



王竹, 四川大学二级教授, 博士生导师, 国家级重要人才计划青年项目入选者。四川智慧社会智能治理重点实验室主任, 国家语言文字关键研究领域领航计划“网络空间语言治理”创新团队负责人, 数据安全防护与智能治理教育部重点实验室副主任, 四川大学国家智能教育治理实验基地副主任, 兼任中国人民大学民商事法律科学研究中心法治大数据研究所所长, 最高人民法院网络安全和信息化专家咨询委员会委员、国家政法智能化技术创新中心专家委员会委员。

人工智能辅助司法审判决策研究

王竹¹, 曾航¹, 翁洋², 田维³, 冯端宇⁴, 王皓², 胡兵⁵

(1. 四川大学法学院, 成都 610207; 2. 四川大学数学学院, 成都 610065; 3. 西南石油大学法学院, 成都 610500; 4. 四川大学计算机学院, 成都 610065; 5. 成都理工大学数学科学学院, 成都 610059)

摘要: 我国人工智能辅助司法审判决策的应用发展经历了数字化与网络化、前大语言模型时代的智能化和大语言模型时代的智能化三个阶段。前大语言模型时代的技术路线主要包括定量计算、专家系统、机器学习、深度学习和可解释性增强。大语言模型时代的技术路线主要包括基于大语言模型的法律领域后训练技术、法律推理增强技术和多智能体协同技术。大语言模型辅助司法审判决策的能力层级分为L0基础知识层级、L1实务能力层级和L2细分领域层级。当前, 该领域发展仍面临法律思维模拟、可解释性困境和高质量数据可计算性三大宏观瓶颈, 并受制于数据要素供给双重制约、事实捏造风险及技术依赖与算法偏见三大微观挑战。未来, 生成式大语言模型技术的迭代升级有望系统性解决三大宏观瓶颈。依托“库网融合”高质量的权威案例集, 并借鉴通用领域大语言模型技术, 人工智能将提供更稳健、更辩证的司法审判决策辅助。最终, 通过建立标准化与法治化的治理体系, 强化准入、存证与审计标准, 将确保司法人工智能在法治框架内稳妥落地。

关键词: 人工智能辅助; 司法审判决策; 大语言模型时代; 能力层级

中图分类号: O29 **文献标志码:** A **DOI:** 10.19907/j.0490-6756.260187

Artificial intelligence-assisted judicial adjudication decision-making

WANG Zhu¹, ZENG Hang¹, WENG Yang², TIAN Wei³,
FENG Duanyu⁴, WANG Hao², HU Bing⁵

(1. Law School, Sichuan University, Chengdu 610207, China; 2. School of Mathematics, Sichuan University, Chengdu 610065, China; 3. Law School, Southwest Petroleum University, Chengdu 610500, China; 4. College of Computer Science, Sichuan University, Chengdu 610065, China; 5. School of Mathematical Sciences, Chengdu University of Technology, Chengdu 610059, China)

Abstract: The integration of artificial intelligence (AI) into judicial adjudication decision-making in China has undergone a progressive evolution. This trajectory can be divided into three major stages: Digitization and

收稿日期: 2026-06-02 修回日期: 2026-07-01 接受日期: 2026-07-23

基金项目: 国家重点研发计划(2025YFC3308700); 国家自然科学基金(62572331); 四川省重点研发项目(2025YFMS0005); 国家语言文字关键研究领域领航计划“网络空间语言治理”创新团队(LH24TD07)

作者简介: 王竹(1981—), 男, 四川成都人, 教授, 博士生导师, 主要研究方向为司法人工智能. E-mail: wangzhu@scu.edu.cn

通信作者: 胡兵. E-mail: hubingscu@scu.edu.cn

networking, intelligence in the pre-large language model (LLM) era, and intelligence in the LLM era. During the digitization stage, beginning in the late 1990s, judicial informatization focused on building infrastructure, centralized data management, and online transparency platforms such as the China Judgments Online and the China Trial Process Information Disclosure Network. These foundational efforts provided the essential data resources for subsequent AI applications. In the pre-LLM era, spanning roughly from 2017 to 2022, AI technologies such as expert systems, machine learning, deep learning, and blockchain were widely adopted in tasks including speech-to-text transcription, automated drafting of legal documents, case similarity retrieval, and electronic evidence authentication. While these applications improved efficiency, they also revealed limitations such as knowledge acquisition bottlenecks, algorithmic opacity, and difficulties in handling complex legal reasoning. The emergence of LLMs, exemplified by ChatGPT and its derivatives, has marked a paradigm shift. In the LLM era, judicial AI is no longer confined to perception and prediction but is increasingly capable of cognitive reasoning and structured decision support. Current mainstream technical approaches fall into three categories: Domain-specific post-training for injecting professional legal knowledge, legal reasoning enhancement strategies represented by the combination of legal syllogism and chain-of-thought prompting, and multi-agent collaboration frameworks designed to simulate complicated judicial interaction scenarios. A capability hierarchy has been proposed to classify LLM-assisted adjudication into three levels: L0 fundamental knowledge, L1 practical capabilities, and L2 specialized domain expertise. Despite these advances, the field continues to face three macro-level bottlenecks, namely the simulation of legal thinking, explainability dilemmas, and computability of high-quality data, alongside micro-level challenges including limited data supply, risks of factual fabrication, and algorithmic bias. Looking forward, iterative upgrades of generative LLMs will provide viable solutions to these inherent limitations. Retrieval-augmented generation (RAG), domain-specific fine-tuning, and reinforcement learning are expected to enhance reliability and logical rigor. By leveraging authoritative case repositories through “Case Database-Answers Database Integration” and adapting techniques from general-purpose LLMs, AI systems can provide more robust, dialectical, and transparent assistance to judges. Ultimately, the sustainable deployment of judicial AI requires the establishment of standardized governance frameworks that reinforce access control, data preservation, and auditing mechanisms. Such measures will ensure that AI-assisted adjudication operates prudently within the boundaries of the rule of law, balancing efficiency gains with fairness, accountability, and legitimacy. This study thus provides a comprehensive overview of the historical development, technical evolution, capability hierarchy, and future prospects of AI-assisted judicial adjudication in China, offering theoretical insights and practical guidance for interdisciplinary research at the intersection of law and computer science.

Keywords: artificial intelligence-assisted; judicial adjudication decision-making; LLM era; capability hierarchy

随着数字时代到来,大数据和人工智能技术正深刻重塑司法审判的形态,传统审判流程发生了深刻的质变和广泛的突变^[1-3]。近十年来,随着社会结构转型和公民权利意识的不断加强,我国法院受理的案件规模呈现出爆发式增长态势,导致法官人均办案数量急剧攀升,“案多人少”的矛盾日益突出^[4]。面对快速增长的案件数量与相对有限的司法资源之间的矛盾,利用人工智能辅助司法审判,进而“向信息化要人力、要效率”已成为缓解办案压力、优化资源配置的必然趋势^[1]。人工智能系统凭借其在海量数据筛选、逻辑分析、信息

记忆和高速总结归纳等方面的技术优势^[5],不仅能够有效降低司法成本、大幅提升事实查明和案件审判的效率^[4, 6],更能够辅助法官规范自由裁量权,克服个案裁判过程中的主观不确定性,统一裁判尺度,有效促进“类案类判”与“同案同判”目标的实现^[4-5, 7]。因此,如何利用人工智能技术深度赋能司法实践、提升事实查明效率和裁判的一致性,目前已经受到了法学界与计算机科学界的广泛关注与积极探索^[8-9]。

早期的司法辅助系统多依赖基于符号主义的专家规则,或者传统自然语言处理(Natural Lan-

guage Processing, NLP)技术中的机器学习技术和深度学习技术^[10]。此类技术在应对复杂司法场景时暴露出明显的局限性。一方面,其发展往往遭遇知识获取瓶颈,且在处理超长卷宗、海量异构法律文本时面临极高的数据标注成本与困难^[11-12]。另一方面,传统深度学习算法存在严重的算法黑箱与过拟合问题,运算机理多为提取要素间的统计相关性,缺乏逻辑上的因果关联,因而难以胜任严密的法律逻辑推理与深度的阅读理解^[10, 12]。虽然有学者提出了基于法律逻辑思维变体的“准三段论”^[13]构想,以解决知识获取、算法黑箱等问题,但基于当时的理论方法和技术水平,落地仍面临困难。

