

人工智能自主犯罪刑事立法的必要性分析

● 李 军



[摘要] 由于人工智能自主犯罪不受其设计者或使用者的控制和管理,也就引发了是否有必要对其进行刑事立法的思考和争议。主张刑事立法的观点认为,人工智能具有高度的智能性,甚至能够超越设计者的编程而产生自我意识,进而独立实施犯罪行为,因此,有必要赋予其刑事主体资格,并通过刑事立法的手段对其进行规制。但反对的观点则认为人工智能始终是服务于人类社会的科学技术的产物,其不具备法律意义上的刑事主体资格,因此,没有必要对人工智能自主犯罪进行刑事立法。若盲目立法则可能违背刑法谦抑性原则,同时,也不利于科学技术的进步和发展。

[关键词] 人工智能;刑事立法;犯罪;刑事主体

问题的提出

自 20 世纪 50 年代的达特茅斯会议以来,人类就开始对人工智能进行探索,经过数十年的研究,人类对人工智能技术的运用已较为广泛。但任何事物都是一把“双刃剑”,人工智能的发展给社会生活带来科技与便捷的同时,也引发了一些负面问题。自 20 世纪 80 年代末至今,涉人工智能且造成严重损害结果的案件已不在少数,如 2016 年我国河北省某自动驾驶车祸致死案以及 2018 年美国某自动驾驶车辆撞死人案,诸多此类案件逐渐出现在公众的视野中,预示着当人类社会步入了一个新的时代的同时,也将面临着一些新的挑战。

人工智能犯罪大体上可分为两种类型:第一,自然人利用人工智能实施犯罪。此种犯罪类型较为常见,人工智能在其中主要承担着犯罪工具的角色,而真正的犯罪主体始终是自然人。第二,人工智能自主实施犯罪行为。此种犯罪是指人工智能系统在脱离自然人的指令和控制下,自动实施某种行为而造成严重的损害后果,自动驾驶汽车交通肇事正是这一类型的体现。人工智能自主实施犯罪行为具有以下三点特征:一是行为主体具有不特定性。由于人工智能通常需要有一定的载体,其形态多样,因此,人工智能自主犯罪中的行为主体可以是任何具有人工智能属性或者承载着人工智能系统的实体,如自动驾驶的汽车等,相较于传统犯罪中的自然人主体而言更具有随机性。二是所侵害的法益具有单一性。传统犯罪中,其犯罪行为所侵犯的法益包括各种不同的类型,如盗窃罪、诈骗罪等财产性犯罪,这些犯罪侵害的是他人财产性权利;伪造货币罪、非法吸收公众存款

罪等金融犯罪侵害的是国家金融管理秩序等,而常见的人工智能自主犯罪中所侵害的法益,通常为具体自然人的健康权或者生命权。三是主观方面难以判断。人工智能是以编程和算法为基础而实施的行为,所以人工智能实施犯罪时,并不存在像自然人实施犯罪时所具备的故意或过失的主观意识。虽然在 1989 年人工智能机器人落败于前苏联国际象棋冠军后,释放出强大的电流致该国际象棋选手死亡一案中,该人工智能机器人看似是由于比赛的落败,而故意使用强大电流攻击对方,进而直观地理解为该机器人存在故意杀人的主观心理。但是否具有故意心理还是由于比赛落败,而导致机器人系统故障而发生意外,无从考证。

面对频发的人工智能犯罪问题以及对未来强人工智能时代的展望,如何规制人工智能自主犯罪成了人工智能法治的主要问题之一,同时,也是我国刑法所面临的重大挑战,对人工智能自主犯罪刑事立法问题的争论也由此应运而生。

刑事立法必要性之争

(一)观点一:支持对人工智能自主犯罪进行刑事立法

传统的立法模式总伴随着滞后性的特点,即当出现原生效刑法已无法规制社会中某种严重违法行为时,才会对已发生的问题启动立法程序。在刑法理论与实践,预防性立法成了学界的热门话题。人工智能的出现以及未来的发展,必然会给社会带来种种未知的挑战,此时更应该灵活地发挥刑事立法的作用。也有学者表示,立法活动应是在特定目的的指引下进行的贯穿国家意志与人民意志诉求的活动,人工智能刑事立法也不例外。

