

生成式人工智能的责任认定研究

孙文波

(中国社会科学院大学 法学院, 北京 102400)

摘要:生成式人工智能的多场景适用正呈现出不可阻挡的嬗变趋势,而在其实现规模化应用的同时却伴生着多维风险,风险“进化”的损害则又会引致归责问题。以往归责路径中,人工智能的侵权责任由提供者、管理者或使用者承担。如今,生成式人工智能存在成为法律拟制主体的可能,因此其责任认定应有别于传统认知论。建议以赋予生成式人工智能有限法律人格为逻辑起点、以围绕“生命周期”分阶段认定责任为逻辑顺延、以处财产责任抑或增添责任类型为逻辑闭环,来促进人工智能从“可控制”发展至“可规制”。

关键词:生成式人工智能; ChatGPT; 有限法律人格; 财产责任; 辅助第三人

中图分类号: D923 **文献标志码:** A **文章编号:** 1672-349X(2025)02-0070-10

DOI: 10.16160/j.cnki.tsxyxb.2025.02.010

Research on Responsibility Attribution of Generative Artificial Intelligence

SUN Wenbo

(School of Law, University of Chinese Academy of Social Sciences, Beijing 102400, China)

Abstract: The multi-scenario application of generative artificial intelligence (AI) is showing an unstoppable trend of transformation. However, while achieving large-scale application, it is accompanied by multi-dimensional risks, and the “evolution” of these risks leads to issues of responsibility attribution. In previous attribution paths, the tort liability of AI was borne by providers, administrator or users. Now, generative AI has the potential to become a legal person, so its responsibility attribution should differ from traditional cognitive theories. It is suggested to start with the logic of granting generative AI limited legal personality, extend the logic of phased responsibility attribution around the “life cycle”, and close the logic loop with property liability or the addition of liability types, thereby promoting the development of AI from “controllable” to “regulatable”.

Key Words: generative artificial intelligence; ChatGPT; limited legal personality; property liability; auxiliary third person

一、问题的提出

科技跃迁的齿轮持续转动,人工智能已经历从分析式人工智能到生成式人工智能的发展

流变。生成式人工智能 (Generative Artificial Intelligence, GAI) 是指基于算法、模型、规则生成文本、图片、声音、视频、代码等内容的虚拟实

收稿日期: 2024-09-08

作者简介: 孙文波(2001-), 男, 山东枣庄人, 博士研究生, 研究方向为数字法学。

体^[1]。ChatGPT(Chat Generative Pre-trained Transformer)是生成式人工智能应用的典型例证和重要标志。基于大量商业利益的导向,生成式人工智能的迭代周期不断缩短。与此同时,不断优化的生成式人工智能也具备了多场景适用的能力,其规模化应用正呈现出不可阻挡的嬗变态势。纵观历史演进,技术革新的规模化应用总是呈现双面性,即机遇与挑战并存。就生成式人工智能而言,在目前以及可预见的未来,以ChatGPT为代表的生成式人工智能规模化应用后,极大可能会引致数据泄露、算法歧视、生成物权属不明等一系列现实风险,对当前法律的部分基本制度框架构成泛在威胁^[2],甚至扰乱社会正常秩序和降低人类生活幸福指数。

风险与治理规则之间存在精密的耦合性,当社会被不确定性的人为风险所支配时,必然就会出现对治理规则的讨论,而治理规则的面向之一乃是谁为风险“进化”的损害“买单”,即谁是侵权行为的责任主体^[3]。在面对新生不确定性风险所产生的损害时,现行治理规则能否在其射程内解决新兴领域问题则值得省思。如果可以,这将大幅减少风险治理成本。反之,现行规制路径可能会失灵,生成式人工智能所引发的侵权行为由谁来承担责任以及怎样承担将变得扑朔迷离。目前来看,既往规制路径中,由于传统人工智能并不具有法律主体资格,因此其责任认定均围绕服务提供者、管理者或使用者展开。但相比传统人工智能所具有的分析、判别与决策等工具性特征,生成式人工智能更能在技术和应用层面体现出自主性特质,对国家和社会发展具有重要价值。假使生成式人工智能存在被赋予法律主体的可能,那么一味地要求服务提供者、管理者或使用者承担责任就似乎不妥。鉴于此,亟须分析解构生成式人工智能侵权行为的责任认定问题,通过廓清其责任认定路径,促使人工智能从“可控制”发展至“可规制”。为避免责任认定问题的讨论形成“空中楼阁”而毫无益处,需要明确生成式人工智能会引致的风险样态,并以此为基础去探究

风险“进化”损害的责任认定问题。

二、生成式人工智能应用的主要风险分析

(一)数据安全和算法偏见

1. 数据安全

从某种意义上说,数据处理具体包括收集、存储、使用、公开、销毁等阶段。然而数据处理并非都具有完整的上述过程,而是会随着场景应用迁移而发生变化^[4]。以ChatGPT被“喂养”的过程为例,数据处理主要涵盖数据收集、使用和销毁三个阶段。

