

# 历久弥坚的 AI 基础设施

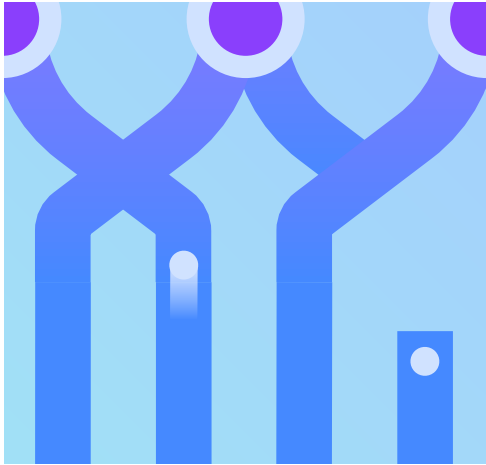
驾驭速度、规模和信任

# 目录

打造坚实 AI 引擎 .....	2
构建灵活 AI 架构 .....	4
筑就可信 AI 基石 .....	10
聚势而强，共赢 AI 未来 .....	14
行动指南 .....	18

## IBM 如何提供帮助

IBM 所提供的 AI 基础设施，不仅具备安全、可扩展、开放与法律保障等核心特质，更能完美支撑高度动态且要求严苛的 AI 工作负载；同时，它更提供了一条可持续的 AI 运营之道，有效实现成本、风险与能耗的三重降低。要了解更多信息，请访问 [ibm.com/solutions/ai-infrastructure](https://ibm.com/solutions/ai-infrastructure)



## 摘要

成功的 AI 架构，应如精密引擎般运作：计算驱动、数据赋能、治理保稳。

- 组织普遍认同面向 AI 的混合基础设施，然理念先行，实践滞后。

70% 的高管坦言，合适的混合基础设施有效兼顾性能与成本，然而，仅有 8% 认为，其现有基础设施能完全满足 AI 需求。

- AI 要想规模化，信任是前提，治理是根基。

83% 的高管表示，有效治理是 AI 基础设施的成功根基；但仅 8% 落实风险管理，尽显知行鸿沟。

- 合作伙伴是 AI 成功的关键，然须谋定而后动，以战略眼光进行甄选。

高管们指出，合作未能充分释放价值，是制约 AI 基础设施投资回报的首要因素。然而，症结不在合作模式本身，而在于组织如何甄选与管理合作伙伴。

# 引言

## 打造坚实 AI 引擎

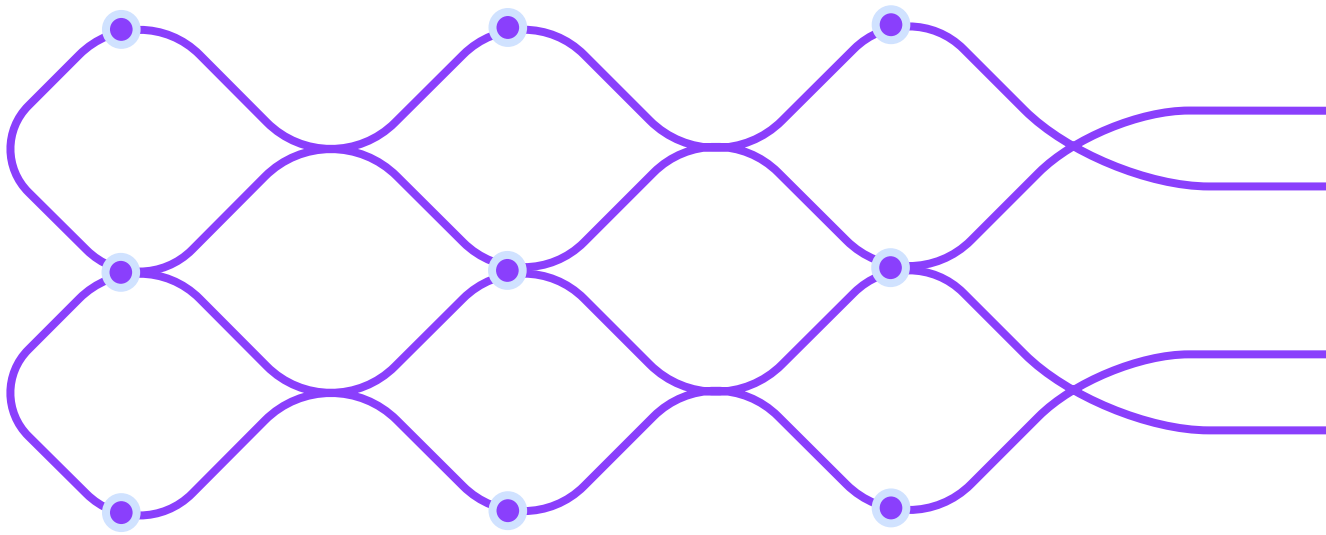
AI 已在董事会议程中成为战略焦点，却未能带来匹配的商业成效。IBM 商业价值研究院 (IBM IBV) 调研揭示，过去两年启动的 AI 项目，有近 40% 止步于试点阶段，未能全面推广。<sup>1</sup> 然而，AI 之困不在试点：底层架构割裂，制约了 AI 落地。

成功的 AI 架构，应如精密引擎般运作：计算驱动、数据赋能、治理保稳。然 AI 真正上阵时，隐患尽显：数据碎片化导致无法有效驱动 AI 模型，治理标准不一致使风险失控，基础设施平台难以承载现实工作负载。当 AI 迈向具备独立决策能力的智能体时代，组织必须以更专业、更协调的基础设施为其保驾护航。

为此，组织正重金押注，以求破局。IBM IBV 最新调研显示，1200 位最高管理层中，三分之二正加码基础设施投入，2025 年预算平均增幅近 19%。

但 AI 的成功无法靠预算堆砌，而是依靠系统性的建设与能力积累。

分析表明，具备高成熟度与基础设施就绪度的组织，共享一套共通基因，助其实现价值驱动的 AI 应用。（详见第 17 页观点：“AI 落地的必备要素”）。这些核心能力，正是打造“有设计、有集成、有成效”的 AI 基础设施蓝图。



**本报告聚焦驱动 AI 成功落地的三大核心要素：**

- 灵活架构如何助力 AI 贯通全域数据流。
- 内生信任机制如何破壁 AI 治理困局。
- 战略生态如何破局 AI 人才与能力瓶颈。

报告最后，提出了切实有效的“行动指南”，为拥有不同 AI 成熟度的组织提供可落地路径，助力构建驱动创新与成果转化的基础设施。

“AI 要成功，先定向，后行事。基础设施，只是赋能工具。”

**Adnan Kashwani**

阿联酋 e& 公司，端到端云与基础设施管理，副总裁

# 灵活架构，AI 之基石

即便组织战略方向明晰，但 AI 落地仍受制于基础设施短板。62% 的高管立志三年内实现跨环境 AI 部署，但仅有 8% 的高管认为，现有基础设施能完全满足 AI 需求。这一巨大落差，实为战略天花板，而非单纯技术鸿沟。