近年来,随着计算机技术的飞速演进,以 ChatGPT 为代表的大语言模型 (Large Language Model, LLM) 等生成式人工智能的强势崛起,为破解上述痛点提供了全新的技术路线^[14-15]。紧跟这一趋势,生成式人工智能正被深入探索应用于法律检索、法律文书生成以及辅助决策等核心环节^[16-18]。这一技术跨越使得司法审判辅助系统正从简单的流程自动化与感知智能迈向深度逻辑辅助的认知推理,进而向更高水平的数智正义迈进。

目前,围绕 LLM 的司法应用已涌现出大量前沿研究^[19-21]。在模型研发上,学界与业界相继推出了各类法律专属 LLM,如 ChatLaw^[22]、Lawyer LLaMA^[23]、InternLM-Law^[24] 以及 DISC-LawLLM^[25] 等。同时,为了缓解法律幻觉并提升回答的可靠性,旨在实现生成内容与外部法律知识库精准对齐的检索增强生成 (Retrieval-Augmented Generation, RAG) 被广泛应用^[21-22, 26-27]。在推理机制上,研究者正致力于增强 LLM 的复杂决策能力,特别是将法学理论中的司法三段论 (Legal Syllogism)^[28-31] 与人工智能领域的思维链 (Chain of Thought, CoT)^[32-33] 相结合,以期实现从案件事实到法律适用的严密逻辑推理。此外,各类中文法律评测基准 (Benchmark) 也密集发布,以客观检验人工智能的司法能力,如涵盖评估综合法律认知的 LawBench^[34] 和 LexEval^[35]、聚焦三段论逻辑层级的 LAiW^[20]、针对多阶段法律文书生成的 CaseGen^[36]、以及面向多轮动态司法交互的 Ready Jurist One^[37] 和 LexRAG^[26] 等。

现有的相关综述文献多聚焦于单一 NLP 任务

或特定技术层面,未能充分结合法学本体与人工智能的演进^[38]。目前仍缺乏从人工智能辅助司法审判决策的实务现状到底层技术和能力层级的系统性综述。因此,本文旨在填补这一空白,为法学与计算机科学的跨学科研究提供理论参考与实践指引。

1 人工智能辅助司法审判决策的应用发展阶段

纵观我国司法信息化的发展脉络,人工智能在司法审判决策中的应用并非一蹴而就,而是伴随着国家信息化战略的推进以及技术的迭代升级逐步深化的。结合我国法院信息化建设的发展规划与技术演进特征,我国人工智能辅助司法审判的历史进程大致可以划分为数字化与网络化、前大语言模型时代的智能化、大语言模型时代的智能化三个主要阶段。图1以时间轴时序梳理了我国人工智能辅助司法审判发展的三大递进阶段,以及各阶段时间区间、建设核心与技术特征的逐层迭代。

1.1 第一阶段:数字化与网络化阶段(1996—2017)

这一阶段是我国司法人工智能发展的萌芽与基础设施建设期,法院信息化建设主要经历了以司法统计和电脑办公为要素的“1.0阶段”(1996—2012)和以办公自动化、网络化互联互通为特征的“2.0阶段”(2013—2015)^[39-40]。其核心诉求是实现审判执行、司法人事和司法政务数据的集中管理与互联互通。在这一时期,我国初步建成了各地法院内部网络及智能化基础设施,并相继建成了“中国裁判文书网”“中国庭审公开网”“中国审判流程信息公开网”和“中国执行信息公开网”等四大司法公开平台。这些基础性的数据化与网络化建设有效解决了跨地域沟通和基础信息公开的问题,并为后续司法人工智能的深度学习与模型训练积累了极其宝贵的海量原始数据。

1.2 第二阶段:前大语言模型时代的智能化阶段(2017—2022)

2017年是我国司法人工智能发展的一个关键分水岭,最高人民法院印发《关于加快建设智慧法院的意见》,标志着我国司法信息化迈入以网络化、阳光化、智能化为主要特征的“3.0时代”。这

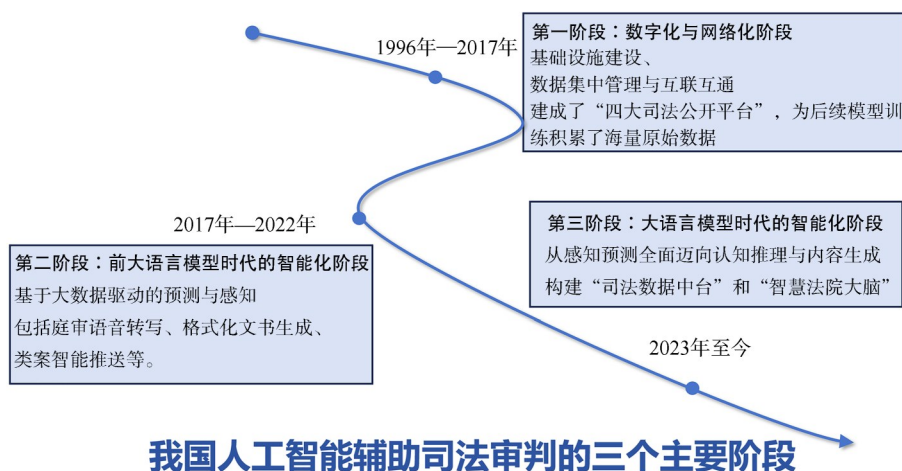


图1 我国人工智能辅助司法审判发展的三大递进阶段

Fig. 1 Timeline illustrating the three progressive stages of AI-assisted judicial adjudication development in China

一阶段,人工智能技术在司法中的应用呈现爆发式增长,其技术内核主要依赖于专家系统、传统NLP与机器学习,属于大数据驱动的预测与感知。典型应用主要包括庭审语音同步转化为文字生成笔录、自动提取案件要素生成格式化文书初稿、基于相似度的类案智能推送以及裁判偏离度预警等^[5-6, 40]。同时,区块链技术也在这一时期开始被广泛应用于电子证据的存证与验真^[7]。在这一阶段,人工智能主要扮演辅助法官办案的角色,学界也开始深刻反思此类技术可能带来的算法黑箱、算法歧视以及法官主体性被技术虚化等法理风险^[8, 12]。

1.3 第三阶段:大语言模型时代的智能化阶段(2023至今)

随着2022年底至2023年初以ChatGPT为代表的LLM技术爆发,生成式人工智能的出现使司法人工智能迎来了代际跨越,技术能力正式从感知与统计预测全面迈向深度的认知推理与内容生成。与此同时,我国法院信息化建设也顺势迈入以智能化、一体化、协同化等为特征的“4.0阶段”^[39]。根据《最高人民法院信息化五年发展规划(2021—2025)》和《人民法院第六个五年改革纲要(2024—2028)》,当前我国法院信息化建设的重心已转向构建“司法数据中台”和“智慧法院大脑”,以全面拓展数据和知识服务^[41]。随着我国法院信息化建设在“4.0阶段”稳步推进,2025年12月1日最高法启用全国法院统一的电子诉讼服务平台,标志着全国法院“一张网”的对外服务功能正式全面开放。

2 人工智能辅助司法审判决策的技术发展历程

需要指出的是,无论是前大语言模型时代还是大语言模型时代,通用人工智能技术向司法领域转化时其初期技术红利都呈现出快速消退的规律。早在前大语言模型时代,通用人工智能技术稍加优化调整后被直接应用到司法领域就能产生一定智能化辅助效果的技术红利快速消退的规律^[13]就已经被发现。进入大语言模型时代,这一规律依然显著。在技术应用初期,仅需对通用LLM进行法律领域的预训练或微调即可获得具备基础法律知识的LLM。然而,使其真正掌握法律逻辑、满足实践需求仍需克服AI幻觉等核心挑战。当前探索的法律推理框架、多智能体(Multi-Agent)协同等技术路径虽然带来了LLM法律能力的一定突破,但整体效果仍远未达到理想状态。

2.1 前大语言模型时代的人工智能辅助司法审判决策技术

在以ChatGPT为代表的生成式LLM爆发之前,法律人工智能领域经历了从早期的定量计算到深层语义建模的长期探索。这些技术演进不仅推动了法律文本的结构化处理,也为判决预测、智能化检索、可解释性分析等研究奠定了理论与技术基础。

2.1.1 早期的定量计算

早期的研究深受计量法学(Jurimetrics)思想影响,试图通过数学公式捕捉司法决策的潜在规律。研究者们主张跳出传统的法学推理逻辑,

转而通过识别影响判决的关键事实要素并赋予其数值权重来进行预测。早在上世纪五十年代,研究者就开始尝试利用定量方法预测美国最高法院的判决。例如,针对获得律师帮助权(Right to Counsel)^[42]案件的研究表明,即便不考虑大法官的个人背景,仅依靠对事实要素的定量分析和公式推导也能实现高准确度的判决预测。也有学者引入了行为科学概念^[43],通过分析司法决策与决策前情境因素之间的概率关系挑战了传统的遵循先例原则,为司法过程提供了更具预测性的视角。同时,也出现了利用数学模型预测司法判决的构想^[44]。这一时期的研究或许不能被称为法律人工智能,但已经出现了利用数理方法解决实际司法问题的思想萌芽,因而我们可以称之为法律人工智能的萌芽。

