刑事侦查中人工智能的应用:实践样态、风险挑战与发展策略

金益锋[1,2],马忠红[1*]

1. 中国人民公安大学侦查学院,北京 100038
2. 公安部鉴定中心,北京 100038

摘要　概述了刑事侦查中人工智能应用常见的实践样态,表明基于人工智能的视频侦查、个体识别、数据检索、情报挖掘、辅助审讯、证据评估等,为侦查机关带来多维锁定、情指一体、全链支撑的模式变革;分析了刑事侦查中人工智能应用所伴随的黑箱效应、精准偏差、算法僭越、全景敞视、侦查密室等一系列风险挑战;提出要锚定以人为本与人机共融的基本立场、推动数据驱动和知识指引的融合、增强刑事侦查中人工智能应用的伦理适配、优化刑事侦查中人工智能应用的法律体系等建议。

关键词　人工智能;侦查;技术效应;人机共融;知识指引

人工智能的快速发展与应用已造成社会政治、经济、文化的不断变化和波动[1],人工智能语境下犯罪的结构与态势开始剧烈变化,已经逐渐脱离传统犯罪的运行逻辑,传统犯罪呈现出明显的智能化特点,以网络犯罪、数据犯罪为代表的人工智能新型犯罪也层出不穷,传统犯罪的智能式异化与智能式新型犯罪的常态化演变并存,突显出人工智能背景下犯罪治理的难题[2]。侦查和犯罪两者之间属于互动博弈的关系,犯罪结构与态势的智能化演变,必然要求侦查具备更高的智能化水平,否则将不能实现发现、控制、揭露以及证实犯罪的目标[3]。

侦查部门作为多元职能融合机构,具有侦查、预防、服务职能[4]。侦查作为一种国家为公民提供的公共服务产品,在人工智能背景下其社会效果取决于侦查决策的智能水平;侦查是连接犯罪与刑罚的纽带,是犯罪治理预防的重要组成部分,侦查与

收稿日期:2023-01-13;修回日期:2023-02-17
基金项目:公安部科技强警基础专项(2021JC17);国家自然科学基金项目(52275535);中央级公益性科研院所基本科研业务费专项(2022JB040)
作者简介:金益锋,博士研究生,研究方向为侦查学、智能侦查、法庭科学,电子信箱:jinyifeng@cifs.gov.cn;马忠红(通信作者),教授,研究方向为侦查学、智能侦查,电子信箱:mzh2615@163.com
引用格式:金益锋, 马忠红. 刑事侦查中人工智能的应用:实践样态、风险挑战与发展策略[J]. 科技导报, 2023, 41(7): 15-27; doi: 10.3981/j.issn.1000-7857.2023.07.002

人工智能的结合符合国家人工智能发展规划的要求。对人工智能之于侦查的研究,不仅是一个关乎侦查技术变革与功能定位的时代题目,也是一个关系刑事司法以及社会综合治理未来走向的核心命题[5],故需要对刑事侦查中人工智能的实践样态进行梳理,总结面临的风险挑战,提出发展策略。

1 刑事侦查中人工智能应用的实践样态

为有效应对复杂多变的犯罪形势,刑事侦查对技术发展与革新有着天然的敏感性,人工智能的技术价值和刑事侦查的实战需求有着高度的契合性,在"发展优位""效率优先"理念的推动下,人工智能已快速渗透到刑事侦查的所有重要领域。

1.1 多维锁定:基于人工智能的视频侦查和个体识别

人工智能背景下的视频侦查能够实现对人、车、机、网、卡、物等多种目标轨迹的自动关联和追踪,智能摄像头开始嵌入到整个智能警务体系中,传统视频侦查的应用领域不断地拓展和深化,跨模态视频片段检索功能进一步扩增了视频侦查的威力[6]。随着视频监控与人工智能、物联网、云计算等信息技术进一步融合,其将变为智慧城市建设与智慧社会治理的重要智能终端以及数据来源,基于人工智能的视频侦查将逐步具备预警和预测功能[7]。

指纹、DNA、人脸识别等生物特征是侦查机关最常用的个体识别手段,其在人工智能的助力下出现了质的转变。中国的指纹自动识别系统(AFIS)于20世纪80年代初开始研发,21世纪以来相应系统广泛服务于侦查办案实践,伴随数据库容量的极速增长,基于特征点标划的指纹识别技术的识别准确率和有效性遭遇瓶颈,近年来以卷积神经网络、自编码器、卷积自编码器为代表的深度学习技术让指纹识别算法进入了全新模式,颠覆了侦查界对指纹识别技术原有的认知[8]。基于自适应小波框架和主动式深度学习研发的"云痕"智能指纹识别系统,能够让单枚指纹图像的表征信息量百倍于以往指纹识别算法,不需要人工标注特征即可实现即提即比,其比对精度在20亿量级的数据库中开展比对测试未出现衰减现象,比对精度显著胜过基于特征点标划的指纹识别系统,服务实战效果更好[9]。

基于深度学习技术,如Fast R-CNN系列方法、级联CNN系列方法以及SSD系列方法等,人脸识别技术具备主动学习、自动检测、智能检索等功能[10],能够完成人脸的监控、辨识、跟踪全流程应用[11],其在案件侦破、命案攻坚、寻找被拐儿童等专项行动中都发挥了重要的作用。2017年全国公安机关依托"依图"人像大平台,侦破案件10万余起[12]。北大弑母案嫌犯和命案逃犯劳某枝被抓就是人脸识别技术应用的典型案例。在人脸评测数据基准库(Labled Faces in the Wild, LFW)的测试中[13],基于深度学习的DeepID2人脸识别方法将LFW上的识别准确率刷新到99.15%,成为首个超过人类识别准确率(97.53%)的方法[14]。Jonathon等[15]专门针对人脸识别算法与法庭科学专家之间识别人脸的准确性进行研究,指出算法的识别准确性已经超过了鉴定专家,算法与专家协作识别时准确性最高。

Marciano等[16]将机器学习引入基因分型中,通过将静态峰阈值转变为动态峰阈值,有效地适用于复杂情景下的DNA分型检验,有助于解决假阳性假阴性的问题。DNA数据库中不仅有DNA分型数据,还有关于案件的时空、类别与手段等信息,同时还包括涉案人员的民族、地域、行为等数据[17],对DNA数据库关联数据的深度挖掘是增强服务侦查破案的关键一环。机器学习在数据挖掘方面具有分类、聚类、关联、分析、预测等明显的优势,在DNA大数据挖掘中能获得良好的应用效果,如宏观层面通过DNA数据能精准地确定个体的民族、语系、地理等相关信息,微观层面可以通过DNA数据有效刻画确定个体的容貌特征、特殊性状等[18],Hwa等[19]基于机器学习聚类的方法使用一组SNP(常染色体,性染色体,线粒体)对18个已经降解的DNA样本进行祖先推断,准确率高达94.4%[19]。在人工智能的助力下,侦查部门能够通过轨迹溯源实现案前锁定,通过人脸识别、步态识别等实现案中锁定,通过DNA、指纹等实现案后锁定,进而形成

