
司法人工智能的技术效应与应用边界

陈灵峰¹

(湖南大学, 湖南 长沙 410082)

【摘要】: 当前, 人工智能主要应用于信息的收集与转换、裁判决策辅助和审判监督管理三类司法场景之中。司法人工智能具有提高司法效率、提升审判质量、规范办案流程之功效。司法人工智能的技术效应存在限度, 对内, 存有法律内在实质正义实现的隐忧; 对外, 存在高质量数据收集的壁垒。究其原因, 从法理维度来看, 司法理性与技术理性之间的矛盾难获弥合; 从实践维度来看, 司法工作人员主体性的导向与智能技术的应用难寻平衡; 从技术维度来看, 机器的算法黑箱障碍难以攻克。基于此, 有必要设计优化司法人工智能的应用边界, 在基本立场层面, 尊重司法价值功效, 司法逻辑主导算法逻辑; 在路径选择层面, 智能辅助人类, 强调司法工作人员的主体地位; 在技术保障层面, 以算法和数据为着力点, 改良司法人工智能整体生态。

【关键词】: 人工智能 技术效应 司法裁判 司法理性

综观既有司法人工智能的理论研究, 不难发现, 其大多着墨于我国司法人工智能的进程发展、法理基础、伦理价值、路径难题、辅助审判等方向, 而较少聚焦于司法人工智能的技术效应界限与应用边界的主题。本文将以前人工智能应用于司法领域的实践状况为基准, 阐明司法人工智能的技术效应存有限度, 从法理、实践、技术三个维度考察司法人工智能技术效应有限性之成因, 即司法实践无限复杂性与算法技术有限单一性之间存在的鸿沟, 进而为司法人工智能的应用边界提供设计优化, 以期实现人工智能与司法的高效协同合作。

一、善假于物：司法人工智能的发展现状

近年来, 伴随智慧司法改革的深入展开, 我国司法人工智能的应用热潮此起彼伏, 大有赶超域外之势¹。以各地法院推广的智能化辅助办案系统为代表, 如上海的刑事案件智能辅助办案系统(206 系统), 浙江的审务云司法大数据平台, 江苏的同案不同判预警系统, 河北的智审 1.0 系统, 贵州的智能模拟判决系统², 这些司法智能系统展示了人工智能技术于我国当前司法实践中应用的最新样态。以下将从类型归纳、价值效用与技术手段三个层面剖析我国司法人工智能的现状。

(一) 类型归纳解构

目前, 人工智能技术多应用于三类司法场景之中: 信息的收集与转换、裁判决策辅助、审判监督管理。

从信息收集与转换的角度来看, 人工智能技术将诉讼信息进行收集、汇总并转换为电子数据的形式。以电子卷宗随案生成系统为例, 纸质卷宗在 OCR 自动识别、信息项智能回填等技术的辅助下得以转换成电子文档以方便各方查阅。又如, 以自然语言处理技术(NLP)为核心的庭审语音识别系统能够区分诉讼各方参与主体的角色, 自动提取各自陈述内容进而转换形成对应的文字内

¹作者简介: 陈灵峰, 男, 湖南大学法学院博士研究生。

²基金项目: 国家社科基金项目“基于民事二审判决书的法官裁判影响因素实证研究”(项目编号: 17CFX064); 湖南省研究生科研创新项目“违约金酌减预测研究”(项目编号: CX201813173)

容³。

从裁判决策辅助的角度来看，人工智能技术以实现精准推送类案、预测实体裁判，进而协助审判决策为目标，依托海量的裁判数据信息，对其加以提取、分析，从中进行分析解构，探寻案件之间的潜在关联。以河北法院的智审 1.0 系统与贵州法院的量刑偏离度分析系统为例，前者能够自动推送法条、案例，比对案例并一键生成文书，后者可以将目标裁判同既有裁判进行比对，以分析目标裁判的偏离度。

从审判监督管理的角度来看，人工智能技术从实体与程序两个方面实现对审判活动的监督管理。以江苏南京法院的同案不同判预警系统与河北法院的审判风险防控系统为例，前者具有预测裁判结果、计算裁判结果偏离度并预警之功能，后者能够将案件整个流程拆分为百余个风险点并予以自动监测，同时在此系统的帮助之下，河北全省超审限案件大幅减少，由 8000 余件降至 200 余件。

