

涉人工智能犯罪的刑法规制

李天智

山东建筑大学法学院, 山东 济南

收稿日期: 2024年3月21日; 录用日期: 2024年4月7日; 发布日期: 2024年5月10日

摘要

随着社会科学技术不断进步, 人类对于人工智能产品的利用越来越多, 涉及人工智能的犯罪也会越来越多。人工智能分为弱人工智能和强人工智能, 我们讨论的涉人工智能犯罪刑法规制包括了上述两种情形。目前涉人工智能犯罪的类型有三种: “工具利用型”、“犯罪对象型”、“人工智能失控型”, 都对现有的刑法带来了不小的冲击。我国的刑法应当进行前瞻性刑法立法: 增设新罪名、完善现有罪名, 还应当考虑赋予强人工智能刑事责任主体地位。唯有如此, 我们才能够更好的利用刑法, 实现保护法益的目的。

关键词

人工智能, 人工智能犯罪, 刑法规制

Criminal Law Regulation on Crimes Involving Artificial Intelligence

Tianzhi Li

School of Law, Shandong Jianzhu University, Jinan Shandong

Received: Mar. 21st, 2024; accepted: Apr. 7th, 2024; published: May 10th, 2024

Abstract

With the continuous progress of social science and technology, human beings are using more and more artificial intelligence products, and there will be more and more crimes involving artificial intelligence. Artificial intelligence is divided into weak artificial intelligence and strong artificial intelligence, and the criminal law regulation of artificial intelligence-related crimes which we discuss includes the above two situations. At present, there are three types of artificial intelligence-related crimes: “tool utilization type”, “criminal object type”, and “artificial intelligence out of control type”, which have brought no small impact on the existing criminal law. China’s criminal law should carry out forward-looking criminal law legislation: adding new charges, improving existing

文章引用: 李天智. 涉人工智能犯罪的刑法规制[J]. 法学, 2024, 12(5): 2746-2751.

DOI: 10.12677/ojls.2024.125389

charges, and should also consider giving strong artificial intelligence the subject status of criminal responsibility. Only in this way can we make better use of the criminal law and realize the purpose of the protection law.

Keywords

Artificial Intelligence, Crimes Involving Artificial Intelligence, Criminal Law Regulation

Copyright © 2024 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

随着社会生产力的不断发展，新兴技术不断涌现。以蒸汽机为代表的第一次工业革命，使机器可以代替手工劳动；以电力为代表的第二次工业革命，使人类进入了“电气时代”；以计算机为代表的第三次科技革命，改变了人类日常生活的各个方面。新技术、新事物层出不穷，法律也要及时更新，尤其涉及新兴技术犯罪，需要刑法进行合理的规制。第三次科技革命不断发展，人类在信息技术的应用上取得了重大突破，尤其是人工智能技术，人工智能技术的迅速发展与法律的滞后性之间的矛盾被进一步放大，涉及人工智能犯罪的刑法规制值得我们进行研究。

人工智能，是指其具备部分和人类一样的智能性，但它归根结底是人类的创造物。目前对于人工智能大体可以分为“弱人工智能”和“强人工智能”两种。弱人工智能是指不具有“自主意识”的人工智能，它实施的数据处理或者行为都是在研发者或者使用者的支配之下。强人工智能是指未来可能会出现，能够具有完全“自主意识”的人工智能体，它所实施的行为完全依照其自主意愿，能够超出设计和编程的范围。我们人类社会目前所利用与使用的，大多是不能进行自主学习的“纯机器”，和不具有自主意识、只能在编程范围内进行一定程度学习的“弱人工智能”。

