

维贝克技术人工物道德能动性思想的批判性阐释

曾庆娣, 陈心滢

(广州中医药大学 马克思主义学院, 广东 广州 510006)

摘要:技术人工物深度嵌入人类生活,催生了技术道德能动性议题。维贝克突破传统主客体框架,提出技术物通过行为调节参与道德实践,以意向性、自由与责任构建为分析框架,揭示技术设计通过价值嵌入实现道德物化,形成人技协同行动网络。但其理论存在三重局限:首先,将技术物对道德行为的工具性贡献等同于本体论层面的道德能动性,导致道德地位泛化;其次,消解技术人工物与自然物、基础技术与复杂技术的本体差异,模糊能动性边界;第三,忽视简单工具与高度先进技术人工物在道德影响机制上的层级差异。这些矛盾根源于后现象学方法论局限,过度依赖个体意识分析而缺乏技术生态维度。未来研究需超越个体意识层面,构建包含技术生态分析的伦理认知框架。

关键词:技术人工物;道德能动性;意向性;自由;道德责任;技术调节

中图分类号:B82-057;G301

文献标识码:A

文章编号:1008-7699(2025)03-0010-09

随着技术人工物深度嵌入人类生存境遇,人与技术的共生关系已突破传统主客体框架,催生出关于技术道德能动性的哲学重构。自温德尔·沃勒克(Wendell Wallach)与科林·艾伦(Colin Allen)于 2009 年提出人工道德主体(artificial moral agents,简称 AMAs)概念以来,^[1]学界围绕技术是否具备道德地位形成两大理论阵营:功能主义路径主张通过算法模拟人类伦理决策,构建具有“准道德能力”的技术系统,^{[2]195-204}而本质主义立场则坚持道德主体性必须基于意识与情感等人类特有属性。^[3]这种根本性分歧在近年衍生出三类研究进路:其一,戴维·J. 贡克尔(David J. Gunkel)通过解构“人类例外论”,将动物权利哲学扩展至智能机器,提出基于关系网络的道德地位判定模型;^[4]其二,尼克·博斯特罗姆(Nick Bostrom)基于技术能力进化论,预言超越人类道德地位的超级智能体可能通过内在属性获取伦理优先权;^[5]其三,以彼得-保罗·维贝克(Peter-Paul Verbeek)为代表的技术调解(technological mediation)理论,试图超越二元对立,将道德重新定义为“人一技交互实践中涌现的混合现象”。

在此学术脉络中,维贝克的技术调解理论具有特殊地位。其突破性在于将道德能动性解构为设计者意向、技术物引导与用户实践的建构过程,这既回应了艾伦等学者提出的“道德图灵测试”困境,又能为化解格兰特·G. 蒂加德(Grant G. Tigard)等学者提出的“责任鸿沟”提供新的分析工具。然而该理论面临三重质疑:技术意向性是否具备伦理本体论地位、混合能动性如何兼容个体自由意志以及调解过程的责任维度缺失问题。本文正是通过意向性、自由、责任三重维度,批判性分析维贝克理论合理内核,试图建立超越人类中心主义与技术决定论的新型分析框架,为人技伦理研究开辟实践哲学的新路径。

一、技术调解视域下技术人工物道德能动性问题再审视

“道德能动性”概念史可追溯至亚里士多德对“实践理性”的奠基,经康德确立为理性主体的自由意志与责任承担能力,其核心始终指向意向性、自主性与伦理反思的三重统一。20 世纪技术哲学转向中,布

收稿日期:2024-11-28

基金项目:广州市哲学社会科学共建课题(2022GZGJ106);广东省哲学社会科学规划项目(GD25CMK02)。

作者简介:曾庆娣(1980—),女,广东梅州人,广州中医药大学马克思主义学院副教授,法学博士。

鲁诺·拉图尔(Bruno Latour)突破人类中心主义,提出“非行动者”概念,为技术物的能动性提供理论入口。维贝克进一步将道德能动性重构为人技交互网络中涌现的实践能力,试图消解主客体对立。在这一转向中,维贝克提出的技术调解理论具有承前启后的意义。他突破传统的主客体二元对立框架,将技术人工物视为人类道德实践的共构者,主张道德能动性产生于人与技术的交互网络。然而,维贝克的创新亦暴露理论裂隙:传统道德能动性要求的意向性、自由等要素,在技术人工物情境下需重新校准标准,例如,自动驾驶系统的意图是否可类比人类意向性?算法决策的自由如何区别于程序确定性?当自动驾驶系统介入伦理决策、社交算法重塑价值判断时,如何确立技术人工物在道德实践中的主体性边界?这需要回归道德能动性的理论本源,在哲学传统与技术现实的张力中重构概念框架。

