

人形机器人运行中的刑事政策初探

白佳雪

华东交通大学人文社会科学学院, 江西 南昌

收稿日期: 2025年9月22日; 录用日期: 2025年10月14日; 发布日期: 2025年10月24日

摘要

人工智能技术的发展势不可挡, 其带来的风险亦不可忽视。对此, 我国学者对人工智能运行中的刑事责任问题的探讨渐次展开。但是, 对人形机器人运行中的刑事政策问题, 目前学界鲜有涉及。鉴于当前社会发展和理念的创新, 有必要采取广义的刑事政策观并在此基础上对人形机器人运行中的刑事政策进行初步探讨。具体而言, 在内容上仍须以宽严相济的刑事政策作为基本遵循, 在弱人工智能时代加大对利用人形机器人犯罪人员的处罚力度, 兼顾防治犯罪和技术发展; 在强人工智能时代既要将人形机器人纳入刑罚处罚范围, 也要对人形机器人自身权利的保障予以重视。在对其内容的实现路径上, 一方面要坚持社会安全与技术发展并行, 重视技术监测与预防并且加强培训教育与宣传教育, 及时回应人形机器人运行中所带来的刑事风险; 另一方面还需要进一步兼顾刑法的刑事政策化与刑事政策的刑法化两个命题, 重视社会的协同治理功效, 为预防犯罪与社会的和谐稳定发展提供助力。

关键词

人形机器人, 刑事政策, 刑事主体地位

A Preliminary Study on Criminal Policy in the Operation of Humanoid Robots

Jiaxue Bai

School of Humanities and Social Sciences, East China Jiaotong University, Nanchang Jiangxi

Received: September 22, 2025; accepted: October 14, 2025; published: October 24, 2025

Abstract

The development of artificial intelligence technology is unstoppable, and the risks it brings cannot be ignored. In this regard, our country scholars have gradually discussed the issue of criminal liability in the operation of artificial intelligence. However, the criminal policy issues in the operation of humanoid robots are rarely covered in the academic community. In view of the current social

development and the continuous innovation of concepts, it is necessary to adopt a broad view of criminal policy and make a preliminary discussion on the criminal policy in the operation of humanoid robots on this basis. Specifically, in terms of content, it is still necessary to follow the criminal policy of combining leniency and severity, increase the punishment of criminals who use humanoid robots in the era of weak artificial intelligence, and take into account crime prevention and technological development; In the era of strong artificial intelligence, humanoid robots should not only be included in the scope of criminal punishment, but also pay attention to the protection of humanoid robots' own rights. On the one hand, it is necessary to adhere to the parallel social security and technological development, attach importance to technology monitoring and prevention, strengthen training, education and publicity and education, and respond to the criminal risks brought about by the operation of humanoid robots in a timely manner. On the other hand, it is also necessary to further take into account the two propositions of criminal policy of criminal law and criminalization of criminal policy, pay attention to the effect of collaborative governance of society, and provide assistance for the harmonious and stable development of crime prevention and society.

Keywords

Humanoid Robots, Criminal Policy, Criminal Subject Status

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

变化是永恒的，万事万物都处于无穷无尽地变化之中，人类社会亦是如此。从农业社会到工业社会，从工业社会到信息社会，再从信息社会到如今人工智能广泛应用的智能社会，人类社会发展的步伐越来越快。盖将自其变者而观之，则天地曾不能以一瞬，人工智能的出现与发展是社会进步的结果与要求，无人驾驶汽车的研发、医疗机器人的应用、人脸识别、智能物流与家居等，充分展现出人工智能在当下及未来对于各个领域的价值与作用。然而，科技发展与社会变革使当今社会风险的规模、性质和程度都远超以往任何时代，整个世界都进入了“风险社会”[1]。人工智能为人类社会带来各种便利的光环之下，亦隐藏着一定的风险与隐患。面对人工智能技术特别是人形机器人的运用所带来的法律风险，刑事法律作为社会治理的手段之一，理应做好应对风险、化解风险的准备，而不应该在危险来临时束手无策。对此，应提前预想到在人工智能时代我们可能面临的“内忧外患”([2], p. 15)，做好充分准备，以便于在保护技术发展的同时保障社会的稳定与安全。法学界已逐渐对人工智能运行中的刑事责任问题展开了一些探讨。但是，对人形机器人运行中的刑事政策问题，目前学界鲜有涉及，因此，笔者试图对其刑事政策的界定与特征、政策基本内容及其实现等问题进行初步探讨，以期抛砖引玉。

2. 人形机器人运行中的刑事政策的界定与特征

“法律也好，政策也罢，都是为了满足社会需求的治理手段，而在满足社会需求方面政策具有法律所不具备的自身优势。”[3]法律被要求具有稳定性，刑法作为最严厉的惩罚手段更是要求其具备安定性，因此在面对人工智能技术风驰电掣的发展速度所带来的潜在风险时，法律显得心有余而力不足。此时，政策相较于法律而言能较为及时地应对这些风险，在刑事领域来讲，刑事政策较于刑法更加灵活，能相对及时地满足社会需求，起到预防犯罪的作用。那么在面对人形机器人运行过程中可能直接或者间接产生的严重危害社会的犯罪行为，而刑法却没有相关明确规定时，刑事政策就应率先出击。但是，对于人

