



AI MARKETING / CAPITAL ALLOCATION
FLIGHT DECK

AI营销预算与 投资组合指南

把九类支出、五段总成本、三层资金池与证据调仓接成一套可比较、可追责、可退出的经营资本系统。

看见真实总成本

九类账户 × 建设、运行、控制、变更、退出

按证据管理仓位

种子、验证、核心、底座、观察与退出六类位置

用单位经济性比较场景

同时报告经营贡献、质量、资产、学习与风险

让预算持续调仓

月观察、季调仓、年重置连接采购与治理

出版说明

竹势 AI 营销智库出品

本报告面向企业董事长、CEO、CFO、CMO、CIO/CDO、业务单元负责人、营销运营负责人、采购与财务 BP、AI 项目负责人。报告提供预算账户、总拥有成本、资金池、单位经济性、投资组合、证据门、分摊与治理方法，用于设计企业自己的资本配置制度。由于企业完整预算、合同条款与内部成本通常不公开，本报告不提供虚构的市场均值、统一预算比例或行业通用 ROI；涉及企业案例时，均标明结果披露主体、测量边界与不可外推之处。

版权归竹势 AI 营销智库所有。公开引用请注明报告名称、出品方与版本时间。本文档不构成会计、税务、法律或证券投资意见。

执行摘要：董事会应掌握的十个判断

竹势 AI 营销智库出品

CAPITAL ALLOCATION RULE / 资本配置原则

AI 营销预算不是工具采购清单，而是对经营假设、学习速度、共享能力和下行风险的分阶段承诺。

钱花在哪里

证据成熟到哪

风险由谁承担

何时调仓退出

判断一：AI 营销预算的对象不是工具，而是一组尚待验证的经营假设。预算应随着证据成熟度释放；一次性购买多年许可，会把技术不确定性、采用不确定性和业务不确定性提前固化成沉没成本。

判断二：模型调用费通常只是可见成本。数据清理、系统集成、人工复核、质量评测、日志监控、流程变更和退出迁移，往往决定真实总拥有成本。

判断三：探索、共享能力和规模化运营应使用不同资金池。探索资金追求学习速度与损失上限；共享能力资金追求复用和治理；规模化资金追求稳定单位经济性和运营责任。

判断四：单位经济性必须落到一次合格任务、一个客户或一个端到端闭环。只看每千字成本或单次调用价格，会把低质量、返工、人工接管和转化损失排除在外。

判断五：高价值场景不一定优先。投资仓位由经营贡献、证据成熟度、复用性与风险共同决定；依赖未满足的高价值项目，应先投资底座或缩小范围。

判断六：加仓、持有、重构和退出必须由证据触发。战略必做项目也需要重构路径、限制损失和明确接管责任，不能以“战略”名义豁免审查。

判断七：中央平台承担可复用底座与治理，业务单元承担采用和经营结果。早期先 showback 建立透明度，待计量稳定与可控性成熟后再 chargeback。

判断八：低首价可能隐藏最低消费、专有工作流、数据迁移、人工服务、评测、监控和退出成本。采购比较必须使用同一工作负载、同一质量门和同一退出情景。

判断九：评测、日志、人工复核、合规和事故响应不是纯粹管理费用，而是压缩损失分布尾部的控制投资；控制强度应与场景风险等级匹配。

判断十：年度预算负责承诺边界，季度组合会负责调仓，月度观察负责识别异常。三者必须连接采购、绩效、风险与运营接管，形成持续运行的资本配置机制。

董事会问题	错误做法	本报告建议的决策对象
今年买哪些 AI 工具？	按厂商和部门列采购清单	为哪些经营假设分阶段承诺多少资金
哪个项目 ROI 最高？	使用未经控制的局部效率数字	比较单位经济性、证据、复用、依赖与风险
费用由谁承担？	全部压给中央或平均分摊	按可复用底座、使用量、可控性与结果责任分摊
何时停止？	合同到期才复盘或只看沉没成本	预先定义加仓、持有、重构、退出证据门
如何防止失控？	增加审批层级或完全禁止	按风险等级配置评测、复核、日志、准备金和接管机制

