

人工智能产品发展风险抗辩制度的理论冲突与平衡协调

胡元聪

(西南政法大学经济法学院, 重庆, 401120)

摘要: 为根据人工智能产品的特点变革我国产品责任制度以最大限度防范、降低和合理分配人工智能产品的发展风险, 确立人工智能产品发展风险抗辩制度势在必行。但是也必须对该制度面临的诸多理论, 包括正外部性与负外部性理论、成本与收益理论、安全与效用理论、权利与效益理论、公平与效率理论、激励与约束理论之间的冲突进行协调, 以实现激励人工智能产业良性发展与约束人工智能产品异化倾向之间的平衡, 缓解人工智能产品生产者和消费者之间冲突带来的张力, 促进我国人工智能战略产业的高质量发展, 进而确保国家战略与包容发展、商业利益与社会责任的平衡, 为人工智能的“中国之治”贡献力量。

关键词: 人工智能产品责任; 发展风险抗辩制度; 理论冲突; 平衡协调

中图分类号: DF414

文献标识码: A

开放科学(资源服务)标识码(OSID)

文章编号: 1672-3104(2020)06-0075-15



一、问题的提出

人工智能产业已经成为全球性的战略竞争性产业。我国在最大程度激发人工智能战略产业发展之巨大潜力之时, 也必须预判人工智能产品带来的挑战及其风险。因此, 需要在人工智能产品生产者和消费者之间寻求一种权利和义务的平衡, 以缓解双方冲突带来的张力。而产品发展风险抗辩制度作为一种创新制度, 恰好具有此项功能。产品发展风险抗辩制度大致包括基本制度如技术标准制度、时间节点制度、主体资格制度和请求时效制度等, 以及配套制度如跟踪观察制度、勤勉警示制度、产品召回制度等。但是, 在理论研究方面, 该制度在全球学界却饱受争议, 大家莫衷一是, 争论焦点的实质在于是该维护生

产者还是消费者利益。在实务运用方面, 大部分国家的法律虽然确认了发展风险抗辩制度, 但在实际应用中都严格限制它的适用。这就体现出各个国家的立法者既想阻止严格责任的绝对化, 又想防止发展风险抗辩滥用的价值追求。进入人工智能时代, 基于人工智能产业的战略地位, 人工智能产品发展风险抗辩制度再次成为热点问题。为根据人工智能产品的特点革新我国产品责任制度以防范、降低和合理分配人工智能产品发展风险, 确立人工智能产品发展风险抗辩制度势在必行。因此, 该制度的适用在人工智能时代具有特殊的意义, 但是其仍然会面临诸多的理论冲突, 需要进行平衡协调。在这些理论冲突中, 一些是基于生产者和消费者之间矛盾张力而产生, 也有一些是基于国家干预和社会稳定考量而形成。从法律价值的角度看, 人工智能产品发展风

收稿日期: 2020-03-16; **修回日期:** 2020-05-03

基金项目: 中国法学会项目“人工智能产品责任之发展风险抗辩制度研究”(CLS(2018)C45); 西南政法大学校级重点项目“人工智能风险治理的财税法机制研究”(2018XZZD-10); 西南政法大学党的十九届四中全会精神研究阐释专项课题“数据的私法规制研究”(2020XZYB-05); 智能司法研究重庆市2011协同创新中心项目“智慧庭审视域下的异步审理研究”(ZNSF2020Y03)

作者简介: 胡元聪, 四川南江人, 法学博士, 西南政法大学经济法学院教授, 博士生导师, 西南政法大学人工智能法律研究院研究员, 西南政法大学激励法学研究中心主任, 主要研究方向: 经济法学、人工智能法学、数据法学、区块链法学, 联系邮箱: 782918534@qq.com

险抗辩制度涉及多个价值目标,但是没有一个是应当得到无限承认和绝对保护的,否则会走向极端。正如一个旨在实现正义的法律制度,会试图在自由、平等和安全方面创设一种切实可行的综合体和和谐体^[1],因此就需要进行多重价值目标的平衡和协调。本文基于研究者的不同学科和不同立场,从正外部性与负外部性理论、成本与收益理论、安全与效用理论、权利与效益理论、公平与效率理论、激励与约束理论方面,阐述人工智能产品发展风险抗辩制度的理论冲突与平衡协调问题,从而缓解人工智能产品生产者和消费者之间冲突带来的张力,实现国家战略与包容发展、商业利益与社会责任的平衡。

二、正外部性与负外部性理论

(一) 正外部性^①理论

对产品发展风险抗辩制度适用与否的争论者往往持两种不同观点,即基于生产者对消费者的正外部性而主张适用,基于消费者对生产者的正外部性而主张不适用。就人工智能产品而言,其正外部性具体表现为如下方面。

一是人工智能产品生产者为消费者的正外部性。第一,人工智能产品生产者为消费者带来直接的正外部性。如通过人工智能产品生产者的生产,可以改变消费者的生活方式,从而实现智能化、便利化、精细化。易言之,智慧城市、智慧家居、智慧生活已经成为主流,智能手机、智能指纹锁、智能语音识别机器人、无人机、无人船、无人艇等产品的出现改变了传统的生活方式。正是人工智能生产者不断研发生产新的产品,才能使消费者切实享受到智能技术带来的社会福利。第二,人工智能生产者也从其他方面间接为消费者带来正外部性。党的十九大报告指出,要加快建设制造强国,加快发展先进制造业,推动互联网、大数据、人工智能和实体经济的深度融合。因此,人工智能生产者无疑是人工智能产业发展的探索者,其通过新兴技术与智能设备

促进我国传统产业升级转型,为制造业转型赋能,从而促进实体经济高质量发展。而经济发展的最终成果必然从整体上惠及消费者,从而使其成为人工智能产业发展的间接受益者。

二是人工智能产品消费者对生产者的正外部性。第一,人工智能产品消费者的消费需求是生产者积极进行创新不断改进技术的重要动力。消费者通过消费人工智能产品,将消费需求传递给生产者,生产者就会积极研发,改进技术,进一步迎合消费需求并从中盈利。因此,从消费决定生产和消费促进生产的角度来看,消费者会刺激生产者的生产动力,给其生产上的盈利心理带来显著的正外部性。第二,消费者对人工智能产品的巨大消费需求会客观促进生产者积极改进产品质量和性能。随着人工智能产品日益普及,人工智能产品生产者越来越多。而消费者的消费喜好会推动诸多生产者尽力提高自己的产品效用以相互竞争,逐利市场。这种健康的竞争会促使生产者为了吸引消费者而不断加大投入,在竞争中持续提升产品质量,以更好地提升消费者的体验,增加消费者的信任。因此,消费者对人工智能产品的消费客观上会带来促进生产者通过合法竞争积极生产更多更好产品的正外部性。

(二) 负外部性理论

与前述相对应的是,学者们基于生产者对消费者的负外部性而主张不适用,基于消费者对生产者的负外部性而主张适用。就人工智能产品而言,其负外部性具体表现为如下方面。

一是人工智能产品生产者为消费者的负外部性。第一,消费者的隐私容易遭到泄露,导致隐私权被侵害。因为人工智能产品深度依赖大数据、互联网和区块链技术,其在被使用中可能使消费者的个人隐私更容易被收集、泄露、窃取、贩卖和利用,在生活中被网络攻击而致权益受损。如商家已经充分利用商业信息数据作为赚钱工具,各种建模和算法很可能巧妙地“将你的生活转化成他人的商机”^[2],甚至“某些老练的法人实体可能正在利用这些算法为自己谋利”^[3]。第二,消费者容易遭受技术性歧视,导致平等权被侵害。人工智能产品在被使用过程中会不断搜

