中外法学 Peking University Law Journal
Vol. 37, No. 2(2025)pp. 344－362

论人工智能致损的特殊侵权责任规则

林洹民*

摘 要 人工智能侵权责任认定与承担是人工智能立法的重点与难点之一。产品责任路径并不能有效回应人工智能应用侵权引发的过错与因果关系证明、新型损害界定以及责任主体确定三大挑战。人工智能侵权责任立法应在过错责任的基础上,修补式配置专门条款,纾解被侵权人在责任成立与承担上的举证压力。被侵权人只有借助对人工智能研发记录、活动日志等文件的访问,才能克服信息不对称带来的举证障碍。立法应规定证据开示规则,规定人工智能相关主体在一定条件下的信息披露义务,为法院发布书证提出命令提供实体法基础。人工智能时代对虚拟损害的救济,与其扩张物质性损害的边界,不若摒弃精神损害赔偿的严重性要件,转采"显著性"标准。为了减轻人工智能商品消费者的因果关系证明负担,立法应确定特定条件下的因果关系推定规则。当损害确定地因义务违反行为而发生,只是责任源头难以查清时,同一商业技术单元成员应对损害承担连带责任。

关 键 词 人工智能侵权 产品责任 过错责任 证据开示 虚拟损害

一、问题的提出

随着人工智能技术的跨越式发展与产业化应用,人工智能致损事件也在不断增加。〔1〕负责任的人工智能发展需要强有力的前瞻性治理。如果没有一个明确、清晰的人工智能侵权责任框架,就不会有真正的人工智能安全。人工智能具有自主性、网络化、不可预测和数据依

* 上海交通大学凯原法学院副教授。本文系国家社科基金重大项目"新一代生成式人工智能发展的法治问题研究"(项目编号:24&ZD135)的阶段性研究成果。

〔1〕 经济合作与发展组织(Organisation for Economic Co-operation and Development, OECD)在2023年11月启动人工智能事件监测系统。截至2025年2月13日,该系统已记录14000多起事故,其中在我国监测的事件约600起。"OECD AI Incidents Monitor (AIM)," https://oecd.ai/en/incidents, last visited on 13 February 2025.

赖等特征。一旦发生人工智能侵权，被侵权人难以有效证明过错、因果关系、损害等要素，甚至难以识别真正的责任主体。〔2〕人工智能应用侵权在归责与归因层面均呈现出较强的不确定性，已经成为阻碍人工智能应用的三大障碍之一。〔3〕《民法典》中的交通事故侵权责任、医疗侵权责任等规则在革新后可以成为调整无人驾驶汽车、医疗人工智能等特定人工智能应用的规范基础。〔4〕若人工智能被应用到核设施、航空器等高度危险作业，当然也有《民法典》第1236条以下高度危险责任的适用。只是针对特定场景的责任规则不能一般性地调整不同类型的人工智能侵权活动（如数字助理、人脸识别、智能家居调控系统、生成式人工智能侵权等）。如何设计一套全面的人工智能侵权责任规范框架，实为人工智能治理的荦荦大端。

一种思路是扩张产品责任的边界，要求人工智能软件提供者、销售者对因人工智能系统缺陷引发的损害承担无过错责任。欧盟2024年10月修订通过、12月生效的《缺陷产品责任指令》[Directive (EU) 2024/2853]第4条第(1)项就明确将软件列为产品，前言第13条更是清晰地指出“软件是一种适用无过错责任的产品，无论其供应或使用方式如何，也无论软件是通过设备存储、通信网络或云技术访问，还是通过软件即服务模式提供”。据此，人工智能系统成为欧盟产品责任的适用对象。但是，产品责任路径未必能有效解决人工智能应用引发的法律适用难题，对人工智能侵权行为的调整需要更为全面的框架设计。本文将在批判人工智能侵权产品责任路径的基础上，对人工智能应用带来的法律适用难题进行整体性观照，进而引出化解法律适用障碍的若干基础性规则，探索建构人工智能特殊侵权责任的基本框架。考虑到我国正在筹备起草《人工智能法》，人工智能致损的侵权责任条款设计是立法的重点和难点之一，研讨如何在我国未来《人工智能法》中设计相关条款，也是本文的写作目的之一。〔5〕

二、人工智能产品责任路径的规范乏力

产品责任属于无过错责任，较之一般过错责任似更能保护人工智能相对人的合法权益。欧盟修改扩张产品责任的适用范围，是对智能时代侵权法挑战的积极应对。然而，产品责任路

〔2〕 See Expert Group on Liability and New Technologies—New Technologies Formation, “Liability for Artificial Intelligence and Other Emerging Digital Technologies,” 2019, pp. 19-24, https://data.europa.eu/doi/10.2838/573689, last visited on 13 February 2025.

〔3〕 European Commission, Impact Assessment Report: Accompanying the Document Proposal for A Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence, SWD (2022) 319 final, p. 7.

〔4〕 参见冯珏：“自动驾驶汽车致损的民事侵权责任”，《中国法学》2018年第6期，第109—132页；冯洁语：“人工智能技术与责任法的变迁——以自动驾驶技术为考察”，《比较法研究》2018年第2期，第143—155页；郑志峰：“诊疗人工智能的医疗损害责任”，《中国法学》2023年第1期，第203—221页等。

〔5〕 国务院自2023年开始已连续两年将人工智能立法列入立法工作计划。全国人大常委会于2023年9月发布的《第十四届全国人大常委会立法规划》也将人工智能立法列入第一类项目（条件比较成熟、任期内拟提请审议的法律草案），并指出“推进科技创新和人工智能健康发展……要求制定、修改、废止、解释相关法律，或者需要由全国人大及其常委会作出相关决定的，适时安排审议”。

径未必能有效化解人工智能应用带来的过错与因果关系证明、新型损害认定与责任承担主体确定三大难题。

(一) 过错与因果关系证明困难

源于人工智能系统具有自主性和网络性,人工智能侵权的过错与因果关系证明困难重重。①人工智能系统具有一定的自主性,当发生人工智能侵权时,难以判断系统研发者、提供者、应用者等相关主体是否具有过错。以往计算机系统未按照预期运行的,系统会被认为出现错误。以深度学习算法为代表的人工智能系统具有自我学习与自我进化能力,可以根据并非完全预先给定的规则做出决定或采取行动。当人工智能系统生成不合期待的结果,并不能当然认定系统出现错误。以生成式人工智能为例,生成式人工智能完全可以从海量数据中挖掘新的语言规律以进行合乎逻辑的"捏造",即产生所谓"幻觉"。生成式人工智能侵犯他人名誉、个人信息的,我们无法确定损害是由人工智能相关主体违反义务造成的,还是人工智能自我成长的结果。〔6〕②人工智能活动具有网络性,难以确定行为与损害之间是否存在因果关系。人工智能输出结果是由算法、算力、数据、人机交互等多种因素共同作用造成的。一旦损害发生,我们往往难以确定是何种因素决定性地导致损害。以智能家居为例,智能设备感应失灵可能是由算法错误、训练数据瑕疵、实时数据感应设备故障、远程数据处理与传输低效、第三方网络攻击等因素单独或叠加导致的,难以确定哪一个或几个因素起到实质性作用。如果无法确定损害发生原因,我们将难以判断行为与损害之间是否存在因果关系。

产品责任路径并不能有效缓解原告的过错与因果关系证明压力。在产品责任路径下,被侵权人仍需证明被告行为与己方损害之间存在因果关系(《民法典》第 1202 条)。相较而言,产品责任以"缺陷"替代"过错",似可减轻被侵权人的过错证明负担。欧盟《缺陷产品责任指令》前言第 2 条就指出,之所以将产品范围扩张至软件,是为了借助无过错责任来解决现代技术发展带来的损害分担问题。然而,缺陷认定同样具有过错色彩。对于缺陷的判断目前主要存在"消费者合理期待"与"风险—效益测试"两套判断标准。①在消费者合理期待标准下,生产者仅对可预见的危险负责。欧盟新修订的《缺陷产品责任指令》第 7 条第 2 款 b 项和 d 项均强调可预见性对于评估缺陷的意义。但可预见性本就是判断注意义务的核心标准。〔7〕在产品责任路径下,被侵权人仍应证明存在合理期待。但是,民事主体很难充分了解人工智能系统的运行机制,未必会产生合理期待。例如,消费者是否应期待人工智能和普通人一样出错,亦或人工智能是否应表现地比人更优秀,均不存在确定的答案。②风险—效益测试同样具有过错判断色彩。风险—效益测试是一种量化分析方式,高度依赖于高质量数据的积累和有效的评估模型。人工智能技术在被投入应用前,并不存在可供评估风险、收益的数据与模型。〔8〕在这一背景下,或可借助美国《侵权法(第二次)重述》第 2 节(b)条提出的替代设计标准判断是否存

〔6〕 Vgl. Käde Lisa/Stephanie von Maltzan, Die Erklärbarkeit von Künstlicher Intelligenz (KI), CR 2020, S. 66 f.

