

人机协同审判模式中司法 责任的智能化因应

陈子君*

摘要 在司法审判中,人工智能凭借其高效便捷的优势,广泛适用于证据认定、法律推送、裁判预测等审判活动及审判监管中。但这也使司法责任面临主观责任认定模糊、责任主体界分不明和责任推诿风险扩大的难题。究其原因,技术局限使审判过程存在因果判断与相关性判断的逻辑抵牾,技术人员和监管者对裁判的隐性嵌入消解了法官的主体地位,并且智能审判也面临效率与公正的价值张力。为此,司法责任应在法官、技术人员与监管者之间,设计、应用与运维之间进行均衡配置。应以技术性正当程序构建沟通协商空间,提升追责过程的清晰度;以伦理责任建立司法可信生态,提高智慧司法的公信和权威。同时根据不同场景,对技术影响审判公正的程度进行风险分级,进而基于比例原则明确各主体的技术和伦理责任,促进风险和责任的功能适配。

关键词 司法责任 智能审判 技术性正当程序 伦理责任

一、问题的提出

司法责任制是指行使司法权的主体根据法律和政策规定,对司法行为及其产生的结果是否承担责任以及承担何种责任的制度。^{〔1〕}这一制度改革是权力与责任相统一原则下的重要制度设计,在保障法官主体地位的同时,通过责任监督促进依法裁判。司法改革和信息化是推动人民法院工作发展的“车之两轮、鸟之双翼”。在智慧法院建设中,制度既有的结构对技术运用效能的直接影响也要求制度随着技术的运用而不断完善调适。^{〔2〕}类案智推等智能应用使案件裁判过程呈现出人机结合的特点。在事实认定和法律适用阶段的实质性嵌入使智能技术一定程度上分享了法官的审判权,影响了法

*北京交通大学法学院讲师、法学博士。本文系国家社会科学基金青年项目“企业合规行政监管的法治化研究”(项目编号:23CFX007)的阶段性成果及上海交通大学首届“面向青年”法学论坛青年教师组优秀奖获奖作品。

〔1〕 参见胡仕浩:《关于全面落实司法责任制综合配套改革的若干思考》,载《中国应用法学》2019年第4期,第20页。

〔2〕 参见邱泽奇:《技术与组织的互构——以信息技术在制造企业的应用为例》,载《社会学研究》2005年第2期,第33页。

官的主体地位。司法责任制应结合人机协同智能审判的特点予以制度因应,以实现权力与责任的平衡。

目前,学界围绕智能审判中司法责任制的挑战进行了系统性反思。^{〔3〕} 如有观点主张,司法智能化应用的过程性留痕和节点监督将结果意义上的司法责任转换为程序意义上的司法责任,可解决司法责任制改革放权之后的监督问题。^{〔4〕} 也有学者对智能审判中的人机关系进行反思,如类案智推环节,若法官因遗漏类案而被强制问责,则可能因这一“工具化对象”而陷入主体性焦虑中。^{〔5〕} 对此,有学者主张引入算法责任,并类推适用侵权法上的产品责任。同时由于刑事司法的算法产品涉及重要的人权问题,算法责任机制应与准入制度挂钩,根据算法表现进行动态评级,调整法院的公共服务购买清单。^{〔6〕} 当算法偏见造成负面后果时,通过立法和公开政策制定来启动对算法设计者和侵权者的追责。^{〔7〕}

现有研究分析论证了人工智能影响审判过程和裁判结果的法理路径。这为人机协同智能审判中司法责任的系统性、针对性研究奠定了基础。人工智能时代的责任机制是复合型问责式,责任实现机制拆分为解释说明与后果承担两个交互作用的维度,责任类型也包括法律责任和伦理责任。^{〔8〕} 智慧法院场景中,智能技术对审判模式产生全方位影响,尤其是已经具备类人化意识与行为能力的生成式人工智能、代码、算法和数据已经被用于预测司法判决的结果,算法逻辑已经可以渗透进入判决的体系架构之中。^{〔9〕} 由此,智慧法院中的司法责任不同于一般场景下的责任追究,需保障权责一致与法官审判自主权的价值均衡。^{〔10〕} 本文将对人机协同智能审判中司法责任的风险予以反思,在此基础上实现司法责任的制度因应。

二、智能审判中司法责任认定的风险挑战

(一) 主观责任的认定模糊

归责原则是确定责任承担的标准和根据。现行规定主要采用主观过错的归责原则。2015年《最高人民法院关于完善人民法院司法责任制的若干意见》第2条第6款规定了“主观过错与客观行为相结合”的原则,^{〔11〕} 关注是否存在“故意违背法定程序、证据规则和法律明确规定违法审判

〔3〕 参见季卫东:《人工智能时代的司法权之变》,载《东方法学》2018年第1期,第132页;张凌寒:《智慧司法中技术依赖的隐忧及应对》,载《法制与社会发展》2022年第4期;卞建林:《人工智能审判的责任解构与制度应对》,载《法治社会》2023年第5期;张琳琳:《人工智能司法应用的责任归结困境与解决路径》,载《法商研究》2023年第5期。