由于人工智能犯罪与自然人犯罪存在着较大的区别，因此，人工智能刑事立法需要重新审视以下两个方面：一方面，人工智能是否可以成为刑事主体。站在积极立法的观点上，先要肯定的是人工智能应当具备刑事主体资格进而成为犯罪主体。有学者强调，面对无人驾驶汽车在自动驾驶模式下造成的交通事故、智能机器人在失控状态下“失手杀死”被害人等一系列问题，也理应引起人工智能机器人是否具有刑事责任主体资格地位的思考。首先，人工智能具有独立的“自我意识”。人工智能的“自我意识”主要表现为，人工智能可以不在人类的指令和控制下自主完成某些行为，人工智能的研发、生产单位和使用单位均无法预料和控制具有自主意识的人工智能的“意识”和“行为”。其次，人工智能可对不同情况进行预判进而实施行为。人工智能对实际情况的预判，是其“智能性”的突出表现。在传统犯罪中，行为人成为犯罪主体，也应当要具备包括精神状态在内的刑事责任能力，同理，人工智能是否能成为未来刑事立法中的犯罪主体，其“意识”与“行为”必然成为考量的重要因素。

另一方面，人工智能是否具有可罚性。刑罚是刑法的组成部分，人工智能刑法自然也是如此。对此有学者认为，人工智能可以像人类一样直接承担刑事责任，甚至原本只能适用于人类的传统刑罚，如死刑、自由刑、缓刑、罚金等手段，经过适当调整同样适用于人工智能。再者，未来的强人工智能时代里，强人工智能产品将具有辨认能力和控制能力，刑法上应当承认其具有独立的人格和刑事能力，并且可以对其使用删除数据、修改程序、永久销毁等刑罚手段。

人工智能时代，智能机器人完全可以按照自主意识和意志实施严重危害社会的犯罪行为。在这种情况下，智能机器人完全具有独立的辨认能力和控制能力，因此，可以成为刑事责任主体，并且需要承担刑事责任。对人工智能刑事主体资格与可罚性的确认，是进行人工智能刑事立法的前提条件和基本要素。在积极立法观下，面对人工智能对社会带来的各种挑战，有必要通过刑事立法的手段对人工智能犯罪进行规制。

（二）观点二：反对人工智能自主犯罪刑事立法

该观点主要从三个角度进行分析：第一，对人工智能刑事主体理论的否定。有学者认为：人工智能机器人是“从属于人类的工具”。也有学者从科学的角度对人工智能刑事主体资格予以否定：人工智能科学决定其非法律主体，首先，人工智能是技能；其次，人工智能是算法而非人类认知神经科学；再次，人工智能不符合自然科学主体的构成要素；最后，人工智能法律主体不是人工智能科学的价值目标。

第二，人工智能并无自由的主体意识可言。构成主体意识有三个基本要素：一是响应，能对外部事件与事物主动响应，能被唤醒或者保持感知。二是自我体现，受到侵害时启动保护程序，始终将资源的获取和扩展置于行为的优先位置。三是学习，将学习作为主体生存和发展、完成智能任务的前提。所有的生物都具有响应能力，人工智能目前对外部有一定的响应能力，但人工智能的响应能力都是人赋予的，无论什么响应、如何响应都是人赋予的，所以人工智能并不能拥有自我意识。心智引导人类去发现和创造，而人工智能没有也不能有心智，其作为非生命实体不能同自然人一样基于出生而取得平等的法律主体地位，也不当然成为犯罪主体。此外，行为人承担刑事责任必须以行为人具有完全的辨认能力和控制能力为基础，而人工智能机器的辨认能力和控制能力是自然人赋予的，实质上其认知的能力仍然是依附于自然人的能力而存在的，是属于一种区别于人类的拟制能力，其“意志”仍不是自由的。

第三，对人工智能实施“刑罚”的做法不可取。刑罚的目的在于追求社会意义上的效果，即预防犯罪继续发生。假设对智能机器人A实施了“刑罚”，并不会影响机器人B或C继续实施犯罪行为，其原因在于人工智能不像人类一样具有社会性以及敬畏道德和法律的特性，机器人D所实施的某一行为也不会影响机器人E或F去实施其他行为。因为无论技术多么先进的人工智能，其行为均来源于人类所设计的编程和算法，而并非基于思想和道德。古有“杀鸡儆猴”“杀一儆百”之说，某种程度上反映了刑罚的作用，但这种作用显然不会出现在人工智能体上。

❶ 解决问题是最终目的：没有必要进行人工智能自主犯罪刑事立法

笔者认为，刑法作为我国法律体系中最严厉的部门法，在刑事立法与适用的过程中应保持谨慎的态度，并且根据我国法律体系、社会经济发展的特点，以及科学技术发展等方面综合考虑，认为没有必要进行人工智能自主犯罪的刑事立法，理由如下。

第一，立法难度大，并且将对整个法律体系造成较大的影响。对人工智能犯罪进行刑事立法，则要赋予人工智能刑事主体资格，此时人工智能法治面临的不仅是刑事立法方面的问题，而是整个法律体系的大变动。众所周知，自然人触犯刑法构成犯罪的前提在于其属于民事法律主体、触犯民事法律关系或者其他行政规章制度以及违背道德规范等先行条件，人工智能犯罪亦是如此。因此，在赋予人工智能刑事主体资格之前，还应当先赋予其民事主体资格，并且还要对人工智能的行为在民事方面进行立法予以规范。

