亚马逊云科技

亚马逊云科技白皮书

亚马逊云科技人工智能、机器学习和生成式 AI 云采用框架

亚马逊云科技人工智能、机器学习和生成式 AI 云采用框架：亚马逊云科技白皮书

目录

亚马逊云科技人工智能、机器学习和生成式 AI 云采用框架

加速云驱动的人工智能转型

发布日期：**2024 年 2 月 13 日**（*文档修订记录*）

在本白皮书中，我们对亚马逊云科技人工智能、机器学习和生成式 AI 的云采用框架进行了概述。该框架描述了一种思维模型，供致力于用人工智能创造商业价值的企业参考。框架中描述了客户企业在提升人工智能和机器学习能力的过程中所经历的发展历程。我们提炼了一系列基础能力，用以构建这一发展历程，帮助企业提高人工智能成熟度。最后，我们概述了这些基础能力的目标状态，并解释如何逐步培养这些能力，在此过程中创造商业价值，从而提供规范性的指导。

人工智能概述

人工智能 (AI) 是一个宽泛的领域，旨在构建或至少模仿能够执行传统意义上需要人类智能才能完成的任务的智能机器。这些任务可能包括从理解自然语言和视觉感知，到决策和解决问题等各个方面。许多人工智能系统的一个共同点是追求概率结果——本质上是生成高置信度的预测或决策，通常能够反映人类判断的复杂性。这样一来，这些系统就可用于自动化或增强知识型工作。

现如今，很大一部分人工智能建立在机器学习 (ML) 的基础上，后者是人工智能的一个分支，专注于开发使计算机能够从数据中学习并基于数据做出决策的技术。机器学习模型不依赖显式编程，而是从样本中归纳总结，使其在众多应用中具有高度的通用性。机器学习领域的各种技术包括深度学习，后者是一个专门的细分领域，旨在利用多层神经网络来分析数据中的复杂因子。深度学习特别擅长处理图像和文本等非结构化数据，并在图像和语音识别等诸多复杂任务中取得了突破。

生成式 AI 给深度学习带来了一项全新的能力，使人工智能能够生成或创作可能具有原创性的新内容。由于能够生成模仿人类思维和推理能力的输出，这门创新的分支学科越来越多地获得人们的认可。算力的增长、数据的可用性和算法创新使生成式 AI 成为了可能，为从娱乐和艺术到科学研究的广泛应用铺平了道路。

总的来说，这些分支学科和技术反映了人工智能层级化却又相互关联的发展趋势，每个层级均致力于开发能够自主执行日益广泛的任务的系统。人工智能的应用和能力有望持续快速拓展，成为我们日常生活不可或缺的一部分，同时也将成为解决复杂问题的重要工具。

“生成式 AI 以鲜有的创新方式激发了人们的想象力。生成式 AI 已经彻底出圈，不再局限于研究人员和开发者的小圈子，从增强消费者体验到解决复杂的企业问题，它展示了无所不包的应用潜力。无论是生成类似人类创作的文本，借助 AI 协助程序员生成代码片段，还是通过智能聊天机器人实现客户互动的自动化，它似乎带来了无尽的可能。除这些应用领域以外，生成式人工智能还充当了催化剂，重新构想了技术如何以史无前例的可扩展性、定制化和智能化融合来增强人类的能力并拓展我们的边界。现如今我们已经站在了大规模采用的边缘，这项技术的潜力不仅在于更高效地完成任务，更在于从根本上重新定义各行各业的可能性。”

——Amazon 首席执行官安迪·贾西 (Andy Jassy)

注：

展望未来，“人工智能 (AI)”一词将作为涵盖其所有子学科的总称使用。当提及 AI 的特定领域时，我们会使用生成式 AI 或机器学习等具体名称，以便与更宽泛的人工智能领域区分开来。

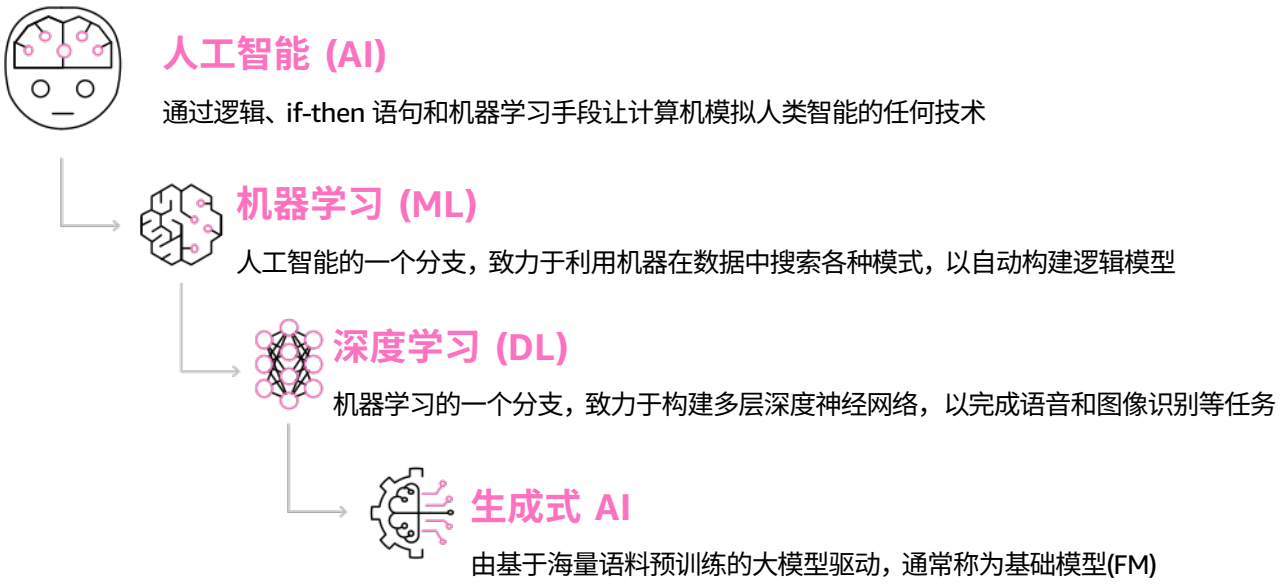

图 1：人工智能、机器学习、深度学习和生成式 AI 的分类

亚马逊云科技 CAF-AI 概述

亚马逊云科技人工智能、机器学习和生成式 AI 云采用框架 (CAF-AI) 既是您开启人工智能、机器学习和生成式 AI 历程的起点，也是指导您持续前进的指南。该框架旨在为您在这些专业领域的中期规划和战略提供启发和指导。在团队内部讨论以及与同事、亚马逊云科技合作伙伴协作时，都可以将其作为人工智能战略探讨的参考资料。

您可能会重点关注并优化特定阶段的技能，也可能会使用整个文档来评估成熟度，指导近期需要改进的领域。具体取决于您在 AI 发展历程中所处的阶段。CAF-AI 是一份不断完善和更新的总结，也是一份企业采用人工智能时需要考虑的所有事项的索引，致力于帮助您超越单一的概念验证 (POC)。我们的目标是为客户提供与亚马逊云科技云采用框架 (CAF) 一致的规范性指导，以便他们成功实施人工智能。在一系列基础企业能力的基础上，亚马逊云科技 CAF 提供了规范性指南，全球数千家企业已成功利用这一指南来加速其云转型历程。

在亚马逊云科技 CAF-AI 中，我们仍然依赖这些基础能力，但我们也丰富了其中诸多能力，使其囊括人工智能所要求的变化。此外，我们还确定并增加了企业在人工智能历程中应考虑的新的基础能力。

亚马逊云科技 CAF：云采用框架

过去十多年来，亚马逊云科技构建了亚马逊云科技云采用框架 (CAF)，为客户的云采用战略奠定了坚实的基础。在该框架的发展过程中，我们在很大程度上避免将其局限于特定技术，而是超越了云本身，以确保来自不同行业的广大客户都能采用其中的洞察和思维模型。然而，人工智能是一种全新的技术，对所有垂直领域和大多数客户均产生了巨大的影响。我们构建了 CAF-AI，旨在帮助我们的客户利用云技术来加速 AI 采用历程。

您是否实现了卓越架构？

亚马逊云科技卓越架构框架旨在帮助您了解在云端构建系统时所做决策的利弊。该框架基于六大支柱，您可以学习设计和运营可靠、安全、高效、经济且可持续系统的架构最佳实践。利用 亚马逊云科技管理控制台 中免费提供的 Well-Architected Tool，通过回答每个支柱的一系列问题，根据这些最佳实践评估您的工作负载。

在《机器学习剖析》中，我们重点阐述了如何在亚马逊云科技云中设计、部署和构建您自己的机器学习工作负载。《机器学习剖析》对亚马逊云科技卓越架构框架中描述的最佳实践提供了补充。

如需获得更多云架构专家指导和最佳实践（参考架构部署、图表和白皮书），请参考亚马逊云科技架构中心。

人工智能云转型价值链

人工智能已从小众技术一跃成为了功能强大且应用广泛的业务能力，机器学习现在正推动新一轮的创新浪潮。在这股创新浪潮中，数据是发明的源泉，而机器学习赋予了企业一种全新的能力，不仅能够描述过去，还能预测未来，并制定有意义的行动计划。由于这种能力能够影响所有市场和企业，各行各业都在加大对人工智能的投入。这种投资可以通过提升客户洞察力、提高员工工作效率和加速创新来建立竞争优势。这一趋势的驱动力来源于人工智能在跨越垂直和水平用例的广泛问题空间中的适用性。

值得注意的是，能够运用人工智能的业务问题空间并非单一的功能或领域，而是在所有业务功能和所有行业领域均有巨大的潜力。在人工智能的确能带来经济效益的市场，它有望重塑竞争格局。对于数十年来一直无法以经济高效的方式解决的问题，或者无法通过人工智能以外的技术手段解决的问题，人工智能能够提供相应的解决方案和解决路径，因此其带来的业务成果可能会产生深远的影响。

举例来说，大型人工智能模型涌现出了的一种新的能力，即在几乎没有额外数据输入的情况下实现特定领域的功能，正使各企业为之震撼，并帮助企业实现差异化能力。这主要归属于生成式 AI 这一领域，目前已经产生广泛的关注度和对技术的想象。然而，这类模型的开发、应用及调优可能是一项非常复杂的任务。

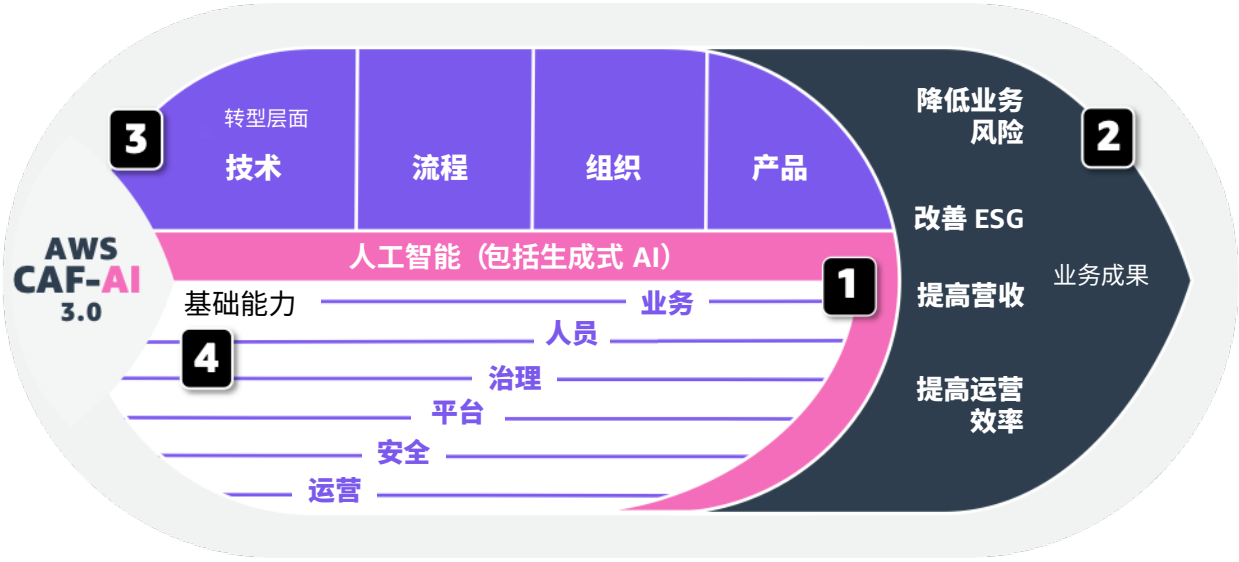

图 2：亚马逊云科技 CAF-AI 的转型价值链（粉色和品红色标注的部分均为我们在此基于原始云采用框架构建的维度）。

面对不断变化的市场格局和快速发展的人工智能领域，上图为人工智能的采用提供了一种思路。

1. 人工智能为您的企业带来新的能力。
2. 有了这些新能力，您和您的企业就能努力创造切实的业务成果。成果可能多种多样，例如降低业务风险（如检测生产线上损坏或有缺陷的零件），改善环境、社会和治理 (ESG) 绩效（如自动汇总并标记环境保护合规报告），增加新的营收来源和现有的营收（如向客户推荐个性化产品和服务），或者提高运营效率（如将差旅收据分类并映射到内部预订代码）。然而，实现这些业务成果取决于您采用人工智能的能力。
3. 要采用人工智能，您的企业需要实现至少四个层面的转型：
    a. **技术：**侧重于构建技术能力，再实现人工智能的使用和采用。
    b. **流程：**侧重于利用人工智能的力量实现业务运营的数字化、自动化、优化和创新。
    c. **企业：**您的业务和技术团队需要协调一致，通过人工智能为客户创造价值并实现您的战略意图。
    d. **产品：**利用人工智能的能力建立新的价值主张（产品、服务）和营收模式，以此重塑您的商业模式。
4. 要实现这些层面的转型和人工智能应用的落地，取决于您在业务、人员、治理、平台、安全和运营方面的基础能力。

