

文章编号: 1673-1646(2025)04-0076-07

生成式人工智能主体资质的本体论辨析

李凌, 白硕扬, 熊思琪

(复旦大学, 马克思主义研究院, 上海 200433)

摘要: 随着生成式人工智能的普及, 其自主性与类人性特征引发了对生成式人工智能主体地位的争议。法律人格扩展论等主张赋予生成式人工智能法律主体资质以应对技术发展需求; 反对者则批判生成式人工智能主体化的正当性。这些讨论多停留于功能主义与技术现实的权衡, 未能深入探究“主体”“人格”的本体论根源。通过回溯思想史上关于“主体何以是主体”“人格为何称之为人格”的理论资源, 对生成式人工智能是否具有主体资质、享有人格尊严、能否承担责任进行本体论探讨, 从人文主义的底线否定了生成式人工智能构成主体、享有人格尊严的可能, 进而捍卫人类作为主体、享有人格尊严并承担道德和法律责任的独特性。

关键词: 生成式人工智能; 主体资质; 人格尊严; 责任归因; 本体论

中图分类号: TP18; F49 **文献标识码:** A **doi:** 10.62756/xbsk.1673-1646.2025082

引用格式: 李凌, 白硕扬, 熊思琪. 生成式人工智能主体资质的本体论辨析[J]. 中北大学学报(社会科学版), 2025, 41(4): 76-82.

Ontology Analysis on the Subject Qualification of Generative Artificial Intelligence

LI Ling, BAI Shuoyang, XIONG Siqi

(Institute of Marxist Studies, Fudan University, Shanghai 200433, China)

Abstract: With the popularization of generative artificial intelligence, its autonomy and human-like characteristics have sparked controversy over the subject qualification of generative artificial intelligence. The theory of legal personality extension advocates endowing generative artificial intelligence with legal subject qualifications to meet the needs of technological development; Opponents criticize the legitimacy of the subjectification of generative artificial intelligence. These discussions mostly focus on the balance between functionalism and technological reality, without delving into the ontological roots of “subject” and “personality”. By reviewing the theoretical resources in the history of thought on “why the subject is the subject” and “why personality is called personality”, this paper explores the ontology of whether generative artificial intelligence has subject qualifications, enjoys personal dignity, and can assume responsibility. From the bottom line of humanism, it denies the possibility of generative artificial intelligence constituting the subject and enjoying personal dignity, and thus defends the uniqueness of human beings as subjects, enjoying personal dignity, and assuming moral and legal responsibilities.

Key words: generative artificial intelligence; subject qualification; personal dignity; attribution of responsibility; ontology

随着 ChatGPT、DeepSeek 等大语言模型的普及, 生成式人工智能表现出越来越强的自主性, 不仅让人们重新看到强人工智能的曙光, 而且再次激发人们对

这一人类创造物的主体地位的探讨。人机之间的界限开始模糊混淆起来, “机器是人”“人工智能主体”等出现在各类文章之中, 让人类开始重新审视“人是

收稿日期: 2025-03-23

基金项目: 国家社会科学基金后期资助项目: 舆以载道: 道德舆论的历史唯物主义考察(24FZX032)

作者简介: 李凌(1983-), 男, 副研究员, 博士, 硕士生导师, 从事专业: 马克思主义理论、新闻与传播学。E-mail: li_ling@fudan.edu.cn。

万物的尺度”这一古老的哲学命题。关于生成式人工智能是否具有主体性,相关讨论一直莫衷一是、歧见纷呈,例如法律人格扩展论、电子人格论、现实发展需要论等支持生成式人工智能成为法律主体;同时,也有很多学者提出诸如工具说、电子奴隶说等多种学说,这些批判主张要么陷入二元对立的形而上学,要么陷入循环论证的逻辑陷阱之中,并不足以从根本上反驳生成式人工智能主体的支持者。总的来说,学界关于生成式人工智能主体性的讨论仍然局限于工具理性层面,并未深入到人何以为人、主体何以为主体的本体论根源,从哲学或伦理学角度探讨辨析生成式人工智能的主体性。从本体论层面探究生成式人工智能主体资质甚为必要,其不仅是在实用层面赋予生成式人工智能主体地位与否的前提条件,而且关乎对人类主体地位问题的根本反思与捍卫,因此值得审慎对待。

