人工智能代理现状

我们调查了 1,300 多名专业人士（从工程师和产品经理到企业领导和高管），以了解 AI 代理的现状。深入研究数据，我们将分析 AI 代理目前是如何被使用的（或未被使用的）。

介绍

2024 年，人工智能代理不再只是小众产品。各行各业的公司都越来越重视将代理纳入其工作流程 - 从自动化日常任务到协助数据分析或编写代码。

但幕后到底发生了什么？人工智能代理是否发挥了其潜力，还是只是另一个流行词？谁在部署它们，又是什么阻止了其他人一头扎进去？

我们调查了 1,300 多名专业人士，以了解 2024 年人工智能代理的状况。让我们深入了解以下数据。

洞察

首先，代理人是什么？

在 LangChain，我们将代理定义为使用 LLM 决定应用程序控制流的系统。就像自动驾驶汽车的自主性级别一样，代理能力也存在一系列范围。

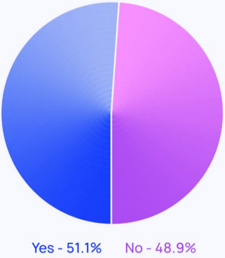
采用代理就像掷硬币一样——但几乎每个人都有计划

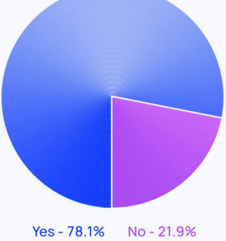
代理竞赛正在升温。在过去的一年里，许多代理框架都获得了极大的欢迎——无论是使用 ReAct 来结合 LLM 推理和表演，还是多代理编排器，或者是像LangGraph这样更易于控制的框架。

关于代理的讨论并非全都是 Twitter 上的炒作。大约**51% 的受访者**目前在生产中使用代理。当我们按公司规模查看数据时，中型公司（100-2000 名员工）在将代理投入生产方面最为积极（占 63%）。
令人鼓舞的
是，**78% 的公司有积极的计划**很快将代理投入生产。虽然很明显人们对 AI 代理的需求很强烈，但实际的生产部署对许多人来说仍然是一个障碍。

贵公司目前有
代理商在生产吗？

您目前是否正在开发代理并计划将其投入生产？





LangChain State of AI Agents - 2024

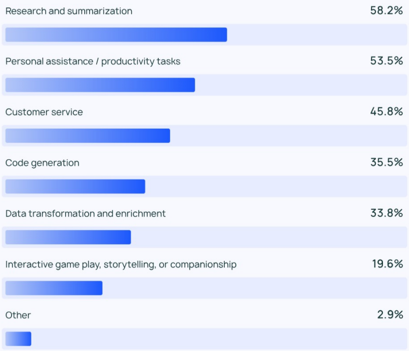
我们还看到，公司正在从简单的基于聊天的实现转向更先进的框架，这些框架强调多代理协作和更自主的功能。（请参阅下面的"新兴主题"部分了解更多信息。）

虽然科技行业通常被认为是早期采用者，但对代理的兴趣正在所有行业中日益增长。在非科技公司工作的受访者中，90%已经或正在计划将代理投入生产（几乎相当于科技公司，为89%）。

领先代理用例

人们使用代理来做什么？代理既处理日常任务，也为知识工作打开了新的可能性。代理的主要**进行研究和总结（58%）**，其次是简化任务以提高**个人工作效率或提供协助（53.5%）。这些都**表明人们希望有人（或某物）为他们处理耗时的任务。用户不必筛选无尽的数据进行文献综述或研究分析，而是可以依靠人工智能代理从大量信息中提炼关键见解。同样，人工智能代理通过协助安排和组织等日常任务来提高个人工作效率，让用户可以专注于重要的事情。效率的提高不仅限于个人。客户服务（45.8%）是代理用例的另一个主要领域，帮助公司处理查询、排除故障并加快跨团队的客户响应时间。

您认为当今代理商最适合执行哪些任务？

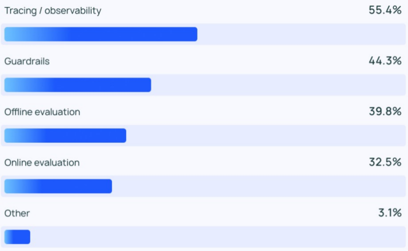


LangChain State of AI Agents - 2024

谨慎行事：需要追踪和人工监督来控制特工

能力越大，责任越大——或者至少需要为代理设置一些制约和控制。**跟踪和可观察性工具**位居必备控制列表之首，可帮助开发人员了解代理的行为和绩效。大多数公司还采用护栏来防止代理偏离航向。

您对代理商采取了哪些控制措施?