想要成功采用人工智能，您需要规划您的这段历程：

- 从您对人工智能能力的认识开始反推。
- 明确您在不同阶段预期达到的业务成果。
- 规划您的企业必须经历的业务转型。
- 发展推动这一历程的基础能力。

您的 AI 转型历程

任何大规模的技术采用计划都是一场漫长的征程，尤其是在采用 AI 等快速演进的技术时。尽管每个企业都有着其独特的技术转型和采用历程，但我们已经观察到了成功采用 AI 的模式。因此，为了帮助客户降低这一过程中的风险，亚马逊云科技 CAF-AI 凭借数千名客户的经验，编写了以下最佳实践观察报告。尽管如此，每个企业在 AI 领域的探索历程仍然是独一无二的。

在踏上或推进您的 AI 转型历程时，请考虑以下四个关键要素，如图 3 所示：

1. 历程的目的地，即您想要实现的**业务成果**，并以此为起点进行反推。
2. 作为历程的驱动力，**AI 飞轮**是一个良性循环，其中初始的优质数据（及时、相关、有价值且有效的数据）被用于训练或微调 AI 系统，然后由该系统产生预测性结果。这些预测性结果对业务成果产生积极影响，进而促进与客户建立更多或更深层次的关系，从而激发产生更多或更优质的数据（网络和飞轮效应）。
3. 您的**数据和数据战略**是保持 AI 飞轮运转的源动力。
4. 您的**基础能力**决定了 AI 采用的成败。

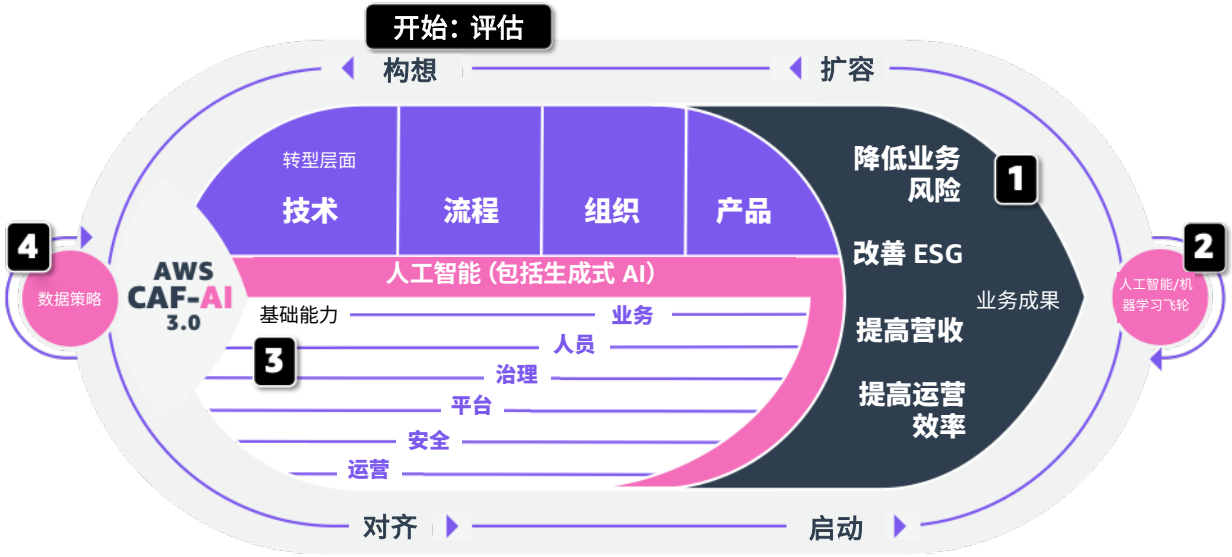

图 3：亚马逊云科技 CAF-AI 云转型历程

在开启这一历程时，请遵循迭代和渐进式改进原则。我们还建议您与您的亚马逊云科技联系人（例如您的客户支持团队）进行沟通，从而获得亚马逊云科技机器学习战略顾问、企业战略顾问和机器学习顾问的协助。在完成初步评估后，技术采用周期将开始，具体包括以下四个阶段：

- **构想阶段：**这一初始阶段主要聚焦于构想 AI 如何助力加速您的业务成果，即根据您的业务目标，对转型机会进行识别并进行优先排序。将您的转型计划与关键利益相关方（即能够影响并推动变革的高级管理人员）以及可衡量的业务成果关联起来。在早期阶段，请务必明确这些计划和机会所依赖的数据资产和数据来源。从机会出发，反向追溯数据需求。
- **对齐阶段：**这一阶段侧重于基础能力的构建。识别跨企业的依赖关系，并明确利益相关方的关注和挑战。与其他技术相比，AI 的采用更是一项跨职能工作。因此，在构想阶段设定的目标上进行内部对齐至关重要。这有助于您制定提升云和 AI 整体就绪度的策略，确保利益相关方的认同和持续支持，并推动相关的企业变革管理活动。
- **启动阶段：**这一阶段重点是交付从早期概念验证到生产部署的试点项目，展示增量业务价值。这些试点项目应对企业和业务产生显著影响，并从应用 AI 中获得实质性的效益。无论成功与否，这些试点项都能为您未来的发展方向提供借鉴。吸取试点的经验教训，有助于您在全面扩展至生产环境之前调整战略和方法。
- **扩展阶段：**这一阶段侧重于在生产环境中扩展试点项目，以实现广泛且持续的价值。这里的“扩展”不仅指扩大解决方案或计划的技术能力，还包括它们在业务和客户中的影响力。此活动可将您的业务活动转化为客户价值。

在这些周期中进行迭代时，要认识到单个周期内可实现的极限。拥有雄心壮志并设定远大目标固然十分重要，但试图在一个周期中完成所有事情可能会导致企业内部产生挫败感。因此，也请务必将宏大的愿景拆解成多个务实、可行的小目标及其可衡量的关键绩效指标 (KPI)。这样，每迈出一步都能让企业更接近目标。不要试图一蹴而就，而要在 AI 转型历程中逐步发展基础能力，提高 AI 就绪度。

基础 AI 能力

要在 AI 转型历程中持续迭代，需要在业务、人员、治理、平台、安全和运营方面具备采用 AI 的基础能力。“基础能力”指的是企业通过流程部署资源（如人员、技术及其他有形或无形资产）以实现特定目标的能力。下图列出了与 AI 采用高度相关的基础能力（以粉色标注），而灰色部分则表示在 AI 采用过程中保持不变的现有 CAF 能力。

图 4：亚马逊云科技 CAF-AI 基础能力

例如，在业务视角章节中提及的产品管理能力。尽管产品管理能力对于成功开发基于云的产品必不可少，但在云端 AI 服务方面，产品管理的实施方式有很大不同。在本文的后续章节中，我们将指出 AI 采用过程中的偏差和特定需求。其他能力请参阅亚马逊云科技云采用框架的原始文档。这些能力分别由哪个管理层级的利益相关方负责，这取决于具体的企业情况。通常，多个利益相关方会对一项或多项能力有共同的兴趣。为了帮助您更好地浏览本文档，我们列出了与各个视角相关的典型利益相关方：

- **业务视角：**这一视角有助于确保您的 AI 投资加速您的数字化和 AI 转型目标，促进业务成果转化。我们丰富了这一视角中的诸多能力，阐释如何让 AI 成为核心驱动力，降低风险，提升客户产出和成果，从而有效制定 AI 战略。典型利益相关方包括首席执行官 (CEO)、首席财务官 (CFO)、首席运营官 (COO)、首席信息官 (CIO) 和首席技术官 (CTO)。
- **人员视角：**这一视角作为 AI 技术与业务之间的桥梁，旨在培养持续成长和学习的文化，让变革成为业务的常态。我们通过关注在 AI 时代对未来竞争优势影响最大的能力来扩展亚马逊云科技 CAF：合适的人才、他们所使用的语言，以及将这些人才凝聚在一起的文化。典型利益相关方包括首席人力资源官 (CHRO)、首席信息官 (CIO)、首席运营官 (COO)、首席技术官 (CTO)、云总监，以及其他跨职能的企业领导者。

- **治理视角：**这一视角有助于您统筹 AI 计划，在最大化企业利益的同时，将转型相关的风险降至最低。我们特别关注风险的变化特点，以及 AI 开发和扩展相关的成本。此外，我们为这一视角引入了一个新的 CAF-AI 能力：负责任地使用 AI。典型利益相关方包括首席转型官、首席信息官 (CIO)、首席技术官 (CTO)、首席财务官 (CFO)、首席数据官 (CDO) 和首席风险官 (CRO)。
- **平台视角：**这一视角能够帮助您构建一个可扩展的企业级云平台，既能运行 AI 驱动或增强的服务和产品，又能开发新的定制化 AI 解决方案。我们丰富了这些能力，以阐明 AI 开发与典型开发任务的不同之处，以及从业者该如何适应变化。典型利益相关方包括首席技术官 (CTO)、技术领导者、机器学习运维工程师和数据科学家。
- **安全视角：**这一视角帮助您实现数据和云工作负载的保密性、完整性和可用性。我们的报告分析主要依赖亚马逊云科技 CAF 的最佳实践，但进一步扩展了如何推断可能影响 AI 系统的攻击向量，以及如何通过云来应对这些攻击向量的方法。典型利益相关方包括首席信息安全官 (CISO)、首席合规官 (CCO)、内部审计负责人以及安全架构师和工程师。
- **运营视角：**这一视角帮助您确保云服务，尤其是 AI 工作负载，能够满足业务需求。我们提供关于如何管理运营中的 AI 工作负载、如何保持它们的运行状态以及如何确保可靠的价值创造的指导。典型利益相关方包括基础设施和运营负责人、机器学习运维工程师、站点可靠性工程师以及信息技术服务经理。

上述每个视角都存在自然或逻辑的顺序。您可以按照这一顺序来提升能力，为您的 AI 转型历程确定当前的行动领域。下图展示了一个示例顺序，以及与经验丰富的 AI 战略实施者共同进行的评估。这有助于确定企业已具备的能力及其成熟度。

图 5：亚马逊云科技 CAF-AI 基础能力（按照成熟度和演进排序）

业务视角：AI 时代的 AI 战略

云技术赋予了企业加速创新的能力，而 AI 和机器学习等新技术范式实现了全新的企业能力、产品和服务。多年以来，复杂的决策过程、非结构化的决策信息数据或不断变化的决策环境等业务问题，难以通过计算机科学的方法来解决。

机器学习领域的最新进展已经改变了这一现状。现如今，那些需要机器进行视觉识别、理解语言、从历史数据中学习并预测结果的问题，突然之间都有了解决方案。这些新兴且触手可及的机器学习能力，正在挑战成熟企业长期以来的市场假设，比如回避驾驶辅助和自动驾驶的汽车公司。因此，业务视角关注的是能够直接帮助企业充分利用这些用例的能力。