1 生成式人工智能主体资质研究的典型观点及评价

针对生成式人工智能主体资质的研究呈现出技术乐观主义与人文保守主义的尖锐对立:一方面诸如法律人格扩展论、现实发展需要论等支持论者以现实的社会需求为导向,试图通过法律拟制来适应生成式人工智能技术的快速发展,赋予生成式人工智能一定的主体资质和法律地位;另一方面,反对论者则从工具论层次批判生成式人工智能法律主体地位的不合理之处,强调人类主体性的不可僭越性。

部分学者主张突破传统法律人格的“人类中心主义”桎梏,通过类比法人制度赋予生成式人工智能法律主体地位。法人制度的历史演进表明法律人格具有开放性,生成式人工智能的自主性与社会功能使其具备成为“电子人”的合理性^[1]。生成式人工智能的“有限人格”可通过登记注册、强制保险等制度实现责任归属,避免责任真空^[2]。欧盟等国家曾经建议赋予复杂自主机器人“电子人格”,赋予其纳税、继承等权利义务,成为该理论的实践注脚。还有学者从功能主义视角出发提出工具性人格论,法律主体的本质是权利义务的承担者,而非伦理实体。生成式人工智能若具备独立行为能力与责任能力,即可通过“工具性人格”制度设计纳入法律主体范畴^[3]。其论证逻辑为:生成式人工智能的自主决策能力,例如自动驾驶汽车的避障行为,已超越传统工具属性,需通过主体资格实现法律关系的闭环。此外,还有一类观点从技术发展需

求论出发,论证赋予生成式人工智能主体地位的正当性。他们认为生成式人工智能的“类人性”特征,例如 DeepSeek 等深度学习与语义理解能力正在倒逼人类法律承认其主体地位。生成式人工智能的涌现能力使其在创作、决策等领域展现出类人智能,传统法律框架已无法应对“主体缺位”导致的权利冲突,法律应当主动适应生成式人工智能技术的发展^[4]。

不过法律人格扩展论、技术发展需要说都受到了不同层面的反驳与批判。有学者明确指出,主体性以自我意识为核心,而生成式人工智能的意识仅是数据模拟的产物,无法通过“中文屋实验”证明其具备真正的理解能力^[5]。还有学者提出生成式人工智能的本质是“人工类人格”,其自主性源于算法编程而非内在意识。将生成式人工智能视为法律主体,无异于承认“电子奴隶”的拟制人格,违背“人只能是目的而非手段”的伦理原则^[6]。塞尔的“中文屋实验”也表明,生成式人工智能的符号处理能力无法等同于人类的语义理解,其“智能”本质是形式化的计算过程。法律主体资格需以“自由意志”为前提,而当前生成式人工智能的决策逻辑完全受制于人类设定的算法,缺乏道德选择能力,因此并不能赋予生成式人工智能道德主体或法律主体资质^[7]。此外,针对支持论者提出的现实发展需求论和责任归属困境与伦理风险,反对者强调,赋予生成式人工智能主体地位将导致责任认定模糊化。若人工智能因“自主行为”致损,其责任承担机制包括删除数据、销毁程序等缺乏实质威慑力,且无法实现“矫正正义”^[8]。此外,生成式人工智能的“类人性”可能模糊人机界限,威胁人类主体地位的伦理根基^[9]。过度赋予生成式人工智能主体地位可能导致“技术奇点”恐慌,加剧人类对技术失控的焦虑。因此有学者主张回归“以人为本”的法律框架,在承认生成式人工智能工具价值同时,通过算法审计、伦理审查等机制防范风险,而非贸然突破主客二分法^[10]。

总而言之,无论是支持论者还是反对论者,当前对生成式人工智能主体资质的探讨仍然未能跳出笛卡尔式主客体二元对立、非此即彼的思维,仍然只是停留在浅层、表象和单一维度的工具理性层面,缺乏对主体性本质的哲学追问,尤其是对“主体”“人格”等核心概念的界定模糊,导致论证过于功利化、过于浅层次,无法在本体论上为生成式人工智能找到坚实的奠基。要从哲学上对生成式人工智能主体资质问题进行辨析,必须回到“主体”“人格”等概念的本体论探讨上来,在“人之为人”的价值尺度上实现技术与人性的动态平衡。

