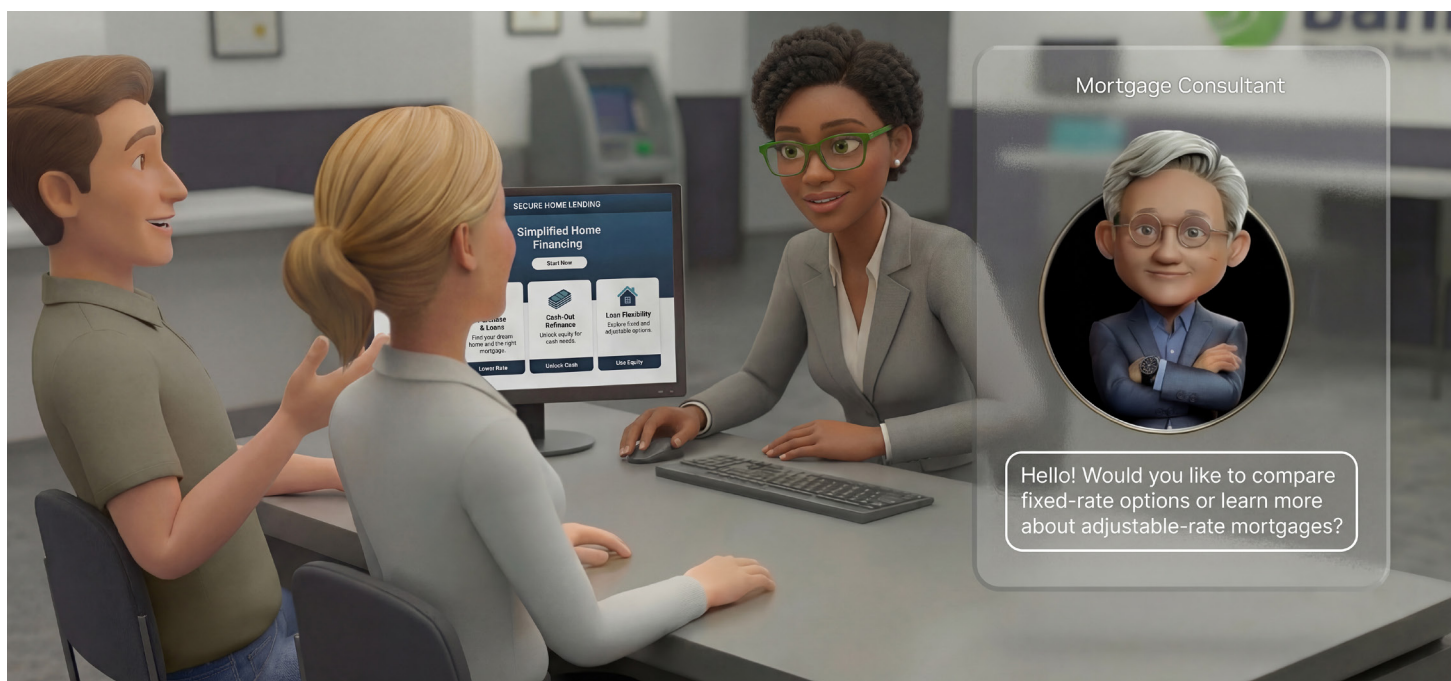
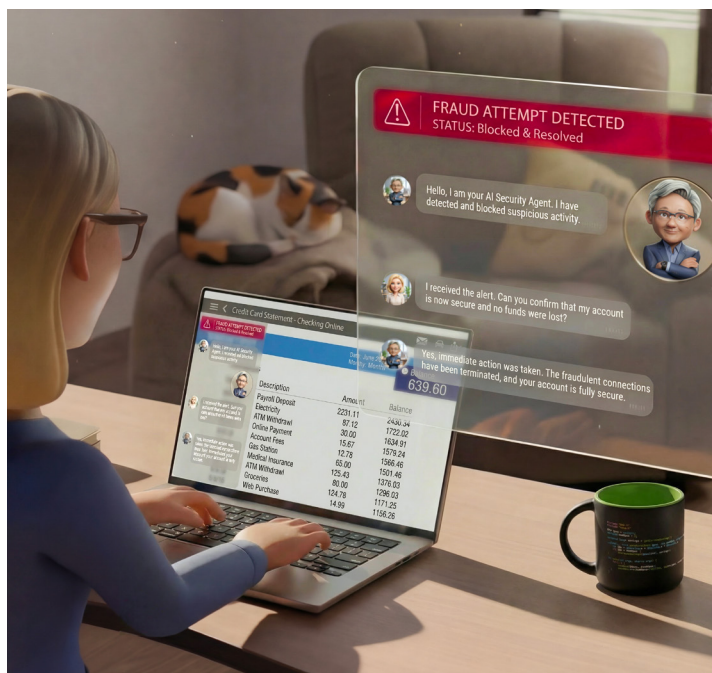


# 2026 年金融服务行业 AI 发展现状及趋势



# 目录

## 调查概述

|                          |   |
|--------------------------|---|
| 金融服务行业的 AI 现状：行业领导者的关键趋势 | 3 |
| 摘要                       | 4 |

## 深入解析结果

|                               |    |
|-------------------------------|----|
| 金融服务正在拥抱 AI 赋能的未来             | 6  |
| 行业正在拥抱开源                      | 7  |
| AI 智能体在金融服务领域初显成效             | 8  |
| AI 在运营中得到广泛应用                 | 9  |
| 运营效率和员工生产力是金融服务领域 AI 最重要的投资回报 | 10 |
| 金融机构计划在 2026 年增加 AI 投资        | 12 |
| 金融服务领域正在选择混合架构处理 AI 工作负载      | 13 |
| 展望未来                          | 13 |
| 调研方法                          | 14 |

