arXiv:2603.01718v1 [cs.DL] 2 Mar 2026

2022年后时期稿件长度、研究团队规模和国际合作的变化：来自*PLOS ONE*的证据

Yossi Ben-Zion[1*]、Eden Cohen[1] 和 Nitza Davidovitch[2]1*物理系，巴伊兰大学，以色列拉马特甘，52900。[2]教育系，阿里埃勒大学，以色列阿里埃勒，40700。

*通讯作者。电子邮箱：benzioy@biu.ac.il;

摘要

自2022年末以来，大语言模型（LLM）已迅速渗透到学术写作中。我们利用2019年至2025年间发表的109,393篇研究论文的完整总体，考察2022年前后总体层面的结构性发表指标，包括全文稿件长度、作者团队规模、参考文献数量和跨语言合作。*PLOSONE*的多学科范围和一致的编辑框架，使我们能够在较长时期内于统一条件下进行跨领域比较。稿件长度显著增加：非洲作者的增幅为14.8%，亚洲作者为11.7%，英语母语作者（NES）为5.3%，使字数差距缩小了39%。更引人注目的是，非英语母语作者（NNES）的作者团队规模从6.54名作者降至6.06名，降幅为7.3%；其与NES合作者的合作比例也从17.8%降至12.2%，降幅为36%；与此同时，NES作者在团队规模和合作率两方面均保持稳定。各组的参考文献数量均小幅且一致地增加。这些发现表明，2022年后的工具可能正在重塑的不仅是科学的写作方式，还有共同开展写作的人员构成。

关键词：稿件长度、研究团队规模、科学合作、署名模式、文献计量学

1 引言

数十年来，英语作为国际科学出版物主要语言的主导地位，塑造了全球知识生产中的参与格局。尽管大多数发表论文的科学家都是非英语母语者（NNES），但被索引的科学产出绝大部分以英语撰写（Kourilova-Urbanczik，2012年；W.Liu，2017）。非英语母语学者在写作和修改上投入的时间要多得多（Amano 等，2023年；Ram´ırez-Casta˜neda，2020年），且因语言相关因素面临更高的拒稿率（Lenharo，2023年）。

大型语言模型（大语言模型），尤其是 ChatGPT 于 2022年末公开发布，恰逢广泛易得的学术写作起草与语言润色工具问世。实验性证据表明，生成式人工智能能够提高写作生产力，并改善人们对产出质量的评价（Noy &Zhang，2023年）。文献计量分析记录了大语言模型辅助写作在各科学领域的迅速普及，其中非英语母语学者的采用尤为显著（Filimonovic, Rutzer,Macher, &Weder，2025年；Liang 等，2024年；J. Liu, He,Zheng,Bu,& Ni，2025年）。双重差分分析进一步显示，2022年后非英语母语者撰写的摘要词汇复杂度有所提高（Lin, Zhao, Tian,&Li，2025年）。

文献计量研究采用稿件长度、作者团队规模、参考文献数量和合作模式等指标，来描述科学生产的结构（Moed, Gl¨anzel,& Schmoch，2004）。2022年后的时期是否与这些结构性出版特征的变化相关，以及此类变化是否因英语母语者与非英语母语学者而异，尚未得到研究。这些指标的变化可能会影响科学体系中的参与程度与分层状况。

我们利用一个大规模文献计量数据集，分析2022年前后稿件长度、作者团队规模、参考文献数量和国际合作的趋势。我们不尝试检测人工智能生成文本，也不推断文档层面的工具使用情况。相反，我们考察从完整全文记录中提取的客观出版指标。我们的分析聚焦于多学科期刊*PLOS ONE* 在2019年至2025年间发表、且截至数据检索时已被收录至2025年第四季度的全部研究论文，以便在一致的编辑框架内进行跨领域比较。我们将2022年视为扩散前的最后一个基准年份，并将此后出版特征与之比较，以评估不同语言群体随时间发生的变化和差异。

据此，我们提出三个研究问题：

•RQ1：2022年之后，稿件长度是否发生变化？不同语言背景的变化趋势是否有所不同？

• RQ2：同一时期内，作者团队规模和国际合作模式是否发生变化，尤其是在非英语母语学者群体中？

• 研究问题 3：不同语言群体的参考文献数量是否呈现出不同的时间趋势？

2 方法

2.1 数据与样本

PLOS ONE 的入选有三个原因：其多学科范围便于在统一的编辑条件下进行跨领域比较；其开放获取政策和公共 API 提供了覆盖整个研究期间的完整全文记录；此外，其投稿量始终较高，确保了所考察的每一年中不同语言和地理亚群体都有充足的样本量。

该数据集包含2019年至*2025*年间发表在 *PLOS ONE* 上的109,393篇研究论文，这些论文在数据检索时已被索引，截至2025年11月。表1报告了各年份的文章数量。仅保留被归类为研究论文的文章；更正、撤稿和勘误均被排除，因为其词数反映的是修改后或撤回的内容，而非完整稿件。

通过*PLOS ONE* 公共 API，使用自定义 Python 脚本以编程方式获取文章。对于每个日历年，首先通过初始查询返回符合条件的文章总数。随后，以每次请求1,000条记录的批次获取包括数字对象标识符、标题、文章类型和主题分类在内的元数据。之后，以每次请求四篇文章为一组下载全文 XML 文件；选择这一批次大小是为了平衡检索速度与服务器负载。

文章类型通过 API 元数据和 XML 文件共同核验；当 API 元数据缺失或不一致时，以 XML 文件作为备用依据。

表 1：*PLOSONE* 按年份（2019—2025年）发表的研究论文数量。

Year	文章数量（N）
2019	15,300
2020	16,015
2021	16,094
2022	15,584
2023	13,814
2024	16,203
2025	16,383
总计	109,393

2.2 稿件长度

词数根据每篇文章的全文 XML 计算，涵盖前置内容、正文和后置内容。提取的文本经清理，去除残留标记后，先合并为一个字符串，再进行分词。

词被定义为由空白分隔且包含两个或更多字符的词元。此阈值可排除单字符格式伪影（例如表格边框字符），否则这些字符会使计数虚高。省略单字符词（例如，

“I”、“a”）对组间比较的影响微乎其微。词数包括参考文献列表。

本研究评估了多种不同的计数方法，考察是否纳入参考文献和单字符词元。所选方法与人工计数结果的偏差最小。我们从随机抽取的 100 篇文章中进行验证，将结果与 Microsoft Word 及多个独立的在线字数统计工具进行比较。观察到的偏差可以忽略不计。

2.3 参考文献

通过统计各篇文章全文 XML 中参考文献部分列出的书目条目，得出参考文献数量。

2.4 作者语言背景分类

从 XML 记录中提取作者姓名和机构隶属关系。通过唯一的方括号标识符（例如，`[aff001]`）将每位作者与其机构隶属关系关联起来，以便在作者有多个机构隶属关系时准确分配。部分 XML 文件中出现的编辑隶属关系在分类前已识别并删除。

使用 GPT-5 mini 通过 OpenAI API 对语言背景进行分类。对于每篇文章，模型都会收到作者名单及其对应的机构隶属关系，并被要求回答三个是非问题：（1）第一作者是否隶属于英语为母语者占多数国家的机构；（2）末位作者是否符合这一条件；（3）是否至少有一位作者符合这一条件。英语为母语者占多数的国家定义为美国、英国、爱尔兰、澳大利亚、新西兰和加拿大。当作者有多个机构隶属关系时，只要其中任何一个符合标准，该文章就编码为 1。每次 API 调用最多分批处理 150 篇文章。完整提示词见附录 A。

在所有分析中，文章均依据第一作者（英语母语者或非英语母语者）的语言背景进行分类；第一作者通常是研究的主要贡献者（Larivi`ere etal., 2016; Tscharntke, Hochberg, Rand, Resh, & Krauss, 2007）。在数据集中的文章里，92.8%的第一作者和最后一作者属于同一语言群体，这表明按最后一作者分类会得出实质上相同的结果。

