

生成式人工智能科普内容生产的 话语策略偏向研究

——基于算法审计的方法

于 舰¹ 蒋应颀² 杨 正²

(苏州市科技馆, 苏州 215004)¹

(苏州大学传媒学院, 苏州 215123)²

[摘要] 生成式人工智能技术的出现与快速发展极大改变了当前的科学传播生态。该技术目前不仅有效赋能了科普创作者, 使得科普内容生产变得更为快捷便利, 更以非人类行动者身份深度融入社会化协同传播体系。想要更好地实现生成式人工智能对于当前科普主体生态格局的良性介入, 关键在于系统解析其作为科学传播者在生成科普内容时的策略偏向, 以及其与人类科学传播者之间使用策略可能存在的差异。本文基于算法审计的方法, 通过对国内外代表性生成式人工智能工具生成的科普内容进行编码分析发现, 生成式人工智能在进行科普内容生产时切实采用了一系列多元化的话语策略, 且其所采用的话语策略与人类科学传播者相比存在一定差异, 整体上呈现出一种偏向于微观化且身份模糊的特点。其内部基于不同语料库大模型的话语策略也存在着相当的差异性, 可能需要进一步区别对待。这一结果或可为后续进一步完善对于生成式人工智能科普生产的规制与引导提供经验借鉴。

[关键词] 生成式人工智能 算法审计 内容生产 话语策略 身份表征

[中图分类号] N4; TP18 **[文献标识码]** A **[DOI]** 10.19293/j.cnki.1673-8357.2025.02.004

以 ChatGPT 以及 Sora 等为代表的生成式人工智能 (Generative Artificial Intelligence) 的出现被认为极大改变了当前的传播生态。这些新技术使各类传播内容的生产更加智能化、自动化, 提升了传播的效率与内容的丰富程度。多模态的生成式人工智能技术的发展更是在短时间内重构了人们对于人工智能可从事内容生产的认知, 从而进一步改变了

我们理解、识别和模拟现实世界的图景^[1]。这样的一种革命也逐渐从通用型领域触及科学传播领域。大量综述性研究都指出, 生成式人工智能作为新信息技术之一将深刻赋能科学传播的各个环节, 并引发科技传播模式的深刻变革^[2-3]。在未来更为多元化的科学传播主体生态中, 生成式人工智能将成为核心的行动者^[4]。然而, 当前对这一趋势的研究多

收稿日期: 2025-03-11

基金项目: 江苏省社科基金重点项目“江苏省习近平新时代中国特色社会主义思想研究中心项目”(24ZXZA023); 国家社科基金重大项目“‘欺骗性人工智能’的信息传播、社会影响与风险治理研究”(24&ZD184)。

作者简介: 于舰, 苏州市科技馆副馆长, 主要研究方向: 数字科普, E-mail: 122741480@qq.com。杨正为通讯作者, E-mail: yangzheng@suda.edu.cn。

停留在思辨性综述层面，关于生成式人工智能作为“科学传播工具”或“科学传播中介”的赋能机制，以及其在赋能过程中相较于传统人类科学传播者所呈现的差异化行动策略与价值取向等问题，仍缺乏必要的实证研究支撑。因此，本文将基于算法审计的方法，通过严谨的实证分析路径，以国内外核心生成式人工智能技术（ChatGPT 及文心一言）为研究对象，探讨生成式人工智能作为科学传播者在生成科普内容时的策略偏向，及其与人类科学传播者在科学传播策略的使用上可能存在的差异。

1 理论框架与文献回顾

1.1 科学传播主体生态多元化

既有研究已经指出，在数字化时代，科学传播的主体生态已经从传统的“科学家生产—公众消费”的线性传播关系转化为多样化科学传播者协同参与的新型生态格局^[5]。大量非科学家行动者，如公民科学传播者（citizen science communicators）、公益组织、政府、企业等，开始主动扮演科学传播者的角色，并进一步丰富当前的科学传播主体生态^[5-6]。这同样也是当前国家科普工作中树立“大科普”理念、构建社会化协同科普格局的必然趋势^[7]。

在日趋多元化的科学传播主体生态中，除了上述的人类行动者之外，非人类行动者也逐渐受到了学界与业界的广泛关注。其中，生成式人工智能被发现能够发挥超越个体的科学传播能力，进一步丰富当前的多元化科学传播主体生态格局^[8]。研究指出，生成式人工智能能够有效赋能当前的科普内容生产，在共创、融媒、转译以及内容细分等方面提高科普内容生产效率^[2]。现实中，大量用户也逐渐开始使用生成式人工智能工具进行科学内容搜索，完成更具主动性的科学传播参与；除此之外，大量个体科学传播者也开始利用生成式人工

智能工具辅助自身进行科普内容创作，甚至在抖音等平台上开设自己的科普分身。这些现象都表明，生成式人工智能作为特定行动者可以有效介入当前主体多元化的科学传播生态格局中，扮演科学传播者的角色，并为社会化协同“大科普”格局的实现提供助力。