准备度差距的根源，是组织仍试图沿用在过去的IT环境下行之有效的传统规划方法，来构建复杂的 AI 能力。既定方法难以匹配 AI 的动态特性，大规模推理产生脉冲式算力需求，模型训练依赖广域性计算架构，而全流程资源消耗则呈现持续高压状态。

“AI 竞赛胜负未定，让系统具备自我进化能力，可随时切换、灵活应变，以应对技术更迭与成本波动。”

**Hauke Stars**

大众汽车集团，首席数据官

AI 时代，组织亟需重构基础设施规划思维，但前行之路受制于两大核心难题。

**挑战一：如何在牺牲安全或治理要求的前提下，让分散的数据为 AI 所用？** 企业数据的分散格局，既是历史遗留的结果，也源于隐私与安全策略的复杂性。高管表示，近三分之二的企业数据仍留驻本地，余下分布于私有云、公有云等混合环境中。数据碎片化正成为 AI 发展的致命瓶颈：多位高管一致指出，数据可访问性与可用性不足，是 AI 投资失利的核心症结。68% 的组织正以数据结构方案打通数据孤岛，但仅 42% 对自身基础设施应对 AI 模型的算力与数据需求有信心。

**挑战二：负载动态变化难以预估，如何构建可伸缩的技术弹性？** 大型语言模型 (LLM) 主导当下，但 AI 的未来属于多模型共融<sup>2</sup>，通用与专用并举。目前，已有 43% 组织采用混合模型策略，根据业务需求战略性组合定制模型、大型语言模型及小型开源模型。多模型并存带来动态负载变化，资源需求随时波动。这种动态变化构成完整的负载脉络：从高强度脉冲式的训练作业，到需持续稳定的低压推理流。然而，仅 49% 的高管相信，能够在未来三年内集成更小的开源 AI 模型，更少数 (46%) 认为，能实现大规模实时推理的高强度需求。传统方案导致组织陷入典型困境：选择资源错配的过度配置，或转向由供应商管理容量风险的即服务模式。

“数据是 AI 价值的基石，若无法顺畅连接、轻松调用，AI 系统终将形同虚设。”

Hiroshi Okuyama

Yanmar Holdings，首席数字官兼董事会成员

## 打造灵活 AI 基础设施的四大核心要素

高管们正面临架构选择的战略分水岭。若缺乏计算、数据与模型三位一体的科学布局战略，组织必将陷入架构困局，被迫采用性能不达标且资源错配的解决方案。

为突破 AI 瓶颈，必须从四个维度构建系统化的基础设施方案：

- **专业计算资源。**传统 CPU 虽能处理常规任务，AI 工作负载却以跨越技术分水岭，迈向专用加速器时代。当 GPU 在训练领域树立通用标准时，而模型推理——即使用已训练模型进行预测的过程——则往往需要另一类专用硬件，它们专为优化吞吐量、成本与能效而设计。
- **先进存储系统。**高吞吐文件系统、可扩展对象存储及并行化数据仓库，共同构成 AI 数据架构的三大基石，有效应对海量训练集与复杂访问模式的挑战。
- **高速网络架构。**为满足分布式训练与推理的海量数据交换需求，AI 工作负载必须构建于专为高带宽、超低延迟设计的网络体系架构之上。当模型步入巨量化时代且实时响应成为刚需，网络架构已与原始算力共成 AI 发展的双翼，缺一不可。
- **智能调度层。**容器统筹平台、专用模型服务架构与工作流管理系统，三者构成 AI 集群的智能调度中枢，实现异构工作负载的精准资源调控。缺失强健的统筹框架，再卓越的独立组件都会陷入“1+1<2”的系统效能陷阱，难以支撑企业级部署的整体要求。

## 混合范式：重塑基础设施的边界逻辑

四大要素唯有协同一体，才能发挥最大价值；而这离不开混合架构，即连接本地、私有云与公有云的统一体系。70% 的高管确认，基于工作负载优化的混合策略，已成为推动成本与性能双重跃升的战略最优解。混合架构解锁的核心能力矩阵：

- 本地数据处理：AI 在数据所在之处运行，既降成本，又避风险。
- 智能法规遵从：自动匹配不同地区和行业的监管标准。
- 增强安全管控：守护 AI 资产安全，抵御模型篡改与数据攻击。
- 动态环境匹配：让 AI 的每个阶段在最优环境中运行，实现算力与性能的完美契合。



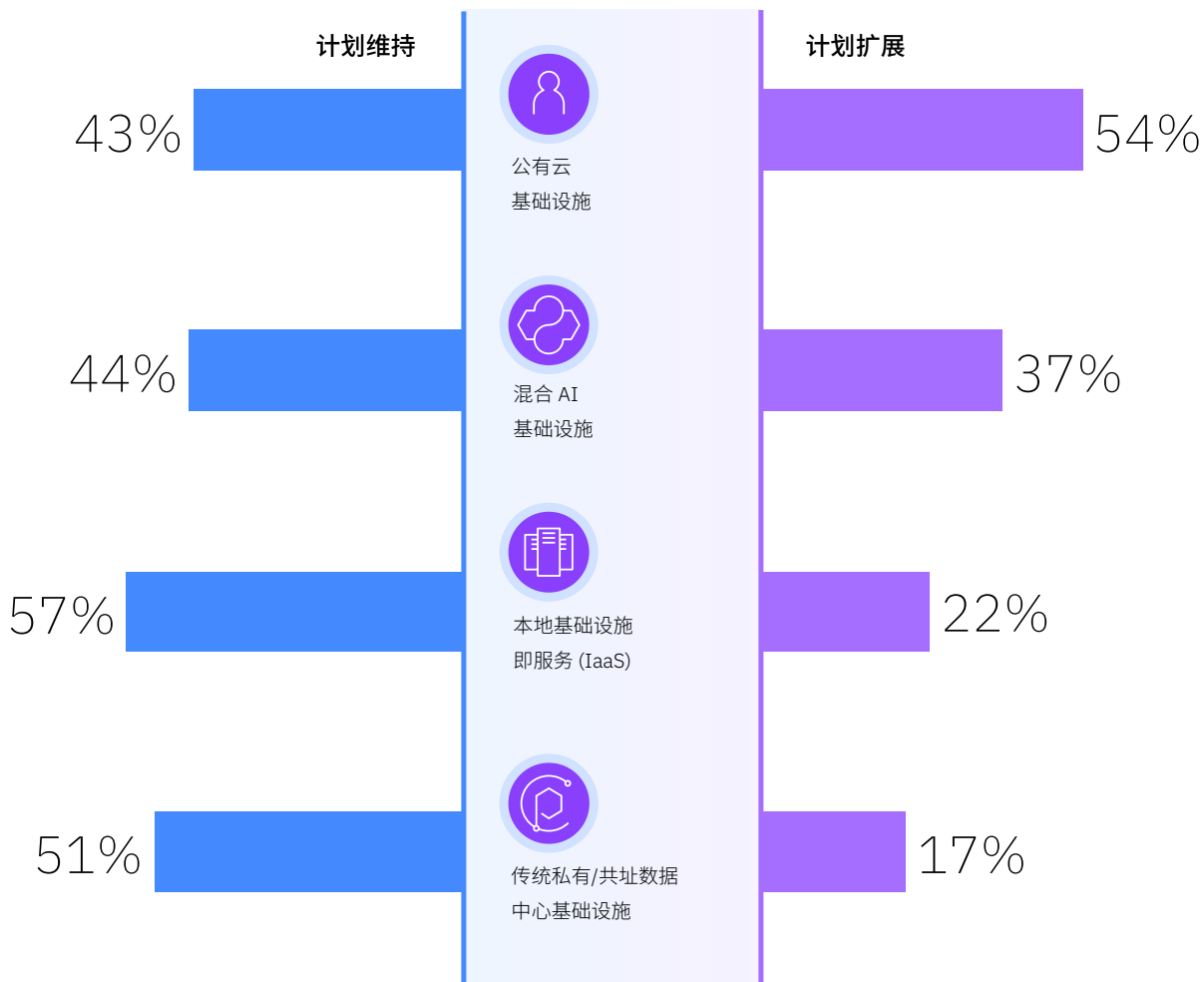
面向特定场景的方案强调，AI 训练与推理对基础设施的需求截然不同。训练阶段得益于大规模并行算力，推理阶段则讲究高成本效益、低延迟的解决方案，并需要更贴近数据源和用户。混合基础设施让组织为 AI 生命周期的每个阶段匹配最优环境。