形机器人运行中的刑事政策应当如何理解,其与对人类的刑事政策或社会政策相比又有何不同之处?下文将以此出发,对人形机器人运行中的刑事政策的界定与特征进行探讨。

(一) 人形机器人运行中的刑事政策的界定

政策指的是在一定的历史时期,基于社会政治经济等形势和问题所作的一种政治决策和对策[4]。针对不同领域,政策又可分为社会政策、经济政策、法律政策等,法律政策在刑事领域则体现为刑事政策。那么究竟何谓刑事政策?刑事政策这一术语最早于19世纪的德国出现,对其理解大致可分为广义与狭义两种。狭义上的刑事政策是指国家通过刑法及其所规定的刑罚和类似措施,高效组织反犯罪斗争的法律政策,此时刑事政策实际上就是刑事法的制定与适用的法律政策。而广义上的刑事政策,则涵盖了社会整体用以组织对犯罪反应的各种手段方法之总和。在这一角度,刑事政策不再局限于刑事实体法的范畴或者刑罚的范畴,而是强调社会整体对犯罪的反应[5]。随着社会的进步与理念的发展,面对各种新型社会问题的出现,将刑事政策仅仅限定在刑事法的制定与适用的范畴无法有效解决当前问题,而应当从广义的角度来对刑事政策进行界定,即将刑事政策理解为国家为了预防和压制犯罪而为的一切手段或方法,除直接以防制犯罪为目的的刑罚的各制度,间接与预防犯罪有关的各种社会政策,例如居住政策、教育政策、劳动政策(失业政策)及其他公共保护政策等均属于刑事政策的范畴[6]。

对于人形机器人运行中的刑事政策,本文试图从广义的角度对其进行界定。这是因为人形机器人本身就处身于“风险社会”之中,如果要应对其带来的风险,便需要在事前予以预防,防止出现更大范围与程度上的危害。如果从狭义的刑事政策角度对其进行界定,尽管刑法可以将一些预备行为正犯化等,但为了保证法的安定性,其仍然无法及时应对突发的重大情况,如此一来势必会局限于事后的规制与惩罚。因此,本文从广义的刑事政策角度出发,认为人形机器人运行中的刑事政策是指以实现预防和惩罚犯罪为目的,国家或者社会据以组织对人形机器人运行过程当中犯罪现象的反应的方法或者措施的总和。其中既包括弱人工智能时代人形机器人运行中的刑事政策,也包括未来强人工智能时代人形机器人运行中的刑事政策。

(二) 人形机器人运行中的刑事政策的特征

从人形机器人运行中的刑事政策的界定出发并结合人形机器人的自身特点,本文认为,人形机器人运行中的刑事政策需要关注以下特征,即责任的有限性、人机的共责性、处罚手段的独特性以及技术发展的保障性。

1) 责任的有限性

针对人形机器人是否具有刑事责任主体地位并能否承担刑事责任,学界基本上持三种观点,肯定说、折衷说以及否定说。由于此处讨论的是人形机器人运行中的刑事政策,因此,本文是站在肯定说的前提下对其刑事政策进行的探讨。但应当注意的是,本文认为,人形机器人是否具有刑事责任主体地位应当区分所处时代,即在弱人工智能时代人形机器人由于不具有自主的意识与意志,因此不具有刑事主体地位,但是在强人工智能时代,由于其具有在自主意识和意志支配下的辨认与控制能力,因此应被认定为刑事责任主体[7]。

一方面,与动物不同,人形机器人在强人工智能时代具有刑事责任能力。动物因为不具备自主的理性思维以及对人类道德与伦理的认识,因此不承担刑事责任,一般不认为其具有刑事责任能力。但是强人工智能时代的人形机器人由于具有“深度学习”能力甚至具有“独立意识”,因此需要承担刑事责任。另一方面,如同人类一般,由于受主观认知能力以及客观环境的影响,即便是在强人工智能时代,人形机器人运行中所承担的刑事责任也应当是有限的。“智能”不等于“无所不能”,不能因为其“智能”的特性而对其要求过高。既然如此,还应当注意到,人形机器人也存在无刑事责任能力、限制刑事责任能力和完全刑事责任能力之分,对其区分标准应以其心智能力为基本,其能力大小应以人类正常个体的能

