

生成式人工智能产品责任承担研究

陈传强

青岛大学法学院, 山东 青岛

收稿日期: 2025年2月19日; 录用日期: 2025年3月12日; 发布日期: 2025年3月24日

摘 要

相较于传统物质性产品, 生成式人工智能的特殊性使其难以被现行的产品责任所涵盖, 致使民事责任承担机制的失调。物质形态并非判断生成式人工智能产品属性的唯一标准, 从目的论角度出发, 在现有产品责任框架内探讨生成式人工智能的责任承担问题, 是合理划分举证责任、弥补损害、规制生成式人工智能发展的可行路径。生成式人工智能在产品责任适用中应当准确认定“生产者”角色, 对通识的网络服务提供者概念进行修正, 并制定明确的质量标准, 结合警示说明义务和风险补救义务综合评判。本文以产品责任为基点, 推动科技发展与立法更新的协调发展。

关键词

生成式人工智能, 产品责任, 设计缺陷, 损害赔偿

Research on the Responsibility Undertaking of Generative Artificial Intelligence Products

Chuanqiang Chen

School of Law, Qingdao University, Qingdao Shandong

Received: Feb. 19th, 2025; accepted: Mar. 12th, 2025; published: Mar. 24th, 2025

Abstract

Compared to traditional material products, the particularity of generative artificial intelligence makes it difficult to be covered by current product liability, resulting in an imbalance in the mechanism of civil liability. The material form is not the only criterion for judging the attributes of generative artificial intelligence products. From the perspective of teleology, exploring the issue of responsibility for generative artificial intelligence within the existing product liability framework is

a feasible path to reasonably allocate the burden of proof, compensate for damages, and regulate the development of generative artificial intelligence. Generative artificial intelligence should accurately identify the role of “producer” in the application of product liability, revise the concept of general network service providers, and establish clear quality standards, combined with the obligation of warning instructions and risk remediation obligations for comprehensive evaluation. This article takes product liability as the basis, promotes the coordinated development of technology and legislative updates.

Keywords

Generative Artificial Intelligence, Product Liability, Design Defects, Compensation for Damages

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

近年来,以 ChatGPT 为代表的生成式人工智能快速发展,对传统的社会生活方式和法律秩序产生了巨大冲击。当前,生成式人工智能(Generative Artificial Intelligence)被应用于各类创作场景中,人工智能生成内容在许多作品领域逐渐成为产业创作新常态[1]。与传统的判别式模型相比,生成式人工智能利用其复杂的算法和模型,对大规模数据深度学习,以创造出原创内容。其最显著的特点在于,生成式人工智能文字、图像、艺术创作等领域突破了原有计算机网络的限制,可以产生数据世界中前所未有的数据信息。从 2022 年底到 2024 年初,在生成对抗网络和变分自编码器等其他技术算法的加持下,生成式人工智能的可产出内容迅速从 ChatGPT 的文字、图表,发展到了 Dall·E3 的图像以及 Sora 的视频层面,不断颠覆着人们对生成式人工智能技术的认知。

2. 生成式人工智能概述

2.1. 生成式人工智能界定

生成式人工智能(Generative Artificial Intelligence, GAI)是指能够自主地创造出新的内容或数据的人工智能系统[2],其拥有复杂的算法和大模型,与传统的判别式模型相比,生成式模型所产生的数据源于对现有数据的二次加工,这也使得生成式人工智能在文字、图像、艺术创作等领域突破了原有计算机程序的限制,可以产生数据世界中前所未有的数据信息。

2.2. 生成式人工智能带来法律风险

生成式人工智能发展至今,带来巨大科技变革的同时也酝酿出了难以估量的法律风险。具体而言,这些风险主要表现为“数据安全风险”“科技伦理风险”“犯罪工具风险”“网络信息内容安全风险”“侵害著作权、肖像权风险”等[3]。尽管有学者指出,这些法律风险作为目前数字法治领域热议的数据安全问题、个人信息保护问题以及算法歧视问题,应当统一纳入到人工智能法律领域之下,但就当下而言,法律为维持社会公众对其运行的一般预期,必须保持其内在的统一性[4],因此应当将其纳入现有法律规范体系中,避免过度立法造成的法律规范与现实的割裂。但不可否认的是,随着算法技术的不断发展,“生成式人工智能”这一说法日益偏离其最初的本意,逐渐演化成为一种具有全新形态、规则、理论的新生事物。如今,探讨生成式人工智能的法律地位及法律风险应当把握其固有特点,使其与现有