如此一来，“混合”不仅是一个架构描述，更是一套深思熟虑的战略，旨在跨环境统筹调度 AI 能力。对此价值的共识正推动持续投资：超过三分之一 (37%) 的组织计划在未来三年内扩展其混合战略（见图 1）。

图 1

### 前瞻布局：组织制定均衡之策，让基础设施战略与 AI 发展脉搏同频共振。

未来三年内，计划在各领域维持或扩展其战略的高管百分比



---

## 观点

# 迎战 AI 的“能耗洪流”：如何实现高效与可持续

AI 的能源需求正呈爆炸式增长。<sup>3</sup> 对此，86% 的组织表示，已在基础设施决策中纳入碳足迹考量，78% 已将可持续深度融入其 AI 战略。

业界专家正全力攻坚破局：通过“小而精”的 AI 模型，能实现在削减 90% 能耗的同时，保持相当性能。<sup>4</sup> 具备功耗上限控制功能的硬件，可在系统层面降低约 15% 的整体能耗。<sup>5</sup> 优化模型训练流程，可带来多达 80% 的能耗节省。<sup>6</sup>

节能带来双赢：既减碳降耗，又降本增效，推动 AI 经济模式更具可持续性。AI 越“饥渴”，效率越关键。今日精通能源优化之道的组织，必将掌握未来 AI 的主导权。

## 案例研究

# 昆士兰大学借助高性能数据存储，加速 AI 驱动的研究进程<sup>7</sup>

昆士兰大学科研实力雄厚，研究范围覆盖生物医学、农业与食品、纳米技术、工程、环境科学、社会科学与数字人文等多个领域。其研究高度协同且注重实效，聚焦健康、可持续发展、科技与社会等全球挑战，依托多所顶级科研机构持续创新。为满足 AI 研究需求，昆士兰大学使用前沿的异构超算系统，而此类高性能计算 (HPC) 系统必须配备极速、可扩展且灵活的数据存储能力，从而实现“随时访问、无处不在”。

该校研究计算中心 (RCC) 致力于打造一体化高性能存储架构，为全校科研提供高效的数据采集与分析支持。该系统以“数据不动计算动”为核心理念，在应对数据量、速度和多样性指数级增长的同时，确保数据的逻辑单一性，并使研究人员随时随地可用。

RCC 携手 IBM 业务合作伙伴 OneTeam，共同打造高性能数据结构平台 MeDiCI（大都市数据缓存架构），通过 IBM Storage Scale 实现集中化管理与主动文件调度。MeDiCI 平台整合多类型平台与数据资源，打破存储孤岛，构建统一智能的数据架构。此方案实现了元数据的自动抓取与汇聚，让研究人员在跨校合作中，能无缝调用分布在不同计算集群的科研数据。

该存储系统支持高达 40 GB/s 的吞吐量，并拥有 EB 级扩展潜力。其混合云模型实现了元数据毫秒级响应，结合智能纠删码技术，全面保障存储系统的可靠性、可用性与性能。

统一的数据结构，让科研团队实现前所未有的运算速度，大幅缩短从数据到发现的周期。以阿尔茨海默症研究为例，团队借助该平台，将项目运行时长缩短了约 74%。新存储阵列仅用两小时便实现了投资回报，其性能改进为全校的医学影像研究人员每周节省了数百小时的处理时间。

# 筑就可信 AI 基石

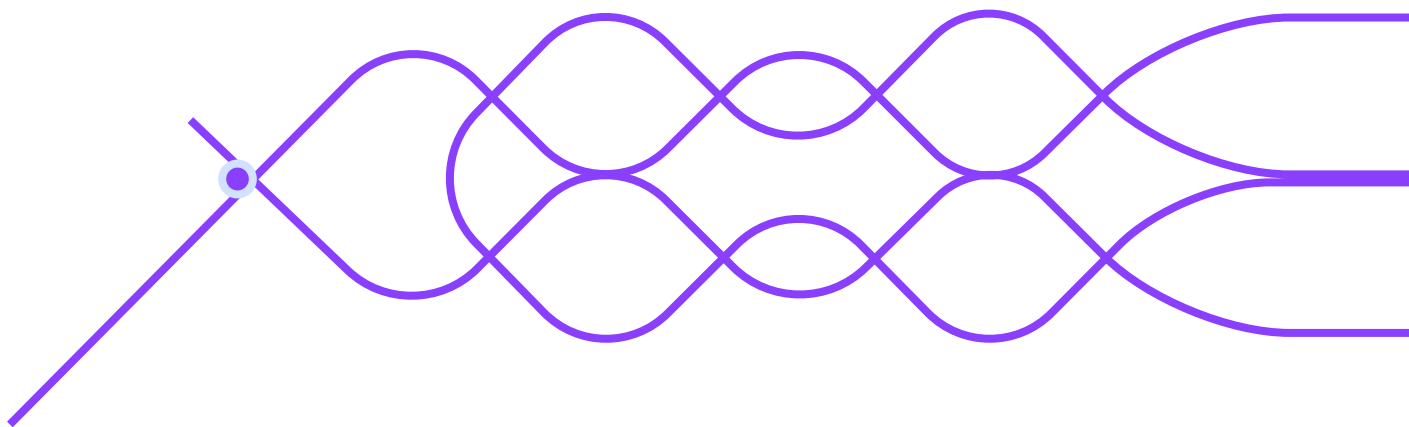
面对飞速发展的 AI 技术，治理与安全框架常常滞后。许多组织抢占 AI 先机，却在事后才补建治理与安全体系。延迟治理不仅放大了合规、隐私与模型安全风险，更可能引发伦理失范，最终动摇品牌信任、侵蚀客户忠诚。