(一)技术调解理论对技术人工物道德能动性问题的探索

近几十年来,技术伦理学领域围绕技术人工物是否具有道德能动性的问题展开激烈争论。维贝克对此提出创新性见解,批判传统伦理学将技术视为人类主体附属工具的观点,揭示其忽视技术潜在道德影响力的局限性。他指出,“其根源在于现代主义,以及现代主义对人类主体和非人类客体的彻底分离。”^{[6]22}通过观察技术中介现象,他指出技术人工物通过调解作用积极介入并重塑人类行为与价值观,这种动态关系在固守主客二分法的人本主义伦理学框架中难以被充分概念化。维贝克将问题根源追溯至现代主义对人类主体与非人类客体的割裂,认为这种形而上学的二元对立遮蔽了技术人工物蕴含的道德维度。为此,维贝克主张“拓宽伦理学的人本主义基础”,建构承认物质性道德形态的技术调解理论,赋予技术人工物某种程度的道德能动性。^{[6]22}

这一理论建构可置于技术哲学两大经典问题视域中审视:其一,人类在多大程度上塑造了技术产品和技术过程?其二,技术产品和技术过程以何种方式塑造人类的行为和经验?维贝克采取中间立场,既否定技术完全受控于人类意志的极端工具论,也反对将技术视为独立实体的绝对决定论。他通过吸收决定论与实体主义的部分洞见,确立技术人工物作为道德行动主体的独特地位,形成兼具批判性与建设性的技术调解视角。该视角着重揭示技术如何影响人类生存体验并重构世界认知,其方法论融合伊德的后现象学、早期技术哲学及技术研究的经验转向,发展出双重分析路径:解释学路径聚焦技术人工物作为感知媒介如何改变现实呈现形态,存在主义路径则关注技术嵌入人类行动决策与生存实践的过程,探讨技术如何成为存在方式的构成性要素。这种双重路径突破传统伦理学的认知框架,为理解人技关系的伦理维度开辟了新的理论空间。

(二)技术调解理论对道德能动性概念的内涵拓展

在技术调解理论框架下,维贝克围绕两大核心议题展开探讨:一是如何将技术人工物纳入道德共同体,二是如何在人技互动中保留人类道德能动性的伦理空间。他认为,“这种扩展伦理学的最重要前提是道德共同体的扩展,即将非人类实体及其与人类的联系都纳入到道德共同体之中。”^{[6]39-40}他主张通过扩展道德共同体范畴,将非人类实体及其与人类的联结纳入伦理考量,强调技术人工物通过调解作用构建人与环境的关系,并潜移默化地塑造行为模式与感知经验。以驾驶汽车为例,技术使用过程中人的注意力通过汽车延伸至道路环境,只有当技术出现故障时才聚焦于工具本身,这揭示技术人工物具有双重调解维度——解释学维度通过调节感知诠释重构现实呈现方式,存在论维度通过介入行为实践重塑人类存在形态。

为实现对技术道德作用的公正审视,维贝克对道德能动性进行根本性重构。他提出两条论证路径:一是,孤立的技术或人类均不具独立道德能动性,这种属性仅存在于二者共同构成的新型混合实体中,将能动性源泉从人类转向人技互动过程;二是,在拓展技术人工物能动性内涵的同时,不是消解人类的主体地位,而是建立二者互补共生的关系。这种重构既承认技术对道德实践的积极影响,也促使人类重新审视自身在技术演进中的定位,揭示人技能动性的非同一般性与相互依赖性。

维贝克试图打破主客二元论,拒绝将道德能动性简单归属于人类或技术人工物,转而强调二者通过相互构成与共同塑造形成不可分割的整体。他借助“混合”(hybrid)概念阐释道德行为模式,以枪支射击

为例指出能动性应归于“枪手”这一人技集合体,而非单独的人或枪,即“枪手,作为人类和非人类元素的混合体”。^{[6]64} 他继续指出,“枪和人形成了一个新的实体,这个实体进行射击”。^{[6]64} 这种分析延续了拉图尔对传统二元划分的批判,聚焦元素间动态交织的过程。维贝克进一步论证,技术并非被动工具,而是通过塑造道德决策环境、影响判断形成与执行,成为道德行动的积极参与者。技术调解理论由此揭示人技互动网络的复杂性,展现道德主体性与技术人工物的深度融合与共生状态,为理解伦理实践的物质性维度开辟新路径。

(三)技术调解理论深化道德能动性内涵的意义分析

维贝克借鉴拉图尔行动者理论重构技术人工物道德能动性概念,但两者对人类与非人类关系的理解存在本质分歧。拉图尔主张以对称性视角消解人类与非人类的先验界限,强调事物同样具有“行动”能力,旨在论证非人类实体成为道德行为者的可能性。维贝克则通过后现象学方法超越这种对称性原则,认为无需否定人技差异,而应聚焦二者的相互作用与共同构成。他提出,人类与技术人工物通过动态交织形成“混合实体”,后人类主义伦理学的关键并非否认二者间的区别,而是反对将二者绝对分离。维贝克认为,技术伦理的核心问题不在于技术能否成为与人类同等的道德行动者,而在于其具体道德角色及调解作用的运作机制,这构成技术人工物道德能动性理论的根本挑战。

