Gartner

《网络安全中的AI：明确战略方向》

减少干扰、有效管理风险并全面释放AI技术的价值。

摒弃盲从，最大限度发挥AI在网络安全中的价值

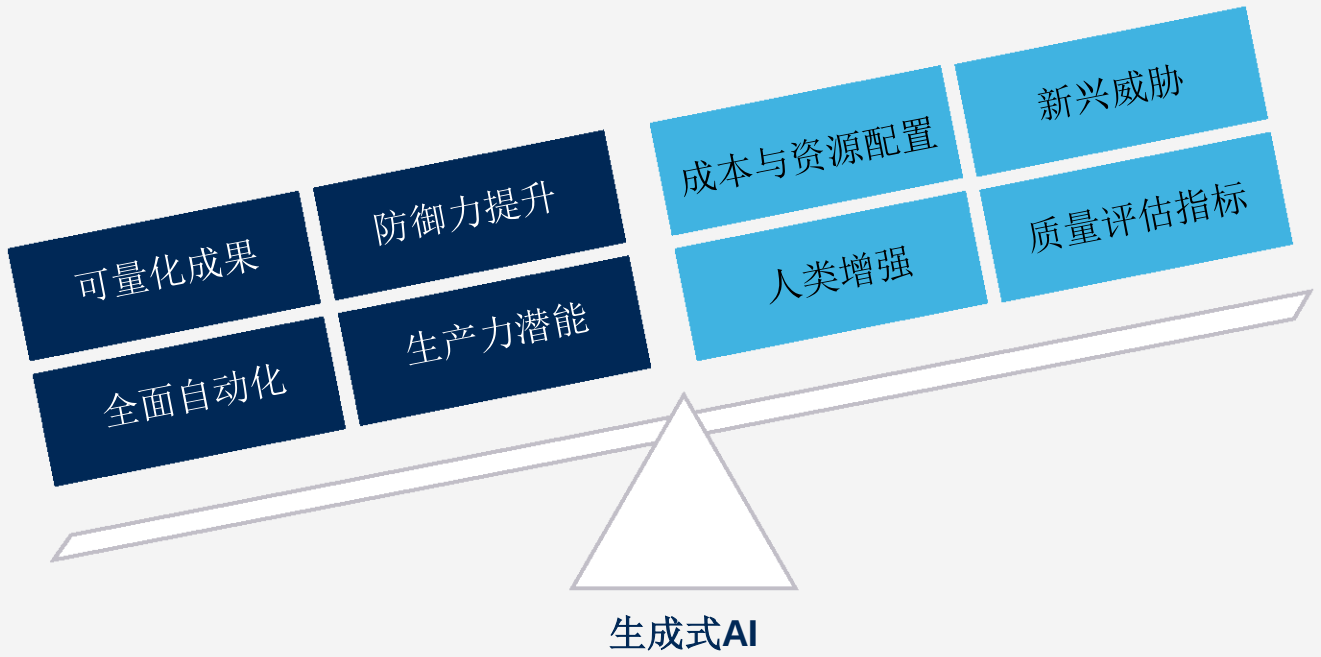
围绕AI及生成式AI（GenAI）的热潮正迅速改变传统商业运作模式，同时也为本就充满挑战的安全环境带来了全新的风险和挑战。风险不断加剧，而AI的潜力却尚未完全释放。

不过，昨日的颠覆蕴藏着明日的机遇。在外界炒作的背后，AI蕴含着深远的战略价值，值得深入挖掘。

AI无疑将重塑企业的运营方式，包括安全领域。随着AI带来的挑战日益凸显及其应用的不断成熟，企业应重点关注以下需求：

- 精准评估AI的影响
- 优先聚焦关键风险领域
- 深化AI价值挖掘
- 预见未来趋势

平衡网络安全现实需求与GenAI愿景



来源：Gartner

精准评估AI的影响

我们的分析显示，**近90%的企业**仍停留在研究或试点生成式AI（GenAI）的阶段，其中绝大多数尚未落实AI TRiSM（信任、风险与安全管理）的技术手段或政策。这一现状正在重塑安全生态。领导者们面临多重影响：

