

关于大语言模型和科普创作的若干问题

史 军*

(玉米实验室科普工作室, 天津 301700)

2025 年, 随着 DeepSeek 等新的大语言模型的推出, 关于其在社会各领域应用的讨论, 特别是关于大语言模型在内容创作领域的讨论不断升温。对于科普创作者来说, 最受关注的问题是: 大语言模型的应用究竟能不能替代人类作者? 在人工智能时代, 科普创作者又该何去何从?

一、大语言模型会取代科普创作者吗?

当前, 科普创作形式多样, 除了以文本为主的文章或图书外, 还有短视频、音频、漫画以及舞台剧等多种形式。本文讨论的科普创作主要是科普文章和科普图书, 因为文本是最基础的表达形式, 其创作逻辑和其他形式也有相通之处。

大语言模型的快速发展为科普创作提供了一种强力工具, 与此同时, 也对创作者提出了新的要求, 带来了新的挑战。大语言模型可以分析和总结科学信息, 并以预设的逻辑和形式(包括文字、图片和视频)进行呈现, 以达到向受众人群传递信息的目的。通过自然语言处理技术, 大语言模型可以适配不同受众的认知水平, 提供个性化的知识传播方案。

在实际工作中不难发现, 大语言模型收集整合数据的效率远远高于人工, 并且大语言模型还能发现看似不相关的“独立事件”之间的关联性, 为创作者提供新的创作视角。

大语言模型创作的质量取决于投给模型的“语料”(数据)。就目前大语言模型生成的内容来看, 这些内容通常不是模型自己创造出来的, 而是对既有数据库和网络数据的深度整合; 就其给出的观点而言, 基本也是对已有学术观点的呈现或整合, 或者只是在阐释方法上有些改进。有时, 会出现一些颇具“创造性”的内容或观点, 但很多都是“知识幻觉”, 甚至是错误的。

在现阶段, 大语言模型就像神笔马良手中的画笔, 本质上是强大的工具, 并不是替代人类的内容生产者。要想生产优质的科普内容, 主导者仍然是人。

二、利用大语言模型开展科普创作的核心是什么?

什么样的科普书才是好的科普书? 数据精确、通俗易懂这些都是必要条件。其实, “有问

*通信作者: 史军, 玉米实验室科普工作室创始人, 研究馆员, 研究方向为植物学和科学传播。sj027@sina.com。

题”的科普书才是好的科普书。一部优秀的科普书，一定是围绕一个有趣且有意义的问题展开。

今天，知识更新的速度已经远超人们的想象。以生物学领域为例，30年前，人们认为DNA上那些不能编码蛋白质的片段（non-coding DNA）就是“垃圾DNA”；30年前，科学家还在努力破解人类全基因组精细图谱……30年来，科学家完成了首个完整的人类基因组测序，发现那些“沉默”的基因片段是非常重要的开关^[1]。

如今，以教授知识为核心的教育很难跟上科学技术的发展速度。面对迅速变化的科学世界，科普作家的创作还有意义吗？答案肯定的。

科普创作的重心并不是一味地输出知识，而是倡导科学方法、传播科学思想、弘扬科学精神^[2]，让受众具备用科学方法来分析和解决问题的能力。比如，对于“人是从猿猴变化而来的”这一说法，科普创作者可以引导受众思考“动物园里的猴子为什么没有变成人”，并基于现代生物进化理论，给青少年读者讲一个关于人和猿有共同祖先的故事——在食物、气候等条件为变量的自然选择过程中，人和猿走上了不同的演化道路。在讲述过程中，除了普及生物学知识，更重要的是让受众理解如何基于生物进化理论去看待、分析问题。

好的科普作品应当展示科学研究历程，而非科学知识的简单汇集。创作一本优秀的科普图书，首先需要确立一个核心问题，并且这个问题是需要经过资料收集、分析整理才能够清楚作答的。图书内容应当围绕核心问题层层展开，引导读者随着作者的思路去了解、剖析和学习这个问题，经过探索最终获取答案。

大语言模型可以协助科普创作者梳理、明确核心问题，让科普作品在创作之初就带有优秀

基因。学习掌握科学知识、科学方法、科学思维和科学精神，需要进行长时间的反复训练，学习如何去收集资料、如何分析和解决科学问题并提升思维层次。带有明确问题意识的优秀科普作品是重要的学习素材，它们给读者提供的价值远超过科学知识本身，从“授人以鱼”变成“授人以渔”^[3]。