为突破困境,维贝克提出道德能动性构成论,强调人类与技术人工物在能动性构建中扮演不同角色。在《关于技术的道德意义的一些误解》中,他指出,“调解理论只是承认人类与人工物在道德能动性的构成中可以扮演不同的角色。”^{[7]85} 他明确指出,调解理论的核心在于揭示人技因素在道德能动性构成中的差异化贡献,而非争论技术是否具备固有道德属性。这种视角将道德责任视为人技互动中动态分配与共同塑造的产物,而非仅依附于人类主体的内在特质。然而,维贝克未明确阐明“构成”的具体机制,导致出现遗留问题:道德能动性究竟是被赋予的固有属性,还是多重因素互动中涌现的现象?这一理论空白引发了对能动性本质的重新审视,要求进一步探索人技要素如何通过具体实践形成伦理关系网络。维贝克的突破性在于将伦理分析从实体归属转向过程性构成,但其未完成的机制阐释仍构成技术调解理论深化的关键路径。

二、技术人工物道德能动性的三维构成解析

为克服传统伦理学理论中道德能动性单一赋予人类个体的理论局限,维贝克重构意向性与自由的概念框架,这种重构旨在揭示道德能动性如何在人类与技术人工物的交互作用中动态生成,为理解技术调解环境下的道德责任分配提供了新的理论框架。维贝克对道德能动性的传统必要条件进行了细致的区分,这些条件通常归结为:意向性—形成意向的能力—实现其意向的自由。^{[6]70} 维贝克通过后现象学方法重新诠释了这两个标准,将它们的应用范围拓展至技术人工物领域。基于这一理论框架,维贝克深入探讨了技术人工物在何种情境下、以何种方式展现出道德能动性的特征,并进一步引申出关键问题:技术人工物应当在何种意义上被赋予责任或承担何种形式的责任?

(一)技术人工物的意向性:道德能动性的基石探讨

在技术调解理论中,意向性作为道德能动性的核心维度,不仅塑造主观意愿,更深刻影响人与世界的交互方式。维贝克区分了伦理学与现象学两种意向性概念:前者指向形成意图的能力,后者则植根于存在论与具身性视角,强调人类对现实的指向性。当技术成为人与世界的中介时,其作用不仅在于辅助人类解释世界,更在于通过这种解释框架影响行动决策。维贝克提出,“可以‘具有’拉丁文‘intendere’一词字面意义上的意向性”,“intendere”的意思是“指导”“指导一个人的路线”“指导一个人的思想”。^{[8]73-103} 技术人工物虽不具备意识,却可通过“intendere”(拉丁语意为“指导”)的意向性形式引导人类行为与体验。这种分析聚焦技术如何作为中介重构人类对世界的理解逻辑与行动逻辑,揭示技术伦理责任的深层内涵——其价值不仅在于功能属性的道德评判,更在于重塑人类认知与道德实践的底层结构。

为阐明技术调解意向性的多元形态,维贝克提出三种意向性模式。首先是中介意向性,他继承并拓展伊德的理论,强调技术通过辅助现实解释间接塑造人类意图,例如,眼镜矫正视觉缺陷后重构用户的空间感知与行为选择。其次是复合意向性,聚焦技术设计对意图的激发机制,以劝说性技术为例,维贝克指出:“人们使用这类技术的意图与其说是‘中介’的,不如说是‘诱导’的,它们不是‘使用’技术的结果——相反,技术利用人类在这里完成他们的工作。”^{[7]83}意向性超越了单纯的中介角色,转而呈现出一种诱导性。最后是赛博格意向性,揭示人机深度融合下的新型意向生成机制,以深部脑刺激技术(DBS)为例,当电极植入大脑直接调控神经活动时,技术不再仅是外部中介,而是与人体共同构成具有内意向性的混合实体,其决策过程既非纯粹生物性也非机械性,而是两者交织的产物。这三种模式共同勾勒出技术从工具性中介到构成性要素的演进图谱,凸显意向性从被动调节向主动生成的范式转变。

维贝克的意向性理论遭到技术哲学家马丁·彼得森(Martin Peterson)的质疑。彼得森承认技术能构建新现实体验,但质疑其道德属性的明晰性:若仅以现实重构能力界定道德意向性,则自然物例如攀登马特洪峰引发的震撼体验也应被赋予伦理意义,这显然模糊了技术人工物的特殊性。对此,维贝克指出该批评隐含激进人本主义偏见,忽视技术调解对人类道德意向的构成性作用。他强调,即使技术人工物设计初衷不包含明确伦理导向,但它通过重塑感知框架与行动可能性,深度参与道德判断的形成。以超声波技术为例,其医学诊断功能虽非为影响生育决策而设计,却在实践中将胎儿健康问题从“命运安排”转化为“伦理抉择”,重构了生育实践的道德图景。这种中介作用超越设计者与使用者的直接意图,展现技术人工物作为道德催化剂的核心价值:它们通过改变现实的呈现方式与互动条件,重塑道德实践的认知基础与行动路径。