直接且紧急

- 第三方应用对敏感数据的非受控使用
- 版权侵权导致的品牌声誉风险

直接但被夸大

用于激发市场关注的GenAI早期夸大宣传

间接且令人担忧

- 潜在隐私风险及威胁行为者引发担忧
- 业务实践变革带来新的攻击面

间接且潜在

- GenAI的持续应用需求推动安全举措与时俱进
- 新法规与合规要求迫使安全团队提前布局
- 潜在的技能短缺与人才挑战带来不确定性

明确战略方向

引入生成式AI将需要全新或调整过的治理框架，并制定明确的网络安全战略路线图，确保涉及AI的问题得到全面考量。

企业的AI治理范围应根据其成熟度进行定制，但每个企业都应专注于以下三条并行的路线图：

1. **调整应用安全策略以支持AI部署**
   在保障安全开发实践的基础上，防范系统运行与开发生命周期中出现的新攻击面。整合隐私增强技术，评估新引进GenAI技术对应用安全的影响。
2. **将新AI技术整合到网络安全战略中**
   将AI当前和未来的潜在影响纳入三年战略路线图
3. **将AI要素纳入全面的风险管理框架**
   随着企业技能需求的不断演变，评估标准、风险登记册及威胁暴露面也将持续调整。

网络安全领导者对GenAI应用的三大风险关注点：

第三方对敏感数据的获取

GenAI应用与数据泄露

错误决策

来源：Gartner

实施AI信任、风险与安全管理（AI TRiSM）框架

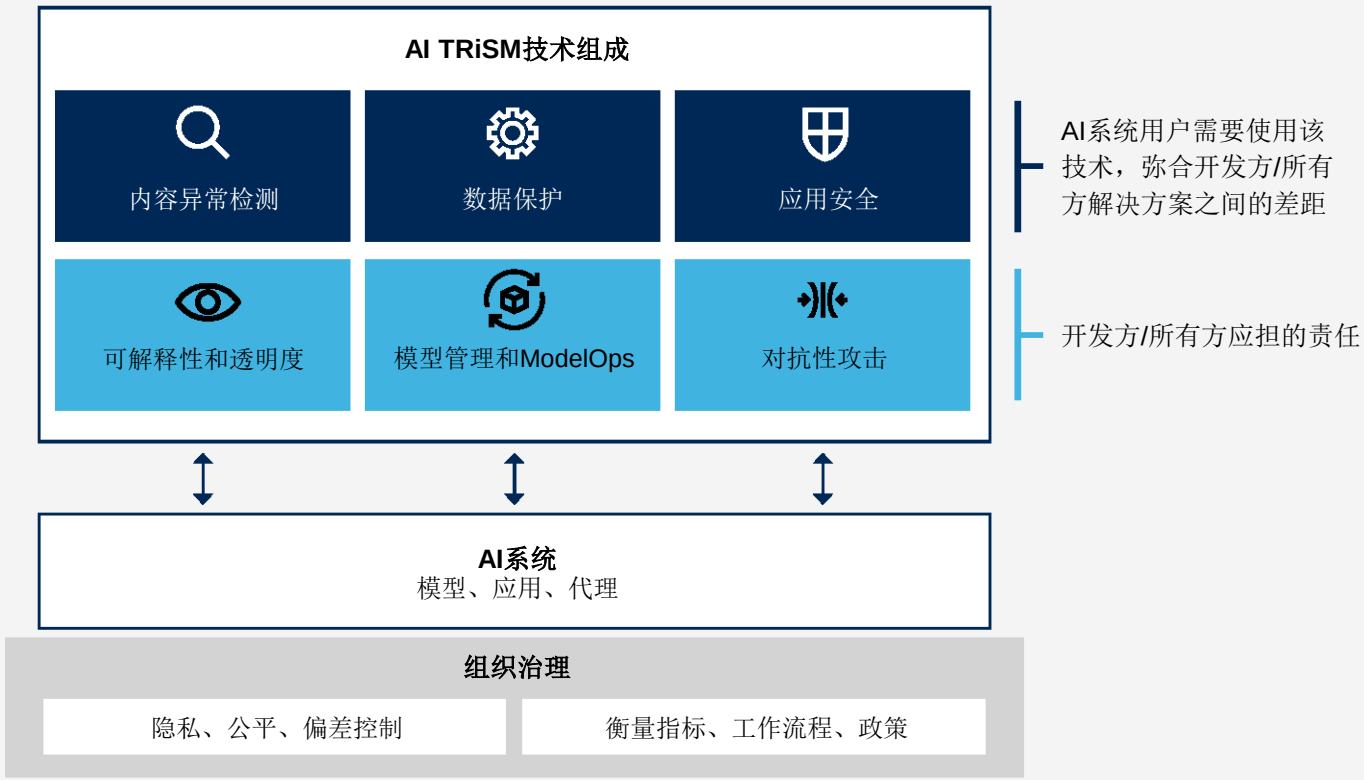
外部托管的大语言模型（LLM）及其他GenAI模型，因其应用流程、数据处理和存储难以被企业直接掌控，显著加剧了潜在风险。