〔7〕 刘文杰:"论侵权法上过失认定中的'可预见性'",《环球法律评论》2013 年第 3 期,第 73—85 页。

〔8〕 参见林洹民:"论人工智能立法的基本路径",《中国法学》2024 年第 5 期,第 85 页。

在缺陷:如果市场上存在更为安全的合理的替代设计,就可以认定产品存在缺陷。但若如此,将导致赢家通吃的局面,不利于中小人工智能企业的发展。欧盟《缺陷产品责任指令》第 7 条第 3 款就明确规定,不得仅以市场上存在较优的产品,就将特定人工智能产品视为存在缺陷。是否存在合理的替代设计,应结合在不损害产品实用性的情况下消除产品危险的可能性、用户的风险控制能力、通过警示控制风险的可能性以及制造商的损失分散可能等多种因素综合判断。〔9〕上述因素同时也构成法学家判断行为人是否应承担注意义务的参考框架。美国学者由此认为,风险—效益测试本质上是在适用过错责任。〔10〕可见,对产品的判断均应证明生产者违反合理的注意义务,此与认定行为人是否具有过错殊途同归。〔11〕源于缺陷认定同样涉及对注意义务的判断,对人工智能侵权活动适用产品责任,并不能有效减轻被侵权人的举证压力。

(二)新型损害体系归入障碍

如何救济新型损害,是人工智能时代侵权法的重大难题。人工智能侵权引发的新型损害主要包括两类:一是基于数据处理产生的歧视、操纵等;二是人工智能活动造成的数据毁损或污染。①歧视、操纵与控制等不利益难以被认定为侵权法上的损害。损害分为财产上损害与非财产上损害,前者指丧失可转让和可替代的物质法益,包括物的损害以及为恢复健康而支出的费用;后者是指不能用金钱衡量的个人生活价值遭受的伤害,如名誉受损遭受精神痛苦等。〔12〕人工智能活动侵犯生命健康权、有形财产权等绝对权的,损害不证自明。人工智能系统不当处理个人信息的,可能使得民事主体遭遇歧视、遭受操纵与控制等不利益。〔13〕只是被侵权人通常没有遭受现实的财产减少,很难被认为遭受物质性损害。民事主体当然可以主张非物质性损害,但若没有造成“严重”的精神损害,就很难得到法院的支持(《民法典》第 1183 条)。②人工智能活动直接造成数据删除、损坏、污染的,被侵权人很难证明存在损害。鉴于目前市场上并不存在个人信息付费模式,个人要证明受损的个人信息具有现实的、确定的财产价值并不容易。〔14〕在司法实践中,有的判决以损害不存在为由,拒绝支持个人信息主体的损害赔偿请求;即便是支持损害赔偿的判决,也倾向于回避损害是否存在的问题,仅以“个人信息具

〔9〕 See John W. Wade, “On the Nature of Strict Tort Liability for Products,” *Mississippi Law Journal*, Vol. 44, No. 5, 1973, p. 825.

〔10〕 See Clayton J. Masterman and W. Kip Viscusi, “The Specific Consumer Expectations Test for Product Defects,” *Indiana Law Review*, Vol. 95, No. 1, 2020, pp. 183, 185.

〔11〕 Jean-Sébastien Borghetti, “Taking EU Product Liability Law Seriously: How Can the Product Liability Directive Effectively Contribute to Consumer Protection?” *French Journal of Public Policy*, Vol. 1, 2023, p. 34.

〔12〕 (德)埃尔温·多伊奇、汉斯—于尔根·阿伦斯:《德国侵权法——侵权行为、损害赔偿及痛苦抚慰金》(第 5 版),叶名怡、温大军译,刘志阳校,中国人民大学出版社 2016 年版,第 229 页。

〔13〕 参见林洹民:“个性化推荐算法的多维治理”,《法制与社会发展》2022 年第 4 期,第 163—166 页。

〔14〕 参见张新宝:“‘普遍免费+个别付费’:个人信息保护的一个新思维”,《比较法研究》2018 年第 5 期,第 2 页。

有财产价值”为由,酌定损害赔偿数额。〔15〕一种思路是借助对存储介质的保护救济数据权利人的损失。但当介质所有者与数据权利人并非同一主体时,就难以满足数据权利人的诉求:数据权利人无法直接提起诉讼并获得相应的损害赔偿。〔16〕被侵权人如若主张精神损害赔偿,将同样遭遇需要证明存在“严重精神损害”的困境。在现有侵权规则框架下,人工智能系统删除家庭照片、旅游视频、聊天记录等,会给民事主体带来困扰,但未必会造成严重的精神损害。

人工智能侵权的产品责任路径并未为救济新型损害扫清障碍。①产品责任并非针对歧视、操纵等不利益而设,关注重点在于人身损害与财产安全。欧盟《缺陷产品责任指令》前言第 24 条第 1 款明确指出纯粹经济损失、侵犯隐私或歧视等不适用该指令。我国《民法典》的产品责任规则虽然没有限制损害赔偿范围,但也没有为人工智能侵权引发的非物质性损害提供新型规范依据。在这一背景下,被侵权人即便主张产品责任,如果不符合《民法典》第 1183 条规定的严格的精神损害赔偿要件,也无法获得有效救济。②产品责任能否救济数据损失仍有争议。欧盟《缺陷产品责任指令》在第 6 条第 1 条 c 款规定“非用于专业目的数据破坏或毁损”构成损害。欧盟将产品责任的救济范围扩张至数据,并未获得一致认可。〔17〕且不论欧盟的创新是否合理,在我国法体系上,数据得否成为财产权的客体并未有明确答案。在《民法典》制定过程中,曾经有过将数据资产和网络虚拟财产作为一种新型知识产权客体的思路,但旋即受到激烈反对而未果。《民法典》后来在第 127 条简单地做出了开窗式的规定,即“法律对数据、网络虚拟财产的保护有规定的,依照其规定”,从而预留下制定专门规则的空间。〔18〕我国目前并未有专门的数据立法。当现行法未提供清晰的指引时,即便对人工智能侵权适用产品责任,仍难以摆脱数据损害认定困境。

(三)责任承担主体难以确定

源于人工智能活动的复杂性,责任主体常隐藏在智能生态系统内部,使得人工智能侵权责任承担成为法律适用难题。人工智能产业链条绵长,一项产品或服务涉及主体众多,包括设计者、提供者、使用者、第三方服务商、感应器制造商等复数主体。人工智能链条上任何一处出现纰漏,均可能导致损害。例如,系统更新对于支持系统稳定、修补系统漏洞极为重要,但若更新不当反倒会降低系统性能,导致出现故障(bug)。〔19〕又如,目前的人工智能活动表现出较强的人机交互性,人工智能系统并非完全依赖于硬件和软件自动运行,而是在系统与人类操作者的互动中实现预设目标。人机交互下的人工智能活动,一方面通过人类参与(human in the

〔15〕 参见江苏省南京市中级人民法院民事判决书,(2014)宁民终 5028 号;北京市朝阳区人民法院民事判决书,(2018)京 0105 民初 9840 号;北京互联网法院民事判决书,(2019)京 0491 民初 6694 号等。

〔16〕 参见纪海龙:“数据的私法定位与保护”,《法学研究》2018 年第 6 期,第 77—78 页。

〔17〕 Vgl. Georg Borges, Die Haftung für Software und KI-Systeme nach der neuen Produkthaftungsrichtlinie, CR 2025, S. 1 f.

