

刑事诉讼中人工智能证据的可采性困境与应对

□ 高宇石

刑事焦点

人工智能证据正以效率革命的姿态渗透刑事司法全流程，同时对以真实性、关联性、合法性为核心的传统证据制度构成巨大挑战。随着《人工智能安全技术框架》2.0版将技术内生安全风险细化至算法模型与数据层面，以及《最高人民法院关于规范和加强人工智能司法应用的意见》确立透明可信、辅助审判等核心原则，我国已初步构建人工智能治理的制度矩阵。但是刑事司法领域尚未建立与之匹配的审查技术规范与程序机制，这种规制滞后性正在加剧技术应用与程序正义之间的张力。亟须从类型识别、困境剖析与制度建构三个维度进行探索。

一、人工智能证据的类型特征

人工智能技术在司法领域推进融合。由算法系统自主或半自主生成的信息正逐步进入刑事诉讼证据体系。所谓“人工智能证据”，是指依托机器学习、深度神经网络、自然语言处理等AI技术，对原始数据进行分

析、推理、合成或预测后形成的、用于证明案件事实的材料。区别于传统电子数据，人工智能证据的核心特征在于其非直接感知性与算法中介性。简言之，证据内容并非源于人类感官记录或设备被动采集，而是经由算法模型“再加工”甚至“再创造”而成。据此，当前司法实践中的人工智能证据可以大致分为三类：

第一类是感知增强型证据。此类证据以智能传感设备为基础，通过AI算法对原始感知数据进行识别、标注或结构化处理。典型如人脸识别系统自动生成的身份比对报告、智能交通监控系统输出的车辆轨迹分析图、无人机搭载AI视觉模块识别的现场异常行为标记等。其识别结果高度依赖训练数据质量与算法精度。

第二类是分析推理性证据。该类证据通过对海量结构化或非结构化数据进行建模分析，形成具有解释性或预测性的结论。例如，在经济犯罪案件中，检察机关利用AI资金穿透工具生成的涉案账户资金流向图谱；在毒品犯罪侦查中，基于通信大数据构建的社交关系网络图；又如法院委托第三方机构使用风险评估模型出具的被告人再犯可能性评分。此类证据常因算法黑箱而难以被控辩双方有效质证。

第三类是内容生成型证据。以生成式人工智能为代表，AI已能生成

文本、语音、图像甚至视频内容。例如，智能语音助手记录并转写的嫌疑人对话摘要、AI根据案情描述模拟生成的作案过程动画、大语言模型辅助撰写的法律意见书或证人陈述草稿等。此类证据最具争议性，其内容可能包含虚构、幻觉或偏见，且难以追溯原始信息源，真实性与可靠性面临严峻挑战。

上述三类证据虽然形态各异，但是共通之处在于其生成过程脱离了传统人和物的直接关联，转而依赖算法这一技术中介。这使得其证据属性既不同于书证、物证，也超越了一般电子数据的范畴。

二、人工智能证据可采性的困境

尽管人工智能证据在提升司法效率、辅助事实查明方面展现出潜力，但其进入刑事诉讼场域后，与现行证据规则和程序原则之间产生了深刻张力，集中体现为真实性认定难、合法性边界模糊、关联性判断失准等多重困境。

首先，真实性审查面临算法黑箱障碍。刑事诉讼法第五十条规定，证据必须经过查证属实，才能作为定案的根据。然而，人工智能证据的真实性高度依赖其底层算法逻辑与训练数据质量，而多数商用或内部部署的AI系统具有封闭性和不可解释性。此类黑箱输出使得传统质证、认证机

制难以有效运行，动摇了证据裁判原则的根基。

其次，合法性判断缺乏明确规范指引。人工智能证据涉及数据采集、模型训练、自动推理等多个环节，其中任一环节若违反法定程序，均可能构成非法证据。例如，公安机关在未取得搜查令的情况下，利用爬虫工具从社交平台批量抓取嫌疑人历史言论用于行为画像，该行为是否属于刑事诉讼法第五十六条所禁止的以非法方法收集证据？在实践中相关司法解释尚未就人工智能介入取证的程序边界作出细化规定，导致合法性审查缺乏统一标准。

再次，关联性与证明力评估易受算法偏见干扰。人工智能证据常以统计概率、风险评分或模式匹配形式呈现，其与待证事实之间的逻辑链条往往间接且脆弱。而刑事诉讼法第五十五条强调重证据，重调查研究，不轻信口供，对间接证据的证明力提出更高要求。人工智能证据若缺乏透明的推理路径，极易沦为技术包装下的主观臆断。此外，人工智能证据的专业性与复杂性也加剧了控辩双方的信息不对称。被告人及其辩护人往往既无力理解算法逻辑，也难以获取训练数据或模型参数，导致质证权、辩论权形同虚设。

