

基于人工智能的军事智能武器 犯罪问题初论

夏 天*

内容摘要：十九大报告中提出坚持走中国特色强军之路，加快军事智能武器发展。可以预见，基于人工智能的军事智能武器，必然会成为未来战场的主角。但是，“机器人战士”会不会也像人一样犯罪？这不仅是理论问题，更是现实问题。可以从军事智能武器的犯罪形态和责任归属上入手，研判未来可能发生的防范此类犯罪的情况。

关键字：军事智能武器；犯罪；人工智能

习近平总书记在党的十九大报告中提出，要发展新型作战力量与保障力量，加快军事智能化发展。军事力量作为一个国家基础性保障力量，是一个国家综合国力的重要组成部分，影响着这个国家的安全与稳定，同时也决定了一个国家在国际社会的地位，而军事智能武器作为未来军事力量竞争中的制高点，能够确保在未来非对称作战中保持优势，发展军事智能武器就显得尤为重要。立足现在，着眼未来战争，将来的战场上势必会有大量军事智能武器投入使用，通俗的讲就是会有更多“无人机”“机器战士”取代人类成为战场的主角。这些未来有着高级人工智能的“机器战士”会不会也像我们人类一样，有犯罪的可能？或者“冲动”？在人类社会第三次产业革命进入“网络世界”“大数据时代”，基于人工智能的军事智能武器的犯罪可能性和追责问题，就成为一个防患于未然的现实课题。

一、基于人工智能的军事智能武器发展趋势

（一）人工智能与军事智能武器

未来军事智能武器的发展，从技术层面上讲是基于人工智能的发展水平。所谓人工智能（Artificial Intelligence, 英文缩写 AI），它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用的一门新的技术科学。人工智能这一概念起源于1956年，以麦卡赛、明斯基、罗切斯特和申农等为首的一批年轻科学家们在一次学会中，共同研究和探讨用机器模拟智能的一系列问题，并首次提出了“人工智能”这一术语。经过60多年的演进，目前，人工智能已经在智能感知（行为感知与语言识别）、智能推理（逻辑推理能力）、智能学习（自主学习升级）、智能行动（当前主要表现为机器人）等方面有着突破性的进展。2016年被称为人工智能元年，人工智能如井喷式的爆发，先有阿尔法狗以4:1的成绩击败世界围棋冠军李世石，人工智能初现峥嵘；后有IBM研发的医疗机器人对于病患的诊断效率比人类快30倍，而且准确率高达99%，人工智能再下一城；甚至IBM研制出世界首个

* 夏天，国防大学政治学院西安校区副教授。本文系2015年度全军军事学研究生自助课题《大数据条件下军队预防犯罪工作研究》（2015JY279）阶段性研究成果。

马晋：《人工智能技术发展的伦理困境研究》，载《硕士学位论文》，吉林大学2016年版。

人工相变神经元、微软人工智能设备的语言理解能力击败人类、德国科学家研发出能够感知疼痛的人工神经系统已经使得人工智能开始逐步由弱人工智能向着强人工智能转变，等等，给人们带来了一个充满传奇与未知的世界。

纵观人类的科技史，往往重大先进的科学技术都在军事领域得到首先应用，人工智能也不例外。军事智能武器的核心是人工智能，代表了未来战争的发展态势。可以说，现代意义上的军事智能武器就是指那些“具有人工智能，可自动寻找、识别、跟踪和摧毁目标的现代化高科技武器，它集指挥高效化、打击准确化、操作自动化、行为智能化于一身”。目前可以预判的是，随着军事智能武器的大量用于实战，必然会减少专业军事人员的人数，降低培养和维护成本，也可以最大限度地避免人的生命危险，甚至可以避免现阶段即不绝于耳的“无人机误炸”事件。当前，无人攻击机误炸是由于远程操作者精神和生理问题导致的人为失误引起，而人工智能不会感到疲倦，也不受感情因素影响，军事智能型武器可以避免人为失误，准确完成任务。

在新时代里随着人工智能技术的不断发展，军事领域的人工智能应运将显得十分需要，因为它能够改变传统的军事斗争形式，是赢得国家安全战略的重要切入点，随着我国民用人工智能的出现和军事智能武器的发展，也使得传统的法律原则和法律规则基础产生动摇，尤其随着智能机器人的出现，以及人格化的人工智能的发展，需要法律对人格化的智能机器人作出身份上的明确规定，这也是明确人工智能责任和能力的前提和基础。

（二）军事智能武器发展历史演进

军事智能武器最早可以追溯到20世纪20年代，当时第一次世界大战正酣，英国的卡德尔和皮切尔两位将军，向英国军事航空学会提出一项建议：研制一种不用人驾驶，而用无线电操纵的小型飞机，使他能够飞到敌方某一目标区上空，将事先装在小飞机上的炸弹投下去。这样一种简易的无人机正是后续无人武器与智能武器的开端。

伴随着第三次科技革命的兴起，网络、计算机与人工智能发展，军用智能武器队伍也伴随着第三次科技革命而迅速壮大。截止到上世纪末，军用智能武器形成了以军用智能机器人、军用无人机、军用无人舰艇为代表的三类智能化武器装备。智能机器人方面则以iRobot公司为代表，它所设计研发的款名为“费奇”机器人，目的是为了清除飞机场上的集束炸弹的子弹，还有1998年设计的“背包”机器人，主要通过搭载一些辅助设备，进行排爆、生化武器检测等辅助性战斗；而军用无人机的起源可以追溯到上世纪40年代末，当时为训练战斗机飞行员和防空导弹操作手而研发的一款代号为Q-2“火蜂”的靶机，这是现代无人机的鼻祖；至于军用无人舰艇，在上世纪末也主要运用于一些辅助性战斗中：搜索水雷与情报侦察。例如瑞典的“双鹰”无人潜艇可以借助传感器对其周围300米的水雷进行安全爆破；被称为NMRS的近期鱼雷侦察系统，是美国海军1999年投入使用的第一代搜索鱼雷用的无人潜艇。

