

论刑法介入人工智能领域的时度

肖 晶

摘 要 | 最近人工智能的刑法规制成为一个热议的话题，学者们甚至热忱的讨论了人工智能刑法的一些细节问题，似乎建立人工智能刑法体系已经迫在眉睫。但是事实上，刑法应该等待更合适的时机再介入对人工智能领域风险的管制，以期避免阻碍人工智能技术和产业创新；维持中国在以人工智能为代表的高技术领域的“管制洼地”优势；以及防止在人工智能这样一个新兴领域制定缺乏现实基础的盲目的刑法条款。刑法介入人工智能领域管制的最佳时机无疑是等待人工智能产业发展成熟之后，而不是在其还在快速发展的初期。

关键词 | 人工智能；刑法；管制

作者简介 | 肖晶，湖北武汉人，湖北师范大学经济管理与法学院讲师，法学博士。

Copyright © 2022 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



一、刑法介入人工智能领域的基本时度

近两年来，刑法学界对人工智能问题的兴趣突然大增，笔者以“人工智能+刑法”为关键词，在知网中搜索，发现近两年学界集中发表了121篇以人工智能的刑法问题为主题的文章。这些文章，有的主张要创制新的罪名和刑罚，来规则人工智能领域的犯罪行为^[1]。有的主张重构刑事责任理论，将具有独立意识的人工智能作为刑事责任的主体^[2]。有的主张在现行的刑事立法框架下，通过对事故责任犯罪和过失犯罪扩张解释，来规制人工智能领域犯罪^[3]。

从这场讨论中不难发现，刑法学界对人工智能风险的刑法规制的研究有很强的紧迫感。学者们认为对没有独立意识的弱人工智能犯罪风险的刑法规制已经迫在眉睫；即使对拥有独立意识的强人工智能的犯罪行为的刑法规制也是在可预见的未来必须

[1] 王燕玲：《人工智能时代的刑法问题与应对思路》，《政治与法律》2019年第1期。

[2] 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》，《法学》2018年第1期。

[3] 姚万勤：《对通过新增罪名应对人工智能风险的质疑》，《当代法学》2019年第3期。

加以考量的事情^[1]。在人工智能时代应当“秉持一种前瞻性的刑法理念”，提前预测人工智能时代可能发生的刑法风险，做好刑法应对^[2]。刑法学界这种紧迫感源自法学家面对飞速发展的科技产业的社会责任与学术担当。但是，却忽视了对刑法过早介入人工智能领域可能带来的风险。对何时才是刑法介入人工智能领域的最佳时机这一问题，缺乏深入的反思。

刑法一旦介入某一社会领域的管理，必然意味着要在这一社会领域设置特定的行为禁区 and 边界，区分可以自由实施的行为，和立法者严令禁止的行为，并且为被禁止的行为配置严厉的刑罚。这种以刑罚为后盾的行为禁区，本质上是一种极其严苛的国家管制^[3]。刑法要介入传统法律经验有失灵风险的新兴科技领域，必须保持足够的谨慎，才能将损害该领域创新发展的风险降到最低。因此，刑法介入人工智能领域的时机，就成为一个值得慎重考虑的问题。

要寻求刑法介入人工智能领域的最佳时机，必须平衡人工智能产业发展对创新和自由的追求、社会对预防人工智能犯罪风险的管制需求和刑法在人工智能这一新兴领域发挥效能的极限这三个要素。

人工智能作为一个新兴产业，对创新的追求要高于其他任何需求。而要实现产业创新能力最大化，就必须建立一个最大自由，最少管制的产业环境。因此，人工智能产业自身发展需要创新和自由。但是，人工智能技术确实存在被用于犯罪活动，侵害法益和社会秩序的外部风险，为了控制人工智能犯罪风险，社会要求对人工智能产业予以必要的管制。其中最严厉的管制方法无疑是利用国家刑罚权禁止人工智能领域某些特定活动。所以，从社会稳定和预防人工智能犯罪的角度，当然是刑法介入管制越广泛、严厉，越有利。最后，人工智能领域中刑法效能能够发挥的极限，本质上涉及刑法在人工智能领域的有效性和经济性问题。例如哪些行为（活动）适合作为犯罪来处理，哪些智能主体能够成为刑法上的主体，哪些手段能够作为刑罚适用。如何划定犯罪、主体和刑罚方式的范围，才能保证打击该领域犯罪效能的最大化，同时又不超过刑事司法体系的能力极限，不过度消耗司法资源。这三个要素的消长，决定了刑法介入人工智能领域的时机。

