

# 人工智能在司法领域面临的困境和规范化革新路径

段泓玉

大连理工大学公共管理学院, 辽宁 大连

收稿日期: 2023年7月20日; 录用日期: 2023年9月5日; 发布日期: 2023年9月12日

## 摘要

人工智能在司法领域的推行与应用是我国司法领域的整体趋势。智能司法既可以突破物理条件的限制实现高效率、多场景的司法实践, 也可以为法官的自由裁量提供参考, 促进司法走向可视化的正义。但是, 智能司法在运行中也存在诸多局限性, 如算法歧视与黑箱、数据采集的有限性、现实应用的机械性。因此, 需要明确智能司法辅助而非替代的基础定位, 培养综合人才促进算法开发, 设立完备程序规范算法运行。

## 关键词

智能司法, 算法歧视, 同案同判, 价值评判

# The Dilemma and Standardized Innovation Path of Artificial Intelligence in the Judicial Field

Hongyu Duan

School of Public Administration and Policy, Dalian University of Technology, Dalian Liaoning

Received: Jul. 20<sup>th</sup>, 2023; accepted: Sep. 5<sup>th</sup>, 2023; published: Sep. 12<sup>th</sup>, 2023

## Abstract

The implementation and application of artificial intelligence in the judicial field are the overall trend in the judicial field of our country. Intelligent justice not only can break through the restrictions of physical conditions to achieve high efficiency and multi-scene judicial practice, it can also

provide a reference for the discretion of judges, also can provide reliable reference for judges to exercise discretion, and promote the judicial toward visual justice. However, there are also many limitations in the operation of intelligent justice, such as algorithm discrimination and black box, the limitation of data collection and the mechanical nature of practical application. Therefore, it is necessary to clarify the basic positioning of intelligent judicial assistance rather than replacement, cultivate comprehensive talents to promote algorithm development, and set up complete procedures to standardize algorithm operation.

## Keywords

Intelligent Justice, Algorithm Discrimination, Same Case and Same Judgment, Value Evaluation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

人工智能司法日益受到顶层制度设计者的重视,自2017年杭州互联网法院设立方案被通过以来,我国人工智能司法的试验性探索便在各如火如茶地展开。人工智能与司法实践的结合为司法制度的发展提供了新的可能,也为司法公正的实现提供了新的路径。但是,各地试验中遇到了各样的问题,导致地方人工智能司法系统真正推广使用的并不多。因此,如何正确审视人工智能与法律结合带来的机遇与挑战,探索人工智能司法的未来发展道路成为一项重要的实践课题。

## 2. 人工智能司法的发展现状

智能司法是指人工智能对既有的司法裁判进行深度学习,对同类案件的处理共性进行概率建模,辅助法官或者自动进行审判。智能司法既可以突破物理条件的限制实现高效率、多场景的司法实践,也可以对法官进行自由裁量提供可靠借鉴,促进了司法走向可视化的正义。

### 2.1. 为程序公正带来可能路径

随着社会进步与人们法律意识的逐步增强,各级法院受理的案件大幅增长,司法领域出现了有限的司法资源难以满足人们日益增长的司法救济需求的矛盾,单纯增加司法审判人员难以从本质上有效提高司法效率,智能司法的出现为解决这一困境提供了途径。

在立案程序中,人工智能为司法程序注入了新的活力,立案程序由审核制转变为自助化或半自助化的登记制,极大的提高了立案效率。例如,上海市浦东新区人民法院采取扫码立案制度,平均立案时间减少45分钟。在审判程序中,智能图像识别与智能语言识别技术能将纸质资料电子化,人工智能还可以进一步将电子化资料进行提取、分类与归档,自动生成标准化文书为法官分担大量司法基础性工作。人工智能还为司法技术带来了许多可以节点控制与提高办案效率的革新。例如,区块链将电子数据与行为可以进行全节点见证,促进智能司法过程中全要素数据的可视化。司法人工智能创新出的“弹幕”电子送达方式,即法院的消息不会受其余任何软件的影响,被送达人只有查阅送达内容后,才可以对手机的其余权限进行解锁。智能司法确保了信息传达的高效率与有效性,但其对自然人其余权利产生的关联限制是否超过了合理范围仍然值得探索。

