



论 AI 法官审判的制度冲突及其调和

莫 皓

摘要:随着大数据、人工智能技术在司法活动中的运用,理论界开始探讨 AI 法官审判的可能性。然而,阻碍 AI 法官进行审判除了技术层面的因素外,更重要的原因可能是其与现行司法制度的理念存在冲突。具体而言,AI 法官审判的逻辑与当前的审级制度、开庭审理制度、证据制度以及调解制度均存在冲突。这其实源于 AI 法官审判在“植入”司法体系时,与司法的多元功能、制度特征、价值取向以及法院角色定位等内生系统并不兼容。AI 法官审判难以与现有法院制度体系相适配,其未来发展的基本思路应该是推动制度与技术的双向调和,并采取根据案件类型与运用场景进行多维度分层式的运用路径。

关键词:AI 法官;智能裁判;算法决策;司法制度

DOI: 10.13734/j.cnki.1000-5315.2024.0320

收稿日期:2024-08-12

基金项目:本文系中央高校基本科研业务费专项基金项目“总体国家安全观视阈下数据跨境流动法治体系”(3300224643)、四川省社科重点研究基地纠纷解决与司法改革研究中心项目(2023DJKTb10)的阶段性成果。

作者简介:莫皓,男,四川宜宾人,法学博士,西南民族大学法学院讲师,E-mail: mohao1220@163.com。

在大数据、人工智能技术快速发展的背景下,语音识别、自然语言处理等一系列已在其他领域大显身手的智能技术开始在司法活动中得到运用,并且随着技术的发展迭代呈现出替代性趋势。如果说在萌芽期的弱人工智能阶段,技术只能在司法活动中承担一些外围、简单的辅助工作,如类案检索与推送、量刑辅助与预警等,那么,在成熟期的强人工智能阶段,技术或许并不甘于配角地位,由 AI 法官而非人类法官坐堂判案可能走向现实。

自大数据、人工智能技术运用于司法活动以来,理论界便展开了激烈的探讨,并取得了基本共识:当前的人工智能成果尚不能取代法律人进行独立实践,在司法活动中,仍然需要人类担任法官、检察官、律师等角色。这主要缘于两方面因素的制约。一是基于技术发展程度的局限。想要替代人类在司法活动中的角色,当前的技术基础尚有局限,因为人工智能裁判是一个长期且极为耗费资源的工程,需要更多既懂法律又懂人工智能的复合型人才加入,也需要能够提供 AI 法官深度学习所需的海量数据与资料^①。二是基于司法伦理的局限。诉讼当事人可能难以接受由机器而非人类对自身利益作出裁判,尤其是在作为裁判依据的算法并不公开透明的情况下。当然,这种司法伦理问题也部分缘自技术因素,例如算法黑箱、算法可验证性等问题。因此,当下的关注点主要集中在 AI 法官审判的技术障碍上,认为阻碍 AI 法官审判的主要因素是技术发展的滞后。笔者认为,这种论断存在一叶障目、过于强调技术难度而忽视了 AI 法官审判的制度障碍。这并非否定技术发展的重要性,而是相较于技术障碍,制度障碍或许更难以得到解决。首先,技术发展的迭代性高而制度的稳定性强。大数据、人工智能的技术发展日新月异,从兴起到全面发展也不过短短数十年,并且随着技术的成熟,其发展迭代的速率将成倍提升。相较而言,制度的发展变化较为缓慢,并且不是线性上

^①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,第 8 页。

升而是迂回发展,往往存在“发展—往复—再发展”的反复过程,尤其是带有稳定性乃至保守性特征的司法制度,其变革更非易事。其次,技术成效具有易检验性而制度试错成本高。技术发展的成效如何,往往一目了然,易于检验,而制度变革的有效性检验则相对缓慢,且试错成本较高,非经长期的论证、试点不会轻易迈出脚步。最后,技术原理具有共通性而制度一般不易被复制。世界各国的法律制度各不相同,法律移植虽然并不鲜见,但成功的法律移植也并非易事,法律制度深嵌于不同的政治经济与社会制度当中,法律制度的单一移植成活率较低。而技术移植则不同,即便各国存在技术发展能力的差别,但因技术没有国别障碍,所以,世界各国的先进技术都可以被吸收运用。

可见,阻碍 AI 法官审判走向实践的根本因素是制度,即技术的快速更新与制度的缓慢改进之间还不匹配。本文立足于分析 AI 法官审判与具体的司法制度之间存在的冲突,并试图提出调和的建议。

一 AI 法官审判的制度冲突

(一) AI 法官审判导致审级制度的功能落空

审级制度,是指对案件依法经过几级法院审理便应当终结、法院所作出的裁判即发生法律效力之制度设计。我国确立了两审终审的审级制度,其中二审制度被认为是纠正一审判决偏误以彰显司法正义的充分体现。正如我国台湾地区学者林钰雄所指出的,“审级制度之目的,在于求取审判结果之正确与适当”,“藉由审级制度之运作,上级(诉)审得以撤销、纠正下级审法院违法或不当之审判,减少下级审法院擅断、误判的机会,并进而保障被告之诉讼权利”^①。在司法功能上,审级制度存在纠正一审错误判决与救济当事人权利的两大功能,尤其在刑事审判中,由于控辩双方诉讼能力与诉讼地位的差距较大,更需要审级制度来保障诉讼的公正性。除此之外,“现代刑事审级制度尽管在纠正司法错误、监督司法行为方面起着较大作用,但该功能只是审级制度构建原理中的一个方面,甚至是其他制度功能(如维系权力自身合法化)的副产品”^②。审级制度实现了职能划分,旨在维系各级别审判职能之间的有机分工,一般而言,级别越高的法院越重视法律适用的统一性。同时,审级制度实现了科层管理机制,上诉机制和审级制度乃是一种保障司法统一、贯彻司法政策的集权化监督机制。比较法学家夏皮罗就曾指出:“上诉本质上是中央对地方司法官员行使中央集权化的监督权力的一种方法。”^③这种管理性功能正是通过将审理程序进行分离和拉长,实现司法活动的秩序性和权威性。

