

人工智能司法应用的现实困境与治理路径研究

朱元浩

浙江农林大学集贤学院, 创新创业学院, 浙江 杭州

收稿日期: 2026年4月20日; 录用日期: 2026年5月13日; 发布日期: 2026年5月26日

摘要

人工智能在司法领域广泛应用促进“智慧司法”建设取得重大进展,但是也带来算法黑箱、数据偏见、隐私安全以及伦理失范等问题,在一定程度上造成了司法公正和技术赋能之间的紧张关系。本文基于我国司法信息化发展的背景出发,采用“技术-制度-价值”的三位一体视角,总结我国人工智能司法应用情况、相关法律规定及代表性案例,借鉴美国、欧盟、韩国等国家和地区经验基础上,对算法公开度、制度契合度、价值取向等问题进行探讨。本研究试图寻求符合司法规律的人工智能运行逻辑,以期对智能化法律适用进行规范化指导并给出可行的操作方式,服务于我国社会主义现代化建设和世界范围内的司法智能化治理体系建设。

关键词

智慧司法, 人工智能, 算法透明, 法律规制, 司法公正

Research on the Realistic Challenges and Governance Paths of Artificial Intelligence in Judicial Applications

Yuanhao Zhu

Jixian Honors College, College of Innovation and Entrepreneurship, Zhejiang A&F University, Hangzhou
Zhejiang

Received: April 20, 2026; accepted: May 13, 2026; published: May 26, 2026

Abstract

The wide application of artificial intelligence in the judicial field has facilitated significant progress in the construction of “smart justice”, but it has also brought about issues such as algorithmic black

boxes, data bias, privacy security, and ethical deviations, which to some extent have caused a tense relationship between judicial justice and technological empowerment. Based on the background of the development of judicial informatization in China, this paper adopts a trinity perspective of “technology-system-value”, summarizes the application of artificial intelligence in the judicial field in China, relevant legal provisions, and representative cases. Drawing on the experiences of countries and regions such as the United States, the European Union, and South Korea, it discusses issues such as algorithm transparency, system compatibility, and value orientation. This research attempts to seek the operational logic of artificial intelligence that conforms to judicial laws, in order to provide standardized guidance for intelligent legal application and feasible operational methods, and serve the construction of socialist modernization in China and the global judicial intelligent governance system.

Keywords

Smart Judiciary, Artificial Intelligence, Algorithm Transparency, Legal Regulation, Judicial Fairness

Copyright © 2026 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

伴随着人工智能的发展，世界各国司法机关正在发生一场深刻变革。近年来，《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》¹《法治中国建设规划(2020—2025 年)》²等文件均提出要加强“智慧司法”，最高人民法院也将人工智能视为司法体制改革重要工具[1]。从智能审判辅助系统到类案推送平台，从语音识别庭审记录到区块链存证验真，人工智能已广泛应用于司法各环节中并大大提高了司法工作效率以及公开程度。据不完全统计，2023 年全国各级人民法院所使用智能系统的每日受理案件数量超过三十余万起，而自动撰写裁判文书的比例达到了六成以上。国际上，美国 COMPAS 量刑评估系统、欧盟 AI4JUSTICE 项目、韩国 LONBOT 法律咨询机器人等都体现出了司法与人工智能融合发展的世界潮流。

但是，人工智能应用于司法活动并不是一蹴而就。从技术上来说，算法不透明造成“黑箱裁判”，损害司法裁判的公开性和可追溯性。从制度上来看，目前我国的相关法律法规对于人工智能的监管尚处于空白状态，如算法责任承担、证据资格认可、个人信息安全等问题都未有合理解决办法。从价值观角度而言，技术理性与司法公平的关系日趋紧张，为了追求高效的算法可能会压缩法官自由裁量权，降低司法裁判人文关怀以及道德内涵。最高人民法院 2022 年调研报告显示，在 78% 的基层法院中，智能系统不能满足他们的要求，而如何实现算法决策与法官自由心证之间协调的问题也未能解决。

目前相关研究大多集中在技术层面或者宏观理论阐述上，而忽视司法场景特点。因此，本文从人工智能应用于司法领域所带来一系列问题出发，提出一个由“技术 - 制度 - 价值”构成三位一体的研究思路，在此基础上参考国内外有关事例及统计数据，对算法公开度、制度契合度以及司法价值取向等问题进行讨论，以期找到契合司法规律的人工智能运作方式，从而解决技术赋能与司法公平之间的矛盾，推动中国特色社会主义法治国家建设和世界范围内司法智能化治理体系建设[2]。

¹https://www.gov.cn/xinwen/2021-03/13/content_5592681.htm。

²https://www.gov.cn/zhengce/2021-01/10/content_5578659.htm。

2. 人工智能司法的规范现状与应用场景

2.1. 规范现状

目前,我国的人工智能应用于司法裁判尚处于不断发展过程中,在此背景下,最高人民法院近年也陆续发布了一系列相关规定,例如《人民法院在线诉讼规则》《人民法院在线调解规则》《人民法院在线运行规则》等。前两项规则主要针对在线诉讼和在线调解的案件类型、条件以及程序等方面进行规定,而2022年出台的《人民法院在线运行规则》是在已有智慧法院建设成果基础上对于系统的建设和技术支持、使用方式及运行模式等问题作出补充和完善[3]。这三份文件各有侧重、相互补充,促使形成独具中国特色的在线司法新模式。另外,《最高人民法院关于规范和加强人工智能司法应用的意见》也于2022年发布,在其中人工智能被视为一种辅助工具,明确了其应用的基本原则以及分步实施计划等,对人工智能应用于司法工作起到很好的指导作用[4]。

从立法角度,虽然我国尚未出台专门针对司法人工智能的法律规定,但是人工智能法草案已经连续两年被列入国务院年度立法工作计划之中,表明国家对于推进该领域立法具有极大的积极性。而在今年3月召开的全国两会上就有人大代表提出由全国人大对当前我国人工智能司法应用情况以及所遇到的问题进行研究,制定出《人工智能司法应用促进法》《司法AI伦理准则》等相关法律法规来规范人工智能的应用,使人们更加清楚地了解人工智能可以使用的范围以及不能使用的范围,明确相关的程序,划分责任,加强监管,使人工智能健康发展。如表1所示,从2016年开始,我国出台了一系列规章制度不断推进“智慧法院”的发展,为人工智能在司法中应用创造了良好的外部条件。而随着人工智能与司法相结合程度越来越密切,必然会有更加完善的相关法律法规来管理这个新生事物。

