

热与冷：中国法律人工智能的再思考^{*}

左卫民

内容提要：法律人工智能界出现了话语与实践、域内与域外的冷热差异。中国在顶层设计、官方政策、学术界研究方面对人工智能的重视程度超过不少国家，但却在司法实践运用效果层面遇冷。而在域外尤其是美国，尽管政府层面的支持力度及辐射规模似乎远不及中国，但地方实践却有亮点。话语层面的“冷热不同”缘于中国与域外国家在人工智能上的认识差异。实践层面“冷热不同”的原因在于前期准备条件的不同、研究方法上的差异、法律数据充分性的不同以及法律人工智能的技术瓶颈。未来，中国的法律人工智能应当由冷转热，由热转实，持续强化法律人工智能的实践能力，通过引入统计学、计算机科学的人才从根本上改变法学院的基因，并通过开设法律大数据、法律人工智能等相关课程，实现真正的跨学科互动与交流。

关键词：人工智能 大数据 司法公开

左卫民，四川大学法学院教授。

随着人工智能技术的发展和进步，人工智能在世界与中国范围内的讨论愈来愈热。2018年6月，Open AI宣布，在Dota2游戏五对五的规则下，人工智能击败业余人类玩家，并准备对专业玩家团队发起挑战。而在一年前，人工智能还只是在一对一的规则下击败人类玩家。^[1]与此同时，美国IBM公司旗下的人工智能“Project Debater”与以色列辩论冠军奥瓦迪亚以及另一名人类辩论员扎福瑞尔，分别就“政府应资助外层空间探索”和“人类应使用远程医疗”问题展开了两场辩论。^[2]辩论过程中，“Project Debater”通过云端技术搜集报纸、学术论文等资料并将其转化为论述，同时分析对手的演讲并提出反驳。虽然“Project Debater”在第一场比赛中落败，但第二场辩论结果显示，机器人以九票优势

^{*} 感谢洪凌啸同学、叶燕杰同学、朱奎彬副教授、王禄生副教授对本文提出的修改意见。

[1] 参见杨吉：《机器人还不会有“普鲁斯特记忆”》，《上海证券报》2018年8月29日第8版。

[2] 参见李忠东：《人工智能“辩论员”登台亮相》，《检察风云》2018年第15期，第34-35页。

领先人类辩手。种种迹象表明,人工智能将继续“侵入”人类的各个领域,那么法律界尤其是中国法律界做好相关的准备了吗?

一 热与冷:法律人工智能界的悖论现象

(一) 话语与实践的冷热差异

1. 官方政策层面之“热”

面对来势汹汹的人工智能“革命”,中国政府对人工智能的支持力度空前强大。纵观世界,可以说尚未有哪个国家像中国这样通过官方的顶层设计来支持司法和公共安全领域大规模地运用人工智能技术的。2017年3月5日,李克强总理在政府工作报告中明确指出支持人工智能发展,强调“加快新材料、人工智能、集成电路、生物制药、第五代移动通信等技术研发和转化,做大做强产业集群”。^[3]这是中国向世界发出的明确信号——中国将成为人工智能技术的先锋。同年7月20日,国务院印发了《新一代人工智能发展规划》,提出“构筑我国人工智能发展的先发优势,加快建设创新型国家和世界科技强国”,^[4]力图在新一轮国际科技竞争中掌握主导权。^[5]另外,以笔者参与的科技部项目评审为例,其中很多内容都涉及到庭审的信息化,并且,项目的资金支持力度也很惊人。

在司法领域,人工智能的运用尝试更是如火如荼。在2016年第三届世界互联网大会智慧法院暨网络法治论坛上,最高人民法院院长周强表示:“将积极推动人工智能在司法领域的应用。”^[6]在地方层面,科技公司开始不断与各地法院联手,大力推进司法领域的人工智能应用,例如,科大讯飞与三省一市(江、浙、沪、皖)高级人民法院签署了《推进长三角地区三省一市“人工智能+法院”深度战略合作协议》,共同推进“人工智能+法院”在长三角地区的应用。^[7]此外,由上海市高级人民法院研发的“上海刑事案件智能辅助办案系统”(即206系统)成为一项标志性工程。^[8]

2. 学术界研究层面之“热”

受司法实务界的影响,法学界对法律人工智能的关注和研究热度也与日俱增。在中国知网上以“法律”“人工智能”为关键词搜索,可以发现,2016年前每年相关主题的文献仅为10-15篇,很多年份甚至是个位数。进入2016年,文献数量猛增为62篇,2017年更

[3] 李克强总理代表国务院在十二届全国人大五次会议上所作的《政府工作报告》,《人民日报》2017年3月17日第1版。

[4] 《国务院关于印发新一代人工智能发展规划的通知》,http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm,最近访问时间[2018-10-26]。

