

# 智能化再犯风险评估司法应用的法律 风险及其防范<sup>〔\*〕</sup>

狄小华

(南京大学 法学院, 江苏 南京 210093)

〔摘要〕遏制犯罪必须首先有效防治再犯,而再犯风险评估是精准防治再犯的前提。再犯风险评估理论与技术虽然经历了近百年的发展,形成了五代再犯风险评估工具,但仍面临着影响再犯风险因素不全、权重不准以及操作不便等难题。信息技术特别是人工智能的快速发展,一方面为实现更为准确的评估和在此基础上的精准施策创造了条件,但另一方面也带来了技术偏见、信息保密等风险。为此,智能化再犯风险评估的研发和应用,需要协调人文与技术关系,形成产、学、研、用战略合作,并通过理论模型优化、规范研发、算法监督等有效防范智能化评估带来的风险。

〔关键词〕再犯风险;智能化评估;应用风险;防范对策

DOI:10.3969/j.issn.1002-1698.2020.05.007

## 一、引言

“凡事预则立,不预则废”。犯罪一旦发生必然造成对国家、社会和个人的危害,因此预防一直被人们看作是应对犯罪的治本措施。相对于初犯预防,正在服刑罪犯或刑满释放人员的再犯预防不仅价值更高,而且可行性更强,因此一直备受学界的关注,并深刻影响着各国的刑事司法。<sup>〔1〕</sup>自刑事实证学派代表人物提出人身危险性概念以来,兼顾报应与预防的“一体刑”不仅已经成为国际刑事立法的趋势,而且已经成为我国

刑事立法的现实。与报应之刑由犯罪分子已经实施的客观犯罪事实来决定不同,预防之刑是由主观性很强的犯罪分子的人身危险性所决定的,因此,如何评估犯罪分子的人身危险性或再犯风险,<sup>〔2〕</sup>并不断提高这种评估的精准性和可靠性,成为刑事司法必须解决的难题,也因此成为犯罪学、法学、心理学等多学科研究的重点课题。

从满足适用预防刑的人身危险性评估到满足再犯风险管理的再犯风险评估,从以心理学、犯罪学等单一学科理论为依据到以综合多学科的专业评估理论为依据,一百多年来,再犯风险

作者简介:狄小华,南京大学法学院教授、博士生导师,南京大学犯罪预防与控制研究所所长,研究方向:犯罪学、刑事法学。

〔\*〕本文系国家重点研发计划资助项目“假释、暂予监外执行、刑释人员犯罪预防支撑技术与装备研究”(2018YFC0831100)的成果之一。

评估经历了四代历史演进:无确定评估指标的一代经验性评估(1.0)、有确定静态指标因素的二代精算评估(2.0)、有确定静态和动态指标因素的三代精算评估(3.0)、有确定评估指标因素且由多元主体参与评估,并服务于风险管理多元目的的四代系统评估(4.0)。随着认知神经科学的发展,通过检测脑电及人体多巴胺、血清等物质变化评估再犯风险的第五代评估工具(5.0)也已经出现。<sup>[3]</sup>值得注意的是,再犯风险评估在西方国家尽管发展到了五代,评估的精确性和可靠性有了很大提高,但仍难以达到司法公正的要求,所以在司法实践中的应用并不普遍。直到20世纪中后期,再犯风险评估在网络、大数据和人工智能等技术的支持下,由于更为精准和便捷,因此得以在美国等西方发达国家广泛应用。

人身危险性概念诞生于犯罪学,后来才成为刑法学共用的一个概念。在我国,犯罪学虽然在20世纪80年代逐步从刑法学中独立出来,但仍被列为刑法学下面的二级学科。在我国,对人身危险性的刑法学研究一直占据主导地位,不仅深刻影响了犯罪学对人身危险性的理解,而且严重阻滞了我国再犯风险评估理论和技术的研发、应用。相对于西方百年历史、五代发展,我国的再犯风险评估理论研究和技术开发,由于起步晚、发展慢,目前仍处于二代到三代之间,因此难以为我国再犯预防特别是刑罚个别化提供强有力的技术支撑。随着信息技术的迅速发展,特别是在国家以重点专项(司法专题任务)资助智能化再犯风险评估技术研发的大背景下,我国的再犯风险评估发展正迎来新的机遇。但要实现后来者居上,则需要充分认识智能化再犯风险评估应用的潜在风险,在研发过程中做到未雨绸缪。

## 二、智能化再犯风险评估司法应用的特征

再犯风险评估的司法应用经历了由人工操作到计算机辅助的发展,随着人工智能的兴起,目前正处于智能化研发和应用的新阶段。美国是将再犯风险评估智能化最早的国家,根据该国

学者 Sarah L. Desmarais 和 Jay P. Singh 2013 年对美国广泛使用的 19 种再犯风险评估工具的智能化管理情况进行调查和评估,发现有 12 种属于第四代,4 种属于第三代,3 种属于第二代,再犯风险评估工具的“智能化”趋势明显。<sup>[4]</sup>再犯风险评估的智能化虽然不是指称新一代再犯风险评估工具,但它极大地提高了评估的准确性、可靠性和操作的便捷性,因此有力地促进了再犯风险评估在司法领域的广泛应用。不同于非智能化再犯风险评估的司法应用,智能化再犯风险评估的司法应用具有鲜明的技术、运行和实践特征。

### (一)智能化再犯风险评估司法应用的技术特征

不论是经验性临床评估,还是结构化的系统评估,都离不开具有临床经验或具有专业知识的专家。<sup>[5]</sup>与此不同,智能化再犯风险评估主要是由评估系统或机器自动进行的,涉及到一系列新技术的应用,其中最关键的技术是人工智能。而人工智能最核心的又是算法,具体可分为两大类:一类是工程学方法,主要从逻辑层面设计并编写算法;另一类为模拟法,更多通过模拟人和动物机体所用的方法来设计算法。以工程学方法实现的人工智能由于其内在算法是固定的、机械性的,因此它具有明显的局限性,即随着算法逻辑层次的不断深入,会因耗费的人力、时间、经济等成本过大而导致算法无法深入。模拟法包括机器学习和遗传算法,前者虽然被科技界普遍认为是实现人工智能的根本途径,但与后者的相互渗透和结合也已经成为趋势。目前的机器学习需要先建立起机器学习模型,并由人搜集模型需要的大量数据来喂饲,因此本质上仍是一种机器的被动学习。这种缺乏自主学习的人工智能再犯风险评估系统的评估准确性和可靠性,一定程度上取决于喂饲数据的数量与质量。与机器的被动学习不同,机器的自主学习尽管前期也需要经历被动学习阶段,但最终可以自主根据需要在环境里选择需要的数据,并发展出自主编程能力而不断自我进化。<sup>[6]</sup>具有自主学习、自主编程

