

司法大数据与人工智能技术应用的风险及伦理规制

王 禄 生^{*}

摘要:与传统司法信息化相比,司法大数据与人工智能技术的应用呈现出数据前置性、算法依赖性、自我适应性与领域限定性的技术特征以及范围全面性、功能根本性、地位关键性与态度开放性的时代特征。上述特征与司法场景特性的融合交叠可能诱发司法固有属性被消解、法官主体地位被削弱、司法改革目标被替代和司法改革结果失控等风险。为了有效防范风险,在司法大数据与人工智能技术应用中必须遵循以保障司法固有属性为终极目标、以强化法官主体地位为根本出发点、以工具主义为功能定位、以比例原则推动审慎创新的伦理规范。

关键词:司法改革 大数据 人工智能 技术伦理

我国司法系统对大数据与人工智能技术的应用肇始于2013年前后。在最高人民法院的推动之下,相关应用在各地迅速落地运行。^①随着前沿技术、司法属性与时代需求的共同卷入,司法大数据与人工智能技术逐步呈现出传统信息技术以及通用大数据与人工智能技术所不具备的新特征。这些新特征在给司法系统带来动力的同时,也与司法的固有属性产生不同程度的冲突,继而带来新风险。受制于风险的潜在性、非指涉性和模糊性,现有的法律规范和通用技术伦理无法有效覆盖。在此背景下,认清司法大数据与人工智能技术应用的特征,进而塑造新型技术风险观和新型领域伦理观就成为亟待解决的理论问题。

目前,学界有关大数据与人工智能技术应用的风险分析主要集中在通用领域。^②对于由司法场景的特殊性所导致的司法大数据与人工智能技术应用的新特征、新风险和新伦理这一关键问题,理论的探讨相对不足,深入技术逻辑的分析更为缺乏。本文尝试对此展开深入技术逻辑的研究,并具体分解为3方面的问题:第一,现阶段司法大数据与人工智能技术应用与传统司法信息化以及通用大数据与人工智能技术相比具有什么显著特征?第二,上述特征与司法场景的特殊性相结合产生了何种风险?第三,对于这种风险为何现有法律规范与通用技术伦理无法充分覆盖?这又需要我们用何种领域伦理规则予以应对?

一、司法大数据与人工智能技术应用的特征

从近5年司法信息化的角度观察,司法人工智能技术虽然是在司法大数据技术之后被提出的,但实际上在科技发展历程中出现得更早。现阶段人工智能技术的诸多司法应用,如诉讼策略辅助、案件结果预测等,在20世纪80年代末90年代初就已面世。例如,20世纪90年代研发的案件推理系统就可以从案件

^{*} 东南大学法学院研究员、东南大学人民法院司法大数据研究基地研究人员

基金项目:国家社会科学基金项目(17ZDA131)、最高人民法院检察理论研究课题(GJ2017C09)、东南大学优秀青年教师教学科研资助计划

^① 截至2018年3月,全国86%的法院建成高度信息化的智能化诉讼服务大厅,包括智能导诉机器人等一批应用投入使用。参见刘子阳、王芳:《铆足马力创新发展深化智能应用》,《法制日报》2018年3月13日。

^② 马长山教授提出人工智能技术可能给社会带来伦理风险、极化风险、异化风险、规制风险与责任风险。参见马长山:《人工智能的社会风险及其法律规制》,《法律科学》2018年第6期。

中提取事实,然后为双方当事人形成辩论要点与反驳观点。^①令人遗憾的是,类似研究很快遭遇“知识获取瓶颈”。原因在于,司法人工智能技术应用的成败取决于能否获得海量法律知识以及能否将这些海量知识转换为计算机可识别的知识。^②而当时可供训练的法律数据量与计算条件否决了很多今天流行的司法人工智能算法。神经网络之父加拿大多伦多大学教授杰弗里·欣顿曾感叹:“回顾过去,失败只是源于缺少数据量和计算能力。”^③囿于短期内不可摆脱的桎梏,司法人工智能与其他领域人工智能技术一起在20世纪90年代之后进入低谷。近年来,得益于大数据技术对“知识获取瓶颈”的突破,司法人工智能技术取得飞速发展。^④也正因为大数据技术的基础作用,本轮司法人工智能技术以“概率”为中心,包括面向预测与预警的各种算法和模型。这与前一阶段司法人工智能技术围绕“逻辑”(法律表达与推理)展开有所不同。考虑到“司法大数据”与“司法人工智能”在技术逻辑方面的相似性与承续性,笔者将其整合为“司法大数据与人工智能技术”这一关键词,并从整体上切入观察。

(一)技术特征

数据前置性、算法依赖性、自我适应性是司法大数据与人工智能技术相较于传统信息技术的特征,而领域限定性则是其区别于通用大数据与人工智能技术的关键所在。具体而言:

1. 数据前置性

结合场景来说,司法大数据与人工智能技术是按照既定目标,对判决书、案件卷宗、庭审视频等非结构化、半结构化数据进行探索和分析,揭示隐藏的、未知的规律性,并进一步将其模型化,从而为司法提供支持的技术总称。其实质是通过数据挖掘发现深层司法规律进而加以运用的过程。与传统数据分析相比,数据挖掘是在没有明确前提预设下的信息挖掘和知识发现。^⑤有学者将这种没有理论预设、由数据到知识的进路称作人类知识发现的第四种范式。^⑥以判决预测为例,其技术逻辑是从半结构化的判决书大数据中自动抽取情节和判决结果,在此基础上挖掘特定情节组合的判决规律,进而对具备类似情节的目标案件进行预测,是典型的“数据挖掘→规律发现”的过程,而非“规律预设→数据验证”的过程。可见,司法大数据与人工智能技术的实质是建立了一种基于海量数据挖掘的认知范式,数据具有绝对的前置性。司法大数据与人工智能技术应用往往不再来源于它的基本用途,而更多源于它的二次挖掘。许多大数据收集的时候并无意用作其他用途,而最终却产生了很多创新性的用途。^⑦

2. 算法依赖性

大量数据并不产生价值,因此需要技术解决方案,其中以算法尤为重要,而数据挖掘就是通过算法从海量数据中搜索隐藏其间的信息的过程。以类案推荐为例,其技术逻辑在于自动抽取每个案件的情节以形成司法大数据仓库,随后将目标案件的属性标签与原有司法大数据仓库中所有案例的属性标签进行比对,在计算相似度之后反馈给用户。影响推荐准确度的关键在于相似度算法,目前较为流行的“K最近邻

^① 参见於兴中:《人工智能、话语理论与可辩驳推理》,载葛洪义主编:《法律方法与法律思维》(第3辑),中国政法大学出版社2005年版,第122~123页。

^② See Henry Prakken, AI & Law on Legal Argument: Research Trends and Application Prospects, 5 SCRIPTed: A Journal of Law, Technology and Society, 449-451(2008).