在数据收集阶段,风险主要是数据收集是否合法合规的问题。具体而言,服务提供者需要专门收集数据去“喂养”生成式人工智能,数据收集的合法合规问题因此而产生。就数据来源而言,数据可能来源于个人抑或非个人。当数据来源于个人时,数据收集会涉及个人数据的保护问题,例如存在非法使用个人数据、不当使用个人数据以及个人数据泄露等风险。当数据来源于非个人时,企业或者平台已经公开的数据就会被类型化识别为原始数据、数据集合和数据产品三种类型。其中,原始数据价值有限,并不需要考虑权益问题;数据集合应受正当竞争行为的约束;数据产品被视为财产权的客体,受到物权规则的保护^[5]。而未公开的数据,则主要涉及商业秘密保护问题。特别要注意的是,在收集国家所公布的数据时,要按照一定的规范和程序进行操作,尤其是当数据跨境流动时不得有损国家的信息安全;而未公布的数据,通常不得进行非法收集。

在数据使用阶段,风险主要涉及数据处理的机密性、脱敏性问题。首先,收集到数据后必须经过一定的技术处理,使其合法合规后才能使用;其次,基于商业利益的考量,数据使用者还会挑选出被隐藏的和有价值的信息。在数据信息分析处理过程中,敏感信息、私密信息若未完全脱敏、脱密就存在被泄露和公开的风险。此外,数据是可流动的,不同服务提供者还可能会根据各自所需进行数据的共享交换,达

到双方利益交换和增值的目的,以此促进合作与共进。对于一些小型服务提供者之间的共享交流来说,数据在相互流转之间非常容易被公开,导致存在数据公开和泄露风险。

在数据销毁阶段,主要是需要对储存在大语言模型里的数据进行及时删除、封存等操作,以保证数据安全。从使用者角度看,对于未提供过服务的生成式人工智能,其储存的数据为原服务提供者输入的数据,这对使用者并不产生安全风险。只有提供过服务的生成式人工智能,才会产生额外的风险。从服务提供者角度看,对于未提供过服务的生成式人工智能,无论是服务提供者不再使用还是未来再利用,都要保证数据处理的合法合规,以此杜绝和减少安全风险的产生。而对于提供过服务的生成式人工智能,服务提供者需要对使用者信息和生成式人工智能自动生成可识别使用者的信息进行脱密、脱敏处理,以防止数据泄漏和被不正当利用。

2. 算法偏见

算法在生成式人工智能运转中同样扮演着重要角色,其能够改变数据的产出样态、组织表达以及流转情境^[6]。换言之,生成式人工智能需依靠算法生存。而实践中,算法会产生不同维度的偏见^[7],依次是被附加的初始偏见、自主学习后的显性偏见、进化迭代后的隐性偏见。正是这些偏见的产生,才会导致服务行为出现偏差。

首先,算法是人为编排的结果,带有个人偏好的价值观念,容易引起算法的初始偏见。不同历史时期、不同国别的人的价值观念会存在较大偏差。即使是同一时期抑或同一国别的人,由于生活环境的不同,其接受的价值观念也会不可避免地存在差异。在对算法进行编排时,人类大脑中的固有偏差和偏好便会发挥作用。虽然服务提供者或设计者通常尽力去避免价值观念的传递,但是其认知的偏差和偏好仍会导致算法模型的认知偏差和偏好,严重时便会产生偏见,甚至是歧视。

其次,数据是社会产物,沾有人类气息和社

会风气,在一定程度上算法模型“食得”海量数据之余可能会生成算法偏见,即显性偏见。通常而言,人类的多维价值情绪无论好坏均会衍生到数据中,这就意味着被“喂养”的数据本身附着着人类的善恶价值观,而算法通过自主学习也会产生善恶价值观。因此,显性偏见具有一定的生存空间。

再者,即便生成式人工智能服务提供者和设计者没有偏见,数据本身也不附带偏见,但算法模型迭代优化后却有可能制造偏见^[8],即隐性偏见。算法模型在为精准服务人类而筛选人类偏好时,也会不断迭代优化,此过程导致当前算法越发复杂,其在自我迭代的同时伴随着算法黑箱,即使是设计者也可能难以进行解释^[7],隐性偏见便由此产生。