(二) 价值效用解构

综合考察人工智能技术应用于司法领域的价值功效，其主要发挥三个层次的作用：司法效率的提高、审判质量的提升、办案流程的规范。

从司法效率提高的维度来看，人工智能技术具备模式化、自主化的特点，凭借此种优势能够很大程度上缩短立案、审理、裁判等诉讼流程的时间，减轻司法工作人员非诉讼核心事务的负担，减少案件审议、决策时间，降低司法资源的消耗，进而最大限度地提高司法效率⁴。以河北省为例，其智审 1.0 系统先后协助司法工作人员办案 76 万余件，生成裁判文书超过 510 万份，为司法工作人员减少近 30% 的案头工作量。

从审判质量提升的维度来看，作为保证司法审判质量功能的代替物乃我国人工智能应用于司法领域的内在定位。现阶段，人工智能技术对于审判质量的提升主要表现为三个方面：一是辅助司法工作人员审查证据、理解适用法律；二是通过架构判决偏离度预警机制，实现裁判标准的统一；三是设置案件裁量基准，指导法官自由裁量权的合理行使。以上海法院的 206 系统为例，其在投入使用后效果明显，1 个月内提供证据指引超过 2600 次，发现瑕疵证据 48 个。

从办案流程规范的维度来看，庭审自动巡查系统、同案不同判预警系统、案件智能管理系统为人工智能技术规范司法工作人员的办案流程提供具体技术支持。除此之外，相较于过往的“人监管人”，司法人工智能对于司法工作人员工作的监管日趋自动化、流程化、智能化。以浙江法院的庭审智能巡查平台为例，其每年巡查案件庭审超过 20 万件，对于司法工作人员工作行为的规范大有裨益⁵。

(三) 技术手段解构

人工智能技术应用于司法领域在技术层面上表现出两个显著特点：以数据为基础，以算法为核心。

在数据的基础性层面，质言之，人工智能技术于司法领域的应用乃是基于海量司法数据的挖掘而进行，考察各地既有司法人工智能系统可知，现阶段人工智能技术与司法数据关联性紧密，且前者对后者的依赖程度较高，甚至可以毫不夸张地说，司法数据乃司法人工智能的“生命之源”。以裁判结果预测与偏离度预警系统为例，其功用之发挥建立在收集、分析大量司法数据的前提之下，进而找寻相似案件情境之中的判决思路与内在逻辑，以预测目标案件的裁判结果。需特别指出的是，受目前技术水平所限，不同地区法院之间数据的互通较为不畅，各个司法智能系统之间多为隔断状态。

在算法的核心性层面，所谓算法，乃是指计算机用以解决问题所遵循的明确策略指令，当前司法人工智能系统一般通过概

括、分析、归纳过往司法决策规律，提炼司法决策模型的方式，以预测司法裁判，这表明算法在司法人工智能之中居于核心位置。换言之，算法在很大程度上决定了司法人工智能的效用，其为司法人工智能的运行提供指令。以类案推荐系统为代表，相似度算法乃决定案件推送精准度之关键，倘若嵌入算法难以实现目标案件情节与数据库中既有案件情节的精确比对，相应智能系统则无法发挥功用。需要指明的是，在目前的司法实践之中，考虑到商业秘密的保护与算法本身的复杂性，对于算法的运行原理，司法智能系统开发者很少予以说明，而司法机关也较少关注算法本身，对司法智能系统背后的运行机理几无涉猎。

二、适用有度：司法人工智能的技术效应存在限度

毋庸置疑，伴随人工智能技术的日渐成熟，其于司法领域之中的应用逐步深入，给智慧司法改革创造了前所未有的机遇，并拥有广阔的发展前景。但是，值得注意的是，在享受司法人工智能便利的同时，也应当理性认识其所致的技术效应有限，警惕技术至上主义乃至“神化”技术，切勿将司法简化为数据与算法。简言之，过度适用并夸大司法人工智能的技术效应，不仅有悖于司法的内在价值，更会对司法实践产生消极影响。

（一）“内忧”：法律内在实质正义实现的隐忧

法律的核心目标在于通过个案裁判实现正义⁶，正义亦是中国特色社会主义法治的精义与灵魂之一。在法治视角下，正义体现的乃是一种恰如其分的责任分配，与行为人的个体行为相对应，依照表现形式与实现目的的标准，正义被分为形式正义（程序正义）与实质正义⁷。尽管形式正义在法治国家建设中扮演着外在程序保障的角色，但仅就当事人而言，实质正义的实现方为其终极诉求。除法律制度的有效设计外，人工智能技术亦可借助算法程序以削减司法工作人员的失误与恣意，进而确保和增强实质正义的实现。前文已述，在大数据、云计算等技术的支撑下，现有的司法人工智能系统具有发现瑕疵证据、搜寻推送类案、提供裁判参考建议、裁判结果偏离预警、自动生成法律文书等功能，这种现代化技术所具备的客观优点对于司法工作人员弱性的克服效果明显。