# 调查概述

## 金融服务行业的 AI 现状：行业领导者的关键趋势

AI 革命才刚刚开始，但已经对金融机构产生了深远的影响。从改变算法交易到加速文档处理和分析，从使用基于 Transformer 的支付基础模型重塑欺诈检测，到利用编码智能体实现传统代码的现代化，AI 已成为推动行业发展的巨大助力，为增长、生产力提升和成本管理开辟了新的途径。

在第六份年度金融服务领域 AI 现状报告中，NVIDIA 探讨了 AI 对行业的影响、AI 如何改变行业的运作方式，以及各种规模的组织如何开始扩展新的 AI 能力。

生成式 AI 自 2022 年引入市场以来，在金融服务领域中的使用逐年增长，61% 的受访者表示在 2025 年使用或评估过这项技术。随着关注点转向投资回报明确的特定业务用例，例如代理式 AI，这一演进趋势仍在持续。

AI 得到如此快速采用和投资的原因显而易见：领先机构的高管已公开承认其可观的投资回报率，特别是在运营效率和员工生产力方面。毫不意外地，近 100% 的受访者确认他们在 2026 年的 AI 支出将增加或保持不变，其中 44% 的受访者表示将增加 10% 以上。

2026 年金融服务领域 AI 现状报告的受访者人数创下了该调查的历史新高。该报告探讨了 AI 在金融服务领域日益增长的应用情况、主要的 AI 用例、金融机构的目标和挑战、对解决方案和基础设施的投资计划、影响和投资回报、基础设施和发展趋势，并在今年新增了一个关于 AI 智能体采用和使用情况的章节。

## 摘要

现在正是投资金融服务领域 AI 的时机，73% 担任领导职务的受访者认为，AI 对其公司未来的成功至关重要。为什么？因为 AI 从数据中创造智能，而没有哪个行业比金融服务更擅长或更依赖将智能变现。

以下是今年报告中的其他一些主要亮点。

# 73%

担任领导职务的受访者认为，AI 对其公司未来的成功至关重要。

## AI 已从试点转向普及

# 65%

的受访者表示他们的组织正在积极使用 AI，这一比例较 2024 年的 45% 有所上升。

# 42%

的受访者表示他们的组织正在积极使用或评估代理式 AI。

AI 将大数据的强大能力与先进的机器学习技术 (从深度学习和强化学习到生成式模型和预测模型) 相结合，推动金融服务行业不断涌现各种各样的用例。积极使用 AI 的组织比以往任何时候都要多，约 90% 的受访者表示，他们的组织正在积极使用或评估 AI 解决方案或试点项目。代理式 AI 是下一个前沿，42% 的组织表示，他们已经在运营工作流程中使用或评估 AI 智能体。

## 开源的重要性

企业级 AI 需要针对专门用例进行调优的模型，而托管服务和通用大语言模型 (LLM) 无法处理这些用例。正因如此，各组织纷纷转向开源模型来构建针对特定用例进行微调的解决方案。总体而言，84% 的受访者认为开源软件的重要性为“中等”到“极其重要”，其中近一半 (48%) 的受访者表示开源软件为“非常重要”到“极其重要”。

大型组织，如银行和其他金融机构，正在将最重要的 AI 用例从托管服务转向开源基础模型。随着推理模型日趋先进以及每 Token 成本持续上升，通过自主拥有模型而非依赖第三方来降低持续成本的经济压力不断加剧。这种方式不仅能以更低的总成本提供更准确、更有用的结果，还能确保专有数据的安全，防止企业价值外流。

# 84%

的受访者表示，开源软件对其组织 AI 战略的重要性为“中等”到“极其重要”。

## AI 的投资回报率显而易见

**89%**

的受访者表示，AI 正在帮助提高年收入并降低年成本。

**52%**

的受访者认为提高运营效率是 AI 在过去一年中改善业务的主要方式之一。

AI 正在影响组织的营收和利润。83% 的受访者表示在 AI 用例中看到了投资回报，其中文档处理、文档管理和客户体验方面的解决方案尤为突出。此外，89% 的受访者报告称，AI 既增加了年收入，又降低了年成本。超过一半 (52%) 的受访者认为提高运营效率是 AI 在过去一年中改善业务的主要方式之一。AI 还提高了员工生产力，48% 的受访者表示这是他们业务获得的最大好处之一，这一比例高于去年的 22%。

## 数据分析是 AI 的核心驱动力

**68%**

的受访者表示，数据分析是其组织的首要 AI 工作负载，这比 2024 年上升了 11 个百分点。

**61%**

的受访者认为生成式 AI 是其首要 AI 工作负载之一，这一比例较 2024 年的 52% 有所上升。

数据分析的重要性涵盖了从金融科技 (fintech) 和资本市场到消费金融的整个金融服务领域。数据分析支持一系列业务流程和用例，从处理支付交易以进行欺诈检测，到分析结构化和非结构化数据以支持投资研究和投资组合管理。

增长最快的 AI 工作负载一直是生成式 AI。在中国以外地区，生成式 AI 的占比从 2023 年的 50% 增长到今年的 73%，巩固了基础模型作为推动全球金融服务和商业运营 AI 发展的关键引擎的地位。

## 混合架构显著增加

最近几次调查数据中的一个明显趋势是，组织越来越多地为其 AI 工作负载部署混合架构，而不是仅限于本地或云端。最主要的原因是经济因素，混合方式能够在不同的环境中运行不同的工作负载，从而优化成本。由于数据分散在不同位置，利用一个适用于所有环境的软件平台简化数据科学工作流程是关键。