三、如何理解大语言模型在科普创作中的价值？

尤瓦尔·赫拉利（Yuval Harari）在《人类简史》（*Sapiens: A Brief History of Humankind*，中信出版社2017年版）中阐述了一个颇为独特的观点。赫拉利认为，人类的大脑并不是为了处理精确的数字而生的。人类在演化的漫长时间里，所面对的大多数是图像信息和社会关系信息。进入农业社会之后，数字才变得越来越重要，但存储和处理数字信息不是人类大脑擅长的工作^[4]，人类用创造工具的能力来弥补人脑在直接处理海量数字信息时的不足。

数字系统的建立以及电子计算机的出现，让人类有了处理复杂数字问题的机会。比如，专业研究人员可以利用电子计算机等工具计算出行星的轨道，计算出飞船太空旅行的路线，甚至可以推算出太阳系外正在发生的大事件。大语言模型的出现，让普通公众拥有信息处理层面的强大工具。

不同的大语言模型进行推理的方式不同，但是它们都可以帮助科普创作者实现三个层面的工作。一是高效收集和整理资料。大语言模型拥有极强的收集和整理资料的能力，可以节省科普创作者的大量时间。但需要注意的是，创作者要为大语言模型提供尽可能精准的“语料”供其学

习，与此同时，创作者仍然需要去精读高质量的文献原文，以加深自己对内容的理解。二是发现话题之间的相关性，带来新的思路。基于大语言模型的运行逻辑，它们不难找出一些隐含在不同科学问题之间的相关性，这对于启发创作者的思路大有裨益。三是拓展创作形式。大语言模型不仅可以帮助创作者润色语言、调整文字风格，甚至可以帮助创作者突破原有的内容形态，不再拘泥于文字、绘画和视频的界限，将内容用最适合的方式表现出来，创作者在创作过程中不再被表现技法所限制，可以尽情释放自己的创意。

需要说明的是，有些创作者在尝试过大语言模型之后，认为对其内容生产没有帮助。实际上，如果仅提供一个简单的问题，就期待大语言模型创造出一个完美的作品，这样的尝试肯定是无效的。创作者必须清醒地认识到大语言模型的特性，并将其结合到自己的创作过程中。

四、如何应对大语言模型对科普创作者提出的挑战？

随着大语言模型辅助创作的推广，单纯的文字写作、绘画以及视频剪辑工作很大可能会被模型取代，与核心内容有关的创意、想法将成为稀缺资源。想要成为一个出色的科普创作者，关键在于能够提出让受众眼前一亮的科学问题。

对于经历过很多科研训练的创作者而言，

可以把科研逻辑运用于大模型辅助科普创作当中。科研训练本身就是一个建成世界观的过程，构建一个完整且自洽的世界观对于科普创作者分析和阐释科学问题大有裨益。比如，所有的生物学问题都可以归结到生物进化问题上，那么要想做好生物学科普，就需要理解生物进化的理论，以及不同理论之间的差异和生物进化理论的发展过程，而不是简单停留在“自然选择”和“用进废退”这些词汇的表层意思上。

要做到这点，最简单的方法就是阅读经典著作。作为一个生物学工作者，应当深入阅读《物种起源》（*The Origin of Species*）等经典著作，并且能够理解达尔文阐释的概念，以及论证的逻辑和证据。这不仅仅是一个构建世界观的过程，也是一个学习讲述科学故事技巧的过程。利用大语言模型，构建属于自己的知识体系，赋予科普作品鲜明的个人色彩，是当下所有科普创作者必须做的事情。

现阶段，大语言模型产生幻觉这个问题依然突出，大语言模型产生的非虚构类作品，需要创作者仔细核查，以保证其精确性。前文提到，创作者需要自行阅读学习高质量的文献原文，让自己具备判断大模型生成内容正确性的能力。作为一种强大的内容生产工具，大语言模型可以辅助科普创作者，提升科普创作效率，新时代的科普创作者应积极拥抱新技术，探索出人机结合的新的创作道路。

参考文献

- [1] PENNISI E. ENCODE Project Writes Eulogy for Junk DNA[J]. *Science*, 2012, 337(6099): 1159–1161.
- [2] 中华人民共和国科学技术普及法 [EB/OL]. (2024–12–25) [2025–05–07]. https://www.gov.cn/yaowen/liebiao/202412/content_6994555.htm.
- [3] 王大鹏. 愿景与门道：40 位科普人的心语 [M]. 南京：江苏凤凰科学技术出版社，2023.
- [4] 尤瓦尔·赫拉利. 人类简史 [M]. 北京：中信出版社，2017.

（编辑 / 邹 贞 齐 钰 李 莹）