但是，正如世界上没有完全相同的两片树叶，每一个个案都有自身的个性色彩。个案正义、实质正义与温度正义乃当事人之追求，而人工智能技术可提供的却是机械且缺乏情感的正义。换句话说，人工智能依据既有的司法案例，诉诸形式逻辑或推理法则，以作出判断评价，而所谓实质正义，则是司法工作人员以案件客观事实和所呈现的证据为依据，综合案中情节，运用法律规则、经验法则和业务技巧，在良知、敬畏之心与公平价值的指引下，针对个案反复思索、多方权衡并精心钻研而推导出的结果。即便不同案件表现出相似的客观事实，行为人的主观心理状态同样也值得司法工作人员考量，这就彰显了司法工作人员的人文精神，远非人工智能所能及，同时因人工智能主观感情缺失，故其更是难以洞悉司法裁判背后蕴藏的修复被破坏社会秩序与恢复被践踏社会正义之价值⁸。除此之外，因为司法裁判让人信服的内容乃是其所指向的生动事实与鲜活生活，并非法律规则本身，所以，司法裁判孕于司法工作人员日常生活经验和智慧的累积之中，深受社会具体生活环境的影响，究其根本，司法裁判唯有与社会现实生活相适应，才能最大程度上获取民众的认可。诚如美国联邦最高法院大法官卡多佐所言，对民众的需求作出回应是司法的职能之一，同时民众的需求也促使司法智能不断丰富⁹。虽然实际上，诸如证据收集、裁判参考、类案推送、文书生成等算法被嵌入至司法人工智能系统，但其并不关涉社会现实环境。相较于司法人工智能所作出的格式统一、千篇一律的裁判文书，当事人更为接受的是蕴含深刻思考、高超法律技艺与浓厚人文关怀的法理释明¹⁰。概而言之，考求个案的特殊性、差异性与鲜明性即实质正义本身所指，只有司法工作人员的“人脑”裁量智慧才能助其一臂之力，或许人工智能的智力水平高于人类，但其终究无法掌握法律内在实质正义的真谛。

（二）“外患”：高质量司法数据收集的壁垒

近年来司法数据井喷式的增长为司法人工智能的蓬勃发展注入了源源不竭的力量源泉，如前所述，司法人工智能的应用以算法为内核，但其不论被嵌入何种算法，均得依托于大量司法数据尤其是大量优质司法数据以实现功效。故而，充分且全面的司法数据收集乃是最大限度发挥司法人工智能作用的必要条件。但我国现存的司法数据尚难以满足司法人工智能高速增长的需要。

第一，既存司法数据体量不够，难以适应智慧司法快速发展的需求。

现阶段，“中国裁判文书网”上公布的裁判文书构成了我国司法数据的主要来源，并以此作为我国司法人工智能开发的基础要素。但据有关学者统计，线上公开的裁判文书仅占我国全部司法裁判数量的一半左右，另外 50% 的文书信息难以以为民众所知悉。并且，从时间维度来看，2013 年之前的生效裁判文书网上公开的数量有限，以 1996—2000 年为代表，仅能从“中国裁判文书网”上检索出 96 份刑事案件裁判文书。故此可见，由于时间变量检验的匮乏，仅依据既有裁判文书，不同历史背景下相对稳定的司法经验难以被有效识别、总结。此外，考察公开裁判文书记载的内容可知，合议庭或审委会的讨论、合议以至决策等映射司法裁判过程的信息并不展现于公开的裁判文书之内，其仅记录简略案件事实与判决结果。总之，无论就时间还是内容而言，公开的司法裁判文书为司法人工智能系统提供的司法数据皆比较有限，如此势必会消解司法人工智能发挥积极功用的根基，进而有碍其初衷——为司法决策提供精确的智能化辅助之实现。