基础能力	**解释**
战略管理	借助 AI 和机器学习解锁新的商业价值。
产品管理	管理数据驱动和 AI 增强或驱动的产品。
业务洞察	利用 AI 的能力回答模糊问题，或根据历史数据进行预测。
产品组合管理	明确可行的高价值 AI 产品和计划并确定优先级。
创新管理	挑战长期以来的市场假设，为现有业务带来创新。
新能力：生成式 AI	利用大型 AI 模型的通用能力。
数据变现	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
战略合作伙伴关系	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
数据科学	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*

战略管理

借助 AI 和机器学习解锁新的商业价值。

机器学习可以催生新的价值主张，进而推动业务成果的提升，例如降低业务风险、增加营收、提高运营效率以及改善环境、社会和治理 (ESG) 绩效。因此，您应当首先为您的 AI 技术采用计划定义一个以业务和客户为中心的核心目标，并为其制定一个切实可行的逐步推进战略。在制定任何采用战略时，必须确保其基于可实现的（短期且可衡量）或有抱负的（长期且难以衡量）业务影响，充分利用 AI 带来的新能力。制定策略时要全面考虑采用 AI 的短期和长期影响。

从现有的业务和客户问题出发，反推 AI 能够如何对这些问题产生影响。在逐步确定 AI 机会的优先级时，须关注如何以及哪些数据能够推动系统能力的提升。从一开始就考虑任何机器学习产品或服务中的数据飞轮效应，即新数据推动系统改进，进而扩大客户基础，反过来又增加了企业能够从中受益的数据量。

在构建这种数据飞轮时，您需要思考所获取的数据是否能为您的价值主张提供一道防御性壁垒（即稀缺且成本高昂的资源）。鉴于 AI 技术已经对市场格局产生了广泛的影响，我们必须考虑到，在不久的将来，客户很可能会对您的产品和服务能力提出更高的期望，而 AI 能力正是这些期望的一部分。

针对每一个 AI 机会，我们需要评估是否需要构建全新的 AI 系统、调优或调整现有系统，还是直接采用现有的 AI 系统。例如，如果您期望利用基础模型涌现出的无所不包的新能力，但缺乏从头开始创建它们的能力，那么您应专注于根据您的特定需求进行系统定制。如果您的目标是创建一个推动业务发展的特定领域通用系统，那么您应更多地投资数据基础建设。

产品管理

管理数据驱动和 AI 增强或驱动的产品。

构建和管理基于 AI 的产品可能是一项重大挑战，因为 AI 系统的开发和生命周期与传统软件和云产品有所不同。无论是开发，还是运营以及持续创造基于 AI 的任何产品的成果（如直接预测），都存在潜在的高成本不确定性，需要特定的应对策略。

在构建 AI 产品或将 AI 嵌入产品时，应从客户和用户预期的价值增益出发反向追溯，将可衡量的业务代理映射到 AI 系统可以支持、丰富或自动化的各个决策点。对于每一个决策点，均需在机器学习解决方案领域内定义潜在的指标（例如，在金融领域，检测欺诈交易的价值增益如何转化为预期的货币收益，以及相关的机器学习驱动的交易分类器的准确性或召回率），并明确对应的机器学习问题（如分类问题、意图提取问题、生成式 AI 等）。这些明确定义的机器学习问题及其各自的解决方案共同构成了机器学习可为您的产品带来的价值增益。

至关重要的是，这些机器学习解决方案对您和您的产品提出了特定的数据要求，因此必须挖掘每项解决方案的 4V 数据特征。在自下而上构建这一知识体系的过程中，请确保将业务、数据、管理层和机器学习领域的利益相关方纳入解决方案的评估中。机器学习产品将数据、特定领域的专业知识和技术融为一体，形成了一个可以预测，有时还能提供指导意见的系统。因此，数据、业务领域知识和技术这几个方面的人员都必须参与其中，缺一不可。请通过适当的生命周期管理来铺就基于 AI 的产品演进之路，考虑用户如何与基于概率的 AI 系统输出进行交互（例如，在系统置信度较低时优雅地处理失败情况），并评估您的解决方案在被采纳后可能产生的影响，以确保负责任地使用 AI。

在正确界定您的产品的机器学习能力范围并提升 AI 产品管理能力方面，有几个关键问题至关重要。例如，采取实验性、通常有时间限制的方法来降低机器学习组件的风险，并从一开始就考虑如何将这些实验中的学习成果转化为生产级系统。同时，这也意味着需要在系统信息流中设计反馈循环（或明确防止其发生），从而通过如数据网格（或数据区域）和数据湖架构等技术，以及团队和产品组之间的知识传递（例如通过 SageMaker Model Cards 实现），让更广泛的企业能够基于其他机器学习系统的输出构建新的 AI 产品。

业务洞察

利用 AI 的能力回答模糊问题，或根据历史数据进行预测。

商业智能 (BI)，主要涵盖描述性分析和诊断性分析，通常是企业采用 AI 的起点。然而，除描述性和诊断性分析以外，机器学习还赋予了预测性乃至指导性能力，这两者共同构成了 AI 发展的道路。关键要认识到，分析与商业智能部门的范围，与企业对 AI 驱动部门所期望的有所不同。

如今，许多企业需要领域专家 (SME) 来筛选见解，并从数据中找出某些观察结果的原因（即“为什么”）。然而，通过 AI 技术的运用，商业智能开始辅助这些领域专家，通过*识别“为什么”和“如果会怎样”*来为他们提供新的见解，进而融入他们的思考过程。数据和 AI 因此一跃成为预测性决策的驱动力。

从商业智能实践过渡到 AI 驱动的实践，全面提升分析层次时，突破瓶颈的有效方法是，利用诊断性分析算法来找出影响问题陈述的关键变量或根本原因。企业成熟度分析不应局限于各个部门，要考虑如何促进成熟企业与不太成熟企业之间的交叉融合，加速您的 AI 历程。

在转型初期，一个行之有效的方法是，建立一个与您的云计划密切相关的分析卓越中心（不一定专门针对 AI）。这样的卓越中心 (COE) 可通过普及 AI 的使用来提供数据驱动的预测和分析，从而创造即时价值，推进您的宏伟目标。最重要的是，养成使用 AI 来指导重大业务决策的习惯，因为这将促使员工认识到 AI 对实际业务成果的价值。

产品组合管理

明确可行的高价值 AI 产品和计划并确定优先级。

机器学习计划的挑战在于，必须在不牺牲长期价值的前提下展示短期成果。在最坏的情况下，短期思维可能导致技术性的 AI 概念验证 (POC) 仅停留在技术阶段，因为它们过于关注与业务无关的技术细节。在明确机器学习计划和产品、确定优先级以及实际落地时，您的首要目标必须是实现可衡量的业务成果。

关键是从小目标着手，达成这样的目标可以增强企业内部的信心，让员工认识到 AI 可以在业务的其他领域发挥作用。同时要考虑，您正在通过多个 AI 项目和产品解决哪些更大的客户和业务问题，并将其整合成一个产品组合，由低层次的项目为高层次的项目提供支持。某些 AI 能力无法一蹴而就，而是需要在彼此的基础上构建。例如，在金融行业，在向客户推荐新产品之前，您必须能够对当前重要的内容进行分类，因此交易分类是下一步最佳报价行动的前提。您的产品组合中的每一层都应该为整个企业创造额外的价值。

接下来，在这个产品组合中引入 AI 飞轮设计，通过产品组合提供的价值推动业务成果，而这些业务成果反过来又能够产生并创造更多的数据，使产品组合自身受益。这一飞轮不必局限于单一产品层面，而是可以贯穿整个产品组合。随着产品组合的发展和扩大，确定对外采购还是自主开发的优先级变得至关重要。要克服“非我发明”的情结，充分利用外部现成的解决方案。

为此，应当提前而非事后才去探索市面上已有的用例和解决方案及其成熟度。同时还应调查哪些解决方案需要定制建模，并通过选择合适的 AI 产品和云环境来提高 AI 人才的工作效率。应意识到，单单在技术层面上管理您的组合，就已经是一项复杂的任务。为确保稀缺的 AI 人才保持较高的工作效率，您需要果断大胆，并克服分析性瘫痪。

最后，随着您的组合的增长，企业内部越来越多的部门开始使用 AI，请确保您的业务部门、团队以及您所依赖的亚马逊云科技合作伙伴之间能够进行高效协作（请参阅 Amazon DataZones、Amazon Redshift 和 Amazon CleanRoom）。

创新管理

挑战长期以来的市场假设，为现有业务带来创新。

如本部分引言所述，机器学习为企业带来了全新的能力。在许多情况下，这些能力可能会颠覆现有的业务和价值链。各行各业都已经见证和感受到了这种通用技术的力量，因为 AI 研究的长远目标就是复制或者至少模仿人类智能。以往只有人类才能完成的知识性工作、处理复杂信息、推理洞察并采取行动，现在均可通过先进的基础模型和生成式 AI 来实现。在您的创新路线图和创新管理实践中，可通过切实可行的短期价值主张来接轨 AI 研究的这一中长期目标。

为此，首先要从内部和外部两个方面入手，探索不断变化的客户期望和需求。CAF-AI 提出的业务成果可指导您识别这些需求和期望。分析采用驱动或融合了机器学习的产品价值链，区分三种创新：通过流程改进等降低成本；通过产品改良提高营收和利润的创新；通过提供创新产品和服务开辟全新营收渠道。

利用机器学习，将之转化为内部利益相关方和外部客户的独特优势。将机器学习与自动化相结合，解锁新能力、增强现有能力并减少工作量。挖掘并深入开发您所访问的数据中蕴含的特定领域知识。为您的 AI 系统设计一个良性的数据价值链，以持续创造价值。一些基于机器学习的产品只有经过时间的积累和迭代才能不断完善，您的创新周期可能比某些公司习以为常的周期更长，但不必因此气馁。在为基于机器学习的产品逐步构建起单一产品线的同时，也要将数据提升为价值创造过程中的头等要务，构建供内部使用的数据产品，从而为整个企业的创新铺平道路。

除了这种自上而下的创新管理方法，还要在内部的 AI 倡导者中开展自下而上的运动。这些倡导者可以是业务主管、产品经理、技术专家，也可以是企业高管。要在宏伟目标和可实现目标之间取得平衡。普通软件系统和软件环境主要靠获取更多用户来提升自身的价值，而机器学习系统的价值主要取决于提高其效率的数据。因此，管理 AI 创新就是要落实数据策略，而不只是仅将历史数据存档。随着可管理、可访问的高质量、高价值数据在整个企业中不断积累，您的 AI 创意和项目将变得极具吸引力。

新能力：生成式 AI

利用大型 AI 模型的通用能力。

AI 技术的总体目标是构建高通用性的系统，并能以极低的成本应用于诸多复杂的问题空间。在这项工作中，生成式 AI 便是一个非常强大的分支。这种 AI 技术能够生成新内容和新创意，包括生成对话、故事、图像、视频和音乐等。生成式 AI 由基于海量数据进行预训练的超大模型（通常被称为基础模型，FM）提供支持。这些基础模型的潜力在于能够跨越不同领域和任务实现泛化。这些基础模型将以某种方式影响您的企业和业务，因为它们能够极大地降低知识性工作的成本。在计划采用这一强大的 AI 技术分支时，您需要考虑三个因素。在构建此类基础模型时，您是否需要：

1. 从零开始，专门为您的业务量身定制？
2. 微调预训练模型，利用其已经习得的能力？
3. 直接采用供应商提供的现成基础模型，无需进一步微调？

在这三者之间做出选择是至关重要的。正确的选择取决于您的业务场景。通常，要真正释放这些大型模型的价值，就意味着要用您在特定领域的数据为模型提供上下文（第 2 种选择），再将模型应用于各种任务。这是因为，经过预训练的大模型已经涌现出了一些新的能力（如推理能力），而要从零获取这些新能力（第 1 种选择）需要付出高昂的成本。因此，在使用基础模型和生成式 AI 时，可利用经过预训练的模型的适应能力和从少量数据（甚至零数据）中学习的能力。

对于许多企业而言，这意味着需要针对自身的业务问题选择合适的基础模型，再利用特定领域的数据或客户专属数据对这些模型进行定制（例如指令调优和少样本学习）和微调。和其他 AI 系统一样，生成式 AI 系统和基础模型的效能和差异化能力，很大程度上取决于您的数据策略和数据飞轮。无论您选择哪种方式，都应重视所使用的数据质量，因为数据会影响到模型在生产环境中的行为，而且要为生成式 AI 系统建立防护措施是非常困难的。