了从案前到案后的多维锁定侦查模式。

1.2　情指一体:基于人工智能的数据检索与情报决策

侦查人员通过对已有的各种专业信息数据库、社会行业的数据库或者互联网开发资源等进行数据检索,以获取各种关联数据和信息,支撑侦查决策[20]。伴随“金盾工程”推进和公安大情报系统的建设,目前中国公安机关内部运行的信息系统已超过7000个,拥有数百亿条基础数据[21]。基于自然语言处理与深度学习的智能数据检索服务改造了原有基于关键词的检索服务[22],能够为侦查人员展现最具价值和相关性的信息,与传统数据库检索仅呈现一大堆检索结果不同。

电信诈骗、集资诈骗、网络传销、内幕交易等网络型、涉众型犯罪具有明显的网络化、数字化、集群化、智能化特征,加强情报挖掘是打击此类犯罪的关键抓手,也是公安机关推进“大情报、小行动”警务变革以及“信息化建设、数据化实战”发展战略的客观需要[23]。大数据具有“4V”特征,侦查原有的情报挖掘与决策模式面临情报需求识别率偏低、跨种类数据整合多、全周期沉浸情报决策交互复杂等一系列问题,存在一定实用性挑战。利用人工智能算法创建集数据情报汇聚、清洗、集成、归约、关联聚类等智能数据情报挖掘技术,形成“数据情报+智能算法+侦查智慧”的情报决策应用模式,降低侦查人员出现“情报失误”“认知偏差”“决策失败”等问题的概率,快速准确掌握涉案犯罪情势与发展态势,提高情报决策的收敛性与聚集性,提升利用数据情报进行侦查预测的精准性,及时评估与修正基于数据情报的侦查实施方案,实现具备即时、动态、空间矩阵、预警层次、分析实施等数据情报智能“云”应用模式,突显侦查数据情报智能挖掘与决策的应然价值和实然效果[24]。侦查机关依托智能警务系统或平台,能有效汇聚各类数据、信息,根据精准实时的情报,显著提升扁平化指挥水平,推动情指一体化建设,实现侦查行为从被动应对向主动处置转型。

1.3　全链支撑:基于人工智能的辅助审讯与证据评估

审讯中大量重复性工作可以由智能笔录系统来完成,能让侦查人员更加专注于审讯策略的应用和审讯效果的评估。语音和语义识别技术能自动记录审讯过程并生成相应的笔录。侦查人员能在问话结束后立即与被讯问对象核实笔录,减少歧义,提升真实性和准确性。基于人工智能的审讯感知系统已完成初步构建,其能利用智能设备采集目标对象生理与行为的信号,建立涵盖面部温度和表情、肢体动作、语音情绪等生理信息及其情感指向相关特征的多模态数据库,实现审测一体化应用[25]。在讯问场景下对犯罪嫌疑人进行情感计算、心理刻画、认知分析,实时向侦查人员推送审讯策略与技巧,进而辅助审讯人员实现人机协作的智慧审讯。

证据是刑事诉讼的灵魂,是实现司法公正的基石,侦查是查明案件事实真相最主要的途径,但是侦查构建的案件事实最终需要经受法庭审判的检验,其中最重要的标准就是其是否达到“事实清楚、证据确实充分”的证明标准。虽然证据的评估更多依靠人的主观评判,但并不意味着人工智能在此无用武之地。如基于人工智能研发的“206”系统能提供统一的证据标准和规则指引,单一证据核查校验[26],能自动提取各个证据的核心元素,遵循特定的证据规则和逻辑规则进行审查、推理、比对,以求确认现有证据是否已构建完整的证据体系[27]。在人工智能的助推下,侦查机关可以将庭审所要求的证据规格和标准实质性地落实在侦查阶段所有证据收集过程中,为侦查证据指引、程序审查、质量审核等提供全链支撑。

2　刑事侦查中人工智能应用的风险挑战

2.1　黑箱效应:人工智能算法自有局限对侦查的“投射”影响

算法是实现人工智能的核心,通过将问题情境和问题要点分别转变或者抽象为限定条件与计算变量,利用数学模型拟合整个问题,进而通过数学计算求解问题答案,是一种机械式运算过程的体现[28]。从技术逻辑角度看,人工智能算法模型的运行实质上是一种相关性预测,通过辨别性或者生成

性方法,利用分析前例来估计和推断当前或者未来相关变量的取值范围[29]。人工智能算法这一本质决定其具有“黑箱”效应、虚假相关、歧视偏见等,而且这些特性都会“投射”影响侦查实践。

人工智能算法类似于一个“黑箱”的自主决策系统,人类无法有效解释算法是如何产生自主决策结果。有学者认为最精确的模型具有内在的不可解释性和复杂性,即不可解释性是实现决策精确度的必要条件[30]。但是,算法“黑箱”全面应用于刑事侦查领域,使人们难以对其有效的评估、质疑、监督。如利用人工智能犯罪风险评估工具对潜在目标进行评估时,决定侦查决策结果的不再是刑事侦查和刑事诉讼规则,而是算法代码。当算法编程工程师将已有既定规则转换成代码时,不可避免地要基于信息技术和算法实现对相应规则进行调整或修改,然而侦查人员和公众对此并不知晓,无法对嵌入到人工智能侦查决策系统中的规则或算法进行可责性、准确性、透明性等审查[22],这使侦查活动面临着巨大的信任危机,无法满足“看得见的公正才是真正的公正”这一现代司法所追求的公开透明原则,同时也给现代法治所提倡的“释法析理”精神带来挑战[31]。

与此同时,算法训练构建中变量的随机性可能引致模型中变量的虚假相关。人工智能算法实质上基于大数据“喂养”训练,进而“发现”不同变量之间的相关性,但其这种相关性的真假有时不易确定,尤其在关键因素或变量缺失抑或对样本总体变化显失关注的情况下更难判定[32]。算法的变量之间可能存在伪相关,即因一个中介变量存在使两个本无相关性的变量具有统计意义的相关关系[33]。前端输入数据之间关系并不明晰,构建的人工智能算法模型也可能归入部分伪相关,进而引发相关性滥用的风险,通常数据越多涌现重复伪相关的可能性越大,出错的概率也就越高[34]。此外,算法“黑箱”效应相伴的算法歧视偏见也对侦查带来风险和挑战,人工智能侦查模型开发、运行和应用阶段都可能产生偏见。开发阶段存在代码缺陷、结构缺陷、规则缺失、数据不完整或数据选择偏差问题,运行阶段存在数据使用错误、算法使用错误、任务配置偏差以及系统安全漏洞等问题,应用阶段存在错误解读、错误应用、错误决策的风险[30]。

2.2 精准偏差:人工智能侦查特征识别的准确度存有罅漏

基于大数据学习建构的人工智能算法模型,能够涵盖影响结果的各种相关变量,相应参数设置基本上也符合样本数据的规律,预测结果亦有相对较高的准确性[32]。但是人工智能算法这种统计学意义上的“精准性”和“科学性”延及到微观的侦查领域可能就谈不上精准和科学了。

以人脸识别为代表的生物特征识别与认知计算存有“决策误区”,目前最先进的人脸识别技术在大规模测试数据集中其人脸辨识的准确度已经达到99.9%,但是其在侦查实际应用中依然面临许多挑战。在大规模人脸数据库布控应用场景中,虚警率与漏报率的尖锐冲突依然存在:可允许的虚警率通常极低,但此时系统的漏报率却可能非常高。相反如将漏报率设置为非常低(如<0.1%),虚警率可能会高到应用无法接受的程度。2018年美国国家标准与技术研究院发布了全世界人脸识别算法测试(face recognition vendor test)结果,在条件高度可控的签证照片应用场景中排名第一的算法错误接收率千万分之一时错误拒绝率仅为0.4%,在日常照片应用场景错误拒绝率达到5.2%,在不可受控的侦查场景下错误接收率万分之一时错误拒绝率达到40%~50%[35]。