最后，人工智能证据在可采性层面所面临的困境，本质上是技术

逻辑与司法逻辑的冲突。若不及时构建适配其特性的程序规制体系，不仅可能侵蚀证据裁判原则，更可能动摇刑事诉讼中的人权保障底线。

三、构建人工智能证据审查的规范路径

面对人工智能证据带来的可采性挑战，不能因噎废食，亦不可放任自流。应在坚持以审判为中心、证据裁判、程序正义等基础上，立足我国司法制度实际，构建“分类识别—分层审查—权利保障”三位一体的人工智能证据审查规范体系。

第一，确立分类分级的准入审查机制。鉴于人工智能证据类型多样、风险程度不一，应实行差异化审查标准。对于感知增强型证据，重点审查算法模型的准确性、误识率及数据采集的合法性。对于分析推理性证据，应要求提供算法逻辑说明、训练数据来源及验证测试结果。对于内容生成型证据，原则上不得单独作为定案依据，须有其他证据印证，且需严格审查其生成过程是否存在幻觉。可借鉴《最高人民法院关于规范和加强人工智能司法应用的意见》，对高风险人工智能证据设置更严格的准入门槛。

第二，完善真实性与合法性的程序化验证规则。建议在现有电子数据审查规则基础上，增设人工智能证据专项审查要素。一是算法可解释性要求，对于影响定罪量刑的关键人工智

能证据，应由提供方说明核心算法原理、输入输出逻辑及误差范围。二是数据溯源与完整性保障，通过区块链、可信时间戳等技术固定训练数据来源、模型版本及生成日志，确保全过程可追溯。三是合法性前置审查，明确禁止使用未经法定程序获取的数据训练司法人工智能模型，严禁以技术便利替代搜查、扣押等强制措施审批程序。必要时，可引入第三方技术机构进行独立验证，提升审查公信力。

第三，强化当事人权利保障机制。人工智能证据的使用不得减损被告人的辩护权与质证权。一方面，应细化刑事诉讼法第一百九十七条专家辅助人制度，明确其有权查阅与人工智能证据生成相关的核心参数（涉密除外）、参与算法说明听证，并就技术缺陷提出质疑。另一方面，建立人工智能证据披露义务，控方或法院拟采纳人工智能证据的，应提前向辩方提供技术摘要、验证报告及专家联系方式，保障辩方充分准备质证。对于重大、复杂案件，可探索设立技术争点整理程序，由合议庭先行确定人工智能证据争议焦点，避免庭审流于形式。

总之，对人工智能证据的规制，不是要阻挡技术进步，而是要将其纳入法治轨道。通过清晰的规则、透明的程序和充分的权利保障，实现技术赋能司法与程序守护公正的有机统一。

（作者单位：北京理工大学法学院）

问题探讨

交易习惯视角下金融诈骗罪的出罪适用

□ 孟令航

金融市场的有序运行一定程度上依赖于经长期积淀形成的交易习惯，其合法性在民事领域已获得明确认可。然而，在刑事司法实践中，部分符合行业惯例且具有社会相当性的行为，往往因形式上符合欺诈特征而面临不当入刑的风险。这种忽视交易习惯出罪功能的倾向，破坏了法秩序的统一性，更与司法目标相悖。因此，亟须正视交易习惯的法理价值，通过类型化建构厘清金融诈骗罪的人罪边界，实现刑事评价与行业实践的精准衔接，在维护金融安全的同时护航经济高质量发展。

一、现实之需：交易习惯与形式主义入罪的冲突

交易习惯作为商事活动中长期积淀形成的非正式规范，其规范价值已获学界与实务界的广泛重视。有商法学者通过系统梳理司法裁判，提炼出189项典型商业行规；地方层面如广东民商法学会亦组织专家开展系统性调研，形成专项报告以证实习惯对经济的积极效用。然而，刑事司法实践中涉及“交易习惯”能否从宽处罚或出罪的争议案件数量正呈爆发式增长。笔者通过北大法宝检索发现，2015年至2020年此类刑事案件从215件飙升至567件，5年增长率超160%，且近年仍维持高位。数据背后折射出在商事犯罪认定日益复杂的当下，以“交易习惯”作为罪轻或无罪抗辩的案件日益增多。面对“形式入罪”的泛化风险，最高人民法院已关注到机械司法的弊端，并呈现出对交易习惯实质判断的裁判倾向。例如在“赵某诈骗再审案”中，赵某因在钢铁贸易中4次提货后未及及时交回结算单，且存在对账争议被控诈骗。原审法院忽略了交易双方“多次提货、滚动结算”的长期惯例，将赵某依循习惯的未交单行为被机