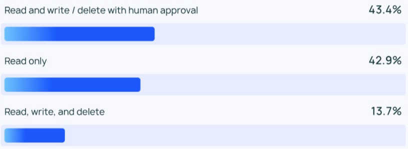
LangChain State of AI Agents - 2024

在测试LLM申请时，人们更多地提到**离线评估（39.8%）**而不是在线评估（32.5%）。这可能说明了监控实时性能的难度。在填写的回复中，许多公司还请了人工专家手动检查或评估回复，以增加一层预防措施。

尽管

人们对代理表现出兴奋之情，但大多数人在允许代理自由发挥到何种程度方面采取了更为保守的态度。很少有受访者允许他们的代理自由读写和删除。相反，大多数团队要么允许**只读工具权限**，要么**要求人工批准更重要的操作**，例如写入或删除。

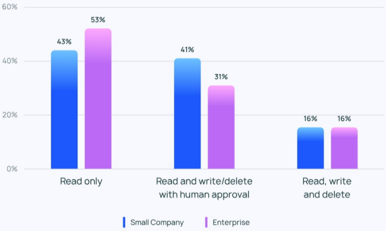
你的代理有什么样的工具权限?



LangChain State of AI Agents - 2024

在代理控制方面，不同规模的公司也会有不同的优先级。不出所料，大型企业（2000多名员工）更加谨慎，严重依赖“只读”权限来避免不必要的风险。他们还倾向于将护栏与**离线评估**相结合，以便在客户看到任何响应之前，在预生产中捕捉回归。

按公司规模划分的工具权限

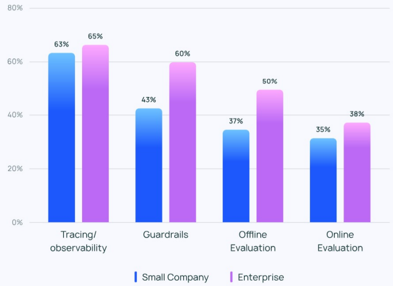


LangChain State of AI Agents - 2024

与此同时，小公司和初创公司（员工人数少于100人）更注重跟踪，以了解其代理应用程序中发生的事情（而不是其他控制

措施）。从我们的对话来看，小公司往往专注于运输，只需查看数据即可了解结果；而企业则全面实施了更多控制措施。

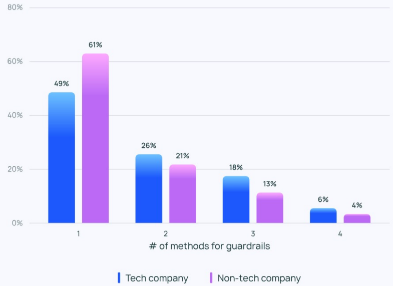
代理商按公司规模控制



LangChain State of AI Agents - 2024

虽然非科技公司和科技公司受访者的代理采用率相似，但在那些在生产中使用代理控制的受访者中，科技公司更有可能使用多种控制方法。51 **% 的科技公司受访者目前使用 2 种或更多种控制方法，而其他行业的受访者**只有39%。这表明科技公司在构建可靠代理方面可能走得更远，因为控制是高质量体验所必需的。

用于控制或护栏的方法数量



LangChain State of AI Agents - 2024

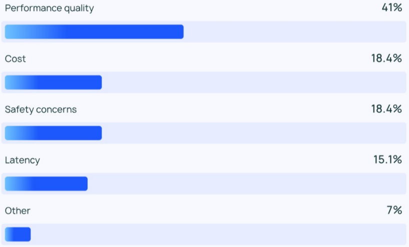
代理商投入生产的障碍和挑战

保持法学硕士 (LLM) 申请的高质量（无论答复是否准确或是否符合正确的风格）并不容易。

性能质量是受访者最关注的问题——其重要性是成本和安全等其他因素的两倍多。

代理使用 LLM 来控制工作流程的固有不可预测性带来了更多的错误空间，使得团队很难确保其代理始终提供准确、符合上下文的响应。

在生产中投入更多代理的最大限制是什么？

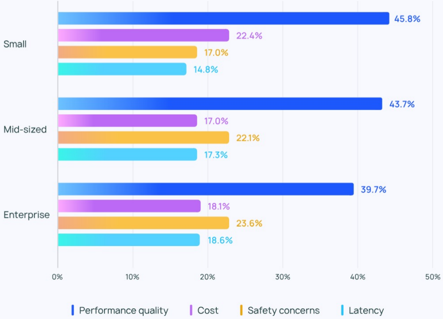


LangChain State of AI Agents - 2024

尤其是对于小公司来说，性能质量远远超过其他考虑因素，45.8% 的受访者将其列为首要考虑因素，而成本（第二大考虑因素）仅占 22.4%。这一差距凸显了可靠、高质量的性能对于组织将代理从开发转移到生产的重要性。