能力的人工智能再犯风险评估系统,不仅评估效率更高,而且评估的精准性更高。目前应用于再犯风险评估的机器学习是这样一种技术:通过对已有数据的训练,发现这些反映个人生物、心理、社会等特征的数据和再犯之间的联系。而所谓的机器学习算法不过是一系列指令,一种有效的搜索联系的方法。<sup>[7]</sup>由此,目前再犯风险评估的智能化仍处于机器被动学习阶段。

智能化再犯风险评估可能运用的技术非常广泛。为了促进公共安全,美国司法研究所(National Institute of Justice, NIJ)在2016年4月发布的一份公告中宣称,NIJ对研究、发展、评估以下技术提供资助:生物技术、防护技术、监控技术、计算机网络技术、矫治设备和技术、地理空间和犯罪地图技术、管理技术等。<sup>[8]</sup>同样为了公共安全,根据国科发资[2018]9号《科技部关于发布国家重点研发计划公共安全风险防控与技术装备重点专项(司法专题任务)2018年度第一批项目申报指南的通知》(以下简称《通知》),我国在未来三年,要通过产、学、研、用联合攻关,形成一批具有中国特色、引领世界司法技术和装备发展的先进技术成果,初步形成智能化再犯风险评估平台,具体涉及:研发面向服刑人员改造的心理表征和行为模式智能采集技术与装备;研究基于监管改造大数据的服刑人员矫正质量置信度评估技术;研发服刑人员循证矫正策略智能推荐技术与装备;研究基于社会状态监测大数据的假释、暂予监外执行人员的矫正智能决策技术;研究基于要素关联的假释、暂予监外执行人员改造质量评估与风险预警技术等。

## (二)智能化再犯风险评估司法应用的运行特征

智能化再犯风险评估的运行过程通常分为三个阶段:第一阶段为数据喂饲。大数据是智能化再犯风险评估的“原材料”,原材料质量越高,数量越大,则评估越是精准。目前美国最常用的第四代COMPAS评估工具,不仅为暴力、再犯、脱离监管、社区矫正失败提供总体的风险评估,

还为矫正罪犯提供犯罪必要因素的概况,包括罪犯的犯罪历史、需求评估、犯罪态度、社会环境以及其他因素,如社交失败、犯罪机会、犯罪个人因素和社会支持等方面的信息。<sup>[9]</sup>这些信息需要通过输入与罪犯有关的大数据来获得,而有关罪犯的大数据有的来自司法大数据,有的来自生活大数据,有的来自对他们的实时监控、监测,还有的则来自后台的测验,如让嫌疑人、被告或罪犯回答相关问题。第二阶段为数据吞吐。计算机通过被动学习或自主学习,寻找影响再犯的关联因素,并计算关联因素影响再犯的权重。如COMPAS评估的风险因素超过了20个大类,包括认知习惯、犯罪同伙、犯罪参与情况、犯罪机会、犯罪个性特征、犯罪态度、暴力倾向、家庭情况、经济状况、违法历史、暴力历史、就业状况、住所情况、社会适应性、社会环境、社交状况、瘾品依赖、教育状况、测谎等。而根据社会学习、亚文化、控制/抑制、反社会人格、犯罪机会等不同理论,依据现有数据的统计,区分性别、犯罪类型等不同群体,“智能化”地确定再犯因素和影响权重。第三阶段为结论输出。智能化再犯风险评估能够根据不同的评估需要,智能化地给出评估结论。与二代、三代再犯风险评估主要根据制定的量表,对相关的项目进行评分,最后汇总分析不同,COMPAS能结合量表及其他多元数据,与选定的对照组进行对比。

从智能化再犯风险评估的运行过程不难看出,智能化评估尽管也是由人根据评估的需要所设计,有时还需要通过人工输入数据,但其整个运行的过程是由系统或机器自动操控的,由此,其输出的结论虽然较非智能化再犯风险评估的准确性和可靠性更高,但无法如人工操作,知晓评估的过程和依据,并对评估的结论作出解释。

## (三)智能化再犯风险评估司法应用的实践特征

自19世纪以来,刑法(罚)经过近代和现代改革,经由威慑时代到等价时代,再到矫正时代,目前正处于兼顾法律报应与犯罪预防的折中时

代。<sup>[10]</sup> 刑事实证学派倡导的刑罚个别化,虽然没有能够主导当代刑事立法,但已经成为各国刑事立法的有机组成部分。再犯风险评估作为落实刑罚个别化的技术保障,它的智能化发展正从三个方面改变着再犯防治特别是刑罚个别化的实践。

### 1. 应用范围由狭窄到广泛

依据人身危险性决定罪犯的刑罚适用和执行带有很强的主观性,因此如何降低评估的或然性,提高评估的准确性和可靠性,也就成了再犯风险评估司法应用必须解决的难题。依赖个人临床经验和直觉的经验性评估,由于评估准确性低、可靠性差和专业人员缺乏等,因此主要应用于有限的缓刑、假释等个案。随着精算评估特别是兼有静态和动态因素的精算评估的准确性和可靠性的提高,再犯风险评估开始扩大应用于预防刑的适用,但只有少数具有专门知识的专家才有资格出具这种专家评估意见,因此,实际应用案件仍极其有限。服务于多元目的,并由多个评估主体参与评估的结构化的系统评估,应用领域虽然由适用缓刑、假释和量刑扩大到了行刑,但由于涉及到更为复杂的影响再犯因素,即使由计算机辅助操作,需要耗费大量人工的调查和数据输入仍限制了其大范围的应用。而再犯风险评估的智能化,一方面,通过大数据挖掘能够获得较抽样调查(小数据)研究更为丰富的影响再犯的因素和更精准的权重,因此能够极大地提高评估的准确性和可靠性并促进刑事司法的公正性;<sup>[11]</sup> 另一方面,伴随着信息技术的发展,特别是5G时代的开启,具有高速度、泛在网、低功耗、低时延特征的5G平台,将进一步提升使用的便捷性,为此,其在刑事司法的应用必将进一步扩大。

美国是智能化再犯风险评估应用最早、目前实际应用最广泛的国家。自美国缓刑之父约翰·奥古斯塔斯于1841年在审判前保释第一位被告人以来,服务于刑罚个别化的风险评估技术就与之形影相随。伴随着再犯风险评估的智能化,美国再犯风险评估已经扩大到假释指南的编制与适用、判决指南的编制与适用、判决释放决定、社