集个人信息，基于这些信息可能导致人工智能在决策中产生算法歧视、算法偏差甚至算法错误的问题，其不确定性的危险在于没有人能确保该算法设计准确。这些违背“算法正义”价值的行为会导致特定人群受到不公正的待遇。同时，代码编写的背后是生产者的商业利益在推动，而消费者则处于话语权缺失状态和弱势地位，只有选择“yes”或“no”的权利，这无疑会或隐或显地制造和加剧社会等差与不公平，从而偏离了“代码正义”的价值目标^[4]。第三，人工智能将使消费者面临更多威胁与未知风险，导致知情权被侵害。如在法律适用领域，法律案件预测的弊端在于可能扭曲当事人(人工智能产品的特殊消费者)的诉讼行为，带来新的偏见，比如算法黑箱。即人工智能的算法本身就面临诸多“技术陷阱”^[5]或“算法后门”，甚至被一些研究者认为是在黑箱中运作。因此，人工智能不成比例地强化了一部分企业的力量，使其成为站在人工智能发展前沿的大数据和算法技术掌控者，同时也使越来越多的人变成了难以保护自己隐私和自由并面临失业风险的弱者^[6]。并且，随着科技的加速进步，强弱悬殊越来越大，强者为所能为、弱者受所必受的局面，就会成为普遍现象^[7]。这样，本来就是弱者的消费者的地位将变得更加弱势，导致“传统的消费者不会回来了”^[8]。

二是人工智能产品消费者对生产者的负外部性。第一，生产者的科技创新热情可能遭受打击。因为人工智能产品生产者的初期发展本身就面临行业发展风险大、投入高而产出慢的问题。如果消费者滥用责任诉讼，过度主张自己的权利，会给人工智能产品生产者带来巨大的压力，使其在生产研发人工智能产品方面瞻前顾后，趋于保守，而不敢大胆地投入新产品的研发生产，如此就会打击创新热情。第二，生产者的生存环境可能遭遇变故。针对人工智能产品，如果坚持按照无过错责任原则来处理产品缺陷造成的损害，会使人工智能产品生产者很容易面临巨额赔偿，甚至会因多数赔偿案件而导致破产。这样会导致人工智能产品生产者怠于开发新型技术生产科技产品，从而最终阻碍社会良性进步和人类

多元文明的形成。人工智能产品有其特殊性，生产者没有能力及时预见某些缺陷，即便尽到了足够谨慎义务，也很难完全了解复杂的算法过程。在这种情况下，规定生产者不得主张抗辩，必然会恶化企业生存环境。第三，生产者的融资环境可能遭受不利。例如消费者过度强调自己的消费权利，会引起较多责任诉讼。而生产者又容易因为无过错原则承担巨额赔偿，这种情形也会给人工智能产品生产者的融资环境带来负外部性。因为投资人会基于人工智能产品生产者容易因为巨额赔偿破产的风险而不愿意进行风险投资。如此一来，必然会影响到生产者的融资业务，从而不利于其吸引投资，扩大生产。

(三) 正外部性与负外部性理论的冲突与协调

随着科技与社会的交互程度不断加深，典型的科技创新在促进社会进步的同时产生了许多负外部性影响^[9]。进入人工智能时代，人工智能产品生产者会对其消费者产生正外部性和负外部性。反过来，人工智能产品消费者也会对生产者产生正外部性和负外部性。而主张适用产品发展风险抗辩制度者看重的是生产者对消费者的正外部性和消费者对生产者的负外部性，反对适用产品发展风险抗辩制度者看重的是生产者对消费者的负外部性和消费者对生产者的正外部性。因此，要让人工智能产业向着符合社会发展的需要方向发展，就需要通过合理适用产品发展风险抗辩制度，对人工智能产品生产者与消费者之间的双向正外部性进行法律激励，同时对人工智能产品生产者与消费者之间的双向负外部性进行法律约束，从而平衡人工智能产品生产者与消费者之间的利益，最终促进人工智能产业的健康发展。

其一，通过法律对人工智能产品生产者和消费者的双向正外部性进行激励，从而鼓励技术创新以实现平衡。一方面，为了给生产者发展人工智能产业保驾护航，应该积极发挥消费者对生产者显著的正外部性激励作用，从而激发人工智能产品生产者的创新积极性，引导生产者正当竞争，改进技术，优化产品，促进人工智能产业良

性发展,不断从国内市场走向国际市场。另一方面,还要充分发挥生产者对消费者的正外部性影响,使人工智能产品等创新成果真正惠及全体人民,不断提高人工智能产品消费者的消费水平和质量,增进消费体验,提升消费幸福感。因此,适用人工智能产品发展风险抗辩制度会起到包容和激励生产者技术创新的作用。比如针对区块链技术,凯文·沃巴赫就认为,过度或不成熟地适用严格的法律义务都会阻碍创新,摒绝利用技术达成公共政策目标的机会,对于像区块链这些新兴技术,努力洞察、善加利用才是最为明智的态度^[10]。反过来,对生产者合适的激励又能对消费者的消费起到激励作用。

其二,通过法律对人工智能产品生产者和消费者的双向负外部性进行约束,从而确保人工智能产业健康发展以实现平衡。一方面,要通过出台适当的法律制度充分保障人工智能产品的安全,如强化算法透明原则,消除算法歧视和算法霸权,增进人工智能产品生产者的社会责任意识等,同时,合理适用产品发展风险抗辩制度,弱化并尽可能消除人工智能产品生产者为消费者带来的负外部性。如学者所言,在自动驾驶汽车行业,如果不要求自动驾驶汽车的生产者承担符合公平正义观念的责任,消费者可能遭受负外部性,其会怀疑生产者为了产业利益而将未臻成熟的自动驾驶汽车推向市场,使其购买或使用自动驾驶汽车的信心受到打击,从而反过来又给生产者带来负外部性,导致消费者对自动驾驶汽车采取敬而远之的态度,这会影响产业的发展^[11]。另一方面,要通过进一步完善相应的法律制度,适度实施人工智能产品发展风险抗辩制度,限制适用严格责任原则,防止消费者滥用责任诉讼,减少消费者对生产者的不当负外部性,从而促进人工智能产业的健康发展,为我国发展人工智能战略产业保驾护航。

三、成本与收益理论

(一) 企业利益视角的成本与收益理论

毋庸置疑,人工智能产品生产者在经济活动

中必然会进行成本与收益的分析,而产品发展风险抗辩制度的适用与否直接决定着其成本与收益的大小。成本与收益是经济学中的一对典型概念,也是典型的经济分析工具。在市场经济条件下,任何一个经济主体在从事经济活动时,都要从投入与产出关系进行一个尽可能科学的评估,从而对具体经济行为的经济价值进行衡量。易言之,企业从事市场经营活动带有强烈的自利性,其以追求利润最大化为目的,因此进行充分的成本与收益分析是企业进行市场活动的前提。企业往往会预先以货币单位为基础对投入与产出进行预判,从而提前计划,合理安排投入。利润则是指总收益与总成本之间的差额,根据经济学研究规律可知,利润最大化原则指的就是边际收益等于边际成本。因为若边际收益大于边际成本,企业就会增加产量继续盈利,若边际收益小于边际成本,则意味着亏损,企业就需要调整产量。就人工智能产品研发生产而言,当企业可以持续盈利时,企业才会进一步加大研发力度并改进技术,从而提高产品质量。易言之,只有在私人收益大于私人成本时,企业的生产研发动力才会进一步被激发,人工智能产业才能够持续发展。然而,人工智能产品牵涉太多大数据及算法问题,研发风险较高,若不允许企业进行产品发展风险抗辩,意味着企业必须投入更多的私人预防成本来防止其无法预见的风险,从而增加企业的私人边际成本投入,同时还要承担事后产品失败的赔偿责任,最终可能导致企业因为风险太大,投入过高,无法达到利润最大化而只有进行保守研发。如此导致的结果是不利于激发人工智能产品生产者的研发动力,从而影响该产业的良性发展。因此,基于利益的考虑,企业肯定支持适用产品发展风险抗辩制度。