2. 涉人工智能犯罪定义

我们研究涉人工智能犯罪，既包括弱人工智能有关的犯罪，也包括强人工智能有关的犯罪。既包括在人工智能的研发以及使用过程中有可能涉及的犯罪，也包括未来出现的强人工智能的“自主”犯罪。如在人工智能研发过程中，研发者出于故意，制造一种犯罪工具，用于实现研发者的犯罪目的；或者人工智能的使用者，借助人工智能技术或者人工智能的行为，去实现使用者的犯罪意图，达成犯罪目的；在未来也有可能出现强人工智能，能够跳出编程之外，自主实施犯罪行为或者与使用者、其他人工智能一起进行犯罪。目前我们科技水平越来越高，对于人工智能的利用也越来越频繁，研究涉人工智能犯罪的种类以及规制就显得十分必要。

3. 涉人工智能犯罪的种类

目前对涉人工智能犯罪我们可以将其划分为三种不同的类型，分别是“工具利用型”、“犯罪对象型”、“人工智能失控型”。

3.1. 工具利用型

“工具利用型”人工智能犯罪，是指犯罪分子通过各种手段，把人工智能当作实现其犯罪目的的工具

具，造成了危害他人、国家、社会的不正当结果。在这种犯罪实施的过程之中，人工智能仅仅是犯罪分子所依赖的工具。目前常见的工具利用型人工智能犯罪是研发者或者使用者，在研发或者使用的过程中，利用人工智能实施犯罪行为。

研发者如果实施工具利用型人工智能犯罪，绝大部分是由于主观故意，而不可能是过失。首先，对于研发者处于直接故意的主观心理。研发者完全有可能为了一己私利或者某种不正当的意图，制造出一种具有特定效果的人工智能产品，达成自己的犯罪目的，侵犯某种法益。对于这种情况，我们可以清楚的判断出，此时人工智能产品只是研发者的犯罪工具，只需要按照刑法对其定罪、量刑便可。比如，研发者为了实施自己的杀人目的，制造出一种智能杀人机器，杀害了某个人，这种情况我们完全可以对其按照故意杀人罪定罪量刑。其次，对于研发者间接故意的心理状态。存在这样一种情况，研发者明知自己的人工智能产品没有达到某种安全标准，可能会造成危害社会的结果，放任其流向市场。对于这种情况，研发者处于一种间接故意的心理状态，可以说此时如果发生了侵害法益的结果，研发者要承担责任，这个时候我们可以定生产、销售伪劣产品罪还是定故意杀人罪或危害公共安全罪；或者为了规范人工智能这种高科技产品的研发与生产，刑法新增一个有关人工智能的罪名、规定新的入罪标准进行处罚，从而达到更好的保护国家、社会、第三人的效果。如果任由研发者故意实施与人工智能有关的犯罪，后果是不可想象的。

同样，对于使用者来说，利用人工智能犯罪也只存在故意的主观心理，也有巨大的社会危害性。使用者利用人工智能进行金融交易，实现比其他交易更为迅捷的目的，从而谋取不正当利益，这种利用人工智能违反公平、操纵市场的行为往往比普通的破坏金融系统的犯罪带来更大的危害，给金融秩序带来巨大冲击；还存在自动驾驶技术的使用者，在程序的范围内，设定导航路径，制造汽车炸弹炸毁特定的场所、伤害不特定的公众，给社会安全带来巨大的危害的情况。上述使用者利用人工智能犯罪都出于故意的心理，对使用者本身利用传统罪名进行定罪处罚便可进行规制。

3.2. 犯罪对象型

“犯罪对象型”人工智能犯罪是指犯罪分子把人工智能系统当作其犯罪目标，通过破坏、侵入等手段，使其不能够正常运作，从而实现犯罪分子的犯罪意图。我们似乎很难发现犯罪对象型人工智能犯罪种研发者的身影，因此我们主要讨论使用者即行为人的利用行为的危害。