[5] 参见李克强总理代表国务院在十二届全国人大五次会议上所作的《政府工作报告》,《人民日报》2017年3月17日第1版。

[6] 邓恒:《人工智能技术运用与司法创新》,《人民法院报》2017年12月14日第7版。

[7] 参见《讯飞携手长三角法院共同推进“人工智能+法院”》,http://www.ah.xinhuanet.com/2018-06/07/c_1122949446.htm,最近访问时间[2018-10-26]。

[8] 参见严剑漪:《揭秘“206”:法院未来的人工智能图景——上海刑事案件智能辅助办案系统164天研发实录》,《人民法治》2018年第2期,第38-43页。

是达到 288 篇,法律人工智能的专著或专题研究类成果也开始出现。^[9] 此外,近年来国家重点研发计划、国家社科基金等重大科研项目中法律人工智能选题的出现与增多也预示了未来相关研究的方向。

通过对现有文献的梳理,笔者发现,目前关于法律人工智能的研究主要集中于以下几个方面:其一是探讨人工智能对整个法律体系、法律制度的影响。^[10] 其二是探讨人工智能所涉及的法律资格、主体的问题,例如是否应当将法律主体资格延伸至人工智能。^[11] 其三是探讨人工智能对部门法的可能影响,例如,人工智能时代侵犯财产犯罪的刑法适用、人工智能担任公司独立董事的法律问题^[12] 以及智能机器人的刑罚体系^[13] 等。与此同时,与法律人工智能相关的研讨会也开得不亦乐乎。一时间,法学领域的人工智能研究炙手可热。

3. 司法实践运用效果之“冷”

尽管在话语层面,我们不断宣扬人工智能已经成功在司法领域开疆拓土,其中的典型代表如上海刑事案件智能辅助办案系统、江苏的智慧审判系统等。但这并不意味着司法领域的人工智能运用已经获得了实质性突破。事实上,对于中国司法系统斥巨资力推的大数据及法律人工智能技术在司法实践中的运用效果并不理想。笔者在 S 省 C 市有关法院调研类案类判系统时曾了解到,一些法官表示法院提供的类案类判系统存在明显问题:首先,无法提供精准的类案。其次,类案范围不够广阔,局限于河南、山东等省份的类案。最后,类案的时间不延续,往往集中于若干年份,但在其它年份却销声匿迹。这就使得官方提供的类案类判系统无法真正起到参考作用,很多法官不得不连接外网寻求企业搜索平台的帮助。^[14] 另外,就上文提到的江苏智慧审判系统的应用情况来看,似乎也不太理想,江苏基层法院的部分法官甚至表示并未使用该系统。此外,上海的智能辅助办案系统也难称“完美”。对此,论者曾评价道,上海政法系统研制的“刑事案件智能辅助办案”系统主要着眼于证据的合法性自动审查,侧重于形式审查方面,而“在实质审查方面还有一些障碍”。^[15]

上述事实表明,目前法律领域的人工智能运用,在实践层面尚与司法当局主导的“话语热”存在相当落差。

[9] 例如,乔路、白雪著:《人工智能的法律未来》,知识产权出版社 2018 年版;[美]瑞恩·卡洛、[美]迈克尔·弗兰金、[加拿大]伊恩·克尔著:《人工智能与法律的对话》,陈吉栋、董惠敏、杭颖颖译,上海人民出版社 2018 年版。

[10] 例如,有学者指出,在人工智能时代,法律功能的独特性面临着丧失的危险。法律的规范性期望被认知性期望取代,法律被代码/算法取代。参见余成峰:《法律的“死亡”:人工智能时代的法律功能危机》,《华东政法大学学报》2018 年第 2 期;李晟:《略论人工智能语境下的法律转型》,《法学评论》2018 年第 1 期。

[11] 参见吴习或:《人工智能的法理与司法应用——论人工智能的法律主体资格》,《浙江社会科学》2018 年第 6 期。

[12] 参见林少伟:《人工智能对公司法的影响:挑战与应对》,《华东政法大学学报》2018 年第 3 期;江必新、郑礼华:《互联网、大数据、人工智能与科学立法》,《法学杂志》2018 年第 5 期;吴允锋:《人工智能时代侵犯财产犯罪刑法适用的困境与出路》,《法学》2018 年第 5 期。

[13] 对于智能机器人的刑罚体系,有论者认为必须坚持罪刑相适应原则、以刑罚目的为导向原则、刑罚节俭性原则。适用于智能机器人的刑罚可以有三种,即删除数据、修改程序、永久销毁。参见刘宪权:《人工智能时代刑事责任与刑罚体系的重构》,《政治与法律》2018 年第 3 期。