〔4〕 参见王亚新:《信息化浪潮中的司法改革——机遇与挑战》,载《法治现代化研究》2018年第2期,第17页。

〔5〕 参见陈洪杰:《从技术智慧到交往理性:“智慧法院”的主体哲学反思》,载《上海师范大学学报(哲学社会科学版)》2020年第6期,第88页。

〔6〕 参见江溯:《自动化决策、刑事司法与算法规制——由卢米斯案引发的思考》,载《东方法学》2020年第3期,第86页。

〔7〕 参见魏斌:《智慧司法的法理反思与应对》,载《政治与法律》2021年第8期,第124页。

〔8〕 参见赵景琛、张凯:《人工智能问责制的法理界定与优化路径》,载《北京航空航天大学学报(社会科学版)》2025年第4期,第44—46页。

〔9〕 参见袁曾:《生成式人工智能的责任能力研究》,载《东方法学》2023年第3期,第26页。

〔10〕 参见宋远升:《司法责任制的三重逻辑与核心建构要素》,载《环球法律评论》2017年第5期,第77页。

〔11〕 相关文件,参见最高人民法院关于印发《法官惩戒工作程序规定(试行)》的通知(法〔2021〕319号);中共中央办公厅、国务院办公厅关于印发《保护司法人员依法履行法定职责规定》的通知(2016年7月21日)。第19条规定:“司法人员的个人信息受法律保护。侵犯司法人员人格尊严,泄露依法不应公开的司法人员及其近亲属信息的,依照法律和相关规定追究有关人员责任。”

的,或者因重大过失导致裁判结果错误并造成严重后果的”。《法官法》第46条明确规定了九项违法审判行为。其中第5项明确了“重大过失+裁判结果错误+严重后果”的认定模式。其他八项对违法行为进行具体列举。这反映出制度设计遵循的是“故意+行为”与“重大过失行为+严重后果”两组相互嵌套的责任认定标准。^[12]同时,《人民法院审判人员违法审判责任追究办法(试行)》第2条也规定:“人民法院审判人员在审判、执行工作中,故意违反与审判工作有关的法律、法规,或者因过失违反与审判工作有关的法律、法规造成严重后果的,应当承担违法审判责任。”^[13]可见,虽然过错程度不一,但现有法律规定均将主观过错作为责任认定的主观构成要件之一。

智能审判为主观过错提供了技术性规避空间,认定过程更为模糊。主观过错认定模糊来源于算法黑箱、算法歧视等技术局限与法律逻辑的内在冲突。由于算法黑箱和不可解释性等技术局限,法官借助智能应用辅助进行法律适用的过程无法被解释,自然也无法对其裁判过程作出完全的说理论证。由此增加了法官行为与案件质量之间因果关系的判断难度。正如在自动驾驶场景中,算法黑箱导致了主观过错认定困境和责任真空。2019年的特斯拉车祸表面看是自动驾驶技术的失误,实则揭示了人工智能与人类责任边界的模糊,即究竟是工程师未预置极端场景训练的过失,还是因传感器硬件缺陷导致的产品责任?此时技术缺陷下的产品责任可能会弱化主观过错的存在。而在智慧法院场景中,不同于类案检索,类案推送的智能属性是在自动化的基础上利用大数据的分析挖掘,实现了个性化定制,由“照我说的做”到“自己看着办”的转变。类案推送的运行路径是在导入案件卷宗后,系统通过案情匹配,自动筛选相似案例,根据相似度、同类判决两个因素进行排序,结果反馈为案情相似度降序和同类判决集合,进而通过定量分析,了解裁判趋势。^[14]在这一过程中,智能司法系统仅单纯以分词的词频作为概率模型的产生基础,并不区分其对应的法律效果,这增加了司法数据的模糊性与不确定性。^[15]因此,类案推送的语义理解机理暗含的是交互技术、庞大数据和机器训练。这一过程的不可解释性使主观责任的认定过程难以追溯。法官可能以智能化应用的推送结果作为自身不具有主观过错的事由,导致难以将其认定为重大过失,或者因类比适用符合形式正义的要求而免于责任承担。因此,智能应用中的算法黑箱、算法不可解释等技术特征和算法运行方式弱化了案件裁判过程的透明度,导致主观过错认定模糊。

(二) 责任主体的界分难题

智能辅助办案、全程监督预警等多场景应用使智能审判呈现出“智能辅助+人工审查+审判管理”的决策特征。在技术人员和管理者可能直接或间接地参与审判过程和影响裁判结果的情形下,如何界定责任主体及其责任范围是另一难题。

首先,技术人员的角色由工具提供者到规则塑造者拓展。“代码即法律”的理念是通过技术架构与法律规则的深度融合,将法律规则转化为可执行的算法代码,实现法律执行的自动化、标准化与不可篡改性。智能应用的节点固化、自动生成等特点使审判程序在一定程度上脱离法官的主导。实践中,法官审判行为被分解为立案、庭审、文书生成、归档等具体环节,并全部内嵌于数字化平台中,诉讼材料、证据、笔录等原生数据通过电子卷宗随案同步生成技术被实时采集。例如,杭州互联网法院上线区块链智能合约司法应用,通过打造网络行为“自愿签约、自动履行、履行不能

[12] 参见方乐:《法官责任制度的司法化改造》,载《法学》2019年第2期,第157页。

[13] 参见《人民法院审判人员违法审判责任追究办法(试行)》。

[14] 参见陈琨:《类案推送嵌入“智慧法院”办案场景的原理和路径》,载《中国应用法学》2018年第4期,第92—93页。

[15] 参见陈亮、徐明:《从数据到知识:智能司法基础设施的困境反思》,载《交大法学》2022年第3期,第147—148页。

智能立案、智能审判、智能执行”的全流程闭环,提升智能合约执行效率,实现网络数据和网络行为的全流程记录、全链路可信、全节点见证、全方位协作。^[16] 动态知识图谱与场景化的法律大模型使得法律规则向算法逻辑延伸。一切自动化算法决策,看似屏蔽了“主观人为”,但其形成决策的知识、逻辑、边界和价值基准是被预置的,^[17]即在技术人员主导下通过智能技术对人类认知进行化约。因此,技术人员可以通过算法设计和数据驱动的动态调整而改变模型运行方式,借助“代码”实现“不在场”的“在场”。^[18]