自古至今都存在盗窃等侵犯财产权的行为，然而随着科技发展，这种侵权行为出现了新的形式，侵犯财产权的方式发生了改变。ATM机的原理为人工智能系统代替人类处理支付业务，也会像普通的业务员一样成为侵犯财产罪的对象，针对ATM机侵犯财产的案件，就是犯罪对象型人工智能犯罪^[1]。与之类似的是，近些年热度持续升高的自动驾驶技术，似乎也更容易成为不法分子的攻击对象，带来巨大利益的同时，也给人类带来了潜在的风险。不法分子完全可以通过远程手段，侵入一家或者数家自动驾驶系统公司的系统，达到一种自我满足，继而可能引发一系列的犯罪行为。这种侵入人工智能系统的行为并不完全都是为了达到某种杀人等犯罪目的，这种人工智能系统有时候本身就具有价值，受到法律的保护，这种“犯罪对象型”人工智能犯罪与“工具利用型”人工智能犯罪并不完全相同。

行为人在侵害某种本身就具有价值的人工智能系统时，往往侵犯的就不单单是一个人的安全和利益，一个智能系统多数情况下是不特定的人共同使用，或者说这个智能系统能够保护不特定多数人的利益。行为人这种侵入一个智能系统的行为毫无疑问危及了不特定多数人的法益，涉及的范围大、程度广，往往不是工业社会能够达到的规模，刑法对于此种类型的人工智能犯罪必须加以规制。

3.3. 人工智能失控型

“人工智能失控型”人工智能犯罪，是指人工智能产品自身发生了故障导致危害社会结果的出现或

者具备“自主意识”的强人工智能自主实施了危害社会的行为。

对于第一种人工智能产品自身故障，这往往是由于研发者代码、程序错误或者是使用者不当使用造成的。人无完人，研发者在编写人工智能系统时，有可能会出于过失导致危害的发生，往往不可能是出于故意。对于过失犯罪，我们要求行为人“应当预见可能发生危险而没有预见”或者“已经预见到可能发生危险却轻信自己能够避免”，以其违反了一定的注意义务为前提。比如，研发者在研发无人机的过程中，对于无人机的结构承载力计算失误或者相信这样不会导致危险的发生，从而导致了危害结果，这就是典型的过失犯罪。虽然说，研发者没有利用其产品进行故意犯罪的目的，但是最终还是由于自己对产品的过失，导致自己的智能产品侵害了法益。当然如果这种注意义务，是研发者拼尽全力、在当时技术背景下无法预见到的，那么我们也不能对研发者有过高的要求，应该把这种危害结果归因于意外事件，也就不需要对其进行刑罚。同样的使用者在使用过程中也有可能没有注意遵守某些规定，从而使危害结果发生，对此不再过多赘述。人工智能产品是新一代科技产品，也不能任由这种过失犯罪率持续增加，否则会带来大于一般的犯罪行为的危害。

对于第二种人工智能失控型，我们往往指的是强人工智能，超出了人类的控制范围，做出一些侵犯法益的行为。强人工智能发展到一定程度，无论是否以实体形式出现，它带来的危害远远超过单个自然人尽其所能造成的危害。当强人工智能拥有自主意识之后，强人工智能也可能会选择和人类共同进行犯罪活动，也有可能选择和其他强人工智能联合犯罪。对于上述三种潜在的巨大风险，刑法上有必要进行理论研究。

4. 涉人工智能犯罪的刑法规制

4.1. 进行前瞻性刑法立法

前瞻性刑法立法又可以称作是预防性刑法立法，是为了防范风险、维护安全，让刑法在技术上超前干预、先发制人，防止严重犯罪的发生的一种立法方法。我国预防性刑法立法的领域主要存在于恐怖犯罪、网络犯罪、食品安全犯罪、公共安全犯罪、环境犯罪等可能侵害国家法益、社会法益或侵害不特定人、多数人人身财产安全的犯罪[2]。上述犯罪所侵害的法益属于重大集体法益，有前置化保护的必要性。

