

# 司法精神病鉴定的人工智能技术应用

王迎龙<sup>1</sup>, 宋业臻<sup>2</sup>

(1. 北京工商大学 法学院, 北京 100048; 2. 青岛认知人工智能研究院, 山东 青岛 266073)

**摘要:**司法精神病鉴定实践中的主要问题表现为鉴定意见的信度不足。这一方面源于涉及精神医学、心理学、法学在内的司法精神病学本身属于交叉学科, 学科间话语系统融合存在困难; 另一方面, 司法精神病鉴定的对象本身存在一因多果、多因一果、多因多果方面的复杂性。随着人工智能技术在执法、司法领域的普及应用, 以及算法、算力、数据的不断成熟, 遵循新型技术嵌入精神病学、心理学而产生新型理论的“一级嵌入”与新型理论嵌入司法精神病鉴定实践“二级嵌入”的“双嵌入”路径, 采用“分块开发”“系统整合”“流程整合”三大操作程序, 并借鉴人工智能技术中的自然语言处理技术、心理计算技术和智能辅助决策技术资源, 能够直接回应并一定程度解决司法精神病鉴定实践中信度不足的问题。这为人工智能辅助鉴定系统的研发, 赋能司法精神病鉴定实践提供了可能。

**关键词:**人工智能; 司法精神病鉴定; 技术路径; 信度

**中图分类号:**DF794 **文献标志码:**A **文章编号:**1009-1505(2022)01-0064-10

**DOI:**10.14134/j.cnki.cn33-1337/c.2022.01.007

## 一、引言

根据法律规定, 经法定程序鉴定患有精神疾病的被告人, 可以不负刑事责任、从轻或者减轻处罚。<sup>①</sup>但司法精神病鉴定在实践中产生了适用标准不统一、诊断不规范甚至“被精神病”等乱象, 导致

**收稿日期:**2021-08-20

**基金项目:**教育部人文社会科学基金研究青年项目“精神病人强制医疗解除程序的中国模式及其完善”(17YJC820048)

**作者简介:**王迎龙, 男, 北京工商大学法学院副教授, 法学博士, 主要从事刑事诉讼法学研究; 宋业臻, 男, 青岛认知人工智能研究院研究员, 心理学博士, 主要从事人工智能心理计算研究。

<sup>①</sup>《刑法》第18条规定:“精神病人在不能辨认或者不能控制自己行为的时候造成危害结果, 经法定程序鉴定确认的, 不负刑事责任, 但是应当责令他的家属或者监护人严加看管和医疗; 在必要的时候, 由政府强制医疗。间歇性的精神病人在精神正常的时候犯罪, 应当负刑事责任。尚未完全丧失辨认或者控制自己行为能力的精神病人犯罪的, 应当负刑事责任, 但是可以从轻或者减轻处罚。醉酒的人犯罪, 应当负刑事责任。”

人们对于精神病鉴定公正性问题的担忧。<sup>①</sup>其主要难点问题表现为“精神病的鉴定、认定在司法实践中的混乱情况”“多次鉴定的不一致性较高,由此引发较严重的重复鉴定、多次鉴定”<sup>[1]</sup>。有学者根据部分地区和机构的统计数字推测,存在多次鉴定的案件中,不同鉴定意见之间的不一致率在30%左右<sup>[2]</sup>。从认知科学与临床精神病学角度来看,这一难点问题可以归纳为司法精神病鉴定的“信度不足”<sup>②</sup>问题。

司法精神病鉴定的信度不足问题,主要基于两方面原因产生。一方面,跨学科话语系统差异造成了学术沟通断层。司法精神病学涉及法学、医学、心理学等多种学科,多学科合作本身就面临着极大的困难与挑战。诺伯特·维纳(Nobert Wiener)就指出:“如果一个不懂数学的生理学家和一个不懂生理学的数学家合作,那么这个人不会用那个人所能接受的术语表达自己的问题,那个人也不能用这个人所懂得的任何形式来做出自己的回答。”<sup>[3]</sup>由于法学、医学、心理学所关注的基本问题以及解决这些基本问题的思维路径完全不同,故在学科间合作共同解决一个实践问题时,会出现种种概念、方法、路径、结论上的龃龉。另一方面,司法精神病鉴定对象本身具有复杂多变性,且目前可用于司法精神病鉴定的手段有限,即“迄今对多数精神疾病的诊断,仍然缺乏精密的、客观的理化检验方法或手段,主要还是依据病史和精神状况检查所见,即临床综合征来确定”<sup>[4]</sup>。面对复杂对象时缺乏充分的检测、识别技术手段,使得司法精神病鉴定实践变得更加困难重重。