资本配置驾驶舱的核心仪表

驾驶舱不是一张财务看板，而是一套决议系统。它把九类账户、五段 TCO、三层资金池、单位经济性、组合仓位、依赖图、证据门、风险准备金和治理节奏连接起来。任何场景只有同时回答“为谁创造什么经营结果、证据有多成熟、共享能力能复用多少、损失如何封顶、谁接管运营”五个问题，才有资格进入组合。

仪表	回答的问题	最低数据要求
九类账户	钱花在何处，是否被藏在其他科目	供应商、成本类型、业务场景、责任中心
五段 TCO	全生命周期承诺是多少	建设、运行、控制、变更、退出成本
场景成本卡	一次合格任务或闭环的成本是多少	工作量、用量、人工、质量、返工、异常
组合仓位卡	应加仓、持有、重构还是退出	贡献、证据、复用、风险、依赖
依赖关系图	底座缺口会阻塞哪些场景	数据、接口、权限、评测、运营依赖
准备金表	尾部损失由谁承担	风险等级、暴露量、控制成熟度、接管成本

第一章 预算对象：经营假设、学习速度与下行风险

竹势 AI 营销智库出品

经营假设	
客户结果 · 单位经济性 · 责任 owner	
01	账户 九类资源
02	证据 五层成熟度
03	仓位 六类组合位
04	风险 损失上限
05	决议 加仓 / 持有 / 重构 / 退出

1.1 工具采购清单掩盖了真正的资本承诺

传统预算把软件许可、云资源、外包服务和培训分散在不同科目，管理层看到的是一组供应商和合同，却看不到这些支出共同支持哪一个经营假设。AI 营销尤其容易出现这种错位：内容团队购买生成工具，数据团队采购模型平台，IT 建设知识库，业务单元再购买代理商服务。每笔支出看似合理，合并后却可能服务同一个低价值任务，也可能没有任何一方对最终经营结果负责。

董事会应把每项投入重写为可证伪命题，例如：“在品牌合规率不低于现状、人工复核时间下降 40% 的前提下，销售线索培育闭环的单位成本可在两个季度内下降，并且知识产权可由三个业务单元复用。”命题必须包含对象、基线、质量门、时间窗、结果指标和边界。预算不是对技术能力的信任票，而是对该命题的分阶段承诺。

1.2 资源配置、交易成本与实物期权的连接

Drucker 将管理者的核心责任之一放在资源配置与机会选择上。对 AI 营销而言，稀缺资源不仅是现金，还包括业务专家时间、数据授权、变更承受力和管理注意力。Coase 与 Williamson 的交易成本和资产专用性提醒管理层：低价外购并不自动优于自建，合同谈判、接口适配、监督、数据迁移和专有流程会形成持续成本；自建同样可能形成对少数工程人员和内部架构的锁定。

实物期权提供了更合适的早期投资逻辑：先购买有限损失的学习权，而不是一次性购买全面规模化。小额试验的价值不只来自直接收益，还来自减少不确定性、发现依赖、验证采用和保留未来选择。修正点在于，企业并非所有投入都可自由等待。强监管控制、数据安全和基础审计可能是不可选投入；此时仍可优化路径和节奏，但不能以短期回报不足为由取消。

1.3 案例：P&G 与哈佛商学院的随机现场实验

主体与时间：Harvard Business School、Wharton 等研究者与 P&G 开展的“The Cybernetic Teammate”现场实验，研究于 2024 年形成工作论文并持续公开讨论。业务情境：791 名 P&G 专业人员执行真实的新产品创新任务。具体动作：随机分配个人或团队、使用或不使用生成式 AI，比较输出质量、时间与跨职能协作。作用机制：AI 提供知识补充与快速方案生成，使个人在部分任务上达到团队式能力。可核验结果：研究报告显示，使用 AI 的个人可在更短时间内产出更高质量方案，且跨职能差异缩小。归因限制：实验任务有明确边界，结果不能直接外推到持续运营、品牌安全、系统集成或真实收入。管理启示：探索预算应购买“任务证据”，不能把单次实验效率直接资本化为年度经营收益。[S1]

