

人造的 智力 2024 年指数报告



Stanford University
Human-Centered
Artificial Intelligence

简介

2024 年人工智能指数报告

欢迎阅读第七版人工智能指数报告。2024 年指数是我们迄今为止最全面的指数,恰逢人工智能对社会的影响力达到前所未有的重要时刻。今年,我们扩大了研究范围,更广泛地涵盖人工智能的技术进步、公众对该技术的看法以及围绕其发展的地缘政治动态等基本趋势。

该版本包含比以往更多的原始数据,介绍了对人工智能培训成本的新估计、对负责任的人工智能前景的详细分析,以及专门讨论人工智能对人类的影响的全新章节。
科学和医学。

AI 指数报告跟踪、整理、提炼和可视化与人工智能 (AI) 相关的数据。我们的使命是提供公正、经过严格审查、来源广泛的数据,以便政策制定者、研究人员、高管、记者和公众对复杂的人工智能领域有更全面、更细致的了解。

人工智能指数被全球公认为人工智能数据和见解最可信、最权威的来源之一。先前版本已被《纽约时报》、彭博社和《卫报》等主要报纸引用,已积累了数百次学术引用,并被美国、英国和欧盟的高层政策制定者引用等地方。今年的展会在规模、规模和范围上都超越了以往的历届展会,反映出人工智能在我们所有人的生活中日益重要的意义。

联合董事致辞

十年前,世界上最好的人工智能系统无法以人类的水平对图像中的物体进行分类。人工智能在语言理解方面存在困难,并且无法解决数学问题。如今,人工智能系统在标准基准上的表现通常超过人类。

到 2023 年,进展将会加速。GPT-4、Gemini 和 Claude 3 等最先进的新系统具有令人印象深刻的多模式功能:它们可以生成数十种语言的流畅文本、处理音频,甚至解释模因。随着人工智能的进步,它越来越多地进入我们的生活。各家公司都在竞相打造基于人工智能的产品,而人工智能也越来越多地被大众所使用。但当前的人工智能技术仍然存在重大问题。它无法可靠地处理事实、执行复杂的推理或解释其结论。

人工智能面临两个相互关联的未来。首先,技术不断改进并得到越来越多的应用,对生产力和就业产生重大影响。它可以有好的用途,也可以有坏的用途。在第二个未来,人工智能的采用受到技术局限性的限制。无论未来如何发展,各国政府都越来越担心。他们正在介入以鼓励积极的一面,例如资助大学研发和激励私人投资。各国政府还致力于管理潜在的负面影响,例如对就业的影响、隐私问题、错误信息和知识产权。

随着人工智能的快速发展,人工智能指数旨在帮助人工智能社区、政策制定者、商界领袖、记者和公众应对这一复杂的局面。它提供了持续、客观的快照,跟踪几个关键领域:人工智能能力的技术进步、推动人工智能开发和部署的社区和投资、对当前和未来潜在影响的公众舆论,以及为刺激人工智能创新同时管理其风险和挑战而采取的政策措施。通过全面监测人工智能生态系统,该指数成为了解这一变革性技术力量的重要资源。

在技术方面,今年的人工智能指数报告称,2023 年全球发布的新大型语言模型数量比上一年翻了一番。三分之二是开源的,但性能最高的模型来自具有封闭系统的行业参与者。Gemini Ultra 成为第一个在大规模多任务语言理解 (MMLU) 基准测试中达到人类水平表现的法学硕士;自去年以来,基准性能提高了 15 个百分点。此外,GPT-4 在综合语言模型整体评估 (HELM) 基准 (包括 MMLU 等评估) 中取得了令人印象深刻的 0.96 平均胜率分数。

消息来自 联合董事（续）

尽管全球人工智能私人投资连续第二年下降,但生成式人工智能投资却猛增。财富 500 强企业的财报电话会议中提到人工智能的次数比以往任何时候都多,新的研究表明人工智能切实提高了工人的生产力。在政策制定方面,全球立法程序中对人工智能的提及从未如此频繁。美国监管机构在 2023 年通过的人工智能相关法规比以往任何时候都多。尽管如此,许多人仍对人工智能产生深度伪造品和影响选举的能力表示担忧。公众对人工智能的了解越来越多,研究表明他们的反应很紧张。

雷·佩罗和杰克·克拉克
AI 指数联席董事

十大要点

1. 人工智能在某些任务上击败了人类,但并非在所有任务上都击败了人类。人工智能在多个基准上已经超越了人类的表现,包括图像分类、视觉推理和英语理解方面的一些基准。然而,它在竞赛级数学、视觉常识推理和规划等更复杂的任务上落后了。

2. 工业界继续主导人工智能前沿研究。2023年,工业界产生了51个值得注意的机器学习模型,而学术界仅贡献了15个。2023年,产学研合作产生了21个值得注意的模型,再创新高。

3. 前沿型号变得更加昂贵。根据AI Index估计,最先进的AI模型的训练成本已达到前所未有的水平。例如,OpenAI 的 GPT-4 估计使用了价值 7800 万美元的计算来进行训练,而 Google 的 Gemini Ultra 的计算成本为 1.91 亿美元。

4. 美国领先中国、欧盟和英国,成为顶级人工智能模型的主要来源。2023 年,61 个著名人工智能模型源自美国机构,远远超过欧洲联盟21个,中国15个。

5. 严重缺乏对LLM责任的稳健和标准化的评估。

人工智能指数的新研究表明,负责任的人工智能报告严重缺乏标准化。

包括 OpenAI、Google 和 Anthropic 在内的领先开发人员主要根据不同的负责任的 AI 基准测试他们的模型。这种做法使得系统地比较顶级人工智能模型的风险和局限性的工作变得更加复杂。

6. 生成式人工智能投资猛增。尽管去年人工智能私人投资总额有所下降,但生成式人工智能的资金却大幅增长,比 2022 年增长了近八倍,达到 252 亿美元。生成式人工智能领域的主要参与者,包括 OpenAI、Anthropic、Hugging Face 和 Inflection,都报告称已进行了大量融资。

7. 数据显示:人工智能提高了工人的生产力并带来更高质量的工作。2023 年,多项研究评估了人工智能对劳动力的影响,表明人工智能使工人能够更快地完成任务并提高产出质量。这些研究还证明了人工智能有潜力弥合低技能工人和高技能工人之间的技能差距。尽管如此,其他研究警告说,在没有适当监督的情况下使用人工智能可能会导致性能下降。

十大要点（续）

8. 得益于人工智能,科学进步进一步加速。2022年,人工智能开始推动科学发现。然而,2023 年将会推出更重要的与科学相关的人工智能应用程序

从让算法排序更加高效的 AlphaDev,到促进材料发现过程的 GNoME。

9. 美国人工智能法规数量急剧增加。美国人工智能相关法规的数量在过去一年和过去五年中显着增加。2023年,人工智能相关法规数量达到25个,而2016年仅有1个。仅去年一年,人工智能相关法规总数就增长了56.3%。

10. 全球各地的人们更加认识到人工智能的潜在影响,也更加紧张。

益普索 (Ipsos)的一项调查显示,去年,认为人工智能将在未来三到五年内极大影响他们生活的人比例从 60% 上升到 66%。此外,52% 的人表示对人工智能产品和服务感到紧张,比 2022 年上升了 13 个百分点。在美国,皮尤研究中心的数据显示,52% 的美国人表示对人工智能感到更担心,而不是兴奋,这一比例高于 2022 年的 37%。