

人工智能时代的司法公正： 价值效用与风险防范

李 婷

内容提要 随着各行业智能化水平不断提升,我国司法系统在面对时代的剧烈变化时,展现了对智能化建设和司法改革的魄力与决心。人工智能可以在“类案类判”和“及时公正”等维度促进可视化的司法公正,但同时也在与司法活动交叉结合的过程中产生了诸如数据算法缺陷弱化公正基础、技术的价值中立性无法真正满足公民个案公正的诉求、技术的工具性消解司法公正本意的风险。为应对以上风险,可以采取大数据与算法规范、制定人工智能融入司法的伦理规范以及将法律论证嵌入司法人工智能系统等措施,在展现司法人工智能优势的同时,实现司法公正的最大化。

关键词 人工智能 司法公正 算法 风险防范

李 婷,南京信息工程大学法政学院副教授

一、研究背景

科技的迅猛发展使各行业各领域的信息化、智能化水平不断提升,我国司法系统在面对时代的剧烈变化时,展现了信息化、智能化建设和司法改革的魄力与决心。理论界对于人工智能和司法互动的研究,最早可以追溯到 1970 年 Bruce G. Buchanan 与 Thomas E. Headrick 共同发表的文章“Some Speculation about Artificial Intelligence and Legal Reasoning”,该文重点研究了人工智能系统在法律推理过程中的应用和问题^[1]。我国理论界对计算机数据系统参与司法过程的最早设想可以参考龚祥瑞、李克强于 1983 年发表的《法律工作的计算机化》,该文预见了我未来法律工作计算机化的必然趋势以及法学研究和法律工作现代化、信息化的必要性。在之后的理论研究中,“网络化”“信息化”“电子

本文为江苏省社会科学基金项目“新就业形态下劳动者权益的法律保障研究”(21FXB009)、2019 年南京信息工程大学教学改革研究课题“科学主义理念下‘三元区分’法学教育模式的构建”的阶段性成果。

[1]参见 Bruce G. Buchanan, Thomas E. Headrick, "Some Speculation about Artificial Intelligence and Legal Reasoning", *Law Review*, 1970, 23(1).

计算机”等词语也越来越多地出现在法学研究领域。随着人工智能技术的完善和成熟,理论界许多学者从各个角度对人工智能参与司法展开热烈讨论。例如季卫东尝试从司法权转变的角度探讨人工智能在司法过程中扮演的角色^[1],郑戈以“算法”为切入点,强调人在算法中的地位和作用^[2],左卫民基于中外司法人工智能的对比,对AI法官在我国的应用前景做了展望^[3]。

在实务中人工智能的发展也取得了长足的进步。最高人民法院于1999年颁布《人民法院五年改革纲要》,明确提出了信息技术、在线审判等关键词,并以此作为司法现代化改革的重点步骤。之后的种种延续性政策也反映了我国对相关技术保持着积极态度。当前,在我国大力推进“智慧司法”这一背景下,除了大家熟知的文书制作电子化(即将证据材料等纸质卷宗以电子文书的形式呈现)、司法信息数据化(即网上办案平台和在线庭审等)、电子数据的生成和审查(即文字、图像和声音的识别技术)等辅助性司法工作以外,各级“智慧法院”在立案、审判与结案环节都努力促成人工智能与司法活动的结合。在立案环节,人工智能系统可以整理案件事实并对其进行初步“画像”;在审判环节,人工智能系统可以提取案件争议焦点、生成庭审提纲;而在结案环节,人工智能可以自动生成裁判文书、提供预测结果与量刑建议,甚至划定量刑标准、发出量刑预警^[4]。

学者们对人工智能介入司法提高司法活动效率、解决“案多人少”矛盾方面的积极意义几乎没有争议,有疑义的是,司法人工智能能否实现司法公正这一目标?又可以在何种范围和程度上促进司法公正?面对司法人工智能的迅速发展,部分学者对人工智能司法的前景持相对乐观的态度,认为如果可以进一步升级计算机硬件和算法,司法人工智能在未来将获得更广泛的使用。同时,他们还认为,在司法独立与公正价值方面,人工智能司法相较于法官司法更具优势。因为人工智能司法可以毫无偏见、毫无感情地运作,排除影响人类法官司法裁决的一切干扰^[5]。