**47%**

的受访者表示其组织使用了混合架构，这一比例较 2024 年的 26% 有所上升。

# 深入解析结果

## 金融服务正在拥抱 AI 赋能的未来

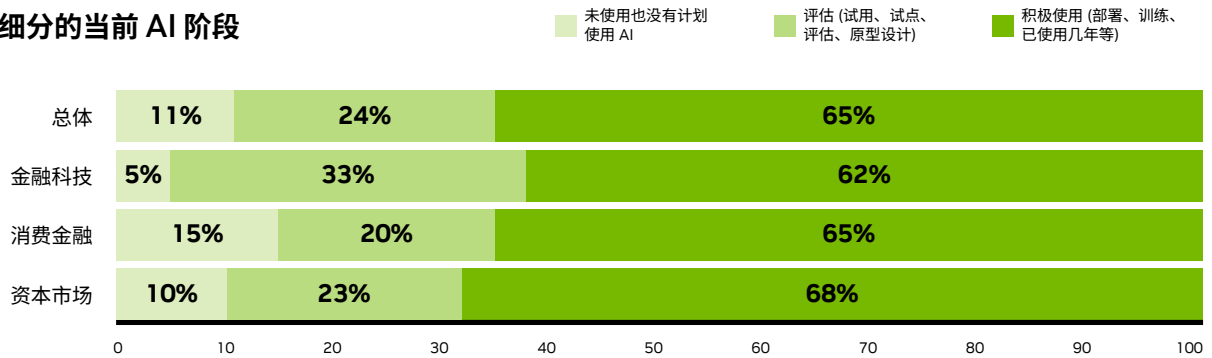
随着 AI 在商业生态系统中日趋成熟，各公司正在从评估阶段转向积极利用并扩展 AI，将其纳入关键业务运营。65% (近三分之二) 的受访者表示，他们的组织正在积极使用 AI，这一比例较 2024 年的 45% 有所上升。各公司正在积极将试点项目扩展到生产环境，24% 的公司表示今年将评估 AI 试点项目，较去年的 50% 有所下降。

特别是，在员工人数超过 1,000 的大型公司中，76% 的受访者表示正在积极使用 AI。金融服务领域内不同细分行业的采用情况相当均衡，超过 60% 的金融科技、消费金融和资本市场受访者表示，他们的组织正在积极使用 AI。

# 76%

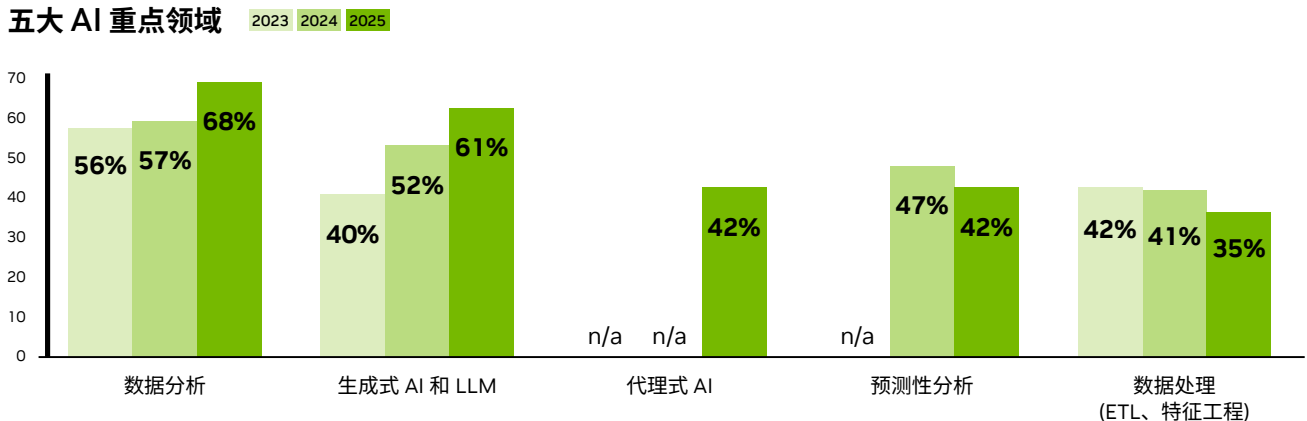
的大型公司 (员工超过 1,000 名) 正在积极使用 AI。

### 按行业细分的当前 AI 阶段

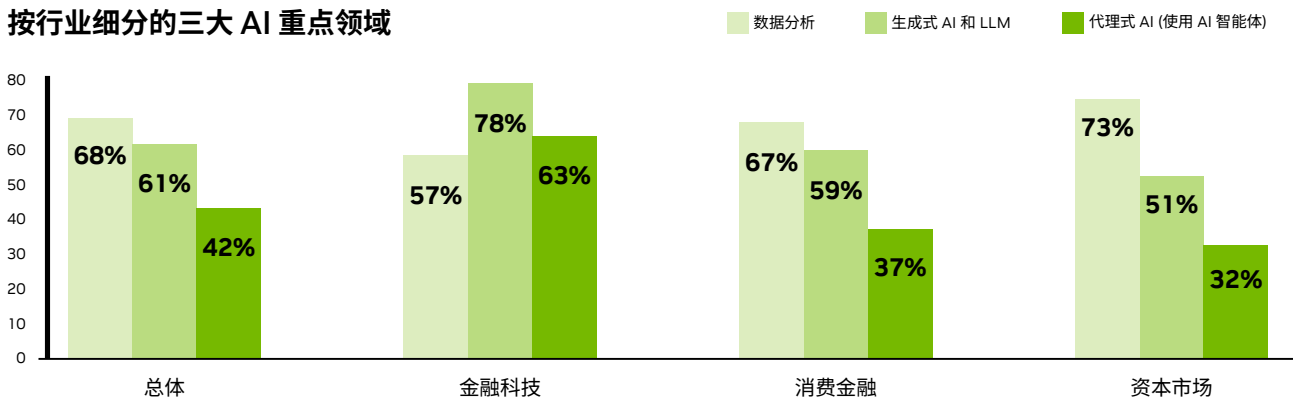


就总体使用情况而言，68% 的受访者表示，数据分析是金融服务机构使用或评估的首要领域。生成式 AI 以 61% 的受访者比例排名第二，随后是代理式 AI 和预测分析，各占 42%。数据处理，例如特征工程和提取、转换和加载 (ETL) 操作，占比 35%。