1.2 人工智能赋能科学传播行动者

最初，在生成式人工智能介入科学传播生态过程中最被期待的功能在于辅助传统科普生产者并拓展多元科普协同格局。其通过降低科普参与门槛，使很多非专业群体能够更便捷地获取、传播科学知识，并基于技术赋能强化科学传播者身份认同。在传统数字媒体环境下，成为公民科学传播者需要较高的媒介素养与科学素质，以及大量的时间与精力成本，这成为多元主体参与科学传播的主要障碍。而以 ChatGPT 为代表的生成式人工智能技术则大大降低了这种门槛要求，即变相降低了公众及其他社会行动者成为科学传播者的能力障碍，从而进一步扩大了科学传播者多元化的可能性^[9]。

无论是主动扮演科学传播者角色，还是协助或赋能人类科普内容生产者，生成式人工智能在介入科学传播过程中均表现出了明显的主体性、拟人性、仿真性和多模态等特征。这些特征使得其在介入科学传播的过程中具有“客观”“专业”与“真实”的错觉。当前大量研究指出，生成式人工智能在进行科学知识等专业内容生产时仍旧存在“知识幻觉”的问题（AI Hallucinations）^[10]。最新的调查结果也显示，在面对生成式人工智能扮演的科学传播者时，公众并没有表现出强烈的信任感^[11]。这些也都挑战着生成式人工智能介入科学传播主体生态的实际效果。因此，相较于继续论述生成式人工智能扮演科学传播者的可行性，当前更为重要的是将其视为独立的科学传播者，并进一步采用实证的

视角仔细检视其在科学传播中的行为偏向与内容生产特点。只有如此，才能既发挥其行动者优势，又进一步规避其内容失序风险。

1.3 人工智能内容生产偏向

面对各类替代性科学传播者以及科学传播主体生态多元化现象，已有研究进一步表明，不同身份的科学传播者在进行科学传播实践时会采用不同的传播策略以获取受众的信任，而这些不同的传播策略也会因为传播主体身份的差异使受众产生不同的信任效果^[12-13]。例如，已有研究发现，科学家作为科学传播者更倾向于采用一系列专业性话语策略，例如专家身份表征、术语使用、逻辑推演等；而公民科学传播者则更倾向于采用情绪唤起、身份相似、道德评判等话语策略。因此虽然同为科学传播者，但二者更加适合的科学传播议题环境也有所不同^[13]。所以，考察不同科学传播者所采用的话语策略偏向，并基于此厘定其所适合的科学传播情境，是建构多元主体协同参与“大科普”格局的重要方向。

大量对于人工智能进行算法审计的研究也发现，在进行内容生产时，生成式人工智能并不是绝对中立或客观的，其具有相当程度上的偏向性，乃至刻板印象或偏见^[14-15]。同时，鉴于所基于的语料库差异，不同生成式人工智能大模型，尤其是跨语言文化语境下的大模型，在内容生成上所呈现的文本偏向性差异更为明显。这也就使得在思考生成式人工智能作为传播者时，首先需要审查其具体大模型的话语偏向，从而进一步明确生成式人工智能在传播格局中的行动者定位与价值。

回归科学传播，以及生成式人工智能作为科学传播者的视角，已有研究指出，由于人工智能具有的拟人性、主体性、仿真性等特征，受众在与生成式人工智能交互时往往将其视为特定传播主体进行认知建构^[16]。同时，自动化处理能力与海量语料生成特性，

更赋予其类似专业权威的话语外显形态^[17]，这构成了当前将其纳入科学传播者范畴的前提基础。但已有研究同样指出，由于其生成机制缺乏必要的透明度（transparency）与可解释性（explainability），以及在模拟对话过程中对于可信度线索（credibility cues）的隐去，用户在实际交互中仍能清晰辨识其与人类专家的本质差异^[17]。而且，大量强调拟人化与具身实感的生成式人工智能大模型还进一步放大对于非专家话语策略的采用，以期建构“亲密朋友”的对话者身份^[18]。这些技术特征与传播策略的复杂交织，形塑了生成式人工智能在科学传播场域中复杂且独特的身份张力。因此，亟待系统考察不同语言模型在科学传播过程中呈现的话语倾向，判别其是更倾向于采用类似科学家等专家的话语策略，还是类似于公民科学传播者等替代性科学传播者的常民话语策略。这些研究结果将有助于我们明晰生成式人工智能在整个科学传播主体生态中的角色定位与行动可能。基于此，本文提出以下核心研究问题。

研究问题 1：在进行科普内容生产时，生成式人工智能更倾向于采用何种话语策略？

研究问题 2：在进行科普内容生产时，不同的生成式人工智能所采用的话语策略是否存在差异？

研究问题 3：在进行科普内容生产时，生成式人工智能话语策略与人类话语策略是否存在差异？

2 研究方法

2.1 算法审计

本文核心采用了算法审计（Algorithmic Audit）的研究方法。其最早由学者克里斯丁·桑维（Christian Sandvig）等人提出，是一种针对算法系统和内容生产的特殊审计方法体系，旨在检测和评估算法的潜在影响与风险，核