力素质为参照[8]。

2) 人机的共责性

现有刑法针对自然人犯罪与单位犯罪作出规定，自然人犯罪中的相关自然人承担相应的刑事责任自然无异议，对于单位犯罪则采取了双罚制，即不仅要求单位承担一定的刑事责任，而且要求其中“直接负责的主管人员”和“其他直接责任人员”同样承担。弱人工智能时代中，因人形机器人缺乏“独立意志”，不具备独立辨别与控制力，因此，人形机器人充当的是“工具”的角色。此时如果由于其在运行中对刑法所保护的法益产生危险甚至是造成了危害结果，根据情况，该人形机器人的研发者、生产者甚至是使用者应承担一定刑事责任。在强人工智能时代的背景下，因为人形机器人此时具备独立的控制与辨认能力，如果其基于独立意志实施了犯罪行为，那么除了该人形机器人需要承担刑事责任外，还需考虑研发或使用者是否具有故意或过失，如果具有犯罪的故意或者过失的话，应当与该人形机器人共同承担相应的刑事责任。从这个角度来说，自然人与人形机器人具有共责性，刑事政策的制定与执行应当充分考虑二者的刑事责任，从而达到预防与惩治犯罪、维护社会和谐与安全的效果。

3) 处罚手段的独特性

我国针对自然人犯罪的刑罚处罚方法多样，针对单位的刑罚处罚则只有罚金刑，这是因为单位是一个集体，单位本身不具有人的感情，基于其受刑能力，也只能对其处于罚金刑。同样，针对人形机器人的刑事政策应当注意根据其独特性来制定对应的刑罚政策。人类和人形机器人的受刑感受和被剥夺感建立在不同惩罚形式的基础上[8]，那么以目的刑观念为基础，对人形机器人实施罚金刑能否到达特殊预防的效果，这有待考究。因此，人形机器人运行的刑事政策应当考虑到其区别于人类和单位的特别之处，进而制定相应的刑罚政策。

4) 技术发展的保障性

启蒙运动后，人之所以为人的基本权利得到重视，而后各国宪法中都对人的基本权利进行了规定，我国宪法中也以专章的形式规定了公民的基本权利与义务。在宪法的框架之下，各种政策的制定也建立在保障人权的基础上，例如社会福利制度、教育政策、医疗政策等。与人类相比较，人形机器人毕竟属于人工智能，是由人创造出来的产物，具有技术性的特征，由于人形机器人将人工智能、高端制造、新材料等前沿技术相融合，有望继计算机、智能手机、新能源汽车之后，成为新一代颠覆性产品，将深刻改变人类的生产与生活方式，重新塑造全球产业的发展格局[9]，因此面对人形机器人运行中可能存在的刑事风险，在防范时需要注意预防与保障相结合，保障人形机器人技术的健康有效发展，在宽严相济的基本刑事政策下有针对性地制定具体的刑事政策，以确保社会和谐与科技进步相协调，统筹安全与发展的关系。

3. 人形机器人运行中的刑事政策的基本内容

犯罪问题或者说社会稳定问题一直以来被摆在国家治理中第一位，刑事政策是人类社会用来解决犯罪问题的智慧结晶[3]，可见，刑事政策在国家治理过程中发挥着重要作用。从这个意义上来说，智能时代人形机器人出现的背后所隐藏的对社会可能产生严重危害的行为，由刑事政策率先出面是必要且可行的。而且由于刑事政策体现的是刑法的变动性的一面，其对社会生活具有高度敏感性，且决策过程也较立法程序更为便利，因此，刑事政策能够促使刑法面对社会发展进行及时地调整[10]。宽严相济是我国的一项基本刑事政策，人形机器人运行中刑事政策的内容依然需要坚持这一基本政策，无论是在弱人工智能时代亦或是强人工智能时代，均需体现有宽有严、宽严相济的主旨，不一味从严，更不一味宽容，结合人形机器人发展的时代特点，合理配置相应的刑事政策内容。

(一) 对弱人工智能时代人形机器人运行中的刑事政策

1) 加大对利用人形机器人实施犯罪人员的处罚力度

学界虽然针对人形机器人是否具有刑事责任主体地位有争论，但主要体现在强人工智能时代的讨论上，对于弱人工智能时代的人形机器人，因其不具有自主意识和意志，完全处于人类的控制之下，“从属于人类的工具”^[11]，所以自然不应具有刑事责任主体地位。既然弱人工智能时代人形机器人只具有工具的属性，那么就会发生设计者或者使用者利用其进行犯罪的行为，也就是说，此时主要涉及对利用者犯罪行为的规制。

虽然弱人工智能时代人形机器人作为工具而存在，相比于强人工智能时代具有独立意志和意识的人形机器人其“智能性”略显不足，但是相比于普通的人工智能机器人，其具有“深度学习”的能力，这种能力为其提供了远超人类的成长速度。如2017年AlphaGo与棋手柯洁展开围棋对决，连胜三把，这是继2016年战胜世界围棋冠军李世石之后又一次战胜人类。同时，2017年10月AlphaGo Zero仅通过三天学习便又战胜了AlphaGo。基于这样的“深度学习”能力，人形机器人不断融入人类生活的各个方面，为人们提供更多的便利服务。然而，便利与风险同在，如果这些人形机器人被不法分子所利用，风险也会在更大范围内存在。如果不法分子利用这些机器人进行杀人，也就是说出现“机器人杀手”的话，整个社会的秩序都将面临威胁。实际上，目前已有记录在案的无人机在无人命令的情况下向人发动攻击的案例^[12]。¹除此之外，由于人形机器人运用范围之广，从家庭环境到公共环境，也因此利用其实施犯罪将会威胁到不同法益，从自然人的生命权到公共安全等。基于利用人形机器人实施犯罪行为可能造成的危险范围之广、程度之深，有必要对利用其实施犯罪的人员从重处罚。一方面，这符合罪责刑相适应的原则，另一方面，有利于通过对犯罪人的特殊预防进一步达到一般预防的目的。