### 五大 AI 重点领域



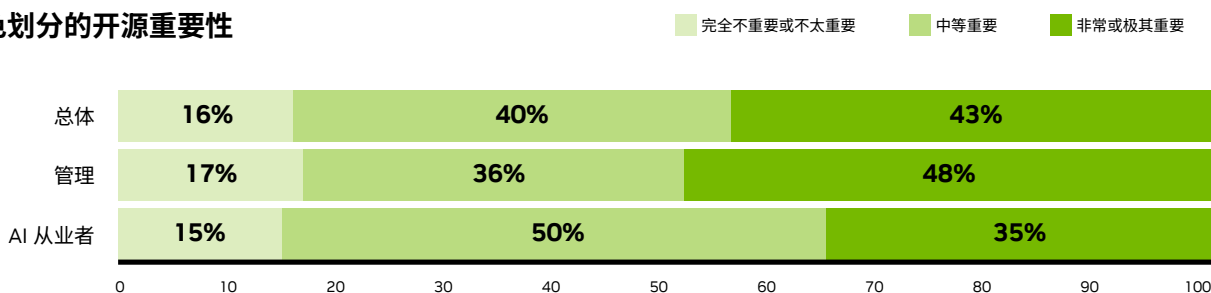
## 按行业细分的三大 AI 重点领域



## 行业正在拥抱开源

随着 AI 用例及其要求越来越契合组织的个性化需求，AI 解决方案的工具和开发也变得更具针对性。例如，开源模型和软件通常最适合构建特定 AI 应用。受访者对此表示认同，其中的 40% 认为开源软件对其组织的策略中等重要，43% 认为非常或极其重要。

## 按角色划分的开源重要性

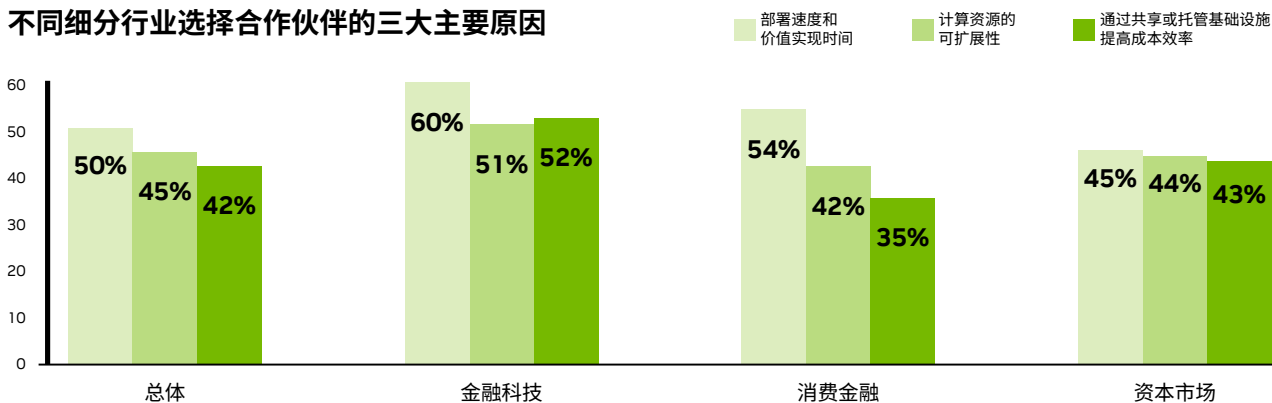


为了构建 AI 解决方案和支持计算基础设施，金融机构依赖于第三方合作伙伴。总体而言，44% 的受访者表示，他们会根据特定项目的需求聘请第三方，另有 31% 的受访者表示他们高度依赖第三方提供 AI 基础设施。

75% 的受访者表示，最常见的第三方合作伙伴是云服务提供商。其次是独立软件供应商 (占比 31%)，再次是原始设备制造商 (占比 25%)。排名前五的还有托管服务提供商 (占比 19%) 以及全球系统集成商 (占比 18%)。

50% 的受访者表示，部署速度和优化价值实现时间是聘请第三方的首要原因。其次是计算机资源的可扩展性 (占比 45%) 以及通过共享或托管基础设施实现的成本效率 (占比 42%)。