由于这三个要素介入时机一定是在长时间动态

平衡中不断试错来寻找的，不可能有一个明确的时间节点。唯一能够确定的事情是，无论现在还是可预见的将来，都不是刑法介入人工智能领域的最佳时机。因为目前正值人工智能产业发展初期，其自身发展对创新和自由的需求远比对建立规制的需求强烈；再者，社会公众对人工智能产业的态度也是非常乐观的，更愿意享受其带来的经济社会发展利益，而对其社会风险并没有过分担心。在这样一种产业和社会氛围下，刑法过早介入人工智能领域，大肆强化其管制，不利于人工智能产业发展，也违背了国民的期待，与刑法的谦抑品格格格不入。另一方面，人工智能产业发展不成熟，很多风险和危害行为很可能随着人工智能技术和产业发展而迅速自动消失，不具有持久性，不值得刑法立即介入加以处理。且该领域的自发规则发展并不完善，刑法贸然介入，会有可能导致刑事立法与人工智能领域自发规则冲突，损害刑法的运行效能，甚至导致刑法被自发规则阻格，而难以实施。

二、刑法介入人工智能领域的时间限度

（一）防止刑法阻碍创新

如前所述，刑法本质上是一种极其严苛的国家管制手段，施于以安定和秩序为主要诉求的领域（比如传统的财产权、人身权保护），可以收到较好的社会效益；但是如果施之于人工智能这类以创新为主要驱动力的社会部门，则有可能窒息创新活动，迟滞技术发展，带来阻碍社会发展的负效应。

刑法在人工智能领域对创新的伤害表现在增加新型人工智能开发的法律风险，阻碍新型人工智能的设计、开发、投产和应用；增加新型人工智能开发者、生产者、销售者和使用者的合规审查成本。

[1] 刘宪权、林雨佳：《人工智能时代技术风险的刑法应对》，《华东政法大学学报》2018年第5期。

[2] 刘宪权、房慧颖：《涉人工智能犯罪的前瞻性刑法思考》，《安徽大学学报（哲学社会科学版）》2019年第1期。

[3] Posner R A, "An economic theory of the criminal law," *Journal of the American Pharmaceutical Association* 85 (1985): 1193.

从而抑制新型人工智能产品的开发、生产与应用。具体而言：

通过刑事立法追究设计者的刑事责任，无疑会导致设计者在创新人工智能产品设计的过程中自我抑制，为了规避刑法风险，将原本处于处罚模糊地带或者虽然能带来巨大社会利益但是有可能触犯刑法的人工智能新想法、新设计直接掐灭在摇篮阶段，直接导致社会失去获得提高生产效率的新技术的机会。将人工智能作为刑事法律主体，会迫使设计者为了规避刑事法律风险，投入额外的时间、精力开发各种规范人工智能行为的系统，导致成本增加或者投产周期延长，导致新技术的运用成本和投产周期增加，阻碍技术发展。

追究使用者责任同样会对人工智能技术创新造成妨碍。如果使用人工智能技术不仅不能减轻使用者的注意义务，反而会增加额外的负担和刑法风险，使用者就不会愿意购买这样的技术，无疑会对人工智能技术的推广产生不利影响。最典型的例子就是汽车自动驾驶技术。如果使用自动驾驶技术，不能减轻汽车驾驶员在行驶中的注意义务，反而要求驾驶员在自动驾驶模式下仍然要时刻对交通事故的发生保持警惕，一旦在自动驾驶模式下发生了交通事故，仍然要追究驾驶员的过失犯罪责任。那么绝大部分驾驶员肯定不会乐意为自动驾驶汽车买单，自动驾驶技术的推广就会受到影响。