为了解决司法精神病鉴定信度不足的问题,学界进行了一系列相关研究,主要从以下两种角度切入:一种是从法律程序的角度,对司法精神病鉴定的相关规范进行研究。研究重点包括“精神障碍者强制收治制度”困境<sup>[5]</sup>、国外相关制度借鉴<sup>[6]</sup>、从刑事特别程序角度分析精神疾病患者的强制医疗制度的“保安处分”属性<sup>[7]</sup>、“保安处分”制度适用中的“适当性原则”<sup>[8]</sup>、强制医疗制度的启动模式<sup>[9]</sup>、强制医疗程序中的证据法问题<sup>[10]</sup>,等等;另一种是从诊断与鉴定技术的角度,对司法精神病鉴定的相关方法进行研究。研究重点包括鉴定病例回溯分析<sup>[11]</sup>、针对病例刑事责任能力判别的专门分析<sup>[12]</sup>,以及从域外经验<sup>[13]</sup>、方法论<sup>[14]</sup>、罪犯反社会人格调查<sup>[15]</sup>等角度来研究犯罪危险性(也被称为“人身危险性”)的评估问题。

然而,无论是从法律角度对司法精神病鉴定的规范、程序进行研究,还是从诊断与鉴定技术的角度进行研究,由于学科间话语差异较大,难以找到有效的跨学科沟通交流机制,以及精神病的诊断与鉴定技术受客观技术条件所限,导致上述两方面的研究都难以直接解决司法精神病鉴定意见缺乏信度的问题,难以为提升司法精神病鉴定意见的信度提供切实可行的技术支撑。基于此,本文首先围绕法学与精神医学间的话语差异与精神医学诊断鉴定技术限制导致的鉴定意见的信度不足的问题,分析其背后的复杂成因;其次论述目前人工智能技术如何应用于司法精神病鉴定以及通过人工智能技术解决这些问题的可能性;最后提出如何运用人工智能技术赋能司法精神病鉴定以及提高司法精神病鉴定信度的技术路径。

---

<sup>①</sup>例如江苏省南通市的“5·28亲姐妹硫酸毁容案”中,犯罪嫌疑人王某曾接受南通市精神司法鉴定委员会委托的南通市通济医院、江苏省司法鉴定委员会、司法部上海鉴定科学技术研究所法医精神病学研究室三家鉴定机构的鉴定,但是得出了完全不同的鉴定意见。

<sup>②</sup>此处使用的“信度”概念,来自自然科学、社会科学实证研究领域,指的是采用同样的测量方法,对同一个研究对象进行重复测量时,其所得到的结论相一致的程度。信度分析常见的方法包括重测信度、复本信度、分半信度等,司法精神病鉴定的信度主要指的是重测信度(Test-retest Reliability),即在不同的时间点先后进行两次测量,两次结果之间存在差异的程度。

## 二、难点分析：司法精神病鉴定的信度问题

### （一）法学与精神医学的学科话语差异

司法实践中,司法精神病的鉴定意见往往造成诸多争议,其重要原因在于上文提到的司法鉴定的信度不足问题,这一问题产生的根源在于法学与精神科临床医学的学科间话语差异。我国《刑法》第18条对精神病人在“不能辨认或者不能控制自己行为”条件下实施危害行为造成危害后果的情形进行规定,所关注的要点在于精神病人在实施危害行为时的“精神状态”,而“精神状态”则关联刑事责任能力的判断问题,即需要通过法定程序进行精神鉴定才能作出这一判断。刑事责任能力指的是“辨认能力与控制能力,即行为人实施危害行为时,能够理解自己行为的性质、后果以及社会政治意义,并能够控制自己行为并对此承担刑事责任的能力”<sup>[16]</sup>。我国《刑事诉讼法》第五篇“特别程序”中专章对“依法不负刑事责任的精神病人的强制医疗程序”进行了规定。强制医疗程序在本质上属于保安处分措施,“对实施了危害行为的精神病患者适用的旨在隔离排害和强制医疗的刑事实体措施,其目的在于消除精神病患者的人身危险性、防止再犯,达到社会防卫的目的”<sup>[17]</sup>,保安处分措施的适用“必须达到监管具有持续危险人的目的”<sup>[18]</sup>。从司法证明的角度来看,在强制医疗程序中,需要证明的要件事实在于“被告人的行为是《刑法》所规定的犯罪行为并造成危害”“被告人为不负刑事责任的精神病人”以及“被告人为对社会有现实危害性的人”<sup>[19]</sup>。综合上述话语体系,法学领域所关注的主要概念包括精神病人的刑事责任能力、辨认控制能力、人身危险性等,其中辨认控制能力和人身危险性是法学领域关注的两个最为核心的问题。

在精神病临床医学领域,精神病的诊断一般包括精神状况检查(Mental State Examination, MSE)、病史采集、躯体与神经系统检查以及辅助检查等流程,其中精神状况检查是《精神卫生法》《执业医师法》中规定的“诊断必须以患者的检查结果为依据”内容。而精神病临床检查的原则被归纳为“SSD原则”,即症状学诊断(Symptom)、临床综合征诊断(Syndrome)以及疾病分类诊断(Diagnosis),其中症状学诊断主要对患者进行原发性症状与继发性症状的区分,以进行症状严重等级的判断;综合征诊断与分类诊断则要考虑所有的假设诊断,并优先考虑等级较高的病症,并遵守“马和斑马”诊断原则,即优先考虑常见病与多发病,而非优先考虑疑难病与罕见病<sup>[19]</sup>。同时,在分析病因的时候,采用的是“素质”(Predisposing)、“诱发”(Precipitating)、“持续”(Perpetuating)的“3P因素分析”,全面梳理为什么患者是易感人群、为什么该疾病会发作以及为什么持续难以恢复这三方面的问题。综合上述话语体系,精神病临床学领域所关注的问题是症状、分类诊断、病因。而法学领域关注的辨认控制能力在临床领域一般被称为“自知力”,法学领域关注的“人身危险性”在临床领域一般被纳入辅助诊断与筛查程序,作为其中一个部分进行评估。