Table 1. The policy evolution and core content of the application of artificial intelligence in criminal judicial adjudication
表 1. 人工智能在刑事司法裁判中应用的政策演进与核心内容

时间	会议/文件	具体内容
2016年12月	《“十三五”国家信息化规划》 ³	推进“智慧法院”建设,打造“智慧检务”
2017年5月	全国法院第四次信息化工作会议	加快“智慧法院”建设,推进审判体系和审判能力现代化
2022年12月	《关于规范和加强人工智能司法应用的意见》	进一步推进人工智能与司法工作的有机结合,大力加强“智慧法院”建设

2.2. 应用场景

2.2.1. 法律信息检索

当前信息环境日趋繁杂,法律领域的条文规范、裁判案例及学术文献呈现爆发式增长,这给行业从业者、科研人员以及普通民众获取精准信息造成了显著阻碍。人工智能技术的迭代升级,正深刻重塑传统法律检索模式,其蕴含的应用价值值得高度关注。传统法律检索依赖人工输入关键词完成匹配,此种模式存在两大突出局限:检索效率低下,面对海量法律数据库,人工逐一审阅既耗费时间,又占用大量精力;匹配方式僵化机械,若用户提问不够精准或句式较为复杂,极易造成关键信息的遗漏。依托自然语言处理技术的人工智能,恰好弥补了这些不足,能够精准捕捉用户提问的核心诉求,即便用户采用日常口语或复杂句式表达,系统也能准确解读其真实意图。

³https://www.gov.cn/zhengce/content/2016-12/27/content_5153411.htm。

人工智能检索最大的优势，远远不是简单地查找文件，而是在大数据中进行深层次的数据分析以及知识组织的能力。同类案件在不同地方法院的不同判决思路，分析地域对于判决的影响；对案件中的合同条款进行法理上的定义及使用范围，并引用专家学者的观点来加深对问题的理解；比较同一类型案件在不同时间点上的判决倾向，可以清楚地看到法律的发展变化。利用各种渠道获取的信息，可以使用户对相关问题有更全面的认识，在案件审判或学术探讨中起到良好作用，有利于法律工作者做出合理正确的决定。

法律体系一直在不断发展变化之中，新的法律法规颁布，新的司法解释发布，大量的新判例出现等都对检索工作提出了新要求，而人工智能由于其高效的数据处理能力可以做到随时连接到权威法律网站上，一旦有新内容产生就可以立即进行抓取并更新。这对律师等法律工作者来说是有利的，它可以让他们更快速地了解最新的法律规定，从而避免因为引用陈旧的规定而导致败诉的情况发生。

2.2.2. 诉讼文件管理

司法人工智能在诉讼材料管理工作中的应用发挥重要作用。对于材料分类、归档工作，利用图像识别以及自然语言处理技术，相关系统可以准确辨别起诉状、证据材料、答辩状、判决书等不同类型的诉讼文书，并根据案件编号或者当事人信息等主要特征将所有相关的文件自动放入相应案件的电子卷宗内，极大提高工作效率及准确性。

信息提取及分析，在此阶段，AI 可以从大量的诉讼文件中准确提取出当事人的名字、身份证号码、联系方式、案件基本情况、诉讼请求等内容并将其转化为结构化的数据以方便案件管理系统进行使用[5]。通过对证据材料进一步的研究，该系统还可以大致判断这些证据对于案件事实或者诉讼请求的相关性大小从而帮助律师等专业人士梳理证据链。

在材料审查及校对过程中，系统可以依照司法文书格式要求检查文档格式是否符合要求，对于不符合要求部分进行标注并提出建议。通过对不同文件中的相同信息进行比较核实，可以很快发现其中可能存在问题以及不合理之处，从而提高司法文书质量把控智能化水平。在涉外案件处理中，该系统具有语言规范化处理以及多种语言互译功能，可以对文书中出现的一些语病、错别字等进行纠正，实现多种语言之间的相互转换，有利于跨境司法合作和沟通交流。

2.2.3. 智能预测与决策支持

数字化、智能化进程不断推进使法律大语言模型在智能预测以及辅助决策方面优势得到充分发挥，正在逐步改变法律行业的作业方式和思维方式。这些模型可以研读大量过去案件并总结经验，从而给现在案件提供一定参考。通过对当前案件与以往案件在案件事实、争议焦点等方面相似程度进行比较，模型可以预测出判决，帮助律师等专业人士。不同于以往依靠经验和直觉进行预测的方法，这种方法是大量的司法判决文书、法律条文以及司法解释为依托，在大量学习之后找出决定案件判决的关键因素及其内在联系。而随着模型结构不断完善以及训练数据不断增加，该模型对于预测准确率和可信度有望进一步提高，从而更好地服务于司法工作。

在智能辅助决策方面，法律大语言模型可以给法官检察官等司法人员提供全面、多元的意见建议。对于疑难复杂的案件，模型可以在短时间内提出多个可能解决方案并在法律合法性、风险、投入产出等方面进行评估比较，形成可供选择的决策报告。同时考虑行业发展趋势以及法律法规变化趋势，模型还可以为企业的中长期发展给出法律上的预见性和指导性意见，帮助企业依法依规经营发展。而在司法机关内部，智能化预测和辅助决策也可以促进资源合理配置。法院可以利用该模型对案件进行初步判断，预判案件难易程度，合理分配审判资源，将有限的人力物力用于解决疑难案件上。对于事实清楚、争议不大的案件，模型可以提供一个标准处理意见，提高审判效率，解决案件积压问题。同时这也有利

于整个司法体系运行效率提升，在一定程度上也做到了司法公平正义和效率二者之间的统一。

3. 人工智能司法应用的现实困境

我国的人工智能裁判系统的相关研究开展的时间相对较短。上世纪八十年代，有人提出将人工智能技术应用于法学的设想，这为我国的法律专家系统的出现奠定了基础；1993年，有研究人员开发出一个基于刑法判决的专家系统，它具有查询法律依据、对犯罪进行定性和量刑建议等作用；进入新世纪以来，我国学者对于人工智能在法律领域的应用的研究逐渐增多，在所发表的文章中约有80%是在2016年之后发表[6]。而在我国司法实践上的人工智能的研究方面，例如类案检索、信息提取、罪名预测等均有所突破。北京法院智能研判系统“睿法官”是利用人工智能和大数据技术识别案件中重要信息，为法官提供案件情况分析、相关法律法规推荐、裁判参考等服务，提升工作效率。2018年1月5日，最高人民法院智慧法院导航系统以及类案智能推送系统开通，在此系统当中设有案情事实、法律适用等内容的模板供使用者查询类似案例。上海高院与科大讯飞合作开发的上海刑事案件智能辅助办案系统以云计算、人工智能为基础，具有证据规则指导、证据审查、智能化开庭等多种功能，近两百个具体项目，大大提高了刑事案件处理速度及质量[7]。