(二) 国家干预视角的成本与收益理论

国家在运用制度干预经济活动时,也必然进行成本与收益的分析,而产品发展风险抗辩制度的适用与否直接决定着社会成本与社会收益的大小。法律经济学是以经济学的理论和方法去研究法律规则、制度、创新及效率的边缘学科。成本收益也是法律经济学中常用的概念,波斯纳曾

说“法律的经济分析”主要表现为用成本与收益理论分析法律制度、法律适用及法律行为的成本与收益。从人工智能产品发展风险抗辩制度本身来分析，该制度主要体现了国家基于整体社会成本与收益考量而进行法律制度干预。具体可以从两方面进行分析：一是采用严格责任而不允许进行产品发展风险抗辩对生产者影响的分析。在此理论指导下，一旦出现人工智能缺陷产品致害，生产者必须承担责任，则其会增加保险购买数额而降低尽力预防风险发生的动力。易言之，只要出现人工智能缺陷产品致害就必须承担全部责任，生产者就会通过购买大量保险而减少预防的私人成本投入。这样，最终市场上反而更容易出现缺陷产品。如此一来，生产者虽然通过保险转移了风险成本，但缺陷产品流入市场后却增加了整体社会成本(如社会治理成本)。若国家基于成本与收益的考量，通过更加合理的制度措施进行干预，如确立人工智能产品发展风险抗辩制度，那么生产者有了免责机会，就会恪尽职守，采取各类预防措施。因为一旦允许适用免责抗辩，生产者需要承担关于“预防达到科技标准”的举证责任，从而产生一系列举证成本，这样反而会激励企业认真做好质量检测，积极预防风险，从而减少缺陷产品的出现机会，最终降低社会成本。二是支持适用发展风险抗辩制度对消费者影响的分析。如果支持人工智能产品发展风险抗辩制度，则会促使消费者在使用人工智能产品过程中尽到足够的谨慎注意义务。因为，一旦出现人工智能缺陷产品致害就由生产者承担责任，消费者就可能不会主动履行使用过程中的谨慎注意义务，由此可能带来更多的消费事故。反之，允许生产者享有产品发展风险抗辩权就可能改变这种现状。一旦消费者因为自身使用不当而受到损害，就有可能承担不利诉讼后果，如此一来，也会减少社会成本。综上，确立国家干预视域下的人工智能产品责任发展风险抗辩制度，体现了对社会成本与收益的考量，国家适度的法律制度干预确实可以从整体上减少社会治理成本，从而降低预防成本，最终增加社会整体收益。因此，基于干预视角的考虑，国家支持修正适用产品发展

风险抗辩制度。

(三) 企业利益和国家干预视角的成本与收益理论的冲突与协调

人工智能产品的生产与消费必须在考量如何适用产品发展风险抗辩制度的基础之上，解决私人和社会视角下成本与收益的冲突问题，从而实现平衡协调。成本收益分析法现在被广泛应用于各大领域，用来理性分析某一事件或行为的发展路径，从而找出最优方案。但成本与收益本身有价值冲突的一面。虽然都是成本与收益，但是成本与收益也有私人与社会之别，即存在私人收益与社会收益，私人成本与社会成本。因此，私人与社会的成本与收益的冲突体现为，企业利益视角主要是基于私人收益与私人成本考虑，国家干预视角主要是基于社会成本与社会收益考虑。是故，国家干预视角下的产品发展风险抗辩制度的适用旨在实现社会成本与私人成本的平衡，以及实现社会收益与私人收益的平衡。因为，当产生社会成本时，就需要公法的介入。具体到人工智能产品领域，有如下表现。

其一，生产者在研发生产人工智能产品时会增加私人成本投入，需要制度进行平衡。生产者的私人成本包括技术成本、设备成本以及检测成本等创新成本和常规成本。精细的私人成本投入必然会对产品的性能以及安全性做到最大程度的保障。但私人成本投入过高，必然会降低企业的私人收益，尤其是让生产者对其完全无法预见的产品发展风险或者缺陷投入过多的私人成本进行预防，必然会给生产者带来不小的成本负担，从而降低私人收益。因此，发展风险抗辩制度的适用可以适当解决生产者此时面临的成本过高的窘境。易言之，该制度的适用除了在生产者与消费者之间实现利益平衡之外，也能够确保在生产者内部达到私人成本与私人收益方面的平衡，从而实现企业利益的目标。

其二，生产者在研发生产人工智能产品时以过度追求自身利益最大化为目的，需要制度进行平衡。生产者投入私人成本越低，在其他条件不变的情况下，意味着私人收益越大。但是，过低的私人成本投入难以保证产品质量以及安全性

能,所以,在发展人工智能产业、生产人工智能产品时,需要关注私人和社会的成本与收益关系,理性协调二者之间的冲突。此时,很有必要通过私人和社会的成本收益分析,寻求最优方案,在确保私人合理收益的前提下,尽可能优化产品性能,从而保证人工智能产品为消费者带来更多福利。产品发展风险抗辩制度的合理适用,既可以平衡生产者与消费者之间的利益,也可以在一定程度上确保在生产者私人成本收益与社会成本收益之间的平衡,从而实现国家干预的目标。

四、安全与效用理论

(一) 安全理论

人的安全乃是至高无上的法律,安全是人工智能时代的核心法价值^[12]。今天,处在风险社会中的人们比以往任何时期都更加关注自身安全和社会安全。有学者认为,产品发展风险抗辩制度的适用可能降低生产者的责任,不利于保障消费者的安全,进而不利于建立一个更安全的社会。产品安全是指产品在使用、储运、销售等过程中,保障人体健康和人身、财产安全免受伤害或损失的能力。产品安全涉及消费者的生命和财产安全,而保证产品安全是生产的第一要义,也是国家、企业及社会关注的重点。易言之,产品必须具备确保基本安全性的最低标准。具体到一个国家,必须合理考虑法律该如何确立产品发展风险标准,因为立法者所树立的安全规则标准成为风险分配的重要界限,其直接涉及风险分配的公平与否^[13]。基于人工智能产品的高科技属性,其安全性更应该被强调和保障。该种强调与保障可以从国家层面、企业层面和社会层面的现有制度措施来进行分析。一是国家层面的产品安全保障。就国家层面而言,我国制定了许多法律条款保障产品质量安全以预防安全事故发生。例如我国《刑法》第146条从处罚措施角度规定了确保产品质量安全的两个方面的标准:在生产方面,包括符合保障人身、财产安全的国家标准、行业

标准,涉及的产品是电器、压力容器、易燃易爆产品以及其他产品。在销售方面,包括符合保障人身、财产安全的国家标准、行业标准,以及造成严重后果的处罚措施。我国《产品质量法》第26条第1款规定生产者应当对其生产的产品质量负责并对产品质量的具体标准,包括保障人体健康,人身、财产安全的标准进行的细化规范。该法第41条还规定了生产者应当承担赔偿责任的具体情形。二是企业层面的安全保障。就企业层面而言,人工智能产品安全也是企业关注的重点。企业通过强化质量意识教育机制、完善岗位质量负责机制、加强内部的监督机制等制度措施确保产品质量安全。企业还可以通过组建科研团队,设立内部专业检测流程进行层层把关,以保障产品质量安全。同时,人工智能产品生产企业还应该不断提高科技实力以增加科技生产能力、产品监测能力和跟踪观察能力,从而确保人工智能产品的相对安全。三是社会层面的安全保障。就社会层面而言,社会各行业组织成立相关行业协会并确立行业标准,同时设立专门监督机制以监督企业产品生产的安全性。人工智能产品研发具有风险特殊性和不确定性,而人工智能产品本身又对社会消费者影响广泛,因此,在研发生产人工智能产品时,更应该重点关注其安全性能,做好预防检验检测工作,减少安全隐患和防范科技风险。总之,安全理论重在从生产者的义务和消费者的权利视角进行考量。因此,其主张不适用产品发展风险抗辩制度,即使适用也不能对产品安全打折扣而损害消费者的权益,否则违反正义原则。

