

# 以人为本的价值重构： 生成式人工智能嵌入下的媒介生态优化路径

武沛颖<sup>1</sup> 刘 茜<sup>2\*</sup>

(1. 清华大学新闻与传播学院, 北京 100084; 2. 北京师范大学新闻传播学院, 北京 100875)



**摘要:**【目的】针对生成式人工智能给人类主体地位带来的冲击, 本文试图构建一种“以人为中心”的制度性实践框架。【方法】本文梳理了人本理念的理論起源与多重意涵, 指出其需从工具理性上升到系统治理层面。但理念落实在技术实践中存在多重结构性张力, 包括平台逻辑与人本价值的冲突、责任模糊与伦理缺位的制度设计问题, 以及技术进化与人类本质的存在论冲突。【结果/结论】提出通过四方面措施解决上述问题, 包括重构以人为本的价值取向, 打造人机耦合的智能治理系统, 构建多元共治网络, 建设情境化治理体系, 将人本理念从伦理愿景转化为制度实践, 以实现人与技术的良性互动。

**关键词:** 生成式人工智能; 人类主体性; 生态优化; 人本理念; 人机耦合

**中图分类号:** G203

**文献标识码:** A

**本文编号:** 1671-0134 (2025) 09-27-06

**DOI:** 10.19483/j.cnki.11-4653/n.2025.09.004

**本文著录格式:** 武沛颖, 刘茜. 以人为本的价值重构: 生成式人工智能嵌入下的媒介生态优化路径

[J]. 中国传媒科技, 2025, 32 (9): 27-32.



人工智能  
解析助读

## 1. 问题的提出

作为媒介技术演化的重要节点, 生成式人工智能正深度嵌入信息生态之中并重塑内容生产机制。从大语言模型到图像生成工具, 再到多模态音视频生成系统, 人工智能生成内容已经从边缘进入核心并深刻影响人类表达方式, “人”在智能语境中的位置也随之发生转型: 从内容创作者变为提示词提供者, 从前端的表达主体到后端的输出监督者。人的角色正在被技术系统与平台结构不断重构, 这并非简单的分工变化, 而是关乎权力协商、价值承载与制度边界的深层调整。

从长远来看, 智能媒介强大的自动化能力和关于计算、排序和分配的规则与价值导向可能超出人类控制, 影响社会价值观与社会秩序。<sup>[1]</sup> 面对技术给人类带来的主体性挑战, “以人为本”“人类中心”的类似表述成为模型设计、政策文件与伦理倡议中的高频词汇。但在实践层面, 这一正当的价值立场却面临显著的张力与悖论: 现有的“人本”话语往往停留在用户体验或功能优化的层面, 难以触及生成机制、价值对齐与制度安排等更为根本的媒介逻辑问题。所谓“以人为中心”的构想, 在缺乏明确的责任链条、生成透

明性与价值协商机制的情况下, 极易沦为象征性的技术承诺, 甚至在某些语境中, 构成对算法逻辑主导地位的掩饰。

因此, 本文的论述重点并不在于对生成式人工智能技术工具性的讨论, 而致力于提出一个更具结构性与制度性的问题: 如何在智能技术的冲击下, 确立人类的主体性地位? 如何将“以人为中心”的构想从一种伦理性愿景转化为制度性实践框架? 这一转向, 既关涉模型设计与技术治理, 也要求对技术、社会和文化等多元要素进行系统重构。

## 2. 智能语境下的人本理念

### 2.1 “人本”概念的理论起源与多重意涵

“人本”(human-centered)并非一个在技术发展中自发生成的理念, 而是长期以来人类在面对技术系统演化过程中的回应性建构。在技术发展的早期阶段, 以人为本的理念主张在技术设计中尊重人的认知与行为决策方式, 反对技术中心主义和系统封闭性。然而, “人”不能被简化为“用户”。Gasson (2003) 提出, 真正以人为本的系统设计必须从问题定义、组织结构与社会实践层面出发, 超越以工具使用效率