从上述话语分析中可以看出,法学领域关注的刑事责任能力与人身危险性同临床医学领域关注的症状与综合征存在学科间话语差异,两个学科之间的思维路径与所关注的问题也存在较大差异。但是,法学领域尤其是司法裁判过程中,对精神病人责任能力的判断极大地依赖于医学领域的专业判断,加之法学领域对临床医学诊断逻辑与主要概念的生疏,造成“司法人员对鉴定意见往往一味地接受,极少进行法定的证据能力与证明力的审查,司法机关对精神医学鉴定的采信率高达90%以上”<sup>[20]</sup>。由此,精神医学领域对精神病人的自知力与人身危险性程度的判断,就成为影响司法精神病鉴定准确性与信度的一个关键因素。

## (二) 司法精神病鉴定的方法及其不确定性

实践中,公安司法人员对于犯罪嫌疑人、被告人是否患有精神疾病的判断高度依赖于司法精神病鉴定意见。问题在于,在精神医学实践领域,一般采用临床症状观察的方式进行自知力判断,采用各项测验进行人身危险性评估,这些方式本身即带有极大的主观性和不确定性。精神医学临床测验一般包括常规心理测验,如 PANSS 精神症状评定量表、BPRS 简明精神病评定表、SCL-90 症状检查表、MMPI 明尼苏达多项人格测验、CGIS 临床总印象表、CIDI 复合性国际诊断交谈检查、SCID 复合性国际诊断交谈检查、MINI 简明国际精神障碍交谈检查表等;还包括专门性疾病筛查测验,如针对双相情感障碍的 MDQ 心境障碍问卷、HCL-32 轻躁狂症自评;以及包括暴力风险评估测验,如 Broset 暴力清单测验、HCR-20 敌意临床风险测验、OAS 外显攻击量表、VRS 暴力风险量表。尽管已经存在多种类型、多种形式、多个代际的风险评估工具,但是仍然难以有效预测精神病人的人身危险性,究其根本存在以下三个方面的原因。

其一,人身危险性的内涵复杂性与生成机理的多因性。人身危险性属于刑法学与犯罪心理学领域中的常见概念,刑法学话语中人身危险性指的是“罪犯再次实施危害行为的可能性”<sup>[21]</sup>,犯罪心理学中人身危险性则不仅包括再次实施危害行为的可能性,还涵盖了初次实施危害行为的可能性;此外,刑法学话语中人身危险性一般指向除了行为人之外他人的危害行为危险性,而犯罪心理学话语中人身危险性也宽泛地涵盖了自伤、自杀、自残等指向自身的危害行为危险性。可见,人身危险性本身的内涵与外延是极为复杂的,除此之外,其成因也极为复杂,存在一果多因、多因多果的复杂作用机制<sup>[22]</sup>,即人身危险性的生成机理具有多因性。举例而言,暴力风险在司法精神病鉴定实践中主要指的是精神病人做出暴力行为的可能性,就其成因来看,包括至少以下三种原因:第一种是基于幻觉、妄想的暴力行为,精神病人处于幻觉或者妄想状态下,尤其是被害妄想状态下,会时常做出基于被害妄想的防御与抵抗行为,在外部表现为随意打杂物品与殴打他人的暴力行为;第二种是当人类在面对刺激较大、距离较近的威胁事件时,均会产生的“反应性攻击”行为<sup>[23]</sup>,但是部分精神病患在运动控制方面存在障碍,则可能在产生“反应性攻击”意向与行为时,无法控制自身的行为运动、行为幅度,从而产生“程度超限”的暴力行为;第三种仍然是在出现负面威胁与刺激情况下,即“反应性攻击”的类型,只是正常人类并不会因为该强度的负面刺激存在极大的威胁性而启动攻击应对反应,而存在情绪控制能力障碍的精神病患则错误地认为该强度的负面刺激威胁性较大从而产生暴力应对行为。上述第一种类型一般对应精神分裂障碍患者,第二种类型一般对应伴随运动障碍的患者,第三种类型一般对应情感型障碍患者。从疾病类型、症状表现、风险行为三者的对应关系来看,其间的关联性极为复杂,因果关联也具有多态性与复杂性。

其二,既有评估工具在形式上的局限性。既有的评估工具在形式上以问卷量表为主,其中问卷量表又可以进一步分为自评问卷量表与他评问卷量表。问卷量表形式的评估工具本质上是基于经验归纳法形成的标准化检查表,即通过样本分析与统计检验,找出与风险行为密切相关的预测变量,并对预测变量的重要程度进行排序与加总,根据评估参与者在问卷量表上的得分估计其发生风险行为的可能性。但是,无论是自评还是他评问卷量表,本质上均难以改变其依靠测评参与者自我经验观察与自我体验报告的数据生成方式,即问卷量表实质上仍然是按照“主观主义”路径实现数据采集,并非完全达到了“客观主义”路径数据采集要求。“主观主义”路径在数据采集上很可能受到测评参与者“社会赞许效应”以及主观观察、主观体验自带偏差导致的数据偏差,从而极大影响测评结果的真实性与准确性。

