



驯服AI代币陷阱

运营总监和采买总监如何控制增长最快的外部支出类别

OLIVER WYMAN
A MARSH BUSINESS

核心论点

人工智能正成为一个独特且快速增长的外部支出类别，正在改变着每家组织。残酷的现实是，大多数公司尚未准备好管理这场转型。

全球范围内，收养（或指“采用”）AI技术已呈加速态势。当前的核心挑战在于如何扩大AI的规模应用，同时避免其成本失控。这需要建立跨多个成本项目的可见性、控制力及商业纪律，这些成本项目包括多个业务单元试点、SaaS附加组件、模型订阅、使用量、云基础设施以及转型预算，还有AI赋能服务合同。

过去的经验表明，这一转型可能会带来痛苦。许多迹象表明，人工智能可能正沿着与云服务采用类似的路径发展。与供应商的持续讨论让人联想到十年前的情景。支出增长迅速，且管理严重不足。尽管人工智能的实施尚处于早期阶段，但关键依赖关系和经济表现恶化可能最早只需几年时间。预期的收益或许无法完全

具体来说，尤其是目前这些收益是基于假设而非可衡量的。

尚未全无希望。Oliver Wyman近期强调了CFO和CPO如何利用人工智能来优化其职能内的采购流程和运营模式。采购部门通过实施人工智能将获得的知识，应能使CFO和CPO在人工智能的采纳中发挥核心作用，帮助整个组织将同样的透明度和纪律应用于所有人工智能成本。

若要企业将 AI 的潜在用例优势转化为实际影响，AI 管理至关重要。这需要让 AI 成本透明化，管理需求，通过模型分层来控制访问权限和消耗，协商供应商经济条款，并确保捕获生产力提升带来的收益。

如今，人工智能的成本往往失控——组织内没有人掌握全局，也无法真正掌控其真实成本。挑战在于，在不扼杀创新的前提下，如何控制人工智能快速增长的成本。

早期解释：AI实际上成本是多少？

根据 Gartner® 咨询公司发布的报告¹，预计到2026年，全球人工智能支出总额将达到2.52万亿美元，同比增长44%。我们觉得这一数字融合了多种性质差异显著的要素，而这些要素很大程度上是由外部供应商驱动的。

我们常常将AI成本误认为是仅限于token的费用。实际上，总AI成本公式融合了其他影响因素，这些因素会显著影响总成本线，尤其是在AI应用规模扩大之后。除了访问和使用成本之外，总AI成本还包括安全与赋能成本，嵌入在其他要素中的供应商溢价（例如SaaS附加服务和云服务），以及专业服务（参见附表1）。

图1：总AI成本公式

总计 人工智能成本	=	访问	+	使用	+	赋能	+	供应商 保费
		谁能 访问人工智能？		多少人工智能 用户消耗？		需要什么 让人工智能安全运行 规模化吗？		供应商提供什么？ 内置人工智能收费 现有产品 服务？
内容		Seats, subscriptions, enterprise licenses		Tokens, application programming interface calls, agent actions		Cloud compute, integration, orchestration, security, governance, implementation services		AI powered capabilities embedded in SaaS, cloud, professional services, and IT contracts
趋势		More than a three-fold, year-over-year increase in employee AI subscriptions		Overall increase due to sharp volume increase, 并非完全抵消 by the token unit price reduction (for foundation models)		Sharp increase, correlated to number of use cases and the size of the user population		Already visible price increase trends, highly likely to rise higher in the future

奥利弗·怀曼分析

¹ Gartner新闻中心，Gartner®表示，2026年全球人工智能支出总额将达到2.5万亿美元，2026年1月。Gartner是Gartner, Inc. 和/或其附属公司的商标。

透明度正在降低，因为提供 AI 基础模型的供应商正将其定价模式从纯粹的按座次访问模式转向混合式变现。访问通常通过订阅来定价，而规模使用则通过代币、API 调用、代理、上下文和计算费用进行变现。对于购买者而言，这意味着 AI 支出不再像固定的 SaaS 许可证那样，而更像是一种云服务式的消费类别。

对于企业客户而言，转向混合收入模式最重要的因素是混合定价模式所带来的商业曝光机会。企业需要意识到，虽然他们可能通过座位费支付访问权限，但一旦AI嵌入工作流程，他们也可能会产生可变使用成本。

基础模型相对的“黑箱”性质意味着，组织可能仅对AI定价结构中嵌入的溢价有有限的理解。这种不透明带来的风险可能导致成本控制失效，因为它限制了买方有效质疑已发生成本的能力。