\* 为通信作者

**基金项目:** 本文受中央高校基本科研业务费专项资金资助——北京师范大学青年拔尖人才培养项目, 项目编号: 1253200018。

为核心的用户中心逻辑,转向以社会角色、价值认同与协同行动为基础的人本中心逻辑。这一观点强调,设计不应仅是技术问题的解法,而应成为社会问题的建构过程。<sup>[2]</sup>基于类似批判视角,Norman(2005)提出对“以人为中心的设计”(HCD)范式的反思,转而倡导“以活动为中心的设计”(ACD)。ACD认为,技术不仅服务于人类活动,更会反过来重塑活动本身;因此,优秀的设计应优先理解活动结构,而非一味“迎合用户意愿”。这一观点揭示了技术与人的互动并非单向适配关系,而是动态协同与结构性塑造的过程。

<sup>[3]</sup>Oviatt(2006)将认知负荷理论引入解释性人工智能设计,强调通过控制信息量、增强情境关联与行动导向,实现对用户认知能力的适配,推动以认知友好为核心的本可解释AI系统的构建。<sup>[4]</sup>从价值哲学层面,Buchanan(2001)指出人本设计更应是一种服务于人类尊严与人权的伦理承诺,设计者的责任在于反思并塑造社会制度与文化实践。<sup>[5]</sup>

进入人工智能时代后,“人本”理念逐渐从操作性维度的用户友好,演化为规范性维度的人机对齐。例如,He等(2017)指出,人本智能机器人的核心任务是理解人类、与人协同、适应人类社会行为规则,需要具备以人为中心、认知能力、交互性和自适应性等个性化等特征。<sup>[6]</sup>Riedl(2019)提出人本人工智能(Human-Centered AI, HCAI)必须同时具备两个维度:一是AI理解人类,包括对社会文化语境、常识知识与行为模式的建模能力;二是帮助人类理解AI,即通过提供可解释性、透明性与可追溯性信息,增强用户对系统运行逻辑的认知与控制。<sup>[7]</sup>Wei Xu(2019)从人机交互角度提出HCAI应建立在三大支柱之上:伦理对齐设计、体现人类智慧的技术增强,以及以“人因(Human Factors)”为核心的界面与交互设计。其中,“人因设计”强调解释性、可理解性与可用性的统一,推动AI系统从“工具”转向“协作体”。<sup>[8]</sup>Ozman等(2023)主张构建HCAI框架,其核心理念是尊重人类的尊严、权利与需求,使AI系统增强而非取代人类智能。推动人本AI发展中有六大挑战:设计可控系统、确保责任性、实现可解释性、提升可审计性、保障道德性、推进以人为本的创新文化。<sup>[9]</sup>

综上所述,“人本”理念在早期主要聚焦于用户适配与认知友好,随后扩展至伦理规范、制度责任与社会价值的系统性维度。然而,当生成式人工智能技术开始参与语言生成、内容创作与价值建构,人与机器的关系也愈发复杂,传统的人本理念面临重构挑战。在开放性、交互性更强的语境中,“以人为中心”不再是简单的技术适配问题,而成为关涉制度权力、文化主体性与社会治理的新议题,亟待重塑理论基础与实践机制。

## 2.2 智能技术嵌入下的“人本”重构

目前,生成式人工智能已经从类似计算机的中介工具转变为传播主体,<sup>[10]</sup>逐步介入知识生成、内容叙事与价值表达的流程中,成为意义建构的共创者。而且,生成式人工智能作为平台化、算法驱动的智技术系统,其嵌入形式已不再是单点工具的接入,而是一种系统级的信息基础设施重组。<sup>[11]</sup>因此,人本理念的嵌入实践需要面对平台治理、伦理合规等更深层次的制度性问题,必须从工具理性范畴上升为一种系统治理与协同机制的重构理念。