## 不同细分行业选择合作伙伴的三大主要原因



## AI 智能体在金融服务领域初显成效

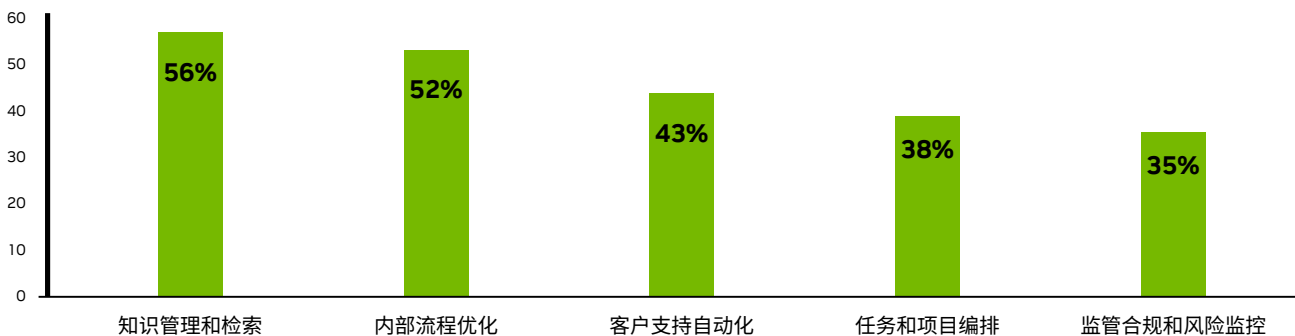
2026 年 NVIDIA AI 现状调查中新增了一个问题，涉及代理式 AI 在组织运营中的应用。AI 智能体是先进的 AI 系统，旨在根据高级目标自主进行推理，制定计划并执行复杂任务。

如上所述，42% 的调查受访者表示，他们的组织目前正在使用或评估代理式 AI。在这一数字中，近半数 (49% 或调查受访者总数的 21%) 表示，其组织内部已部署智能体。另有 41% (受访者总数的 18%) 表示将在未来一年内部署 AI 智能体。

在使用或评估代理式 AI 的组织中，有 56% 将知识管理和检索视为 AI 智能体的首要用例。其次是内部流程优化 (占比 52%) 以及客户支持自动化 (占比 43%)。

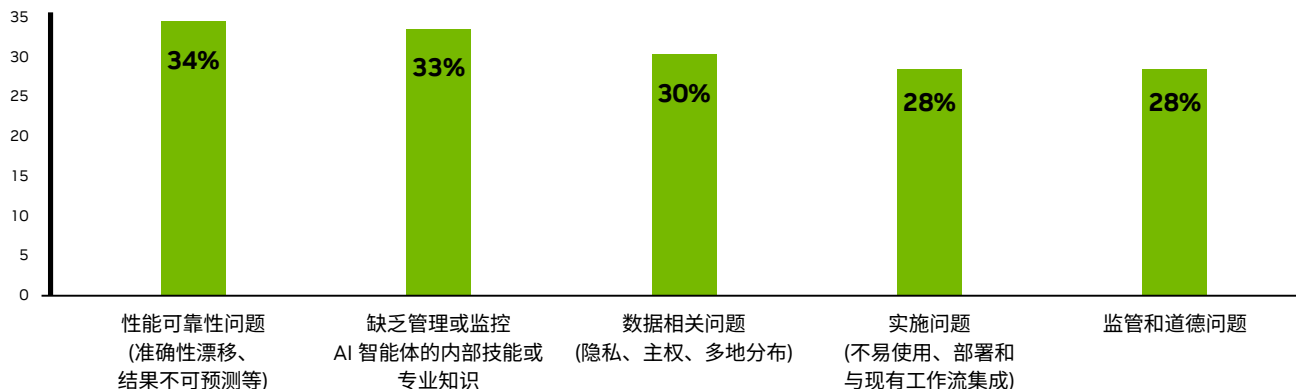
知识管理和检索是正在使用或评估代理式 AI 的组织最关注的用例。

## 使用 AI 智能体的五大 workflow



代理式 AI 仍处于企业部署的早期阶段，其面临的挑战与大多数新技术在被引入时所面对的难题如出一辙。性能和可靠性问题是调查中最多的回答，34% 的受访者选择了这个选项。其次是缺乏管理和监控 AI 智能体的内部技能或专业知识 (占比 33%)、数据相关问题 (占比 30%) 以及实施问题 (占比 28%)。