对于管理层而言，混合定价模式带来了复杂的挑战。要有效管理成本，就需要针对每个组成部分采取不同的控制措施。许可证需要权限管理。代币和API调用需要消耗治理。基础设施成本管理需要架构和云端的纪律性。数据使用管理需要执行权利和再利用规则。服务成本管理需要控制范围。最后但同样重要的是，要探究基础模型中内含的溢价，需要应对商业挑战。

人工智能的管理方式需要与以往的技术实施略有不同：

- 人工智能成本需作为独立的支出类别进行管理，而非仅仅是组织内一系列孤立采购的总和
- 由于该挑战具有全局性，首席采购官需要协助塑造人工智能支出类别的议程

买家越快将人工智能视为一个完整的支出类别越好，因为人工智能的使用范围正在急剧增长，而供应商格局和定价模式都在快速演变。在不久的将来，影响、塑造和控制支出的机会窗口可能会关闭。

当人工智能支出失控时

人工智能（AI）支出目前正随着使用量的指数级加速而快速增长，这对代币价格产生了混合效应，并且由于按使用量计价已成为常态，现在给客户带来了更大的通胀风险敞口。

随着越来越多的用户获得访问权限，越来越多的工作流消耗着代币和计算成本，并且越来越多的供应商引入了AI赋能的定价，这一趋势将继续下去。由于商业案例假设了难以验证的生产力效益，因此情况变得更加具有挑战性。

人工智能定价仍然不透明。 当前价格演变呈现非对称性，人工智能使用单位价格下降的同时，接入和嵌入式价格却在上涨。

当前，在组织迅速投身于人工智能及其成本，且所处环境变化极快的形势下，大多数企业尚未实施充分的成本控制也就不足为奇了。这项任务十分紧迫。

控制成本需要关注五个主要挑战：

1 低能见度

人工智能成本高度分散——涉及预算、供应商、试点项目、续订和使用模式。

2 弱控制

需求与消费的增长速度往往快于批准、融资或价值追踪机制。此外，供应商发票并不总是准确的，可能导致不合理的额外成本。

3 规格管理不善

用户获得超出其使用场景所需的高级功能、更高型号等级或访问权限。

4 有限的谈判筹码

买方往往缺乏基准、使用预测、成本效益分析或可信的替代方案。

5 过时的定价模式

传统的定价结构往往无法体现由人工智能驱动的生产效率和基于消费的成本。或绩效激励，阻止了买方充分实现人工智能带来的经济效益

在多个方面，人工智能在云计算方面的现状，呼应了以往企业实施云计算时的经验。

维度	云体验	今天的人工智能
消费动态	Cost accrued per compute hour, scaling after a single provisioning decision	Cost accrues per token and per action, scaling after a single access approval
影子收养	Capacity was accrued locally, outside procurement's line of sight	Pilots and subscriptions run on business unit budgets and developer cards
定价不透明度	Complex, non-comparable SKUs, 进行基准比较很困难	Non-standard pricing models, with similar workloads varying several-fold, 取决于配置
供应商依赖性	Hyperscaler lock-in through data gravity and migration cost, adding to underlying sovereignty risks exposure	Small number of Foundation Model providers, data gravity, and multi-year AI-inclusive SKUs Highly dependent on US-based players, with emerging Chinese alternatives

另一方面，有两点使得人工智能的治理比云计算曾经面临的要困难得多。第一个区别在于，人工智能使用所产生的大量成本仍然是不透明的。在许多情况下，嵌入基础模型中的费用上涨永远不会体现在消费账单上，因此无法受到用于控制云支出的类似监控机制的管理。

第二个方面源于使用模式的差异。云优化计算用例已经得到应用。相比之下，人工智能能力正被应用于新的用例，其效益尚待捕获。

将人工智能支出控制在可控范围内

将AI支出置于控制之下的第一步是认识到，AI需要被视为一个完整的对外支出类别，而不仅仅是在软件预算中的一条项目，或是一个临时的IT项目。作为一个支出类别，它需要以结构化的方式进行管理——也就是说，作为一个经常性的、多供应商的成本，一个具有高度战略价值且渗透到企业所有职能的成本。

从以往实施新IT的经验中可以吸取的教训是，不要一开始就集中控制支出。相反，应在不必要的成本和采购习惯嵌入组织结构并固化为制度之前，首先在整个组织中建立必要的商业纪律。

AI类别管理手册：规模化中的纪律

管理人工智能支出需要采用不同于许多其他成本类别的不同策略。

人工智能在其本质上是跨职能技术。由于其固有属性，人工智能的实施必然涉及企业多个职能部门。这一成本类别不可能有单一的所有者。

采购职能在此处尤其相关（创建商务结构），与独立业务（产生需求）以及IT（作为技术平台和仪器的创造者）并列。财务提供最终的成本控制。

采购的角色具体且关键，但目前却常常被严重低估。它是唯一掌握必要商业和供应链市场情报的职能，这些情报对于满足其他职能的要求以及以合同和商业条款的形式交付成果至关重要。但这并不意味着采购有权独自掌控人工智能。采购的角色是拥有跨职能运营模式中的商业维度，这能发挥其在市场情报方面的优势。

