

人工智能时代犯罪新危害的多维审视与法律应对

——基于技术演进与刑事治理的双重维度

郭艳东¹, 黄美潼²

1 广州新华学院法学院、广东省网络与电子商务法学研究会, 18818875461@139.com

2 广州新华学院法学院, 3273880447@qq.com

摘要: 人工智能技术在深刻赋能社会发展、提升效率与生活品质的同时, 也催生了犯罪形态的智能化演变, 引发了传统犯罪升级与全新犯罪类型涌现的双重挑战。一方面, 网络犯罪、盗窃、诈骗等传统犯罪因 AI 技术的融入, 呈现出自动化、精准化与隐蔽性增强的新危害; 另一方面, 深度伪造、算法欺诈与操纵等基于 AI 技术特性的全新犯罪类型, 对公民人格权、财产权及社会秩序构成了前所未有的威胁。面对这些挑战, 当前司法实践在证据收集与认定、法律适用、跨区域管辖等方面存在明显困境。为有效应对, 本文从完善专项立法、强化技术侦查手段、推动跨学科协同治理等维度提出了法律应对之策, 并对未来技术发展可能带来的新型风险进行了前瞻性分析, 以期构建适应人工智能时代的犯罪防控体系提供理论参考。

关键词: 人工智能, 深度伪造, 犯罪防控

1 引言

当今时代, 人工智能(AI)已进入人们生活的各个方面, 从人们的日常生活--家居、交通再到我们的专业领域如医疗诊断、金融风控等等, 几乎涵盖了人们生活的主要方面。的确, 在当今社会中人工智能已为人类带来了前所未有的便利, 显著提升了人们的工作效率和生活品质。然而, 伴随着人工智能的迅猛发展和广泛应用, 犯罪活动的方式也随之发生了变化。传统意义上的犯罪受到人工智能的影响, 呈现出新的形式和更大的危害。新型犯罪日益增多, 特别是随着深度伪造技术、AI 换脸、智能模拟技术等的不断进步, 犯罪活动展现出智能化、高科技化的新特征和新形态。因此, 探讨研究人工智能时代犯罪的新危害有着十分重要的现实意义, 基于此提出一点法律应对之策, 实有必要。

2 人工智能时代传统犯罪形式的演变及新危害

2.1 网络犯罪的演变与新危害

人工智能技术应用到网络犯罪中, 让网络犯罪变成更加具有自动化、智能化的特点。如犯罪分子能用机器学习算法来进行全网扫描, 通过人工智能准确发现系统的漏洞, 并实施相应的攻击等行为, 这种网络犯罪手段明显具有更强的隐蔽性, 因为人工智能系统可以实现自动完成犯罪任务, 其行为很难被传统安全防护体系识别出来, 同时还能有效提升犯罪效率, 相对于传统的以人

作者简介: 郭艳东, 河南林州人, 法学博士, 广东金融学院与广州新华学院联合培养硕士生导师, 广州新华学院法学院讲师, 广东省法学会网络与电子商务法学研究会常务理事, 主要研究方向刑法学、金融不良资产。

黄美潼, 广东中山人, 广州新华学院法学院 2024 级涉外法治特色班本科生。

本文是广东省高等教育学会“十四五”规划 2024 年度高等教育研究课题青年课题“高等教育人工智能应用的法律风险及应对研究”(编号为 24GQN48)的阶段研究成果。

工操作为主的网络犯罪行为,在时间效率上,人工智能驱动型的网络犯罪能够在较短的时间内完成针对更多目标的攻击行为,这样的变化使得网络犯罪危害着广大民众的个人隐私、财产安全以及扰乱网络空间正常的秩序:一方面人们在网络上所留下的各种信息都极有可能成为不法分子的敛财之源;另一方面网络犯罪的强大渗透力与不可见性使得网络空间秩序已经到了一种难以为继的地步,严重地破坏了人们的工作环境,生活氛围。以上危害,已经严重侵扰到人们的生活安宁及财产安全等,随着网络跨境电诈犯罪的演化,甚至有侵害公民的人身、名誉权益之虞。