如表2所示，我国许多地方的人民法院已经出台各类具有自身特点的人工智能辅助系统，在类案推送、证据指导、量刑建议等各个方面都实现了多样化布局，是人工智能刑事司法应用取得一定成绩的表现。

Table 2. Artificial intelligence application systems and core functions in the field of criminal justice adjudication

表 2. 刑事司法裁判领域人工智能应用系统及核心功能

法院	系统名称	涉及刑事司法裁判领域的具体功能
北京高院	“睿法官”审判智能辅助系统	相关法律法规查询及案例检索、生成案情“画像”、量刑预测，提取案件重要信息，进行案件情况智能化分析以及法律条文推送和判决参考等。
上海高院	上海刑事案件智能辅助办案系统	证据标准以及证据规则指引、单一证据核验，智能辅助开庭等功能在内的26个功能点，近100个子功能点。
河北高院	“智审1.0”审判辅助系统	刑事类案推送、类案检索、一键生成司法文书
贵州高院	大数据智能分析平台	精确推送典型案例、裁判结果参考
江苏高院	同案不同判预警平台	类案推送、量刑建议与偏离预警
最高人民法院	智慧法院导航系统和类案智能推送系统	查找相似案件

但是，随着人工智能的应用越来越广泛，在各个方面的争议也随之而来，必须引起重视。从技术方面来说，由于算法“黑箱”的存在使得司法裁判过程中使用的算法不可见，从而导致司法裁判过程中的逻辑链条难以被追踪，程序正义受到严重的冲击；从制度方面来看，目前我国的相关法律法规对于人工智能的监管远远跟不上其快速发展的步伐，在算法的进入门槛、责任划分以及个人信息安全等方面都没有具体可行的规定，事实上存在着一种“规制空白”。从价值角度而言，以效率为先的技术思路正在挤压司法公平正义的基本原则，法官的自由裁量权受到由算法推荐的统计数据挤压，在判决时的人文关怀以及道德考量被代码所取代，技术赋权与司法公平正义之间出现的矛盾已经成为一个不得不面对的问题。

3.1. 技术：司法算法化的透明性悖论

随着人工智能技术广泛应用于司法实践中，“黑箱裁判”的问题由于算法缺乏可解释性而成为技术赋权与司法公信力之间的难题。这主要是由机器学习算法本身决定的：神经网络采用多层非线性映射进行特征学习的方式导致结果难以被追查，在大量的参数进行加权运算之后得到的结果是无法用简单的线性关系进行还原的。但是，司法活动天然要求裁判逻辑符合“事实认定-法律适用”的线性论证结构，在案件事实与法律结果之间须有可以验证的推导过程。而从根本上说，这是一场“概率思维”与“规范思维”的对抗。如果法官基于不知其来源而信服的算法提出的建议作出判决，则不但背离了程序正义的要求“理由公开”，而且由于算法隐含的偏见或者数据错误造成实质上的误判。

那么，可解释的人工智能是否可以解决这一问题？部分学者认为，事后的解释性本身并不能满足司法真诚所要求的主观方面的要求——即法官不仅要能够对算法的结果进行说明，而且还要使自己的思维方式与其价值观相一致。因此，算法黑箱的问题并不是简单的技术上的完善就可以克服，它反映了一个更为根本的问题即机器逻辑与人的判断之间存在着不可逾越的认知鸿沟。

技术应用的弊端并不仅限于“黑箱”。更进一步说，在司法领域中数字技术可以运用范围、其自身的可信度及其对于司法人员的影响等问题都是值得思考的问题。

第一，算法有其自身的适用范围。算法的优点是规则明确、操作简便快捷，但是它的运行机制本身是抽象的形式化符号运算，在家事争议、邻里纠纷等涉及人情世故、道德情操的问题上，算法不能分辨也无法解决那些不能量化的方面——感情、目的、人情世故往往是决定判决的重要因素，但是这些却是无法用数据建模来表达的内容。这说明技术手段在法治领域内是有一定局限性的[8]。

第二，算法系统自身具有误差可能。任何一个算法模型都允许一定程度上的误差，在大模型技术飞速发展的今天，算法偏见以及“幻觉”的出现不可避免。而目前在司法智能化过程中，法官阅卷的工作量可下降 80%，案件审理的时间可以节省大约三分之二。但是需要警惕的是，如果整个系统的准确率达到了 99%，那么剩下的那 1% 的错误落到每一个具体的案子上，对于当事人就是 100% 的不公。如果算法的结果关乎一个人的人身自由甚至是财产权益，那么这个技术上的错误就会变成一个真实的法律风险。

第三，技术运用可能导致司法主体出现认知依赖。由于司法大模型的能力不断增强，法官可能会越来越依赖于算法给出的结果。当判决建议是由机器自动生成并且大多数时候都是正确的，则法官进行深思熟虑的主动性就会降低，在具体案件事实认定以及是非曲直判断上也容易被简单的算法结果所掩盖。更重要的是，算法本身就有一定的价值取向及其排序机制，在不知不觉间会对法官的思想观念和审判取向造成影响，从而损害法官的独立性和自主裁量权。

这说明，算法问题不是简单的“看不清”的黑箱，而是“算不了”的能力限制、“算不准”的固有风险，“算了之后”对人的判断能力的削弱等多重问题交织在一起，使技术赋能与司法公正间的关系变得更加微妙。

3.2. 制度：滞后的司法调适困局

人工智能技术的风险具有潜在性、模糊性和难以量化的特征，从而导致传统的法律法规对于智能司法的应用存在着制度上的滞后性问题。目前世界上所有的国家都还没有形成对于人工智能技术有效的法律约束力，在已经推出相关立法措施的欧洲联盟里也是这样，他们在是否要对基础模型进行监管的问题上也各执一词，在“风险等级划分”、“透明度要求”等方面存在着严重的矛盾冲突，《欧盟人工智能法案》能否真正落地存疑[9]。

我国也存在制度供给不足的问题。2022 年最高人民法院出台《关于规范和加强人工智能司法应用的

意见》⁴，更多地从原则层面对人工智能的应用进行引导和规制，“辅助审判”以及“伦理先行”的思想被提出，但是缺乏细化的操作性规定。如，对于算法系统有无资格限制？算法升级换代之后是否需要再次审查？法官背离算法意见时应当满足什么条件等，在现有规定中均未涉及。但是2023年七部门发布的《生成式人工智能服务管理暂行办法》⁵，虽然也提出分类分级监管理念，但是由于其主要针对服务提供者，因此对于司法机关内部使用专业的系统无法起到有效规制作用。⁶