法律制度适配的同时，探索立法更新的未来趋势。

3. 生成式人工智能产品责任适用困境

3.1. 生成式人工智能法律责任主体争议

生成式人工智能诞生之初，以郭少飞教授为代表的部分学者建议，鉴于生成式人工智能展现出了强大的类人力能力，可探索依据主体论的观点，授予生成式人工智能以法律主体地位，使其独立承担由其引发的侵权责任，从而不再以产品责任的视角分析侵权问题，并深入探索了生成式人工智能的自有财产范围、准财产责任等问题[5]。但随即便有学者发文反对，主张生成式人工智能并非具有侵权责任中的主体地位，不应被赋予法律主体资格[6]。认同该观点的学者们从“拟制说”出发，分析了生成式人工智能与法律中现有的拟制主体的不同，譬如，白玉杰和张昱提出法律制度的核心始终在于以人为中心，而人工智能本质上只是一种机器，不具备作为法律主体的资格，再处理侵权纠纷时仍应优先适用产品责任。

3.2. 生成式人工智能“产品与服务”之争

在生成式人工智能“非主体”说的情形下，如何界定生成式人工智能的法律地位是明确侵权行为发生后请求权基础的法律前提。随着社会经济的发展，无体物作为财产登上历史舞台，这对传统产品的概念提出了挑战[7]。对此，学界主要也存在两种观点。

一方观点认为，生成式人工智能应当被视为一种网络服务，其本质是算法提供者将其编写的特定程序上传至允许多数人访问的网络空间，并为多数人提供一定的算法服务，帮助解决复杂问题，由于其不具备任何物理形态，难以依靠产品质量相关法规划定合理评判标准。因此，由其造成的损害应当根据一般侵权责任的规定进行补偿。

另有观点认为，由于生成式人工智能大多以租赁、购买会员等方式获得在一定时间内的使用权，无法像物理产品一样获得民法意义上的所有权，因此在形式上更加契合大众对于网络服务的概念。从责任承担的角度来说，如果适用一般侵权责任的过错原则，生成式人工智能使用者往往难以承担相应的举证责任，在证明生成式人工智能存在过错方面具有先天不足，将生成式人工智能拟制为类似于商业软件之流的信息产品，参照产品责任的严格责任原则进行规制更有利于处理数字时代、智能化时代的侵权问题。

3.3. “产品缺陷”认定困难

从规范层面看，产品责任适用的前提是产品存在缺陷，唯有缺陷产品造成他人损害，生产者才承担产品责任。但对于生成式人工智能来说，其产出内容源于模型运算以及数据汇总，相较于传统产品来说，其运行有相当程度的不可控性，生产者对其控制力明显减弱，加之算法程序的不稳定性，生成式人工智能是否存在“产品缺陷”难以界定。在此条件下，生成式人工智能的侵权行为如何认定是由设计之初的“产品缺陷”导致，需要法学界和技术界的共同协商。

由于生成式人工智能的数字性、智能化，有关“产品缺陷”的认定应当适当打破对传统产品的认知，在立法层面适当扩充服务提供者的跟踪、观察、管理义务，将服务提供者的后续不作为行为也界定为生成式人工智能的“产品缺陷”，以此完善生成式人工智能的责任划分。

4. 生成式人工智能产品责任归责路径

《民法典》第1202条规定：因产品存在缺陷造成他人损害的，生产者应当承担赔偿责任。由此，生成式人工智能的产品责任适用应当解决两个问题，一是有关“生产者”的认定，二是有关“产品缺陷”的判断。