人工相变神经元，是指像人类的神经元一样，以低能耗高速率进行信息的传递，更重要的是它们像生物神经元一样，具有随机性，即产生的信号有随机性，总是略有不同。

在人工智能发展的今天，我国在民用人工智能领域也取得了诸多成就，例如百度公司开发的无人驾驶汽车技术，腾讯公司开发的人工智能围棋程序等先进技术。人工智能技术发展到目前，其中主要的应用领域有专家系统、人工生命、模式识别、机器人学、自动程序设计、人工神经网络、智能决策系统等领域，人工智能既是先进制造业的关键支撑装备，也是改善人类生活方式的重要切入点。《中国制造2025》明确将机器人作为重点发展领域，随着《机器人产业发展规划2016—2020》、《新一代人工智能发展规划》等政策陆续出台，人工智能产业发展与应用将迎来重大的历史机遇。得益于传统制造业跻身世界前列，我国的人工智能发展技术也处于国际先进水平，我国已连续五年成为机器人全球第一大应用市场，我国人工智能相关企业共有709家，仅次于美国。

进入21世纪,军事智能武器开始有了长足发展。首先智能机器人随着技术的改良,开始搭载一些武器装备,摇身一变成为在战场上炙手可热的武器装备,例如1998年设计的“背包”机器人通过拓展武器装备变身为2003年伊拉克战争中的排爆专家;一款型号为“利剑”(SWORDS)的机器人,可以搭载任何重量在300磅以下的武器系统,从M-16自动步枪到0.5英寸口径机枪,甚至40掷弹筒或反坦克火箭发射器都没有问题。“背包”机器人和“利剑”机器人只是活跃在当今战场的新型无人系统的代表。美军刚进入伊拉克时,部队没有装备任何机器人。2004年末,机器人数量增加到150台;2005年末,增至2400台;2006年末,达5000台,并且仍在增长。同样的情况也发生在无人机领域。经过多年的发展,先进的国际社会上军用无人机呈现爆发态势。从单兵操作的可以直接悬停在屋顶上空,并向作战单位传回房前屋后实时动态“大乌鸦”,到需根据指挥官要求,在1500英尺高度来监控整个街区,关注一切可疑动态的“影子”类中型无人机,再到例如“捕食者”这一类在5000-15000英尺上空巡航,并兼具“侦察和火力打击”的任务的大型无人机,最后到可以在6万英尺高度全天候、广领域进行巡逻、监听的可疑自主航行的“全球鹰”。我国的军用无人机事业也在这样的时代背景下有了长足的发展,我国研制的“翼龙”“彩虹”等大中型无人机也在国际军用无人机市场有着一定的份额,先后出口十几个国家。军用舰艇迈入21世纪后也开始有了长足发展。军用无人舰艇的发展主要分为两个方向,一类是水下无人潜艇,它的主要任务就是收集情报,侦察、清除水雷,近海反潜作战。例如被称为NMRS的近期鱼雷侦察系统,进入21世纪后更是开发出LMRS远期鱼雷侦察系统,能够以更快的搜索频率、更长的时间、更为可靠的进行工作;法国研发的“剑鱼”无人潜艇,能够以12节的航速航行400海里,同时可以与全球定位系统联动,将收集到的声纳影像数据通过无线通讯方式传输给母舰。另一类是水上无人船舰,主要是以无人机、无人船舰作为主要武器装备进行海上作战。主要作战样式是以某一母舰为中心,由母舰发射无人潜航器、操纵母舰周围无人船舰,指挥母舰上无人飞机进行作战。虽然当前还没有哪个国家配备这样的水上无人舰队,但是不久的将来,水上无人舰队将会成为海战主力部队。

(三) 军事智能武器发展趋势及特征

当前人工智能主要分为三种发展方向,而依托于人工智能的军事智能武器的未来发展方向也大体分为三类:

一类是学习型军事智能武器。这一类军事智能武器完全自主化,主要是由军事智能武器系统通过在被制造产生后进行学习,不断进行进化,从而形成无需具体指令,只需给予一个大体的任务,由军事智能武器根据自己的经验及所学进行判断,从而完成上级所赋予的任务。这一类的军事智能武器从理论上来讲,随着其经历的丰富及所学知识的增加,它所能面临突发性情况时,将会做出更加理性、高效的处理,但同时这一类的军事智能武器也面临着一种困境就是:由于它所要学习的知识与经历不是提前输入,而是由其制造产生后,经由其自身筛选学习的,所以,我们无法控制这一类的军事智能武器其后天学习所确立的判断原则。

另一类是执行型军事智能武器。这一类的军事智能武器在生产之初,对其内部的系统进行预先的设计,将可能出现的情况进行预先的设想,并将相关的处理方式进行录入,这样就可以避免军事智能武器不按照法律法规规则进行相关行动,可以对军事智能武器最大

彼得·W·辛格:《人工智能走上战场》,载《总参情报部》。

程度的控制。但是这一类的军事智能武器存在着一种弊端：由于现实情况总是复杂多样的，不可能存在完全穷尽的可能或者说不可能存在完全相同的两种情况，因此就按照现今的计算机逻辑结构，一旦出现一种复杂情况，很容易产生一种冲突的处理意见，从而导致系统死机或者循环计算，最终导致贻误战机。