83% 的高管达成战略共识，AI 治理已成刚需，其中道德与责任标准、隐私与数据安全、以及透明度和可解释性被共同认定为核心支柱（见图 2）。但仅 8% 真正将 AI 治理机制嵌入体系。此外，数据隐私漏洞、安全机制缺失与合规性挑战，正共同构成 AI 基础设施价值实现的三大梗阻。当战略认知超前而就绪度滞后，其间裂痕已转化为真金白银的代价。

要跨越这道裂痕，必须建立全新的破局范式。我们的研究表明，领先组织始终将治理内嵌于 AI 基础设施之中，这一理念我们称之为“可信设计”。该理念以体系化视角审视 AI 安全，在应用层、数据层、模型层与使用层构筑起安全控制与治理机制的双重防线，形成对传统风险与未知威胁的立体防御。<sup>8</sup> 面对全球监管环境的剧变，推行这一战略已从“最佳实践”升级为“生存必需”。

当组织在其运营中规模化部署 AI 智能体时，治理复杂度急剧上升：重复任务、模型割裂、敏感数据外泄等风险接踵而至。组织必须建立集中化的智能体管理平台，实现对合规、可信智能体的统一治理，并实时掌握交互与使用情况。

可信设计推动安全范式迁移：从被动信任转为持续验证，让安全成为可度量的运行状态。这一路径要求部署零信任框架，配备明确的安全防护机制与持续验证流程。可信设计推动安全范式从机密性、完整性与可用性 (CIA) 向未来威胁防御演进，为应对生成式 AI 洪流、自治系统集群及量子计算冲击构建前瞻性屏障。<sup>9</sup>



安全优先的理念必须全面融入治理体系。当治理融入架构核心，它将从“附属功能”转变为“增长引擎”，助力快速部署，增强系统韧性，同时确保数据、模型与伦理标准的可控与可信。

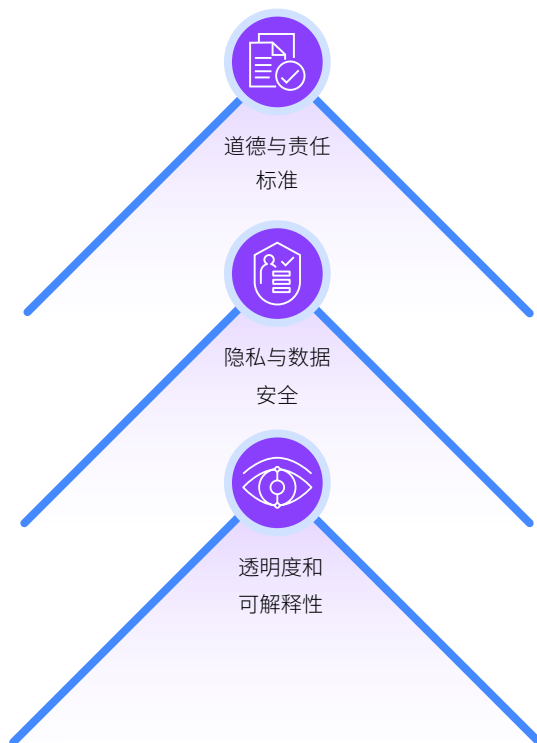
“在全球化进程加速的当下，数据治理的复杂性与日俱增。要应对这种变化，我们需要重新设计管理体系和技术底座，以支撑跨区域、跨规模的治理需求。”

**James Felix**

第一三共株式会社，首席数字转型官

图 2

高管构建 AI 治理体系时，始终聚焦于构建信任的核心基石。



## 可信基础设施的关键要素

让安全与治理融入 AI 生命周期的每个环节，相辅相成、协同运作。

### 安全与完整性

- **安全为核。**借助强健的网络与访问控制、数据加密及智能威胁检测，全面防护数据流、模型服务与 API 接口。
- **防护如一。**统一的安全协议与治理机制，确保 AI 工作负载无论在云端、本地还是边缘环境都能获得一致防护。

### 可追踪性与可溯源性

- **数据模型可溯。**借助数据沿袭与模型注册体系，实现自动化版本控制，为每一项 AI 决策创建不可篡改的审计追踪。
- **审计全覆盖。**对系统访问、模型修改及数据使用进行详尽记录，确保问责透明与合规可证。

### 监控与自适应

- **监控持续化。**实时监督模型漂移、性能偏差与异常行为，及时检测并纠正偏离。
- **策略动态调。**借助模块化控制机制，灵活调整以满足不断变化的合规要求，避免系统重构。

安全为盾，治理为纲，共筑可信根基，  
护航 AI 征途。

## 案例研究

# 巴西银行以“责任与信任”为引擎，推动 AI 创新与伦理并进<sup>10</sup>

秉承创新传统，巴西银行将 AI 视为当前必须抓住的战略机遇，而非未来才需应对的趋势。巴西银行作为肩负监管责任与社会使命的公共金融机构，必须制定清晰战略，以责任为先、透明为本地应用 AI。AI 在关键银行业务中快速普及，也带来了信任、问责与监管合规方面的全新挑战。为应对这些挑战，巴西银行与安永和 IBM 深度合作，共同打造面向未来的 AI 治理体系。

合作伙伴分工互补：安永构建契合银行体系的治理框架，IBM 则以技术平台将战略转化为现实，推动 AI 生命周期中的自动化、透明化与持续监管。

该框架为 AI 生命周期建立了清晰的角色分工与责任边界，并引入基础模型的严格评估机制，包括基准测试与漏洞检测。平台具备实时监控、主动告警及自定义指标功能，形成贯穿全流程的统一管控与信任体系。

战略与技术的协同创新，让巴西银行能够在高安全性环境中，兼顾伦理与信任，负责地扩展 AI 能力。主要成果包括：

- 实现 AI 生命周期自动化治理，降低人工干预成本，加速业务落地。
- 借助实时监控与模型可解释性，实现网络风险和合规性管理。
- 结构化评估基础模型，确保安全性与伦理合规。
- 可追溯性与版本管理体系，确保责任清晰与合规透明。
- 持续监测 AI 表现，实现快速纠偏与动态优化。