^③ 转引自李彦宏等:《智能革命:迎接人工智能时代的社会、经济与文化变革》,中信出版社2017年版,第17页。

^④ 当前司法人工智能应用的核心思路被人工智能专家概括为“大数据、小任务范式”(big data for small task,也称作“鸚鵡模式”)。换言之,本文论证的司法人工智能实际上属于“大数据智能”的流派。然而,“大数据智能”并非当前实现人工智能的唯一方向。“小数据、大任务”(也称作“乌鸦模式”)的人工智能研发已经展开。我国《新一代人工智能发展规划》就将人工智能分为大数据智能、跨媒体智能、群体智能、混合增强智能、自主智能等领域。美国资深人工智能学者钱德拉塞克拉曾谈到:“最近几年,人工智能在很大程度上集中于统计学与大数据。我同意由于计算能力大幅提高,这些技术曾经取得过某些令人印象深刻的成果。但我完全有理由相信,虽然这些技术还会继续改进、提高,总有一天这个领域(指AI)会对它们说再见,并转向更加基本的认知科学研究。”参见周志华:《机器学习》,清华大学出版社2018年版,第i-ii页。

^⑤ 参见周英、卓金武、卞月清:《大数据挖掘:系统方法与实例分析》,机械工业出版社2016年版,第7页。

^⑥ 参见黄茂荣:《大数据如何看待理论、因果与规律——与齐磊磊博士商榷》,《理论探索》2016年第12期。

^⑦ 参见[英]维克托·迈尔-舍恩伯格、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2014年版,第197页。

算法”(KNN)已被证明能够大幅度提升类案推荐的准确性。^①然而,正如算法领域“没有免费午餐”(No Free Lunch)定理强调的那样,不存在能够在所有可能的分类中性能均为最优的算法,^②因此,在司法大数据与人工智能技术开发中针对同一主题也需要反复尝试不同的算法并寻求最优解。

3. 自我适应性

传统算法就是一系列指令,告诉计算机该做什么,将数据输入计算机,然后得出结果。它在更大程度上可以被视为解决特定任务、硬编码的程序。司法大数据技术的应用就需要广泛借助传统的固定算法。与之形成鲜明对比的则是在此轮人工智能热潮中大放异彩的机器学习算法,尤其是其中的深度学习算法。此类算法并不是固定的,而是具备自我构建能力,能够随着计算、挖掘次数的增多而不断自动调整算法的参数,使挖掘和预测结果更为准确。^③从这个角度出发,算法是否具有自我适应性就成为区分司法大数据与司法人工智能的核心标准。目前在量刑预测开发中广泛应用基因算法、卷积神经网络、循环神经网络等深度学习算法,可以通过量刑预测结果与实际量刑结果的比对,对量刑模型采用的算法进行自评估,由此形成一个自适应优化的闭环。

4. 领域限定性

司法大数据与人工智能技术并非通用技术在司法场景中的平移运用。实际上,司法场景的特殊性提出了独特的技术需求,也由此形成了诸多具有领域特色的技术。这主要可以从两方面展开:其一,司法领域用语呈现出明显的语义鸿沟。然而,当前人工智能领域的自然语义处理技术主要是围绕日常用语展开训练,其与司法领域中法言法语的适配性较弱。因此,就需要开发法律领域的专属自然语义处理技术,深入到语义链的层面实现法言法语与日常用语在语义空间的协同映射与对齐。这绝非现成技术的平移运用,而是一种结合领域需求的技术创新。其二,相较于其他领域而言,司法领域的知识本体构建具有维度多、属性多、要素多的复杂性特点。以刑事司法领域为例,本体就是将刑事司法的关键问题,如罪行、罪犯特征、从轻情节、从重情节等概念及其关系以计算机可以理解的语言加以表示。有学者研究显示,刑事案件领域本体构建的属性维度是医疗诊疗领域本体构建的2.5倍、属性多样性是其2倍,如仅故意伤害罪就有包括作案工具在内的43个主要维度,每个维度的属性值更是多样。^④通用领域的知识本体常常在“向量空间”表示,^⑤也就是通过一组数据记录信息。考虑到司法场景的多维度、多属性的特征,“向量空间”表示势必需要丢失大量有效数据,此时就必须借助“张量空间”,以高维数据表最大限度地记录司法本体信息。由“向量”到“张量”也绝非既有技术的平移运用,而是一种突破瓶颈的技术创新。

(二)时代特征

得益于顶层的大力推动,司法大数据与人工智能技术的应用不再是某一片断、局部场景的小范围运用,而是呈现出范围全面性、功能根本性、地位关键性与态度开放性的时代特征。

1. 范围全面性

或许是因为管理工作及管理者更易于接受现代信息技术,也或是因为司法(审判)方式有保守性而较难因应现代技术迅速改变,总之,我国司法信息化建设在前一阶段的重要切入点在于审判管理方面而非审判本身。^⑥具体而言,主要围绕“法院综合信息管理系统”展开,对立案、分案、审理、审批等各环节实现动态管理,使得案件信息录入、审限管理、程序运行等方面更加系统、规范、客观、真实。^⑦与以往不同的是,

① “K最近邻算法”(K-Nearest Neighbor, KNN)广泛被应用于相似度的计算,其在类案推荐中的详细技术逻辑,参见余贵清:《基于规则和案例推理集成的刑事案件量刑决策支持研究》,电子工业出版社2016年版,第108~115页。

② 参见姚海鹏、王露瑶、刘韵洁:《大数据与人工智能导论》,人民邮电出版社2017年版,第29页。

③ 参见涂子沛:《数据之巅:大数据革命,历史、现实与未来》,中信出版社2014年版,第268页。

④ 以“伤情”这一维度为例,包含的属性值就有轻微伤、轻伤、重伤等不一而足。参见余贵清:《基于规则和案例推理集成的刑事案件量刑决策支持研究》,电子工业出版社2016年版,第94页、附录E。

⑤ 简单来说,标量(scalar)是一个数,向量(vector)是一列数,矩阵(matrix)是一张二维的数据表,而张量(tensor)则是三维数据表。

⑥ 参见左卫民:《信息化与我国司法——基于四川省高级人民法院审判管理创新的解读》,《清华法学》2011年第5期。

⑦ 参见李万涛:《我国法院信息化建设研究》,载王小林主编:《信息化时代司法公开的逻辑与进路》,人民法院出版社2015年版,第520页。

当前阶段的司法大数据与人工智能技术已经不再局限于传统流程管理,而是扩展到“司法管理”领域之外的司法公开、诉讼服务、案件审判、判决执行等业务环节。从司法公开环节的精准推送、^①诉讼服务环节的结果预测,^②到案件审判环节的证据统一化辅助、^③判决执行环节的执行风险智能预警,^④再到司法管理环节的案件量预测^⑤与案件偏离度预警,^⑥都打上了司法大数据与人工智能技术的烙印。用官方的话语来表达就是“全业务”“全流程”与“全方位”嵌入司法大数据与人工智能技术。

2. 功能根本性

传统信息技术更大程度上是基于提升效率和降低成本的功能考量而被引入司法领域。尽管彼时确有部分论证中提到信息化技术能够提升司法公正,但更多是从间接的角度出发的,如信息化技术通过提升司法效率使当事人更快地获得诉讼结果,进而有助于实现司法公正。在上述话语中,技术并非司法公正本身,而只是对司法公正起到间接促进作用。但在现阶段,随着量刑辅助、证据统一化辅助、案件偏离度预警等技术的提出、开发和应用,司法大数据与人工智能技术不仅被赋予了传统意义上的提升司法效率、降低司法成本的意涵,还被赋予直接优化审判体系、提升司法能力、改善审判质效、实现同案同判、维护司法公正的功能。换言之,在小数据时代,信息技术在司法场景中的功能是工具性的,而在大数据与人工智能时代,信息技术在司法场景中的功能则具备了根本性。