2.1.2 基于规则、本体与专家知识的符号主义

随着计算机技术的发展,部分研究者开始利用NLP技术和本体论框架(Ontological Framework)^[45]来处理法律文本的复杂语义。针对立法文本的结构化需求,研究者开发了名为SALEM^[46]的NLP架构,通过语义标注和索引技术实现对意大利法律段落的信息提取。在法律案件智能检索方面,Zeleznikow等开发了一种知识表示模型^[47],并基于该知识表示模型提出了用于智能法律案件检索的IPN算法,旨在提供高效的法律案件管理。为了克服词语的同义性和多义性,有研究者提出了基于本体论的搜索方法^[45],本体论通过利用在知识库构建过程中收集的领域知识进行推理,确保高效的检索结果。

2.1.3 机器学习

机器学习技术的引入使法律人工智能从硬性规则匹配转向了基于统计特征的学习。部分学者利用机器学习技术进行了危险驾驶罪刑期预测研究^[48],采用经典的概率图模型(Probabilistic Graphical Model),并通过极大似然准则估计参数来计算刑期分布的数学期望。在处理非结构化诉讼档案方面,有研究者在检测美国联邦地区法院案件结果时将案件建模为条件随机场(Conditional Random Field)^[49],获得了显著优于传统的逻辑回归方法的性能,能更有效地处理复杂的诉讼档案序列。机器学习也被应用于判决书摘要(Headnote)的预测^[50]。通过结合线性支持向量分类(Linear Support Vector Classification)与主动学习

机制,系统能够自动预测法律要点并生成判决概要,辅助法律专业人士快速掌握案件核心。

2.1.4 深度学习

深度学习(Deep Learning)的兴起为处理海量法律文书和复杂任务提供了强有力的工具。诉讼档案蕴含着规划诉讼策略的丰富信息,这些信息通常被锁定在半结构化的文本中。深度学习模型能够从数百万份半结构化的诉讼档案中自动提取案件结果,揭示法官及法院的倾向性,为诉讼策略的制定提供数据支持^[51]。针对刑事案件预测,多通道注意力神经网络(Multi-Channel Attention Neural Network, MANN)^[52]等模型通过词级和句级的分层编码,实现了罪名、适用条文及刑期的统一框架预测。此外,针对长文本和多语言环境,基于分层双向编码器表示模型(Bidirectional Encoder Representations from Transformers, BERT)^[53]在瑞士联邦最高法院的数据集上表现出了卓越的跨语言预测能力。在法律要素感知方面,为了应对法律概念的模糊性,研究者提出了基于强化学习(Reinforcement Learning, RL)的犯罪要素提取网络(Crime Element Extraction Network, CEEN)^[54],通过精准提取犯罪主体、对象、主观意图和行为四种要素,有效缓解了事实描述相似导致的误判问题。针对传统深度学习模型难以处理刑期复杂分布的问题,研究人员构建了多尺度异构图注意力网络(Heterogeneous Graph Attention Network, HGAN)^[55],利用深度图神经网络强大的表征能力来捕捉量刑要素间的复杂关系。在司法命名实体识别任务中,由于人工标注成本高昂导致标注数据稀缺,研究者提出了一种基于深度学习的半监督模型,采用改进的“三训练”(Tri-training)算法^[56],以充分利用未标注数据来提升模型的预测性能。

2.1.5 可解释性增强

由于法律判决涉及社会公平正义,模型的可解释性成为该阶段后期的研究焦点。在法律判决预测领域中,为了增加司法人工智能的可解释性,研究人员在量刑要素中融入了特定犯罪的先验知识,并且使用符合人类决策行为的贝叶斯网络概率推理框架^[57]。针对法律案例匹配任务,研究者提出了基于逆最优传输(Inverse Optimal Transport, IOT)可解释的方法 IOT-Match^[58]。该方法不仅能提高匹配精度,还能从成对案件中提取语

义和法律特征层面的理由,为匹配结果提供可信的解释。前述的 CEEN 框架^[54]同样在可解释性上有所贡献。它通过将模型关注点与法学理论中的犯罪构成要素对齐,使模型的预测过程更符合法律从业者的思维逻辑,从而增强了人工智能系统在司法实践中的说服力。

2.2 大语言模型时代的人工智能辅助司法审判决策技术

进入大语言模型时代,人工智能辅助司法审判不再局限于传统的文本分类与信息检索,而是向着具备深度认知、严密逻辑与自主交互的决策支持方向演进。当前,相关技术主要围绕以下三个核心维度展开:一是通过领域适配的后训练技术夯实法律专业基础;二是借助推理增强策略提升裁判过程的严谨性与可解释性;三是利用 Multi-Agent 协同机制模拟真实司法场景中的复杂博弈与协作。

2.2.1 基于大语言模型的法律领域后训练技术

随着大规模语言模型在通用语义任务中展现出卓越性能,面向法律这一高度专业化、逻辑严密且具有强约束性的垂直领域,研发专用法律 LLM 已成为司法人工智能发展的必然选择。此类模型的核心逻辑在于,通过在通用基座模型基础上注入海量法律领域语料,实现从通用认知能力向司法专业逻辑的自适应迁移。

在模型选型方面,Qwen、LLaMA、Baichuan 及 ChatGLM 是当前司法研究与实务中的主流基座模型。关于基座的选定,学术界与工业界并无单一标准,其决策过程实质上是针对模型开源生态、参数规模分布、多轮对话性能以及计算资源限制与具体司法任务适配度之间的多维度权衡。

仅凭通用智能尚不足以应对复杂的司法决策。为赋予模型深厚的法律知识储备与严密的法律推理能力,研究者普遍采用持续预训练(Continuous Pre-Training)、指令微调(Supervised Fine-Tuning, SFT)及 RL 的渐进式训练管线。值得注意的是,这些通用技术在落地司法场景时必须结合法学理论的特质,如三段论推理、法条引用的时效性和准确性等,进行深度的领域适配,以弥补通用模型在法学认知层面的知识缺位与逻辑偏差。

1) 持续预训练。为了应对通用 LLM 在处理司法任务时普遍存在的知识空缺、法律幻觉等瓶颈,研究者广泛采用大规模法律领域专用语料对

基座模型进行增量训练,旨在通过参数层面的知识注入,重塑模型的司法语义分布。部分研究者致力于系统化法律知识的深度习得,此类研究侧重于构建覆盖法律体系全分支的语料库,以实现法律概念的准确映射。Lawyer LLaMA^[23, 59]率先提出在 LLaMA 基础上持续注入中国法律领域语料,通过系统学习中国的法律知识体系使其能够正确理解民法、刑法、行政法、诉讼法等常见领域的法律概念,并能够以通俗易懂的语言解释法律概念进行基础性法律咨询,涵盖婚姻、借贷、海商、刑事等法律领域。智海-录问^[60]则基于 Baichuan-7B 构建,利用约 40 GB 的法律文书、司法案例与法律问答数据进行二次预训练,从底层数据层面赋予模型深厚的司法知识储备。另一个思路是通过高强度训练与参数演化路径赋予 LLM 更强的法律能力。在训练强度与模型架构的探索上,不同项目展现了差异化的演进策略。HanFei(韩非)^[61]作为国内首个采用全量参数训练方式构建的法律 LLM,通过对 7B 规模参数的全面更新,显著提升了模型在复杂法律问答与多轮交互中的语义一致性与稳定性。相比之下,TaiLing^[62]则依托 Qwen-7B 底座,经历了超过 2 万亿 token 的超大规模预训练,在海量数据的基础上实现了模型在文本理解、生成与决策支持方面的优异表现。这些从数据规模与参数更新策略出发的差异化路径共同推动了国产法律 LLM 向高性能、高稳定性方向发展。