为评估分类准确率，作者对随机抽取的100篇文章进行了独立编码，并将结果与模型输出进行比较。一致率为100%。这一高一致率有两个原因。首先，分类任务的歧义相对较小：通常可以直接根据隶属机构文本判断其是否属于英语占多数的国家。其次，如附录 A 所述，提示词经过反复迭代完善，直至形成稳定的表述。在整个开发过程中，均进行了人工检查。尤为重要的是，验证样本是在最终提示词确定后抽取的，从而确保所报告的一致率反映的是样本外表现，而非针对样本内数据进行优化的结果。

为进一步评估基于大语言模型的分类结果的稳健性，我们于2025年12月1日使用分析时可通过 OpenAI API 获取的三个 GPT-5 模型变体（`GPT-5`、`GPT-5-mini` 和 `GPT-5-nano`），对随机抽取的1,000篇文章重新进行了处理。选择这些模型，是因为它们代表了研究开展时 OpenAI 语言模型中最先进、最新的版本。三个模型的分类结果完全一致（一致率为100%）。为便于核验和审计，在后处理之前保存了原始输出。

2.5 地理分类

地理分类借助并行的大语言模型提示（附录 A）进行，并应用于作者姓名和机构隶属关系。对于每篇文章，模型识别了（1）第一作者所在的洲，（2）末位作者所在的洲，以及（3）以下五个地区中是否各自至少有一位作者：亚洲、欧洲、非洲、美洲和其他（包括澳大利亚和新西兰）。为与语言背景分类保持一致，地理区域依据第一作者的机构隶属关系确定。

当一位作者拥有跨越不同大洲的多项机构隶属关系时，采用最先列出的隶属关系所对应的洲。

2.6 领域分类

主题分类以层级格式从*PLOS ONE* API检索获得，其中每个主题标签表示从宽泛领域到具体子主题的路径（例如，/医学与健康科学/传染病/细菌感染）。对于每个主题标签，仅保留顶级领域。在整个研究期间，共观察到十一个不同的顶级领域。

为给每篇文章指定一个主要领域，我们采用了年度加权程序。对于每一年 y，令 N_y 表示文章总数，N_{iy} 表示至少出现一次领域 i 的文章数量。每个领域的年度权重与其普及程度成反比：

$$w_{iy} = \frac{N_y}{N_{iy}}. \tag{1}$$

对于在年份 y发表的文章 d，计算其在各领域 i 上的加权得分：

$$s_{iy}(d) = f_{iy}(d) \cdot w_{iy}, \tag{2}$$

其中，$f_{iy}(d)$表示领域 i 在该文章主题标签中出现的频次。保留重复的主题标签，因为重复出现反映了某一领域对文章的表征程度。随后，将主要领域指定为：

$$\mathrm{Field}(d) = \arg\max_i s_{iy}(d). \tag{3}$$

此方法会提高较少见领域的权重，同时通过标签频率保留文章内部的主题侧重。

3 实证策略

由于稿件字数呈右偏分布，所有关于稿件长度的主要分析均基于经过 $\log_{10}$ 转换的词数进行。图表则采用原始尺度呈现，以便于解读。

稿件长度的时间差异（***2019—2025***年）。
使用对 $\log_{10}$（词数）进行的单因素方差分析评估不同发表年份之间的差异。采用经 Tukey 校正的估计边际均值计算两两事后比较。

作为稳健性检验，我们还在原始词数尺度上进行了 Kruskal–Wallis 检验，并进行了 Dunn 事后比较（Bonferroni 校正）。这些非参数检验得出了实质上相同的结论，因此未予报告。

第一作者语言背景的异质性（***2019—2025***年）。
为评估时间模式是否因第一作者语言背景而异，我们对 $\log_{10}$（词数）进行了双因素方差分析，纳入年份和第一作者组别的主效应及其交互作用。事后对比采用估计边际均值计算。

按洲划分的异质性（***2022***年与***2025***年）。
对于洲级分析，我们将样本限制在2022年和2025年，并对 $\log_{10}$(词数) 进行双因素方差分析，纳入年份、洲组别及其交互作用。我们计算了经 Holm 校正的事后对比；推断主要基于交互效应。

按主要领域划分的异质性（***2022***年与***2025***年）。
我们使用2022年和2025年的子样本评估领域异质性，对 $\log_{10}$(词数) 进行双因素方差分析，纳入年份、主要领域及其交互作用。我们计算了经 Holm 校正的事后对比；推断主要基于交互效应。

双重差分回归（***2022***年与***2025***年）。
我们使用2022年和2025年的子样本，通过线性回归模型估计不同第一作者语言背景的稿件长度在2022年至2025年间的差异化变化。我们估计了：(i) 仅纳入年份的基线模型；(ii) 纳入年份 $\times$第一作者交互项的双重差分模型；以及 (iii) 在此基础上加入主要领域固定效应和年份与领域交互项的完整模型。

模型分别基于原始词数尺度和 $\log_{10}$(词数) 进行估计。不同模型设定的结果在实质上相当。为与分布特性和方差分析保持一致，本文仅报告 $\log_{10}$ 结果。

作者团队规模和参考文献数量（***2022***年与***2025***年）。
作者团队规模与参考文献数量采用带对数链接的负二项回归进行建模，模型包含年份与第一作者语言背景之间的交互作用。指数化系数以发生率比报告。

与***NES***合作者的合作（***2019—2025***年）。
在NNES第一作者的文章中，至少有一名NES合作者的概率采用逻辑回归建模，模型纳入发表年份指标，并以2019年为参考年份。结果以优势比及置信区间报告。

3.1 软件与缺失数据。

所有统计分析均在R（4.5.0版）中进行。线性模型和方差分析模型使用R基础函数估计；负二项回归使用`MASS`软件包估计；事后对比使用`emmeans` 软件包计算。图在 *Python* 中生成（使用`pandas` 和`matplotlib`）。结果变量或关键预测变量存在缺失值的观测值已从相应模型中排除（完全案例分析）。

4 结果

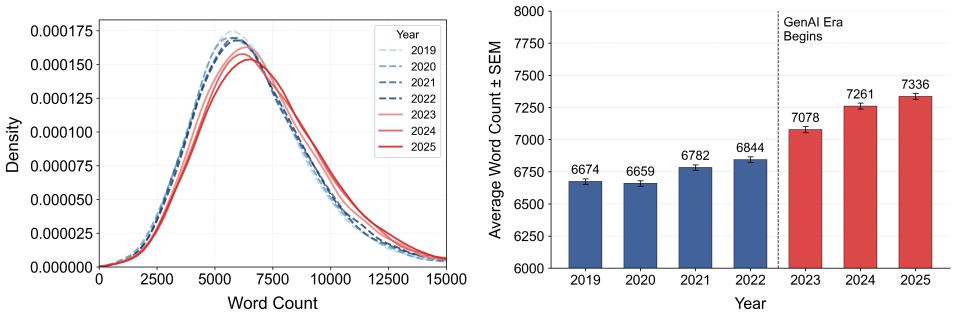
4.1 稿件长度的时间变化（2019—2025年）

2019年至2025年间，文章长度呈非线性增长。由于词数分布右偏，所有统计分析均基于经对数$_{10}$转换后的数值进行；为便于解读，图中展示的是原始词数尺度上的数值（图1）。

图1A 展示了文章长度分布的两个明显时期：2019年至2022年间，各分布大体重叠；而从2023年起，分布明显向右偏移，表明与2023年之前相比，稿件长度整体更长。

单因素方差分析证实，发表年份对 $\log_{10}$（词数）有显著影响（$F(6,\ 109,\ 386) = 145.80,\ p < .001$）。事后比较（Tukey HSD检验）显示，2019年与2020年之间没有差异（$p = .897$）。2020年至2021年间词数有所增加（$p = .001$），但2019年至2021年的累积差异未达到统计显著性（$p = .082$）。2021年至2022年间未观察到变化（$p = .373$）。2022年至2023年间出现显著增长（$p < .001$），这是研究期间观察到的最大年度变化；随后，2024年相较于2023年进一步增加（$p < .001$）。2024年与2025年之间未发现显著差异（$p = .318$）。

与这些趋势一致，文章的平均长度从 2023年开始大幅增加（图 1B）。到 2025年，文章平均比 2019年长约 662个词，且从 2022年起与 2019年基线的所有比较均具有统计学显著性（均 $p < .001$）。