近年来,北京互联网法院的电子诉讼平台、浙江“移动微法院”、广州中院的“律师通”和“移动执行”手机 app 等纷纷上线运营,这些小程序可以让人们克服物理条件的限制,实现了司法在虚拟与现实双重空间的运行[1]。例如上海“206 系统”涵盖了刑事案件从立案侦查、审查起诉到判决刑罚以及刑满释放等一系列司法流程,让利害关系人可以足不出户参与在线庭审、调解或执行,形成了灵活便捷的多样司法环境。同时上海“206 系统”拥有全面的自动审查功能,即会对公安机关的逮捕条件、全案证据与检察机关和司法机关的类案推送、办案程序合法性等进行自动审查,智能司法的自动运行机制可以大大提高司法效率。

## 2.2. 为实体公正提供可靠依据

我国长期的司法实践中,出现了大量相同或类似案件在不同的法院产生不同判决的情况,中共中央办公厅于 2017 年印发的《关于加强法官检察官正规化专业化职业化建设全面落实司法责任制的意见》指出要善于运用智能司法的辅助功能[2],促进个案实体公正的实现。各级法院积极进行类案类判智能技术的探索,例如,最高法院开发裁判剖析大数据引擎与同案智推大数据引擎,为法官行使自由裁判权提供类案借鉴[3]。

具体而言,人工智能从类案推送、量刑辅助与偏离预警三大功能来辅助实现个案实体公正。类案推送是指人工智能提取案件要点,在数据库中搜索汇总类似判决,推送给法官凝聚着群体智慧的裁判思路。量刑辅助是指人工智能将识别到的案件要点投入到其通过深度学习后构建的算法中,进而向法官提供基础的量刑建议。偏离预警是指当法官做出了与算法计算结果相差较大的审判结果时,人工智能会给出不同程度的预警。人工智能的三大功能从理性的角度为法官提供凝聚着法官对法律的均值理解,提供了法官对案件的均值判断思路,促进法官以统一的司法尺度做出科学理性的审判,有效防止自由裁判权的滥用。

## 3. 人工智能在司法领域存在的困境

### 3.1. 算法的歧视与黑箱

代码作者将法律规定转化为代码,代码编写的过程杂糅着代码作者对法律规定的个人偏好和价值判断,编写的代码中甚至可能夹杂着企业对于自身利益的考量、政治因素等,即代码作者站在特定角度对法律规定进行了二次的解释。例如,通常代码是在司法机关的主导下进行编写,群众无法实现其参与、监督的权利,代码可能会自然的偏向司法机关的立场,不可避免的导致司法权范围的扩张。经济发展的不平衡也会导致不同地区数据收集水平不同,同类事实在不同经济水平的地区会产生不同的处理结果,不全面的数据会导致算法计算结果在特殊地域出现不合理的情形。

由于代码由特定的人群进行编写,群众不能参与和监督代码的创制过程,导致了算法黑箱产生。人们只能单纯接受来自算法自动执行后生成的结果,没有途径去了解算法的运作,也没有方法去救济算法单一计算后的结果,给算法霸权的产生提供了空间。即便算法决策中出现不公正的因素,算法黑箱使群众陷入被动的算法霸权之中,可能会引发社会性危机。同时,人工智能虽然可以进行诉讼流程节点控制与完美预测,但是没有缺陷的自动化程序是不存在的,程序员的疏忽导致的系统错误是无法规避的。人工智能的自动执行暗中加强了算法产生结果的权威性,人们对算法运算的逻辑进行抗辩、算法结果的正确性进行质疑的难度增加,事后救济中对算法运行进行解释的需求随着人工智能的应用逐步增强。但是,决策算法的透明化面对着算法作为商业秘密私利和公利的冲突,还有算法如何转化为可理解的自然语言的技术难题。因此,算法运行的解释与算法执行后的救济均应是智能司法应用中值得注重的问題。算法黑箱剧了个人与算法控制者之间信息与权利不均衡的现象,司法机关运用算法将个人置于控制之下,而