(二) 效用理论

效用理论也被称为消费者行为理论。从该理论的表面看,其是研究消费者消费行为的理论,而事实上,其对生产者的生产行为产生决定性作用。学者普遍认为,即使适用发展风险抗辩制度也要在追求产品效用最大化的同时确保产品更安全。效用可以分为总效用与边际效用,二者都是针对消费者消费商品而言的。总效用是指消费者在一定时期内消费商品所获得的效用总和;边际效用是指消费者在一定时间内增加单位商品

所引起的总效用的增加量。产品的效用是生产者在从事市场经济活动中关注的重点，因为产品研发生产对消费者产生的效用越大，资源利用率就越高，企业就更有发展动力。相反，产品研发生产对消费者产生的效用越小，资源就愈加无法物尽其用，而资源配置效率低下造成的社会资源浪费又会阻碍企业发展。所以，人工智能产品生产者是注重效用的企业。人工智能产品研发过程中涉及精密的算法及庞大冗杂的数据信息，因此企业在进行人工智能产品研发时本身就已经调用了大量的社会资源，投入了许多经济成本。同时，由于该行业不确定性非常明显，风险较高，若不适用产品发展风险抗辩制度而仍采用严格责任，要求生产者对倾尽预防手段依然不能发现的未知缺陷承担责任，则会降低生产者发展人工智能战略产业的积极性。同时，生产者经济投入过大，产出的产品对消费者效用较小，也会造成社会资源配置效率低下以及资源的严重浪费。因此，激励人工智能产品生产者发展就要考虑企业的经济投入与效用产出，从法律制度上给予激励从而降低生产风险，如此以调动企业的研发活力。效用理论重在从生产者的利润追求和消费者的福利保障视角进行考量。因此，其主张适用产品发展风险抗辩制度，通过激发人工智能产品生产企业活力，以提高社会整体效率，从而最终提高消费者的福利。

（三）安全与效用理论的冲突与协调

如何理解安全与效用理论的冲突？对生产者而言，一方面要追求确保消费者的最大化效用，这样才能使自己的产品畅销，另一方面要追求产品的安全，以免给自己带来赔偿危机。对消费者而言，一方面要追求生产者的产品能够最大化自己的效用，这样才能使自己支出较少而收益更多，另一方面又要期待产品的安全，以免产品给自己带来人身和财产安全风险。为了合理适用产品发展风险抗辩制度，体现实质的公平与正义，我们需要在其内在诸多价值之间进行协调以寻找合理的平衡点，从而构建内部统一的法的价值体系。安全与效用若能同时兼顾，必然是理想状态，但事实上二者往往相互冲突，不可兼得。

因为安全与效用成反比关系，于生产者而言，基于成本的考虑，其追求过高的安全创新必然会降低效用创新，而一味追求效用，也会使消费者安全度降低。于消费者而言，基于支出的考虑，其追求过高的安全性必然使得支出增加，而一味追求效用，就只会买到安全性能降低的产品。虽然如此，法律依然需要寻找二者之间的平衡点，因此选择适用能够使安全与效用达到均衡点的产品发展风险抗辩制度，对人工智能产品的生产者与消费者都相对公平。在产品责任领域，法律既要强调消费者安全权的享有，也要强调生产商生产积极性的保护^[14]。发展人工智能产业要做到安全与效用两手抓，既要考虑产品安全性，又要关注消费者的效用最大化，二者不可偏废。

其一，确保人工智能产业良性发展，注重产品安全需要制度协调。人工智能社会也是风险剧增的社会。风险常态化、复杂化、多样化导致社会需要安全保障方面的制度建设^[13]。随着人工智能产品的自主性不断提升，其在一定范畴内对于现实物理世界的控制也会引发民众对于安全的担忧。因此，人工智能产品安全是消费安全的基础与前提，而产品安全意味着生产者必须投入更多私人成本以及足够的谨慎度来积极预防可能发生的危险。但是，现实中存在的许多无良生产者为了追求极端利润，不择手段降低生产成本，牺牲产品安全，不仅扰乱了市场秩序，也严重损害了其他生产者的权利。因此，在激励人工智能产业发展的同时，一定要注重产品质量安全保障。要通过法律制度保障产品安全，细化法律义务与惩罚条款，提高生产者的产品安全意识，督促生产者谨慎对待产品质量安全，尽可能减少人工智能缺陷产品致害事件的发生。因此，适用产品发展风险抗辩制度时必须注意“度”的把握。尤其是，还要针对人工智能产品不安全致人和物的损害问题，积极构建赔偿机制和基金筹集机制，健全消费者保护相关制度，保护人工智能产品消费者的合法权利，确保消费安全。比如对于安全性的强调，可以鼓励而非阻碍自动驾驶汽车产业甚至是人工智能产业的发展，也只有自动驾驶汽车的安全性得到极大强调的情况下，该产

业的发展才是可持续性的^[11]。

其二, 确保人工智能产业良性发展, 注重产品效用需要制度协调。安全与效用可以相辅相成, 发展人工智能产业必须要最大程度地协调二者矛盾。一是边际效用递减原理启示生产者要不断进行科技创新, 研发出不同的人工智能产品以实现多样化, 从而促进消费者效用的增加。二是尊重消费者主权, 即生产者要根据消费者需求效用进行研发。消费者购买人工智能产品, 旨在实现效用最大化, 且产品的效用越大, 消费者愿意支付的价格越高, 此时, 生产者与消费者能够实现双赢。三是满足特殊消费者的效用需求。人工智能产品生产者在开发新产品时, 可以部分定位于特殊消费者, 根据其爱好来开发产品, 从而实现消费者效用的提升。易言之, 生产者在保护消费者安全的同时, 也要充分关注产品效用, 以保持消费者欲望对生产者创新活力和生产积极性的激励。随着人工智能产业被世界各国定位于国家战略产业, 就要求我们考虑确立人工智能产品发展风险抗辩制度, 允许生产者对于超出科技水平的产品缺陷的安全性进行抗辩免责, 减少超出生产者预防能力的损害赔偿情形, 同时考虑建立事后救济制度, 合理分配损害, 补偿无辜消费者的损失以寻求二者之间的利益均衡点。事实上, 现代侵权法还创造了社会安全保障义务, 一方面维系了过错责任的权威性, 另一方面又起到了风险分散的作用^[15]。合理适用产品发展风险抗辩制度, 在注重效用的同时兼顾安全, 达到效用与安全的最佳结合, 既促进人工智能行业的快速发展, 又维护消费者的消费安全权利, 实现生产者与消费者的利益均衡, 从而缓和二者矛盾。

五、权利与效益理论

(一) 权利理论

产品的严格责任理论根源于消费者权利理论, 是在对生产者功利主义等观点批判的基础上发展起来的, 代表人物有德沃金等。权利理论以实现个人权利为宗旨, 认为应该把个人权利作为