第二章 九类预算管理账户

竹势 AI 营销智库出品

软件

许可与席位

模型

调用与推理

算力

容量与流量

数据

授权与质量

内容

素材与权利

集成

接口与异常

培训

任务胜任度

运维

监控与接管

顾问

实施与转移

九类账户是管理视图，不要求与法定会计科目一一对应。其目标是把分散支出重新归集到场景、能力和责任中心，并允许同一张发票拆分到多个管理账户。每个账户至少记录：场景、资金池、成本阶段、固定/变动属性、可复用范围、责任人、合同期限和退出条件。

账户	包括什么	常见漏项	关键控制指标
1 软件	SaaS 许可、席位、插件、工作台、API 管理平台	闲置席位、环境重复、附加模块	活跃席位率、功能重叠率、单活跃用户成本
2 模型	输入输出用量、微调、托管推理、专用模型许可	重试、长上下文、代理循环、缓存策略	每合格任务模型成本、失败重试率
3 算力	GPU/CPU、存储、网络、容器、推理集群	峰值冗余、跨区流量、闲置资源	利用率、单位工作负载成本、承诺覆盖率
4 数据	采集、清理、标注、授权、脱敏、主数据、保留	历史数据整理、版权许可、删除与归档	可用数据覆盖率、质量缺陷成本
5 内容	素材采购、版权、拍摄、编辑、翻译、人工创作	品牌审核、版本管理、渠道适配	单合格资产成本、复用率、下架率
6 集成	接口、连接器、身份权限、流程编排、测试	接口变更、异常处理、沙箱与回滚	每接口维护成本、失败率、变更前置期
7 培训	角色训练、操作演练、管理者辅导、认证	新人培训、业务规则更新、岗位再设计	关键任务采用率、胜任达标时间
8 运维	监控、支持、事件响应、容量、版本管理	夜间值守、知识更新、人工接管	可用性、事件恢复时间、接管率
9 顾问	诊断、方案、实施、治理、评测、持续优化	驻场协调、供应商管理、迁移支持	成果验收率、内部能力转移率

2.1 Activity–Based Costing 的当代应用

Kaplan 与 Cooper 的作业成本思想适合解决“共享资源如何落到真实活动”的问题。模型、数据平台和运营团队常被平均分摊，导致高消耗场景获得隐性补贴，低消耗场景承担不合理成本。企业应识别成本动因：模型以合格任务、有效上下文长度和重试次数为动因；人工复核以风险等级、异常率和复杂度为动因；数据治理以数据域、变更频率和质量缺陷为动因。管理账户可以服务决策，即使财务报表仍按软件、服务或人工分类。

2.2 案例：腾讯营销服务的 AI 投入与成本边界

主体与时间：腾讯控股 2025 年年度业绩及 2026 年第一季度业绩。业务情境：微信生态广告推荐、闭环营销和广告创作。具体动作：升级 AI 驱动的广告推荐模型，扩展闭环营销能力，并增加 AI 相关设备投入。可核验结果：腾讯披露 2025 年营销服务收入同比增长 19% 至人民币 1450 亿元；2026 年第一季度营销服务收入同比增长 20% 至人民币 382 亿元，并说明广告表现和定价改善。同期营销服务毛利率受到 AI 设备折旧和相关运营成本影响。归因限制：收入增长同时受用户参与、广告负载、行业广告需求等因素影响，不能全部归因于 AI；披露的是平台经营结果，不是广告主单位经济性。管理启示：AI 预算必须同时追踪收入贡献和设备折旧、运营成本，避免只记录模型费用。[S2][S3]