人员视角：AI 优先的文化和变革

可靠、可重复地采用 AI 创造价值，不仅仅是一项技术挑战。任何 AI 计划的成败都取决于保驾护航和推动落地的人员。虽然 AI 作为一种通用技术将影响各行各业，但只有员工接受 AI 的能力，企业才能取得成功。考虑到优秀 AI 系统的落地需要利益相关方、业务部门和实践部门之间的协作，这一点尤为重要。

人们常常谈论 AI 取代人力劳动的潜力，而实际上，AI 更多地扮演着丰富、补充甚至赋能人类工作的角色。虽然某些领域有望实现自动化，但当前的 AI 技术主要在于帮助人类完成在人看来特别复杂的任务。我们观察到，倡导 AI 优先的企业降低了运营成本、增加了营收，员工也得以从事更具挑战性、更有意义的工作。凝聚企业力量、培养合适的人才，并在发掘有价值的商业问题时使用相同的语言，是这一视角的重点所在。*文化为上*——在采用 AI 的历程中更是如此。这一视角包括下表所列的七项能力。典型利益相关方包括 首席信息官 (CIO)、首席运营官 (COO)、首席技术官 (CTO)、云总监以及跨职能部门领导者和企业领导者。

基础能力	解释
新能力：机器学习熟练度	建立共同的语言和心智模型。
人才转型	吸引、启用和管理 AI 人才——从使用者到构建者。
企业对齐	加强和依托跨企业协作。
文化演进	*文化为上*——在采用 AI 的历程中更是如此。
变革型领导	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
云熟练度	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
企业设计	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*

新能力：机器学习熟练度

建立共同的语言和心智模型。

人工智能和机器学习的界限和语义范畴并没有良好的界定。二者还承载着各种心智模型和情感解释。因此，从内部统一利益相关方对这些术语的理解至关重要。要传播对这些术语内涵大体一致的认识，进而确定对其感兴趣的利益相关方，即潜在的内部 AI 倡导者。

一旦第一层解读在企业内部普及开来，接下来就需要着手解决第二层，即更为技术性的解读问题：AI 项目及其需求在术语使用和重视程度上可能存在差异。从产品管理实践到工程和数据科学实践，各方需要就有效合作所需的共同理解达成一致。一个行之有效的方法是，定义不同实践之间的衔接词汇，例如如何在机器学习中衡量成功，以及如何在业务领域中衡量成功。

通过机器学习熟练度和机器学习文化培训来实现对齐，因为这有助于您在整个企业中获得支持。在助力业务负责人适应机器学习用例的独特方面以及设定客户期望方面，这种理解可能会变得至关重要。

最后，无论是在企业内部还是对于客户，都要考虑如何才能以最佳的效果传达 AI 的产出成果。考虑到客户的心智模型和对术语的解读有所不同，要让 AI 系统体面地面对失败，同时维持客户的信任度，是一项极具挑战的任务。使用合适的语言并具备合适的熟练度，不仅能提高沟通效率，还能降低构建有悖客户利益的系统的风险。

人才转型

吸引、启用和管理 AI 人才——从使用者到构建者。

吸引、留住和再培训能够推进 AI 战略的人才，是 AI 成功的关键因素之一。AI 战略的成功需要多个角色的参与。有些角色可以外包，而有些角色只能由内部员工担任。首先，您的 AI 战略领导者需与您的业务紧密联系，从内部推动价值创造。这个角色很难外包给第三方公司。

招聘或培养诸多亟需的角色来支持这些领导者，从而取得 AI 采用的成功：

- 技术人才（如数据科学家、应用科学家、深度学习架构师和机器学习工程师）。
- 管理路线图并识别需求的非技术产品人才（如机器学习产品经理、机器学习策略师和机器学习布道师）。

招聘战略应与整体的 AI 战略和目标保持高度一致：

- 资深的博士生可能适合雄心勃勃的大型科研项目，但最好能与那些同业务联系紧密的同事（如机器学习策略师）形成互补。
- 让部分现有人才转型到 AI 岗位，有利于 AI 在整个企业中的采用。
- 如果您计划基于成熟的解决方案、基础模型，或者引入企业能力范围之外的 AI 成果来构建 AI 能力，招聘机器学习工程师和深度学习架构师是较为明智的选择。

除了这支内部团队之外，建议您尽早与合适的亚马逊云科技合作伙伴合作，避免您的 AI 计划无法落地。人才匮乏时，要对外传播 AI 愿景，启动能产生成果、吸引新人才的项目。从一开始就要认识到留住 AI 人才的困难，因为这类人才历来都是供不应求。另一个因素是，现实世界的 AI 与通常驱使人才涉足 AI 领域的学术工作有着显著的不同。要应对这一差异，尽可能创造机会，让您的 AI 专家开展合作、出席会议并撰写白皮书。

然而，人员流失总是在所难免。要具备灵活性，建立招聘流程，及时补充人才，在人员流失时保持资源到位。我们在 CAF-AI 的其他部分中提到的流程，对于在面对人员流失时保持业务稳健至关重要。要为 AI 员工持续提供再培训机会、学习在 AI 领域表现出色所需的新技能。这种方法的另一个优点是，员工既可以积累深厚的业务知识，又能够执行项目。最后，要认识到 AI 领域的人效比要高于其他领域。优秀的小团队通常比大团队表现得更好，因为这类工作更多是智力性工作，而非机械性工作。

企业对齐

加强和依托跨企业协作。

当 AI 成为各企业的首要考虑因素时，第一步通常是成立一个自成一体、被赋权的、独立运作的单位，以此传播和传递 AI 的价值和知识。AI 卓越中心便可发挥这一作用，招聘和培养专注于 AI 的团队。要确保企业的汇报线与 AI 战略的利益相关方对齐，并缩短高层的汇报线。这是为了确保在需要时能够快速做出决策和变更，并让新团队找到自己的节奏。同时，关键是要将此类卓越中心的激励措施与您的战略、业务和您的客户（最重要的维度）对齐。一个常见的错误是，一手建立起来的 AI 团队无法创造业务价值。

随着时间的推移，您的人才转型应能够让您企业中的更多人和其他构建者有效地使用卓越中心和现有的 AI 服务，并有效地开展协作。要杜绝“非我发明”的心态，如果云上已有解决方案可满足您的业务需求，企业就不必从零开始构建。确保您的卓越中心和人才培养一种工程思维，认识到维护不同系统的成本，并建立机器学习运维最佳实践，从而在文化中引入 DevOps 思维。随着此类部门、其他内部构建者和 AI 人才的发展，应培养数据驱动的产品思维来推动您的数据飞轮。不仅要让企业内不同业务部门共享和管理数据，还要打造一个充满活力的数据产品生态。但是，不要为了数据产品本身而构建数据产品。

文化演进

文化为上——在采用 AI 的过程中更是如此。

发展 AI 优先的文化是一个漫长而富有挑战的过程，因为这通常需要打破旧有的心智模型。在传统的云开发和软件开发中，文化焦点在于赋权构建者将复杂的规则和系统编写成代码。而 AI 更多地依赖于这样一种文化：寻找正确的输入，以生成期望的输出。为避免以技术为中心的文化，需要拥抱这样一种心态：构建者、企业和其他利益相关方要基于业务机会和客户需求着手解决问题，再进行反推，直至解决诸多 AI 挑战。

反推意味着预先确立业务环境变化的预期结果，再思考“要实现这一变化需要做什么”。在某种程度上，这就是 AI 系统的构建方式：定义预期的输出，然后寻找能产生该输出迹象的输入。

基于这种价值驱动的思维模式，关注构建 AI 优先文化的基础要素：

- 将试验心态与敏捷工程实践相结合
- 跨团队和跨业务部门的协作与依赖
- 自下而上和自上而下地发掘 AI 机会
- 以客户价值为导向设计全面兼容的 AI 采用方案

您可以通过以下方式开始培养 AI 优先的文化：

- 鼓励构建者敢于尝试 AI 系统，不是为了尝试而尝试，而是因为构建 AI 系统本身就需要不断探索，找出可行的解决方案，避免走进死胡同。采用路径明晰的现有 AI 服务有助于降低风险。

  在鼓励尝试的同时，根据 AI 的不确定性调整敏捷思维方式。要认识到，面对复杂项目时，您无法可靠地预估所要投入的时间和工作量，因为许多业务价值较高的复杂 AI 问题尚未解决。在这种情况下，要加倍投入那些有望产生最大客户价值的项目。

- 拥抱这样一种文化：各团队以数据作为纽带，共创价值。不要建立脱离业务的数据科学团队，而要营造一种能驱动协作飞轮的文化。
- 倡导这样一种文化：能在企业各个层面发现、认可和实现价值。这包括领导层要激励和提拔敢于挑战现状的员工。
- 营造这样一个环境：对 AI 影响和应用的关注不只是纸上谈兵，更要落实到决策过程中。

治理视角：管理 AI 驱动的企业

管理、优化和扩展企业的 AI 计划是治理视角的核心。将 AI 治理纳入企业的 AI 战略中，对于建立信任、大规模部署 AI 技术，以及攻克挑战、推动业务转型和增长至关重要。通过推动一致性，AI 治理能够促进与企业目标的对齐，并确保 AI 技术的运用符合伦理规范并得到有效的管理。为此，AI 治理框架在企业中创建一致的实践，以应对企业风险、符合伦理规范的部署、数据质量和使用，甚至是监管合规性，并管理 AI 工作负载的不同成本模式。制定一套可扩展的 AI 部署流程和标准，能够帮助企业将 AI 计划从个别业务部门推广至整个企业，进而在更大范围内创造持久的业务价值。

建立 AI 治理实践需要与企业的 AI 战略紧密协调。第一步是确认所有关键的利益相关方，并组建一个由多个业务单元的代表组成的团队。该团队的职责如下：

- 定义治理目标，包括合规目标与伦理目标，并识别具有潜在风险的领域。
- 制定涵盖数据、透明度、负责任 AI 和合规性的政策和指南。
- 确立相应的机制，以监控 AI 系统、性能、合规性和偏差，并根据预定义的阈值确定需要采取的行动。
- 持续修订成果和现有政策，以确保与业务目标和 AI 安全对齐。

在治理视角中，我们针对治理过程中遇到的挑战描述了若干解决方案，并介绍了一项新的能力：负责任地使用 AI。这是未来在 AI 领域获得竞争优势的决定性因素。

基础能力	**解释**
云财务管理	在云端规划、测算和优化 AI 使用成本。
数据策展	基于数据目录和数据产品创造价值。
风险管理	利用云服务来缓解和管理 AI 固有的风险。
负责任地使用 AI	通过负责任地使用 AI，持续推动 AI 创新。
规划和项目管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
数据治理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
效益管理	*该能力未针对 AI 进行增强，请参阅亚马逊云科技 CAF。*
应用组合管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*

云财务管理

在云端规划、测算和优化 AI 使用成本。

在云端管理 AI 项目需要针对训练和推理的成本结构进行规划。在为单个项目制定预算，以及为 AI 计划制定整体的拨款预算时，就需要提前考虑这一点。关于 AI 生命周期中的这种成本结构，这里提供了一个*锯齿形*成本（或者称为低/高/低/高成本阶段）的示例：

- 初期您可能面临较高的成本，以建立构筑解决方案所需的数据，或者提高数据的质量。但如果您的数据已准备就绪，初始成本可能很低。接下来是一个变数较大的概念验证阶段。
- 虽然大多数 AI 相关的概念验证项目的算力成本可能较低，但有一些技术方法可能会迅速让成本变得十分高昂，例如（在生成式 AI 的背景下）训练大型模型，或者为特定领域的机器学习模型进行持续的重新训练。在此类情况下，您可利用专门打造的 AI 硬件来降低成本，比如由 Amazon Trainium 提供支持的 Amazon Elastic Compute Cloud (Amazon EC2) Trn1 实例，或是由 Amazon Inferentia2 提供支持的 Amazon EC2 Inf2 实例。如果您拥有合适的人才、 AI 服务和亚马逊云科技合作伙伴，您可借助他们的专业知识来评估用例不同阶段和整体 AI 战略所需的资源。如果可行，核算一下对一个机器学习指标进行渐进式改进需要投入的成本，以此决定如何优化您的投资。
- 系统完成第一次迭代后，下一个阶段（即打造最小可行产品）可能需要较高的成本；例如，推广系统的能力，或者获取对于用户采用至关重要的边缘案例和长尾数据。如果您的用例需要用到生成式 AI 的能力，您可以直接使用或微调基础模型，这将对成本产生显著而积极的影响，因为您的供应商已经承担了模型的初始训练成本（例如，Amazon Bedrock Titan 基础模型）。
- AI 模型部署完毕后，推理本身在很大程度上取决于请求量，在很多情况下，推理的成本是比较低的。否则，您可采用专门构建的 Amazon Inferentia 架构。在这一阶段，监控模型指标并标记漂移，可提醒您发生的变化，并确定是否有必要重新训练您的算法。在云端，您可以利用扩展资源的低成本优势。与此同时，在整个 AI 生命周期中，跟踪成本并标记所有的资源和机器学习工作负载也很重要。