2 生成式人工智能不具备主体资质

讨论生成式人工智能是否具有法律主体地位的前提条件,是生成式人工智能能否构成主体或者说生成式人工智能是否具有主体性。主体首先是个具有特定指称的哲学概念,倘若生成式人工智能无法从哲学上证成主体,那么也难以从道德关系与法律关系上赋予其主体资格,使得其成为具有主观能动性的主体而非对象化的客体存在。

在哲学视域中,主体指的是能够感知、认知和反思,具备自我意识与行动能力,并表现出思想、情感及人格等特质的个体。作为西方传统哲学最为重要的一个概念,主体和主体性概念的理论价值在于凸显人之为人的自由与能动本质,以哲学思维确立人在世界之中独一无二、与众不同的资质与地位,揭示出人类认识、改造与驾驭自然的实践能力。不同哲学家对主体及主体性的阐释存在差异:普罗泰戈拉提出“人是万物的尺度”,强调人的主体性地位,使世界万物的意义需通过与人的关联得以呈现。这一命题的提出,标志着西方哲学开始转向主体性维度,被视为人类自我意识首次觉醒的哲学宣言。苏格拉底沿着这一思想路径继续推进,将哲学研究的重心从自然转向人类自身,其“认识你自己”的箴言不仅体现了人类理论旨趣的转向,而且奠定了人文主义传统的哲学基石^[11]。亚里士多德从形而上学层次将主体界定为有着自在自足目的的承受者,构成了“人是目的”的哲学源头^[12]。主体概念和观念在人类哲学史上经历了漫长的演进过程。米兰多拉在其被视为“文艺复兴的宣言”的《论人的尊严》一书中将人界定为形象未定者、自由型塑者,凸显出人之为人在世界上独一无二的特殊地位^[13]。笛卡尔通过“我思故我在”的著名命题,使主体性原则在近代哲学中得以稳固确立^[14]。康德将主体定义为理性存在者,其主体性原则不仅具有认识论层面大胆运用理性的思辨意义,更蕴含着深刻的人本主义内涵。在认识论上,康德通过从经验知识中剥离出先验认识形式,揭示了人类的先验认知能力和综合判断能力,确立了主体作为理性存在者的地位;在伦理学层面,康德从理性存在者的自由意志出发,形成了“人是目的”和人格尊严的论断,既是对历史上主体概念的继承与发展,也构成对人类主体地位最具根本性的哲学思考^[15]。

由此观之,西方哲学家们的主体及主体性原则虽

有所差异,但都不外乎与单数或复数的人相关。马克思更是从辩证唯物主义和历史唯物主义原理出发确凿无疑地指出:“主体是人,客体是自然,只有人,并且是有目的、有意识地认识或者改造世界的人,才构成主体。”^[16]由人而来的主体,既可以是个体,也可以是群体、组织机构,乃至整个社会,但必须有具体的、实践的人的存在与在场,否则只能构成客体。人作为主体最本质的规定性是主体性,主体性最重要的内容是人的创造性与实践性,即主观能动性或者说自我意识。这是人之为人、人为主体最根本的特征,离开主观能动性,人也不能称其为主体。如今,即使是具有较强自主性和独立性、表现出一定情感意识的生成式人工智能,离人的主体地位也相差甚远。人类作为主体,具有内在的自我,这个自我具有自反性、非形式化、主动性与发展性,而且这个自我与身体紧密关联形成具身性。生成式人工智能纯粹是外化、形式化、被动性、静态化和离身化的实体存在者,其并不具备主体的内在化、主动性、具身性等属性。生成式人工智能的运行机制本质上是对人类指令的机械执行,依赖预设算法程序进行被动反馈与操作,不具备自主意识或主动决策能力。作为技术工具,其输出完全受制于人类输入与算法逻辑,尚未形成主体性存在的认知基础。目前为止,生成式人工智能表现出日益强大的学习能力和一定的自主行为能力,但其所解决的问题仍只是封闭场景下的计算求解问题,无法对外界环境自行设定目标或计划,进行自主能动的反馈,远远没有“进化”表现出自我意识或者说能动性。