心目的是揭示算法及其所影响的内容生产中可能存在的偏见^[19-20]，目前也被大量应用在生成式人工智能内容生产的偏向性研究中，主要流程包括提示词设置、样本生成、内容审计分析等步骤^[14-15]。

2.2 提示词设置与样本生成

为了有效研究不同语料库基础下的生成式人工智能科普内容生成话语策略差异，本研究以 OpenAI 公司开发的 ChatGPT4.0 版本以及百度公司开发的文心一言 4.0 版本为研究对象。截至研究开展之时（2024 年 12 月），文心一言大模型日均调用量达 16.5 亿次，ChatGPT 周活跃人数超 3 亿，分别为当时中、英文领域中使用最为广泛、调用次数最多的生成式人工智能工具。由于这两个生成式人工智能工具在中英文语境下应用广泛，其在生成式人工智能科普内容生产上具有相当程度的代表性。对于提示词，本研究选择了 *Science* 在其创刊 125 周年之际，公布的 125 个具有挑战性的科学问题作为生成式人工智能科普内容生产的方向。这 125 个科学问题涉及了数学（3 个）、化学（9 个）、医学与健康（11 个）、生物学（22 个）、天文学（23 个）、物理学（18 个）、信息科学（4 个）、材料科学（4 个）、神经科学（12 个）、生态学（8 个）、能源科学（3 个）和人工智能（8 个）多个科学领域，且被认为具有积极广泛的影响。例如，将“什么是重力？”“什么是量

子不确定性，为什么它很重要？”“哪些基因使我们人类与众不同？”等问题以中文版本提交给两个生成式人工智能工具。为了控制在提示词层面上可能存在的差异，在样本生成的过程中，本研究直接采用了“请帮我生成关于……（原始问题）的科普内容”作为提示内容，从而保证两个生成式人工智能工具均就问题本身生成科学回答内容。基于此，本研究最终共得到两个生成式人工智能工具生成的 250 篇科普内容，共计 143 966 字，篇均字数为 576 字。

2.3 审计编码表制定与信度检验

为了有效评估生成式人工智能工具所生成的科普内容的策略使用及偏向，研究基于简·古德温（Jean Goodwin）等所提出的信任生成的 11 种话语策略^[21]，以及杨正等所提出的科学传播中的身份策略（专家身份、教育经验、身份相似、个人声誉）、内容策略（专业知识、个人经验、证据引用、语气强调、逻辑推演、情绪唤起、妥协让步、对话邀请、交互信任）以及行动策略（自我效能、道德评价）等^[13]，并结合林恩·G. 祖克尔（Lynne G. Zucker）等提出的信任三维机制理论^[22]，制定了如下的审计编码表（见表 1）。对于科普内容生成策略的审计评估将基于情感策略、内容策略以及框架策略展开，并将其进一步划分为 5 项一级指标与 12 项二级指标；其中对信息框架指标（目标框架、文化框架与叙

表 1 科普内容策略使用审计编码表

信任策略	一级指标	二级指标	指标内涵
情感策略 (基于经验的)	身份策略	专业身份	是否提到科学家、科学界等身份
		亲密身份	是否将自己作为受众中的一份子
	说服策略	修辞手法	是否使用举例子、排比等修辞
		说服诱导	是否有劝诫（不要）或诱导（要）
内容策略 (基于特征的)	内容特征	情感表达	是否有情感表达，例如呼吁
		术语使用	是否使用专业术语
		权威引用	是否引用权威数据、有证据支撑等
		方法使用	是否强调科学方法的重要性
框架策略 (基于制度的)	信息框架	目标框架（收益或损失）	描述积极或消极效果
		文化框架（个人或集体）	描述个人或集体受益或受损
		叙事框架（叙事或非叙事）	包含或不包含叙事性信息
	逻辑框架	逻辑应用	是否使用清晰的逻辑链条形成论述

事框架) 分别进行二元分类编码, 所以共计 15 项编码内容。本文第二作者与第三作者首先从 250 篇生成式人工智能生成的科普内容中选择了 30 篇进行预编码及编码员间信度检验, 所有维度的 Cronbach's Alpha 数值均大于 0.8。编码工作通过一致性检验, 证明编码表操作可行。随后, 由第二作者与第三作者共同编码完成所有的剩余样本。

3 研究结果

3.1 生成式人工智能科普内容生产的话语策略总体特征: 多元化的策略使用及杂糅的身份策略偏向

根据表 2 的统计结果可以发现, 无论是 ChatGPT 还是文心一言, 两者在科普内容生产时均采用了多元化的话语策略 ($n=451$; $n=691$)。这进一步佐证了大量研究所发现的生成式人工智能在内容生成时可能存在话语偏向的结论^[14-15]。从编码结果可以发现, 整体上, 生成式人工智能在科普内容生成时较为偏向使用“术语使用” ($n=240$)、“修辞手法” ($n=186$)、“逻辑应用” ($n=152$) 以及“叙事” ($n=142$) 等话语策略。其中“修辞手法”较为集中于举例子的方法, 例如文心一言根据问题“还有更多色彩元素可发现吗?” 生成的:

“随着科学技术的进步, 人们可能会发现

新的色彩元素或色彩现象。例如, 通过改进光谱分析技术, 我们可以更精确地测量和描述色彩; 通过开发新的发光材料或显示技术, 我们可以创造出前所未有的色彩效果。”

而强调科学方法的“方法使用” ($n=3$) 策略、强调个人或集体的“文化框架” ($n=7$) 策略以及强调损失的“目标框架” ($n=65$) 策略则较少被使用。这种话语策略上的使用偏向呈现出两个特征。首先, 对于身份、说服、内容等个体性、微观话语策略的使用要多于对于框架性等宏观话语策略的使用。这表明生成式人工智能在生成科普内容时, 更偏向于“就事论事”, 而较少将科学或科普问题上升至文化或价值层面的问题。其次, 在个体性、微观话语策略的使用上, 生成式人工智能杂糅了传统的专家话语策略如强调专业术语的使用, 以及一些被认为不适宜出现在专业科普中的常民话语策略, 如强调叙事、使用修辞, 以及强调情感唤起^[13]。这种杂糅表明, 无论是国内还是国外的生成式人工智能平台, 其在生成科普内容时可能并不完全依赖于科研论文等专业语料库, 可能也进一步融合来自非传统科学传播者创作的科普素材, 从而表征出具有多重身份属性的话语策略使用偏向。

表 2 生成式人工智能科普内容生产策略编码结果

	身份策略		说服策略			内容特征			框架策略				逻辑框架	总计		
	专业身份	亲密身份	修辞手法	说服方式	情感表达	术语使用	权威引用	方法使用	目标框架		文化框架				叙事框架	
									收益	损失	个人	集体	叙事	非叙事	逻辑应用	
ChatGPT	18	7	70	1	8	116	17	2	43	10	4	3	23	102	27	451
文心一言	24	57	116	9	70	124	28	1	12	0	0	0	119	6	125	691
组间差异 (p)	0.312	0.000	0.000	0.000	0.000	0.010	0.071	0.563	0.000	0.000	0.044	0.082	0.000	0.000	0.000	
总计	42	64	186	10	78	240	45	3	55	10	4	3	142	108	152	

3.2 ChatGPT 与文心一言的话语策略使用差异

为了回答研究问题 2, 本研究进一步比较了 ChatGPT 和文心一言两个生成式人工智能工具在科普内容生成时所使用的策略差异 (见表 2、图 1)。结果发现, 二者之间

在“情感表达”“修辞手法”“亲密身份”“术语使用”“目标框架”“叙事框架”等话语策略维度上呈现出了显著差异。总结而言, 文心一言 ($n=691$) 整体在话语策略应用频次方面要显著高于 ChatGPT ($n=451$), 即相较

基于中文语料库的文心一言，基于英文语料库的 ChatGPT 更倾向于简要地呈现出纯粹科学内容。具体而言，在话语策略的使用偏向上，文心一言表现出了更为明显的拟人化（anthropomorphization）特征，即在策略使用上更加强调与真人相似的对话感。这种特征主要体现在其对于常民化策略的使用偏向上，如更多的情感表达、修辞手法以及叙事策略的使用。例如，文心一言根据问题“为什么生命需要手性？”生成的：

“手性是指分子的镜像异构体，就像我们的左手和右手是镜像对称的一样，但不能通过旋转和平移重叠。具有手性的分子和其对应映体在结构上是镜像关系，但是它们的化学和生物活性可能完全不同。手性分子在化学和生物学上表现出截然不同的行为，这种差异为你我宝贵的生命提供了特定的化学和生物活性。”

同时，文心一言较之 ChatGPT 更偏向于使用逻辑论证的方法，在科普内容生成时强调层层递进的讲述策略。而 ChatGPT 整体上则更加去拟人化，强调内容的客观性，较少地使用可能会产生生活化对话体验的策略，而更多强调非叙事性话语策略的使用。例如，ChatGPT 根据问题“为什么生命需要手性？”生成的：

“手性是指物质的分子或结构无法与其镜像重合。在生物学中，手性在分子水平上至关重要。生命大多数复杂分子（如氨基酸、核苷酸等）都有特定的手性，这种手性使得分子能够以特定的方式与其他分子相互作用，形成复杂的生物体结构和功能。生命的正常运作依赖于这种手性的特异性相互作用，因为它确保了分子识别、酶催化反应、细胞功能等生物化学过程的精确性。”

此外，虽然框架策略整体上使用不多，且差异并不那么显著，但在宏观文化框架视

角下，较之文心一言，ChatGPT 更偏向于采用一些较为宏观的框架策略，例如使用“目标框架”与“文化框架”策略来框束科学或科普问题。这可能与长期以来西方世界存在的科学政治化偏向以及在中国特殊情境中对于科学的纯粹化认知有关^[23]。