(a) (b)

图 1：**稿件长度趋势（2019—2025年）**。**（A）**各年份的词数分布（单位：词）。虚线表示2019—2022年；实线表示2023—2025年。**（B）**文章平均长度（单位：词，± 均值标准误）。竖直虚线表示2023年观察到的显著增长开始的时间。统计检验基于对数$_{10}$转换后的数值；图中数值按原始词数尺度呈现。

4.2 差异化加速与语言趋同

汇总趋势掩盖了作者语言背景方面的显著异质性。因此，我们根据第一作者是英语母语者还是非英语母语者进行分层分析。

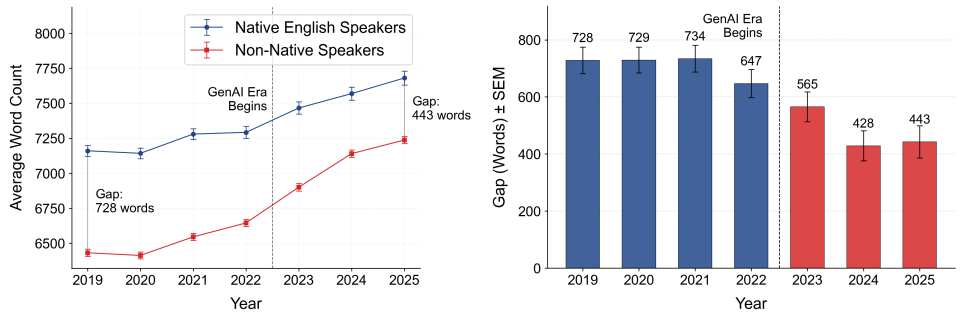
以对数$_{10}$（词数）为因变量、年份和第一作者语言背景为因素进行的双因素方差分析显示，年份具有显著主效应（F(6，109，375) =147。25, $p < .001$），语言背景也具有显著主效应（F(1，109，375) $= 1$, 129。45, $p < .001$），且交互作用显著（F(6，109，375) $= 11.01$, $p < .001$），这表明不同群体的稿件长度随时间变化的趋势存在差异。

图2A 展示了各组的文章平均长度随时间的变化。NES 在每个时间点撰写的稿件都更长（所有事后检验均 $p < .001$）。2019年，基线差异为728词（7,161词对6,432词）。尽管差距始终显著，但其幅度总体有所缩小：对数$_{10}$ 尺度上的估计差异从2019年的0.049降至2025年的0.028。

2022年后，两组的文章长度都增加了，但增速不同。NES 的文章长度从2022年的7,292词增至2025年的7,680词（+5.3%），而 NNES 则从6,645词增至7,238词（+8.9%）。因此，原始差距从2019年的728词缩小至2025年的443词，降幅为39%。图2B 展示了英语母语者（NES）与非母语英语作者（NNES）之间文章长度的年均差异。

4.3 文章篇幅增长的区域差异（2022年 vs. 2025年）

为评估稿件长度增长的区域差异，我们采用以年份和洲为因素、对词数取 $\log_{10}$ 后进行的双因素方差分析，比较了 2022年和 2025年各主要地理区域群组的平均词数。分析显示，年份的主效应显著（F(1, 31,931) $= 227.17$, $p < .001$），洲的主效应也显著（F(4, 31,931) $=$



(a) (b)

图2：按作者背景划分的文章长度趋同。（A）2019年至2025年英语母语者（蓝色）和非英语母语者（红色）的文章平均长度（词数，± 均值标准误）。2022年之后（垂直虚线），两组的文章长度均有所增加，其中非英语母语作者的绝对增幅更大。两组之间的差距从2019年的728词缩小至2025年的443词。**（B）**七年期间英语母语者与非英语母语者之间词数的平均差异（±均值标准误）。2019年至2022年间，两组之间的差距稳定在700词以上，2022年之后则显著下降。这一下降与生成式人工智能时代的开始相吻合（垂直虚线），差距在2024—2025年缩小至最低水平。

155。53, $p < .001$），交互作用也显著（$F(4, 31{,}931) = 8.90$, $p < .001$），这表明各区域的变化幅度有所不同。

事后比较显示，五个区域组从2022年到2025年均呈现显著增长（均为 $p < .05$）。增幅最大的是非洲（非母语者），从6,033词（$N_{2022} = 1$，312）增至6,929词（$N_{2025} = 1$，227；+14.8%；对数$_{10}$ 差异 = 0.058，$p < .001$）；其次是亚洲（非母语者），从6,347词（$N_{2022} = 5$，452）增至7,090词（$N_{2025} = 8$，273；+11.7%；对数$_{10}$ 差异= 0.046，$p < .001$）。欧洲（非母语者）从7,303词（$N_{2022} = 3$，293）适度增加至7,860词（$N_{2025} = 2$，511；+7.6%；对数$_{10}$ 差异 = 0.029，$p < .001$）。英语母语者从7,292词（$N_{2022} = 4$，796）增至7,680词（$N_{2025} = 3$，653；+5.3%；对数$_{10}$ 差异 = 0.023，$p < .001$）。增幅最小的是美洲（非母语者），从6,951词（$N_{2022} = 715$）增至7,255词（$N_{2025} = 709$；+4.4%；对数$_{10}$ 差异 = 0.022，$p = .018$）。

组间比较显示，随着时间推移，若干地区之间的差距有所缩小。2022年，英语母语者与亚洲（非母语者）之间存在差异（log$_{10}$ 差异 = 0.059,$p < .001$），而在2025年，这一差异有所减小（log$_{10}$ 差异 = 0.036, $p <$。001）。同样，2022年亚洲（非母语者）与非洲（非母语者）之间存在差异（$p <$。001），但2025年则不存在差异（$p = .413$）。基线词数较低的地区在2022年后增幅更大，因此到2025年，各地区稿件平均长度的差异有所缩小。这些地区增长模式及其带来的地区间差异缩小情况见图3。

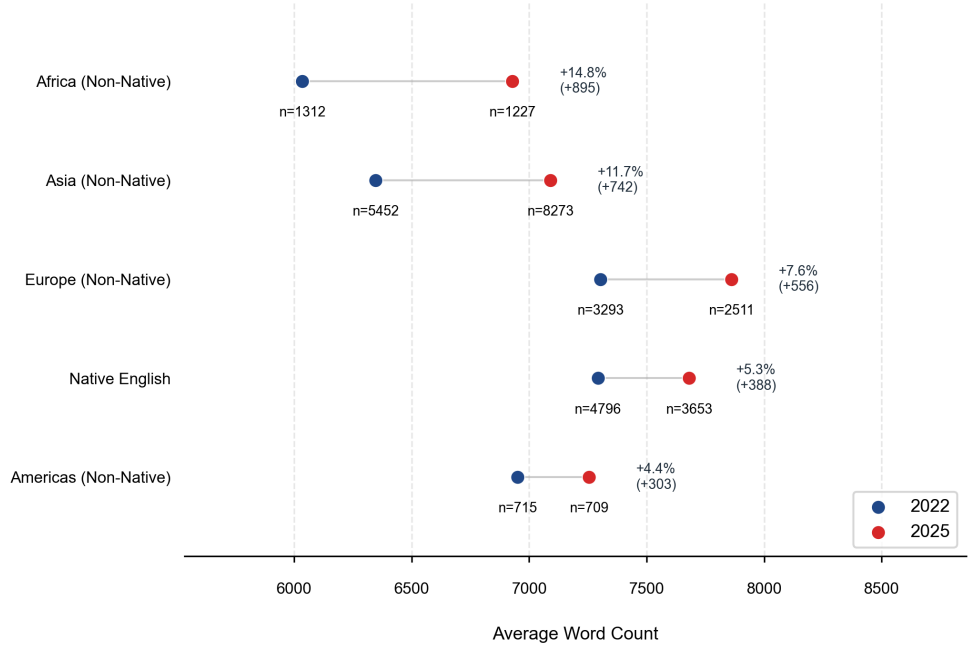


图 3:稿件长度增长的区域差异（2022年与2025年）。哑铃图展示了各区域2022年（蓝色）和2025年（红色）的平均词数。百分比表示这两年间的相对增长，且 n 表示每个区域分析的稿件数量。