更需警惕的是，目前司法领域的人工智能应用大多由法院和技术公司共同研发而成，在不同的技术提供商所具备的技术能力、数据资源、开发经验等因素都不尽相同的情况下，各个系统在功能设计、运行效果以及适用范围方面也良莠不齐，而且相互之间又没有统一的技术标准和接口规范，因而造成了一家建设多家使用、重复投入、系统互不相通的局面越来越严重。

与此同时，在数字警务、数字监察等领域中，大量地用于内部事务处理、风险管理等方面的应用型算法被大量应用。这些“公共算法”虽然广泛参与到政府权力运作当中，但是长期以来一直处在不受监管的状态下，成为事实上的监管真空地带以及治理孤岛。如何收集数据合法合理？如何设计公平公正的算法？如何保证算法能够被有效回溯等，由于缺乏相应的法律规定以及有效的监控手段而难以解决，存在极大的数据被滥用、算法歧视以及权力失控的风险。

3.3. 价值：算法效率崇拜与司法正义隐没

随着人工智能技术推动司法效率变革，在效率价值和公正价值之间矛盾已经成为当代司法所面临主要道德难题。而这一问题根源在于技术工具有理性和司法价值理性之间断裂：当算法以案件数量、平均审期以及结案率等可以度量标准作为改进方向时，它自身工程技术思路容易把司法工作简单化成一种工业化流水线作业，“看得见的效率”排挤掉“看不见的正义”[10]。

法官在智能系统实时推送类案压力之下，“依法裁判”逐渐变成“依数据裁判”。类案推送机制依据《关于统一法律适用加强类案检索的指导意见》而得以制度化，初衷是实现同案同判，但是当算法根据以往的数据进行总结得出“多数意见”，实际上是以经验概率代替对具体案件的裁量[11]。“自由心证”的范围被算法推荐的概率所挤压，在需要法官仔细权衡“合理怀疑”，有道德难题的“情有可原”，以及体现司法温情的“酌定量刑”等因素都消失在精确计算之中。

更深层次的问题是，司法裁判正在从一种“实现正义的艺术”沦为一种“解纷流水线上的工艺”。技术系统注重形式化的法律因素以及推理过程，而忽视或者不能理解社会伦理情感，在“于欢案”等涉及伦理困境的情况下，算法不能理解“母亲受到羞辱”的情形下的道德常识，只能根据要素匹配给出一种形式化裁量结果⁷。中国的司法一直有伦理法的传统，强调法理情的结合，对于实质正义的要求不可能转化为程序代码。人机协同所面临的根本问题并不是要让机器更加“智能”，而是在技术帮助下人的判断仍然具有独立性——技术是加强人们对法治的认识工具，而不是掩盖司法初衷屏障。

效率崇拜对司法公正的危害，在更深层次上体现为一种权力运作的方式——当数字技术使司法机关拥有从未有过的积极作为的能力，而缺乏对其制约的情况下，这种积极作为就有可能变成司法职能异化以及司法权扩张。

⁴2022年《关于规范和加强人工智能司法应用的意见》第5条规定：坚持对审判工作的辅助性定位和用户自主决策权，无论技术发展何种水平，人工智能都不得代替法官裁判，人工智能辅助结果仅可作为审判工作或审判监督管理的参考，确保司法裁判始终由审判人员作出，裁判职权始终由审判组织行使，司法责任最终由裁判者承担。

⁵https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm。

⁶2023年《生成式人工智能服务管理暂行办法》第2条规定：行业组织、企业、教育和科研机构、公共文化机构、有关专业机构等研发、应用生成式人工智能技术，未向境内公众提供生成式人工智能服务的，不适用本办法的规定。

⁷参见山东省高级人民法院(2017)鲁刑终151号刑事附带民事判决书。

数字司法能动不同于传统的司法能动。它来源于国家层面的数字化发展战略，在平台整合、数据汇聚以及算法作用下，司法机关拥有了前所未有的执行力，使工作方式发生全方位、立体化的数字化改造。但是，技术融入越深入也就意味着权力行使隐秘程度越高。更为重要的是，基于数据共享、数据分析和用户画像，司法机关可以掌握人民群众的行踪轨迹、社会交往甚至情绪状态——这是当下数字检察、数字法院等工作常态。但是，这类新形式的执法行为并没有清晰界限，也缺少统一、权威规定来加以限制。更重要的是，在注重融入社会治理一盘棋的过程中，一些司法机关利用数字化方式过度介入社会治理过程，出现角色前移、功能扩张等问题，甚至有的情况下在同一事件中既充当执法者又担任裁判员，造成“裁判员”与“运动员”一身二任，背离司法中立原则。

归根结底，数字赋能是为了提高司法工作效率，并不是可以成为变通规则、缩短流程、扩大权力以及突破当事人权益的一种工具[12]。如果过度收集、汇总并使用公民身份信息、活动轨迹甚至情绪等，一味进行大数据画像以及算法“穿透”的技术革新，则会对个人信息、数据安全乃至基本人权造成严重危害。技术运用要有限度，权利行使要有限度——这应当是数字司法发展中不能逾越的法律红线。

4. 人工智能司法应用的比较考察

4.1. 域外各国的应用现状

4.1.1. 美国的人工智能司法应用

美国人工智能司法应用呈现多样化尝试以及极大争议局面，在量刑、风险评估等方面都有广泛应用。而在量刑和风险评估方面，最具有代表性莫过于北点公司研发的 COMPAS 系统。它基于年龄、性别、犯罪记录等 137 个指标来预测被告人再犯可能性大小从而辅助法官量刑参考。但 ProPublica 在 2016 年一篇报道中指出，COMPAS 对于非裔美国人错误标签率为 4.1%，而同样比例对于白人仅为 2%。这引发了对算法偏见以及司法公平问题的广泛讨论[13]。然而因为 COMPAS 是一个商业软件，其源代码是保密的，所以被告不知道它是如何工作的，也就不能挑战该评分系统的公平性，这就体现了算法“黑箱”与司法公开之间的矛盾。

在刑事司法技术创新上，芝加哥警局与伊利诺伊理工大学共同研发的犯罪预测模型利用社会学的方法找出有暴力倾向的人群；Cadre 实验室开发的一种枪击声检测软件可以准确区分不同类型的枪支发出的声音；而来自雪城大学的研究人员则开发了一种基于机器学习的 DNA 混合物分析程序，该程序将人的经验和计算机的能力结合起来，极大地提高了对复杂 DNA 样本进行分离的能力。