# 人才与伙伴聚势而强， 共赢 AI 未来

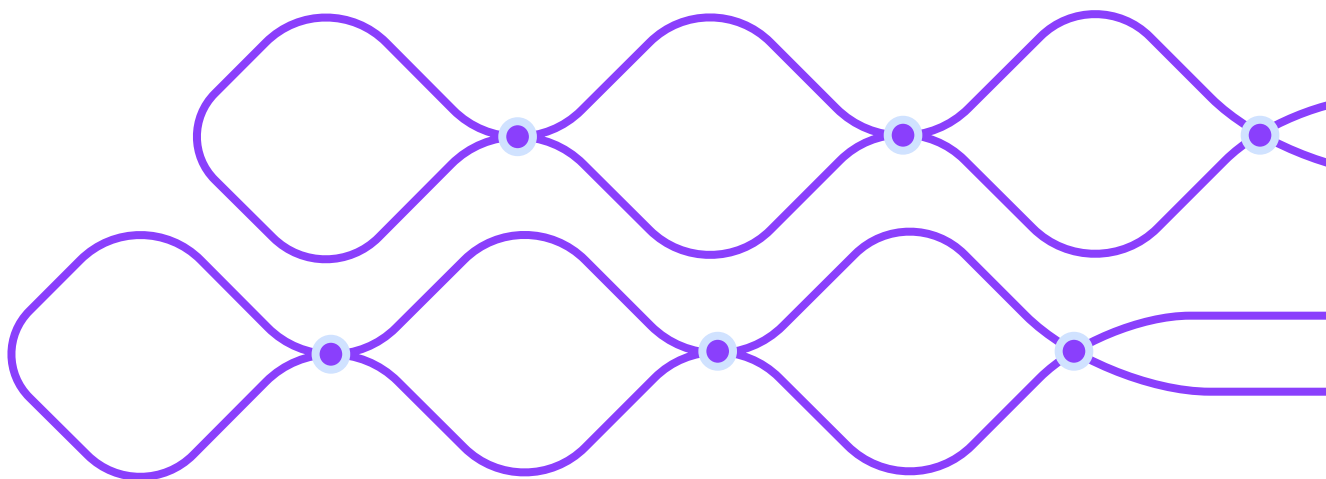
制约 AI 突破的关键要素早已转移：从技术壁垒转向人才壁垒。虽有 87% 的组织加大 AI 人才投资，但仍有三分之二坦言其人才体系仍处于早期建设阶段。人才短板带来连锁反应：基础设施建设放缓、方案难以扩展、创新受阻、竞争机会白白流失。

征途未知技术领域，唯有汇聚专业智慧的共同体，方能筑起最可靠的风险防线。研究表明，建立 AI 卓越中心 (CoE) 能显著加速组织成熟度与基础设施的完善，但目前仅 38% 的企业具备此能力。

阿联酋 e& 公司副总裁 Adnan Kashwani 指出：“面对 AI 等新兴技术，集体决策机制具有天然优势，技术日新月异，无人能独揽全部智慧。跨部门常设委员会机制，通过融合多元视角实现决策最优化，显著降低试错成本。”

大型企业依托 CoE 实现 AI 协同，而规模较小的组织同样可借助设立 AI 推广大使、虚拟社区或跨部门团队实现共创。结构可以多样，但成功的关键在于：让业务与技术形成战略共识、共享知识、统一标准，并系统培养跨职能能力。在 AI 生态日新月异的时代，唯有人才能够驾驭复杂，保持创新的持久动能。





## 供应商为何必须向战略伙伴转型

为弥合 AI 技能鸿沟并强化内部能力，几乎所有组织都在寻求外部合作：99% 的企业拥有至少一个 AI 生态合作伙伴，68% 正在积极探索或发展外部 AI 合作关系。但仅 5% 将合作关系融入运营体系，超过三分之一的组织仍将合作伙伴视为供应商资源池，而非战略联盟。

这种交易型合作模式造成了巨大的鸿沟。尽管外部合作是构建 AI 能力的关键加速器，但高管们普遍指出，合作执行效果未达预期是 AI 基础设施投资未能实现价值的主因。其症结不在于合作本身，而在于企业如何甄选、管理与融合这些外部能力。

“真正的伙伴既懂技术，也懂文化。既了解我们的通信底蕴与本地特性，又懂如何在这一生态中放大创新。”

**Adnan Kashwani**

阿联酋 e& 公司，端到端云与基础设施管理，副总裁

图 3

### 组织依据战略愿景选择伙伴， 却常陷于执行困局。

#### 远见导向

聚焦面向未来的基础设施，赋能尖端 AI 能力

62%

选择能支撑 AI 计划的前瞻性  
战略

56%

选择性能为先的基础设施以  
支持高负载 AI 工作

45%

选择具备弹性扩展能力的 AI  
解决方案，以满足工作负载  
与组织增长需求

#### 务实导向

聚焦运营效能，立见投资回报

40%

选择能精简 AI 模型全生命  
周期流程的基础设施

36%

选择能提供高速性能且停机  
时间最短的基础设施，以确  
保运营的连续性

31%

选择具备有竞争力的定价与  
明确投资回报的 AI 基础设施

多数组织在选择 AI 伙伴时，更看重其战略愿景与 AI 基础设施的前瞻性设计（见图 3）。但他们常忽视那些决定落地成效的关键能力：如高效的模型开发与流程部署，以及即插即用、稳定可靠的基础设施性能。

理想的合作伙伴，不仅懂技术，更懂业务：清晰数据流向、决策机制与运营节奏。专注细分行业的智能体供应商正崛起，他们以预配置平台为各行业量身定制解决方案，加快了 AI 落地。兼具业务理解与技术能力的合作伙伴，才能带来卓越成果。

## 观点

# AI 落地的必备要素

组织的运营成熟度与基础设施就绪度，如何共筑组织 AI 未来的胜负手？IBM 商业价值研究院 (IBM IBV) 揭示核心规律：AI 运营成熟度越高，基础设施就绪度越强的组织，不仅具备支撑生产级 AI 的技术能力，也更能把握未来的智能化机遇（详见第 20 页：“调研和分析方法”）。

成熟度体现“管理力”，衡量组织在政策、流程、人才与风险管理等方面构建 AI 体系的完善程度。就绪度体现“技术力”，衡量组织是否具备扩展算力、整合模型、分配负载与确保合规的能力，以应对未来三年的 AI 挑战。双维度领先的组织，年营收增长平均高出同行 9%。

研究发现，三类关键能力构成这两大维度的进阶引擎：奠定根基的基础能力、打通关键的赋能能力、以及激发突破的加速能力，三者共同形成驱动 AI 转型的动力系统。这些能力实为 AI 规模化的通行证，若缺失任一维度，企业都将困于试点孤岛，难以抵达规模化应用的彼岸。

### 基础能力：

- 构建分布式 AI 平台，灵活应对不同数据安全级别与算力场景。
- 建立健全的风险与治理框架，保障 AI 的伦理、公正与可追溯性。
- 搭建战略合作生态，促进 AI 创新共创与能力共建。