人工智能侦查中运用生物特征识别技术,通常以概率或相似分数来判断两个比对目标是否同一,而相应的概率和相似分数是统计数据而生成,但是统计数据的可能性仅仅是对某一类事物或某一群体的通常描述。基于人工智能的生物特征识别技术在利用大数据样本训练构建算法模型时结合侦查领域的特殊性需求和刑事鉴定专家的经验性知识,让其在同一认定中完成从相关关系向“形式”因果关系转化,然而这一转化过程中并未考虑特征价值权重概念,同时加之相关关系较弱导致转化为“形式”因果关系时充分性和说理性不强,进而造成进行同一认定时存在准确度的罅漏。如人脸识别技术对于同一性质或层级的特征所赋权重存在差

异,造成输出结果不一,这也与刑事证明所强调的同一性标准相矛盾,相同输入获取明显差异的输出。这个问题的本质在于有关生物特征分类、特征价值的评估原则与方法等同一认定理论的知识和原理未能有效嵌入人工智能侦查模型中[36]。

2.3　算法僭越:人工智能侦查的技术权力冲击侦查权力

伴随着人工智能系统不断自我学习,对特定领域的知识掌握愈加全面,其分析决策能力将胜过人类,进而促使人类对机器决策愈加依赖。当侦查领域中人工智能系统被授予决策权后,在除去安全风险问题之外,首先要考虑的一个伦理问题就是人工智能系统是否具备决策资格,技术权力对侦查权力的冲击是人工智能侦查下一步发展无法回避的挑战[37]。

权力的实质是一种存在于一定社会关系网格中的支配力量。技术是人们开展理性认识的手段和工具,技术本身与权力无关,在具体应用过程中才与权力产生关系。人工智能广泛应用于侦查,对侦查产生直接影响甚至形成支配之势,此时很难将技术看作是纯粹的认识工具,其伴随有一定的权力属性[38]。与传统技术明显不同的是人工智能从诞生之初就展现出显著的权力属性,其超越狭义层面的技术工具论概念,呈现为一种对应用场域的支配权[39]。人工智能侦查场域中侦查同样受技术权力的影响甚至支配。人工智能模型利用其技术优势,根据大数据运算配置侦查资源,以不易察觉的方式转变为侦查规则、规范影响着侦查人员的认知与行动,成为刑事侦查中一种新兴的技术权力。人工智能这种技术权力具有内容上的专业性和形式上的隐匿性,以所谓"客观、理性、高效"的技术权威干涉甚至支配着侦查决策和行动,加之其具有分散化与无边界化的特性,使这种技术权力的干涉和支配无处不在,无远不届[40],导致侦查人员不断地将自己的权力单向让渡于技术,随着人工智能不断优化发展及其在侦查领域多维度全方位的渗透,人工智能的技术权力范围还在持续扩张[41]。

技术权力的介入让侦查主体基于专业知识和个人经验构建而成的侦查权力逐渐流失,同时造成侦查主体高度依赖人工智能手段,忽视甚至漠视原有的侦查手段和措施,传统的工作与沟通技巧不断弱化,不能良好地深入到群众中开展侦查工作以获取特定高价值的信息和情报,无法构建基于具身认知的侦查隐性知识体系,侦查人员的侦查认知和决策能力不断弱化,进而促使其越加依赖人工智能技术,形成技术权力不断扩展而侦查主体能力不断弱化的恶性循环[3]。在人工智能为侦查人员"技术赋能"的同时也伴随着"技术索权"[42],当人工智能在侦查活动占据绝对主导地位时,容易造成侦查过程的机械化、程式化,侦查治理效果冰冷化,弱化侦查人员的感知与创造能力,破坏社会的"文化"层面架构,导致侦查人员与社会交互关系愈加"冰冷""僵硬""功利",无法发挥人文关怀和本原价值等法理精神[43]。

2.4　全景敞视:人工智能侦查对个人权利的弥散性侵害

传统侦查权力运行机制是"命令-封锁"的规训方式,数字化背景下侦查机关深度融合人工智能、大数据与传统侦查手段,形成了一种"全景敞视"的规训机制,其通过体系化的精细策划,使权力以一种看似可见但实质又无法可见的方式间接完成对权力对象的规训。侦查部门基于大数据智能合成平台可以有效共享公安机关内部各警种所汇集的公民个人信息,并且经过技术加工后锁定犯罪嫌疑人、固定与提取证据,从而降低了传统侦查中对嫌疑人身体的过度依赖,同时相当程度上改变了原有基于现场"面对面"调查取证的权力运行方式。人工智能侦查以一种虚拟化与间接化的方式让权力关系弥散分布于社会之中,确保权力效应可以抵达最细微、最偏远的角落[44]。

借助科技外衣,人工智能侦查更具有潜匿性和欺骗性,其通过技术并作为技术实现自我巩固与扩张,而且为自己进一步拓展统治权力供应足量的合法性[45]。例如现阶段的刑事案件办理中,人们对公安机关持续广泛且有深度的个人信息采集行为赋予足够理解,为了抓捕重特大犯罪嫌疑人或逃犯,侦查机关在未经严格审批的情况下,采取视频追踪、人脸识别、网络侵入、信息深度挖掘等各种可能损害到普通民众个人隐私的技术手段时,民众出于

对社会安宁稳定的美好祈望而有大可能对其给予宽容和谅解[46]。大众潜意识认为多数人利益与少部分人权益是一种厚此薄彼的“零和关系”,而不是一种同向发展的“正和关系”,进而同意侦查机关为了维持权力的有效运转能够忽视甚至可以牺牲少部分人的合法权益[47]。

人工智能全面介入侦查领域促使侦查权的辐射区域表现出明显的泛化趋势[48]。“权力-权利”关系处于显著失衡状态,警察可以不间断、全方位、隐匿式地对个体进行监测控制,然而普通民众对相应行为知之甚少抑或毫不知情,公民的社会生活彻底处于被“监视”状态,毫无私密可言[47],人工之智能侦查实现了福柯所担心的“少数人甚至一个人能够在瞬间看到一大群人”的全景敞视主义[44]。可以预见随着技术的优化,人工智能侦查这种全景敞视“规训”还有很大可能不断扩张,致使公民个人地位持续式微,形成个人权利被弥散性侵害甚至消解的溢出效应。