AI 法官审判则导致上述审级制度的多种功能落空。其一,导致司法纠错与救济功能丧失。AI 法官作出的判决是基于算法模型,这种判决能运用在司法实践中的前提条件是算法具有科学性,且结论具有唯一性与可检验性。然而这种算法结论的唯一性可能成为一柄双刃剑,原因在于当前由人类法官进行裁决的审理制度并不绝对趋向结论的唯一性,甚至二审制度的存在本身就代表着对于一审结论的可推翻性。但算法模型却不同,不同算法之间在技术层面无高下之分,原则上在一定范围内各级人民法院适用的算法应该具有同一性,并不存在适用于一审法院/基层法院的初级算法和适用于二审法院/高层级法院的高级算法之分,事实上,当前的技术研发也不存在按法院层级进行划分的研发路径。试想,如果上下级法院使用的算法模型之间存在技术先进差异,则下级法院所用的算法会彻底失去可信度,二审法院的纠错与救济功能也会失去意义,因为此时一审判决的不可信将导致二审变为实质上的“一审程序”。如若上下级法院之间使用的是同一种算法,输入相同事实得出的结论也将会完全一致,此时二审结果必然与一审重合,纠错与救济便无从谈起。可见,算法的技术优势反而使得审级制度在审判层面的纠错功能丧失。其二,导致管理性功能不复存在。一方面,从职能分工角度看,运用审级制度实现一审二审各有侧重的职能划分,在某种意义上是审判资源有限而裁判需求高速增长背景下的应对举措,司法资源的有限性归根结底是源于审判人员能力与精力的有限性,所以需要调整优化资源配置。而 AI 法官审判在此方面的优势尽显无遗,可以源源不断产出裁判结果的精力优势与不断学习迭代提升的算力优势,使其无须错位配置便可实现同一机器既可查明案件事实也可兼顾法

① 林钰雄《刑事诉讼法(下册 各论编)》,中国人民大学出版社 2005 年版,第 216 页。

② 林喜芬《我国刑事审级制度功能考辩与变迁改良》,《东方法学》2009 年第 5 期,第 52 页。

③ 马丁·夏皮罗《法院:比较法上和政治学上的分析》,张生、李彤译,中国政法大学出版社 2005 年版,第 55 页。

律适用争议。另一方面,从科层管理角度看,审级制度是一种保障司法统一、贯彻司法政策的集权化监督机制,主要是为了防止下级法院擅用自由裁量权进行恣意裁判,在这点上,不同层级法院的 AI 算法具有一致性,不存在 AI 擅权的可能性;同时,算法是由研发者预先设置的,在设置之初相应的司法政策等信息均已纳入算法模型,必然会遵循相应的政策理念;并且,从深层次看,科层制的出现主要源于人的“不可靠性”,是为了实现对人的管理而创造出来的管理机制,既然作为管理对象的“人”已不存在,科层制的生存土壤自然也不复存在。

(二) AI 法官审判导致开庭审理制度的虚化

开庭审理的意义之一在于规范审判信息获取渠道。庭审过程中的信息流动模式是:原告方提供信息、被告方提供信息→法官接收信息(法官提出信息需求后原被告双方再反馈)→法官对信息进行综合认定并进行自由心证→法官得出裁判结果。可以发现,传统开庭审理信息流动呈现出多方性特征,信息的主要提供者是原被告双方,作为信息主要接收者的法官,也会择机提出相应需求推动双方再次提供审理信息。这种信息提供具有即时性,双方在同一时空轮替进行,保证了信息提供机会的对等性与相对真实性。

AI 法官审判的出现使得庭审中的信息流通机制前置化、书面化,后续的开庭审理程序似乎无必要进行。具体而言,AI 法官审理的信息流动模式看上去仍是“原被告双方提供信息→AI 接收信息→AI 进行综合判断并输出裁判结果”,但差异在于,这种信息的流动可以不再具有同步性与筛选性,即原被告双方可以异步提供相应的审理信息,并且这种信息缺少开庭程序对于关键信息(案件争议焦点)的筛选。正如有研究所总结的:“传统司法场景中法官和当事人的沟通是动态的时间轴,是过程性的;而机器的交互是法官和机器间发生的静态的点,是结果性的。”^①可以发现,整体而言,AI 法官审理的信息交换方式较为粗疏,传统更为精细化的信息筛选机制主要是为了聚焦争议焦点、提升庭审效率,而 AI 的高算力看似可以保障在繁多信息中也能实现有用信息的快速筛选,然而,这种信息流动机制的转变,可能带有隐藏的“信息歧视”,不加限制的信息提供机制,使得信息提供主体的搜集、分析与供给信息能力更为重要。尤其是在刑事诉讼中,控方具有天然优势,证据搜集能力本就较弱的辩方在新形势下更难以与能调动国家资源的控方相提并论,这将加剧控辩地位的不平等,所以,AI 法官审判在带来信息处理优势的同时也会附带隐性歧视。

就我国司法实践而言,以刑事诉讼为例,刑事二审程序长期存在书面审、虚置化的现象,而 AI 法官审判的出现,可能使这种现象持续乃至进一步恶化,毕竟虚置化的前提也是存在实体的开庭场景,而 AI 审理连虚置的“过场”都不用经历。