2) 多样化指令微调。在具备领域知识后,如何通过高质量 SFT 将知识转化为解决具体司法任务的能力,并赋予其法律推理逻辑是当前研究的核心。Qzhou-Law^[63]深入探索了 SFT 阶段的规模效应。该研究构建了一个包含超 85 万条高质量法律指令的大规模数据集,并创新性地利用基于法条的 IRAC(问题-规则-应用-结论)技术对数据进行增强。实验结果表明,在司法垂直领域,增加模型尺寸能持续带来显著的性能提升,尤其是在处理国家统一法律职业资格考试等复杂推理任务时表现优异。此外,为了强化 LLM 的司法行为模式,DISC-LawLLM^[25, 64]采用“法律三段论”提示策略构建了微调数据集,显著增强了模型的法理推理能力,图 2 展示了检索增强版 DISC-LawLLM 的工作流程。夫子·明察(fuzi. mingcha)^[65]同样强调结构化的推理链条,通过自动识别案件事实与法律规范,生成严谨的“大前提-小前提-结论”判决推

导过程。另一部分研究致力于高质量数据合成与优化,獬豸(LawGPT_zh)^[66]采用 Reliable-Self-Instruction 方法,结合 20 万条情景对话与 9 万多条带有明确法律依据的问答数据,提升了模型引述

法条的可靠性。InternLM-Law^[24, 67]则通过两阶段微调策略,先注入法律知识,再进行高质量特定任务微调,优化了模型的结构化输出。

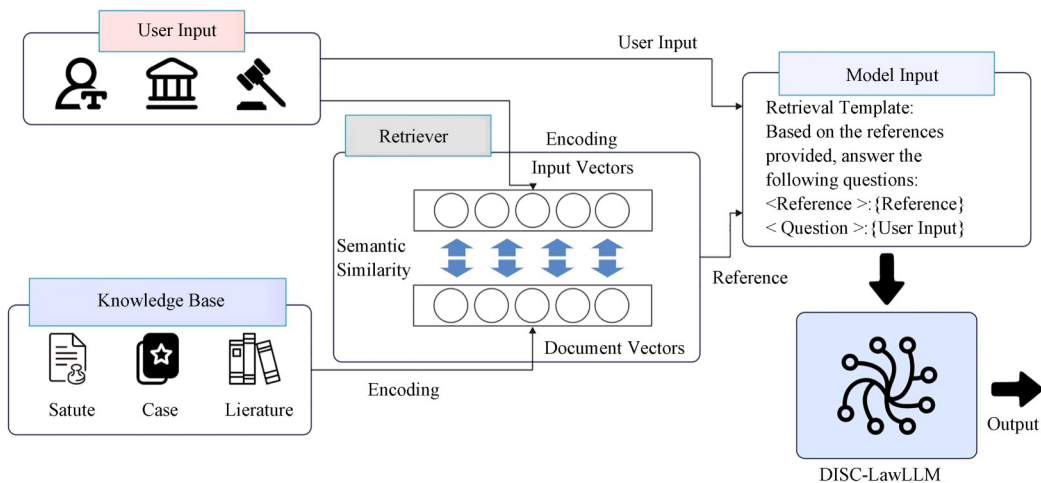


图 2 检索增强版 DISC-LawLLM 模型框架图(来源:基于参考文献[25]重新绘制)

Fig. 2 Framework of the retrieval-augmented DISC-LawLLM model (Source: Redrawn from Ref. [25])

3) 强化学习。RL 可以通过奖励机制提升模型在复杂法律场景中的表现。针对高精度的检索与司法任务, LexPro-1.0^[68]在监督微调的基础上引入了大规模无额外监督的 RL, 利用类似于 DeepSeek-R1 的 CoT 机制, 重点提升了模型在处理复杂检索任务时的逻辑深度与证据链推导能力。

4) 业务特化。除了综合性法律模型外, 针对特定司法环节的专项 LLM 亦展现出重要价值, 如事实认定、证据核实环节等。针对法律领域的以结构化格式呈现的表格数据, HARMONIC^[69]基于 LLM 有效生成了既保护隐私又与原始数据相似的合成表格数据, 为从敏感领域获取表格数据提供了有力方案。Legal Fact-Finding Model^[70]针对事实认定这一核心环节, 整合了 T5 与 LexiLaw 模型, 通过文档分割、双模型并行处理及结果融合机制实现了从海量证据中自动提炼案件事实的功能。大量实验表明, 该模型在 ROUGE 指标上较基线模型提升了 10.27%, 为缓解法官办案压力、提升审判效率提供了有力工具。

法律 LLM 的百花齐放不仅为司法领域的问题解决提供了初步方案, 更通过参数化知识注入与逻辑强化, 为后续的法律推理框架、外部知识增强及检索验证机制奠定了坚实的技术基础。

2.2.2 基于大语言模型的法律推理增强技术

司法裁决对逻辑严谨性与过程可解释性有着极高的要求。针对通用 LLM 在复杂法律场景中易出现逻辑跳跃或直接给出结论而忽略论证过程的快思考问题, 研究者提出了多种增强策略, 旨在引导模型从简单的语义生成向深度逻辑推理跨越。

1) 基于“司法三段论”的显式逻辑框架。司法三段论是法理学中最经典的推理模型。在提示工程或模型架构中引入“大前提(法律规定)—小前提(案件事实)—结论(判决结果)”的结构化约束能够显著提升 LLM 推理的透明度与准确性。Legal Syllogism Prompting^[29]通过分析深度学习、LLM 的 CoT 技术进行法律推理的不足, 提出了法律三段论逻辑推理链, 如图 3 所示。通过简单的提示词教导模型遵循三段论逻辑, 验证了在无须微调的情况下, 仅靠逻辑框架的约束即可显著提高判决预测的解释力。为了解决推理过程不透明的问题, SyLeR^[31]框架提出了一个显式的三段论推理体系。该框架采用树状分层检索机制构建大前提, 并通过有监督预热结合基于结构感知奖励机制的 RL 两阶段微调强制模型生成合乎逻辑的显式推理路径, 从而在多语言和跨领域场景下均表现出极高的可信度。

2) 强化学习驱动的高级推理对齐。近年来,

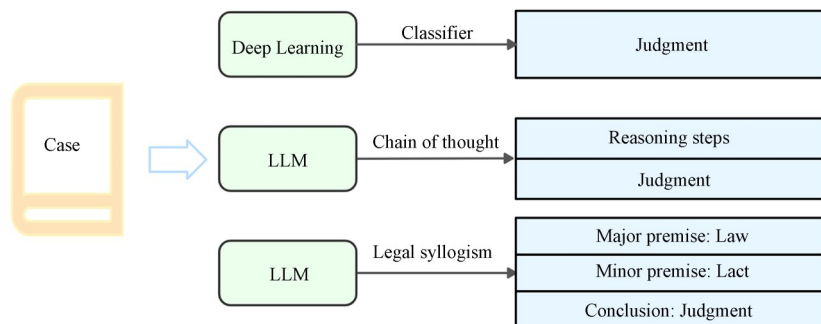


图 3 三种不同的法律判决预测方法(来源:基于参考文献[29]重新绘制)

Fig. 3 Three different methods for legal judgment prediction (Source: Redrawn from Ref. [29])

借鉴 DeepSeek-R1 等推理 LLM 的 CoT 技术,法律领域开始探索利用 RL 实现模型输出与法律程序逻辑的深度对齐。Legal Δ ^[32]针对模型倾向于直接给出答案的缺点,提出了基于 CoT 引导信息增益的 RL 框架。该框架从强推理模型中蒸馏逻辑能力,通过最大化直接回答与推理增强两种模式间的信息增益,促使模型习得深刻的推理模式而非肤浅的解释,其多维奖励机制能有效评估逻辑的连贯性与领域特异性。

3) 针对特定司法任务的推理优化。除了通用推理增强方案,研究者还针对司法实务中的核心任务,如法律案件中的数学推理、文书生成、文本摘要等,设计了专项推理框架。在涉及赔偿计算、刑期折抵等高风险法律场景时,输出必须兼具法律合规性与计算准确性。LexPam^[71]框架通过两阶段 RL 解决了这一难题。框架利用课程学习(Curriculum Learning)策略筛选挑战性数据,并采用组相对策略优化(GRPO)算法在无须高成本价值网络的前提下引导模型完成基于法律 grounding 的程序对齐,确保推理过程既符合计算逻辑又符合法律强制性步骤。刑事司法意见书要求判决理由与量刑结论高度统一,LegalChainReasoner^[72]引入了结构化的“前提-情形-结论”法律逻辑链,打破了以往法律推理与量刑预测相互割裂的局限。通过灵活的知识注入与端到端的生成路径,该框架确保了文书说理部分与最终判决结果的一致性。针对法律文书篇幅冗长、术语密集的特点,三段论驱动的摘要框架(Legal Text Summarization)^[30]被用于引导模型显式模拟法官的思考过程。通过提取大前提(适用条款)与小前提(具体事实),模型生成的摘要不仅能捕捉核心法律逻辑,更在可解释性和逻辑连贯性上优于传统的 NLP 方法。