〔18〕 参见黄薇主编:《中华人民共和国民法典侵权责任编解读》,中国法制出版社 2020 年版,第 407—410 页;龙卫球:“数据新型财产权构建及其体系研究”,《政法论坛》2017 年第 4 期,第 65 页。

〔19〕 Wolfgang Wurmnest, in: Münchener Kommentar zum BGB, 9. Auflage 2022, BGB § 308 Abs. 4 Rn. 9.

loop)防止系统失控,另一方面也使得人工智能被滥用的风险增加。〔20〕医疗人工智能可能被使用者修改参数或者通过特定群体的医疗数据进行训练,更容易出现假阳性而非假阴性等情况。只是当发生医疗事故时,被侵权人无法确定是人工智能系统自身设计出现问题,还是后续的使用行为导致损害。〔21〕因为无法确定真正的责任人,被侵权人的损害赔偿诉求很难获得法秩序的支持。

《民法典》产品责任主体范围较窄,如若扩张责任主体,则等同于让复数主体承担无过错连带侵权责任。《民法典》将产品责任主体限定为生产者与销售者(第1203条)。当发生人工智能侵权时,被侵权人可以直接向生产者或销售者问责,似可破解损害分担难题。〔22〕传统产品往往是线性生产链的产物,即产品零部件的次级供应商将零部件出售给生产商,生产商制造并将之投入市场。因为消费者无法直接接触零部件生产商,消费者应向将部件"捆绑"组成产品的终端生产者请求损害赔偿,此为产品责任限定责任主体的合理性所在。与之不同,人工智能产品具有非线性特点。消费者在购买硬件后,可自行购买智能软件,并接受第三方提供的后台支持。在数字化世界,产品会被"拆分",这使得价值链条并非线性,而呈现出一种网络性结构。产品状态的改变,一方面使得终端生产者无法如同传统产品生产者那样充分控制产品的生产与使用,由生产者对损害承担无过错责任未必妥适;另一方面也意味着消费者可以直接选择软件提供者、第三方服务商并决定是否进行升级或更新等,消费者也应为自身的决定承担不利后果。〔23〕欧盟《缺陷产品责任指令》仍坚持无过错责任,但又不能忽视智能产品的网络性特征,最终只好采取扩张责任主体的策略。欧盟将产品责任主体扩张至零部件生产者(第8条第1款),其中提供导航服务、健康监控等关联服务的主体也被视为零部件生产者(前言第17条);源于用户可以直接选择服务提供者并在一些情况下更改系统原有设计与功能,《缺陷产品责任指令》又不得不指出,对产品做出实质性更改的自然人或法人也应被视为生产者(第8条第2款)。于是,欧盟产品责任主体包括产品生产者、零部件生产者(软件生产者、关联服务提供者)和进行实质性修改的用户。又根据该指令第12条,多个生产者应对损害承担连带责任。据此,欧盟新的产品责任本质是通过将相关主体都界定为生产者并承担连带责任的方式解决责任分担问题,原有限定责任主体、由终端生产者或销售者承担无过错责任的制度设计落空。网络型商品中的任何一方主体均无法控制商品的表现,简单地罗列可能的责任主体并要求其对

〔20〕 Miriam C. Buiten, "Product Liability for Defective AI," *European Journal of Law and Economics*, Vol. 57, Issue 1-2, 2024, pp. 239, 262.

〔21〕 See Taro Makino et al., "Differences Between Human and Machina ne Perception in Medical Diagnosis," *Scientific Reports*, Vol. 12, No. 1, 2022, pp. 1-13.

〔22〕 参见戴昕:"无过错责任与人工智能发展——基于法律经济分析的一个观点",《华东政法大学学报》2024年第5期,第47页;徐伟:"生成式人工智能服务提供者侵权归责原则之辨",《法制与社会发展》2024年第3期,第190—204页。

〔23〕 Vibe Ulfbeck, "Product Liability Law and AI: Revival or Death of Product Liability Law," in Ernest Lim and Phillip Morgan (eds.), *The Cambridge Handbook of Private Law and Artificial Intelligence*, Cambridge: Cambridge University Press, 2024, p. 221.

损害承担连带责任，未必合理。

综上所述，产品责任路径不能有效解决人工智能侵权引发的过错与因果关系证明、新型损害归入以及责任承担主体确定三大难题。我国不应受到布鲁塞尔效应的影响，亦步亦趋地效仿欧盟，通过改革产品责任的方式回应挑战，而应回归问题本身，在责任成立与承担层面有针对性地设计破壁性规则，减轻被侵权人的举证压力，救济被侵权人的损失。

三、责任成立：认定人工智能侵权的举证便利性规则

人工智能活动具有自主性、网络性、数据依赖性等特征，引发诸多举证难题。与其修改《产品质量法》、借助产品责任调整人工智能活动，不若直接“对症下药”，在《人工智能法》中配置信息披露规则、损害与因果关系认定规则，解决侵权责任成立证明难题。

（一）过错责任下的信息披露义务

考虑到人工智能应用带来的过错认定障碍，欧盟尝试借助作为无过错责任的产品责任降低被侵权人的举证难度。但是，无过错责任并不符合私法自治理念，对其适用应慎之又慎。下文将首先讨论人工智能应用侵权的归责原则，在此基础上分析应如何降低被侵权人的过错证明难度。

1. 人工智能侵权责任的过错归责原则

人们容易直观地接受一种观点：人工智能应用带来新型风险，法秩序应对引发风险的人工智能活动适用危险责任。人工智能活动引发的风险未必属于民法上的危险。民法上的危险，典型如《民法典》第 179 条规定的“消除危险”，针对的是具体事实之危险，且危险是现实存在的、迫切的，对他人人身、财产安全造成现实的威胁。〔24〕《民法典》中的产品责任、高度危险责任等均针对的是具体的危险而非抽象的风险。与之不同，当我们讨论新技术产生的风险时，多指涉的是一种集合意义上的抽象的事件，而非个别、具体的危险。〔25〕不能简单地以人工智能应用带来新型风险，就强调应对人工智能活动适用危险责任。即使是高风险人工智能系统的研发与应用，也仅在能清晰地界定具体的、现实的高危险行为前提下，始得适用危险责任。若抽象地认为“对高风险人工智能活动适用危险责任”，将因无法界定高风险人工智能活动的类型与范围而在事实上欠缺可执行性。

对人工智能侵权适用无过错责任，未必有利于降低事故发生频率，反倒可能增加总体事故成本。主张适用无过错责任的重要理由是，无过错责任能够提供可预期的外部环境，激励行为人采取合理的预防措施最小化社会成本（事故成本、信任环境、预防成本等）。〔26〕无过错责任

〔24〕 黄薇主编：《中华人民共和国民法典总则编解读》，中国法制出版社 2020 年版，第 581—582 页。

〔25〕 周学峰：“论人工智能的风险规制”，《比较法研究》2024 年第 6 期，第 46 页。

〔26〕 See Hans-Bernd Schäfer and Claus Ott, *The Economic Analysis of Civil Law*, Cheltenham: Edward Elgar Publishing, 2022, pp. 149-150.

