

Snowflake 的 AI 演进：以数据优先战略推动创新升级

关于 Snowflake

美国 Snowflake Inc. 成立于 2012 年，总部位于美国蒙大拿州博兹曼市，是一家提供数据云平台的云数据存储公司，其致力于将数据整合到单一来源，以此推动有意义的业务洞察、构建数据驱动的应用程序，以及共享数据和数据产品。该公司支持多种数据工作负载，服务于广告、金融、医疗等多个行业。2024 年，Snowflake 公司营收为 28.06 亿美元，员工人数约为 7000 人。

Snowflake 公司在数据平台方面提供了丰富多样的服务和产品。其核心的云数据平台采用云原生架构，支持多云部署，具备分离计算和存储、多集群共享数据、支持多种数据格式以及数据共享和即时访问等特性，可承载多种数据工作负载。在此基础上，Snowflake Data Marketplace 作为集中式的数据集市，方便企业进行数据集的买卖；Snowpark 为开发者提供了在平台上构建和部署 AI 模型的工具；Snowflake Open Catalog 能助力企业管理数据架构需求；Document AI 可从非结构化数据中提取有价值的洞察；Snowflake Horizon Catalog 具有新威胁预防和安全监控功能；存储生命周期策略帮助组织管理存储成本和合规性。此外，Snowflake 的平台还支持各类数据的统一存储和分析，能与多种 AI 框架和工具无缝集成，服务于众多行业。



转型中的 Snowflake：以数据优先的方式拓展 AI 业务

Snowflake 公司的 AI 战略以数据优先为核心，充分利用其在数据管理方面的优势，推动先进 AI 能力的发展。在竞争激烈的数据和 AI 平台市场中，Snowflake 公司的战略有三个关键要素使其脱颖而出。首先，该公司凭借其作为数据仓库提供商的基础，提供完全集成的数据和 AI 平台。通

过优先管理结构化和非结构化数据，Snowflake 公司在统一的生态系统中实现了 AI 驱动的分析、机器学习（ML）工作流程和高级处理。其次，Snowflake 公司运用先进技术拓展其 AI 能力，其中包括使用 GPU 加速机器学习工作负载；利用 Snowflake 容器服务（SCS）实现高效的模型部署；以及通过 Snowpark，支持使用 SQL、Python 和 Java 进行无缝 AI 开发。第三，Snowflake 公司通过 Cortex 展开开源 AI 合作，不断增强其生态系统，整合来自 Meta、Hugging Face 和 Mistral 的模型，为自然语言处理、预测分析和自动化提供支持，这一切都在一个以数据为核心的安全框架内进行。通过将数据作为 AI 发展的基础，Snowflake 公司在确保生态系统的安全性、性能和管理规范的同时，实现了高效的规模拓展。

客户对可扩展、可管理且具有实际应用价值的的需求，仍然是推动 Snowflake 公司发展的关键因素。该公司管理和协调不同类型的能力被多次强调。例如，日产利用 Snowflake 平台对数百万来自多个市场的客户资料进行分析。这一举措展示了 Snowflake 公司应对复杂大规模数据挑战的能力，同时为决策提供了具有实际应用价值的见解。

生成式 AI：释放聊天机器人之外的无限潜能

Snowflake 公司的可扩展性不仅体现在性能方面，还通过成本与价值的合理匹配，为 AI 的广泛应用发挥了关键作用。该平台的随用随付定价模式，能够根据 AI 应用的动态需求进行调整，尤其适用于生成式 AI（GenAI）和预测建模等资源密集型应用场景。这种灵活的模式使企业能够高效地拓展 AI 工作负载，并降低了实施先进 AI 解决方案的门槛。

在“Snowflake 生成式 AI 日”活动中，Snowflake 公司展示了生成式 AI 在传统应用（如聊天机器人和内容生成）之外的巨大潜力。例如，Snowflake 公司的合作伙伴西格玛计算公司（Sigma Computing）展示了 Snowflake 平台如何将 Salesforce 的原始数据转化为具有实际应用价值的见解。这种由 AI 驱动的分析不仅帮助西格玛的商业领袖们做出更明智的决策，还减少了手动数据准备所花费的时间，从而更快地得出更有依据的结果。

然而，随着企业对生成式 AI 应用的规模扩展，他们面临着数据偏差、知识产权风险和 AI 伦理等方面的挑战。为了赢得客户的信任，供应商在设计 AI 解决方案时必须考虑到治理、公平性和透明度，以确保 AI 的负责任部署。客户需要实施强有力的数据治理措施，仔细监控数据，避免产生不准确或带有歧视性的结果。

黄金数据集与 AI 发展的未来

“Snowflake 生成式 AI 日”活动中强调的一个新兴趋势是黄金数据集，即针对生成式 AI 应用场景优化的结构化和非结构化数据的精选集合。这些数据集在 Snowflake 平台的加持下，能够帮助企业实现更准确、更具影响力的 AI 成果。此外，Snowflake 公司对文本到语言提示功能的关注，通过减少对复杂 SQL 查询的依赖，简化了数据交互过程，这体现了其改善用户体验的决心。借助 Snowflake 公司的通用搜索（Universal Search）功能，客户可以根据数据质量和在工作流程中的使用情况，在其账户中识别数据集，从而创建经过优化的——即“黄金”数据集。