其次,智能技术成为审判监管者对审判内容和过程进行实质性监管的重要技术工具。审判管理智能化的实现方式之一即以案件审理流程为着力点构建过程约束机制。^[19] 实践中,案件监管平台与法官办案系统相互连接,监督主体可以实时动态“观看”案件裁判过程。法官操作界面同步显示审限倒计时、案件评查反馈等信息,而院庭长可实时查看审判案件的流程状态。^[20] 例如盐城市法院系统将刑事、民事、行政案件从立案到归档拆分为诉前调解、立案登记、开庭审理、裁判阶段、判后跟踪五大模块,每个模块下设若干可锁定、可预警的节点。^[21] 根据影响审理效率的重点因素,审判执行的流程节点可被分为已有明确规范的硬性节点和没有规范但容易造成审理拖延的软性节点,并分级分析管理。^[22] 这进一步细化了审判监管的颗粒度,并基于审判权运行过程而动态呈现。技术构建了自动运行、可持续化、隐蔽性的权力监管机制。

(三) 责任推诿的风险扩大

在智能辅助中,人机协同智能审判以“智能辅助+法官人工审查”为运行方式。在事实认定环节的争议事实和争议焦点归纳中,智能化平台先归纳无争议事实或校验证据的缺失性或合规性,在此基础上法官再进行人工审查。此时法官仍享有自主选择或添加的裁量权。在法律适用环节,现有规范对法官的智能检索行为课予了明确且严格的义务和责任负担。2020年《最高人民法院关于统一法律适用加强类案检索的指导意见(试行)》第3条规定,承办法官依托中国裁判文书网、审判案例数据库等进行类案检索,并对检索的真实性、准确性负责。截至2023年10月24日,我国现行有效法律共有299部。^[23] 行政法规和各类规章更是不胜枚举。此外,还有大量类案、学理等说

[16] 参见《杭州互联网法院区块链智能合约司法应用上线 存证量超19亿条》,载中华人民共和国最高人民法院网站2019年10月28日, <https://www.court.gov.cn/zixun/xiangqing/194221.html> # : ~ : text = %E6%9D%AD%E5%B7%9E%E4%BA%92%E8%81%94%E7%BD%91%E6%B3%95%E9%99%A2%E5%8C%BA,%E8%AF%81%E3%80%81%E5%85%A8%E6%96%B9%E4%BD%8D%E5%8D%8F%E4%BD%9C%E3%80%82。

[17] 参见齐延平:《数智化社会的法律调控》,载《中国法学》2022年第1期,第97页。

[18] 参见[美]劳伦斯·莱斯格:《代码2.0:网络空间中的法律》,李旭、沈伟伟译,清华大学出版社2018年版,第363页。

[19] 参见樊传明:《审判管理的运行体系和制度逻辑》,载《法学研究》2024年第5期。

[20] 参见《人民法院司法改革案例选编(八)【案例133】浙江省高级人民法院:制度建设和科技创新双管齐下 完善“四类案件”全流程在线监督管理体系》,载最高人民法院网站2020年7月10日, <https://www.court.gov.cn/zixun/xiangqing/240641.html>。

[21] 参见李江蓉:《智能管控审判流程节点的实践与探索》,载《人民司法》2022年第13期,第74—78页。

[22] 参见《人民法院司法改革案例选编(八)【案例128】北京互联网法院:构建三维立体式审判监督体系 全面加强在线诉讼案件质量管理》,载最高人民法院网站2020年7月10日, <https://www.court.gov.cn/zixun/xiangqing/240641.html>。

[23] 参见《现行有效法律目录(299件)》(截至2023年10月24日十四届全国人大常委会第六次会议闭幕 按法律部门分类),载中国人大网2023年11月2日, http://www.npc.gov.cn/c2/c30834/202311/t20231102_432724.html。

理依据需要法官掌握。^[24] AI模型辅助运用端中的智能化应用可以直接将与案情相关的法律规范和类案推送给案件裁判法官。面对海量法律法规、类案裁判的处理、分析和运用,法官在借助智能技术提高效率的同时,也需要通过实质性审查保证公正审判。

然而,智能化应用为提高审判能力和审判效率的同时,存在“智能辅助+法官放弃裁量”的决策风险,由此导致法官避责行为的产生。如美国密西西比州的一名联邦法官被指控,其依靠人工智能起草的一项裁决因明显事实不准确而损害当事人利益。^[25] 这源于智慧司法中的技术依赖,即当司法决策依赖技术系统生产、提供信息,并将信息结构化与要素化时,机器决策开始侵蚀人类决策,导致技术依赖生成。^[26] 目前,尽管法官享有自主选择或添加的裁量权,但在案多人少的审判压力下,法官往往会直接接受智能化应用作出的结果,由此导致以智能辅助“替代”法官行为。当审判主体和决定者难以特定时,权力边界也就变得模糊不清,司法问责制就容易流于形式,至少推卸责任的可能性被大幅度扩充。^[27] 同时,制度规范又进一步提高了对类案推送结果的认可度,塑造了公正性外观。2018年《关于进一步全面落实司法责任制的实施意见》明确,各级人民法院应当在完善类案参考、裁判指引等工作机制基础上,建立类案及关联案件强制检索机制,确保类案裁判标准统一、法律适用统一。存在法律适用争议或者“类案不同判”可能的案件,承办法官应当制作关联案件和类案检索报告,并在合议庭评议或者专业法官会议讨论时说明。实践中,考虑到照搬类案省时省力,不会引发数据异常报警,且能避免领导质疑和当事人上访的风险,会被绝大多数法官欣然接受。^[28] 因此,法官可能将机器的提示或校验作为增强自身决策合理性的理由,由此扩大了责任推诿的风险。