虽然可以使得负外部性内部化，但正外部性不会全部回流给参与者。〔27〕即便承担无过错责任的民事主体积极提升系统安全水平，如果人工智能活动链条中的其他相关主体怠于作为，也无法有效地控制危险。人工智能活动涉及算法、算力、数据等多个要素。人工智能系统控制者无力充分审查从第三方获得的训练数据、传感数据是否具有瑕疵；当存在算法交互时，也没有人能确保算法设计准确。〔28〕如果损害难以避免，理性的责任主体将怠于通过技术与管理手段降低事故发生概率，转而倾向于通过价格机制分散事故成本，最终导致事故成本上升、社会总体福利下降。

与之相较，过错责任能够全面评价民事行为，激励相关主体采取措施防止损害发生。过错责任能够细致评估各方行为，筛选出应被问责的主体。以大语言模型为例介绍之。①在设计层面，人工智能系统由基础大语言模型和具体模型结合构成。通过海量数据训练出的大模型提供底层逻辑支持，具体模型则通过“术业有专攻”的专业优化训练，适配众多的具体行业和场景。在这一协同应用关系中，系统引发的风险具有叠加性。只有适用过错责任，使得每个民事主体为自己的过失行为负责，才能恰当地追究不同上下游开发者的法律责任。②在应用层面，人工智能大模型使用者可进一步细分为前端使用者与后端运营者，前者有权决定并使用人工智能系统，后者持续定义人工智能的技术特征并提供必要的服务。前端使用者虽然具有所有者或保管者身份，但人工智能系统的运行离不开后端运营商的服务，且后者从持续提供服务中获益。在过错责任模式下，法律人将在整个应用脉络中寻找责任人，要求相关主体根据自身的过错程度承担损害赔偿责任。法经济学研究也指出，当侵权纠纷中相关主体的补充过错(complement care)较之替代过错(substitute care)更为普遍时，亦即损害并非由单一行为造成时，无过错责任容易降低整个社会对人工智能应用的注意标准；与之相反，过错责任将激励每一个相关主体采取“相辅相成的努力”防止损害的发生。〔29〕

过错推定责任可能使中小人工智能企业陷入责任困境，并非立法的理想选择。过错推定责任作为过错责任的一种类型，既有助于减轻原告的举证压力，又不至于对人工智能企业责之过苛，似可成为调整人工智能活动的折中方案。我国《个人信息保护法》第69条第1款规定的个人信息侵权规则也采过错推定责任。立法若对人工智能侵权适用过错推定责任，也能与《个人信息保护法》的规定保持一致。然而，人工智能产业链条绵长，一项产品或服务涉及主体众多，相关主体若均需自证无过错，将产生巨大的交易成本。鉴于人工智能系统具有自主性与不透明性等特性，人工智能活动参与方证明自身无过错的难度极高，将产生寒蝉效应，阻碍中小企业进军人工智能行业。这一担忧并非杞人忧天。我国个人数据交易之所以不畅，过错推定

〔27〕(瑞)米里亚姆·布伊滕、(比)亚历山大·德·斯特里尔、(德)马丁·佩茨：“人工智能责任的法律与经济学”，张韬略、陈沪楠译，《上海政法学院学报》2024年第4期，第141页。

〔28〕(美)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》，任莉、张建宇译，人民邮电出版社2017年版，第169页。

〔29〕See Miriam Buiten, Alexandre de Streel and Martin Peitz, “EU Liability Rules in the Age of Artificial Intelligence,” 2021, pp. 43-46, https://cerre.eu/wp-content/uploads/2021/03/CERRE_EU-liability-rules-for-the-age-of-Artificial-Intelligence_March2021.pdf, last visited on 13 February 2025.

责任是重要原因之一。当数据交易侵犯个人信息权益时,数据需方若无法自证已经充分履行个人信息保护义务,将承担相应的法律责任。这使得数据需方惮于进行个人数据交易。〔30〕若对人工智能侵权适用过错推定规则,中小企业可能因过错推定责任而心生忌惮。过错推定责任将成为人工智能活动的"紧身衣"。对于新技术研发与应用,妥适的方式是采用"反应性管制"模式,即借助过错责任规则加强与业界沟通,在判断是否有过错过程中形成对行为的共同理解,逐渐形成新的标准与规范。〔31〕在人工智能发展初期,不应贸然采纳过错推定责任,不妨留待产业发展与经验累积后再考虑是否推行之。

保险制度往往成为对侵权行为适用严格责任的重要理由,但在人工智能发展初期,保险机制并不能充分发挥作用。无论是无过错责任还是过错推定责任,若可配之以有效的保险制度分散成本,也具有采纳的合理性。遗憾的是,目前市场上缺乏建立人工智能侵权保险模型的有效数据。人工智能属于新兴技术,相关应用尚未全面推广,损害赔偿案例较为稀缺。无论是推广自愿保险(市场驱动)还是强制保险(国家监管),由于缺乏事故与损害的统计数据,都无法进行有效的风险边际计算。〔32〕考虑到人身财产损害的赔偿数额较高,保险公司会更倾向于对人工智能活动收取较高的保险费用。欧盟委员会的评估报告指出,严格责任模式下的人工智能企业应支付的保费每年将增加 35%。〔33〕大型企业或可承受较高的保险费用,但潜在创业者将因过高的保费而对是否从事人工智能研发与应用举棋不定。由是观之,保险机制不仅难以有效分散风险,反倒可能成为阻碍中小企业进军人工智能行业的制度壁垒。

综上,人工智能侵权仍应采过错归责原则,且不应贸然规定过错推定责任。至于人工智能特性引发的过错证明难题,法秩序应设计专门的信息披露规则予以化解。

2. 便利被侵权人举证的证据开示规则

在过错责任模式下,原则上应由原告举证证明被告具有过错。在人工智能侵权纠纷中,原告要证明被告具有过错并不容易。人工智能活动记录是帮助查清真相的重要媒介。只是被告常以涉及企业商业秘密为由,拒绝人工智能系统相对方查阅信息的请求。

为了解决证据偏在问题,司法机关创设书证提出命令制度,但运行并不顺利。在现代型诉讼中,关键型证据多由被告掌握,证据偏在现象明显。原告因不掌握关键证据,难以提起诉讼,即便提起诉讼也常遭遇败诉。〔34〕书证提出命令制度应运而生。2015 年发布的《最高人民法院关于适用〈中华人民共和国民事诉讼法〉的解释》(法释〔2015〕5 号)第 112 条正式规定了文书提出命令制度;2019 年修正的《最高人民法院关于民事诉讼证据的若干规定》(法释〔2019〕19 号)第 45 条至第 48 条将书证提出命令制度予以具体化与细致化。当出现证据偏在时,法

〔30〕 参见林洹民:"数据产权登记的私法定位与制度设计",《法商研究》2024 年第 5 期,第 93—94 页。

〔31〕 See William McGeveran, "Friending the Privacy Regulators," *Arizona Law Review*, Vol. 58, No. 4, 2016, pp. 997-1003.

〔32〕 See Michael G. Faure, "Economic Criteria for Compulsory Insurance," *The Geneva Papers on Risk and Insurance: Issues and Practice*, Vol. 31, No. 1, 2006, pp. 149-168.

〔33〕 European Commission, supra note 3, p. 54.