算法却很难助力个人私权利的保护，导致了新的不公平问题萌芽。

### 3.2. 数据采集的有限性

一个案件中会涉及到时间、地点、自然人主客观状态等一系列复杂的数据，如果想对一个类型的案件进行算法建模，便需要对若干个案件中的海量数据进行汇总分析。例如，上海“206系统”为了构建盗窃罪的量刑模型，对全国30万个盗窃案件的量刑判决进行了人工智能的深入学习。但是，要将定罪量刑的算法模型覆盖到整个法律体系之中是极具挑战性的。

一方面，智能司法是在中国裁判文书网的数据基础上进行深度学习的，而中国裁判文书网是自2014年才开始进行裁判文书的数据统计，网上收集到的数据不足审结文书的百分之五十[4]，其中也不乏存在一些具有制作不规范、错误或因为法律规则变动、位阶变化而无法被利用的法律文书。同时，公开的数据只有判决书，没有可以体现法官逻辑思路的合议庭笔录、会议报告等，可能无法包含案件法律思路和所有的细节，对类似案件的审理会引发一些错误的指导。

另一方面，人工智能不能单独完成对司法要素高效率与高准确率的标记与提取，现有的自动提取与人工提取两种方式都需要人工的检验确认，人工智能不能自动对不确定的信息进行处理，存在一定的技术壁垒。例如，现有的自然语言识别技术无法满足将法律语言转化为数字代码，因为法律语言不同于自然语言是一种对人行为的概括性描述。

### 3.3. 现实应用的机械性

人工智能缺少人类生活经验与常识，有时候不能够理解人们对话中的真实含义，具体来说它只能对大数据进行统计与预测，没有语言沟通系统。例如，当运用智能语音识别系统时，法官或者律师进行诱导性提问，人工智能并不能理解真正含义，只能机械的进行记录，而此时需要的是书记员的技术性处理。斯蒂芬·图尔敏认为法律议论本质是通过说理和情感的共鸣达成共同的理解，反过来又进一步促进沟通。这种对话过程电脑化的关键是梳理原告与被告之间争论点与妥协点的话语活动的通信协议系统[5]，而人工智能恰恰欠缺这种具有较高理解力的语言处理系统，只依靠智能司法进行审判时难以进行法律议论。人工智能通过深度学习既有案例对未来的案件进行评判，但是社会的快速发展激发了新的思维与行为模式的快速产生，旧的算法应用于新型案例会产生处理偏颇，人工智能缺乏生活经验的缺点将会被放大。

司法人员可能会基于人本身懈怠与推卸责任的心理，只对案件的审判付出最小的精力成本，即对人工智能计算出来的结果盲从，从而导致机械司法。法官的主观能动性与自主决策性会在不断的司法实践循环中遭遇劣币驱逐良币的困境，案件的审判结果愈加向人工智能的算法结果靠拢，法官为了追逐表面的公平正义逐步丧失独立审判的地位，而被人工智能裹挟住，只能做出同案同判的僵化审判。法官的审判中也不能再发现法律的漏洞，无法起到对法律的反哺作用。同时，有学者认为：“严格意义上的同案并不存在，致力于结果平等的同判主张的理论逻辑是非常脆弱的[6]。”在现实应用中，一类案件的处理需要设计多种不同的算法来适应我国司法系统中不同的审级与地域之间的差异性。运用多套算法可以保证不同审级之间的监督功能，帮助应对不同地域的经济文化差异，但是应用多套算法会导致全国司法的统一性受到挑战[1]，差异性与统一性存在较大矛盾。

同时，案件中价值的表现是多样的，可能会涉及自由、平等、效率、公正等多维度不同的方面。价值不仅在不同的人心中有着不同的价值位阶排序，在不同的法中价值的次序有所差别，所以人工智能以最大公约数的方式去试图解决复杂疑难案件中价值冲突的法律问题存在建模的逻辑障碍。复杂的社会环境中会出现多样的情理问题与海量的司法场景，各个机关没有统一的数据处理与交换方式，所以人工智