区监管级别分类和监狱监管级别分类。<sup>[12]</sup> 在过去20多年中,美国法院已经利用COMPAS软件辅助处理超过100万宗案件。<sup>[13]</sup> 而在刑罚执行中,伴随着循证实践的兴起,并在美国循证刑事政策的引导下,<sup>[14]</sup> 从风险确定、动态预警到依据证据进行精准干预,智能化再犯风险评估在循证矫正中起到了不可或缺的作用。我国再犯风险评估基于监狱行刑个别化需要而产生,自20世纪90年代司法部监狱管理局与中国心理学会法律心理专业委员会合作研制出《中国罪犯心理测试量表个性分测验》以来,服务于监狱刑罚执行的评估工具,从江西赣州监狱和中央司法警官学院教授联合开发的“狱内危险性评估系统”到浙江十里丰监狱和浙江省司法警官学院联合开发的“罪犯自评系统”,从司法预防犯罪研究所与上海青浦监狱等合作研究的“再犯风险评估智能平台”到中国政法大学犯罪心理学研究中心与北京心法科技有限公司等单位共同开发的“非接触式服刑人员风险动态实时评估系统”,不仅影响因素由静态到动态,而且越来越具有智能化。伴随着宽严相济刑事司法政策实施和进一步的法律化,从强制措施适用到量刑,再到刑罚执行,乃至刑满释放以后帮扶,再犯风险评估在我国再犯防治中的应用正呈现出迅速扩大趋势。<sup>[15]</sup>

### 2. 应用方式由公开到隐秘

相对于已然犯罪事实,作为未然犯罪事实的再犯风险评估的结论带有较强的主观性,但这种主观判断的结论也是建立在客观依据之上的,并不是评估者个人的臆断。不论是经验性的评估、量表式的精算评估,还是结构化的系统评估,从对评估对象信息的收集到量表的选用,从评估考量的因素到每个因素的影响权重,从评估报告的呈现到接受他人质疑等,专业人员的评估不仅过程是公开的、透明的,而且结果也是可以接受质疑的或能够解释的。智能化再犯风险评估尽管也会有人工输入的数据和通过以往评估工具所获得的数据,但系统或机器如何挖掘数据并最终形成评估结论,完全属于“黑箱操作”。即使是了

解算法步骤的专家,也无法解释它依据什么得出最后的评估结论。而这种算法黑箱还导致应用者无法通过解释来回复他人对评估结果的质疑。

### 3. 应用效力由可以到应当

随着评估技术的迭代演进,再犯风险评估在刑事司法中的应用不仅范围逐步扩大,而且法律效力也由“可以参考”逐步发展为“应当参考”。由专业人员凭借经验、直觉,或同时借助量表测量所形成的非智能化再犯风险评估结论,在法律上具有间接意见证据的性质,起初只是法官裁量刑罚可以参考的因素之一。而评估结论的证明力则需要通过庭审质证程序,由陪审团来决定。但自1987年美国联邦判决委员会颁布第一部判决指南以后,再犯风险评估的应用在采纳判决指南的美国各州日益规范。根据2011年美国判决委员会发布的判决指南,联邦法院的判决表由犯罪的严重程度和犯罪前科记录纵横两个维度构成,而判决表横轴中所列的六类程度不同的风险级别正是再犯风险评估技术应用的结果。<sup>[16]</sup>随着智能化再犯风险评估的应用,法官在量刑前,被告都必须接受再犯风险评估。法官尽管还需要结合其他因素综合考量被告的风险等级,但系统或机器输出的结论客观上会对法官的最终判断产生更大的影响。

在我国,1980年施行的第一部《刑法》第六十七条有关适用缓刑“确实不致再危害社会的”规定,第七十三条有关适用假释“不致再危害社会”的规定,以及第一部《刑事诉讼法》第四十条有关“采取取保候审、监社居住等方法,尚不足以防止发生社会危险性”逮捕条件的规定,虽然都涉及到人身危险性,但并没有规定如何调查和评估,并由此导致司法实践中一直以“主观恶性”来衡量人身危险性。但随着宽严相济刑事司法政策和认罪认罚从宽制度的落实,有关社会调查和社会危险性评估开始出现在刑事立法中。<sup>[17]</sup>不仅如此,2012年修订的《刑事诉讼法》第二百六十八条还只是规定“公安机关、人民检察院、人民法院办理未成年人刑事案件,根据情况可以对

未成年犯罪嫌疑人、被告人的成长经历、犯罪原因、监护教育等情况进行调查”,但在2015年发布的《反恐怖主义法》第三十条则规定:“对恐怖活动罪犯和极端主义罪犯被判处徒刑以上刑罚的,监狱、看守所应当在刑满释放前根据其犯罪性质、情节和社会危害程度,服刑期间的表现,释放后对所居住社区的影响等进行社会危险性评估。”从可以调查评估到应当调查评估的转变,表明再犯风险评估已经成为刚性需求,而具有精准、便捷特点的智能再犯风险评估的研发,正是为了满足这种公正性要求极高的刚性风险评估需求。

### 三、智能化再犯风险评估应用的法律风险

再犯风险评估的智能化可以提高评估的准确性和可靠性,并促进刑罚个别化,提高再犯防治的实际效果。但正如乌尔里希·贝克指出:“任何科技技术的运用在给人们带来便利之时,也会带来风险与挑战,那些生产风险或从中得益的人迟早会受到风险的报应。风险在它的扩散中展示了一种社会性的‘飞去来器效应’,即使是富裕和有权势的人也不会逃脱它们。”<sup>[18]</sup>伴随着再犯风险评估智能化研发和应用,我国的智能化再犯风险评估不可避免地会面临以下法律风险:

#### (一) 算法歧视引发的刑事司法不公

公正是司法的灵魂!智能化再犯风险评估由于既避免了依赖评估者经验进行评估的主观性,又克服了依据抽样数据进行分析的局限性,因此,能够极大地提高评估结论的准确性和可靠性,并进一步促进司法的公正。然而,再犯风险评估的智能化及其应用是一个不断发展和规范的过程,不论是作为算法“饲料”的数据,还是作为算法本身,都可能产生偏见和歧视,并引起新的司法不公,这在弱人工智能时代更为突出。

#### 1. 数据问题引起的歧视

智能化与非智能化再犯风险评估最根本的区别,就在于智能化再犯风险评估具有深度学习的能力。而这种深度学习的能力由于依赖于数据

为基础的算法运行,因此,代表性数据数量的大小和错误数据的情况直接影响到算法运行结果,并在一定程度上决定着智能化再犯风险评估的准确性和可靠性。

我国再犯风险评估虽然发展快,但起步较晚,其智能化开发和应用:一方面会面临代表性数据规模小的问题。经过多年不懈努力,我国公安、检察院和法院虽然已经形成了庞大的数据库,<sup>[19]</sup>但与再犯风险评估相关的代表性数据不仅数量不大而且要素不全。我国的刑罚虽然是兼顾报应与功利的“一体刑”,但司法实践中影响预防刑的主要是犯罪目的、动机、犯罪后的态度等因素,并不是经综合评估的再犯风险高低。<sup>[20]</sup>为此,在我国报应性司法模式下,不论是侦查、起诉还是审判都不会深入分析犯罪嫌疑人、被告人或罪犯的犯罪原因,并将影响犯罪的相关信息载入各种案卷材料或统计表格之中。即使是以改造人为宗旨的监狱和社区矫正部门,也由于重管理轻改造、矫正等原因,大多没有形成针对犯罪、矫正需求进行精准施策的科学矫正机制。根据司法部《社区矫正管理系统数据采集结构规范》(以下简称《规范》),目前社区矫正部门采集的社区矫正数据尽管多达 37 大类 510 项,但再犯风险评估所需要的数据项却严重缺失。如精神心理状态类因素在各国再犯风险评估中被排在犯罪情况之后,属于第二位影响再犯,且涉及具体项最多的一类因素,<sup>[21]</sup>但在《规范》中却只涉及“对社会的心态”“对生活是否有信心”“心理健康状况”三项因素。影响再犯的数据是通过人为设计收集的,这些收集到的数据本身没有问题,但它只能局部反映再犯情况。在代表性数据不够大的情况下,算法在根据现有缺少代表性的数据属性,推断目前尚不知的属性过程中,可能通过冗余编码,从没有明确收集的数据中推断出受保护的敏感数据。如种族信息是被保护的敏感信息,一般算法不能依据种族确定一个人再犯风险的高低,但在深度学习的算法中,大数据仍能够通过其他维度线索,如姓名、地