2.5 侦查密室:人工智能侦查应用法律规制的阙如与责任主体的模糊化

人工智能技术的开放性和不明确性使得人工智能侦查很难被清晰地归入现有法律框架下特定种类的侦查措施,现有侦查措施体系只能解释部分人工智能侦查行为,目前法律文本明显滞后于人工智能时代下侦查措施的实际发展[49]。人工智能侦查和《刑事诉讼法》规定的搜查、勘验检查、技术侦查等强制性侦查措施均有所联系同时又有区别,人工智能侦查中的数据收集和搜查在实施对象上存在重叠;人工智能侦查对网络数据的采集和固定与勘验检查的线上模式具有明显的共性;技术侦查具有对象特定性、技术性、秘密性和内容高度隐私性的特点,人脸识别通过布控能像技术侦查一样实现对特定目标的跟踪,在数字化条件下人工智能侦查通过社交软件和通信工具比传统技术侦查更容易实现监控,在一定程度来说人工智能侦查与技术侦查这一强制性侦查措施具有最高的相关性[50]。刑诉法针对强制性侦查措施均有相应的法律规定和要求,如勘验检查对象要求与犯罪相关,需要出具证明文件和见证人在场;对于技术侦查措施的使用要求更为严格,适用的案件范围有严格限定,事先需要经过严格的审批并有明确的适用期限要求。由于人工智能侦查法律属性的模糊性以及文本法律自身的滞后性,其应用几乎无需司法机关审核,实质有效的程序性限制也极其缺乏[51]。人工智能侦查措施无法被现行法律的授权框架所覆盖,更遑论对其进行精密化层级式的规制[52]。此外,侦查权“自我管理”的行政化运行方式亦有强化滥权风险之嫌,技术的“黑箱效应”持续增强“侦查密室”效应,让人工智能侦查成为一种“超级侦查权”[46]。

在正式、统一以及成熟的理念和制度没有完全形成时,定位于辅助地位的人工智能侦查沿袭传统科层制组织结构的宰制,在一定程度上进一步强化了侦查组织对侦查人员的统辖,进而卸载了侦查人员的主体性[53]。当人工智能侦查因算法偏见或歧视等做出错误的侦查决策,由于技术壁垒和算法“黑箱”等因素的影响,侦查人员通常很难识别其中的错误,进而最终导致冤假错案时,此时责任主体是谁?由于人工智能技术嵌入到侦查决策形成过程中,起着“第三方”隔离作用,传统的“人-人”关系形式被“人-技术-人”的关系形式所替换[54],致使各关系主体、行为原因与后果的直接联系等均被人工智能技术所隔离,造成原有的法律责任认定和归责机制难以奏效[55]。

3 刑事侦查中人工智能应用的发展策略

面对人工智能全面浸透侦查,我们既无法回避和无视它,也无法否定与阻止它[56]。更加合理的进路是从实践出发,积极探索运用技术、法律、伦理等多维度的发展策略化解吸纳风险。

3.1 基本准则:锚定以人为本与人机共融的立场

“人是目的而非手段。”[57]在充分吸纳人工智能技术为侦查所带来便利性的同时,需要反思应用中的风险性和客观化,权衡价值理性与工具理性之间的关系,明确“以人为本”的基本立场,为人工智能侦查的“技术权力”划定边界,以此减少和规避风险,确保其有序运行[31]。“人”自始至终是侦查最核

心的要素。首先基于哲学视角从技术和人两者的终极关系看,人工智能本质上是属人的。尽管人工智能可以做许多人类无法完成的事情,但是人类创造与发明技术的根本目的是为人类的生存与发展服务的,技术无法独立于人类而存在[58],“技术作为人类的一种存在方式,是与人类相伴而成的”[59]。侦查中人工智能的本质是为了辅助侦查人员发现和证明犯罪而存在,是侦查的工具和手段。其次从侦查实践看,人工智能即使善于数学计算、逻辑思维甚至可以具备人工情感与情感计算等,但无法置身于错综复杂的犯罪事实中去抽丝剥茧和运筹帷幄,更无法像人类一样去综合思考和判断侦查活动中涉及的司法价值、法律意义、刑事司法政策的运用效果等[25]。再者基于对技术“霸权”的警惕,在侦查技术高度智能化的背景下,技术官僚化愈加明显,侦查人员的主体性作用不断被弱化,只有锚定“以人为本”立场,才能有效抑制 “一切皆可计算化”所造成的技术理性操纵与人文价值流失的冲击[3]。

算法是以数据之间相关性为基础,运用概率运算进行预测和分析。对于回溯性刑事侦查来说,人工智能可以帮助侦查人员在更大盖然性的方向上进行调查取证,但是基于纯粹相关性的运算和推演难以完成对确定性、唯一性犯罪事实的认知、查证与构建[60]。现有的人工智能侦查系统设计和实现主要是从技术角度出发,常常忽视了无法全部模型化或者定量化的非技术因素(人)的介入影响,更多的是将人视为系统的“看客”即纯粹的使用者[61]。人工智能技术需要与侦查人员的智慧进行融合,积极构建人机共融体,充分运用侦查人员的隐性知识,在相关性数据的基础上查找因果关系,完成对犯罪事实的证据收集与固定,实现唯一性证明,达到刑事诉讼证明标准[3]。人机共融体的核心仍然是人,人工智能以其强大的能力助力侦查人员但并不能取代侦查人员,在不断扩大人工智能应用的同时积极注意其风险控制[62]。

3.2　技术路径:推动数据驱动和知识指引的融合

3.2.1　建构数据和知识双重驱动的人工智能侦查模型

目前人工智能的学习模式主要是采用数据驱动形式,以深度神经网络为代表的人工智能算法取得了显著进展,然而数据驱动形式的人工智能在数据获取、可解释性、鲁棒性等方面均存在不足[63]。在追求因果关系溯源的侦查领域中,这种不足更为突出。强化数据驱动与知识引导是未来人工智能2.0非常重要的特征[64]。现有的人工智能侦查模型是一种单纯数据驱动模型,其参数设置与计算仅取决于训练数据,缺乏逻辑知识和科学定律的支撑,领域核心知识,引入和驱动不足,致使模型的可解释性和契合性存在短板。为此需要在现有算法中引入先验假设、逻辑规则、科学定理等知识,构建数据与知识双重驱动的人工智能侦查模型。在数据驱动模型的基础上采用知识迁移等方法[65]融入常见的侦查相关知识,如逻辑知识、科学定律知识、特定行业知识、专家经验知识等。

以往的逻辑知识一般以专家库的形式表现,其所能模拟呈现的知识与符号仅是人类知识的极小部分,在应对现实世界刑事侦查所遇到的复杂问题时,若仅应用单一谓词形式则无法有效解决侦查场景任务。现阶段逻辑知识正向知识图谱方式转型,知识图谱能够表达实体的描述性知识以及实体之间关系的语义网络,基于知识图谱的自动问答与推荐系统等技术目前已在不同领域有所应用;例如针对电信诈骗案,侦查部门利用通话数据构建知识图谱进而挖掘此类案件中的线索[66-67],下一步在人工智能侦查模型研发过程中需要将知识图谱的语义信息嵌入深度学习,将关系知识引入到算法模型并运用知识约束来实现人工智能侦查的协同优化[68]。

3.2.2　提高算法的可解释性和透明度

人工智能侦查算法的可解释性是保障个人权利,维护社会信任,推动技术创新和深度应用的共同需求,是提高模型变量参数合理化、促进智能侦查决策过程公正化、实现侦查推理科学化的重要前提[69]。目前对提高“算法可解释性”“算法透明度”技术路线主要可以归结为“解锁I-T-O三阶段”与“逆向工程学”。这两种技术路线都是以“看进去”的思路应对算法黑箱,但是囿于商业机密要求、专业技术门槛、公民算法素养等制约,这种策略落地存在困难,同时容易导致对算法“黑箱效应”复杂机