[14] 参见左卫民:《如何通过人工智能实现类案类判?》,《中国法律评论》2018 年第 2 期。

[15] 参见刘品新:《大数据司法的学术观察》,《人民检察》2017 年第 23 期。

(二) 域内与域外的冷热

与域外相比较,中国在顶层设计方面对人工智能的重视程度远远超过世界其他国家。但是,涉及到具体的、深度的人工智能运用时,却又是“上热下冷”,形成冰火两重天(尤其是司法领域)的景象。而在域外(尤其是美国),尽管政府层面的支持力度及辐射规模远不及中国,但地方实践却大相径庭。例如,美国一些州的法院已经开始有规模地将法律人工智能技术应用于审前保释或判后假释的风险评估之中。^[16]但是,这种做法引起了争议,由此还产生了著名案例,威斯康星州诉卢米斯一案(Wisconsin v. Loomis)即是这方面的典型。在该案中,卢米斯认为美国威斯康星州法院使用“风险评估工具”(Correctional Offender Management Profiling for Alternative Sanctions)智能量刑系统,将再犯可能性危险等级认定为“高风险”,有违客观性与中立性,要求美国联邦最高法院对此做出裁决。然而,2017年6月,联邦最高法院却拒绝受理该案件。^[17]笔者推测,美国联邦最高法院之所以拒绝受理该案,原因之一或许是因为最高法院的九位大法官对人工智能是否能够应用于司法实践尚无足够的信心,对人工智能在司法实践中运用的程度及范围尚缺乏把握。

除了上文提到的“风险评估工具”之外,美国的一些州也正在使用法律人工智能工具来确定被告人的刑期长短。这种“风险评估工具”是在参考了数十年的量刑案例后所设计的算法。^[18]该算法设置了所在居住地、受教育水平、关系亲密人群的犯罪状况、中学前是否有过乘坐飞机旅行的经历等权重模块,以此评估被告人在一定时期内重新犯罪的可能性。^[19]有人更是通过机器学习的方法研究了新奥尔良地检署 1988 - 1998 年十年间 145000 名被告人的 280000 起案件,建立了被告人的再犯可能性模型,分析何种案件、什么样的人会得到法院的减刑与保释。^[20]在《通过算法测算被告的再逮捕率与人类检察官大不相同,算法还能识别出干扰因素》一文中,论者即提出,如果算法得当,使用该模型可以降低 5% - 9% 的再犯危险,^[21]并有效区别出人类检察官所蕴藏的主观因素。在此意义上,法律人工智能可以作为一个基本的分析工具。其算法相比检察官的主观判断更加客观、公正。

美国康奈尔大学、哈佛大学、斯坦福大学等高校的一批学者,以乔恩·克莱因伯格

[16] See Christopher Slobogin, Risk Assessment, in *The Oxford Handbook Of Sentencing And Corrections* 196, 203 - 05 (Joan Petersilia & Kevin R. Reitz eds., New York: Oxford University Press, 2012). at 200.

[17] See Michelle Liu, Supreme Court Refuses to Hear Wisconsin Predictive Crime Assessment Case, <https://www.jsonline.com/story/news/crime/2017/06/26/supreme-court-refuses-hear-wisconsin-predictive-crime-assessment-case/428240001/>, 最近访问时间[2018 - 10 - 26]。

[18] 参见左卫民:《关于法律人工智能在中国运用前景的若干思考》,《清华法学》2018 年第 12 期,第 108 - 124 页。

[19] 参见左卫民:《关于法律人工智能在中国运用前景的若干思考》,《清华法学》2018 年第 12 期,第 108 - 124 页。

[20] 参见 Amaranto, et al., Algorithms As Prosecutors: Lowering Rearrest Rates Without Disparate Impacts and Identifying Defendant Characteristics “Noisy” to Human Decision-Makers (June 25, 2017), Available at <http://dx.doi.org/10.2139/ssrn.2993003>, 最近访问时间[2018 - 10 - 26]。

[21] 参见 Amaranto, et al., Algorithms As Prosecutors: Lowering Rearrest Rates Without Disparate Impacts and Identifying Defendant Characteristics “Noisy” to Human Decision - Makers (June 25, 2017), Available at <http://dx.doi.org/10.2139/ssrn.2993003>, 最近访问时间[2018 - 10 - 26]。

(Jon Kleinberg)、希马比都·拉卡拉胡立尔(Himabindu Lakkaraju)、莱斯科韦茨(Leskovec)、塞德希尔·穆来纳森(Sendhil Mullainathan)等领衔,利用决策树、迭代算法等机器学习算法,分析了美国法院 15 万余份重罪案件的承办法官所做的保释或假释决定,^[22]研究美国法官在给予嫌疑人保释或假释决定时究竟考虑了何种因素,哪种因素最为关键,哪些因素为人类法官所过度考虑以及保释决定是否正确等问题。最后,该研究团队发现:独立的机器学习算法的预测成绩要优于人类法官的判断,特别是当法官不当运用了“不具有观察性”因素的情况下,尤其如此。该研究团队进而提出,如果推广基于机器学习预测的保释或假释规则,就能使美国在犯罪率没有任何增加的前提下,将监禁人数减少 25%,或在不改变监禁人数的前提下,将犯罪率减少 20%,这等同于增加了 2 万多警力。^[23]