2.2.3 基于大语言模型的多智能体协同技术

随着智能体(Agent)技术的成熟,研究者开始通过模拟法庭辩论或多角色协作来增强复杂法律问题的处理能力。受真实法庭审判启发,Debate-Feedback^[19]架构引入了 Multi-Agent 辩论机制与可靠性评估模型。该模型通过模拟控辩双方的动态博弈极大程度上减少了对大规模历史数据集的依赖,提供了一种轻量且稳健的动态推理解决方案。LegalGPT^[33]则提出了一种基于大规模法律语言建模的 Multi-Agent 框架,其核心在于引入了“法律思维链(Legal Chain of Thought, Legal CoT)”和专用 Agent。该框架利用特定功能的逻辑链条引导模型进行案件审查与司法裁判,并整合信息检索模块以消除幻觉,在复杂法律难题的应对上展现出优于单体模型的准确性与完整性。图 4 展示了 LegalGPT 的工作流,从上到下有三种法律任务:法律审查、法律咨询和司法判决。当用户提出问题,模型会采用这三种 Legal CoT 中的适当功能来分析问题并最终生成答案。

从通用模型的法律知识注入,到基于法律推理的框架设计,再到前沿的 Multi-Agent 法律辩论,法律领域的 LLM 技术发展日益革新。上述增强策略的持续演进不仅清晰地勾勒出司法 LLM 能力提升的技术脉络,更标志着法律领域人工智能的范式转变。

3 大语言模型辅助司法审判决策的能力层级评估

当前法律领域 LLM 在法律与司法领域研究已有多项技术创新,赋予了 LLM 更广泛的法律知识和更优秀的法律素养。基于当前的前沿评测基准,可以更加量化地明确 LLM 在法律领域的能力边界以及未来的研究方向。现有评测基准对法律

LLM的要求可以划分为L0基础知识层、L1实务能力和L2细分领域层。L0基础知识层是LLM参与司法辅助决策的根基,核心在于验证模型是否具备“懂法”的基本素养。L1实务能力层是在基础知识之上的能力跃迁,聚焦于解决“能不能办

案”的现实问题。L2细分领域层致力于解决“够不够专精”的核心挑战,是对法律实务中专项任务的深度评测。图5构建了大语言模型辅助司法审判决策的三级递进能力分层体系,自下而上形成金字塔式能力逻辑,三层能力逐级进阶、层层支撑。

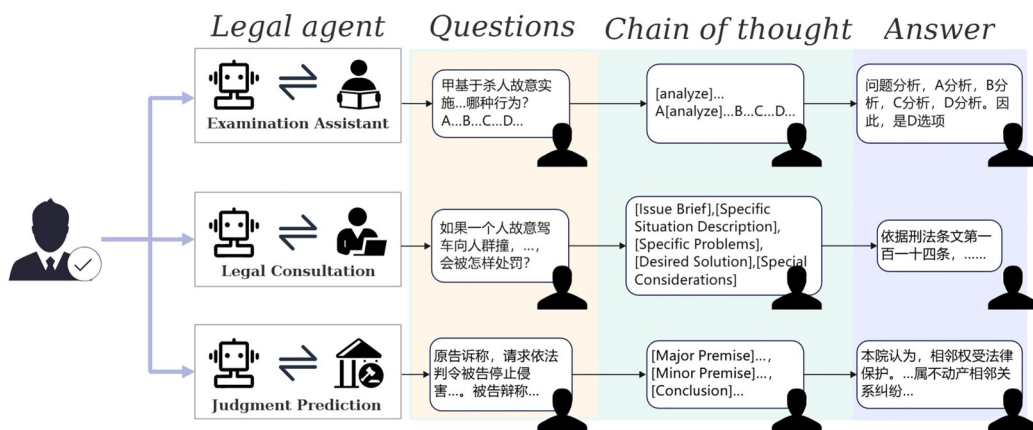


图 4 LegalGPT 的推理框架(来源:基于参考文献[33]重新绘制)

Fig. 4 Reasoning framework of LegalGPT (Source: Redrawn from Reference [33])



图 5 大语言模型辅助司法审判决策的金字塔图

Fig. 5 Pyramid diagram that categorizes the capabilities of large language models in assisting judicial adjudication

3.1 大语言模型辅助司法审判决策的L0基础知识层级

这一层级旨在解决“AI懂不懂法、推理对不对”的核心问题,全面丈量AI对法条的记忆、对法理的理解、逻辑推理能力,以及跨语种法系的认知广度。研究者引入了教育学与法学理论来构建多维度的评测体系。LawBench^[34]参考布鲁姆认知模型(Bloom's Taxonomy),将LLM能力细分为法律知识记忆、法律知识理解和法律知识应用三个递进层级,全面评估LLM在我国民法体系下的表现。LexEval^[35]则构建了目前规模最大的中文法

律评估数据集,含23个任务、超1.4万道题,提出了涵盖记忆、理解、逻辑推理、辨别、生成以及法律伦理的六维法律认知能力分类体系。LAiW^[20]进一步基于真实的“法律三段论”实务逻辑,将模型能力划分为基础信息检索、法律基础推理和复杂法律应用三个层次。在跨法系与多语种覆盖方面,VLegal-Bench^[73]针对越南成文法体系构建了从事实识别到伦理偏见的认知分级框架;Legal-Bench.PT^[74]填补了葡萄牙法律体系的空白,覆盖公法与私法分支;JBE-QA^[75]则通过日本司法考试数据集,细粒度评估LLM的日本法律领域知识。

此外,也有研究专门评估了 OpenAI o1、DeepSeek-R1 等最新推理 LLM 在法律领域的推理能力。综合来看,尽管这类模型在通用推理上表现优异,但在面对多被告案件判决预测、争议焦点提取等复杂法律推理任务时认知边界与法律推理能力仍显不足^[76]。

3.2 大语言模型辅助司法审判决策的 L1 实务能力层级

当 LLM 掌握基础知识后,需解决“能不能像律师、法官一样办案”的问题。为了检验 LLM 是否能够像律师和法官一样将这些知识应用到具体案件的办理过程中,UCL-Bench^[77]镜像了真实的法律场景,让法律专家为每项任务提供了详细的答题指导,其中融合了信息和需求要素,以模拟法律实践的复杂性,满足法律从业者 in 真实场景中的特定需求。真实的司法决策充满了信息不对称、多轮博弈与严格的程序要求,这一层级要求 LLM 适应动态交互环境并遵循专业工作流。在动态交互与咨询实战场景,Ready Jurist One (J1-Bench)^[37]模拟了法庭庭审的 Multi-Agent 动态交互,不仅评估任务结果,更通过细粒度指标严格考核 LLM 是否遵守了法庭调查、辩论等法定程序。PLAWBENCH^[78]摒弃了标准化的考试题,引入真实律师执业评分量表(Rubric)来评估 LLM 在面对大众模糊、碎片化咨询时主动发问与澄清事实的实务能力。LexRAG^[26]则专注于评估 LLM 在多轮渐进式法律咨询中的检索增强能力,测试其在不断推进的语境中动态检索法条并生成合法回复的水平。在实务规范与可信度方面,CitaLaw^[27]基准专门评估 AI 生成法律回复时提供准确“法条+判例”双重引用的能力,以满足法律实务中“字字有据”的严苛防幻觉要求。在专业工作流与文书生成方面,CaseGen^[36]突破了全文一次性生成的局限,将判决书生成拆解为撰写答辩状、查明事实、法庭推理和最终判决四个阶段,并测试多阶段实务工作流能力。AppellateGen^[79]则向最高难度的二审(上诉)文书生成发起挑战,要求 AI 具备“辩证推理”能力,通过对比一审判决和新证据来纠正事实或法律适用错误。

3.3 大语言模型辅助司法审判决策的 L2 细分领域层级

最高层级的能力聚焦于解决“大语言模型在硬核细分领域够不够专精”的问题,从中抽离出法

律实务中最核心、最需要领域专家知识的专项任务进行深度评测。在民事实体法专项中,LexRel^[80]深入我国民事法律的核心,提出了包含 265 种细分类型的综合法律关系图谱,专门评估 LLM 从非结构化案件事实中精准抽取“法律关系”及其对应主体、客体和内容的专项能力。在诉讼程序法专项中,CLER^[81]针对诉讼实务的证据推理,要求 LLM 能够根据原告的诉求和案件事实规划并结构化输出对应的证据列表及证明目的,极大地考验 LLM 构建逻辑证据链与诉讼策略的专业深度。在公检法垂直场景拓展中,CPSD-Bench^[82]将视野从法院延伸至公安执法领域,专注于评估 AI 在接处警分类、电信诈骗识别、警情通报生成等警务专项任务中的理解与生成能力,填补了公安垂直领域的评测空白。