区等,来标识特定的人种,从而形成某种种族的人更容易犯罪的歧视。由此,在代表性数量不够大的情况下,甚至会出现智能化的评估结论还不如非智能化评估结论准确或可靠的问题。

另一方面还会面临错误数据的问题。大数据是对过去社会运行的记录,本质上是社会的一种镜像。再犯风险评估的大数据主要由评估对象犯罪前、刑事司法中和刑满释放后的数据所构成,不仅社会的偏见和歧视,可能是他们初犯的重要影响因素,而且刑事司法及其犯罪的标签效应还会进一步加深社会对他们的偏见和歧视,并成为他们再犯的重要原因。当现实社会生活中的偏见和歧视反映到大数据中,或收集数据方法欠科学而产生的问题数据,通过输入成为再犯风险评估算法的“饲料”时,这种错误的信息可能成为“训练数据”。起初,算法可能只是简单地复制这些历史数据中的偏见和歧视,而后,则可能从受歧视的人群样本中收集数据并作出推断,从而形成算法输出的结果的歧视。

## 2. 算法问题引起的歧视

算法虽然是一种技术,但并非绝对“中立”而公正,相反可能因“歧视”而带来不公。“算法是由人类来编写完成的,那么人类的价值观会嵌入到这些软件当中。他们经常也必然在使用数据的过程中带入人类的偏见。”<sup>[22]</sup>算法作为从体量大、种类多、速度快且时效高的低密度数据中获得高价值数据的最重要的技术,具有两个方面的特征:一方面,算法以效率为导向。信息是决策的基础,一个正确的决策有赖于正确、详细和全面的信息。从经验性评估到精算式评估,再到依托大数据的智能化评估,再犯风险评估的历史演进,始终伴随着评估所需信息量的扩大。为了高效地处理纷繁复杂的信息,人类总是采用图式分类的方式简化自己的认知过程,由此,“通过图式对事物进行归纳是人认识外在世界的基本方式”,<sup>[23]</sup>但值得注意的是“类型化”这种认知方式,在方便人们认识世界以改造世界并提升生活质量的同时,也容易造成对某类事物或群体的刻

板印象,并因此成为人类偏见和歧视产生的根源。挖掘大数据过程中的深度学习,与人类的“类型化”认知方式具有相似之处,有三个主要环节:第一是用大量未标注样本逐层提炼,无导向自动形成特征;第二是用一些已标注样本对特征分类,并根据分类结果调整系统参数,以优化系统在区分不同类别信息上的性能;第三是用系统未见识过的样本数据检验系统学习效果。<sup>[24]</sup>也正是这种“类人”思维的算法设计,使得算法在追求效率的过程中,也容易犯人类“类型化”思维所具有的毛病。

另一方面,算法影响相关性判断。预测一个人的再犯可能性是再犯风险评估的目的所在,然而,一个人犯罪的原因是多方面的,且是多层次的。与利用犯罪发生变化的因果关系预防再犯不同,由于算法追求的是相关性,而不是因果性,因此智能化再犯风险评估是利用相关性分析,即要通过挖掘大数据找到影响再犯的预测因子,并确定其影响再犯的权重来预测再犯可能性的。而事物之间联系的普遍性,决定了与再犯相关的因素几乎涉及罪犯、社区矫正对象和刑满释放人员自身和所处环境的方方面面。某种因素与再犯高度相关,并不能解释为这种因素就是引起再犯的原因,更不能归纳出具有该种特征的群体更容易犯罪。但现实生活中,人们时常将这种相关关系当作因果关系,并形成偏见和歧视。

### 3. 算法霸权引起的不公

偏见和歧视是任何一个社会都客观存在的现象,但由算法引起的偏见和歧视经技术“中立”的包装、技术霸权的助力,常以极其隐秘的方式存在,因此难以准确识别,并进行有效监督。伴随着刑事普通程序正当化改革,不论是犯罪事实的认定还是法律的适用,我国刑事诉讼越来越重视中立、公开、平等、交涉等正当程序要求。智能化再犯风险评估的应用,虽然为刑罚个别化提供了依据,但也容易因算法霸权而引起不公。首先,“算法黑箱”引起的监督难。智能化再犯风险评估从数据输入到结果输出,中间必须经历一

个人们难以洞悉的数据吞吐过程。在这一过程中,算法是如何自动化地挖掘数据并形成个别化评估结论的,人们既无法观察得到,更难以说得清楚。由此,导致这种不透明的算法成为监督的盲区,各种算法歧视不仅难以及时发现,而且可能被更深地掩盖。其次,“算法保护”引起的监督难。算法是人工智能的核心技术,也是在大力发展人工智能过程中法律保护的重点。任何有关再犯风险评估的新算法,都可以获得知识产权保护。而这种为促进技术发展而实施的法律保护,也会进一步增加算法监督的难度。第三,“相关关系”引起的解释难。算法以效率为导向,并以相关分析确定影响再犯因素和影响权重,难以做到绝对精确。再犯风险高低的评估结果虽然不是直接作为刑罚个别化的依据,但仍实质影响嫌疑人、被告人的诉讼权利与刑罚,影响罪犯、社区矫正对象处遇和刑释人员帮教。评估结果如果既无法公开检验,又难以解释说明,那么很容易引起质疑,并影响刑事司法结果的可接受性。被称为人工智能刑事司法应用第一案的美国威斯康星州诉艾瑞克·卢米斯案[Statev. Loomis, 88IN. Wisconsin. 2d749(wis. 2016)],正是由此引起的。卢米斯因偷窃被枪击者抛弃的汽车而被警察误当作枪击者予以逮捕,并受到与驾车枪击有关的五项刑事指控,卢米斯承认了其中两项较轻的指控。卢米斯回答了再犯风险评估工具COMPAS所呈现的一系列问题,并被COMPAS评定给出再犯风险等级为“高风险”。法官参考了COMPAS评估结论及其他众多因素,并将COMPAS评估结论作为卢米斯量刑前的调查报告(PSI)的一部分,最终判处其六年监禁和五年监外执行。卢米斯以法庭严重依赖COMPAS系统进行判案侵犯其宪法权利为由提起上诉,2016年7月美国威斯康星州最高法院驳回了卢米斯的请求,2017年6月美国最高法院也拒绝了提审该案的请求。<sup>[25]</sup>智能化再犯风险评估在刑事司法领域的应用,在美国通过判例在事实上获得了承认,但也引发了人们对智能化再犯风险评估

