

人工智能时代人的主体性困境及其应对路径

常博文

北京联合大学马克思主义学院, 北京

收稿日期: 2026年7月5日; 录用日期: 2026年7月27日; 发布日期: 2026年8月6日

摘要

人工智能的迭代升级正深刻重塑人类的生产生活方式、思维模式与社会关系,它在解放人类体力与脑力、提升社会生产效率的同时,也对“人作为主体”的核心地位产生了前所未有的冲击,引发了人的主体性困境。这一困境突出表现在人的认知能力退化、实践能力减弱以及情感能力钝化,其深层缘由在于技术资本逻辑驱动下的异化、数据时代的认知重构以及人类自身在技术浪潮中的价值迷失,但人工智能也蕴含着增强主体性的潜能,从个体自觉、技术伦理与社会教育三个层面探讨主体性复归的可行路径,以期呼唤在拥抱技术红利的同时坚守人的主体地位,实现人与人工智能的良性共生。

关键词

人工智能, 主体性, 技术依赖

The Dilemma of Human Subjectivity in the Era of Artificial Intelligence and Its Countermeasures

Bowen Chang

School of Marxism, Beijing Union University, Beijing

Received: July 5, 2026; accepted: July 27, 2026; published: August 6, 2026

Abstract

The iterative upgrading of artificial intelligence is profoundly reshaping humanity's modes of production and daily life, thinking patterns and social relations. While liberating human physical and mental labor and boosting social productivity, artificial intelligence exerts an unprecedented impact on the core standing of humans as subjects, giving rise to the dilemma of human subjectivity. Such a dilemma is prominently manifested in the degradation of human cognitive capacity, the

文章引用: 常博文. 人工智能时代人的主体性困境及其应对路径[J]. 哲学进展, 2026, 15(8): 202-210.

DOI: 10.12677/acpp.2026.158381

decline of practical ability and the numbness of emotional competence. Its underlying causes lie in alienation driven by the logic of technological capital, cognitive reconstruction in the data age, and the loss of human values amid the technological wave. Nevertheless, artificial intelligence also contains the potential to strengthen human subjectivity. This paper explores feasible approaches to the return of human subjectivity from three dimensions: individual self-awareness, technological ethics and social education. It aims to call for adherence to the human subject status while embracing the dividends of technology, so as to realize sound symbiosis between humans and artificial intelligence.

Keywords

Artificial Intelligence, Subjectivity, Technological Dependence

Copyright © 2026 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

技术进步在解放人类体力与脑力、大幅提升社会生产效率的同时，也对“人作为主体”这一地位产生了前所未有的冲击。个体对技术的依赖程度持续加深，如今，思维惰化、实践弱化、情感钝化等倾向日益凸显，人的主体性消解风险不断加剧，这已成为智能时代亟待回应的重要理论与现实命题。现有研究普遍认同人工智能对人的主体性构成多重挑战，这些研究多从技术伦理和认知科学的角度进行探讨，也有学者尝试用马克思主义异化理论来构建分析框架。尽管相关研究已取得丰硕成果，但仍存在可深化空间：一是对困境的内生成机制剖析不够深入，多停留在技术表现层面，未能充分结合马克思主义主体理论进行系统性阐释；二是对现实困境的把握不够全面，对认知、实践、情感三重维度的内在逻辑关联挖掘不足；三是提出的应对路径偏于宏观，缺乏分层落地的体系化举措。基于此，本文立足马克思主义理论视域，系统梳理主体性困境的具体表现，揭示其深层生成逻辑，并从个体、技术伦理、社会教育层面构建复归路径体系，以期在拥抱技术红利的同时守护人的主体地位，推动实现人机良性共生。