首先,新型人本理念应建立在人之为人的存在性立场之上,突破当前以个体适配性为导向的工具理性逻辑。人在系统中的角色不应仅被设定为高效使用者或反馈提供者,而应被重新理解为制度的承载体、意义的生成者与伦理的判断者。真正的人本设计不应只关注用户使用是否流畅,而应回到更深层次的问题:技术为何存在、技术如何改变人的处境、技术系统是否尊重人的主体性与不可替代性。这一转向意味着,“人”不再是系统效率优化的参数,而是技术运行所服务的目的本身。

其次,人本理念必须回应当下人机能力边界日益模糊的现实张力。生成式AI在语言生成、图像创作与判断预测等方面的高度自主性,已在多个领域与人类专业性产生交叉甚至冲突。在此情境下,人本不应再被简化为人控制AI,而应被理解为一种协商性机制的设定:技术可以发展,但人的认知边界与社会优先性必须得到明确的制度性保障。这要求人本理念转型为一套具有明确分工、角色定位与伦理承载机制的人机协作规范。

最后,人本理念的真正考验在于其是否能够承担起公共维度的治理责任。生成式AI系统并非仅服务于个体用户,而已在社会层面影响公共议题建构、信息流动机制,改变社会认知结构。因此,真正的人本立场必须从个体导向跃迁为公共伦理与制度协商,确保系统嵌入公共价值判断,保障表达多元、避免信息操控、捍卫集体福祉。这也意味着,个体与技术的关系不再是私人层面的使用工具问题,而是公共空间治理与媒介生态正义的问题。

## 3. 结构张力:生成式人工智能中的人本困境

尽管“以人为中心”已成为生成式人工智能系统设计与伦理治理中的主流话语,但这一理念在平台架构、算法伦理与制度设计中频繁遭遇实践层面的深层阻碍,构成了一系列结构性张力。

### 3.1 系统目标的张力:平台逻辑与人本价值的冲突

生成式人工智能的普及似乎带来了技术乐观主义者推崇的民主化使用:无论是个人用户、基层机构还是中小型创作者,都拥有了调用大模型参与内容生产的平等权利。然而,这种表面的技术开放性掩盖了其



背后的结构性集中趋势。AIGC系统的核心能力并非来自用户自身，而是依赖于极少数科技平台所掌握的大模型基础设施。而且，这些大模型的训练依赖海量数据，这些数据本质上是将人类实践活动的数字化痕迹简化后得到的特征向量，其背后的人类情感、隐私、身份意义被剥离。平台掌握着对模型训练语料、推荐机制、风险词库与内容审核规则的全部决定权，而用户则处于信息不对称的被动地位，既无法知情，更无从介入，也缺乏有效的协商与监督渠道。这种封闭式的运行模式构成了AIGC系统的显著特征，使其演变成为一种技术与平台的权力复合体，以人为本的理念只能作为外部话语存在，难以真正嵌入其制度结构与流程逻辑之中。

在这种技术架构下，用户的表达能力被预设于特定模型框架之中，形成了一种“拟开放性”的表达空间。尽管生成式人工智能声称赋予用户前所未有的表达自由和内容创造能力，个体在使用大模型进行创作或对话时，必须适应一种被模型“听得懂”的表达方式，而非完全出自其自身语境与经验的自然表达。用户的使用行为实则被模型预设的语料边界、提示工程规则与平台审核机制所规训和形塑，用户的主体性也被压缩为一组高质量提示词的提供者。这种新的知识生产方式并非传统意义上的人类之“思”，会引发人类思维逻辑机器化，以及知识权威和知识标准混乱等潜在后果。<sup>[12]</sup>其导致的结果是：一种新的中心化结构在去中心化的外观之下悄然成型。虽然AIGC平台鼓励个体使用模型进行创作、表达、工作与思考，但实际上，用户的能动性被重构为在平台提供的语言结构中选择，而非真正的意义生成主体。这种再中心化的结构性安排强化了平台对话语的控制权，使内容生产的自由化进程反而演变成为一种平台化治理的深度扩展。平台不再只是内容的分发者，而是通过模型训练、提示词引导、反馈机制等方式，直接参与表达的生成，从而在无形中重塑了语言实践的社会基础。