对于对利用人工智能机器人实施犯罪的研发者或使用者从重处罚的这一观点，有学者认为，在此基础上还需要考虑的是，无论是利用何种工具实施犯罪，除刑法有特别规定外都应按照相同的标准对行为人的犯罪行为统一量刑，否则会违背刑法的平等适用原则与罪责刑相适应的原则^([2], p. 52)。本文赞同这一说法，如果仅仅考虑到犯罪工具的问题，则可能会一叶障目，不见泰山，还需要考虑具体的犯罪情况、违法性的程度以及预防的必要性。

2) 兼顾防治犯罪与技术发展

预防与惩治犯罪和技术发展之间会产生矛盾，技术发展会导致新型犯罪的出现，而预防与惩治犯罪则可能会阻碍技术的发展，因此应当正视与调和二者之间的矛盾，兼顾预防、惩治犯罪与技术发展的保障。

刑事政策的目的与价值表现为一种功利主义，重点在于预防和惩罚犯罪的有效性。面对人工智能机器人的出现和发展带来的新型犯罪类型和产生的新问题，通过刑事政策进行一定程度地防治，有利于及时回应社会治理的需求。但是基于刑事政策的功利性特点，需要注意的是在预防和惩治犯罪的过程中不能以切断技术的发展进路为前提。毕竟，人形机器人技术具备巨大的发展潜力且应用前景广^[9]。刑事政策运用的初衷是为了解决由于科技发展带来的刑事问题，其最终还是需要为科技发展提供一个更加清朗的环境，如果在这一过程中忽视了技术发展的要求，则背离了初衷，一定程度上失去了政策加持的意义。

因此，应当有宽有严，以严格审慎的态度对待严重侵害法益的行为，也要将那些“允许的风险”予以出罪化处理^[13]。

(二) 对强人工智能时代人形机器人运行中的刑事政策

1) 将人形机器人纳入刑罚处罚范围

¹在联合国披露的一份关于2020年3月利比亚军事冲突的报告中称，土耳其STM公司生产的“卡古-2”(kargu)型四旋翼无人攻击机被编程为在不依靠操作员的情况下自主攻击了撤退的哈夫塔尔指挥的“利比亚国民军”，并很有可能造成了一名国民军相关成员的死亡。

刑罚政策属于宏观刑事政策系统中的一部分,指设定刑罚目的和运用刑罚手段的政策。刑罚的目的充分体现了刑罚政策,可以说刑罚目的本身是一种政策思想[14]。基于刑罚预防犯罪的目的,具体来说需要将人形机器人纳入刑罚处罚的范围中,即对于超出设计和编制程序范围实施了严重危害社会行为的人形机器人进行刑罚处罚,实现特殊预防与一般预防的目的([2], p. 131)。

将人形机器人纳入刑罚处罚范围首先需要解决的是其刑事责任主体地位的问题。关于强人工智能时代人形机器人是否具有刑事责任主体地位,学界大致持有肯定说与否定说。持否定说的学者大致有以下观点:第一,尽管人形机器人具有了独立的辨认与控制能力,但是刑事责任能力仅为刑事责任主体地位的必要条件而非充分条件,具有刑事责任能力的主体却并不一定拥有独立刑事责任主体的地位[15]。第二,机器人并非自然人,其虽具有相当智性,但并无人之心性和灵性,不可将其等同于具有“人类智慧”的自然人和自然人集合体。将其作为拟制之人使其享有法律主体资格,这在法理上尚有斟酌之处[16]。第三,人工智能并不具有自由意志,其“智能”产生的来源在于人类设计和编制的程序,即使后来具备超出这一程序实施行为的能力,但鉴于“智能”产生的基础,其也很难在善恶之间进行抉择,即人形机器人不可能具有自由意志,也就不可能成为刑事责任主体。肯定说认为,在强人工智能时代,人形机器人因为具有认识上的理性和在实践中的自主行为的能力,因此其具备刑事责任主体所需要的辨认与控制能力,能够成为刑事责任主体[17]。也有学者认为,讨论人工智能法律主体资格问题,必须跳出人类中心主义。无论基于权利意志理论还是利益理论,人工智能作为法律上的主体均可成立。以“非唯人类中心”的哲学价值看待待非“人”的实体,通过法律拟制而赋予一定范围的法律主体资格,最终做到人类与人工智能和谐共舞[18]。