2. 人工智能时代人的主体性困境

人工智能在重塑生产生活范式的同时，对“人作为主体”的地位产生了前所未有的冲击，引发了人的主体性困境，突出表现在人的认知能力退化、实践能力减弱以及情感能力钝化。

(一) 认知能力的退化

人工智能在人类认知领域造成的冲击尤为显著。人工智能(AI)的精准与高效正在使人类大脑的某些基础功能疏于运用而日渐退化。

首先，记忆与知识内化能力显著弱化。人类在与自然互动的过程中，是通过认识、利用和改造自然规律，逐步深化对自然的认知与理解，这一过程正是人的主体性和创造性的彰显。而如今搜索引擎和大型语言模型提供的“即问即答”的服务使充满活力的现实世界变成了冰冷的数据，人得到的只是被机器处理过的信息罢了[1]。这种获取方式虽然高效，却剥夺了人在知识建构过程中至关重要的深度思考和记忆巩固环节。海德格尔在《技术的追问》中提出，现代技术的本质并非工具性的手段，而是一种“座架”(Gestell)，它“促逼”着人将世界和人自身都解蔽为“持存物”(Bestand)——即可随时调用、随时优化的资源[2]。当知识被转化为可即时调取的数据，人与世界之间原本那种需要时间沉浸、需要亲身经历的理

解方式便被取代了。记忆不是简单的信息储存，而是人在时间中与世界建立意义关联的过程。

其次，人们的批判性思维与独立判断力也面临着退化的危险。如马尔库塞所揭示的，当个人将强加于己的存在视为自己的发展和满足时，“异化了的主体被其异化了的存在所吞没”[3]。算法系统基于用户的偏好构建出了信息茧房，这使我们不断接触并强化已有观点的单一内容，难以发现多元、异质的信息。信息茧房的产生导致个体所接触的视野日益狭隘，从而萎缩了其批判性思维与独立判断能力[4]。在这一情况下，面对 AI 提供的现成答案或解决方案，人们更倾向于直接采纳而非质疑和验证。这种思维上的惰性使得我们逐渐丧失了对信息真伪、方案优劣进行独立审视和批判分析的动力与能力。

(二) 实践能力的减弱

马克思曾指出，“从前的一切唯物主义(包括费尔巴哈的唯物主义)的主要缺点是：对对象、现实、感性，只是从客体的或者直观的形式去理解，而不是把它们当作感性的人的活动，当作实践去理解。”([5], p. 499)实践是人类能动地改造世界的社会性物质活动，更是主体保证其自身本质、实现自我发展的根本路径。

当下人工智能重塑了脑力劳动模式，带来诸多便利，但高度智能化工具包办各类实操、思考环节，持续消解人的实践参与过程。当 AI 在大量实践场景中得到应用，人的实践主体性可能因“一键代劳”的便利而面临被弱化的风险。生活中一个典型例证就是空间导航能力的下降，车载导航让车主只需被动跟随系统指示几乎不用动脑思考路线。这导致了人们对导航的依赖度不断增加，而主体的空间感知和路径规划等实践能力减弱。当人们过度依赖智能导航时，一旦遭遇卫星信号微弱或地图更新不及时而致使导航出现失误的状况，驾驶员可能会因无法适应而陷入混乱，甚至可能引发交通事故。这生动说明了过度依赖如何削弱了我们在真实环境中应对挑战的本能。同样，在工业、农业乃至家居生活中，人工智能系统代替了人的直接操作。工人不再需要在生产线上精细调节，农民无需凭实践经验判断灌溉时间。这些 AI 的应用在提升效率的同时，也使得宝贵的实践技能和直观经验难以传递和积累。

(三) 情感能力的钝化

马克思指出，人的本质“在其现实性上，是一切社会关系的总和”([5], p. 501)。人的本质是在现实生活中通过与他人、社会的各种交往互动逐步形成和体现的，脱离了社会关系就无法形成完整的人性。如今人工智能催生了新型社会关系，当人类的感知向数据迁移时，虽然强化了社会整体流动性，却也侵蚀着传统社交环境，导致人际交往变得复杂而脆弱。