弗吉尼亚理工大学发现分析中心的学者则通过数据驱动结构的机器学习,分析了美国最高法院的以往裁判,并以此预测未来判决。人工智能通过分析裁判文书,计算与争点关系最密切的词语并设置相关权重,进而结合大法官的投票行为预测文字背后大法官们的重点关注所在。研究发现,人工智能不仅准确发现了法官之间不同的裁判观点与立场,还对大法官的中间妥协派做了区分。基于此,人工智能对美国最高法院未来裁判的预测准确率可达 79.46%。^[24]美国芝加哥的伊利诺理工大学与南德克萨斯法学院利用 1791 至 2015 年的美国最高法院数据库,合作开发了一种算法。该算法再现了从 1816 年到 2015 年美国最高法院法官的 28000 项决定和 240000 次投票,正确率分别达到 70.2% 和 71.9%,这高于法学家们 66% 的预测准确率。^[25]

此外,美国的法律人工智能技术还广泛地应用于对商业合同文本的审查上。例如,美国法律人工智能平台 LawGeex 与斯坦福大学、杜克大学法学院和南加州大学的法学院教授开展合作,让法律人工智能程序就四小时审查五项保密协议(NDA)与 20 名有经验的人类律师展开竞争。主要考察如何就仲裁、关系保密和赔偿的准确界定问题。实验结果显示,在准确性方面, LawGeex 以 95% 的准确率完胜人类律师的 85% 准确率。在速度方面,法律人工智能完成任务仅需 26 秒,而人类律师则需 92 分钟。^[26]英国律师事务所 Berwin Leighton Paisner 也在特定的产权纠纷案件上使用了人工智能系统,该系统从英国土地注册局出具的官方所有权契据中提取数据并核查细节,从而在房地产案件中准确提供关

[22] See Kleinberg, et al., Human Decisions and Machine Predictions (February 2017), *NBER Working Paper* No. w23180.

[23] See Kleinberg, et al., Human Decisions and Machine Predictions (February 2017), *NBER Working Paper* No. w23180.

[24] 参见 Mohammad Raihanul Islam, et al., What AI can Tell Us about the U. S. Supreme Court? Available at <https://the-conversation.com/what-ai-can-tell-us-about-the-u-s-supreme-court-55352>,最近访问时间[2018-10-26]。

[25] See Katz DM, Bommarito MJ II, Blackman J (2017) A general Approach for Predicting the Behavior of the Supreme Court of the United States, *PLoS ONE* 12 (4): e 0174698.

[26] 参见 Harry Pettit, The AI that is Faster and More Accurate than Top Human Lawyers When Spotting Legal Issues in Contracts, Available at <http://www.dailymail.co.uk/sciencetech/article-5435479/The-AI-outperforms-human-lawyers.html>,最近访问时间[2018-10-26]。

于产权所有人的法律提示。^[27] 加拿大科技公司开发的 ROSS 法律人工智能项目利用 IBM Watson^[28] 的平台进行技术开发与升级,能够识别用户自然语言问题(例如,公司破产之后是否能够继续经营),并给予精准、快速的回答。截止 2016 年 4 月,ROSS 就获得了超过 20 个客户,包括贝克豪思和瑞生等全球知名的律师事务所,并且开始向国际领域扩展。

我们从上文分析可以看出,国内和域外在法律人工智能的话语—实践—设计—应用领域呈现出“倒挂”现象。在中国,我们注重顶层设计,但对司法一线人员在法律人工智能技术应用方面的真正需求却知之不多,实践中真正有用并投入常态化运营的法律人工智能系统并不普遍。而在国外,尽管政府层面的顶层设计远逊于中国,但在司法过程中的应用(包括研究)却显然更为扎实。所以,就当前来看,尽管我国在话语层面不断地烘托出一片法律人工智能应用欣欣向荣的景象,但深入审视却发现,很多宣传资料上的预设并未落地生根,实际上,我们基本上还没有使用能够实现裁判预测功能的法律人工智能系统与工具。

二 冷与热:原因何在

上文分析表明,同域外一些国家相比,中国法律领域的人工智能技术应用存在着令人惊讶的“倒挂”现象。那么,在对待与运用法律人工智能的态度与实践中,为何会呈现出明显的“冷热不同”?进一步而言,影响法律人工智能话语与实践冷热不同的因素究竟为何?