### 3.2 制度设计的张力：责任模糊与伦理缺位

在传统传播活动中，创作者、编辑者与传播者三者的角色相对清晰，责任归属具有可追溯性。但在AIGC系统中，生成内容往往是由用户输入提示词、模型生成文本、平台再分发的复合性协作结果。由于模型本身并不具备道德判断能力，其输出内容一旦造成误导、歧义或伦理后果，责任分配便在用户、平台与模型之间形成灰色地带。平台通过服务协议规避风险，将生成责任转嫁至用户，而用户则因技术黑箱与生成不可控性丧失有效判断力。在此结构中，“人”逐渐失去作为道德主体的清晰位置，伦理责任转化为平台治理机制中的一组可配置参数，造成技术过程与伦理判断的脱钩。这种责任的非人格化不仅削弱了人的行动后果的意识，也动摇了社会系统中基于责任划分所

建构的伦理秩序。这种责任模糊性进一步延伸至制度规范的空白地带。现行法律体系对AIGC的规制仍停留在“事后追责”层面，缺乏对生成过程中伦理嵌入的前瞻性设计。

更深层次的原因在于，生成式人工智能的技术逻辑与伦理道德的本质属性存在冲突，伦理规范依赖情境化判断、价值权衡、责任自觉等非量化维度，而AIGC的制度设计则倾向于将伦理问题转化为可编码规则（如禁止生成的关键词库、内容审核的量化指标）。人机对齐仍存在多重阻碍：首先，价值对齐范式主要采取了表征主义和行为主义的AI方案，难以精准捕捉和编码复杂的人类价值观。<sup>[13]</sup>“人”的表征权并不属于广泛社会主体，而是属于模型训练过程中对数据进行抽象平均化、去语境化、服务于商业逻辑的、扁平化的“人”。这种机制将多样性压缩为可编码的技术参数，而非将人类纳入协商性伦理共识的生成过程。其次，大模型并不具备真正的伦理理解能力，其所谓“对齐”仅是通过统计规律与用户反馈的概率分布来近似模拟伦理判断，结果是伦理被工具化、价值被算法化。最后，当前的对齐机制排除了用户在表达偏好与价值权衡中的真实参与，用户被技术系统重新定义为数据输入源，失去了在系统层级上进行伦理协商与规则共建的主体地位。制度设计对量化规则的过度依赖，本质上是对伦理问题复杂性的简化，最终造成制度合规性与伦理正当性的割裂。

### 3.3 存在论张力：技术进化与人类本质的冲突

尽管当下人工智能的主体性仍是不完整、不充分的，但智能机器对人类主体性的反噬风险已然出现，技术逻辑的泛化使得人在认知、判断、决策等方面的自主性越来越多地被机器削弱。<sup>[14]</sup>这一技术逼近人类的态势，逐渐引发一种更为根本性的结构性张力，即技术进化与人类本质之间的存在论冲突。这不再是工具理性层面的效率优化问题，也不仅是伦理治理层面的价值协调问题，而是直接触及人何以为人、人是否仍为尺度这一终极命题。

思维的可模拟性与主体性的危机构成了当前存在理论张力的关键问题。现有的生成式人工智能技术已经具备对客观世界进行模仿、仿真和拟像的完整媒介再现体系，这意味着它已经获得了一种与人类想象力极为相近的文化势能。<sup>[15]</sup>人类长期以来以语言和意识为核心建构的主体性地位便受到质疑。技术复制“思”的能力，正在冲击笛卡尔“我思故我在”的哲学基石。随着人机融合深度推进，一旦技术系统不仅协助人思维、辅助人表达，而是直接参与人类意识与生理的耦合，人类将不可避免地进入一个后人类语境。人类自现代性以来确立的主体地位将遭遇前所未有的挑战。智能技术打通了有机物人类肉身与无机物机器的二元区隔，创造了人一机合一的新型主体以及人一机互嵌的社会系统，将人与环