在确立成本可视化机制之后，分析数据、训练和推理的成本随时间的变化就变得至关重要。各类问题（文本、预测、文档处理等）会层出不穷，它们的初期成本并不高，但随着数据量的增加，成本会呈线性增长。还有一些依赖音频和语音数据的 AI 问题，它们的启动成本较高，即使是在概念验证阶段也需要明确的目标，以免产生意外的费用。将您的 AI 愿景与业务目标对齐，应当能指导您如何确定工作范围；建立一套机制来权衡模型成本和模型性能，这对于维持正投资回报率至关重要。此外，数据获取成本在很大程度上受到企业围绕其数据流程建立的机制的影响。针对新数据和主数据的获取建立一个标准流程，这是降低成本的关键，将数据保存为 AI 可用的格式亦是如此（减少复制/读取/复制或抽取、转换、加载 (ETL) 的需求）。而通过治理良好的数据服务和零 ETL 模式，您可在云端解决所有这些挑战。

此外，始终将您的 AI 计划与潜在的业务目标联系起来。倘若涉及新的营收来源，就要假设有多少收益与哪些成功标准相关，并将业务价值转换为您的 AI 指标。如果未能认识到负责任地使用 AI 的必要性所带来的潜在成本，就可能会低估 AI 系统的总体成本，因此在评估成本时务必将这一点考虑在内。由于这一点很重要，我们在治理视角的内容后面新增了负责任地使用 AI 这一能力。

数据策展

基于数据目录和数据产品创造价值。

您获取、标记、清洗、处理、交互数据的能力，将加快您的进度，缩短价值实现周期，并提升模型的性能（如准确性）。当模型因准确性而停滞不前时，可考虑回过头来丰富、增加或改进输入算法的数据。这通常比单靠重构模型或竭力榨取模型百分之一的性能要容易。

以机器学习为中心的数据收集对于实现您的 AI 路线图至关重要，您应该和其他领导应思考以下问题：“我们能否通过普及数据访问和使用来推动 AI 创新？”“我们是否将数据视为一种产品？”“在整个企业中能否发现我的数据？”这些问题的答案通常不是非黑即白，而是介于两者之间。但关键是要记住，一切都是为了强化一种文化：将数据视为现代发明的起源。将数据等同代码看待，视其为业务的重中之重，而非事后诸葛。

数据质量评估及围绕治理制定的规则既能加速数据的利用，也可能阻碍所有进展。平衡这两方面的需求，并使用恰当的工具使整个企业都能进行创新至关重要。为数据集指定直接负责人或数据管理员，这有助于构建稳健的数据生态系统。从小处着手，再持续扩展您的数据网格，这样可以保持数据飞轮的持续运转。确保不同类型的用户能够通过不同的方式访问和发现数据。这种方法让您能更全面地了解环境中正在进行的工作，避免出现数据治理框架外未经许可私自开展数据运维 (DataOps) 的现象。

易于使用、人类可读的数据存储库、数据目录和数据字典，可为企业的数据资产提供一个集中有序的数据和元数据仓库，确保不同技能水平的团队都能发现、理解、协作处理数据，并开始利用数据创造业务价值。这大大加快了针对其他用例所需的额外投资成本做出决策的速度。提升数据潜力的方法多种多样，例如购买外部数据源，通过机器学习算法增强或创建合成数据，通过众包团队来标注内部数据，甚至改变业务实践以自动生成和捕获数据。确立决定何时使用每种资源的良好实践，这一点非常重要。

风险管理

利用云服务来减少和管理 AI 固有的风险。

虽然每项新技术都会带来一系列新风险，但由于 AI 模型的非确定性，管理 AI 系统设计和开发过程中的风险以及 AI 部署、长期运营和应用中的风险充满挑战。其中就包括一些财务风险。首先，在开发过程中要考虑沉没成本的风险，因为 AI 开发项目的结果难以提前保证（优化系统输出与专门构建系统以实现该目标存在本质的区别）。需要确立可靠的实践，比如使用模型卡和对抗性输入等手段，并建立可靠的机制，比如概念验证、最小可用产品和最小可行产品，以降低和控制风险。

其他风险则属于法律和伦理范畴。这些风险既包括由当地立法机构分类的风险，例如欧盟界定的风险，也包括 AI 本身固有的风险，例如隐藏的反馈循环、未校准输出的误解，以及可能对不同人群产生负面影响的意外结果。同时，还需考虑其在专业领域、企业乃至社会层面的使用和影响（例如，回音室效应或对客户行为的长期影响）。如需了解更多信息，请参阅《负责任地使用 AI》。

应优先开发和采用在必要时（不限于安全关键环境）可约束系统的安全措施和架构。确保子系统故障不会传播并加剧下游的 AI 系统问题。思考哪些主题是相关的，比如可解释性、透明性和可诠释性。管理这些风险时，不仅仅针对单一受 AI 影响的决策或行动，而是要贯穿整个流程或更大的系统运作中。要认识到数据和现实世界概念随时间漂移可能给系统带来的长期挑战，并致力于加固系统以防范恶意行为者（参见安全视角：AI /机器学习系统的合规性与保障）。最后，不要低估在某些领域将人工智能系统提升到与人类同等水平所面临的复杂性挑战。

负责任地使用 AI

通过负责任的 AI 实践，持续推动 AI 创新。

直到最近，许多企业在开发人工智能解决方案时，往往只专注于技术层面以及追求特定的业务目标，而忽视了负责任地使用这项强大的新技术。然而，人们日益认识到 AI 系统是基于海量数据进行学习的，但学习结果并不总是符合预期。这使得关注负责任的 AI 实践变得至关重要。负责任的 AI 实践是促进 AI 持续创新的关键，并确保在符合伦理、透明、无偏见的前提下开发、部署和使用 AI 解决方案。随着 AI 应用范围的扩大及其影响力的增加，这一点变得尤为重要。因此，在 AI 项目的整个生命周期中，尤其是在初期阶段，就应当考虑并解决负责任地使用 AI (RAI) 的问题。

您应成立一个由多个业务部门的代表（如研发、人力资源、多元化与包容性、法务、政府与监管事务、采购以及公关部门）组建的 AI 治理委员会，与 AI 领导团队紧密合作或加入其中，以确保 AI 解决方案对员工、客户和整个社会安全无害。该委员会应负责监督和指导 AI 技术的开发、部署和使用，确保其是符合伦理的、负责任的，并负责推动与行业法规保持一致并遵守 AI 相关的立法。随着时间的推移，您应考虑负责任 AI 对设计、开发和运营的影响。您应思考您的系统如何影响个人、特定的人群、用户、客户以及整个社会。鉴于 AI 在云端快速扩展的能力，您需要考虑如何融入关键的负责任 AI 维度，比如可解释性、公平性、治理、隐私、安全、健壮性及透明度，并考虑技术如何影响不同的文化和人口结构。将负责任的 AI 理念作为您的 AI 愿景的重要组成部分。这包括制定深思熟虑的原则和指导方针，阐明如何负责任地使用 AI，以及 AI 将如何影响您的计划。特别是，需要纳入算法公平性、多元化和包容性以及偏见检测。

尽可能在设计上将可解释性融入 AI 生命周期中，并确立识别和发现预期和非预期偏见的实践。可考虑使用合适的工具来帮助您监控现状并提示风险。利用最佳实践来推动负责任使用 AI 的文化，构建或利用系统来协助您的团队检查这些因素。虽然在算法投入生产前采取负责任的 AI 实践会产生前期成本，但从中长期来看是值得的，因为这可以减轻人工智能可能带来的负面影响。特别是当您计划构建、微调或使用基础模型时，要了解新出现的关注点，比如模型 “幻觉”、版权侵权、模型数据泄漏和模型 “越狱” 等。务必询问原始供应商或提供商是否其开发过程中采取了负责任的 AI 方法，并了解其具体实施细节，因为这会直接影响到您的业务案例。

> **注：**
>
> 亚马逊云科技负责任地使用 AI 团队针对这一课题撰写了一份白皮书。

平台视角：AI 基础设施和应用

随着 AI 和机器学习算法及其用例的进步，用于运行这些算法的系统和流程可能很快就会过时。正如在任何高效的制造流程中一样，您需要为 AI 开发构建系统和平台，以确保产出统一、稳定的产品。这里所说的产品，实际上是由算法驱动并为企业创造价值的成果。打造一个与您的基础能力相匹配的平台，有助于塑造竞争优势并加速创新步伐。一个能够降低风险的平台应当具备可靠性、可扩展性，并且能够兑现其承诺——提供基础能力，这些基础能力用于支撑与本文其他视角一致的长期业务价值。

支持 AI 的平台需遵循一系列设计原则，确保各组件目的明确，意图一致，并随着时间的推移涵盖机器学习生命周期的方方面面。其核心在于管理和访问分布式及受治理的数据，这些数据需按照满足个别消费者特定需求的方式进行准备和提供。此外，平台还需支持通过端到端的综合开发体验来开发新型 AI 系统。充分利用现有的 AI 能力和基础模型也是至关重要的。一旦这些模型经过训练，就可以通过编排、监控并随后分享以集成到应用、系统或流程中，以供下游消费者使用。这些活动由平台赋能团队监督，他们持续根据收到的反馈进行迭代，以实现持续改进。

基础能力	解释
平台架构	实现可复制的 AI 价值的原则、模式和最佳实践。
现代应用程序开发	构建架构卓越且 AI 优先的应用程序。
AI 生命周期管理和机器学习运维	管理机器学习工作负载的生命周期。
数据架构	设计符合预期目的的 AI 数据架构。
平台工程	构建具有增强功能的 AI 环境。
数据工程	为 AI 开发实现数据流的自动化。
预配和编排	开发、管理和分发获得批准的 AI 产品。
持续集成和持续交付	加速 AI 的发展。

平台架构

实现可复制的 AI 价值的原则、模式和最佳实践。

随着机器学习开发日臻成熟、从研究驱动型的技术走向工程实践，可靠地、可重复地基于其应用创造价值就变得越来越重要。平台架构的目标是，综合考虑不同的 CAF 视角的输入，设计一个与业务目标相契合的基础架构，确保 AI 生命周期的采用和赋能。首先，要了解平台利益相关方的成熟度和能力，以及他们对于机器学习技术栈的需求：您是否试图启用预构建的现成 AI 服务、低代码和自动机器学习功能，让非专业人士也能访问 AI？还是希望支持专业人士在其 AI 开发生命周期中使用和定制机器学习框架、直接访问基础设施？特别是当您涉足生成式 AI 领域时，这些问题对平台架构有着重大的影响。可从三个层面考虑 AI 相关的具体需求：

1. **计算层：**AI 的训练和推理可能会对硬件有很大的需求，可能需要大量的算力资源（用于基础模型）。除了消费保护措施外，性价比也是为您的企业设定标准的关键因素之一。可考虑采用性价比优于传统 CPU 或 GPU 的专用硬件，以降低成本。
2. **机器学习和 AI 服务层：**规划您的平台如何支撑机器学习与 AI 服务的开发、部署与迭代过程。机器学习服务需赋能技术专家群体，例如，进行定制模型的训练或调优（如基础模型），而 AI 则应确保能够便捷地调用模型与功能（如生成式 AI 领域训练成本高昂的中大型基础模型）。尽管这种区分并非总是泾渭分明，但各类需求存在差异。
3. **消费层：**此层面向您的 AI 能力的下游用户。既可以简单到一个仪表板应用，也可以复杂到通过 Prompt 工程对基础模型进行增强，或是利用特定的生成式 AI 架构，比如检索增强生成 (RAG) 应用等。