(一) 话语层面为何“冷热不同”

上述现象的产生,主要原因在于中国与域外国家在人工智能上的认识差异。可以说,在法律人工智能顶层设计上,域内域外所体现出来的冷热不均是国内外在对待人工智能态度问题上的直接传导。尽管域内域外均将人工智能技术应用视为下一个社会变革与经济增长的引擎,但在对待人工智能的态度上仍有明显区别。相较而言,西方国家对于人工智能更为警惕,总体上呈现出一种“保守倾向”,而中国则积极得多。

事实上,在域外,在人工智能萌芽的初始阶段,即 20 世纪四五十年代计算机与自动化曙光初现时,人们即开始担心人与机器的相互关系问题了。而当 AlphaGo 开始在围棋界大杀四方后,这种忧虑再次被传递到了社会的各个角落,人们开始担心自身的地位以及未来是否会被机器所取代的问题。斯坦福大学人工智能与伦理学教授杰瑞·卡普兰甚至指出,除了体力劳动者,银行职员、文员、记者、会计、律师都要承受被取代的风险。^[29] 史蒂

[27] 参见《英媒:人工智能正在取代初级律师》,http://www.xinhuanet.com/world/2017-09/20/c_129708856.htm,最近访问时间[2018-10-26]。

[28] 美国 IBM 公司开发的 Watson 是能够使用自然语言回答问题的人工智能系统。2011 年 Watson 一鸣惊人,在一档问答类综艺节目“危险边缘”中打败了人类选手。其后,IBM 将 Watson 平台向各研发主体开放,用以在各垂直领域(如医疗、法律等)开展研究。本报告中的多个法律人工智能机器人平台都是在 Watson 的基础上进一步开发的。如今,Watson 已经被运用到超过三十五个国家十七个产业领域。

[29] 参见杰瑞·卡普兰:《人工智能即将颠覆这 8 大行业》,《科技中国》2016 年第 6 期,第 20-25 页。

芬·霍金、埃隆·马斯克等反人工智能派也再三强调,要警惕人工智能的风险。霍金称:“人工智能在不远的未来,可能会成为人类的一个真正威胁。”^[30] 马斯克更明确表示,“我们需要十分小心人工智能,它可能比核武器更危险”。^[31] 《未来简史》的作者尤瓦尔·赫拉利更是预测,人类将失去其神圣地位,成为机器人所圈养的动物,并可能被机器人随意屠宰。^[32] 西方国家的民众普遍将人工智能与《终结者》等科幻电影中的邪恶形象联系在一起,因此对人工智能有着天然的警惕。

所以,尽管在实践层面西方国家正在不断提升法律人工智能技术的运用范围与深度,但在国家的顶层设计方面却要慎重得多。西方法治发达国家的权利意识浓厚也是对人工智能运用保持警惕态度的重要原因之一。由于担心人工智能的运用会引发一系列的侵权行为以及当下作为权利维护手段的法律所无法解决的问题,西方国家普遍对人工智能尤其是数据的利用心存芥蒂,并设置了较多限制。例如,2018年6月22日,美国联邦最高法院以五对四的投票比支持对公民移动电话隐私权的保护,认为执法单位长时地掌握公民个人手机位置信息的行为,应受到美国宪法第四修正案的规制,并应受到较为严格的限制。^[33]

相反,不像域外一些人那样,我国不少人对于人工智能并未怀着如此之深的恐惧和警惕,民众普遍将人工智能技术视为“为我所用”的工具,而对其不利的一面则抱持乐观的心态。当前中国法律人工智能在实践中遇冷,某种程度上更多地是技术而非观念所致。在国家的顶层设计上,中国更是将人工智能视为下一个赶超西方发达国家的重要途径,因此在高层推动上不遗余力,从而进一步提升了人工智能在话语层面的热度。另外,在法律规制方面,中国对人工智能所设置的进入门槛较低,相应限制也较少,虽然正在对人工智能的伦理问题进行研究,但一定的滞后性反而为人工智能的政策选择与发展战略提供了有利条件。

(二) 实践层面为何“冷热不同”

1. 前期准备条件不同

中国在建设人工智能上没有“历史包袱”,但是正因如此,中国在人工智能技术、人才团队、资金投入上的前期准备远远不如西方发达国家。西方发达国家研究法律人工智能技术的历史已达数十年,而中国在这方面则刚刚起步。在美国,早已成立了专门的法律人工智能协会研究“法律+人工智能”问题。在人才资源方面,美国的 Google、Deepmind 等人工智能团队在大规模资金的支持下迅速成长,以年甚至月为周期推动人工智能行业不断突破。而在中国,法律人工智能像是一个“新生儿”,从官方到学者接触法律人工智能

[30] 牛禄青:《人工智能将引发“无用阶层”?》,《新经济导刊》2017年第12期,第54-59页。

[31] 江晓原:《人工智能:威胁人类文明的科技之火》,《探索与争鸣》2017年第10期,第18-21页。

[32] Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow*, New York: Harper Collins Publishers Inc, 2016: 318-321.