(二) 知识产权风险

生成式人工智能的兴起给诸多产业带来了挑战,由其创新特质所引发的知识产权风险尤为值得关注。生成式人工智能的生产流程包括内容自动化处理分析、智能化润色加工、多维化转换以及创造性生成,会直接改变内容生产模式与内容供给范式,其生成物极其具有特殊性^[9],因此,其生成物的知识产权保护以及权利归属问题备受热议。OpenAI 公司在 ChatGPT 的《共享和发布政策》中曾提到,与 ChatGPT 共同创作的内容归属于用户,但同时也要用户要求在作品中必须以一种任何读者都不会错过的方式来清楚地揭示 ChatGPT 的作用^[10]。不过用户应该揭示到何种程度以及如何揭示并未有过多交代,若引致实践中的做法不一,就会使得生成物的权利归属处于游离状态。然而,在我国并未形成关于生成式人工智能权利归属的统一看法。有的学者认为,生成式人工智能输出阶段的生成内容具有作品的基本要素特征,应受到知识产权保护,但其权利应由使用者进行掌控^[11];而有的学者则认为,从产生过程分析,生成式人工智能生成物无法成为知识产权的客体,应当归属于公有领域^[12]。

进一步而言,生成式人工智能还会存在侵害知识产权却难以认定之嫌疑。如今大量事实

表明,生成式人工智能生成物具有侵害第三人著作权的可能。譬如,目前大多数作品发布途径均包含网络发布,很大程度上难以有效规避信息网络传播的侵权风险^[13]。而生成式人工智能生成内容也难以成为例外,因为生成式人工智能在输出内容的全过程具有“自主意识”,能够对标人脑逻辑回路来摄取与生成数据^[9]。相比分析式人工智能,生成式人工智能由于自身特性,其所涉及的著作权侵权问题更为复杂。

(三) 衍生风险

通常来说,人死亡后会存在生前利益归属不明、责任承受不清等风险,那么生成式人工智能“死亡”即自动或者被动结束服务后,是否也会衍生出一系列的安全风险?

就利益归属而言,首先需要明确的是,生成式人工智能是否可具有利益? 本文认为生成式人工智能至少可具有财产性收益。在生成式人工智能提供服务时,其可以向使用者进行收费从而提供额外的特殊服务,抑或让使用者购买类似于会员资格才能提供完整服务,此时财产性收益可部分归于生成式人工智能本身。有学者还曾提出,“当人工智能具有财产后,通过建立储备基金系统,人工智能作为一个整体拥有一个独立单一的资金池,以备在承担赔偿责任时使用”^[14]。值得关注的是,生成式人工智能生成物若能受到知识产权保护,就会产生知识产权收益,从而存在财产性收益来源。然而,当生成式人工智能具有财产性收益时,在其结束服务后,其财产利益归属几何? 如果可归属于某一个人或者组织,那么当生成式人工智能具有巨大财产性收益时,其随时都可能有“生命危险”。

此外,在生成式人工智能侵权后被动或者自动结束应用程序之余,责任是否需要被承受? 如果生成式人工智能产生侵权后仍然存在,学界观点如下:有学者认为,“目前,以 ChatGPT 为代表的生成式人工智能侵权责任领域尚不具备赋予其主体地位的条件,将其确定为一种人工智能服务更加适宜”^[15];有的学者则认为,生成式人工智能作为主体的本体表现俨然已成,

“人工智能侵权,首先由其自担客观责任”,其次“人工智能开发者或制造商、用户、受害人、第三方等对侵权发生具有原因力,应依可归责事由,厘定责任主体、责任形态”^[16]。可以发现,上述论断对于是否承认生成式人工智能的法律主体性进行了界分:一是在将生成式人工智能看作服务型工具基础上,进而认定服务提供者、管理者抑或使用者责任;二是赋予生成式人工智能法律主体地位,以此为基点分别认定生成式人工智能、服务提供者和使用者责任。类推之,在生成式人工智能结束应用程序后,以是否承认其法律主体性而演化为两条路径:一是当生成式人工智能仅为服务型工具时,由服务提供者、管理者抑或使用者承担责任即可;二是当生成式人工智能具有法律人格时,类比现行法律主体死亡规则,需寻找其责任承继者。

三、传统人工智能的责任认定路径

就以往规制路径而言,鉴于传统人工智能的非法律主体性,其自身并不承担侵权责任。即使行政责任、刑事责任与民事责任的认定存在差异,也均并非由人工智能承担责任。法律责任的追究首先涉及民行刑责任在整个法律体系中的基本定位和顺序安排^[17],因此需在廓清何种情形会引致民事责任、行政责任和刑事责任的基础上,再去探讨人工智能的责任认定路径。

(一) 责任规制

法律是行为规制法,民事责任、行政责任与刑事责任是法律规制的三种主要方式。从狭义上看,三大部门法中的责任性质是不同的,即侵权、违法和犯罪。但责任规制的目的却是相同的,即保护法益。法益,顾名思义,是指法律所保护的权益。“法律最重要的目的与任务,就是保护法益。”^[18]然而,不同法律所保护的法益是存在差别的,侵害不同的法益会造成责任性质的不同。因此,考究不同法律所保护的法益自然就成为廓清传统人工智能所引致责任性质的标准。

