising lawyers, defendants, and victims.

Chinese AI systems on sentencing come in three dimensions. Some AI systems, such as the "similar case" system developed by the Chinese Supreme Court, offer the judge information about sentences in cases that are similar to the one before him or her. AI systems like "Little Judge Bao" go further and suggest particular sentences based on factors the judge has chosen. The third type of systems warn judges of obvious disparities of their proposed sentence from other cases stored in the system's database.[32] The core of these systems is algorithms based on big data. They use mathematical modelling[33] in analyzing past cases and decisions, comparing their textual similarities, extracting factors relevant to sentencing, weighting these factors, and quantifying them. The algorithms categorize these data and produce a list of similar cases or propose a sentence by comparing factors embodied in earlier cases and identifying pertinent sentencing rules. In this way, the same input is designed to lead to the same output.[34] AI is believed to be free from personal bias and irrelevant influences; AI systems thus appear to be perfect tools for implementing the principle "same case, same sentence".[35] In China, the application of AI is in general regarded as positive[36] and is encouraged by policy makers. Although judges are not obliged to adopt the sentences produced by AI and some use the system only as a database, it is foreseeable that more AI systems will be developed in Chinese criminal justice and will be used more frequently.

II PROBLEMS AND CONCERNS REGARDING USING AI IN SENTENCING

New technologies are a double-edged sword. The possible impact of AI used in sentencing should therefore be studied thoroughly before

[32] Wenjie Feng (冯文杰), "Double-Construction of the Fairness of Artificial Intelligence in Sentencing (人工智能辅助量刑公正取向的双重构建)" (2020) 163 *JOURNAL OF EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY: SOCIAL SCIENCE (*华东理工大学学报·社会科学版) 114, 119.

[33] Sarah Valentine, "Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control" (2019) 46 *FORDHAM URBAN LAW JOURNAL* 364, 365.

[34] Ma, *supra* note 25, 31.

[35] Huang (黄春燕), *supra* note 8, 143.

[36] Zheng, *supra* note 25, 678.

it is implemented. Policy makers should, in particular, take into account several lines of criticism that have been raised by scholars.

2.1 *Criticism of Guidelines*

A general criticism acknowledges that the guidelines reduce the risk of arbitrary sentencing but suggests that they are too rigid because they oblige judges to impose sentences in a rather narrow range if certain conditions are fulfilled. As an indirect consequence of such detailed rules, judges tend to ignore circumstances relevant to sentencing but not mentioned in the *Guidelines*. One example is the case of Zhao Chunhua.[37] She had a stall in a park where tourists could shoot at balloons with toy guns to win prizes. Ms. Zhao was charged with illegal possession of guns. The first instance court, which applied the law "correctly", imposed three and a half years of imprisonment, which was obviously disproportional given her harmless intention and behaviour.[38] The *Guidelines* may thus lead to unfair sentences by restricting the discretion of judges.[39]

Moreover, some Chinese authors have raised objections to the goal behind the *Guidelines*, namely, "same case, same sentence". Looking for the "same case" has been called meaningless because two different cases can never be exactly equal, just as leaves of a plant are never identical. Therefore, the slogan, "same case, same sentence", has been criticized for lacking clear standards and has been called as a "false proposition"[40] and "a fictional myth of the rule of law".[41] Critics claim further that in deciding whether two cases are "the same", it is necessary to consider factors beyond the criminal act itself, including the purposes of punishment. The CCL mentions several purposes of punishment, including retribution (Art. 61

[37] (2017) Jin 01 Criminal Final No.41 ((2017) 津01刑终41号).

[38] The appeals court upheld the conviction but reduced the sentence to three years imprisonment with probation; (2017) Jin 01 Criminal Final No.41.

[39] Wenhua Peng (彭文华), "The History and Current Situation of the Federal Sentencing Guidelines and New Trends of Sentencing Reform in U.S. (美国联邦量刑指南的历史、现状与量刑改革新动向)" (2015) 6 *JOURNAL OF COMPARATIVE LAW (*比较法研究) 92, 104–5.

[40] Bensen Li (李本森), *Three Points on Reform of Sentencing Standardization (*量刑规范化改革的 '三点论'), in: Jinghai Shi (石经海) and Jinsong Lu (禄劲松) (ed) *RESEARCH ON SENTENCING (*量刑研究), Vol. 1 (Law Press China 2014), p. 7.

[41] Shaohua Zhou (周少华), "Like Cases Be Treated Alike: A Fictional Myth of the Rule of Law (同案同判:一个虚构的法治神话)" (2015) 11 *LAW SCIENCE (*法学)131, 140.

CCL[42]), education (Art. 72 CCL[43]), and prevention of crime. Equalising sentences only with regard to the criminal act and the harm resulting from it would ignore the preventive and educative purposes of the law.

2.2 *Misinterpretation of Cases*

Since AI can only work with factors which are already in its database, it may misidentify "same" cases, and if courts base their sentences on new factors, AI needs some time to integrate these factors, so in the meantime its results are inaccurate.[44] Moreover, not only the number of common factors but also the weight granted to them is essential for determining the similarity or relevance of cases.[45] It is impossible to fix the weight of each factor in advance because that weight is necessarily case-related.[46] The question of which factor is decisive in a case depends upon a comprehensive understanding of the whole case, which algorithms currently cannot provide.[47]

AI also has problems in "understanding" decisions correctly.[48] The flexibility of the Chinese language exacerbates this problem. It is common that the same legal concept is described differently in Chinese judgements. For example, "confession" as a mitigating factor

[42] Art. 61 CCL: "……the sentence shall be imposed on the basis of the facts of the crime, the nature and circumstances of the crime, and the degree of harm to society, in accordance with the relevant stipulations of this law."