其一，表现在道德责任感的模糊化。道德判断与责任担当是人的主体性的重要维度，真实的伦理选择包含着主体对价值的权衡和责任意识。人工智能的广泛应用，将鲜活的道德困境转化为数据标签下的规则匹配与概率运算，原本需要人主动承担的道德决策过程，被简化为对算法输出的被动执行，个体的道德主体性被技术逻辑不断稀释。在互联网内容审核、智能政务服务等场景中，从业者高度依赖 AI 的分类判定与标准化流程，面对涉及权益边界、公序良俗的复杂议题时，常放弃独立伦理判断，直接以算法标准为处置依据；出现决策偏差时，又易将责任推诿给算法设定与系统规则，消解自身应尽的道德义务。这种将道德判断权让渡给技术的行为模式，使伦理责任在技术黑箱中弥散，最终形成“责任主体性真空”[6]，人的道德责任感也在对算法的路径依赖中逐步钝化。

其二，表现在情绪感知能力的钝化。虚拟助手和社交机器人能带来个性化、无压力的互动体验，像《恋与制作人》《光与夜之恋》这类高沉浸互动恋爱“纸片人”游戏，通过预设剧情和情感反馈机制，让游戏里的“恋人”能满足用户的情感需求，但同时也失去了现实中人际交往真实情感的复杂性与不确定性。人工智能虽然能够凭借强大的计算能力模拟人类情感，但这种模拟本质上是机械的、基于数据逻辑的，仅仅是对人脑的简单模仿，始终缺乏真实的情感体验。长此以往，这种单一的情感体验可能会削弱

人类对真实情感的敏感度，使人类陷入虚拟感官体验的困境当中。

3. 人工智能时代人的主体性困境产生缘由

人工智能对人的主体性造成的多重冲击，其深层缘由既需从马克思异化理论加以检视，也需引入当代技术哲学的视角。技术资本逻辑下的异化、工具理性的越界以及人类自身在技术浪潮中的价值迷失，三者形成闭环式的影响链条，共同加剧了人的主体性的消解。

(一) 技术资本逻辑下的异化新形态：从物质劳动到数字劳动

马克思主义所揭示的科技异化现象在人工智能时代以更隐蔽、更深入的形式展现，技术发展受到经济效率最大化逻辑的强烈影响，被置于追求利润的工具性体系之中。马克思在《1844年经济学哲学手稿》中揭示了异化劳动的四重规定：劳动者同劳动产品相异化、同劳动过程相异化、同人的类本质相异化，以及人与人相异化[5]。

在人工智能时代，这种异化在数字劳动中呈现出不同于物质劳动的新形态。在平台经济中，外卖骑手、网约车司机等新型就业群体看似拥有了自由自在的表象，实则被囚禁在算法的牢笼中。配送时效、接单量、用户评分等数据指标构成了无形的数据约束，劳动者的劳动时间、劳动强度被算法精准计算并压缩至极限，沦为算法系统中的被动执行者。与马克思所描述的工厂劳动不同，在工业时代，工人直接面对机器体系的物理强制，劳动异化是可以感知的外在压迫；而在数字劳动中，劳动者往往不占有自己生产的数据产品，这些数据被平台聚合、分析并转化为利润，劳动成果以不可见的方式与劳动者相对立。更重要的是，算法系统将劳动过程高度标准化，使劳动者丧失了根据经验与判断进行自主调整的空间。

技术哲学家唐·伊德的后现象学为理解这种关系提供了微观视角。他提出了“人-技术-世界”的四种关系模式，包括具身关系、诠释关系、他者关系和背景关系[7]。在智能系统中，算法越来越以一种“他者关系”出现，成为与人互动的准主体，而人则在这种关系中逐渐习惯于听从技术的指令。伊德指出，技术并非中立的工具，它本身具有“技术意向性”。当劳动者通过应用程序界面感知工作世界时，算法已经预先框定了可选项和优先级，人的意向性被技术的意向性所取代。

马尔库塞在《单向度的人》中进一步指出，现代社会的异化不再仅仅表现为物质匮乏和强制劳动带来的痛苦，而是演变成为一种在技术舒适中隐藏的精神禁锢[8]。资本借助人工智能构建了“技术舒适区”，便捷的智能服务、精准的消费推送、个性化的虚拟社交，让人们在物质与精神的双重满足中丧失了批判意识与创新能力。当智能助手替代了信息检索，大数据推荐承包了审美选择。这种在“舒适异化”中形成的精神禁锢，比马克思时代的物质匮乏型异化更难突破，因为它以满足需求的表象掩盖了剥夺主体性的本质，使人们主动认同并接受这种异化状态。