(三) AI 法官审判导致证据制度与证据规则不适配

证据是审判活动的核心要素,无论是由人类法官进行审判,还是由 AI 法官进行审判,都无法脱离对证据的需求,但算法技术的特性可能使得当前的证据制度与规则发生重大变化。

AI 法官审判使证据制度存在法定证据主义风险。关于证据制度奉行的基本原则,历史上存在法定证据主义与自由心证主义之分。法定证据主义是指法律预先规定各种证据的证明力和意义,法官只能机械地按照法律规定的条件来衡量证据的可靠程度,并作出判断结论,而不能按照对证据的认识来分析和判断。自由心证主义是指证据的取舍及其证明力由法官根据自己的理性和良心作自由判断,形成确信并依此认定案情。由此可见,法定证据主义与自由心证主义的重要区别在于法律是否事先规定各种证据的证明力,而现代证据法基本上不规定证明力的审查判断。然而,人工智能技术在证据中的运用使得当前自由心证可能遭受法定证据主义的侵袭,具体表现在下述三个方面。首先是证明规则的预先赋予。法定证据制度的典型特征是预先指定每个证据的证据能力,法官只能据此作出判决,无法对单个证据的证明力进行自主判断。而 AI 法官的运行方式就是由研发者预先设定相应的算法,使其能够抓取同类案件相应的情节,并在此基础上进行证据运用的判断。其次是证明自由度的丧失。在法定证据制度中,“只要法官把起诉方提交的证据加在一起可以构成一个完整的证明,他就必须做出有罪判决;如果不能构成一个完整的证明,他就必须做出无罪判决。无

^①周翔《司法人工智能对裁判说理的辅助价值和实现路径》,《法学杂志》2024 年第 1 期,第 115 页。

论个人内心对具体证据的确信程度如何,法官在审判中必须严格遵守上述规则”^①。这种格式化的证明方式,恰好与 AI 法官区别于人类法官自由心证的证明程序具有相似性。最后是证明过程、证据标准带有数字化运算特征。AI 法官通过数据分析和算法得出结论的过程,虽然可以提供大量有用的信息,但也可能忽略某些无法量化或数字化的因素。例如,在某些案件中,证人的证言、物证的鉴定等,可能无法完全用数字方式来衡量其真实性和可信度。此外,数字化运算的结论可能存在误差或偏见,尤其在数据模型本身存在缺陷或数据不完整时,AI 法官的分析结果可能会出现偏差。而关于证明标准,有学者指出:“‘数据化证据标准’成为了智慧司法的基础性环节。证据标准的数据化,是指将证据标准转化为由数字化符号逻辑表示的计算机可以运算、推理的证据标准体系,从而实现个案要素信息的结构化。”^②这种数据化的证明标准带有明显的客观化色彩,域外“无论是‘内心确信’还是‘排除合理怀疑’,都是着眼于认识主体的主观思维。因此,这种证明标准可以称作一种主观标准,适用这类标准所产生的认识也不免带有一定的主观性”^③。我国目前建立起的是主客观相统一的证明标准,其中主观标准为内心确信和排除任何合理怀疑,具体表现为司法人员在排除任何人为和非人为干扰的情况下感到对事实认定确有把握。有学者针对上海 206 系统适用的研究指出,证据标准指引建立在证据规则(包括国家层面的证据规则和地方性证据规则)的基础上,由于证据规则受法定证据主义影响,因此,证据标准化的推进不可避免地会加剧法定证据主义的倾向。“同时,过分精细化的证据标准也会导致证明标准的客观化,实质性地侵蚀自由心证主义”^④。

AI 法官审判在证据规则方面导致了证人出庭规则的架空。证据发展的历史为证人构建了诸多证据规则,典型的如交叉询问规则,此类复杂的证人出庭规则旨在通过要求证人亲临庭审现场,展现证人在面对控辩双方询问时的眼神、语气、下意识肢体动作等,帮助审判人员判断证人证言的真实可信度。若将人类法官替换为 AI 法官,证人出庭规则将无用武之地。一方面,证人的证言可以通过书面形式庭前提交,导致证据形式的书面化,但书面证言在审判适用中存在“扭曲作证者本意”,“妨碍了诉讼对方的质辩”,“排除了法官直接辨识证言真意及其可靠性的可能”,“可信性保障条件不足”等弊端^⑤;另一方面,评判证人出庭时的神态表现是法官依据人类共有的生理表现特征而进行逻辑判断的依据,例如紧张时会脸红、回答磕磕绊绊等,但其并非一种硬性指标,需要结合现场情况综合判断。申言之,证人回答不流畅并不意味着一定存在伪证情形,也可能是因为记忆模糊导致的,因此,AI 法官无法模仿人类法官进行此种真伪判断,即便将一些典型特征提前设置为伪证信号,也可能因为过于机械而无法察觉出真实情况。

(四) AI 法官审判缺乏调解制度需要的博弈与共情能力

作为解纷方式的一种,调解在审判活动中广泛存在并发挥着重要作用。在我国司法制度中,调解制度具有悠久的历史传统与实践生命力。法院调解兴起于 20 世纪 40 年代的陕甘宁边区,确立了“就地审判、不拘形式、联系群众、解决问题”为特点的马锡五审判方式,以此为基础的法院调解作为民事诉讼法的基本原则确定了下来。2010 年最高人民法院印发的《关于进一步贯彻“调解优先、调判结合”工作原则的若干意见》中,提出了“调解优先、调判结合”的调解工作原则,显示出调解在我国民事审判工作中的突出地位。尤其是在贯彻落实习近平总书记关于“要坚持把非诉讼纠纷解决机制挺在前面”的重要指示精神的大背景下,调解更是法院在审判活动中需要着重考量的因素。