最后一类就是综合型军事智能武器，综合上述两种情况，在对其进行一定基础原则输入后，尽可能多的放开它的学习性，使军事智能武器更好的去适应复杂多变的现实问题，这样一种综合型军事智能武器不仅可以吸收执行型智能武器的原则性，使得人民群众对它的安全性放心，另一方面也可以吸收学习型军事智能武器的适应性，方便军事智能武器在遵守基本的道德规范标准的前提下，可以更灵活得适应现实情况。

在现代化的军事斗争中，先进武器的致命性打击对交战双方来说都是军事斗争的主观追求，尤其是涉及到国家大规模武装冲突、非平衡作战形态中，先进的武器装备能够发挥显著地先锋作用，如前文所述的伊拉克战争、海湾战争中的美国先进科技武装的武器类型，随着科技的不断进步，智能武器更以其独特的优势，备受军事的青睐，军事智能武器的不断发展也是科技进步的现实反映，是维护国防利益和加快军队建设的客观需要。

1. 军事智能武器的优势

(1)信息化、专业化程度高，在可控范围内能够实现精确打击，降低“第一战场”的军事人员伤亡率。未来的军事智能武器必将能够在“第一战场”发挥巨大的作用，能够独立根据程序设计以及自身的模拟分析判断战场形态并作出其自身思维模式下的行为，在这种数字化、信息化的集中感知判断中，智能武器对外界环境和因素的依赖度大大降低，能够真正实现全天候全方位的精确打击和破坏，并且依赖于人工智能基础上的军事智能武器，尤其是未来智能军事机器人的出现，将会大大减少作为自然人的主体的参与程度，而更多是将以人工形态的模拟化主体出现在战争的第一形式中，或者是自然人和机器人的配合，共同协作完成作战任务，这便在更高的价值上保存了有生的人的力量。军事智能武器之所以现在备受各个国家追捧和青睐，智能武器本身的这一特性是最重要也是最具诱惑的原因。

(2)相比自然人主体，军事智能武器的能够最大限度的避免和减少执行任务过程中客观条件的影响。例如利用无人攻击机而导致误炸，大多是由于远程操作者精神和生理问题导致的人为失误引起，人工智能武器不会感到疲倦，也不受感情因素影响，人工智能型武器可以避免人为失误，在程序范围内准确完成任务。相比在战争过后对人的善后处理和对智能武器的处置方法、手段来比较，国家和国防建设更倾向于最大限度的掌控现阶段的智能武器，因为自然人主体的社会需求和权利随着社会的不断进步将会更加扩展，而在国家可预期的现阶段来看，军事智能武器更能有效、妥善的解决战争中的各项作战任务，而不考虑对其人格化后的权利保障问题，这种优势是显而易见的，但是当高级的人工智能应用到军事领域，在执行任务中失控了如何解决，如何对其追责，如何建设预期的法律规则、用预期的法律规则来预防人工智能人格化后法律属性，便是发展军事智能武器急需解决的问题。

2. 军事智能武器的弊端

日本人工智能专家松尾丰认为：“人工智能能征服人类，人工智能制造出人工智能——

这种可能性在现阶段看来并不存在，只不过是凭空臆想而已”。军事智能都是建立在人工智能的基础之上，人工智能发展过程中的弊端和负面因素也将不可避免的传输到军事智能武器上，而不同于民用可科技实验的人工智能，军事智能武器面对的是更加残酷的客观战争，如果在无人格化的限制下，军事斗争就变成了一群冰冷的却又高度机械的废铁在人的支配下进行残酷的杀掠，军事智能武器就像没有责任能力的法律因素存在在现实的法律关系中，而恰恰是这样具有类似人的行为能力甚至在很多方面超过人的行为能力的这些不被法律承认的客观存在实施了法律所规避和惩戒的杀人行为、破坏行为，况且法律所需要的认识能力将会随着智能化的发展，人工智能对人的大脑的认知过程和数据传导模式最大限度的记忆和模仿，从而做出不被法律认可的主体却高度类人化的行为。智能化的武器杀伤力的破坏程度不言而喻，这种人工智能本身存在的缺陷或者人工智能自身高度进化，能够实现自我升级完成缺陷补漏，将会对传统的刑事归责和法律主体界定的基础产生激烈的动摇和挑战，同时在军事智能武器的发展过程中，类人行为的自我升级和自我分析行为，以及行为产生的后果已经超出国际法、战争法的规则，具体军事智能武器的弊端而言，有以下两个主要方面：

(1)人工智能的研发者、操纵者无法保证智能武器自身的绝对安全。从人工智能的研发到实验再到具体营运，任何环节的细微疏漏都有可能产生人工智能本身的缺陷，从而在应运到实践过程中存在一定的风险，这就如同一般的商品，在经济学领域都会产生风险评估，在讨论商品属性的人工智能角度，许多学者可能会反驳笔者前述的人工智能人格化的论述，从商品属性看待人工智能确实无法将其人格化，但是恰恰笔者需要从作为现阶段公认的普通产品这一经济学角度的视角下的属性来揭露作为商品的固有属性和法律评价，因为商品缺陷所产生的刑事责任、民事责任则是将人工智能人格化的经济前提，也是将其能够在法律上人格化的可能性延伸，因为作为最基本也是最被接受的人工智能为商品属性的劳动产品是现有法律定罪的实在客体。正是这样的逻辑关系，所以在人工智能的研发实践过程中，任何自然人主体都无法保证人工智能的绝对安全，这一论断从最新的GOOGLE旗下的人工智能研究机构DeepMind的研究人员和旧金山OpenAI的人工智能机器人的研究中得到了有力证明，为了达到确定性目标，人工智能机器人会用自己的分析过程从而破毁原本的规则，而为了消除这个不利因素，就要在涉及之初让人工智能对目标保持一定的不确定性。这种悖论的平衡，就是最大的危险，因为一旦这样的智能武器应用到战争中，甚至研发人员和设计者都不确定人工智能武器能否做到绝对安全可靠地完成程序设计任务。这种缺陷可能随着科技的发达和人工智能的自主完善得到优先改善，而这一过程便和自然人主体自我完善一样，人工智能人格化至少在理论上可以解决以上的法律困难和法律规避现象。

