

论人工智能证据的独立性

郭雪慧^{*}

内容提要 人工智能证据在司法实践中的应用日益增多,但学界对其是否应成为独立的证据类型存在较大争议。人工智能证据是指,以具备自主数据处理能力的人工智能系统为核心生成主体,由算法驱动自主或半自主生成,承载能够证明案件事实的客观化信息,并服务于诉讼证明活动的证据材料。将人工智能证据归入书证、鉴定意见、电子数据等现有证据类型均存在理论与实践层面的困境。人工智能证据在生成主体、生成机制和表现形式等方面具有区别于传统证据的本体独特性,且对传统证据的准入规则、审查规则等形成系统性冲击。因而,从归类验证、本质剖析与规则研判三个层面分析,确立人工智能证据的独立证据地位具备理论依据。构建体系化的人工智能证据规则,应统筹修订三大诉讼法,将人工智能证据增设为独立证据类型;构建以合法性、相关性与可靠性为核心的证据准入标准,并设置附条件准入的例外规则;建立分类审查规则、人机协同纠错规则、可采性验证规则等证据审查规则。

关键词 人工智能证据 独立性 证据类型 证据准入 证据审查

目 次

- 一、引言
- 二、人工智能证据的概念厘定
- 三、人工智能证据独立性的理论证成
- 四、人工智能证据独立性的规则因应
- 五、结语

^{*} 中国民航大学法学院教授。

一、引言

人工智能已超出传统辅助工具的范畴,具备自主数据处理、逻辑分析乃至结论生成的能力。在司法实践中,随着人工智能技术的飞速发展,证据生成方式发生根本性变革。例如,在民事纠纷案件审理过程中,人工智能合同审查系统能够自动甄别出格式条款中无效的内容,并生成具体的审查意见;图像识别技术可以精准比对出侵权产品和原产品之间的差异,并快速形成溯源报告。人工智能不再是单纯辅助司法人员办案的工具,其正在逐步升级为证据生成的主体。在这一现实背景下,人工智能系统通过自主分析、研判、合成等方式生成可用于证明案件事实的材料,使得人工智能证据开始出现在各种诉讼程序中。^①但是,司法机关在面对这类新型证据时常陷入定性模糊、筛选识别困难、审查规则缺失等困境。^②要解决当前难题,首先必须回答人工智能证据能否成为独立法定证据类型的问题。

学界对人工智能证据独立性的认知分歧较大,形成了以下代表性观点:一是主张人工智能证据不具有独立性,应将其归入现有证据类型。有学者认为,可以将人工智能证据归入电子数据的范畴,主要依据在于人工智能证据以数字化形式存在,这契合电子数据载体特征。^③另有学者认为其为一种新型意见证据。^④二是主张人工智能证据在生成机制和证明逻辑等方面具有独特性,应成为独立的法定证据类型。^⑤然而,主张将人工智能证据归入现有证据类型的研究未能充分回应人工智能证据和传统证据在生成主体、形成过程等方面存在的本质差异。主张人工智能证据应成为独立证据类型的研究未能深入、完整地探讨人工智能证据为何具备独立的证据地位。

我国现有证据体系的构建逻辑与规则设计均植根于传统证据生成机理,凸显出证据生成过程中人的主观能动性。而人工智能所具备的自主性与黑箱性,使得人工智能证据与现有的各类证据在生成机理和存在形态上存在本质差异,现有证据规则难以契合人工智能证据的本质特征。^⑥本文以人工智能证据的独立性为研究核心,按照“概念厘定—理论证成—规则构建”的逻辑展开,力图从本体论层面揭示人工智能证据独立性的理论依据,并构建一套与之相适配的证据规则体系,为推动人工智能证据同司法公正的深度匹配提供理论支撑。

① 参见谢登科:《可信时间戳认证电子数据的法律效力与审查认定》,载《法律适用》2025年第7期,第76页。

② 参见杨波:《程序性事实观下刑事证据制度改革路径探究》,载《法学论坛》2025年第1期,第35页。

③ 参见熊晓彪:《生成式人工智能证据认定的困境与规范进路》,载《法律科学》2025年第1期,第84页。

④ 参见余鹏文:《刑事诉讼中人工智能证据的法律性质和运用规则》,载《中国刑事法杂志》2024年第5期,第36页。

⑤ 参见马国洋:《论刑事诉讼中人工智能证据的审查》,载《中国刑事法杂志》2021年第5期,第161-162页。

⑥ 参见张吉豫:《赋能型人工智能治理的理念确立与机制构建》,载《中国法学》2024年第5期,第67页。

二、人工智能证据的概念厘定

探讨人工智能证据的独立地位,首先需要厘清其概念,具体分析人工智能证据的内涵与外延。基本属性对应人工智能证据的内涵,主要形态以及与其他新兴证据之间的关系对应人工智能证据的外延。

(一) 人工智能证据的基本属性

人工智能证据的基本属性包括主体属性、生成属性、内容属性与目的属性。其中,主体属性明确人工智能证据的核心生成主体为人工智能系统;生成属性说明人工智能证据的生成过程系由算法驱动的自主或半自主生成;内容属性指明人工智能证据的内容为能够证明案件事实的客观化信息;目的属性则确立人工智能证据以服务诉讼证明为导向。具体论述如下:

一是主体属性。主体属性是对人工智能证据进行界定的首要前提,核心争议集中于人工智能证据的生成主体是开发或者操作人工智能系统的人,还是系统本身。对这一问题的回答直接决定了人工智能证据与传统证据的本质分界。传统证据的生成主体都是人,如书证制作者、证人证言提供者、鉴定意见出具者和专家辅助人等,证据生成过程无不体现着人的主观意志和认知能力,责任追溯也是围绕人展开。^⑦而人工智能证据呈现出明显的“非人格化”特征,人工智能系统往往扮演核心生成角色。人工智能证据的生成主体是具备自主数据处理和结论生成能力的人工智能系统。但强调人工智能系统具有核心生成主体地位,并不否定人类在人工智能证据生成过程中所发挥的辅助作用。人类在其中的作用主要体现在:开发与维护人工智能系统、输入初始数据、设置参数,进而为运算提供基础条件;筛选系统输出的结果,并将符合诉讼需求的材料纳入司法程序。将人工智能系统视为核心生成主体,也为后续责任追究与审查判断明确了对象。司法机关不仅需要审查输入数据的合法性、真实性和关联性,更要审查人工智能系统的技术可靠性和运作规范性。^⑧

二是生成属性。生成属性是人工智能证据与传统证据相区别的主要特征,关键区别在于人工智能证据的生成由算法驱动,而非由人类主观意志直接主导。从生成方式上看,人工智能证据生成可以分为自主生成和半自主生成两类。但不管何种类型,算法驱动这一核心特征始终不变。自主生成指人工智能系统在接收到初始数据后,无需人类进行进一步干预便可以通过算法模型完成对全部数据的分析和结论生成。比如,在网络诈骗案件中,人工智能反诈骗系统通过自主分析海量的交易数据和通讯数据,识别出具有诈骗特征的交易模式,并生成与涉案人员相关联的分析报告,全程由算法驱动,无需人类实时参与。半自主生成指人工智能完成初阶信息处理后,需要人类接续完成进一步的信息比