一般而言,成功的调解需要具备两方面因素。一是调解人的共情能力。与判决这样的强制性处理机制不同,调解更类似于软性处理机制,调解结果的达成依赖于参与方的协商。在法院调解中,法官的身份转换为不具有强制力的第三方调解人,如何在产生冲突的双方当事人之间斡旋,其中一项重要的基础在于与当事人的共情能力。在诸多调解成功的案例中,尤其是涉及婚姻家庭事、个人名誉等与自然人个人利益高度相关的纠纷,当事人除了有寻求物质利益的诉求,还隐含着内心倾诉、情感共鸣等精神性需求。因而,法官在进行调

① 何家弘《对法定证据制度的再认识与证据采信标准的规范化》,《中国法学》2005 年第 3 期,第 145 页。

② 张凌寒《智慧司法中技术依赖的隐忧及应对》,《法制与社会发展》2022 年第 4 期,第 182 页。

③ 龙宗智《我国刑事诉讼的证明标准》,《法学研究》1996 年第 6 期,第 121 页。

④ 熊秋红《人工智能在刑事证明中的应用》,《当代法学》2020 年第 3 期,第 87 页。

⑤ 龙宗智《论书面证言及其运用》,《中国法学》2008 年第 4 期,第 129 页。

解时,实质上扮演了精神抚慰者的角色,与当事人共情是让其能够接受调解方案的前提条件。二是调解的博弈性。法庭调解本质上是一种博弈行为,其本身没有“唯一解”,而是通过双方协商达成“共同认可”。“在作为中立第三方的法官(或调解员)主持下,双方当事人就矛盾纠纷进行协商、互让互惠,最大效度地解决有争议的问题,实现各自权益的最佳分配”^①。可见,调解是通过话语博弈的方式最终实现权利分配。这期间,需要当事人与法官各方通过语言沟通的方式进行,因其本质是当事人权利的让渡,需要极为高超的调解技巧,法官往往会采用面对面同时调解与一对一单独调解相结合,以及对裁判结果作预测分析等方式对当事人进行劝导。这需要法官具备丰富的语言学知识、心理学知识,否则,单纯以宣扬立法的普法式调解往往难以如愿。

共情与博弈是人类所擅长的两种能力,其基础是人类的情感与利益取舍,难以进行数字量化,也没有固定的标准与模式,所以难以由机器习得。因此,在进行调解方面,人类法官有着 AI 法官无法替代的优势。

二 影响 AI 法官审判在我国适用的内生因素

AI 法官审判与司法制度的冲突归根结底是其在“植入”司法体系时产生的“排异反应”,而这与司法的功能、司法的历史传统、司法的价值取向及法院的角色定位等因素密切相关。

(一) AI 法官审判影响判决在司法中的多元功能

一般认为,判决是整个司法程序中最为重要的一环,判决使得诉讼两造的争议有统一定论,最终起到定分止争的效果。除了终结诉讼争议的功能外,判决实际上还具有其他功能,包括如下两个方面。

一方面,通过判决融入社会主义核心价值观。有研究指出,“核心价值观融入司法裁判的途径共有三种:作为裁判理由、作为裁判依据和作为价值冲突的解决基础”^②。2021年2月,最高人民法院印发《关于深入推进社会主义核心价值观融入裁判文书释法说理的指导意见》,并先后发布三批“弘扬社会主义核心价值观典型案例”,这表明,法院判决在培育和践行社会主义核心价值观方面起着引领、规范和保障作用。据统计,2022年,全国法院在法律框架内运用社会主义核心价值观释法说理的一审民事案件达38.25万件,比2021年增长21.66%^③。在当前裁判文书公开普及化的背景下,裁判文书所承载的评价指引功能进一步凸显。一份好的裁判文书往往寓理于法,充分体现社会主义核心价值观的精神和要求,以法服人,以情感人,传递社会正能量,进而获得最大化的社会认同。

另一方面,通过判决实现教化功能。中国司法传统中有一种“教化”意识,或者说“教化”是中国古代司法的一项重要功能。司法实践因而伴生了“宣教”功能,通过调解、审理、判决等方式来宣传儒家的道德观念,并使当事人和旁观者受到教育感化,此亦即“寓教于审”或“寓教于判”。古代法官的“书判”往往是一些将道德训诫与刑罚威胁相结合的判词,充分发挥了司法的教化功能。部分地方法院在裁判文书中间或者后面增写“法官寄语”、“法官后语”,有研究便认为,“法官寄语作为裁判文书说理的有效延展,是对法官调解工作的高度凝练,这种书面化了的道德感化语言,不仅能够填补裁判文书规范化的不足,还能充分发挥其亲情感化功能,拉近法律和当事人之间的距离”^④。

然而,这种解决纠纷之外的教化功能却非 AI 法官所长,其系统设置的核心目标在于通过事实判断解决争议,算法难以理解纠纷解决之外所蕴含的教育意义和价值,即便算法开发者尝试在 AI 法官的判决中加入社会认同的价值观念,但这种由机器所进行的道德倡导与教化能对当事人产生的影响可能也较为有限。

(二) AI 法官审判弱化了司法活动的仪式性特征

伯尔曼认为:“法律像宗教一样起源于公开仪式,这种仪式一旦终止,法律便丧失其生命力。”^⑤司法仪式为庭审规划出一个剧场化的空间,人们在剧场之内活动时,必须遵守已经预设的制度、规范和程序。“庭审活

① 柯贤兵、廖美珍《法庭调解话语博弈交际研究》,《外语学刊》2011年第5期,第70页。

② 雷磊《社会主义核心价值观融入司法裁判的方法论反思》,《法学研究》2023年第1期,第10页。

③ 孙航《明规则 破难题 扬正气 树新风——最高法民一庭相关负责人就第三批人民法院大力弘扬社会主义核心价值观典型案例答记者问》,《人民法院报》2023年3月2日,第3版。