〔34〕 参见肖建国:"现代型民事诉讼的结构和功能",《政法论坛》2008 年第 1 期,第 112—123 页。

院根据一方当事人申请，向控制书证的对方当事人或者第三人发出提交书证的裁定；书证控制人若拒绝执行，将承担不利法律后果。如果能在人工智能侵权案件中适用书证提出命令，法院将根据原告的申请裁定是否要求被告披露技术文件、系统日志等信息，帮助查清事实，减轻原告的举证压力。然而，实践中法官发出书证提出命令的情况极少，“书证提出命令司法适用率低、准许率低”。〔35〕这一现象也具有一定的合理性，毕竟书证提出命令制度是对“谁主张谁举证”原则的修正。我国学者也强调，一般性的事案释明并不适合我国的诉讼实践，应将书证提出义务限定在法律规定的情形下。〔36〕如果没有法律的明确授权，司法机关多不敢发布书证提出命令。

我国未来的人工智能法应规定特定条件下的证据开示规则，为法院发布书证提出命令提供明确的规范依据。在我国现行法体系下，即便两造之间存在合同关系，被告也不负有解释损害发生原因的义务。消费者有权了解商品和服务的相关信息（《消费者权益保护法》第8条第2款），远程销售或金融服务经营者还应额外提供安全注意事项和风险警示、售后服务、民事责任等信息（第28条）。但是，消费者无权了解行为人的日常经营和运行情况。只有通过《人工智能法》创设证据开示规则，为司法机关发布书证提出命令提供实体法依据，才能有效解决人工智能侵权中的证据偏在问题。欧盟《缺陷产品责任指令》第9条就专门规定证据开示规则，允许被侵权人提出申请并由法院判断是否核准披露证据，从而推动自下而上的人工智能善治。我国未来的人工智能立法可借鉴欧盟的立法经验，允许人工智能活动相对人在特定条件下向法院提出申请，由法院发出书证提出命令，要求相关主体提供活动记录；如果义务人拒不提供相关信息，即推定其具有过错。

证据开示规则毕竟是对经典证明责任理论的偏移，惟有满足特定条件始得适用。首先，原告申请公开的应是法律明确要求被申请人记录并保存的信息。人工智能活动参与方不负有一般性的说明义务。被侵权人不能漫无边际地要求人工智能系统提供者或使用者披露所有信息。被侵权人仅应请求开示被申请人有义务记录的信息，以核验被告是否履行数据管理、人工监督义务、合规评估等法定义务。其次，被侵权人应提出合理的事实和证据，证明存在“合理怀疑”。证据开示并非是被侵权人的权利，人工智能相关主体是否有义务提供证据，应由法院予以判断。被侵权人应证明存在合理怀疑，但因证据偏在无法进一步查清真相。法院接收并审查被侵权人的申请，裁定是否发出书证提出命令。只有对证据开示规则进行必要的控制，才能防止造成滥诉，严重干扰人工智能企业的正常经营。最后，只有在原告已尽一切适当之努力仍无法获得足够证据情况下，才能请求法院发布书证提出命令。证据开示规则旨在平衡两造之间的诉讼力量，如果原告可以轻易获得相关证据（如被告使用开源算法），不应额外增加被告的

〔35〕 参见潘剑锋、牛正浩：“书证提出命令程序性制裁理论检视——以商业秘密侵权诉讼为切入”，《政法论丛》2021年第5期，第61—71页。

〔36〕 参见陈杭平：“‘事案解明义务’一般化之辨——以‘美国事证开示义务’为视角”，《现代法学》2018年第5期，第159—169页；曹建军：“论书证提出命令的制度扩张与要件重构”，《当代法学》2021年第1期，第128—139页。

负担。

(二) 精神损害赔偿范围拓展

如何救济新型损害,是人工智能侵权案件处理的又一难点。一方面,我国未来的《人工智能法》不应忽视虚拟空间中日益增长的损害赔偿诉求;另一方面,立法也应避免过于苛责,致使中小人工智能企业动辄得咎,损及行为自由。

1. 物质性损害体系归入困境

在对物质性损害的判断上,差额说一直居于通说地位,客观损害说、规范损害说构成对差额说的必要补充。差额说下的损害是指财产的实际状态与损害事件未发生时财产状态之间的差额;〔37〕客观损害说与规范损害说则在承认差额说基本内涵的基础上,分别强调单个、具体之物的损害以及规范目的重要性。〔38〕无论采何种损害赔偿学说,在数据被滥用或破坏情况下,都很难证明存在物质性损害。

数据被不当处理的,即便民事主体可能遭受歧视、被操纵等风险,因民事主体的财产状态并未发生变动,难以被认定存在物质性损害。为了解决这一难题,学理上有以“风险作为损害”的革新损害概念的观点。〔39〕但也有学者指出欠缺确定性的风险不能被归为损害。〔40〕风险不同于危险,前者具有抽象性、不确定性,后者则更为具体与紧迫。如果没有遭受现实的损害或支出防止损害发生的费用,仅存在风险而非危险,司法机关难以认定存在“实际损害”,至多只能认可存在“对未来损害的猜测”。根据差额说,此时并不存在现有状态与应有状态之间的价值变动差额;无论是根据客观损害说还是规范损害说,因为不确定风险并不贬损财产的实际使用状态,也不存在第三方介入控制损害的情况,民事主体并未遭受客观损害和规范损害。是以,民事主体仅仅遭受歧视、被操纵等风险的,难以证明存在值得法律救济的物质性损害。实际上,法秩序对风险活动多借助事前管控手段予以调整。风险表现为去个体化特征,亦即不是从个人损害救济的角度而是从整体可控的视角调整风险行为。〔41〕对于产品安全风险,我国出台专门的《产品质量法》事前地控制风险,在存在现实的损害后才适用《民法典》中的产品责任规则。同理,对于人工智能活动引发的歧视、被操纵等风险,更为便捷的路径是借助监管部门发布的《生成式人工智能服务管理暂行办法》《互联网信息服务算法推荐管理规定》等管理性规定予以规范。

数据被删除、污染或毁损的,并不能表征存在物质性损害。其一,数据被破坏并不当然意味着存在物质性损害。数据的价值是在使用过程中体现的,并不存在客观的市场价格,贴现收

〔37〕 王泽鉴:《损害赔偿》,北京大学出版社 2017 年版,第 63 页。

〔38〕 参见曾世雄:《损害赔偿法原理》,中国政法大学出版社 2001 年版,第 124—128 页;徐建刚:“《民法典》背景下损害概念渊流论”,《财经法学》2021 年第 2 期,第 31—45 页。

〔39〕 参见谢鸿飞:“个人信息处理者对信息侵权下游损害的侵权责任”,《法律适用》2022 年第 1 期,第 23—36 页;朱晓峰、夏爽:“论个人信息侵权中的损害”,《财经法学》2022 年第 4 期,第 51—66 页。

〔40〕 参见程啸、曾俊刚:“个人信息侵权的损害赔偿责任”,《云南社会科学》2023 年第 2 期,第 99—110 页。

〔41〕 刘刚编译:《风险规制:德国的理论与实践》,法律出版社 2012 年版,第 196 页。

入法、成本法、市场参照法、多维度定量评估等传统定价方法对确定数据价格均存在较大局限。〔42〕因数据是否具有价值尚不确定，被侵权人难以主张现有状态与应有状态间存在差额，也无法证明使用价值受到损害（客观损害说）或在规范层面存在损害（规范损害说）。其二，数据之上承载个人信息权益，数据被破坏属于民事权益受到侵犯，不能同时被认定存在损害。《民法典》第1165条第1款规定："行为人因过错侵害他人民事权益造成损害的，应当承担侵权责任。"《民法典》将"民事权益"与"损害"并列。由是观之，民事主体证明民事权益受到侵害后，还应证明遭受损害。当数据受到破坏时，并不能当然推定存在损害。《个人信息保护法》第69条第1款同样将"个人信息权益"与"损害"并列，表明个人信息主体在证明个人信息权益受到侵犯后，仍应证明存在损害。是以，数据受损也不能构成损害本身，被侵权人仍应证明存在物质性或非物质性损害。

2. 由"严重"转向"显著"的精神损害赔偿

与其扩张物质性损害的边界，不若摒弃精神损害赔偿认定的严重性标准，便利被侵权人提起精神损害赔偿之诉。学者革新物质性损害的重要原因是，只有严重的精神损害赔偿（如精神类疾病）才受到侵权法的救济。被侵权人因数据处理行为遭受歧视、操纵，或者对其具有明显意义的数据被破坏的，往往会产生恐惧、痛苦、愤怒等情绪。只是若未达至严重精神损害，被侵权人就难以提起精神损害赔偿之诉。法秩序之所以严格限制精神损害赔偿的适用，在于担心精神损害赔偿被滥用。〔43〕在人工智能活动明确违反《个人信息保护法》等法律的情况下，要求被告承担损害赔偿责任，并不会导致责任泛滥，毕竟被告的行为本就具有可谴责性。欧盟《通用数据保护条例》第82条第1款就没有将非物质损害的适用限定在"严重精神损失"条件下，而是指出"因违反本条例规定遭受物质或非物质损害的任何人有权从控制者或处理者处获得赔偿"。欧盟法院在2024年最新的判决中也指出，原告主张非物质损害赔偿的，只要证明被告进行了原告明确反对的数据处理行为使其丧失了对数据的控制权即可，无需证明损害达至超过一定程度的严重性。〔44〕美国法院在判断是否存在严重精神损害时，也不再要求出现身体症状，只要被告的行为足够极端就可以推定原告遭受严重的精神痛苦。〔45〕将精神损害赔偿的适用僵化限定在"严重精神损害"条件下，已经有些不合时宜。《民法典》第1183条虽然规定只有对严重精神损害才能进行赔偿，但立法释义也指出，"对严重的理解，应采容忍限度理论，即超出社会一般人的容忍限度，就认为是严重"。〔46〕我国人工智能立法不妨摒弃认定精神损害赔偿的严重性要件，适当扩张精神赔偿的范围，消除新时代侵权法救济不足之弊端。