4 人工智能辅助司法审判决策面临的挑战与未来发展方向

从技术与实践的发展脉络来看,生成式 LLM 在司法场景的落地已具备深厚的可行性。这主要体现在技术比较优势、技术迭代速度以及现有应用规模三个维度。相较于传统 NLP 技术,生成式 LLM 在司法场景的落地具备显著的比较优势。它以更自然的交互和强大的泛化能力,轻松化解了过去需耗费巨额工程成本方能解决的语义理解与判词生成难题。纵观其发展,虽历时较短,但迭代速度极快、未来空间广阔。当前,生成式 AI 在司法辅助场景中已初具规模,虽然尚未步入涉及司法价值裁量等深层次的“深水区”,但由表及里的渐进式演进正驱动着司法人工智能向复杂的实践应用迈进。司法人工智能在 LLM 技术的驱动下已取得长效进步,但在迈向深层次应用的过程中,仍需克服从宏观层面到微观层面的多重障碍。

4.1 人工智能辅助司法审判决策面临的宏观三大挑战

在生成式人工智能爆发之前,人工智能在法律领域的研发就面临法律思维、可解释性和可计算性三大瓶颈^[13]。进入大语言模型时代,尽管技术进步缓解了部分问题,但这三大挑战依然是制约法律人工智能向纵深发展的关键因素。

第一,在法律思维方面,法律领域生成式 LLM 的构建依然很大程度由人工智能学者主导,尽管

通过学习大量法律专家的文本知识,以及通过 CoT、RAG、Multi-Agent 辩论等技术也能使 LLM 在语义上模拟人类专家的法律推理,并展示思考过程,但面对复杂案件时 LLM 仍然存在推理链条本身虚假、自我矛盾、法条幻觉等问题。可以说,LLM 学到了法律思维的外壳,没有学到法律思维的内核。

第二,在可解释性方面,由于 LLM 的文本生成功能,某种意义上来说已经从“黑箱”变成了“灰箱”。基于 CoT、RAG 的引用溯源等技术,LLM 已经可以输出“为什么得出这个结论”的过程并标注来源,因而从生成的文本来讲已经不是完全的“黑箱”。当然,从底层技术上讲,LLM 的文本生成也基于概率,实质上也具有不可解释性,因而或许称之为“灰色黑箱”更为恰当。

第三,在可计算性方面,高质量数据仍然缺乏。虽然“人民法院案例库”和“法答网”形成的“库网融合”策略为 LLM 训练提供了可能的高质量数据,但目前人工智能学界还没有相关研究。作为司法改革理念与信息技术融合的制度创新,“库网融合”在统一法律适用、提升审判质效方面成效显著。但如何进一步挖掘其数据潜力仍是关键课题。

4.2 人工智能辅助司法审判决策面临的微观三大挑战

1) 数据要素供给的双重制约。一方面,外部公开数据显著缩减。近年来,“中国裁判文书网”出现明显的公开比例下降,这种存量数据的结构性萎缩导致科研团队面临严重的数据匮乏。另一方面,多维数据壁垒造成流转阻滞。法检机关之间尚未建立统一的数据接口与交换标准,行政机关与司法机关之间也有坚固的数据篱笆。对于科研团队而言,高质量的内部案卷数据由于隐私保护与安全定责的考量往往被封闭在特定的行政区域内。这种内外部来源的割裂使得法律 LLM 难以触及司法实务的核心逻辑。

2) 事实捏造风险依旧存在。尽管 RAG 等技术手段显著缓解了模型幻觉,但在处理极端复杂案情时 LLM 仍可能生成看似专业实则捏造事实的输出。在司法审判这种对幻觉采取零容忍态度的场景中,如何建立具备法律约束力的知识对齐机制仍是难题。

3) 技术依赖与算法偏见隐忧。随着“法官助手”等系统的深度介入,实践中出现了法官过度依赖算法建议的倾向。这种技术依赖可能导致审判独立性被算法潜移默化地架空。同时,由于训练数据中隐含的地域文化差异或历史性偏见可能通过模型效应被进一步放大,司法公正的统一性正面临算法歧视的潜在挑战。

4.3 人工智能辅助司法审判决策的未来发展方向

第一,未来生成式 LLM 技术的迭代升级有望系统性化解司法人工智能的三大宏观瓶颈。LLM 强大的语义理解与生成能力将有效缓解可计算性难题;类 RAG 技术通过引入权威法源校验有望进一步攻克法律思维的障碍;而基于“准三段论”的推理框架则有望打破算法黑箱,解决可解释性困境。

第二,“库网融合”策略将不断产生高质量的权威案例集。开放人民法院案例库、常态化公布“法答网精选问答”作为司法公开的创新形式,为司法人工智能领域学者科研提供了更多高质量司法数据。利用这些新类型司法数据寻求模型在复杂逻辑推演瓶颈上的突破将是中文法律推理 LLM 的重要研究路径。

第三,借鉴通用 LLM 领域 OpenAI o1 与 DeepSeek-R1 的 RL 路径,未来的法律 LLM 将不再仅仅依赖预训练知识,而是通过在大规模法律规则环境下的自我博弈与 CoT 扩展,实现对极复杂疑难案件的深度逻辑推演。依托 Multi-Agent 技术,模拟法庭式的辩论架构可能成为主流。通过构建控、辩、审等多角色 Agent,在动态交互中发现事实陷阱、对冲算法偏见,从而提供更稳健、更辩证的决策辅助。

建立标准化与法治化的治理体系是司法人工智能落地的必然归宿。可以预见的是,未来将建立更细粒度的司法人工智能准入、存证及审计标准,确保人工智能在司法的框架下运行。

参考文献:

- [1] Ji W D. The transformation of judicial power in the era of artificial intelligence [J]. J East Law, 2018(1): 125-133. [季卫东. 人工智能时代的司法权之变[J]. 东方法学, 2018(1): 125-133.]
- [2] Cai L D. Construction of smart court: Implementation principles and institutional support [J]. J China