[43] Shizhou Wang (王世洲), "Modern Theories on the Purposes of Punishments and Chinese Choices (现代刑罚目的理论与中国的选择)" (2003) 3 *RESEARCH ON LAW (*法学研究) 107, 123.

[44] Sun, *supra* note 5, 89.

[45] Id 90.

[46] Guodong Zhang (张国栋), "Research on Application Limits of Artificial Intelligence in the Criminal Justice Field (人工智能在刑事裁判领域应用限度研究)" (2020) 113 *JOURNAL OF BEIJING UNIVERSITY OF CHEMICAL TECHNOLOGY (SOCIAL SCIENCE EDITION) (*北京化工大学学报(社会科学版)) 63,65.

[47] Arpan Mandal, Kripabandhu Ghosh, Saptarshi Ghosh and Sekhar Mandal, "Unsupervised Approaches for Measuring Textual Similarity between legal Court Case Reports" (2021) 29 ARTIF INTELL LAW 417, 418. See Jie Feng (冯洁), "Challenges of Artificial Intelligence to Theories on Trial: Response and Limitation (人工智能对司法裁判理论的挑战:回应及其限度)" (2018) 21 *JOURNAL OF EASTERN CHINA UNIVERSITY OF POLITICAL SCIENCE AND LAW (*华东政法大学学报) 21, 30.

[48] Lusheng Wang (王禄生), "Technological Obstacles of Judicial Big Data and the Development of Artificial Intelligence (司法大数据与人工智能开发的技术障碍)" (2018) 20 *CHINA LAW REVIEW (*中国法律评论) 46, 46.

can be expressed as the Chinese word "confession"("坦白") or as "making incriminating statements voluntarily", or other similar expressions. AI is unlikely to understand that these expressions denote the same concept. Moreover, AI programmers cannot "teach" AI in advance what expressions have the same meaning because they cannot foresee usage in each judgement. Another risk of misinterpretation has to do with the lack of punctuation in Chinese texts. For example, if A, B, C, and D each represent one Chinese character they will appear as "ABCD" without space between two characters in the judgement. If the judges mean to write the two words "AB" and "CD", the combination "BC" can accidentally have a different meaning, which AI may register as relevant, thus arriving at a wrong result.

To sum up, the quality of AI's interpretation of Chinese judgements is not yet satisfactory. Consequently, judges have often complained that AI systems refer them to cases that are irrelevant to them.[49]

2.3 *Failure of AI in Making Value Judgements*

Imposing a sentence does not work like a mathematical equation. To devise a fair sentence, judges must invariably make value judgements. For example, a perpetrator's blameworthiness depends not only on his motives but also on the impact of the crime on society, which in turn must be assessed with regard to the social values protected by criminal law.[50] Moreover, the "softer" goals of the criminal justice system, such as dignity, equity, and mercy, also require individual value judgements.[51]

Currently used algorithms based on mathematical modelling are unable to accommodate value judgements, which cannot well be quantified.[52] AI systems therefore either screen out value issues or simplify them and interpret them as factual issues.[53] "Disturbing" value factors filtered out by algorithms, however, can be highly rel-

[49] Zuo, *supra* note 23, 28.

[50] M Schwarze and JV Roberts, *Reconciling Artificial and Human Intelligence*, in: Jesper Ryberg and Julian V. Roberts (eds) *SENTENCING AND ARTIFICIAL INTELLIGENCE*, Oxford University Press 2022, p. 208.

[51] Andrea Roth, "Trial by Machine" (2016) 104 *GEO. L.J.* 1245, 1247.

[52] Jinghui Chen (陈景辉), "The Legal Challenges of Artificial Intelligence: What Is the Start? (人工智能的法律挑战:应该从哪里开始?)" (2018) 5 *JOURNAL OF COMPARATIVE LAW (*比较法研究), 136, 141.

[53] Sun, *supra* note 5, 95.

evant to sentencing. Lacking an effective way to properly identify and assess value issues, AI may produce "justice" only on a formal level and miss the substantive questions.[54] As a consequence, AI may arrive at inappropriate sentences, as in the case of the lady with the balloon stand.[55] A human judge could easily avoid such incorrect results, whereas AI is limited to providing sentences that are proper according to its own system. AI sentencing may work in cases not involving value issues – but such cases are rare.

2.4 *Possible Bias of Algorithms*

In criminal justice, factors like gender, age, and socioeconomic status may be regarded as "preexisting biases in underlying data".[56] If factors are discriminatory, the results "will be doing nothing more than reinforcing the existing……bias in the criminal justice system".[57] For example, if an AI system takes gender into account when evaluating past sentences, it is likely to find that male defendants are given longer sentences than female defendants for violent crimes. If AI's recommendations followed this pattern and judges complied with the recommendations, this would lead to more decisions imposing harsher sentences on males just because they are males. Gender as a preexisting bias factor would thus be reenforced in the criminal justice system. The same applies to other factors with potential for discrimination based on belonging to a particular group.

This concern has also been voiced in China. AI systems for sentencing in China have mainly been developed by private companies,[58] either hired by courts or on their own initiative. Companies usually claim trade secret protection for their algorithms and refuse to disclose algorithms to courts. Judges therefore cannot know what factors were used and defined as relevant to sentencing and how much weight is given to each factor.[59] Some factors, such as gender, may

[54] *Ibid*.

[55] Zhao, (2017) Jin 01 Criminal Final No.41 (n 21).

[56] Ric Simmons, "Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System" (2016) 4 *MICH. ST. L. REV.* 947, 980.

[57] *Ibid*. See also JD Humerick, "Reprogramming Fairness: Affirmative Action in Algorithmic Criminal Sentencing" (2019) 4 *COLUM. HUM. RTS. L. REV. ONLINE* 213, 244.