^⑦ 参见韩旭、王剑波:《刑事庭审质证运行状况实证研究——以100个庭审案例为样本》,载《法治研究》2016年第6期,第52页。

^⑧ 参见张霞:《资金分析报告司法审查的实践困境与规范进路》,载《法学》2025年第10期,第151-153页。

对和筛选,从而形成最终结果。例如,在知识产权侵权案件中,人工智能图像识别比对系统自主提取侵权产品与原产品图像特征并完成比对、生成初步结果,后续由专业人员进一步比对特征的关联性和准确性。

三是内容属性。内容属性是证据的核心属性,对于案件待证事实的认定主要是依托证据所载信息完成的。客观性是采信证据的主要标准之一,也是司法人员审查证据的主要内容。如果经审查发现某个证据不具备客观性,则法官不能将其作为定案依据。人工智能证据的内容应界定为可证明案件事实的客观化信息。因其客观化程度更高,人工智能证据在生成过程中受到人类主观干预较少,更能反映客观事实。人工智能证据的呈现形式多种多样,包括文字、图像、音频、视频等,既可以是文字分析报告和数据表格,也可以是图像对比结果和行为轨迹图,还可以是语音分析报告和动态分析视频。司法机关在审查人工智能证据时,需要全面核查其内容的关联性、合法性和可靠性,以确保其可以作为定案依据。

四是目的属性。目的属性也是对人工智能证据进行界定的重要维度,其核心在于人工智能证据的生成和提交必须以进行诉讼证明为目的。如果人工智能系统生成的材料并不是用于证明案件事实,此类材料就不属于人工智能证据的范畴。例如,在诉讼活动中,智能辅助系统所生成的仅用于提供线索、风险提示与信息整合的材料,不属于人工智能证据范畴。数字法院依托大数据技术开展关联案件筛查、虚假诉讼识别工作,相关输出结果仅作为案件线索、办案参考或监督材料,其功能在于通过人机协作提升办案效率,不属于直接用于诉讼证明的人工智能证据。

人工智能证据的四项基本属性彼此关联、缺一不可,共同框定了人工智能证据的核心内涵,即人工智能证据是指,以具备自主数据处理能力的人工智能系统为核心生成主体,由算法驱动自主或半自主生成,承载能够证明案件事实的客观化信息,并服务于诉讼证明活动的证据材料。该内涵为后续探讨人工智能证据的独立地位奠定了概念基础。

(二) 人工智能证据的主要形态

随着人工智能技术在司法领域的运用,证据形态也随之多元化。结合当前司法场景中人工智能技术的应用现状,可以将人工智能证据的主要形态划分为智能分析报告类证据、智能识别结果类证据、智能预测和溯源类证据、智能数据合成类证据等,这些形态共同构成人工智能证据的多元化图景。不同形态的人工智能证据,其生成逻辑、呈现样式、证明力各有不同。

一是智能分析报告类证据。这类证据是我国司法实践中较为常见的人工智能证据形态,是指人工智能系统在接收原始数据后,依托算法模型完成分析、归纳与总结,最终生成以文字或表格形式表现出来的分析报告。其中,智能轨迹分析证据是由智能轨迹分析技术生成,可用来证明案件事实的研判意见。^⑨例如,行为轨迹智能分析报告是人工智能系统对涉案人员或物品的轨迹数据进行分析后所产生的报告。在刑事案件侦查过程

^⑨ 参见马国洋:《人工智能证据运用规则研究》,中国政法大学出版社2024年版,第104页。

中,办案机关会调取监控录像、卫星定位记录等各种原始电子材料,并交由人工智能进行轨迹推演与预测,还原涉案人员行动路线、涉案物品流转的全过程,从而形成最终的分析文书。在“张某走私、贩卖、运输、制造毒品案”中,法院就采纳了行为轨迹智能分析报告,将得出的最终结果用于对被告人行为轨迹进行认定。^⑩

二是智能识别结果类证据。各类人工智能识别结论在锁定行为人身份和追查案件线索时都具有极高的价值。图像人工智能识别结果,是依靠图像比对算法,根据案件中相关图片和监控录像内容完成拆解、比对和甄别后形成的结论。近年来,人脸识别结论被广泛应用于刑事侦查领域。^⑪在处理民事侵权责任纠纷时,可以对涉案侵权商品和正版产品影像素材作详细比对,标注出两者区别和重合之处,并出具比对意见,进而为法院认定侵权事实提供支撑。语音人工智能识别结果往往依托语音解析算法对案件相关录音文件进行拆解,能够精准地检测音频文件,判断其是否存在伪造、修改痕迹,从而为法官核实证据真伪提供有力支撑。比如,通过声纹鉴定实现对声音真实性的证明。^⑫

三是智能预测和溯源类证据。智能预测和溯源类证据多数是由人工智能系统自主生成的,呈现形式包括智能预测报告和智能溯源报告等。智能预测报告是人工智能系统依托历史数据与算法模型,对未来可能发生的案件事实或者发展趋势进行预测而形成的报告。比如,通过对罪犯个人情况和历史数据进行分析,对其再犯可能性形成风险评估报告;^⑬办案人员对罪犯相关数据搜集后,借助算法对数据加以解析,可以形成对犯罪风险概率的研判意见^⑭。智能溯源报告是人工智能系统通过数据溯源和物品溯源等技术,追溯涉案物品来源、流转过程或者涉案数据生成、传输、存储情况而形成的报告。

四是智能数据合成类证据。这类证据是人工智能系统对数据加以合成、重构而形成的服务于诉讼证明活动的新形态材料。这类证据的生成多数采用半自主模式,需要人类对输入的多源数据进行筛选、整理,再由人工智能系统完成合成和重构。根据合成内容的不同,可将其分为智能数据重构结果、智能图像视频合成结果等具体类型。智能数据重构结果是指人工智能系统将不完整、碎片化或者相互矛盾的数据进行整合与复原后所形成的证据材料。智能图像视频合成结果则是指人工智能系统对图像和视频进行合成、生成或者深度加工后所形成的证据材料,比如,基于深度合成技术而生成的涉案场景还原影像等。这类证据能够弥补传统取证手段的局限性,以更加直观的形式呈现案件事实,但因其生成过程涉及对原始数据的加工与改造,容易引发对人工智能证据真实性的质疑。基于此,审查时需要重点核查合成过程的技术参数和原始数据的留存情况。