3. 地位关键性

在法院信息化 1.0 与 2.0 时代,^⑦科学技术在司法现代化中主要是以辅助的性质出现,并处于边缘地位。现阶段,以大数据与人工智能为代表的先进信息化技术的地位则由边缘向中心过渡:信息化技术被官方视为破解司法改革难题的重要手段,从而被赋予了与司法改革同等地位的时代意义。这突出地表现在各层级领导的话语表达中。习近平总书记曾指出:“把深化司法体制改革和现代科技应用结合起来,不断完善和发展中国特色社会主义司法制度。”^⑧此外,中央政法委书记郭声琨也明确肯定了科学技术的关键作用,指出“要坚持科技引领、信息支撑,把现代科技应用作为政法工作现代化的大战略大引擎,不断提升新时代政法机关的核心战斗力,推动新时代政法工作跨越式发展”。^⑨同样是对科学技术的强调,最高人民法院周强院长在多个场合都强调“司法改革与信息化建设是人民法院事业的车之两轮、鸟之两翼”。^⑩

4. 态度开放性

一般认为,司法具有被动的属性。对于前沿技术司法一般奉行谨慎甚至保守的态度,通常需要在技术

① 《最高人民法院关于加快建设智慧法院的意见》提出“深度分析用户诉讼行为,挖掘用户个性化需求,精准推送司法公开信息”。其实质是在借助网页浏览记录进行“用户画像”的基础之上分析当事人的关注点,并推荐符合当事人需求的司法公开信息。

② 2018年1月,科技部发布了国家重点研发专项“公正司法与司法为民关键技术研究与应用示范”专题第一批指南,拟投入4.5亿进行16项司法大数据与人工智能技术的研发,其中就包括诉讼结果的预测。上述应用的技术逻辑通常是从海量裁判文书中抽取情节和裁判结果,通过机器学习等形成量刑预测模型,并在新的案件出现时自动提取情节、匹配模型进而做出判决结果预测。

③ 《人民法院信息化建设五年发展规划(2016—2020)》第30条规定“建设基于大数据智能服务的审判支持系统”。以证据统一化为例,上海与贵州法院通过海量卷宗的挖掘,构建刑事案件证据模型,把统一的证据标准镶嵌到数据化的程序之中,系统会自动预警证据中存在的瑕疵。参见余东明:《上海应用“人工智能”办案防范冤假错案 全国首个“智能辅助办案系统”问世》,《法制日报》2017年7月11日;汤婷婷:《“智慧法院”高效公正》,《贵州日报》2017年7月11日。

④ 通过被执行人的“人物画像”(社交关系、财产状况),预警财产转移风险,形成精准化执行策略的应用也逐步在全国各地试点运行。参见徐隽:《司法大数据 这样助力审判执行》,《人民日报》2017年5月24日。

⑤ 基于对大数据的挖掘,可以形成案件发展趋势的模型,进而对当前乃至未来一段时间内犯罪方式的演化、审判态势的发展提供预测和管理决策。浙江省高级人民法院就通过数据挖掘形成危险驾驶罪高发地段的地图并预测醉驾的数量。参见高敏:《高院牵手阿里“智慧法院”充满魅力》,《浙江法制报》2015年11月25日。

⑥ 参见丁国锋:《八种“机器人”助力苏州法官办案》,《法制日报》2017年4月15日;顾元森:《江苏将推广“智慧法院”庭审效率可提高30%》,《现代快报》2017年4月15日。

⑦ 法院信息化1.0以“数字化”为核心,关注的是纸质文档向电子文档的转换;法院信息化2.0以“网络化”为核心,关注的是法院工作方式由线下向线上转换;法院信息化3.0则以“智能化”为核心,是在数字化与网络化基础之上通过大数据与人工智能技术实现智能化。

⑧ 习近平:《坚定不移推进司法体制改革》,《新华日报》2017年7月11日。

⑨ 郭声琨:《以习近平新时代中国特色社会主义思想为指导 奋力开创新时代政法事业新局面》,《公安日报》2017年11月1日。

⑩ 周强:《以习近平总书记系列重要讲话指导司法改革和司法实践》,《人民法院报》2015年7月3日。

充分成熟并经过反复论证后才正式采用。^①正因如此,我国法院的信息化在较长一段时间内并未与其他行业保持同步。从时间来看,国家信息化建设起步于20世纪90年代初期,而在法院层面,信息化建设则是在1996年才正式起步。^②2000年前后,全国范围内开始推动电子政务网络化建设,^③而法院系统则是在2009年前后才开始全面推行网络化。^④与之前司法信息化形成鲜明对比的是,当前司法系统对大数据与人工智能技术报以极为开放的态度,相关技术的推动与国家的信息化建设高度同步。自2013年之后,最高人民法院密集出台了一系列旨在推动应用前沿科技的信息化建设规划,包括《人民法院信息化建设五年发展规划(2016—2020)》《最高人民法院信息化建设五年发展规划(2016—2020)》和《最高人民法院关于加快建设智慧法院的意见》。此外,在《最高人民法院关于全面深化人民法院改革的意见——人民法院第四个五年改革纲要(2014—2018)》65项改革任务中也有35项不同程度地依赖信息科技手段。^⑤

二、司法大数据与人工智能技术应用的风险

我国司法大数据与人工智能技术应用的技术特征与时代特征相互嵌套、叠加,在给司法大数据与人工智能技术应用带来动力的同时,也可能因为“技术异化”而给司法领域带来不容忽视的风险。^⑥

(一)司法固有属性被消解

信息技术对司法的固有价值尽管不容忽视,却也未能完全避免技术本身的优势在司法领域的不适应性和排斥性。也就是说,大数据与人工智能技术在其他场景的优势恰恰可能成为颠覆传统司法格局的潜在力量。^⑦这种风险在理论上可以从技术的本质属性来解释。按照法兰西学派的经典观点,技术代表着新的知识,而知识与权力之间是“维系与生产”的关系。权力与知识结盟,二者互相蕴含。^⑧特定的知识必然产生特定的权力,而特定权力的运作需要特定知识的支撑。新知识的产生和壮大会导致依托其的新权力发展壮大。^⑨“在信息时代,知识的问题比过去任何时候都更是统治的问题。”^⑩也就是说,在大数据、人工智能技术与司法领域融合这一现象的背后,根据“知识—权力”的理论框架,反映出的是新兴技术代表的技术知识与传统司法场景中的法学知识的“碰撞”过程,而知识的“碰撞”几乎必然导致权力的“碰撞”。^⑪尽管新兴技术介入司法场景并非才发生的现象,但与传统信息化技术相比,司法大数据与人工智能技术应用呈现出数据前置性与算法依赖性等技术特征。这形成了“算力即权力”的新型权力形态,^⑫它可能将司法纳入到更隐蔽、更泛在、更多元的技术权力的侵蚀之中。技术权力以科技为手段,以全流程司法数据的获取为基础,借助代码、算法和架构将支配深入到司法运作各个环节的毛细血管中。例如,2018年1月,在科技部发布的国家重点研发计划“公正司法与司法为民关键技术研究与应用示范”专项(第一批)的16项司法大数据与人工智能技术的研发中,就有相当部分涉及对司法全流程的数据解构与智能管控,如“支持面向审判各环节的审判行为合法合规监督和偏离预警,预警准确率不低于95%”。在此之前,部分省份