2.2 其他传统犯罪的演变与新危害

和传统犯罪一样,人工智能也为一些违法犯条例带来了新的手段与方式,例如:盗窃罪可以通过智能设备实现远程盗窃,比如:通过攻击并控制系统家中的智能家居系统来获取家中物品,或者是破解别人家的门锁实施入室盗窃。再比如诈骗犯罪,在人工智能发展进程中通过分析个人的信息、行为习惯等信息后可以更加准确地制造出骗局从而实施诈骗行为,并且受害人的数量也是成倍数的增长。同时,通过人工智能的设备进行远程的犯罪活动,不受地点以及时间和空间的制约,在进行犯罪的时候的影响力也会扩大,并且所带来的危害也会越来越大,对人们的财产造成一定的损害的同时也对社会的安稳及群众的安全感产生不良的影响。让群众感到自己的财产和生命安全都会出现问题,无法营造出一个和谐的社会环境。

3 人工智能催生的全新犯罪类型及危害

3.1 网络犯罪的演变与新危害

人工智能技术升级迭代带来的是算法犯罪化程度加深且呈现规模化的态势,其中典型的犯罪形式为算法欺诈和算法型市场操纵。在算法欺诈中,犯罪主体运用机器学习模型从用户的多维信息中提取特征,并依据消费者的消费行为、经济能力构建了更为精准的骗术模型。首先,犯罪主体运用机器学习算法建立虚假的投资平台,并利用用户画像数据发现目标用户并为其推送给具有更大概率会投资的项目,而且这种诈骗的方式也有极高的迷惑性以及较高的成功率^[1]。其次,针对市场操纵类的犯罪来讲就是使用了高频交易算法,利用其同时处理海量的数据并根据事先设置好的交易方案作出的买卖股票指令。也有一部分人在短期区间内把大量的证券产品集中在某一时点买或卖,其实质是做相反的行为以获利,在期间这种突然出现的异常大量操作可以引发的价格的变动对于正常人的影响就更大^[2]。

3.2 基于深度伪造技术的犯罪

深度伪造技术的技术特性决定其犯罪形态与其紧密相关,在特定语境下广泛应用到侵犯人格权和破坏社会秩序的犯罪当中。从技术实现途径看,基于生成对抗网络的面部表情合成系统能够对某一类人的容貌特征进行仿真模拟,并且通过结合语音克隆形成相关音频视频材料具有很强的欺骗性^[3]。由于此类伪造内容用作诽谤手段,以具体的犯罪主体使用场景还原术,将目标人物置于具体的违法情境中,直接造成被害人人身和社会评价的大范围降低。从公共领域角度看,其以虚假新闻生产、不实信息传播为主要行动对象的新型犯罪模式已经出现并不断升级,对于民众的认知和判断构成了直接的影响,也在实际上对信息传播生态系统的可信度形成致命的打击^[4]。

4 人工智能时代犯罪新危害在司法实践中的应对难题

4.1 证据收集与认定困难

人工智能犯罪证据数字化特质极为显著,以电子数据形式为主,如程序架构代码、网络交互日志等数字痕迹^[5]。具有极易被改动毁坏的特点,行为人可以采用技术性的手段对一些数据的核心字段进行无痕迹的篡改,使整条证据链条遭到严重破坏;在证据移送时由于存储介质具有分布

式的特点,传输路径是网状的拓扑结构,因而取证需要覆盖大量的终端设备和异构的网络节点,这就将证据移送的物理边界进行了有效的扩大。至于司法认定环节,一方面因为人工智能系统的运行机理非常专业,所以大量技术内容是普通法官无法知晓的,因此造成了技术上的认知鸿沟。无法完成对证据的内容解析及关联性的判断。另一方面,基于之前所述的技术性危害,在此过程中也会对人的能力造成一定的影响,使得人工智能证据没有完全得到生效证明力的结论,会对案件实体的判定带来一定影响。

4.2 法律适用模糊

现实生活中的人工智能犯罪目前并未获得相关法律约束,从规范结构上讲,并不存在类型化的罪名体系和阶梯式的量刑规则;从司法实践中来看,还无法就技术研发者、系统操作者以及智能体本体作为责任主体的选择达成统一,目前还没有确定责任归属以及各主体之间责任分配的具体准则。这种规范上的缺位也使司法机关在遇到此类案件的时候,难以找到适当的法律依据来作为裁判根据,在一定程度上影响了司法的统一性和权威性^[1]。