(三) 人工智能证据与其他新兴证据的辨析

伴随社会发展与技术革新,与大数据、人工智能等相关的新类型证据相继出现。根

^⑩ 参见云南省高级人民法院(2020)云刑终734号刑事裁定书。

^⑪ 参见李胥:《论刑事侦查人脸识别结果的证据化运用》,载《证据科学》2023年第5期,第584页。

^⑫ 参见谢登科:《电子数据的技术性鉴真》,载《法学研究》2022年第2期,第210页。

^⑬ 参见江潮:《自动化决策、刑事司法与算法规制——由卢米斯案引发的思考》,载《东方法学》2020年第3期,第76-88页。

^⑭ 参见陈永生:《大数据预测警务的运作机理、风险与法律规制》,载《中国法学》2024年第5期,第173页。

据司法实践中的具体情形和学界现有争议,当前和人工智能证据密切相关的其他新型证据主要包括算法证据和大数据证据。^⑮明确这两者和人工智能证据的关系,是厘清人工智能证据外延必须回应的问题。人工智能证据、算法证据和大数据证据三者并非简单的并列关系或者互斥关系,而是呈现出交叉或包含的逻辑关系。

算法证据并非独立于人工智能证据的新型证据,其核心是用算法运算结果证明案件事实,生成过程依赖于算法驱动,生成主体多是具备自主运算能力的人工智能系统,内容往往为算法运算所产生的客观信息。在司法实践中,由算法生成的涉案人员关联分析报告以及侵权行为识别报告等,都属于人工智能证据的范畴。如果仅体现算法本身,比如算法代码、算法流程图,而没有通过人工智能系统运算生成可以证明案件事实的信息,也没有用于诉讼证明,则既不属于算法证据,也不属于人工智能证据。

大数据证据与人工智能证据的关系是互不包含却交叉重叠。大数据证据的核心是凭借海量数据分析结果,以证明案件事实。^⑯当前主流的人工智能模型往往依赖对大规模数据集合进行学习、识别与计算而生成输出内容。^⑰当大数据分析由具备自主运算能力的人工智能完成,满足人工智能证据的主体属性、生成属性、内容属性和目的属性时,这类大数据证据就属于人工智能证据。例如,人工智能反诈骗系统通过分析海量交易数据生成的诈骗模式报告,兼具人工智能证据和大数据证据两种证据属性。一些大数据证据可以通过人工分析或者简单的数据处理工具生成而无需人工智能系统自主运算,因其未契合人工智能证据的基本属性,而只属于大数据证据。

三、人工智能证据独立性的理论证成

判断一种新兴证据是否应成为独立的证据类型,可以依次展开以下论证:一是归类验证,判断能否把它划入现有证据类型,如果全部无法适配,就说明归入现有证据类型的路径行不通;二是本质剖析,对难以归类的根源进行探究,系统梳理新兴证据类型和传统证据类型存在的本质区别;三是规则研判,分析这类本质差异是否让传统证据规则无法适用。以上三项论证逐层推进,可完整证成人工智能证据的独立性。

(一) 人工智能证据归属现有证据类型的困境

人工智能证据这一新型证据形态出现后,司法实践中已经形成了书证归类、鉴定意见归类和电子数据归类等多种归属方案。^⑱现有证据类型与人工智能证据存在差异,强行归类难免陷入“削足适履”的困境。

^⑮ 参见前注④,余鹏文文,第36-37页。

^⑯ 参见王燃:《大数据证明的机理及可靠性探究》,载《法学家》2022年第3期,第58-59页。

^⑰ 参见游劝荣:《人工智能赋能新时代审判工作高质量发展的机遇、挑战及应对》,载《中国应用法学》2025年第3期,第9页。

^⑱ 参见前注④,余鹏文文,第39页。

1. 书证归类模式的困境

书证归类模式通常是将人工智能生成的分析报告、识别结论等文本类材料纳入书证的范畴。其核心依据是两者都是以文字、符号等形式承载着证明信息,人工智能证据契合书证以内容证明案件事实的核心特征。但通过深入剖析可以发现这一模式存在着严重的逻辑断裂,人工智能证据与书证的本质属性存在冲突。

从本质属性上看,书证的核心特征在于生成过程蕴含着人类的主观意志,记录了人类的思想或者行为。根据我国的证据法学理论,书证指以文字、符号、图形等记载内容、表达思想,以证明案件事实的书面文件或其他载体物品。人工智能证据的内容是人工智能系统对数据的分析和呈现,并不是人类思想或者意志的直接记录。^{①⑨} 尽管生成文本的形式和书证相类似,但其生成主体是人工智能系统,而不是人,内容来源是算法运算,而非人的主观意志,这和书证的本质属性存在差异。将人工智能证据归入书证实质上混淆了文本形式和本质属性的关系。书证归类模式还面临责任主体资格缺失的难题。书证制作人需要具备相应的民事行为能力,其行为后果可以归责于制作人。如果将人工智能证据归为书证,一旦人工智能证据存在虚假或者错误,无法像书证那样追溯具体制作人的责任,这就使得责任追溯机制失灵,不利于对当事人合法权益的保护。

从生成逻辑上看,书证的形成具有一次性和稳定性的特征,内容生成后往往较为固定,修改必留痕迹,这也是对书证进行审查时判断其真实性的的重要依据。人工智能证据的生成则呈现动态性和可重复性的特征,即便参数和版本固定,随着训练数据更新,系统分析的结论也可能发生变化,这种动态生成逻辑和书证稳定生成逻辑是相悖的。在司法实践中,书证归类模式还会引发证据审查标准适用错误,进而影响司法公正。书证审查的重点是形式真实性和内容关联性,核心方法包括对原件进行比对、询问制作人、鉴定笔迹等。将书证审查标准适用于人工智能证据可能导致两种后果:一是使审查流于形式,只关注文本的完整性,而忽视算法本身的可靠性和数据真实性等核心问题;二是错误要求人工智能证据满足制作人签字和单位盖章等书证的形式要件,导致具备证明价值的证据被排除。

2. 鉴定意见归类模式的困境

鉴定意见的核心在于人的主观专业性判断,人工智能证据的核心则在于算法的客观运算,两者生成逻辑存在本质差异。鉴定意见通常为具有专门知识的人运用专业知识、经验和技能对案件中的专门问题进行鉴别和判断后得出的意见,是鉴定人主观判断的体现。鉴定人在判断过程中表现出个性化和可解释性特征。即便针对同一专业问题,具有不同专业知识背景和经验积累的鉴定人可能会得出不同的意见。^{②⑩} 而且,鉴定人能清晰地解释判断过程、依据和理由。