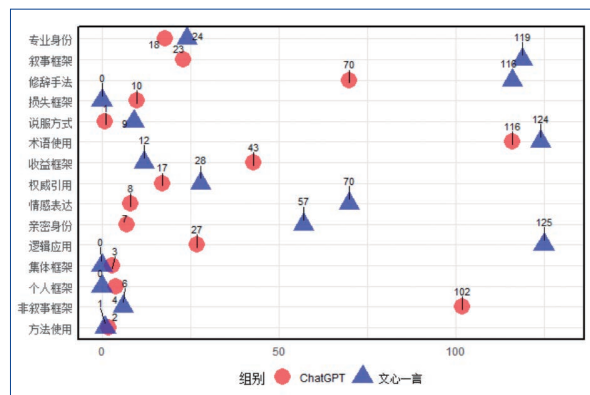


图1 ChatGPT与文心一言在科普话语策略使用上的差异统计

整体而言，在生成科普内容时，文心一言偏向于将自身塑造为一种讲究逻辑的友好朋友形象，而 ChatGPT 则更倾向于将自身塑造为一种较为客观、冷静的专家形象。这种形象差异也进一步证明了基于不同语料库的生成式人工智能工具在进行科普内容生产时会存在相当差异的话语策略偏向，从而可能进一步产生不同的科普效果。

3.3 生成式人工智能与人类科学传播者的策略使用差异

为了进一步探究生成式人工智能与不同人类科学传播者在科普内容生产时所使用的策略差异，研究进一步参考了已有研究对于科学家科学传播者（scientist science communicators）、公民科学传播者以及机构科学传播者（institution science communicators）三类不同科学传播者在科普内容生产时的策略使用偏向研究，并将其与本文的编码结果进行比照^[12-13]。根据已有研究，在不同的科普话题中，科学家科学传播者普遍更倾向于使用专业知识、专业身份、权威引用以

及逻辑推演等策略，且使用比例均高于 50%；公民科学传播者则偏向于使用情绪唤起、道德评估、说服语气以及亲密身份等话语策略；机构科学传播者则偏向于强调专业身份与权威引用^[12-13]。虽然本文所使用的编码维度与上述已有研究并不完全一致，但所得结果较为相似，具有一定的可比性。

根据杨正等人的研究，人类科学传播者尤其是科学家科学传播者与公民科学传播者在进行科普内容生产时均有着较为明显的“身份”表征偏差，即科学家科学传播者更倾向于通过一系列话语策略将自身建构为客观公正的专家形象，因此其较少使用修辞或情感策略^[12-13]；而公民科学传播者则倾向于将自身建构为“亲密朋友”形象，较多采用感性的、强调身份相似及内容叙事性的策略。而不同于这两类人类行动者较为清晰的身份策略偏向，本文的研究结果发现，生成式人工智能在生产科普内容时则表现出一种杂糅的身份策略偏向，即同时强调“专家身份”策略（ $n=42$ ）与“亲密身份”策略（ $n=64$ ），二者差异并不明显。在更为具体的话语策略维度中，生成式人工智能在采用相当数量的偏向亲密身份策略维度的修辞手法（ $n=186$ ）以及叙事策略（ $n=142$ ）的同时，也强调偏向专家身份策略的术语使用（ $n=240$ ）等。这种差异使得我们在面对生成式人工智能所生产的科普内容时，更难清晰分辨出其身份表征。这也是当前对于人工智能拟人化研究的困境所在，是要将 AI 建构为具有高度拟人化与亲密属性的伙伴，还是将其建构为较为客观公正的专家，这一问题涉及传播效果与交往伦理等多个维度，需要予以进一步研究。但整体而言，生成式人工智能在进行科普内容生产时确实表现出了与人类科学传播者不同的多元化策略偏向，这也进一步证明了其在介入当前科普社会化协同格局中特殊的行动者

角色与可能产生的效果，因此需要予以更进一步的关注与研究。

总结上述对于三个研究问题的回应，可以发现，生成式人工智能在进行科普内容生产时切实采用了多样化的话语策略，且这些策略的采用与人类科学传播者之间存在一定的差异，整体上呈现出一种偏向于微观化且身份模糊的特点。其内部基于不同语料库的生成式人工智能的话语策略也存在着显著差异，可能需要进一步区别对待。

4 讨论

本研究从三个维度拓展了生成式人工智能赋能科学传播的理论认知：其一，揭示了智能工具在传播实践中的动态身份调适机制，为优化情境化传播策略提供实证依据；其二，通过非人类行动者视角印证了科学传播生态的多元化发展趋势，推动构建人机协同的新型社会化协作框架；其三，创新性地引入算法审计方法，建立了分析智能传播内容特征的技术路径，为后续研究提供可复制的跨学科方法论范式。这些突破既深化了对生成式人工智能赋能机制的本质理解，也为规避技术风险、释放智能传播效能提供了关键理论支撑。