(二) 数据时代的认知重构与工具理性的越界

尤瓦尔·赫拉利(Yuval Noah Harari)在《未来简史》中指出：“数据的流动量已经大到非人所能处理的地步，人类无法再将数据转化为信息，更不用说转化成知识或智能。于是，处理数据的工作应该交给能力远超人类大脑的电子算法。”[9]当今信息的洪流正以指数级速度增长，个体在数据迷宫中愈发失去方向感。

首先，数据资源的集中化导致个体沦为“数据生产者”而非“数据的主人”。人工智能的运行依赖海量数据的喂养，而这些数据的核心来源正是人类的日常行为。大型科技公司通过垄断数据资源，构建起强大的数据解释权，而个体却对自身数据的流向、用途缺乏知情权与控制权。当个体的思想、行为甚至情感都被量化为数据标签，人的主体性便被简单转化为“数据画像”，人不再是完整的、能动的主体，而是成为供算法分析的“数据原料”，这种数据异化直接侵蚀了人的主体性根基。

其次，算法系统会导致认知固化，进而压缩个体自主判断的空间。哈贝马斯在《交往行为理论》中区分了两种理性类型：工具理性以效率最大化为准则，适用于处理人与自然的关系；交往理性以相互理解为旨归，适用于人与人之间的沟通与协商。现代社会的一个核心问题，是工具理性从经济和技术领域溢出，侵入本应由交往理性调节的生活世界——哈贝马斯称之为“系统对生活世界的殖民”[10]。在人工智能时代，算法推荐系统正是这种“殖民”的典型载体。搜索引擎的排序不由知识的价值决定，而由商业优化逻辑决定；购物平台的推荐不由使用者的真实需要决定，而由消费预测算法决定。当算法替代了个体的自主筛选与判断，人们就会逐渐丧失对多元信息的接触能力和对复杂问题的思辨能力，认知活动从“主动建构”退化为“被动接收”，认知主权的让渡直接导致主体性的弱化。

再者，信息过载引发的认知负荷，进一步催生了个体的认知惰性。心理学研究发现，人的认知资源是有限的。当面对海量超出自身处理能力的信息时，人们很容易出现焦虑、疲惫等心理状况。在这种情况下，人们往往会选择最省力的认知方式，也就是依赖算法给出的现成答案。这种“认知惰性”并非个体的主观缺陷，而是数据时代的普遍困境。当人工智能能在瞬间提供人们需要花费数小时筛选信息才能获得的结论，在权衡付出与收获的考量下人们主动放弃自主思考。长此以往，个体的深度思考能力、批判性思维逐渐退化，形成“越依赖算法，越缺乏自主认知能力；越缺乏自主认知能力，越依赖算法”的恶性循环，最终导致认知层面的主体性丧失。

(三) 人类自身在技术浪潮中的价值迷失

技术对人的主体性产生侵蚀，而这种侵蚀最终要得到人的主观认可才会真正实现。在人工智能时代，一些个体由于过度崇拜技术，导致价值观出现偏差，陷入了“效率至上”的技术误区，这便是主体性被削弱的内在原因。

首先，当下社会的功利主义倾向与即时满足心理，驱使个体为换取技术红利主动让渡自身主体性。现代社会生活节奏快、竞争激烈，人们对效率和便捷有着极高追求，而人工智能恰好能满足这种需求。在“时间就是金钱”的功利主义价值观主导下，人们倾向于以最低的成本获取最大的收益，而深度思考、实践探索、情感交流等需要花费时间与精力的活动则被视为低效行为。