人工智能证据的生成源于算法驱动,系统本身并不像人一样具备主观判断能力,最

^{①⑨} 参见葛翔:《司法实践中人工智能运用的现实与前瞻——以上海法院行政案件智能辅助办案系统为参照》,载《华东政法大学学报》2018年第5期,第74页。

^{②⑩} 参见程军伟:《专家辅助侦查制度的域外考察与中国镜鉴》,载《中国司法鉴定》2026年第3期,第19页。

终的输出结果是基于输入的数据和算法模型运算的产物。同时,受算法黑箱的影响,系统也无法像鉴定人那样清晰而准确地对整个运算过程进行解释。^②相较于人类可认知的鉴定意见,人工智能证据生成过程让人类难以充分理解。实践中,鉴定意见的出具需要依赖鉴定人,人工智能系统则不符合鉴定意见的主体要件。对鉴定意见的审查主要包括询问鉴定人、通知鉴定人出庭作证、补充鉴定或者重新鉴定等。对于人工智能证据的审查,无法要求人工智能系统出庭作证,其开发人员或者操作人员即便出庭,也往往无法完整、详细地对算法运算过程和核心原理进行解释,质证难以有效开展。对人工智能证据的审查则需要重点关注算法模型的可靠性、运算过程的规范性、输入数据的合法性与真实性,以及系统版本的稳定性等。对人工智能证据的可靠性存在争议时,传统意义上的补充鉴定和重新鉴定制度也无法适用。鉴定意见在法定情形下可以进行补充鉴定或者重新鉴定,补充鉴定需要鉴定人对原意见进行完善,重新鉴定需要其他鉴定机构重新作出鉴定。但是,人工智能证据生成依赖于特定算法和系统,事实上无法由其他鉴定机构通过相同或不同方式重新生成结论。

3. 电子数据归类模式的困境

现阶段认可度最高、使用范围最广的归类方式就是把人工智能证据归类为电子数据。但电子数据的核心功能只是把人类行为和客观情况转化为数字形式留存,生成过程需要人为操作,其本身不具备自主运算和推导生成新信息的能力。如果直接将人工智能证据归入电子数据的范畴,会使得单纯数字化记录和算法自主运算得出数字化结论这两者的边界模糊,忽视人工智能系统自主产出内容这一人工智能证据的独有属性。

电子数据和人工智能证据在主体属性和生成属性方面存在本质区别。电子数据由自然人操作设备生成,而人工智能证据的生成主体为人工智能算法系统,整个运算和生成过程都带有自主研判的属性。在日常场景中,发送邮件、线上交易等人为操作电子设备的行为,都会同步产生电子数据,这也是电子数据较为常见的产生途径。而人工智能证据的生成则往往呈现自主特征,系统可以主动地对输入的原始数据进行加工、处理和分析,最后输出结果,整个过程无需人类实时介入。这种被动记录和主动生成的本质区别决定了两者在生成逻辑上的根本不同。

审查标准的适配性问题在电子数据归类模式中同样突出。电子数据审查的核心在于真实性和完整性,^③即数据是否存在伪造、篡改的情形,能否完整地反映事实^④。我国相关司法解释针对电子数据审查制定了相应的规则,涵盖对证据生成、收集、存储、传输过程的合法性审查,伪造、篡改痕迹的核查以及与其他证据之间的相互印证等。如果对人工智能证据适用电子数据的审查标准,则缺少要求审查算法、模型和训练数据的相关规则,会使对人工智能证据的核心审查要素缺失,难以保障其可靠性。此外,将人工智能证据归入电子数据还会引起证明力判断标准的错位。对电子数据证明力的判断需要关注

^② 参见熊秋红:《人工智能在刑事证明中的应用》,载《当代法学》2020年第3期,第88页。

^③ 参见吴洪淇:《电子数据完整性的法律定位与理论反思》,载《国家检察官学院学报》2024年第1期,第146-160页。

^④ 参见王芳:《数字证据的性质及相关规则》,载《法学》2004年第8期,第72-79页。

关联性和真实性,结合其他证据最终形成完整的证据链。人工智能证据的证明力大小不仅取决于关联性和真实性,还取决于算法的科学性和准确性。^{②④}即便输入数据真实且与案件事实相关,如果算法存在漏洞,也会严重影响输出结果的证明力。

(二) 人工智能证据的本体独特性

在对人工智能证据各种归类路径进行逐一排除后,需要探寻人工智能证据无法被纳入现有证据类型背后深层次的原因。这就需要进入第二层论证,即对人工智能证据的本体独特性进行分析。基于上文关于人工智能证据基本属性的论述,可以进一步从生成主体、生成机制、表现形式等方面论证人工智能证据的本体独特性。

1. 生成主体的非人格化

传统证据的生成均以人作为核心主体,例如,鉴定人、证人均拥有主观认知。传统证据的信任基础也始终围绕着人展开,即便对于物证、书证等客观证据,信任的建立也离不开人的解读和判断,本质上仍然是对人类理性的依赖。与之不同的是,人工智能证据呈现鲜明的非人格化特征,其核心生成主体是具备自主运算能力的人工智能系统,人类只是负责人工智能系统开发、原始数据输入与参数调试等辅助工作。但人工智能生成主体的非人格化并不意味着责任虚化。人工智能证据的开发方、使用方、部署方都要承担相应的责任。在对人工智能证据进行审查时需要理清各主体的责任边界,从而避免因主体缺位而使质证无从展开。^{②⑤}

2. 算法驱动的生成机制

传统证据的生成围绕“人—行为—事实”的逻辑展开,无论是书证、物证、证人证言、鉴定意见还是电子数据,生成机制都以人类主观意志为核心,本质上是对人类行为、思想或者客观案件事实的记录与呈现。而人工智能证据的生成则遵循“技术—算法—事实”这一全新的逻辑链条,其以算法作为核心驱动力,基于预设算法模型而展开深度的分析与逻辑推理,进而形成结论,并以文字、图形、视频等多元形态自主呈现,实现了从传统证据被动记录事实到主动分析数据并输出结论的本质转变。^{②⑥}人工智能的输出结果是算法对海量数据进行自主运算的产物,而不是对人类意志或者客观事实进行简单记录。^{②⑦}人工智能证据具有动态性和迭代性的特征,体现了证明逻辑的动态变化,这和传统证据生成后内容固定、证明逻辑单一的特征形成了明显区别,也进一步凸显了人工智能证据在生成机制层面的本体独特性。

3. 载体与内容的深度融合

判断一种证据能否成为独立的证据类型,载体和内容的关联性是重要的参考标准。对此,可以从以下两个层面进行分析:一是形式层面,考虑证据内容能不能脱离原始载体完成复制、流传和存储;二是实质层面,考虑载体与内容是否必须依托特定技术条件。