## AI 智能体面临的五大挑战

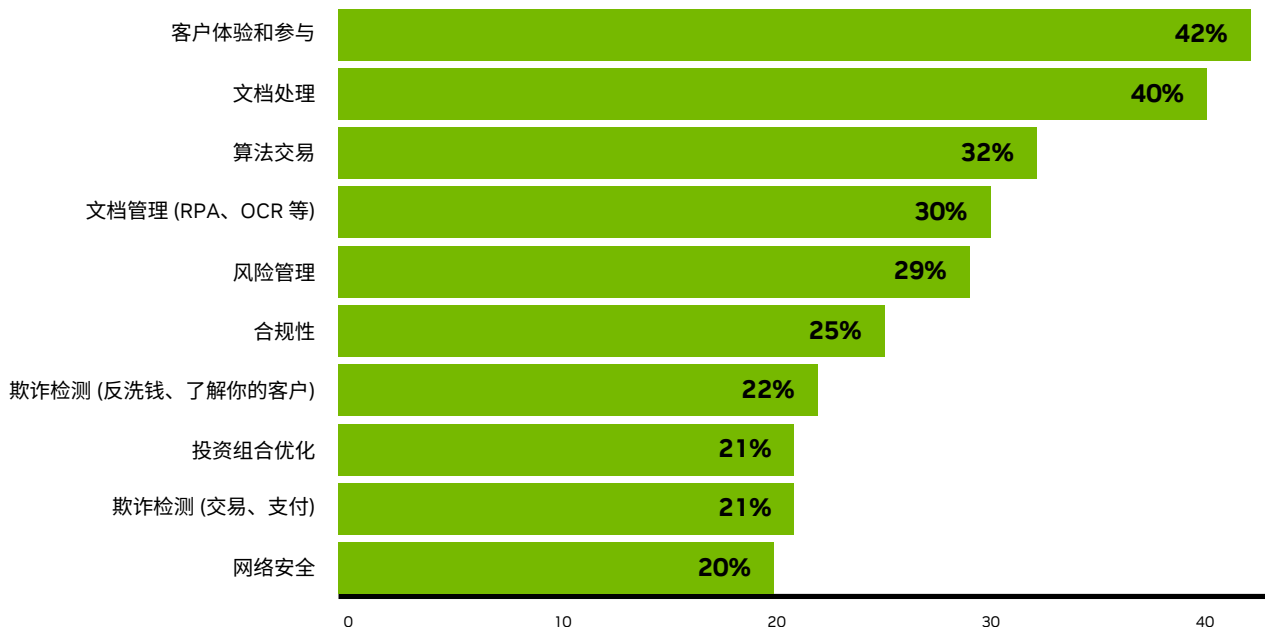


## AI 在运营中得到广泛应用

金融服务是一个大量涉及语言和数字的行业，非常注重文档处理和数值分析。这些特性使该行业非常适合采用 AI 来处理广泛的流程和用例。调查结果反映了这一现实，报告的 AI 用例广泛分布在组织运营的各个方面。总体而言，42% 的受访者表示，金融服务行业中最主要的 AI 驱动用例是客户体验和参与。其次是文档处理 (占比 40%)、算法交易 (占比 32%) 以及文档管理，包括机器人流程自动化 (RPA) 和光学字符识别 (OCR) (占比 30%)。

行业最主要的 AI 用例  
是客户体验和参与。

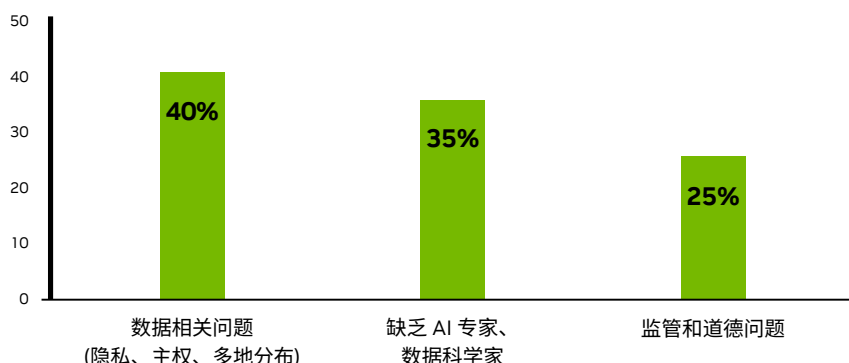
## 2025 年十大 AI 用例



组织在使用 AI 方面面临的<sup>最大挑战与数据相关</sup>——40% 的受访者提到了这一点，这一比例较 2024 年的 33% 有所上升。关键问题包括隐私、主权和数据分散在不同位置。35% 的受访者认为第二大挑战是 AI 专家和数据科学家短缺。

数据问题是行业采用 AI 的最大挑战。

## 2025 年的三大 AI 挑战



今年调查中的一个显著趋势是，金融服务机构似乎在利用数据进行模型训练和提高准确性方面取得了重大进展。将足够规模的数据用于训练和提高准确性不再是行业的主要挑战之一，认为这是主要挑战的受访者比例从 2023 年的 49% 下降到 2024 年的 31%，到 2025 年仅有 16%。

在过去一年中，各机构在利用数据进行模型训练方面取得了显著进展。

## 运营效率和员工生产力是金融服务领域 AI 最重要的投资回报

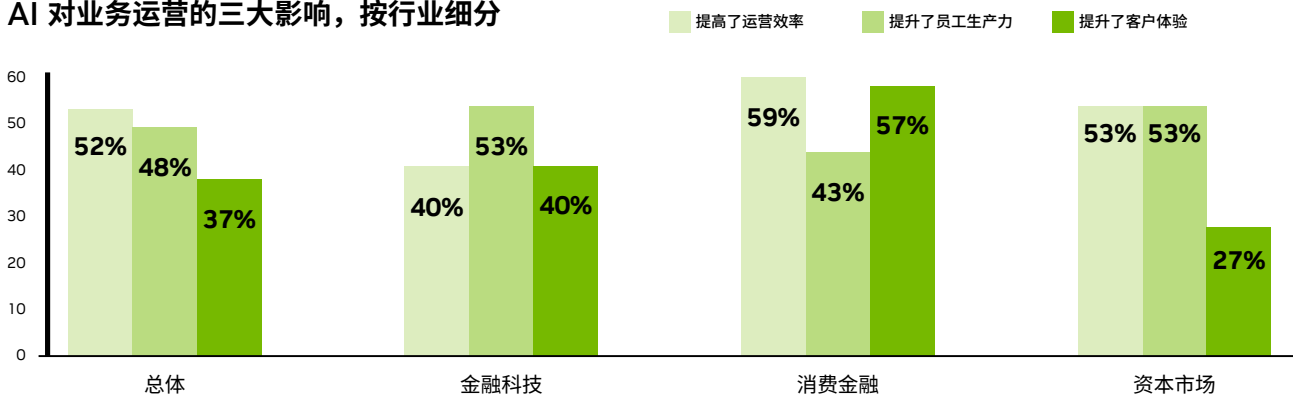
AI 正在对金融机构的利润产生积极影响，89% 的受访者表示，AI 既帮助增加了收入，又降低了年成本。影响可能非常显著，64% 的受访者表示 AI 帮助其收入增长了超过 5%。61% 的受访者表示，AI 帮助降低了同等比例的成本。

这种影响对于小型公司和大型公司同样适用，63% 的大型公司认为 AI 帮助其收入增长了超过 5%。特别是，75% 的金融科技受访者表示，AI 帮助其收入增长了超过 5%。