[58] Cui, *supra* note 24, 100. Inviting private companies to develop AI for the judicial system has been supported by the President of the Chinese Supreme Court. Zheng, *supra* note 25, 683.

[59] Zheng, *supra* note 25, 683.

increase the accuracy of algorithms[60] but may at the same time institutionalize gender-based discrimination. Private companies may nevertheless program AI to consider such factors to match precedents better with the targeted case and thus to increase the accuracy of their products. With this business incentive in mind, private companies may even ignore the instructions given by the courts and adopt those biased factors "secretly". This cannot be prevented unless companies must disclose the algorithms used. Moreover, technological companies tend to invest more resources in more profitable areas than criminal justice. As a consequence, the quality of algorithms designed for the judiciary cannot be guaranteed.[61]

It should be noted, however, that the company producing "Little Judge Bao" claims on its website that only factors provided in official sentencing guidelines are programmed into its algorithms and that factors such as gender and socioeconomic status are not used. This policy reflects the principle that only legislatures and courts should determine which factors should be considered in sentencing and which factors should consequently be included in AI algorithms.

Summing up, AI needs proper standards if it is to provide useful clues for sentencing decisions. Moreover, it should be kept in mind that uniformity of sentences is not the only criterion for fairness in sentencing, which should also allow for individualization.[62] Proper sentencing involves common sense, wisdom, and a concept of justice, which exceeds the capacity of AI.[63]

2.5 *Low Quality of Judicial Data Used by AI*

2.5.1 *Incomplete Judicial Data Sets*

In 2014, the Chinese Supreme Court introduced the database "China Judgements Online",[64] where a large number of Chinese court decisions can be found. Almost all AI systems used in the Chinese judicial system have been developed from this database. However, the number of court decisions collected in "China Judgements Online" is probably only half the number of all decisions made by Chinese

[60] Simmons, *supra* note 56, 970.

[61] Zheng, *supra* note 25, 684.

[62] Huang, *supra* note 8, 138.

[63] Anne von der Lieth Gardner, *AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING* (MIT Press 1987), pp. 59–60.

[64] Homepage https://wenshu.court.gov.cn/ accessed 21/03/2022.

courts each year. For example, according to the working report for 2020 issued by the Beijing High Court, Beijing courts in 2020 closed 836,514 cases,[65] but only 611,483 decisions can be found on "China Judgements Online".[66] The difference can be explained by the fact that some judgements are not permitted to be published for reasons of privacy, trade secrets, national secrets, involving minors, or having been resolved by settlement; serious duty-related crimes are also normally not published.[67] Moreover, courts have some discretion as to submitting cases to "China Judgements Online".[68]

The percentage of published decisions among all judgements in criminal matters differs among individual provinces. Beijing courts in 2020 decided 18,703 criminal cases,[69] 15,540 (83.1%) of which can be found in "China Judgements Online".[70] By contrast, Tibet courts closed 31,701 cases in 2020, among them were 2907 criminal cases.[71] Only 1172 (40.3%) of these cases were published in "China Judgements Online".[72]

These statistics show that provinces are not represented equally in "China Judgements Online". China is a big country with diverse cultures and various ethnic groups. If only relatively few cases decided in areas with large minority populations, such as Tibet, are published, they may be superseded by a large amount of data from other areas. As a consequence, the special social values and traditions of regions like Tibet may be disregarded in the algorithms of AI. Given the differences in reporting, AI may conclude that people from

[65] See https://www.bjcourt.gov.cn/article/newsDetail.htm?NId=175002668&channel=100001012 accessed 21/03/2022.

[66] See https://wenshu.court.gov.cn/website/wenshu/181217BMTKHNT2W0/index.html?pageId=01f6fd40936e5b4c366ec5e7d671eef6&s8=02 accessed 21/03/2022. This number includes decisions in all instances and rulings on procedural issues.

[67] Weimin Zuo (左卫民), "Towards Legal Research with Big Data (迈向大数据法律研究" (2018) 40 *CHINESE JOURNAL OF LAW (*法学研究)139, 142.

[68] However, the Beijing High Court claims that it publishes 99.9% of those cases which can be published. See https://www.bjcourt.gov.cn/article/newsDetail.htm?NId=175002668&channel=100001012 accessed 21/03/2022.

[69] Working report for 2020 issued by Beijing High Court: https://www.bjcourt.gov.cn/article/newsDetail.htm?NId=175002668&channel=100001012 accessed 21/03/2022.

[70] https://wenshu.court.gov.cn/website/wenshu/181217BMTKHNT2W0/index.html?pageId=01f6fd40936e5b4c366ec5e7d671eef6&s8=02 accessed 21/03/2022.

[71] https://www.thepaper.cn/newsDetail_forward_11028621 accessed 21/03/2022.

[72] https://wenshu.court.gov.cn/website/wenshu/181217BMTKHNT2W0/index.html?pageId=01f6fd40936e5b4c366ec5e7d671eef6&s8=02 accessed 21/03/2022.

certain provinces commit more crimes because more criminal decisions from this province can be found in its database, thus AI may increase sentences for people from this province. In addition, local parliaments may set up different amounts of monetary damage as thresholds for the incrimination of economic crimes, depending on the average income of local inhabitants.[73] However, if AI follows the *Guidelines* and uses only local data for calculating sentences,[74] the goal "same case, same sentence" on a national level is missed. There are reports on conflicting results on the same case produced by different AI systems adopted in different provinces.[75]

2.5.2 *Problematic Judicial Data*

Algorithms work accurately only if they are based on accurate data. Yet, the quality of decisions published in "China Judgements Online" cannot be guaranteed.