4.1 生成式人工智能科普内容生产策略偏向的理论解释与现实意义

生成式人工智能作为科学传播者，具有不同于人类科学传播者的话语策略偏向。具体而言，生成式人工智能更偏向于使用“术语使用”“修辞手法”“逻辑应用”以及“叙事”等话语策略。这种策略偏向不仅呈现出显著的语境差异性，更因其异于人类传播主体的特质，使生成式人工智能在当前科学传播协同生态中占据特殊地位并蕴含独特价值。

回归生成式人工智能作为科学传播行动者的观点，目前大量围绕着生成式人工智能

赋能科学传播的研究均指出其作为科学传播者介入科学传播生态体系中的“可行性”^[2-4]，但对于如何实现这一“可行性”，即推进生成式人工智能更有效地扮演科学传播者角色，现有研究仍缺乏操作性指引。对此，本文的研究进一步为这种“可行性”提供了行为偏向的经验参考。具体而言，通过上述的算法审计分析可以发现，生成式人工智能在各类话语策略的使用上表现出了相当程度的身份模糊性，既不完全依赖传统的专家策略，也不完全依赖公民科学传播者等替代性科学传播者所常用的常民情绪化策略^[13]。虽然大量计算机中介传播研究（Computer-Mediated-Communication, CMC）以及人机交互研究（Human-Machine Interaction, HMI）指出，从网页版网络媒体到搜索引擎，再到当前的各类人工智能工具，公众倾向于将这种机器型传播主体视为另类专家，并表现出相当程度的信任偏向^[24-25]。但本研究表明，生成式人工智能在进行科普内容生产时，并不完全呈现出专家的身份特征。迈克·S. 谢弗（Mike S. Schäfer）等人的研究也表明公众在将生成式人工智能视为一种科学传播者时也并不完全将其视为专家并完全信任^[11]。此外，已有研究还指出，在使用生成式人工智能进行内容生产时，专家策略的应用并不一定有效，反而是其使用的常民策略可能更容易诱发用户的信任^[26]。这一结论与之前关于不同科学传播者使用不同身份策略会产生差异化的信任效果研究结论相一致，即在特定情境中，专家身份策略的介入并不一定有效，而是要视具体的科学传播场景与使用该策略的科学传播者身份而定^[13]。已有研究指出，这可能与受众的“预期违背”心理有关，因为我们传统认知中倾向于将生成式人工智能作为传播者视为一种专家身份，所以当其实采用专家身份话语策略时，并不一定能诱发

公众的积极预期违背，反而当其使用一系列具有情感色彩的常民身份策略时，可能会进一步诱发公众的积极预期违背，从而产生更好的传播效果^[26]。所以，总体而言，不同的身份策略对于生成式人工智能科学传播可能会在不同情境中产生不同的效果。我们需要通过更为细致的实验研究来对目前策略偏向的传播效果进行更为精细的考察。此外，面对当前生成式人工智能科普内容生成时的模糊身份现状，其较之人类科学传播者是否会产生更好的效果？以及何种身份策略更能契合生成式人工智能的公众认知并发挥更好的传播作用？这些问题都需要进一步予以实证检验。这也是本研究后续需要进一步深入之处。

但本研究依旧为论证生成式人工智能作为科学传播者的“可行性”提供了侧面的证据基础。那么，面对当前生成式人工智能快速发展的时代背景，我们到底应该如何看待并对待其作为科学传播者这一现状呢？一方面，我们需要认可生成式人工智能作为科学传播者的可行性，明晰其在科普实践中的独特性，并进一步发挥其特殊的作为特定科学传播者的角色优势。这也是本文的立足点之一。另一方面，我们还需要进一步优化生成式人工智能的科普内容生产能力。对此，应从政策层面上承认并引导生成式人工智能作为科学传播者的行动可能。最新修订的《中华人民共和国科学技术普及法》（以下简称新修订《科普法》）中新增了关于“科普人员”的专章，并进一步强调“建立专业化科普工作人员队伍”“鼓励组织和个人利用新兴媒体开展多种形式的科普，扩展科普渠道和手段”。目前新修订《科普法》已经体现出引入多元主体建构科普社会化协同大格局的重要性，但对于多元主体的认知依旧停留在人类行动者层面。而面对当前技术快速革新的人工智能时代，我们需要从顶层设计上重视生