即便是由企业内部托管与管理的本地模型，在缺乏完善安全和风险管理措施的情况下，同样面临挑战。

AI信任、风险与安全管理（AI TRiSM）框架可帮助企业管理相关风险。该框架通过控制措施和信任机制，持续提供：

1. 内容异常检测
2. 数据治理与保护
3. 降低应用安全风险

AI信任、风险和安全管理技术



来源：Gartner

优先强化GenAI应用的安全防护

为Web、SaaS、云基础设施即服务（IaaS）和平台即服务（PaaS）构建坚实的基础安全控制体系，并进一步加强GenAI应用的保护。

Web与SaaS应用的使用

- GenAI可接受使用政策（AUP）
- 用于验证、审批及上线SaaS应用的安全需求清单
- 保护公有云中敏感数据的安全标准
- 安全服务边缘（SSE）产品，保障Web和SaaS的安全使用

云托管企业应用

- 公有云安全使用标准
- 云和web应用安全技术
- 自建应用的安全防护能力
- 机器人检测机制，确保GenAI应用的访问权限仅限人类用户
- 对内外部API端点的保护能力

聚焦三大关键风险领域

GenAI在提升效率和生产力等方面展现出巨大潜力，但也带来了三类全新的风险挑战：

内容异常检测

- 不当或恶意使用
- 幻觉
- 生成内容可能包含不准确、非法、侵犯版权或其他破坏性输出

数据保护

- 数据泄露
- 内容与用户数据受威胁
- 隐私与数据保护政策治理
- 数据隐私影响评估
- 区域监管合规

应用安全

- 对抗性攻击
- 向量数据库攻击
- 黑客入侵

明确战略方向

AI TRiSM是一项跨部门协作，AI、安全、合规及运营团队应紧密配合，推进新的AI TRiSM举措。以下是实施步骤：

- 组建专门的工作小组或设立专责部门，全面负责AI TRiSM的管理与执行。
- 在企业内部推动协作，整合最优工具集，将其纳入完整的AI TRiSM战略框架。
- 制定并落实可接受使用政策，构建系统化的用户应用审批和使用情况记录流程。
- 根据设定的目标持续监控应用使用情况，实时调整使用参数。