First, some decisions published in "China Judgements Online" applied the law wrongly and therefore should not be considered. For example, there is a special phenomenon in Chinese criminal justice called "hard action" ("严打"). The central government sometimes orders the police and prosecutors to take more severe actions against certain types of crime. For example, "hard action" was ordered against organized crime in 2018.[76] During the period of "hard action", many judgements on the targeted crimes are handed down, and judges tend to impose more severe sentences than usual, which cannot

[73] Para. 2 of Art. 1 of Interpretation of the Supreme People's Court and the Supreme People's Procuratorate on Several Issues concerning the Application of Law in the Handling of Criminal Cases of Theft (最高人民法院、最高人民检察院关于办理盗窃刑事案件适用法律若干问题的解释) provides: "The higher people's courts and the people's procuratorates of all provinces, autonomous regions and municipalities directly under the Central Government may, in light of the economic development status of their respective regions, and in consideration of the social security situation, determine, within the scope of the amounts specified in the preceding paragraph, specific amount standards for their respective regions, and report them to the Supreme People's Court and the Supreme People's Procuratorate for approval." A list of incriminating thresholds for theft in different Chinese provinces can be found: http://www.gztingjun.com/m/view.php?aid=615 accessed 14/09/2021.

[74] Para. 1, No.4 of the *Guidelines* provides that judges should consider the economic situation and the imposed sentences for similar cases in their area. This means that sentencing for similar cases in different areas can differ.

[75] Zuo, *supra* note 23, 28–9.

[76] Notice on Special Combat against Criminal Clan (关于开展扫黑除恶专项斗争的通知) http://www.gov.cn/zhengce/2018-01/24/content_5260130.htm accessed 14/09/2021.

be regarded as representative. If AI learns from such "special" judgements, the influence of these excessive sentences remains in the system even after the termination of a "hard action" campaign.

Second, published decisions do not always present a complete picture of a case. Some district courts edit the original judgements for publication, deleting paragraphs which may cause protest.[77] Factors with an impact on the sentence may be included in those deleted paragraphs. Moreover, judges may not mention all factors they consider in making sentencing decisions,[78] such as local protectionism,[79] criminal policy, political elements such as "hard action", opinions of the public, or instructions from higher judges.[80] Some of these factors are even prohibited from being considered. Some discriminatory factors, such as gender and employment status, may also play a role in decision-making but will not be mentioned.

Third, it is not rare that different sentences are imposed in similar cases. Some of these conflicting decisions can still be found in the database and may confuse AI.[81]

Fourth, many decisions do not include any legal reasoning, which obviously decreases the value of a decision for the machine-learning of AI. A criminal judgement may just describe the facts and list the evidence considered and close with the statement, "Defendant A committed XX crime. In accordance with Art. XX of Chinese Criminal Law, A is sentenced to XX years of imprisonment." Courts give no reasons why a provision of the CCL is to be applied and what impact particular circumstances have on the sentence. Moreover, circumstances mentioned in the decisions are, to a large degree, limited to the ones provided by the CCL and sentencing guidelines, and there is no individualized reasoning on their application.[82] If aggravating or mitigating circumstances exist, judges tend to just

[77] Zhang, *supra* note 46, 66.

[78] Weimin Zuo (左卫民), "Several Considerations on Prospect of the Application of Legal Artificial Intelligence in China (关于法律人工智能在中国运用前景的若干思考)" (2018) 12 *TSINGHUA LAW REVIEW (*清华法学)108, 114–7. See also Zheng, *supra* note 25, 679.

[79] Zuoxiang Liu (刘作翔), "Criticism on Local Protectionism in Chinese Judiciary (中国司法地方保护主义之批判)" (2003) 1 *RESEARCH ON LAW (*法学研究) 83, 90–4.

[80] Zhang, *supra* note 46, 67.

[81] Id 66.

[82] Xi Si (司旭) and Jin Wang (王进), "Research on Reasoning on Sentencing in Criminal Decisions (我国刑事判决书量刑说理问题研究)" (2018) 159 *JOURNAL OF SHANDONG ACADEMY OF GOVERNANCE (*山东行政学院学报) 76, 78.

mention them before listing the applicable legal norms but again without giving reasons. Many judgements list the same circumstances or even text blocks for explaining the sentence. It is very rare for courts to explain the amount of fines or the length of probation. An empirical study on 75 criminal judgements showed that fines were imposed in 73.3% of the sample, but there were no reasons given for the use and the amount of the sanction.[83] Only 27 out of 75 judgements (36.0%) were found to have "sufficient reasoning on sentencing".[84] Furthermore, Art. 201 of the *Chinese Criminal Procedure Law* (2018) gives binding effect to sentencing recommendations made by prosecutors in plea bargaining cases; judges "shall in principle" follow these recommendations.[85] In pilot cities, 53.7% of cases were resolved through plea bargaining between 2016 and 2018,[86] and judges followed the sentencing recommendations in more than 90% of these cases. In the first half of 2020,[87] the percentage of plea-bargaining cases increased to 82.2%. Accordingly, more and more decisions just state that "the sentencing recommendation made by the prosecutor is appropriate……"[88] or "the bench confirms the sentencing recommendation……"[89] without any independent reasoning. Decisions without reasons or with block texts on sentencing contribute little to improving the accuracy of algorithms for sentencing because they provide no useful information to sentencing AI.

---

[83] Yueqin Jiao (焦悦勤), "Survey on Reasoning on Sentencing and Research on the Reform Approach (刑事判决书量刑说理现状调查及改革路径研究) (2016) 34 *HEBEI LAW (*河北法学) 75, 79.

[84] Id 77. However, the author did not explain his standards on determining "sufficient reasoning" or "insufficient reasoning".

[85] Jiahui Shi (石家慧), "Reconsideration of the Role of Prosecutors in the Chinese Plea Bargaining System: A Comparative Perspective" (2021) 10 *CHINESE STUDIES* 88, 90.