根据《中华人民共和国民法典》第 2 条可知,民法所调整的乃是平等主体之间的人身关

系和财产关系,所保护的乃是平等主体之间的民事权益。行政法所侧重的则是防止危害、管理公共危险和增进公共利益^[19]。而依照《中华人民共和国刑法》(以下简称《刑法》)的规定,符合犯罪构成要件需要进行制裁的,依法追究刑事责任。具体到在传统人工智能所引致的损害中,如果引致的损害打破了平等主体之间的民事权益平衡,就会产生民事责任;如果不当地损害公共秩序、增加公共危险和减少公共利益,就会产生行政责任;而如果符合犯罪构成要件需要进行刑事制裁的,则应当依法追究刑事责任。值得注意的是,刑法作为法益保护的最后补充性手段,在民法与行政法规制不足的前提下,即使传统人工智能引致的损害与犯罪构成要件存在些微差别,也可参照刑法规范对其进行规制。

(二)责任认定

1. 民事责任

首先,对于传统人工智能引发的侵权责任,如果是由人工智能设计者的设计缺陷造成的,那么此时可适用产品责任^[20]。其次,在不存在设计缺陷的情况下,如果侵权是由管理者的不当管理造成的,那么由管理者承担责任^[21];如果是使用者未履行谨慎义务而引致了传统人工智能侵权行为的发生,那么则由使用者承担侵权责任^[22]。再次,在传统人工智能侵权行为产生原因不明的情况下,可由人工智能设计者、管理者承担相关侵权行为的连带责任^[23],当然为避免责任分配不均,如果存在使用者,那么使用者也需要承担连带责任。

2. 刑事责任

单位也并非人类主体,传统人工智能刑事责任的认定可对比刑法对单位犯罪的评定。从入罪角度看,根据单位犯罪机理可知,单位“具有自己的整体意志,从而也具有自己的犯罪能力和刑事责任能力,它是作为一个有机整体实施犯罪和负担刑事责任的”^[24]。反观传统人工智能,其并不具备犯罪能力和刑事责任能力,仅是人类所观察的客体。从刑罚角度看,单位通过自主经营可以获得相应财产,财产处罚是可进行的,而传统人工智能并非一般意义上的生

命物种,难以获得如自然人一般的人身自由,并获得独立的财产^[25],因此对其无法进行处罚。综合来看,传统人工智能通常被视为犯罪的工具抑或辅助手段。

3. 行政责任

传统人工智能在行政领域作为一种技术而使用,主要还存在公共行政技术责任^[26]。此时,责任主体是行政主体,并非传统人工智能。就行政主体所负义务而言,首先,无论行政主体是否使用传统人工智能,行政主体都应当具有保护行政相对人免受损害的义务。其次,假如行政主体同为使用者,那么行政主体要具有审慎使用传统人工智能的注意义务;假如行政主体与使用者并非一体,那么行政主体对使用者的合理使用具有监督义务。而具体承担何种责任,则主要包括两个方面:从内部视角看,存在技术应用责任;从内外交叉视角看,存在信息安全责任^[26]。

(三)责任竞合

通常来说,责任存在竞合的可能。在传统人工智能引致损害的特殊情形中,民事侵权、行政违法和刑事犯罪三者界限有时也并非泾渭分明,而是存在责任竞合的情况。责任竞合主要发生在民事侵权与行政违法、民事侵权与刑事犯罪、行政违法与刑事犯罪以及三者之间,是指传统人工智能引致的损害可能被涵摄在多重责任之下。毋庸置疑的是,责任是可以竞合的。然而存在疑问的是,责任竞合样态下是否可以并罚?就民事侵权与行政违法、民事侵权与刑事犯罪而言,毫无疑问是可以并罚的;但就行政违法与刑事犯罪以及三者之间而言,是否可以并罚则存在疑问。本文认为如果刑法条文与行政法条文中具有对应的内容,那么此时将不可能存在责任竞合事项,更遑论并罚。例如,同为传统人工智能所引致的网络侮辱诽谤行为,治安处罚与刑罚入罪之间的区别主要在于情节,当情节严重时便构成侮辱诽谤罪^[17]。而当无对应内容时,则存在责任竞合的可能,不过如果当事人的违法行为已构成犯罪并被追究刑事责任的,一般不再作行政处罚,最终体现为刑事处

罚^[27]。但是在一些特殊情况下,也应当并罚。譬如,甲在公共道路上驾驶自动驾驶汽车(L3级^①),车辆突然急转向前行驶,速度高达规定时速的160%,撞到了乙,且乙死亡。此时,由于甲未及时对自动驾驶汽车进行降速干预,所以根据《中华人民共和国道路交通安全法》第99条第1款第4项和第2款规定可知,存在对甲吊销驾照的可能;而同时甲又违反交规且将乙撞死,根据《刑法》第133条规定,则构成交通肇事罪。