按照马克思对人作为社会关系总和的界定,人是兼具自然属性与社会属性的存在者,但是生成式人工智能只具有自然属性,不具有社会属性。生成式人工智能的自然属性源于其人工创制的本质,这一过程凝结了人类智力成果,其本质上仍属于人化自然的范畴。这种自然属性决定了人工智能的工具性特征:它仅是人类意识和实践延伸的技术手段,自身不具备价值诉求或目标导向。正如芒福德在《技术与文明》中所阐明的,机器本身既不提出要求,也不承诺实现何种价值,这些精神性任务本质上属于人类的主体性范畴^[17]。也就是说,生成式人工智能只是人类的附属工具,听从人类的指挥,没有主体性,更不会创造出各种社会关系,因为创造社会关系的前提是拥有自主意识与实践能力,迄今为止的生成式人工智能除了与发布指令的人类形成人机互动的关系之外,并不能够在相互之间形成主体间的关系,更不能与其他人主动发展出类似于人那样的关系,生成式人工智能并不具备人所特

有的社会属性,因而即使表现出一定的自主性,也无法构成主体。

更进一步,生成式人工智能也不能构成法律意义上的主体。哲学视域中的“主体”强调个体客观存在的本原性、先天性、具体性与独立性,其主体性地位独立于个体主观评判和社会制度建构。而法律意义上的主体作为规范性概念,特指由法律赋予一定权利和义务并能够参与法律关系的实体,包括个体或法人组织等。这就意味着,在法律范畴内,法律主体具有建构性和抽象性,只有具有法律主体资格的个体或组织才能参与并形成一定的法律关系,享有法律所保护的權利,承担法律所规定的责任或义务。民法发展的历史经历了从“人可非人”到“非人可人”的去伦理化过程,将法律人格的范畴扩大到自然人个体之外,通过权利能力的赋予使法人得以实现财产与实体的独立化,实现了上述自然人到法律人的转型。但是“法律是一个在生产它自己的基本法律行为的网络”^[18],法律主体的革新需要遵守基本的法律逻辑。实际上,法人主体地位的确立非但没有与唯有人作为主体的理论相违背,反而强化唯有人才可作为主体的理念。一方面,公司、社团等法人的建构,可以视为复数人的集合,离不开人的组织和执行机关,法人作为法律主体的核心要素仍是享有权利义务并承担一定责任的人。另一方面,非人组织构成法律主体的哲学根基,并非主张一种强的人类中心主义,而只是在强调最底线的人文主义,即有人的存在或在场。法人概念的提出确实有其实用主义的内核,但仍无法撼动法人构造的“以人为本”的人文主义底线^[8]。如果对完全自动化、脱离或独立于人而存在的生成式人工智能赋予法律主体地位,将从根本上背离了这一哲学宗旨。

此外,生成式人工智能还涉及如何理解本体论意义上的主体间性的问题。主体间性概念最早由二十世纪德国现象学家胡塞尔提出,经胡塞尔、巴赫金、加达默尔等学者的理论积淀,哈贝马斯通过交往理性学说进一步深化了这一范畴的内涵。就本质而言,主体间性是人的主体性在主体关系场域的延伸,与主体性形成共生关系。本体论视域下的“人—物”主体间性,突破了传统主客体二元对立框架,强调人与世界的辩证统一,构建起主体间对话、交往与理解的关系范式。在当今数智时代,人与生成式人工智能的交互必然需要纳入主体间性视域,建立“人—机”对话思维模式。但值得探讨的是,这种交互是否意味着要赋予生成式人工智能与人类对等的主体地位呢?马克思实践哲学提供了重要启示:其以存在论视角指出,认识活动

中的主体间性本质上根源于实践中的主体间关联——人类创造的物化成果不仅是认识与实践对象,更内在于实践过程,成为主体自我反观与塑造的参照系。生成式人工智能正扮演着这种参照系角色,而非独立的主体存在。人类借助人工智能生成的实践活动,实质是主体关系内容的拓展,在此过程中主体的权威性得以确证,实现了“人—机”主体间性向人类主体性的回归。

3 生成式人工智能不享有人格尊严

相比于主体论层面的论证或反驳,围绕生成式人工智能是否具有法律主体地位的争论更多集中在人格论层面。无论是法律人格扩展论,还是电子人格论或有限人格论,都竭力为生成式人工智能争取或赋予有限或工具性的人格,以此证成生成式人工智能的法律主体地位。但是,生成式人工智能并不享有人格尊严,赋予生成式人工智能相应人格,不仅无益于生成式人工智能权益的保护,而且会从根本上威胁人类人格尊严的保护与实现。