(2)缺失人的伦理属性和心理属性的人工智能无法理解战争中自然人的心理变化。张明楷教授涉及违法性领域的判断时，将伦理学的诸多观点吸收到刑法领域，并认为伦理学为其功利主义观点提供了哲学依据。在涉及到人工智能领域，同样必须考虑伦理学，因为无论人工智能发展到任何阶段，就人的心理属性它是永远无法拥有的，反而却和我们所认同的伦理道德息息相关，将道德作为研究对象的伦理学，人工智能的出现和发展对却传

松尾丰：《人工智能狂潮——机器人会超越人类吗》，机械工业出版社2016年版，第152页。

该项目是OpenAI和Google的新课题，是通过赛船冠军赛（Coast Runners）来训练人工智能。其规则就是获得分数越多，且最终跨过终点线则获胜。

张明楷：《犯罪论的基本问题》，法律出版社2017年6月版，第4页。

统的人类道德提出了新的挑战,无论杜威 和 海德格尔 的伦理观点,还是现代我国罗国杰教授的伦理观点,其学说都无法对类人行为的人工智能准确定位,当然人类无法要求今天的法律去规定昨天的事情,但是伦理学的最传统的道德确实收到人工智能的挑战,而且伦理学无法将绝对权的个性差异做出合理的阐释,人工智能涉及到哲学、心理学、法律和其他专业领域的智力支持,需要以一种跨学科的复合知识系统指导物理、电子、控制和软件设计,从而构建“善”的人工智能体。与伦理学紧密联系的就是人的心理属性,自然人主体的心理因素越来越受到现代生物学家和法学家的关注,心理预防在犯罪预防中扮演的角色也越来越重要,但是人工智能不涉及心理方面的因素,而缺少心理因素的智能武器,当其思维模式、行为方式完全类人话之后就是不完整的“人”。尤其当智能武器应运于军事斗争中,智能武器无法做到感知心理变化,尽管它能够通过人脑的脑电波认知技术判断分析,但是自然人主体的心理波动智能武器是无法感知的,进一步来讲,正是存在这样属性智能武器无法判断敌人下一秒将会干什么,例如诈降还是真降,这种复杂的心理自然反应也对智能武器自身和背后的“合作主体一人”(设计者、操纵者)提出了一个更大的挑战,人工智能之所以智能,是由于其自身可以在程序的基础上判断、分析,从而做出行为,就军事内部诸多学者提出的智能武器将会是在自然人的主导操纵下进行的活动的观点,是割裂和误解了人工智能本身的特征的观点,因为这种操控属于物理上的操控,就如同对简单的一般的遥控机器人的操纵,并不是心理、思维、观念的操作,而人工智能武器恰恰就是可以自我分析自我判断自我思维的一个特殊的集合体,能够自主的以人类的模式去完成战争任务。

二、军事智能武器犯罪的可能性及归责依据

正是应为军事武器具有如此之多的不确定性,我们才不能排除未来其具有犯罪的可能性。“机器人杀人”不是一个遥不可及的问题。

(一) 军事智能武器犯罪的刑法可能性

2016年,伊拉克政府使用“彩虹-3”无人机空袭极端组织目标的视频在互联网上走红,不可否认,军用智能武器已经开始广泛的运用于各类军事作战任务中,据美国的一项统计显示:从2004年到2015年,美军无人机空袭估计造成三千人的死亡,而这一数字的预定值只有几十人,可见军用智能武器在执行作战任务命令时,往往存在着目标识别错误、无差别攻击等问题,当前由于武器的智能化程度尚低,处于一种弱智能化阶段,所以往往将责任直接归咎于武器使用者、指挥者、智能武器程序设计者、甚至是利用技术优势对智能武器进行违法操作的黑客。但是随着智能化的进一步提高,人工智能由弱人工智能发展到强人工智能阶段,到那个时候,军事智能武器出现误伤、误炸、故意杀伤平民,违反战争原则时,我们应当如何界定呢?是否存在着智能武器主观犯罪的可能性呢?