① 美国联邦最高法院首席大法官罗伯茨认为,尽管当今世界科学技术的发展突飞猛进,但法院应该像执着的乌龟一样回应。参见[美]罗伯茨:《美国联邦最高法院2014年年终报告》,黄斌、杨奕编译,《人民法院报》2015年1月16日。

② 参见蔡长春:《“智慧法院”为司法事业插上腾飞翅膀》,《法制日报》2016年4月13日。

③ 参见张毅:《促进电子商务健康快速发展外经贸部成立国际电子商务管理司》,《人民日报》(海外版)2001年8月31日。

④ 2009年8月19日,上海市第一中级人民法院启动了全国首家“网络法院”,从而使整个诉讼过程具备了网络化。参见《首家网上法院上海开庭 具备完善诉讼过程》,《人民日报》(海外版)2009年8月21日。

⑤ 参见中国社会科学院法学研究所国家法治指数研究中心、中国社会科学院法学研究所法制指数创新工程项目组:《中国法院信息化第三方评估报告》,中国社会科学出版社2016年版,第5~7页。

⑥ 关于科技异化的相关论述,参见李桂花:《科技哲思——科技异化问题研究》,吉林大学出版社2011年版,第182~183页。

⑦ 参见王琦、安晨曦:《时代变革与制度重构:民事司法信息化的中国式图景》,《海南大学学报》(人文社会科学版)2014年第5期。

⑧ 参见[法]米歇尔·福柯:《规训与惩罚》(修订译本),刘北成、杨远婴译,生活·读书·新知三联书店2017年版,第29页。

⑨ 参见[法]米歇尔·福柯:《规训与惩罚》(修订译本),刘北成、杨远婴译,生活·读书·新知三联书店2017年版,第251页。

⑩ [法]让-弗朗索瓦·利奥塔尔:《后现代状态:关于知识的报告》,车槿山译,生活·读书·新知三联书店1997年版,第17页。

⑪ 参见王禄生:《大数据与人工智能司法应用的话语冲突及其理论解读》,《法学论坛》2018年第5期。

⑫ 参见郑戈:《算法的法律与法律的算法》,《中国法律评论》2018年第2期。

已展开类似的技术运用,并被认为有助于履行审判监督管理权。^①也有法院利用智能的视频技术对法官庭审质量进行全面、实时的评查,以实现规范庭审流程、提高法官庭审驾驭能力、提升庭审质量的目标。^②用官方的话语来表达,司法大数据与人工智能技术可以用数据打造一套静默化、自动化、可视化的全流程监控系统,构筑“数据铁笼”、实现“科技控权”。^③可见,相关技术运用的实质是将司法的全流程都纳入技术、算法、模型的泛在监控之中。在此过程中,司法应当具有的独立特性就可能在技术权力的“肢解”中不断消解。

更进一步,与传统信息化技术在法院定位的边缘性和工具性相比,在本轮司法大数据与人工智能技术的推进过程中,技术应用具有的时代特征导致围绕技术知识和法学知识而形成的技术权力与法律权力之间的冲突将范围更广、程度更深。一言以蔽之,当前司法大数据与人工智能技术应用特征的潜在结果是司法作为一个系统的独立存在的正当性将不断被侵蚀。司法毫无限度地向大数据与人工智能技术开放的直接结果,是极大地增加了司法被技术取代以及司法权沦为一般国家权力的潜在风险。

(二)法官主体地位被削弱

司法大数据与人工智能技术应用过程中的数据前置性和算法依赖性的技术特征会造成削弱法官主体地位的潜在风险。从法学思想史来看,“法官才是法律世界的王侯”。^④法律通过法官的适用而形成对世界的规制/治理力量。然而,数据前置性和算法依赖性的技术特征使得司法大数据与人工智能技术的应用极易形成“数据主义司法观”——一种将法官司法过程数据化并以数据主导司法的理念。^⑤由此,司法场景中的一切事物皆可被数据化并纳入到非预设的算法挖掘轨道之中。它不仅反映了人们相信司法大数据与人工智能技术可以解决一切司法问题的信念,而且还表达了人类对未来生活司法模式的理解。在此过程中,算法并不把法官预设司法场景的特殊主体来对待,而是作为可计算、可预测、可控制的客体。^⑥“数据主义司法观”的一个潜在结果是导致司法决策由“以法官为中心”转向“以数据为中心”,“司法智识”不再为法官所独享。也正是从这个角度看,传统信息技术对司法场景的影响总体上是外部的、局部的、工具意义上的,而司法大数据与人工智能技术则可能是内部的、全面的、本体意义上的。

此外,科技哲学认为,技术具有(隐性)操纵性等综合特征。^⑦法兰克福学派代表人物马尔库塞就认为,科学技术同意识形态一样,具有明显的工具性和奴役性,具有统治人和奴役人的社会功能。^⑧法官越是依赖技术就越有可能受技术“操纵”。与传统信息化技术形成鲜明差异的是,当前司法大数据与人工智能技术应用具有的时代特征客观上为法官构建了一个依赖技术的制度环境——利用技术被视为进步的标志,反之则被视为落后。由此,法官越依赖科学技术,其越有可能被“奴役”而被削弱甚至丧失主体性。技术运用可能在形式上和短期内强化法官能力,但长远来看则可能形成一种技术对法官的隐性操纵。

实际上,法官中心地位的弱化并非纯技术哲学层面的思考。从已经投入使用或者正在研发的相关应用来看,其理念已经体现了法官决策权由部分让渡到完全让渡、由非核心决策权的让渡到核心决策权的让渡的发展历程。以证据统一化为例,部分法院通过海量卷宗的挖掘,构建刑事案件证据模型,把统一的证据标准镶嵌到数据化的程序之中,系统会自动预警证据中存在的瑕疵。^⑨法官根据智能提示判断预警的

① 参见丁国锋:《八种“机器人”助力苏州法官判案》,《法制日报》2017年4月15日;顾元森:《江苏将推广“智慧法院”庭审效率可提高30%》,《现代快报》2017年4月15日。

② 该系统被称作“庭审智能巡查系统”或“庭审智能督查系统”,在搜索引擎上键入上述关键词即可发现已有产品面世。

③ 参见阎凌欣、詹旋江:《智慧法院:释放不一样的司法红利》,《福建日报》2018年4月21日。

④ [美]德沃金:《法律帝国》,李常青译,中国大百科全书出版社1996年版,第361页。

⑤ “数据主义”(data-ism)是一种视一切为数据并以数据为中心的世界观,大数据技术的发展极大地推动并强化了“数据主义”。参见高兆明:《“数据主义”的人文批判》,《江苏社会科学》2018年第4期。