4.1. 生成式人工智能“生产者”角色判定

生成式人工智能产品责任归责，首先要解决的问题是责任主体的认定问题。类比传统产品责任的责任主体，生成式人工智能的“生产者”可以笼统概括为包括程序设计、算法编写、远程服务在内的“服务提供者”，但如何准确认定“服务提供者”当下仍缺乏明确的标准。

首先，生成式人工智能服务提供者不包括大模型网络介入、传输、储存服务的提供者。尽管在形式上看使用者是通过网络服务者提供的服务连接到了大模型算法的使用，但对于生成式人工智能运行来说，这些角色很难产生足以干涉程序运行的不利影响，看似与生成式人工智能的运行密不可分，但却在损害发生和责任划分上存在巨大差别。网络服务提供者为用户提供了“信息选择权”，允许使用者自由决定想要获取的信息，而生成式人工智能服务提供者则为用户产出了“唯一性答案”，对用户的影响力可见一斑。对于网络服务传输、储存者来说，其也无法控制传输、储存的内容，因此不宜将其与大模型“服务提供者”混为一谈。从 Open AI 公司在使用协议中自认的免责条件来看，有关 ChatGPT 生成内容的免责条件与网络通信服务提供者、系统缓存服务提供者的免责条件也大不相同，这也印证了生成式人工智能规则制定者也对大模型服务者和网络服务者做出了区分。

其次，生成式人工智能服务提供者也不得认定为训练数据信息提供者。尽管大模型的训练离不开海量的信息数据，这些数据的倾向、密度、真实性对大模型的输出内容具有绝对性影响，但在发生诸如侵犯著作权等实体权利的案件时，不能将相应的责任归结于原始数据的上传者。原因在于，数据提供者虽然提供了训练数据，但缺乏对于训练结果的控制力，鉴于大模型算法的复杂性，也就不宜对其“控制力”以外的法律责任负责[8]。目前，学界已经有了将网络服务者分为“互联网信息服务提供者”和“深度合成服务提供者”的划分理论，基于此，有学者指出，对以 ChatGPT 为代表的大型语言模型，应采用“深度合成”作为法律概念[9]，用以划分一般的网络服务者和大模型服务提供者。

4.2. 生成式人工智能产品缺陷识别

4.2.1. 设计缺陷

通常认为，设计缺陷是指产品的设计，如产品结构、配方等存在不合理危险。但对于生成式人工智能来说，设计缺陷则更加地抽象、隐晦。生成式人工智能的技术基础是深度学习算法，例如变分自编码器、生成对抗网络等，如果设计者的设计参数不当或算法调优不足，就有可能导致输出内容偏见化或伦理风险化。因此有学者认为，人工智能产品的核心是算法系统，这使得软件越来越成为设计缺陷判断的重点[10]。其次，生成式人工智能的训练过程需要大量的训练数据，训练数据的质量和多样性也决定着大模型的输出能力。生成式人工智能文本生成由其自然语言模型所决定，而自然语言模型在本质上取决于算法选择以及用于模型训练的庞大数据库[11]。未对训练数据进行核实的行为、过度使用带有偏见性的训练数据的行为，同样可以判定生成式人工智能存在设计缺陷。因此，在认定生成式人工智能大模型是否存在设计缺陷时，应当从技术标准和道德规范两个角度评判大模型的设计训练过程，关注其算法结构以及训练数据的选择是否合理，数据生成过程是否会升高不合理的风险。

4.2.2. 警示说明缺陷

警示说明缺陷也称为销售缺陷[12]。根据我国《消费者权益保护法》第 18 条规定，生产者、经营者负有提供符合安全要求的产品以及准确表明使用危险的义务。违反这一法定义务的，就构成警示说明缺陷。

对于生成式人工智能来说，由于其算法模型的变动性和不确定性，生成内容难以完全避免偏差、歧视及误导性、侵权性，具有较大的局限性，因此在生成式人工智能运行规则设计之初，应当将警示说明义务作为先前义务纳入到产品设计之中。在生成式人工智能服务运行中，设计者、生产者提醒使用者避