[33] Mazza, Sandy, U.S. Supreme Court Rules against Cellphone Searches without a Warrant, Available at <https://www.questia.com/article/1P2-36752439/u-s-supreme-court-rules-against-cellphone-searches>, 最近访问时间[2018-10-26]。

几乎都始于围棋比赛中 AlphaGo 取得对人类的碾压性胜利之后。即使逐渐兴起对法律人工智能的关注和研究,但是研究进路基本上仍依循传统的研究范式,即思考如何以其身处的部门法规制人工智能,而非以透彻地了解人工智能技术的基本原理及其特征为目的,更遑论成立专门研究法律人工智能的协会了。此外,在人才团队建设上,由于资本的作用,法律领域很难吸引专业的人工智能人才入行。最好的人工智能人才往往优先考虑互联网企业或金融行业,少数投身于法律人工智能行业的要么是迫于无奈,要么是将其视为暂时的栖身之所,真正兴趣使然的则少之又少。

2. 研究方法上的差异

在研究方法上,中国的法律人工智能研究远远落后于域外国家。从研究传统上看,域外国家更为重视法律实证研究,对于法律与科技的交叉研究有着长期的学术传统和坚实的研究基础。反观中国的社会科学尤其是法律界,其主流的研究方法仍然具有诸多文化、抽象化、思辨化的特征,而缺少实证化、技术化的研究方法的运用。从 2015 年前相关法律人工智能的论文数量屈指可数上就可窥见一斑。可以说,中国的法律学界向来缺乏运用科学技术方法模拟法律、运用法律的思维,在资金上和技术上也缺乏对人工智能运用的投入,进而导致整个中国法律界事实上都缺乏法律人工智能研究的土壤。而在西方的研究体系中,法律人工智能已经开始在商业化运营的大环境之下迈入了良性循环和发展之路。

尽管在媒体与学术会议上,我们在不断地讨论、推广与使用法律人工智能,但仔细审视就会发现,我们主要探讨的仍然是如何规制人工智能这一新事物,而不是从技术逻辑切入当下人工智能的最新发展以及人工智能的核心——算法方面讨论人工智能对旧有法律制度、体系所形成的冲击。申言之,我们的知识路径依然停留在法学圈子范围内,真正了解、熟悉人工智能发展现状与趋势的人并不多,也缺少从人工智能角度阐释法律未来发展方向的相关作品和思考。

另外,中国的法学家在研究工具上大大落后。受制于学科训练背景,我们的法学家通常既不熟悉统计学的相关知识,对最新的机器学习方面的相关知识一般也所知甚少。而在域外,统计学的方法几乎是一项基本的研究技能。笔者在参加欧洲第二届法律实证研究年会时发现,相关学术会议在正式开始之前,会用一个整天的时间对与会人员进行会前培训,其主要目的就是提升参会者用统计学的方法进行实证研究的能力。培训会议上分布着各个年龄段的学者,无论是老教授还是学术新人,他们都有意愿与需求学习统计学的方法。这说明,国外法学界的研究范式正在发生根本性的改变——法学教授开始学习数学、统计学、机器学习、数据科学等基本方法,这些工具构成了域外学者进行法律实证研究的基础。

并且,这一影响已经开始在域外的学术界结出了一定的学术成果,其学术影响力在不断扩散与增加。除了一些典型的论文成果,国外的法学期刊杂志越来越接受那些通过标准的统计学方法进行法学定量实证研究的文章。然而,这种类型的论文在国内尚难以被主流的法学期刊杂志所接受。由此可见,我们缺少基础性的工具与思想支撑。在中国,目

前尚未看到如欧洲、美国那样的基于先进算法的法律研究文章,甚至连通过统计学方法进行法学研究的论文也数量寥寥。或者即使使用了统计学方法,也只是简单地使用描述性统计学的方法,而缺少(使用例如回归等方法)分析性统计方法的文章。如果新一代的学者们既不懂统计学的相关知识,也不了解人工智能中的机器学习算法,也就无法撰写如国外学者那样的文章,更有可能在最新的国际学术竞争中落于下风。在这样的环境下,要想深入研究法律人工智能问题是不可能的。

因此,我国法学界迫切地需要来一场法学知识的范式革命,我们不仅需要将法学研究由书斋法学向实证法学推进,更要将文本法学向数据法学转变。可以说,如果法学的知识范式没有革命性的变化,那么,法律人工智能的技术应用在很大程度上就不可能获得成功。

3. 法律数据的充分性不同

法律人工智能在司法领域的逐步推进是一个漫长而艰巨的过程。从现阶段看,即使是域外国家,司法实践中所使用的预测性法律人工智能也有范围的限制,例如,美国学者主要通过人工智能机器学习算法研究,预测美国最高法院而不是美国全部法院的判决结果。这一现象的出现有其背后缘由。美国法院分为联邦与州两级法院系统,不同州之间的法律体系差异也很大,许多州在法律数据的公开上并不完善,很多时候甚至没有记录完整的庭审过程。能够做到全程记录法庭庭审过程的法院是美国最高法院,这也就解释了为何美国法律人工智能对审判结果的预测首先围绕美国最高法院展开。近年来,随着法律人工智能的发展,判决结果预测开始不断向联邦最高法院之外的联邦法院扩展,比如知名的预测平台 Premonition。^[34]