虽然质量仍然是企业的首要考虑因素，但安全问题对这些必须遵守法规并更谨慎地处理客户数据的大公司来说也很普遍。

部署代理的障碍（按公司规模划分）



LangChain State of AI Agents - 2024

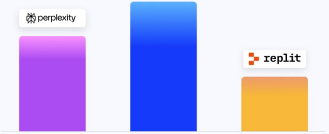
挑战并不止于质量。从书面回复来看，许多人对构建和测试代理的最佳实践感到不确定。特别是，两个主要障碍脱颖而出：**知识和时间。**

- **知识**：团队经常难以掌握使用代理所需的技术知识，包括针对特定用例实施它们。许多员工仍在学习诀窍，需要提高技能才能有效利用人工智能代理。
- **时间**：构建和部署所需的时间投入非常大，特别是在尝试确保代理可靠运行时——这可能需要调试、评估、微调等。

经纪人成功案例：Cursor 抢尽风头

最热门的人工智能代理应用程序

CURSOR



LangChain State of AI Agents - 2024

在我们的调查中，Cursor是讨论最多的代理应用程序，紧随其后的是**Perplexity**和**Replit**等重量级应用程序。Cursor **是**一款人工智能代码编辑器，可帮助开发人员通过智能自动完成和上下文帮助编写、调试和解析代码。Replit 还可以通过设置环境和配置来加速软件开发生命周期，并让您在几分钟内构建和部署功能齐全的应用程序。Perplexity 是一个人工智能答案引擎，可以使用网络搜索和链接源来回答复杂的查询。这些应用程序正在突破代理功能的界限，表明人工智能代理不再是理论上的——它们正在解决当今生产环境中的实际问题。

人工智能代理采用的新兴主题

从我们的书面回复中，我们看到组织在将 AI 代理引入其工作流程时面临着许多不断变化的期望和挑战。人们对 AI 代理的这些功能表示钦佩：

管理多步骤任务

代理具有更深入的推理和上下文管理能力，从而使其能够处理更复杂的任务。

自动执行重复任务

人工智能代理继续被视为自动化管理任务的关键，它可以让用户腾出时间进行更有创造性的问题解决。

任务路由和协作

更好的任务路由可确保正确的代理在正确的时间处理正确的问题——尤其是在多代理系统中。许多人想知道如何有效地协调任务并跨代理网络进行协作。

类似人类的推理

与传统的法学硕士不同，人工智能代理可以追溯他们的决策，包括时间旅行、审查和根据新信息修改过去的决策。

但对于构建代理的团队来说，也有一些挑战需要考虑。这包括：

- **理解代理行为的障碍。**几位工程师写信谈到了他们向公司其他利益相关者解释人工智能代理的能力和行为的困难。有时，一些额外的步骤可视化可以解释代理响应发生了什么。其他时候，LLM 仍然是一个黑匣子。可解释性的额外负担留给了工程团队。

尽管面临挑战，但以下领域仍然充满活力：

对开源 AI 代理的兴奋

人们对开源人工智能代理表现出明显的兴趣，许多人认为集体智慧可以加速代理的创新。

期待更强大的车型

许多人都在等待人工智能代理的下一次飞跃，即由更大、更强大的模型驱动，以便代理能够以更高的效率和自主性处理更复杂的任务

结论

整合人工智能代理的竞赛已经开始，因为公司已经开始重塑工作流程并设计未来，而法学硕士则负责改善决策和提高人类生产力。

然而，尽管兴奋之情高涨，但公司也意识到他们必须谨慎行事，播下正确的控制以驾驭新的用例和应用程序。团队既渴望又谨慎，他们尝试使用框架来尝试保持代理响应的高质量和无幻觉。

展望未来，能够破解可靠、可控代理代码的公司将在下一波人工智能创新中占据先机，并开始为智能自动化的未来制定标准。

方法论

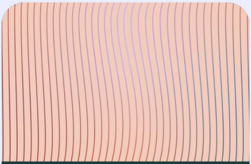
五大行业：

- 技术（60% 的受访者）
- 金融服务（占受访者的 11%）
- 医疗保健（占受访者的 6%）
- 教育（占受访者的 5%）
- 消费品（4%）

公司规模：

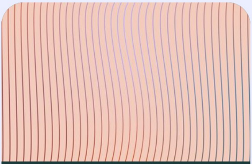
- <100 人（占受访者的 51%）
- 100-2000 人（占受访者的 22%）
- 2000-10,000 人（占受访者的 11%）
- 10,000 人以上（占受访者的 16%）

查看相关资源



LangGraph：构建可控代理

“In the Loop”：关于人工智能代理的博客系列



5 个特工突围故事

评估

检索

更新日志

LangSmith 信托门户

YouTube

社区

营销资产

LangChain

- 所有系统均正常运行

隐私政策

服务条款