能的深度学习存在无法穷尽各类情理样本的现实困境。即便人工智能通过深度学习构建了一个算法模型，也不能代替现实中法官自由裁量权的行使，即不能用人工智能计算出来的概率去代替法官根据实际案件进行的权宜平衡，算法提供的计算结果只是可参考性而非决定性。法官在进行价值评判时可以充分发挥其灵活性与创造性，而人工智能只能通过算法的计算对价值做出是或否的机械化评价，很可能在评判中丧失人文关怀。因此，大多数国家仍将办案人员作为审判中心，人工智能作为辅助系统，来维护人文匠心与公平正义。

## 4. 人工智能在司法领域的规范化革新路径

### 4.1. 明确人工智能司法的基础定位

智能司法到来的价值意义不在于替代法官的工作，而是有效的实现司法智慧，更好的保障公民利益，实现法律面前人人平等。智能司法只是一种提供辅助作用的手段，例如人工智能提供的相似案例只是参考而不是根本性决定。因此，明确人工智能在司法审判中的运用界限对实现人工智能司法高效运行起着重要作用。司法实践中审判工作可以分为信息收集整理与审判决策两部分，人工智能可以独立承担前期信息处理方面的工作，也可以辅助法官进行审判。审判中的事实认定环节需要法官的逻辑思维、价值判断与经验法则等观念系统的做出判断，这种集合情理与法制的复杂判断难以单纯用算法来进行替代，智能司法在弱人工智能时代只能限制其作为辅助审判的手段。

而2020年最高人民法院印发的《关于深化司法责任制综合配套改革的实施意见》提出要逐步将人工智能运用于审判阶段，发挥其自动生成法律文书的功能。随着人工智能技术的逐步成熟，智能司法逐步进入审判中心环节，要将自动生成的法律文书责任到法官个人，才能有效避免法官审判权的缺位。同时，也可以对人工智能设置熔断机制来防止其对司法权的僭越，即当遇到特定主体或事项等时，及时做出风险反应来避免人工智能引导的审判盲从。具体而言，人工智能是否能单独审判案件应充分考虑案件的繁琐程度、被告人的意愿、被告人的认罪情况、价值判断的复杂程度等[7]。

### 4.2. 培养复合型人才促进算法开发

算法歧视与黑箱问题的出现反映了人工智能时代复合型人才之匮乏，智能司法的高效应用需要大量精通算法与理解法律的复合型人才。但是，传统的法律教学模式与司法培养方式，使法律人才难以摆脱单一学科的禁锢，对人工智能等科学技术可能只存在概念上的了解。而人工智能领域的人才往往又难以摆脱利益的驱使，相较于高利润行业很少选择投资于法律行业，这就造成了智能司法领域复合型人才缺失的现实困境。有多少人工，才能有多少智能，缺乏复合型人才完成数据提取、整合等基础性工作，将会对智能司法的智能程度进行折损。司法人员基于对人工智能的不了解，可能会产生抵触或过度依赖两种极端心理，良好的人机互动关系也难以建立，第三方的介入也会增加智能司法的应用风险，法律人才的单一培养方式亟需变革。

突破法学单一学科培养的现状，由计算机、法学等理论与实践专家共同制定交叉学科综合性人才的培养方案，为法学生增设计算机基础原理、算法原理等相关课程，鼓励高校与大型科技公司进行教学与实习合作，促使法学生在培养法律逻辑思维的同时掌握计算机编程技术。也要广泛接纳不同学科的学生跨专业进行法律学习，为法律行业的发展储蓄多样化人才。变革原先的司法培养方式，增设对人工智能、大数据、区块链等技术准确定位与时代意义的考察，在司法机关内部培养一批优秀的综合性人才进行数据的提取与汇总、算法的编写与校正，促进算法的开发。也要定期对司法人员进行培训，统一司法机关对时代高速发展变化下的智能司法新情势的理解，帮助司法人员与时俱进的理解算法，更好的发挥出算