4.4 文本扩展的学科异质性（2022–2025）

2022年至2025年间，稿件长度的增长幅度在不同研究领域间存在显著差异。以年份（2022年与2025年）和研究领域为因素，对$\log_{10}$（词数）进行双因素方差分析，结果显示年份具有显著主效应（$F(1,\ 31,\ 893) = 236.50$, $p <$。001），研究领域也具有显著主效应（$F(10,\ 31,\ 893) = 188.92$, $p < .001$），且交互作用显著（$F(10,\ 31,\ 893) = 10.18$, $p < .001$），表明变化幅度因学科而异。

事后检验Tukey比较显示，各领域的增长存在显著差异（图4）。医学与健康科学增幅最大（对数$_{10}$ 差 $= 0.056$，$p < .001$；$+13.6\%$），其次是工程与技术（对数$_{10}$差 $= 0.042$，$p < .001$；$+9.8\%$）和人物与地点（对数$_{10}$ 差 $= 0.035$，$p < .001$；$+8.6\%$）。物理科学（$+7.9\%$；$p < .001$）、研究与分析方法（$+8.0\%$；$p < .001$）、生物学与生命科学（$+7.2\%$；$p < .001$）以及计算机与信息科学（$+7.6\%$；$p < .001$）也出现了显著增长。到2025年，计算机与信息科学的均值词数在所有领域中最高，每篇文章超过8,000词。

相比之下，生态学与环境科学（$p = .955$）、社会科学（$p = .776$）、地球科学（$p = .465$）或科学政策（$p = .161$）均未检测到显著增长。

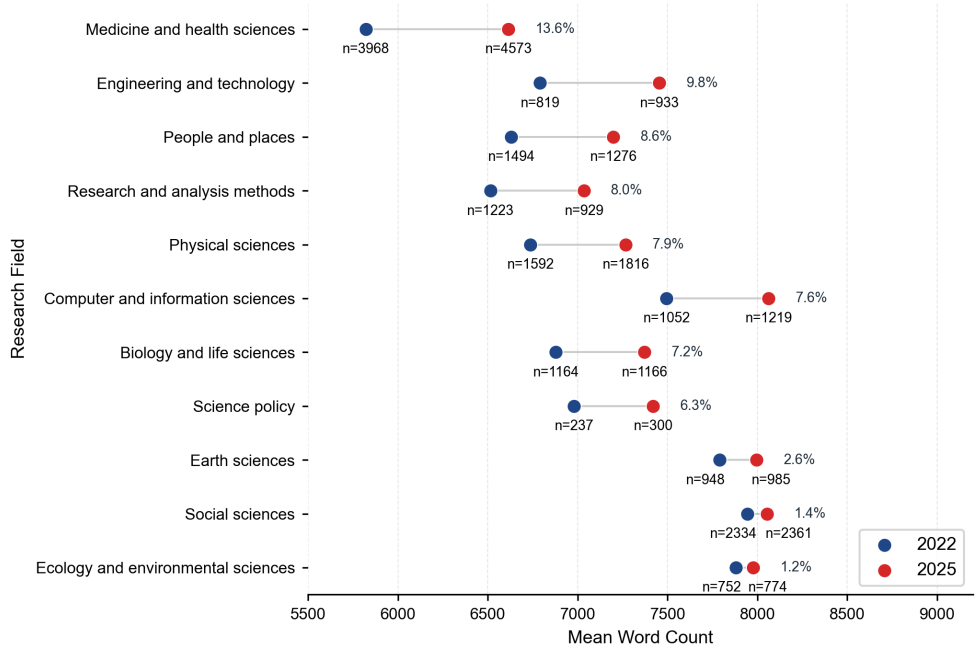


图4：不同研究领域文章平均长度的变化（2022年与2025年对比）。哑铃图展示了各研究领域2022年和2025年的平均词数（以词为单位）。数据点代表年度均值，水平线段表示两年间的差异。每个线段右侧标注的百分比表示该领域词数的相对增长。各年份的样本量（n）也一并显示。

4.5 文章长度的双重差分回归

我们针对$\log_{10}$(词数)估计了一个双重差分（DiD）回归模型，对比2025年与2022年，并纳入作者语言背景和研究领域。模型设定如下：

$$\log_{10}(\text{词数}_i) = \alpha + \beta_1 \text{时间}_i + \beta_2 \text{母语背景}_i + \beta_3 (\text{时间}_i \times \text{母语背景}_i) + \sum \delta_k (\text{时间}_i \times \text{研究领域}_k) + \varepsilon_i \tag{4}$$

其中，时间表示2025年（相较于2022年），Native表示第一作者的语言背景，交互作用项用于捕捉各组之间的差异化增长。共计

$N = 31$，在排除缺少领域分类的记录后，纳入了915篇文章。结果见表2。

模型1估计了2022年至2025年之间的无条件时间效应，并显示文章长度总体显著增加（$\log_{10}$ 系数 $= 0.03$，$p < .001$；$R^2 = 0.007$）。

模型2引入了语言异质性。英语母语者表现出更高的基线词数（0.04 $\log_{10}$ 单位，$p < .001$），相当于大约多出9–10%的词数。交互项表明，相对于非英语母语者，英语母语者的增长较慢（交互作用 $= -0.01$，$p = .005$；$R^2 = 0.014$）。

模型3纳入了领域特异性效应和交互作用。解释方差大幅增加（$R^2 = 0.072$），总体时间效应降至0.01，且不再具有统计学显著性（$p = .161$）。领域特异性增长集中于医学（0.05，$p < .001$）、工程学（0.04，$p < .001$），以及包括物理科学、以人为本的研究、生物学、方法和计算机科学在内的一组领域（均约为0.03，均 $p < .01$）。政策（0.02，$p = .199$）、地球科学（0.00，$p = .695$）和生态学（-0.00，$p = .873$）未显示出显著的领域特异性增长。

重要的是，在完整模型中，语言交互作用仍具有统计学显著性（交互作用 $= -0.01$，$p = .040$），这表明在控制学科差异后，英语母语者与非英语母语者之间的趋同仍然存在。

4.6 作者团队规模趋势（2019—2025年）

对2019年到2025年作者团队规模进行的纵向分析显示，英语母语者与非英语母语者之间的趋势存在分化（图 5）。为适当地对这一计数结果进行建模（该结果存在过度离散，即方差超过均值），我们采用了带对数连接函数的负二项回归，而不是标准方差分析。模型将第一作者语言背景（英语母语者与非英语母语者）、发表年份（2022年与2025年）及二者的交互作用作为预测变量。排除作者人数超过100人的文章后，共获得 $N = 31,963$个观测值。

模型显示，第一作者身份与年份之间存在显著交互作用（$\beta = -0.078$，$SE = 0.015$，$z = -5.24$，$p < .001$），表明英语母语者与非英语母语者第一作者的作者人数随时间的变化不同。使用估计边际均值进行的事后检验表明，2022年，由英语母语者第一作者主导的出版物（均值 $= 6.61$）与由非英语母语者第一作者主导的出版物（均值 $= 6.54$；比率 $= 1.009$，$p = .366$）之间的团队规模没有显著差异。相比之下，到2025年，英语母语者第一作者的出版物作者人数显著高于非英语母语者第一作者的出版物（均值分别为 $= 6.61$ 和 $= 6.06$；比率 $= 1.091$，$p < .001$）。

值得注意的是，这一分化是由NNES团队规模的选择性缩减所驱动的。对于由NES第一作者署名的出版物，2022年与2025年之间未观察到显著变化（比值 $= 0.999$，$p = .956$）。然而，对于NNES第一作者，作者均值数量从2022年的6.54显著下降至2025年的6.06（比值 $= 1.081$，$p < .001$）。图5 中的纵向数据表明，2019年至2022年期间，两组的团队规模均相对稳定；但在2022年这一临界点之后，发展轨迹出现了明显分化。