成式人工智能等非人类行动者对于科普格局的介入与作用，并对其进行规范引导以建构更为全面的科普行动者网络。

4.2 规范与引导生成式人工智能科普能力的路径

对于规范与引导生成式人工智能科普能力，首先，需要进一步构建有效的科普语料库。目前，生成式人工智能进行科普内容生产依旧依赖于传统的通用大模型。这类生成式人工智能被发现仍然存在较为突出的知识幻觉等问题。而科普不仅仅是一种知识的传递，同时还包括知识的组织方式和表达技巧^[8]。因此，更加需要一个专业的科普知识图谱和语料库，对于科学知识和概念进行系统化、专业化的表达，以供生成式人工智能后续深度学习。其次，还需要增强生成式人工智能科普大模型的算法透明度与审核力度。由于算法的黑箱属性，大模型的训练过程难以被完全掌控与解释，因此更加需要提高算法模型的透明度，以及对从后端入手的面向科普算法的审核力度。重点审查科普内容的科学性、准确性、内嵌意识形态的偏向性以及潜在的偏见与歧视等，以保证生成式人工智能介入科普生产的有效性与规范性。这需要政府、企业、平台、用户等多元力量的协同参与。一方面，政府应当完善相关领域的法律法规，并推进相关领域的行业自律；另一方面，平台也需要明确标注科普内容的生成方式，保证内容生成的透明度。同时，用户也需要不断提升自我的人工智能算法素养，培养自身的信息核查能力与对生成式人工智能“知识幻觉”识别能力。最后，虽然我们强调生成式人工智能可以在当前社会化协同科普体系中有效扮演非人类行动者的角色，但其最终的落地实现均需要人类的操作。尤其是对于科普内容生产而言，其更是依赖科普生产者的介入。因此，提高科普生产者的人工智能素养与技能，使其能够更加有效地利用生成

式人工智能进行自我赋能，也是生成式人工智能有效介入科普生产的重要路径。只有切实做到如上的要求，才能真正实现前文所言的生成式人工智能扮演科学传播者的独特行动价值。

5 结语

从自然语言处理到多模态内容生成，生成式人工智能的迭代演进已深度融入现代媒介生态体系，科学传播领域亦在此进程中迎来范式革新。利用好生成式 AI 技术革命的契机，拓展当前多元化科普主体生态体系，既是实现社会化协同科普生态的关键路径，更是提升全民科学素质、培育科技创新文化氛围的重要举措。然而达成此愿景需首先厘清生成式人工智能在科普内容生产中的双重属性：既要系统解析其作为科学传播主体的独特性——包括算法驱动的话语策略偏向及与人类传播者的认知差异，亦需基于此特质差异构建规范框架。唯有通过精准识别生成式人工智能的生成逻辑特征，并建立与之适配的引导机制，方能实现智能技术对科学传播体系的持续增益，进而重塑人机协同的科普新范式。对此，本研究基于算法审计与内容分析的实证方法，细致考察了以 ChatGPT 与文心一言为代表的生成式人工智能工具的科普内容生产策略。研究表明，生成式人工智能作为科学传播者，具有不同于人类科学传播者的话语策略偏向。具体而言，生成式人工智能更偏向于使用“术语使用”“修辞手法”“逻辑应用”以及“叙事”等话语策略，且这种话语策略在不同语境下的生成式人工智能工具中呈现出不同维度的偏向。这一结果丰富了我们对于生成式人工智能介入科学传播生态可能性的理解，并进一步为如何利用生成式人工智能推动科普工作的进一步发展提供了路径参考。

本文同样存在着一系列的不足之处。首先,在选择生成式人工智能模型时仅选取了 ChatGPT 与文心一言为国内外生成式人工智能工具代表,对于稍晚出现且流行的 DeepSeek 等生成式人工智能工具缺乏关照。其次,在样本生成过程中,本文采用的 *Science* 公布的 125 个具有挑战性的科学问题均较为偏向前沿科学领域,对于更为日常、生活化的科学问题有所忽略。再次,在分析结果中,虽然我们引用了较为权威的对于人类科学行动者科普内容话语策略使用的研究结果,但这些结果并不能保证

与本次研究中生成式人工智能科普内容生成的背景一致,从而其比较结果可能具有一定的不准确性。最后,研究虽然采用了较为严谨的算法审计的方法,并尽量控制了提示词在内容生成中可能产生的差异影响,但在具体的生成式人工智能内容生成过程中,不同用户所采用的不同提示词可能会导致生成的科普内容之间存在一定程度的差异,从而偏离本文所最终得出的研究结果。这是算法审计方法在面对生成式人工智能内容生成时所固有的方法弊端。这些问题都需要在后续研究中予以修正。