维贝克的回应凸显其理论的核心突破:道德意向性并非人类主体的独占领域,而是人技交互网络的涌现属性。技术人工物通过调解认知框架、扩展行动可能、重构实践情境,持续参与道德意义的生成与分配。这种视角既避免将技术简化为被动工具,又规避赋予其拟人化主体地位,转而聚焦动态交互中伦理关系的物质性建构。理论挑战在于,如何在承认技术调解力量的同时,厘清人类责任与技术影响的边界,这要求超越传统的主客二分,在混合实体的构成性过程中重新定位道德能动性。这就指向技术伦理学未来发展的关键方向:构建既能解释技术物质性调解机制,又能为伦理判断提供清晰的理论范式。

(二)技术人工物的自由:元能动性的形式分析

在技术调解理论的语境下,自由作为能动性不可或缺的组成部分,对于理解技术人工物如何影响人类行为及其道德责任具有重要意义。道德能动性要求行动者不仅具备明确的行动意向,还须享有实现这些意向的自由。依据以赛亚·伯林(Isaiah Berlin)的划分,自由可区分为消极自由与积极自由:前者强调免于外部强制或干涉的状态,后者侧重于个体拥有并执行自主选择的能力。面对现代技术的深入渗透,一个常见的担忧是,技术调解下的行为可能丧失其自主性与自由性,从而质疑其作为道德能动性体现的合法性。这种观点倾向于将技术视为一种外在的、削弱个体自由的力量。维贝克却认为,柏林的积极自由概念能够更恰当地阐释技术人工物与自由之间的关系。

维贝克将自由重新定义为行动者与技术关联的特征,而非独立存在的无拘状态。他指出:“我们不应把不受(技术)影响作为道德能动性的先决条件,而应把自由重新解释为行动者与决定他或她的事物相联系的能力……技术‘本身’不可能是自由的,但人类也不可能是自由的。自由是人类与技术关联的一个特征。”^{[6]60}技术作为行动媒介非但不限制自由,反而通过扩展选择范围与行动手段,使特定情境下的道德判断成为可能。传统自由观念将技术视为对立物,维贝克则强调技术调解构成能动性的新基础,技术既不替代人类决策,也不削弱自主性,而是通过重塑行为可能性为道德实践开辟新路径。这种自由概念突破传统范畴,将技术调解功能视为能动性的本质展现:其道德意蕴不在于赋予独立行动能力,而在于动态调节道德行为与决策过程,形成“人—技共构”的伦理实践格局。技术由此从压制性工具转化为自由实现的构成性条件,推动伦理分析从主客对立转向关系性生成。

为了阐明技术人工物道德能动性的特殊形式,维贝克通过明确区分“自由”与“自主”概念,得出技术人工物的自由是“关系自由”的推论。根据德沃金的梳理,“自主的核心理念来自希腊语 autos(自我)和 nomos(规则或法律)的词源。Autonomia 一词最初用于描述城邦,在城邦中,公民制定自己的法律,不受其他主导力量的控制。”^{[9]12-13} 这一概念后来被应用于个人,意指个人的自决或自治。在道德哲学领域,个人自主的概念被理解为界定自由道德主体的特征。传统研究聚焦心理动机却忽视技术调节,这就凸显了维贝克技术人工物自由概念的独特之处。维贝克指出,自主概念在技术语境中是有问题的。许多行动和决策都是被技术调节了的——这并不意味着我们不再是我们所做的东西的主人,而是意味着我们不再是唯一的主人。在技术调解理论的语境中,绝对自由不可能是道德能动性的先决条件,自由不是通过摆脱技术的影响而产生的,而是通过与技术建立关系而产生的,自由的这种形式可称之为“关系自由”。

对于人的主体性而言,究竟什么是技术人工物的关系自由?在福柯主体构成的实践论基础上,维贝克推出人类通过自由使用技术人工物实施行动或做出决定从而构成自身道德主体性的论断。他指出:“从道德能动性的主流观点来看,自我实践可被视为一种元能动性——旨在塑造自身能动性的能动性。”^{[6]87} 人与技术人工物互动实践中建构的关系自由,体现的是塑造人类主体性的元能动性。维贝克多次指出,重新定义道德能动性是在伦理学中公正对待技术道德意义的唯一途径,他的伦理方法侧重于发展人类与技术中介的自由关系的能力,即实现“关系自由”。“关系自由”的理念表明,在维贝克那里,技术人工物自身不可能是自由的,人类自身也不可能是自由的,真正的自由在于:技术人工物与人类主体形成统一体,这种统一体构成了自由的居所。技术成为扩展人类自由边界、深化主体体验的积极力量。