理认知的幼稚化、简单化。以“可接受性”“可理解性”为中心的算法解释策略更加适合于以因果关系为目标的人工智能侦查模型,更具有可操作性,其坚持用户导向,把社会接受效果与民众相对理解当作衡量准则,不执着于从数据库到计算模型、从输入到输出的绝对公开,强调从“数据主体”出发对算法透明度和可解释性加以“厚描述”[70]。

要实现对一套复杂系统的理解,对其逐一拆解、理解其中的每个部分不一定最佳选择。理解应是基于“互动”基础,借助语音交互、接触式互动甚至游戏互动等不同交互形式实现“深度描述”,让信息接受者或系统使用者对算法运行及其对信息流的影响形成实质理解[71]。算法研发机构和侦查机关应定期向民众公布智能模型的技术逻辑和特征,运用可视化方式向公众展示人工智能侦查中数据采集、算法决策、结果输出的基本流程,让民众能够了解、评估并解释算法模型的中间状态、功能逻辑、潜在风险等[72]。

3.2.3 强化概率方法在侦查推理的应用

人类所处的世界充满不确定性,这种不确定性是“世界本身的随机性与非决定性导致的结果”[73]。侦查过程同样充满不确定性与博弈性,在不确定性中寻找相对确定性是提高算法可解释性的要求之一,也是人工智能侦查下一步发展的重要方向。

“理性”追求适当地使用系统性理由,谋求以最佳的可能方式实现有效决策,包含认知、实践与评价等3重理性目标,且需要具备想象力、信息处理、评估、选择、执行等5项能力[74]。概率推理正是具有这5项能力的理性决策过程,数学概率是一种适度理性的典范,其可以基于不完整、不全面的知识或信息解决日常生活中的难题[75]。为此,概率推理作为应对不确定性问题并能有效实现精准决策的方法,其具备结构化的逻辑推理和决策框架、数字化的信念表达和事实推论以及科学化的信息处理与信念结合机制等特征,与人工智能技术有天然契合性,是实现人工智能侦查溯因推理,提高可解释性的重要方法。例如运用基于结构化评估似然比的概率推理,揭示侦查中相关证据证明力在事实推理进程上的传递与合取机制,然后根据有关概率公理创建能够在单链多级推理链条和收敛结构中明确证据证明力与实现可解释决策的算法模型[76]。

当然,应用概率方法建构人工智能侦查模型也面临着挑战,主要表现为概率数值的确定与证据聚合难题,贝叶斯网络与富兰克林原则为解决相应难题提供了进路。贝叶斯网络可以有效融合主、客观概率,提供统一框架以消除概率数值的评估分歧;而富兰克林原则可以保障合适参考类的理性选择,解决客观概率赋值的数值难题。为此,基于贝叶斯公式的一致性与规范性以及借助严格的形式化概率推论能够解决证据聚合难题,说明概率方法是不确定性条件实现计算推理的可靠方法,这为面向侦查实践构建具有可解释性的人工智能推理模型提供了理论支持[77]。

3.3 制度保障:强化技术伦理与法律规制的配套

为应对人工智能侦查所带来的风险调整,在制度方面需要立足中国实际,充分借鉴域外经验,探索建构以伦理为前导的社会规范调控机制和以法律为主导的风险规制体系[56],为人工智能侦查的发展和应用提供法治化的制度保障,让其兼顾打击防范犯罪与保障人权权利,平衡实体真实和程序正义[78]。

3.3.1 增强刑事侦查中人工智能应用的伦理适配

合适的伦理规约不仅不会阻碍科技的发展,反而起到积极的推动作用。相较于法律的“刚性”监管,对侦查中人工智能的应用与发展进行伦理角度的“软治理”,具有显著的理论和现实意义[79]。

1)充分借鉴国内外算法伦理研究成果,引入侦查“数字人权”理念。“正义是社会制度的首要价值”[80]。基于过去科技背景构建的伦理制度不一定适合人工智能时代的伦理需求,更谈不上是完全正义的[81],刑事侦查是关乎社会正义的重要支柱,这对侦查中人工智能应用的伦理构建提出了更高的要求。中国先后出台了《关于加强科技伦理治理的意见》《人工智能应用准则》等关乎算法的伦理准则,美国出台了《国家人工智能研发战略计划:2019更新版》《关于算法透明度和问责制的声明》等规范,欧盟也发布了一系列规范如《可信赖人工智能伦理指南》《人工智能道德守则》,人工智能侦查的

研发和应用应充分借鉴以上相关伦理规范,结合侦查应用场景的特点,形成人工智能侦查伦理框架,厘清与规范其应用范围和边界[82]。

目前中国刑事侦查程序中有关人权保障的研究和实践仍主要局限于人身权、财产权等传统人权范畴。传统的人权保障思路与数字化、智能化时代下强调保护“数字人权”的理念严重不符。“数字人权”被认为是“第四代人权”,其以数据与信息为载体,是智慧社会中人的数字化生存样态与发展需求的基本权利,涵盖数据信息的自主权、知情权、表达权、公平利用权、隐私权等[83]。“数字人权-数字治理”是“人权-治理”框架在数字空间中的扩展[84],刑事侦查需要从“数字人权”出发,就如何更加合理重新划分任意性侦查和强制性侦查开展学术研究和理论探讨。对可能侵害“数字人权”的初查措施与侦查措施如何开展有效的分类和规制,进行细致、谨慎的探索[85]。

2)在人工智能侦查算法开发中嵌入伦理设计。相较于民用领域人工智能的应用,侦查领域人工智能的应用更具伦理风险性,在算法模型开发伊始就需要融入伦理规范[86]。价值敏感性设计理论要求将人类价值尤其是伦理道德价值和行为方式等内嵌到设计里,以求消解技术设计和伦理价值关切间的差异[87],其强调人和算法的互动关系,要求将算法的应用情景与设计有机融合[88]。基于价值敏感设计理论,人工智能侦查算法设计首先需要对其所涉及的权力、隐私、权利、透明、公平、安全、责任等相关伦理道德价值概念进行识别、判断与界定;其次充分运用定性和定量方法对算法设计开展技术和价值维度的评估,为相关伦理道德价值提供经验数据的支撑,同时在经验调研基础上开发者需要重点思考如何处理自身持有的道德选择、价值判断等,以及权衡侦查应用中可能遇见的各类不同价值之间的冲突或矛盾;最后根据已明确的相应伦理价值开展具体的算法设计、性能与风险分析[89]。

3)创建多方主体联合参与的伦理治理模式。应对人工智能侦查所带来的风险挑战关键在于有效调节人类社会和警务信息技术两者间的互动关系,让信息技术和整个现代社会处于共存、共融、共生的态势[90]。政府、公安机关、科技单位、民众等主体均以直接或间接的方式参与人工智能侦查的建设和应用,其技术风险规避亦需要前述多元主体的共同参与治理才能实现良好的效果,为此需要创建多元主体共同治理模式[91]。政府部门要加强政策制定和宏观引导;公安机关要坚守权力-权利制衡底线,规范人工智能技术的应用,减少对公民个人权利的侵害;科技企业要提升技术人员的伦理价值意识,在算法设计中自觉遵守伦理规范,同时企业要转变“唯技术效益”论的人才评价体制,将产品的伦理价值和影响归入考核体系并加大权重,引导技术人员主动认可科技伦理[72];公民作为人工智能侦查体系中的相对弱势的一方,需要进一步树立数据意识[92],提高参与监督的积极性,运用法律武器监督人工智能侦查的运行,坚决保护个人信息权和隐私权[91]。