对于任何给定的外部支出类别，人工智能支出优化都需要做到购买更少、更好、更便宜。在实施人工智能以及用例和代理上线运行时，有效协调运用这三种杠杆是控制成本的关键。

1. 规格和访问权限管理

在使用方面，规格和访问控制是成本控制的核心；这是抑制过度配置的途径。在做出任何采购决定或授予访问权限之前，需要建立适当的访问级别和规格。

为有效控制人工智能成本而建立必要框架，需要首先在组织及职能层级识别并明确所需的人工智能能力水平，其次需建立处理非监管人工智能使用、消除过度配置、工具重复以及不合理使用扩张的方法。

一旦标准规范达成一致，就需要建立治理机制，以强制执行约定的访问和使用模型，并确保采购部门签订的合同能够反映公司在内的实际AI使用情况。

2. 消费监测

企业应当且将部署人工智能解决方案。由于人工智能成熟后运营成本未必能稳定；因此必须持续监控并积极管理。因此，每个人工智能应用案例都应通过组织的正常投资审批阶段 gate，并对预期的构建成本以及预期的长期持续运营成本进行明确评估。

监控必须在代币被消耗之前开始。每种材料使用案例、代理、试点和嵌入式功能都应向其所有者、业务目的、资金来源、供应商、预期生命周期成本、消耗模型和价值假设进行注册。一旦部署，应在个人、团队和应用层面监控消耗，以识别低效的使用模式并寻找优化成本的机会。

此外，采购部门将参与协商解决方案，以实现针对每种用途智能地选择“足够好”的模型（例如，为简单任务选择开源或经济型模型，为大型/复杂请求选择前沿模型）。

采购的角色很明确：随着AI使用模式每天都在演变，商业协议应定期审查，以确保价格反映实际消耗，而不是最初做出的静态假设。

3. 商业谈判

采购部门是关键领域，在此领域可以显著影响整体人工智能成本。

当权力大多集中在供应商一方时，CPO们将不得不努力工作，以避免供应商锁定和艰难的谈判。

COO和CPO可以通过确保支出透明度、供应商细分、基准开发、使用承诺、价格保护、生产力回报、基于价值的SLA、续订控制、审计权以及退出条款来实现这一点。

最顶尖公司的首席产品官（CPO）还可以利用与其最相关的早期用例，在与基础模型（LLM）供应商的谈判中创造商业优势，因为这些供应商如今正在寻找有价值的客户证明案例或用于营销目的。

CFO和CPO现在应该做什么

对管理层来说，好消息是，尽管人工智能可能与以往的技术实施有很大不同，但同样的规则依然适用。

在应对人工智能时，管理层需要应用以往应对其他类别时所采用的最佳实践。整个团队都需要参与其中。采购部门将发挥关键作用，同时业务用例负责人、人工智能卓越中心或首席人工智能官、IT部门以及财务部门也将参与其中。

该团队的目标需要确保遵循以下六项最佳实践。

确保对现有情况保持真实透明

第一步至关重要——必须确保对谁向谁购买什么、出于什么目的以及以什么价格进行购买完全透明。现实情况往往比想象中更令人不快。购买图谱常常揭示存在数十份零散的合同，包括通常带有AI组件（而无人意识到这一点）的SaaS订阅，以及签署时未包含退出条款的各种多年期承诺。

在类别存储库中创建一个AI分类法

正如我们所强调的，人工智能需要在组织内被视为一个独立的类别。它不仅仅是软件、咨询服务或IT基础设施；它同时具备这些属性，尽管单一类别都无法充分捕捉人工智能的独特性质。将人工智能作为一个类别来创建，并非采购部门的独有责任，而需要整个团队采用跨职能的方法共同思考。

新AI类别将包含子类别、评估标准和绩效指标。分类行为既是管理行为，也是技术行为。

建立内部专业知识——或快速获取外部资源

采购团队可能需要提升技能。为Foundation Model谈判合同与购买ERP许可证截然不同。AI定价模型极其复杂（涉及token、API调用、微调及计算），而服务水平协议（SLA）目前尚未标准化。这次谈判至关重要，因为虽然数据机密性至关重要，但通常的条款默认对己方不利。由于市场高度集中导致前所未有的权力失衡，此类谈判不可避免地非常艰巨。采购团队可以通过培训、定向招聘，或与专注于技术采购的公司建立合作伙伴关系来提升技能。