免潜在风险,向使用者警示生成内容的可预见危险性,并且提示正确的使用方法。若缺少有关风险警示,如“生成内容仅供参考”“网络数据存在偏差”等,可能导致用户误信生成结果的,便可视为生成式人工智能存在产品设计缺陷,对于造成的损害结果,可按照产品责任的构成要件以及归责原则加以规制,承担相应的赔偿责任。

4.2.3. 风险补救缺陷

对于人工智能产品来说,许多产品在投入使用时,受限于当时的技术条件和社会环境、法律制度等因素,设计缺陷并不能及时显现,而是随着技术的进步,逐步暴露出新的问题和风险[13]。在设计之初防范无法预见的潜在风险、确立补救措施,对于具有深度学习能力的生成式人工智能来说尤为重要。

当下,生成式人工智能的一大优势在于,在其运行过程中,会对收集到的网络数据以及使用者提供的数据进行深度学习,以此实现大模型自身的不断更新。但伴随而来的,是生成式人工智能长期存在的持续风险。因此,生成式人工智能的设计者应当在设计之初设置完善的长期监督机制以及风险补救措施,对于暴露出的风险不应视若无睹,而应当及时对模型进行修正、剔除风险数据甚至对大模型进行重训,以最大程度防范、降低生成式人工智能长期运行过程暴露的风险。作为具有强大算力且能够同时服务大量用户的网络服务,生成式人工智能风险补救措施的缺位,应当视为存在重大设计缺陷,对于潜在风险未采取补救措施或补救措施不及时,应当适用产品责任有关规定。

4.3. 生成式人工智能产品责任下的损害赔偿

4.3.1. 生成式人工智能损害的特殊性

生成式人工智能作为当下人工智能产品的高阶形态,在算法设计、数据训练、深度学习、用户输入等众多因素的影响下,生成式人工智能的输出过程具备“黑箱”特征,其算法过程被设计者以商业秘密的形式予以保护,这也导致其生成内容具有秘密性以及高度的不确定性。在这种不确定性的影响下,有关侵权行为的行为过程及损害结果变得难以预测和防范,且多为非物质性损害,这使得生成式人工智能的损害赔偿问题有待深论。

首先是损害结果的不确定性,包括损害发生的不确定性和损害范围的不确定性。在网络技术的加持下,具有侵害性的生成内容能够以极快的速度进行生成、传播。当侵害发生之后对损害结果进行评估时,生成式人工智能的不可控性使得损害发生的源头难以溯及、损害时间难以准确认定、损害范围通常也难以估量,传统的民事责任规范在面对这种大规模损害结果时显得力不从心。

其次是责任主体难以界定。对于生成式人工智能来说,从形式上看,生成内容是由大模型独立运算生成的,算法程序全程自动运行;但究其实质,生成内容的输出是设计者、提供者、使用者等多方提供数据共同运算下的结果。侵权的输出结果实际受到基础算法、训练材料和输入行为三个要素的影响,这三个要素可能独立导致侵权结果,也可能相互作用累积引起侵权,这使得具体侵权案件的归责难度加大[14]。因此,当损害发生时,偏差性的生成内容是由设计算法导致,还是训练数据、使用者具体指令导致通常难以划分。

同时,由于设计者、提供者、使用者之间相互补充相互作用的数据关系,各方与损害结果之间的因果关系也难以界定,对于损害发生的具体原因难以通过司法手段进行证明,各方之间证明责任的分配也存在法律漏洞,容易导致举证责任相互推诿的情况。综上所述,以现行侵权责任法律规范,应当如何划分各方的民事责任承担比例仍较为困难。

4.3.2. 现行法下的损害赔偿

现行民事法律规范下,生成式人工智能所造成的损害可以依照《民法典》第1179条人身损害赔偿责

任、第 1183 条精神损害赔偿责任、第 1184 条财产损害赔偿 responsibility 进行规制。

4.3.3. 惩罚性赔偿的适用

有关生成式人工智能如何适用惩罚性赔偿的细则应当作为未来立法更新的重点。

《民法典》第 1207 条规定，由于产品缺陷导致他人健康严重受损的，可以适用惩罚性赔偿。该条款作为实物性产品的责任追偿请求权基础，在司法实践中已经较为完善，但对于“产品拟制”下的生成式人工智能类产品来说，抽象的算法程序与现实损害间的因果关系难以判定，如何确定“设计缺陷”与“造成他人死亡或者健康严重损害”之间的法律关系，以及如何认定“生产者”存在“明知”产品缺陷的证明方法有待法条的进一步明示。由于生成式人工智能损害的非物质性，在适用惩罚性赔偿时，可以适当突破“健康严重损害”的约束，扩大惩罚性赔偿的适用范围。