笔者认为，要避免刑法对人工智能领域创新的阻碍，最有效的方法就是推迟和减少刑法在该领域的介入和干涉。那种寄希望于刑法能够谨慎介入，有效控制刑法在人工智能领域的妨碍作用的观点过于天真，因为如前所述，立法者不是全知全能的主体，不可能事先预料到人工智能可能取得突破的领域。立法者有可能认为他只在某一个很小的问题上设置了障碍，但是这个小小的限制可能就把人工智能领域一个惊天动地的突破掐灭在摇篮之中。因此，在人工智能发展还不成熟，还需要该领域从业人员和全社会共同大胆尝试，不断试错的时代，刑法最好的态度就是待在一旁观望。等到该领域发展成熟起来，创新的作用开始逐渐让位于规制确立和利益分配的时候，刑法再介入反而是一个比较合适的时机。

（二）避免丧失“制度洼地”优势

现代经济的竞争，说到底是科技产业的竞争，

而且这种竞争不是单纯科研能力的比拼，而是对新技术从研发、生产到销售的全产业链的掌控能力的比拼。在资本和技术全球化的时代，特定产业会随时逃离束缚其发展的地区，向更利于产业发展的“洼地”转移和富集，从而为这些“洼地”创造机会和财富^[1]。我国改革开放初期承接欧美国家产业转移，创造中国奇迹就是一个明显的例子。在这些洼地中，有一类称为“制度洼地”，既制度对产业的管制最为松弛的地区，对管制特别敏感的行业会从管制严格的地区向这一类洼地转移。

人工智能这一类依赖创新能力的产业对管制就特别敏感。如果某一地区管制特别严格，伤害到人工智能产业创新，人工智能产业就会寻找管制较为宽松的“制度洼地”，为自身发展创造良好的外部条件。例如现代药物研发就逐渐向实验动物权利保护比较宽松，药物实验伦理管制较为松弛的地区集中^[2]。而我国在科技产业创新管制方面，较之西方国家具备较为宽松这一天然优势。完全能够成为人工智能等高新产业所寻求的“管制洼地”，在新的一轮产业竞争和产业转移中，成为人工智能为代表的新型产业的聚集地，从而在新的产业竞争中一举超过其他国家，实现后发先制的比较优势。

因为一方面我国科研伦理和科研法律规范发展都较晚，对科研行为的限制没有其他国家那么严格。这就使得我国以人工智能为代表的高科技产业能够在研发、生产和运用方面获得更多的自由，不会像其他国家那样受到法律和伦理规范的过度束缚。另一方面，我国民众对科学创新活动有着更高的同情和支持度，对科研活动更看重其带来的生产和生活便利，更热衷于享受科研成果对生活质量的提升；而对科研活动所带来的伦理问题和社会风险，持一种更开明，更宽容的态度，较少的向西方消费者那

[1] Brandt L. Thun E, "The fight for the middle: upgrading, competition, and industrial development in China," *World Development* vol. 38, no. 11 (2010): 1555-1574.

[2] Benatar S R. Singer P A, "A new look at international research ethics," *Bmj* vol. 321, no. 7264 (2000): 824-826.

样去主动审查产品背后的科研伦理行为和潜在危害后果。这使得我国科研活动的开展,较之其他国家更少社会舆论的阻碍,科研成果的推广也更加容易。同时,这也意味着我国在以人工智能为代表的新兴科技领域进行立法管制的需求,远没有其他国家那么急迫。总之,我国法律和社会较之其他国家对于新兴科技领域的管制和审查没有那么严格,对创新行为的限制较少,有利于以人工智能为代表的新兴科技成果在我国研发、生产和投入使用,这会成为我国在新一轮高科技产业竞争中不可多得的一项优势。

但是我国要在纷扰世事中保持这一优势却并不容易,实现这一点要求立法者在采用刑法这类强管制措施的时候慎之又慎,不要受到本国舆论和国际压力的影响。对于人工智能发展中遇到的刑法问题,总的立法原则应该使宜慢不宜快,宜轻不宜重。其他国家由于技术法律体系相对比较完善,在人工智能产业发展还不成熟的时候,其刑法就开始对相关研发、生产、运用行为进行规制。面对这种状况,我国立法者不宜跟风,反而应当推迟相关立法,利用时间差,为我国人工智能创新发展创造相对宽松的监管环境。人工智能产业一旦感到这种刑事管制上的区别对其创新活动的不同影响,就会逐渐退出那些刑事管制严格的地区,向我国集中,以求得相对宽松的管制环境。那么我国就能够利用各国先进的发展成果,推动我国人工智能产业发展弯道超车。等到我国人工智能领域逐渐发展成熟,创新不再是产业驱动力的时候,我国立法者可以在这一领域重新补上刑法规制这一课。