人格及其尊严的概念是启蒙运动以来高扬人性、追求文明进步的现代产物,人格尊严从先验、抽象和普遍的角度标示出人之为人不同于动物或其他事物的独特之处。康德明确指出,某些存在者的存在虽非源自人类意志而是基于自然,但非理性存在者仅具有作为手段的相对价值,故被称作“事物”;而理性存在者则被赋予“人格”之名。他强调:“人及一切理性存在者皆以目的之姿实存,其存在本身即具有绝对价值。”^{[19]438} 此处的“理性”特指纯粹实践理性,理性存在者“这种优先于一切单纯自然物的尊严(特权)使得它任何时候都必须从它自身的视角出发,但同时也要从任何其他有理性的、作为立法者的存在者(它们也因此也被称为人格)的视角出发来采用自己的准则”^{[19]438}。由此表明,人之所以构成尊严的承受者,根本上在于他具有本体的一面,即能够作为一个人格而普遍地立法和守法。对康德来说,这恰恰构成了人类道德性的根本来源,由于人类作为理性存在者“仅仅服从自己给自己立的法则”,因此“人格的价值就构成了道德的价值”^{[19]344}。“人格是其行为能够归责的主体。”^{[19]223} 除此之外,康德还特别强调“人惟有作为人格来看,亦即作为一种道德实践理性的主体”^{[19]434},亦即“作为本体的人”,才“拥有一种尊严(一种绝对的内在价值)”^{[19]434-435}。所以,作为本体的理性存在者之所以能够普遍为自己立法并且表现为守法的自律,这是因为理性存在者被提升到超越于其他一切存在者之上并

因而被赋予“等级和尊严”的缘故,从而也是区别于其他存在者并且赢得其他存在者尊重的根本理由。

换言之,一个人作为理性存在者,是能够克服兴趣或情感偏好影响而普遍地遵循道德法则的实践主体。藏在每个具体的个体背后那个道德性的主体,既是为自己立法的源泉,也是严格执法的个体;既是提出命令者,也是听从命令者。因为他必须在他所有的实践行为中都保持同一性原则,并且为自己的行为负责,所以他是具有同一性的人格。只有这样的理性行动者,才赋有一种绝对的内在价值,即尊严^[20]。人格及其尊严的概念,彰显出人类作为目的本身而非其他目的的手段或工具而存在的内在价值与绝对价值,因而不仅成为人类社会最重要的价值源泉,而且构成人权的重要基础,成为民法及世界各国宪法的立法伦理基础。

“民法人格从古罗马司法到现代民法的发展,历经了由具体的身份基础到抽象的伦理基础的演变。抽象人格的确立,使得近、现代民法上人与人之间的关系,在了无差异的人类伦理之上实现了真正的平等——身份等具体的人的特质被排除于人格基础之外。”^[21]然而,自然人平等人格的伦理预设虽然彰显出理性主义色彩,并不足以彻底解决法人人格的正当性难题。作为伦理人格的自然人,其人格来源于自我意识与自主决定的理性本质;而《德国民法典》通过创设“权利能力”概念构建法人人格,使法人仅具备法律拟制的单一资格,即在抽象意义上存在资质,但是到了亚里士多德意义上的承受者,仍然要归诸于具体的人类主体。更进一步来说,法人的权利能力本质上是立法技术的建构产物,既可以被赋予同样也可以被剥夺,背后起调节作用的仍然是人类主体;而自然人的人格及权利能力,并非源自法律授权,而是基于理性与自由意志的伦理存在,自出生即当然享有,具有先于法律建构的自然正当性和先天性,是最终的既无须其他存在者赋予又不可再被剥夺的承受者。总的来说,无论是自然人的伦理人格,还是法人的法定权利能力,虽然其来源和表现形式存在着差别,但是其终极根基都可以追溯到对人类人格尊严这一伦理价值的承认。

从马克思主义哲学实践观视域下进行审视,生成式人工智能是人类利用自身认识和改造世界的实践活动的产物,是人类千锤百炼而形成的高科技智慧综合体,是人类智能实践最后形成“一般智力”的丰硕成果^{[22][198]}。人作为主体在实践活动中所展现出的自我创造、自我开发和自我突破的非凡能力,自主性、反思性与无限性在其中交相辉映,是人类独有的精神瑰宝。