(三)技术人工物的道德责任:混合能动性的意义阐释

道德能动性是道德责任的先决条件,因而技术人工物道德能动性的第三个构成维度涉及道德责任议题,关键在于厘清人类与非人类实体之间道德能动性的混合如何重塑道德责任。在维贝克那里,技术人工物参与了人类的道德实践,意向性和自由是属于人与技术人工物混合实体的。这种混合道德能动性的观点让研究者感到困惑,因为在哲学传统中,道德能动性一直是属于人类的概念,大多数研究者倾向于将自由和意向性不可分割地统一于人类之中,道德责任通常被归因为道德行为人。基于这样的理解,维贝克的观点将导致难以赋予人类相关责任的重大理论困难:人类究竟在何种意义上承担由技术人工物诱发的行为负责?技术人工物自身是否也要担负相应的道德责任?为回应这些问题,维贝克认为需要区分两种责任:因果责任和道德责任。

具体而言,因果责任关注事件发生的直接原因链,即某人在因果关系上成为某事件或状态的原因时,就需承担因果责任,但这并不直接等同于道德责任的判定。道德责任更侧重于行为主体的目的性、自由意志及行为后果的道德评价。也就是说,道德行为并不自动等同于承担道德责任,而技术人工物在特定情境下,依据其影响行为与决策的方式,可能间接承担某种形式的道德责任,尽管这种责任的本质与人类的道德责任有所不同。因此,维贝克的混合道德能动性观点不仅促使我们重新审视道德责任的传统分配模式,还启发我们探索技术与人类共生时代背景下道德责任的新维度与新边界。

在有关道德责任的判断中,行为是否由行为者自主做出是一个关键的要素。因此,有关道德责任的争论就聚焦到一个核心问题:“什么使得一个行为成为行为者自己的行为?”^{[10]137} 对这个问题的回答,直接关系到维贝克有关技术人工物道德能动性的论证是否成立,关系到对道德能动性和道德责任的真正本质的理解。一般说来,传统哲学对自主性的研究并没有深入到技术层面,并未关注对行为者的行动和决策起作用的技术维度。在高度技术化的社会中,人的行动和决策受到技术架构的影响很大。由前面有关维贝克区分“自主”与“自由”的相关论述可推论出,承担道德责任的是人类与技术人工物的混合实体,但显然维贝克并不认为道德责任在人类与技术人工物之间平等分配,也没有得出结论认为技术人工物要担负道德责任。如果放弃道德能动性与道德责任的关联,那么就需要阐明技术人工物为何不需要承担责任,但维贝克并没有对此给出证明。对于这些批判性质疑,一个合理的解释是:在维贝克那里,技术人工

物的能动性,只是一种以人类与技术人工物混合的能动性为基础的派生的能动性。有待进一步追问的是,承认这种混合实体具有道德能动性,进而推论出其需承担道德责任,有何意义?

维贝克对此做出回应。在他看来,承认人类与技术人工物混合实体的道德责任,其意义在于强化人类主体的责任。人类不是在真空中承担责任的,而是在以技术人工物为中介的境况中承担责任。正如超声波的例子所显示的那样,产科超声不会强迫父母在怀唐氏综合症孩子时堕胎,人类主体的道德力量来自于通过批判性地审察技术如何塑造选择情境和解释框架,使得主体为技术调解下的能动性承担责任成为可能。技术人工物从根本上塑造了人类主体的责任感以及如何承担这种责任的方式。更为关键的是,在维贝克看来,技术人工物通过影响我们对道德价值观的定义和体验来调解我们的价值框架,而承认技术人工物的这种道德作用,使得人类主体能够以负责任的态度来指导技术设计、实施、使用和评估。“只有当我们理解技术在实际中的确切角色时,我们才能围绕技术重组实践。看到技术的道德意义让我们更有责任感,而不是更少。”^{[7]85}

尽管伊德提出质疑,主张不宜过分夸大技术人工物蕴含的道德意义,强调需正视技术人工物性质的多元性以及设计者意图的不明确性,进而指出技术调解效应具有难以预见的特性。维贝克在肯定伊德分析合理性的同时,进一步提出,尽管人类与技术人工物的互动关系确实展现出不可预见性,但这并不意味着人类主体完全无能为力。当前技术发展轨迹明确展示了人类具备通过“向善”设计技术人工物来施加影响的能力,这一观点揭示了探索技术人工物道德责任归属的复杂性,我们需积极承担起通过设计促进技术向善的责任。他强调,正是基于人类能够有意识地引导技术向有利于社会福祉的方向发展的能力,探讨技术人工物的道德责任才显得尤为迫切和必要。因此,在承认技术调解作用是不可完全预测的同时,更应聚焦于如何提升人类在设计过程中的道德意识与责任感,以确保技术发展的成果能够最大化地服务于人类的共同利益。