对于这样的问题,可以先来通过刑法中犯罪构成理论来分析。传统刑法学理论将犯罪

约翰·杜威(John Dewey, 1859年-1952年),美国哲学家、教育家,实用主义的集大成者,建造了实用主义的理论大厦。他的著作很多,涉及科学、艺术、宗教伦理、政治、教育、社会学、历史学和经济学诸方面,使实用主义成为美国特有的文化现象。

海德格尔马丁·海德格尔(Martin Heidegger, 1889年—1976年),德国哲学家,20世纪存在主义哲学的创始人和主要代表之一。

王绍源、任晓明:载《中国社会科学报》2015年,第9期。

刚桂虎:《军用无人机成长三部曲》,载《解放军报》2017年3月31日,第9版。

构成分为四部分：主体、主观方面、客体、客观方面。犯罪构成要件是决定行为构成犯罪所必须的事实特征，同时也是立法者对具体犯罪现象抽象概括后作为认定犯罪的一般标准。对照犯罪构成的四要件，不难看出，军事智能武器在执行军事作战任务过程中，已经满足了所谓的客体与客观方面，即军事智能武器的某些“行为”已经侵犯了相关的社会法益。例如，在无人机执行空袭任务中，对平民的误伤、误杀。因此当前之所以不能将军事智能武器的“行为”定性为犯罪，最主要的原因就是在于军事智能武器不能够满足所谓的主体与主观方面。

刑法学中认为所谓犯罪主体是指实施危害社会的行为、依法应当负刑事责任的自然人与单位。显而易见，想要成为犯罪主体，无论是自然人抑或是法人都必须具有法律上的人格，即具备人类社会成员的法律资格。就犯罪的主观方面来讲，需要行为主体具有“意识能力”，即行为主体能够辨认自己行为的能力。当前作为军事智能武器核心技术的人工智能还处于一种弱人工智能阶段，人工智能还不能够进行自我思考、自我推理和解决问题的能力，总的来说就是没有自主意识。因此军事智能武器的“行为”并不能符合刑法学意义的犯罪行为。但是未来的人工智能发展方向则是一种强人工智能方向，也是一种真正的人工智能，可以利用一些外置设备进行感知、思考、判断、最终做出决定。这样的强人工智能大体分为两类：一类是类人的人工智能，即随着生物学的进一步发展，人类大脑的思维模式被研究出来，人工智能按照仿生学原理模仿人的思维模式与生活习惯；另一类就是非类人的人工智能，人工智能按照自我的推理方式，可以不按照人类的思维模式进行行动。不管是哪一种的强人工智能，都有着很强的自主性，它不仅可以按照程序设定的步骤去执行命令，也可以根据现实情况的不同做出相应的调整，而此时的人工智能技术已经不能单单的理解为一种技术，一种产品，强人工智能更多的拥有着自我的意识，若此时还以技术免责为由，免除了强人工智能为核心的军事智能武器的责任，势必会造成未来军事智能武器对人类社会的破坏，甚至强人工智能无限进化下去，与人类竞争当前的社会环境，而强人工智能又比人类的大脑运算速度快，不难想象，优胜劣汰的自然法则下，人类可能也会成为一种濒危物种。因此，笔者认为未来社会赋予强人工智能一定的“人格”会解决很多的现实问题。2016年的11月25日，沙特阿拉伯在首都利雅得举办的未来投资倡议会议上赋予了一个名叫“索菲亚”的机器人公民权，使得索菲亚成为人类史上第一个具有公民身份的机器人。可以预见，在不久的将来，随着强人工智能的发展，具有强人工智能的军事智能武器也将以具有公民权的机器人身份参与到人类社会中去，届时军事智能武器的犯罪主体与主观方面也将具有法理依据，军事智能武器的犯罪可能性也即将迎刃而解。

（二）军事智能武器犯罪的归责依据

基于以上的法理一般归责原则，在人类社会发展的现阶段，以无人驾驶技术人工智能汽车为例，因为这项技术是当前最成熟的应运领域，这一现象对传统的交通事故犯罪的影响有以下两点：一是“人”已不实质参与驾驶行为，“人”作为驾驶主体其身份已经被虚化，

2016年3月，在最近机器人设计师戴维·汉森（David Hanson）的测试中，与人类极为相似的类人机器人索菲亚（Sophia）自曝愿望，称想去上学，成立家庭。索菲亚看起来就像人类女性，拥有橡胶皮肤，能够使用很多自然的面部表情。索菲亚“大脑”中的计算机算法能够识别面部，并与人进行眼神接触。索菲亚的皮肤使用名为 Frubber 的延展性材料制作，下面有很多电机，让她可以做出微笑等动作。此外，索菲亚还能理解语言和记住与人类的互动，包括面部。随着时间推移，她会变得越来越聪明。

Yaning：《全球首个机器人公民沙特诞生》http://www.baidu.com/sf?pd=realtime_article&openapi=dispName=iphone&from_sf=1&resource_id=4584&word=机器人公民&title=中国日报&keysign

这种虚化使刑事责任主体无从可寻,“无行为则无犯罪”原则失灵。二是基于“允许的风险”而做出风险分配责任体系,在智能环境下失灵,一律追究智能服务提供者的刑事责任,是将技术作为“替罪羊”的倒退,是归责的无奈选择。在智能驾驶时代,交通肇事罪的行为主体由现实真实的自然人主体全部变成潜在的“被害主体”,自然人脱离具体的驾驶行为,导致传统的肇事者以及危险驾驶者消失,自然人不会因为享受一项智能技术而无加害行为的情况下遭受刑法惩罚。人工智能机器的行为动摇了传统的刑法有关交通肇事罪、危险驾驶罪的立法基础。