另一方面,面对那些必须通过刑事立法手段规制的人工智能犯罪行为,也应当比其他国家在立法层面要宽松,法网要稀疏,刑罚要轻缓。相同性质的行为,能够不纳入刑法调整的,尽量不要纳入刑法调整。例如各种人工智能产品致害行为,能够使用民法或消费者权益保护法调整的,尽量不要纳入刑法规制。对于纳入刑法规制的人工智能犯罪行为,其刑罚也应当较其他国家类似行为更轻,其行为构成的范围也应当尽可能更窄。如此,才能体现我国保护和鼓励人工智能产业发展的诚意,吸引更多相关研发、生产和投资平台到我国开展创新活动,推动我国人工智能产业形成超越其他国家的竞争优势。

(三) 避免立法过于超前失去现实基础

立法者的认识能力存在局限性,不可能准确预测人工智能为代表的新兴科技产业的发展趋势。科技发展史一再表明,科技产业的发展带有明显的偶然性,不可能服从人类的计划和安排,也难以通过现有的趋势加以预测。立法者难以预测人工智能产业发展的趋势,更不可能预料到与人工智能产业相关的犯罪行为会以何种方式存在,又如何影响社会生活。立法者如果仅仅基于眼下对人工智能产业发展前景盲人瞎马似的认识,就匆忙设计出相关刑事法律制度,最终只会失去其存在的现实基础,而沦为无法落实的“纸面上的法律”。

立法者无法预料到人工智能技术的发展程度,不能理解在未来人工智能是如何产生、存在和发挥作用的,也就不可能在刑法上准确定义人工智能主体,更不用说为其设置特定的刑罚。例如有学者主张给人工智能主体设立一种“永久销毁”的刑罚措施^[1]。但事实上,基于互联网的特性,和人类的技术能力,想要完全彻底的销毁一种网络程序是不可能的。不要说人工智能,即使面对没有“智能”的计算机病毒,也没有那个机构敢于宣布彻底销毁了一种特定的计算机病毒在网络中的所有备份。因此,“永久销毁”这种类似于“死刑”的刑罚对于人工智能技术是没有意义的。在人工智能技术真正发展到强人工智能之前,任何立法者都不可能依赖在书斋中苦思冥想设计出的一套“人工智能刑事主体”制度和符合人工智能特性的刑罚方法。只能等到人工智能技术发展成熟,将它的面貌完全展现在立法者面前,立法者才会有足够的现实依据去建立这些法律制度。

立法者也不可能预料到人工智能发展过程中哪些风险需要刑法特别关注,哪些风险能够通过技术发展来解决,不需要刑法特别的关注。有可能当下觉得必须通过刑法立法加以解决的那些问题,比如具有反人类意志的“邪恶”人工智能问题等,在人工智能技术发展成熟之后,完全可以通过技术手段加以解决,不需要刑法特别的关注。核电技术便是如此。核电安全更多依赖安全保障技术的发展和安

[1]王燕玲:《人工智能时代的刑法问题与应对思路》,《政治与法律》2019年第1期。

全操作规范的完善,没有任何一个国家在建立核电站的时候,会把核电安全的希望,寄托在刑罚的威慑作用上^[1]。也许等到人工智能技术发展成熟的时候,立法者会发现,人工智能安全问题极有可能和核电安全一样,基本依赖技术手段就可以解决。立法者既不需要为了人工智能制定专门的刑事法律来保障其安全,也不需要改革现有的刑事立法体系和法律理论来适应人工智能产业的发展。仅仅将现有的职业过失犯罪和责任事故犯罪体系自然的扩展到人工智能领域,就能解决人工智能安全问题。