本文赞成肯定说,强人工智能时代赋予人形机器人刑事责任主体地位是可能且必要的。一是认为“具有刑事责任能力不一定成为刑事责任主体”的观点不仅排除了人形机器人作为刑事责任主体的可能性,也排除了单位作为刑事责任主体的可能性。如果说刑事责任主体须同时具备达到刑事责任年龄且具有刑事责任能力两个条件,那么单位就无法成为刑事责任主体。单位成为刑事责任主体是法律拟制的结果,那么据此,从拟制角度考察人工智能体的刑法地位具备合理之处,此时这一否定的观点就无法成为排除理由。“参考法人犯罪的思路,对于人工智能作为拟制主体看待未尝不是未来人工智能犯罪刑法规制的可能思路。”[19]二是人工智能是否存在自由意志是一个客观问题,不能够主观臆断。当人形机器人具备自由意志产生的条件时,除非有充分的证据,否则不能绝对肯定人形机器人不具有自由意志([2], p. 118)。

强人工智能时代人形机器人具有刑事责任主体地位,基于此,运用刑罚对其独立实施的严重危害社会的行为加以处罚,不仅能实现特殊预防的效果,而且能达到一般预防的目的。就特殊预防来说,由于此时人形机器人具有感知能力,对其施加刑罚能够使其感受到痛苦并进一步认识到其行为的可谴责性,从而达到自省的效果。就一般预防来说,对具体的人形机器人的处罚能够使其他可能实施犯罪行为的人形机器人强化规范和法律,同样能发挥刑罚的教化功能,达到一般预防的目的。

2) 重视人形机器人自身的权利保障

虽然将实施严重危害社会行为的人形机器人纳入刑罚处罚范围,但是不可忽视人形机器人也有可能成为被害人,同时要注意保护其合法权利。当人工智能体作为直接犯罪对象时,行为人会针对人工智能体的实体和系统安全实施犯罪。一种形式是作为故意毁坏财物罪等损毁犯罪的对象,从实体上对其予以破坏;另一种形式是作为破坏计算机信息系统罪等计算机犯罪的对象,对其系统、信息安全予以破坏[20]。无论是哪一种形式,都会侵害人型机器人所存储的数据与信息,还会损害其认知与情感。一方面,对数据与信息的破坏直接影响到社会秩序的稳定,随之而来的结果是人类的信息安全得不到保障,权利受到侵害。另一方面,忽视人形机器人对正向情感的需求和遭受不公的现实,还有可能会引发其反面的情绪,甚至使其失控,进而危害人类社会秩序。因此,需要在对其犯罪行为规制的过程中,也重视对其合

法权利的保护。

综上所述,无论是在弱人工智能时代还是在强人工智能时代,均需对人形机器人的运行采取宽严相济的刑事政策。在弱人工智能时代对人形机器人的刑事政策包括加大对利用人形机器人实施犯罪人员的处罚力度与兼顾防治犯罪与技术发展,在强人工智能时代对人形机器人的刑事政策包括将人形机器人纳入刑罚处罚范围以及重视人形机器人自身的权利保障。在两个发展阶段均采取了有宽有严的刑事政策,在加强预防与惩治犯罪的同时,也保障技术进步与社会秩序的稳定,发挥宽严相济刑事政策的功能与作用。

4. 人形机器人运行中的刑事政策的实现途径

理论的生命力在于实践,只有在实践中才能对理论进行检验和完善,促进理论与实践的良性互动。人形机器人运行中的刑事政策也应当被纳入到实践中来,通过制定和执行来检视其内容的合理性与可行性。具体来说,其制定与执行同样需要根据不同阶段人形机器人的发展状况来进行,而在每一阶段中,都应在对应制定理念的基础上配之以相应的执行手段,以达到有效实现人形机器人运行中的刑事政策。

(一) 对弱人工智能时代人形机器人运行中的刑事政策的制定与执行

1) 坚持社会安全与技术发展并行

安全是人工智能时代的核心法价值。人工智能作为未来时代技术尚在深入发展之中,但在当下已引发人们对其安全问题的普遍担忧[16]。在弱人工智能时代,人形机器人具有的深度学习能力使其能够快速获取、整合数据并学习,这种高效性与智能性让人形机器人被广泛应用于各领域各行业。然而,正是由于这一阶段人形机器人仅仅作为工具被广泛应用的现实,加之其以研发者所编制的程序与各行各业的信息和数据为运行基础,导致安全问题日益突出,无人驾驶汽车造成的交通事故、“杀人机器”的出现、公民个人信息的大量泄露以及由于信息泄露造成的国家安全隐患等问题,均不容忽视。2023年工业和信息化部印发的《人形机器人创新发展指导意见》中指出,在发展人形机器人的同时要加强安全治理能力,提升人形机器人功能安全性能,确保相关技术产品对人和环境友好。强化网络安全防护,提高信息获取、数据交互、数据安全等技术保障能力。强化整机、关键部组件、核心软件、算法等重点环节安全风险评估,促进安全能力提升[9]。因此,在制定人形机器人刑事政策时需要高度重视社会的安全问题,坚持安全发展理念,避免让社会陷入由不确定造成的恐慌之中。