当然,更多学者对人工智能司法的态度较为谨慎。左卫民教授认为,尽管智能化技术已经为突破司法服务的困境提供了可能性,但与域外人工智能的运用相比,我国人工智能还处于“弱形式”阶段,还未运用在审判决策的重要环节,尤其是未在量刑环节发挥实质性作用。加上我国人工智能司法是基于知识图谱式算法而缺少定量实证方法的使用,计算模型是规范取向而非经验模式,“最终产出的不是实践理性的AI法官,而是被‘训练了的’AI法官”^[6],这样的AI法官无法真正实现实质正义的司法价值。雷磊教授从实体公正与程序公正的区分入手,结合人工智能运行的基本原理与人工智能技术的封闭性、被动性和价值中立等特点,认为人工智能介入司法在实现司法实质公正方面具有不可克服的缺陷^[7]。

基于此,我们需要思考的问题是:司法人工智能的运用现状与乐观主义者的预期是否吻合?就目前而言,“弱形式”的人工智能实践在何种范围与形式上能够促进司法公正价值,以及面临何种障碍与挑战?思考人工智能介入司法大趋势的当下,追求司法公正这一古老话题能否在新技术革命的激励下获取新能量,科技的变革如何与人们对崇高价值的追求产生良性互动,是本文写作的主旨。

二、人工智能促进司法公正的价值效用

在法律与诸多价值交融的过程中,公正始终被认为是与法律联系最为密切的价值。在研究法律

[1]参见季卫东:《人工智能时代的司法权之变》,《东方法学》2018年第1期。

[2]参见郑戈:《人工智能与法律的未来》,《探索与争鸣》2017年第10期。

[3]参见左卫民:《AI法官的时代会到来吗——基于中外司法人工智能的对比与展望》,《政法论坛》2021年第5期。

[4]孙庆春:《人工智能司法决策研究》,重庆大学博士论文2021年,第85页。

[5]参见David J. Walton, *Litigation and Trial Practice in the Era of Big Data*, *Litigation*, 2015, 41, pp.55-58.

[6]左卫民:《AI法官的时代会到来吗——基于中外司法人工智能的对比与展望》,《政法论坛》2021年第5期。

[7]雷磊:《司法人工智能能否实现司法公正?》,《政法论丛》2022年第4期。

应当体现哪些价值或法律应当保护哪些价值的问题上,公正“占有最为显赫之地位”^[1]。习近平总书记就法治问题谈道:“法治不仅要求完备的法律体系、完善的执法机制、普遍的法律遵守,更要求公平正义得到维护和实现”^[2]。因此,谈论人工智能对司法活动的影响首先就是要理清人工智能与司法公正之间的关系,人们在对新时代、新技术感慨万千的同时也要守住法的核心价值,这也是本文的基本论调。

司法公正可以从广义和狭义两个维度上来解释。广义上的司法公正包含司法程序的各个领域和各个方面,与司法的权威性、司法活动被社会伦理的认同程度、司法制度的宏观构架,以及司法程序的合理性相关^[3]。由于广义的司法公正范畴过大,本文在讨论人工智能对司法公正的影响时,将采用狭义的司法公正观,即主要讨论司法裁判活动是否符合公正的价值诉求。而(狭义)司法公正又包括两种类型:司法结果具有公正价值和司法程序具有公正价值。前者——实体公正是司法公正的根本目标,后者——程序公正则是司法公正的重要保障^[4]。

1. 程序公正维度:类案类判

在司法公正价值的众多表现中,“依法裁判”是其核心内容。这也就是说,法官必须依照事先颁布的一般性规则来处理案件,即适用“可视化”的规则进行“类案类判”,人工智能对司法公正的良性影响在此处可供发挥的空间很大。

其一,法律要求普遍性。依照人工智能算法的逻辑,每一种结果的产出总是伴随着相关性的依据,而这种依据作为数据已经提前输入系统中。一些给当事人带来不利影响的标签可以在这种运作过程中达到去标签化,这使得具体案件的个人能够真正被当作一般人来对待,同时对某些人的优待因为缺乏相应的依据(特权显然是被排除在外的标签)而得到限制。其二,法律要求稳定性,这是“类案类判”最直接的体现。人工智能在司法中一个重要的活动就是为法官提供决策服务,即通过对数据库的比对,推送案情相似的判决,这使得法官的恣意受到限制。相似案件判决结果相差过于悬殊将使当事人在上诉的过程中获得更多的依据,法官也需要考量明显不合理判决所带来的问责风险。