中心, 关心对个体权利的尊重和实现。他们同时认为公共利益要真正获得正当性和政治认同, 只有在公共利益能够满足平等地落实为个人权利时, 才能被社群成员普遍接受。德沃金是一位典型的权利论者, 他认为法官在审理复杂案件时是依靠权利理论而不是经济分析证明其决定的, 法院在决定是否采用某个学说的时候考虑的是公平而不应该是功用^[16]。德沃金崇尚自然权利, 认为个人权利是与生俱来的, 具有优先性和至高无上性, 在进行利益衡量时应该注重个人权利的保障。在德沃金看来, 不能因为可能带来社会整体的利益而牺牲个人权利, 即使政府认为取消该项权利对大多数人来讲是好的, 这种做法也不可行。因为个人权利是一种压倒性或排他性的法律理由, 足以压倒社会利益考量。因此, 他反对功利主义式的公共利益观, 为个人权利赋予了强硬意义。理查德·赖特(Richard Wright)也指出, 侵权法的理论应当奠定在康德的平等自由基础上。国家的首要目的是实现每个人的平等自由权利^[17]。而康德认为, 一个人在打算对别人作不兑现的诺言时就看得出来, 他这是把别人仅仅当作自己的工具, 而不同时把他当作自在目的^[18]。这实际上是要求生产者承担起尊重他人(当然包括消费者——笔者注)而不仅仅是将他人作为自己牟利的工具的道德义务。权利理论为实行产品严格责任提供了理论支撑。普儒瑟(William Prosser)则进一步发展了严格责任, 他提出四点理由: 消费者有权得到最大限度的保护; 受到伤害的消费者难以承担举证责任; 生产者投放产品到市场时起着关键作用; 生产者最能通过价格和保险在一般公众中分配风险。综上, 从权利理论视角来看, 针对人工智能产品的研发应该实行严格责任, 以着重保障我国《消费者权益保护法》规定的消费者的九大权利。只要缺陷产品进入市场并给消费者带来人身安全或者财产损害, 生产者就应该承担严格的赔偿责任。因此, 这里的权利理论主要是站在消费者权益保护视角来谈的, 其主张不适用产品发展风险抗辩制度给生产者免责。

(二) 效益理论

效益理论主要是以追求效益为最大目标, 其

考虑生产者的经济效益目标。因此，确定产品责任主要考虑效益最大化，认为要在有利于资源配置效率化、经济效益最大化的前提下确定产品责任。效益理论强调以功用和效益为标准去衡量责任划分。效益理论的代表为经济分析学派，他们认为人类是以“自我利益”的最大化实现者为前提而形成的，法官和立法者应该充分考虑财富效益最大化原则。成本收益分析以及相应的数学计算是经济分析学派提倡的基本方法。20世纪中期发展起来的经济分析法学将经济学定理用于法律制度及法律规则分析，进一步论证了效益理论，其中以科斯和波斯纳为代表。科斯第二定理主要谈到要以资源配置最大化为考量的重要标准之一，其提出，如果交易的费用不断上涨，那么资源配置就会产生不同的社会效益。这就要求在立法和执法时，应考虑对初始权利的界定，要有利于资源配置的效益最大化。波斯纳则提出要关注交易活动的经济效益并注重效率，他还提出资源流转定律，因为其有助于实现经济效益最大化。与效益理论紧密相关的论述还有以休谟等为代表的功利主义理论。他们认为衡量正确和错误的尺度就是最大多数人的最大幸福，其与经济分析学派追求财富最大化的观点相似。具体到人工智能产品，适用产品发展风险抗辩制度下的侵权责任机制可以协调各方利益，使各方在风险分配机制中实现最大效益^[13]。因此，在效益理论视角的前提下，纯粹的、极端的产品严格责任并不可取，应该允许生产者享有对产品发展风险进行抗辩的权利，从而考虑人工智能产品生产者的效益目标。因为在风险社会中，侵权责任作为风险分配的模式之一，侵权法成为风险分配的基本手段，必然要求侵权法在宏观制度、微观责任机制上进行调整^[13]。总之，这里的效益理论是站在生产者视角来谈的，其主张适用产品发展风险抗辩制度。

（三）权利与效益理论的冲突与协调

事实上，权利理论与效益理论之间的矛盾和张力客观存在。就人工智能产品来说，代表这两种理论的各自主体——生产者与消费者之间的利益冲突是该张力的客观基础^[16]，即这两种理论站

在不同主体的角度追求各自的价值。但是，当最高层次原则发生矛盾时，则应当相互协调、互不让步^[19]。因此，从根本上说，效益与权利的关系并非水火不容。实际上，它们既是对立矛盾又是相辅相成的关系。效益与权利都是反映不同社会主体(这里是指生产者和消费者)需要的价值，在一定条件下，权利是效益的表现形式，效益则是权利的具体内容。

其一，权利理论注重通过法律规范实现每一个社会主体的权利(这里特指消费者的权利)，需要协调以进行平衡。其把尊重个人权利放在中心位置，旨在维护公平，实现正义，保障社会稳定和秩序。这里是指，该理论更偏向于对消费者等弱势群体利益^[16]的考量。但是，毋庸讳言，此时也不能走极端，即为了确保消费者的权利就完全忽略生产者的权利(具体体现为生产者效益)，而只需要进行权利和义务的倾斜性配置，就可以适度保障消费者的权利。如果过度保护消费者的权利，最终的后果是生产者不愿意进行科技创新，产品的更新换代变得缓慢，消费者可以选择的产品也会变少，这样一来，最后失去福利的仍然是消费者。因此，从本质上来说，适度适用产品发展风险抗辩制度还是有利于消费者的，这就为缓解生产者与消费者之间的矛盾和对立提供了理论基础。

其二，效益理论运用经济学的方法研究如何提高一定社会主体(这里指生产者)的经济效益，需要协调以进行平衡。正义提出了这样一个要求，即赋予人的自由、平等和安全应当在最大程度与共同福利相一致^[1]。这就强调了赋予权利也是有限度的，必须与社会共同福利进行协调。效益理论旨在增加社会总体财富(社会福利)或一部分人的物质财富(个人福利)，这里主要是指增加生产者的物质财富，从而增加社会总体财富。就支持适用产品发展风险抗辩制度而言，该理论更加偏向于生产者一边。但是，在效益理论指导下的生产者如果将成本收益分析运用到极致，而忽视自己应当承担的社会责任，包括对消费者的责任以确保消费者的消费权利得到充分保障，那么此时的效益也不可能维持长久，最终生产者会

被市场中的消费者淘汰出局,这样受到损害的还是生产者自身。因此,适用产品发展风险抗辩制度也不能完全偏向生产者,否则从本质上来说也会损害生产者的利益。总之,权利理论与效益理论虽然侧重点有所不同,但在确定产品责任时都起到了重要作用,因此应当做到两者平衡而不走极端。适用人工智能产品发展风险抗辩制度时,也应当充分考虑权利与效益的关系,在人工智能产品生产者和消费者之间寻求适当的平衡,从而公平合理地确定和分摊人工智能产品的发展风险责任。

六、公平与效率理论

(一) 公平理论

公平理论是严格责任理论的逻辑起点之一。公平是法的重要价值之一,亦是正义的衍生物。谈及正义论,不得不提到亚里士多德、康德及罗尔斯等重要人物。亚里士多德提出的正义包括分配的正义和矫正的正义。分配的正义强调资源、权利、义务、责任之间的分配公正。当分配的公正原则被违反时,矫正的公正就开始发挥作用。就产品缺陷责任而言,生产者明知产品存在缺陷依然违反规定将其投放市场,根据矫正的公正理应承担赔偿责任。康德的正义思想主要体现于他的义务论中,他认为义务就是绝对命令,违反法权义务就是不正义的,只要侵犯了别人的法权就应当受到正义的惩罚^[20]。康德认为每个人都应有尊重别人的道德义务,包括不把别人贬低作为自己牟利的工具,违反该义务的行为也是不正义的^[18]。因此,就本文探讨的主题而言,生产者也应该承担尊重消费者而不仅仅是将其作为牟利工具的义务,就产品缺陷致害行为理应承担赔偿责任,这样才能维持应有的公平。罗尔斯正义理论提出两大正义原则即平等自由原则和差别原则,并认为第一原则优先第二原则,在保障平等自由与经济效益之间只能二选一时,必须优先满足前者,在权利得到保障之后才可以去谈经济效益。在罗尔斯看来,在一个民主平等的社会里,