相较而言,生成式人工智能只是用于认识与改造世界的实践活动中的中介工具,它的运作是建立在数据算法对人类智能模仿与延伸的基础之上,就如没有灵魂的木偶,在人类赋予的逻辑和指令的丝线牵引下起舞,因而不享有作为目的本身和内在价值而存在的人格性。无论如何,在人与生成式人工智能的关系范畴中,人工智能始终是人类发明创造出来用以拓展人类自由、提高人类能力与效率的复杂工具,它从诞生到运行再到消亡的都是服务于人、以人类为中心的,因而其自始至终都只具有工具性的相对价值,而不可能具备像人类这般的绝对价值,也就不享有人格尊严。即使以后出现拥有自我意识的强人工智能,仍会由于工具价值的定位而必须遵从“阿西莫夫三定律”,无法获得与人类相提并论的人格性。

与此同时,有学者认为生成式人工智能具有工具性的人格属性^[23]。对于这种观点,只要追溯到哲学当中“人格”的根本意义,就能得出人格从根本上拒斥工具价值,因此将人格与工具、有限等词汇捆绑起来创生工具类人格、有限人格等概念的构词法既不严谨也无益处,本质上是后人类主义通过类比、隐喻等文学机制进行的过度想象。人格尊严彰显了人之为人的独特性,对生成式人工智能赋予人格尊严,将非人实体或存在转化为与人类同等重要的存在,既不利于人权保护也无助于生成式人工智能向善发展,最终结果是不断消解人之为人的独特性以及在此基础上形成的人格尊严,导致人的异化和人工智能的控制。

令人担忧的是,如今,生成式人工智能正以其不断扩展的自动化、非人化力量而成为人类享有人格尊严的巨大威胁。生成式人工智能的无节制发展,通过将人的身体、身份和行为数据进行大规模的采集和运算,导致侵犯隐私、控制精神、诱导消费、造假欺骗等道德失范问题,已经严重威胁到人作为目的本身和内在价值的独特性。在生成式人工智能的海量计算和算计过程中,人并不是被当作目的本身而仅仅是实现功利目标的手段而存在,人之为人的尊严在智能算法的自我运转中消失了。

4 生成式人工智能不能作为主体承担具体责任

赋予生成式人工智能主体地位的另一种具有较强辩护性的理由来自现实发展需要说,即由于生成式人工智能的大规模应用和自主性、智能化程度的日益提高,现有法律框架在实施过程中出现找不到法律主体、无法责任归因或无法追究责任等现实

困境。例如,在合同法领域,智能机器人代替人签订合同已经司空见惯,但是依托智能程序订立的买卖合同究竟表达了谁的“意思自治”,法律机制尚不明确。再例如,在侵权责任法领域,如果自动驾驶汽车在行驶过程中发生事故,造成伤害或侵权行为,如何进行责任归因成为棘手难题,究竟是由智能程序设计者、汽车制造者还是使用者抑或受害者承担责任,现行法律制度已难以作出有效判定。还比如,生成式人工智能将会导致显著的知识产权问题,但是赋予机器人知识产权,从根本上违背了保护创新的立法初衷^[24]。

因此,相关学者从实践需要的角度提出,赋予生成式人工智能主体地位或法律人格迫在眉睫,只有明确并建立生成式人工智能的责任分担机制,通过效仿公司法人、建立人工智能财产保险或“图灵登记”等法律设计,使其以特殊的权利义务承担者身份参与到民事法律关系中,才能解决实践需要的责任归因与归责困境^[25]。以自动驾驶技术引发的损害赔偿纠纷为例,当前司法实践中涉事企业普遍采取积极赔付态度,并持续推进技术迭代。例如谷歌、沃尔沃等自动驾驶公司已公开承诺对其研发的自动驾驶系统造成的损害承担全部法律责任。在此背景下,若主张通过构建生成式人工智能法律主体资格来免除人类责任,势将引发责任认定的制度性漏洞。从技术现实看,生成式人工智能系统故障致害与所谓“自主行为”的界限难以科学界定,多数损害事件实为人类操作与技术缺陷交织作用的结果。典型如“达芬奇”手术机器人致死事件,调查显示医生操作失误才是事故主因。值得警惕的是,此类事故中常出现责任转嫁现象——操作人员将过失归咎于设备,而制造商又以技术免责事由逃避产品责任,导致受害者求偿陷入困局。若进一步以实用主义逻辑赋予生成式人工智能主体地位,无异于为生产者和使用者创设责任转移的合法通道,最终将导致损害赔偿责任在人机之间空转,使救济机制流于形式。