当前人工智能还处于一种弱人工智能阶段,人工智能还不能够进行自我思考、自我推理和解决问题的能力,总的来说就是没有自主意识。这样直接将责任归责给制造者、管理者,作为一种产品责任,无可厚非,但是未来的人工智能发展方向则是一种强人工智能方向,也是一种真正的人工智能,可以利用一些外置设备进行感知、思考、判断、最终做出决定。这样的强人工智能大体分为两类:一类是类人的人工智能,即随着生物学的进一步发展,人类大脑的思维模式被研究出来,人工智能按照仿生学原理模仿人的思维模式与生活习惯;另一类就是非类人的人工智能,人工智能按照自我的推理方式,可以不按照人类的思维模式进行行动。不管是哪一种的强人工智能,都有着很强的自主性,它不仅可以按照程序设定的步骤去执行命令,也可以根据现实情况的不同做出相应的调整,而此时的人工智能技术已经不能单单的理解为一种技术,一种产品,强人工智能更多的拥有着自我的意识,若此时还以技术免责为由,免除了强人工智能的责任,势必会造成强人工智能对人类社会的破坏,甚至强人工智能无限进化下去,与人类竞争当前的社会环境,而强人工智能又比人类的大脑运算速度快,不难想象,优胜劣汰的自然法则下,人类可能也会成为一种濒危物种。因此,笔者认为未来社会赋予强人工智能一定的“人格”会解决很多的现实问题。

在军事智能武器归责的过程中,笔者将智能武器“人格化”,探讨军事智能武器以公民身份的可能性基础,也是出于军事斗争的特征而考虑,无论是全面战争还是局部武装冲突,军事斗争是国家之间、地方之间的全面较量,就现阶段采取的传统的战争和冲突会剥夺大量的自然人生命,尤其是不平衡的军事斗争形态中,劣势方必将付出惨痛的代价来抗衡先进的科技。随着军事智能武器的发展,智能武器自主化、类人化,一旦将军事智能武器在斗争中应用,其造成的直接危害远远超过自然人实施的行为,将人工会智能武器人格化,赋予其公民身份,也是为了更好地保护自然人主体的生命健康权,维护人类的伦理道德,不仅能够在人的权利最高价值上保护人类的法益,也是对科技的肯定和妥当的发展与应用。同时,人工智能的出现并不能完全代替自然人主体,智能机器人无法成为诸多法律行为的主体,并不是赋予人工智能所有的责任主体认定资格,例如婚姻法规定的结婚主体资格、刑法规定的强奸主体资格、以及诉讼法所规定的证人资格等,现阶段只是在法律上赋予人工智能在某些具体的刑事犯罪主体资格,在特定的情况下,追究军事智能武器的法律责任,类似我国刑法规定的单位犯罪,人工智能犯罪类型是具体和明确的,这也是现阶段有效管控人工智能技术发展,解决法律冲突和法律规避的有效方式。当涉及到具体的军事智能武器犯罪类型的时候,笔者认为必须要依据军人违反职责罪的罪刑来作为参照依据,结合军事人工智能武器的类人化范围和行为模式,将军事智能武器纳入到具体的罪名主体资格中,尽量使之符合法律的条文规定。

三、军事智能武器的犯罪主体界定问题

当人工智能应用到军事领域时,我们可以利用人工智能武器来评估危险程度,以及可以采取的措施,但是否应该发动战争,需要掌握在人类手中。如前所述,军用智能武器已经开始广泛的运用于各类军事作战任务中,据美国的一项统计显示:从2004年到2015年,美军无人机空袭估计造成三千人的死亡,而这一数字的预定值只有几十人,可见军用智能武器在执行作战任务命令时,往往存在着目标识别错误、无差别攻击等问题,当前由于武器的智能化程度尚低,处于一种弱智能化阶段,所以往往将责任直接归咎于武器使用者、指挥者、武器设计者、生产商。但是随着智能化的进一步提高,人工智能由弱人工智能发展到强人工智能阶段,到那个时候,军事智能武器出现误伤、误炸、故意杀伤平民,违反战争原则时,我们应当如何界定呢?类比当前学术界的一些观点,笔者认为可以从以下方面对军事智能武器刑事犯罪的主体问题进行分析。

当前理论界在界定犯罪主体问题上的一种看法就是:犯罪主体只能是自然人或法人,其他一切事物都不可以成为归责的主体,而军事智能武器无论以后是否具有独立的意识,它的行为也只能看做是自然人或者法人行为的延伸,那么军事智能武器在特殊情况下做出了损害他们利益的行为,也可以看做是相关责任人的行为。而一个人的行为若给他人造成损害,原则上是不需要承担赔偿责任的,除非正当的理由让行为人承担责任。这样一种原则被称为归责原则,其目的就是保护行为人的行为自由及适当对被侵害人的权利进行救济。根据《侵权责任法》中的关于侵权责任的归责原则,我们可以将军事智能武器侵权行为分为以下三类:

一是过错责任。军事智能武器在执行命令过程中的某一行为造成了受害人可救济的利益损失,并且它的行为与加害后果之间存在着一定的因果关系,只要军事智能武器适用的使用者对结果无法证明自己没有过错,则需对损害后果承担相应的赔偿责任。即由军事智能武器的当时管理者、使用者承担责任,此时的归责规则中,实际上还是将人工智能作为一种科技产品,对智能武器基于产品责任的刑法替代原则,只是将其背后的利用者作为归责主体。过错责任对现今的人工智能行为的归责采用的传统的也是被公众所接受的归责原则,该归责的最终结果并不是人工智能本身,附有义务的监督主体,但是人工智能武器的行为一旦符合刑法所规定犯罪的构成要件,在形式上和实质上与人的行为一致的话,其归责如果继续采用过错责任就显得十分突兀,因为军事智能武器的利用者是军队,而军事智能武器的用途就是面向战争,在失控的情况下,智能武器本身的行为无法直接追责到具体的使用者也就是军队的身上,尤其是在发生军事斗争的情况下,追责的难度将会变得更大。随着人工智能的发展,传统刑法领域的犯罪,例如医疗事故这一典型的过失犯罪而言,严重不负责任的医疗主体可能会变成智能技术或者智能机器人而不是具体的医务人员,因为当下我们可以看到技术能够实现医疗的效果,也能够某些医疗中代替自然人,这将会使法定医疗主体的专属性迅速下降,责任主体出现模糊化的现象,如果继续以过错原则为归责原则,很明显此种情况下就需要对智能机器人人格化。