在搭建平台的过程中，需细致分析行业特有的法律要求，这些要求对数据管理、模型开发流程及部署均有影响（例如数据的强制性分类），并据此设定相应的防护措施。要投入时间明确各项标准，例如关于数据隐私和数据治理的标准，并分发给下游团队供其使用。接下来，简化合规环境和基础设施的配置，从而加速 AI 新用例的开发和部署。通过了解您的团队可能如何使用“人在回环”和“人机监督”功能（它们是 AI 工作流程中重要的检查点），为您的平台整合反馈回路。最后，确定机器学习特有的监控需求，比如在模型行为变化时进行偏见检测、可解释性分析并安排人工复审。

设计模块化的 AI 价值链至关重要，因为它能支持独立扩展与更新。这种模块化方法有助于加速数据标注流程，并明确划分不同组件的所有权和责任归属。在选定标准化的云原生解决方案时，必须综合考虑成本、可靠性、可恢复性及性能等因素。所有这些最佳实践以及设计指南和标准，都应发布至一个中心知识库，以供企业内所有实践者访问。实施反馈机制及衡量平台采用度的指标，能够为您的 AI 项目持续提供洞见，助力您做出明智的决策。

现代应用程序开发

构建架构卓越且 AI 优先的应用程序。

> **注：**
>
> 《亚马逊云科技卓越架构框架——机器学习剖析》为工作负载和架构设计模式和最佳实践提供了权威资料。

随着 AI 技术的成熟，它深刻影响着应用开发的方方面面：

1. AI 增强型应用开发：通过 AI 技术提升软件开发生命周期 (SDLC) 的效能。利用 AI 服务和工具赋予应用生成及自动补全特性，或通过识别潜在的代码问题来简化审查流程，同时通过确保开发过程高效无误，实现性能和测试的自动化。从创意构思到软件维护，全面重塑软件开发生命周期的各个环节。
2. 将 AI 作为产品的差异化要素：将 AI 融入软件之中，不仅能提升用户体验，甚至可成为价值主张的核心。AI 能够增强软件的功能性，确保其紧密贴合用户的实际需求与期望，最终打造出深受用户欢迎的产品。在开发此类应用时，需考虑数据在系统中的流转方式、如何影响 AI 系统、产生何种输出、消费者和客户如何解读这些输出，以及这些输出如何进一步生成可用于迭代的新数据。在进行架构决策时，应以 AI 领域成熟的设计原则为基准。
3. AI 模型开发：在将 AI 融入软件开发的过程中，考量改造现有模型、利用开源方案或构建定制化解决方案变得尤为重要。随着现代应用程序开发的不断演进，掌握 AI 技术已成为日常开发不可或缺的一环。针对特定使用场景，您或许需要更高程度的个性化定制，即运用特定的数据对模型进行微调，以确保模型能够精准适配您的需求。

针对这三个方面，考虑如何将应用程序和开发流程分解为更小、更易管理的部分。将微服务或多元模型方法与敏捷实践相结合，以此提升灵活性，加快交付速度，更有效地应对变化。此方法在 AI 开发中尤其有益，因为 AI 开发需要大量的迭代测试、实验及优化。需在开发团队中树立清晰的认识：用户及客户对 AI 系统的感知的确存在差异，而且众多用户缺乏与这些系统进行有效交互的心智模型。也就是说，与客户和用户直接交互的 AI 应用都将从对其用户体验 (UX) 的重新审视中直接受益。

AI 生命周期管理

AI 生命周期管理分为架构视角与工程视角，这两个视角随企业能力的成长而逐步完善。

架构视角侧重于 AI 生命周期管理的设计、规划和概念层面。管理机器学习工作负载的生命周期是一项复杂任务，需要综合性的方法。它有三个重要组成部分：

1. 识别、管理和交付业务成果和客户价值。
2. 构建和发展 AI 解决方案的技术组件。
3. 整个生命周期中 AI 系统的运维，也称为机器学习运维 (MLOps)，对于更大的模型而言则称为基础模型运维 (FMOps)。

鉴于这三个组成部分均较为复杂，我们在《卓越架构框架：机器学习剖析》中提供了详尽的指导。不同的 AI 策略对这三个组成部分会有不同的着眼点。例如，如果您的总体目标是通过定制模型来推动新产品开发，那么您对生命周期管理的看法将不同于借助公开可用的服务来提高内部运营效率这一策略的看法。无论采取何种方法，都应采用集中化的存储库和版本控制系统来存储 AI 工件，并跟踪模型谱系和数据谱系。

工程视角着重于 AI 生命周期管理的实施与运作。为简化这一流程，实施机器学习运维实践至关重要，以实现 AI 模型部署和监控的自动化，减少人工干预，提升可靠性，缩短部署时间，并增强可观察性。确保遵循一套明确的流程来管理 AI 生命周期，涵盖从构思到部署再到监控的全过程。该流程应包括数据收集和存储、模型训练和部署、模型监控和评估（CAF - AI 的运维章节），以及性能监控等步骤。这有助于尽早发现缺陷，支持模型的持续演进。最后，要建立一个自动化的框架，重新训练您的 AI 模型，例如，在性能下降或有新数据送达时进行重新训练。

为了更好地了解您相对于行业最佳实践所处的现状，可借助亚马逊云科技合作伙伴或亚马逊云科技评估您的机器学习运维成熟度，并基于机器学习运维和生命周期框架做出决策。这些流程和标准是防止系统仅依赖于机构知识的最佳防范手段，有助于减少 AI 技术债务。数据团队通常过分关注硬性的机器学习指标，而忽视了这些指标如何影响业务指标，这是生命周期管理不足的表现。无论采用何种路径，都要确保您为机器学习运维建立的流程和标准是可重复的。这些机器学习运维最佳实践还有助于确保您的科学团队不会因建模而疲惫不堪，而是专注于成果，避免因大量并行实验而分心。

数据架构

设计和发展符合预期目的的 AI 数据架构。

数据是 AI 技术的关键。随着数据类型和数据量的爆炸式增长，传统的数据架构亟需变革。特别是，AI 对存储、管理和分析提出了新的需求，以应对其日益增长的复杂性，因为 AI 正逐渐成为商业决策的核心。记住，AI 工作负载不仅需要大量的数据，还需要多样化的优质数据来进行模型训练和验证。由于这些数据来自多个来源，通常具有不同的格式和结构，传统数据架构受数据传输和数据类型方面的限制，往往无法有效管理如此多样化和大规模的数据。因此，要深入研究现代数据架构的演进。这些架构将数据湖、数据仓库和其他专用的数据存储结合在一起，减少了治理的复杂性，同时实现了数据的传输，这是 AI 的一个关键层面。

在当今的企业中，三种架构成为了主流：数据仓库（吞吐量经过优化的结构化仓库）、数据湖（从各种数据孤岛聚合数据，并充当中央数据库的作用）和业务应用专属仓库（NoSQL 数据库、搜索服务等等），每种架构均支持不同的用例。然而，在这些仓库存取数据可能具有挑战性而且代价昂贵。因此，随着数据传输对于 AI 系统来说越来越重要，您需要增强架构以满足数据传输的需求：

- **由内向外：**数据最初从各种来源（数据库和结构良好的电子表格等结构化数据；或媒体和文本等非结构化数据）汇总到数据湖中。随后将数据的一个子集传输到专用的存储载体中，以便用于特定的分析任务，比如搜索分析或构建知识图谱。
- **由外向内：**数据起初存放在适合特定应用的专用存储中。例如，为了支持在云端运行的游戏，应用可能会使用特定的存储载体来维持游戏状态和排行榜。此类数据随后被迁移至数据湖中，以便开展更全面的分析，以提升游戏体验。
- **外围：**这涉及在专用数据存储载体之间传输数据，例如从关系型数据库迁移到 NoSQL 数据库，以满足诸如报告仪表板的特定需求。

为了保持 AI 团队的高速运转，需要以可行且无缝的方式实现这种数据传输。随着 AI 技术的快速发展，具备这种灵活性至关重要。由于数据在 AI 领域至关重要，数据几乎等同于机器代码，AI 和数据架构之间的界限变得日益模糊。现代数据架构使得企业能够将数据本身视为一种产品。现代数据架构并非静态结构，而是设计成流转的结构，能随着新数据类型和技术的涌现而适应变化。因此，要研究各种新兴的数据架构原型，如现代数据架构、分布式数据网格和数据集市等，并构想一个统一的平台或生态系统，以容纳所有类型的数据。最后，定期反思当前架构，预先考虑访问模式和需求，并选择适合目的的架构。制定计划，确保您的数据集易于发现、记录完整且易于理解。建立元数据原则或数据文档化标准，用以描述数据，包括数据含义、与其他数据的关系、来源、用途和格式。

平台工程

为 AI 构建具有增强功能的合规环境。

云技术从根本上改变了企业使用先进 AI 基础设施和服务的方式。通过普及 AI 的使用，企业可简化其 AI 工作流程，并利用规模经济带来的巨大优势。因此，设计合理的 AI 平台可让您的 AI 团队以更低的成本实现更多成果。要相应地设计您的平台，为不同的利益相关方（如开发人员、数据团队和运维人员等）提供简化和抽象，减少他们的认知负担，同时增强其工作方式的创新能力：

- **AI 服务：**通过简化平台与开箱即用的 AI 服务之间的连接来赋能您的团队，考虑到预构建模型和特定应用场景，并直接融入现代数据架构。
- **机器学习服务：**在云端，开发人员可使用专为 AI 应用的开发和部署而设计的特殊环境。在考量 AI 模型的训练和部署时，此类托管型机器学习服务就变得不可或缺。它们能高效处理机器学习系统工程中固有的、复杂且耗时的流程。借助这些服务，您将为 AI 团队重新分配宝贵的时间，投身于更具战略意义的项目。
- **机器学习基础设施：**通过托管平台中高度专业化的底层 AI 基础设施，为您的团队减轻繁重的负担，从而赋能您的团队。请记住，AI 团队的赋能通常不在于拥有基础设施，反而经常因基础设施而受限，导致业务价值无法实现。

云端的主要优点之一是其实现常规任务自动化的能力。尽可能地实现机器学习平台任务的自动化，因为它可加快流程、减少人为错误，并确保一致性。您的 AI 解决方案越复杂，专属的机器学习运维实践的相关性就越强。从一开始就要在您的平台中融入特定的 AI 监控工具。这些工具会跟踪 AI 工作负载的性能、针对其运作提供有价值的洞见、帮助及早识别问题。反馈机制会影响模型微调和超参数配置。通过实时监控工作负载，企业能够更好地保障其 AI 应用处于最佳性能，并能够迅速解决出现的任何问题。

尽管云端提供了极大的灵活性，但采取防护措施至关重要。通过实施指导原则或限制条件作为管控手段，确保开发者在既定的最佳实践和安全参数范围内工作，从而降低风险并确保负责任地使用资源。构建一个安全网，既要鼓励创新，也要确保创新活动绝不损害企业的安全性、合规性或性能标准。

数据工程

为 AI 开发实现数据流的自动化。

由于数据是任何 AI 战略和开发过程的第一要素，数据工程就显得极为重要。它不应是事后考虑的环节，而应成为企业和团队内随时可用的能力。由于数据被用于主动塑造 AI 系统的行为，所以正确开展数据工程至关重要。数据准备工具是开发过程的重要组成部分。虽然实践本身并没有根本性的改变，但其重要性和持续演进的需求却日益增加。考虑将数据管道和实践直接整合到 AI 开发过程和模型训练之中，通过精简无缝的预处理实现这一目标。可考虑从传统的提取、转换和加载 (ETL) 过程转向零 ETL 方法。通过此种数据工程方法，可减少数据实践和 AI 实践之间的障碍。赋能 AI 团队将不同来源的数据整合成一个单一的、统一的视图，使之成为一种自助服务能力。配合可视化工具和技术，帮助 AI 和数据团队以可视化的方式探索和理解其数据。

尽可能确保数据的准确性、完整性和可靠性。在工作流程中专门设计用于机器学习的数据模型或转换（标准化、一致且文档齐全），以促进数据的有效处理与管理。这将显著提高 AI 应用的性能，并减少开发过程中的障碍。