没有理由要求一些人为了另一些人的需要而放弃自己的权利。正义的原则表明,社会基本结构中的人都有不把彼此当作手段而是作为目的本身来看待的愿望^[21]。罗尔斯的正义理论为确立产品责任提供了理论基础,正义理论提倡保障个人权利,保障平等自由,保障公正公平,从而为实行产品严格责任提供了支撑。在该理论的核心宗旨指导下,产品责任的确立应该集中保障个人权益,保障作为弱势一方的消费者。因此,当出现缺陷产品时应实行严格责任,由生产者承担赔偿责任,不允许生产者进行免责抗辩,以此充分维护消费者的个人权利及利益。总之,在人工智能产品发展风险责任中,此理论偏向于反对适用产品发展风险抗辩制度,否则对消费者而言不公平。

(二) 效率理论

公平与效率都是法的重要价值,与公平理论相对应的效率理论主要指的是法的效率价值。法律的传统使命是为财富交易行为提供公平的保护,但法律的当代使命发生了嬗变,是在基于传统使命的前提下,实现资源的最大限度的优化使用和配置。其追求社会成员在公平分享“蛋糕”的基础上进行资源优化配置,充分发挥法律的作用而把这块“蛋糕”做大。公平是法的古老价值,而效率是法的新兴使命。在一定程度上,效率甚至决定着公平的质量及其由实然向应然迈进的速度。20世纪的经济分析法学派试图探寻法律运行的经济逻辑。其基于经济学是研究资源分配的科学,而法律也是一种资源,因此法的创制与执行都应以最大限度地增加社会财富为出发点和追求目标,在进行法律制度设计时,要在成本收益分析的基础上,充分考虑效率价值目标,设置有利于资源充分利用、效率最大可能发挥的制度^[22]。著名的汉德公式提出:对社会而言,预防的努力和损害的挽回都需要付出成本,而哪种成本较低,就是有效率的理性选择。即有效的分担机制在于通过低成本来实现风险转化,也就是通过社会视角实现风险在更大范围内的转移和分摊^[13]。如果不支持发展风险抗辩制度,生产者付出成本的提高就成为必然性事件;而支持发展风

险抗辩制度，产品出现不可发现的缺陷且对消费者致害则是偶然性事件。整个工业界为防范发展风险所必然付出的成本，显然高于消费者偶然损失的成本。因此，在人工智能产品的发展风险责任中，效率理论偏向于支持适用产品发展风险抗辩制度，因为这样才是有效率的。

（三）公平与效率理论的冲突与协调

人工智能产品的生产与消费必然面临公平与效率的冲突问题。因为，从消费者视角来看，公平更重要。而从生产者视角来看，效率更重要，他们都有相应的理论作支撑。但是，追求纯粹的公平与效率实不可取。所以，必须进行公平与效率之间的协调。具体可以从公平与效率的经济学视角和法学视角进行分析。

其一，经济学视角下公平与效率理论的冲突协调。从经济学视角分析，公平与效率相互依存且相辅相成。一方面，效率是公平的物质保障，因为效率意味着生产力的提高，带来物质财富的增加，继而促进社会公平目标的实现。另一方面，公平的规则可以营造良好的社会氛围^[23]，调动生产者的积极性，从而有助于提高经济效率，促进社会生产力的发展。因此，要坚持公平优先，兼顾效率，既要注重公平，又要体现效率，这就要求务必平衡二者之间的关系，从而促进社会发展。产品发展风险抗辩制度既是公平的体现，也是效率的体现。它的适用摒弃了纯粹公平、权利和安全的考量。发展风险抗辩，可以说是在产品责任领域为数不多的可供生产者用来自我保护的工具体，因此是对严格责任原则适用存在问题的冷静反思，也是对严格责任原则的否定之否定。即在风险社会下，侵权法应适应风险社会治理而采纳分配正义作为正义基础^[13]。也正如冯巴尔教授所言，“我们有必要把产品责任法拉回到侵权行为法的一般原则中考察：从一个制造者那里期待的行为模式是怎样的”^[24]。

其二，法学视角下公平与效率理论的冲突协调。从法学视角分析，公平与效率都是法的重要价值，二者不分轻重且相辅相成。公平与效率可谓法的双翼，使法运行于社会的理想状态便是公平与效率的最佳平衡。一方面，法律是实现公平

正义的重要保障。任何时期的法律都是以保障公平和促进公平的发展为最终目标。法学上的公平，既包括形式公平也包括实质公平，法必须发挥其特殊功能，协调不同的利益群体的要求，使之达到社会发展所要求的公平状态^[22]。另一方面，效率决定着公平的质量及其由实然向应然迈进的速度。法的效率价值着力解决资源配置问题，通过法律制度与法律规则促进社会财富积累，促进社会经济效益提高，最终促进公平目标的实现。产品发展风险抗辩制度既要体现公平价值也要体现效率价值。因此，产品责任法在确定产品责任时，应该衡量公平与效率二者之间的关系并合理划分责任，对人工智能产品发展风险抗辩制度考量时，也应当注重两者的平衡，即在处理公平与效率的关系时，根据实际情况寻求最优路径。

七、激励与约束理论

（一）激励理论

我国最早关于“激励”一词的使用是在《史记·范雎蔡泽列传》之中，即“欲以激励应侯”，意为激发砥砺，促人振奋。而现在一般是指激发鼓励之义。倪正茂教授给出的定义是“激发、鼓励人们为一定行为，以实现满足需要的愿望”^[25]。付子堂教授认为，法律应该是具有激励正外部性功能的一套作用机制，即通过法律设立一种激励机制，鼓励个体作出法律所要求和期望的正外部性行为，最终实现法律所设定的整个社会关系目标的要求，取得预期的法律效果，实现理想的法律秩序^[26]。本文主要从法学视域进行界定，即主要是从权利、义务和责任视角进行界定，通过权利赋予、义务减免和责任免除等对正外部性行为供体进行鼓励之意。换言之，法律激励是指在法律关系主体实施某种正外部性行为之后，获得的一种肯定评价以及由此而带来的某种利益。法律激励的目的在于鼓励、引导法律关系主体主动、积极实施正外部性行为。前文已述，人工智能产品生产者与消费者之间具有双向的正负外部性。

在激励理论视域下,主要探讨人工智能产品的生产者与消费者之间的双向正外部性。在人工智能产品生产消费过程中,人工智能产品的生产者会对消费者产生正外部性。相反,消费者也会对人工智能产品的生产者产生正外部性。从法律的公平与正义的视角来看,必须对这两种正外部性进行积极的激励和促进。其中,适用产品发展风险抗辩制度,就能够对生产者产生激励作用;不适用产品发展风险抗辩制度,就能够对消费者产生激励作用。

(二) 约束理论

与激励理论相对应的是约束理论。前文已述,外部性有正负外部性之分。人工智能产品生产者与消费者之间具有双向的正负外部性。在约束理论视域下,主要探讨人工智能产品生产者与消费者之间的双向的负外部性。在人工智能产品生产者的生产经营过程中,人工智能产品的生产者会对消费者产生负外部性。相反,消费者也会对人工智能产品的生产者产生负外部性。从法律的公平与正义的视角来看,必须对这两种负外部性进行约束和抑制。其中,不适用产品发展风险抗辩制度,就能够对生产者产生约束作用;适用产品发展风险抗辩制度,就能够对消费者产生约束作用。因为任何新生事物的产生必然带来一些问题,现如今,发展人工智能产业已经成为我国的战略性目标,因此我们必须着力解决人工智能产业发展过程中的负外部性,规范约束人工智能产品生产者的行为,同时采取措施约束消费者的不当行为,促进人工智能产业的健康发展。在约束人工智能产品的生产者方面,可以让其承担相应的社会责任,包括法律责任和道德责任,避免随着机器在普通人的生活和工作中的普及,通过机器使行为和责任脱钩日益成为普遍现象。越来越多的人将失败的责任转嫁到机器身上^[2]。除此之外,也可以制定一系列的生产标准及规范,完善相应的法律制度,为人工智能产品生产者的生产研发行为提供统一标准,有限度地适用发展风险抗辩制度,从而约束其经营行为,确保人工智能产业发展的规范性以及合法性,尽力消除其负外部性,从而促进人工智能产业的有序发展。