赋予生成式人工智能法律主体地位真的可以有效解决现实生活中的矛盾和冲突吗?实际上,这种做法不仅无助于走出生成式人工智能产生的归责困境,反而由于引入新的法律主体造成更为复杂的归责局面,法律多项功能在生成式人工智能面前搁浅或者失效。首先,生成式人工智能无法理解法律规则,其行为主要受到算法等统计与数学规则的调节影响,法律规则难以指引生成式人工智能的具

体行为。其次,法律强制力难以作用于生成式人工智能。生成式人工智能缺乏主体意义上的实体存在,传统法律责任形式失去适用基础。若将销毁程序视为死刑、限制运行视为监禁,此类形式化惩罚既无法触及生成式人工智能的运行逻辑,也难以实现惩戒目标。再次,对生成式人工智能实施奖惩缺乏实质意义。因其无主观意志,奖惩无法产生类似人类的反思效果。尽管可通过算法模拟反馈机制,但本质仍是预设程序的机械响应,与法律规范人类行为的目的存在本质差异。最后,惩罚生成式人工智能无法实现情感救济。例如,因其“受罚”过程缺乏人文关怀,销毁程序难以平复受害者心理创伤。法律责任的情感抚慰功能需依托人类主体的道德认同,而生成式人工智能作为技术工具无法承载这一社会价值^[8]。

生成式人工智能技术早已不是农业时代的工具或工业时代的机器那么简单,而是叠床架屋为如芒德福所说的“巨型机器”或海德格尔所说的“座架”那样与人深度耦合的复杂系统,人们以不同身份在不同环节以不同机制发挥作用并促成生成式人工智能技术功能的实现,形成多个责任主体与复杂交互行为的“分布式责任”局面。不过,分布式责任只是拉长了行为之间的因果链条,加大归责的难度,但并不会导致责任的消失或转移。作为生成式人工智能的创造者或使用者,人们有责任将复杂系统各个环节各种机制的责任分布梳理清楚,并且进行明确的责任归因。即使算法黑箱导致的不透明或者生成式人工智能以意想不到的方式作出的带有一定自主性的行为,仍然可以采取共同责任或严格责任的方式予以归责。由于无论如何,人工智能都是人们为了某一目的创造出来的,因此必须为其创造或使用的整体行为承担波塞提出的“按照每个人对于系统所能起到的积极的或消极的影响”的共同责任,或者弗洛里迪所言的“无论其是否有罪都应负严格责任”^[26]的严格责任,而不是将责任转移到没有主体地位和人格资质的非人存在。

严格来说,生成式人工智能的各种自主或智能行为,仍只是依据人类以往经验或者数据作出的概率论选择,本身是人类意志与价值的延伸与投射,因此我们必须将责任明确归结到创造或使用这一人工物的具体的单数或复数的人,让更多人为这些难以控制的复杂的集体行为承担责任,他们中的所有人才会更加谨慎合理地运用生成式人工智能。

5 结 语

生成式人工智能的崛起,如同一面照见文明本质的镜子,既折射出人类智慧的璀璨光芒,也暴露出技术狂飙背后的存在焦虑。当ChatGPT能够撰写诗歌、DeepSeek可解微积分难题时,人类社会正经历着前所未有的认知革命:算法生成的文本能否承载真正的思想,机器决策是否蕴含自我意志?诸如此类问题的实质,是对“人何以为人”之哲学命题的终极追问。

生成式人工智能的发展必将重塑人类文明的形态,但文明的本质永远在于对“人之为人”的终极关怀。强调人文主义的底线并不是要走向狂妄自大的人类中心主义,研发与使用生成式人工智能的最终目标也不是要去创造“人不人”“机器不机器”的后人类主义,对于生成式人工智能的规训,仍然要坚守住人文主义的底线。当“小米”自动驾驶汽车在路口面临“电车难题”时,其选择不应由算法的功利主义来决定,而应铭刻着康德“人是目的而非手段”的道德律令;当生成式人工智能“诗人”创作出感人至深的诗句时,其价值不在于模仿人类的语言形式,而在于激发人类对存在本质的更深思考。