^{②④} 参见前注^{①⑥},王燃文,第63页。

^{②⑤} 参见王莹:《算法侵害责任框架刍议》,载《中国法学》2022年第3期,第178-183页。

^{②⑥} 参见于辉:《案件事实认定中人工智能适配的逻辑方法》,载《政法论丛》2025年第6期,第123-127页。

^{②⑦} 参见刘品新:《论大数据证据》,载《环球法律评论》2019年第1期,第26页。

传统证据的载体和内容,无论从形式上还是实质上看,都是能够相互脱离的。比如,写在纸上的文书、以实物形式呈现的物证,它们所承载的信息本身都具有独立性。一份合同不管是被抄写、拍照或是转换为其他形式,表达的含义都不会发生变化。人们对其进行解读也不需要借助额外的技术工具,载体只是起到单纯的信息存储作用,和其所载的内容并没有深度绑定。电子数据在形式上可以脱离原有载体,通过复制、备份完成流转,这也是所有数字信息的共同特性。查阅电子数据,只需要使用普通电子设备和应用软件就可以完成,设备和软件仅仅起到信息展示作用,载体只是内容的承载媒介,两者也不存在深度的绑定关系。虽然人工智能证据也能够被存储,但是其内容本身与整套人工智能技术深度融合,它不是单纯对过往客观事件作留存记录,而是算法经过大量运算后得出的结论,想要完整地还原并读懂这份证据包含的全部信息就需要关注算法模型。算法模型的选用、算法运行依托的软硬件环境、训练数据集、权重参数和阈值等各类技术指标的改变,都会导致人工智能证据承载内容的核心逻辑和判断结论发生相应改变。缺少任意一项基础条件都无法还原人工智能证据生成的整个过程,也无法精准地解读人工智能证据传递的案件信息。

(三) 人工智能证据运用对传统证据规则的冲击

依托传统证据而构建的部分证据规则已难以适配人工智能证据的发展。人工智能证据对部分证据规则造成冲击,也是将其增设为独立证据类型的重要缘由。

1. 对传统证据准入规则的冲击

我国的证据准入规则对证据合法性、相关性和真实性都有明确的规定。人工智能证据在生成、存储等过程中,技术属性贯穿始终,传统准入规则在判断其合法性、相关性、真实性时存在明显的适配障碍。

首先,传统的证据合法性判断标准难以覆盖人工智能证据生成与收集的全流程。传统的证据合法性审查主要聚焦于收集主体适格性、程序合规性及形式法定性。而人工智能证据生成涉及输入数据采集、算法模型开发和运算过程等多重环节,这些新型要素未被传统证据准入规则涵盖,使得司法机关无法基于传统标准全面校验人工智能证据的合法性。

其次,传统的证据相关性判断标准在人工智能证据适用中存在固有局限。传统证据相关性判断采用实质关联说,即核查证据内容和案件事实的实质性关联和证明价值。人工智能证据相关性则呈现出间接性和复杂性的特征,其输出结果与案件事实的关联需要通过输入数据相关性和算法运算逻辑性等间接层面进行佐证,而无法仅凭内容本身就能直接进行判定。

最后,传统的证据真实性判断标准难以适配人工智能证据的动态生成特征。传统证据真实性审查往往聚焦于证据原件的属性、伪造和篡改的痕迹以及证据之间的相互印证关系等。科学证据的形态并非一成不变,而是随着人类科学技术的不断进步而逐步发展的。人工智能证据生成具有动态性,同一组输入数据在不同算法参数或者系统版本下可能会产生差异化的结果。

2. 对传统证据审查规则的冲击

证据审查规则是司法机关对进入诉讼程序的证据进行审查,以确定其证明力大小和可采性的规范,构成了证据规则体系的核心内容。对人工智能证据的审查需要遵循独立验证要求,同时要制定适配人工智能证据特性的可采性标准。

第一,人工智能证据对传统证据审查规则形成冲击的根源在于,审查对象从传统的单一对象变为多元复合体系。传统证据的审查对象为单一对象,即仅审查证据本身,审查的核心焦点始终围绕证据的形式和内容展开。例如,书证审查聚焦于文本自身的真实性和与案件的相关性,即便在审查过程中有时也会涉及证据的生成过程,但也只是为了佐证证据本身的可靠性,并不会将生成过程列为独立的审查对象。人工智能证据的审查对象是由输入数据、算法模型、运算过程、输出结果共同构成的多元复合体系。实践中,该审查对象常会遭遇多重难题,比如,可供分析的数据样本覆盖面有限,难以实现全样本采集;构建的算法模型存在偏差等。^{②⑧} 实时输入数据是人工智能系统产出人工智能证据的基础材料,材料本身是否真实,所获取的渠道是否合法,直接决定着后续算法所输出的结论是否可靠,这也是办案人员开展证据审查要把关的环节。算法的准确性、透明度和可靠性至关重要,若将人工智能算法引入司法取证中,还需要考虑算法偏见和技术伦理层面的难题。^{②⑨} 整套算法模型是生成式人工智能证据的核心载体。算法模型设计是否科学、合理,运算逻辑是否存在偏见,是证据审查的核心要点。通过算法得出的结论是人工智能证据审查工作的落脚点。最终结论能否与案件事实形成有效的关联,其内容是否客观和真实,是其是否具备案件证明效力的关键。

人工智能证据需要审查的各类要素并不是简单拼凑在一起的,数据输入、算法模型和结果输出等各个环节之间存在着严密的因果关系。审查对象的多层复合属性直接导致其审查思路要与传统证据的审查思路区分开来。传统证据的核查方式相对简单,一般包括直接核对原件、对多份材料交叉比对、进行司法鉴定和询问当事人等多种形式。而审查人工智能证据需要融合多种技术和法律审查手段,不仅要核查基础数据的真伪,评估整套算法运行的稳定性,还要追溯完整的运算流程,判断最终结果和案件之间的关联程度,明确算法是否具备可解释性等。算法具备可解释性不等于必须对外公开算法代码,其核心要求是清晰地梳理整套程序的运行逻辑和运算原理,让办案人员能够理解算法的推演路径。要评判算法模型是否稳定和可靠,需要专业技术人员针对模型设计思路、训练全过程、测试与校验结果展开全方位的专业评估。要核验运算流程是否规范,要借助专业技术手段对全部的运算步骤进行复原,核对每一步操作是否与预设算法逻辑相贴合。^{③⑩} 由此可见,审查人工智能证据时,不能只对单一内容进行评判,必须综合全部关联要素进行分析。人工智能证据和传统证据在核查对象上存在区别,决定了对人工智能证据的运用需要构建专属的审查规范。

^{②⑧} 参见前注①⑥,王燃文,第62-65页。

^{②⑨} 参见郭锐:《人工智能的伦理和治理》,法律出版社2020年版,第176页。

^{③⑩} 参见周翔:《算法可解释性:一个技术概念的规范研究价值》,载《比较法研究》2023年第3期,第196页。

第二,人工智能证据对传统证据审查方法、可采性规则等形成冲击。我国传统证据审查规则依托证据本体属性和生成逻辑,形成了直观审查、比对审查、鉴定审查、询问审查等多元化的审查方法,但传统的审查方法难以适配人工智能证据审查对象多元这一特性。仅采用传统审查方法法官很难理解人工智能深度学习算法的训练过程和运算逻辑,难以判定是否存在算法漏洞,也就无法对输出结果进行科学评估。审查方法适配不足使得司法机关无法对人工智能证据开展全面有效的审查。人工智能证据在可采性方面也存在挑战。如果证据不具有可采性,将不能对裁决发挥作用。^{③①}人工智能在实际应用中曾暴露出的算法黑箱问题,影响了人们对人工智能技术的信任度。^{③②}人工智能证据采信的逻辑应是技术理性验证,也就是说,需要借助技术审核、专家评估、算法溯源等手段,对输入数据的真实性进行验证,构建科学的算法和运算规范,从而增强输出结果的可信度。^{③③}