人工智能参与司法另一显著特征就是它能够预设一套智能化的程序系统,在这样一个程序下,人工智能充当的角色与其说是决策的辅助者不如说是管理者和监督者。司法人工智能融入司法改革有其法理意蕴,“数字正义”是司法人工智能与司法改革相融合的法理表达^[5]。多元化的在线纠纷解决模式构建了更加简便、快捷、高效的法律程序以“接近正义”(access to justice)^[6]。信息技术的发展,加之与互联网和司法活动相结合,使司法的价值更多地体现为线上的可视化“数字正义”。司法人工智能借助新科技的技术优势,让科学技术的“法外”形态介入司法活动,充分挖掘和释放司法实践的程序正义。人工智能技术的司法应用,开启了从“接近正义”迈向“可视正义”的历史进程^[7]。

2. 司法效率维度:及时公正

波斯纳的经济分析法学给司法正义带来的启发集中在他的《正义/司法的经济学》一书中。书中特别提到了“法本身——包括它的规范、程序和制度,都在于促进效率的实现”,从而提出了“财富最大化”理论。同时,针对批评者认为波斯纳在正义和效率之间忽视了正义,他回答道:“财富最大化之所

[1]哈特:《法律的概念》,张文显、郑成良等译,中国大百科全书出版社1996年版,第155页。

[2]中共中央宣传部:《习近平新时代中国特色社会主义思想学习纲要》,人民出版社2019年版,第103页。

[3]姚莉:《司法公正要素分析》,《法学研究》2003年第5期。

[4]何家弘:《司法公正论》,《中国法学》1999年第2期。

[5]魏斌:《司法人工智能融入司法改革的难题与路径》,《现代法学》2021年第3期。

[6]参见M. Ethan Katsh, Orna Rabinovich-Einy, *Digital Justice: Technology and the Internet of Disputes*, Oxford University Press, 2017, p.40.

[7]马长山:《司法人工智能的重塑效应及其限度》,《法学研究》2020年第4期。

以是一种更可得到辩证的道德原则,还在于它为分配和校正正义提供了一个坚实的基础。”^[1]波斯纳把效率看作正义的一个层面的观点提醒我们不得不重视经济效益在司法过程中的内涵。它要求尽量减少司法资源的浪费,以有限的司法资源投入获得最大的司法效果反馈,从而达到司法资源优化的目的,人工智能在此处依然发挥着重要的作用。

首先,作为司法决策的辅助者,人工智能并不像人类那样由于精力的有限、阅读内容数量的局限而容易发生错误的判断,它只需按照既定的算法,采用高效数据计算的方式去运作。即便是人数再多、时间跨度再复杂的案件,人工智能也很难产生错误。这样法官可以将有限的精力投入到真正需要自由裁量的领域,案件的产出符合公平正义又符合社会期望的可能性大大增加。其次,有限地使用强人工智能可以使得司法资源得到最大化的合理分配。这里的强人工智能是相对于弱人工智能而言的,理论和实务界对司法人工智能的定位多为辅助性角色,毕竟技术只能是人的工具,而不能代替人。但从节约司法成本的角度来看,对于某些案件事实清晰、争议不大,适用简易程序、速裁程序以及调解结案的案件,人工智能或许可以通过适当扩大自主性,减少对此类案件的过度纠缠,这对于那些亟需处理和社会影响重大的案件来讲,反映出某种“财富最大化”的正义。

2021年《最高人民法院工作报告》显示,在信息化建设与司法改革相融合的背景下,2020年全国智慧法院建设成果充分显现。如裁判文书和在线庭审直播方面,“裁判文书上网公开1.2亿份,庭审直播1159万场”,智慧法院在线执行方面,“网络查控案件1464.5万件,网络拍卖成交金额4027亿元,同比均有大幅增长”;也有一些较难量化的成果,如“疫情期间,无接触式诉讼服务广泛应用……审判执行不停摆、公平正义不止步”,以及“通过构建中国特色、世界领先的互联网司法新模式,为全球互联网法治发展积极贡献中国方案”^[2]。在探讨基于司法效率的公正问题时,虽然我们必须正视司法成本与效益不能完全通过“数字”体现出来的现实,但也不能放弃人工智能对司法效率推动的研究探寻。