预配和编排

开发、管理和分发获得批准的 AI 产品。

由于 AI 系统在不同的开发和部署阶段对基础设施的需求会显著变化，预配和编排在现有的云策略中值得重新审视。了解您在 AI 转型历程中所处的位置，及其与机器学习运维成熟度的关系。考虑到您的消费者、数据工程师、数据科学家、开发人员和业务分析师在履行其角色时，有着不同的需求和要求。要找出方法，为不同的用户（尤其是技术知识有限的用户）提供自助式 AI 环境配置。这一点可通过创建已获平台架构批准的目录、组合和产品来加以实现。目录可分发给终端用户，其中的产品可用于消费。产品可定义为基础设施即代码，并可通过个性化的门户网站来部署，或通过符合（由平台团队管理的）企业政策的持续集成和持续交付管道加以部署。一个常见的应用场景是，建设一个个性化的门户网站，为数据团队提供预定义的notebook和计算资源，以便他们能够快速针对新的业务问题进行试验，而不必等待平台团队预配资源。对于需要一整套工具的数据科学家等高级角色，可将目录配置为部署整个 AI 环境，包括提供基础模型加速器访问权限。

要考虑到 AI 模型的训练或调优步骤可能需要高性能计算，并使用符合预算和治理约束条件的预批准服务来实现自动化预配。尽可能使用 API 和框架级别的自动化和编排功能。设计用于管理 AI 工作负载部署的各种机制，并简化底层基础设施的构建。

持续集成和持续交付

加速 AI 的发展。

在 AI 技术的背景下，针对持续集成和持续交付有两种截然不同的观点：第一种是尽可能实现模型开发和部署过程的自动化和强化，例如开发定制化模型的过程。第二种是把 AI 技术本身作为 DevOps 实践的一部分，利用它简化持续集成和持续交付。

对于第一种观点，企业可针对 AI 模型的部署和测试实现自动化，赋能团队以云端速度进行创新。在定制化模型的情形下，目标是实现 AI 工作负载部署和管理的自动化，同时托管复杂的工作流，如数据处理、模型训练、模型评估、后处理、模型注册和模型部署等。在实现 AI 开发过程自动化的时候，您会用到专用于机器学习管道的工具，以及传统应用开发中常用的方法和工具。通过合理的架构和蓝图设计，数据科学家能够试验不同的模型，并确保模型在投入生产之前经过了全面的测试。要花时间考虑构建这种能力是否适合您的企业。

可通过了解生产机器学习模型的产出速度、更新机器学习模型的需求，以及用例的关键性和影响来做出决定。随着时间的推移，模型漂移可能会发生，而且时有发生，因此需要考虑可在多大程度上实现验证过程的自动化，比如设定重新训练的阈值。自动化验证会根据预定义的标准去检查模型的性能，如果模型性能超出了可接受的阈值，则触发自动重新训练或回滚至先前版本。最后，通过整合人工反馈和自动化模型验证、模型测试和重新训练等任务，可重复性提高了 AI 工作负载的可靠性，为数据科学家和工程师腾出了宝贵的时间，让他们专注于更重要的任务。通过整合这些方面，企业能够以高成本效益的方式迭代 AI 模型，同时确保即使数据和需求发生变化，模型及其封装的 AI 系统仍保持相关性和有效性。

对于第二种观点，将 AI 本身用于与 AI 有关或无关的开发运维活动，丰富开发流程，并在适当的情况下使用生成式 AI。商业价值的显著增长往往来源于 AI 直接应用于开发流程之中。因此，探讨利益相关方如何在其技术工作流程中采纳 AI 至关重要。这可以意味着利用 AI 分析工作负载中的异常，通过 AI 优化代码级性能，或是根据开发者的 Prompt 生成代码。在此过程中，始终确保针对开发运维的 AI 应用做好企业级准备，并将安全牢记在心。

安全视角：AI 系统的合规性和保证

安全是亚马逊云科技的重中之重，所有客户，无论规模大小，都能从亚马逊云科技对安全基础设施和新服务的持续投资中获益。对于正在开发 AI 亚马逊云科技工作负载的客户来说，安全是亚马逊云科技整体解决方案不可或缺的一个组成部分。生成式 AI 是扩展基础模型以实现业务成果的关键推动力，有多种方式可以创建生成式 AI 工作负载。在 AI 的方方面面整合安全性和隐私性，这对于实现业务成果的整体成功至关重要。有关 AI 应用的基本业务案例是解决具体的业务问题，这些问题范围广泛，从日常生产力任务的简单自动化到包含敏感数据的复杂医疗或财务决策。可运用风险管理技术来实施本视角中定义的安全和隐私功能，以满足您的业务需求。

基础能力	解释
漏洞管理	持续识别、分类、修复和减少 AI 漏洞
安全治理	建立与 AI 工作负载相关的安全政策、标准和指南以及相关的角色和职责
安全保障	根据 AI 工作负载的监管及合规要求，应用、评估和验证相关的安全和隐私措施
威胁检测	检测和减少 AI 工作负载中潜在的、与 AI 相关的安全威胁或意外行为
基础设施保护	确保用于运行 AI 工作负载的系统和服务的安全
数据保护	保持对用于 AI 开发和使用的数据的可见性、安全访问和控制
应用安全性	在 AI 工作负载的软件开发生命周期过程中检测和减少漏洞
身份及访问管理 (IAM)	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
事件响应	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*

漏洞管理

持续识别、分类、修复和减少 AI 漏洞。

AI 系统可能存在技术相关的漏洞，您需要对此有所了解，例如 Prompt 注入、数据投毒和模型逆向等漏洞。所有 AI 系统的三大关键要素是输入、模型和输出。为了减轻工作负载的潜在漏洞，可以采用以下最佳实践来保护这些组件：

- **输入漏洞**与所有包含模型入口点的数据有关。此类输入可能是目标模型和分布漂移的来源，在此情况下，恶意行为者可能会试图随着时间的推移逐步影响决策过程，或者故意对特定数据引入隐藏的偏差或敏感信息。通过数据质量自动化和持续监控来强化这些输入。模型滥用是 AI 解决方案中因 Prompt 注入导致的一种漏洞，因为数据和指令相互交错。另外需要特别注意的是，基础模型越狱领域发展迅速。需执行输入有效性验证，将数据与指令隔离，遵循最低权限原则，将大语言模型 (LLM) 的访问权限限定在特定授权范围内。避免访问会广泛影响运维的系统命令、可执行文件和日志操作。
- **模型漏洞**与利用模型对真实世界或所见数据的误读有关。可通过威胁建模来减少有记录的已知威胁，从而增强您的模型。在使用商用生成式 AI 模型时，审查其数据来源、模型微调的使用条款、以及可能源于模型本身或第三方库使用的漏洞，这些都可能对您产生影响。验证是否对模型目标及其结果进行了监控及其是否随着时间的推移保持一致性，以免发生模型漂移。
- **输出漏洞**与长期同系统交互有关，这可能允许推断出关于您的模型输入和属性的关键信息，通常称为数据泄露。对生成式 AI 而言，需验证其输出是否经过清理而不是直接使用，以减少跨站脚本漏洞和远程代码执行的风险。以上只是您需要为工作负载考虑到的几个漏洞。虽然并非所有 AI 系统都会受上述漏洞的影响，但请警惕与您的特定工作有关的风险。定期进行测试、亚马逊云科技 Game Day 和桌面推演，以验证按操作手册规定的补救措施的有效性。

安全治理

确定与 AI 工作负载相关的安全政策、标准和指南以及相关的角色和职责。

确保针对内部和外部托管的商用或开源模型的使用制定了明确的政策。同样，如需使用商用生成式 AI 模型，请考虑企业敏感数据泄露到商用模型平台的风险（请参阅数据保护能力）。了解适用于您所在行业或企业的 AI 相关资产、安全风险和合规要求，有助于确定安全工作的优先级，为指定角色分配足够的安全资源，并提供透明度。

AI 带来的风险可能会造成影响深远的后果，包括隐私泄露、数据篡改或滥用以及决策失误。采取稳健的加密措施、多因素身份验证、持续监控并与风险容忍度框架保持一致（例如 NIST AI RMF），对保障 AI 环境的完整性和安全性至关重要。

为您的工作负载的三大关键要素持续提供指导和建议：

- **输入**——确定数据源和 AI 使用的审批人。在审批过程中，需要全面评估数据相关的风险因素，包括数据的分类或敏感程度、数据集内是否存在受监管的数据、数据的来源和时效性，以及处理数据的合规性和法律依据。为管理风险，需评估用于获取输入数据的机制。评估时应考虑数据源的可信度、数据的获取方式，以及数据存储和安全措施等因素。验证数据源的数据分类是否与解决方案的分类一致，例如不允许在公用 AI 解决方案上处理机密数据。
- **模型**——确定创建和训练模型的角色及其责任。确定与模型发布的作者、审批者和发布者对应的相关角色。为管理风险，需评估模型训练机制，包括所涉工具和个人，以避免有意或无意引入漏洞。评估模型架构是否存在影像输出结果的漏洞。确保任何模型的故障模式都能达到模型关闭或进入安全状态的效果，以避免数据泄露。
- **输出**——确定已建立的输出的生命周期管理。制定分类标准，密切关注可能包含不同数据集或不同分类数据集的输出结果。为管理风险，建立适当的保护和保留控制措施，根据个人识别信息 (PII) 等重要性和敏感性对您的数据分类并定义适当的访问控制。确立数据保护控制和生命周期管理政策。确立遵守隐私法规和其他合规要求的健全的数据共享协议。

安全保障

应用、评估和验证符合 AI 工作负载的监管和合规要求的安全和隐私措施。

您所在的企业和您服务的客户需要对您采取的控制措施抱有信任和信心。随着客户和用户日益意识到 AI 系统存在的安全风险和潜在滥用问题，他们对 AI 系统在安全合规方面达到更高标准的诉求也与日俱增。在设计、开发、部署和监控解决方案时，应优先考虑网络安全，满足监管要求，并切实有效地管理 AI 特有的安全风险，同时也要符合您的业务目标和风险承受能力。由法律专家、合规专家、数据科学家和 IT 专业人员进行全面监督，通力协作，透明运作，有助于验证全方位的保障措施。采取测试程序和补救流程也不失为积极主动的安全保障举措。持续监控和评估您的工作负载的三大关键要素：

- **输入**——由于模型训练和分析通常需要大量的数据，您需要验证输入的数据类型是否与模型的目标和输出结果一致。建立审计机制，以了解既定控制框架的遵循情况。
- **模型**——确认用户了解哪些 AI 使用方式是符合企业政策且可接受的。。采取政策和控制措施，以验证企业是否了解 AI 能否适用的对应场景。确定审计机制，以识别模型的数据使用方式，以及 AI 功能在企业内部的使用情况。
- **输出**——确定可接受的输出使用标准，同时注意数据可在何处重复使用或被重新引入其他 AI 模型。建立审查输出数据的发现或审计机制，以验证生成的数据是否不会泄露敏感信息或受监管数据，无法被用于推断或重构此类数据。建立验证输出真实性和来源的机制，尤其是针对医疗诊断等可信度要求极高的领域。

保护个人隐私要求严格遵守道德和法律准则，以防止未经授权的访问、数据滥用或披露。在发挥 AI 潜力的同时尊重隐私权，有助于建立公众信任，让大众从 AI 能力中受益。参见《亚马逊云科技卓越架构框架》中的《MLSEC-05：保护敏感数据隐私》。建立透明度和知情同意等机制。将数据保留限制在功能所需的范围内，并订立数据共享协议。再次强调，需考虑与工作负载三个关键组成部分相关的隐私要求：

- **输入**——验证您了解如何使用受隐私相关法规（例如《通用数据保护条例》、《加州消费者隐私法》、《儿童在线隐私保护法》、《个人资料保护法》）约束的数据以及处理数据的法律依据。请考虑到数据驻留地以及存储和处理数据的地点。为每次使用受监管数据建立隐私影响评估 (PIA) 或类似的流程。
- **模型**——在训练或调整模型时，需考虑是否存在处理数据的法律依据，以及是否能证明数据主体的透明度。确定与潜在的模型数据泄露有关的隐私影响评估或类似流程。
- **输出**——考虑受监管数据是否用于训练其他模型，以及个人数据的二次使用是否受限。建立一个完成*删除权或遗忘权*类型的请求的机制。确定审查输出数据的发现或审计机制，以验证所生成数据不能用于推断或重新生成先前已消除身份识别信息的数据。