第三章 五段总拥有成本

竹势 AI 营销智库出品

BUILD 建设 需求 · 数据 · 集成
RUN 运行 用量 · 人工 · 支持
CONTROL 控制 评测 · 日志 · 合规
CHANGE 变更 升级 · 迁移 · 再培训
EXIT 退出 导出 · 重建 · 过渡

五段 TCO 把采购价格扩展为生命周期承诺。企业应在立项时估算区间，在试点后更新实测，在规模化前签署运营和退出责任。任何只报建设费或年度许可费的方案，都不能进入组合比较。

阶段	核心成本	典型决策问题	责任人
建设	需求、数据准备、集成、配置、测试、初始培训	达到首个真实闭环还需投入什么	业务 owner + IT/数据
运行	许可、模型、算力、内容、人工复核、支持	每个合格单位的稳定成本是多少	运营 owner + 财务 BP
控制	评测、日志、审计、合规、红队、事件响应	损失尾部是否被控制在容忍度内	风险/法务/安全
变更	模型升级、接口变化、业务规则、再培训、迁移	变化发生时由谁出钱、多久恢复	产品/平台 owner
退出	数据导出、工作流重建、并行运行、合同终止、知识转移	能否在可接受时间和成本内离开	采购 + IT + 业务

3.1 TCO 公式与情景法

建议使用三种情景：基准、压力与退出。基准情景反映预期使用量和质量；压力情景加入调用增长、人工接管率上升、接口频繁变化和监管要求；退出情景模拟供应商不可用、价格调整或战略切换。TCO 不是单一精确数，而是一个随证据收窄的区间。早期最大的价值在于识别成本驱动因素，而不是追求小数点后的准确。

推荐公式：五段 TCO = 建设一次性成本 + 运行期固定成本 + 单位变动成本×业务量 + 控制成本 + 预期变更成本 + 退出准备与迁移成本。对于低概率高影响事件，除期望值外还要单列最大合理损失和恢复时间，因为均值无法表示尾部风险。