第二，现有司法数据的质量难以与司法人工智能的快速发展相匹配。

司法人工智能的不断精进既要求司法数据的量大，更要求司法数据的质高，假使不能保证司法数据的质量，纵然拥有海量的司法数据，亦无法推动司法人工智能深层次的进步。针对当前公开司法数据的准确性问题，有学者指出，民众所接受的司法信息可能存在“表象性”现象，即公开的司法数据有部分可能是已经过人为制作的标准化信息，以证成其所载司法决策的正确性，距离真正影响司法决策作出的“实质信息”尚有差距¹¹。同时，现有公开的司法裁判还存有诸如案号重复、信息冗杂等影响司法数据准确性的情形。此外，当前公开司法数据结构化水平参差不齐的问题，多表现为部分裁判文书中同类事项表达不一、法律术语使用粗糙，例如，“自首”“主动投案”“如实供述罪行”等多种表达均表示被告人具有自首情节这一意思，但基于当前司法人工智能的技术水平，其准别识别该种非结构化的信息表达存在困难，因此不利于司法规律的提炼。一言以蔽之，囿于司法数据数量与质量的固有缺陷，司法人工智能精确归纳司法经验的难度大为增加，以致于精准长效且普遍适用的司法人工智能机制尚难以被架构。

三、成因探究：司法实践的无限复杂性与算法技术的有限单一性

现代科学技术在推动社会进步的同时，亦会带来多重风险与挑战，甚至有成为最大风险源之势¹²，人工智能技术作为现代科学技术的一种亦不例外，其虽然给司法智慧化改革提供了源泉动力，但也不可避免地使司法领域面临潜在风险。诚如前述，司法人工智能的技术效应存有限度，并非绵延不绝，一味夸大更是会加剧风险。究其原因，乃是算法技术的有限单一性与司法实践的无限复杂性之间存在张力，以下将从法理、实践、技术三个维度分别予以详述。

（一）法理维度：司法理性与技术理性之间的矛盾难获弥合

技术算力与人的理性往往具有目标一致性，二者交融互通构成技术理性¹³。所谓技术理性，通常表现为一种追求效率、效益提升的人类智慧与能力，相对忽视人文关怀与价值情怀，规范性、高效性、功用性乃其典型特征。作为技术理性的样板范例，人工智能技术基于已有司法判例，以探索司法工作人员的逻辑思维与决策过程为路径，进而作用于司法领域。但是，司法理性的内在要义在于，包括司法裁判在内的司法决策的作出并非单纯的理性思维过程，其更多乃是依靠司法工作人员的认识、心性、德行并结合案件发生的客观现实环境，同时以法律规则为准绳而适用，在此过程中，司法工作人员的技艺能力能够很大程度上捏合舒张法理与情理、法理与个案正义之间的冲突¹⁴。换言之，司法理性以追求法律效果与社会效果的有机统一为目标，努力实现法律说理性、个案特殊性与司法工作人员主观合理性之巧妙结合。而前文已述，人工智能乃是机械模仿人类智性以为行为，故其在确保个案于法有据的同时又兼顾人性情理几无可能，并且，司法人工智能的标准化与程式化还可能会使条文背后的法律光辉与人性精神被遮蔽，不利于法治精神的彰显。

此外，针对社会不公，为民众提供一个释放诉说、寻求救济的途径乃蕴含于现代司法诉讼之中的重要效能之一，即疏导功能

¹⁵。民众与司法沟通的桥梁纽带可能会由于司法人工智能而被影响，因为机器程序无法解读民众的情感需求与非理性诉求。同人工裁判相比，尽管于法律条文的找寻适用层面，人工智能或许比司法工作人员更为准确，判决结果亦合乎规则，但受困于机器的形而上学与自身参与度的降低，民众内心未必能够接受人工智能所作出的司法裁判，以致于其可能难以切身感受公平正义，造成司法公信力的削弱。

(二)实践维度：司法工作人员主体性的导向与智能技术的应用难寻平衡

司法工作人员的工作任务乃以语言表达为媒介处理社会关系中的人际纠纷，此一性征决定其无法完全转化成程序而交由计算机运作。同冰冷的机器相比，人类拥有丰沛情感，能够有机联结各种错综复杂的社会关系，同时面对不同情境可以自由转换为相应主体角色。司法过程的真实奥义绝不是教条的三段论演绎所能展现，司法裁判更非仅靠程序公式便能推导，本质而言，司法过程可谓“法—理—情”多维图谱的构建，意涵丰盈：科学的法律规则适用、累积的司法实践经验、充沛的社会道德情感。总而言之，作为现代科学技术精进的代表性产物，人工智能技术即便可以在提升司法效率、合理配置司法资源方面发挥作用，但也不能本末倒置，因此保证司法工作人员在司法工作中的主体地位实为必须。