综上所述,虽不同面向的责任规制存在差异性,甚至在某些情形中会出现责任竞合而达到并罚样态,但基于传统人工智能的非法律主体性,其自身均并不承担责任。

四、生成式人工智能成为法律拟制主体的可能性

生成式人工智能不同于传统人工智能,其可成为法律拟制主体。现代社会中,生物人从诞生那一刻起,便具有了法律主体资格。而非自然人取得法律主体资格则需要进行法律上的拟制。换言之,非生物人需要满足一定的条件才可成为法律主体。法律拟制主体是人类自身为满足特定利益而进行的特殊性建设活动,其核心目的是去维护人类在现代社会中的生存和发展。因法人和生成式人工智能同为非生物,都与人类自身特定利益存在相应关联,既然法人可以成为法律拟制主体,那么考量法人成为法律拟制主体的动机和原因,生成式人工智能也将可能成为相对应的法律拟制主体。本文认为生成式人工智能需要满足以下必要条件才有可能成为法律拟制主体:首先,对国家发展有益;其次,具有一定社会价值;再次,耦合人类可持续发展。

生成式人工智能可促进国家高质量发展。如今高质量发展已然成为全面建设社会主义现代化国家的首要任务,技术创新驱动战略的重

要性日益凸显。生成式人工智能作为当下技术创新最先进的典型代表,对于促进国家高质量发展具有巨大作用。在国家安全领域,虽生成式人工智能可能成为国家核心数据和重要数据被窃取的“隐秘出口”^[28],但其所具备的顶尖算法和强劲算力使其能够在些许时间内给出高质量且具有差异的安全防御方案,为国家实施一系列安全措施提供加速度。在数字政府建设层面,相比分析式人工智能采用“判别式建模”(Discriminative Modelling),生成式人工智能最大的不同是采用“生成式建模”(Generative Modelling),它可创造和训练与样本中数据类似的新数据^[29],进一步提高智能量级,并且可以应用于不同的数字服务中,为数字政府建设提质增效。此外,技术革新还会直接影响一国经济的发展,而经济优势也会使一国在国际中的竞争力呈跃升趋势。譬如,ChatGPT的诞生使得美国科技影响因子处于世界领先地位。

生成式人工智能具有强大的社会价值。就数字乡村建设而言,“基于数字乡村建设与数字技术之间的密切关联,其为生成式人工智能发挥社会价值提供了新的作用场域”^[30]。目前,我国乡村快递物流服务质量并未得到完全改善,现有农村物流体系效率低的弊端阻碍了农村“双向物流”的发展^[31],与我国乡村快递物流范式所勾勒的理想蓝图存在差距。通常理解,快递物流效率提升无非就是准点、安全抵达,尤其是对于新鲜且易腐蚀的物品,因而高效率的配送至关重要。而生成式人工智能以其具备的顶尖算法、海量数据以及强劲算力可以为衡量不同利益相关者的目标(低碳排、低风险、低成本、高顾客满意度)提供定量研究的依据,以解决冷链生鲜产品“最初一公里”的衔接、“最后一公里”配送路径的优化以及配送过程中的智能调度、智能配载和能源优化等问题^[32],可以说,大幅提升乡村快递物流的配送质量并非“天马行空”。

^① L3级(有条件自动驾驶)由系统完成大多数驾驶操作,必要时向驾驶人提出系统请求,驾驶人需要提供适当的干预,同时在驾驶过程中尽到安全注意义务。参见宋宗宇、林传琳《自动驾驶交通事故责任民法教义学解释进路》,载《政治与法律》2020年第11期,第150-160页。

生成式人工智能能促进人类可持续发展。可持续发展本质上是指人类生存发展的需要可以自由地得到满足,并且可以进行多轮迭代。而人类的需要具有多样性和层次性,底层的需要是出于人类作为动物的本能,但高层次的需要却并不是天然形成的,而是要依靠人类自身不断地发掘和构建^[33]。人工智能的出现和发展正是人类生存发展需要得到自由满足的缩影。1956年达特茅斯学院人工智能学术会议召开,给予“AI”这一术语正式的身份,人工智能就此诞生^[34]。以此为时间节点,人类开启智能化时代。近些年,在物联网、大数据、云计算、区块链、脑机接口等科技以及经济社会发展强烈需求的共同推动下,人工智能得到快速发展^[35]。在上述过程中,人类生存发展的需求促使人工智能诞生与发展,人工智能的发展则又满足于人类的生存发展,两者在一定程度上处于互予状态。而2023年成功出圈的生成式人工智能则更是两者互予状态的典型例证。以人类教育为例,教育治理现代化的根本目的是实现人的现代化,而生成式人工智能可“赋能教育治理现代化以先进性、科学性、人本性相统一的价值体系为指引”^[36]。具体而言,海量数据的众合、顶尖算法的集成和强劲算力的叠加可以满足教育应用场景的先进性;而教育应用场景的先进性则可以满足教育治理的准确度和高效性,进而提升教育生态的科学性。为更好地促进教育进步、提升人类受教育水平、促进人的全面自由发展,生成式人工智能的优化迭代也会被持续向前推进。