[86] Yunteng Hu (胡云腾), *INTERPRETATION AND APPLICATION OF CHINESE PLEA BARGAINING (*认罪认罚从宽制度的理解与适用) (People's Court Press 2018), pp. 271–8.

[87] Zhaokun Shi (史兆琨), "Several Indexes on Procedural Supervision increase to Provide Legal Protection (多项诉讼监督指标"止降转升", 努力为大局稳定提供法治保障)", *Daily of Prosecution (*檢察日報) (Beijing, 24/07/2020) https://news.sina.com.cn/c/2020-07-24/doc-iivhuipn4773042.shtml accessed 14/09/2021.

[88] For example, (2021) Liao 0682 Criminal First Instance No. 188 ((2021) 辽0682刑初188号).

[89] For example, (2021) Liao 0106 Criminal First Instance No. 619 ((2021) 辽0106刑初619号).

Decisions on sentencing can also quickly become obsolete, either because the statute is amended or because courts interpret its provisions differently. This problem is especially dramatic due to the fast changes in Chinese society and economy in the past decades. For example, the CCL was amended eleven times in the past twenty years and four times in the last ten years. In addition, many judicial explanations were issued, amended, and invalidated by the Chinese Supreme Court. The cases decided according to old versions of legislation or judicial explanations can be regarded as "invalid" and should be excluded from the database of AI. One good example of a new interpretation is the self-defence clause in the CCL, which was regarded as "dead" for a long time; persons defending themselves were often given disproportionally severe sentences.[90] Recently, two cases on self-defence attracted public attention and were discussed intensively.[91] In response to public concern, the Chinese Supreme Court encouraged lower courts to use the self-defence clause more liberally. As a result, more decisions applying the self-defence clause have occurred recently. The older cases thus have little value and may even confuse the algorithms.

2.6 *Reduced Acceptance by the Public*

The objections listed above focus on the legitimacy and accuracy of AI in sentencing. Another important problem to be considered is whether the use of AI in criminal justice can gain support from the public.[92]

Especially in the common law world, theorists claim that procedural justice is not about "actual" fairness but about the people's perception, namely, whether the public and the participants of the process *believe* the procedure to be fair.[93] The outcome of a fair procedure is normally considered as fair and receives high acceptance.[94] Since people are more likely to believe that a procedure in

[90] Zhang, *supra* note 46, 66.

[91] Jiahui Shi (石家慧), "Self-defence in German Criminal Law (德国刑法中的正当防卫制度)" (2018) 6 *CHINA REVIEW OF ADMINISTRATION OF JUSTICE (*中国应用法学) 173,173.

[92] Ric Simmons, "Big Data and Procedural Justice: Legitimizing Algorithms in the Criminal Justice System" (2018) 15 *OHIO ST. J. CRIM. L.* 573, 573.

[93] Id 574.

[94] Tom R. Tyler, *WHY PEOPLE OBEY THE LAW* (Princeton University Press 1990), p. 109.

which they participate is fair,[95] better participation in criminal justice is essential for acceptance by the public. This consideration suggests an interest in greater transparency of the decision-making process.[96] Although Chinese criminal procedure, which is organized in line with the inquisitorial principle, emphasizes the search for an "objectively" fair sentence, Chinese authorities also recognize that it is important to make the fairness of the process perceptible in order to preserve the credibility of the legal system and to increase the public's acceptance of court decisions.[97] Therefore, the Chinese Supreme Court in 2013 sought to increase the transparency of criminal justice through a project called "Judiciary under Sunshine" to improve judicial transparency,[98] for example, by introducing live-streaming of trials.[99]

The extensive use of AI may, however, reduce the current level of participation and transparency in criminal justice. The "black box" of algorithms is one of the key problems in AI's decision-making process.[100] AI translates all activities at the trial, such as arguments, cross-examination, or a defendant's sincere apology, into a single item on the list of sentencing factors. The weight of these factors is calculated by AI algorithms, which are not under the control of judges. Although arguments of parties may not have a great impact even when judges make the decision, parties feel that they are being

[95] Simmons, *supra* note 92, 579–80.

[96] Ric Simmons, "Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System" (2018) 52 *UNIVERSITY OF CALIFORNIA DAVIS LAW REVIEW* 1067, 1087.

[97] Section 4 of *Dec*isions *on Several Important Issues on Promoting Rule of Law*(中共中央关于全面推进依法治国若干重大问题的决定) mentions that "……try to make the public feel fairness and justice in every case". http://www.gov.cn/zhengce/2014-10/28/content_2771946.htm accessed 14/09/2021. For a similar opinion see Tyler (n 94) 107, stating that "fair procedures can act as a cushion of support when authorities are delivering unfavorable outcomes."

[98] See https://www.chinacourt.org/article/detail/2016/11/id/2336215.shtml accessed 14/09/2021.

[99] Some authors argue that live-streaming of trials has a negative effect on truth-finding. Weimin Zuo (左卫民), "Rethinking Live-stream of Trials: From the Perspective of Judicial Transparency (反思庭审直播 – 以司法公开为视角)" (2020) 9 *POLITICAL SCIENCE AND LAW (*政治与法律) 91, 98.

[100] Johannes Kaspar, Katrin Höffler, Stefan Harrendorf, "Datenbanken, Online-Votings und künstliche Intelligenz" (2020) 32 *NEUE KRIMINALPOLITIK* 35, 51–2.

heard and are thus more likely to accept the final decision.[101] If sentencing means "judges hear, but AI decides",[102] participation in the trial loses its meaning since no relevant interaction with the real decision-makers takes place. Consequently, the public and all parties of the criminal process, including judges, may lose trust in the fairness of the procedure. Moreover, if judges regard their role in sentencing only as marginal due to the involvement of AI, they may even lose any motivation to listen to what participants argue at trial, which in turn worsens the situation.