表2：文章长度对数$_{10}$的分层双重差分回归结果，2022–2025

变量	模型1	模型2	模型3
常数项（基线）	3.80***	3.79***	3.86***
时间（2025年相对2022年）	0.03***	0.04***	0.01
NES	–	0.04***	0.04***
时间 ×原生	–	−0.01**	−0.01*
Baseline Field Differences in 2022 (vs. Social Sciences)			
医学	–	–	-0.14***
工程学	–	–	-0.07***
生物学	–	–	-0.06***
计算机科学	–	–	-0.02**
物理科学	–	–	-0.07***
以人为本的研究	–	–	-0.08***
方法	–	–	-0.09***
政策	–	–	-0.06***
地球科学	–	–	-0.01
生态学	–	–	0.00
Field-Specific Growth (Interaction with 2025)			
时间 ×医学	–	–	0.05***
时间 ×工程学	–	–	0.04***
时间 ×生物学	–	–	0.03**
时间 ×计算机科学	–	–	0.03**
时间 ×物理科学	–	–	0.03***
时间 ×以人为本的研究	–	–	0.03***
时间 ×方法	–	–	0.03**
时间 ×政策	–	–	0.02
时间 ×地球科学	–	–	0.00
时间 ×生态学	–	–	−0.00
领域固定效应	No	No	Yes
观测值	31,915	31,915	31,915
R^2	0.007	0.014	0.072

* $p < .05$, ** $p < .01$, *** $p < .001$。参照类别：社会科学（领域）、非母语（语言）。

4.7 2022年后的国际合作转变

我们使用二元逻辑回归分析了由NNES署名的76,338篇稿件中的合作模式，并以2019年作为参考年份（表3）。结果变量表示是否至少有一名英语合作者参与。

相较于2019年，合作在2020年和2021年有所下降，2022年部分恢复，随后自2023年起大幅减少。到2025年，非英语母语作者与英语合作者合作的可能性比基线低36%（$OR = 0.643$，95%置信区间 $[0.60, 0.69]$，$p < .001$）。

图6 展示了逐年的变化轨迹，突显了2022年后趋势的转变。

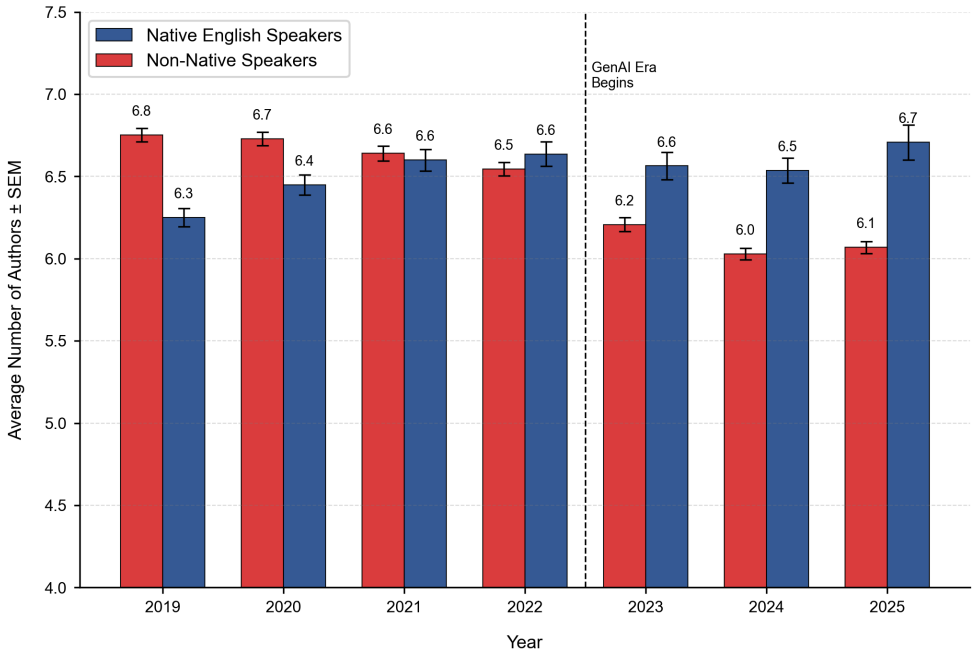


图 5：2019年至2025年按第一作者语言背景划分的作者团队规模变化。该图展示了七年期间英语母语作者和非英语母语作者每篇出版物的平均作者人数。误差线表示均值标准误。y轴截断于4。统计比较聚焦于2022年与2025年，结果显示存在显著的交互作用（$P < 0.001$）；其中，非英语母语作者的团队规模有所减少。

4.8 参考文献数量趋势（2022年与2025年对比）

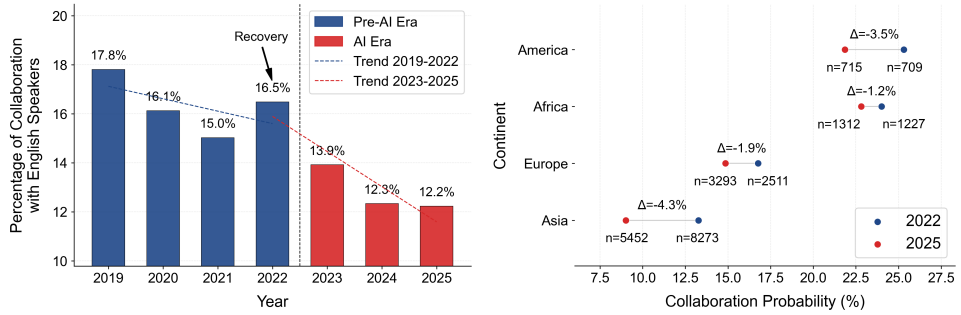
为补充对文本产出的分析，我们考察了英语母语者与非英语母语者的每篇出版物参考文献数量是否存在不同变化。鉴于参考文献数量为非负整数且存在过度离散，我们采用了带对数连接函数的负二项回归模型。因变量为参考文献数量，自变量包括第一作者语言背景、发表年份（2022年与2025年）及其交互作用（$N = 31,364$；因缺失值排除603个观测值）。

模型显示，第一作者身份具有显著的主效应（$\beta = -0.078$，$SE = 0.008$，$z = -9.84$，$p < .001$）：NNES第一作者发表的论文平均参考文献数量较少（发生率比 $= 0.93$）。年份也呈现显著的主效应（$\beta = 0.037$，$SE = 0.010$，$z = 3.72$，$p < .001$），表明与2022年相比，2025年的参考文献数量总体有所增加（发生率比 $= 1.04$）。第一作者身份与年份之间的交互作用不具有统计学显著性（$\beta = 0.004$，$SE = 0.012$，$z = 0.31$，$p = .755$），表明随着时间推移，各组之间的参考文献数量并未出现差异性变化。

表3：逻辑回归：与英语合作者合作的概率

Year	N	优势比	95%置信区间
2019年（基线）	10,217	1.00	–
2020	10,611	0.888	[0.83, 0.95]
2021	10,896	0.816	[0.76, 0.88]
2022	10,788	0.911	[0.85, 0.98]
2023	9,469	0.747	[0.69, 0.81]
2024	11,627	0.650	[0.60, 0.70]
2025	12,730	0.643	[0.60, 0.69]

参考年份：2019年。结果：至少有一位英语合作者。



(a) (b)

图6：**国际合作趋势（2019—2025年）。(A)** 非英语母语作者与英语合作者的逐年合作概率。该图显示，2020—2021年有所下降，2022年部分回升，自2023年起显著下降。**(B)** 该图展示了2022年和2025年不同地区非英语母语作者的合作概率。到2025年，各大洲均呈现一致下降。各分段上方的百分比（Δ）表示概率的绝对变化，n表示每年的样本量。