⑥ 参见郑戈:《算法的法律与法律的算法》,《中国法律评论》2018年第2期。

⑦ 参见王魏芳:《警惕技术权力这把“双刃剑”》,《社会科学报》2016年3月3日。

⑧ 参见陈爱华:《哈贝马斯科技伦理观述——评哈贝马斯〈作为“意识形态”的技术与科学〉解读》,《伦理学研究》2007年第3期。

⑨ 参见余东明:《上海应用“人工智能”办案防范冤假错案 全国首个“智能辅助办案系统”问世》,《法制日报》2017年7月11日;汤婷婷:《“智慧法院”高效公正》,《贵州日报》2017年7月11日。

准确性从而做出下一步的决策。在此过程中,法官的决策权已经部分让渡给智能系统。与证据预警类似,类案推送、量刑辅助等技术应用在逻辑上都包含不同程度的决策权让渡。又如,部分法院推出了基于复杂程度来识别案件的“案件繁简分流智能化平台”。此类平台可以按照既有规则智能地将案件区分为繁案、普案、简案 3 类,并自动分配给相关法官审理。应用该类技术对案件复杂度的判断已经隐含着法官非核心决策权的完整让渡。再如,还有法院正在开发轻罪速裁智能办案系统。根据设想,由于速裁案件通常案情简单、法律关系明确,可以模式化处理,即由智能系统根据起诉书自动抽取相关情节、识别证据要素、模拟量刑并自动生成判决书。^① 该系统研发的目标是批量处理简单刑事案件,使法官有足够的精力审理复杂案件。尽管上述设想具有正当目标,但如果该系统在技术指标上达到一定的水平并使法官产生系统依赖时,法官的核心决策权就可能在潜移默化中让渡出去,“数据决策”将最终替代“法官决策”。从现有发展形势来看,法官将全部决策权转移给算法的风险虽为时尚早,但部分决策权转移的风险已迫在眉睫。

(三)司法改革目标被替代

司法大数据与人工智能技术应用的特征还可能导致司法改革陷入“技治主义”的路径,司法改革的复杂性被化约为简单计算,司法改革的难题被通过技术性手段回避,最终导致司法改革发生“目标被替代”的风险。

其一,司法大数据与人工智能应用过程中的数据前置性和算法依赖性强化了“技治主义”的改革路径。所谓“技治主义”兴起于 20 世纪 30 年代的美国,强调在政治决策和社会管理过程中实行科学原则、技术手段和数量方法。^② 该理念致力于将政策推行、社会改革等过程全面改造为类似于物理学领域中的计算分析的规程,包括公共利益等在内的各种因素都被视为某种客观存在,相关决策就是通过计算和模型推理来发现客观存在和揭示一般规律的过程。^③ 司法大数据与人工智能技术的数据前置性、算法依赖性与“技治主义”的主张不谋而合,推动了司法改革“技治主义”的产生。具体而言,就是把司法改革的诸多内容视为精密的、决定论的、可预测的对象,把司法大数据与人工智能技术视为精确测量改革对象的工具。“技治主义”司法改革路径的潜在风险是把相对丰富的司法改革问题简化为机械主义的计算,大大降低决策者对司法改革难度的预期。此种趋势在本轮司法大数据与人工智能技术的应用中已然出现,如通过“案件偏离度”的智能计算来解决同案同判和司法公正的问题、通过“量刑辅助系统”来解决法官独立审判问题、通过“法官饱和工作量大数据测算”来解决员额法官的动态调整问题。虽然,上述指标的提出和计算大都经过相应论证,也可以在一定程度上反映司法改革的关键问题,但是司法改革是一项系统工程,单一技术指标的提出和运用并不能从根本上解决与司法公正、独立审判等密切相关的诸多体制、机制难题。

其二,司法大数据与人工智能技术地位的关键性,甚至是与司法改革等量齐观的时代特征可能进一步固化“技治主义”的改革路线,并造成司法改革目标被替代的潜在风险。长期以来,受制于各方面的局限,我国司法改革一直奉行技术主义的改良路径,即在司法改革过程中聚焦于中观乃至微观的机制、程序方面的技术性调整。司法大数据与人工智能技术的兴起及应用,一方面,潜在地降低了部分决策者对改革难度的预期;另一方面,为此前一以贯之的技术主义改良路径提供了正当化的话语支撑。原有的不触及体制的技术主义改良路径与前沿技术紧密地结合,构筑了新的改革话语的正当性。前文论及“功能根本性”时提及的改革决策者期冀通过技术来优化审判体系、提升司法能力、维护司法公正的尝试就是明证。此时,司法改革将可能失去整体推进的布局,陷入局部切入;体制与机制改革中的难点问题被替换为通过“细枝末节”的技术设计来调试工作机制和改进工作方法。^④ 久而久之,转型期司法改革的矛盾就可能被压缩在依

^① 类似的“智慧公诉辅助办案系统”在检察机关已经投入使用,两者的技术逻辑是一致的,只是场景由法院替换为检察院。参见范跃红、关巧巧:《办理“醉驾”案驶入“快车道”》,《检察日报》2018 年 6 月 13 日。

^② 参见刘永谋:《论技治主义:以凡勃仑为例》,《哲学研究》2012 年第 3 期。

^③ 参见向玉琼:《论公共政策的“公共性”》,《浙江社会科学》2016 年第 2 期。

^④ 关于中国司法系统的“技术性改革”,可参见宋远升:《技术主义司法改革与法治现代化》,上海人民出版社 2017 年版,“序言”第 1~2 页。

托前沿技术更容易解决的范围内,寄希望于通过技术应用来缓解司法改革,尤其是司法体制改革的压力,从而产生司法改革的目标替代效应——当司法改革遭遇体制难题和瓶颈时,决策者可能就倾向于诉诸大数据与人工智能技术,从“技术”的视角切入进行改革,而不是试图通过体制改革予以回应。

(四)司法改革结果失控

技术的本质是在“确定性终结”基础上的复杂性探索,其结果具有不确定性。事实上,此种不确定性也正是推动技术进步的关键所在。但是,司法场景对稳定性的追求与技术的不确定性之间势必产生冲突,并诱发预期之外的伴生风险。并且,司法大数据与人工智能技术应用的技术特征与时代特征又强化了上述不确定性,使得稳定性与不确定性的抵牾加剧,相关司法改革面临结果失控的潜在风险。