境（自然物、人造物）构造为连接、转换、生成的动态整体系统，从而引发人类文明的更新。无论是社会的“拟现实化转向”，还是当前人工智能媒介实施的人类主体性领域，都引发了人类社会一系列重大问题。<sup>[16]</sup>

从这一视角看，所谓“人本”人工智能不仅是一种设计目标或治理规范，更应是一种关于人类本体边界的时代性回应。在技术不断侵蚀人类定义边界的进程中，唯有重新界定“人”的核心，不以工具使用能力为标准，而以伦理判断、情感共鸣、责任承担与意义生成成为内核，才能在后人类语境中重构人之为基础。否则，以人为本将沦为空洞修辞，掩盖人类主体在技术系统中的逐步消解。因此，存在论张力不仅是一个哲学问题，更是生成式 AI 时代人类制度构建、社会组织与技术设计必须共同回应的深层结构。人本 AIGC 的未来，归根结底取决于我们如何回答那个被 AI 重新激活的古老问题：“人”究竟意味着什么？

#### 4. 人机共治：迈向人本导向的智能生态建构

生成式人工智能技术的快速发展，一方面反映了平台化进程日益加深，平台影响逐渐超出其产业范围和技术范畴而向社会化延伸。另一方面也表现出强大的型构能力，不仅在不同的制度化情境中介入社会行动者的生产、生活实践，也将自身建构为具有一定自主权的、拥有特定规则与资源的独立制度。智能平台与其他社会制度逻辑之间的互动与张力，推动了社会 and 文化的转型。<sup>[17]</sup> 因此，必须通过全局性的制度设计进行生态改革，使技术进化与人类社会和伦理道德协同演化，以实现人与技术之间的深度共生与良性互动。为此，本文提出以下四方面的优化策略。

##### 4.1 重构价值取向：锚定以人为本的根本价值

生成式人工智能的技术进步并非纯粹中性，而是深嵌于特定权力结构、经济利益与文化语境之中的价值体系重构过程。在技术的发展过程中，我们已观察到以效率、规模和算法最优解为导向的技术逻辑不断侵蚀人类社会中原有的伦理底线和主体能动性，进而引发深层次的人本困境。因此，必须将以人为本的根本价值作为技术发展的起点与归宿，通过明确的价值重构，纠偏当前智能技术的工具理性异化趋势。

重构价值取向的首要任务是确立“人的尊严、能动性与创造性”是作为制度与技术设计的核心标准，避免以技术效率遮蔽人类发展的整体目标。一方面需要超越工具理性框架，从本体论层面回应“人为何不可被技术替代”的根本问题。尽管生成式 AI 可模仿人类语言与行为，甚至在特定任务中展现超越人类的能力，但其并不具备真正的情境理解、道德判断与价值反思能力。人类的决策并非单纯逻辑演算的结果，而是嵌入具体文化、情感、伦理与历史语境之中。在新闻、教育、司法等关键领域中，人的判断之所以不可或缺，正是在于人类能

对复杂语境进行价值权衡、伦理考量与责任承担。因此，任何将 AI 全面替代人的技术幻想，都是对人类复杂性与独特性的极端简化与规训。重构价值取向意味着要在技术系统设计中保留人类不可替代的制度接口，如人工终审机制、人类解释责任与道德裁量空间，并在设计层面尊重人之为人的基本存在。另一方面，以人为本的价值并不意味着简单地以用户为中心，而是应构建面向公共利益与社会整体福祉的价值导向。在数据驱动的商业模型中，用户的短期偏好常被等同为人本价值的代理，但这种算法偏好往往是被平台不断训练与操控的结果，缺乏对公共价值、弱势群体与长远发展的反思能力。因此，重构价值导向应在个体权利与集体利益之间寻求平衡，推动生成式 AI 回归公共性逻辑。此外，重构价值取向还意味着从宏观结构上反思生成式 AI 背后的治理逻辑与制度格局。当前 AI 发展仍以大型科技公司为核心节点，资源、算力、数据与话语权高度集中，导致人本价值在技术设计初始阶段便已被边缘化。因此，以人为本的实现也需与技术去中心化的实践相结合，推动多元主体的参与机制与价值协商机制建设。包括公民社会、学术机构、地方政府与专业社群在内的广泛主体，应在 AI 伦理标准制定、技术评估与风险反馈中占据实质性位置。