其三,既有评估工具在逻辑上的局限性。既有评估工具的开发主要是医学与心理学研究领域的

学者,评估工具的开发以探索个体差异性层面的风险预测因素为主。以暴力风险评估工具为例,在VRAG(Violence Risk Appraisal Guide)工具中,VRAG评估工具主要对评估参与者是否存在精神障碍与人格障碍,以及其幼年时期的生活状态、是否存在异常行为史等预测因素进行研究,预测其出现暴力行为的可能性<sup>[24]</sup>。在逻辑上,既有的评估工具几乎均是遵循这一评估逻辑,从评估参与者个体差异性角度出发寻找预测因素,从而对评估参与者是否可能出现暴力行为进行预测。但是在司法精神病鉴定实践中,不能仅仅关注评估参与者在人身危险性方面的个体差异性,还需要关注个体内差异性,也就是“在什么时候/什么情况下”作为评估参与者的精神病患可能出现危害性行为。当前主流评估工具在逻辑上缺乏这一设计。

综上,一方面,法学与精神医学在学科之间存在较大话语差异,法学领域重点关注精神病患者的刑事责任能力、辨认控制能力与人身危险性,但是在精神医学领域重点关注患者的症状、分类诊断与病因,并且除了“自知力”对应“辨认控制能力”之外,两个学科尚未找到其他具有共识性的沟通交流概念;另一方面,由于精神障碍本身的成因与机制复杂性、评估工具形式缺乏客观性以及评估工具难以做到对患者个体内差异性的追踪测量评估等因素,共同导致司法精神病鉴定的信度不足问题。

### 三、技术前瞻:人工智能应用于司法精神病鉴定的有效性

#### (一) 人工智能技术的实践发展

在执法、司法实践中,人工智能技术已经取得了广泛的应用。具有代表性的如在公安部门主导建设的“雪亮工程”基础上,进一步增加各类具有如人物身份识别、车辆信息识别、数据汇总决策、智能研判功能的智能系统,以增强公安机关在犯罪嫌疑人发现、证据采集等多个方面能力;又如在《人民法院信息化建设五年发展规划(2016—2020)》中提出到2020年深化完善人民法院信息化3.0版建设任务,基本建设“智审、智执、智服、智管”的智慧法院体系<sup>[25]</sup>。目前,智慧法院建设项目相关的人工智能系统已经实现了执行信息化、电子卷宗随案同步生成、增强电子诉讼服务能力以及提供司法大数据服务等方面的效果。审判机关通过网上法庭、庭审直播与电子签章这三项信息化基础应用,皆在实际工作中起到了提高审判效率的作用<sup>[26]</sup>。在司法权运行体制的建构中,杭州等地最先成立的“互联网法院”已经作出了重要的探索<sup>[27]</sup>。除此之外,以人工智能技术为支撑研发的新型智能测谎系统、犯罪危险性智能化评估系统也逐渐在实践中开始广泛应用<sup>[28]</sup><sup>98</sup>。总言之,在执法、司法实践中,人工智能技术已经开始全面涉入执法信息化、司法信息化建设中,逐渐开始发挥重要支撑与服务作用。

另外,当前人工智能技术的快速发展主要依靠三大方面的技术基础——算力、数据、算法。从算力角度来看,随着芯片技术的日益成熟完善,行业界逐渐开发出了适用于人工智能深度学习模型的智能芯片,在硬件结构上能够适用于处理深度学习模型与算法的运算方式。此外,随着类脑智能研究的发展,芯片行业针对传统冯·诺依曼结构计算机存在的“内存墙”问题,进一步在人类大脑结构启发下,研制出了模拟人类大脑生物结构与生理运行机制的类脑芯片。例如IBM公司在2014年研制的True North类脑芯片、斯坦福大学电子与微系统实验室所研制的人工突触芯片,可用于支撑手写字体的识别;中国科学院自动化研究所类脑智能中心、类脑信息处理(BRAVE)在生物神经结构启发下完成的相关研究。从数据角度来看,当前以深度神经网络为核心的人工智能技术,尤其是其中应用在工业实践中最广泛、最成熟的监督学习技术,需要大量的标注数据。随着近年来“众包”模式的发展以及专门化数据公司的发展,面向各行各业、各种专门化需求的标注数据库日渐完善起来,为深度学习模型提供了足够的样本学习资源,从而加快了以监督学习技术为主的人工智能技术向应用领域转化的速度。而且,随着人工智能自主学习、深度学习