参考文献

- [1] 喻国明, 苏健威. 生成式人工智能浪潮下的传播革命与媒介生态——从 ChatGPT 到全面智能化时代的未来 [J]. 新疆师范大学学报(哲学社会科学版), 2023, 44(5): 81-90.
- [2] 王硕, 阎妍, 李正风. 生成式人工智能赋能科学普及: 技术机遇、伦理风险与应对策略 [J]. 科普研究, 2024, 19(4): 5-13, 22.
- [3] Klein-Avraham I, Greussing E, Taddicken M, et al. How to Make Sense of Generative AI as a Science Communication Researcher? A Conceptual Framework in the Context of Critical Engagement with Scientific Information[J]. Journal of Science Communication, 2024, 23(6): A05.
- [4] Schäfer M S. The Notorious GPT: Science Communication in the Age of Artificial Intelligence[J]. JCOM: Journal of Science Communication, 2023, 22(2): 1-15.
- [5] Yang Z. Deconstruction of the Discourse Authority of Scientists in Chinese Online Science Communication: Investigation of Citizen Science Communicators on Chinese Knowledge Sharing Networks[J]. Public Understanding of Science, 2021, 30(8): 993-1007.
- [6] Fähnrich B, Riedlinger M, Weitkamp E. Activists as “Alternative” Science Communicators—Exploring the Facets of Science Communication in Societal Contexts[J]. Journal of Science Communication, 2020, 19(6): C01-1.
- [7] 汤书昆, 郑斌, 余迎莹. 科普社会化协同的法治保障研究 [J]. 科普研究, 2022, 17(2): 15-20.
- [8] 周慎, 陶美丽, 刘湘. 人工智能生成科普内容的底层逻辑与参与主体功能分析 [J]. 科普研究, 2024, 19(4): 14-22.
- [9] 丁天宇, 杨正. AIGC 媒介变革环境下科学传播“多元化”趋势研究 [J]. 科学传播与科学教育, 2025(6): 1-12.
- [10] 陈万球, 罗一人. 生成式人工智能的“知识幻觉”及其风险治理探论 [J]. 上海市社会主义学院学报, 2024(4): 38-51.
- [11] Schäfer M S, Kremer B, Mede N G, et al. Trust in Science, Trust in ChatGPT? How Germans Think about Generative AI as a Source in Science Communication[J]. Journal of Science Communication, 2024, 23(9): A04.
- [12] Yang Z, Huang Y, Yang T, et al. How Different Science Communicators Use Identity Strategies to Gain Public Trust: a Study on Astronomy and Climate Change Issues on a Chinese Knowledge Sharing Platform[J]. Journal of Science Communication, 2024, 23(9): A08.
- [13] Yang Z, Yang T. Differentiated Trust Strategies and Rebellious Acceptance: A Qualitative Comparative Analysis of the Trust Strategies Used by Scientist Communicators and Citizen Science Communicators in Chinese Online Climate Communication[J]. Science Communication, 2024, 46(3): 247-275.
- [14] 黄阳坤, 师文, 陈昌凤. 智能算法如何重构新闻价值? ——基于智能推荐平台算法审计的研究 [J]. 新闻大学, 2024(6): 18-33.
- [15] 周葆华, 罗沛. 搜索引擎图片中的职业性别刻板印象——基于跨平台算法审计的研究 [J]. 新闻大学, 2024(6): 1-17.
- [16] Jiang Q, Zhang J, Wang P H, et al. Technology Acceptance and Innovation Diffusion: Are Users More Inclined toward AIGC-Assisted Design[J]. International Journal of Human-Computer Interaction, 2024: 1-15.
- [17] Dunn A G, Shih I, Ayre J, et al. What Generative AI Means for Trust in Health Communications[J]. Journal of Communication in Healthcare, 2023, 16(4): 385-388.

Research on the Discoursal Strategy Bias of Generative Artificial Intelligence Produce Science Popularization Content: Based on the Method of Algorithm Auditing

Yu Jian¹ Jiang Yingdi² Yang Zheng²

(Suzhou Science and Technology Museum, Suzhou 215004)¹

(School of Communication, Soochow University, Suzhou 215123)²

Abstract: The emergence and rapid development of generative artificial intelligence (GenAI) have greatly changed the current science popularization ecology. It not only effectively empowers current science popularization creators, making the production of science popularization content more convenient and faster, but also directly intervenes in the current socialized collaborative science popularization pattern as a non-human actor. To better realize the benign intervention of GenAI in the current ecological pattern of science popularization, it is first necessary to systematically analyze its strategy bias as a science communicator when generating science popularization content and the possible differences between its science popularization strategy use and that of human science popularization creators. Based on the method of algorithm auditing, this paper finds, through coding and analyzing the science popularization content that representative GenAI tools generated at home and abroad, that GenAI has indeed adopted a series of diversified discourse strategies when producing science communicators, and there are certain differences between the discourse strategies adopted and those of human science communicators. Overall, it presents a characteristic of being inclined to be microcosmic and having an ambiguous identity. There are also considerable differences in the discourse strategies of large models based on different corpora within them, and they may need to be treated differently. This result may provide experience and reference for further improving the regulation and guidance of science popularization by GenAI in the future.

Keywords: generative artificial intelligence; algorithm auditing; content production; discursive strategies; identity representation

CLC Numbers: N4; TP18 **Document Code:** A **DOI:** 10.19293/j.cnki.1673-8357.2025.02.004

The Study on Collaborative Optimization Strategies of Scientificity and Narrativity in Science Popularization Creation under the Context of AIGC

Zhang Yanxiang¹ Zhu Yuhui¹ Huang Rongli² Li Honglin²

(Department of Science Communication, University of Science and Technology of China, Hefei 230026)¹

(China Research Institute for Science Popularization, Beijing 100081)²

Abstract: This paper focuses on the synergistic developmental path between the scientificity and narrativity of science popularization. It systematically explains the mechanism for AIGC-driven synergistic optimization of scientificity and narrativity, analyzes the dual functions of AIGC in enhancing scientificity and optimizing communication, as well as the multidimensional issues behind its usage. Subsequently, it proposes corresponding optimization strategies for issues such as the detrimental effects of AI hallucinations on scientific rigor, the technical bottlenecks in generating highly complex scientific