Therefore, if Chinese authorities wish to preserve public trust in the criminal justice system, they need to guarantee the same level of transparency and participation for AI-supported procedures as in the traditional process.

III THE ROLE OF AI IN FUTURE SENTENCING

Given all these problems and shortcomings, should AI still be considered a useful tool for sentencing? As mentioned above, Chinese authorities are very keen to promote the use of AI in criminal justice[103] and consider AI systems as an important tool for realizing the goal of "same case, same sentence". Given the current political atmosphere, it is not possible to reverse this trend. The primary issue thus can only be how to regulate the use of AI in sentencing to ensure a better, or at least not a worse, sentencing system.

3.1 *To Guarantee the Final Decision of Judges and to Use AI as a mere "Assistant"*

Chinese authorities encourage judges to use AI, but they do not explain what role AI should play and to what degree AI should be used. Such questions should be first answered when discussing the application of AI in sentencing.

It is a fundamental principle of modern legal theory to exclusively entrust the judge as a neutral party with making decisions on con-

[101] *See, e.g.*, E. Allan Lind, Ruth Kanfer, P. Christopher Earley, "Voice, Control, and Procedural Justice: Instrumental and Noninstrumental Concerns in Fairness Judgements" (1990) 59 *J. PERSONALITY & SOC. PSYCHOL.* 952.

[102] Yujie Zhang (张玉洁), "Judicial Application on Sentencing Algorithms: Logic, Difficulties and Procedure Responses (智能量刑算法的司法适用: 逻辑、难题与程序法回应)" (2021) 3 *ORIENTAL LAW (*东方法学) 187, 194.

[103] Zheng, *supra* note 25, 678.

viction and sentencing. Art. 131 of the *Chinese Constitution* grants courts the authority to "exercise judicial power independently, in accordance with the provisions of law, and not subject to interference by any administrative organ, public organization or individual." This principle is a cornerstone of the criminal justice system, which gains a new dimension in that AI should not impair the independence of judges in sentencing. AI should only play a role as an "assistant" to judges instead of taking over their job. This is also the current practice, which is supported by the common view of Chinese scholars and practitioners.[104] Although some writers strongly support the use of AI in sentencing,[105] judges should not be legally obliged to implement the results produced by AI. Leaving the final decision on sentencing in the hands of human judges is the best weapon against an unwarranted interference of AI and for protecting the independence of judges.

To serve this purpose, the judicial reform goal of "same case, same sentence" should not be interpreted rigidly or be used as a ground for restricting the discretion of judges to an undue degree. The individualization of sentences based on the facts of each case should be the result of a thorough assessment of "blameworthiness through a combination of complex fact-finding, equitable discretion, and mercy."[106] In doing so, judges can turn to AI for advice by searching similar cases with the help of keywords and use suggested decisions as a point of reference in order to prevent obvious disparities. The alarm system mentioned above has the same purpose. In such a case, AI functions more or less as a "legal database" or data analyst. Judges can review the outcomes worked out by AI to decide whether those cases are really "similar" or have any relevance to the case at hand. Judges should, nevertheless, remain free to decide whether they follow the suggestions of AI or prefer to decide otherwise.

---

[104] Weidong Ji (季卫东), "Changes of Judicial Power in AI Era (人工智能时代的司法权之变)" (2018) 1 *ORIENTAL LAW (*东方法学) 125,132.

[105] Hang Zhen (甄航), "Artificial intelligence intervention in sentencing mechanism: dilemma, orientation and deconstruction (人工智能介入量刑机制:困境、定位与解构), *JOURNAL OF CHONGQING UNIVERSITY (SOCIAL SCIENCE EDITION) (*重庆大学学报(社会科学版)) (first published online 18/12/2020). https://doi.org/10.11835/j.issn.1008-5831.fx.2020.12.003. The author argued that AI should "regulate" the behavior of judges which means that judges must give explanations when they do not want to follow the outcome offered by AI.

[106] Roth, *supra* note 51, 1247.

Judicial discretion is indispensable because only judges can properly understand the social values behind a case. Compared to the issue of conviction, which is a decision on facts, sentencing requires legal judgement to decide on issues such as blameworthiness, equity, mercy, and human dignity.[107] Therefore, relying on AI, which cannot "understand" these issues, to ultimately decide on sentencing undermines the necessary human element in criminal justice. Research has shown that the application of AI in any field tends to dehumanize the decision-making process,[108] and dehumanization in criminal justice is more harmful than in other areas. Therefore, it is essential to grant judges the authority to make value judgements in sentencing.[109]

Moreover, to prevent judges from becoming psychologically dependent on AI results, AI should not be designed to directly suggest a number of years or months of imprisonment. Judges are not always confident in their own decisions, especially when they reach a different conclusion from the one suggested by AI, which is generally believed to be more accurate.[110] To avoid psychological stress stemming from the need to decide on other people's fate, judges may wish to just follow the suggestions from AI, disregarding their responsibility for independent decision-making. They can then blame AI for any wrong decisions. By contrast, if they do not follow the suggestions from AI, they may be criticized more harshly than in a system without AI. This risk is extremely high in the Chinese judicial system, where individual judges do not enjoy independence. If a Chinese judge imposes a lighter sentence than the one suggested by AI, he or she may have to give a detailed explanation; the judge can even be disciplined, fired, or accused of corruption. To avoid that risk, judges in China are likely to just follow the suggestion made by AI without asking questions. Therefore, AI should never present a particular sentence for the judge to impose. Instead, AI could suggest "similar" cases, calculate a range of sentences based upon big data from previous cases, or present the possible impact of aggravating

[107] Simmons, *supra* note 96, 1096.