三、人工智能介入引起的司法公正风险

不可否认,人工智能介入司法可以显著提升司法信息化水平,推进司法审判能力和司法体系的现代化,对于促进司法程序公正、提高司法效率具有重要的意义。但与此同时,我们不得不重视其在价值判断中的局限性。人工智能是基于庞大的数据库、强大的数据整合分析能力、不断优化的算法推理能力等为司法裁判者提供决策参考。但现有人工智能运行中司法数据收集处理环节标准的缺失、算法法律地位的界定不明以及缺乏有效约束机制等都将弱化司法公正的技术基础;同时,人工智能作为一项智能技术,无法具备人类法官的理性思维能力,也无法准确把握法律规范背后深刻的立法价值基础,而这种价值的中立性恰恰是人工智能司法无法真正实现司法实质公正的核心因素所在。

1. 技术维度:数据算法缺陷弱化司法公正基础

任何一种科技的发展都必然伴随着一定的风险,风险防范对于司法公正的实现具有至关重要的影响。人工智能的核心部分是它的数据和算法,这两者直接决定了人工智能所提供的司法结论,必须重视这方面的建设工作。

其一,数据的数量质效不高,明确化和层次化不足。司法人工智能必须建立在足够的有效数据基础之上,精确的判断也必须以大量优质的司法数据作为基础。然而我国法律数据信息离完备还有一定差距,相应的法律数据质量也参差不齐。我国仅仅是法律规范数据建设就任重而道远,此类基础性的数据由于过于复杂,其明确性、合理性大大降低。同时,法律术语往往一词多义并需要加以解释。

[1]波斯纳:《正义/司法的经济学》,苏力译,中国政法大学出版社2002年版,第69页。

[2]周强:《最高人民法院工作报告》,2021年3月8日,<http://www.court.gov.cn/zixun-xiangqing-349601.html>。

如果在对数据进行学习时不加以明确区分,便会造成结果上的偏差。此外,由于参与诉讼的主体并不能准确表达每种术语,在进行数据处理的过程中要先把主体描述的语言转化为人工智能可以识别的语言,这就无形中加大了人工智能学习的难度,并且极易产生混淆。除了法律术语,法律规则和法律原则在效力层级上的差异加大了数据检索的困难。人工智能在这方面的判断很可能产生数据上的交叉重合,而这种数据的不当应用很有可能造成案件性质和结果的偏差与疏漏。

其二,算法“黑箱”易导致司法隐形歧视。算法作为技术的核心,科技公司并不会去主动公开算法的运行方式;与此同时,司法工作人员很少会对这方面的技术有深刻的理解,他们在判断人工智能系统合适与否时,更多地是看其是否有助于高效办案,这就使得算法“黑箱”成为一种可能。当司法工作人员习惯于享受算法带来的便捷却又缺乏对算法的理解,人对算法的控制便处于一种不可知的状态,这种状态使得人工智能带来了程序的公开却又使得程序的运行原理被另一种帷幕所笼罩。同时,数据的累计和关联性的不断强化,使不受监督的算法很可能在不断的学习过程中产生标签化的歧视,因此中立的人工智能可能因为算法的“黑箱性”变得不中立,这种歧视往往处于很隐蔽的状态。绝大多数案件也许不容易产生明显的误判,但这种潜移默化的处理方式却可能使得弱势群体的权益得不到合法的保护,这种隐形歧视往往引导社会对一部分被贴上标签的群体产生排斥,而更不利于真正地解决社会中特定群体性问题。

其三,技术缺陷易导致司法缺乏实质合理性。就目前人工智能的发展状况来看,司法领域主要使用的技术为知识图谱(mapping knowledge domain)+深度学习。所谓“知识图谱”,简单来讲就是以可视化的方式将知识资源的关联性表现出来。这种方式对数据的完备程度非常依赖,可以说如果一个案件相关的因素储备得越多,那么所得出的结论也就越符合大部分法官将做出的判断。对于案件争议焦点较少以及一些基础性的结论,人工智能确实能够很好地作出反应,并且准确性一般不会有问题。但是并不是所有案件都能够从有限的知识资源中发掘判断思路。尤其是在不同的裁判依据产生竞合的情况下,法官会依据道德、政策、社会等多方因素考量,而并非将每一项裁判依据进行完整的阐明。这种情况下人工智能算法的学习将不能准确反映现实裁判的思路。究其根本原因还是法官裁判和人工智能裁判特点的差异性:法官的裁判在依据事实情况和条文规定并在两者之间往返查看的基础上,还需要通过辩论、论证以达到说服其他听众的目的;人工智能的判断仅仅按照相关度进行,虽然从数据层面上看具有结论的合理性,但缺乏了说服的过程,而这种过程正是算法设计的困难之处,最终将导致判决的正当性、合理性遭受削弱。