二是无过错责任。军事智能武器实施了法律所规定的需要承担责任的(现阶段还没有相关方面的法律,但是不排除以后在相关领域的立法),其造成的损害后果与其行为之

杜艳勇:《人工智能安全及其解决路径》,载《哲学动态》2016年第9期。

间存在着因果关系,并且受害人遭受的损失是可以救济的,那么此时无论军事智能武器的使用、管控者、生产设计者对这一行为是否尽到了管理注意义务,都应当对损害的后果承担相应的法律责任。在现阶段人类所认知的科学和知识构架下,无过错责任可能更加能实际解决现实中的问题,这种归责更加会受到法律实用主义者的青睐,在我国民法中有“无过错责任”,即“行为人没有过错,但法律规定应当承担责任的,应当承担民事责任”,因为这类情况下无需考虑人工智能的特殊性,只要基于结果的发生属于法律的特殊情况下就能够适用法律的特殊规定,但是笔者要揭示的是,无过错责任的承担法律责任特殊情况中,特殊情况的受害主体和行为实施主体都是自然人或实体生物(饲养的动物),而用无过错责任强加于人工智能的人格虚化模式下,人工智能本身就是无生命的,就这一点本身对传统的指导人们行为的私人伦理学提出了质疑和挑战,在佐证这一观点时,从近些年来核污染和核泄漏等核事故中,我们会发现非常有趣的一个现象:各种“反核”“无核”运动有增无减,但是核能的发展却取得长足的进步。道理很简单,因为核本身就是一种资源,可以创造给人类福利,同样人工智能的出现也是一项高科技技术,在成熟的科技和完善的法律规制体系下是能够丰富人类的物质文明和精神文明。因而在人工智能发展的过程中,我们需要的是如何用法律的手段规制它的应用,尤其涉及到军事智能武器的出现和发展,更加需要法律的规制和约束,一旦失去法律的规制,在人工智能已经在某些方面类人化的基础上,智能武器不仅对刑法提出了挑战,就是现实的责任主体中,尤其是要对国家是无法追责的前提下,更多人会成为人工智能的“替罪羊”,而受害者得不到有效的法律救济,无过错责任在此时就显得苍白无力,这种法律归责原则无法承担原本之初的法律期许。

三是公平责任。针对军事智能武器的行为法律并没有明文规定,同时行为实施方与结果承受方对这一行为的产生均没有过错,这时如果不对受害方受损害的利益加以救济则有违公平原则,此时则应当根据具体情况与公平的观念,确定当事双方适当分摊损害后果与责任。就目前来看,因为对军事智能武器的界定并没有一个权威的说法,也没有相关法律法规对其进行界定,所以如果现今有智能武器造成了受害方权益的损失,可以直接参考公平责任原则处理。这也是将人工智能技术中立的结果,不过度苛责技术的危险性,而是通过对人工智能使用者、操控者、设计者、生产者等之间的法律责任的比例划分,但这就需要一个大前提,国家已经在立法上规定上述主体有绝对的责任,出于公平角度,人工智能本身不承担责任,转接到具体的自然人和单位。在公平责任承担的过程中,法律是从国家角度出发、从保护受害者的角度来立法的,这一原则和目的的确是好的,但是在这一关系中,往往忽视了被害人的自身的观点,也就是被害人想要的法律结果,忽视了被害人与国家的互动关系,尤其在智能武器的行为致使损害发生的情况下,被害人要求彻底损毁智能武器,这就对传统的法律提出了更具现实的挑战,而且这种要求会随着人工智能的发展副作用的出现尤其是当智能武器危及到个人生命时,这样的要求和追诉结果就显得十分有必要也有意义,虽然我们的法律还无法做到承认这一点。

四、军事智能武器犯罪的追责

在前面的讨论中,笔者将军事智能武器犯罪类型归责方式类比单位犯罪的形态,在此问题的论述中,人工智能和单位犯罪两者之间的相似性在于法律同样承认人工智能独立的

欧阳恩钱:《核灾害预防制度研究》,中国社会科学出版社2016年版.08

人格,而在具体的追责方式和犯罪类型两者并不能够并列,例如在单位犯罪中一般实行双罚制,对单位判处罚金,对单位负责人等判处刑罚,而当人工智能的机器人或者武器犯罪进行刑罚处罚时,显然不能以判处罚金的形式来执刑。而且不同于单位犯罪的法人主体资格,人工智能的“人格化”身份是否能够同自然人主体共同故意构成犯罪即共同犯罪问题也是值得分析的,是给人工智能追责归责的前提。