等方面不断成熟与发展,人工智能不再只是人脑的自然延伸状态,而有了属于自己的自运转系统,从而开始拥有自己的知识创造性与创新性<sup>[29]</sup>。有些学者认为在强智能时期,人工智能甚至将拥有自己独立的意识与意志<sup>[30]</sup>。从算法角度来看,无论图像识别算法、自然语言处理算法,还是推荐算法,均在人工智能领域的不断突破下取得了显著的能力提升;从应用层面来看,从机器学习到深度学习的算法跃升,使得原先很多难以计算的关联关系逐渐被突破,使得人工智能算法在应用层面的能力实现了一次跨越式升级。

## (二) 精神医学领域的人工智能技术应用

在精神医学领域,人工智能技术目前已经进行了大量的实验工作,并取得了不俗的实践应用效果,主要体现在以下两个方面:

一方面,当前人工智能技术主要应用于精神障碍的辅助诊断与风险筛查领域,在基础医学研究中也进行了大量应用。例如在抑郁障碍的辅助诊断中,已经采用了话语行为、眼部活动、头部姿态多模态融合计算技术,以支持向量机为计算模型的抑郁障碍的辅助诊断,在随机对照实验中取得了测试准确率90%以上的检验效果<sup>[31]</sup>。在抑郁障碍的基础医学研究中,为了区分重度抑郁障碍(Major Depression Disorder, MDD)和双极抑郁障碍(Bipolar Depression Disorder, BD),采用前馈搜索策略的支持向量机模型(SVM-FoBa)进行脑影像学的数据分析,由此找到区分两者的生物标记物,在样本量为69人的临床试验中,取得了区分度大于92.07%的检验效果<sup>[32]</sup>。此外,在基础医学研究中也使用了大量人工智能技术,典型的如在类脑智能技术研究中,使用了基于IGIO突触晶体管进行突触抑制性模拟的技术,以人工神经形态芯片模拟人类神经环路运行机制,对重度抑郁障碍进行研究<sup>[33]</sup>。

另一方面,人工智能技术分支中的计算机视觉技术、计算机语音识别技术、计算机自然语言处理技术和类脑芯片技术也广泛应用于精神障碍基础研究与辅助诊断中。计算机视觉技术主要应用于对患者的神经影像学资料进行分析,辅助临床医生作出更加精确化的诊断。如在精神障碍分类诊断过程中,传统诊断方法在双相情感障碍与抑郁障碍的区分上存在较大困难,80%的双相情感障碍被误诊为抑郁障碍<sup>[34]</sup>。但是在多模态融合的计算机视觉技术辅助下,对脑神经影像学的分析精确度得到了大大提升,尤其是在融合低频波动分数振幅(FALFF)、脑灰质体积以及以体素为基础的形态测量学数据基础上,对神经环路异常的探测与辅助诊断准确率能够提升到90%以上。此外,计算机视觉技术还应用于患者外部症状表现的采集分析,尤其是在抑郁障碍患者的非语言行为特征的采集分析方面,使用计算机视觉技术对患者的面部活动与表情特征进行采集分析,能够准确地区分抑郁障碍与非抑郁障碍的患者,使用计算机视觉技术对患者的头部姿态与头部活动进行采集分析,建立抑郁障碍判别模型以及对患者的注视轨迹与眼部活动进行采集分析,建立抑郁障碍判别模型,均取得了良好的检测准确率。除了计算机视觉技术之外,计算机语音识别技术与自然语言处理技术也在精神障碍筛查与辅助诊断中发挥着重要作用,尤其是计算机自然语言处理技术,在互联网论坛、推特、微博上进行语言数据采集与智能分析之后,能够对社会公众中存在抑郁情绪与抑郁障碍风险的人群进行精确定位,快速在人群中筛查出具有风险的个体<sup>[35]</sup>。

总体而言,无论从人工智能技术在司法实践领域中的广泛应用,还是从人工智能技术在精神医学实践领域中的技术迭代与不断完善,都为人工智能技术在司法精神病鉴定领域的应用提供了支撑。有待解决的问题是,应当选取人工智能技术的哪个分支,以何种方式切入司法精神病鉴定领域,真正有效地赋能司法精神病鉴定实践。

## 四、嵌入路径:人工智能技术赋能司法精神病鉴定

人工智能技术赋能司法精神病鉴定实践,需要紧密围绕解决司法精神病鉴定意见信度不足的问题

题,针对法学与医学学科间话语差异、精神病诊断与鉴定测量工具方法局限性两方面成因,引入直接相关的技术体系,开发切实有效的智能系统,从技术与方法角度为解决司法精神病鉴定意见信度不足提供支撑。下文着重阐述人工智能赋能司法精神病鉴定需要遵循怎样的原则、应当采用怎样的技术路径设计智能系统以及哪些技术与解决该问题密切相关。