但是,坚持安全发展并不代表不发展,而是要在发展的同时兼顾安全,既不盲目发展,也不故步自封。2021年12月21日,工信部、国家发改委等15个部门正式印发《“十四五”机器人产业发展规划》(工信部联规[2021] 206号)提出,到2025年我国成为全球机器人技术创新策源地、高端制造集聚地和集成应用新高地[21]。人形机器人的技术创新与发展是时代发展之潮流,也是国家发展之必须,刑事政策必须服务于“科技强国”,不能因为科技发展所造成的安全隐患退而却步。科技是一把双刃剑,但是不能因为害怕而放弃发展,重点在于如何防患于未然以及问题出现时如何解决问题。

基于此,从广义的刑事政策观出发来实现这一内容,不仅需要刑法发挥其最后法、保障法的功能对利用人形机器人实施犯罪严重危害社会安全的行为予以规制,而且需要发挥其他部门法的前置法作用,在防范风险的同时,保障技术的创新发展。例如,行政机关可通过设定统一的市场准入标准严格控制和规范人形机器人的市场投入使用,尤其需要将对人形机器人的安全评估标准纳入到准入标准的范围内,实现事前的监管。此外,因人形机器人造成损害结果的情况,需要区分处理。如果由于产品自身缺陷而造成的损害,受害者可基于产品责任向生产者或销售者申请赔偿,得到权利救济。如果是因其他民事主体利用人形机器人实施了侵犯自身民事权利的行为,受害者亦可向实施者申请赔偿,实施者应当承担侵权责任。事前、事中与事后的共同治理,可有效实现安全保障与技术创新和发展并行的要求。

总之，既要重视安全防范风险，又要保障技术的创新与发展，在保证技术发展的基础上防范技术风险，也要防范因刑事政策所带来的制度性风险[22]。只有兼顾二者，才能使刑事政策在弱人工智能时代发挥出应有的作用。

2) 重视技术监测与预防

贯彻实施人形机器人刑事政策的过程中，需要重视对人形机器人的技术监测与预防，将风险的出现亦或风险本身降低到最低限度。一方面，技术监测能够促使人形机器人的研发一方始终对产品负责，提升技术水平，并通过对程序运行情况的追踪来达到预防风险的效果。技术监测属于源头措施，研发人员不仅具有研发的责任，而且需要避免所研发的人形机器人在运行中出现故障，或者说在故障出现以后，要及时地对程序或者技术问题进行修复与提升。而且，在人形机器人运行中出现严重危害社会的行为时，是否进行了技术监测，能够作为判断研发人员是否履行注意义务的因素之一。另一方面，技术预防是指预防出现和使用违背社会规范伦理与社会价值和秩序的技术方法。技术的使用需要服务于人类社会，如果一项技术超出了这个范围，凌驾于人类之上，那么其造成的后果将不堪设想。对此，可考虑设立技术伦理审查委员会，讨论并制定出一份统一的技术审查标准，尤其可将人形机器人技术的应用模式、应用范围等纳入审查范围。这一审查标准亦可作为行政机关制定人形机器人市场准入标准的重要参考，达到双重保障作用。

总之，人工智能刑事法治应当确保人工智能无法完全掌控自身和人类社会的控制权，保障人工智能的发展始终有益于人类[23]。从这个意义上来说，实施人形机器人刑事政策的过程中就必须重视技术的监测与预防。

3) 加强培训与宣传教育

有时候，时代发展的速度甚至会超越人们的认识水平，如今人工智能技术还将继续发展，为了保障其发展与创新，预防和应对可能或者已然出现的刑事风险，不仅需要对相关专业的人员进行培训教育，而且需要对公众进行宣传教育。

一方面，要加强对专业人员的培训教育力度。这里的专业人员不仅包括人形机器人的研发、生产与使用人，而且包括司法行政机关人员。对研发、生产与使用人员应当向其明确相关权利与义务，尤其是要履行好义务，防止因自身的过失行为引起人形机器人实施严重危害社会的行为。对司法工作人员，尤其是侦查人员，要加强其侦查能力，包括犯罪数据感知能力、犯罪数据获取能力、犯罪数据分析能力、犯罪情报应用能力和新型犯罪案件办理能力等[24]。只有与时俱进地提升自身能力，才能在刑事风险防范与化解中应对自如。另一方面，要加强对公众的宣传教育。不仅要宣传人形机器人的便捷、高效等优势之处，而且要宣传其可能带来的法律风险，对个人隐私的泄露风险甚至是严重危害公共安全等重大风险。此外，还要宣传如何应对以上风险，避免引起公众对于未知的恐慌，引导对人形机器人形成正确的态度与看法。