就案件审理过程而言，其乃是司法工作人员理解重构案件事实的过程。在这一过程中，司法工作人员得先后历经筛选证据、认定事实之步骤，同时，面对各方主体提供的信息，司法工作人员基于自身实践经验与专业知识进行理解、组织、再解释，进而决定采用何种方式将这些信息提交给数据处理系统，辅之以逻辑演绎，最后形成判决。通常而言，最佳判决由司法工作人员架构的“最佳案情理解模式”与“最佳证据识别模式”通力合作，共同完成¹⁶，而这有赖于司法工作人员精湛法律专业技艺的发挥。但是，伴随司法人工智能的大幅应用，司法工作人员仅需将案件信息输入至相应智能系统，便可得出处理意见，在司法系统“案多人少”的背景之下，司法工作人员可能难有动力对机器所作裁判进行反思，长此以往，司法人工智能可能将有抑制司法工作人员思维拓展、业务能力施展，并滋生其惰性之虞。

除此之外，司法权的特征也决定了司法工作人员的不可或缺性，其要求司法工作人员高度职业化，兼具法学素养与法律知识，并恪守法律职业人员伦理准则，此非司法人工智能所能替代¹⁷。与此同时，相较于错案责任追究制，司法责任制的问世丰富了司法工作人员司法责任承担的内蕴，其强调司法工作人员应审慎独立行使职权，并为之负责¹⁸。实践中，当裁判因司法人工智能提供的信息有误而出错时，司法责任如何分配成为焦点，据此，与传统人工裁判模式不同，司法工作人员可能以错误系信息提供者所致进行辩解，拒绝担责。考虑到司法人工智能还处于初级阶段，司法工作人员在司法决策过程中仍占主导，故而前述抗辩理由尚不成立。但需注意的是，随着人工智能于司法领域应用程度的加深，“机器责任”的话题将无法回避，其对司法责任制度的冲击已近在咫尺。

(三)技术维度：机器的算法黑箱障碍难以攻克

现阶段，有关人工智能的理论研究大致可以分为五类派系：符号学派、联结学派、进化学派、贝叶斯学派、类推学派。作为主流方向，符号学派以演绎推理为内在逻辑，而经验概括总结则是联结学派的范式方法。具言之，在智能预测裁判的方向上，符号学派是将法律条文转化为对应的规则程序，并据此执行演绎作出预测，而依托海量司法数据进行归纳加工以预测则是联结学派的作法¹⁹。如前所述，司法人工智能的运行机理本质上乃是发现已有数据规律并写为算法，希望其依法律人的思维模式进行思考，以期使算法程序的纯理性克服人工偏差。但事实上，实践中的司法人工智能可能会与其希冀背道而驰，就对司法不公正的影响而言，相较于人为偏见，算法偏见更加深刻，与专家鉴定意见相似，支配司法判决生成的算法作为“专门知识”同样超出了司法工作人员的认知范围，并拥有权力背书。作为“专门知识”的一种，算法只要对司法决策产生影响便获取了与之相对应的符号权力，由此演化成“知识—权力”的二元关系，继而可能会致使算法偏见如明火执仗般地映射于司法裁判之中。

此外，决策过程透明度较低乃司法人工智能的另一缺陷²⁰。考虑到模型程式的高度专业性，以及数据收集、汇总、加工、处理的流程复杂，纵使程序开发者本人也难以明晰依托海量数据迭代算法背后的运作过程，只知其行为结果²¹，职是之故，司法人

工智能的运行展示出繁杂性、封闭性、模糊性等表征。而与之相异的是，现代司法要求公开透明，除裁判结果外，证据审查、案件事实认定、法律条文释明、证据链条形成等关联裁判决策作出的信息均应公开。考察司法实践中出现的“难办案件”²²，多是因部分司法工作人员于裁判过程之中说理不明所致，因此难以消除民众内心的困惑，给予民众满意的答复。相较于形式上不偏不倚的“公正判决”，逻辑通畅、内容翔实的说理更能为民众展示法治全貌，加之法律的有限性，部分司法判决结果可能难以让各方当事人所满意，此时愈加彰显裁判说理的重要意义，这样才能从根本上平息诉讼，使当事人服从判决。而被嵌入司法人工智能之中的算法显然无法胜任详细阐明说理的角色职能。如此，司法人工智能决策过程的隐秘性与司法的公开透明原则形成悖论，人工智能技术的算法黑箱犹如一把达摩克利斯之剑悬于司法决策上空。