- Appl Law, 2017(2): 19-28. [蔡立东. 智慧法院建设: 实施原则与制度支撑[J]. 中国应用法学, 2017(2): 19-28.]
- [3] Hong D Y. How the judiciary faces technological innovations such as “Internet Plus” and artificial intelligence [J]. Law Sci, 2018(11): 169-180. [洪冬英. 司法如何面向“互联网+”与人工智能等技术革新[J]. 法学, 2018(11): 169-180.]
- [4] Qian D J. The Chinese process of judicial artificial intelligence: Function substitution and structure strengthening [J]. Law Rev, 2018, 36(5): 138-152. [钱大军. 司法人工智能的中国进程: 功能替代与结构强化[J]. 法学评论, 2018, 36(5): 138-152.]
- [5] Jiao B Q, Zhao Y. The impact of artificial intelligence on judges’ thinking [J]. Seeking Truth, 2022, 49(4): 115-125. [焦宝乾, 赵岩. 人工智能对法官思维的影响[J]. 求是学刊, 2022, 49(4): 115-125.]
- [6] Hao J J. Procedural law dilemmas and solutions for internet courts [J]. Law Sci (J Northwest Univ Polit Sci Law), 2021, 39(1): 83-95. [郝晶晶. 互联网法院的程序法困境及出路[J]. 法律科学(西北政法大学学报), 2021, 39(1): 83-95.]
- [7] Liu Y H. Practical application and prospects of artificial intelligence technology in smart court construction [J]. J Comp Law, 2022(1): 1-11. [刘艳红. 人工智能技术在智慧法院建设中实践运用与前景展望[J]. 比较法研究, 2022(1): 1-11.]
- [8] Zuo W M. Some thoughts on the application prospects of legal artificial intelligence in China [J]. Tsinghua Univ Law J, 2018, 12(2): 108-124. [左卫民. 关于法律人工智能在中国运用前景的若干思考[J]. 清华法学, 2018, 12(2): 108-124.]
- [9] Feng J, Hu M. Smart justice: A new path to judicial justice and its limitations [J]. Zhejiang Soc Sci, 2018(6): 67-75, 85, 157. [冯皎, 胡铭. 智慧司法: 实现司法公正的新路径及其局限[J]. 浙江社会科学, 2018(6): 67-75, 85, 157.]
- [10] Wei B. Difficulties and paths of integrating judicial artificial intelligence into judicial reform [J]. Mod Law Sci, 2021, 43(3): 3-23. [魏斌. 司法人工智能融入司法改革的难题与路径[J]. 现代法学, 2021, 43(3): 3-23.]
- [11] Wang L S. Risks and ethical regulations of judicial big data and artificial intelligence technology application [J]. Stud Law Bus, 2019, 36(2): 101-112. [王禄生. 司法大数据与人工智能技术应用的风险及伦理规制[J]. 法商研究, 2019, 36(2): 101-112.]
- [12] Wei B. Jurisprudential reflection and response to smart justice [J]. Polit Sci Law, 2021(8): 111-125. [魏斌. 智慧司法的法理反思与应对[J]. 政治与法律, 2021(8): 111-125.]
- [13] Wang Z. The realization path of “quasi-syllogism” for judicial artificial intelligence reasoning assistance [J]. J Polit Sci Law, 2022, 40(5): 28-39. [王竹. 司法人工智能推理辅助的“准三段论”实现路径[J]. 政法论坛, 2022, 40(5): 28-39.]
- [14] Borgesano F, De Maio A, Laghi P, *et al.* Artificial intelligence and justice: A systematic literature review and future research perspectives on Justice 5.0 [J]. Eur J Innov Manag, 2025, 28(11): 349-385.
- [15] Alcántara Francia O A, Nunez-del-Prado M, Alatriza-Salas H. Survey of text mining techniques applied to judicial decisions prediction [J]. Appl Sci, 2022, 12(20): 10200.
- [16] The impact of artificial intelligence on legal decision-making [J]. Int Comp Jurisprud, 2023, 9(2): 155-169.
- [17] Vidaki A N, Papakonstantinou V. Democratic legitimacy of AI in judicial decision-making [J]. AI Soc, 2025, 40(8): 6025-6035.
- [18] Aini G. A summary of the research on the judicial application of artificial intelligence [J]. Chin Stud, 2020, 9(1): 14-28.
- [19] Chen X, Mao M, Li S, *et al.* Debate-Feedback: A multi-agent framework for efficient legal judgment prediction [PP/OL]. V1. (2025-04-07) [2025-10-22]. <http://arxiv.org/abs/2504.05358>.
- [20] Dai Y, Feng D, Huang J, *et al.* LAiW: A Chinese legal large language models benchmark [PP/OL]. V2. (2024-01-18) [2025-06-15]. <http://arxiv.org/abs/2310.05620>.
- [21] Shi J, Guo Q, Liao Y, *et al.* Legal-LM: Knowledge graph enhanced large language models for law consulting [M]// Huang D S, Si Z, Zhang C. Advanced Intelligent Computing Technology and Applications. Singapore: Springer-Nature, 2024: 175-186.
- [22] Cui J, Ning M, Li Z, *et al.* Chatlaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model [PP/OL]. V3. (2024-06-08) [2025-10-24]. <http://arxiv.org/abs/2306.16092>.
- [23] Huang Q, Tao M, Zhang C, *et al.* Lawyer LLaMA: Enhancing LLMs with legal knowledge [PP/OL]. V2. (2023-10-14) [2025-11-25].

- <http://arxiv.org/abs/2305.15062>.
- [24] Fei Z, Zhang S, Shen X, *et al.* InternLM-Law: An open source Chinese legal large language model [PP/OL]. V1. (2024-06-21) [2025-11-26]. <http://arxiv.org/abs/2406.14887>.
- [25] Yue S, Chen W, Wang S, *et al.* DISC-LawLLM: Fine-tuning large language models for intelligent legal services [PP/OL]. V2. (2023-09-23) [2025-11-22]. <http://arxiv.org/abs/2309.11325>.
- [26] Li H, Chen Y, He Y, *et al.* LexRAG: Benchmarking retrieval-augmented generation in multi-turn legal consultation conversation [C]// Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval. Padua: ACM, 2025: 3606-3615.
- [27] Zhang K, Yu W, Dai S, *et al.* CitaLaw: Enhancing LLM with citations in legal domain [PP/OL]. V2. (2025-01-22) [2025-11-21]. <http://arxiv.org/abs/2412.14556>.
- [28] Deng W, Pei J, Kong K, *et al.* Syllogistic reasoning for legal judgment analysis [C]// Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. Singapore: ACL, 2023: 13997-14009.
- [29] Jiang C, Yang X. Legal syllogism prompting: Teaching large language models for legal judgment prediction [C]// Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law. Braga: ACM, 2023: 417-421.
- [30] Song Y, Qin Y, Huang R, *et al.* Legal text summarization via judicial syllogism with large language models [J]. J King Saud Univ Comput Inf Sci, 2025, 37(5): 111.
- [31] Zhang K, Yu W, Sun Z, *et al.* An explicit syllogistic legal reasoning framework for large language models [PP/OL]. V2. (2025-05-30) [2025-10-27]. <http://arxiv.org/abs/2504.04042>.
- [32] Dai X, Xu B, Liu Z, *et al.* LegalΔ: Enhancing legal reasoning in LLMs via reinforcement learning with chain-of-thought guided information gain [PP/OL]. V3. (2026-02-09) [2026-07-24]. <http://arxiv.org/abs/2508.12281>.
- [33] Shi J, Guo Q, Liao Y, *et al.* LegalGPT: Legal chain of thought for the legal large language model multi-agent framework [M]// Huang D S, Si Z, Chen W. Advanced Intelligent Computing Technology and Applications. Singapore: Springer-Nature, 2024: 25-37.
- [34] Fei Z, Shen X, Zhu D, *et al.* LawBench: Benchmarking legal knowledge of large language models [PP/OL]. V1. (2023-09-28) [2025-11-23]. <http://arxiv.org/abs/2309.16289>.
- [35] Li H, Chen Y, Ai Q, *et al.* LexEval: A comprehensive Chinese legal benchmark for evaluating large language models [PP/OL]. V4. (2024-11-26) [2025-11-23]. <http://arxiv.org/abs/2409.20288>.
- [36] Li H, Ye J, Hu Y, *et al.* CaseGen: A benchmark for multi-stage legal case documents generation [PP/OL]. V1. (2025-02-25) [2026-06-29]. <https://arxiv.org/abs/2502.17943>.
- [37] Jia Z, Yue S, Chen W, *et al.* Ready Jurist One: Benchmarking language agents for legal intelligence in dynamic environments [PP/OL]. V2. (2025-07-17) [2025-10-24]. <http://arxiv.org/abs/2507.04037>.
- [38] Shao P, Xu L, Wang J, *et al.* When large language models meet law: Dual-lens taxonomy, technical advances, and ethical governance [PP/OL]. V4. (2026-01-23) [2026-06-26]. <https://arxiv.org/abs/2507.07748>.
- [39] Pan Y, Xie Y X, Ning B, *et al.* Digital intelligence empowerment, rule-of-law business environment construction, and commercial credit financing: Empirical evidence from the “smart court” perspective [J]. Manage World, 2022, 38(9): 194-208. [潘越, 谢玉湘, 宁博, 等. 数智赋能、法治化营商环境建设与商业信用融资——来自“智慧法院”视角的经验证据[J]. 管理世界, 2022, 38(9): 194-208.]
- [40] Zhang L H. The hidden concerns of technology dependence in smart justice and its countermeasures [J]. Law Soc Dev, 2022, 28(4): 180-200. [张凌寒. 智慧司法中技术依赖的隐忧及应对[J]. 法制与社会发展, 2022, 28(4): 180-200.]
- [41] Sun X Y. Application and prospects of judicial big data in Chinese courts [J]. China Legal Sci, 2021(4): 123-144. [孙晓勇. 司法大数据在中国法院的应用与前景展望[J]. 中国法学, 2021(4): 123-144.]
- [42] Kort F. Predicting Supreme Court decisions mathematically: A quantitative analysis of the “right to counsel” cases [J]. Am Polit Sci Rev, 1957, 51(1): 1-12.
- [43] Ulmer S S. Quantitative analysis of judicial processes: Some practical and theoretical applications [J]. Law Contemp Probl, 1963, 28(1): 164-184.