但是，技术发展的同时也带来深层次伦理困境。算法不公开使得司法判决无法接受社会监督，商业秘密所包裹的算法原理也侵蚀程序公正性。而在算法整个生命期内都存在数据偏差的问题，在以前判例中系统性歧视有可能通过算法被重复并放大力度，这就是所谓的“算法即偏见”。另外，大规模使用生物特征识别技术也会给个人造成隐私权侵害的可能性[14]。如何找到技术效率和权利保护之间的平衡点，是美国司法 AI 运用需要面对的一个问题。

4.1.2. 欧盟的人工智能司法应用

在欧盟司法领域，人工智能的应用主要集中于法律检索、案件预测分析、风险评估以及文件审核等方面，比如 ROSS 采用自然语言处理提高法律搜索速度，Lex Machina 基于大量数据对专利诉讼结果进行预测，但是这都是一些问题。由于人工智能作出决定具有“黑箱”性质，造成裁判不公开，从而损害当事人对于司法公平的信心，同时也带来责任问题——如果人工智能出现错误，则开发者、使用者(法官或者律师)以及机器之间的责任分配尚无法律规定。

为了克服这些问题，欧盟试图用一种规范性的方式去影响技术的发展方向，但是这并不够，它缺乏

强制性和完整性。比如,《欧洲司法伦理宪章》虽然提出了“透明度”、“非歧视”,但是对于如何监督这些原则的实施并没有相关规定;而跨境司法项目 e-CODEX 旨在通过标准化的技术规定使各成员国的司法信息系统互相连接起来,但是在实践中由于对于某些概念的理解不同(例如,“送达程序”的定义在不同的国家有所不同)使得意大利米兰法院的案件数量仅相当于德国的一六十分之一。此外,一些国家已经开始采取行动来解决这个问题,比如,在 2018 年,意大利议会通过了一项法案禁止使用全自动化的裁判系统并且规定必须有人类进行审查。法国《数字共和国法》还规定,在公布司法数据之前必须对其是否经过去标识化进行风险评估。有学者提出,未来需要制定更加严格的跨境规则:一是要求算法提供“可解释性”的报告提高透明度;二是界定开发者的责任以及使用者的责任;三是对法官等从业者开展有关人工智能的风险评估的技术伦理培训[15]。

同时,也有具体案例表明技术的应用并非一帆风顺。虽然 e-CODEX 项目使欧盟内的支付令(EOP)可以电子化传递,但由于各成员国司法体制不同带来的障碍使得该系统的使用率不尽如人意,例如在德国由于集中管辖方式更容易普及,而在意大利因为法院之间管辖权划分造成该系统的利用率较低。

4.1.3. 韩国的人工智能司法应用

韩国司法领域的人工智能应用具有明显的“政策导向型”,其发展很大程度上依赖于国家数字化发展战略的整体规划。2021 年,韩国政府出台《AI 国家战略 2.0》提出打造“智能司法系统”,由法院行政处负责研发法律文书自动化工具以及案件预测模型等,目的是利用技术提高司法效率和可获得性。该战略基于韩国的信息通信技术基础,在全国范围内推广司法智能化,自上而下地进行。

从技术角度来说,韩国已实现多层次、全方位司法人工智能应用格局。法务部开发了 LONBOT 法律咨询机器人运用自然语言处理技术,向民众提供免费的基础法律咨询服,降低民众获取法律信息成本。另外,由首尔大学与法院联合研制 AI 量刑辅助工具,在大量收集 2010 至 2020 年所有刑事案件判决结果后进行训练,使用 LSTM 模型预测刑期范围,预测正确率为 78%,用于帮助法官量刑时做到同案同判。另外,韩国电子诉讼平台还具有 AI 文书生成功能,在诉讼材料提交、文书起草方面给予辅助,促进诉讼活动信息化、规范化。

但是,技术发展的同时也带来一系列伦理问题以及制度反思。而学者们最关心的是缺少算法透明度所造成的隐患,在判刑预测系统中,过去的判例中就存在着性别或者地区等方面的歧视,如果不加以甄别并且纠正,那么算法就会把这些歧视性的东西吸收并且扩大。另外就是对于法官自由裁量权的保护问题也是人们关注的重点问题。韩国司法政策研究院 2023 年的一项实证研究显示有 76%的法官反对 AI 参与到最重要的决定当中来,他们认为人工智能只能负责收集证据、撰写文书等工作而不是裁量,而且这个界限必须要明确[16]。为解决此问题,韩国在法律方面也有所行动,《AI 算法透明性法案》的出台即是对算法作出决定的过程予以法律上的规范。

目前韩国相关研究集中在技术应用优化、数据治理以及人机协同等方面。高丽大学的研究者认为,判决文书匿名化及算法训练之间存在根本矛盾,在保证当事人隐私及维护数据质量方面如何找到一个折衷方案是数据治理面临的问题所在。而在比较法视野下看,相比于中国,韩国对于技术落地速率和技术伦理的要求有更高的妥协程度,在大力推广司法人工智能的同时也较早地开始进行伦理审查以及立法工作,体现出一种未雨绸缪的态度。韩国的经验也说明了司法 AI 的发展离不开技术进步的支持,但是也需要有相应的制度伦理相配套,在效率与公平、赋权与制约之间取得恰当的比例才是技术与法律的良好结合。

4.2. 域外现状对我国的启示借鉴

一是强化算法可解释性与数据治理。目前司法算法主要是基于深度学习模型,“黑箱”性质与司法

论证的公开性相悖。解决这个问题要在技术上寻求突破，在技术层面研发可解释的人工智能，设计出可以被理解的算法模型，让其作出的决定可以表现为一种“事实认定-法律适用”的链条式的论证过程。此外还要有一套涵盖数据收集、标签化、训练等各个环节的管理制度，对于训练的数据进行偏误识别以及净化工作，避免历史上存在的潜在不公正被算法反复强化。

二是加快制定专门性立法，明确算法责任归属。当前相关规定大多仍处于原则性规定上，无法解决因算法进入司法而导致的责任空白问题——如果裁判错误是基于算法提议作出，则应向谁追责：开发者、使用者抑或是法院？这就需要制定专门法律法规来确定算法使用门槛、升级评价机制以及法官偏离算法建议所必须遵守程序等。在此之上建立跨学科伦理审查机构，在涉及个人自由或者重大财产利益案件中对于算法应用进行事前把关，以避免出现大规模歧视以及“算法黑箱”现象。