四、人工智能证据独立性的规则因应

人工智能证据的运用对传统证据规则造成了冲击,传统的证据准入规则、审查规则等无法适配其本体独特性。要解决这些问题,就必须突破传统证据规则的束缚,赋予人工智能证据独立的法定证据地位,并构建适配其本体独特性的证据规则体系。当然,构建人工智能证据的证据规则,并非全盘推翻现有的证据规则体系,而是要在坚守一般规则的前提下,针对人工智能证据的独有属性,修正传统证据规则存在的不适之处。

(一) 人工智能证据的独立定位

明确人工智能证据为新的独立法定证据类型,是相应规则体系构建的逻辑起点。唯有赋予人工智能证据独立的法定地位,才能充分尊重其本质属性。这一制度设计基于人工智能技术的发展和现实司法需求,将填补新型证据形态的规制空白。

结合我国现行法律体系和司法实践难题,明确人工智能证据独立的法定地位,应采用统筹修订和分层细化相结合的实现路径。从宏观层面而言,应统筹修订三大诉讼法的相关条款,以明确人工智能证据的独立法定地位,结合刑事诉讼、民事诉讼、行政诉讼的共性需求和证据规则要求,增设人工智能证据这一独立的证据类型。这种统筹修订方式能够化解不同诉讼程序中对人工智能证据地位认定的分歧,避免在规则适用上产生争议。在微观层面,需要借助司法解释和专项规范进行分层细化,完善具体适用依据。三大诉讼法的修订主要解决人工智能证据的法定地位与核心定义问题,却难以覆盖收集、固定、举证、质证、认证全流程的具体规则。对此,可以制定人工智能证据专项适用规则,重点明确取证主体资质、收集提取程序、固定保全方式以及质证等相关内容。

在明确人工智能证据法定地位的过程中,需同时兼顾人工智能技术快速迭代的特

^{③①} 参见张保生主编:《证据法学》(第4版),中国政法大学出版社2023年版,第21-22页。

^{③②} 参见苏宇:《算法规制的谱系》,载《中国法学》2020年第3期,第167-169页。

^{③③} 参见前注^{③①},马国洋书,第259页。

性,妥善处理法律稳定性与技术发展性的关系。专项规范可以采用原则性规定加动态调整模式,在人工智能证据的核心定义与基本规则设定上采用抽象概括表述,避免对技术细节作出过于具体的限定。同时建立规则动态调整机制,赋予司法机关结合技术发展与实践需求细化补充规则的权限,通过出台指导性案例、修订专项规则等方式动态优化相关规则,以确保证据规则既保持相对稳定,又能及时回应技术发展带来的新挑战。

(二) 人工智能证据的准入规则

证据准入是证据材料适用于司法证明的第一道门槛,其核心功能在于排除不具备证据能力的材料。对于人工智能证据准入资格问题,宜采用合法性、相关性、可靠性三重审查维度。合法性审查要聚焦证据收集、固定流程的法定性,这是所有证据审查时不可或缺的基础要件。相关性审查则指向证据与案件事实的实质关联性,这是证据发挥证明效力的前提。可靠性审查聚焦人工智能证据生成全流程的合规性和结果真实性,是人工智能证据具备合法性和可采性的关键。三者紧密衔接、互为支撑,共同构成人工智能证据准入资格的审查框架。

1. 合法性审查

对人工智能证据的合法性审查主要围绕收集、固定、保管等程序的合法性展开,核心是确保收集流程符合法律规范的规定。在收集主体合法性层面,刑事诉讼公诉案件和检察公益诉讼中的证据收集主体分别为公安机关、检察机关等具有法定侦查或调查权力的机关,其他单位与个人未经法律授权不得擅自开展取证工作;民事诉讼、行政诉讼、法律规定的机关和有关组织提起的民事公益诉讼中,当事人可自行收集证据,但必须严守法律划定的边界。若人工智能证据涉及国家秘密、商业秘密、个人隐私,取证主体还须具备对应资质,并严格履行保密程序。证据收集程序合法性要求遵循先固定原始载体、后提取证据内容的原则;针对云端存储等无法固定原始载体的情形,应当采用在线提取、镜像备份等技术手段固定证据,同步制作详尽提取笔录,完整记载提取时间、提取地点、操作流程、提取人员、见证人等全部信息。严禁采用未经授权侵入信息系统、非法获取他人数据等违法手段取证,通过此类非法方式获取的人工智能证据应当依法予以排除。证据保管程序合法性要求建立专门化证据保管制度,明确保管主体、保管方式与保管期限,并配套加密、异地备份等技术防护措施,防范证据出现篡改、丢失、信息泄露等问题。

2. 相关性审查

对人工智能证据相关性进行审查时,宜采用“实质关联+证明价值”双重标准。^④评估证据的证明价值,也就是评估证据对案件事实的证明作用大小。证明价值极其微弱的,即使存在实质性关联,也可以酌情不予采纳,这将有力提升诉讼效率。法官可以通过经验推论完成对事实的认定,最终作出中立的裁判。^⑤在人工智能证据相关性审查的过程中需要避免陷入技术崇拜的误区,不能仅凭其技术含量高就认定其与案件事实具备相

^④ 参见前注⑨,马国洋书,第259页。

^⑤ 参见前注⑩,张保生主编书,第19页。

关性,而要结合案件的具体情况进行综合判断。对大数据分析报告等复杂的人工智能证据,可以借助专家辅助人判断其与案件事实的关联程度。

3. 可靠性审查

可靠性是指单一证据或证据来源本身值得采信的属性。^{③⑥}可靠性审查是人工智能证据资格准入的核心之一,也是人工智能证据审查区别于传统证据审查的关键所在,需要从数据可靠性、算法可靠性和系统可靠性三个层面全方位构建可靠性审查体系。

