
美国斯坦福大学发布《2026年AI指数报告》

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/39375.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

美国斯坦福大学发布《2026年AI指数报告》

。近日，美国斯坦福大学以人为本人工智能研究所发布了《2026年AI指数报告》。这份423页的年度报告对人工智能（AI）在能力、投资、就业、应用等方面进行了系统梳理。

中美差距几乎消失

最引人关注的发现之一，是中美两国在AI模型性能上的差距已几乎消失。

报告指出，自2025年初以来，两国模型已多次交替登顶性能榜单。过去一年来，这一差距虽有波动，但始终保持在个位数。2025年2月，中国“深度求索”公司的DeepSeek-R1曾短暂追平美国最佳模型。

截至2026年3月，美国顶尖模型Claude Opus 4.6的Elo评分为1503，而中国顶尖模型紧追其后，差距仅有2.7%。

报告指出，美国在基础模型创新、资本投入和算力基础设施上仍保持领先。美国拥有5427个数据中心，是其他国家的10倍以上。与之同步的是，其能源消耗量也位居世界第一。

相比之下，中国在科研产出与产业应用方面表现更为突出。报告显示，中国在论文发表量、引用次数以及专利产出方面均处于领先地位，在工业机器人（物理AI领域）部署数量上也位居世界前列，占全球装机量的54%。越南媒体vietnam.vn网站指出，这反映出中国AI技术在制造业领域的快速发展，也体现了中国在AI领域的长期投资战略以及政府、企业和科研机构之间的紧密合作。

美国AI人才吸引力下降

长期以来，美国是全球AI人才的“引力中心”，但这一优势正在快速消退。

报告指出，自2017年以来，移居美国的AI学者数量下降了89%，且仅在过去一年中就锐减了80%。

尽管美国仍是全球AI研究人员最多的国家，但支撑这一优势的“人才流入”正在枯竭。对于那些习惯于顶尖研究者自然流向美国公司的企业来说，这一假设已不再稳固。

与此同时，全球AI竞赛已不再是部分大国之间的“擂台赛”，韩国已成为全球“创新密度”最高

的国家，人均专利申请量领跑世界。

随着各国竞相布局，“主权AI”已成为许多政府的首要政策考量。过去一年，一些欧洲和中亚国家对其AI基础设施进行了大量投资，使得拥有“国家支持的超级计算集群”的国家达到44个。但报告警告，南美和中东国家在AI基础设施上落后，可能引发新的“数字鸿沟”。

AI“偏科”严重

报告指出，生成式AI在3年内达到了全球53%的人口使用率，普及速度快于个人电脑和互联网。然而，AI能力的进步并非均衡发展。

谷歌的Gemini Deep Think模型在2025年国际数学奥赛中，以35分（满分42分）的成绩夺得金牌，能在4.5小时内用自然语言完成端到端的解题。然而，在专门测试时钟识别的ClockBench上，顶尖模型正确读取指针时钟的概率仅为50.1%，而人类为90.1%。

“AI能赢得数学奥赛金牌，却仍然无法可靠地读取时间。”报告用“锯齿状智能”来形容这种能力分布的不均衡。

在其他领域，进步同样显著。AI智能体处理现实世界计算机任务的成功率，从18个月前的12%跃升至2026年3月的66%，距离人类表现仅差6个百分点。在网络安全任务上，AI智能体解决问题成功率从2024年的15%飙升至93%。

然而，从数字世界跨入物理世界，AI的能力急剧衰减。机器人在软件模拟环境中的成功率可达89.4%，但在真实的家务任务（如叠衣服、洗碗）中，成功率骤降至12.4%。可见，整洁的实验室与凌乱的家居环境之间，隔着一道难以逾越的“鸿沟”。

AI对就业的影响显现

AI对就业的影响正在显现。报告发现，生产率提升明显的领域，往往也是初级岗位减少的领域。

在美国，22—25岁的软件开发人员就业率自2024年以来下降了近20%，而年长从业者数量增加。类似情况也出现在客户服务领域。企业调查显示，计划中的裁员规模已超过近期实际裁减幅度。

然而，当就业市场对年轻人的“大门”开始收窄时，教育体系却尚未教会他们如何与AI共处。报告发现，4/5的美国高中生和大学生已经在使用AI完成学业任务，主要用于研究、论文编辑和头脑风暴，但只有一半的中小学制定了AI使用政策，仅6%的教师表示这些政策是清晰的。

全球范围内，超过90%的国家已提供计算机科学课程，但AI教育推进较慢。中国和阿联酋已开始将AI纳入必修课程。

最强AI模型透明度下降

当AI能力越来越强，我们对它的了解却越来越少。报告中的“基础模型透明度指数”平均得分，从上一年度的58分骤降至40分。IBM以95分领跑，而xAI的Grok和Midjourney仅得14分。

能力最强的模型往往披露的信息最少。超过90%的知名AI模型由私营公司创造，但许多公司不再

披露训练数据集大小、参数数量或训练时长。去年发布的95个知名模型中，有80个未公开其训练代码。

报告还披露了AI的环境成本。xAI公司训练Grok 4模型估计产生了72816吨二氧化碳当量，相当于17000辆汽车一年的排放量。AI数据中心的电力容量已达29.6吉瓦，接近纽约州峰值用电需求。仅GPT-4o的推理用水量，就可能超过1200万人的饮用水需求。

作者：张佳欣 来源：科技日报

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发