人工智能时代，仅仅依靠事后处罚的方法无法规制难以遇见的风险。我国刑法目前对于法益的保护有两种方式，一种是对实害犯进行处罚，处罚对法益造成实害结果的行为，惩罚犯罪、保护法益；另一种是对危险犯进行处罚，在行为对法益造成危险时就处罚，防止保护的法益遭受实际侵害。传统刑法是以实害犯为核心，以实害的发生作为可罚性的界限，这种保护方式契合了传统社会的犯罪简单化、直观化特点[3]。人工智能领域的信息系统更加智能化、敏感化，其处理的数据更加海量，因此人工智能领域犯罪的发生更加难以防范、危害后果更加难以计算，当然也就不能适用传统刑法等到后果发生才予以规制的做法。我国刑法在网络犯罪方面采取了预备行为实行化、帮助行为正犯化的预防性立法，其目的就是要实现早发现、早预防，不是在出现了现实直接的危险或实害结果才处罚，而是提前规制或者降低入罪门槛[4]。所以，注重事前预防人工智能领域的侵害风险，将刑法对人工智能犯罪评价的“起刑点”前移，才能实现对涉人工智能犯罪的有效预防和打击。

第一，适度设立新罪名，精准打击涉人工智能犯罪。无论是“强人工智能”犯罪，还是“犯罪对象型”犯罪，我们难以找到合适的法律依据，使行为人承担罪责刑，有时候无法用现有的罪名去进行预防与惩罚。对于这种新事物的出现，完全可以参考汽车、飞机等新技术出现而设立的新罪名，从而对于这种具有不同于其他类型的犯罪行为进行妥当且有效的规制，确立适当的罪状、罪名、法定刑。

第二，立足于刑法的谦抑性，完善现有罪名。我们社会目前应用的人工智能产品，大多是弱人工智能产品，因此涉及的人工智能犯罪，也多为“犯罪工具型”，少数为“犯罪对象型”。对于“犯罪工具

型”人工智能犯罪，尽管危害结果的发生，是借助人工智能产品来完成的，但此时的人工智能产品，并不具备独立的自主意识，仅仅能够按照编程的内容进行或活动，危害结果不能够归责于人工智能产品，其仅仅是犯罪分子的工具罢了。在研究弱人工智能犯罪的归责上，应当准确的在研发者或者使用者身上分配刑事责任。由于弱人工智能产品和非智能时代的其他产品一样，都是人类的工具、制造物，所以现有刑法中的一些罪名有时候能够或多或少的加以规制，比如生产不符合安全标准的产品罪、非法利用信息网络罪、帮助信息网络犯罪活动罪等罪名。

不得不说，人工智能产品又不同于其他的非智能产品，有的人工智能产品的应用范围广、保护的法益更为重要，关乎不特定多数人的利益，仅仅套用刑法现有的罪责体系无法准确进行刑事责任的承担，我们需要做的就是能够在能够利用现有罪名进行规制的情形下，对罪名进行完善，结合涉人工智能犯罪的特点，进行入罪标准的降低，或者对于与人工智能有关的犯罪根据其危害程度进行加重处罚、法定刑升格等。

4.2. 赋予强人工智能刑事责任主体地位

毫无争议的是，我们当今处于弱人工智能时代。但随着人类社会日新月异的变化、技术突飞猛进的发展，一部分学者认为强人工智能时代必然会到来，且并不遥远；当然也有一部分学者认为，强人工智能不可能出现，人类也不会允许有超越人类智能体的出现。持有两种不同观点的学者自然对赋予强人工智能体的刑事责任主体地位问题上存在肯定论与否定论之争。我认为，虽然目前的科技水平无法开发出强人工智能，但是就近些年人类文明的发展速度来看，不远的将来，强人工智能的出现是必然的，我们现在进行赋予强人工智能刑事责任主体地位的研究也是十分有必要的。