其次，社会评价体系因技术化发展出现异化，致使人的自我认同过度依赖技术反馈。在数字时代，点赞数、播放量、评分高低、粉丝数量等量化指标，慢慢取代了传统的道德评价和能力认可，成为衡量个人价值的主要标准。人们的行为不再以实现自我目标为追求，而是以获得系统的正向反馈为导向，为了获得更多点赞而刻意迎合算法偏好去创作内容，为了提高评分而放弃自身原则满足用户需求，为了积累流量而不断生产同质化内容。这种价值评价的技术化，使个体的自我认同完全依附于技术系统的反馈，人不再追问“我是谁”“我想要什么”，而是纠结于“算法如何评价我”“数据如何定义我”。

此外，人类对“确定性”的本能追求，也促使其向人工智能让渡决策权。现代社会的复杂性与不确定性，使人们面临着前所未有的选择焦虑，而人工智能提供的“精准预测”“最优方案”恰好满足了人们对确定性的需求。海德格尔曾警告，座架的真正危险不在于机器本身，而在于人可能丧失“更为原初的解蔽方式”，即那种在诗、艺术和沉思中与世界相遇的方式。当医疗诊断、投资理财、职业选择等重大决策都依赖人工智能的建议时，那么个体就丧失了承担决策风险的勇气与能力，将自身的命运交给算法掌控。真正的主体不仅能享受技术带来的便利，更能承担决策的责任、面对未知的挑战，而不是在不确定性面前退缩，将主体地位让渡给技术。

4. 人工智能增强主体性的潜能及其审视

如果仅仅聚焦于困境与异化，容易陷入单向度的技术悲观主义，仿佛技术的发展必然以人的主体性丧失为代价。这种叙事忽视了人-技术关系中同样存在的积极可能。海德格尔在《技术的追问》中引用

了荷尔德林的诗句：“但哪里有危险，哪里也生出拯救。”([2], p. 37)。正视人工智能增强主体性的潜能，是为了更全面地理解人 - 技术关系的复杂性，使应对路径的构建不至于陷入“拒绝技术”的简单化立场。

(一) 认知维度的增强：从思维外包到思维升级

大语言模型在信息检索、模式识别和大规模文本处理方面具有人类无法比拟的速度和广度。一个人穷尽数月才能读完的文献量，模型可以在数秒内完成初步梳理。这种能力如果被恰当使用，其效果不是取代人的思考，而是将人从烦琐的信息搜集和初步筛选中解放出来，使人能够将更多精力投入到假设生成、理论建构和深度反思这些更具创造性的认知活动中。认知科学家安迪·克拉克(Andy Clark)和大卫·查尔默斯(David Chalmers)提出的“延展心智”(Extended Mind)理论认为，人类的认知从来就不是封闭在颅骨之内的孤立过程，而是不断延伸到外部工具和环境之中的[11]。

但技术对主体性最终是增强还是消解，并非由技术本身决定，而是取决于人与 AI 的互动模式。如果研究者在与 AI 的互动中始终保持问题意识，那么 AI 就只是在替代认知过程中那些机械的、不需要创造力的环节，人的判断力、批判意识和整合能力不仅不会被削弱，反而会因精力释放得到强化。2023 年，哈佛商学院的一项实验研究发现，使用 AI 辅助工作的咨询顾问在分析性任务上的表现显著优于未使用 AI 的对照组。研究同时发现，AI 并非在所有任务上都有效，当任务需要更复杂的判断或 AI 的训练数据中缺乏相关信息时，依赖 AI 反而会降低表现[12]。这说明，人与 AI 的关系不是简单的替代或增强，而是一个需要被持续调试和反思的动态过程。使用者的判断力和批判意识，正是在这个过程中决定 AI 走向“增强”还是“替代”的枢纽。

(二) 实践维度的辅助：具身技能的新可能

在实践应用中，人工智能更多是作为技能补充，而非完全替代人力。以外科手术机器人为例。达芬奇手术机器人系统能够将外科医生的手部动作转化为更精细、更稳定的器械操作，同时提供高清三维视野。在这一人机协作过程中，医生的决策主体地位并没有被取消，但借助机器人系统，医生能够完成此前难以实现的高难度手术，其专业能力的边界实际上被扩展了[13]。在这一情境下，医生的“手感”和临床判断力依然是手术的核心，技术只是让这些能力的表达获得了更大的精度和范围。