通用搜索功能确保了无论用户的技术专业水平如何，都能有效地利用 Snowflake 公司的能力进行 AI 开发、分析和决策。然而，构建和维护黄金数据集面临着巨大的挑战。对于许多企业来说，大规模整理和清理数据需要先进的治理框架和专业团队，以确保数据的质量、相关性和准确性。缺乏这些能力的企业可能难以从其 AI 模型中获得有意义的见解。此外，黄金数据集中的错误或不一致可能会导致有偏差的结果，从而削弱人们对 AI 驱动决策的信任。

简化用户交互

“生成式 AI 日”活动中强调的另一个主题是 Snowflake 公司对提高用户可访问性的关注。通过在其数据和 AI 平台中融入文本到语言提示功能，Snowflake 公司降低了可能缺乏 SQL 或其他编程语言专业知识的用户的技术门槛。这一功能确保了非技术用户也能有效地与平台进行交互，使数据驱动的见解在不同团队中都能得到应用。

2025 年预测：从实验到企业级 AI 的转变

Snowflake 公司预测，随着企业在 2025 年从实验性试点全面转向成熟的企业级 AI 解决方案，AI 的应用将发生重大转变。根据 TBR 在 2024 年下半年进行的云基础设施与平台客户研究调查结果显示，由于成本限制和对新兴技术缺乏专业技术知识，生成式 AI 解决方案在 2025 年的应用进展可能比预期更为缓慢。尽管存在这些挑战，Snowflake 公司预计 AI 的可观测性将推动 AI 的应用，因为企业越来越需要优先考虑投资回报率的衡量、部署的可靠性和法规合规性。在演讲中，发言者讨论了 Snowflake 公司的关键 AI 进展，如嵌入模型以增强包括 GPT 模型在内的大语言模型的性能，如何实现特定任务的定制化、提升多语言能力和优化整体模型性能。Snowflake 公司的平台通过 Snowflake Notebook 和 Snowflake 容器服务（SCS）等容器化运行时环境，为这些工作提供支持，这些工具为 AI 开发提供了可扩展且高效的手段。贝希托尔德强调了强大的数据集在支持生成式 AI 和传统机器学习模型方面的关键作用。Snowflake 公司的平台解决了数据安全、治理和可访问性等关键挑战，确保企业能够在从医疗保健到制造业的各个行业中自信地部署 AI 解决方案。

安全、治理与容器化：建立对 AI 的信任

安全和治理成为 Snowflake 公司 AI 战略的核心主题。随着企业越来越多地集成多个平台和环境，数据泄露和违规的风险也在增加。Snowflake 公司的治理方法包括围绕保护云配置、验证模型访问和监控运行时环境制定最佳实践，以确保其 AI 解决方案具有可扩展性、安全性，并符合不断变化的法规要求。例如，OM1 公司对 Snowflake 平台的使用展示了容器化系统如何简化治理流程，提高可扩展性和效率。通过利用这些系统，Snowflake 公司确保客户能够放心地部署 AI 解决方案，因为他们知道他们的数据和模型得到了保护。

尽管 Snowflake 公司做出了努力，但大规模管理安全和治理仍然是一个持续的挑战。在金融或医疗等监管严格的行业运营的客户，可能需要额外的定制化服务，以确保他们符合严格的监管

要求。此外，扩展治理框架以适应快速发展的 AI 应用场景，可能会给 Snowflake 公司的平台和资源带来压力。在保持创新的同时，提供一致的企业级支持，对于 Snowflake 公司应对这些挑战至关重要。

Snowflake 公司的路线图：在满足企业需求的同时推动创新升级

展望未来，Snowflake 公司将继续专注于扩展其集成数据和 AI 平台，同时保持其可扩展性、灵活性和可观测性的核心优势。该公司在弥合结构化和非结构化数据之间差距的能力，再加上其在用户体验、嵌入模型和 AI 可观测性方面的投资，将使其在新一轮 AI 创新中占据领先地位。然而，Snowflake 公司的成功将取决于它能否在创新与治理之间取得平衡，确保企业在满足合规要求的同时，能够应对独特的数据挑战。通过专注于赋能用户、简化 AI 部署和推广先进技术，Snowflake 公司将能够更好地满足快速变化的市场需求。

结论

Snowflake 公司的发展历程体现了其通过数据优先战略推动 AI 发展的决心。通过应对现代数据生态系统的复杂性，并使其平台与新兴的 AI 趋势保持一致，Snowflake 公司已在 AI 领域确立了重要地位。这一战略重点不仅推动了数字转型，还塑造了市场竞争格局，对合作伙伴、竞争对手和技术供应商都产生了影响。然而，Snowflake 公司在集成多样化的数据环境和大规模维护数据质量方面面临挑战。随着 AI 应用的扩展，对 GPU 等重要基础设施的投资需求也可能成为障碍。随着生成式 AI 和 AI 可观测性的不断发展，Snowflake 公司的集成平台有能力支持合作伙伴和利益相关者应对行业转型的下一阶段。通过提供可扩展且安全的 AI 工作流程，Snowflake 公司正在帮助他们应对跨行业大规模采用 AI 所面临的挑战。

Nokia 作为电信运营商数据共享平台的主要合作伙伴，也在积极探索开源社区和联合开发，持续扩展数据符合和 AI 能力，使电信运营商用户能够更轻松部署和定制数据服务和 AI 模型。同时，Nokia 也重点关注 AI 的可观测性和可解释性，为电信运营商提供跟踪性能、优化结果和确保投资回报率的工具，建立数据治理框架确保法规合规性，保护电信运营商的 AI 数据和模型。同时致力于简化用户体验，使非技术用户更容易使用 AI，这符合新的行业标准。通过降低技术门槛，使电信运营商更多业务部门和企业客户能够利用 AI，并鼓励市场朝着更用户友好的解决方案进行创新。