只要个人数据对民事主体显而易见的重要，能引起广泛的认同，就应推定精神损害存在。

〔42〕 许伟、刘新海："中国数据市场发展的主要障碍与对策"，《发展研究》2022年第7期，第47页。

〔43〕 参见黄薇，见前注〔18〕，第79页。

〔44〕 GP v. juris GmbH, Case C-741/21, ECLI:EU:C:2024:288[2024].

〔45〕 （美）丹·B. 多布斯：《侵权法（下册）》，马静、李昊、李妍、刘成杰译，中国政法大学出版社2014年版，第718—719页。

〔46〕 参见黄薇，见前注〔18〕，第79页。

随着社会变迁与价值观念的变化，精神损害赔偿范围的扩大已成为现代民法发展的重要课题。〔47〕《德国民法典》第 651n 条第 2 款规定，如果旅行失败，旅游消费者可以就浪费的休假时间要求适当的金钱补偿。在旅行失败情况下，消费者因没有如愿休假可能产生不满、愤懑与失望等复杂情绪，但未必会遭受严重的精神损害。尽管如此，德国法学认为，给予消费者精神损害赔偿，并不会引起质疑，反倒容易引起大众共鸣。〔48〕我国对于适用精神损害赔偿一直持谨慎态度，但也逐步认可因性别、残疾或疾病等原因遭受就业歧视时的精神损害赔偿。〔49〕是否支持精神损害赔偿，关键不在于精神损害是否严重，而在于精神损害是否足够显著能获得社会一般人的认同。〔50〕当求职者因性别、疾病等遭受就业歧视时，法院支持精神损害赔偿的，容易获得广泛的支持。同理，当人工智能系统不当处理或破坏数据时，法院是否应支持精神损害赔偿，也应结合数据的典型意义进行判断。家庭照片、旅游视频等个人数据对民事主体的情感价值显而易见。对此类数据的滥用或破坏难以为民事主体所容忍，法院认可精神损害赔偿的，也多不会遭遇质疑。与之不同，购物记录、浏览记录等个人数据虽然同样受到《个人信息保护法》等法律的保护，但对民事主体的价值相对有限。除非民事主体能够证明上述数据存在特定意义，否则对数据的滥用或破坏不应导致精神损害赔偿。至于精神损害难以计算的问题，并非数据类侵权独有的问题。2020 年修正的《最高人民法院关于确定民事侵权精神损害赔偿责任若干问题的解释》（法释〔2001〕7 号）第 5 条规定六种参考因素，可为必要的指导。

争议在于，是否应将精神损害赔偿限定在行为人具有故意或重大过失的条件下。一种观点认为，附着人格权益的个人数据类似于具有人身意义的特定物，被侵权人应仅在行为人具有故意或重大过失的前提下才能提起精神损害赔偿（《民法典》第 1183 条第 2 款）。从规范目的上看，《民法典》第 1183 条第 2 款之所以要求行为人具有故意或重大过失，在于“法律不能一般性地期待侵权人认知特定物对被侵权人具有人身意义”。〔51〕个人数据不同于特定物，前者人身意义明显，后者的人身属性则具有一定的隐蔽性。侵权人借助人工智能处理个人数据的，理应认识到可能侵犯人格权益，此时并不存在保护行为人合理期待的问题。是以，将精神损害赔偿的适用限定在行为人具有故意或重大过失情况下与立法精神并不契合。

（三）消费者保护目标下的因果关系推定

如何证明被告行为与原告损害之间存在因果关系，是人工智能侵权纠纷解决的又一难点。人工智能活动具有网络性，复数主体的行为相结合共同推动人工智能应用。即便被侵权人指出侵权人的特定行为可能使得人工智能作出致害输出，但若无法达致高度盖然性标准，就难以

〔47〕 参见洪国盛：“民法典的精神损害赔偿体系——以功能主义为视角”，《法学研究》2024 年第 4 期，第 94—112 页。

〔48〕 Vgl. Klaus Tonner, in: Münchener Kommentar zum BGB, 9. Aufl., 2023, § 651n, Rn. 62 f.

〔49〕 参见北京市朝阳区人民法院民事判决书，（2008）朝民初字第 06688 号；北京市第三中级人民法院民事判决书，（2016）京 03 终 195 号；广东省广州市中级人民法院民事判决书，（2016）粤 01 民终 10790 号。

〔50〕 参见彭诚信：“数字法学的前提性命题与核心范式”，《中国法学》2023 年第 1 期，第 102 页。

〔51〕 参见黄薇，见前注〔18〕，第 80 页。

被认为完成举证义务。〔52〕

一种思路是通过完全的举证责任倒置破解人工智能侵权中的因果关系证明障碍。《民法典》第1230条就规定应由污染环境方证明行为与损害之间不存在因果关系。立法释义指出，环境污染损害具有长期性、潜伏性、持续性、广泛性等特点，且容易出现多因一果的现象，判断是否存在因果关系必须具备专门知识，因此有实行因果关系举证责任倒置的必要性。〔53〕人工智能致损同样具有持续性与广泛性，致害过程也具有高度专业性与复杂性等特点，查清因果关系较为不宜，似可通过因果关系举证责任倒置的方式减轻被侵权人的举证负担。欧盟就曾设想在人工智能侵权中推行完全的因果关系推定，但旋即放弃。完全的因果关系举证责任倒置虽然可以救济被侵权人的损失，但也将产生显著的负面经济影响，如增加人工智能企业法律成本、导致滥诉等；长期看也不利于增加消费者整体福利，毕竟企业成本上涨最终会体现到产品或服务价格上，同时也将减少消费者免费接触新产品的机会。〔54〕妥适的方式是将举证责任倒置限定在特定条件下，以避免过度阻碍人工智能产业发展。

法秩序应确立企业与消费者之间发生人工智能侵权纠纷时的因果关系推定规则。人工智能生态系统呈现出高度网络化与专业性特征，在人工智能生态系统外的消费者很难理解系统运行逻辑，难以证明企业行为与己方损害后果之间存在因果关系。相较而言，人工智能企业处于证明因果关系不存在的有利位置。法秩序应将举证责任分配给最有可能查清真相的一方。欧盟《缺陷产品责任指令》第10条第4款规定，当法院确定地发现，因果关系因科技或组织的复杂性等因素难以被查清时，基于保护消费者的考虑可推定因果关系成立。当人工智能应用侵犯消费者合法权益时，应推定因果关系成立，转由被告证明因果关系不成立，以此化解举证难题，增加公众对人工智能应用的信任。〔55〕与之不同，如果是企业之间或个人之间的纠纷，仍应适用“谁主张谁举证”的经典证明责任规则。在个人之间的人工智能侵权纠纷中，并不存在偏袒某一方的必要；在企业间的相关纠纷中，原告同样未必处于举证弱势地位。以企业间的环境侵权纠纷为例，原告较之被告可能更有能力证明因果关系的存在。在“通诚煤炭储运有限公司与益强包装制品有限公司环境污染损害赔偿纠纷案”中，益强公司主张煤炭厂产生的灰尘对其生产的药品包装产品造成污染，导致其医药产品出现质量问题。生产医药产品的益强公司更有能力证明灰尘对其产品质量的影响，但因该案属于环境侵权纠纷，法院要求被告通诚煤炭储运有限公司承担举证责任，并承担举证不能的不利后果。〔56〕立法一刀切地逆转因果关系证明，可能有失偏颇。同理，企业间发生人工智能侵权纠纷的，原告相较于被告也未必处于举证的不利地位。只有当人工智能侵权纠纷发生在企业与消费者之间时，才应推定因果

〔52〕 参见张卫平：《民事诉讼法》（第六版），法律出版社2023年版，第286页。

〔53〕 参见黄薇，见前注〔18〕，第246—247页。

〔54〕 European Commission, Commission Staff Working Document Impact Assessment Report Accompanying the Document Proposal for a Directive of The European Parliament and of the Council on Liability for Defective Products, SWD/2022/316, pp. 46-54.