械认定为“隐瞒真相”的欺诈手段。但最高人民法院再审该案时指出，“提货与付款不一对应”系双方认可的交易习惯，赵某据此持续付款证明无非法占有目的，遂改判无罪。然而，这一案例虽然确立了实质判断的导向，但如何在个案中精准识别习惯与犯罪的边界仍缺乏统一标准，易导致类案处理偏差。

相比于一般商事犯罪，金融诈骗罪往往更具专业性与复杂性，导致部分实质上符合行业惯例的行为，更容易因形式上具备“虚构事实”的外观而面临入罪风险。例如，在大宗商品贸易中，行业长期形成“延迟付款”的习惯，若在融资中存在欺诈外观，后仅因回款周期长导致资金暂时未到位，便比对金融诈骗罪构成要件，极易误判。再如，建筑与经销行业中普遍存在的“滚动核销”习惯，即多项目间资金混同使用、不依单笔业务逐一对应结算的模式，也常引发定性争议。在“王某票据诈骗无罪案”中，王某与某公司长期家具经销合作中，双方以此模式结算多个项目，王某伪造公司印章私自支取工程款，形式上具备“虚构事实、侵占款项”的欺诈表征，一审认定其构成票据诈骗罪与伪造公司印章罪并罚；但二审认为该行为本质是双方长期交易默契，王某无非非法占有目的，不构成金融诈骗。若司法机关忽视其背后的交易习惯，将此类旨在通过“时间换空间”维持经营的行为，因没有及时交付款项或货物，就一律认定为金融诈骗，不仅会误伤正常经营的企业，更会挫伤市场主体的交易信心。

更为关键的是，忽视交易习惯的出罪功能，会破坏法秩序的统一性。民事领域早已明确认可交易习惯的法源地位。原合同法设有11个条款、民法典以14处条文明确“习惯”的适用规则。实证数据显示，2017年10月至2020年12月的196份有关交

易习惯的裁判样本中，91.8%聚焦买卖、保险等金融商事纠纷，印证了其在金融领域的重要价值。根据法秩序统一性原理，若前置法已承认某种交易习惯的合法性或容忍性，刑法作为“最后手段”便不应将其评价为犯罪。否则，会导致同类行为在民刑领域评价截然对立，既让市场主体无所适从，也动摇了司法公信力的基础。

二、金融诈骗罪通过交易习惯出罪的法理思考

在当前经济形势下，尊重交易习惯对金融诈骗出罪，契合宽严相济的刑事政策。在经济上行期，金融活力较强，司法侧重金融安全；但在近年来经济增速放缓、民营企业面临转型挑战的背景下，为贯彻落实中央“六稳”“六保”工作要求，司法机关需从单纯打击转向服务发展。宽严相济政策在当下的具体展开，便是要求司法机关应重视交易习惯在金融诈骗认定中的实质出罪功能。这种理念在2018年发布的《最高人民法院、最高人民检察院关于办理妨害信用卡管理刑事案件具体应用法律若干问题的解释》中已有明确体现，该解释规定“有效催收”需两次间隔30日，并要求“非法占有目的”需综合多情节判断。其实质便是加重银行义务，尊重适当延迟还款的习惯。若无视企业融资中普遍存在的适度美化报表等符合行业惯例的行为，一律认定为金融诈骗，将阻碍经济复苏。

交易习惯承载着特定领域的行业共识，其出罪逻辑契合社会相当性理论对实质违法性的阻却。金融服务面向实体经济，不同行业的风险特征催生了特定的交易习惯。部分行为虽具欺诈外观，但符合行业内部的风险共担共识，应被排除在犯罪圈之外。以科技行业为例，因研发具有高度不确定性，银行适度放宽融资条件本质上是对“试错”的容忍，若仅因研发失

败无法还款，便将科技融资认定为欺诈，有悖行业共识。再如，房地产行业受回款周期长影响，形成依赖商票流转支付的结算习惯。房企在资金紧缺时开具商票，形式上极易被解读为“虚构偿债能力”的空头支票，但实质上是行业内为维持资金链连续性，基于未来回款预期而形成的融资惯例。上海票据交易所数据显示，2023年6月至11月持续逾期的承兑人中，房地产行业占比高达64.9%。这表明此类“拖欠”往往是受行业周期性特征影响而形成的结算惯例，而非恶意骗取，全盘否定不利于金融风险的软着陆。社会相当性理论认为，形式上符合构成要件的行为，若未突破社会常规秩序，则不具备实质违法性。此类融资行为虽存在形式上的不实，但属行业普遍接受的常态，未突破社会常规秩序，不具实质违法性。因此，不能简单以“形式欺诈”倒推犯罪，而应通过交易习惯认定其缺乏实质违法性，从而排除金融诈骗罪的适用。