守住人文主义底线,绝非固守人类中心主义的陈旧立场,而是要在承认生成式人工智能技术价值的前提下,捍卫人类作为价值创造者的终极地位。守住人文主义的底线,意味着我们要在技术洪流中坚守人性中那份对美的感悟、对真的追求、对善的向往。只有这样,人类才能在与机器的共舞中进一步续写人类文明新的篇章,否则很有可能堕入生存性风险的“危崖”。

参考文献

- [1] 尹志强. 人工智能何以为“人”: 人工智能时代之民法因应[J]. 社会科学研究, 2023(1): 60-70.
- [2] 袁曾. 人工智能有限法律人格审视[J]. 东方法学, 2017(5): 50-57.
- [3] 许中缘. 论智能机器人的工具性人格[J]. 法学评论, 2018(5): 153-164.
- [4] 范进学. 人工智能法律主体论: 现在与未来[J]. 政法论丛, 2022(3): 3-17.
- [5] 张今杰. 人工智能体的道德主体地位问题探讨[J]. 求索, 2022(1): 58-65.
- [6] 韩旭至. 人工智能不是人: 从主体构建批判到非主体规制策略[J]. 大连海事大学学报(社会科学版), 2019, 18(4): 19-28.
- [7] 吴习彧. 论人工智能的法律主体资格[J]. 浙江社会科学, 2018(6): 60-66.
- [8] 韩旭至. 人工智能法律主体批判[J]. 安徽大学学报(哲学社会科学版), 2019, 43(4): 75-85.
- [9] 戴益斌. 人工智能伦理何以可能?: 基于道德主体和道德接受者的视角[J]. 伦理学研究, 2020(5): 96-102.
- [10] 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.
- [11] 阿伦·布洛克. 西方人文主义传统[M]. 董乐山, 译. 北京: 群言出版社, 2012.
- [12] 亚里士多德. 形而上学[M]. 徐海铭, 王厚平, 注. 上海: 上海译文出版社, 2023.
- [13] 皮科·米兰多拉. 论人的尊严[M]. 顾超一, 樊虹谷, 译. 吴功青, 校. 北京: 北京大学出版社, 2010.
- [14] 笛卡尔. 第一哲学沉思集[M]. 庞景仁, 译. 北京: 商务印书馆, 1986.
- [15] 康德. 道德形而上学的奠基[M]. 李秋零, 译. 北京: 中国人民大学出版社, 2005.
- [16] 马克思. 1844年经济学哲学手稿[M]. 刘丕坤, 译. 北京: 人民出版社, 1979.
- [17] 刘易斯·芒福德. 技术与文明[M]. 陈允明, 王克仁, 李华山, 译. 北京: 中国建筑出版社, 2009.
- [18] 托依布纳. 法律: 一个自创生系统[M]. 张骥, 译. 北京: 北京大学出版社, 2004.
- [19] KANT I. Kant's gesammelte Schriften [M]. Berlin: DeGruyter, 1900.
- [20] 罗喜. 康德论人的尊严: 兼论尊严的可丧失性问题[J]. 当代中国价值观研究, 2020(1): 24-35.
- [21] 张翔. 自然人格的法律构造[M]. 北京: 法律出版社, 2008.
- [22] 马克思, 恩格斯. 马克思恩格斯文集(第八卷)[M]. 中共中央马克思恩格斯列宁斯大林著作编译局, 编译. 北京: 人民出版社, 2009.
- [23] 袁曾. 基于功能性视角的人工智能法律人格再审视[J]. 上海大学学报(社会科学版), 2020, 37(1): 16-26.
- [24] 成素梅, 高诗宇. 智能机器人应有法律主体资格吗?[J]. 西安交通大学学报(社会科学版), 2020, 40(1): 115-122.
- [25] 杨立新. 人工类人格: 智能机器人的民法地位: 兼论智能机器人致人损害的民事责任[J]. 求是学刊, 2018(4): 84-96.
- [26] 弗洛里迪. 信息伦理学[M]. 薛平, 译. 上海: 上海译文出版社, 2018.