预计到2026年，采用AI TRiSM控制措施的企业将至少**降低50%**因不准确或不合法信息导致的决策失误风险。

来源：Gartner

CIO 如何最大程度发挥 GenAI 的潜力

GenAI有望大幅改变安全机制和业务流程。为了实现价值最大化，CIO应优先关注以下事项：

- **盘点、监控并有效管理**第三方GenAI应用及其功能的**使用情况**
- **更新供应商及技术选择标准**，积极应对隐私保护、版权合规、可追溯性及可解释性等挑战。
- **优化AI应用及数据安全策略**，以应对由新攻击面带来的安全风险。
- **在将GenAI纳入网络安全战略前，开展概念验证**，着眼于提升人类工作效能，而非完全替代人类操作。
- **密切监控威胁态势的变化**，如现有安全控制的有效性与检测精度正在下降。确保获得精准且及时的威胁情报，尽管对未来GenAI攻击场景的预判回报可能有限。

2025年，生成式AI将导致用于保障网络安全的资源大幅度激增，**应用和数据安全方面的支出将增加15%以上。**

来源：Gartner

CISO如何最大化挖掘GenAI的潜力

以下是CISO在全面挖掘GenAI价值时应优先考虑的事项：

- **类似于评估其他安全工具的方式，全面评估GenAI技术**是否会利用敏感数据带来新的风险。
- **明确“优良”标准**，用以衡量AI在提升现有安全指标的同时，避免引入新的风险因素。
- 从安全运营和应用安全的特定小范围用例入手，**试用现有安全供应商提供的新增功能**。
- 开发 LLM 和 GenAI 驱动的自有应用，或引入类似的第三方应用时，需全面**遵循 AI TRiSM 框架**。
- 为企业范围内 GenAI 应用带来的直接影响（例如隐私保护、知识产权保障、AI 应用安全）和间接影响（其他部门如人力资源、财务、采购等使用 GenAI）**进行团队培训，做好充分准备**。

预计到 2028 年，生成式增强技术的广泛应用将大幅缩小技能差距，**届时50% 的初级网络安全职位无需专业教育背景也能胜任。**

来源：Gartner

明确战略方向

接下来：

- **系统评估**AI技术，定义符合企业需求的“优质”标准。
- **优化**检测与响应能力，应对未知和复杂威胁。
- **投资**风险暴露管理与威胁情报，以精准识别高优先级威胁。

未来12个月内，**约三分之一（34%）**的企业计划部署GenAI。

来源：Gartner

决定AI战略制定与执行的关键领导角色

CIO/技术负责人

作为CEO、同行和董事会寄予厚望的关键角色，CIO需制定清晰的AI战略（或任命AI团队负责人），并成功达成以下目标：

- 为全企业设定AI战略目标，明确AI应用场景，并量化相关收益与风险
- 统筹业务与技术团队，重塑组织能力，构建全面支持AI应用的运营环境
- 指派AI团队负责人，整合多方创意，推动创新

CISO/安全主管及团队

网络安全领导者需确保网络安全和数据隐私深度嵌入企业AI战略，并达成以下目标：

- 全面监督安全与风险管理计划的执行。
- 预判并有效应对数据泄露、版权侵犯等潜在风险。
- 持续优化技能储备，确保应对新兴威胁的能力与时俱进。

CDAO/数据与分析主管及团队

数据与分析（D&A）领导者需主导数据整理工作，为企业AI战略奠定基础：

- 确定增强数据分析与数据管理的AI应用场景。
- 在现有数据与分析框架的基础上，建立面向AI的数据治理政策。
- 利用AI开发全新数据价值来源。
- 完善数据整理，为AI部署做好准备。