3.2 FinOps 对 AI 成本的修正

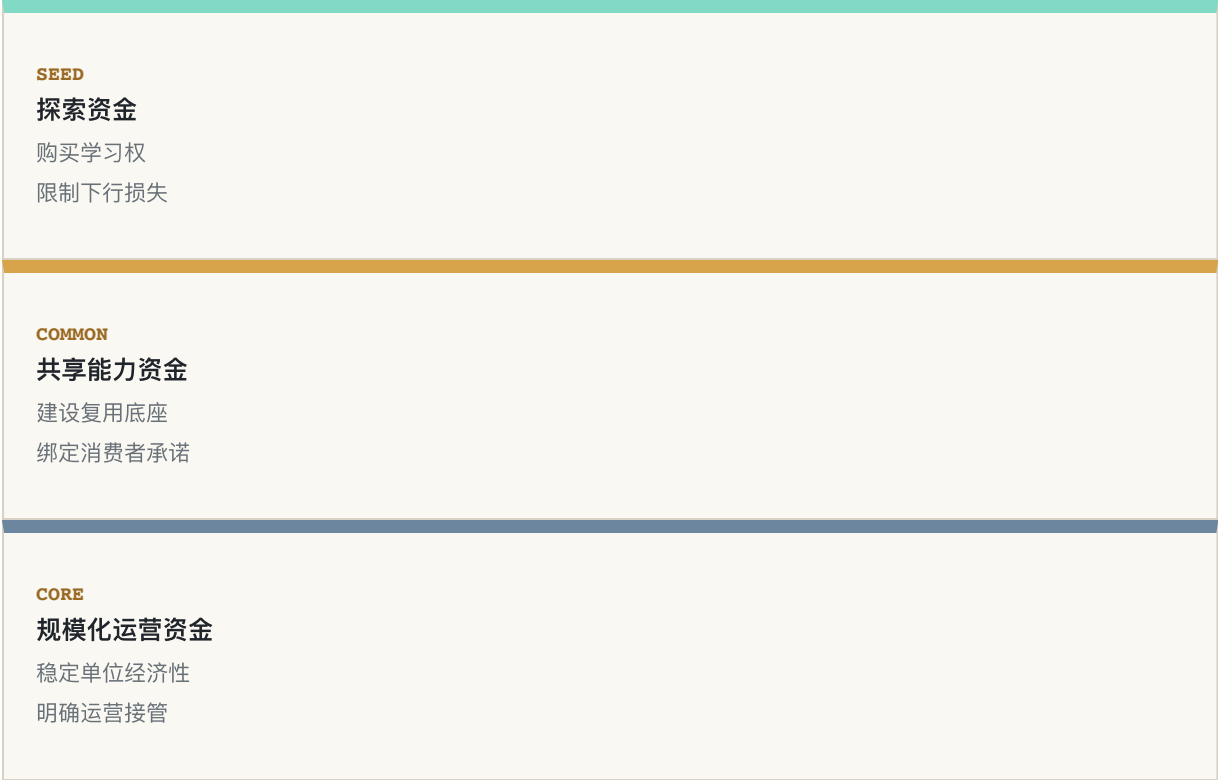
FinOps Foundation 2025 年报告显示，管理 AI/机器学习支出与获得单位经济性显著上升为实践重点；其 AI 成本估算指南强调从开发、试点到生产阶段使用不同估算方法。AI 工作负载与传统云资源不同：请求成本受上下文、输出长度、推理深度、重试、代理循环、缓存和模型路由共同影响；质量不足会把节省的调用费转化为返工和业务损失。因而成本优化必须在质量门之后进行。[S4][S5]

3.3 案例：Klarna 客服自动化的收益与反转

主体与时间：Klarna 于 2024 年披露其 AI 客服助手上线后的运行结果，随后在 2025 年公开强调将恢复更多人工服务选择。业务情境：高频消费者服务与交易支持。具体动作：让 AI 助手处理多语言客服会话，并将部分工作从外包和人工团队转移。可核验结果：公司曾披露助手处理约三分之二客服对话、平均解决时间下降，并估计年度利润改善；这些数据由公司测量。后续管理层承认过度强调成本可能影响服务质量，提出增加人工服务。归因限制：利润估计不是审计后的单场景净现现金流，且客户体验、长期留存和复杂问题成本难以完全计入。管理启示：规模化决策必须把质量、人工接管和品牌影响纳入运行与控制成本，不能只把自动化率视为成功。[S6][S7]

第四章 三层资金池与拨款权

竹势 AI 营销智库出品



资金池	目的	典型周期	拨款证据	主要决策权	停止条件
探索资金	验证问题、任务、采用与风险	2-12 周	基线、可测任务、损失上限、学习计划	业务 sponsor + 轻量评审组	无法获得关键证据、依赖不可控
共享能力资金	建设数据、集成、评测、权限、知识等底座	季度至年度	跨场景需求、复用承诺、架构与治理标准	CIO/CDO + CFO + 业务代表	复用不足、替代方案更优、维护责任缺失
规模化运营资金	稳定运行并扩大业务量	年度承诺、季度调仓	单位经济性、SLA、采用、风险、运营接管	业务 owner + CFO/CMO	单位经济性恶化、质量失守、需求消失

4.1 March 的探索与利用：不同资金池不能用同一回报要求

March 区分探索与利用：探索产生新知识但回报不确定，利用把已有能力转化为稳定产出。若财务部门要求所有探索项目立即达到规模化回报，企业会只选择容易测量的局部自动化；若共享能力没有复用责任，则会形成昂贵平台。三层资金池的关键不是预算标签，而是不同的决策权、证据标准、损失上限和停止条件。