三、人机协同智能审判的应用特征与责任优化

司法审判的人机协同模式要求合理界分人类智能和机器智能的责任承担,并穿透至技术背后的监管者和技术人员,明确司法责任中不同主体间的责任配置和承担方式。以下本文将结合人机协同智能审判的特征探究司法责任的优化路径。

(一) 智能辅助定位中的人机责任界分

人工智能在司法裁判中仅具有辅助性。2022年12月9日最高人民法院发布的《关于规范和加强人工智能司法应用的意见》既明确了促进人工智能同司法工作深度融合的路径规划,也指出了人工智能辅助审判的功能定位。^[29] 人机责任界分应促进司法人工智能积极应用与保障公正安

[24] 参见高尚:《赋能还是解构:数字时代司法如何说理》,载《探索与争鸣》2024年第2期。

[25] See Eric Hastings, *Mississippi Judge Retracts AI-Flawed Ruling Over Fake Citations*, WebProNews (1 August 2025), <https://www.webpronews.com/mississippi-judge-retracts-ai-flawed-ruling-over-fake-citations/>. See Gary Marchant, *Judge Accused of Using AI to Issue Garbled Ruling; If You've Asked ChatGPT a Legal Question, You May Have Accidentally Doomed Yourself in Court*, Bloomberg Law (30 July 2025), <https://news.bloomberglaw.com/us-law-week/judges-who-benefit-from-ai-technology-must-avoid-its-hazards>.

[26] 同前注[3],张凌寒文,第186—187页。

[27] 同前注[3],季卫东文,第132页。

[28] 参见徐骏:《智慧法院的法理审思》,载《法学》2017年第3期,第61页。马靖云:《智慧司法的难题及其破解》,载《华东政法大学学报》2019年第4期,第114页。

[29] 《关于规范和加强人工智能司法应用的意见》提出,到2025年,基本建成较为完备的司法人工智能技术应用体系,为司法为民、公正司法提供全方位智能辅助支持。到2030年,建成具有规则引领和应用示范效应的司法人工智能技术应用和理论体系,为司法为民、公正司法提供全流程高水平智能辅助支持,应用效能充分彰显。

全有序的价值平衡。

首先,法官责任的认定应当遵循责任主义的归责原则。责任制度的根本目的是扩大负责(守法)行为的概率,而不是对不负责任(犯法)的行为施加法律责任。^[30]以自我决定为根据并可以对决定过程予以解释的行为自我答责是责任主义原则的基本原理。^[31]易言之,人类只对通过自由意志而选择的行为或自己的决断承担责任。为此,司法责任的要义之一在于责任的可预见性,即法官对自己履职的所作所为了然于心,且能够对将来可能要承担的责任之有无、份额、轻重等有所预见。^[32]但在人工智能参与下的裁判活动,由于算法的不可解释性,当法官自身也无法对其作出合理解释时,让法官承担智能裁判的不利后果不仅无法实现预防违法审判行为的制度功能,同时会因过重的责任负担,抑制法官进行案件裁判的主动性和创造性。同时,司法智能化建设积极推进具有深刻的现实原因:一是科学应对统一法律适用改革目标与法官专业业务水平参差不齐的现实困境;二是有效平衡案多人少的审判压力和司法高效的司法改革目标。但目前我国智慧法院仍处于初级阶段,存在实践应用不充分、系统融贯性不畅等问题,^[33]这需要有效协调法官责任与技术固有局限或发展不充分之间的关系。以过错程度为责任认定标准有助于实现权责统一,并促进智慧司法的发展适用。

其次,基于产品责任区分人机责任,并明确法官责任的合理范围。《中华人民共和国民法典》第1202—1204条分别将产品生产者、销售者和有过错第三人作为责任主体。由于人工智能并不具有人格属性,应由人工智能的生产者、销售者、设计者、管理维护者和数据提供者,以及有过错的第三人承担相应的产品责任。在具体适用上,因司法智能化应用仅具有辅助审判的功能,自然人法官应对审判的公正性、合法性,即是否有“司法过错”承担首位责任。认定方式采用过错推定原则,即先推定法官存在审查过失,除非其能证明已尽合理注意义务。人工智能不独立承担法律责任,仅作为法官免责或减责的“事由”,且以“合理不可识别”为边界。^[34]若法官已尽合理审查义务仍无法发现错误,可免责;若凭一般认知即可发现错误但仍采信,则法官有重大过失,不得免责。

(二) 权力分散化下的责任均衡适配

智能技术拉近了审判者和监管者、社会公众之间的距离,表现出短距式且多中心监督的特点。责任的均衡适配旨在充分考虑制度设计、技术约束、主体能力等因素的影响,确保从技术研发到裁判落地的全链条中,权责清晰、风险可控、公正可感。

首先,法官、技术人员和监管者的责任均衡适配。生成式人工智能的发展使其具备了类人类的意志特征,如认知能力、实践理性能力和价值判断能力,推动了智能审判的人机深度融合。^[35]智能审判可能呈现出法官、技术人员和审判监管人员的“合作”。而司法责任制改革的价值目标是完善新型审判权力运行机制,实现司法为民和公正司法,切实落实“让审理者裁判,由裁判者负责”的要求。司法责任的智能化应该构建以审判者为主体的责任机制,合理界分法官、技术人员和审