从风险分配的角度看，交易习惯产生的风险属于金融交易中的“被允许的风险”，而非刑法意义上的“不法风险”。金融交易本质是风险与收益的共生，交易习惯承载的是行业固有风险。从被害人角度而言，其对于交易习惯可能产生的风险往往是知情且接纳的。因为习惯的产生就是双方在博弈中选择的一种风险分担方式。以恶意透支型信用卡诈骗为例，银行事后多会允许延期、展期，体现了其为维持客户资源而对风险的接纳。有学者通过研究2014年至2022年337个信用卡诈骗不起诉和无罪样本案件发现，行为人到案后能归还欠款的共284个，占比达84.3%，其中银行多对行为人表示谅解。这在一定程度上体现出金融领域适度延期还款符合行业习惯，未产生需要刑法规制的实质风险。实际上，金融领域有充分的风

险措施来消解交易习惯中的风险。例如，在抵押贷款业务中，企业虽可能为满足形式审查而虚构购销合同，但银行会要求企业提供足额的土地或厂房作为实物抵押。在这种情况下，银行的债权已被锁定在安全范围内。因此，在担保物真实足值且足以覆盖债权的情况下，即便借款理由存在形式瑕疵，因未产生实质性的资金安全风险，不应认定为金融诈骗。

三、金融诈骗通过交易习惯出罪的建议

首先，对于“借新还旧”这种常见信贷习惯，出罪边界在于资金是否投入实体经济以修复偿付能力。若资金用于清偿旧债及维持经营闭环，即便续贷手续存在夸大资产等瑕疵，也应认定该形式欺诈旨在通过审查，因缺乏非法占有目的而不构成金融诈骗罪。例如，“边某贷款诈骗案”中，边某在贷款无法偿还后，与银行达成合意进行续贷，进行欺诈骗取贷款。法院认为，信贷双方明知续贷资金用于填补前期贷款空缺，为化解不良贷款而作虚假贷款，所贷款项始终在金融机构控制之下，不能认定借款人主观上具有非法占有目的。反之，若资金脱离实际经营、依赖“以债养债”空转，则属金融诈骗的延续。例如，在“王某集资诈骗案”中，行为人虚构项目募集资金用于非经营性用途，无实体经济支撑却通过新增借款支付本息、扩充资金池，所谓“借新还旧”仅为骗局表象，法院认定其构成集资诈骗罪。此类案件中资金流向的非生产性已切断了交易习惯的正当性，行为入主观上已无修复法益的意愿，客观上亦丧失了清偿的可能。

其次，延迟交付作为大宗商品等行业的常见交易习惯，在适用金融诈骗罪时应区分延迟原因，判断其是否契合行业惯例本质。一方面，如果金融业务中存在欺诈，之后延迟还款的原因是回

款周期慢等惯例导致的资金链断裂，应该认可其交易习惯的正当性，慎重以金融诈骗罪定罪。例如，在“黄某票据诈骗二审改判无罪”中，上诉人与被诉单位之间长期以签发延期支票的方式支付货款，这种做法已成为双方认可的交易习惯。当黄某再次签发延期支票时，虽然形式上存在“空头支票”的外观，但实质上符合双方既往交易模式，且其向被害人表示待工程款收回后一并归还，一直有协商还款事宜。二审法院基于该行为符合双方长期交易习惯这一事由，撤销原判，宣告无罪。此类判决是对行业运行规律的尊重，也契合金融交易“风险共担”的本质逻辑。反之，若行为人已实现资金回笼或自身具备足额偿还能力，却以回款周期长等原因作为理由不予还款，甚至通过转移资产、隐匿行踪等方式逃避债务，则完全背离交易习惯“客观履行障碍”的条件，本质是借行业惯例作为犯罪伪装，其主观恶性与行为危害性均已符合金融诈骗罪的构成要件，理应依法追究刑事责任。

法治不仅意味着严格执法，也意味着对市场规律与商业自治的充分尊重。不仅在金融诈骗犯罪，乃至在经济犯罪中，刑法都应重视交易习惯的实务价值。在商事活动日益频繁的今天，成文规范难以周全。而法规范的形成恰恰是由符合社会利益的行为现象逐渐类型化、习惯化，最终上升为成文规则的过程。随着时间推移，众多契合行业规律的惯往往会演化成规范体系，重视并认可当前符合社会利益的习惯的出罪功能，正是对社会发展渐进性与法规范形成规律的尊重。反之，若机械否定所有交易习惯的正当性，无异于中断潜在的未来规范生成进程，与刑法服务经济社会发展的治理目标相悖。未来，刑法应不断重视习惯的识别并使其类型化，将符合行业共识与社会利益的交易习惯纳入入罪考量。

（作者单位：华东政法大学刑事法学院）