三、技术人工物道德能动性思想的限度及方法论反思

维贝克关于技术人工物道德能动性的深入探讨,为技术伦理学领域开辟了新的概念疆域与价值评判维度,展现了技术人工物如何重塑人类主体的经验结构,揭示了技术人工物与人类主体的互构关系,即技术人工物不仅作为工具被使用,更在一定程度上塑造了人类的行为模式、价值取向乃至道德判断,这种分析框架无疑为理解技术伦理的复杂性提供了新的路径。然而,与任何理论探索一样,维贝克的论证不可避免地面临着内在的理论边界。这些局限性,通过细致审视其所采用的后现象学方法论视角,能够得以更为详尽的分析与反思。

(一)维贝克技术人工物道德能动性思想的限度

首先,维贝克过度扩展了技术人工物的道德地位,将其对道德行为的贡献等同于道德能动性的体现。这一立场在学术界引发分歧。反对者普遍认为,尽管技术人工物在道德实践中占据一定地位并能够影响或促进道德决策与执行,但并不等同于其具备认知能力与道德能动性。认知能力与道德能动性作为道德主体的核心特征,涉及对情境的理解、价值判断以及自主选择与行动能力。相比之下,技术人工物,尽管设计时可能已融入了道德考量,其运作却依然依赖于预设规则或算法,而非自主的道德推理与判断。因此,将技术人工物在道德层面的作用提升至道德能动性的高度,就忽视了能动性概念的深层次内涵及其对主体性的要求。以计算机为例,黛博拉·G·约翰逊(Deborah G. Johnson)认为,“计算机是道德实体,而非道德主体,因而不具有道德能动性。”^{[2]195} 在技术人工物寻求道德主体地位的议题中,一个核心前提在于它们需具备在善恶框架下进行道德判断的能力。

探讨道德判断的根基可追溯至休谟的经验主义与康德的先验主义。康德认为道德判断源于人类理性本质,理性构成其道德能动性核心;休谟则主张道德判断植根于情感,如移情与同情。据此,康德式道德主体需具备理性能力,休谟式道德主体则依赖情感能力。当前技术人工物既无法满足理性前提,又缺

乏情感基础,因此无法成为真正道德主体。尽管技术人工物通过调解行为扩展了道德能动性范围,但其根基仍完全归属于人类。维贝克的理论价值在于揭示技术在现代社会的强大影响力,并呼吁以责任伦理引导技术开发与应用,但其局限在于过度放大技术人工物的道德功能,模糊了人类作为唯一道德主体的根本属性。

其次,维贝克取消了技术与技术人工物、技术人工物与自然物之间的区别,陷入泛化道德能动性的理论风险。一方面,在维贝克技术人工物道德能动性的相关论述中,并没有区分技术与技术人工物,而事实上技术不能等同于技术人工物。人工物指实物,而技术是人工物、社会实践、社会关系和知识体系的组合。^{[2]197} 如果没有知识体系、社会实践和人际关系,技术人工物是不存在的。只有在人类社会实践的背景下,技术人工物才会被制造、采用、传播、使用并具有意义。技术人工物通过社会实践而产生,作为社会实践的一部分被人类使用,只有在其被认可和使用的特定语境中才能彰显意义。维贝克模糊技术与技术人工物的区别,可能导致泛化道德能动性的理论后果。

再次,维贝克的理论模糊了技术人工物与自然物的界限。例如,原始猎人投掷的木棍既是自然元素,又是技术工具;现代生物技术产物如转基因食品与药品,更呈现了自然与人工相互渗透的复杂关系。这些案例挑战了传统分类框架,使自然物与人工物的边界趋于消解。然而,若彻底放弃自然与人为的区分,不仅会导致评估人类行为及其环境影响的标准缺失,还将削弱技术伦理判断的精确性,使技术发展方向失去必要约束。尽管技术人工物与自然物共同参与塑造人类行为,但若将其完全等同并泛化赋予物体道德能动性,既违背道德责任需精准追溯主体的基本原则,也会阻碍对技术社会后果的深度解析与风险预判。维贝克的洞见揭示了物性关系的动态本质,但其理论激进性可能掩盖技术伦理中人类主体的核心地位。

最后,维贝克的理论框架在探讨技术人工物的道德能动性时,未能充分区分其内部多样性的层次,特别是忽略了简单技术人工物(如铅笔刀)与高度先进技术人工物(如自动驾驶汽车)之间的本质差异。这种差异提示我们,复杂技术系统能够展现出更为显著的潜在道德主体性特征,从而促使我们构想一个多维度的理论谱系,以精准映射自然物、人类与各类技术人工物之间的智能与道德能力分布。在此理论谱系中,自然物与人类作为两个极端,分别代表了未经人工干预的自然状态与高度发展的自主道德判断与行动能力。二者之间的广阔地带由不同等级的技术人工物所填充,其中,人工智能机器人作为新兴且具有重大影响的实体,尤为值得深入探究。

至关重要的是,机器的意向性不应被视为一个抽象或难以捉摸的“黑箱”概念,而应视为衡量其是否具备真实智能的关键先决条件。随着智能技术的不断演进,预见高度智能化的人工智能将趋向形成成熟的道德能动性成为可能。这一前景要求我们在讨论技术人工物的道德能动性时,必须采取一种差异化的公正视角,正视并尊重各类技术人工物之间存在的显著差异。