2. 伦理维度:技术的中立性无法满足公民个案正义的诉求

司法活动与立法活动不同,法官在适用法律规则的同时,又对法律进行了“逐案塑造”^[1]。“法学所关心的不仅是明确性及法的安定性,同时也致力于在具体的细节上,以逐步的工作来实现‘更多的正义’。”^[2]因此,法官仅仅依照法律的“一般性规则”进行“类案类判”还不足以体现司法实体公正的全部。司法裁判中的实体公正涉及实质价值或道德考量,是与一般正义、抽象正义相对应的个案正义与具体正义。当然,个案中的正义并非法官的个人公正观,它是受到一定社会或地区主流价值观或公正观限制的,是一种“同理心正义”——一种多元的、动态的正义^[3],难以明确表述为或“编码”为一套明确的规则。在此,个案情境非常重要,它决定了不同情境下法官的同理心,从而也造就了司法公正呈现出愈发复杂的内涵。

[1] 克里斯蒂安·施塔克、赵静:《论法律的价值证立之必要性》,《苏州大学学报(社会科学版)》2022年第1期。

[2] 卡尔·拉伦茨:《法学方法论(第六版)》,黄家镇译,商务印书馆2020年版,第253页。

[3] 雷磊:《司法人工智能能否实现司法公正?》,《政法论丛》2022年第4期。

其一,司法人工智能无法避免裁判中的价值偏见风险。从表面上看,司法人工智能裁判将统一的法律规则、类似的案件事实输入智能计算系统,应当能够避免法官先入为主的偏见。但事实上,司法人工智能本身并无法真正达到我们希望的客观与中立。司法人工智能在展开司法裁判工作之前,会深入学习和挖掘先前司法裁判的数据,并以此为基础进行司法预测。在这个实践过程中,人工智能既有可能学到先前司法裁判的有效经验,也不可避免地被动学习到先前司法裁判中带有法官价值偏见的非合理经验,即“偏见进,则偏见出”^[1]。

其二,司法人工智能缺乏主观能动性,无法及时根据个案场景、情境或社会发展需求“因时制宜”“因地制宜”“因事制宜”。而蕴含于其中的司法智慧、权益考量和共情能力,正是司法活动中实质公正的重要表现。司法裁判并不仅仅是将案件事实与法律规则简单叠加,而是需要裁决者充分发挥司法裁判的溢出性效用,能动地展开适当的裁量,以契合司法改革的时代目标和司法实质公正理念。然而“无须满足任何更多的条件即可生成‘是’与‘否’的二元选项代码”^[2]的司法人工智能,并不具备展开价值判断和规则创新的能力。司法人工智能的价值中立性和被动性特征使得其应用领域被限缩在那些是非对错明晰、知识需求单一、潜在模式简易可辨的场景之中^[3]。

其三,现代司法诉讼的重要功能之一,是为民众提供一个释放诉说、寻求救济的途径,即司法的疏导功能^[4]。由于人工智能介入司法活动,原本沟通民众与司法的桥梁——法官的作用被弱化,生硬的程序算法无法解读民众的切实情感需求与非理性诉求。与人类法官裁判相比,人工智能在寻找配比法律规则方面或许更为准确快速,裁判程序与结果也合乎规则,但缺乏疏导功能的司法活动会使得民众对司法的参与感降低,从而引起民众内心对司法公平正义感受的弱化,造成司法公信力的削弱。

3. 法理维度:技术的工具性消解司法公正本意

前面讲到,人工智能引入司法的一大优势就是其作为一种智能技术,并不像人那样会面临各种情感、利益上的影响,而具有相当的客观性、中立性。这里所说的对司法中立产生的风险不是从算法运行程序本身而言的,更多地还是由上述算法黑箱、工具性局限所带来的实质结果上的公正风险。人工智能程序的开发本身不是由法院自身独立完成的。现实中,绝大多数法学教育是较少涉足自然科学和信息技术的,法学家也很少能够精通这一领域。由于法院自身缺乏人工智能技术专业人员,多数人工智能法律系统都是由专业科研技术人员研发的,技术人员自身原有的价值偏见等便有可能被写入司法人工智能算法中。