4.2 防止搭便车、重复建设与预算藏匿

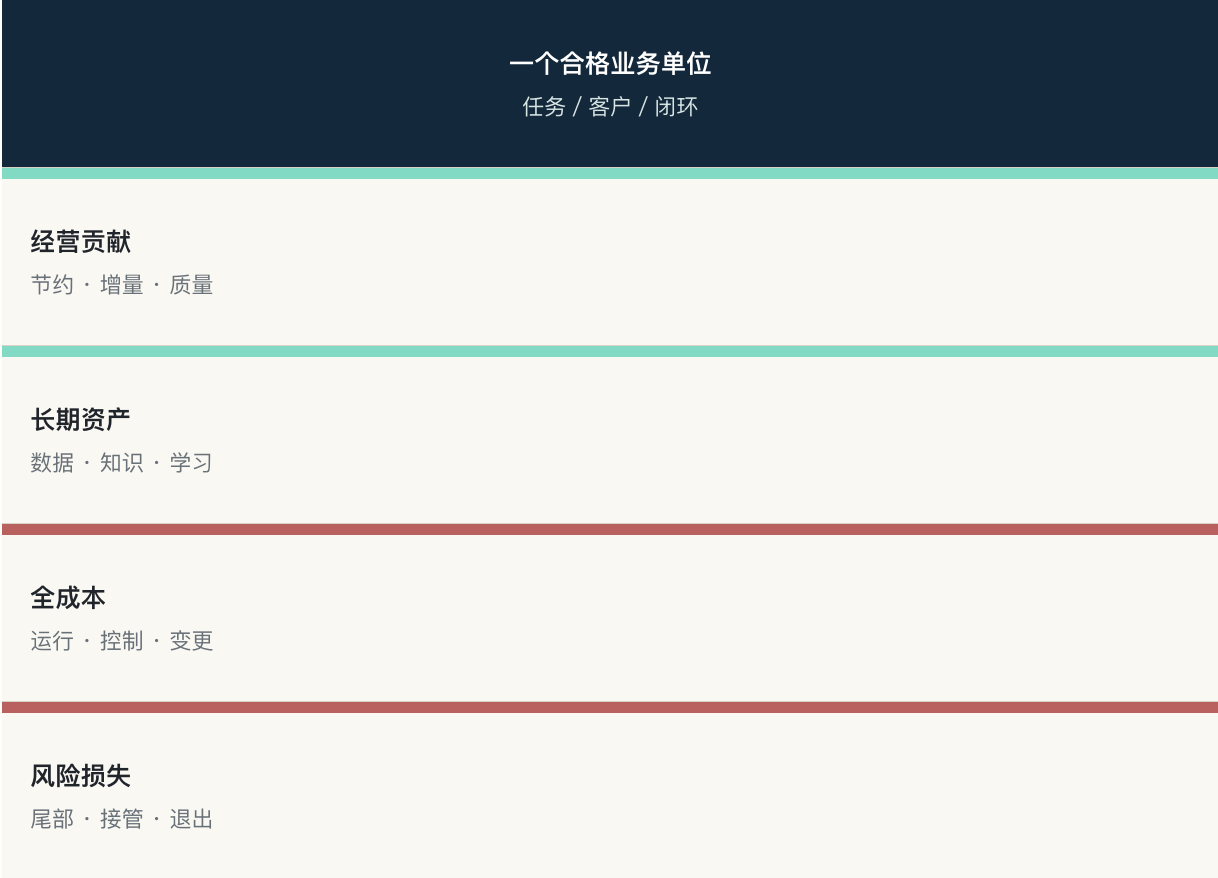
共享能力最容易出现两种问题：业务单元希望中央承担全部成本，却不承诺采用；中央平台以“战略底座”名义持续扩张，却不证明复用。治理上应采用能力赞助人制度：每项共享能力至少有两个真实场景承诺使用，明确首批消费者、服务目录、成本动因、SLA 与退役条件。重复建设通过架构例外机制处理：业务可申请例外，但必须披露增量价值、兼容与退出成本。预算藏匿则通过九类账户和合同拆分识别，例如顾问费中隐藏长期人工运营，云合同中隐藏最低消费。