维贝克提出的泛化假设:所有人类与技术人工物的混合体均同等地具备道德能动性,这种假设不仅忽略了技术进步的多样性,还可能误导我们对复杂技术人工物道德地位的准确评估。相反,我们应秉持开放态度,为探索人类与非人类实体间互动关系的深层道德意蕴预留足够的理论空间,以便更全面地理解和应对技术发展带来的伦理挑战。

(二)维贝克技术人工物道德能动性思想的方法论反思

维贝克关于技术人工物道德能动性的理论框架,其局限性深刻植根于其采用的后现象学方法论之实证性倾向。后现象学方法聚焦于人类与具体技术人工物之间交互作用的直接经验,致力于剔除主观臆断与预设理论框架对技术人工物认知的干扰。然而,此方法论的局限在于它倾向于将技术人工物的本质简化为其实用功能,忽略了技术人工物超越其实用性的内在自主力量与潜在影响。

在题为《什么是后现象学》的演讲中,伊德认为后现象学具有三个重要特征:其一,在后现象学方法中,现象学成为实用主义的补充或丰富。其二,在后现象学中,实用主义因现象学而得到加强。其三,后

现象学关注技术与人类和世界的相互关系,尤其聚焦于对技术的中介作用进行经验分析。因此,维贝克将后现象视为一种经验哲学的典型形式,认为“后现象学研究总是将实证工作作为哲学反思的基础”。^{[11]21} 他重提现象学的口号“回到事物本身”,并且认为,西方理论界对事物的性质评价过低,“哲学没有认识到事物及其物质性的意义”。^{[12]2} 现象学的口号表明,现象学的意图是通过引入物本身作为哲学思考的一个重要部分,也表明现象学有别于人类中心主义和唯我论视角中仅仅关注主体的方法。胡塞尔将“物”视为“现象的基本结构和表象可能性条件”的一部分。^{[13]94} 毋庸置疑,现象学与后现象学有着密切联系,但越来越多的文献提出了后现象学是否足够“现象学化”的问题。^{[14]1501} 这是对后现象学基本原理的质疑,其焦点在将两个框架联系起来的纽带。有的研究者认为,后现象学中的经验要素过于扎实^{[15]6};而有的研究者则认为,后现象学为我们提供了一些核心现象学概念的深刻诠释^{[16]2}。另外,许多研究者也对当代技术哲学中超验元素的作用以及经验转向的作用提出了质疑。

如果说,海德格尔、乔纳斯或埃吕尔等思想家所主张的“古典技术哲学”关注的是使现代技术成为可能的历史和超验条件,那么可以说,芬伯格、哈拉维、伊德、维贝克等思想家所主张的新技术哲学,关注的则是伴随技术文化发展而来的技术与社会的共同演变。后现象学非常强调方法的经验性。维贝克将这种新型技术哲学的方法论诉求概括为:我们需要一种基于“实际技术”的“经验主义”方法,而不是应用“预设”且过于抽象的“传统”方法。但是,维贝克后现象学理论所声称的“面向事物本身”中的“事物本身”不是“现象学意义上的”,而是“经验上的”。事实上,现象学是以现象学还原方法而非以经验的方式去分析经验中被给予的事物。实际上,在维贝克的经验理论视野中,技术的中介力量被简化为技术的实用功能。

维贝克以实用主义的方式理解和对待技术人工物,把技术人工物的技术性简化为使用意义上的实用性,即简化为技术在人类语境中的有意义的功能,这就可能导致无法合理地对待技术本身塑造人类与世界之间关系的能力。后现象学的理论重心是研究人与技术关系的中介问题,这种哲学并没有把技术本身当作关键的要素给予足够的关注,因为技术被简化为主观或客观两个极点,没有成为关系的第三极。建立关系的两极是主体和世界,技术则依附于这两者,技术虽然重要但仅仅是辅助角色,与此同时,技术只能根据主体和世界这两极产生的关系场域来定位自己。如果后现象学理论只是把技术视为“中介”,就会错过技术本身。根据后现象学的观点,意向性代表了技术对现实某些部分的指向性,而另一方面,中介发生在“人—技术”关系中,通常表现为一种转换活动。即是说,意向性提供了一个方向,中介则根据意向性提供的方向改变体验。从这个意义上,我们可以认为,意向性先于中介而存在,中介是以意向性为前提的。技术的意向性是技术本身的一个更基本的要素,如果不理解这一点,就无法把握技术的本质。此外,技术意向性概念在后现象内部造成了一种概念上的张力。一方面,后现象学接受技术有其意图,另一方面,它又坚持认为技术已被“人—技术”关系所定义,因此也被其中介活动所定义。显然,这两种观点不能同时成立。要解决这个张力,就不能仅仅依赖后现象学的观点,而是需要对技术意向性进行更深入的探究,可能的路径就是关注技术本身。