制造业也出现了类似的变化趋势。如今，在引入人机协作机器人的生产线上，工人得以从传统流水线上的重复体力劳动中解放出来，转而负责监控、诊断和优化生产流程。这种转变对工人提出了更高的要求：需要理解生产过程的全貌，需要在异常情况发生时做出判断和调整。从马克思的视角来看，当工人不再只是“机器的附件”，而是重新获得了对劳动过程的全局性理解和控制时，劳动所蕴含的主体性和创造性因素就在更高层次上得到了复归。当然，能否实现这种转变，取决于生产组织方式的设计者是否真正为工人留出了判断和决策的制度性空间。技术本身并不必然带来解放，但它提供了可能性。

另一个有力的例证来自残疾人辅助技术领域。智能导航虽可能削弱普通人的空间记忆，但对于视障人群而言，结合 AI 环境感知的辅助设备显著提升了独立行动能力。这种辅助不是替代了一个原本完好的能力，而是补足了一种因身体条件而受限的能力，实质上扩展了使用者主体性实践的范围。它表明，当技术补偿了原本存在的功能性缺失时，依赖技术恰恰意味着更大程度的自主。

(三) 情感维度的支持：技术化情感的边界与可能

对于那些因各种原因面临真实社交困难的群体，技术提供的情感支持具有不容忽视的积极意义。以独居老人为例，日本和欧洲的一些养老机构已经引入了能够进行基础对话、播放音乐、提醒用药的社交机器人。研究表明，这些机器人虽然无法替代真实的人际交往，但能够在一定程度上减轻老人的孤独感[14]。对于社交焦虑障碍者，基于 AI 的对话练习工具提供了一个低风险的训练环境，帮助他们逐步重建社交信心。这些应用场景中，AI 并不是在取代人的关系，而是在为那些暂时难以获得充分人际连接的人搭建一个过渡性的支撑结构。

在心理健康领域，AI 辅助工具正逐渐展现出其重要价值。借助认知行为疗法，AI 聊天机器人可以随时为用户提供情绪疏导服务，帮助用户在情绪出现波动时及时获得干预和调节。它们虽然不能替代专业的心理咨询师，但作为专业服务的补充和前端入口，这些工具显著降低了心理健康服务的门槛，使更多人能够在需要时获得及时的帮助[15]。

5. 人的主体性复归的路径探讨

人工智能对人的主体性的双重效应表明，技术发展离不开人文价值的规约与引导。推动人工智能健康发展，须以以人为本、智能向善为核心原则，立足主体性守护与发展的双重目标，从技术伦理、社会教育与个体实践三个维度构建系统性应对路径，确保技术进步始终锚定人的发展需求，构建人机协同共生的良性格局。

（一）技术与伦理层面：守护主体能力的底线

2025 年 8 月，《国务院关于深入实施“人工智能+”行动的意见》明确提出：“推动模型算法、数据资源、基础设施、应用系统等安全能力建设，防范模型的黑箱、幻觉、算法歧视等带来的风险，加强前瞻评估和监测处置，推动人工智能应用合规、透明、可信赖。” [16]

第一，建立强制性的“算法审计”制度。算法审计是由独立第三方对算法系统的公平性、透明性、可解释性及社会影响进行系统性评估，并将结果向公众披露。从哈贝马斯的视角来看，算法审计的核心功能在于：将封闭的算法逻辑重新拉回公共商谈的程序之中。当算法用于内容推荐、信用评估等公共领域时，工具理性已侵入了本应由交往理性主导的领域。算法审计要求算法的设计者必须以交往理性可理解的语言，向利益相关者说明：模型的优化目标反映了谁的偏好？被算法排除的选项是什么？受影响者拥有什么救济权利？[17]这一过程就是把工具理性的技术系统重新置于交往理性的程序约束之下。