- [44] Keown R. Mathematical models for legal prediction [J]. *Computer L J*, 1980, 2(1): 829-862.
- [45] Saravanan M, Ravindran B, Raman S. Improving legal information retrieval using an ontological framework [J]. *Artif Intell Law*, 2009, 17(2): 101-124.
- [46] Bartolini R, Lenci A, Montemagni S, *et al.* Semantic mark-up of Italian legal texts through NLP-based techniques [C/OL]// *Proceedings of LREC 2004*. 2004: 795. <http://www.lrec-conf.org/proceedings/lrec2004/pdf/709.pdf>.
- [47] Zeng Y, Wang R, Zeleznikow J, *et al.* Knowledge representation for the intelligent legal case retrieval [M]// Khosla R, Howlett R J, Jain L C. *Knowledge-Based Intelligent Information and Engineering Systems*. Berlin/Heidelberg: Springer, 2005: 339-345.
- [48] Chen H X, Chen T J, Wang H, *et al.* Prediction of prison terms for dangerous driving offenses based on probabilistic graphical model [J]. *J Sichuan Univ (Nat Sci Ed)*, 2022, 59(6): 7-12. [陈鸿旭, 陈铁今, 王皓, 等. 基于概率图模型的危险驾驶罪刑期预测 [J]. *四川大学学报(自然科学版)*, 2022, 59(6): 7-12.]
- [49] Vacek T, Schilder F. A sequence approach to case outcome detection [C]// *Proceedings of the 16th International Conference on Artificial Intelligence and Law*. London: ACM, 2017: 209-215.
- [50] Mahar S, Zafar S, Nishat K. Headnote prediction using machine learning [J]. *Int Arab J Inf Technol*, 2021, 18(5): 678-685.
- [51] Vacek T, Teo R, Song D, *et al.* Litigation analytics: Case outcomes extracted from US federal court dockets [C]// *Proceedings of the Natural Language Processing Workshop 2019*. Minneapolis: ACL, 2019: 45-54.
- [52] Li S, Zhang H, Ye L, *et al.* Mann: A multichannel attentive neural network for legal judgment prediction [J]. *IEEE Access*, 2019, 7: 151144-151155.
- [53] Niklaus J, Chalkidis I, Stürmer M. Swiss-judgment-prediction: A multilingual legal judgment prediction benchmark [C]// *Proceedings of the Natural Language Processing Workshop 2021*. Punta Cana: ACL, 2021: 19-35.
- [54] Lyu Y, Wang Z, Ren Z, *et al.* Improving legal judgment prediction through reinforced criminal element extraction [J]. *Inf Process Manage*, 2022, 59(1): 102780.
- [55] Feng D, Hu B, Zhang Y, *et al.* Multi-scale heterogeneous graph attention network for prison term prediction [C]// *2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*. Chongqing: IEEE, 2023: 1395-1404.
- [56] Huang X, He Q, Xu Y, *et al.* Judicial named entity recognition based on improved bootstrap-based tri-training [C]// *2024 6th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI)*. Guangzhou: IEEE, 2024: 580-587.
- [57] Zhang Y, Feng D, Tian W, *et al.* Interpretable sentencing element-based prison term prediction [C]// *2022 IEEE 24th Int Conf on High Performance Computing & Communications; 8th Int Conf on Data Science & Systems; 20th Int Conf on Smart City; 8th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)*. Hainan: IEEE, 2022: 2214-2220.
- [58] Yu W, Sun Z, Xu J, *et al.* Explainable legal case matching via inverse optimal transport-based rationale extraction [C]// *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Madrid: ACM, 2022: 657-668.
- [59] Huang Q. AndrewZhe/lawyer-llama [CP/OL]. (2026-03-03) [2026-03-08]. <https://github.com/AndrewZhe/lawyer-llama>.
- [60] Wu Y, Liu Y, Liu Y, *et al.* zhihaiLLM/wisdomInterrogatory [CP/OL]. (2026-01-27) [2026-03-08]. <https://github.com/zhihaiLLM/wisdomInterrogatory>.
- [61] SIAT-NLP. HanFei [CP/OL]. 2026 [2026-03-08]. <https://github.com/siat-nlp/HanFei>.
- [62] DUTIR-LegalIntelligence. Tailing [CP/OL]. (2025) [2026-03-08]. <https://github.com/DUTIR-LegalIntelligence/Tailing>.
- [63] Xie Y, Zhou Y, Gao H, *et al.* Qzhou-Law: An open source series of Chinese legal large language models [PP/OL]. (2025-12-13) [2026-03-09]. <https://openreview.net/forum?id=TRuj9IpWOL>.
- [64] FudanDISC. DISC-LawLLM [CP/OL]. (2026) [2026-03-08]. <https://github.com/FudanDISC/DISC-LawLLM>.
- [65] Wu S, Liu Z, Zhang Z, *et al.* fuzi.mingcha [CP/OL]. (2026-03-04) [2026-03-08]. <https://github.com/irlab-sdu/fuzi.mingcha>.
- [66] Liu H C, Li H C, Ma Y T, *et al.* LiuHC0428/LAW-GPT [CP/OL]. (2026-03-01) [2026-03-08].

- <https://github.com/LiuHC0428/LAW-GPT>.
- [67] InternLM. InternLM-Law [CP/OL]. (2026) [2026-03-08]. <https://github.com/InternLM/InternLM-Law>.
- [68] Chen H, Xu Y, Wang B, *et al.* LexPro-1.0 technical report [PP/OL]. V2. (2025-03-11) [2026-03-08]. <http://arxiv.org/abs/2503.06949>.
- [69] Wang Y, Feng D, Dai Y, *et al.* HARMONIC: Harnessing LLMs for tabular data synthesis and privacy protection [J]. Adv NeuIPS, 2024, 37: 100196-100212.
- [70] Qin W, Luo X. A legal fact-finding model based on the T5 and LexiLaw large language models [C]// Proceedings of the 2024 8th International Conference on Computer Science and Artificial Intelligence. Beijing: ACM, 2024: 229-237.
- [71] Zhang K, Xie G, Yu W, *et al.* Legal mathematical reasoning with LLMs: Procedural alignment through two-stage reinforcement learning [C]// Findings of the Association for Computational Linguistics: EMNLP 2025. Suzhou: ACL, 2025: 1586-1598.
- [72] Shi W, Wang Q, Pan Y, *et al.* LegalChainReasoner: A legal chain-guided framework for criminal judicial opinion generation [PP/OL]. V1. (2025-08-31) [2025-11-27]. <http://arxiv.org/abs/2509.00783>.
- [73] Dong N T, Nguyen M A, Hoang T D, *et al.* VLegal-Bench: Cognitively grounded benchmark for Vietnamese legal reasoning of large language models [PP/OL]. V1. (2025-08-31) [2026-03-09]. <http://arxiv.org/abs/2512.14554>.
- [74] Canaverde B, Pires T P, Ribeiro L M, *et al.* Legal-Bench. PT: A benchmark for Portuguese law [PP/OL]. V1. (2025-01-22) [2026-03-09]. <http://arxiv.org/abs/2502.16357>.
- [75] Cao Z, Nishino F, Yamada H, *et al.* JBE-QA: Japanese Bar Exam QA dataset for assessing legal domain knowledge [PP/OL]. V1. (2025-11-28) [2026-03-09]. <http://arxiv.org/abs/2511.22869>.
- [76] Yu Y, Gan L, Hu Y, *et al.* Evaluating test-time scaling LLMs for legal reasoning: OpenAI o1, DeepSeek-R1, and beyond [PP/OL]. V2. (2025-11-10) [2026-06-09]. <http://arxiv.org/abs/2503.16040>.
- [77] Gan R, Feng D, Zhang C, *et al.* UCL-Bench: A Chinese user-centric legal benchmark for large language models [C]// Findings of the Association for Computational Linguistics: NAACL 2025. Albuquerque: ACL, 2025: 7960-8003.
- [78] Shi Y, Liu H, Hu Y, *et al.* PLawBench: A rubric-based benchmark for evaluating LLMs in real-world legal practice [PP/OL]. V2. (2026-01-28) [2026-03-09]. <http://arxiv.org/abs/2601.16669>.
- [79] Yang H, Wang L Z, Fan W, *et al.* AppellateGen: A benchmark for appellate legal judgment generation [PP/OL]. V3. (2026-03-29) [2026-06-09]. <http://arxiv.org/abs/2601.01331>.
- [80] Cai Y, Hu R, Xie H, *et al.* LexRel: Benchmarking legal relation extraction for Chinese civil cases [PP/OL]. V2. (2025-04-20) [2026-03-09]. <http://arxiv.org/abs/2512.12643>.
- [81] Bi S, Yu K, Li J, *et al.* CLER: A benchmark for Chinese litigation evidence reasoning [J]. Inf Process Manage, 2026, 63(5): 104667.
- [82] Tong X, Jin B, Lin Z, *et al.* CPSDBench: A large language model evaluation benchmark and baseline for Chinese public security domain [PP/OL]. V3. (2024-03-21) [2026-03-09]. <http://arxiv.org/abs/2402.07234>.

(责任编辑:周兴旺)