

热与冷：中国法律 人工智能的再思考

□ 左卫民

四川大学 法学院 四川 成都 610044

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，可以说尚未有一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。在司法领域，人工智能的运用尝试更是如火如荼。2016年11月17日，在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指出：“将积极推动人工智能在司法领域的应用。”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。

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。

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，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的突破。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运用效果似乎并不理想，一些司法产品甚至被长期搁置以至于难觅踪迹。

（二）域内与域外的冷热

相较于域外，中国在顶层设计方面对人工智能的重视程度远超世界其他国家。但涉及具体的、深度的运用，却又是“上热下冷”，呈现出冰火两重天。而在域外，尽管政府层面的支持力度及辐射规模似乎远不及中国，但地方实践却大相径庭。

可见，国内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。

二、冷与热：原因何在

在对待与运用法律人工智能的态度与实践中，为何呈现出明显的“冷热不同”？进一步而言，影响法律人工智能话语与实践冷热的因素究竟为何？

（一）话语层面为何“冷热不同”

笔者认为，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。尽管域内域外均将人工智能视为下一个社会变革与经济成长的引擎，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”。而中国要明显积极得多，明显是一种乐观主义。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方面设置了较多限制。相反，我国不少人对于人工智能并未怀着像域外一些人那样如此之深的恐惧，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始终抱持乐观的心态。而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推进，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”，但是也正因如此，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之久，而中国在这方面则是刚

刚起步。

2. 研究方法上的差异。首先,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人文化、抽象化、思辨化的特征,而缺少实证化、技术化的研究方法引入。其次,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。国外法学界的研究范式正在发生根本性的改变——法学教授开始努力学习数学、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所接受。由此可见,我们缺少基础性的工具与思想支撑。因此,法学界迫切地再次需要一场法学知识的范式革命,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,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。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,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。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,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。

由此可以看出,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,并且目前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。

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司法数据在总量上比之域外

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,这为我们提供了极为重要的司法数据来源。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。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。

当然,尽管中国在裁判文书网建设方面效果突出,甚至许多域外学者通过中国的裁判文书网研究相关司法问题。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。此外,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。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。因此,在数据具有结构性缺陷、数据并非关键数据、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来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,我们应当保持充分的冷静与警醒。

4. 法律人工智能的技术瓶颈。人工智能的本质问题是围绕数据、算法、模型展开的技术问题,人工智能科学更多地应被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。目前市面上所使用的技术绝大多数是一种通用的共性技术,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。

综上,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,背后映射出深刻的社会、法律、思想上的差异。在未来,中国的法律人工智能要由冷转热,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的所有领域都由冷变热,由热变实。因此,中国的法律人工智能依然有漫长而艰巨的道路需要跋涉。

■ 《环球法律评论》2019年第2期,约20000字