赋予强人工智能刑事责任主体地位契合刑法的理念。刑法上的刑事责任主体具有辨认和控制自己行为的能力，而我们在这里研究的强人工智能的前提就是具有完全“自主意识”，能够按照自己的意愿进行活动。既然强人工智能具备了刑事责任能力，按照刑法上的“无行为无责任”理念，强人工智能就应当为自己可能实施的犯罪行为负责。目前，大多数国家的刑法都是奉行“以人类为中心”的理念，拒绝将其他的非人类事物纳入处罚。早在古代社会，就有对石头、动物等非人类事物进行处罚的先例，只不过后来，人们发现处罚这些事物似乎也不会起到良好的效果，既不会有一般预防作用，又不会有特殊预防作用，才以人类为中心进行规制。随着社会的发展，“单位”这种经济社会的产物，也成为了除人类之外的刑法所规制的对象。单位本质上是由许许多多自然人进行运转的事物，但单位犯罪是不同于自然人犯罪的，单位可以成为刑法规制的对象，恰恰打破了这种“以人类为中心”的刑法传统。虽说法人是全体自然人的共同意志，其行为也是单位内部人员意志的加和所在，人工智能更是由人类创造的，其累积的知识和伦理道德与人类更为相似，可以说强人工智能是世界上最类似于人的主体^[4]。强人工智能和公司法人相比更应该成为刑法规制的对象。

赋予强人工智能体刑事责任主体地位能够更好的实现刑法的任务和机能。一方面，赋予强人工智能体刑事责任主体地位，可以通过规制强人工智能的行为来实现刑法的任务和机能。刑法的威慑力靠的是刑罚，通过让犯罪人承担一定的刑事责任，达到惩罚犯罪、保护法益的功能。强人工智能是具备自主意识的，目前刑法上虽然没有合适的刑罚进行处罚，但赋予其刑事主体地位之后，我们可以考虑新增删除数据、修改程序、永久销毁等刑罚，达到惩罚的目的。另一方面，赋予强人工智能体刑事责任主体地位，可以通过规制人类的行为实现刑法的任务和机能。强人工智能体具备刑事责任主体地位，可以让我们反思，是否伤害强人工智能体的行为，同样要得到处罚。毕竟，强人工智能和我们人类一样，具有情感和道德，受到伤害后可能会产生过激的行为，从而侵害其他法益。对强人工智能也施加保护，更加符合人类的利益。

5. 结论

新技术快速发展,人工智能应运而生,给人类带来巨大便利的同时,也给人类社会带来了潜在的风险。刑法作为保护人类利益的最后一道防线,也要随着时代、社会生产力的变化而不断调整其内容,适应时代的需要。人工智能作为一种快速发展且广泛应用的技术,与之相关的犯罪越来越多,“犯罪工具型”、“犯罪对象型”、“人工智能失控型”作为常见的三种涉人工智能犯罪,有些与传统的犯罪并无差别,有些结合了人工智能时代自身的特点。涉人工智能犯罪往往更具隐蔽性,波及范围、危害程度也比一般的犯罪要高出许多,这就要求刑法进行提前考量,既对涉人工智能犯罪已经表现出的风险进行规制,也对潜在的风险进行部署,达到惩罚犯罪、保护法益的目的。

参考文献

- [1] 朱新超. 涉人工智能犯罪的刑法规制研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨商业大学, 2022. <https://doi.org/10.27787/d.cnki.ghrbs.2022.000552>
- [2] 李文吉. 论人工智能时代的预防性刑法立法[J]. 大连理工大学学报(社会科学版), 2020, 41(5): 96-103. <https://doi.org/10.19525/j.issn1008-407x.2020.05.012>
- [3] 郎胜. 我国刑法的新发展[J]. 中国法学, 2017(5): 23-46. <https://doi.org/10.14111/j.cnki.zgfx.2017.05.002>
- [4] 王一霖. 人工智能犯罪刑事责任主体资格的探究[D]: [硕士学位论文]. 哈尔滨: 黑龙江大学, 2021. <https://doi.org/10.27123/d.cnki.ghlju.2021.000424>