其一,司法系统与科技服务提供者之间交流存在障碍。司法人员在对人工智能算法设计提出要求时往往很难精确地表达服务需求,以目前的学科交叉程度来看,也只能通过一些宽泛的语句进行沟通。这样不充分沟通下产生的业务系统很可能背离了司法主体的初衷,或者说仅仅在效率、数据汇总等弱人工智能应用层面能够产生良好的互动。也许法院本身需要的系统是按照某种价值评判标准去运行的,而现实却可能反映不出司法过程的内涵。沟通的失误便使得人工智能系统算法设计从一开始就没有契合司法的特点,从而失去其客观中立的特点。

其二,智能系统拥有者易对司法结果做出干扰。随着信息时代的发展,科技本身成为一种专利,数据资源本身成为一种财产。基于这种特性,智能系统和数据往往都不会公开,这无形中形成一种“黑箱”,上文主要是对这种“黑箱算法”的隐蔽性所带来的未知从微观角度进行分析,这里更多地是从

[1] S. G. Mayson, "Bias in, Bias Out", *The Yale Law Journal*, 2019, 128(8), pp.2122-2473.

[2] 卢克·多梅尔:《算法时代:新经济的新引擎》,胡小锐、钟毅译,中信出版社2016年版,第220页。

[3] 王文玉:《司法人工智能的可能空间、内在限度与前景展望》,《东北大学学报(社会科学版)》2022年第3期。

[4] 强世功:《法律人的城邦》,上海三联书店2003年版,第23页。

宏观角度来探究。当今时代,拥有核心技术以及大量的数据,就有可能拥有足够大的能量去影响社会、改造各种社会关系。在司法领域,掌握着技术、数据的主体在本身陷入到纠纷时,由于其本身对算法的熟悉,在应对各种法律风险时,显示出较强的预测能力,这种优势地位的形成很难让相对方感受到真正的司法中立;更有甚者,由于算法实际上难以做到公开,那么对算法的篡改或者说有倾向性的引导可能会对司法结果产生干扰。

其三,不加限制的智能技术将会使司法逐渐工具化。我们在法治建设中强调对法的公平正义、司法独立健全的呼吁,这种呼吁可以为公民树立一种“法律信仰”,正所谓“没有信仰的法律将会退化成僵死的教条”^[1]。也正因如此,学界和实务界对司法人工智能的发展持谨慎态度,他们普遍认为不加限制的技术发展以及过于激进的应用会使得司法逐渐工具化,这种工具主义将会对司法的独立性带来冲击。因为工具必然导致信仰的弱化,而这种弱化会使得技术、数据的实际控制者将司法过程理解为简单的判决工厂,从而崇尚技术的独立价值,贬低司法的独立价值。司法机关若陷入工具主义思维,过分注重数据、算法的管理功能,则通过对数据的操控和调整便可能实质上使得法院的数据资源获取受到有意识地限制,法官的自由裁量权被过度削弱,从而弱化司法的独立性,独立性的减弱直接使得公正裁判的基础遭到威胁。

四、人工智能影响司法公正的风险防范

人工智能技术运用于司法领域势必会带来双重影响,既给司法改革带来了前所未有的发展机遇,同时也会引起司法不公的风险。基于此,必须构建人工智能影响司法公正风险的防范体系,尽可能地预防和应对人工智能介入司法引起的不良后果,推动司法改革向更深、更远的方向进行。

1. 强化司法大数据管理,规范算法决策

其一,强化司法大数据管理。大数据的共享和流通不可忽视,司法人工智能在进行学习和关联性排列的过程中亟需充足的优质数据,前文已经详细论述过。这里主要强调全国公、检、法系统的数据互通与及时更新。通过共享和数据的筛查可以最大限度地减少由数据内容的瑕疵所带来的判决纰漏。数据库就像是技术的养分,只有定期的清理、排列,才能使得其生成的内容富有实践指导意义,而首先要做的就是将法院裁判文书数据及时上传。据统计,中国裁判文书网上的裁判数据与全国法院系统数据总量相差甚远,并且除了上传文书,对文书的数据格式统一化也必须同步进行。在确保法院系统数据充实的前提下,司法系统外的其他部门也应当对数据进行甄别和引进。在这个信息爆炸的时代,数据的来源不应当过于闭塞,而应当加强合作构建共享型数据库,在此基础之上对数据的研究与学习将更有参考价值。

