简化医疗报告并改善医患沟通的全新人工智能工具

马德里卡洛斯三世大学（UC3M）的一个研究团队评估了多种人工智能（AI）语言模型，旨在简化医疗报告，将其转化为针对非专业公众的易懂语言。这项研究将帮助患者更好地理解自身的健康状况，并准确遵循医嘱进行治疗。

该研究已发表于《Scientific Reports》期刊，切入点是当今社会普遍存在的一个难题：许多人难以理解自己的医疗报告，从而影响了其自我管理健康的能力。研究合著作者、UC3M计算机科学系教授帕洛玛·马丁内斯（Paloma Martínez）解释道：“我们观察到 46% 的欧洲公民在理解医疗文件方面存在困难。因此，我们开发了一套有助于改善医患沟通的工具。为解决这一问题，语言模型可以生成两个版本的文档：一个供医生使用，另一个则专门面向患者，使其能够理解报告中出现的缩写或复杂句式。”

研究团队指出，这项研究成果的发布非常及时。2027年1月2日，皇家法令 Real Decreto 707/2026 将正式生效。该法令于9月3日在官方公报（BOE）上公布，批准了首个认知无障碍基本条件法规，规定了提高公众相关文件、手续和服务可读性的要求，并将影响司法、医疗、就业、数字服务和交通等领域的公共管理部门。

本研究的另一位合著作者、同样担任UC3M计算机科学系教授的鲁尔德斯·莫雷诺（Lourdes Moreno）表示：“‘通俗语言’（Plain language）是一个标准化的方法论框架，旨在确保信息易于检索、理解和使用。它不仅意味着缩短句子或简化词汇，更在于以连贯的方式组织内容、调整专业词汇，并根据用户的具体信息需求提供明确引导。”

研究人员分析了由安达卢西亚家庭与社区医学学会提供的 50 份临床病例医疗报告。每份报告均由认知无障碍专家根据通俗语言国际标准 ISO 24495-1:2023 进行三份独立的改编，共计 150 篇改编文本。这些原始文本及其改编版被用于模型的微调与评估。

研究人员评估了四个语言模型系列（Llama 3.2 1B 和 3B、Salamandra 2B 和 7B、RigoChat-v2 7B 以及 Medical-mT5 3B）在简化程度、语义相似度、事实一致性、可读性以及环境影响方面的表现。结果表明不存在完美的单一模型：例如，Llama-3.2-3B 在效率、鲁棒性和综合性能方面表现最为均衡；而当优先考虑输出质量和可读性时，RigoChat-v2-7B 则表现突出。

对生成文本的分析揭示了所有四个模型系列中普遍存在的共性错误：一是过度简化（删除患者需要了解的临床细节）；二是临床幻觉（捏造或篡改原始报告中的数据）；三是不连贯性（出现断句或跳跃，增加阅读难度）。帕洛玛·马丁内斯强调：“正是因为存在这类错误，在医疗健康这样敏感的领域中，人脑监督仍然不可或缺。”

研究团队已将所有材料向科学界开放，以推动可重复性研究。“我们的模型完全开源。我们所开发的一切——包括数据集、微调后的模型和提示词（prompts）——均可在 GitHub 和 Zenodo 等公共存储库中获取，以便同行进一步扩展我们的研究，”帕洛玛·马丁内斯总结道。

该项目是在 HumanAI（*Human-Centered AI: Towards Accessible and Personalized Language Models*）框架下开展的，属于西班牙国家科学、技术与创新研究计划（PEICTI）项目，由西班牙国家研究署提供资助（PID2023-148577OB-C21）。更多信息请访问：https://humanai.uc3m.es/

参考文献:

Martínez, P., Sánchez-Gómez, J. M., & Moreno, L. (2026). Benchmark of small language models for plain-language simplification in Spanish clinical texts. *Scientific Reports*. DOI: https://doi.org/10.1038/s41598-026-65740-w

e-Archivo de la UC3M: https://hdl.handle.net/10016/51142

视频: https://youtu.be/cF0DYHY07KU



关于人工智能如何改善医患沟通的图片。图片来自Cottonbro，来源：Pexels.