首先,数据可靠性是人工智能证据可靠性的基石。与合法性审查往往聚焦收集、固定与保管等程序层面的合法性不同,数据可靠性审查针对的是作为人工智能证据生成基础的输入数据本身质量,其中包括数据来源的可靠性、真实性、准确性、完整性与代表性等,程序合法只是数据来源可靠的一项基础要求。就数据来源而言,输入的数据必须合法,必须通过合法途径收集,不存在非法获取的情形。比如,医疗人工智能系统的数据需要来源于经过患者授权的合法医疗数据,而非通过非法途径获取的隐私数据。同时,数据的真实性和准确性也至关重要,要考虑到数据存在出错的可能性。对此,需要审查输入数据是否存在篡改、伪造、错误等问题。第三方数据平台提供的数据,应当来源于资质与信誉良好的平台,必要时要求提供方出具真实性证明文件。^{③⑦}在数据完整性和代表性方面,要确认输入的数据是否完整覆盖相关领域的信息,是否存在缺失或者片面性问题。

其次,算法可靠性是人工智能证据可靠性的核心。算法是人工智能系统生成结果的逻辑基石。^{③⑧}人工智能系统录入数据与最终输出结果之间,存在难以被人类完整认知的隐藏层级,这也是学界所指的算法黑箱。^{③⑨}算法黑箱的核心问题在于信息不对称和不公开,算法黑箱特性增加了人工智能证据审查的难度。^{④⑩}对此,需要采用“形式审查+实质核验”的双重方式。其中,形式审查要求证据提供者提交算法的设计原理、逻辑框架和参数设置等说明性文件,审查是否存在明显的逻辑缺陷或者偏见。在对人工智能证据进行实质核验时,可以采用“第三方测试+交叉测试”的模式,由具备法定资质的第三方技术机构测试算法正确性和稳定性,同时,将系统输出结果与其他证据进行交叉验证。对涉及商业秘密的算法,可以采用“有限披露+保密审查”的方式,由法官在保密环境下审查,或者委托第三方机构进行保密性测试,这样既保障了算法的保密性,又确保了审查的有效性。

最后,系统可靠性是人工智能证据可靠性的保障。人工智能系统的硬件设备、软件环境和运行状态都会影响最终生成结果的可靠性。对系统可靠性进行审查,需要重点关注硬件配置是否达标、有无故障、性能是否正常等问题,同时要关注软件环境是否安全稳

^{③⑥} 参见前注^{③①},张保生主编书,第28页。

^{③⑦} 参见陈瑞华:《实物证据的鉴真问题》,载《法学研究》2011年第5期,第131页。

^{③⑧} 参见前注^{③⑤},马国洋文,第171页。

^{③⑨} 参见徐凤:《人工智能算法黑箱的法律规制——以智能投顾为例展开》,载《东方法学》2019年第6期,第79页。

^{④⑩} 参见王聪:《“共同善”维度下的算法规制》,载《法学》2019年第12期,第69页。

定,是否存在恶意病毒等安全隐患,审查系统是否经过相关质量认证与检测,是否符合行业或者国家安全标准,等等。

4. 附条件准入的特殊情形

在一些特殊情形下,人工智能证据材料虽然无法完全满足上述准入标准,但满足特定条件时,也可以允许其作为适格证据进入诉讼程序。这一规定不仅能充分发挥人工智能证据的证明作用,也能防范潜在风险。在紧急情况下收集的证据,比如在重大刑事案件中为了及时侦破案件而紧急提取的人工智能证据,尽管没有完全遵守法定的收集程序,但并未严重侵犯公民合法权益,这种情形下的人工智能证据可以作为证据使用。附条件准入的人工智能证据,需要明确准入条件以及后续审查要求,通过补充程序弥补现存的缺陷。

机器本身并不涉及主观诚实与否的问题,不过其运行所依托的算法程序却可能潜藏各类瑕疵,包括算法偏差、算法歧视和算法不透明带来的黑箱难题。^⑪对于瑕疵人工智能证据材料的后续审查,要求证据提供者在规定的时限内补充提交第三方测试报告和算法说明等可靠性材料。无法证明补充材料具有可靠性,或者通过验证发现其存在严重缺陷时,该证据材料应当予以排除。此外,附条件准入的人工智能证据在庭审中需要接受严格的质证和审查。

(三) 人工智能证据的审查规则

证据是串联过往事实与当下案件事实的核心纽带,在诉讼中,纳入审查认定的有效证据越充分,办案人员精准认定事实的可靠程度就越高。^⑫根据人工智能证据的独特性,人工智能证据审查规则的构建主要可以从以下方面展开:

1. 建立分类审查规则

根据人工智能证据的外在表现,可以将人工智能证据分为智能分析报告类证据、智能识别结果类证据、智能预测和溯源类证据、智能数据合成类证据四类。这种划分方式便于法官和侦查人员快速对不同人工智能证据载体进行区分。但是,仅仅依靠外在形式分类不足以支撑司法机关从规则维度审查人工智能证据的真实性和可靠性。基于此,可以根据人工智能证据在生成阶段的人为介入程度,进一步将上述四类人工智能证据分别归入自主生成型、辅助生成型和技术处理型三大类别,构建一套能够适应算法潜在风险并区分审查强度的分类审查机制,化解审查对象多层复合性带来的难题。

自主生成型人工智能证据的生成无需工作人员持续介入,由人工智能系统自主运算所产生的资金往来分析证据、涉案人员行踪追溯证据,以及全程依靠人工智能系统独立运算得出的人脸比对证据等,都属于这一类。辅助生成型人工智能证据是由办案人员完成对案件信息的筛选与整理,并设置对应的运行参数,再交由人工智能系统处理所生成的。图像合成、多维度数据重组后所形成的证据,大多属于这类证据。技术处理型人工智

^⑪ 参见汪庆华:《算法透明的多重维度和算法问责》,载《比较法研究》2020年第6期,第163-166页。

^⑫ 参见张保生:《事实、证据与事实认定》,载《中国社会科学》2017年第8期,第117页。

能证据是借助人工智能工具完成声纹匹配、破损图像修复等技术操作所获得的证据。对于自主生成型人工智能证据的审查,核心在于生成全过程无人干预,应采用严格的可靠性审查规范。对于辅助生成型人工智能证据的审查,审查重点应聚焦人类干预的合法性、合理性,同时兼顾人工智能系统的可靠性,必要时需要干预主体出庭说明干预过程与理由。针对涉及重大利益的案件,法院可主动委托第三方机构对人类干预的合理性与系统可靠性进行验证。对技术处理型人工智能证据的审查,其可靠性主要取决于原始证据的合法性、真实性以及技术的可靠性,需要采用原始证据与处理技术相结合的审查模式。

2. 建立人机协同纠错规则

人工智能证据并非由特定人员直接出具,传统证据制作者自行核查并把关的操作方式很难适用。为此,必须从制度层面建立人机协同的纠错机制,以防止因算法失误导致的裁判偏差。在审查过程中,工作人员需要留意原始数据中潜在的地域、性别等各种带有歧视倾向的内容,必要时可要求举证方详细说明,确认系统是否已完成相应的偏差修正。针对涉及定罪量刑、重大财产归属等核心事实的人工智能证据,应该增设技术复核流程,委托具备资质的第三方机构或者专业技术人员再次对人工智能证据涉及的算法逻辑和运算结果进行复核,若复核结论与原结果存在出入,则不能直接采信最初的人工智能证据。