应用风险的深入思考。<sup>[26]</sup>

## (二)智能评估引发的侵犯公民隐私权利

隐私权是人格权的一种,涉及到信息隐私、空间隐私和自决隐私等广泛领域。各国法律将隐私权作为基本人权加以保护,目的是要保持人的心情舒畅和维护人格尊严。保护罪犯隐私权,一方面是严格执法的要求。罪犯因自己的犯罪而被剥夺或限制人身自由,他们在监禁期间的信件、邮包等要受到检查,电话、会见等也要受到监听,但这只是基于监管安全和矫正的需要,对他们的隐私权加以必要的限制,并不意味着他们没有隐私权。另一方面是改造罪犯的前提。隐私权是否受到尊重直接影响到罪犯的自我评价,而一个人的自我评价又关乎其自尊与自信。罪犯因犯罪标签的影响多表现为自卑,并因此普遍存在内在改造动力不足的问题。为此,要把罪犯改造为守法公民,首先得帮其克服自卑,重拾自信。尊重罪犯的人格尊严,依法保护其隐私,又是其重拾自信的前提条件。

具有精准推荐功能的智能化再犯风险评估系统,离不开评估对象的大数据。数据越大、质量越高,则经挖掘获得的价值就越大。然而,大数据的收集、存储、传输等整个过程都会面临数据泄密的风险。首先,数据收集时的泄密风险。非智能化再犯风险评估收集的信息有限,且主要由个别专业人员完成,即使涉及评估对象的隐私,但泄密的风险也较低。但大数据的收集具有范围广、来源多、内容杂等特点,不仅涉及隐私信息多,而且收集环节和经手人员多,因此,泄密的风险也随之增高。其次,存储和传输环节的泄密风险。非智能化再犯风险评估的信息主要以纸质形式,即使现在普遍采用数据存储,但计算机通常不联网,或只与监狱系统内网相连,所以较少存在受到黑客攻击而产生数据泄密的问题。但智能化再犯风险评估由于依赖于大数据分析,需要与外网连结,因此在存储和传输涉及隐私信息数据时,更容易受到黑客攻击而出现泄密。第三,评估结果应用过程中的泄密风险。数据是信

息的载体,信息则是数据的内涵。对评估对象大数据的挖掘,能够获得一个人不愿公开的个人隐私信息,如人格缺陷、心理问题、行为习惯,甚至再次犯罪的风险高低等。这些隐私信息虽然对保证刑事诉讼顺利进行、监管安全、精准预防等具有特殊的价值,但在评估结果应用过程中也会面临泄露隐私的风险。

智能化再犯风险评估造成罪犯、社区矫正对象或刑满释放人员的隐私保护难题,一方面,会加剧他们的自卑心理,弱化甚至使其丧失羞耻心,并最终失去改过自新的内在心理动力。另一方面,隐私的泄露还会加剧社会对他们的偏见和歧视,进一步增加其融入社会,重新开始新的社会生活的难度。更重要的是这种偏见和歧视,还可能进一步演变为对某一类群体的偏见和歧视。

## (三)精准预防面临的人权保护问题

根据科技部2018年发布的《通知》中有关再犯预防研发目标,我国正在研发的是具有精确评估、动态预警和精准推荐功能的智能化再犯风险评估系统。再犯风险评估虽经历了多代演进,但非智能化再犯风险评估由于评估结果的精准性不够高,无法对再犯犯因和预防需求进行自动化检测,又难以对罪犯进行循证矫正,因此主要用于量刑参考或用于罪犯分类。伴随着智能化再犯评估系统的研发,建立在精准评估基础上的动态预警和精准预防将成为现实。而具有精准特点的预警、预防也容易引起新的人权保护问题。

世上没有天生的罪犯!犯罪是一个人的生物因素、环境因素和心理因素相互作用的结果,而每个人犯罪的原因又各不相同,因此,犯罪预防的实际效果取决于预防措施针对性。根据犯罪心理内化与外化的原理,影响初次犯罪的因素可按其所起的作用区分为三类:导致犯罪心理形成的致因、推动犯罪行为的心因和促使犯罪心理外化的诱因。而依据这些犯因是否可以改变和改变的难易程度,又可以区分为不变因素、难变因素和易变因素。与影响初次犯罪的因素不尽相同,罪犯、社区矫正对象或刑满释放人员再

次犯罪的影响因素,不仅包括致因、心因和诱因等初犯犯因因素,还包括针对犯因所采取的应因因素。应因也即应对犯罪的措施或因素,包括对正在服刑罪犯、社区矫正对象的管控、矫治和帮助,对刑满释放人员的帮扶和有限管控。确定再犯的致因、心因与诱因是为了更好地采取应因措施,而应因则既可以起到管控不变或可以改变的致因、心因与诱因的作用,又可以通过矫治、帮助或帮扶起到弱化或消除可以改变的影响再犯的致因、心因与诱因的作用。应因虽然是为积极干预犯因而采取的措施,但也构成了影响罪犯、社区矫正对象或刑满释放人员的因素。缺乏科学的应因措施,不仅难以发挥其管控、减弱或消除犯因的积极作用,而且可以成为再犯的犯因。

犯因包括生物因素、环境因素和心理因素,应因则涉及管控措施、矫正措施和帮助或帮扶措施。智能化再犯评估系统的精准推荐功能,其实就是根据循证原理,依据已有证据在犯因与应因之间寻找最佳匹配关系。<sup>[27]</sup>然而,应因措施既可以针对环境也可以针对人,既有赋权性的帮助或帮扶措施,也带有强制性的矫正措施,更有涉及限制甚至剥夺人身自由的管控措施。不论是对物还是对人,一旦涉及适用强制性的措施,如果没有法律明确授权,都会引起侵犯人权的质疑。对再犯预防主要涉及三类对象,即正在监狱服刑的罪犯、正在社区接受矫正的矫正对象和刑满释放人员。现代监禁刑以剥夺人身自由为内容,对正在监狱服刑的罪犯采取矫正和管控措施,只要遵守监狱法的规定,不存在侵犯人权的问题。社区矫正由于涉及管制这种非监禁刑罚的执行,缓刑、假释这类附条件不执行刑罚的监督措施的执行,以及暂予监外执行这一监禁刑罚的非监禁执行方式,因此,根据预警而对再犯危险性高的社区矫正对象,是否可以采取剥夺人身自由的管控措施,不仅缺乏法律明确规定,而且存在理论上的争议。<sup>[28]</sup>而对重新获得自由的刑满释放人员来说,如果智能化评估系统预测其存在再犯的高风险,那么采取限制或剥夺人身自