四、综合进路：司法人工智能应用边界的设计优化

前文已经提及，人工智能技术应用于司法领域宛若一柄双刃剑，给智慧司法改革创造机遇的同时，也诱发了诸多现实风险，司法人工智能的技术效应存有限度。然而，伴随科学技术进步的日新月异，不论各界对司法人工智能的态度如何，司法智慧化转型已是大势所趋，不可逆转。基于此，必须设计一套司法人工智能应用的规则体系，划定其应用边界，以最大限度地规避司法人工智能的风险，进而推动司法人工智能的有序运行。

（一）基本立场：尊重司法价值功效，司法逻辑主导算法逻辑

作为现代司法的价值导向，公平正义不仅是司法裁判具有公信力的正当性基础，还是司法人工智能获得约束力的应然要求，并且，于个案之中实现公平正义乃现代司法的内在逻辑。故此，人工智能技术应用于司法领域基于司法数据编写算法应以司法逻辑——追求公平正义为主导²³。具言之，首先，成为拥有内在法律品格的公正裁判者应为设计司法人工智能算法的逻辑起点，于此，其能够正确适用法律作出科学裁判，尽可能地摒弃一切可能引发司法不公的人为干扰因素，以做到同案同判。在此同时，具备审理疑难案件的能力被赋予为司法人工智能的应有之义，据此其能够识别法律漏洞并予以填补，继而进行翔实的论证说理²⁴。其次，不同于法律信息检索系统以便捷为开发目标，司法人工智能以前述司法逻辑为意旨进行构筑。正所谓公平正义的实现以司法理性为必要，而司法逻辑则是践行司法理性的手段。就司法逻辑本身而言，其含括多种意蕴，存有不同理解，且个中理解并无优劣，仅依据是否合乎目的而进行适当选择²⁵。理论上，纯粹理性的逻辑推理能够很大程度上忽略外界环境的干扰，以增加司法裁判的确定性，进而使法的安定性实现最大化。然而，客观实践中的案件事实千差万别，法律条文的旨意包罗万象，以致于司法决策过程充满多变性，在此语境下，被嵌入司法人工智能的算法逻辑必须以司法逻辑为出发点，保持稳定与理性客观，最大限度发挥其辅助司法之价值功能。

除此之外，司法逻辑还包含实现个案公正与普遍公正平衡之意。普遍公正乃司法的生命之源，表现出抽象性特点，与之相反，个案公正通常为具象，各具特色的个案公正共同构成了普遍公正，普遍公正又指引个案公正的实现，二者相辅相成。具体来说，不同表现形态的个案公正共同勾画了普遍公正的谱系，而普遍公正又给个案公正提供了对照，司法工作人员于此之中来回穿梭以寻求普遍公正与个案公正之平衡，此亦即司法逻辑的另一重语境²⁶。立基于此，被嵌入司法人工智能的算法应尊重个案情节法律认定与法律规则解释明的结合，以此为逻辑，判断目标案件与过往案件之异同，在统一法律文本适用的前提下，促成个案公正的达成。与此同时，在保证法律条文适用正确的基础之上，司法人工智能算法还应加大释法说理的逻辑比重，重视先前司法裁判的社会影响效果，并将其纳入至算法编写的考量因素之内，尽可能为司法工作人员决策的作出提供科学理性参考。即借助技术但不对其形成依赖，最终促使个案中实现具体公正与一般公正的协调一致。

（二）路径选择：智能辅助人类，强调司法工作人员的主体地位

首先，针对司法人工智能于司法活动之中的角色，应秉持“辅助地位”的理念定位，而就司法人工智能的发展，则可采取“有限智能化”的态度推进。作为现代科学技术的前沿方向代表，人工智能技术应用前景广阔，强化其于司法领域之中的应用能够使司法智慧化改革的推动获益良多。具体来说，应从实际出发，以回应司法实践、直击司法痛点、尊重司法规律为导向，开发、