在做出任何承诺之前，建立资格认定流程

由于其持续的使用成本，人工智能采购需要严格管控。任何超出贵组织定义的阈值的新人工智能支出，都需要经过最低限度的采购资格流程。该流程应确定首选供应商并评估替代方案，评估模型运行地点（内部或供应商平台）的影响，分析关键合同条款（数据所有权、可移植性、审计），并验证其整个生命周期的商业模式。

这对于避免被锁定在单一供应商的合同中至关重要，该合同可能导致数月内出现难以控制的价格上涨。

准备进行AI商业模型的周期性重置

商业架构需要保留透明的基座服务价格，公开AI使用情况，并奖励经过验证的效率提升。

目标并非仅仅是强迫低价，而是要确保买方能获得更好的性能，并公平地分享生产率增长的收益。

与此同时，人工智能赋能的配送尚处不稳定状态。更高的吞吐量可以通过集成工作、治理、质量保证和人工监督来弥补。因此，生产力应进行衡量，并逐步纳入商业条款，而非事先假定。

定期重新定价或调整商业结构，就变成了一个核心流程，买方将寻求获取供应商效率提升所带来的公平份额。

表明立场 人工智能治理

越来越多的公司正在成立人工智能委员会，甚至任命首席人工智能官，他们加入执行委员会。

为确保成本控制，采购部门应在制定构建、采购或与AI供应商合作的战略决策时，积极向相关治理机构建言献策。这包括参与识别战略供应商和制定多年预算的决策。

结论

企业有一个短暂的机会窗口来实施控制不断上涨的AI成本的杠杆。这个窗口已经开始关闭。

人工智能因其独特性及其潜在影响，需要被视为一个独立的支出类别。它是少数几个呈指数级增长的支出领域之一。不断上升的成本是由快速增加的用户数量和使用案例、更高的使用强度以及嵌入供应商溢价相结合所驱动的。

CPO们需要帮助他们的组织应对这个新兴的AI市场，一个仍在发展和变化中的市场，其中合同标准尚未确定。尽管AI定价仍然远非透明，但目前谈判商业条款的机会有限。随着市场在未来两到三年内成熟，

经过数年发展，其（市场或行业）可能会变得更加集中和标准化。因此，购买者应预期商业模式将更加成熟，且在影响价格或合同条款方面缺乏灵活性。

采购职能在公司实现有效的AI治理方面可以发挥关键作用。在欢迎围绕AI的创新的同时，采购也有助于确保AI的实施尽可能稳健、经济高效和财务可持续。强大的需求和规范管理是完成这项任务的核心。

人工智能支出不仅是最快增长的成本类别，也带来了迄今为止对首席运营官、首席产品官以及整个组织能否有效控制不受约束的支出的最大考验。当然，人工智能供应商押注人工智能支出将保持不受约束。

威灵顿咨询公司（Oliver Wyman），Marsh公司（纽约证券交易所代码：MRSN）旗下业务，是一家管理咨询公司，凭借深厚的行业洞察力、大胆的创新精神以及一种能够化繁为简、协同合作的方法，帮助组织应对其最具定义性的转型时刻。Marsh是全球在风险、再保险和资本、人力资源与投资以及管理咨询领域的领导者，为130个国家的客户提供建议。公司年营收达270亿美元，拥有超过9.5万名同事，Marsh致力于通过视角的力量，帮助客户建立信心，实现繁荣发展。

欲了解更多信息，请访问 oliverwyman.com，或在领英（LinkedIn）和推特（X）上关注我们。

欧洲
+44 20 7333 8333

美洲
+1 212 541 8100

亚太地区
+65 6510 9700

印度、中东和非洲
+971 (0) 4 425 7000

作者

蒂博尔·罗谢

Partner, Performance Transformation Europe
thibault.rochet@oliverwyman.com

版权所有 ©2026 OliverWyman

版权所有。未经奥利弗·怀曼（Oliver Wyman）书面许可，本报告不得全部或部分复制或转载。
且奥利弗·怀曼对第三方在此方面的行为概不负责。

本报告中的信息及观点由 Oliver Wyman 整理。本报告不构成投资建议，亦不应被视为投资建议或替代专业会计师、税务、法律或财务顾问的咨询。Oliver Wyman 已尽合理努力使用可靠、最新及全面的信息和分析，但所有信息均不附带任何明示或暗示的保证。Oliver Wyman 不承担更新本报告中信息或结论的责任。对于因依据本报告或本报告中引用的任何报告或信息来源所采取或未采取的任何行动而导致的任何损失，或因任何后果性、特殊或类似损害（即使已被告知此类损害的可能性），Oliver Wyman 均不承担责任。本报告不构成买卖证券的要约或招揽买卖证券。未经 Oliver Wyman 书面同意，本报告不得出售。