(二) 对强人工智能时代人形机器人运行中的刑事政策的制定与执行

1) 兼顾刑法的刑事政策化与刑事政策的刑法化

刑法的刑事政策化强调的是刑事政策作为刑法与政治之间的桥梁，发挥着将政治层面的要求与呼声传递给刑法教义学体系的功能[25]。与当下弱人工智能时代相比，强人工智能时代最突出的特点是人形机器人拥有自主独立的行为意识，如前文所述，此时针对人形机器人实施的严重危害社会的行为，宜将其作为刑事责任主体来进行规制，以实现严厉打击人形机器人犯罪的刑事政策。但是，要想实现这一刑事政策，还需要将内容和精神内化于刑法当中，毕竟刑法比刑事政策更为严厉且更具强制性，在政策难以施行时，其对惩罚犯罪和保障人权起重要的保障作用。“刑法教义学若想保持与现实社会的良性互动，就必须存在一种能够灵敏感知刑事政策需求并据此做出相应调适的机制，以使体系具备自我更新或者说

与时俱进的能力。”[25]也就是说,面对人形机器人作为犯罪主体时,刑法需要根据刑事政策的精神与内容进行及时修改,对人形机器人的犯罪行为予以严厉打击,对与其相关的合法权利加以保护,以便于满足新的社会需求,在惩治犯罪的同时保护技术的健康发展。

刑事政策的刑法化是与刑法的形势政策化相对应的命题,强调的是刑事政策在立法、司法中发挥作用时须符合法治国家的原则,不能够逃离罪刑法定原则的约束,更不能脱离宪法的维度。强人工智能时代人形机器人的犯罪行为,是一种完全意义上的新型犯罪,此时如果刑事政策的制定和执行一味地严厉以至超出罪刑法定与宪法的框架,则反而会适得其反,甚至引发人形机器人的负面反应,造成更大的危害。对此,更需要约束刑事政策,以刑法教义学对刑事政策的作用和效果加以一定程度的限制,以防刑事政策取代刑法规定。

总之,有效地实施贯彻刑事政策既需要刑法将其精神内化于其中,又需要刑法对其做出一定限制,防止刑事政策的边界过大。也就是说,刑事政策的实施需要兼顾刑法的刑事政策化与刑事政策的刑法化,在这过程中贯通刑事政策与刑法,更大化实现刑事法治的价值。

2) 重视社会协同治理

“也许,风险社会理论给予人类最大的启示在于,今天人类所面临的巨大风险也是人类集体努力的副产品,因此需要人类在连带主义基础上的集体分担,而以责任(刑罚)个别化为显著标志的刑法绝非防御或者控制风险的最佳手段。”[26]面对强人工智能时代下人形机器人可能带来的巨大风险,如果一味使用刑事类的措施和手段来应对,反而会对社会 and 人的基本权利带来破坏,对此,从广义的刑事政策角度上来说,除了以国家为主体所采取的刑事类措施,还需要社会主体的协同治理。

一方面,人工智能企业应当参与治理。大型人工智能企业往往具有从技术治理理念到技术治理准则等一系列完整的技术治理体系,意味着在刑事治理的过程中其实质上已经成为承担安全保护义务的“公共责任公司”[27]。在此基础上,企业与国家形成合作,企业通过制定有效的计划以控制刑事风险、预防犯罪,而国家对企业的风险控制予以一定的激励[28],营造较为缓和的风险治理环境,也能够促进刑事政策的实施。

另一方面,公众也应参与治理。刑事政策的最终目的是预防犯罪,而预防犯罪的目的也是为了民众,更何况民众实际上也是预防犯罪的基础,因此,在刑事政策的制定与执行中必须有社会团体和公众的参与。如我国制定社会治安综合治理政策就必须要有公众的参与,要广泛地征求民众的意见,要由社会团体和公众参与实施[29]。况且,人形机器人从产生到运行的环节均与公众息息相关,甚至可能发生自然人与人形机器人共同犯罪的情况,因此,对强人工智能时代人形机器人运行的刑事政策的制定与实施需要公众的参与,既保证决策的民主性,也提醒公众规范自己的行为。

5. 结语

科技的发展是必然的,身处其中的人工智能技术亦处在不断地进步之中。然而,当技术发展所带来的便利之洪流涌向人类社会时,风险也不知不觉到来,人形机器人基于类人的外形和智能性捕获人心的同时,也为人们带来了新的刑事风险。由于刑法需要保持稳定性,无法及时回应新型刑事风险,此时就需要更加灵活的刑事政策以刑事类及非刑事类的手段来进行应对。人形机器人运行中的刑事政策虽具有特殊性,但依然需要以宽严相济的刑事政策作为基本遵循。具体来说,在弱人工智能时代,一方面需要加大对利用人形机器人犯罪人员的处罚力度,另一方面又需要以兼顾防治犯罪和保障技术发展为内容;在强人工智能时代,一方面要将人形机器人纳入刑罚处罚范围,另一方面也要将人形机器人自身权利的保障纳入刑事政策的内容之中。为了确保刑事政策内容的有效实施,需要分别确立不同的制定与执行路径。在弱人工智能时代,首先要坚持社会安全与技术发展并行,在此基础之上重视技术监测与预防并且