由的管控措施,更是于法无据,容易引发侵犯人权的质疑。更重要的是,引起再犯的犯因中还可能涉及到对家长、老师等自由公民的干预,此时任何带有强制性的干预都可能会面临法律问题。1994年一起儿童受到性侵害的案件,催生了美国严厉的反性侵害法律“梅根法”,<sup>[29]</sup>并引发了人们对防卫社会与保障人权的激烈争论。2016年浙江省慈溪市检察院牵头法院、公安、司法等机关出台了《性侵害未成年人犯罪人员信息公开实施办法》,这一被称为慈溪“梅根法”的出台,更是引起了中国学者的质疑与担忧。<sup>[30]</sup>

根据我国现行刑事法律规定,不论是适用剥夺或限制人身自由的强制措施,还是适用剥夺或限制人身自由的自由刑,都必须以存在阻碍刑事诉讼顺利进行的行为,或有充分和确实的证据证明实施了犯罪行为为前提,单纯依据一个人的再犯风险高,不能对一个人采取限制或剥夺人身自由的任何措施。而对具有高风险再犯可能性的特定对象进行精准干预,虽然既有利于防卫社会,也利于防止其个人再犯,但依据未然犯罪事实决定一个自由公民的人身自由的剥夺与限制,容易面临“无罪也罚”的不公质疑。与针对特定对象的信息公开不同,根据预警而采取的强制性应因措施,不仅涉及的是不特定的自由公民的人身自由权利,而且涉及的是对一个人最为重要的人身自由的剥夺或限制,因此,如何在防卫社会与保障人权之间权衡与选择,是智能化评估带给刑事司法研究的新课题。

#### 四、智能化再犯风险评估应用的风险防范

再犯风险评估智能化最核心的技术是算法,而“算法的实质是逻辑,以此形成人们基于使用必须遵守的规则。这些规则就像法律,只是法律还需要司法机关的维护来确认辩解”。<sup>[31]</sup>如同法有“善”与“恶”之分,算法也有“先进”与“落后”、“适合”或“不适合”之别,要防范再犯风险智能化带来的风险,需要从理论、技术与规范三个方面加以防范。

### (一)理论模型的优化

第一代到第四代再犯风险评估虽然经历了由经验到精算、静态到动态、个别到综合等发展,但都是根据“刺激—反应”(“S-R”)行为模式所确立的评估模型,即通过检测评估对象对典型刺激所作反应,并比对群体常模以推测评估对象的再犯风险。这种再犯风险评估虽然也能够获得犯因及其犯因性需求,但这种犯因及其犯因性需求不仅不够全面,容易导致重要犯因及其犯因性需求的遗漏,而且无法对犯因及其犯因性需求进行动态监测,并根据变化进行及时、精准干预。

再犯风险评估的智能化是为了实现再犯风险的动态预警、改造或帮扶质量的评估、改造或帮扶方案的精准推荐,若仍以“S-R”行为模式构建智能化再犯风险评估理论模型,将面临两方面的问题:一方面,以再犯群体的标准确定个体再犯风险高低所面临的评估准确性问题。再犯原因既有共性也有个性,以共性原因代替个性原因不仅会影响评估的准确性,也会影响预防的个别化效果。另一方面,以群体或个体过去的再犯因素确定评估对象现在的风险,也存在着难以适应不断变化的环境、个体情况而导致的不当干预问题。而根据“S-O-R”原理,再犯风险评估既要关注外在环境中的不良刺激,又要关注不良环境因素刺激下的个体不良心理的形成,还要重视已经形成的个体不良心理如何在不良环境刺激下外化为犯罪行为。而犯罪行为的发生又会进一步强化犯罪心理,并增加个体对不良环境因素选择的倾向性。为此,我们可以将影响初犯即初次犯罪的因素,区分为致因、诱因和心因。致因与诱因都是外部不良环境因素,但它们在犯罪发生中发挥着不同的作用。犯罪心理因素是外部不良环境因素与主体遗传、已经形成的不良心理因素等相互作用形成的。犯罪心理的内化与外化是一个可以不断循环的过程,犯罪心理的内化促进或直接推动犯罪行为的发生,而犯罪行为的发生又会进一步强化犯罪心理。<sup>[32]</sup>

构建包括致因、心因、诱因,特别是应因在内

的影响再犯的全要素理论模型,不仅可以更好地适应再犯风险评估智能化后不断拓展影响再犯因子的需要,而且可以在管控、矫正、帮扶等应因因素与犯因变化及其犯因性需求的满足之间建立互动模型,以构建更为科学的改造质量评估和再犯风险预警模型。再犯风险评估的目的决定理论模型的选择,也决定采用什么样的算法。很显然,基于“S-O-R”原理建立的再犯风险评估模型所要求的算法难度更高。

### (二)智能算法的监督

再犯风险评估的智能化已经成为一种不可阻挡的历史趋势,但为让我国再犯风险评估智能化研发和应用少走弯路,有必要从一开始就建立算法监督机制。首先,算法验证。算法的合目的性和先进性虽然一定程度上决定着人工智能再犯风险评估的准确性,但评估结论的准确性由于直接影响罪犯的处遇,因此算法的准确性必须经过验证。对此,有美国学者提出:通过交互式诊断的方式分析人工智能模块的记录情况,可忠实重现特定决策结果做出的计算过程与该过程的执行情况,并辅助确定何种输入特征导致了该特定结果。<sup>[33]</sup>深度学习算法由于无法确定归纳偏置即偏见而存在算法上的黑箱现象,对此贝叶斯统计算法可以通过计算归纳偏置确定有用的算法是否正确。其次,算法开源。智能化再犯风险评估即使准确性高,但如果只能呈现风险高低的结论,却无法对结论进行有说服力的解释,那么不仅检察官、法官或刑罚执行人员不愿用、不敢用,而且当事人乃至公众也会提出这样做是否公正的质疑。为此,一方面,可以对所使用的人工智能算法的关键决策步骤进行开源,并公开算法运行过程中的相关数据。另一方面,对影响再犯的相关因素进行功能性分类,即进一步区分犯因、应因,在此基础上通过构建模拟指标算法的方法实现对算法的检验,并通过功能性分类方式呈现影响因素,以增强评估结论的可解释性。第三,算法核查。算法既可以是人工设定的,也可以是机器自主设定的,但都是为人服务的,因此

必须严格受人的核查。正如凯西所指出的,想要“规管算法,驯服算法”,就要让“算法指出可疑之处,由人类去完成最后的核查”,“它们(算法)的运作必须是透明的:我们必须知道它们接受哪些数据输入,产生什么结果,而且它们必须接受稽查。”<sup>[34]</sup>为此,2017年美国计算机协会为解释人工智能技术的算法歧视问题,发布了关于算法歧视的七项原则,同年还出台了“算法问责法案”。而探索可解释性的算法模型则是另一种有益的探索,由英国学者 Cynthia Rudin 所设计的具有可解释性的算法模型,对累犯预测的准确性与 COMPAS 等黑箱算法已经不相上下。