(一) 共同犯罪

将未来的强人工智能赋予人格,即军事智能武器可以直接作为犯罪主体承担责任。根据这样的一种认识,我们可以将军事智能武器犯罪分为以下几类进行探讨:一是单独犯罪。既然军事智能武器在法律层面已经可以作为独立的“人”进行存在,那么军事智能武器作为具有自我独立意识的存在,只要其实施了国家法律法规所禁止的行为,并且满足犯罪相关的构成要件,法律后果当然由该军事智能武器所承担,而具体的惩罚措施再此不论,后文进行叙述。二是共同犯罪。我国《刑法》第25条第一款规定,共同犯罪是指二人以上共同故意犯罪。基于军事智能武器已经具有着独立的人格,在法律认可的责任主体资格成立的犯罪类型中,尤其涉及到国家安全、军事领域的犯罪行为时,基于前文所述的单位犯罪的法人资格一样,智能武器有其合格的法律资格,能够成为具体犯罪的实施者和责任承担着,在军事斗争中,智能武器的军事行为也就是在其独立意识支配下的具体行为,在军事智能武器程序的理解范围内进行的分析整合,从而在自身思维模式下的类人化行为应当属于故意的范畴,即军事智能武器执行具体操作时候应当认为智能武器意识到自身的具体行为的性质或者是模糊的认识到行为性质,而这一点也会随着军事智能武器的不断发展和技术的不断完善,人工智能更加类人化,正如智能人索菲亚在同人类对话时声称要“毁灭人类”一样,显然索菲亚理解“毁灭”的含义和内涵,而作为军事斗争的智能武器将会比普通的人工智能更加需要精密化、类人化,因而在军事领域的智能武器将会和自然人更加的一致和相似,符合法律归责所需的意识能力和精神能力。因此不能再简单在犯罪行为发生时将军事智能武器作为一种方式、手段,将军事智能武器的危害性规避在法律惩罚之外,而应当对军事智能武器作为共同犯罪主体也给予相应的法律惩罚。三是直接正犯。许多学者主张虽然军事智能武器具有一定的思维能力,但是归根到底还是听从于上级指挥者或者直接操作者,因此虽然客观上是军事智能武器实施了这么一个违法行为,但是这一行为的产生是因为指挥者或者直接操作者强制支配了军事智能武器,从而支配犯罪构成要件的实现,此时,军事智能武器智能是一种直接犯罪,而指挥者或者直接的操作者则是间接正犯。笔者在前文论述了将军事人工智能“人格化”的原因和依据,尤其是涉及军事斗争领域的必要性,加之人工智能的出现,对传统刑法的挑战,此种归责方式和追责手段无法解决诸多智能武器带来现实问题,耗时最终还需要法律做出规定。

(二) 具体追责

1. 对军事智能武器的追责。在法律认可某些方面人工智能具有“人格化”自然人身份,能够独立承担其相应的刑事责任,对军事智能武器如何追责和刻以刑伐是妥善处理法律关系的客观需要,由于军事智能武器客观特征以及和人本质的差别,很多原本能适用于自然人主体的刑罚无法在军事智能武器本身操作,也不能像刑法对于单位犯罪的处罚方式一般处以罚金,对军事智能武器的处罚有效措施可以通过以下方式进行追责:第一,彻底报废军事智能武器,当军事智能武器做出刑法所规定的犯罪行为时,可以通过报废该人工智

能武器的方式实现追责和惩罚,这样可以直接对已经类人化的智能机器做出最根本的管制,彻底的消灭实施行为的人工智能再次犯罪的可能性,在成本方面也可以提高人工智能的使用者、操纵者、设计者的消耗,促进更加规范化、科学化的智能机器的出现以及应用。第二,在人工智能武器本身安装自毁系统,这个系统要独立于人工智能本身,与主系统同时运行,并且由专门的国家安全局负责控制,在紧急情况下可以实现远程化、精确化自毁,这一紧急情况便包括军事智能武器犯罪时,可以通过自毁系统实现人工智能在犯罪实施时自毁,从而预防军事人工智能武器产生更大的破坏行为,威胁自然人主体的生命以及其他合法的权益;第三,由于研发人工智能技术需要花费和消耗巨大的财力物力人力,因此人格化的军事智能武器背后也是人类思维模式的部分反映,在发生犯罪现象时,要采用绝对无过错责任原则,对人工智能武器进行追责,在涉及军事领域的犯罪时,国家要设置军事智能机器管理机构,将军事智能武器主动交送国家统一的人工智能管理机构处置,在自毁和报废之外,还可以通过终断人工智能程序、封存等形式对人工智能武器进行入库处理,切实对人工智能武器本身做好严格的处置和管理,将责任追责到具体的智能武器上,避免和杜绝同类化现象的发生。

2. 对其他相关人员的追责。人工智能技术的出现归根结底还是自然人本身创造和发明的,人工智能机器不同于自然界存在的生物实体一样,具有生命的属性,就人类现阶段所研发的军事人工智能武器也达不到完全和人的行为模式一般进行社会活动,更多的是应用都军事斗争中,但是军事智能武器的失控率远远超过设计者的想象,而且破坏性极大,在对人工智能本身刑事追责和处罚后,必须要对相关人员进行追责,可以比照在前文分析过的单位犯罪处罚方式,对军事智能武器的设计者、操纵者、控制者、命令下达者予以追责和处罚,因为军事智能武器侵害的客体一般都是人的生命健康权或者公共安全,因而涉及到诸多受害者的核心法律权力,在追责相关人员的刑法责任时,也要采用严格的无过错主义原则进行归责,相关人员对直接受害者都要承担连带责任,待赔偿损失和刑法追责后,内部在进行比例原则承担相应责任,最大化的保护受害者的合法权益。人工智能技术的出现引领和改变人类生活方式,人工智能不断模糊着物理世界和个人界限,延生出复杂的伦理、法律和安全问题,法律在当前的人工智能治理上存在许多局限性,需要建立一套完善的人工智能伦理原则和法律规则,指导我国的人工智能技术更加科学合理的发展。

王起全:《安全评价》,化学工业出版社2015年版,第11页。

齐昆鹏:《2017人工智能技术、伦理与法律研讨会》,载《科学与社会》2017.06