概括起来,这一潜在风险是由司法大数据与人工智能技术应用的技术特征与时代特征共同衍生的单一性、建构性与潜在性造成的。^①首先,司法大数据与人工智能技术应用的领域限定性特征必然衍生出技术单一性。司法大数据与人工智能技术要结合司法场景的特殊需求进行专有的技术开发,如本文第一部分提及的语义空间协同映射与对齐、知识本体的张量表示等。这些新的领域技术通常是一个没有先例可循的单一性工程,在开发过程中我们通常需要面临司法场景许多单一的现象,并作出针对性的单一决策。由于没有先例可循,与成熟技术的直接应用相比,司法大数据与人工智能技术可能形成一系列技术风险与法律风险,结果更加不确定、不可预期。例如,在犯罪人再犯几率预测的技术应用过程中,意外地产生了对特定种族、肤色、性别人员的算法歧视;旨在加强公检法三机关办案效率的“公检法智能协同办案系统”可能导致三机关“相互制约”进一步弱化。其次,司法大数据与人工智能技术应用的领域限定性特征导致功能的建构性。科学技术本质上是一种社会建制,其发生作用取决于科学技术子系统与社会其他子系统之间的互动。给社会公众提供定向化与个性化司法公开和诉讼服务的技术开发,其成效不仅受到人物画像等技术先进程度的影响,也可能受到社会公众对相关技术干预隐私的容忍度的影响。司法大数据与人工智能技术由于刚刚出现,因此同其他社会子系统的磨合进度、难度与成熟技术具有明显差异。再次,司法大数据与人工智能技术应用的数据前置性与自我适应性特征极易造成结果的潜在性。科技哲学的通说认为,技术的潜在性主要是因为技术的应用效果与应用领域可能远远超出技术发明时的最初设想;^②技术成果产生后其具体的应用方式和目标指向并不受技术开发时原初目的的制约,出于合理目的开发的技术不能保证技术使用者的动机和使用方式也是合理的。^③以司法大数据与人工智能技术为例,其核心功能的发挥在于对数据的二次开发,由于数据收集者与数据使用者分离、目标与方向之间不可控,因此技术应用结果具有不确定性。例如,最初用以提供个性化服务的当事人信息挖掘也可能在二次挖掘中被适用于执行环节对当事人偿债能力的判断上。此外,司法人工智能技术的自我适应性特征突破了“机器只能做程序员编排它去做的事情”这一局限,如果结合某个具体目标学习足够多的司法数据,那么最终系统会产生什么结果完全不可预见,不受创造者控制。^④这也加剧了司法人工智能技术开发结果的不确定性。

综上所述,司法大数据与人工智能的技术特征使得司法改革的成效具有不确定性。“人类在应用这些科学技术的同时却无法全面预知它的整体效应以及它与其他事物错综复杂的相互联系。”^⑤申言之,我们很难预料会出现何种司法大数据与人工智能技术,这些技术会被应用到司法领域的哪些具体场景,进而又会造成何种影响。因此,部分大数据与人工智能技术在未稳定和未成熟的情况下就率先在对稳定性有较高要求的司法系统内推动,其结果的不确定性将可能给司法带来不可预见的负面影响。与此同时,我国当前司法大数据与人工智能技术应用的时代特征又进一步强化了上述风险。一方面,范围全面性、功能根本

^① 关于现代技术的单一性、潜在性与不确定性的相关论述,参见庄友刚:《风险社会中的科技伦理:问题与出路》,《自然辩证法研究》2005年第6期;李文潮:《技术伦理面临的困境》,《自然辩证法研究》2005年第11期。

^② 参见庄友刚:《风险社会中的科技伦理:问题与出路》,《自然辩证法研究》2005年第6期。

^③ 参见庄友刚:《科技伦理讨论:问题实质与理论自觉》,《观察与思考》2017年第3期。

^④ 参见郑戈:《人工智能与法律的未来》,《探索与争鸣》2017年第10期。

^⑤ 安宝洋、翁建定:《大数据时代网络信息的伦理缺失及应对策略》,《自然辩证法研究》2015年第12期。

性与地位关键性使得结果具有不确定性的司法大数据与人工智能技术进入司法的全流程、全领域、全场景；另一方面，态度开放性可能造成技术风险观缺位并加大了上述风险产生的可能性。

三、司法大数据与人工智能技术应用的伦理规制

当前，司法大数据与人工智能技术的应用尚处于初步阶段，技术对司法的影响主要以正面形式呈现。然而，随着技术应用的不断深入，司法固有属性被消解、法官主体地位被削弱、司法改革目标被替代、司法改革结果失控的风险有可能逐步加剧。为了避免陷入“科林格里奇困境”，^①在司法大数据与人工智能技术发展的初期就应当通过法律规范或技术伦理对其应用进行有效规制。令人遗憾的是，现有法律体系和通用领域的技术伦理对司法大数据与人工智能技术应用失去规制能力。一方面，由于本文探讨的风险尚未完全发生（潜在性）、不针对具体个人（非指涉性）、损害难以证明与量化（模糊性），因此传统民事、刑事、行政责任体系无法有效涵盖；另一方面，由于司法大数据与人工智能技术应用是以正面目标为出发点的，因此传统技术伦理中诸如“不作恶”等要求就对其面临的风险失去评价能力。可见，现有法律规范和通用技术伦理的失范产生了对全新规范体系的需求。然而，作为新兴的技术形态，司法大数据与人工智能技术现正处于创新发展的重要阶段。此时，如果片面依靠法律进行禁止性或强制性规制，那么将可能陷入“管制型”的思维而因噎废食、扼杀创新。因此，通过领域技术伦理规则的构建，培育一种司法大数据与人工智能技术应用的自我约束机制就成为这个时代平衡技术创新与规避风险之间张力的重要手段。从科学哲学的角度看，司法领域技术伦理规则的实质是在司法大数据与人工智能技术的应用过程中嵌入技术伦理价值，并据此制定伦理框架和行为规范，从而规范开发者与使用者的行为，帮助他们清晰地认知在技术应用过程中所应遵循的理念、反思在技术应用过程中所应避免的禁区。具而言之：

（一）以保障司法固有属性为终极目标

“从广义的层次看，伦理问题总是受到一定的文化和实践制约的……技术伦理中的伦理标准位于科学与技术之外。”^②因此，司法大数据与人工智能的技术伦理涉及其与法学伦理之间的互动关系。毋庸置疑，司法有着相对于其他国家权力的独立性和延续性。司法全流程、全领域、全场景贴上大数据与人工智能的技术标签将可能与司法固有属性形成潜在冲突，并削弱司法存在的独立价值。这就要求在司法大数据与人工智能技术开发的过程中，充分尊重司法的固有属性，不触动司法权的被动性、透明性、仪式性、事后性等基本底线，将开发的具体领域与技术应用的程度限制在不消解司法存在正当性的范围内。

第一，司法大数据与人工智能技术的应用要充分尊重司法的被动性。司法的被动性是指“法院不能主动对任何一项社会争端或事项进行裁判活动，它也不能主动干预或介入社会生活，而只能在有人向其提出诉讼请求以后，才能实施司法裁判行为”。^③被动性是司法权区别于行政权的核心特征，也是司法保持中立的根本要求。这意味着，“线索发现”^④、“提前介入”“深挖案源”“保驾护航”等主动启动权力的大数据、人工智能技术与司法权的本质存在不同程度的兼容难题。因此，在上述技术应用的开发过程中，应该遵循这样的原则，即所有“主动/提前介入型应用”应当在纠纷进入法院之后方可适用，不鼓励在纠纷尚未进入司法环节就主动去发现与识别，并进行先期引导。

第二，司法大数据与人工智能技术的应用要充分尊重司法的透明性。与行政活动通常呈现的秘密性和封闭性不同，司法裁判活动应当具有公开性和透明性。其中，透明性是指裁判结论在形成过程、根据和理由方面的“公开”，也即司法裁判的透明。^⑤现阶段，量刑建议是司法大数据与人工智能技术最为典型的

^① 由英国学者科林格里奇在1980年提出，意指一个技术在发展初期若不充分规制，发展壮大后就无法规制。参见陈凡、贾璐萌：《技术控制困境的伦理分析——解决科林格里奇困境的伦理进路》，《大连理工大学学报》（社会科学版）2016年第1期。