四、结论

作为现象学的方法论自觉,后现象学需避免将实证科学成果直接等同于现象学直观,转向建构中介性理论框架以揭示二者深层关联。技术不仅通过实用性功能直接介入意向性结构,更以非透明方式间接调节人与世界的存在论关系,其功能性运作常超越主体意向性边界,甚至突破胡塞尔所言的主体间性范畴,触及先验领域的超越性维度。技术哲学须将这种隐匿的技术意向性及其非透明运作机制纳入本体论考察。应在人机共生的技术哲学视域下重构维贝克的技术人工物道德能动性理论,既回应技术中介化生存的伦理可能性,亦需警惕其过度泛化倾向。这种批判性反思既构成中国特色伦理话语体系应对技术现代性的现实语境,也是人机互构时代重构美好生活范式的理论基础。

参考文献:

- [1] WALLACH W ALLEN C. Moral machines: Teaching robots right from wrong[M]. Oxford: Oxford University Press, 2009: 10.
- [2] JOHNSON D G. Computer systems: Moral entities but not moral agents[J]. Ethics and information technology, 2006 8: 195-204.
- [3] BOSTROM N YUDKOWSKY E. The ethics of artificial intelligence[M]//FRANKISH K, RAMSEY W M. The Cambridge handbook of artificial intelligence. Cambridge: Cambridge University Press, 2014: 316-334.
- [4] GUNKEL D J. The machine question: Critical perspectives on AI, robots, and ethics[M]. Cambridge, MA: MIT Press, 2012: 67-112, 189-230.
- [5] BOSTROM N. Superintelligence: Paths, dangers, strategies[M]. Oxford: Oxford University Press, 2013: 109.
- [6] VERBEEK P-P. Moralizing technology: Understanding and designing the morality of things[M]. Chicago: University of Chicago Press, 2011.
- [7] VERBEEK P-P. Some misunderstandings about the moral significance of technology[M]//KROES P, VERBEEK P-P. The moral status of technical artefacts. Berlin: Springer Netherlands, 2014: 75-88.
- [8] VERBEEK P-P. Morality in design: Design ethics and the morality of technological artifacts[M]//KROES P, VERMAAS P E, LIGHT A, et al. Philosophy and design. Dordrecht: Springer Press, 2008: 91-103.
- [9] DWORKIN G. The theory and practice of autonomy[M]. Cambridge: Cambridge University Press, 1988: 12-13.
- [10] 刘晓飞. 道德责任与多重可能性——法兰克福式反例的启示[J]. 世界哲学, 2016(6): 130-137.
- [11] ROSENBERGER R VERBEEK P P C C. Postphenomenological investigations: Essays on human-technology relations[M]. Kentucky: Lexington Books, 2015: 21.
- [12] VERBEEK P-P. What things do: Philosophical reflections on technology agency and design[M]. Pennsylvania: Penn State University Press, 2005: 2.
- [13] ZAHAVI D. Husserl's phenomenology[M]. New York: Stanford University Press, 2003: 94.
- [14] RITTER M. Philosophical potencies of postphenomenology[J]. Philosophy and technology, 2021, 34: 1501-1516.
- [15] SCHARFF R C. On making phenomenologies of technology more phenomenological[J]. Philosophy and technology, 2022 (3): 6.
- [16] HANFF W A. Weaving science and digital media: Postphenomenology's expanding hermeneutics[J]. AI & SOCIETY, 2023 (6): 2339-2345.

A Critical Interpretation of Verbeek's Thought on the Moral Agency of Technological Artifacts

ZENG Qingdi, CHEN Xintian

(School of Marxism, Guangzhou University of Chinese Medicine, Guangzhou 510006, China)

Abstract: Technological artifacts are deeply embedded in human life, giving rise to the issue of technological moral agency. Verbeek transcends the traditional subject-object framework by proposing that technological artifacts participate in moral practices through behavioral mediation. He constructs intentionality, freedom, and responsibility as an analytical framework, revealing how technological design achieves moral materialization through value embedding, thereby forming a collaborative human-technological action network. However, his theory exhibits three key limitations: First, equating the instrumental contribution of technological artifacts to moral behavior with moral agency at the ontological level leads to the generalization of moral status. Second, dissolving the ontological distinctions between technological artifacts and natural objects, as well as between foundational and complex technologies, blurs the boundaries of agency. Third, his theory overlooks the hierarchical differences in moral influence mechanisms between simple tools and highly advanced technological artifacts. These contradictions stem from the methodological constraints of post-phenomenology, which overly relies on individual consciousness analysis while neglecting the dimension of technological ecology. In the future, researchers must transcend the individual consciousness level and construct an ethical cognitive framework that incorporates technological ecological analysis.

Key words: technological artifacts; moral agency; intentionality; freedom; moral responsibility; technological mediation

(责任编辑:傅游)