3.3.2　优化刑事侦查中人工智能应用的法律体系

1)构建算法引入的事前审查和适度公开制度。侦查中人工智能算法的应用并非是一种纯粹的技术理性,其内在蕴含着侦查部门对社会分化的价值判断,也是对社会不同阶层利益诉求的一种平衡和协调[93]。为此,一要评估其总体目标与应用场景,不同目标和场景决定运行模式和伴随风险性的不同,需要评估引入算法是否符合本地治安形势和侦查需求,是否存在功能过于超前进而存在滥权的风险;二要审查技术参数的效能等级,通过审查算法涉及的数据种类、数据量、分析对象、挖掘的深度与广度等技术参数,评估其技术权力效应和侵权风险等级,由第三方审查机构根据技术清单进行专门审查,出具审查和认证报告[94];三要设置算法公开清单,在平衡算法知识产权保护与民众知情权的需求情况下推动算法适度公开,不强制要求算法全部源代码的公开,但是人工智能侦查开发企业需要公开算法规则、运行逻辑、裁判机制、应用效果、潜在风险、出错概率等公开清单上所要求的信息[95],针对不特定的数据所有人、犯罪嫌疑人、专门机构和司法机关等不同对象公开的详实程度可以有所区别[96]。

2)推进人工智能侦查措施类型化及其阶层式

规制。一是采用实质标准和形式标准对人工智能侦查措施类型化,以是否侵害个人的重要权利或基本权利作为判断任意和强制的实质标准,以强制力、同意、秘密性等作为判断的形式标准[97],将人工智能侦查划分为4种类型,显失干预性措施、轻微干预性措施、紧急态势下干预措施、常规情势下的强制干预措施;二要针对不同类型行为采取阶层式规制,对于常规情势下的强制干预措施采用纸质令状的前置方式;对于紧急态势下人工智能侦查干预措施可以电子令状授权方式,以满足动态性和程序性要求;对于轻微干预性措施,基于现场情势需要在未得前置令状许可的条件下可以先行实施此类侦查措施,但要求在事后必须取得法定机关补发的纸质令状修复其实施过程的形式瑕疵;对于显失干预性人工智能侦查措施,无需令状审查,只需经过事前报备程序便可实施[52];三要设置严格的备案制度,确保人工智能侦查系统每次使用都有记录和备案,严禁不是出于案件办理目的需要的非法越轨使用[98],侦查机关内部需要定期审查使用情况,倒查追责越权使用。

3)设立多元监督与权利救济的机制。一要创建人工智能侦查的检察监督机制,提高检察机关对人工智能侦查行为开展全程监督的能力,特别是要对具有明显强制性的人工智能侦查措施强化监督,检察部门可以适时适情配备具备专业的数字侦查检察官,以便快速纠正人工智能侦查中的越轨违法行为,及时处理公民的申诉和控告,在调查核实后对受侵害公民给予权利救济,在不压制侦查机关人工智能应用的客观需求下,防止技术滥用而过度侵害公民权利[97];二要探索建立第三方监督机制,由国家信息管理和司法行政部门牵头,联合社会力量组建第三方监督和评估主体,定期对侦查部门所用的人工智能模型进行监督和评审并出具报告[53];三要完善权利救济机制,给予侦查行为相对人充分的陈述和申辩权利,允许当事人聘请有专门知识的人作为专家辅助人或司法鉴定人对人工智能侦查模型给予充分的质询,同时将"比例原则"法律理念嵌入人工智能侦查过程中,综合权衡人工智能侦查行为违法的严重程度、犯罪行为的危害程度、有无其他合法侦查措施的替代选择、对个人权利侵害的程度、相关证据对司法公正的影响程度等因素[97],若人工智能侦查措施明显不符合"必要性""妥当性"和"相称性"[98],则其获取的相应证据要作为非法证据予以排除。

4 结论

侦查学是一门应用性、综合性、开放性的学科,犯罪的方式和方法随着社会发展和技术进步千变万化,侦查技术与方法也会随时快速调整[99]。侦查的这种特点决定着其对社会和技术的变化有着高度的敏感性,人工智能在侦查领域的广泛应用是犯罪态势日益智能化、非实体化的必然要求,也是侦查部门顺应时代潮流开展自我变革的现实路径。但就目前人工智能侦查实践样态看,仍处于初级阶段,远未达到充分且成熟的状态,这一过程必然也伴随着技术应用所产生的风险与挑战。为了有效应对人工智能的"双刃剑"效应,科技之矛和法律之盾理应协同推进,充分发挥人工智能技术基础性应用的支撑作用与法律制度导向性的把控规制作用,在有效发挥人工智能技术效能的同时,防范技术诱因驱动下侦查过度扩张与越轨使用造成的侵权行为。在侦查、司法、立法、科技等多方力量的共同努力下,推动科技和理性在人工智能侦查领域相映生辉。

参考文献(References)

[1] 孙笛. 人工智能时代的犯罪风险分析[J]. 中国人民公安大学学报(社会科学版), 2018, 34(4): 11-16.
[2] 张旭, 朱笑延. 弱智慧社会语境下的犯罪治理: 情势、困境与出路[J]. 吉林大学社会科学学报, 2019, 59(1): 19-29, 219.
[3] 周学农, 潘庆娜. 智慧侦查:发展逻辑、突出问题及实现路径[J]. 湖南警察学院学报, 2022, 34(1): 5-15.
[4] 刘为军. 论侦查、预防、服务三元一体的侦查理念[J]. 中国人民公安大学学报(社会科学版), 2020, 36(2): 29-39.
[5] 商瀑. 人工智能与刑事侦查:历史变迁、技术分类及未来展望[J]. 中国人民公安大学学报(社会科学版), 2020, 36(6): 46-54.