### 赋能能力：

- 构建弹性、安全、可持续演进的基础设施，以支撑 AI 创新与扩展。
- 培养专业 AI 人才，管理并优化 AI 基础设施。

### 加速能力：

- 建立 AI 卓越中心 (CoE)，集中专家智慧并规模化推广最佳实践。
- 强化隐私与安全防护标准，保障合规落地。
- 明确目标与指标，量化 AI 价值。

# 行动指南

## 以基础设施， 赋能 AI 创新

今日的基础设施选择，正绘制明天的 AI 蓝图。精于抉择的组织，方能引领行业变革。无论您正处于起步阶段，或是规模化扩展阶段，以下指南，将助您构建坚实基础设施，真正满足 AI 需求。

### 初级阶段组织

应聚焦基础能力建设：数据统一、治理可信、人才共育。

### 1. 设计面向 AI 的混合基础设施

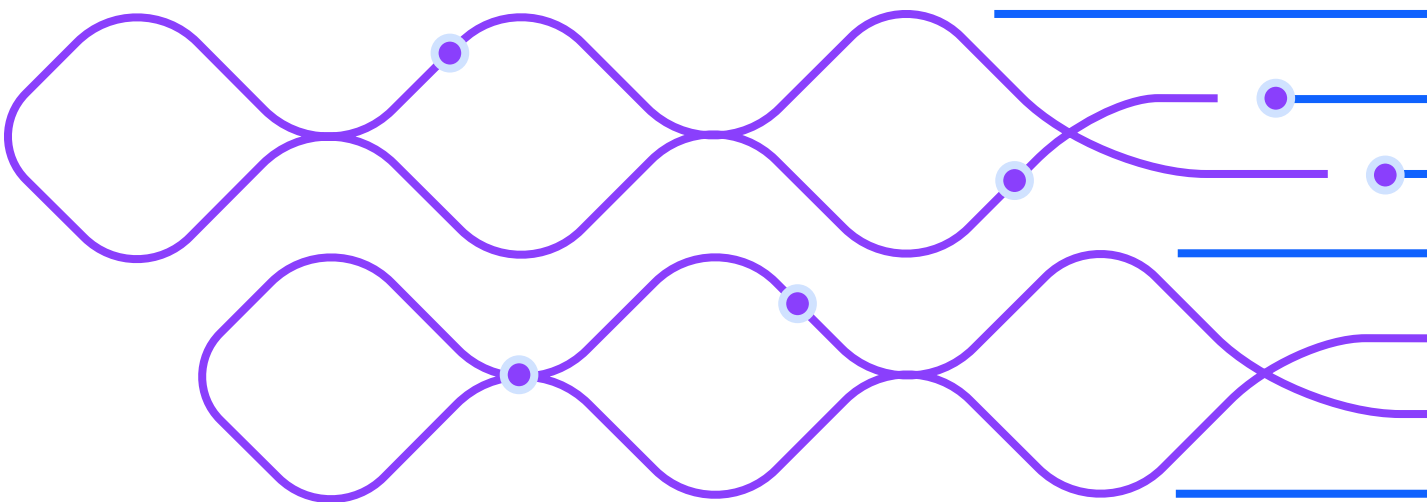
遵循“负载优先”原则，并将其与分布式 AI 平台战略紧密结合。首先诊断现有环境中的数据孤岛与算力瓶颈，找出性能提升的关键点。部署灵活数据结构，整合跨环境数据，实现跨分布式环境的统一数据访问，确保 AI 工作负载不受数据位置限制。构建动态容量规划体系，应对 AI 工作负载的不可预测性。继而依托数据流向、安全等级、延时阈值和算力消耗四大维度，实现 AI 任务与云、本地、端环境的最优匹配。制定技术与业务双维指标，确保基础设施投入与组织战略相辅相成。

### 2. 打造可信 AI 基础设施

秉承可信设计的核心原则，将安全性、合规性与透明性内嵌于 AI 基础设施。构建数据溯源体系与模型登记机制，为 AI 决策建立可审计的轨迹档案。部署持续监测体系，实时发现模型偏差、性能漂移与安全隐患。构建能平滑应对法规调整的架构，确保运营连续与合规安全。

### 3. 规模化 AI 基础设施能力

制定均衡的人才发展路线，清晰区分“自建能力”与“外部引才”方向。设立 AI 卓越中心，打通技术与业务壁垒，推动跨部门协作与经验复用。开展 AI 素养培训，为技术及非技术人员厘清 AI 能力与局限。建立可衡量的评估体系，兼顾团队执行力与伙伴长期价值贡献。



## 成熟阶段组织

打造灵活混合云基础设施、构建制度化治理体系、培育共生型伙伴网络，释放 AI 潜能。

### 1. 打造灵活 AI 基础设施

从人工调度迈向自动化统筹系统，根据性能、延迟与成本动态分配 AI 任务，在云、边缘、本地环境之间智能切换，并针对训练与推理需求分别优化。研发智能体专用统筹框架，以实现多 AI 系统间复杂工作流的协同调度。借助持续反馈机制，实现自我学习与优化，主动提升系统性能与资源效率。此举不仅可降低人工干预风险、加速部署，还能满足生成式 AI 与智能体 AI 对实时响应的高要求。

### 2. 构筑全栈可信 AI 体系

在 AI 开发与部署全流程中内嵌合规审查、偏差识别与可解释机制。借助自动监控与风险预警工具，在规模化运营中持续维系信任，同时不拖慢创新步伐。为自主智能体配置专属监测机制，及时识别异常信息共享或越权操作。部署预测性风险分析体系，精准预判监管趋势，把握政策动向，从而构建信任与创新相互促进、而非彼此制约的文化氛围。

### 3. 人才伙伴网络化协同赋能

突破买卖关系局限，携手懂业务的战略伙伴开展共同创新，让基础设施方案与 AI 运营体系实现基因级融合。发掘深耕特定行业的专业供应商。借助模块化架构与开放 API，加速新 AI 功能的集成与落地。建立聚焦持续学习、跨职能 AI 社群与知识共享的人才计划，培育具备敏捷适应力和复合技能的人才队伍，掌握复杂 AI 基础设施的演进与管理能力。

# 研究和分析方法

IBM 商业价值研究院 (IBV) 于 2025 年第二季度开展了一项全球综合性调研，调研对象为 1200 名直接负责或主导 AI 基础设施战略规划与投资决策的高管。受访者涵盖技术、数据、安全、运营、财务及管理等多维职能角色，包括首席技术官 (CTO)、首席信息官 (CIO)、首席数据官 (CDO)、首席人工智能官 (CAIO)、首席信息安全官 (CISO)、首席执行官 (CEO)、首席财务官 (CFO) 和首席运营官 (COO)。本次调研覆盖 20 个国家、24 个行业，并在不同年收入规模（从 2.5 亿美元至 200 亿美元以上）的企业中实现均衡取样。