[30] [澳]皮特·凯恩:《法律与道德中的责任》,罗李华译,张世泰校,商务印书馆2008年版,第92页。

[31] 参见刘艳红:《人工智能的可解释性与AI的法律责任问题研究》,载《法制与社会发展》2022年第1期,第85页。

[32] 参见贺宁:《法官责任制度的双重逻辑与价值衡量》,载《法制与社会发展》2021年第1期,第133页。

[33] 参见刘艳红:《人工智能技术在智慧法院建设中实践运用与前景展望》,载《比较法研究》2022年第1期,第1—3页。

[34] 同前注[3],卞建林文。

[35] 同前注[9],袁曾文,第25—27页。

判监管者之间的责任边界,以明确审判监督者和信息技术人员介入范围,保障人机协同审判中的法官主体地位。具体地,技术人员应承担技术可靠性的法律责任,而审判者则是裁判责任的“最终承担者”,承担实质审查义务、技术合理使用义务和基于价值平衡的合理裁量义务。审判监管者作为风险防控者,应当进行系统风险防控和正当程序监管。由此促进法律治理与技术治理的二元互补。

其次,从技术设计至应用以及运行维护全过程的责任均衡配置。智慧法院的规范体系建设包括两套规则体系:即技术建设规则和应用风险规则。前者旨在明确各类智能化应用的技术要求和管理要求,如2018年最高人民法院对电子卷宗管理技术的规范与明确,《人民法院在线运行规则》及2021年《最高人民法院办公厅关于推进司法数据中台和智慧法院大脑建设的通知》等。^[36] 后者如《人民法院在线诉讼规则》《人民法院在线调解规则》。2022年的《最高人民法院关于规范和加强人工智能司法应用的意见》也遵循从技术建设标准化到功能应用规范化的设计路径。可见,“智慧法院”是一个具有多层含义的概念,不仅仅是“人工智能审判”。^[37] 同时,在实践效能上,数据、算法等基础设施不仅影响司法大数据的发展,甚至可能造成智能应用无法有效运行,影响案件裁判。^[38] 因此,责任的均衡适配需从审判阶段穿透至智能应用背后的建设、生产环节,^[39]减少因技术错误而造成的司法不公,明确各阶段的权利义务关系。

(三) 以可信生态构建缓解价值张力

可信人工智能是指系统本身在安全性、可解释性、公平性、可靠性、合规性和透明度等各方面具有可信性。^[40] 人机协同智能审判模式与传统审判模式间的价值张力需要建立可信的智慧司法,以调和逻辑和法治逻辑的张力。

可信生态的正义观有助于协调智能审判中法律因果判断与预测性相关性判断间的张力。司法大模型的技术路径包括基于规则的专家系统和基于数据的机器学习。前者是人工智能法律推理知识系统模型的早期探索,将规则编码为一种简单的、程式化的“如果-是”的逻辑结构,模仿人类的推理方式构建一个巨大的“逻辑森林”,让人工智能模仿人的推理方式。优势是可以借助符号从其知识路径中进行解释、说理,提高推理路线的解释性。^[41] 但由于建构耗时较大,并非所有的知识都可以归纳为“逻辑森林”,且难以处理复杂问题,也只能用于简单案件中。^[42] 机器学习则是让机器从大量的数据中学习和提取知识,而无须人工明确编程。其可以处理大规模数据,并具有适应性和学习能力的优势,但存在算法黑箱、歧视等问题。目前主流的人工智能模型大多基于神经网络和深度学习技术进行训练和预测,^[43]基于裁判文书数据来预测和模拟司法裁判。这一技

[36] 参见《最高人民法院关于进一步加快推进电子卷宗同步生成和深度应用工作的通知》(法发[2018]21号),其中明确“做好技术支持。抓好评估考核。各高级人民法院应结合辖区法院工作实际,开展系统功能建设和应用成效两方面的评估考核”。

[37] 同前注[3],季卫东文,第131页。

[38] 参见刘雁鹏:《司法大数据建设的现状、困境与对策》,载《中国社会科学院研究生院学报》2019年第5期,第120—121页。

[39] 同前注[9],袁曾文,第29—30页。

[40] 参见谢思敏、王铮、周擎擎:《科技向善:可信人工智能赋能中国式现代化》,载《中国经济报告》2025年第2期。

[41] See Richard E. Susskind, *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal Reasoning*, 49(2) *The Modern Law Review* 168 (1968).

[42] 参见刘东亮:《新一代法律智能系统的逻辑推理和论证说理》,载《中国法学》2022年第3期,第150页。

[43] 参见衣俊霖:《论人机协同审判的信任构建》,载《中国法学》2025年第2期,第157页。

术特征使审判过程呈现出以“数据相关性”替代“法律因果性”，以“统计规律”覆盖“价值判断”的倾向。如在证据认定的智能预测中，美国法院将不同情形的证据关系和假定解释转换为数量值，并通过一些变量信息来获取其他概率信息，将证据能力中的相关性、真实性、合法性判断转化为概率判断。^{〔44〕} 这一运行方式的正当性隐忧是将司法理性通约为数据运算，侧重于司法活动中隐藏规律的发现，排除价值判断。这不仅导致追责过程的模糊化，也难以实现实质正义。^{〔45〕} 而可信数据空间是基于共识规则，连接多方主体，实现数据资源共享共用的一种数据流通利用基础设施，是数据要素价值共创的应用生态。其强调人权保障、公平原则、反歧视原则等。^{〔46〕} 价值层面的伦理对齐可以对多元主体的行为形成约束，以可信再塑智慧司法的公正性。