法律是从现实社会生活中发现和归纳的准则,而不是立法者超人理性设计的产物。对于人工智能法律这类缺乏历史经验借鉴的新兴科技领域立法尤其如此。立法者试图在人工智能产业发展成熟之前就制定出一套合理有效的刑事法律制度,防患于未然,提前做好应对人工智能犯罪的法律准备。这一想法归根结底反映了立法者理性的无端傲慢和对人类社会自发秩序的无理猜疑。立法者不相信人工智能领域能够依靠自发生成的规制和秩序,避免想象中的风险发生;立法者认为其理性设计远胜人工智能领域自发形成的秩序,足以防范人工智能发展中各种可能毁灭人类的风险。但是历史已经无数次证明,人类理性设计不仅无法防范风险,反而往往成为风险的源头;只有在人类社会自然进程中发展起来的自发秩序和自在规制才是规范人类行为,防范技术和社会风险的良药^[2]。

笔者认为,立法者应该放下对人工智能产业自我纠错能力的无端忧虑,收起对法律理性的盲目乐观,重拾对现实生活和社会自然秩序的敬畏之心。在以人工智能为代表的新兴科技领域,充分尊重刑法的谦抑性,等待人工智能产业自身发展成熟,建立起自身的一套规制之后,再通过立法者特有的理性思维,将其中重要的规则内容吸纳到刑法规范之中,使得刑法立法在人工智能领域运行无碍。

三、刑法介入人工智能领域的时机选择

在可预见的未来,刑法都不宜介入人工智能领域。但是,这并不意味着人工智能领域永远成为刑法发挥作用的禁区。随着人工智能技术和产业不断发展,刑法也应该逐步介入该领域,规制该领域犯罪行为,预防该领域的风险。而刑法介入人工智能

领域的时机,就在于人工智能产业度过发展初期阶段,逐渐走向成熟。到达这一阶段,刑法再介入人工智能领域,不仅不会妨碍人工智能产业的发展,还会收到良好的法律效果和社会效果。

从人工智能产业自身发展角度而言,随着人工智能技术发展日益成熟,行业整体盈利能力下降,各主体间关系由共同扩大生产,做大蛋糕向争夺利益,分配蛋糕的方向转变。创新不再是人工智能产业发展的决定性要素,产业整体的发展方向由锐意进取转向持重守成,早期的无序竞争状态向有序竞争转型。行业中各主体该产业内部也会由注重创新与活力,追求更少的管制和更多的行为自由;转而强调产业内部利益分配和格局调整,注重行业秩序的建立和行业长远稳定的发展。行业会更加重视规则的建立,以抵抗在利益分配过程中出现的不正当竞争行为,确保行业长远健康发展。

通过上述分析不难看出,刑法在人工智能产业发展相对成熟之后再介入人工智能领域对产业发展有很大益处。成熟之后的人工智能产业,管制对产业发展的伤害就不如其成长期那么严重,刑法在此时介入,不会对产业发展产生重大的伤害。刑法在成熟期后介入,选择行业法规中值得附加刑罚者,添设罪名或者将其解释到现有条文之下,有利于帮助人工智能产业确立行业规则的权威性,建立和维护其行业秩序。不仅不会妨碍人工智能产业发展,还对其产业发展有促进作用。

从社会效果而言,在人工智能还没有发展到足够普及的时候,民众对人工智能危害的担忧,仅仅停留在对人类命运的悲观想象,这一类抽象命题之上,相反民众从人工智能发展中获得利益却是现实的,可触摸的。只有到人工智能的应用,普及到如移动互联网技术一样,与民众的生活密不可分的时候,它带给民众生活和其他利益的伤害才会逐步暴露出来,当民众发现经济法和民商法不足以在自身陷入危险时提供足够的救济,才会逐步转向支持对人工智能风险进行刑法上的控制。

[1] Stanton. Neville, *Human factors in nuclear safety* (Taylor & Francis: Cre Press, 1996), pp. 121-123.

[2] Leoni B, *Freedom and the Law* (Liberty Fund Inc, 2012), pp. 45-46.