三是政策推进需平衡效率与公正。智能司法的目标不能简单理解为“案件吞吐量”，而要在宏观上限制公平的价值。即设立“辅助而不是取代”的防火墙，在制度上保障法官对于算法结果有最后决定权以及偏离权，保证自由心证不受概率统计干扰。还要进行面向法官、律师乃至普通民众的数字技能培训，消除由于技术使用而产生的数字鸿沟，避免因知识水平不同造成新的程序不公平。提高效率是为了让正义尽早实现而不是为了牺牲个案的公正性而提高效率。

四是构建法律、技术与伦理动态互嵌的监管体系。中国的司法既有“统一性”，又有“地方性”，因此治理之道在于找到一个合理的中央层面的一体化规则与地方自主探索之间的关系。监管应当具有灵活性，能够对算法进行适时的评价，并及时给出反馈意见，让技术的进步始终处于法律以及道德的评判之中，从而构建出具有中国特色的司法人工智能治理模式[17]。

5. 人工智能司法应用的治理路径

5.1. 技术：从“算法黑箱”到“数据透明”——人机协同的源头治理

司法是书面上法律条文运用的过程，有良好法律但无法执行也会影响法律本身威信及法律效力，相反如果民众相信司法公平公正，那么就on容易选择通过法律途径解决问题，有利于社会稳定和谐，但是使用人工智能进行司法活动，由于司法决策模型建立过程中存在“算法黑箱”，使司法决策系统裁判依据不明晰，进而对司法决策系统公正性、有效性提出质疑。除去由于商业秘密或者国家秘密造成不透明以及由于技术文盲造成的不透明以外，人工智能中算法不透明主要是因为机器学习算法特点所致。

数据是知识形成和文明进步的重要标志，也是人类对客观世界认知的成果。数据既涉及国家和社会公众的数据权利问题，也关乎算法输出结果是否公平公正的问题[18]。机器学习算法由于数据质量的不同会产生不可预测的结果。应用于司法领域的决策系统需要以优质的案例为基础并且代码具有可追溯性。系统的设计师可不断地改进和完善算法及模型，也可以尝试开发新的模型等。并且要有一套完善的技术准则来保证 AI 应用的合法合规性、公开透明性和安全性。在模型训练阶段，应让法官和技术人员一起训练模型，保证数据来源及质量可信，比如就以往案件进行“贴标签”，从中选取合适有效的案件素材等；同时需不断补充司法案例库，紧跟最新法律法规以及判例发展；对于由人工智能参与实质性工作的审判程序，则可设立定期考核制度，在此基础上进行反思并优化相关的人工智能系统及其算法模型。

人工智能的神经网络算法不同于符号算法，它是一种非线性的思维方式，不能通过线性思维对其进行逆向推导。因此，事后再去审查机器是怎么样得出一个司法裁判结果是很难做到准确有效。但是反过来人自身也是这样，判决书也不能把法官作出裁判时所考虑的所有因素都写明。人的审判思路也不能完全记录以便以后检查，只能通过对行为和结果进行倒查并收集相关证据来判断。所以，在构建人工智能司法决策模型时，对于算法黑箱的监管不需要像未来主义文学中那样对未来的人工智能充满焦虑情绪，

而只需要最少限度的技术手段[19]。

解决算法“黑箱”问题不能仅仅停留在对于技术特点的认识及被动接受上，而应该从人机协同的角度出发建立一种全方位治理模式。技术上的治理方式要正视上述提到的三个难题即算法适用范围有限性、系统偏差不可避免性和法官对算法信任所带来的问题，在此基础上寻求可行的技术方案。

一是建立算法适用边界的可识别机制。针对算法在情感纠葛、伦理冲突等情形下存在不足，在系统开发时应加入“场景敏感型”识别环节。利用自然语言处理技术对案件信息进行初步筛选，识别出有可能是带有情感色彩、伦理问题或者是人之常情的案件并将其标注出来，然后由人工进行进一步审理。比如，在家事纠纷、邻里纠纷这类案件中，在推送相关案例以及量刑建议的基础上，要以明显的方式提醒法官注意其中是否存在非量化的成分，防止算法忽视情理的因素。

二是构建误差容错与风险隔离的双重防线。对于算法本身所固有误差问题，特别是其中那“1%”落到具体案件上可能导致的重大不公平，需要有多层次防范措施。一方面，在系统输出环节采取“置信度阈值”方式——如果算法给出建议置信度小于一定值，则不予输出建议而由人进行处理；另一方面，在整个司法过程中设置一个“算法建议复核清单”，让法官在使用较高风险领域(例如量刑建议、羁押必要性审查等)中的算法结果之前逐一检查系统未考虑到的重要因素，做到人机结合共同作出决定。

三是构建算法运行全流程的透明化记录与追溯机制。根据“过程哲学”的理念，把算法参与司法裁判各个环节都纳入可以被记录、可回溯的技术路径中，在数据收集、特征选择、建模计算到最后输出结论的过程中，都有详细的操作日志，使得每个算法提出的意见都可以追根溯源。而这种技术留痕的作用不仅仅在于事后的监管以及责任的认定，更重要的是，在其运行之时就对算法起到一种限制作用——当人们知道自己的行为会被永久保存的时候，往往就会更加注意自己所做的决定是否合理。

5.2. 制度：“规范滞后”到“弹性规制”——分层次适应性法律反映方式

国家的法治建设起着至关重要的作用。而伴随着人工智能时代的到来，人工智能辅助司法裁判的规范化建设也面临着制度滞后的问题。其进入司法领域所带来的一系列问题显然不能仅靠“引导、监督”来解决，我国亟需一部专门针对司法实践中使用人工智能的技术进行规制的法律法规[20]。技术本身所具有的不确定性、模糊性特点以及司法活动中对于确定性和价值判断的需求之间存在着一种冲突，在此背景下需要制定专门法律法规来弥补这种缺陷，以避免因技术不当使用造成对权利侵害、裁判机械化、司法自主性的丧失等问题。

建构法律规制应以确定人工智能的辅助角色为基点。法律应规定，在所有情况下，人工智能都不能取代法官独自作出判决，它所提出的意见仅具有参考意义。还应设立一些针对特定领域的限制性规定以及绝对禁区，在立法过程中要突出司法独立与公正的要求，让法官保持独立、公正的态度，不受人工智能的影响，同时也要保证人工智能的应用公开透明。此外也不能忽视的是，作为公共权力的国家机关可能会通过利用人工智能的技术而扩大自身权限甚至滥用职权，人工智能有被权力化的风险。因此，相应行使监督权的国家机关还需要设置独立的审查机制以及反馈纠正机制对法官运用人工智能情况进行监督与制约以保障司法公平正义。