### (一)“双嵌入”的应用原则

在具体的嵌入路径上,心理计算技术体系可以按照“双嵌入”路径逐渐应用于司法精神病鉴定实践<sup>[28]103</sup>。“双嵌入”路径指的是新技术嵌入传统理论改变传统理论与方法的一级嵌入和新理论与新方法嵌入工作实践改变实践方式的二级嵌入路径。按照“双嵌入”原则的思想,造成司法精神病鉴定意见信度不足的学科话语差异方面的原因以及传统精神障碍诊断与评估方法、工具局限性方面的原因,均源于“二级嵌入”层次的传统理论与传统工作方法方面的局限性。而人工智能技术本身属于一种客观、抽象、工具化的客观实在,借鉴人工智能技术几乎不可能直接解决法学与精神医学之间的话语差异或者直接突破传统精神障碍诊断与评估工具的局限性。那么,首先需要进行“一级嵌入”,将人工智能技术分别引入法学和精神医学领域,依靠人工智能技术分别解决法学与精神医学话语差异和精神障碍诊断与评估工具存在局限性的问题。

在“一级嵌入”层面上,人工智能技术需要分别嵌入法学与精神医学两个领域,解决与司法精神病鉴定问题密切相关但是又无法通过传统方法解决的问题。一方面,法学与精神医学话语差异较大的问题可以提炼为如何实现司法决策与诊断决策的智能化辅助问题。司法决策与精神障碍诊断、鉴定决策均属于决策问题,两者的差异在于决策表征维度差异,也就是作出决策依赖的标准、所考虑的因素、判断分析问题的逻辑推理路径与角度存在差异。人工智能技术在“一级嵌入”层次上,需要将两种学科的决策标准、决策涉及因素、判断分析的逻辑推理路径进行智能化记录、分析与便捷化展示,由此实现对两类决策的分别辅助与沟通。另一方面,精神障碍诊断、鉴定工具存在局限性的问题可以提炼为如何利用人工智能技术增加诊断依赖信息量、提升诊断依赖信息精确度与客观度的问题。人工智能技术在“一级嵌入”层次上,需要利用自身的优势为精神障碍诊断者与鉴定者提供更多维度、多尺度、多方法的信息,提高诊断所依赖信息的精确度与客观度。

在“二级嵌入”层面上,主要考虑人工智能技术分别和法学与精神医学结合之后,由此研发出的智能辅助决策系统的可用性,也从司法精神病鉴定业务实践角度考虑智能系统的可用性与易用性。解决可用性与易用性的问题,需要遵循“分块开发”“系统整合”与“流程整合”的系统设计与研发逻辑。“分块开发”是指技术开发人员与司法精神病鉴定人员进行密切配合,调查研究司法精神病鉴定人员与司法裁判者所关注的影响精神病鉴定的关键指标,并结合技术资源探讨哪些关键指标可以通过已有的技术手段实现,并将其逐个、分块开发实现,通过对实践工作者的调研、需求提炼、模拟设计与功能探讨来满足实践工作者的业务要求。“系统整合”指的是将“分块开发”的系统模块拼装成一个完整系统,重点解决的是业务信息、业务数据的传输与分发的问题。在司法精神病鉴定实践中,哪些信息和数据需要传输给精神病诊断与鉴定部门,哪些信息和数据需要传输给司法裁判部门,哪些数据和信息需要同时在两个部门留存与备份,都是在“系统整合”阶段完成的任务。“流程整合”指的是如何实现智能系统与司法精神病鉴定工作组织与工作机制有效衔接,具体表现为安排哪些部门使用智能系统,是否需要专门岗位与部门使用智能系统,智能系统是否能够简化传统工作流程,以降低实践部门的实际操作负担。

总而言之,在“双嵌入”应用原则指导下,“一级嵌入”的核心理念是依靠人工智能技术解决传统理论与方法难以解决的问题,将司法精神病鉴定实践中传统上几乎不可能解决或者难以解决的问题利

用技术予以化解。“二级嵌入”的核心理念则是将具有更高科学性的研究成果与工作实践结合,不仅要实现人工智能系统的可用性,更要让人工智能系统变得“好用”,切实提升实践部门的工作效率,降低重复工作负担。

## (二) 自然语言处理、心理计算与智能辅助决策技术资源

在“双嵌入”应用原则的指导下,当前与司法精神病鉴定密切相关的人工智能技术资源主要包括自然语言处理技术、普适心理计算技术与智能辅助决策技术。

首先,自然语言处理技术能够对大规模文本进行中心词与关键词识别、摘要抽取与主题抽取,辅助用户进行快速的大规模文本分析与理解、知识梳理、信息检索。如果应用于司法精神病鉴定实践,自然语言处理技术能够辅助精神病诊断鉴定人员与司法裁判人员进行案情线索梳理、关联关系分析、类案推理、类案检索等工作,能够极大地提升鉴定人员与裁判人员对案情的分析与研判的能力,帮助鉴定人员与裁判人员在一系列案情描述、证据中精确、快速地找到与精神病鉴定相关的信息。在国内司法实践领域,目前应用较多的自然语言处理技术案例主要是阿里云与达摩院联合研发的自然语言处理开放平台、睿企科技研发的文本大数据分析系统。

其次,心理计算技术通过摄像机采集人像图片或者视频数据、麦克风采集语音数据、社交网络行为数据、社会行为数据等多种类型的数据并进行智能建模分析,使用户能够随时随地对自我与他人的心理特征、心理过程、心理状态进行全面计算分析。针对临床精神医学实践中常用的风险评估工具中存在的形式问题与逻辑问题,心理计算技术体系一方面将原先“主观主义”的数据采集方式变成了纯粹客观化的数据采集方式,另一方面不再依赖被鉴定、被评估对象主动陈述或者提供资料信息,从而消除了其提供存在偏差信息的可能性。