##### 4.2 技术架构优化：打造人机耦合的智能治理系统

在以人为本的智能技术治理中，传统依赖人工事后审查和外部制度干预的路径，已难以应对生成式人工智能技术在高频调用、跨平台传播与自我迭代中的风险外溢特征。人类监管不仅在速度和规模上具有天然劣势，单一的外部治理也往往滞后于技术演化节奏。在这种情境下，实现技术善治的关键不在于用更多人工去弥补技术缺陷，而致力于借助智能技术自身的计算、感知与推理能力，将人类伦理与社会价值以可计算的形式深度嵌入技术内核之中，构建一种可持续运转的自我监督、自我纠偏与自我优化闭环机制，使技术在生产的同时具备反思与自我修正的能力。

因此，未来的技术架构需要引入独立运行的监督型大模型或智能体，其功能定位不同于直接参与生产的生成式人工智能，而是作为一种“数字监察官”对大规模数据与输出结果进行快速识别与深度分析。在内容生成、加工与分发等关键节点，监督系统需开展多维度检测，精准锁定偏见放大、事实虚构、隐私泄露、语境错位、逻辑矛盾等可能侵蚀公共信任与社会稳定的风险点。该监督架构应配套多模型对抗与交叉验证机制，通过部署不同架构、不同训练语料背景模型，让它们对彼此的产出进行相互审查和结果比对。一旦在生成内容的事实层、逻辑链或情感倾向上发现显著差异或潜在风险，系统可即时标记并触发人类介入，这种机制能有效防止单一模型错误或训练数据偏差所导致的系统性失真。在此基础上，监督系统还需具备



自适应的伦理更新模块,通过持续的机器学习与专家标注动态吸收最新的法律法规、行业规范、社会舆论和国际惯例,并将其转化为模型可执行的价值权重与决策参数。这一过程确保监督系统的价值判断与社会共识保持动态同步,避免陷入规则僵化或道德滞后的困境,也为应对不同文化语境、行业特征与政策环境的差异化要求提供了制度化保障。

值得注意的是,该架构必须坚持“人类在回路中”(human-in-the-loop)的闭环优化模式。虽然监督与改进的初步执行可由智能技术承担,但在价值冲突、伦理灰区或政治敏感的情境下,最终判断权与决策权必须保留在人类手中。监督系统生成的风险提示、异常分析与改进建议应由多元人类主体(包括编辑、法律顾问、伦理委员会及跨学科专家)进行复核与价值权衡,再将经审议的规则和参数反向写入生产系统,实现技术自律与人类监管的协同共治。

#### 4.3 强化层级协同:构建多元共治网络

生成式人工智能技术的迅猛发展,已深刻嵌入国家治理、行业管理与基层实践等多层级社会结构之中,其带来的风险与挑战均呈现出跨领域、跨主体、跨尺度的复杂特性。因此,我们需要综合考虑技术进步、个人利益、社会发展,构建一个涵盖国家监管、行业规范、平台自律、专业社群与用户参与等多元主体在内的多层级协同治理网络,以期建设更加公正、公平和高效的生态环境。<sup>[18]</sup>

在国家层面应发挥顶层设计与价值引领的核心功能,推动生成式人工智能治理进入立法、执法、监督、评估一体化的规范轨道。在技术快速演化与商业实践领先的现实背景下,传统滞后的规制逻辑需被超越,取而代之的是更具前瞻性与系统性的智能治理体系。这一体系不仅要设立明确的法律底线,更应聚焦生成式人工智能特有的文本生成、内容操控与情感诱导等新型能力,对其输出结果的真实性、可追溯性与责任归属制定更具操作性的规范条款。同时,需建立起跨部门合作与协调机制,在中央与地方、政策制定与技术监管之间建立高效联动的层级对话平台,实现规范与实践之间的动态适配。