有关惩罚性赔偿的数额问题也缺乏明确规则。目前，生成式人工智能适用产品责任承担惩罚性赔偿的数额，可以依照《消费者权益保护法》第 55 条的规定，以实际损失作为计算基数，以 2 倍作为计算倍数，确定人身损害赔偿的惩罚性赔偿责任。

5. 生成式人工智能产品责任完善建议

5.1. 明确生成式人工智能的产品属性

目前，《民法典》关于产品责任的条款主要适用于具有物理形态的产品，对于生成式人工智能来说，难以通过传统意义上对“产品”的定义加以解释。这导致在侵权行为发生后，被侵权人只能通过一般侵权责任予以追偿，需要承担相应的举证责任。对于不具有专业计算机知识及算法知识的大众而言，举证能力微乎其微，难以通过法律途径维护己方合法权益。其次，生成式人工智能设计者及服务提供者基于维护商业秘密、技术秘密等原因，也会对大模型的算法予以加密，使得传统手段下的举证能力几乎不可能，即使借助司法力量，也难以通过对大模型算法的鉴定证明其实施了侵权行为。即使证明了生成式人工智能存在算法漏洞、设计缺陷，基于一般侵权责任的法律规定，只能进行损害填平，无法适用惩罚性赔偿，对于科技企业来说，司法约束力微乎其微，难以遏制生成式人工智能侵权行为的发生动因，难以敦促设计者、提供者履行警示、防范义务。

因此，应当以法律的形式明确生成式人工智能产品或服务的“产品”属性，或参照欧洲《人工智能法案》，将具有深度合成能力、高度危险性的人工智能系统纳入产品责任的调整范围，将举证责任交由生成式人工智能的设计者、服务者。当损害发生后，若其无法证明产品没有过错，则应当承担相应的责任，对于造成严重后果的，依法适用惩罚性赔偿。

5.2. 规范生成式人工智能产品缺陷认定

认定生成式人工智能是否存在产品缺陷，也应在立法层面明确生成式人工智能的质量标准。《民法典》及《产品质量法》等法律规范中，有关产品质量的标准大多是物理性的，无法衡量生成式人工智能大模型算法的得失。为了准确认定生成式人工智能是否存在产品缺陷，以及判断损害发生的具体原因，应当结合生成式人工智能算法的透明度、数据来源、训练方式、公平正义原则等要素，制定合理、易懂的缺陷判断标准，减轻司法压力、维护司法正义，为使用者提供有效的司法保护。

同时也应明确生成式人工智能设计者、服务提供者的警示义务、补救义务，将二者规定的合理与否纳入产品质量的评判范围，增强产品设计的透明度，以便增强使用者对生成内容的信任，降低瑕疵生成内容带来的风险。

5.3. 完善损害赔偿机制

对于生成式人工智能来说，其损害方式常见为数据泄露、损害名誉、生成内容侵权等非物质性损害

行为。对于非物质性损害结果来说,难以通过现行规范评估损害发生的范围、程度,也缺乏相应的赔偿标准。面对当下越发严重的生成式人工智能侵权问题,可以参照《民法典》第179条有关承担民事责任的条款制定适当的赔偿规范,明确生成式人工智能侵权行为的损害赔偿方式,适当扩大精神损害赔偿、恢复名誉、停止侵害等方式的适用范围以及使用标准,维护使用者及被侵权人的合法权益。

同时应当强化惩罚性赔偿的适用,明确生成式人工智能适用惩罚性赔偿的请求权基础,对于多次侵权、恶意侵权、造成重大影响的侵权行为适用惩罚性赔偿,并且结合损害范围、恶意程度,设定更高的数额限制,有效遏制生成式人工智能潜在的侵权风险。