影响已经在多个领域显现。当被问及 AI 如何改善业务时，52% 的受访者表示 AI 帮助提高了运营效率。其次是提高员工生产力，持这一观点的受访者占比 48%，高于去年调查中的 22%。提升客户体验位列第三，37% 的受访者如此表示，尤其是金融服务中更面向客户的领域。57% 的消费金融领域受访者报告称，AI 帮助提升了客户体验，而资本市场领域只有 27% 的受访者如此表示。

行业内的大小公司都看到了 AI 带来的积极变化。

## AI 对业务运营的三大影响，按行业细分

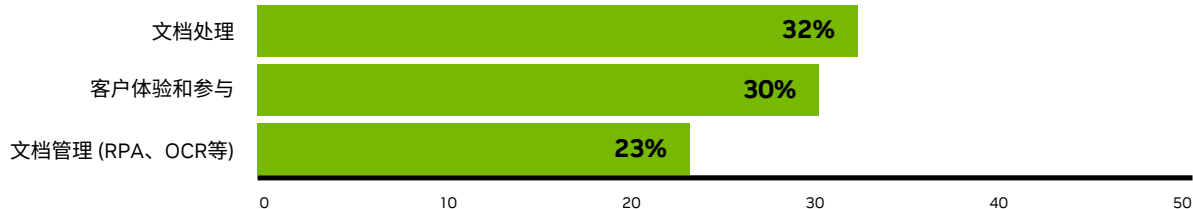


受访者列举了大量已实现出色投资回报率的 AI 用例。总体而言，投资回报率排名前三的用例是文档处理 (受访者占比 32%)、客户体验和参与 (占比 30%) 和文档管理 (占比 23%)。受访者的回答体现了 AI 为业务运营带来的广泛效益，其中 12 个回答的提及率超过 10%。

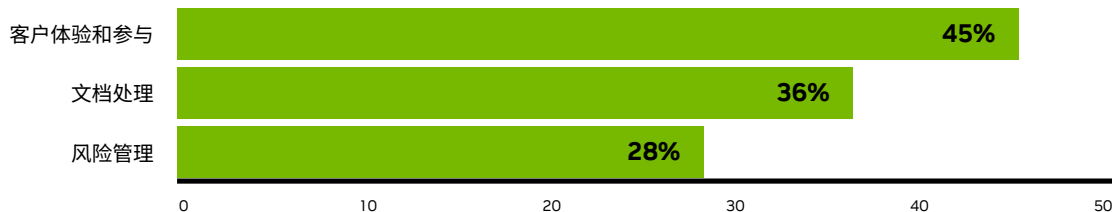
各个行业细分领域也显示出不同的优先级，26% 的资本市场受访者将算法交易列为投资回报率最高的用例之一。在金融科技企业组织中，28% 的受访者认为 AI 驱动的风险管理是投资回报率最高的用例。24% 的消费金融受访者认为欺诈检测和反洗钱是投资回报率最高的用例之一。

## 投资回报率最高的三大 AI 用例，按行业细分

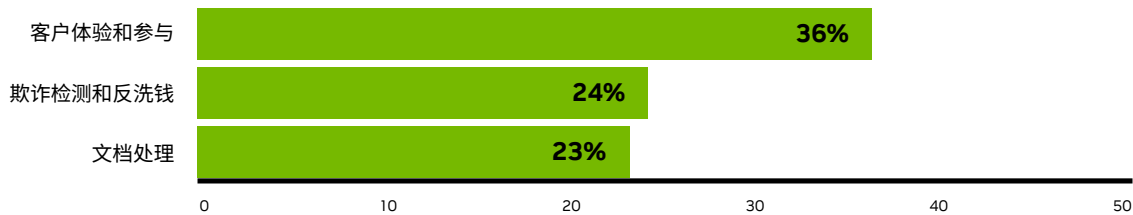
### 总体



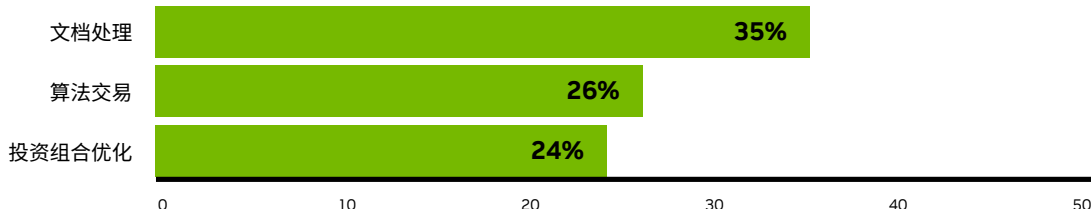
### 金融科技



### 消费金融



### 资本市场



## 金融机构计划在 2026 年增加 AI 投资

由于 AI 已被证明对企业的收入和利润都有影响，对该技术的投资将继续增加。总体而言，近 100% 的受访者表示，其组织在 2026 年对 AI 的投资将增加或保持不变，83% 的受访者表示会增加。近半数的受访者 (44%) 表示，他们的 AI 预算在未来一年的增幅将超过 10%，与 2024 年相比，持有这种观点的受访者比例增加了 14 个百分点。大型公司也计划在未来一年增加 AI 预算，46% 的受访者表示他们的 AI 预算将增加超过 10%。