在人工智能辅助司法过程中，要充分保障人民群众合法权益及隐私[21]。我国也可以参考一般数据保护条例或者隐私权法等相关公民数据立法的经验，出台一部“个人信息保护法”，规定数据权利内容，限定个人信息使用范围，防止个人信息被非法获取与利用。解决制度上存在的规范滞后，需要有层次性、差异化的法律回应方式，在如下几个方面进行制度革新：

一是基于风险评估矩阵的差异化规制。为改变当前规制方式笼统、操作性不强的局面，应构建一个以“权利影响程度”与“决策自动化程度”为维度的司法人工智能风险评估矩阵，将应用场景划分为高、

中、低三个风险层级，并实施差异化的全生命周期监管。具体而言，高风险应用指那些深度参与司法核心判断、可能实质性替代法官裁量权或对公民生命、自由等基本权利产生重大影响场景(如量刑建议、羁押必要性审查、证据链分析等)。此类应用必须实施最严格的监管：在准入阶段，强制要求通过第三方技术审计与独立的伦理审查，并向最高人民法院指定的专门机构进行算法备案，提交包括模型架构、训练数据来源、预期误差率及纠偏机制在内的完整技术文档；在运行阶段，须建立实时监控与定期重评估机制，并确保算法的完全可追溯与可审计。中风险应用指辅助司法事务处理、对当事人权益有一定影响但非决定性的场景(如类案推送、法律文书初稿生成、诉讼材料分类等)。对此类应用，监管要求适度放宽：无需进行事前强制性伦理审计，但须进行算法功能与运行范围备案，并要求系统在输出结果时进行明确的风险提示(如告知法官“该推送结果仅供参考，可能存在选择偏差”)，同时建立定期的事后抽检制度。低风险应用指主要用于内部管理、司法公开或便民服务，基本不涉及裁判权行使的场景。对此类应用，主要采取企业/研发机构自我声明加事后抽查的规制方式，在保障创新活力的同时，通过用户反馈渠道和年度合规报告进行兜底监管。通过构建这一清晰、可操作的风险评估矩阵与差异化监管体系，能够有效平衡技术创新与司法公正之间的张力。

二是建立司法人工智能系统的标准化治理框架。为解决各地法院各自为战、系统之间存在诸多障碍问题，在最高人民法院指导下出台一套全国通用的司法人工智能标准规范，包括数据格式、接口标准、评估方法、安全保障等方面内容。标准化并不是反对地方创新，而是在一个共同基础上允许不同地方进行个性化尝试——各地可以在遵守基本标准基础上根据当地实际情况设计相应功能模块。此外还要有跨系统间互连互通，消除信息孤岛现象，让好的算法能够在更大范围内被利用。而标准化更重要一点是它为算法的可解释性、可追溯性、可审计性奠定了基础，“算法透明”就可以从一个理念变成一种可以实现的要求[22]。

三是将“公共算法”纳入公权力监督的制度框架。对于大量已广泛应用于数字警务、数字监察等领域的算法系统，但由于其处于监管之外而无法对其进行有效约束，有必要将其视为一种“公权力运行载体”，并对其施以比司法人工智能更高或者至少相同程度的管控措施。即要实行公共算法清单制度，在行政机关利用算法进行判断的过程中，必须向公众披露算法所具有的功能、使用的数据以及作出决定的方式及救济途径等内容；尤其是当该决定对公民的权利义务产生影响时，则还应当保障当事人的知晓权以及提出不同意见的权利等。这一制度的理论依据是：当算法参与公权力行使之时，它就是一种公权力，应当受到法律规则以及正当程序的要求，而不能因为商业秘密或者技术复杂等原因而逃避监管。

5.3. 价值：“效率优先”到“公正锚定”的转变——技术理性与司法伦理之间张力解决模式

公正、高效、权威是中国司法现代化目标和特征，在积极促进中国司法现代化过程中、提高中国司法现代化水平的过程中，人工智能应用是一个无法回避的问题。数字技术目的是尽可能快地办理案件，但是对效率过分强调会忽视数字正义，降低司法公信力[23]。为了防止人工智能进入司法后出现效率价值和正义价值背离情况发生，在人工智能技术运用到司法过程中必须坚持司法公平原则进行效率优化，从整体上制定出一套能够协调技术工具理性与司法内在价值之间关系制度规则。

虽然人工智能司法是以算法司法的形式出现，表现为法官后撤的情形，但是后撤并非意味着离场，而是从“幕前主导”变为“幕后监管”。在算法司法环境中，司法人员仍需对算法所给出的答案进行校正。可以成立司法人工智能伦理委员会，在委员会指导下审查、反馈算法结果，相关技术部门根据委员会意见对算法进行纠偏。面对人工智能司法大势所趋的趋势，作为司法主体的广大司法人员不可能被彻底排斥在外。在坚持法官独立裁判这一基本原则的前提下，完善法官的职业能力建设，进行模拟法庭对抗、裁判思维诊断报告等形式的训练，使法官具备对于智能建议与自身判断之间关系的认识，在此基础

上，能否辨别、检验算法给出的结果也是法官技术能力的一个体现。

解决价值层面“效率至上”的困境，需要在技术和司法伦理之间找到一个折衷点，在此基础上从三个方面进行改革：

一是确立法官认知主权与审慎判断的回归机制。对于算法推荐对法官自由裁量权产生的压制作用，应当从制度上规定必须让法官的审慎判断进入决定过程中，在此基础之上可以设置“算法建议偏离备案制度”，即法官采用量刑建议或者类推案件之前，对于有争议之处需要自行进行思考并且写明理由；如果不同意算法建议，则应当在判决书中予以阐述原因，比如个案中存在特殊情况、情理因素或者是道德难题等。这就可以使法官由一个消极的“算法接受者”转变为积极的“算法批判者”，使得法官在面对程式化结果时仍然能够注意到每一个个案的独特性并且进行反思。

二是构建情理要素识别与价值预留机制。由于算法无法感知到人类所经历的情感纠葛、伦理冲突等非量化的方面，在智能化司法裁判辅助系统中应当设置一个“情理预警模块”。利用自然语言处理技术对案件事实先行判断，识别出其中可能存在的人伦常理、道德困境的因素——例如家庭关系中的亲情纽带、邻里矛盾中的乡土人情、紧急情况下的伦理选择等，用明显的方式提醒法官注意。而且在类似案件推荐算法中留有余地供法官作出价值判断的空间，而不是让系统仅仅根据形式上的要素相同就向法官推荐相似案例，忽略了个案中的特殊情况及情理。这种机制技术上是让算法认识算法本身的界限，在程序上确定的是：有一些东西不能被编码为算法，而这些则需要人来进行评判。