可信生态中沟通商谈的程序有助于回应司法高效与司法公正间的价值张力。技术治理遵循的是效率的逻辑，法律治理遵循的是权利的逻辑。传统上审判程序的两造对立、沟通对抗、证据辩论、逻辑推理等将法庭活动作为一种“剧场表演”以促成法律活动的技术化和专门化，增强法律的神圣性和权威。^{〔47〕} 这构成了审判公正的可信外观。然而，在线诉讼、在线调解、证据认定、裁判结果的智能预测解构了传统的审判场域，瓦解了传统审判的仪式感与权威性。同时，审判效率导向下的智慧司法存在明显的工具理性倾向，人类的自我认知存在对技术的过度依赖，甚至被技术所左右的困境。^{〔48〕} 责任推诿等认知偏误的产生要求理性认识智能审判对司法效率的意义及其中的人机协作关系。可信生态要求人工智能在程序上实现透明性、可解释性，保障公民的知情权，让人们了解人工智能是如何运作的。^{〔49〕} 同时，程序性伦理规则是聚焦利益相关者的沟通过程，其通过提出大家觉得公平合理的论辩标准与论辩规则，并在充分沟通后，基于这些论辩标准和规则得出一个评价风险决策合理性的道德规范。^{〔50〕} 这成为算法知情权、程序选择权、人工救济权的制度空间。因此，可信生态中的沟通协作型治理模式有助于塑造程序外观，为价值共识的达成提供数字空间。

四、司法责任智能化因应的制度展开

在智能审判模式中，司法责任的完善应从法官主体地位出发，以正当程序和伦理责任的法治化完善责任类型，并通过风险分级实现责任与场景的功能适配，由此明确人机协同司法责任的制度内涵。

（一）以程序义务提升追责过程的清晰度

2017年《最高人民法院关于加快建设智慧法院的意见》提出要“推动流程再造促进审判高效有

〔44〕 See D. Thompson, *Creating New Pathways to Justice Using Simple Artificial Intelligence and Online Dispute Resolution*, 2(1) *International Journal of Online Dispute Resolution* 4 (2015).

〔45〕 参见雷磊：《完美的效率还是负责任的正义？》，载《东方法学》2025年第4期。

〔46〕 参见姚力博：《论我国信息治理的信任建构及其路径选择》，载《天府新论》2025年第3期，第127—131页。

〔47〕 参见舒国滢：《从司法的广场化到司法的剧场化——一个符号学的视角》，载《政法论坛》1999年第3期，第16—18页。

〔48〕 参见郑戈：《在法律与科技之间——智慧法院与未来司法》，载《中国社会科学评价》2021年第1期，第92页。

〔49〕 参见李功源、刘博涵、杨雨豪、邵栋：《可信人工智能系统的质量属性与实现：三级研究》，载《软件学报》2023年第9期；杨建军：《可信人工智能发展与法律制度的构建》，载《东方法学》2024年第4期，第104页。

〔50〕 参见季卫东、赵泽睿：《人工智能伦理的程序保障——法律与代码的双重正当化机制设计》，载《数字法治》2023年第1期，第64—65页。

序运行……促进程序公正与实体公正的有机统一”。智能审判中司法责任的认定应以程序义务的明确提高追责过程的客观化和清晰度,提升社会公众对智慧司法的信任。技术性正当程序为智能审判中正当程序的落地提供了具体的分析工具和规范基础。

首先,以法官的程序义务弥补技术交互对传统诉讼言辞辩论空间的压缩。“欧美的主观程序正义研究大多强调‘表达机会’的重要性,进而用‘过程控制’解释其中的机制。”^[51]相较于“表达机会”等技术因素,司法人员是否有礼貌、当事人的主体地位是否体现等因素对中国公众的司法信任具有更重要的影响。^[52]一方面,智能审判程序的启动应以当事人的同意为基础。^[53]司法审判的最终目的是保障公民合法权利。司法智能化的平等适用以社会公众具有平等使用现代技术的能力为预设。但实践中诉讼当事人双方的技术运用能力势必存在较大不同。因此,诸如在线诉讼等智能化应用的启动应以当事人同意为前提。由此通过充分汲取当事人的意见,避免技术鸿沟所带来的新的司法不公正,促进接近正义的实现。如《人民法院在线调解规则》第15条规定在线调解前应当征求当事人的意见。^[54]另一方面,法官应对智能应用产生结果承担说理义务。这既体现在裁判文书中的说理,也包括庭审过程中的说理。具体包含三项基本要素:决定是什么、决定是如何做出的以及为什么作出某一决定。^[55]同时,为保障当事人程序选择权的实质性实现,法官应承担相应的释明义务。正如《人民法院在线调解规则》第7条要求法官在在线调解前应当告知当事人在线调解的主要形式、权利义务、法律后果和操作方法等。^[56]因此,智能审判在为法官提供便利的同时,仍需要保障当事人的程序权利,并以法官的程序义务避免当事人因数字能力不足而权利受损。