在行业与平台层面,应构建更具行业自治性与专业响应性的协同机制。平台企业作为智能技术最主要的开发者与提供者,其责任不仅限于合规性执行,更应承担起公共伦理与技术责任的第一责任人角色。行业自律是平台治理体系的核心要素之一,尤其是在缺乏统一法律框架和规范的情况下,通过行业协会、联盟等组织平台的积极合作,可以促进跨平台的数据共享、技术评估与共同标准的制定。平台还应在协同机制中增强与监管机构的沟通与合作,推动行业自律与政府监管之间的有效衔接。监管不仅仅依赖于制定法

规,更应通过动态监管机制,了解技术发展趋势、行业需求与社会反馈,制定出灵活且具有适应性的政策框架。在平台内部,也需要注重对从业者的伦理合规培训,将技术向善理念嵌入到日常技术实践之中。

公众与专业群体的有序参与亦是实现共治格局不可或缺的一环。生成式人工智能的社会影响深度取决于公众能否理解技术逻辑、识别风险边界与参与规范建构。因此,应推动AI素养教育在公民层面的制度化普及,提升用户识别AI生成内容的能力,鼓励其在出现偏差或误导性内容时主动反馈。同时,科研机构、媒体、技术社区等专业群体应构建跨界的知识共享与伦理对话平台,推动技术开发、政策制定与社会实践之间的协同学习与反思共建,真正形成多元主体、分层治理、动态协商的制度场域。

#### 4.4 推进场景落地:构建情境化治理体系

生成式人工智能的应用场景高度多元,不同领域、行业与文化语境下的风险类型、伦理冲突与价值优先级呈现出显著差异。这种差异性使得统一化的治理模式难以全面覆盖实际运行中的复杂局面,甚至可能导致制度与现实脱节,出现形式合规而实质风险依然存在的情形。因此,在以人为本的总体原则下,必须将治理思路由宏观层面的价值引导和技术架构设计,延伸至面向具体使用场景的差异化配置,建立具备动态适配能力的情境化治理体系。这一体系的核心目标在于,通过对不同应用情境的风险结构与价值诉求进行精准刻画,将普遍性原则转化为针对性强、可操作性高的制度与技术安排,从而使以人为本的理念在多样化实践中具备落地能力。

情境化治理的关键前提在于进行系统性的情境识别与映射。不同应用场景中,生成式人工智能可能面临不同的合规要求、伦理风险和社会影响路径,因此有必要在技术部署前开展针对性评估,明确该场景下的核心风险类别、优先应对顺序与可接受的风险容忍度。这种情境识别不仅涉及技术层面的特征分析,还需引入制度、文化与用户行为模式等多维度因素,以确保风险刻画具有全局性与精确性。在此基础上,应建立由领域专家、技术人员、治理机构和公众代表共同参与的情境适配协作机制,通过跨主体协作将风险评估结果转化为可执行的治理规则与技术参数。这一机制应同时具备前置性与持续性,即在系统上线前明确风险控制框架,并在运行过程中进行周期性复核与必要调整,确保治理规则能够适应环境变化与技术迭代。

多场景适配的另一核心要素在于制度的弹性与可调节性。统一且僵化的治理规则容易在具体执行中引发适配性不足的问题,而过度分散的治理又会削弱整体协调性与公平性。因此,需要在制度设计中保留一定的调整空间,使不同场景能够在不背离核心价值原

则的前提下,根据政策环境、社会舆情和技术状态进行细化与修正。这种制度弹性不仅体现在规则的更新性上,也应反映在执行策略的可调度性上,例如根据风险等级、影响范围和主体责任结构,动态调整治理力度与技术介入程度,从而实现治理成本与效果的最优平衡。此外,情境化治理的有效运行离不开多元参与协同决策机制。在多场景适配过程中,不同主体对价值权衡、风险评估和执行优先级可能存在差异,因此需要建立开放的协商平台,将多方意见纳入制度形成与技术部署过程。这不仅有助于提升治理规则的合理性与可接受性,也为风险争议的解决提供了制度化渠道。通过这种多元协同模式,可以在保障治理一致性的同时,保留场景特异性的灵活处理空间。