④ 李新亮《“法官寄语”体现“刚柔并济”》,《人民法院报》2011年11月6日,第2版。

⑤ 伯尔曼《法律与宗教》,梁治平译,中国政法大学出版社2012年版,第25页。

动也是由法官、检察官、律师和当事人等参与角色表演的活动”^①。由此可见,司法的仪式性特征通常具有以下特定的因素:固定的场景、固定的诉讼仪式和固定的话语。然而,AI法官的出现使得这些特定因素开始发生变化。

首先,诉讼场景云端化。有针对远程审判的研究指出,诉讼云端化影响了当事人听审请求权的行使,“传统民事诉讼在审判方式上强调当事人的程序参与和直接言辞,其在传统民事诉讼中均以当事人在审判现场听审为前提。这就是司法裁判的在场性”^②。能够在场庭审,意味着“当事者更有可能接受裁判结论,裁判者也更容易取得当事者的信任。而有了这种信任,裁判结论连同其制作过程本身,就容易让人心服口服”^③。相对于远程审判的影响,在AI法官的审判场景中,线下的庭审空间已不再是开展实质庭审的场所,裁判是由算法在后台推导得出的,即便诉讼当事人与AI法官同处一室,亦与听取双方陈述和辩论的传统法庭相去甚远。

其次,诉讼仪式简易化。诉讼仪式不仅是法律程序的一部分,也是展示程序正义和权威性的重要方式。在法庭上,司法人员身着专门袍服,诉讼参与者按照规定的程序和顺序进行陈述、举证和辩论,这是法庭审理规范的流程和仪式。随着AI法官等现代化技术的应用,这种诉讼仪式正在经历逐渐简化的过程。例如,传统的庭审需要经历开庭、法庭调查、法庭辩论、最后陈述等环节,每个环节都有一定的时间限制和规范,这些规范和仪式可以营造出一种庄重、严肃的氛围,让参与者和旁听者感受到司法的权威性和公正性。而AI法官通过语音识别、自然语言处理等技术,快速分析证据和法律文书,这些环节极易被简化甚至省略,虽然减少了庭审中看似“烦琐”的发问、辩论等程序,但附随的结果可能是弱化了传统诉讼仪式的庄重和严肃,而这些流程,可能寄托着对于公平正义的追求和信仰。“即使人工智能做出的司法裁判也是公平公正的,但因缺少当事人与人类法官沟通的时间和环节,可能导致公众难以切身感受到公平正义”^④。

再次,诉讼语言通俗化。“自然语言和机器语言之间存在着天然的隔阂,自然语言往往依赖于谈话语境、上下文关系以及生活经验的积累与理解,而机器语言则主要通过词向量转换、分词技术等抓取语言要素中的核心语义以理解文本的含义”^⑤。诉讼语言是自然语言中具有特定内涵的、用于司法活动的专门语言类型,具有准确、严谨、规范的特点,能够确保司法活动的专业性。然而,随着AI法官等现代化技术的应用,司法活动中诉讼语言的重要性在逐渐被稀释。具体而言,自然语言与法律语言之间存在着差异,难以完全匹配,为了能让AI法官充分理解案件内容,往往需要将法言法语转化为通俗语言进行表述。例如“公序良俗”,可能存在着“有伤风化”、“不道德”等多种生活化语言表述,人类法官可以准确识别理解这种语言并且进行相应转化,但当前广泛采用的机器图谱式的人工智能无法准确理解,只能通过“喂养”大量与“公序良俗”相关的场景、语词进行语料库补充,这可能导致法律术语的准确性和严谨性受到影响,甚至造成法律概念的模糊与混淆。

(三)AI法官审判与现代司法程序正义的价值理念不符

所谓程序正义,是指法律程序在具体运作过程中所要实现的价值目标。程序正义是一种过程价值,它主要体现在程序的运作过程中,是评鉴程序本身正义性的价值标准。“一项法律程序本身是否具有程序正义所要求的品质,要看它是否使那些受程序结果影响的人受到了应得的待遇,而不是看它能否产生好的结果”^⑥。谷口安平曾言:“满足正当程序要件的程序才是合乎程序正义的程序,反过来说合乎程序正义的程序就是正当程序。”^⑦在司法活动中,程序正义的要件主要表现在以下几个方面。

首先,审判程序的公开透明,即法庭审判过程向公众开放,确保其了解案件的审理过程和结果。而算法

①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》,《政法论坛(中国政法大学学报)》1999年第3期,第16页。

②段厚省《远程审判的双重张力》,《东方法学》2019年第4期,第105页。

③陈瑞华《看得见的正义》,北京大学出版社2013年第2版,第17页。

④张凌寒《数字正义的时代挑战与司法保障》,《湖北大学学报(哲学社会科学版)》2023年第3期,第136页。

⑤王文玉《司法人工智能:实现裁判正义的新路径》,《大连理工大学学报(社会科学版)》2022年第6期,第104页。

⑥陈瑞华《刑事审判原理》(第三版),法律出版社2020年版,第72页。

⑦谷口安平《程序的正义与诉讼》,王亚新、刘荣军译,中国政法大学出版社1996年版,第4页。

自动化决策带有的黑箱特征使得诉讼程序的运行过程不透明,当事人和公众无法对算法运行进行质询和抗辩,更遑论针对算法模型设计可能的缺陷进行监督与救济。不仅如此,算法开发者凭借专利保护、商业秘密保护权利等“合理抗辩理由”,使得算法决策获得了“算法霸权”。