其次,以设计者、研发者等技术人员的程序义务构建不同主体间的沟通协商平台。算法黑箱、不可解释性因沟通协商空间的压缩而造成正当程序原则的削弱。为提高当事人和公众对智能审判程序的接受度,技术人员可在司法智能化应用的建设阶段与法官、社会公众进行相应的沟通协商,从而提高责任认定过程的客观性和清晰度。一是建设阶段的解释适配性:以通俗语言说明算法逻辑。技术性正当程序要求在平台的架构选择和算法设计中融入排除偏见、听取意见等要素,特别是对作出的决定必须说明理由。^[57]技术人员和法官的充分合作有助于智能技术和审判规律的深度融合,提高司法智能化应用的有效性。审判活动是一种高度智能化且包含广泛政策与价值目标实现的活动,很难完全用数据进行准确测量与评价,即使是能够客观测量

[51] 冯健鹏:《基于主观程序正义的技术性正当程序——我国算法治理的一种路径》,载《法制与社会发展》2025年第2期,第154页。

[52] 参见苏新建:《程序正义对司法信任的影响——基于主观程序正义的实证研究》,载《环球法律评论》2014年第5期,第30页。

[53] 参见郭金鸿:《道德责任论》,人民出版社2004年版,第117—122页。

[54] 《人民法院在线调解规则》第15条规定:“人民法院对当事人一方立案前申请在线调解的,应当征询对方当事人的调解意愿。调解员可以在接受人民法院委派调解之日起三个工作日内协助人民法院通知对方当事人,询问是否同意调解。对方当事人拒绝调解或者无法联系对方当事人的,调解员应当写明原因,终结在线调解程序,即时将相关材料退回人民法院,并告知当事人。”

[55] 参见刘东亮、常纯:《技术性正当程序在网络平台治理中的运用——以企业合规管理为视角》,载《西华大学学报(哲学社会科学版)》2023年第4期。

[56] 《人民法院在线调解规则》第7条规定:“人民法院立案人员、审判人员在立案前或者诉讼过程中,认为纠纷适宜在线调解的,可以通过口头、书面、在线等方式充分释明在线调解的优势,告知在线调解的主要形式、权利义务、法律后果和操作方法等,引导当事人优先选择在线调解方式解决纠纷。”

[57] 同前注[55],刘东亮、常纯文。

的部分,往往也只能显示职能是否被充分执行,而无法表明执行的真正价值。^[58] 智慧司法场景下智能化应用的专业性特征需要技术人员和法官专业人员的充分合作。二是运行维护阶段的动态监控与及时响应。数学模型的本质是基于过去的数据来推测未来,基本假设是:模式会重复。^[59] 而法律需要在持续的重复制作中获得持久的生命力。^[60] 数据、算法等智能应用的基础设施应不断回应案件审判的当下价值,保障个案正义的实现。

(二) 以伦理责任建立智慧司法可信生态

相较于程序义务的制度刚性、技术可验证性与规则治理,伦理责任侧重于价值弹性与原则治理。2021年9月25日发布的《新一代人工智能伦理规范》在第五章从伦理规范角度明确了智能应用的使用规范,并提出了5项具体要求:提倡善意使用、避免误用滥用、禁止违规恶用、及时主动反馈和提高使用能力。^[61] 伦理责任建设应构建从设计研发到应用审判的可信智慧司法生态。

首先,明确法官伦理责任以促进司法裁量权的规范化行使。智能技术的伦理准则之一在于超越高效性这一单纯的功利主义、技术理性考虑。基于法官的裁量不同,案件的判决将会存在很大的差异。智能审判的裁量难题使得标准化裁判无法成为现实可能。^[62] 法官的伦理责任要求遵循实质正义原则,结合个案不同情况合理行使司法裁量权,避免裁量懈怠。《伦敦司法伦理宣言》(London Declaration 2010 on Judicial Ethics)更加强调了自由裁量权、尊重与倾听能力、称职与透明度等伦理标准应当成为司法官的共同价值观。我国的《法官职业道德基本准则》也应积极回应智慧法院的应用需求,对公正、廉洁、为民原则予以拓展。以“技术赋能下司法裁量的合理规范运行”为核心逻辑构建适应法官的伦理框架:一是在审判过程中,法官可对司法智能化应用生成的内容进行合法性审查,保障智能技术生成内容的公正性,尤其是在事实认定和法律适用等关系裁判结果公正性的实质环节。二是对裁判结果的实质说理。数据模型均来源于过去的案件事实或要素。这就意味着基于数据模型进行的裁判是过去式的。但“社会现实的变化是生活的法律,回应社会现实的变化则是法律的生命”^[63]。法官应通过裁判说理提高当事人的接受度,实质化解争议。

其次,以技术人员的伦理责任保证智能应用的价值对齐。在技术逻辑和法律逻辑的深度融合中,算法之治需要始终与法治相契合,鼓励内部算法师和外部算法师对其行为承担责任,并培养、发展其责任伦理以及以人为本的基本价值观。^[64] 司法审判同样兼具法理、情理、裁判效果,即需要法律效果、政治效果与社会效果的统一。由此司法裁判过程不仅要严格适用法律,还要有情理的温度和人文的关怀。^[65] “代码即法律”的技术逻辑要求强化数据处理部门、电脑工程师等生产者、研发者的伦理责任,将智慧司法与公序良俗等司法价值的对齐拓展至生产研发阶段,预防技术研发阶段的产品责任风险。基于公平公正原则,强化技术人员在数据库搭建、算