综上所述,生成式人工智能对国家发展有益、具有一定社会价值、耦合人类可持续发展,符合成为法律拟制主体的必要要件,拥有成为法律主体的可能。鉴于此,以往规制路径对于生成式人工智能而言,存在难以调适的空间。因此,需要在已有治理经验的基础上,调整生成式人工智能责任承担体系,确保生成式人工智能始终按照预设路线发展,真正成为“负责任的人工智能”^[37],从而使其从“可控制”发展至“可规制”。

五、生成式人工智能的责任认定路径

(一)逻辑起点:赋予生成式人工智能有限法律人格

生成式人工智能虽可以具有法律拟制主体地位,“但这种法律人格应当是有限度的,并非完全的法律人格”^[2]。生成式人工智能成为法律拟制主体的同时意味着其将获得法律上的权利和义务。权利的享有和义务的承担主要由权利能力和行为能力所反映,而有限法律人格则主要表现为权利能力和行为能力的双重限缩性。原因主要在于生成式人工智能与现存法律主体还存在一些差异。目前来看,生成式人工智能不可能被赋予自然人所特有的全部的法律权利。譬如,婚姻自主权是指可以自主选择婚姻的权利,自主选择婚姻的前提是有存在配偶的可能,而生成式人工智能明显无法寻求自身配偶,也不可能存在配偶。生成式人工智能也并非学界所论的“超人工智能”(Artificial Super Intelligence, ASI)。来自牛津大学的知名人工智能思想家 Nick Bostrom 认为,ASI“在几乎所有领域都比最聪明的人类大脑都聪明很多,包括科学创新、通识和社交技能”^[38]。不过从目前来看,超人工智能时代的奇点还并未完全到来,生成式人工智能无法到达完全等人性的格局,赋予其完全法律人格不甚必要。因此,赋予生成式人工智能有限法律人格是现阶段最佳的利益选择。

(二)逻辑顺延:围绕“生命周期”分阶段认定责任

在赋予生成式人工智能有限法律人格后,责任认定可以围绕其“生命周期”分阶段进行特殊处理。生成式人工智能的“生命周期”可分为提供服务前、提供服务时和结束服务后三个阶段。提供服务前阶段是指生成式人工智能需要经过数据输入、算法训练等程序才可提供服务。提供服务时阶段是指经过一系列的深度进化,生成式人工智能在接收到“信号”后可以自动生成回应。另外,此阶段还包括可以提供服务但

未提供和提供服务后并未被动或者自动结束应用程序的情况。结束服务后阶段是指生成式人工智能在提供服务后被动或者自动结束应用程序。

首先,在生成式人工智能提供服务前阶段,虽然服务模型存在一定的侵权风险,但此阶段并不存在适格的生成式人工智能,类似于人类婴儿,责任承担可参考监护制度进行转置。在无第三方介入的样态下,将侵权责任归属于服务提供者或者设计者则较为合理。倘若侵权行为涉及多方主体,理应按照各自过错划分比例承担相应责任。其次,在生成式人工智能提供服务时阶段,需以一般侵权归责原则予以规制,以此契合其有限法律人格。而“由于过错责任原则的限制会使许多受害人无法得到赔偿,为了维持人工智能提供者与受害人之间的利益平衡,仍应给予受害人以必要的救济,防止其持续地受到损害”^[39]。关于过错如何进行举证的问题,建议适用相对宽松的标准。当生成式人工智能存在不符合“理性人工智能”^①的样态时,可以认定其存在过错。再次,在生成式人工智能结束服务后阶段,基于利益平衡的考量,应适用过错推定原则,衍生侵权责任应由服务使用者承担。而在服务使用者能够证明对损害发生不存在控制力时,则不承担侵权责任,此时责任由社会公众整体承担。此外,我国当前的人工智能发展与应用仍然处于起步阶段,宜对AI企业采取包容审慎的态度,对技术创新给予更大的包容度^[40],因此服务提供者或者设计者一般不必承担衍生侵权责任。

值得注意的是,如果生成式人工智能侵害的也是人工智能,那么责任认定路径可以分为损害传统人工智能和侵害生成式人工智能两种。损害传统人工智能是侵害他人财产,可按照上述路径予以规制。而就侵害生成式人工智能而言,以“生命周期”分阶段认定责任需要经过多次分类,一般较为繁琐。提供服务前和结

束服务后两阶段侵害其他生成式人工智能或者被其他生成式人工智能侵害的概率较小,鉴于此,可以仅规定提供服务时阶段的生成式人工智能侵害其他提供服务时阶段的生成式人工智能的责任认定路径。为避免现阶段赔付难度过高的问题,应适用无过错归责原则。