推广并应用司法人工智能，继而实现“人工智能+司法”的既定效用。此外，肯定司法工作人员的主体地位，司法人工智能的监管由司法工作人员负责，让司法工作人员享有技术的最终决定权，于司法活动之中确立人类终审原则。正如法国学者克罗齐耶所言，相较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²⁷。现阶段，在司法转型时期案件数量增多、复杂程度增大的环境背景下，部分司法工作人员虽然有时可能会因为工作压力基于非理性作出与司法公平正义相违背的裁判，但这绝不能成为人工智能取代司法工作人员主体地位的缘由，反而应将其当作加强司法工作人员主体性培养的动因，以此为契机，注重司法工作人员综合素养的提升与复合型知识的储备，于司法人工智能和人类智力之间，坚定地选择人类作为司法活动的最终决策者，通过人类智慧矫正机器算法所致偏差。概言之，机器不应该主导司法工作人员的思考，而只能为其提供辅助。

其次，与肯定司法工作人员主体地位相对应的是，还应厘清司法人工智能场景中司法责任的承担主体与归责原则，落实司法责任制。司法人工智能在为司法工作人员提供帮助的同时，难以避免地会出现工作失误²⁸，由此产生的瑕疵判决乃至错误判决所生责任应以过错推定为原则，交由司法工作人员承担。理由如下：在司法人工智能应用的初级阶段，司法人工智能于司法工作人员的监控之下运行，对于案件的裁判过程与裁判结果的作出，司法工作人员仍享有最终决定权，以司法工作人员为担责主体，一方面能够有效践行权责一致原则的要求，防止因司法工作人员的倦怠而使司法责任制浮于形式，另一方面还有利于督促司法工作人员积极履职，加强对司法人工智能运行过程和结果的管控。当然，倘若确有证据证明错判系人工智能技术的关联方如开发者所致，司法工作人员对此并无过错，则应当减轻、免除司法工作人员的责任进而重新划定。

(三)技术保障：以算法和数据为着力点，改良司法人工智能整体生态

第一，优化算法。一方面，提高算法的准确性应寓于算法程序的开发编写过程中，尽最大努力提升判断结果的正确性，乃被嵌入司法人工智能的算法所追求之首要目标，亦是对其最为基本的要求。所谓判断结果的准确性，指的是算法执行任意一个合法命令都能得出正确的结果。另一方面，增加算法的透明度，乃被嵌入司法人工智能的算法程序开发编写过程中之另一目标。司法机关在与技术开发公司签订技术外包协议时，应写明就司法人工智能所涉的程序、算法、数据、具体决策，技术开发者得尽到记录与必要释明义务，并且在保护商业秘密的前提下，公开算法具体的决策过程并接受社会监督。此外，还应于算法的开发设计之中防范其歧视性，深度探索算法运行机理，从技术角度对其予以完善，并以风险防范为出发点，消弭因算法偏见所致的隐患。

第二，审查算法。建立专门的算法审查委员会，委员构成包括法律业务专家、技术开发专家、社会公益团体、有关职能部门，于算法程序开发、应用的全过程之中贯穿审查。在算法程序投入应用之前，算法程序的设计开发者应就算法程序的开发原理、运行程式、决策依据向算法审查委员会进行说明。在算法程序应用于司法领域之时，算法审查委员会可随机抽查算法程序的运行情况，结合实践对算法程序进行定期审查检验，并责令算法开发者及时改正错误算法程序。与此同时，拓宽民众对算法的监管渠道，积极接受民众反馈，回复民众有关算法决策的困惑，必要时依据相关法律规定启动问责机制。

第三，挖掘整合当下司法数据和历史司法数据、近期和远期司法数据，多途径收集司法数据，尤其是收集优质司法数据，提升司法数据的量与质。首先，应要求司法机关积极配合开展文书上网公开工作，在规定时间内规范上传裁判文书，增加司法“结果数据”的供给量。其次，除司法判决、裁定外，对于能够反映裁判决策过程的信息，诸如庭审笔录、合议庭笔录、审委会讨论过程等，建议将其纳入信息收集的范畴之内，拓展司法数据的表现形式，为算法编写开发提供更多的“过程数据”。再次，注重司法数据质量的把控，加大司法信息鉴真、结构化工作的实施力度，以关联对比、专家分析、数据清洗、图谱构建等专业技术为手段，科学审慎地对司法数据进行剖析、解构、筛选、校检，以供司法人工智能模型设计参考。

注释：

1 早在数十年之前，域外就已开始尝试人工智能与法律领域的结合，其早期研究重点为如何有效借助人工智能的方法建构基于规则和案例的法律推理模型或专家系统。随着人工智能与法律的深入结合，近些年来域外相关方面的研究愈发细微、具体。例

如,司法人工智能产品研发主体的问题,人工智能产品帮助法官采信科学性专业证言的问题,司法过程中运用私主体开发的人工智能产品所可能产生的问题等。详情参见 Pamela S. Katz, “Expert Robot: Using Artificial Intelligence to Assist Judges in Admitting Scientific Expert Testimony,” Albany Law Journal of Science and Technology, vol. 24, no. 1 (2014), pp. 1-45; Natalie Ram, “Innovating Criminal Justice,” Northwestern University Law Review, vol. 112, no. 4 (2018), pp. 659-724.