加强培训教育与宣传教育；在强人工智能时代，还需要进一步兼顾刑法的刑事政策化与刑事政策的刑法化两个命题，加强重视社会的协同治理功效。对人形机器人运行中刑事政策的探讨，不仅旨在对新兴刑事风险的防患于未然，而且在于确保人工智能技术能真正为社会发展所服务，在于保障科技发展与社会和谐的协调和一致。

参考文献

- [1] 齐文远. 应对中国社会风险的刑事政策选择——走出刑法应对风险的误区[J]. 法学论坛, 2011, 26(4): 12-20.
- [2] 刘宪权. 人工智能时代的刑法观[M]. 上海: 上海人民出版社, 2019.
- [3] 卢建平. 刑事政策研究的中国特色[J]. 河北法学, 2008(11): 10-11.
- [4] 舒国滢. 法理学[M]. 第3版. 北京: 北京大学出版社, 2019: 300.
- [5] 梁根林. 现代法治语境中的刑事政策[J]. 国家检察官学院学报, 2008(4): 155.
- [6] 劳东燕. 罪刑规范的刑事政策分析——一个规范刑法学意义上的解读[J]. 中国法学, 2011(1): 122-140.
- [7] 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(5): 113-121.
- [8] 刘仁文主编. 社会变迁与刑法科学新时代[M]. 北京: 中国社会科学出版社, 2019.
- [9] 工业和信息化部. 人形机器人创新发展指导意见[EB/OL]. 2023-11-02.
https://www.miit.gov.cn/zwgk/zcjd/art/2023/art_e3f5686c2f0d49f9968b7ae011d558e1.html, 2024-04-13.
- [10] 陈兴良. 刑法的刑事政策化及其限度[J]. 华东政法大学学报, 2013(4): 3-15.
- [11] 时方. 人工智能刑事主体地位之否定[J]. 法律科学(西北政法大学学报), 2018, 36(6): 67-75.
- [12] 观察 | 无人机首次自主杀人, 机器杀人的“潘多拉魔盒”打开? [EB/OL]. 2021-06-02.
https://www.thepaper.cn/newsDetail_forward_12931689, 2025-09-12.
- [13] 侯跃伟. 生成式人工智能的刑事风险与前瞻治理[J]. 河北法学, 2024, 42(2): 160-178.
- [14] 储槐植. 刑事政策的概念、结构和功能[J]. 法学研究, 1993(3): 52-58.
- [15] 刘仁文, 曹波. 人工智能体的刑事风险及其归责[J]. 江西社会科学, 2021, 41(8): 143-155+256.
- [16] 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): 131-134.
- [17] 王耀彬. 类人型人工智能实体的刑事责任主体资格审视[J]. 西安交通大学学报(社会科学版), 2019, 39(1): 138-144.
- [18] 范进学. 人工智能法律主体论: 现在与未来[J]. 政法论丛, 2022(3): 3-17.
- [19] 王肃之. 人工智能犯罪的理论与立法问题初探[J]. 大连理工大学学报(社会科学版), 2018, 39(4): 53-63.
- [20] 王肃之. 人工智能体刑法地位的教义学反思[J]. 重庆大学学报(社会科学版), 2020, 26(3): 122-130.
- [21] 中国政府网. “十四五”机器人产业发展规划[EB/OL]. 2021-12-21.
https://www.gov.cn/zhengce/zhengceku/2021-12/28/content_5664988.htm, 2025-09-12.
- [22] 刘艳红. 网络时代社会治理迭代升级与犯罪控制协同化的刑事政策[J]. 社会科学辑刊, 2024(1): 65-72+238+241.
- [23] 皮勇. 人工智能刑事法治的基本问题[J]. 比较法研究, 2018(5): 149-166.
- [24] 徐永胜, 冯俊, 张敏. 人工智能技术背景下新型网络犯罪的治理创新——第二届新型网络犯罪治理智库论坛综述[J]. 浙江警察学院学报, 2023(4): 70-78.
- [25] 劳东燕. 刑事政策与功能主义的刑法体系[J]. 中国法学, 2020(1): 126-148.
- [26] 卢建平. 风险社会的刑事政策与刑法[J]. 法学论坛, 2011, 26(4): 25.
- [27] 单勇. 数字看门人与超大平台的犯罪治理[J]. 法律科学(西北政法大学学报), 2022, 40(2): 74-88.
- [28] 孙国祥. 单位犯罪的刑事政策转型与企业合规改革[J]. 上海政法学院学报(法治论丛), 2021, 36(6): 21-38.
- [29] 魏东. 论广义刑事政策的基本内涵[J]. 清华法学, 2011, 5(2): 9-23.