第二，探索“数据信托”模式。数据信托是由法律授权的独立受托机构，按照信托章程代表数据主体管理和处置数据，确保数据使用符合主体的利益和社会的公平准则[18]。这一机制试图改变当前个体在数据关系中的结构性弱势，通过集体授权和专业代理，使数据主体获得对数据使用的实质性发言权，从而在数据关系中重新锚定主体性。

第三，必须确立“人类最终裁决权”原则，在医疗诊断、司法裁判、军事行动等涉及重大价值的领域，最终决定权必须牢牢掌握在负责任的人类手中。《人工智能安全治理框架》2.0 版本的可信人工智能基本准则中提出人类最终控制权概念：“在人工智能系统关键环节设置人类控制机制，使最终裁决权归属人类，通过设计安全控制阈值、设置安全终止开关、预留人工干预有效窗口等措施，确保人工智能系统能够实现人类预期目标、不会脱离人类监督运行失控” [19]。

（二）社会与教育层面：提供能力提升的支撑

《国务院关于深入实施“人工智能+”行动的意见》明确提出：“要把人工智能融入教育教学全要素、全过程，创新智能学伴、智能教师等人机协同教育教学新模式，推动育人从知识传授为重向能力提升为本转变。” [16]面对人工智能的挑战，教育体系必须进行根本性改革，从知识灌输转向能力培养。

批判性思维的培养不能仅靠呼吁，而需要在反复的实践中落地。以 AI 辅助的批判性思维训练为例：让学生面对一个有争议的社会议题，先独立整理自己的立场和论据，然后对比 AI 生成的论点。训练的核心不在于使用技术，而在于追问——AI 给出的论据有没有可能是幻觉？它的信息来源可能是什么？哪些重要视角被系统性地遗漏了？这一过程将 AI 定位为被审察的“对手”而非服从的“权威”，训练的是信息核实、视角转换和对自己判断过程的反思。这种训练直接回应了认知层面的主体性退化问题。

全社会应倡导“AI 辅助”而非“AI 替代”的风气，明确技术是增强人类能力的工具，目标是“人机协作”，实现 $1+1>2$ 的效果。人工智能是人脑力对象化的产物，这样的性质和地位就决定了它不可能

真正地取代人的主体地位。人工智能将来的每一点进步，都是对人本质力量的再一次确证。唯有保持人工智能与人类智能的辩证张力，实现人机融合、人机协同、人机共生，才能避免陷入对新兴技术进行控制的“科林格里奇困境”¹。

(三) 个体层面：解决能力弱化的源头

人的主体性是一切技术都无法替代的终极价值。面对人工智能时代的挑战，真正的困境不在于技术本身，而在于我们能否保持人的清醒与自觉。马克思曾指出：“动物只是按照它所属的那个种的尺度和需要来构造，而人却懂得按照任何一个种的尺度来进行生产，并且懂得处处都把固有的尺度运用于对象。”([5], p. 163)这表明人的生产活动是自由自觉的，人不会被技术理性彻底裹挟。

首先，个体必须培养批评使用技术的自觉。海德格尔晚年在沉思中提出“泰然任之”(Gelassenheit)的态度：这不是拒绝技术，而是“对物说‘是’的同时也说‘否’”，我们使用技术，但不让技术支配我们的本质。我们要清醒地认识到人工智能是强大的工具，但人才是目的和主体，应有意识地划定技术使用的边界，避免被技术理性裹挟，重建独立批判思考能力。

其次，要主动践行“数字节食”，通过主动的“中辍”[20]行为抵抗技术的宰制。长期沉浸于数字世界会导致真实感知能力淡化、社会关系疏离，要定期从数字世界中抽离，投身于线下劳动、面对面社交和自然体验之中。通过亲身实践重建对真实世界的直接感知和互动能力，弥补少动手和少参与的短板，面对面社交、自然体验可以修复数字疏离带来的情感连接弱化，强化人与人、人与环境的真实情感共鸣。