2 这些代表性的智能司法系统样态大多来源于《法治蓝皮书·中国法院信息化发展报告(2020)》中的介绍、笔者的实地调研与通过百度搜索引擎获取的智能司法应用实例。

3 以苏州法院的庭审语音识别系统为例,其在庭审中普通话识别正确率已接近 100%,稍带口音的普通话识别正确率也达 90%。

4 罗洪洋、李相龙:《智能司法中的伦理问题及其应对》,《政法论丛》2021 年第 1 期。

5 常鑫、张祥龙:《人工智能介入司法的实践检视与风险防范——以江浙沪冀黔五省市的司法实践为样本》,载《司法体制综合改革配套改革中重大风险防范与化解——全国法院第 31 届学术讨论会获奖论文集》(上),人民法院出版社 2020 年版,第 553—554 页。

6 [德]卡尔·拉伦茨:《法学方法论》,商务印书馆 2003 年版,第 340—341 页。

7 肖建国:《程序公正的理念及其实现》,《法学研究》1999 年第 3 期。

8 Anne von der Lieth Gardner, *An Artificial Intelligence Approach to Legal Reasoning*, Cambridge: The MIT Press, 1987, p. 9.

9 [美]本杰明·卡多佐:《司法过程的性质》,商务印书馆 2005 年版,第 7 页。

10 最高人民法院提出,针对当事人争议较大、法律关系复杂、社会关注度较高的一审案件,以及所有的二审案件、再审案件、审判委员会讨论决定的案件,应增强其裁判文书的说理性。

11 左卫民:《关于法律人工智能在中国适用前景的若干思考》,《清华法学》2018 年第 2 期。

12 刘铁光:《风险社会中技术规制基础的范式转换》,《现代法学》2011 年第 4 期。

13 徐骏:《智慧法院的法理审思》,《法学》2017 年第 3 期。

14 顾培东:《社会冲突与诉讼机制》,法律出版社 2004 年版,第 114—116 页。

15 强世功:《法律人的城邦》,上海三联书店 2003 年版,第 23 页。

16 所谓“最佳案情理解模式”,乃是司法工作人员基于对案情的理解而形成,而“最佳证据识别模式”亦是司法工作人员基于对证据事实的合法、真实及关联程度进行识别与认定而形成。

17 陈瑞华:《司法体制改革导论》,法律出版社 2018 年版,第 18 页。

-
- 18 陈卫东：《司法责任制改革研究》，《法学杂志》2017年第8期。
- 19 林艳：《人工智能的符号主义纲领及其困境》，《求索》2019年第6期。
- 20 郑曦：《人工智能技术在司法裁判中的运用及规制》，《中外法学》2020年第3期。
- 21 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》，《东方法学》2020年第3期。
- 22 所谓“难办案件”，指的是案件事实清楚，却没有明确法律条文可以援引，或适用法律结果不合情理的案件。有关“难办案件”的详述可参见苏力：《法条主义、民意与难办案件》，《中外法学》2009年第1期。
- 23 刘立明：《“感受到公平正义”的法治意蕴》，《江苏社会科学》2020年第5期。
- 24 左卫民、陈明国主编：《中国特色案例指导制度研究》，北京大学出版社2014年版，第152页。
- 25 [德]乌尔里希·克卢格：《法律逻辑》，法律出版社2016年版，第17页。
- 26 姚显森：《疑罪从无个案公正实现机制的完善》，《人大法律评论》2015年第2期。
- 27 [法]米歇尔·克罗齐耶：《法令不能改变社会》，格致出版社2008年版，第96页。
- 28 司法人工智能工作失误的情形大致可分为两类：涉及实体事项判断的错误与无关实体事项判断的错误，前者如偏离预警不准确、量刑建议有偏差，后者如电子卷宗生成瑕疵，语音识别不准。