这一无显著交互作用的结果与词数和作者团队规模所观察到的显著差异性变化形成对比，表明文本产出的时间变化并未延伸至引用量。

5 讨论

5.1 稿件长度与趋同模式

生成式语言模型可以从两个相反的方向重塑学术写作。它们既可以通过精简和润色现有文本来压缩散文，也可以降低撰写更多文本的边际成本，从而促进扩张。

哪种效应占主导地位是一个经验性问题。在我们的数据中，占主导地位的模式是扩张。

稿件长度在2019年至2022年间在统计上保持稳定，均值或分布形态均无显著差异。2023年出现急剧增长，为研究期间最大的同比变化，随后在2024年继续增长，并在2025年趋于稳定。尽管这一时间点与生成式人工智能写作工具的扩散相吻合，但仅凭时间上的重合不足以进行因果推断。本研究设计为观察性设计。纳入领域固定效应以及年份与领域交互项，可减轻对跨学科构成变化的担忧，但同期的结构性影响因素无法被完全排除。

总体增长掩盖了有意义的异质性。2022年至2025年间，英语母语作者的稿件长度增加了5.3%，而非英语母语作者增加了8.9%。词数差距在2019年至2021年间一直约为730词，到2025年缩小至442词，降幅达39%。压缩最为明显的时期出现在2023年和2024年，随后趋于稳定。

这种趋同呈现出明显的梯度。基线词数较低的群体，后续增幅更大。非洲机构署名的第一作者增幅最大，达14.8%；其次是亚洲机构署名作者，增幅为11.7%；欧洲非英语母语作者为7.6%，英语母语作者为5.3%。不同学科也呈现出类似的排序：医学与健康科学和工程学的增幅高于社会科学和生态学。

Lin 等人(2025年)记录了2022年之后不同领域词汇变化的显著差异，其中技术相关学科和生物学的影响最为显著，而社会科学或艺术与人文学科则未见显著影响。这种学科异质性表明，生成式工具的采用或有效整合本身可能因领域而异。 Ding、Lawson 和 Shapira (2025年)同样记录了2022年后生成式人工智能研究的扩张在不同学科间极不均衡，计算机科学及相关技术领域的增长尤为迅速。若情况确实如此，那么此处观察到的部分差异，可能反映了不同学科共同体在应用速度或强度上的差异，而非纯粹的语言限制。

与此同时，回归均值可能机械地使起点基线较低的群体产生更强的增长。即便在控制了领域固定效应以及年份与领域交互项之后，双重差分设定中交互项的持续存在表明，趋同并非仅仅是一种统计假象。在完整模型中，主要时间效应在统计上变得可以忽略，而年份与语言背景的交互项仍具有统计显著性（$p = .040$）。因此，稿件长度的增加既反映了领域特定动态，也反映了语言背景效应，而非仅仅是跨学科的简单构成重新分布。尽管如此，该设计并不能确立因果关系。

这些结构模式与近期关于2022年后其他文本维度趋同的证据相吻合。 Lin 等人 (2025年) 报告称，在大规模OpenAlex数据中，非英语母语者撰写的摘要词汇复杂度有所提高；而 Filimonovic 等人 (2025年) 记录了数百万篇文章的文体向英语母语者写作规范趋同的现象。同样，Bao、Zhao、Mao 和 Zhang (2025年) 发现了 ChatGPT 发布后出现的变化

在词汇复杂度、句法简化以及非英语母语作者文体变化加剧方面，arXiv 摘要中也出现了类似变化。尽管这些研究考察的是词汇和文体特征，而非篇幅，但词汇、文体和结构维度上的同步变化表明，2022年后这一时期可能与学术写作的多维变化相关，尤其是在面临较高初始语言障碍的作者群体中。

近期证据也表明，学术界与人工智能相关的行为存在地域异质性。Ben-Zion、Michaeli 和 Finkelstein (2026) 记录了生成式工具问世后，非英语地区的传统学术搜索活动降幅大于美国和英国。尽管他们的研究考察的是信息消费，而非稿件产出，但他们发现的地域梯度与一种可能性相符：应用程度的差异或许促成了本文观察到的模式。这一解读仍属提示性结论，而非因果判断。

稿件长度增加可能反映了报告更加详尽、方法透明度有所提升，或论证更加充实；也可能意味着文本产出的边际成本降低，但实质性贡献并未相应增加。由此引出两个不同的问题。首先，哪些机制推动了不同语言群体和学科群体之间采用程度的差异？其次，文本扩展是否会转化为科学质量或影响力方面可衡量的提升？前一个问题涉及学术生产中的激励结构与限制因素；后一个问题则可通过关联引用影响力等下游指标及其他相关学术影响力衡量指标进行实证检验。

5.2 作者团队规模与国际合作

除了文本扩张之外，生成式语言模型的普及可能会影响科学生产的社会组织方式。如果这类工具降低了用英语起草、修改和润色稿件的成本(Aydin, Y¨ur¨uk, Reiso˘glu, &Goktas，2023年)，它们可能会部分取代以往由承担语言相关职能的合作者所作出的贡献。

结果表明，2022年之后作者团队构成发生了变化。2022年，由英语母语者和非英语母语者第一作者主导的出版物，其团队规模均值没有具有统计学显著性的差异（分别为6.61名和6.54名作者，$p = .366$）。到2025年，这种差异已具有统计学显著性：由英语母语者第一作者主导的出版物，团队规模保持稳定（均值为 $= 6.61$），而由非英语母语者第一作者主导的出版物，团队规模则降至均值6.06名作者（$p < .001$）。2022年至2025年间的减少仅见于非英语母语者第一作者。

国际合作也呈现出类似的模式。我们将分析范围限定为由非英语母语者撰写第一作者的出版物，并对其中至少包含一位英语母语者合著者的概率进行建模。以2019年为参考年份，合作在2020年和2021年有所下降，2022年部分回升，随后从2023年起持续大幅下降。到2025年，与2019年相比，NNES第一作者与英语母语者合著者合作的可能性低了36%（OR $= 0.643$, $p < .001$）。

2020年和2021年的下降与 COVID-19 相关的国际合作中断相吻合（Abramo、D'Angelo 和 Di Costa，2022年；Cai、Fry 和 Wagner，

2021年；Ni& Wang，2025年），而2022年的回升与学术流动恢复相一致。自2023年开始再次下降，发生在不同的时间背景下，仅凭疫情后续影响不太可能完全解释这一现象。还应考虑其他可能的解释。

政策驱动的地缘政治紧张局势可能是跨语言合作减少的原因之一。对美国“China Initiative”（中国行动计划）的文献计量分析显示，美中合著显著减少，并对西方合作网络产生了可测量的溢出效应。该计划于2018年启动，并于2022年正式终止（Geng, Wu, &Chen，2025年）。这些报告的估计结果与一种替代动态相符：与美国和G6合作伙伴的合作减少，同时中国国内合作增加。协作关系的这种重新配置预计会改变国际网络的构成，而不是导致所有地理区域的合作率普遍下降。还有两个模式使单纯以地缘政治解释这一现象变得更为复杂。第一，时间变化轨迹与政策驱动的抑制作用并不一致：在中国行动计划启动后的2019年，跨语言合作率达到整个研究期间的最高值（17.8%）。第二，也是最关键的一点，2022年后的收缩并非仅限于中美轴线；正如我们的区域分析（图6）所示，亚洲、欧洲、非洲和拉丁美洲的合作均持续下降。这一普遍趋势表明，背后的机制具有全球性，而非局部性。尽管仅凭文献计量趋势无法确立直接因果关系，但团队规模同步缩小、母语者合作减少，以及非英语母语作者的稿件长度呈现差异化扩张，这些现象都与生成式写作工具可能产生的影响相符。这些发现表明，技术可能推动了科学合著关系的重组；这一假设值得通过直接调查或文体文本分析进一步检验。

既有研究表明，跨国合作与认知多样性、更广阔的认识论视角以及更高的科学影响力相关（Larivi`ere, Gingras, Sugimoto, & Tsou，2015；Velez-Estevez,Garc´ıa-S´anchez, Moral-Mu˜noz, &Cobo，2022年；Wang, Jesiek, &Zhang，2024年）。如果过去与英语母语者合作者合作，除了实质性的科学贡献外，还承担了起草或编辑语言方面的职能（Lillis&Curry，2006），那么生成式工具如今可能已承担了其中一部分职能。NNES第一作者团队规模的选择性缩减，加上与英语母语者合作同步减少，这与语言合作过去所发挥的功能角色可能正在重组的观点相符。不过，这种解释仍属推测。此项分析无法确立因果关系，也未直接观察工具的使用情况。仍有必要开展持续的纵向分析，以确定这些模式反映的是暂时性调整，还是持久的结构性转变。