### (三)研发与运用的规范

与以往再犯风险评测工具的计算机化不同,再犯风险评估智能化集产、学、研、用于一体,具有以下明显特征:首先,需要不同学科的知识融合。再犯风险评估智能化涉及犯罪学、心理学、法学、计算机、人工智能等众多学科,在学科分化和专业分工越来越细的今天,再犯风险评估的智能化离不开多学科知识的融合和多方面专家与学者的通力合作。其次,需要研发机构与应用部门的战略合作。如果说非智能化的再犯风险评估研发与应用是可以截然分开的,检察院、法院、监狱等部门花钱买产品使用即可,但再犯风险评估的智能化研发由于涉及到只有刑事司法部门才能掌握的数据,因此,没有研究机构与应用部门的战略合作是无法成功的。第三,需要人文与技术的互相协调。乌尔里希·贝克曾说过:“没有社会理性的科学理性是空洞的,但没有科学理性的社会理性是盲目的。”<sup>[35]</sup>现代技术的异化根源于技术理性与社会理性的断裂,而避免或控制智能化再犯风险评估的风险,则需要协调人文与技术的关系。为此,不论是智能化再犯风险技术的研发还是运用,既需要人文精神的滋养,也需要人文道德、法律的规范。

智能化再犯风险评估研发的以上特点,决定了我国的再犯风险评估的智能化研发和运用需要从以下四方面加以规范:第一,规范研发合作,

打造协同关系。再犯风险评估的智能化研发,不论是过去较多使用的“知识图谱”算法,还是目前主流的深度学习和强化学习,抑或未来融入“知识图谱”的深度学习和强化学习算法,如果没有再犯风险评估、罪犯改造等领域专家的深度参与,几乎是不可能完成的任务。然而,刑罚执行的封闭性以及相关信息的保密性与技术的知识产权保护,既会阻碍技术专家对技术研发所需行刑信息的掌握,也会影响同行专家及其智能化再犯风险评估使用者对技术的监督和接受程度。为此,需要通过规范双方行为,有效解决信息保密和知识产权保护难题,形成产、学、研、用战略合作。战略合作既不同于一个简单的局部间的合作协议,也有别于追求各自利益最大化的合作,是一种建立在共同利益基础上的深度合作,具有长期性、整体性。刑事司法部门只有与科研单位建立战略合作,才能从源头消除形成研发所需协同关系的障碍。

第二,规范罪犯改造,形成互动关系。数据作为人工智能的“原材料”,其规模及其质量直接影响到最后的产品质量。<sup>[36]</sup>针对我国再犯风险、改造质量评估所需要素缺失,有必要根据循证理念,规范罪犯改造过程,即不论是一个罪犯从入监或入矫到出监或解矫的整个改造,还是针对罪犯某一具体问题的矫正,都需要通过评估首先确定其犯因和犯因性需求,然后施以针对性的管控、矫正、帮扶等措施,最后还需要进行效果评估。唯有如此,才能形成“好数据”并通过不断积累形成大数据。相对于经验性的传统矫正,科学的循证矫正不仅专业性强,而且资源保障要求高,这对尚未完全解决经费保障的监狱和社区矫正来说,确实面临巨大的挑战。但任何科学技术应用需要有科学的态度,否则会反受其害。为此,要充分发挥产、学、研、用一体研发模式的优势,在研发出满足实践需要的产品的同时,通过研发促进我国循证矫正实践的发展。

第三,规范结果运用,强化交涉关系。平等交涉作为刑事司法正当化的重要特征,又是增强

评估结果可接受性的关键。为此,评估结论不论是作为被告量刑的参考,还是作为罪犯处遇的依据,都需要建立规范的程序。裁判者或执法者不仅不能仅仅依据评估结论决定被告的预防刑或罪犯的处遇,而且不能不经各方质证或充分表达意见,将其直接作为量刑或决定的参考或依据。随着量子科学的发展,科学的确定性观念已经被打破。正如普里戈金所认为,在人类处于一种新的理论的转折的时候,科学就不再是确定的了,而是对现实世界复杂性和不确定性的反映。<sup>[37]</sup>影响再犯风险的因素及其相互关系错综复杂,加上影响再犯因素还存在着个体差异,而人们只能通过外在典型行为推测一个人的再犯可能性,因此再犯风险评估始终具有或然性或不确定性。伴随着人工智能的发展,智能化再犯风险评估虽然有望将评估结论的或然性或不确定性降低到无关紧要的地步,但目前的人工智能不仅对常识的判断具有不确定性,而且作为人工智能重要研究领域的不确定推理,如基于概率的不确定性推理网络——贝叶斯网络,神经网络算法等本身就具有不确定性。正是这种不确定性决定了再犯风险评估结论的司法适用应当确立可反驳原则,并依据这一原则建立由各方参与的平等交涉程序。

第四,规范信息安全,建立互信关系。再犯风险评估的智能化由于有利于刑罚执行部门对罪犯精准施策,对行刑公正、提高预防再犯效果具有重要意义。为此,它本质上是一项利于罪犯重新融入社会的技术,而不是对罪犯实行控制的手段。但与自由公民的生活大数据不同,一方面,再犯风险评估涉及的罪犯数据的搜集由于具有强制性,因此一些数据的可靠性会受到评估对象是否配合的影响。以心理测试为例,若没有罪犯的自愿配合,以通用的自陈式量表所获得的数据完全不可用;另一方面,再犯风险评估的结论具有标签效应。在服刑期间的高风险的评估结论,意味着罪犯可以获得更加精准的管控、矫正与帮扶,因而有利于他们的改造,但这种评估结论也容易成为一种负面标签,并随着他们的出狱

或解矫而成为重新融入社会的障碍。为此,有必要明确再犯风险评估智能化的目的,规范再犯风险评估数据搜集、存储和结果运用,在评估者与被评估对象之间形成信任关系,才能获得更可靠的数据,并使评估结论产生更积极的效果。

## 五、结 语

“法律行业过去 200 年的变化,远不及未来 20 年的变化”,<sup>[38]</sup>伴随着再犯风险评估的智能化发展,我国刑事司法及其再犯预防将迎来新的机遇和挑战。我们既要抓住机遇全力发展我国的再犯风险评估理论与技术,实现后来者居上,又要正确处理人文与技术的关系,坚持再犯风险评估的目的是帮助罪犯和刑满释放人员更好融入社会,而不是主要为了控制。唯有如此,我们才能更好地化解再犯风险评估的智能化可能带来的风险。

### 注释:

[1]再犯通常恶性更大,造成的危害也更为严重。不仅如此,对于一个国家的犯罪率来说,再犯贡献率要远高于初犯。美国学者 Beck & Bemand(1989)通过档案分析发现,5%的犯罪人要承担45%案件的发生责任。Farrington(1996)的研究也显示,在所有案件中,有将近一半是由6%的犯罪人完成的。为此,预防再犯具有不同于预防初犯的特殊价值。更值得关注的是,再犯预防由于对象确定、条件可控等,无论是对罪犯风险评估,还是对其风险的控制更具有实际可操作性。

[2]人身危险性包括初犯的可能性和再犯的可能性。根据风险理论,风险可以分为纯粹风险和投机风险。人身危险性评估强调只有危害机会而无获利机会的纯粹风险,而再犯风险评估则重视再犯既有危害可能也有获利机会的投机风险。与依据人身危险评估决定假释和缓刑不同,再犯风险评估的重点在于通过刑罚个别化对再犯风险进行有效管理,以追求降低甚至消除再犯风险的目标。

[3]何川、马皓:《罪犯危险性评估研究综述》,《河北北方学院学报(社会科学版)》2014年第2期。

[4]Sarah L. Desmarais & Jay P. Singh, “Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States”, April 20, 2019.