三是划定技术赋能与权力制约的法治边界。对于数字司法能动带来权力扩张问题要坚持“技术赋能不过界”的基本原则，在此基础上要实行司法数据收集及使用上的“必要且最小”，不得以智能化为由大规模收集当事人个人信息；对数据画像及算法穿透使用范围予以限制，不得在缺乏相应法律规定情形下对个人进行风险评估或者行为控制。更重要的是，必须由法律确定司法机关在数字化管理中的角色，避免因技术介入造成“裁判员”与“运动员”合一。数字赋能应是为了提高司法履职效率而非成为变通规则、简化程序、扩大权力或突破当事人权利边界捷径——这是一条数字司法发展中必须遵循法理原则也是保证技术理性与司法伦理之间协调基础。

6. 结语

人工智能技术的应用对于司法而言既是促进司法公平正义的重要动力源泉也是对司法公正提出新的挑战。从技术赋能到制度调适从效率追求到价值平衡人工智能并不是一个纯粹的技术手段而是在改变着司法逻辑的一个重要方面。本文基于“技术-制度-价值”的三维视角分析了我国人工智能在司法实践中运用情况及存在的问题以及国外相关经验并据此提出适合我国国情的人工智能司法治理方案。

研究显示，人工智能应用于司法并不只是技术问题，而是一种包含算法公开性、法律规制完备性和司法价值导向性在内的综合性系统工程。技术上的“黑箱”与司法上的“公开”，两者之间本身就存在着天然的冲突；而制度上的滞后又与技术上的快速发展相悖，效率至上的技术思路又与以公平为导向的司法宗旨相左。解决这些问题，就需要在技术上促进可解释人工智能发展的同时，在立法上制定符合司法需要的相关规定，在价值观上坚守司法公正是首要原则。

未来，人工智能与司法融合是大势所趋，但是技术的应用不应以损害司法独立性、公正性和人文温度为前提条件^[24]。只有让算法逻辑与裁判理性进行良好沟通，在提高效率的同时保护当事人合法权益，在促进科技创新的同时遵守相关法律规定，才能更好做到“让人民群众在每一个司法案件中感受到公平正义”^[25]。人工智能助力司法，最终应当是推动法治发展、维护司法本色的力量而不是掩盖人的意志、代替人的责任的“黑箱裁判”。

基金项目

浙江省大学生创新训练项目智慧司法中 AI 技术应用的现实困境与治理路径研究 S202510341150 (浙江农林大学资助项目)。

参考文献

- [1] 高学强. 人工智能时代的中国司法[J]. 浙江大学学报(人文社会科学版), 2019(4): 235.
- [2] 张海波. 推进高质量司法的挑战与应对[J]. 中国应用法学, 2025(1): 8.
- [3] 李声高. 技术程序正当论下的刑事在线诉讼规则研究[J]. 社会科学战线, 2023(7): 218.
- [4] 秦策. 操作性智能标准下的 GPT 参与司法[J]. 中国应用法学, 2023(7): 105.
- [5] 胡向阳, 张巍. 基于大数据的侦查讯问文本数据挖掘与分析[J]. 中国人民公安大学学报(社会科学版), 2019(6): 41.
- [6] 王率先. 论人工智能在法院裁判中的定位——以美国 COMPAS 量刑辅助系统的应用历史为视角[D]: [硕士学位论文]. 上海: 华东政法大学, 2019.
- [7] 宋万年. 基于要素抽取的司法文书智能生成方法研究与系统设计[D]: [硕士学位论文]. 重庆: 重庆大学, 2023.
- [8] 蔡立东, 郝乐. 司法大数据辅助审判应用限度研究[J]. 浙江社会科学, 2022(6): 59.
- [9] 穆捷. 欧盟数字主权理念与实践研究[D]: [博士学位论文]. 北京: 北京外国语大学, 2023.
- [10] 李景豹. 智能革命与法理革新: 大模型技术对基础法理的挑战及其制度因应[J]. 理论导刊, 2025(6): 77.
- [11] 许文妍. 研究如何解决刑事领域的“类案不同判”冲突[D]: [硕士学位论文]. 郑州: 河南财经政法大学, 2023.
- [12] 马长山. 数字司法的法治边界[J]. 东方法学, 2024(4): 132.
- [13] Rigano, C. (2019) Using Artificial Intelligence to Address Criminal Justice Needs. *NIJ Journal*, No. 280, 36-47.
- [14] Ulenaers, J. (2020) The Impact of Artificial Intelligence on the Right to a Fair Trial: Towards a Robot Judge? *Asian Journal of Law and Economics*, **11**, Article ID: 20200008. <https://doi.org/10.1515/ajle-2020-0008>
- [15] Lupo, G. (2019) Regulating (Artificial) Intelligence in Justice: How Normative Frameworks Protect Citizens from the Risks Related to AI Use in the Judiciary. *European Quarterly of Political Attitudes and Mentalities*, **8**, 75-96.
- [16] Choi, J., Lee, S. and Jamal, T. (2021) Smart Korea: Governance for Smart Justice during a Global Pandemic. *Journal of Sustainable Tourism*, **29**, 541-550.
- [17] 晓克. 数字时代的智慧审判研究[D]: [博士学位论文]. 上海: 华东政法大学, 2022.
- [18] 刘艳红. 人工智能司法安全风险的算法中心治理模式[J]. 东方法学, 2024(7): 89.
- [19] 周尚君, 伍茜. 人工智能司法决策的可能与限度[J]. 华东政法大学学报, 2019, 22(1): 63.
- [20] 张玉洁. 人工智能辅助量刑的法律风险与制度建构[J]. 学术交流, 2023(3): 64.
- [21] 张越然. 人工智能介入司法领域的困境及其应对[J]. 边缘法学论坛, 2019(2): 15-18.
- [22] 翟岩, 李小波. 可知的不可知: 算法透明的内涵澄清与“模块化”制度体系构建[J]. 河南社会科学, 2025, 33(6): 91.
- [23] 张凌寒. 数字正义的时代挑战与司法保障[J]. 湖北大学学报(哲学社会科学版), 2023, 50(3): 134.
- [24] 吕士哲. 人工智能司法应用的基本共识、关键分歧与发展趋势[J]. 湖北经济学院学报(人文社会科学版), 2024, 21(9): 102.
- [25] 王新清, 高林. 新时代刑事司法公正观论略[J]. 河南财经政法大学学报, 2024, 39(2): 9.