其次,当事人程序参与权与辩论权的保障。正当程序强调参与性和主体间性,注重程序参与者的主体性地位以及主体间的互动性。这要求程序参与者进行面对面的对话和辩论,当事者以自身的语言和感官为媒介,即时地陈述意见、进行质疑和表达主张,可见,充分的法庭辩论往往是程序正义的构成性条件。但是算法自动化决策流程是瞬时杂糅完成的,不仅压缩了正当程序的自由裁量空间,也限制了告知、说明理由、听证等的实现。并且,人工智能的结果导向式决策可能无法充分地捕捉到法庭辩论的实践性,例如有学者认为,“大数据和人工智能将会变成法庭辩论的‘断头台’,酿成‘一切取决于既定的软件,面对面的对话式论证算不了什么’的氛围”^①。在这种情境中,当事人的程序参与权与辩论权在实质上被消解。

再次,客观中立的裁判权,即裁判者应当保持中立,不受任何不当因素的影响,以无私的态度作出公正的裁决。人们似乎更愿意相信没有情感羁绊的、以特定算法为思考逻辑的人工智能在中立性上远超人类,并且愿意相信算法自动化决策能更好地满足程序的中立性和客观性要求。实际上,对算法自动化决策过程和结果的绝对信任是一种对科学逻辑的误解。在算法系统的设计过程中不可避免地会嵌入某些伦理假设与价值植入,而偏见可能存在于设计者对伦理与价值的理解中,并以隐而不显的方式影响着决策结果,所以,算法并非绝对中立客观的存在。

(四) AI 法官审判难以适应国家建构视野下法院治理的角色定位

党的十八届三中全会提出“推进国家治理体系和治理能力现代化”,开启了从国家建构视野来认知法院职能定位的新视角。国家建构是指在现代化过程中以民族国家为中心的制度与文化整合措施、活动及过程,其基本目标是建立一个合理的、能对社会与全体民众进行有效动员与监督管理的政府或政权体系。“一直以来,中国司法都是实现国家建构的一种重要方式,承担着国家权力下沉,重塑国家权威的政治功能。因此,中国司法具有强烈的‘嵌入性司法’的色彩”^②。法院在国家治理体系中除定分止争之外,还扮演着更多治理性角色,例如促进地方经济发展、推动法律的普及等。总的来说,其角色定位从被动的司法售货机转向进行司法治理的主体。

首先,通过审判活动的治理。这是指法院以司法判决为核心,在审判活动中运用创造性的思维推动法律的发展与变革,从而通过判决的引领达到推动社会管理方式的创新与发展。在英美法系国家,因为有法官造法的传统,审判人员在审判活动中运用创新性思维解决纠纷,并确立新的判例。例如,美国联邦最高法院在1966年米兰达判决中,第一次宣告了刑事被告人在被侦查讯问时拥有律师在场权、沉默权等多项救济性权利。米兰达判决不仅影响到美国联邦与各州的刑事司法,亦漂洋过海影响到其他国家侦讯法制的变革,形成了著名的“米兰达规则”。正如卡多佐所言:“实际上,每个法官都在他的能力限度内进行立法。无疑,对法官来说,这些限度都比较局促。他只是在空白处立法,他填补着法律中的空缺地带。”^③虽然我国并非判例法国家,但仍存在具有创制新的裁判规则并被纳入指导性案例/典型案例的情形,在某一领域内形成参考性价值。

其次,通过执行公共政策的治理。在我国,法院执行公共政策的典型表现有最高人民法院制定司法解释、司法文件等。例如在非法集资活动猖獗、官方打击收效甚微的背景下,出台《最高人民法院关于审理非法集资刑事案件具体应用法律若干问题的解释》,“不但有助于法院审判工作,实际上也有助于民间融资活动的合法开展”^④。有学者认为,最高人民法院已经是一个制定公共政策的法院。“最高人民法院的影响不仅仅局限在各级法院的审判过程,而是已经溢出诉讼过程和法院系统,对社会经济生活产生越来越大的影响”^⑤。在美国,司法机关介入公共政策执行的趋势更加明显,美国法院参与政治决策,在多个方面深入参与政治生

^①季卫东《人工智能时代的司法权之变》,《东方法学》2018年第1期,第132页。

^②郑智航《国家建构视角下的中国司法——以国家能力为核心》,《法律科学(西北政法大学学报)》2018年第1期,第28页。

^③本杰明·卡多佐《司法过程的性质》,苏力译,商务印书馆1997年版,第63页。

^④彭冰《非法集资行为的界定——评最高人民法院关于非法集资的司法解释》,《法学家》2011年第6期,第38页。

^⑤侯猛《司法的运作过程:基于对最高人民法院的观察》,中国法律出版社2021年版,第220页。

活,法官在重大公共事务中成为积极和持续的决策者。“有学者认为,美国所有悬而未决的政治问题最后都会转变成司法问题”^①。

然而,AI法官与这种法院参与治理的原则相违背,AI法官只适合被动坐堂判案的方式,但人类法官的裁判思维可以不局限于机械的法条主义,可以从社会、经济维度考虑,将此种考量和法律原则、司法政策、相关法律规范等有机融合,从而在更高层面满足自治性和安定性。在这种情况下,法官对于个案的处理已经超越了案情、法律本身,更敏锐地捕捉到经济社会条件的变化,并在合理的范围内,通过各种司法技艺调节经济基础和上层建筑之间的不适,而人工智能对这种复杂的场景变换是难以适应的。

三 制度障碍之调和

(一)基本认知:AI法官难以与现有法院制度体系相适应

基于上述分析,笔者认为,导致AI法官审判暂时无法成为现实的更重要原因在于司法制度层面。简而言之,现存司法制度是按照以“人”为管理目标而构建的,而AI法官的运用,会颠覆现有法院系统的运行与管理体系,从而造成变革的代价超过了现有制度的承载力而不得而被放弃。