## 成熟度与就绪度指数

本调研的因变量为综合成熟度与就绪度指数，整合了两个调研度量项：其一评估组织 AI 运营模式在治理、基础设施、人才、流程、合作伙伴关系、风险管理及绩效衡量等维度的成熟度；其二衡量组织在可扩展性、集成性、适应性、实时处理、多环境部署及合规要求等方面调整人工智能基础设施的就绪程度。两项指标最终整合为统一的序数指标，同时代表当前的 AI 能力成熟度和未来就绪度。

## 分析方法

IBM IBV 研究团队采用了描述性统计与多元统计方法的组合。描述性分析汇总了不同角色、行业、地域及组织特征的分布情况，而交叉分析则考察了按行业领域、国家及角色划分的差异。

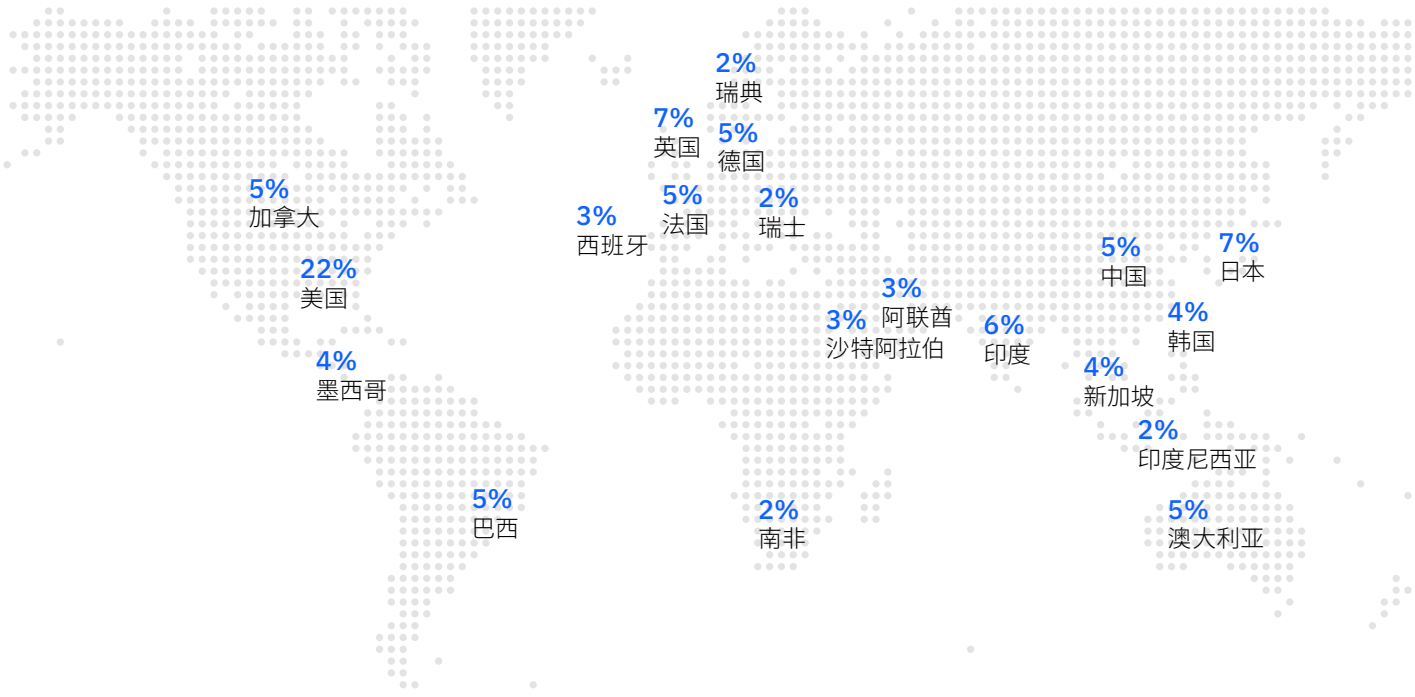
采用有序逻辑回归法进行了高级统计建模，以成熟度和就绪度指数作为因变量。建立了模型以识别 AI 成熟度和就绪度的关键驱动因素。

## 模型验证与显著性检验

拟合优度统计量与似然比卡方检验均证实，两个模型相较于零模型均具有统计学意义上的显著改进 ( $p < 0.01$ )，表明预测变量可靠地解释了综合指数的变异性。多个预测变量均显示出显著性 ( $p < 0.05$ )，这凸显了它们 AI 基础设施成熟度、就绪度及应用成效的影响力。

综合性研究将实地调研与分析方法相结合，深入剖析了领导角色、组织特征与战略实践如何共同推动 AI 基础设施的成熟度与就绪度。该研究结果提供了可操作的洞察，使高管能够将本组织的 AI 基础设施战略与同行进行基准对比，并通过战略性的 AI 实践释放组织潜力。

# 全球样本分布



1,200  
名高管

20  
个国家/地区

24  
个行业

## 关于作者

### *Ric Lewis*

高级副总裁

IBM Infrastructure

[linkedin.com/in/ric-lewis-37713b4/](https://www.linkedin.com/in/ric-lewis-37713b4/)

[Ric.Lewis@ibm.com](mailto:Ric.Lewis@ibm.com)

Ric 的全球职责涵盖四大核心领域：IBM Z® 系统、Power® 服务器、IBM Storage 及软件定义解决方案、IBM Cloud®，以及由 IBM 技术生命周期服务 (TLS) 构建的全球基础设施支持体系。他同时肩负着 IBM 供应链体系的全域战略领导职责。IBM Infrastructure 作为混合云与 AI 战略的核心基石，以顶尖的系统架构、软件定义能力与超大规模集成电路 (VLSI) 创新，为客户提供价值倍增的技术引擎。

### *Hari Kannan*

战略与业务发展副总裁

IBM Infrastructure

[linkedin.com/in/kannanhari/](https://www.linkedin.com/in/kannanhari/)

[Hari.Kannan@ibm.com](mailto:Hari.Kannan@ibm.com)

Hari 主导 IBM 混合云与 AI 的战略规划与生态布局。他广泛携手客户、伙伴及生态供应商，共同识别并破解数字化转型与企业 AI 落地中的核心难题。此前，他曾在 Rambus、Ericsson 和 EMC (现 Dell) 担任领导职务，并在 McKinsey 为客户提供数字化转型与增长战略咨询。

### *Hillery Hunter*

IBM Power 总经理兼 IBM Infrastructure 首席技术官

IBM 院士

[linkedin.com/in/hillery-hunter-97962a14/](https://www.linkedin.com/in/hillery-hunter-97962a14/)

[hhunter@us.ibm.com](mailto:hhunter@us.ibm.com)