4.3 跨区域犯罪管辖难题

除此之外,由于人工智能犯罪地域跨度大,在以互联网为主的网络空间中,犯罪分子可以突破地域障碍,在数域同构状态下在不同司法区实施违法犯罪行为;如犯罪预备阶段发生在甲国,实行行为发生在乙国,结果发生在丙国,发生地点分为三个司法管辖区。空间分布的分散性状况会引发司法管辖权出现重叠与交叉的现象。每一个国家都依据自身的主权原则,主张在自己国家主权所覆盖的范围内,拥有最高层级且最为有效的司法管辖权力,这种基于主权的优先主张使得国际范围内的司法合作面临更大的挑战,合作的难度显著提升。与此同时,不同的国家在其法律体系的构建上存在差异,具体体现在实体法和程序法的规定各不相同。这种差异的存在导致各个地区在法规层面进行信息交流与沟通时,遭遇了不同程度的阻碍,信息无法顺畅地互通有无,进而使得各国之间在司法方面的相互联动难以切实有效地落实到位,影响了国际司法协作的整体效能^[6]。另在我国“一国两制三法系四法域”的实然背景下,不同法域之间也可能面临这样的问题,中国区际刑法学的概念也伴随而兴^[7]。

5 应对人工智能时代犯罪新危害的路径探索

5.1 完善法律法规

人工智能犯罪形态变化催生出对既有法律无法响应的问题,迫切需要建立专项立法,而立法的重点就是如何弥补之前的技术适配瑕疵,即探讨人工智能技术的特殊性对于犯罪行为模式的影响、是否包含具有自主决策能力的智能系统的主体身份等难题,并且,在此基础上还需要明确技术研发者作为行为人承担产品责任的法律属性及其与人工智能系统作为一个行为人的被害人或共同犯人的界限。另外,在今后人工智能的法律规范中应当重点注意填补智能犯罪定罪和量刑的相关规则,将二者规定入刑法学理论体系当中去,并且给以后的司法裁判提供适用的具体参照^[1]。

5.2 提升技术侦查手段

为了针对人工智能犯罪的技术对抗性特征建立专业化技术侦查体系,可把突破目标主要放在算法逆向解析工具、分布式数据追踪系统等两类专门研究突破对象上,研究相关算法逆向解析及分布式数据追踪的技术方法。对于获取证据的环节,需要考虑证据提取时可信计算环境下的证据数据固化和运用区块链存证技术开展证据完整性验证等问题,通过建立犯罪数据特征库和异常行为检测模型提高司法机关发现隐蔽性技术犯罪和固定电子证据的能力和水平^[8]。

5.3 加强跨学科合作

人工智能犯罪治理需突破单一学科认知局限,建立“技术-法律-行为”三维研究范式。在合作框架设计中,计算机学科负责构建犯罪技术特征分析模型,法学领域着重解决新型行为入罪化判定问题,犯罪学则提供犯罪收益成本分析工具^[9]。这种协同机制可通过组建专家论证委员会、建立技术事实查明规则等制度载体实现知识融合,最终形成涵盖犯罪预防、监测预警与损害救济的全链条治理方案^[10]。

6 人工智能时代犯罪新危害的未来趋势预测

6.1 技术发展带来的新风险

人工智能技术向高阶化发展的过程中,产生新的犯罪风险形式。具有自我决策功能的人工智能系统一旦被恶意使用或因系统本身的技术原因出现问题时,就有可能发挥出造成大范围破坏的功能。虽然这类具备较强攻击性的系统可以利用对大数据的挖掘分析功能进行精准化的犯罪策划,例如将针对关键信息基础设施的定向攻击作为重大违法犯罪活动的一种手段,但与此同时也将有可能严重危及到社会运行的根本以及国家安全。另外,系统自身拥有的自主性会导致系统产生无法准确预测的结果,现有的针对犯罪预防方式基于对行为模式的判断将会被完全影响。除此以外,人工智能技术运用到了包括医疗诊断、自动驾驶等关系到民生的重要领域后,如果遭受不合理利用将突破原先传统犯罪的危害界限,进而给人们带来更多的生命伤害和更大的社会恐慌感^[2]。