从政府管制的角度而言,刑法介入也应该在出现了行政措施、经济措施等管制手段无法应对的威胁之后才被期许。例如对于互联网和信息技术的刑法规制,就是一个政府不断发现管制措施失灵,并且不断制定新的刑法条款以期堵塞漏洞的过程。因此,才出现了信息刑法由最早的破坏信息网络犯罪一步一步扩展到包括侵害个人信息和利用网络散布虚假信息滋事等多种犯罪行为的刑法体系。这种逐步扩张的立法方式,体现了政府在网络管制领域的审慎态度,即在新技术发展过程中先给出充分的自由,再尽量求助伤害较小的行政、经济、民事手段加以管制,最后万不得已才动用国家刑罚权加以干预。这种审慎态度,在事实上平衡了网络信息技术创新发展和预防其风险的管制需求。这种审慎的态度应当延续到人工智能领域,并且成为人工智能刑事立法的一个原则,只有在人工智能领域出现了其他干预措施无法控制的风险,国家刑罚权才值得发动。

从刑法立法者自身而言,等待人工智能产业成熟再介入,能够避免刑法失灵的很多尴尬。一方面,随着人工智能产业逐渐成熟,产业活动中的行为模式会逐渐定型,相关犯罪的行为模式也会逐渐稳定下来,不会再出现那种迅速产生又迅速消亡的犯罪行为类型。立法者不必紧跟在变幻莫测的犯罪手段后面疲于奔命,刑法也不会存在法律刚刚颁布,规制的犯罪行为却已经趋于消亡的尴尬境地。相反,立法者可以从容地将已经定型的少数几种性质严重的侵害行为认定为犯罪。既维持了刑事立法的稳定,保障了国民对人工智能领域行为自由的预测可能性;又能准确的打击人工智能领域最严重的犯罪行为,起到维护行业秩序,护航行业发展的作用。

另一方面,随着人工智能产业不断发展,行业规范会逐步建立和发展起来。针对人工智能领域的各种民事纠纷和行政监管活动也会日渐频繁,民法和行政法会率先介入,形成民法和行政法规范。这些规范在人工智能领域发挥调节作用的同时,也受到人工智能领域实践发展的检验,其中不适合人工智能领域发展或者失去调节作用的规范会在实践中

被修改或淘汰,最终留存下来的各种规范,对人工智能领域具有相对较好的规范作用,能够取得较为有利的行业效果和社会效果。这些规范,能够为刑事立法提供必要的指导和借鉴,立法者只要选择其中亟需严厉的管制措施,适合刑法调整的内容吸收到刑法之中,就能事半功倍的建立人工智能领域的刑法规范体系。即节约了刑事立法资源,减少了人工智能这一新兴领域中刑事立法的论证负担和不确定性;又能体现刑法作为第二序位调整规范,保障人工智能领域其他法律和社会规范顺利实施的保护法地位。

四、结语

时至今日,人工智能领域仍然没有取得足以给生活带来实质性改变的商业化应用成果,就是曾经被寄予厚望的自动驾驶技术,也仍然“在路上”。人工智能产业仍然是一个脆弱的新兴产业。更多的保护、更少的干预和限制,才能有效地保障人工智能产业的突破尽快到来。而过多的限制,无疑会人为制造产业发展的困难,阻碍产业创新的实现。

有鉴于此,刑法应当在介入人工智能这类新兴科技产业的时机选择上更加务实和审慎。在人工智能产业发展还非常脆弱,尚无力承担严格管制的时候,刑法贸然介入,有可能阻碍人工智能技术和产业创新,导致我国在长期市场化过程中逐步积累起来的,高技术领域“管制洼地”优势丧失。这样的后果,无论对人工智能产业发展还是对我国实现高技术领域弯道超车的发展目标都是有害无益的。

相反,等待产业发展更加成熟之后,再择机介入对人工智能领域风险的管制。即可以继续保持我国在人工智能领域技术创新方向的比较优势,又能够防止在人工智能这样一个新兴领域制定缺乏现实基础的盲目的刑法条款。无论对国家技术发展还是对法律制度建设都是利大于弊的。刑法介入人工智能领域管制的最佳时机应当是等待人工智能产业发展成熟之后,而不是需要继续放松管制、鼓励创新的产业发展时期。