[5]文姬:《再犯危险性评估在英美法系的应用》,《刑事法评论》(第31卷),北京:北京大学出版社,2012年,第569页。

[6]参见陶阳明:《人工智能两大趋势:自主能力和生物特

性》,《电子技术与软件工程》2018年第17期。

[7] Richard Berk, *Machine Learning Risk Assessments in Criminal Justice Settings*, Springer Nature Switzerland, 2019, p. 170.

[8] National Institute of Justice, Types of Technology and Tools, <https://www.nij.gov/topics/technology/types/Pages/welcome.aspx>, April 20, 2019.

[9] T. Brennan, W. Dieterich & B. Ehret, "Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System", *Criminal Justice and Behavior an International Journal*, 2009 (36), pp. 21 - 40.

[10] 邱兴隆:《刑罚理性评论——刑罚的正当性反思》,北京:中国政法大学出版社,1999年,“引论”。

[11] 参见杨廷超:《机器人法——构建人类未来新秩序》,北京:法律出版社,2019年,第108-109页。

[12][16] 参见曾贇:《风险评估技术在刑事司法领域中的应用》,《刑事法评论》(第32卷),北京:北京大学出版社,2013年,第492-510、493页。

[13]《软件能测保释罪犯是否再犯罪? 科学家提出质疑》,新浪网, <http://tech.sina.com.cn/roll/2018-01-19/doc-ifyqtxcx0087047.shtml>。

[14] 参见 David L. Myers and Jason D. Spraitz, "Evidence - Based Crime Policy: Enhancing Effectiveness Through Research and Evaluation", <http://cjp.sagepub.com/content/22/2/135>。

[15]《刑法》第三十七条“人民法院可以根据犯罪情况和预防再犯罪的需要,禁止其自刑罚执行完毕之日或者假释之日起从事相关职业”的规定、第八十一条假释的规定,《反恐怖主义法》第三十条“安置教育”的规定,以及《刑事诉讼法》第六十七条取保候审条件的规定等,都从预防再犯的角度,提出了对嫌疑人、被告人或刑满释放人员进行再犯风险的评估的要求。

[17] 少年司法中特有的社会调查制度强调要对涉罪少年的成长经历、环境等影响少年犯罪或影响他们融入社会的原因进行调查分析,并根据其犯因和需求选择更适合其回归健康成长之路的处分方式。社会调查制度最重要的还是风险评估,社会调查只是获取评估所需要的信息。

[18][35] 参见[德]乌尔里希·贝克:《风险社会学》,何博闻译,北京:译林出版社,2004年,第39、30页。

[19] 全国3519个法院和9279个人民法庭通过专网实现互联互通,各级法院正以每5分钟一次的频率向大数据管理和服务平台自动汇聚新收集的各类案件数据,智慧法院平台也已经汇集了1亿多件案例数据和2900多万份裁判文书。参见倪寿明:《充分挖掘司法大数据的超凡价值》,《人民司法(应用)》2017年第19期。

[20] 犯罪动机目的及犯罪后的态度尽管是影响再犯的重要因素,但决定再犯风险高低的还有其他更多的因素,如初犯年龄、改过程度、接受矫正情况等。也正因为如此,美国等西方国家量刑都要进行再犯风险评估,而文章所提到的COMPAS评估工具正是智能化再犯风险评估工具。

[21] 参见冯卫国、王超:《中外社区矫正风险评估因素结构

差异研究》,《法学杂志》2014年第7期。

[22] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, MA: Harvard University Press, 2015, p. 38.

[23] 虞嘉琦:《少数族、外部者、边缘人:多元文化社会的偏见与歧视及其教育应对》,《教育学报》2019年第2期。

[24] 参见胡侯立、魏维、胡蒙娜:《深度学习算法的原理及应用》,《信息技术》2015年第2期。

[25] Christopher Markou, "Why using AI to sentence criminals is a dangerous idea?", Available at The Conversation, <http://theconversation.com/why-using-ai-to-sentence-criminals-is-a-dangerous-idea-77734>, April 20, 2019.

[26] 参见朱体正:《人工智能辅助刑事裁判的不确定性风险及其防范——美国威斯康星州诉卢米斯案的启示》,《浙江社会科学》2018年第6期。

[27] 参见夏苏平、狄小华:《循证矫正中国化研究》,南京:江苏人民出版社,2013年,第28-34页。

[28] 参见梁宇栋:《社区矫正的性质之探析》,《社科纵横》2019年第3期。

[29] 根据“梅根法”,性侵犯人释放出狱后,不仅要到所居住的社区登记报备个人相关信息,而且警方会将个人行踪、住址、驾照号码、体貌特征等信息向社区公布,甚至在互联网上予以广而告之,以示警醒。

[30] 参见尤敏:《慈溪版“梅根法”的制度风险——兼评慈溪市〈性侵害未成年人犯罪人员信息公开实施办法〉》,《青少年犯罪问题》2017年第2期。

[31]《算法霸权与公益诉讼》,正义网, [http://tech.jcrb.com/zbdt/201811/t20181109\\_2405098.shtml](http://tech.jcrb.com/zbdt/201811/t20181109_2405098.shtml)。

[32] 参见罗大华:《犯罪心理学》,北京:中国政法大学出版社,2003年。

[33] Samy Bengio, et al., "Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks", <https://arxiv.org/pdf/1506.03099v3.pdf>, April 20, 2019.

[34][美]凯西·欧尼尔:《大数据的傲慢与偏见:一个圈内数学家对演算法霸权的警告与揭发》,许瑞宋译,台北:大写出版社,2017年,第112页。

[36] 参见仇筠茜、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》,《郑州大学学报(哲学社会科学版)》2018年第5期。

[37][比]伊利亚·普里戈金:《确定性的终结——时间、混沌与新自然法则》,湛敏译,上海:上海科技教育出版社,2009年,第5页。

[38][英]理查德·萨斯坎德:《法律人的明天会怎样?——法律职业的未来》,何广越译,北京:北京大学出版社,2015年,第13页。

[责任编辑:邹秋淑]