6.2 应对策略的前瞻性布局

法学研究须建立动态应对技术迭代风险的对策,制度上要实现法律规范和技术的发展同步化,在相关制度规则上预先考虑针对自主智能系统的具体行为存在过错时的规制框架。应建立健全技术防范体系,如研发具有环境适应性的智能监测系统,能够针对复杂算法做出解析并发现系统是否存在异常决策;打造人才储备,在相应制度中设置有关加强技术赋能法学专业人才培养的保障措施,注重培养法学和人工智能交叉复合型人才,以技术事实查明专家库等制度为依托,加强针对该类案件的技术解读,并能对未来可能出现的新型技术犯罪提前做出预警,做好处置准备^[4]。

参考文献:

- [1] 范聪聪.人工智能时代的刑事风险与刑法规制[J].法制与经济,2018,27(8):165-167.
- [2] 宋友艳.人工智能风险的刑法思考[J].经济研究导刊,2020,(22):151-152.
- [3] 程乐.“数字人本主义”视域下的通用人工智能规制鉴衡[J].政法论丛,2024,(3):3-20.
- [4] 芦丹.论人工智能时代的犯罪预防与治理[J].辽宁经济职业技术学院学报.辽宁经济管理干部学院,2022,(4):34-36.
- [5] 刘钊,林晞楠,李昂霖.人工智能在犯罪预防中的应用及前景分析[J].中国人民公安大学学报(社会科学版),2018,34(4):1-10.
- [6] 司伟攀.人工智能刑事责任主体地位及其分配[J].中国应用法学,2020,(6):172-186.
- [7] 郭艳东.中国区际刑法学论述纲要[J].中国式现代化与中国特色社会主义刑法学研究,2025,(3):103-113.
- [8] 刘钊,林晞楠,李昂霖.人工智能在犯罪预防中的应用及前景分析[J].中国人民公安大学学报(社会科学版),2018,34(4):1-10.
- [9] 焦旋,张懿轩,邹逸彬.人工智能时代刑事责任体系的重构[J].东南大学学报(哲学社会科学版),2023,25(S01):109-113.
- [10] 梁颖欣,梁雅淇.人工智能背景下刑法的挑战及应对[J].区域治理,2021,(43):65-67.

A Multidimensional Examination of New Criminal Threats in the Era of Artificial Intelligence and Legal Responses —Based on Dual Dimensions of Technological Evolution and Criminal Governance

Guo Yandong¹, Huang Meitong²

¹ School of Law, Guangzhou Xinhua University, Guangdong Society of Cyber and E-Commerce Law, 18818875461@139.com

² School of Law, Guangzhou Xinhua University, 3273880447@qq.com

Abstract: While artificial intelligence technology profoundly empowers social development, enhances efficiency, and improves quality of life, it has also catalyzed the intelligent evolution of criminal behavior. This has given rise to a dual challenge: the escalation of traditional crimes and the emergence of entirely new criminal types. On one hand, traditional crimes such as cybercrime, theft, and fraud have evolved into new threats characterized by increased automation, precision, and concealment due to the integration of AI technology. On the other hand, entirely new crime types based on AI's inherent characteristics—including deepfakes, algorithmic fraud, and manipulation—pose unprecedented threats to citizens' personal rights, property rights, and social order. In the face of these challenges, current judicial practice faces significant difficulties in evidence collection and determination, legal application, and cross-regional jurisdiction. To effectively address these challenges, this paper proposes legal countermeasures from multiple dimensions, including refining specialized legislation, enhancing technical investigative methods, and promoting interdisciplinary collaborative governance. It also conducts a forward-looking analysis of emerging risks that may arise from future technological advancements, aiming to provide theoretical reference for building a crime prevention and control system adapted to the era of artificial intelligence.

Keywords: Artificial Intelligence, Deepfakes, Crime Prevention and Control