5.3 参考文献数量趋势

参考文献数量呈现出一种对比鲜明的模式。2022年至2025年，每篇文章的参考文献数量略有增加（发生率比 $= 1.04$，$p < .001$），但这一增幅在英语母语作者和非英语母语作者中一致，交互作用不显著（$p = .755$）。这种无显著交互作用的结果可作为内部基准：相比之下，

稿件长度和合作模式因语言背景不同而呈现出差异性变化，参考文献数量则没有。

这一对比表明，上文记录的2022年后变化，具体体现在生成式工具最直接降低生产成本的写作维度上，即文本生成和语言润色，而不是反映出版指标的普遍膨胀。参考文献整合需要识别相关来源，并将研究发现置于现有文献背景之中，因此可能不像散文写作那样直接受到生成式工具的影响。

与此同时，人们担忧大语言模型生成文本中存在伪造参考文献（Chelli 等，2024年），因此未出现差异性变化这一点值得注意：这表明，非英语母语作者对生成式工具的任何采用，都未在总体层面造成可检测的引文实践扭曲。

6 局限性

本研究存在若干值得考虑的局限性。

首先，分析仅限于单一期刊 PLOS ONE。尽管这一设计具有方法学优势，能够确保跨学科和跨时间采用一致的编辑框架，但也限制了研究结果的普适性。需要在选择性程度各异的期刊中进行重复研究，之后才能得出更广泛的结论。

其次，本研究识别出2022年后趋势与结构性出版特征之间的时间关联，但无法确立因果归因。我们无法在文献层面直接测量大语言模型的使用情况，也不存在外部对照条件。

第三，将作者划分为英语母语者或非英语母语者，依据的是机构隶属关系，而非作者实际的语言背景。这一操作化定义引入了系统性误分类：第一语言并非英语、但供职于英语占多数国家机构的研究人员会被编码为英语母语者，反之亦然。

然而，这种误分类更可能削弱而非夸大所观察到的影响：在英语占多数国家的机构中工作的非英语母语者研究人员被编码为英语母语者，从而缩小了所测得的群体差异。因此，报告中的估计值应被解读为所记录模式真实幅度的保守下限。

第四，与英语母语者合著者合作的下降，部分可能反映了某些国家在政策层面对国际研究合作施加的限制，其中最显著的是中国；在我们的样本中，中国作者占亚洲机构隶属作者的较大比例。我们无法量化这一因素的影响程度，而将政策驱动效应与技术驱动效应区分开来，仍是未来研究的一个方向。

7 结论

本研究记录了2022年后科学发表领域同时出现的三项变化：稿件变长，英语母语作者与非英语母语作者之间的字数差距缩小，而非英语母语作者减少了团队规模，并减少了与

英语母语合作者的合作。前两项发现补充了有关词汇和文体趋同的现有证据；第三项此前尚未有研究记录。这一发现与如下解读一致：生成式语言工具可能正在重塑科学文本的产出方式，也在重塑合作结构的组织方式，并在一定程度上替代了以往发挥语言功能的合作关系。在所有指标中，效应最大的群体是原有差距最显著的群体，包括隶属非洲和亚洲机构的作者，以及基线稿件长度较短学科的作者；这与技术满足既有需求的情形相符。

（Lepori、Andersen 和 Donnay，2025年）认为，如果生成式人工智能改变了科学文本的写作方式，那么稿件长度、署名模式和参考文献等常见科学计量指标，可能就不再具有过去那样的确切含义。本文记录的变化与这一担忧相符，而且我们可能已经身处科学文本生产方式转变的过程中，这种转变或许会影响人们对标准出版指标的解读。如果确实如此，2022年之后出版指标的部分变化，可能反映的是科学文本生产方式的改变，而非科学研究内在实质的变化。

这种重组是否会给全球科学体系带来净收益，仍是一个悬而未决的问题。与语言相关的差异缩小，意味着更广泛的获取机会；跨语言合作减少，则意味着更趋封闭。这两种趋势未必相互矛盾。对语言中介的依赖降低，长此以往或许能促成以科学而非语言互补性为基础的新合作。不过，目前的数据仅捕捉到这一转变的初始阶段，其长期影响仍有待持续进行纵向考察。

作者贡献

概念化：Y. Ben-Zion；方法：Y. Ben-Zion；正式分析与调查：Y. Ben-Zion、E. Cohen；撰写——初稿准备：Y. Ben-Zion、E. Cohen、N. Davidovitch；撰写——审阅与编辑：Y. Ben-Zion、E. Cohen、N. Davidovitch；监督：Y. Ben-Zion。

资金

作者声明，在撰写本稿件期间未获得任何资金、资助或其他支持。

利益冲突

作者没有需要披露的相关财务或非财务利益。

数据和材料的可用性

支持本研究结果的数据集已在 Zenodo 上公开发布，网址为 https://doi.org/10.5281/zenodo.18790135

参考文献

Abramo, G., D'Angelo, C.A., Di Costa, F.（2022年）。新冠肺炎危机如何塑造科研合作行为。《科学计量学》，*127*（8），5053–5071,

Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., ... 等人（2023年）。在科学领域，非英语母语者所付出的多重代价。《公共科学图书馆·生物学》，*21*(7)，e3002184,

Aydin, A., Yürür, S.E., Reisoğlu, I., Goktas, Y.（2023年）。学者发表论文时面临的主要障碍及可能的促进因素。《科学计量学》，*128*(1)，623–650,

Bao, T.,Zhao, Y.,Mao, J., Zhang, C.（2025年）。考察 ChatGPT 推出前后学术写作中的语言变化：一项关于预印本论文的研究。《科学计量学》，*130*(7)，3597–3627,

Ben-Zion, Y., Michaeli, O., Finkelstein, N.D.（2026）。从搜索到生成式人工智能查询：全球物理学信息检索趋势的主题与区域差异。*arXiv*预印本 *arXiv:2602.09550*，，

Cai, X., Fry, C.V., Wagner, C.S.（2021）。新冠疫情期间的国际合作：2020年秋季的进展。《科学计量学》，*126*(4)，3683–3692,

Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., . . . 等（2024）。ChatGPT 和 Bard 在系统综述中的幻觉率与参考文献准确性：比较分析。《医学互联网研究杂志》,*26*(1), e53164, Ding, L., Lawson, C., Shapira, P.（2025）。生成式人工智能在科学领域的兴起。《科学计量学》, *130*(9), 5093–5114, Filimonovic, D., Rutzer, C., Macher, J., Weder, R.（2025）。生成式人工智能作为全球科学中的语言均衡器。*arXiv*预印本 *arXiv:2511.11687*, , Geng, H., Wu, Y., Chen, S.X.（2025）。中美政治紧张关系对中国国际与国内科研合作的挤出效应：“中国行动计划”的证据。《科学计量学》, *130*(11), 6191–6226,

Kourilova-Urbanczik, M.（2012）。影响非母语者以英语进行科学报道的一些语言和语用学考量。跨学科毒理学, *5*(2), 105, Larivi`ere, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., Sugimoto, C.R.（2016）。知识生产中的贡献署名与劳动分工。科学社会研究, *46*(3), 417–435, Larivi`ere, V., Gingras, Y., Sugimoto, C.R., Tsou, A.（2015）。团队规模很重要：1900年以来的合作与科学影响力。美国信息科学与技术学会会刊, *66*(7), 1323–1332, Lenharo, M.（2023）。科学领域语言障碍给非英语母语者带来的真正代价。《自然》, *619*(7971), 678–679, Lepori, B., Andersen, J.P., Donnay, K.（2025）。生成式人工智能与科学计量学的未来：当前议题与未来问题。*arXiv*预印本 *arXiv:2507.00783*, , Liang, W., Zhang, Y., Wu, Z., Lepp, H., Ji, W., Zhao, X., . . . 等（2024）。描绘大语言模型在科学论文中日益增长的使用情况。*arXiv*预印本 *arXiv:2404.01268*, , Lillis, T., & Curry, M.J.（2006）。多语学者的专业学术写作：在英语文本生产过程中与读写中介的交互。《书面传播》, *23*(1), 3–35, Lin, D., Zhao, N., Tian, D., Li, J.（2025）。ChatGPT能否成为语言均衡器？量化大语言模型驱动的学术写作词汇变化。*arXiv*预印本 *arXiv:2504.12317*, , Liu, J., He, Y., Zheng, Z., Bu, Y., Ni, C.（2025）。非英语国家及资历较浅科学家的人工智能辅助写作增长最快。*arXiv*预印本 *arXiv:2511.15872*, , Liu, W.（2017）。非英语论文在学术交流中的角色变化：来自科学网三大期刊引文索引的证据。《学术出版》, *30*(2), 115–123,