威胁检测

检测和减少 AI 工作负载中潜在的安全威胁或意外行为。

为加强对任何机器学习或生成式 AI 系统的三大关键组成部分（输入、模型和输出）的保护，可通过以下最佳实践来检测并减少对您的工作负载的威胁：

- **输入——**检测 AI 解决方案的威胁对修复可能影响业务的漏洞至关重要。在将输入数据用于模型训练之前，您需要对其进行清理，以检测并消除潜在的安全威胁。持续跟踪用户会话的输入数据，以检测和减少影响可用性和导致滥用的威胁。
- **模型——**针对 AI 系统执行特定的威胁建模，并开展威胁搜寻演练以检测和减少潜在威胁。更新威胁模型和监控措施，纳入 AI 特有的威胁概念，包括使用意外的用户输入训练模型、用于内容或训练的数据集中毒、隐私泄露以及数据篡改等。关联输入数据和模型使用的数据，以检测异常或恶意活动。
- **输出——**监控偏离模型目标的输出异常，并启用检查功能以检测模型输出中的敏感数据。建立包含适用于您的工作负载的已识别的已知威胁目录。建立自动化测试，以验证检测能力并整合威胁情报，从而提高效率并减少误报。考虑利用威胁情报来提高效率和减少误报。

基础设施保护

确保运行 AI 工作负载的系统和服务的安全。

机器学习运维将开发运维实践应用于 AI 工作负载，而安全措施则需覆盖构成整个环境的基础设施。您的 AI 模型可采用安全端点和 Amazon API Gateway 对模型访问进行速率限制。对于所有内部和外部使用的 API，采用 API 安全最佳实践，并明确创建一个允许列表，以包含来自模型自身 VPC 之外的 API 调用。您可从亚马逊云科技的《Security Reference Architecture》（安全参考架构）入手，根据您的环境建立网络、计算及存储方面的安全控制措施。

模型一般部署在跨网络和跨服务器的多重环境中。这些环境之间的通信传输应采用加密技术予以保护。需对开发和生产环境进行集中配置，并采取由安全管理员独立管理的预防和检测性防护措施。隔离模型训练等敏感任务的开发环境。确保为终端用户提供会话隔离，以保持体验的完整性，防止数据意外泄露。将输出相应和相关会话数据录入单写多读 (WORM) 存储设备中，以满足合规性和故障排查需求。考虑实施模型漏洞悬赏计划，以发现和减少可能导致安全问题的边缘用例。

数据保护

对用于 AI 开发和应用的数据，需要保持可见性、安全访问和控制。

数据保护在整个 AI 开发生命周期中至关重要，需确保安全治理定义的数据保护政策付诸实施，例如《亚马逊云科技卓越架构框架：机器学习剖析》的《MLSEC-07：仅保留相关数据》中提及的内容。如果使用商用模型来开发生成式 AI，请注意直接使用数据作为模型的输入可能会导致敏感信息泄露。同样，让您的专有或自托管模型访问受保护数据也会为升级数据相关权限敞开大门。请因地制宜地评估模型使用和服务条款。在预训练和微调阶段为模型开发收集的数据，其安全性在传输过程中、静态存储时和实际使用时都应得到保障。在进行清理、规范化和转换等数据预处理操作时，请考虑使用数据 Token 化流程，将敏感数据替换为非敏感数据 Token。为模型使用的所有数据源创建验证机制，尤其是对用于训练模型的推理数据。监控并对敏感数据或可能导致敏感级别升级的数据创建警报。利用数据活动监控技术，通过使用情况和频率等指征来检测数据访问模式。避免使用敏感数据训练模型，因为这可能导致模型输出在无意中披露数据（例如推理期间发生数据泄露）。标记和标注所有不同环境中的训练所用数据，并调整数据标记和标注，确保数据标签符合数据分类政策和标准。验证非生产和开发环境的数据谱系和数据访问控制措施是否得当，以防止数据被恶意篡改引发模型漏洞。考虑使用 CI/CD 管道将数据推广到测试和生产环境，以保持数据完整性。记录并屏蔽敏感数据，同时为数据访问创建审计追踪。对敏感数据存储和设计上本不该存储指定数据类别（例如机密）的数据存储采取数据丢失防护技术，并监控敏感数据的意外泄露情况。验证模型输出的数据质量，以建立信任，避免产生幻觉。监控模型输出数据的敏感级别，如果敏感级别上升，则通过编校或隔离响应来触发重新分类。例如，如果有新的数据集用于模型或模型训练，则应验证输出数据是否符合现有的敏感级别。

应用安全

在 AI 工作负载的软件开发生命周期流程中检测和减少漏洞。

核查模型开发者是否在本地环境和 CI/CD 管道中执行 Prompt 测试和其他安全测试用例，以验证模型的使用情况。创建并维护测试用例库，以验证测试覆盖率并实现自动化。在所有开发、测试和生产环境中利用与安全扫描集成的数据和模型管道，将所有模型构件存入安全的仓库中。维护 AI 模型库存，将模型实例分配给具体确定的技术和业务负责人。验证已知的训练模型是否已备份。保留时间点恢复功能，以便受损模型能恢复到已知的良好状态。保护对模型和数据备份的访问，以验证其是否受损，并定期测试模型恢复情况，以便其能够完全恢复到已知的良好状态。需跟踪参数、元数据等与模型和数据开发有关的数据，以确保输出结果的可追溯性并支持其有效性。为数据集和模型单独创建并使用运行手册和测试回滚机制，以便在发生运行或安全事故时执行，为模型提供恢复力。

运营视角：AI 前景的运行状况与可用性

运行机器学习应用对很多客户而言还是一件新鲜事。在 AI 生命周期管理和机器学习运维的 CAF-AI 新功能中，我们已经介绍了应对这一问题的部分观点和指导。此外，其余的主要考虑因素均围绕事件管理和性能展开。为深入探讨 CAF-AI 相关的内容，我们建议参阅《机器学习运维成熟度框架》和《亚马逊云科技卓越架构框架：机器学习剖析》，两者均针对上述挑战给出了详尽说明和最佳实践。

基础能力	解释
事件和问题管理	识别和管理不可预见的 AI 行为。
性能和容量	监控并管理 AI 工作负载的性能。
可观察性	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
事件管理（AI 运维）	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
变更和发布管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
配置管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
补丁管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*
可用性和持续性管理	*该能力尚未进行 AI 增强，请参阅亚马逊云科技 CAF。*

事件和问题管理

识别和管理不可预见的 AI 行为。

AI 系统常用于一个人的知识不足以掌握或解决某个问题时。AI 系统的这种属性让它很难理解系统的整体行为和边缘用例，所以很难预见性能在一定时间后的潜在衰退程度。因此，从业者会利用代理和简化统计来研究 AI 系统。这些简化的 AI 系统视角在观察和监控 AI 时非常关键。这一现象在开发早期阶段已经得到验证，但在 AI 系统实际投入使用时就显得尤为重要。

务必确立一套实践，承认 AI 系统虽经验证但无法完全证实，并且需要持续不断的控制与观察。例如训练服务偏差，即实验室开发的 AI 系统性能与生产环境中的表现有显著差异。必要时，应允许客户和用户标记标记系统输出的不理想或错误结果，并为他们提供直接报告此类事件的渠道。从一开始，就要为数据变化及随之而来的性能漂移、训练-服务偏差、“黑天鹅”事件以及未观察到的数据点做好准备。在系统允许的前提下，提供从容应对失败的方法，报告和响应此类事件，从中吸取教训。请留意系统使用体验不佳的客户和用户，他们的真实使用场景很可能在模型训练数据中没有得到充分覆盖。最终，应预期这些事件会发生，若未曾报告此类事件，反而应当提高警惕。随着您的 AI 系统的规模和复杂程度不断增加，这一挑战将愈发严峻。例如，与简单的决策树相比，基础模型的修正和监控难度明显要大得多。

性能和容量

监控并管理 AI 工作负载的性能。

AI 的开发周期与传统软件截然不同，因此性能和工作负载特性也有所不同：AI 开发初期主要侧重于数据探索，成本和性能要求需要能够适应众多差异巨大的工作负载。这些工作负载通常以实验和训练为主，需要强大的机器、专用硬件和高效的内存架构。您可以利用云计算来支持这些多样化的工作负载，因为云能够动态适应每种工作负载特性，而这些特性在开发生命周期的某些特定点才会偶尔出现。

随着时间的推移，训练和优化后的预处理逐渐占据主导地位，工作负载特征变得更为稳定和可预见。创新速度将受到您适应这一新特征并在这两种状态之间快速、持续切换能力的影响，同时保持开发与生产环境的清晰界限。确保模型构件和支持这些优化工作负载的数据可用来应对潜在的回退需求。一旦模型进入部署和运行阶段，应确保针对非功能性要求（如延迟或吞吐量）进行推理优化，并建立成本控制、性能和容量的监控机制。在 AI 生命周期管理能力中，我们介绍了机器学习运维成熟度模型，可参考它获得更深入的运营洞见。随着时间的推移，多种类型的工作负载特征会交织在一起，与数据科学家在独立开发（通常称为实验室环境）时所经历的往往大不相同。您可以深入研究亚马逊云科技卓越架构框架及机器学习剖析，了解如何在云端设计此类系统的架构。

总结

在本白皮书中，我们对 CAF-AI 进行了概述，介绍了客户如何企业和构建其 AI 历程，成功实现这一目标需要具备的能力，以及迭代这些能力的心智模型。本白皮书提及的基础能力可作为您与 AI 专家进一步研究、学习和探讨相关内容的索引。所有这些能力均与亚马逊云科技云采用框架有关，使企业能够在思考其云迁移历程的同时，也能规划其 AI 发展路径。

贡献者

本白皮书的贡献者名单如下：

- Alexander Wöhlke，生成式 AI 创新中心高级机器学习战略师，亚马逊云科技 CAF-AI 项目主管
- Caleb Wilkinson，生成式 AI 创新中心高级机器学习战略师，亚马逊云科技 CAF-AI 项目主管
- Payal Vadhani，亚马逊云科技专业服务安全总监
- Mayank Jain，亚马逊云科技专业服务首席高级经理
- Michael Sinnwell，亚马逊云科技专业服务高级安全数据分析师
- Mark Lieberg，亚马逊云科技专业服务高级安全顾问
- Matias Undurraga，现代化创新转型项目转型架构师
- Tony Santiago，WW 合作伙伴解决方案架构师，CAF 平台视角主管
- Dr. Saša Baškarada，亚马逊云科技云采用框架全球负责人
- Neil Mackin，机器学习解决方案实验室首席机器学习战略师
- Shuja Sohrawardy，生成式 AI 创新中心高级机器学习战略师
- Emily Soward，亚马逊云科技专业服务数据科学家
- Margaret Sharp，亚马逊云科技专业服务参与安全技术项目经理
- Ana Echeverri，亚马逊云科技全球专家部门 (WWSO) 高级 AI 服务专家，CAF-AI 评估主管
- Phil Le-Brun，亚马逊云科技企业战略部总监

延伸阅读

如需了解更多信息，请参阅：

- 亚马逊云科技云采用框架 (CAF)
- 亚马逊云科技卓越架构框架：机器学习剖析
- 亚马逊云科技卓越架构
- 亚马逊云科技架构中心
- 亚马逊云科技规范性指南
- 亚马逊云科技白皮书和指南

文档修订记录

如需获取本白皮书的更新通知，请订阅 RSS 源。

变更	描述	日期
更新	更新和扩展了概述章节，增加了安全性、平台和治理视角。	2024 年 2 月 13 日
首次发布	本白皮书首次发布	2023 年 5 月 22 日

注：

若要订阅 RSS 更新，您的浏览器必须启用 RSS 插件。

重要须知

客户需自行对本白皮书中的信息进行独立评估。请注意，本白皮书：(a) 仅供参考；(b) 仅代表亚马逊云科技当前产品与实践，如有变更，恕不另行通知；(c) 不代表亚马逊云科技及其附属公司、供应商或许可方做出任何承诺或保证。亚马逊云科技的产品或服务将“按原样”提供，不附带任何明示或暗示的担保、声明或条件。亚马逊云科技对客户的责任和义务受其与客户订立的协议的约束，本白皮书不属于亚马逊云科技与客户之间订立的任何协议，也并非对任何协议的修订。

亚马逊云科技名词解释

如需获取亚马逊云科技最新术语，请查阅《亚马逊云科技名词解释》。