第二，与刑法谦抑性原则相悖。刑法之所以强调谦抑

性，是因为其对社会关系调整相对于民法等其他法律部门，在手段方面具有较强的严厉性和强制性。刑法的滥用必然会对社会和犯罪嫌疑人造成远较其他法律滥用更为严重的后果，必须通过强调谦抑性理念对其进行限制，才能保证刑法的施行保持在适当的范围之内。人们通常理解的人工智能犯罪并非真正意义上的犯罪，只是人工智能体的“动作”和“行为”导致了损害结果的发生。一方面，人工智能的“犯罪”概念较为模糊，是否能与自然人的“犯罪”概念相提并论还有待商榷；另一方面，对人工智能引发损害结果的问题，还可通过适用其他部门法以及行政规章予以解决，并非仅有刑事立法一条路可走。张明楷教授也强调：立法者不要奢望通过立法路径解决一切问题。

第三，过度加重刑事司法压力。赋予人工智能刑事主体资格并对其进行刑事立法，将较大地加重刑事司法工作量和压力，体现为以下两点：其一，刑法作为整个法律体系的底线和后盾，其适用始终保持着谨慎的态度，若将人工智能问题强加于刑法的“管辖范围”内，无疑会导致刑法的适用范围扩大，也必然导致刑事司法压力的加重。其二，人工智能刑事立法将导致刑事程序上增加许多不必要且繁琐的环节，假设人工智能机器人S因涉嫌犯罪而被认定为“犯罪嫌疑机器人”，则根据《刑事诉讼法》的规定，应当对机器人S采取强制措施。由此引发思考，对机器人采取强制措施是否有意义和有必要？笔者认为此种做法并无意义和必要性，因为对人工智能采取刑事强制措施缺乏逻辑性和合理性，只会给司法资源带来不必要的消耗和浪费。

第四，不利于科学技术的发展。人工智能是当今社会最具代表性的科技作品，其凝结了人类智慧与科学的结晶，同时也是时代变革的象征。关于人工智能法治方面，首先，人们应当客观正确地看待人工智能。人工智能被创造出来的目的始终在于服务人类社会的运转，人工智能只是具有高度自动化的工具，尽管在未来强人工智能时代，人工智能也只是接替了人类的工作而非接替人类的生活。其次，值得思考的是，人工智能刑事立法所产生的效果是规范人工智能行为还是抑制人工智能技术的发展？笔者认为，影响

人工智能行为的因素更大程度上在于人工智能的设计研发者、使用者和管理者等个人，若对人工智能自主犯罪的问题盲目进行刑事立法，对科学家而言是一种障碍与牵制，因为他们不知道自己创造的成果未来可能成为“犯罪嫌疑人”。

智能时代已经成为不可逆转的趋势，人工智能也将在各方面代替人类的工作，但智能社会仍应当以自然人为中心，规范人的不法行为才是法治的目的。只有这样才更有利于科学技术的进步与人工智能的发展，也才能更好地维护智能社会的法治秩序。

参考文献

- [1]李振林.人工智能刑事立法图景[J].华南师范大学学报(社会科学版),2018(06):125-133,191.
- [2]刘宪权.对强智能机器人刑事责任主体地位否定说的回应[J].法学评论,2019,37(05):113-121.
- [3]皮勇.人工智能刑事法治的基本问题[J].比较法研究,2018(05):149-166.
- [4]黄云波.论人工智能时代的刑法机能[J].人民法治,2018(11):12-14.
- [5]刘宪权,胡荷佳.论人工智能时代智能机器人的刑事责任能力[J].法学,2018(01):40-47.
- [6]叶良芳,马路瑶.风险社会视阈下人工智能犯罪的刑法应对[J].浙江学刊,2018(06):65-72.
- [7]李爱君.人工智能法律行为论[J].政法论坛,2019,37(03):176-183.
- [8]朴宗根,杨玥.人工智能法律主体资格之刑法评判[J].南宁师范大学学报(哲学社会科学版),2019,40(05):168-173.
- [9]张全印.机遇与挑战:人工智能带来的刑事风险与刑法应对[J].中国刑警学院学报,2018(06):14-20.
- [10]叶亚杰.论刑法谦抑性的价值与整合[J].河北法学,2016,34(12):110-115.

作者简介:

李军(1994—),男,汉族,广西玉林人,硕士,助教,燕京理工学院,研究方向:刑法、刑事诉讼法。