(三) 激励与约束理论的冲突与协调

激励理论与约束理论在多个学科被使用,本文主要从法学学科进行探讨。人工智能产品的生产与消费中存在正外部性和负外部性,必须从法律制度方面进行激励与约束,而既要激励又要约束,且激励与约束的对象不一致,必然存在冲突问题,这些问题必须得到解决,以实现二者之间的协调。可以通过法律责任的分配来激励科学技术研发的热情,尤其是通过对合规抗辩和发展风险抗辩的限制,来责成生产者研发任务的实现^[3]。但是,法律激励与法律约束在一定程度上具有冲突性,因为激励(如适用产品发展风险抗辩制度)意味着积极采取措施,给予主体(如人工智能产品生产者)充分的自由,促进其实施一定的生产研发行为,尽可能减少约束举措,充分调动其生产创新的积极性,而这样会给消费者带来约束。但约束(如不适用产品发展风险抗辩制度)则意味着采取一系列措施干预主体(如人工智能产品生产者)的某些生产研发行为,附加一些条件以及框架约束生产者的行为,减少其活动的自由度,过多的约束举措会阻碍生产者的行为积极性,从而在一定程度上抑制其发展的动力,而这样会对消费者产生激励。因此,此类冲突必须进行协调。

其一,法律与科技关系视野下激励与约束的冲突协调。虽然“法治的进步总是伴随着科技与法律的相互摩擦”^[4],但是科技的发展也离不开法律的激励与约束。一方面,科技的发展需要法律的激励,这在全球各个国家体现得淋漓尽致。如针对美国的专利制度,林肯就指出是“为天才之火浇上利益之油”。另一方面,科技的发展也需要法律制度的约束。具体到人工智能产业发展而言,人工智能产业已经成为全球新一轮战略竞争的重点,鼓励其发展是各国的一致目标。同时,各国纷纷出台相应的规划并投入巨资激励人工智能行业发展,鼓励生产者创新,积极研发人工智能产品。我国亦是如此。但激励并不意味着无底限的自由,需要激励和约束双管齐下,同时把握好适当的度。即一方面采取激励措施充分调动人工智能产业发展的积极性,另一方面也采取举

措适当约束人工智能产业发展的一系列负外部性，促进其规范化发展，推动全球智能社会的进步。毕竟“没有社会理性的科学理性是空洞的，但没有科学理性的社会理性是盲目的”^[28]。这些都需要我们建构一套合理的人工智能产品发展风险抗辩制度，从而对生产者和消费者都起到激励作用。

其二，生产者与消费者之间双向正负外部性视野下激励与约束的冲突协调。人工智能产业在发展过程中，其生产者与消费者之间会产生双向的正负外部性。因此，需要健全针对生产者和消费者的激励与约束相协调的法律制度。一方面，基于人工智能产品生产者对消费者的正外部性，需要对其进行法律激励；基于生产者对消费者的负外部性，需要对其进行法律约束；基于人工智能产品消费者对生产者的正外部性，需要对其进行法律激励；基于消费者对生产者的负外部性，需要对其进行法律约束。但是，激励与约束都不能走极端，而必须保持合适的“度”。另一方面，对人工智能产品生产者进行约束，可以对消费者产生激励；对人工智能产品生产者进行激励，可以对消费者产生约束。反之，对人工智能产品消费者进行约束，可以对生产者产生激励；对人工智能产品消费者进行激励，可以对生产者产生约束。这些都需要我们建构一套合理的人工智能产品发展风险抗辩制度，从而对生产者和消费者都起到激励和约束作用。总之，需要完善针对人工智能产品生产者与消费者的激励与约束协同法律规制，针对人工智能产品合理适用发展风险抗辩制度，从而实现激励创新与约束风险的平衡，最终实现人工智能产业的繁荣与发展。

八、结语

分配正义包括财富分配正义和风险分配正义。人工智能产品发展风险抗辩制度适用与否决定着在生产者和消费者之间关于产品发展风险分配是否正义。目前，人工智能产品发展风险抗辩制度适用与否面临激烈争论而难以达成共识，

这些争论可以用诸多相互冲突的理论进行诠释。这些理论包括但不限于正外部性与负外部性理论、成本与收益理论、安全与效用理论、权利与效益理论、公平与效率理论、激励与约束理论。这些理论的冲突迫切需要进行平衡才有助于问题的解决。这些理论之间也存在着逻辑勾连，如从涉及的学科来看，成本、收益、效用、效率、效益、外部性、权利、公平、安全、激励理论综合涉及经济学、法学、心理学、管理学等学科。从所持的立场来看，这些理论大致可以归纳为从生产者和消费者视角进行立论。具体来讲，正外部性理论、收益理论、效用理论、效率理论、效益理论、激励理论可以解释为重在支持人工智能产品生产者而主张适用发展风险抗辩制度，从而促进人工智能产业的高质量发展。相反，负外部性理论、成本理论、安全理论、权利理论、公平理论、约束理论可以解释为重在支持人工智能产品消费者而主张不适用发展风险抗辩制度，从而确保弱势消费者群体的消费权益。事实上，这些冲突的理论并非方枘圆凿。我们也不能走向任何一个极端，而必须秉持公平正义，在生产者和消费者之间进行充分的协调，从而实现激励人工智能产业良性发展与约束人工智能产品异化之间的衡平，充分缓解人工智能产品生产者和消费者之间冲突带来的张力，在保护消费者权益的前提下促进我国人工智能战略产业的高质量发展，以实现国家战略与包容发展、商业利益与社会责任的平衡，从而为人工智能善治提供中国方案。

注释：

- ① 从经济学视角看，庇古认为外部性指的是经济主体为了实现自身利益而进行一项经济活动，但是在实现该目的的过程中，却对与该活动无直接关系的他人或组织产生了有利或者不利的影 响。其中，对他人产生的有利影响就是正外部性，而产生的不利影响则是负外部性。从法学视角看，正外部性是一个经济主体在行使权利时，将可由自己行使的权利让渡给他人且没有施加任何义务。负外部性是一个经济主体在行使权利时将本应该由自己承担的义务施加给其他经济主体。参见胡元聪《外部性问题解决的经济法进路研究》，法律出版社 2010 年版，第 53 页。

参考文献:

- [1] 博登海默. 法理学: 法律哲学与法律方法[M]. 邓正来, 译. 北京: 中国政法大学出版社, 1999: 281-299.
BORDENHEIMER. Jurisprudence: Legal philosophy and legal method[M]. Trans. DENG Zhenglai. Beijing: China University of Political Science and Law Press, 1999: 281-299.
- [2] 皮埃罗·斯加鲁菲. 智能的本质——人工智能与机器人领域的64个大问题[M]. 任莉, 等译. 北京: 人民邮电出版社, 2017: 170-171.
PIERO SCARUFI. The nature of intelligence——64 big questions in the field of artificial intelligence and robotics [M]. Trans. REN Li, et al. Beijing: Posts and Telecommunications Press, 2017: 170-171.
- [3] 卢克·多梅尔. 算法时代: 新经济的新引擎[M]. 胡小锐, 等译. 北京: 中信出版社, 2016: 144.
LUKE DORMER. Algorithmic age: The new engine of the new economy[M]. Trans. HU Xiaorui, et al. Beijing: Citic Press, 2016: 144.
- [4] 马长山. 智能互联网时代的法律变革[J]. 法学研究, 2018(4): 20-38.
MA Changshan. Legal reform in the era of intelligent internet[J]. Legal Research, 2018(4): 20-38.
- [5] 左卫民. 迈向大数据法律研究[J]. 法学研究, 2018(4): 139-150.
ZUO Weimin. Towards big data law research[J]. Law Research, 2018(4): 139-150.
- [6] 郑戈. 在鼓励创新与保护人权之间——法律如何回应大数据技术革新的挑战[J]. 探索与争鸣, 2016(7): 79-85.
ZHENG Ge. Between encouraging innovation and protecting human rights——How law responds to the challenge of technological innovation of big data[J]. Exploration and Contention, 2016(7): 79-85.
- [7] THUCYDIDES. The War of the Peloponnesians and the Athenians[M]. Trans. JEREMY MYNOTT. Cambridge: Cambridge University Press, 2013: 380.
- [8] 尼克·比尔顿. 翻转世界——互联网思维与新技术如何改变未来[M]. 王惟芬, 等译. 杭州: 浙江人民出版社, 2014: 237.
PILTON NICK. Turning the world around: How internet thinking and new technologies will change the future[M]. Trans. WANG Weifen, et al. Hangzhou: Zhejiang People's Publishing House, 2014: 237.
- [9] 李建军. 技术创新活动的负外部性及其缓冲机制分析[J]. 中国人民大学学报, 2008(1): 125-130.
LI Jianjun. Analysis of negative externalities and Buffer Mechanism of technological innovation activities[J]. Journal of Renmin University of China, 2008(1): 125-130.
- [10] 凯文·沃巴赫. 链之以法: 区块链值得信任吗?[M]. 林少伟, 译. 上海: 上海人民出版社, 2019: 3.
KEVIN WOLBACH. Can blockchain be trusted?[M]. Trans. LIN Shaowei. Shanghai: Shanghai People's Publishing House, 2019: 3.
- [11] 冯珏. 自动驾驶汽车致损的民事侵权责任[J]. 中国法学, 2018(6): 109-132.
FENG Jue. Civil tort liability for damage caused by self-driving cars[J]. Chinese Law Science, 2018(6): 109-132.
- [12] 涂永前. 规制人工智能: 一个原则性法律框架研究[J]. 人工智能法学研究, 2018(1): 119-231.
TU Yongqian. Regulation of artificial intelligence: A study on a principled legal framework[J]. Artificial Intelligence Law Research, 2018(1): 119-231.
- [13] 何国强. 风险社会、风险分配与侵权责任法的变革[J]. 广东社会科学, 2018(3): 228-237.
HE Guoqiang. Reform of risk society, risk allocation and tort liability law[J]. Guangdong Social Sciences, 2018(3): 228-237.
- [14] 张民安. 美国侵权法上的售后危险警告义务研究[J]. 北方法学, 2008(6): 41-60.
ZHANG Minan. Research on the obligation of after-sales danger warning in American tort law[J]. Northern Law, 2008(6): 41-60.
- [15] 朱岩. 风险社会与现代侵权责任体系[J]. 法学研究, 2009(5): 18-36.
ZHU Yan. Risk society and modern tort liability system [J]. Legal Research, 2009(5): 18-36.
- [16] 张骥. 在效益与权利之间——美国产品责任法的理论基础[J]. 中国法学, 1997(6): 103-109.
ZHANG Qi. Between benefit and right—The theoretical basis of American product liability law[J]. Chinese Law Science, 1997(6): 103-109.
- [17] OWEN D G. Philosophical Foundations of Tort Law[M]. Oxford: Clarendon Press, 1995: 165.
- [18] 伊曼努尔·康德. 道德形而上学原理[M]. 苗田力, 译. 上海: 上海人民出版社, 2005: 49.
Immanuel Kant. Principles of moral metaphysics [M]. Trans. MIAO Tianli. Shanghai: Shanghai People's Publishing House, 2005: 49.
- [19] 拉伦茨. 法学方法论[M]. 陈爱娥, 译. 台北: 五南图

- 书出版公司, 1996: 390.
- LARENZ. Methodology of law[M]. Trans. CHEN Ai-e. Taipei: Wunan Book Publishing Company, 1996: 390.
- [20] 舒远招, 吴雪. 从义务论的角度看康德的正义思想[J]. 道德与文明, 2019(1): 58-67.
- SHU Yuanzhao, WU Xue. Kant's thought of justice from the perspective of deontology[J]. Morality and Civilization, 2019(1): 58-67.
- [21] 约翰·罗尔斯. 正义论[M]. 上海: 上海译文出版社, 1991: 196.
- JOHN RAWLS. A theory of justice[M]. Shanghai: Shanghai Translation Publishing House, 1991: 196.
- [22] 齐延平. 法的公平与效率价值论[J]. 山东大学学报(社会科学版), 1996(1): 69-74.
- QI Yanping. Value theory of equity and efficiency of law[J]. Journal of Shandong University (Social Science Edition), 1996(1): 69-74.
- [23] 丁春福, 陈彦超. 新时代: 公平与效率关系解析及政策选择[J]. 黑龙江社会科学, 2018(4): 132-135.
- DING Chunfu, CHEN Yanchao. New era: Analysis of the relationship between equity and efficiency and policy choice[J]. Heilongjiang Social Sciences, 2018(4): 132-135.
- [24] 克里斯蒂安·冯·巴尔. 欧洲比较侵权行为法(下卷)[M]. 北京: 法律出版社, 2004: 339.
- CHRISTIAN VON BAHR. Comparative tort law in Europe (vol. 2)[M]. Beijing: Law Press, 2004: 339.
- [25] 倪正茂. 激励法学要言[J]. 东方法学, 2009(1): 3-17.
- NI Zhengmao. On incentive law[J]. Oriental Law, 2009(1): 3-17.
- [26] 付子堂. 法律功能论[M]. 北京: 中国政法大学出版社, 1999: 68-69.
- FU Zitang. Legal function theory[M]. Beijing: China University of Political Science and Law Press, 1999: 68-69.
- [27] 张玉洁. 区块链技术的司法适用、体系难题与证据法革新[J]. 东方法学, 2019(3): 99-109.
- ZHANG Yujie. Judicial application, system problems and innovation of evidence law of blockchain technology [J]. Oriental Law, 2019(3): 99-109.
- [28] 乌尔里希·贝克. 风险社会[M]. 何博闻, 译. 南京: 译林出版社, 2004: 30.
- ULRICH BECK. Risk society[M]. Trans. HE Bowen. Nanjing: Yilin Press, 2004: 30.

Theoretical conflict and balance coordination of risk defense system in AI product development

HU Yuancong

(School of Economic Law, Southwest University of Political Science and Law, Chongqing 401120, China)

Abstract: In order to change the product liability system in China according to the characteristics of artificial intelligence (AI) products so as to prevent, reduce and distribute the development risk of AI products to the greatest extent, it is imperative to establish the defense system of AI product development risk. However, the conflict between many theories that the system confronts, including positive externality and negative externality theory, cost and benefit theory, safety and utility theory, right and benefit theory, equity and efficiency theory, incentive and constraint theory, must be coordinated, so as to realize the balance between stimulating AI industry benign development and restricting product differentiation tendency of artificial intelligence, to relieve the tension brought about by conflict between producers and consumers, to promote high quality development of AI strategic industry in China, and further to ensure the balance between the national strategy and inclusive development, between business interests and social responsibility, hence contributing to the rule of China and of artificial intelligence.

Key Words: AI product liability; developing the risk defense system; theoretical conflict; balance coordination

[编辑: 苏慧]