[58] 参见龙宗智,孙海龙:《加强和改善审判监督管理》,载《现代法学》2019年第2期,第49页。

[59] 参见宋旭光:《论司法裁判的人工智能化及其限度》,载《比较法研究》2020年第5期,第86页。

[60] [美]卡多佐:《司法过程的性质及法律的成长》,张维编译,北京大学出版社2012年版,第89—90页。

[61] 参见《新一代人工智能伦理规范》第18—22条。

[62] 参见孙海波:《论法律的数字化与司法裁判的标准化难题》,载《行政法学研究》2025年第5期,第30页。

[63] [以]巴拉克:《民主国家的法官》,毕洪海译,法律出版社2011年版,第16页。

[64] 参见[英]维克托·迈尔-舍恩伯格、[英]肯尼思·库克耶:《大数据时代:生活、工作与思维的大变革》,盛杨燕、周涛译,浙江人民出版社2013年版,第223—226页。

[65] 参见罗恬澹、段陆平:《“人工智能法官”的四大伦理障碍》,载《民主与法制时报》2021年10月13日,第3版。

法运行、大模型训练中的平等、无偏见等司法价值运用,以此应对算法黑箱、算法歧视等技术局限。

(三) 比例原则下责任的分级配置

在为技术理性融入法治逻辑和价值关怀的同时,还需要秉持敬畏、开放与包容的态度,即通过法律鼓励创新、包容错误。^[66] 司法责任应基于比例原则实现不同场景下责任与风险的适配,以期达到促进效率与公正的价值平衡。

司法责任应促进不同场景内风险程度与责任内容合乎比例且功能适配。法官裁判责任追究的目的在于:一方面对法官起到警示的作用,间接地发挥防范错误裁判的功能,另一方面是对受害人和民众起到安抚的作用,修复人们对司法的信任。^[67] 人工智能规范体系的特点之一是以场景化应用为规范对象,强调规范与技术、场景的适配性。^[68] 目前,司法智能化应用已广泛适用于案件审理的全过程,包括法律适用、事实审查以及制作裁判文书等各环节,例如实质审判中的法条推送、类案推送、证据校验等,裁判阶段的文书智能生成系统等。司法责任应基于风险强度的不同层级予以类型化。基于智能辅助的应用场景是否直接涉及案件裁判,是否直接影响司法公正,可以区分为决策型的智能辅助和审判事务性工作的智能辅助。^[69] 前者如类案推送等,需要结合智慧司法的专业性特征予以场景化适配。因会对法律适用、事实认定等决策行为产生影响,更关系案件裁判的公正性,前者应用于中高风险场景(如人工智能量刑建议等对案件裁判活动的智能辅助),具体包括自主决策场景(如小额诉讼智能裁判、简单类案的要素式智能审判)和决策支持场景(如量刑预测、证据链分析等)。而后者如电子卷宗、庭审笔录和合议笔录等,更多的是通用性人工智能在司法场景的平移运用,多应用于低风险场景(如语音转写等审判事务性活动的智能辅助)。责任分级的意义在于促进不同场景下责任内容与风险程度的适配。如决策支持场景中,可设置过程性的强制人工复核节点。而在事务性辅助中,如语音转写,以事后抽检为主。

五、结 语

大数据、人工智能技术的迅猛发展给法官高效公正审判、社会公众接近司法提供了助益。智能技术在影响审判活动运行方式的同时,也深刻改变了审判权运行的结构环境。智能技术的加持使技术人员直接或间接地参与审判过程和裁判结果的作出。同时,智能技术的高效性特征使审判监管者的监管呈现出静默化、全流程、细微化、短距性的特点。审判监管权也实质性、全方位地参与案件审判过程和裁判结果的作出。由此,智能技术促使司法审判结构呈现出多元主体协同参与的特点。这需要从整体上理解司法责任制度智能化背后复杂的行动体系结构和主体间关系。司法制度改革中“放权”和“责任”的均衡逻辑为司法责任制的智能化因应提供了完善路径。人机协同智能审判中的司法责任需要合理划分法官、技术人员、监管者之间的责任边界。以权力、责任和保障三者之间的结构均衡为原则明确法官责任,并通过明确技术人员、审判管理者的责任,避免相

[66] 参见张骥:《在技术理性与法治之间——交错时空中的人文主义法学立场》,载《法商研究》2023年第6期。

[67] 参见郭延军:《我国法官裁判责任的追究限度》,载《法学评论》2021年第6期,第21页。

[68] 参见林涸民:《论人工智能立法的基本路径》,载《中国法学》2024年第5期,第90—91页。

[69] 参见左卫民:《从通用化走向专门化:反思中国司法人工智能的运用》,载《法学论坛》2020年第2期。

关主体利用智能技术对法官审判权造成不利影响,由此保障审判中心主义和庭审实质化的实现,探索“负责任”的智慧司法。

Abstract Artificial intelligence empowers judicial trials and is widely applicable in trial activities such as evidence identification, legal application, judgment making, and trial supervision due to its advantages of efficiency and convenience. However, this also poses challenges to judicial responsibility, such as vague subjective attribution, unclear delineation of responsibility subjects, and an increased risk of responsibility shifting. The reason is that technological limitations lead to logical contradictions between causal and correlation judgments in the trial process, and the implicit embedding of judgments by technicians and regulators undermines the judge's subjectivity and faces a value tension between efficiency and fairness. Therefore, judicial responsibility should be balanced between judges, technicians, and regulators, as well as between design, application, and operation. A communication and negotiation space should be established through technical due process to enhance the clarity of the accountability process. A trustworthy judicial ecosystem should be established based on ethical responsibility to enhance the credibility and authority of smart justice. At the same time, risk classification is carried out based on the impact of technology on trial fairness in different scenarios, and then the technical and ethical responsibilities of each subject are clarified based on the principle of proportionality, promoting the functional adaptation of risks and responsibilities.

Keywords Judicial Responsibility System, Smart Adjudication, Technical Due Process, Ethical Responsibility

(责任编辑:樊传明)