(三)逻辑闭环:处财产责任抑或增添特殊责任类型

生成式人工智能的责任承担,可以处财产责任抑或增添特殊责任类型。就前者而言,财产责任是为契合生成式人工智能的有限法律人格,其可适用的基础上文已论述,此处不再赘述。关于责任财产的范围,则由生成式人工智能的自身财产承担,包括但不限于生成式人工智能提供服务所得、接受的赠与等。另外,建立储备基金系统的,可以由基金先行赔付。就后者而言,在刑事犯罪领域,除财产赔偿外,有学者认为针对涉及人工智能的犯罪,有必要在刑法中增设删除数据、修改程序、永久销毁等刑罚^[41]。而这可以为侵权领域提供借鉴,譬如,如果生成式人工智能因侵害自然人人身权益而造成严重精神损害的,在处财产责任的基础上,可以适当增加暂停服务、修改程序或永久销毁等责任承担方式,但是否适用惩罚性赔偿还值得商榷。应当言明的是,基于商业服务的高成本性,一般并不鼓励直接适用永久销毁的责任承担方式。

此外,为打通损害赔偿到位的最终路径,还可为生成式人工智能填补缺失的法律人格,使其真正地向自然人靠拢。建议为其设置“辅助第三人”,其主要职责就是辅助损害赔偿顺利到位。在围绕于生成式人工智能周围的主体中,服务提供者是最佳人选,原因在于服务提供者是生成式人工智能正式接触外界的第一通道,理应承担相应职责。当然,“辅助第三人”并不需要承担生成式人工智能所应该承担的责任。

^① 理性人工智能原则上以具有通常服务模型的生成式人工智能在相同或近似场合均不会产生风险或者损害为衡量标准。

而在“辅助第三人”不适当履行其职责或者恶意损害生成式人工智能权益时,则需要其承担赔偿责任。

参考文献:

- [1] 毕文轩. 生成式人工智能的风险规制困境及其化解:以 ChatGPT 的规制为视角[J]. 比较法研究,2023(3):155-172.
- [2] 袁曾. 生成式人工智能的责任能力研究[J]. 东方法学,2023(3):18-33.
- [3] 王亚琪. 风险社会视域下全球治理的不确定性探析[J]. 东北亚论坛,2022,31(3):44-59.
- [4] 徐双,刘文斌,李佳龙,等. 大数据背景下的数据安全治理研究进展[J]. 太原理工大学学报,2024,55(1):127-141.
- [5] 李健男,高宁宁. 企业公开数据权益保护的复合路径:基于数据形态进化的可能选择[J]. 北京行政学院学报,2023(2):96-108.
- [6] 曹树金,曹茹焯. 从 ChatGPT 看生成式 AI 对情报学研究与实践的影响[J]. 现代情报,2023,43(4):3-10.
- [7] 顾男飞. 生成式人工智能的智能涌现、风险规制与产业调控[J]. 荆楚法学,2023(3):70-83.
- [8] 汪靖. 从人类偏见到算法偏见:偏见是否可以被消除[J]. 探索与争鸣,2021(3):32-34.
- [9] 刘艳红. 生成式人工智能的三大安全风险及法律规制:以 ChatGPT 为例[J]. 东方法学,2023(4):29-43.
- [10] Sharing & publication policy[EB/OL]. (2022-11-14) [2023-07-27]. <https://openai.com/policies/sharing-publication-policy>.
- [11] 吴汉东. 人工智能生成作品的著作权法之问[J]. 中外法学,2020,32(3):653-673.
- [12] 曹博. 人工智能生成物的智力财产属性辨析[J]. 比较法研究,2019(4):138-150.
- [13] 马治国,赵龙. 文本与数据挖掘对著作权例外体系的冲击与应对[J]. 西北师大学报(社会科学版),2021,58(4):107-115.
- [14] 约翰·弗兰克·韦弗. 机器人是人吗?[M]. 刘海安,徐铁英,向秦,译. 上海:上海人民出版社,2018:33.
- [15] 聂童. ChatGPT 生成式人工智能侵权责任研究[J]. 工业信息安全,2023(4):6-10.
- [16] 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.
- [17] 赵宏. 网暴案件中的民行刑责任与一体化衔接[J]. 北方法学,2023,17(5):5-20.
- [18] 袁曾. 人工智能法益位阶研究[J]. 地方立法研究,2019,4(5):73-84.
- [19] 姜涛. 网络暴力治理中刑事责任、行政责任与民事责任的衔接[J]. 法律科学(西北政法大学学报),2023,41(5):102-114.
- [20] 刘小璇,张虎. 论人工智能的侵权责任[J]. 南京社会科学,2018(9):105-110.
- [21] 刘谢慈,贺驰宇. 人工智能侵权责任的主体识别与担责:以风险社会治理为视角[J]. 湖南人文科技学院学报,2023,40(4):16-21.
- [22] 彭中礼. 新兴技术推动法理论变革的因素考量:以人工智能产品侵权责任分配理论为例的反思[J]. 甘肃社会科学,2022(4):115-128.
- [23] 付其运. 人工智能非主体性前提下侵权责任承担机制研究[J]. 法学杂志,2021,42(4):83-90.
- [24] 李冠煜. 单位犯罪处罚原理新论:以主观推定与客观归责之关联性构建为中心[J]. 政治与法律,2015(5):33-46.
- [25] 彭景理. 论人工智能时代智能机器人的刑事责任能力[J]. 大连理工大学学报(社会科学版),2020,41(2):71-79.
- [26] 颜佳华,王张华. 人工智能场景下公共行政技术责任审视[J]. 理论探索,2019(3):88-94.
- [27] 胡建森. 论行政处罚“一事不二罚”原则及其“一事”与“二罚”的认定标准[J]. 法学评论,2023,41(5):29-44.
- [28] 赵精武. 生成式人工智能应用风险治理的理论误区与路径转向[J]. 荆楚法学,2023(3):47-58.