Hillery 专注于制定 IBM Power 及基础设施平台的全球战略，并致力于推动 IBM 基础设施业务的跨职能创新与增长。她曾在 IBM Cloud、IBM Systems 和 IBM Research 担任过多种业务和战略要职，并于 2017 年被任命为 IBM 院士 (IBM Fellow)。

### *Robert Zabel*

副合伙人兼研究负责人

IBM 商业价值研究院

[linkedin.com/in/rzabel/](https://www.linkedin.com/in/rzabel/)

[zabelr@us.ibm.com](mailto:zabelr@us.ibm.com)

Rob 专注于以数据驱动的高管洞察研究。拥有 20 余年管理咨询经验，曾协助多家全球 1000 强企业重塑业务、运营与技术战略。

### *Jana Chan*

管理研究顾问

IBM 商业价值研究院

[linkedin.com/in/janachan/](https://www.linkedin.com/in/janachan/)

[chanja@us.ibm.com](mailto:chanja@us.ibm.com)

Jana 负责混合云与基础设施项目。她具备经济学专业背景及十年以上咨询经验，曾帮助政府及大型机构实现数字化转型。



## 特别感谢

Steve Gessner

品牌与内容战略负责人，混合云与 IBM Infrastructure  
杰出营销专家（DMIP 认证），IBM Marketing

Wayne Hickey

营销副总裁，IBM Infrastructure

Chris Priest

IBM Power 和混合云战略总监，IBM Marketing

Pankaj Srivastava

战略与分析负责人，IBM Infrastructure

### IBM 商业价值研究院：

Sara Aboulhosn，创意副总监

Steve Ballou，研究中心总监

Nathan Boudreaux，设计负责人

Kathleen Martin，高级管理顾问

Thiago Sartori，数据科学家

Joanna Wilkins，编辑主管

## 相关报告

### ***Mainframes as mainstays of digital transformation***

IBM Institute for Business Value. October 2024.

<https://ibm.biz/mainframe-hybrid-cloud>

### ***灵活性架构：混合设计运营模式是未来企业的基石***

IBM 商业价值研究院。2024 年 11 月。

<https://www.ibm.com/cn-zh/services/insights/hybrid-by-design>

### ***打造超级生态：混合设计开启强力生态引擎***

IBM 商业价值研究院。2024 年 11 月。

<https://www.ibm.com/cn-zh/services/insights/generative-ai>

## IBM 商业价值研究院

20 年来，IBM 商业价值研究院一直是 IBM 的思想领导力智囊团。我们提供有研究支持和技术支持的战略洞察，帮助领导者做出更明智的业务决策。凭借我们在商业、技术和社会交叉领域的独特地位，IBV 每年都会针对成千上万高管、消费者和专家展开调研、访谈和互动，将他们的观点综合成可信赖的、振奋人心和切实可行的洞察。需要 IBV 最新研究成果，请在 [ibm.com/ibv](https://ibm.com/ibv) 上注册以接收 IBV 的电子邮件通讯，或通过 <https://ibm.co/ibv-linkedin> 在 LinkedIn 上联系我们。

## 选对合作伙伴， 驾驭多变的世界

在 IBM，我们积极与客户协作，运用业务洞察和先进的研究方法与技术，帮助他们在瞬息万变的商业环境中保持独特的竞争优势。

## 备注和参考资料

1. Blair, Kate PhD, Francesco Brenna, Nicholas Fuller PhD, Brian Goehring, and Matt Sanchez. *From AI projects to profits: How agentic AI can sustain financial returns*. IBM Institute for Business Value. June 9, 2025. <https://ibm.co/agentic-ai-profits>. Unpublished data.
2. “Gartner Predicts by 2027, Organizations Will Use Small, Task-Specific AI Models Three Times More Than General-Purpose Large Language Models.” Gartner news release. April 9, 2025. <https://www.gartner.com/en/newsroom/press-releases/2025-04-09-gartner-predicts-by-2027-organizations-will-use-small-task-specific-ai-models-three-times-more-than-general-purpose-large-language-models>
3. “AI is set to drive surging electricity demand from data centres while offering the potential to transform how the energy sector works.” International Energy Agency. April 10, 2025. <https://www.iea.org/news/ai-is-set-to-drive-surging-electricity-demand-from-data-centres-while-offering-the-potential-to-transform-how-the-energy-sector-works>
4. “AI Large Language Models: new report shows small changes can reduce energy use by 90%.” UNESCO press release. July 8, 2025. <https://www.unesco.org/en/articles/ai-large-language-models-new-report-shows-small-changes-can-reduce-energy-use-90#:~:text=Smaller%20models%20are%20just%20as,models%20for%20all%20their%20needs>
5. “New tools are available to help reduce the energy that AI models devour.” MIT News. October 5, 2023. <https://news.mit.edu/2023/new-tools-available-reduce-energy-that-ai-models-devour-1005>
6. Ibid.
7. “The University of Queensland.” IBM case study. Accessed September 11, 2025. <https://www.ibm.com/case-studies/the-university-of-queensland>
8. Rodgers, Clarke, Moumita Saha, Dimple Ahluwalia, Kevin Skapinetz, and Gerald Parham. *Securing generative AI: What matters now*. IBM Institute for Business Value in partnership with AWS. May 6, 2024. <https://ibm.co/securing-generative-ai>
9. *Cryptographic Inventory: Deriving Value Today, Preparing for Tomorrow*. HSBC, Infosec Global, and Thales. 2025. <https://www.ventures.hsbc.com/-/media/ventures/250602-cryptographic-inventory-deriving-value-today-preparing-for-tomorrow-2025.pdf>
10. “Banco do Brasil leads in innovation with responsible AI and ethical oversight.” IBM case study. Accessed September 11, 2025. <https://www.ibm.com/case-studies/bb-ey>



© Copyright IBM Corporation 2025

国际商业机器（中国）有限公司  
北京市朝阳区金和东路 20 号院 3 号楼  
正大中心南塔 12 层  
邮编：100020

美国出品 | 2025 年 11 月

IBM、IBM 徽标、ibm.com 和 IBM Research 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。以下 Web 站点上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表：[ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)。

本文档为自最初公布日期起的最新版本，IBM 可能随时对其进行更改。IBM 并不一定在开展业务的所有国家或地区提供所有产品或服务。

本文档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据的协议条款和条件获得保证。

本报告的目的仅为提供通用指南。它并不旨在代替详尽的研究或专业判断依据。由于使用本出版物对任何企业或个人所造成的损失，IBM 概不负责。

本文所引案例仅供演示说明，不构成实际业务承诺。实际效果因客户配置与运营条件而异，故无法提供普适性成果承诺。

本文件采用符合森林管理委员会 (FSC) 负责任林业认证的无氯 100% 消费后纸张印刷。制造这种纸张的能源是通过可再生绿色能源产生的。请回收。

扫码关注 IBM 商业价值研究院



官网



微博



微信公众号