^② 李文潮：《技术伦理面临的困境》，《自然辩证法研究》2005年第11期。

^③ 陈瑞华：《司法权的性质——以刑事司法为范例的分析》，《法学研究》2000年第5期。

^④ 《人民法院信息化建设五年发展规划（2016—2020）》第30条要求，建设基于大数据智能服务的审判支持系统，建设线索发现等审判支持应用。

^⑤ 参见陈瑞华：《司法权的性质——以刑事司法为范例的分析》，《法学研究》2000年第5期。

开发领域之一。在其出现前,法官判决需要进行充分的说理和论证以便当事人和公众理解。但是,量刑决策系统并不如此运作,普遍人往往难以理解其算法的原理和机制,因为自动化决策常常是在算法这一不透明的“黑箱”中做出的。实际上,此种借助算法进行预测与预警的决策辅助类应用,除了量刑建议之外,还有诸如被告人再犯风险评估、被执行人转移财产风险评估、虚假诉讼风险评估等,无论是直接运用人工智能系统的推荐结果,还是法官在筛选后有选择地运用决策系统的建议,其本质上都与司法固有的公开属性存在冲突。公众对司法的信任建立在司法透明的基础上:如果这些算法继续保持神秘,那么司法大数据与人工智能技术将很难取得公众的信任。这一信任不仅关乎司法大数据与人工智能的成功发展,而且也关乎司法体系长期、健康的运行。^①可见,在司法领域中,与片面追求算法的准确性相比,算法的透明度和可解释力更加重要。司法人工智能技术较多地借助深度学习算法,但这些算法因自我适应性而缺乏可解释性。以神经网络为例,算法在每个环节的行动都是基于数据的随机选择。因此,下一步决策辅助类技术开发时应当特别关注过程的透明性,尤其要将人工智能技术模拟的决策过程以公众可以理解的方式可视化地呈现出来。这就要求技术研发人员针对司法领域的需求开发具有更强解释力的算法。实际上,提高深度学习算法的可解释性也是前沿技术突破的重要方向之一。

第三,司法大数据与人工智能技术的应用要充分尊重司法的仪式性。司法仪式首先由一系列符号构成,这些符号包括法院的物理环境(法庭、办公大楼等)、法槌、法袍、特定的程序、法庭的装备。上述符号的组合形成了公众对司法的固定认知并成为司法权威的重要来源。目前,司法大数据与人工智能技术开发过程中所推动的虚拟法院(微信调解、远程庭审)^②等改革虽然在很大程度上提升了诉讼当事人获取诉讼服务的便利化程度,但是,由于网络庭审等应用在实体法庭之外展开,缺乏法庭传统布置和服饰而丧失诉讼程序的正统性、庄严性、权威性与严肃性。^③从司法固有的仪式性考虑,虚拟法院的开发应当遵循以下原则:其一,非庭审事项可以尽量通过网络交互的方式实现,如证据交换、案件调解等,而庭审事项原则上仍应在实体法庭中进行;其二,案件事实和证据有争议的案件应慎用虚拟庭审的方式;其三,即使使用虚拟庭审,也要保证宣判环节在实体法庭中进行,从而保证司法的仪式性与严肃性。

第四,司法大数据与人工智能技术的应用要充分尊重司法的事后性。司法只针对已经发生的行为。因此,当前司法大数据与人工智能技术应用过程中基于被告人大数据而开发的再犯风险预警系统、缓刑决策辅助系统等都一定程度上与司法的事后性存在冲突。这是因为,基于未来可能行为之上的惩罚是对公平正义的侵蚀,因为公平正义的基础是人只有做了某事才需要对它负责。基于大数据所导致的犯罪治理活动提前启动与无罪推定的原则存在根本冲突。^④

(二)以强化法官主体地位为根本出发点

针对司法大数据与人工智能技术可能导致的法官主体地位被削弱的风险,其技术伦理应以强化法官主体地位为出发点。

第一,司法大数据与人工智能技术的应用要尊重法官的独立性。提前预警是当前司法大数据与人工智能技术开发的重要方向。尤其在审判执行与司法管理领域已经有相关的应用投入运行。例如,通过对海量裁判文书的量刑学习形成智能模型,当法官判决结果与既往海量文书“历史经验”形成冲突时,系统可以实现对当前判决偏离度的实时预警。此类开发可能出现与法官自由裁量权的潜在冲突,尤其是部分地区在开发过程中规定系统识别超过一定偏离度的案件无法进入下一环节,这实际上就是对法官独立行使职权的一种事前干预,也与员额制改革中要求还权一线法官的理念不完全兼容。对此,可行的改进方法应该是系统自动预警,允许法官自主依据预警对判决进行修正;若法官仍然坚持意见,则允许法官基于裁量

① 参见李本:《美国司法实践中的人工智能:问题与挑战》,《中国法律评论》2018年第2期。

② 远程庭审是传统司法信息化建设的内容。不过,随着司法大数据与人工智能技术的引入,远程庭审开始增加智能模块,包括根据在调案件情况智能生成调解预案、调解书,并给调解员推荐相似案件。

③ 参见侯猛:《互联网技术对司法的影响——以杭州互联网法院为分析样本》,《法律适用》2018年第1期。

④ 参见裴炜:《个人信息大数据与刑事正当程序的冲突及其调和》,《法学研究》2018年第2期。

权做出判决。事后审判监督部门可以在审核之后决定是否启动监督程序。也就是说,预警类的应用开发应当遵循如下原则:对于实体类预警,包括事实认定和法律适用,事前预警应当是建议性的;对于程序类预警,包括证据形式和案件流程,事前预警才可以是干预性的。

第二,司法大数据与人工智能技术的应用要充分尊重法官的亲历性。行政活动具有一定的便宜处分性,行政官员可以听取下级官员口头或书面汇报,对有关事项作出决定,而司法权则要求亲历性,即裁判者要亲自经历裁判的全过程。^① 反观目前开发的旨在减低司法人员工作负担的前端智能过滤系统,则很可能与司法的亲历性产生冲突。例如,部分主体正在开发涉法信访申诉案件智能处理系统,其可以通过证据形式的审查对海量申诉案件进行过滤、筛选与处理。概而言之,系统可自动屏蔽一些缺乏足够依据的申诉案件,将值得处理的案件推送给法官。且不说此类开发还面临诸多技术障碍,仅在功能设定阶段就可发现其某种程度上有“越俎代庖”之嫌,尤其是过滤申诉案件的行为实际上是一种机器决策,与法官的亲历性存在冲突。也就是说,相关的机器决策类应用应当遵循以下几个基本的原则:其一,对于法官的核心决策权,如审判权,以法官亲历为原则,强调法官的最终决定权;其二,对于法官的非核心决策权,如案件繁简分流等,可以适当由机器辅助。

(三)以工具主义为功能定位

针对司法大数据与人工智能技术应用导致的司法改革目标被替换的风险,其技术伦理应以工具主义为功能定位。

第一,在司法改革推进过程中要正确认识“技治主义”路径的利弊得失,避免唯数据主义和片面技术理性主义的误区,将前沿技术在辅助工具的价值层面展开,而不是将司法大数据与人工智能技术应用作为司法改革本身,尤其要避免将司法改革的复杂问题简化为一般的技术性问题。