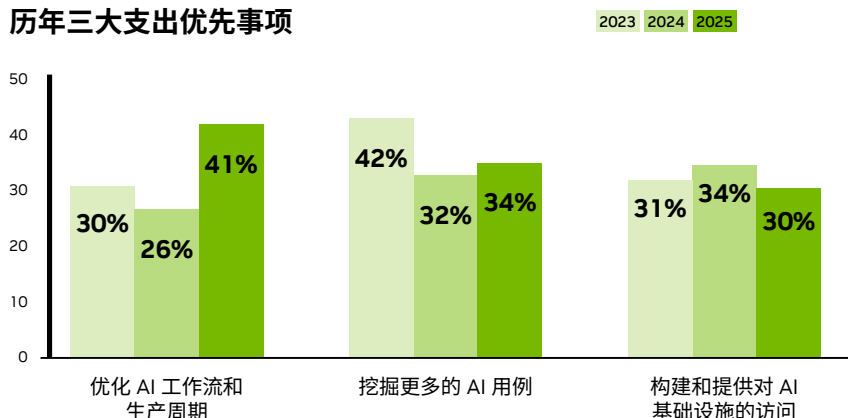
金融机构的 AI 投资将趋向优化现有工作流和寻找这项技术的新用途。总体而言，41% 的受访者表示，优化 AI 工作流和生产周期是他们 AI 支出的首要任务，这一比例较 2024 年的 26% 有所上升。基本上，受访者表示希望改进已经行之有效的方案。接下来的支出优先事项是挖掘更多 AI 用例 (占比 34%) 和构建对 AI 基础设施的访问 (占比 30%)。

AI 正在对金融机构的收入和利润产生积极影响。

调查受访者一致认为，2026 年对 AI 的投资将增加或保持不变。

优化 AI 工作流和生产周期是首要支出优先事项。

### 历年三大支出优先事项



## 金融服务领域正在选择混合架构处理 AI 工作负载

AI 推理是指经过训练的 AI 模型通过对新数据进行推理和预测来生成新输出的过程，该过程实时对输入进行分类并应用所学的知识。本质上，推理是 AI 模型的输出。推理不同于模型训练，因为它是一个持续的过程，且需要成本。每次查询模型时，都会进行推理。

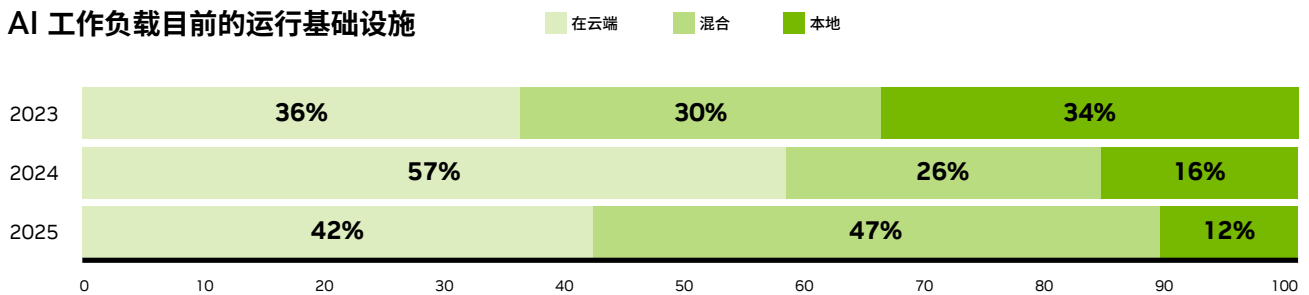
推理是 AI 的一个多方面环节，需要在模型的准确性和速度以及基础设施性能、功耗和整个系统的成本效率之间取得平衡。当被问及组织运行推理时最重要的因素是什么时，37% 的受访者表示模型性能和基准测试是最重要的因素。其次是数据驻留和合规性 (占比 36%)、优化总拥有成本和成本效率 (占比 34%) 以及延迟、准确性和吞吐量 (占比 32%)。

金融机构越来越多地转向采用混合架构 (整合云端和本地系统) 来处理 AI 工作负载。总体而言，47% 的受访者表示其组织使用混合架构，这一比例较 2024 年的 26% 有所上升。混合架构的兴起与独立云和本地架构的减少相关，云端采用率从 2024 年的 57% 下降到 2025 年的 42%。仅本地部署的基础设施从 2024 年的 16% 下降到 2025 年的 12%。

# 47%

的公司现在正在使用混合架构，这一比例较 2024 年的 26% 有所增加。

### AI 工作负载目前的运行基础设施



41% 的受访者表示，运行混合架构的首要原因是在不同的环境中运行不同的工作负载可优化成本。

## 展望未来

展望未来，行业领导者正致力于寻找更多可以让 AI 产生影响的领域，包括欺诈检测、算法交易、风险管理、合规监管、客户参与和文档处理。AI 目标遍布各业务职能，42% 的受访者表示，其首要目标是提高运营效率。排在其后的是创造竞争优势 (占比 34%) 以及提高员工生产力和提升客户体验 (两者均占比 33%)。

在未来一年，金融机构将扩大已证明其价值的用例规模，采用开源工具和模型，并与专家合作伙伴合作构建专门的解决方案。预计其中许多解决方案将以代理式 AI 的形式出现，因为自动化流程的能力日益凸显，而且新方法使得在企业中构建和管理 AI 智能体变得更加容易。这一切都将基于 AI 工厂构建，AI 工厂根据数据制造智能，为支持金融服务中各项功能和业务线的 AI 应用提供动力。

## 调研方法

本调查于 2025 年 8 月至 9 月进行，共收到 839 名受访者的回复。其中管理层 (包括高管) 和 AI 从业者各占 50%，并涵盖多个行业领域，包括资产管理人、资产所有者、经纪交易商、咨询公司、商业和零售银行、金融科技、投资银行、保险、市场数据和交易所、支付以及监管机构。受访者来自各种规模的公司，其中三分之一的年收入超过 2.5 亿美元，另有 40% 拥有超过 500 名员工。在线调查源自 NVIDIA 的分发列表和全球社交媒体，在中国则通过第三方机构进行。

## 准备好开始了吗？

要详细了解领先的金融机构如何使用 AI，  
请访问 [nvidia.cn/finance](https://www.nvidia.cn/finance)