Moed, H.F., Gl¨anzel, W., Schmoch, U. (2004). 定量科学与技术研究手册 （第10卷）。Springer。Ni, R., & Wang, J. (2025年)。应对冲击：疫情对香港科学合作与研究新颖性的影响。《信息计量学杂志》，*19*(2)，101656，Noy, S., & Zhang, W. (2023年)。生成式人工智能对生产力影响的实验证据。《科学》，*381*(6654)，187–192,

Ram´ırez-Casta˜neda, V. (2020年)。英语在科学领域占据主导地位所导致的科学论文准备与发表方面的不利影响：生物科学领域哥伦比亚研究人员的案例。*PLOS ONE*，*15*(9)，e0238372,

Tscharntke, T., Hochberg, M.E., Rand, T.A., Resh, V.H., Krauss, J. (2007)。多作者出版物中的作者顺序与贡献署名。《公共科学图书馆·生物学》，*5*(1)，e18,

Velez-Estevez, A., Garc´ıa-S´anchez, P., Moral-Mu˜noz, J.A., Cobo, M.J. (2022年). 为什么国际合作论文获得更多引用？图书馆与信息科学论文的文献计量分析。《科学计量学》, *127*(12), 7517–7555,

Wang, C., Jesiek, B., Zhang, W. (2024年). 通过战略性科研资助提升国际合作与学术成果：国家留学基金管理委员会资助出版物的文献计量分析。《科学计量学》，*129*(7), 4329– 4351,

附录 A：分类提示词

这两个提示词经过反复试错设计而成，以适应 PLOS ONE XML记录的特定结构，尤其是该期刊独有的基于方括号的隶属关系索引系统。我们对每个提示词进行了测试和完善，直到它能够生成准确且一致的分类结果。所有模型输出均经过人工检查；本文报告的最终提示词表述，是生成分析所用分类结果的版本。文章分批提交给模型处理，这是此类大规模分类任务的标准做法。

7.1 A.1 英语母语者分类

以下提示词已提交至 GPT-5 mini API。

你将收到多篇文章（最多 150 篇）。请针对每篇文章，分别分析其作者单位。

对于每篇文章的作者单位列表：- 每个列表之间用 | 分隔。- 每个列表中的作者以逗号 (,) 分隔。- 每位作者姓名后面的方括号 [] 中标有机构单位代码。- 同一代码也会出现在对应机构单位前面。

重要：机构单位代码在每篇文章中都是唯一的。不同文章可能使用相同的代码（例如 [aff001]），但这些代码指代的机构不同！

请针对每篇文章回答以下 3 个问题：问题 1：第一作者是否来自英语母语者占多数的国家？问题 2：最后一位作者是否来自此类国家？问题 3：是否至少有一位作者来自此类国家？

英语母语者占多数的国家：美国、英国、爱尔兰、澳大利亚、新西兰、加拿大。

重要提示：如果某位作者有多个所属机构（例如，[aff001、aff002]），请全部检查。只要其中至少有一个来自符合条件的国家，答案即为是（1）。

示例：文章 1：作者：Michelle Da Silva Lodge [aff001] |Nick Pullen [aff002] |Miguel Pereira [aff003] |Timothy S. Johnson [aff002]所属机构：[aff001] 学术肾脏病学科，谢菲尔德大学，英国 | [aff002]辉瑞全球研发，马萨诸塞州剑桥，美国 | [aff003]优时比制药，斯劳，英国 第一作者：Michelle [aff001] →英国 → 1末位作者：Timothy [aff002] →美国 → 1至少一位：是 → 1输出：1 1 1

输出格式：仅返回具有以下结构的 JSON 对象："results":{"article": 1, "analysis": "1 1 1"} 请勿包含任何其他文本、解释或 markdown 格式。仅返回该 JSON 对象。

以下是要分析的文章：

7.2 A.2 地理（洲别）分类

以下提示已提交至 GPT-5 mini API。

你将收到多篇文章（最多 150 篇）。请对每篇文章的作者单位分别进行分析。

对于每篇文章的单位列表：- 每个单位列表之间以 |分隔。- 每个列表中的作者以逗号 (,) 分隔。- 每位作者姓名后面的方括号 []中会显示机构单位代码。- 相同的代码也会出现在对应机构单位名称之前。

重要提示：机构单位代码在每篇文章中都是唯一的。不同文章可能使用相同的代码（例如， [aff001]），但它们指代的是不同的机构！

对于每篇文章，请回答以下问题：问题 1：第一作者来自哪个洲？问题 2：最后一位作者来自哪个洲？问题 3：每个洲是否至少有一位作者来自该洲？请按以下确切顺序，用 5 个二进制数字（0 或 1）作答：第 1 位：是否至少有一位作者来自亚洲？（1=是， 0=否）第 2 位：是否至少有一位作者来自欧洲？（1=是， 0=否）第 3 位：是否至少有一位作者来自非洲？（1=是， 0=否）第 4 位：是否至少有一位作者来自美洲？（1=是， 0=否）第 5 位：是否至少有一位作者来自其他地区？（1=是， 0=否）

洲代码：
1 = 亚洲（中国、日本、印度、以色列、韩国、新加坡、泰国、马来西亚、印度尼西亚等）
2 = 欧洲（英国、德国、法国、西班牙、意大利、荷兰、瑞士、瑞典、挪威、波兰等）
3 = 非洲（南非、埃及、尼日利亚、肯尼亚、埃塞俄比亚、摩洛哥等）4 = 美洲（美国、加拿大、巴西、墨西哥、阿根廷、智利、哥伦比亚等——整个北美洲和南美洲）5 = 其他（澳大利亚、新西兰以及任何其他地区）0 = 未知或无法确定

重要提示：如果一位作者有多个所属机构（例如， [aff001, aff002]），请检查所有这些机构。如果所属机构来自不同大洲，则使用最先提到的洲。

示例：
文章 1：作者：Michelle Da Silva Lodge [aff001] | Nick Pullen [aff002] | Miguel Pereira [aff003] |Timothy S. Johnson [aff002]隶属机构： [aff001]学术肾脏病学科，谢菲尔德大学，英国 | [aff002] 辉瑞全球研发，马萨诸塞州剑桥，美国 | [aff003] UCBPharma，

斯劳，英国 第一作者：Michelle [aff001] → 英国 → 欧洲 → 2末位作者：Timothy [aff002] → 美国 → 美洲 → 4各洲分布：亚洲=0，欧洲=1(Michelle、Miguel)，非洲=0，美洲=1 (Nick、Timothy)，其他=0 → 01010输出：2 4 01010

文章 2：作者：Li Wang [aff001] |Sarah Cohen [aff002]所属机构： [aff001] 北京大学，北京，中国 | [aff002] 特拉维夫大学，以色列 第一作者：Li [aff001] → 中国 → 亚洲 → 1 末位作者：Sarah [aff002] → 以色列 → 亚洲 → 1 各洲作者分布：亚洲=1（两位作者均在此洲），欧洲=0，非洲=0，美洲=0，其他=0 → 10000输出：1 1 10000

输出格式：仅返回一个 JSON 对象，结构如下：结果:{"article": 1, "analysis": "2 4 01010"}{"article": 2, "analysis": "1 1 10000"}重要提示：问题 3 必须始终包含恰好 5 个二进制数字（0 或 1），依次表示 [亚洲] [欧洲] [非洲] [美洲] [其他]。不要包含任何其他文字或说明。

以下是待分析的文章：