- [29] 陈永伟. 作为 GPT 的 GPT: 新一代人工智能的机遇与挑战[J]. 财经问题研究, 2023(6): 41-58.
- [30] 王茂福, 严雪雁. 生成式人工智能赋能数字乡村建设: 应用价值、现实梗阻与路径支持[J]. 深圳大学学报(人文社会科学版), 2023, 40(4): 87-96.
- [31] 杨周, 何律琴. 乡村振兴背景下贵州省农村物流体系发展对策研究[J]. 物流科技, 2024, 47(11): 51-52.
- [32] 王术峰, 何鹏飞, 吴春尚. 数字物流理论、技术方法与应用: 数字物流学术研讨会观点综述[J]. 中国流通经济, 2021, 35(6): 3-16.
- [33] 曹昱. 以人类可持续发展为中心的智能化思想分析: 从马克思异化劳动理论的视角[J]. 南京政治学院学报, 2018, 34(4): 36-40.
- [34] 马艳华, 张明军. 人工智能时代人机关系的历史演化与未来走向[J]. 邵阳学院学报(社会科学版), 2022, 21(6): 32-36.
- [35] 顾世春. 斯塔尔人工智能生态系统伦理治理思想研究[J]. 昆明理工大学学报(社会科学版), 2022, 22(6): 36-41.
- [36] 周洪宇, 李宇阳. 生成式人工智能技术 ChatGPT 与教育治理现代化: 兼论数字化时代的教育治理转型[J]. 华东师范大学学报(教育科学版), 2023, 41(7): 36-46.
- [37] 袁曾. 生成式人工智能责任规制的法律问题研究[J]. 法学杂志, 2023, 44(4): 119-130.
- [38] 王春晖. 从弱人工智能到超人工智能 AI 的道路有多长[J]. 通信世界, 2018(18): 9.
- [39] 周学峰. 生成式人工智能侵权责任探析[J]. 比较法研究, 2023(4): 117-131.
- [40] 曹建峰. 迈向可信 AI: ChatGPT 类生成式人工智能的治理挑战及应对[J]. 上海政法学院学报(法治论丛), 2023, 38(4): 28-42.
- [41] 刘宪权. 人工智能时代的“内忧”“外患”与刑事责任[J]. 东方法学, 2018(1): 134-142.

(责任编辑: 于博艺)

(上接第 60 页)

(26)到内室念佛看经, 体格终须是妇女, 交搭更便。(明《初刻拍案惊奇》卷六)

按, “御室”与“御内”同义, 建议《大词典》增补词条“御室”, 可释义为与妻子交合。

三、结语

作为明末清初的孤本杂剧篇目, 前人对《破梦鹃》的词语考释研究尚为空白。本文选取剧中部分需要考释的词语作为研究对象, 检索相关语料, 并结合《大词典》和《汉语方言大词典》中的相关释义对词语进行考释, 然后对《大词典》提出增补词条、补充义项等修订建议, 以期为我国大型语文辞书修订工作提供新的材料。另外, 剧中有些词语颇为费解, 亦有不少因形近而致讹误的例子, 本文亦稍有提及, 以期能为读

者理解文意提供参考, 亦为图书编订者提供新的校注材料。

参考文献:

- [1] 徐立. 古杂剧《破梦鹃》初探[J]. 西南大学学报(社会科学版), 2011, 37(2): 150-155.
- [2] 黄仕忠. 明清稀见孤本戏曲汇刊[M]. 桂林: 广西师范大学出版社, 2014.
- [3] 汉语大词典编辑委员会, 汉语大词典编纂处. 汉语大词典[Z]. 上海: 上海辞书出版社, 1997.
- [4] 许宝华, 宫田一郎. 汉语方言大词典[Z]. 北京: 中华书局, 1999.
- [5] 许慎. 说文解字[M]. 北京: 中华书局, 2016: 116.

(责任编辑: 冯兆娜)