6. 结论

如今,生成式人工智能的发展速度让身处这个时代的每一个人都感到诧异乃至恐惧。如何推动当下的法律制度更新,使其足以回应生成式人工智能所带来的法律挑战,是学界及实务界不得不考虑的问题。

产品责任是民事法律制度中保护个体权益、维护社会秩序不可或缺的一环,对于填平民事损害、减少侵权行为、维护社会公平正义具有重要意义。从目的解释角度来看,作为一项具有高度创新性的人工智能技术,尽管生成式人工智能具有高度的自主性,难以契合立法者在制定产品侵权有关规范时对“产品”的定义,但究其运行来说,目前仍难以摆脱人的控制,具有较强的工具性,可视为被人所使用的“产品”。因此,面对生成式人工智能有关纠纷,依据《民法典》中规定的产品侵权责任未尝不可。生成式人工智能在适用产品责任时,应当准确认定承担责任的“生产者”角色定位,并且制定明确的“产品缺陷”认定标准,合理确定产品责任下的赔偿范围。

未来可进一步探索设立契合生成式人工智能技术特点的纠纷解决机制。由于生成式人工智能技术的复杂性,现有的司法机制难以准确认定责任主体、损害缘由、损害范围等关键因素,容易导致司法的误判以及低效。对此,应当参照互联网在线纠纷解决机制,建立生成式人工智能纠纷解决机制,既要发挥司法机关作为维护社会正义“最后一道防线”的作用,也要发挥好各纠纷解决主体机制之间的协调配合作用^[15]。引入第三方专业技术机构参与鉴定、调查、调解,确保纠纷解决的公信力、高效性、专业性。建立生成式人工智能纠纷解决机制与司法审判的衔接机制,推动纠纷解决机制的司法前置化,以适应未来科技侵权至损新形态,实现科技创新与法律保障的协调发展。

参考文献

- [1] 徐小奔. 论人工智能生成内容的著作权法平等保护[J]. 中国法学, 2024(1): 166-185.
- [2] 李杨. 生成式人工智能风险的法律类型化治理[J]. 延安大学学报(社会科学版), 2024, 46(1): 31-38+57.
- [3] 赵精武. 论人工智能法的多维规制体系[J]. 法学论坛, 2024, 39(3): 53-66.
- [4] 许庆永. 判例权威逻辑之证成[J]. 政法论坛, 2024, 42(5): 103-114.
- [5] 郭少飞. 主体论视域下人工智能侵权责任承担构想[J]. 地方立法研究, 2023, 8(3): 64-80.
- [6] 张高源. 试论生成式人工智能侵权的责任主体——以 ChatGPT 为例[J]. 产品可靠性报告, 2023(12): 120-122.
- [7] 郑志峰. 人工智能产品责任的立法更新[J]. 法律科学(西北政法大學學報), 2024, 42(4): 3-17.
- [8] 徐伟. 论生成式人工智能服务提供者的法律地位及其责任——以 ChatGPT 为例[J]. 法律科学(西北政法大學學報), 2023, 41(4): 69-80.
- [9] 张凌寒. 深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成式人工智能治理的中国路径[J]. 法律科学(西北政法大學學報), 2023, 41(3): 38-51.
- [10] 郑志峰. 自动驾驶汽车产品缺陷的认定困境与因应[J]. 浙江社会科学, 2022(12): 48-58+156-157.
- [11] 刁晓东. 论生成式人工智能的数据安全风险及回应型治理[J]. 东方法学, 2023(5): 106-116.

-
- [12] 杨立新. 人工智能产品责任的功能及规则调整[J]. 数字法治, 2023(4): 27-39.
- [13] 胡元聪. 人工智能产品发展风险抗辩后的损害救济分摊机制研究[J]. 政法论丛, 2020(3): 121-130.
- [14] 陈全思, 张浩, 栾群. 人工智能的侵权责任与归责机制[J]. 互联网经济, 2019(4): 28-31.
- [15] 许庆永. 在线纠纷解决的正义表达[J]. 华东政法大学学报, 2024, 27(2): 83-93.