举例而言,中国心理学会法律心理学专业委员会、中国政法大学犯罪心理学研究中心与青岛认知人工智能研究院联合研发的“非接触实时动态心理评测系统”就是心理计算技术的一个应用示范,系统能够对测评参与者在长连续时间序列的心理状态进行分析,建立其心理表征的电子跟踪记录。对于司法精神病鉴定实践而言,获取被鉴定人在一个连续的长时间段内的心理状态跟踪记录,能够极大地提升鉴定人对被鉴定人症状特征、症状表现、风险等级等诸多方面情况判断的客观性、准确性。在精神障碍分类诊断中,根据DSM-V手册的分类规范,在双相障碍(Bipolar Disorder)中的双相I型障碍、双相II型障碍以及其中的复发性抑郁障碍(Recurrent Depressive)、轻躁狂(Hypomanic Episodes)研究较少,区分较困难<sup>[36]</sup>。早在克雷丕林(Kraepelin)时期,“躁郁症”(Manic-Depressive Insanity)的描述就包含了现在几乎所有的情绪障碍,但是在其中亚型并未作进一步区分<sup>[37]</sup>。根据当前的分类诊断标准,上述几种障碍均涵盖了躁狂、抑郁之间的状态转移,亚型与亚型之间区别仅在于分布时间段、强度与组合方式的差异。若采用传统结构化访谈或者填写量表的方法,几乎很难在一个持续周期内精确采集到患者在躁狂、轻躁狂、抑郁情绪上的状态转移情况,故难以在几种抑郁障碍亚型之间准确进行分类诊断。采用先前所述的心理计算系统,能够每日连续多次对患者进行情绪状态、情绪类型、情绪异常等级的自动采集记录,形成追踪式的情绪电子病历,从而辅助临床精神科医师对患者躁狂、轻躁狂、抑郁状态的分布时间长度、持续时间长度、情绪强度、情绪异常等指标进行精确化、计量化分析,从而为分类诊断提供客观、精确的依据。

最后,智能辅助决策系统主要用于辅助用户作出科学、精准的决策。一方面,智能辅助决策系统为决策者提供更加多元化的决策基础信息,即在决策信息量上增加数量与质量,为决策者提供更多的决策支撑信息;另一方面,智能辅助决策系统能够帮助决策者筛选、过滤无关信息,提取与决策高度相关的信息,在具体决策辅助上降低信息的冗余度,提升决策者加工信息的效率。具体而言,智能辅助决策



系统能够帮助司法裁判人员与精神病诊断与鉴定人员进行跨学科话语沟通,根据用户身份为用户提供对应的决策支持信息,如向司法裁判人员推送与人身危险性、辨认控制能力相关的诊断病历信息与案情信息,向精神病鉴定人员推送患者行为描述、测评结论等信息。

综上,以“双嵌入”原则为指导,遵循“分块开发”“系统整合”和“流程整合”的设计与研发路径,依靠自然语言处理技术、心理计算技术与智能辅助决策技术资源开发一套人工智能辅助精神障碍鉴定的系统,能够赋能司法精神病鉴定实践,一定程度上解决因学科间话语差异较大、精神病评估工具与方法精确性与客观性欠缺而导致的司法精神病鉴定意见信度不足的问题。