## 5. 结语

生成式人工智能不仅重塑了知识生产与传播方式,更深刻挑战了既有的制度边界与人文秩序。在技术逻辑加速推进的时代背景下,回应人本张力的关键,不在于简单的技术修补或规范确立,而在于重建一套以人为中心的价值引导、制度安排与实践路径。本文从目标重塑、技术优化、层级协同与情境适配四个维度提出治理策略,旨在为生成式人工智能的可持续发展提供一套以人本共生为核心的思维框架。唯有在人、技术与制度的深度协同中,确立人的主体地位与价值尊严,人工智能的发展才能真正服务于人的自由拓展与社会整体的有序演进。

## 参考文献

- [1] 方婕妤.人类中心主义与媒介物质性之间的抉择:论唯物现象学的内在矛盾[J].新闻界,2025(1):87-96.
- [2] Gasson, S. Human-centered vs. user-centered approaches to information system design. *Journal of Information Technology Theory and Application (JITTA)*, 2003, 5(2): 5.
- [3] Norman, D. A. Human-centered design considered harmful. *interactions*, 2005, 12(4): 14-19.
- [4] Oviatt, S. Human-centered design meets cognitive load theory: designing interfaces that help people think. In *Proceedings of the 14th ACM international conference on Multimedia*, 2006(10): 871-880.
- [5] Buchanan, R. Human dignity and human rights: Thoughts on the principles of human-centered design. *Design issues*, 2001, 17(3): 35-39.
- [6] He, W., Li, Z., & Chen, C. P. A survey of human-centered intelligent robots: issues and challenges. *IEEE/CAA Journal of Automatica Sinica*, 2017, 4(4): 602-609.
- [7] Riedl, M. O. Human-centered artificial intelligence and machine learning. *Human behavior and emerging technologies*, 2019, 1(1): 33-36.
- [8] Xu, W. Toward human-centered AI: a perspective from human-computer interaction. *interactions*, 2019, 26(4): 42-46.
- [9] Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., & Xu, W. Six human-centered artificial intelligence grand challenges. *International Journal of Human-Computer Interaction*, 2023, 39(3): 391-437.
- [10] 别君华.AI是传播主体吗?人机传播的兴起及超越传播本体论的限度[J].全球传媒学刊,2024,11(3):57-73.
- [11] 陈昌凤,王金鹏.人工智能成为新闻基础设施:技术动因、权力转移与信任重塑[J].新闻界,2025(4):4-13,52.
- [12] 刘海龙.生成式人工智能与知识生产[J].编辑之友,2024(3):5-13.
- [13] 夏永红.人工智能伦理治理范式:从价值对齐到价值共生[J].自然辩证法通讯,2025,47(1):1-8.
- [14] 《国际新闻界》传播学年度课题组.2024年中国的传播学研究[J].国际新闻界,2025,47(1):27-52.
- [15] 常江,王鸿坤.当人工智能拥有想象力:从后人类媒介到新文艺复兴[J].现代传播(中国传媒大学学报),2024,46(10):79-87.
- [16] 孙玮.破域:数字时代的媒介论[J].中国社会科学,2024(6):143-161,207.
- [17] 贾瑞雪.从“平台”到“平台化”:现实基础、理论视野及研究路径[J].新闻与传播研究,2024,31(10):54-70,127.
- [18] 李本乾,刘家辰,孙滢昊.生成式人工智能传播伦理的系统风险及和谐共生治理[J].中国编辑,2025(4):87-94.

**作者简介:** 武沛颖(2000—),女,河南平顶山人,清华大学新闻与传播学院博士研究生,研究方向为智能传播和计算传播;刘茜(1989—),女,湖北黄冈人,副教授,北京师范大学新闻传播学院,研究方向为智能传播、人机传播和计算传播。

(责任编辑:李净)