法的辅助作用。

### 4.3. 设立完备的程序规范算法运行

智能司法的初衷是规范司法机关人员的行为，高效率的实现全国司法统一，但是机械运算反而促使滋生了许多不公正现象，因此应当设计一套完备的程序对算法进行规范。

首先，设置事前规范与评估机制。对各地区的司法机关进行数据标记与提取的内容进行统一规范，去除各个机关的数据互相使用的障碍，使一套算法可以在司法机关内部贯通使用。司法机关也要对不同部门法的价值位阶进行多样化排序，将法律价值按部门法的分类标准设计出多种算法，使司法机关内部的价值评判可以被智能化的统一规范。算法黑箱引起的偏见、霸权等问题是全方位多领域的，这就需要做好科技创新与私权保护的平衡。对算法的编程设置公开评估机制，组织法学专家、技术专家与社会公众等形成评估委员会，以普通公众可以理解的自然语言公开算法决策的流程，如算法中数据的来源、数据的使用、数据所附的运算比例等，并广泛吸收评审委员会对算法的意见。对算法的事前评审，以公开透明的方式实现了公众的知情权，既保障了算法运行的合法性与合理性，也维护了司法权威，使智能司法更快的被人们理解与信任。同时，事前评估可以为仿照欧盟《一般数据保护条例》第 71 条设立算法解释权打下良好基础。

其次，设置算法检验与纠补机制。算法应经过相关领域专家在未投入应用前期反复的审查与检验，确保算法决策的公正合理。在算法运用中定期开展专家审查，使人工智能及时学习新的思想、行为模式与立法动态等，也及时剔除掉因立法变化或监督审查改判等失去效力的案件，确保算法决策时刻处于符合时代发展的状态中。对实践中复杂的案件不能死板的运用标准化的数据模型，要为可能构建新的裁判规则的特殊案例，如孙志刚案，设置纠补机制留下灵活的裁判空间，防止个案的不公平。

最后，设置算法解释制度。这是保障公众知情权与抗辩权的行使，防范和化解算法歧视、机械化等问题的的重要手段。应仿照欧盟《一般数据保护条例》第 71 条的规定设置算法解释制度<sup>[8]</sup>，即“数据主体有权获得人工干预，表达其观点，获得针对自动化决策的解释并提出异议的权利<sup>[9]</sup>。”赋予受到算法决策影响的公众，请求进行算法解释的抗辩权，使其可以在对算法充分了解的情况下对自己的权利进行主张，尽量保证司法机关与公众处于信息与权利均衡的法律地位。具体到实际应用层面，当司法机关在应用人工智能时，应考虑当事人意愿并明确告知当事人使他们充分了解以便提出抗辩权；当算法决策遭受质疑时，算法开发商也应以专家辅助人等方式参与质证，对算法运行的科学合理性给出相应的证据。

## 5. 结语

人工智能司法为司法实现高效率、多场景的可视化正义带来了可能途径，但是人工智能有着其难以回避的局限性，原先的司法体制与智能化的新生模式之间也会不可避免地产生磨合的阵痛。在人工智能与司法实践相结合时，必须以审慎的态度解决公正与效率、裁量与智能之间的矛盾。我国距离实现智能化的司法审判体系还需要理论与实践的进一步积累，当然我们相信综合性人才的不断培养与完备程序的逐步建立会推进智能司法时代的到来。

## 参考文献

- [1] 马长山. 司法人工智能的重塑效应及其限度[J]. 法学研究, 2020, 42(4): 23-40.
- [2] 郭志远, 王胤. 新时代背景下我国司法责任制的完善[J]. 新疆社会科学, 2018(1): 124-133+168.
- [3] 冯皎, 胡铭. 智慧司法: 实现司法公正的新路径及其局限[J]. 浙江社会科学, 2018(6): 67-75+85+157.
- [4] 高鲁嘉. 人工智能时代我国司法智慧化的机遇、挑战及发展路径[J]. 山东大学学报(哲学社会科学版), 2019(3): 115-123.

- [5] 季卫东. 人工智能时代的法律议论[J]. 法学研究, 2019, 41(6): 32-49.
- [6] 周少华. 同案同判: 一个虚构的法治神话[J]. 法学, 2015(11): 131-140.
- [7] 马靖云. 智慧司法的难题及其破解[J]. 华东政法大学学报, 2019, 22(4): 110-117.
- [8] 孙清白. 人工智能算法的“公共性”应用风险及其二元规制[J]. 行政法学研究, 2020(4): 58-66.
- [9] 魏琼, 徐俊晖. 人工智能在行政处罚中的应用及法律规制[J]. 时代法学, 2021, 19(1): 12-22.