#### 参考文献:

- [1]陈卫东,程雷.司法精神病鉴定的基本问题研究[J].法学研究,2012(1):163-178.
- [2]胡纪念,李从培.110例精神病重复鉴定结论分歧原因分析[J].法医学杂志,1995(2):103-104.
- [3]顾凡及,卡尔·施拉根霍夫.脑研究的新大陆:一位德国工程师和一位中国科学家之间的对话[M].上海:上海教育出版社,2019:34.
- [4]李从培.司法精神病鉴定的实践和理论[M].北京:北京医科大学出版社,2000:13.
- [5]叶肖华.论我国精神障碍者强制收治制度的困境与出路[J].社会科学战线,2014(6):198-202.
- [6]胡肖华,董丽君.美国精神病人强制住院治疗法律制度及其借鉴[J].法律科学,2014(3):184-191.
- [7]陈卫东.构建中国特色刑事特别程序[J].中国法学,2011(6):32-42.
- [8]刘夏.德国保安处分制度中的适当性原则及其启示[J].法商研究,2014(2):136-142.
- [9]田圣斌.强制医疗程序初论[J].政法论坛,2014(1):62-72.
- [10]纵博,陈盛.强制医疗程序中的若干证据法问题探析[J].中国刑事法杂志,2013(7):89-95.
- [11]沈志环,徐惠霞,胡永,等.50例精神发育迟滞司法精神医学鉴定的初步分析[J].临床精神医学杂志,1994(2):115.
- [12]吕盼,刘建梅,胡峻梅.3720例刑事责任能力鉴定案例分析[J].华西医学,2014(8):1525-1530.
- [13]倪涓.强制医疗程序中“社会危险性”评价机制之细化[J].法学,2012(11):89-99.
- [14]陈伟.认真对待人身危险性评估[J].比较法研究,2014(5):59-72.
- [15]蒋奖,许燕.罪犯反社会人格障碍的调查[J].中国特殊教育,2007(5):80-85.
- [16]马克昌.犯罪通论[M].武汉:武汉大学出版社,1999:244.
- [17]陈卫东.构建中国特色刑事特别程序[J].中国法学,2011(6):32-42.
- [18]时延安.隐性双轨制:刑法中保安处分的教义学阐释[J].法学研究,2013(3):140-157.
- [19]唐宏宇,方贻儒.精神病学[M].北京:人民卫生出版社,2020:7.
- [20]何恬.重构司法精神病学[M].北京:法律出版社,2008:156.
- [21]曲新久.刑法学[M].北京:中国政法大学出版社,2016:65.
- [22]朱迪亚·珀尔,达纳·麦肯齐.为什么:关于因果关系的新科学[M].北京:中信出版社,2019:98.
- [23]BLANCHARD R J, BLANCHARD D C, TAKAHASHI L K. Attack and Defensive Behaviour in the Albino Rat[J]. Animal Behaviour,1977,25(3):197-225.
- [24]HARRIS G T, RICE M E, QUINSEY V L. Violent Recidivism of Mentally Disorder Offenders;The Development of a Statistical Prediction Instrument[J]. Criminal Justice and Behavior,1993,20(4):315-335.
- [25]继续深化司法改革 加快建设智慧法院——二论深入贯彻学习党的十九届五中全会精神[N].人民法院报,2020-11-06(1).
- [26]谢登科.论在线诉讼中的个人信息保护[J].中国政法大学学报,2022(1):127-137.
- [27]章安邦.人工智能时代的司法权嬗变[J].浙江工商大学学报,2020(4):149-160.
- [28]马皑,孙晓,宋业臻.犯罪危险性智能化评估的理论与实践[M].北京:中国法制出版社,2020.

- [29]王雪乔.人工智能生成物的知识产权保护立法研究[J].湖南科技大学学报(社会科学版),2020(2):96-102.
- [30]刘宪权,张俊英.人工智能时代机器人异化与刑事责任[J].法治研究,2019(4):90-97.
- [31]ALGHOWINEM S, GOECKE R, WAGNER M. Multimodal Depression Detection; Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behaviors[J]. IEEE Transactions on Affective Computing, 2018, 9(4):478-490.
- [32]JIE N F, ZHU M H, MA X Y. Discriminating Bipolar Disorder From Major Depression Based-on SVM-FoBa; Efficient Feature Selection with Multimodal Brain Image Data[J]. IEEE Transactions on Autonomous Mental Development, 2015, 7(4):320-331.
- [33]WANG J B, LI Y X, YIN C Q, et al. Long Term Depression Mimicked in an IGIO-Based Synaptic Transistor[J]. IEEE Electron Device Letters, 2017, 38(2):130-137.
- [34]HIRSCHFELD R M A, VORNIK L A. Bipolar Disorder-Costs and Comorbidity[J]. American Journal of Managed Care, 2005, 11(3):85-90.
- [35]ALGHAMDI N S, HANAN A H M, AJITH A. Predicting Depression Symptoms in an Arabic Psychological Forum[J]. IEEE Access, 2020, 8(1):57317-57334.
- [36]FRANCO B. Bipolar Disorder-Focus on Bipolar II Disorder and Mixed Depression[J]. The Lancet, 2007, 369(3):935-945.
- [37]KRAEPELIN E. Manic-Depressive Insanity and Paranoia[J]. The Journal of Nervous and Mental Disease, 1921, 53(4):324-350.

## Application of Artificial Intelligence in Forensic Psychiatric Expertise

WANG Yinglong<sup>1</sup>, SONG Yezhen<sup>2</sup>

(1. School of Law, Beijing Technology and Business University, Beijing 100048, China;

2. Qingdao Institute of Cognitive Artificial Intelligence, Qingdao 266073, China)

**Abstract:** The main problem in the practice of forensic psychiatric expertise is the lack of credibility of expertise. There are two reasons. First, judicial psychiatry is an integrated discipline which involves psychiatry, psychology and law. Second, the object of forensic psychiatric expertise has the complexity of one cause and one effect, multiple causes and one effect, multiple causes and multiple effects. With the continuous maturity of algorithm, computing power and data, artificial intelligence technology gradually grew in the field of law enforcement and justice. Following the “double embedding” path of “level-one embedding” of new technology embedded in psychiatry and psychology and “level-two embedding” of new theory embedded in forensic psychiatric practice, we can adopt three major procedures of “block development”, “system integration” and “process integration” to develop artificial intelligence. Using the natural language processing technology, psychological computing technology and intelligent decision making technology, we can directly resolve the problem of insufficient credibility in the practice of forensic psychiatric expertise. This provides possibility for the research and development of artificial intelligence system to improve the practice of forensic psychiatric expertise.

**Key words:** artificial intelligence; forensic psychiatric expertise; technology path; credibility



(责任编辑 张伟 郑英龙)