具体而言,现有的法院制度建构实际上深嵌于整个司法体系,而AI法官审判与现有的制度几乎没有适配性。一方面,从司法运行机制角度看,AI法官审判打破了传统的通过面对面获取审判信息作出裁断的运行体制。法院成为一种“线上服务”,这种“线上法院”的模式需要建构起与当前的立案、分案、审判、执行流程完全不同的运行机制。另一方面;从司法管理体制层面看,AI法官审判需要调整现有的法院架构。例如在审判管理方面,是否还有必要设立院长、庭长等职位;如果是以保证裁判的准确性为目的,如何证明其能力高于AI法官;若是以行政管理为目的,则会面临作为管理对象的司法人员的缺失。再如法院内部部门的设置,是否还有必要划分各类审判庭,设置各种专业化审判组织,原则上运用算法的机器人可以胜任任何类型的案件,无须像人类一般术业有专攻。总之,要完成上述这些变动,需要考虑的因素早已超越了审判的范畴,并且这种改革需要付出的代价可能高于其预期的收益。

(二)基本思路:制度与技术的双向调和

制度与技术的双向调和是指司法制度与智能技术在实现自身需要的基础上向彼此的转变与倾斜。当前制度与技术冲突时的解决路径往往呈现出一种单向调和思路,并更多地表现为由制度构建向技术发展革新的调和。例如信息技术的发展催生出线上开庭方式,这种全新的开庭方式引发了法庭审理制度的重大转变,打破了法庭空间场景里开展诉讼活动的历史传统,也导致了诸多相应线下诉讼制度的随之变化,异步审理机制便是典型。当然,也存在极少数由技术向诉讼制度的单向调和,例如语音识别转换的通用技术被移植于庭审场景中,衍生出庭审语音识别系统,并在通用技术的基础上,采用法律词汇数据库更新、法言法语训练等方式,以针对诉讼活动的特点进行适配。双向调和则是站在更加宏观的层面上,更加强调这种转变的同一指向与适配。

首先,树立制度应随技术革新的观念。在传统的观念中,司法制度被视为一种相对稳定和静态的体系,其变革往往被视为不必要或是具有较大的制度代价。然而,这种观念或许本身存在一定误区。作为上层建筑的政治制度,本身便会因为经济基础的变化而处于变动的过程中,而司法制度属于政治制度的重要组成部分,自然会随之革新。有学者对二者关系进行过例证,由科技进步和经济增长带来了法律制度变迁,主要表现为“过去不需要的制度现在需要了(比如专利法和金融法)”,“过去标准较低的制度现在提高了标准了(比如安全标准和刑事司法证明标准)”^②。需要指出的是,司法制度的变革其实是存在现实需求的。一方面,通过变革可以解决现有制度中存在的问题和缺陷,提高司法的效率和公正性。例如有技术对于司法监管制度变革的研究提出,“在过去,纸质卷宗案牍的司法办案方式,因其卷宗调取和使用对司法组织内部成员具有相对公开化特征,这为一些组织内部成员泄露他人案件办理信息提供了现实条件”,“全程留痕的电子信息系统能

^①邱静《美国司法政治化现象探析》,《中国海洋大学学报(社会科学版)》2024年第3期,第90页。

^②桑本谦《法律简史:人类制度文明的深层逻辑》,生活·读书·新知三联书店2022年版,第453页。

够协助司法人员抵制外部权力违规插手、干预司法案件的正常办理,有利于预防司法权力的腐败”^①。另一方面,通过变革可以适应新的社会环境和社会需求,推动社会的进步和发展。例如,随着互联网和社交媒体的发展,人们对于信息公开和透明度的需求越来越高,审判直播与裁判文书公开等司法制度的变革便是适应这种新的社会需求的表现。因此,我们应当转变观念,从司法制度的静态思维向承认司法制度也可以随着技术革新而发展的动态思维转变。

其次,技术移植、技术开发与法律制度形成预先规划。基于技术与制度双向调和的基本立场,二者之间在结合之前有意识提前谋划,有利于技术的发展与法律的进步能够相互促进,而非相互制约。一方面,在技术移植与开发时预先考虑制度的适应性。这要求技术开发者深入理解法律制度的背景与内涵,在技术移植时首先针对技术需求旺盛、程序要求不敏感的法律制度进行试点。例如,对于庭审语言系统的开发便是适宜的尝试,即便效果不尽如人意,也不会对现有法律制度根基、诉讼权利保障产生影响。另一方面,立法者应合理规划法律制度的适应空间。在技术移植和开发的过程中,如果法律制度能够及时地适应技术的进步,将有助于降低技术的开发成本和风险,提高技术开发的成功率和经济社会效益。这要求法律制定者也需要关注技术的发展趋势,了解技术对社会和经济的影响,以便在法律制度的改革和完善中充分考虑这些因素,通过在立法中加入宏观性条款,可根据情势发展再进行细化规范。

再次,从体制机制层面认知技术局限性与改革渐进性。人工智能技术在司法中的应用,虽然近年来得到了快速发展,但在当下仍然面临着许多技术上的局限和挑战。例如,人工智能在处理具有高度规则性的问题上具有优势,而对于需要进行情理推断和价值判断的问题,它并不擅长。又如,人工智能在司法中的应用还涉及数据安全和隐私保护等附随性问题,这需要进一步完善相关的技术和法律制度。同时,司法制度的变革本身也是一个渐进性的过程,难以一蹴而就。技术的革新对司法制度的影响不是简单的“替代”或者“颠覆”,而是一种渐进的“渗透”和“融合”。司法制度的变革需要考虑到社会背景、法律体系、文化传统等多方面的因素,而这些因素的变化都是需要时间和过程的。因此,我们需要对技术的发展保持开放和包容的态度,也需要对司法制度的变革保持耐心和平常心。这意味着在面对司法实践中的难题时,不能仅仅依赖于技术的解决,而需要从体制机制层面宏观考虑,综合多方因素,寻求全面、系统、可持续的解决方案。