[6] 王妍, 詹雨薇, 罗昕, 等. 视频片段检索研究综述[J]. 软件学报, 2023, 34(2): 985-1006.
[7] 庄华. 视频侦查研究进展与未来展望——基于598篇视频侦查相关论文的知识图谱分析[J]. 政法学刊, 2020, 37(6): 19-29.
[8] 吴春生, 李孝君, 吴浩. 基于深度学习的指纹自动识别技术[J]. 刑事技术, 2022, 47(1): 88-95.
[9] 徐杰, 刘哲元, 霍鑫, 等. 人工智能指纹识别技术在警务实战中的应用[J]. 刑事技术, 2021, 46(3): 299-304.
[10] 余璀璨, 李慧斌. 基于深度学习的人脸识别方法综述[J]. 工程数学学报, 2021, 38(4): 451-469.
[11] 张涛. 人脸识别技术在政府治理中的应用风险及其法律控制[J]. 河南社会科学, 2021, 29(10): 44-55.
[12] 张雪. 人脸识别技术在侦查应用中的局限与应对[J]. 北京警察学院学报, 2021(2): 84-88.
[13] Huang G B, Mattar M, Berg T, et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments[R]. Amherst: University of Massachusetts, Amherst, 2007.
[14] 刘琦, 于汉超, 蔡剑成, 等. 大数据生物特征识别技术研究进展[J]. 科技导报, 2021, 39(19): 74-82.
[15] Jonathon P P, Yates A N, Ying H, et al. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms[J]. Proceedings of the National Academy of Sciences, 2018, 115: 6171-6176.
[16] Marciano M A, Williamson V R, Adelman J D. A hybrid approach to increase the informedness of CE-based data using locus-specific thresholding and machine learning [J]. Forensic Science International Genetics, 2018, 35: 26-37.
[17] 刘冰. DNA数据库数据挖掘应用研究[J]. 刑事技术, 2015, 40(5): 345-352.
[18] 刘志勇, 张更谦, 严江伟. 人工智能在法医学中的应用与展望[J]. 刑事技术, 2019, 44(5): 383-387.
[19] Hwa H L, Wu M Y, Lin C P, et al. A single nucleotide polymorphism panel for individual identification and ancestry assignment in Caucasians and four East and Southeast Asian populations using a machine learning classifier[J]. Forensic Science Medicine and Pathology, 2019, 15(1): 67-74.
[20] 任惠华, 金浩波. 人工智能侦查的实践应用与制度构建[J]. 河北法学, 2018, 36(6): 46-54.
[21] 陈刚. 信息化侦查大趋势[M]. 北京: 中国人民公安大学出版社, 2010: 1.
[22] 腾讯研究院, 中国信息通信研究院互联网法律研究中心, 腾讯AI Lab, 等. 人工智能:国家人工智能战略行动抓手[M]. 北京: 中国人民大学出版社, 2017: 271, 243.
[23] 孙静晶. 情报导侦在风险型经济犯罪中的应用[J]. 中国刑警学院学报, 2020(3): 54-61.
[24] 薛亚龙, 罗珂岩, 马麒. 数据情报侦查的循证决策方法[J]. 中国刑警学院学报, 2022(3): 24-34.
[25] 毕惜茜. 审讯中人工智能的应用与思考[J]. 中国人民公安大学学报(社会科学版), 2020, 36(3): 30-36.
[26] 崔亚东. 司法科技梦: 上海刑事案件智能辅助办案系统的实践与思考[J]. 人民法治, 2018(18): 92-97.
[27] 宋家宁. 人工智能辅助侦查的思考——基于价值呈现与适配要求的双重视角[J]. 中国刑警学院学报, 2018(5): 31-36.
[28] 栗峥. 人工智能与事实认定[J]. 法学研究, 2020, 42(1): 117-133.
[29] 郭锐. 人工智能的伦理和治理[M]. 北京: 法律出版社, 2020: 156.
[30] 唐要家, 唐春晖. 基于风险的人工智能监管治理[J]. 社会科学辑刊, 2022(1): 114-124, 209.
[31] 翁晓斌, 吴宇琴. 人工智能司法运用的技术效应与法理审思[J]. 自然辩证法通讯, 2022, 44(8): 98-104.
[32] 聂友伦. 人工智能司法的三重矛盾[J]. 浙江工商大学学报, 2022(2): 66-75.
[33] 肯尼思·D·贝利. 现代社会研究方法[M]. 许真, 译. 上海: 上海人民出版社, 1986: 69.
[34] Calude C S, Longo G. The deluge of spurious correlations in big data[J]. Foundations of Science, 2017, 22(3): 595-612.
[35] 孙哲南, 赫然, 王亮, 等. 生物特征识别学科发展报告[J]. 中国图象图形学报, 2021, 26(6): 1254-1329.
[36] 张超, 吕凯. 人工智能时代的人身同一认定研究[J]. 合肥工业大学学报(社会科学版), 2022, 36(2): 23-34.
[37] 李修全. 人工智能应用中的安全、隐私和伦理挑战及应对思考[J]. 科技导报, 2017, 35(15): 11-12.
[38] 刘永谋. 机器与统治——马克思科学技术论的权力之维[J]. 科学技术哲学研究, 2012, 29(1): 52-56.
[39] 梅夏英, 杨晓娜. 自媒体平台网络权力的形成及规范路径——基于对网络言论自由影响的分析[J]. 河北法学, 2017, 35(1): 36-47.
[40] 王聪. “共同善”维度下的算法规制[J]. 法学, 2019(12): 66-77.
[41] 张淑玲. 破解黑箱: 智媒时代的算法权力规制与透明实现机制[J]. 中国出版, 2018(7): 49-53.
[42] 罗英. 数字技术风险程序规制的法理重述[J]. 法学评论, 2022, 40(5): 151-160.
[43] 李小波, 李文润. 人工智能在警务应用中的异化及其风险规制[J]. 中国人民公安大学学报(社会科学版), 2022, 38(4): 117-129.