域外法律人工智能的另一主要战场是商业合同文本的审查环节。法律人工智能之所以能够在这一领域大显身手,原因就在于商业合同文本在数据方面所具有的独特优势。商业合同文本在社会中的使用量巨大,并且合同文本的内容呈现出半结构化甚至全结构化的样态,很多时候要么是合同模板,要么是在合同模板的基础上进行了一定修改,但基本的结构并未改变。因此,这种规则清晰、数据巨大、结构稳定的文本天然具有适用人工智能的优势。

综上所述可以看出,法律人工智能特别是深度机器学习在域外司法实践中的运用具有相当的潜力,并且在有些方面已经展现出接近甚至开始超越法律人的表现。当然,这一切的前提是充分、准确的司法数据支撑,即无论是司法大数据还是司法小数据,都需要通过深厚、扎实的实证研究来获得,以确保数据的代表性与准确度。唯有如此,才能通过司法数据“喂养”法律人工智能,进而生产出正确的司法结论。或许这就是未来法律人工智能在某些领域与方面超越人类法官的可能之道。

相较域外而言,我国在法律人工智能与大数据研究特别是基于大数据的法律人工智能应用方面虽然较为有限,但在司法数据方面却略胜一筹。这不仅在于我国司法机关的

[34] 相关动态可以参见网站 <https://premonition.ai/>。

司法数据在总量上比域外法院多。更为关键的是,中国的裁判文书网业已成为全球最大的裁判文书公开平台,每天都有不断更新的司法裁判文书上网公开,为我们提供了极为重要的司法数据来源。这些最新的一手数据,甚至吸引了全世界范围内研究司法制度学者的目光。根据统计,有“超过 17.5 亿的访问量来自海外,访问范围覆盖全球 210 多个国家和地区”。^[35] 此外,中国法院还在不断地将法院的历史裁判文件档案电子化,通过自动化的扫描手段将近几十年来的法院纸质裁判文书转化为 PDF 格式保存。笔者在参加由科技部牵头的司法技术项目评标会上也发现,有项目研究者指出,在将纸质文件转化为 PDF 格式时,对图片数据内容的识别准确率可以提升至 80% 以上。目前,我国法院系统正在不断地将过往的司法裁判文件数据化,并且还在努力提高裁判文书上网的数量,可以说,中国的司法数据是较为丰富的。

而在域外,欧洲法院一般不将法院的裁判文书上网,也就不可能具有像中国裁判文书网这样大的数据库。此外,欧盟刚通过的《通用数据保护条例》(GDPR)对数据使用做了相当大的限制,这也在一定程度上阻碍了司法的网上公开。而在美国,以刑事诉讼为例,多年来,美国在州和联邦两级刑事重罪案件中至少有 90% 是通过辩诉交易解决的。^[36] 那么这些案件的司法数据自然无法最终被跟踪、使用。即使剩下的 10% 左右的案件进入了最终的庭审环节,但是英美对抗式的审判模式决定了在陪审团机制下,相关的事实认定与证据使用只涉及支持或不支持,而没有裁判说理的过程,这也决定了在庭审中难以为人工智能系统提供历史性的决策数据。

当然,尽管中国在裁判文书网建设方面效果突出,甚至许多域外学者都通过中国的裁判文书网研究相关司法问题。^[37] 但是,中国在司法大数据运用方面依然存在着许多问题。首先是司法数据过于依赖裁判文书网。尤其值得注意的是,裁判文书网上的司法裁判文书虽然有些案件类型(如醉驾类案件)公开程度较为充分,但有些案件类型(特别是职务犯罪类案件、行政诉讼类案件)的上网比例依然不足,整体上呈现出一种结构性的缺失。因此,如果过分依赖裁判文书网,可能导致我们基于裁判文书网所建构的裁判模式出现系统性偏差。其次,特别需要指出的是,裁判文书中所记载的判决信息并非中国司法决策过程中的关键信息,其只是审判过程中信息含量很少的一部分。很多时候,合议庭、审判委员会中法官的讨论过程、法官判决后所撰写的审判报告或许才记载了判决形成的决定信息。很显然这些文件、信息都尚未上网公开,难以为学者所研究利用。最后,还需要辨明裁判文书上所载的表面信息与案件判决中的实质信息的差异性。网上裁判文书所公开的信息往往只是文字化的表面信息,至于其背后的实质内容为何,隐藏了多少其他的政治、社会因素考量,我们不得而知。因此,在数据具有结构性缺陷、数据

[35] 中国裁判文书网总访问量突破百亿, <http://www.court.gov.cn/zixun-xiangqing-57042.html>, 最近访问时间 [2018-10-26]。