其二,数据资源的安全性也必须作为法律的重点领域加以保护。我国近年的立法不断强调对个人信息的保护,虽然数据的共享和流通是必要的,但是这不能被理解为数据可以随意流转给他人。数据一旦被不法分子或者具有优势地位者利用,司法过程的双方实质力量将会产生巨大差距,因此需要进一步明确数据使用和保存的规范。一方面,数据的使用不可随意。另一方面,要从技术上采用更强力的数据保护举措,防止数据流失所带来的隐患。同时,应提前进行遭受数据攻击的演练,对破坏数据的行为要进行行政法乃至刑法上的严厉处罚,并对数据管理主体建立问责制度。

其三,规范算法决策。首先,需要建立事前评价和事后检查的机制。在将算法应用于司法实践之前,应严格规范数据收集、信息挖掘等一系列程序,以保证算法规则的合理性和合法性。相关专家评

[1]伯尔曼:《法律与宗教》,梁治平译,商务印书馆2012年版,第59页。

审应在仔细研究算法规则后进行公示,使社会各界详细了解算法决策的运作机制,增加算法规则的透明度与合理性,维护司法公信力。在做出算法决策后,还应该建立相应的算法评审和测试机制,保证算法决策机制的高度合理性和关键过程的透明性。算法应用的每一个环节都需要相关领域专家的不断检验和论证,使其尽可能不受无意识的歧视和个人偏见的干扰,保障司法的公正价值。

最后,法律赋予公民监督司法活动的权利,以促进司法程序的公正。但是,在算法“黑箱”的缺陷背景下,算法的隐藏性在一定程度上侵犯了公民的知情权,更勿论对司法活动的监督权。因此,需要从制度上确立与完善算法的解释规则,防止人工智能融入司法实践而带来的不确定性风险。建立算法解释规则的目的是实现决策主体与社会公众的有效沟通,这种沟通可以概括为以下两点:一是实然层面的解释,即算法规则是什么,全面说明人工智能司法决策的规则,消除其中的歧义和冲突。二是应然层面的解释,即规则为什么,为适用规则本身提供充分的理由。

2. 推进“有限智能化”,制定司法人工智能的伦理规范

人工智能技术作为现代科学技术发展的前沿方向之一,强化其于司法领域之中的应用能够大力推动司法智能化改革。但针对人工智能在司法活动中的角色,可采取“有限智能化”态度推进^[1]。具体来说,首先应肯定司法工作人员的主体地位,司法人员应当始终掌握系统运行的整个流程,不应一味追求司法人工智能系统的高度自洽性。根据司法人员和人工智能的能力,结合两类角色在司法活动中的不同特性来分配权限和责任。当智能化系统的预测结果可能与人类法官的决策不相同时,人类法官应享有最终决定权。同时,司法工作人员应当监督人工智能的司法活动。正如法国学者克罗齐耶所言,相比于有关国家理论的建构,人际关系的发展更有利于获取成功,同程序相比,人更加值得信赖依靠^[2]。

前文已述,人工智能介入司法会产生影响司法公正价值的特有伦理问题,需要建立以伦理规范为主的监督机制。司法人工智能本质上是数据和算法的司法活动,而数据和算法又是人工智能研发者的科技成果。因此对人工智能介入司法的伦理规范本质上仍然是对人的规制。算法除了“黑箱”和隐形的“歧视”等自身的问题,还可能包含人工智能研发者的主观恶意。算法的监督除了前述对技术自身的事先事后规制外,还要确立人工智能研发者的伦理行动规范。

对人工智能研发者的伦理行动规范,应当包括尊重基本人权、安全性、可追溯性、包容性、预防原则、隐私权保护、效益最大化和危害最小化等^[3]。人工智能的伦理规范不仅以尊重和保障人类基本权利为价值目标,还要规范参与人工智能各环节工作人员的行为,如研究、设计、生产和使用过程中的人员的行为。同时,需要建立必要的监管机制,使其在可控、低风险的环境中发展。司法人工智能要继续坚持技术中立、价值无涉原则,恪守技术伦理,结合司法改革任务的实际情况,创造人工智能在司法领域的技术应用规则,划定人工智能技术的应用边界和限度,在享受司法大数据红利的同时,保障个人信息安全,维护司法公正。

3. 将法律论证嵌入司法人工智能系统,降低工具主义风险

论证是法律人必备的素养,也是法律人实现结论为人信服所必要的过程。司法过程的法律论证可以被理解为法官通过说理来使其听众确信某种法律评价或者某个案件的合理性。论证真正需要发挥作用的地方往往也是人工智能处理薄弱的环节,即复杂疑难案件中对争议集中点做出裁决的合理性重构。法律论证的内容主要包括论证所要满足的标准、法官对规则解释的必要性以及法律规则、法律原则以及一般道德公理之间的关系。而实现最终论证结果的合理性则需要一套对判决结果的标准