[44] 米歇尔·福柯. 规训与惩罚: 监狱的诞生[M]. 刘北成, 杨远婴, 译. 北京: 生活·读书·新知三联书店, 2007: 155, 244.
[45] 赫伯特·马尔库塞. 单向度的人[M]. 刘继译, 译. 上海: 上海译文出版社, 2006: 144.
[46] 刘小庆. 从“权力监督”到“权利制约”: 大数据侦查法律规制的理性之维[J]. 重庆大学学报(社会科学版), 2022, 28(2): 220-231.
[47] 自正法, 刘小庆. 大数据侦查的本质属性及其溢出效应——基于福柯“规训”理论的分析[J]. 西南民族大学学报(人文社会科学版), 2022, 43(6): 108-113.
[48] 裴炜. 个人信息大数据与刑事正当程序的冲突及其调和[J]. 法学研究, 2018, 40(2): 42-61.
[49] 陈刚. 解释与规制:程序法定主义下的大数据侦查[J]. 法学杂志, 2020, 41(12): 1-17.
[50] 胡铭, 龚中航. 大数据侦查的基本定位与法律规制[J]. 浙江社会科学, 2019(12): 12-20, 155.
[51] 陈华, 唐颖菲. 互联网与人工智能视域下的警察权研究[J]. 北京警察学院学报, 2021(1): 1-7.
[52] 商瀑. 人工智能时代的侦查变革及其法治图景[J]. 中国人民公安大学学报(社会科学版), 2019, 35(6): 65-72.
[53] 张迪. 刑事证明中人工智能的应用: 精准定位、理念反思与路径优化[J]. 华中科技大学学报(社会科学版), 2022, 36(4): 64-73.
[54] 齐延平. 论人工智能时代法律场景的变迁[J]. 法律科学(西北政法大学学报), 2018, 36(4): 37-46.
[55] 齐延平. 数智化社会的法律调控[J]. 中国法学, 2022(1): 77-98.
[56] 吴汉东. 人工智能时代的制度安排与法律规制[J]. 法律科学(西北政法大学学报), 2017, 35(5): 128-136.
[57] 康德. 实践理性批判[M]. 韩水法, 译. 北京: 商务印书馆, 1999: 95.
[58] 陈凡, 程海东. 人工智能的马克思主义审视[J]. 思想理论教育, 2017(11): 17-22.
[59] 程海东, 刘炜. 语境: 技术的现实存在场域[J]. 东北大学学报(社会科学版), 2014, 16(6): 558-562.
[60] 倪春乐. 大数据背景下的侦查创新与现实局限[J]. 公安学研究, 2019, 2(4): 91-104, 124.
[61] 黄孝鹏, 周献中, 陆晓明, 等. 从“以人为本”角度理解和认识决策系统的演化趋势——兼谈人机协同决策系统的构建[J]. 系统科学学报, 2013, 21(1): 35-39.
[62] 王柏村, 彭晨, 易兵, 等. 智能时代的人机共融体: 技术驱动、以人为本——《The Humachine: Humankind, Machines, and the Future of Enterprise》导读[J]. 中国机械工程, 2021, 32(19): 2390-2393.
[63] 金哲, 张引, 吴飞, 等. 数据驱动与知识引导结合下人工智能算法模型[J/OL]. [2022-10-24]. http://kns.cnki.net/kcms/detail/11.4494.TN.20220901.1012.006.html.
[64] Pan Y H. Heading toward artificial intelligence 2.0[J]. Engineering, 2016, 2(4): 409-413.
[65] 张启阳, 陈希亮, 曹雷, 等. 深度强化学习中的知识迁移方法研究综述[J/OL]. [2023-02-16]. http://kns.cnki.net/kcms/detail/50.1075.TP.20230214.1409.028.html.
[66] 凡友荣, 杨涛, 孔华锋, 等. 基于知识图谱的电信欺诈通联特征挖掘方法[J]. 计算机应用与软件, 2019, 36(11): 182-187.
[67] 许振亮, 刘喜美. 电信诈骗研究的知识图谱分析[J]. 中国刑警学院学报, 2017(3): 50-56.
[68] Chen X, Zhang N Y, Xie X, et al. Know prompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction[C]//Proceedings of the ACM Web Conference 2022. New York: Association for Computing Machinery, 2022: 2778-2788.
[69] 刘云. 论可解释的人工智能之制度构建[J]. 江汉论坛, 2020(12): 113-119.
[70] 仇筠茜, 陈昌凤. 基于人工智能与算法新闻透明度的“黑箱”打开方式选择[J]. 郑州大学学报(哲学社会科学版), 2018, 51(5): 84-88, 159.
[71] Resnick M, Berg R, Eisenberg M. Beyond black boxes: Bringing transparency and aesthetics back to scientific investigation[J]. Journal of the Learning Sciences, 2000, 9(1): 7-30.
[72] 林伟. 人工智能数据安全风险及应对[J]. 情报杂志, 2022, 41(10): 105-111, 88.
[73] Weatherford R. Philosophical foundations of probability theory[M]. London: Routledge & Kegan Paul, 1982: 249.
[74] Rescher R. Rationality: A philosophical inquiry into the nature and the rationale of reason[M]. Oxford: Oxford University Press, 1988: 1-11.
[75] Daston L. Classical probability in the enlightenment[M]. Princeton: Princeton University Press, 1988.
[76] 熊晓彪. 概率推理: 实现审判智能决策的结构化进路[J]. 中外法学, 2022, 34(5): 1278-1298.
[77] 杜文静. 法律人工智能概率推理的困境与破解[J]. 学术研究, 2022(4): 29-34.
[78] 王家宁, 程宏斌, 宋国庆. 论大数据智能化侦查应用的特点及其构建[J]. 新疆警察学院学报, 2021, 41(1): 32-39.
[79] 杜严勇. 人工智能伦理风险防范研究中的若干基础性问题探析[J]. 云南社会科学, 2022(3): 12-19.
[80] 约翰·罗尔斯. 正义论[M]. 何怀宏, 何包钢, 廖申白, 译. 北京: 中国社会科学出版社, 2009.
[81] 陈仕伟. 大数据技术异化的伦理治理[J]. 自然辩证法研究, 2016, 32(1): 46-50.

[82] 孟天广, 李珍珍. 治理算法: 算法风险的伦理原则及其治理逻辑[J]. 学术论坛, 2022, 45(1): 9-20.
[83] 马长山. 智慧社会背景下的"第四代人权"及其保障[J]. 中国法学, 2019(5): 5-24.
[84] 苏明, 陈·巴特尔. 数字人权的挑战与治理[J]. 电子政务, 2022(3): 101-112.
[85] 梁坤, 陈易璨. 数字化时代侦查学术研究的发展与前瞻: 2016—2020年[J]. 中国人民公安大学学报(社会科学版), 2021, 37(2): 61-70.
[86] 李婷. 人工智能时代的司法公正: 价值效用与风险防范[J]. 江苏社会科学, 2023, doi: 10.13858/j.cnki.cn32-1312/c.20230207.012.
[87] Cummings M. Integrating ethics in design through the value-sensitive design approach[J]. Science and Engineering Ethics, 2006, 12(4): 701-715.
[88] 刘培, 池忠军. 算法的伦理问题及其解决进路[J]. 东北大学学报(社会科学版), 2019, 21(2): 118-125.
[89] 郭林生. 论算法伦理[J]. 华中科技大学学报(社会科学版), 2018, 32(2): 40-45.
[90] 唐皇凤. 数字利维坦的内在风险与数据治理[J]. 探索与争鸣, 2018(5): 42-45.
[91] 韩春梅, 邱文康, 杨宏基. 人工智能技术嵌入智慧警务的潜在风险与规避[J]. 中国人民公安大学学报(社会科学版), 2020, 36(2): 78-86.
[92] 张宪丽, 高奇琦. 人工智能时代公民的数据意识及其意义[J]. 西南民族大学学报(人文社科版), 2017, 38(12): 211-216.
[93] 蒋勇. 从合规性到正当性:我国警察法治体系的重塑——基于"新行政法"理论的展开[J]. 中南大学学报(社会科学版), 2017, 23(2): 85-93.
[94] 张晓华. 数智时代预测性侦查的算法规制研究[J]. 中国人民公安大学学报(社会科学版), 2022, 38(4): 65-74.
[95] 任颖. 算法规制的立法论研究[J]. 政治与法律, 2022(9): 98-111.
[96] 张可. 大数据侦查措施程控体系建构: 前提、核心与保障[J]. 东方法学, 2019(6): 87-94.
[97] 何军. 数据侦查行为的法律性质及规制路径研究[J]. 中国人民公安大学学报(社会科学版), 2021, 37(1): 78-85.
[98] 梁坤, 周韬. 当前人工智能侦查的应用困境及突破进路[J]. 山东警察学院学报, 2018, 30(3): 51-59.
[99] 马忠红. 刑事侦查学[M]. 北京: 中国人民公安大学出版社, 2014: 3-4.

Application of artificial intelligence in criminal investigation: Practice patterns, risk challenges and development strategies

JIN Yifeng[1,2], MA Zhonghong[1*]

1. School of Criminal Investigation, People's Public Security University of China, Beijing 100038, China
2. Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China

Abstract This paper summarizes the common practice patterns of artificial intelligence application in criminal investigation, mainly including video investigation, individual identification, data retrieval, information mining, auxiliary interrogation, evidence evaluation, etc., which brings multi-dimensional lock, situation and finger integration, and full chain support to the investigation department. Then this paper analyzes a series of risks and challenges accompanied by the application of artificial intelligence in criminal investigation, such as black box effect, precision deviation, algorithm overstep, open panorama, investigation secret room and so on. Some suggestions are put forward, such as anchoring the basic stance of human-oriented and human-computer integration, promoting the integration of data-driven and knowledge-guided methods, enhancing the ethical adaptation of artificial intelligence application in criminal investigation, and optimizing the legal system of artificial intelligence application in criminal investigation.

Keywords artificial intelligence; investigation; technical effect; human-machine integration; knowledge to guide

（责任编辑　卫夏雯）