(三)基本方式:分层式应用路径

在上述基本思路的基础上,AI法官审判的具体运用应该确立分层式路径,可考虑采取案件类型与运用场景的多维度分层。

首先,针对不同的案件,应有选择地使用自动化决策系统。“道德推理和价值判断属于感知和高阶认知的混合智能”^②,现有的技术水平无法保障人工智能具备人类所具有的伦理性思维。因此,一方面,对于无事实争议以及情感因素较少的简单案件,例如小额借贷纠纷案件以及其他简单的民事案件,人工智能可以快速、准确地完成裁判任务,提高效率。因此,在这类案件中,应重点推动 AI 法官的介入并在案件裁判中逐步发挥更加重要的作用。当然,需要更进一步思考的或许是如何筛选、明确这类案件,可行的方式是先从部分案件试行,然后逐步扩展。对于民事案件,可以采用以标的金额大小为主、以权利义务关系明晰程度为辅的方式进行划分。因为适用标的金额大小作为识别案件简单与否是一种客观且较为容易的方式,且我国立法已有相应的基础,小额诉讼一审终审等制度的确立已经为 AI 法官审理此类案件打下了基础。事实清楚、权利义务关系明确这条标准因为带有一定的主观判断前提,通常由法官自由裁量,但可以将是是否存在争议、被告人是否选择答辩等情形作为判断事实是否清楚的前置标准之一,从而将主观判断客观化。对于刑事案件可以采取以刑期为标准的方式进行划分,首先从刑期一年以内且被告人认罪认罚的案件中进行,尤其是在危险驾驶等重点案件中试行。由于当前我国刑事治理的轻罪化趋势,实际上 AI 法官审判介入危险驾驶等轻罪案件已经能够在很大程度上缓解刑事案件堆积的处理难度。对于特定类型案件例如婚姻家庭类纠纷,或者存在重大争议的疑难案件,以及可能判处三年以上有期徒刑的刑事案件(可以综合是否认罪认罚、是否采

^①左卫民、潘鑫《通过技术规制司法:进步与挑战》,《法学评论》2023年第4期,第59页。

^②魏斌《智慧司法的法理反思与应对》,《政治与法律》2021年第8期,第123页。

用普通程序等情形判断),则不宜采用 AI 法官审判。正如有学者所总结的,“对于涉及道德判断和价值衡量的案件、存在重大争议的疑难案件以及具有极高证明标准和极大试错代价的刑事案件,则应始终坚持以法官为中心的审判原则,谨慎使用司法自动化决策系统”^①。

其次,应该构建起针对不同场景的适用方式,使得 AI 法官审判可以在各个环节获得支持,如证据搜集与分析、法律文书撰写、庭审辅助等。当前已有智能系统可以辅助生成不同类型的法律文书,但生成的文书质量良莠不齐,无法帮助法官解决审判中的痛点难点。笔者认为,在此方面,可以采用人机交互模式,AI 法官可以在事实认定和法律适用角度发挥重点功能,而需要融入价值观念或者情理说明的部分,仍由人类法官进行,“法官说理有助于补充智能裁判系统在事实要素与法律构成要件涵摄过程的欠缺,并通过价值判断,实现个案正义并保障法官主体地位”^②,以此达到分工合作、提高效率的目的。

再次,在证据运用层面,AI 法官也可以发挥重要作用。在传统的证据关联性分析中,法官需要花费大量时间和精力去理解证据内容,并手动查找和整理证据之间的联系。然而,利用算法可以大大简化这一过程。通过自然语言处理技术可以对证据文本进行预处理,包括文本清洗和实体识别等操作,将文本转换为计算机可理解的数据格式。然后,利用相似度算法,基于文本之间的语义和结构相似性来计算相似度,从而识别出证据之间的关联性。在司法实践中,已经开始运用自然语言处理技术对证据文本进行处理。例如,智能系统可以自动提取案件涉及的关键词,根据文本内容自动分类,甚至自动抽取证据中的关键信息等。接下来,通过可视化技术,将证据之间的关联性以图形化的方式呈现出来,法官可以更直观地理解证据之间的关联性,从而更准确地判断证据之间的相关性,提高审理效率。此外,智能系统还可以自动对大量证据进行分析并“去粗取精”,“从当事人提供的关于事实的描述材料中‘剪辑’出对裁判而言有价值的、值得保留的事实信息,比如刑事方面应当提取出定罪的事实和量刑的事实,民事方面应当提取出违约/侵权的事实和结果程度的事实”^③。

四 结语

与大多数将以 AI 法官审判为代表的司法人工智能运用替代人类为时尚早归结于技术因素的观点不同,笔者认为,制度的不适配或许才是更为深层却容易被忽视的症结。但是,技术层面的差距犹可弥补,制度层面的鸿沟却往往难以逾越。原因在于制度的不适配可能是发现了症结,却由于制度变革“牵一发而动全身”的特点容易被搁浅。因此,对于 AI 法官审判的适用路径,在提升智能技术的基础上,注重制度的协同转变,才是更加全面、可行的道路。

[责任编辑:苏雪梅]

①郭春镇、勇琪《算法的程序正义》,《中国政法大学学报》2023年第1期,第175页。

②陈子君《智能裁判系统的法律推理逻辑》,《四川师范大学学报(社会科学版)》2024年第2期,第76页。

③周翔《司法人工智能对裁判说理的辅助价值和实现路径》,《法学杂志》2024年第1期,第124页。