企业架构主管及团队

企业架构（EA）领导者需聚焦AI的实际业务价值，并实现以下目标：

- 制定全面的AI基础设施蓝图
- 掌控AI技术架构的投资决策
- 主导AI解决方案的评估与采纳，以实现业务成果转化

软件工程主管及团队

软件工程领导者需深刻理解AI技术对业务的深远影响，并确保实现以下目标：

- 界定集成AI后需实现的业务成果
- 建立AI工程最佳实践，并在企业内全面推广
- 革新产品、服务和用户体验，并在战略路线图中优先考虑AI应用

我们的研究提供了若干关键洞察，旨在**帮助各职能角色有效采取行动，以实现有价值的AI战略成果：**

	1 设定与业务目标高度对齐的AI愿景	2 精准筛选AI用例并部署测试	3 将AI深度整合于技术架构与业务运营
CIO/技术负责人	按照 Gartner 遴选的最佳实践**审慎选择AI工作重点**，聚焦对业务最具影响力的关键指标	根据潜在业务价值与可行性**优化试点业务**，确保在实现战略目标的同时挖掘其颠覆性潜力。	通过指定专门领导、合理分配资源和资金，并建立严格的规章制度和治理框架，**推动AI部署落地**。
CISO/安全主管及团队	利用AI行为模型提升威胁检测与响应能力，**保持对复杂攻击手段的领先优势**。	通过Gartner的网络安全AI用例棱镜工具，深入分析可行性与风险管理，**筛选出最具潜力的AI应用场景**	与AI团队紧密协作，在AI项目的各个开发阶段进行全面的网络安全评估，确保**有效控制AI风险**
CDAO/数据与分析主管及团队	通过量化AI对特定KPI的预期影响，并设定领先与滞后指标，确保进展可控并**与业务目标高度对齐**	通过聚焦业务价值维度、优化用例并推动决策参与，**提升用例优先级的效率与精确度**	借助数据专家的专业能力，增强跨职能团队协同，应用最合适的技术并降低技术债务，以**确保AI交付高效可持续**。
企业架构主管及团队	通过精准识别需深度调研的领域，制定AI战略与计划，**构建高效的AI生态系统**	运用Gartner四步能力建模法，**以战略眼光规划AI项目**，优化AI基础设施	遵循Gartner五阶段AI执行框架，**确保业务目标的实现，并有效规避失败风险**
软件工程主管及团队	通过引入AI增强的软件工程实践，**推动世界级应用开发与运营**	精准识别AI在软件测试中的关键应用领域，如视觉测试等，以此**最大化挖掘AI价值**	结合人类专家与GenAI的优势，深化解决方案的探索与理解，**激发突破性思维**

可执行的客观性洞察

探索其他免费资源和工具：

电子书
使用信息安全项目成熟路线图保护企业商业资产
制定成熟的信息安全项目，有效降低网络安全风险。

研究报告
为中国数据出境安全评估做准备
提前计划，尽早准备和申请安全评估，以防数据传输中断或新项目上线延迟。

专题页面
什么是网络安全
了解网络安全的定义，以及相关重要议题和解决方案。

网络研讨会
2025年网络安全重要趋势
从“赋能技术及业务变革”和“嵌入安全韧性”两个维度，了解2025年的重要安全趋势。。

获得更多Gartner关于AI的深度洞察

构建价值驱动的AI战略，推动企业增长和创新

使用Gartner AI机遇雷达图识别并确定AI用例机遇

AI数据就绪的关键要素

您已经是Gartner客户？
您可在客户门户网站上获得更多的资源。登录

Gartner

通过参与Gartner会议，推进您的企业AI战略

与同行交流宝贵洞察，掌握如何传达关于AI的机遇与风险；制定战略、执行试点并实现规模化；以及统筹管理AI对企业软件、人才与技能、风险、信任与治理的深远影响。

不容错过。

立即查看会议日历，找到适合您的Gartner会议。

→ 查看安全和风险会议

→ 查看CIO与IT高管会议







联系我们

获得可执行的客观性洞察，针对企业最关键事项做出更明智的决策、推动业绩增长。联系我们成为客户：

成为客户

点击了解更多关于Gartner 网络安全领导者的相关信息
gartner.com/cn

您可扫描以下二维码，关注Gartner官方微信公众号：

Gartner