第二,在推动司法大数据与人工智能技术应用于司法改革时要充分回避技术盲区,从而最大限度地发挥技术的工具价值。人工智能还难以应付人类主观意识影响的社会文化和意识领域的各类问题。^② 也就是说,人工智能技术还无法审时度势地处理法律纠纷中常见的利益和人情,也无法有效解决司法推理过程的复杂性问题。^③“数据分析本身无法揭示案件背后复杂的社会关系,容易抹平地区差异,也就无法取代法官在个案特别是疑难案件中的自由裁量、利益平衡以及对公平的判断。在这个意义上,司法经验和以后果为导向的实用主义思维方式比纯粹依赖数据做出的判决更加可靠。”^④更为重要的还在于,法律的适用不仅仅是法条以及判例的适用,在特定情境下,还需要满足不同的功能定位,如倡导某项新的原则、对特定群体进行警示。^⑤ 这显然也是司法人工智能技术无法具备的能力。技术盲区也决定了司法改革中相当多领域问题的复杂性不适宜通过简单的模型化来解决。

第三,在推动司法大数据与人工智能技术应用于司法改革时要充分警惕技术瓶颈,从而形成对技术工具价值的合理预期。受制于领域限定性的技术特征,当前司法大数据与人工智能技术应用在相当一部分领域的专有技术方面仍然面临障碍。从司法知识图谱构建中面临的自动化难题、到司法领域自然语义处理技术研发过程中需要解决的语义鸿沟、再到司法领域模型训练中面临的“垃圾进、垃圾出”悖论都不同程度地正处于技术创新的过程之中。^⑥ 这就决定了在司法领域大数据与人工智能技术应用的具体场景中,要对其可能发挥的工具价值做合理的预期。

① 参见陈瑞华:《司法权的性质——以刑事司法为范例的分析》,《法学研究》2000年第5期。

② 参见吴月辉:《人工智能会取代人类吗》,《人民日报》2017年7月7日。

③ 参见张新宝:《把握法律人工智能的机遇 迎接法律人工智能的挑战》,《法制日报》2017年6月28日;沈寅飞:《人工智能+法律=?》,《检察日报》2017年7月26日。

④ 胡凌:《大数据影响下的法律实践与法学研究》,《光明日报》2014年12月3日。

⑤ 参见高奇琦、张鹏:《论人工智能对未来法律的多方位挑战》,《华中科技大学学报》(社会科学版)2018年第1期。

⑥ “垃圾进、垃圾出”(garbage in, garbage out)是人工智能领域最为重要的难题,在司法领域是指模型训练过程中如何确保训练样本都是正确的判决书(生效判决不一定就是正确判决),其本质是司法人工智能训练数据的可靠性难题。关于司法大数据与人工智能技术面临的障碍,参见王禄生:《司法大数据与人工智能的技术限度》,《中国法律评论》2018年第2期。

(四)以比例原则推动审慎创新

针对司法大数据与人工智能技术可能导致的司法改革结果失控的风险,其技术伦理应以比例原则审慎推动司法大数据与人工智能技术的应用创新。比例原则又可以细分为适当性判断、均衡性判断和必要性判断3个部分:其一,在选定创新场景时需要优先进行适当性判断,即司法大数据与人工智能技术在该场景的应用能否有助于实现改革者的预期。其二,在适当性判断之后还需要进行均衡性判断,即将创新应用可能获得的正面效应(如司法效率与司法公正的提升)与其负面影响(如算法歧视、隐私侵犯、司法不公等)进行比较。考虑到司法领域对稳定性的高度依赖,只有确认正面效应明显高于负面影响时才可以作为创新场景推进司法大数据与人工智能技术的应用。其三,在确定应用司法大数据与人工智能技术之后,还需要进一步做必要性判断。其实质就是以对法院、法官、当事人影响最小、损害最轻的方式推动技术创新,最大限度地降低技术的不确定性,从而使改革不至于陷入失控的境地。下面笔者围绕现阶段已经产生的算法歧视与隐私泄露两大风险来实例解析比例原则的运用。

以面向诉讼当事人网上行为轨迹的司法公开信息精准推送技术为例,^①决定是否要在该场景展开技术创新需要遵循3个步骤:首先,需要判断该技术的应用是否真正有助于当事人获得更加精准的司法公开信息。众所周知,当前司法公开系统并未实现定向化的精准推送,这导致当事人往往需要经过大量信息浏览和筛选后才可能从司法公开平台上获得自己所需要的相关信息,这极大限制了司法公开成效的充分发挥。因此,从适当性判断来看,该应用场景无疑有助于实现改革者优化司法公开的目标。接下来,就需要进一步判定:精准司法公开与当事人隐私侵害,两者孰轻孰重。由于只是对当事人在司法公开网站上的行为轨迹进行整合,其对隐私的损害处于较低水平,因此,该应用场景能够通过均衡性判断,从而具备推进的正当性基础。接下来,还需要在技术研发过程中采用对当事人隐私侵害最小的方式进行,如在数据清洗时要充分脱敏,最大限度地去除可能涉及个人隐私的信息(如姓名、身份证、住址等);在数据挖掘时要在法院内网的环境中进行,避免信息泄露;建立数据保密制度,明确相关人员的责任;在任何时刻都禁止数据挖掘人员实行“反向匿名”等。^②一个相反的例证是未成年人犯罪风险预警技术。该技术虽然有助于实现公安机关犯罪预防的目标,然而基于未成年人的“画像”对不同个体犯罪风险进行预警,可能造成对预警对象的歧视,导致未成年人群体之中不恰当的社会分选,并且容易强化办案机关有罪推定的倾向,因此其无法通过均衡性判断而不具备推行的正当性。

四、结语

从某种程度上讲,本文是给方兴未艾的司法大数据与人工智能运动泼冷水,文中对大数据与人工智能技术在司法场景中运用的分析,有些可能是吹毛求疵甚或杞人忧天。然而,技术伦理从属于哲学,而哲学的社会功能在于批判性,即“防止人类在现存社会组织慢慢输灌给它的成员的观点和行为中迷失方向”。^③此外,本文的反思也是伦理演进和发生的前提。“技术的发展本身会反作用于我们的伦理观念以及对伦理问题的看法,并且已经导致了伦理的变化或者说进化。”^④也就是说,本文提出的司法大数据与人工智能的技术伦理更大程度上是在当前场景下展开的。未来随着相应技术的改进,部分伦理的内容势必发生相应调整。技术伦理与技术发展之间将在相互调试中逐步达到融洽状态。当然,伦理只是规制司法大数据与人工智能技术应用的一种方式而不是全部,其作用的发挥还有待于其他规制措施的协同推动。

责任编辑 王虹霞

^① 2018年7月,科技部发布了国家重点研发专项“公正司法与司法为民关键技术研究与应用示范”专题第二批指南,拟投入1.5亿进行5项司法大数据与人工智能技术的研发,其中就包括司法公开精准推送技术。

^② “反向匿名”是数据挖掘的一种方式,既在数据“匿名化”的前提下通过挖掘反推相应人员的身份。

^③ [联邦德国]马克斯·霍克海默:《批判理论》,李小兵等译,重庆出版社1990年版,第250页。

^④ 李文潮:《技术伦理面临的困境》,《自然辩证法研究》2005年第11期。