[108] Lee Rainie and Janna Anderson, Code-Dependent: Pros and Cons of the Algorithm Age, Pew Res. Ctr. (Feb. 8, 2017) https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/ accessed 14/09/2021.

[109] Stephen E. Henderson, "A Few Criminal Justice Big Data Rules" (2018) 15 *OHIO ST. J. CRIM. L.* 527, 534.

[110] See Grove, *supra* note 2, 19.

and mitigating circumstances. Judges then need to make the final decision on sentencing and take responsibility for it. Moreover, to further reduce psychological dependence on AI decisions, a duty to explain should be imposed on judges even if they follow AI's suggestions. Judges should not be able to simply argue that "AI decided so" but should give reasons why *they* think the imposed sentence is appropriate in the individual case.

3.2 *To Increase the Quality of Judicial Data*

The judicial data used by Chinese AI have various problems: a large number of decisions are unpublished; different locations and different types of crimes are over- or under-represented; differing legal terminologies are being used; many decisions contain no reasons; the sources of decisions are sometimes unclear; and some decisions are conflicting or outdated. Since operating with data of good quality is a precondition of a reliable AI system, the database "Chinese Judgements Online" should be improved and expanded. For example, the Chinese Supreme Court should require courts to publish their decisions completely and continuously with only limited exceptions, such as when cases involve national secrecy, trade secrets, or private interests. The Chinese Supreme Court made one step forward toward promoting good reasoning in 2018 by issuing the *Guidelines on Strengthening and Standardizing the Analysis and Reasoning in Adjudicative Instruments*.[111] The Supreme Court also requires courts on various levels to make the quality of reasoning a factor in the evaluation of the performance of judges. However, since this document applies mainly to reasoning on legal issues and fact-finding, it has only limited relevance to sentencing. In addition, although this document has an internal effect, its wording fails to provide clear standards on what is good reasoning; it is therefore not possible to evaluate the quality of reasoning on the basis of these Guidelines.[112] As a result, its effect is only "advisory". The Supreme Court should consider issuing specific instructions on giving reasons in sentencing, for example, obliging judges to mention all factors considered for the sentence and to explain in detail how these factors have influenced the decision. Only with such information can a judgement be good

[111] 最高人民法院关于加强和规范裁判文书释法说理的指导意见，法发[2018]10号 http://www.court.gov.cn/zixun-xiangqing-101552.html accessed 15/09/2021.

[112] Some courts reward judges if their decisions are selected by the Supreme Court as "Model Cases".

material for training AI. Apart from issuing such guidelines, more systematic efforts are needed to improve the judges' ability on reasoning, e.g., introducing pertinent clinical courses in law schools and offering more professional training to judges. Moreover, the legislature could consider imposing a duty to explain judgements in the *Judge Law of the P.R. China*.[113] That would give the Supreme Court a solid legal basis for promoting good reasoning.

It is a long-term task to train judges and to make them accustomed to giving substantive reasons in their judgements. An effect of training courses cannot be expected within a short time, but the use of AI in sentencing practice cannot wait. Therefore, making the best use of available data is the most practical solution to improving AI within a short time. The department of the Chinese Supreme Court that runs the database "Chinese Judgements Online" should establish a special working committee on AI (hereafter referred to as AI Committee), which could consist of experienced judges appointed by the Supreme Courts, scholars as consultants, and IT engineers as assistants to solve technical problems. The legal experts in this committee should continuously review decisions published in this database to ensure the quality of decisions used by AI. They should mark those decisions which were overturned by higher courts, are outdated because of amendments on legislation or judicial explanations, applied the law incorrectly, or contain no useful reasoning.[114] The programmers of AI who use data from "Chinese Judgements Online" should not import those marked decisions into their products and should frequently update the database of AI.

3.3 *To Improve the Transparency of Algorithms and to Ensure Their Reviewability*

The "black box" of algorithms of AI is one of the main concerns regarding the use of AI in sentencing. Judges have at best a very limited ability to review the maths used in AI algorithms.[115] Greater transparency of algorithms is therefore essential both for substantive

[113] The current version of the *Judge Law of the P.R. China* (中华人民共和国法官法) does not contain any obligation to give reasons for judgments.

[114] These marked decisions, including the decisions applying the law incorrectly, should still remain in the database "Chinese Judgements Online". Although they have little value for AI they could still be used for other purposes, such as academic research and public supervision.

[115] Jason Tashea, "Calculating Crime" (2017) 103-*MAR A.B.A. J.* 54, 58.

and procedural justice.[116] For substantive justice, to disclose what factors are considered by algorithms makes it possible for judges to review whether any of these factors contains a bias, which may lead to unreasonably higher sentences for defendants belonging to certain groups. Regarding procedural justice, a transparent decision-making process of AI combined with the final determination to be made by the judges can strengthen the trust of the public in the fairness of criminal justice and increase the acceptance of decisions by the participants.[117] Imposing a duty on judges to give a substantive explanation would also reduce the "black box" character of sentencing by algorithms.

The transparency of algorithms should be guaranteed in two ways. First, programmers of AI systems should be instructed in clear terms what factors may be embedded in algorithms. Currently, factors used by AI for sentencing mainly reflect those provided in the national and local sentencing guidelines. However, these documents do not cover all offence categories of the CCL. Therefore, the AI Committee should first work out sentencing guidelines for offences not listed in the *Guidelines*. If further studies show that certain factors should not play a role in sentencing, AI programmers should be instructed to rewrite algorithms accordingly. Before courts put the system into use, they should require programmers to disclose the algorithms to them and examine whether their instructions have been followed.[118] AI should also be required to disclose the grounds for each suggestion they make, including the factors used and the weight given to each factor. Only with such information can the judge review whether the suggestion given by AI is proper, whether any additional factors should be considered, and whether the weight given to each factor should be adjusted. This is important for guaranteeing fair sentencing because AI is likely to ignore factors that have not been embedded in its programming[119] and may misunderstand textual similarities between court decisions.[120]

[116] Simmons, *supra* note 96,1087.