〔55〕 Vgl. Dirk Staudenmayer, Haftung für Künsterliche Intelligenz, NJW 2023, S. 901.

〔56〕 参见河南省平顶山市中级人民法院民事判决书，(2015)平民三终字第126号。

关系成立。

需要说明的是,即便适用因果关系推定规则,原告至少应证明被告行为与己方损害之间存在关联性。若原告完全没有指出这一可能性,案件都无法进入诉讼程序。在环境侵权案件中,最高人民法院就指出原告应提供被告行为与损害之间具有关联性的证据,并结合污染环境、破坏生态的行为方式、污染物的性质、环境介质的类型、生态因素的特征、时间顺序、空间距离等因素,综合判断被告行为与损害之间的关联性是否成立。〔57〕在人工智能侵权纠纷中,被侵权人也应证明被告行为与损害之间存在一定程度上的关联性,如存在同批次产品造成损害的报道或并不存在其他异常因素等,以使得案件顺利进入司法程序。

综上所述,我国未来的《人工智能法》应通过专门的信息披露、精神损害认定与因果关系推定等举证便利性规则,减轻被侵权人责任成立证明压力。具体条文可如下表述:

第 n 条 人工智能系统相关主体因过错侵犯他人民事权益造成损害的,应承担侵权责任。

被侵权人已尽一切适当努力仍无法获得证明侵权行为相关证据的,可以向人民法院提出申请,由人民法院裁定人工智能系统相关主体开示记录与保存的相关信息。

第 n+1 条 人工智能系统相关主体侵害自然人人身权益造成明显的精神损害的,被侵权人有权请求精神损害赔偿。

人工智能系统侵犯消费者合法权益的,行为人对行为与损害之间不存在因果关系承担举证责任。

四、责任承担:损害分担的人工智能特殊多数人侵权规则

在化解人工智能侵权责任认定困境后,仍有必要进一步讨论责任承担问题。考虑到人工智能的特性,很可能出现下列情事:当发生人工智能应用致损时,真正的侵权人隐匿在复杂的人工智能活动链条中,最终导致“侵权行为人消失”。〔58〕

共同危险责任规则并不适用于人工智能侵权责任源头不清的情况。《民法典》第 1168 条及其以下规定的多数人侵权规则,本就有便利被侵权人举证的功能。《民法典》第 1170 条规定:“二人以上实施危及他人人身、财产安全的行为,其中一人或者数人的行为造成他人损害,能够确定具体侵权人的,由侵权人承担责任;不能确定具体侵权人的,行为人承担连带责任。”该条旨在解决无法查清侵权人时的责任分配问题,似可适用于人工智能致损源头不明的情况。然而,共同危险责任要求行为人皆为义务违反行为,亦即多个行为人都制造了值得法律谴责的危险,只不过其中一人的危险最终转变为损害。典型例子是,多人向马路方向射击,其中一发子弹击中路人。行为人向马路方向射击,均违反注意义务,行为均具备法律上的可谴责性,此为法秩序严厉对待行为人的法理基础。但在人工智能应用中,参与方的行为并不必然具有可

〔57〕 参见《最高人民法院关于生态环境侵权民事诉讼证据的若干规定》(法释〔2023〕6 号)第 5 条。

〔58〕 Vgl. Spiecker gen. Döhmann, Zur Zukunft systemischer Digitalisierung-Erste Gedanken zur Haftungs-und Verantwortungszuschreibung bei informationstechnischen Systemen, CR 2016, S. 698, 700.

谴责性。如果不能确定参与方是否均进行义务违反行为，就不能仅因无法查清责任源头而适用共同危险责任。

（一）同一商业技术单元连带责任规则

侵权法并非仅根据行为，也可依据特定的利益结合状态，要求相关主体承担法律责任。《民法典》中的雇主责任、监护人责任等，均是因为存在紧密的利益结合关系，雇主或监护人才应为他人行为承担法律责任。《民法典》第1254条规定的高空抛物责任也是一种根据结合状态承担法律责任的侵权类型。只不过建筑物使用人因居住形成一种偶然结合关系，居民之间大概率互相都不熟悉。当难以确定具体责任人时，仅出于松散的邻里关系就要求可能加害的建筑物使用人给予补偿，并不能获得人们的普遍认同。〔59〕与之不同，如果企业通过合同或其他协议结合在一起，有共同的商业利益与技术基础，即便各个商业单元之间仍保持独立，当无法查清责任主体时，要求各个成员共担风险也存在事理上的合理性。

当责任源头无法查清时，立法可要求构成同一商业技术单元的参与者对损害承担连带责任。为了防止责任隐藏于商业链条中，当数个主体在合同或类似基础上紧密合作形成商业技术单元时，若被侵权人可以证明该商业技术单元造成损害，尽管无法证明单元内部哪个具体环节出现问题，仍可以让该商业技术单元成员对损害承担连带赔偿责任。〔60〕例如，制造商A生产的智能安全系统被安装到由B设计的智能家居系统中，该系统又在C研发的人工智能生态系统上运行，三者之间存在合同协议，共同组成一个商业技术单元。当发生一起入室盗窃案时，智能安全系统未启动最终给用户造成重大损失。虽然不清楚故障发生在哪一环节，但确定损失是因智能安全系统未充分发挥功能造成的，此时同一商业技术单元成员就应对损失承担连带损害赔偿责任。市场主体为了共同的商业目的主动结合的，可以通过合同或类似方式控制互动与互操作性风险，并提前就事故成本的分配达成一致。让同一商业单元成员承担连带责任，并不会引发群体性抵触情绪。借助同一商业技术单元这一概念，既可以防止责任漫无边际，又能有效解决无法查清责任源头时的责任分配问题。

要求同一商业技术单元成员对损害承担连带责任，是责任法对人工智能时代商业模式演变的积极回应。随着人工智能、数字化技术对社会的深度浸润，商业模式也在发生变化。介于组织与合同之间的新型商业结构正在形成。按照交易成本理论，如果利用市场协调民事主体之间的交易费用（如发现价格的成本以及谈判费用等）大于利用一体化组织的协调费用，为了降低交易费用，就会出现一体化组织，也即企业。〔61〕在网络信息技术的加持下，市场主体希望最大限度地利用市场与组织的优点，于是出现介于合同与组织的第三种商业模式“混合模

〔59〕 参见张新宝：《中国民法典释评·侵权责任编》，中国人民大学出版社2020年版，第292页；邹海林、朱广新主编：《民法典评注.侵权责任编(1)》，中国法制出版社2020年版，第840页。

〔60〕 See Expert Group on Liability and New Technologies—New Technologies Formation, supra note 2, p. 56.