3. 构建人工智能证据可采性验证规则

人工智能证据资格准入解决的是人工智能证据能否进入诉讼程序的问题,而人工智能证据可采性验证规则解决的则是法院基于技术性验证进一步确认其证据能力的问题。构建人工智能证据可采性验证规则,可以从以下层面展开:

其一,确立独立验证的可采性前提。对人工智能证据的审查,传统审查方法难以直接适用。独立验证要求法庭不依赖于证据提供方的单方说明,而是通过外部手段对人工智能证据的生成过程与结果进行科学核验。独立验证包括数据验证、算法验证与结果验证。数据验证,即对输入数据的来源、完整性和准确性进行独立的核查,以确认数据未经篡改。算法验证要求围绕透明度、合法性、公平和公正等核心价值,对算法模型的运行逻辑进行科学测评,明确其是否存在系统性偏差或者逻辑缺陷。^④ 结果验证,即在相同或者相似条件下,对人工智能系统进行重复运行,以检验输出结果是否具有 consistency。独立验证并不是要求法庭自行开展独立验证,而是要通过可操作的验证程序,使法庭能够对人工智能证据的可靠性形成合理判断。对无法通过独立验证的人工智能证据,原则上应不予采信。

其二,确立技术理性的可采性验证标准。人工智能证据生成过程涉及复杂的技术,只是凭借经验法则难以对其可靠性作出准确的判断。因此,可采性验证要从单纯的经验判断转向技术理性验证模式。^④ 技术理性验证的核心要求是将技术标准转化为法律上可

^④ 参见张欣:《算法影响评估制度的构建机理与中国方案》,载《法商研究》2021年第2期,第109页。

^④ 参见冯俊伟:《机器生成的电子数据之可采性》,载《中国刑事法杂志》2024年第3期,第69-70页。

操作的可采性验证标准。比如,对于大数据分析报告类人工智能证据,可以要求提供方说明分析模型的数据规模和样本所覆盖的范围,以对结论的可靠性程度进行评估。技术理性验证并不是以技术标准替代法律判断,而是为法庭的可采性判断提供科学的技术支撑,是否采信仍然要由法官结合案件的综合情况作出裁量。^{④5}

其三,构建与分类审查规则相对应的分层可采性验证规则。根据人工介入程度不同,人工智能证据可以分为自主生成型、辅助生成型和技术处理型,不同类别的人工智能证据在风险程度上存在差异,对此,可采性验证标准应当有所区分。对于自主生成型人工智能证据,其生成过程无人干预,算法黑箱风险最为突出,对其要适用最为严格的可采性验证标准,要求证据提供方完整披露算法运行原理并且提供第三方验证报告,否则法庭可以不予采信。对于辅助生成型人工智能证据,可采性验证的重点在于人类干预的合法性、合理性,对于人工智能系统的可靠性,可以适当降低对算法完全披露的要求。对于技术处理型人工智能证据,如果原始证据合法,而且技术手段经过充分验证,则可以适当放宽可采性验证标准。

其四,明确不予采信人工智能证据的情形。如果有以下情形,人工智能证据原则上不予采信:一是算法存在系统性歧视或者偏差且无法通过技术修订消除。二是证据提供方拒绝配合独立验证,或者在规定的时限内未能提供必要的算法说明和测试报告等应当提供的材料。独立验证是可采性判断的前提,缺乏验证基础的人工智能证据,无法排除合理怀疑。^{④6}三是人工智能系统的运行环境存在重大的安全风险,可能导致结果或数据被篡改或伪造。四是同一人工智能系统在类似案件中曾出现严重错误,且证据提供方未能证明该问题已经得到有效解决。通过明确证据排除的情形,不仅能够为司法实践提供清晰的指引,而且也能够倒逼人工智能证据提供方提升证据质量。

五、结 语

为紧跟人工智能发展的步伐,需要加紧制定并完善适配人工智能治理的法律法规。^{④7}人工智能技术深度融入当前司法领域,催生出新型的证据形态。人工智能证据能否取得独立的法定地位已经成为证据法学无法回避的时代命题。本文立足人工智能时代发展和司法改革的双重背景,系统地研究人工智能证据的独立性,以期摆脱其在司法实践中的适用困境。人工智能证据具有区别于现有法定证据类型的本体独特性,归入任何现有证据类型都面临着理论逻辑和实践操作上的双重困境,其独立地位具有正当性。需要通过立法明确其独立的法定证据地位,构建以可靠性为核心、差异化为导向、公平性为目标的制度化、体系化证据规则,以实现人工智能技术赋能司法发展。

^{④5} 参见纵博:《人工智能在刑事证据判断中的运用问题探析》,载《法律科学》2019年第1期,第68页。

^{④6} 参见胡铭:《论数字时代刑事证据的三元结构》,载《中外法学》2025年第1期,第57页。

^{④7} 参见黄文艺:《法治规划、建构理性与中国式法治现代化——对党的二十届四中全会精神的法治解读》,载《法学家》2026年第1期,第12页。

人工智能证据的出现既是对传统证据规则体系的挑战,更是推动证据制度现代化的契机。构建人工智能证据规则体系既适应了人工智能证据的自体特性,又契合了司法公正的基本要求,为保障人工智能证据独立地位提供了具体的制度方案。随着人工智能证据规则体系的不断优化,人工智能证据会在各类诉讼案件中发挥越来越关键的作用,有助于实现让人民群众在每一个司法案件中感受到公平正义这一目标。

Abstract: The application of artificial intelligence evidence in judicial practice is increasing, but there is a significant controversy in the academic community over whether it should be regarded as an independent type of evidence. Artificial intelligence evidence refers to evidence materials that are generated by an autonomous data processing system with artificial intelligence as the core, driven by algorithms to autonomously or semi-autonomously generate, carrying objectivized information capable of proving the facts of the case, and serving the litigation proof activities. Assigning artificial intelligence evidence to existing evidence types such as documentary evidence, expert opinions, and electronic data all have logical and practical application dilemmas. Artificial intelligence evidence has unique ontological characteristics in aspects such as generation subject, generation mechanism, and manifestation form that are different from traditional evidence, and it also forms systematic impacts on the admission rules, review rules, and other rules of traditional evidence. Therefore, analyzing from the three aspects of classification verification, essential analysis and rule evaluation, establishing the independent evidence status of artificial intelligence evidence has a theoretical basis. On this basis, artificial intelligence evidence should be added as an independent evidence type; establishing evidence admission standards centered on legality, relevance, and reliability, and setting an exception mechanism for conditional admission; establishing evidence review rules such as classified review and human-machine collaborative error correction, thereby constructing a systematic set of rules for artificial intelligence evidence.

(责任编辑:赵毅宇)