[1]陈灵峰:《司法人工智能的技术效应与应用边界》,《求索》2021年第6期。

[2]米歇尔·克罗齐耶:《法令不能改变社会》,张月译,格致出版社2008年版,第96页。

[3]雷磊:《中国特色社会主义智慧法治建设论纲》,《中共中央党校(国家行政学院)学报》2020年第1期。

或规范,这种规范并不是当下仅依靠“是”或“否”和进行标签化处理的人工智能所能够完全承担的,因此法律论证的任务就在于提供合理性依据并将工具主义倾向遏制住。

其一,最大化尊重“听众”的能动性,维护个案实质公正。在法律论证中有一种修辞学方法,它是与逻辑方法相对应的一种方式,其属于对内部论证方法寻找具有说服力的规则的外部论证方式之一。这种方式注重论述内容的可接受性,突出论证的修辞特征和策略安排,而可接受性则取决于论证的说服力,具有代表性的观点是佩雷尔曼的新修辞学理论^[1]。他将论证与正义论结合起来为我们讨论司法人工智能如何促进正义提供了思路。按照佩雷尔曼的观点,形式逻辑并不能代表正当性,最终案件的价值判断要归结到对话式的讨论,人工智能在作为论证助手的过程中承担着定位论证主体的功能。主体可被分为三类“听众”,即普泛听众(讨论中除自己以外所有人)、实际谈话中的特定听众、思辨过程的听众。论证辅助系统在程序设计中可确定一种时间流程,给予每类听众发表自身观点的时间,保障听众的意见充分录入智能系统之中,最后通过总结听众共识以达成最终价值评判的结果。这使得人的能动性得到最大化尊重,在面临人工智能难以评判的实质公正内容时,通过听众角色的设置可以提升裁判内容的说服力。

其二,规范参与者对话规则,维护司法程序公正。人工智能辅助论证的过程不能仅仅停留在确定时间维度上的对话流程,而必须对对话内容采取一定程度的规制,以保障对话的效率以及结果的正确性。首先,当论证主体提出自己的主张后,需要输入证明内容。缺乏证明支持的主张,在算法程序中应当被评定为低合理性。同时也允许相对方主体的批评和质疑,将这一系列对话内容综合记录于辅助程序中,为运算提供数据支持。其次,约束法律商谈结果的规则。这一规则对解决司法人工智能局限性的启发在于,司法数据的庞杂使得智能系统对各种规范的优先层级较难判断,而此种程序性规则要求证立法律判决必须先从一个总体性规则出发,这个规则在论证辅助系统中占最高优先级。例如在民事案件论辩过程中,如果有制定法、特殊法的规定,就不应当越过这些规定而直接以民法原则和一般法作为论证大前提(作为论证的合理性的支持观点是可以的),否则就意味着对法律规则的违背,那么这样的论辩是不能被评价为合理有效的。总而言之,上述论证方法在智能系统算法中的设立,实质上都是为了保证司法过程中人的存在,当人真正主导了司法过程的合法性、合理性,所谓工具主义对司法主体物化的风险将被控制到最低;同时,数据优势者及算法独裁所带来的负面影响将会在追求公平正义的论辩中得到消解。

五、结语

司法人工智能在目前的学界和实务界的定位倾向于司法辅助角色,显然,综合目前科技发展水平、数据质量结构、法治建设水平等因素,这种看法是恰如其分的,过于激进地对待技术将会给司法公正以及诸多传统价值伦理带来威胁。不过这并不意味着面临技术冲击需要抑制甚至停止它的引进和发展,毕竟在当今信息化和智能化时代,科技生产力越来越成为改革的重点。从经验论来看,智能技术总体而言对人类社会的进步起到积极的促进作用,现实中我国司法系统也一直注重法院的智能化建设。因此在以一种适当的节奏发展司法人工智能的过程中,法律人要保持理性审慎的态度,使目光往返于技术和公正之间。人工智能是方式而不是目的,我们有必要把握这样一个基本的原则,那就是公正永远是司法的目的和归宿,任何智能化技术的发展都应当以此为风向标。

[责任编辑:吴玲]

[1]参见海姆·佩雷尔曼:《新修辞学》,杨贝译,商务印书馆2021年版,第25页。