总而言之，个体层面的主体性复归，是在思想自觉与行动节制的统一中重新确立人对技术的主导权，以人的自由自觉活动对抗技术异化，这是破解主体性困境最本源的路径。

参考文献

- [1] 张可. 人工智能技术发展对人的主体性影响研究[D]: [硕士学位论文]. 重庆: 西南政法大学, 2019.
- [2] 马丁·海德格尔. 演讲与论文集[M]. 孙周兴, 译. 北京: 生活·读书·新知三联书店, 2005: 3-37.
- [3] 赫伯特·马尔库塞. 单向度的人: 发达工业社会意识形态研究[M]. 上海: 上海译文出版社, 2006.
- [4] 王海稳, 王洋. 智能在场与主体回溯: 人工智能时代人的主体性复归的唯物史观审视[J]. 浙江社会科学, 2025(9): 97-107+158-159.
- [5] 马克思, 恩格斯. 马克思恩格斯文集(第1卷)[M]. 北京: 人民出版社, 2009.
- [6] 姚明明, 常嘉洋. 人工智能时代人的主体性转向——基于马克思主义科技观的哲学审视[J]. 浙江社会科学, 2025(7): 127-133+159.
- [7] 唐·伊德. 技术与生活世界: 从伊甸园到尘世[M]. 韩连庆, 译. 北京: 北京大学出版社, 2012.
- [8] 冯讲琴. 浅谈马尔库塞技术理性批判理论——以《单向度的人》为例[J]. 张家口职业技术学院学报, 2019, 32(2): 52-55.
- [9] 尤瓦尔·赫拉利. 未来简史: 从智人到神人[M]. 林俊宏, 译. 北京: 中信出版社, 2017: 335-336.
- [10] 哈贝马斯. 交往行为理论(第一卷)[M]. 曹卫东, 译. 上海: 上海人民出版社, 2018.
- [11] Clark, A. and Chalmers, D. (1998) The Extended Mind. *Analysis*, **58**, 7-19. <https://doi.org/10.1093/analys/58.1.7>
- [12] Dell'Acqua, F., McFowland, E., Mollick, E.R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., *et al.* (2023) Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Harvard Business School Working Paper, 2023. <https://doi.org/10.2139/ssrn.4573321>
- [13] Woo, Y., Hyun, Y.J. and Park, Y.S. (2021) Robotic Surgery: A Review of Current Status and Future Perspectives. *Journal of Robotic Surgery*, **15**, 1-11.
- [14] Pu, L., Moyle, W., Jones, C. and Todorovic, M. (2019) The Effect of Using PARO for People with Dementia: A

¹科林格里奇困境指代技术控制的两难困境，若担心技术的不良后果而过早实施控制，那么技术可能难以取得进步。反之，如若控制过晚，技术已经成为整个经济和社会结构不可或缺的一部分，就可能导致技术走向失控，此时再解决技术引发的不良后果就会更困难。

- Systematic Review and Meta-Analysis. *International Journal of Nursing Studies*, **97**, 10-17.
- [15] Fitzpatrick, K.K., Darcy, A. and Vierhile, M. (2017) Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, **4**, e19. <https://doi.org/10.2196/mental.7785>
- [16] 国务院关于深入实施“人工智能+”行动的意见[EB/OL].
https://www.gov.cn/zhengce/zhengceku/202508/content_7037862.htm, 2026-07-17.
- [17] 哈贝马斯. 在事实与规范之间: 关于法律和民主法治的商谈理论[M]. 童世骏, 译. 北京: 生活·读书·新知三联书店, 2014.
- [18] Delacroix, S. and Lawrence, N.D. (2019) Bottom-Up Data Trusts: Disturbing the “One Size Fits All” Approach to Data Governance. *International Data Privacy Law*, **9**, 236-252. <https://doi.org/10.1093/idpl/ipz014>
- [19] 《人工智能安全治理框架》2.0 版本[EB/OL].
https://www.cac.gov.cn/2025-09/15/c_1759653448369123.htm, 2025-09-15.
- [20] 吴大娟. 智能在场与主体之困: 人工智能时代人的主体性危机与破局[J]. 理论导刊, 2024(3): 74-81.