[36] 参见王禄生著:《美国司法体制的数据观察》,法律出版社 2018 年版,第 183-186 页。

[37] 例如,哥伦比亚大学法学院的本教授通过裁判文书网上河南法院的相关数据研究了中国刑法中的宽大与严惩问题。参见 Benjamin L. Liebman, Leniency in Chinese Criminal Law? Everyday Justice in Henan, 33 *Berkeley J. Int'l Law*, (2015)。

所含并非关键性内容、数据与真实情况背离的状况下,仅仅依靠裁判文书网上的大数据或其他有限公开渠道上的信息,谋求对中国法律运行真实情况的充分认识,显然是不客观、不科学且具有局限性的。因此,对于司法大数据及以此为基础搭建的法律人工智能,应当保持充分的冷静与警醒。

4. 法律人工智能的技术瓶颈

尽管各个领域的学者对于何为人工智能见解不同。有些未来学的学者甚至简单地将科幻电影中的人工智能形象与现实中的人工智能技术运用等同起来,但就其本质而言,人工智能是通过算法所搭建的预测模型。因此,人工智能的本质问题是围绕着数据、算法、模型而展开的技术问题,人工智能科学更多地应归入到以统计学方法为基础的数据科学中。

因此,需要追问的问题是:通用领域的人工智能技术真的适用于法律的独特场景吗?这些技术还面临怎样的瓶颈?目前,中国的法律人工智能界并没有成体系地为“法律”量身定制一整套专有技术。市面上所使用的技术绝大多数是通用的共性技术,而没有充分考虑到法律行业的特殊性。更进一步,在部分法律场景专有人工智能技术研发尝试中也普遍面临瓶颈。以知识图谱构建为例,普遍采用自顶向下的人工构建方式,而自底向上的自动化构建方式则并未取得实质进展。^[38]

总之,中国法律人工智能在话语层面的“热”往往是一种“虚热”,而实践层面的“冷”基本上是一种“真冷”。西方国家法律人工智能实践层面的“热”是“实热”,是相对成功的“热”。中国的法律人工智能在研究思维与实际的运用能力及相关资金投入、技术支撑上,远远不如西方,这也是中国在法律人工智能实践上遇“冷”的根本原因。然而值得庆幸的是,中国对法律人工智能抱有更加乐观、积极的态度与期望,并且,相对于西方国家在司法数据上非记录化、非公开化的趋向,中国在司法大数据的收集、保存、分析与运用上具有较大的优势,中国法院对司法数据的公开也持更为开放与成熟的态度。或许当下中国的法律人工智能不如西方那般运用成熟,但相比西方有限的应用领域,未来中国将具备在更多领域超越西方的潜力。

三 结 语

在法律人工智能领域,中国的“冷与热”与西方的“热与冷”形成了鲜明的反差与对比,这一现象的背后映射出深刻的社会、法律、思想方面的差异。在未来,中国的法律人工智能要由冷转热,并进而由热转实,持续强化法律人工智能的实践能力。与此同时,我们也需要对法律人工智能的局限性有清醒的认识与判断,认识到不可能使法律人工智能运用的所有领域都由冷变热,由热变实。法律人工智能到底是什么,到底做了什么,未来到底怎么做,朝哪个方向前进,这些问题需要长久的探求。因此,中国的法律人工智能依然

[38] 关于法律知识图谱构建的技术瓶颈,参见王禄生:《司法大数据与人工智能开发的技术障碍》,《中国法律评论》2018年第2期,第50页。

有漫长而艰辛的道路需要跋涉。笔者提倡并已经尝试,通过引入统计学、计算机科学的人才从根本上改变法学院的基因,并通过开设法律大数据、法律人工智能等相关课程,实现真正的跨学科的互动与交流。唯有如此,才可能真正迎来大数据和人工智能在中国法律界光辉灿烂的未来。

[**Abstract**] In the field of legal artificial intelligence, there is a “cold and hot difference” between discourse and practice, and between China and foreign countries. China pays more attention to artificial intelligence than many other countries at the levels of top-level design, official policy, and academic research. However, far less attention is paid to this issue at the level of application effects in judicial practice. In foreign countries, especially in the United States, although the support and radiation at the government level seem to be far weaker than those in China, there are many bright spots at the level of local practice. The difference at the discourse level results from the differences in the understanding of artificial intelligence between China and foreign countries. The difference at the practical level results from the differences in preparatory conditions, research methods, adequacy of legal data, and the technical bottleneck of legal artificial intelligence. In the future, China should attach more importance to and continuously strengthen the practical ability of legal artificial intelligence and fundamentally change the “DNA” of law schools by introducing statistical and computer science talents, and offering such courses as legal big data and legal artificial intelligence, so as achieve true interdisciplinary interaction and communication. ?

(责任编辑:雨 沐)