[117] See Lind et al., *supra* note 101.

[118] Hunter, *supra* note 3, 785.

[119] Sun, *supra* note 5, 89.

[120] Mandal et al., *supra* note 47, 418.

A good example of transparency of algorithms is the PSA program used for bail decision-making in the U.S.[121] This system, using a dataset of more than 500,000 cases from multiple jurisdictions in the U.S., discloses all factors it uses and how these factors are weighted and scaled.[122] PSA was proved to lower the rate of pretrial detention by 20% in the city of Charlotte, with no increase in crime or bench warrants, and helped to increase the accuracy of pretrial decisions in the state of Kentucky.[123] Moreover, studies showed that using PSA does not increase racial disparities.[124]

In sum, quality control of AI systems before their use and the reviewability of their outcomes must be guaranteed in order to make AI-supported sentencing sufficiently transparent.

3.4 *To Centralize AI Systems with Quality Control*

Implementing the policy of promoting "Intelligent Courts", Chinese courts seek to use their own AI systems. Especially in highly developed areas such as Beijing and Shanghai, no court wants to fall behind in this "competition". This has the consequence that there exists no centralized system; various courts in different provinces use AI systems with different functions and focuses.[125] This diversity negatively affects the "same case, same sentence" project.[126] The decentralization of AI systems also makes it more difficult to guarantee the quality of AI. It would therefore be better to centralize the AI system in criminal justice under the guidance of the Chinese Supreme Court and the AI Committee suggested above. Experts who operate the "Chinese Judgements Online" would thus have an overview of available information. The Committee would have expert staff to effectively supervise AI programmers. These experts could also examine the algorithms developed by IT firms before they are being put into use. They could use "model cases" to test the accuracy

[121] https://advancingpretrial.org/psa/about/ accessed 14/09/2021. The PSA was developed by Arnold Ventures and more than 40 jurisdictions across the country have implemented the tool. See https://www.nmcourts.gov/court-administration/pretrial-release-and-detention-reform/public-safety-assessment-for-pretrial-release-and-detention/ accessed 14/09/2021.

[122] https://advancingpretrial.org/psa/factors/ accessed 14/09/2021.

[123] Simmons, *supra* note 56, 965–6.

[124] https://advancingpretrial.org/psa/about/ accessed 14/09/2021.

[125] Even the system developed by the Chinese Supreme Court used to refer to "same cases" is not a centralized one and has not been adopted by every court.

[126] Zuo, *supra* note 23, 28–9.

of AI. Moreover, the Committee should prepare a handbook for judges on the use of the system, including information on the rate of accuracy and the risks inherent in the system. Judges could then determine individually to what degree they wish to follow suggestions by AI in their decisions on sentencing. Judges who work in areas with a large portion of minority inhabitants should be instructed that their areas might be underrepresented in the database and that they should take local particularities into consideration when making sentencing decisions. If there are no local conditions to be considered, judges should refer to nationwide data collected by the centralized system. They could still be permitted to determine sentences referring only to cases from their own province. The AI Committee should also offer training courses for judges to familiarize them with the AI system. For continuously improving the system, it should have an "error report" function permitting judges to report to the AI Committee any mistake or inaccurate outcomes.

IV CONCLUSION

Chinese policy makers have introduced AI to reduce judicial bias in sentencing and to implement the goal "same case, same sentence". In the future, AI will play an increasing role in Chinese sentencing practice. However, "the potential for good is huge, but the potential for misuse and abuse – intentional, and inadvertent – may be greater."[127] One should not forget that all AI systems have been developed by humans. By filtering, analyzing and transforming the input data, they "deliver the patterns that preexist in our society, many of which are undesirable and even widely unknown to the members of the society."[128] Moreover, AI lacks the capacity of making decisions based on value judgements, and the non-transparency of algorithms reduces the trust of the public in the fairness of criminal justice. The positive effect of AI on the goal of "same case, same sentence" therefore should not be overstated, and an unlimited application of AI must be avoided. Given that the application of AI in the judicial system is still at the initial stage in China, now is the best time to design the way in which AI is employed for sentencing, before

[127] Rainie and Anderson, *supra* note 108, accessed 14/09/2021.

[128] Kia Rahnama, "Science and Ethics of Algorithms in the Courtroom" (2019) 2 *JOURNAL OF LAW, TECHNOLOGY AND POLICY* 169, 186.

inappropriate use of AI shapes the interactions between the relevant social groups and makes the system harder to deconstruct.[129]

Chinese policy makers should take care to regulate the use of AI in criminal justice and should integrate AI into sentencing only with great caution. AI should be employed not as a decision-maker but only as an "assistant", providing information for judges and aiding them in making sentencing decisions. The final determination should in any event remain in the hands of the judge. Moreover, algorithms should be made transparent so that judges can review their operation. A Committee supervised by the Chinese Supreme Court should be established to guarantee the quality of judicial data on "Chinese Judgements Online" and to operate a centralized AI system on sentencing for the whole Chinese court system. All these measures would help to make the best use of judicial data and to introduce a fair, accurate, and efficient sentencing system.

ACKNOWLEDGEMENTS

I would like to express my special thanks to Prof. Dr. Thomas Weigend at the University of Cologne for encouraging me to initiate this project and for offering valuable comments and feedback on this manuscript.

DECLARATIONS

CONFLICT OF INTEREST

I have no conflict of interest to disclose.

[129] *Ibid*. See also Mike Ananny, "Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness" (2016) 41 *SCI. TECH. & HUM. VALUES* 93, 96.