〔61〕 参见(美)奥利弗·E. 威廉姆森、西德尼·G. 温特：《企业的性质》，姚海鑫、邢源源译，商务印书馆2010年版，第25—26页。

式”(hydrids)。〔62〕在混合模式中,各个成员既相互依存又相互独立,在降低市场寻价成本、违约风险的同时保障合作者的灵活性和创造性。混合模式为科技公司广泛采用,我国学者也适时地强调新时代合同的组织经济功能。〔63〕但是,混合模式有遮蔽侵权行为、掩盖责任人的副作用。以往的产品或服务多是由一个可明确识别的民事主体最终提供,问责机制相对清晰。在混合模式下,人工智能产品或服务均处于智能生态系统中,该生态系统由设计者、提供者、系统更新或升级服务商、后端运营者等成员组成,成员之间存在密切关联与互动,行为隐藏在复杂的生态系统内部。一旦发生侵权情事,系统外民事主体不能清晰地查清致害机制。经济模式进化提升生产效率,但不应让被侵权人承受商业模式改变带来的责任不清风险。如果智能生态系统中的成员可借助复杂生态系统逃脱责任,将会产生逆向激励:本来可以由单一企业完成的任务,会被尽可能地拆分给复数独立成员,从而掩盖因果关系,帮助潜在侵权人逃脱法律责任。

(二)同一商业技术单元成员的认定标准

困难在于如何界定同一商业技术单元的成员。合作紧密的企业可通过较低的交易成本预先达成内部责任分担机制,如签署赔偿条款等,从而避免遭遇“无妄之灾”。即便应对被侵权人承担连带责任,因为各企业已经对责任有所预期与安排,成员对于责任分担不会有太强的抵触情绪。高空抛物责任的问题不会在紧密合作的企业之间发生。反之,如果企业合作并不密切,不同企业在产业链上没有太多沟通机会,沟通责任机制的交易成本较高,无法通过合同或其他方式协调责任分担。一旦损害发生,要求各个企业承担连带责任遭受抵触的可能性就较大。企业如果认为自身并没有过错反倒应为损害赔偿,可能会惮于进一步从事商业行为,最终退出交易市场,这将对企业与市场产生明显的消极影响。〔64〕由是观之,判断企业是否属于应承担连带责任的同一商业技术单元,本质应是判断企业间的结合是否紧密,内部交易成本是否足够低,可以事先形成一套责任分担方案。

法秩序应结合是否存在持续性的合同或类似协议、是否存在商业与技术依存关系等因素,衡量企业间的交易成本,判断是否构成同一商业技术单元。一般看来,民事主体之间仅存在持续性的合同或类似协议的,并不能表明存在相互依存的紧密合作关系。民事主体是否构成同一商业技术单元,还应结合商业经营策略、技术相互依赖与互操作程度以及排他性与否等因素进行判断。多个主体是否在同一品牌下针对同一市场群体进行合作(特斯拉等),是否紧密依存且数据直接联通(如车企与后端数据服务商),系统是否通过专有协议或封闭网络运行(如阿里巴巴与蚂蚁科技集团)等,均将影响对是否构成同一商业技术单元的评估与认定。在前述入室盗窃案中,A、B、C 三家公司通过合同进行持续性合作,且商业依存度极高,离开任何一方的

〔62〕 See Oliver William, “Economic Institutions: Spontaneous and International Governance,” *Journal of Law Economics & Organization*, Vol. 7, Special Issue, 1991, pp. 159, 184.

〔63〕 王利明:“论合同法组织经济的功能”,《中外法学》2017 年第 1 期,第 104—120 页。

〔64〕 See Alan O. Sykest, “The Economics of Vicarious Liability,” *The Yale Law Journal*, Vol. 93, No. 7, 1984, p. 1231.

协作人工智能系统均无法发生作用，此时应认定三者构成同一商业技术单元。如果智能安全系统未启动是因为网络连接暂时中断导致的，受害人就不能要求网络服务商与A、B、C共同承担连带责任。智能家居设备通常不需要通过特定提供商的专门网络连接，消费者也不会认为网络服务商与特定人工智能生态系统之间存在特别合作关系。仅在例外情况下，如智能家居系统与某特定网络服务商之间存在特别安全协议，后者为前者提供专门的网络连接线路，商家也以此为独特的商业优势吸引消费者，才应认定网络服务商加入该智慧生态系统，与其他三方属于同一商业技术单元的成员。

可能的质疑在于，在同一商业技术单元中可能存在强势主体，内部的损害分担协议未必公平。强势企业借助合同建立闭环的智能生态系统，中小企业在进入一段时间后，只能依赖于该系统生存而很难从中退出。此时强势企业将主导整个智能生态系统，并通过合同的方式转移损害赔偿责任。只是这一问题并非智能系统的独有问题。在市场上占据优势地位的传统车企也常通过协议，要求在无法查清缺陷原因的情况下由零部件生产商承担损害赔偿责任，或要求零部件生产商对整车召回或事故承担全部责任。无论内部如何安排，同一商业技术单元仍应对被侵权人承担连带责任。至于内部责任分担协议是否公平的问题，应结合《民法典》中的格式条款规则等相关规范进行审查，防止强势企业不合理地压榨弱势企业。〔65〕

综上所述，法秩序应设计专门的多数人侵权规则，化解责任主体不清带来的损害分担难题，具体条款表述如下：

第n＋2条　因人工智能系统造成损害，无法确定具体侵权人的，由紧密合作、相互依存的同一商业技术单元成员对损害承担连带责任。

五、结　语

面对人工智能应用带来的挑战，法律界不应过于被动，应积极思考妥适的应对之策。将产品责任扩张至人工智能系统，并不能有效解决人工智能应用引发的过错与因果关系证明困难、新型损害难以归入以及责任主体不清等问题。法律人应秉持必要的谨慎，不能为了一时的规范需求，贸然突破概念边界最终损及法的体系性。我国未来的人工智能立法应在维护体系一致性的前提下修补式地配置规则，通过信息披露规则、损害与因果关系认定规则以及特殊的多数人侵权规则，减轻被侵权人在责任成立与承担层面的举证压力。

需要补充说明的是，上述新规则虽然有助于保护被侵权人，但也会被质疑增加企业的法律成本。上述规则侧重于平衡两造之间的诉讼力量，比对被告适用无过错责任要温和很多。无过错责任模式具有极强的标题意义，给市场主体的心理压力极大，容易引发寒蝉效应。若对人工智能系统适用产品责任，一些中小企业可能会因恐惧无过错责任而不敢从事人工智能研发与应用。已经涉足人工智能领域的企业为避免“引火烧身”，也将尽可能地删除开源代码或以

〔65〕 参见林洹民：“数据交易合同的性质认定与规范要点”，《法制与社会发展》2025年第1期，第178—182页。

牺牲学习能力为代价确保系统可控。相较而言，直接针对人工智能不透明性、网络性等特征的新型侵权规则更侧重于程序性的沟通，只要被告协助积极查清真相，尚有可能推驳法律预设的责任承担机制。过错责任下的举证便利性规则将激励责任主体积极发现组织或技术漏洞，更好地实现人工智能安全治理。

Abstract: The determination and assumption of liability for artificial intelligence (AI) infringement are among the key and challenging issues in AI legislation. The product liability approach fails to effectively address three major challenges posed by AI-related infringements: proof of fault and causation, the definition of new types of damages, and the identification of liable parties. AI infringement liability legislation should be based on fault liability while incorporating specialized supplementary provisions to alleviate the evidentiary burden on the infringed party in establishing and assuming liability. The infringed party can only overcome the evidentiary barriers caused by information asymmetry by accessing AI development records, activity logs, and other relevant documents. Legislation should establish evidence disclosure rules and impose information disclosure obligations on AI-related entities under certain conditions, providing a substantive legal basis for courts to issue orders for document production. Regarding the redress of virtual damages in the AI era, rather than expanding the scope of material damages, it would be preferable to abandon the strict requirement of "seriousness" for mental damages and adopt a "significance" standard instead. To reduce the burden of proving causation for AI product consumers, legislation should establish causation presumption rules under specific conditions. When damages are definitively caused by a breach of obligation but the exact source of liability is difficult to ascertain, all members of the same commercial and technological unit should bear joint and several liability for the damages.

Key Words: Artificial Intelligence Tort; Product Liability; Fault Liability; Evidence Disclosure; Virtual Damage

(责任编辑：彭 錞)