4.3 案例：中国工商银行等金融机构的不可选控制投入

主体与时间：招商银行在《2025年中期报告》中披露，截至2025年6月末，已在零售金融、公司金融、风险管理、运营管理等领域上线AI智能助手。业务情境：大型商业银行需要同时服务个人与企业客户，并在高监管环境中处理大量知识检索、业务办理、客户沟通和内部运营任务；单点工具容易形成重复采购、口径不一致和治理成本分散。具体动作：招商银行以企业级共享能力方式建设并推广多领域AI助手，把模型能力接入岗位流程和业务知识，覆盖一线服务与中后台运营，而不是只采购面向个人的通用对话工具。作用机制：共享模型、知识、权限和运营能力降低不同部门重复建设成本；岗位助手把检索、生成与流程提示嵌入真实任务，使技术支出能够按使用场景、节省工时和服务质量进行观察。公开结果：该行披露AI智能助手累计节省475万小时工作量；该数字由招商银行自行测量和披露，报告未公开完整预算、计算口径明细、对照组、质量变化、客户经营增量或现金兑现比例。归因限制：节省工时不自动等于费用下降或收入增加，结果可能同时受到流程改造、数据治理、员工采用和其他数字化投入影响，不能据此倒推出统一ROI。适用边界：适合已有较成熟数据、知识、权限和运营体系的大型机构；中小企业不应照搬平台规模，而应先选择高频任务建立基线。预算管理启示：具备跨部门复用条件的模型、知识、评测和权限能力应进入共享能力资金池；业务部门仍需对采用、质量和工时兑现负责。共享底座的拨款证据应同时包括活跃消费者、合格任务量、单位成本、质量尾部与可兑现结果，而不能只报告累计调用或理论节省。[S31]

第五章 场景单位经济性树

竹势 AI 营销智库出品



5.1 先定义“合格单位”，再谈成本下降

单位可以是一次合格任务、一个被有效服务的客户、一个合格线索、一项完成的营销战役或一个端到端闭环。选择原则是：单位必须与业务需求驱动因素对应、可重复计量、包含质量门，并能连接到价值。内容生成以“被批准并发布、达到品牌和合规门槛的资产”为单位，而不是草稿数量；线索培育以“在时限内完成规定触达并产生可验证下一步”的客户为单位，而不是发送消息数。

价值分支	度量方式	常见误区	修正办法
成本节约	单位人工、外包、工具与返工成本变化	把释放工时等同现金节约	区分可释放、可重配和不可兑现工时
质量	准确率、品牌一致性、合规、客户体验	只看平均分忽略严重缺陷	设置硬门槛与尾部缺陷率
增量经营结果	收入、毛利、转化、留存、速度	把相关性当因果	使用对照、分层试验或保守归因
资产沉淀	数据、知识、连接器、模板、评测集复用	把文档数量当资产	要求可调用、可治理、有人维护
学习价值	减少关键不确定性、发现依赖	把活动量当学习	预先写明待验证假设与决策影响
风险	预期损失、最大合理损失、恢复时间	只看事故发生率	同时计入暴露量、严重度和控制能力

5.2 单位经济性树的推荐公式

单位净贡献 = 可兑现成本节约 + 保守归因的增量毛利 + 质量改善的经济价值 + 资产复用摊销 + 学习期权价值 – 单位运行成本 – 单位控制成本 – 预期风险损失。对于品牌、知识和组织能力等长期资产，不建议强行货币化到虚假精度，可采用“货币结果 + 资产指标 + 风险指标”的三栏报告，并设置明确的后续验证时间。

5.3 场景级成本卡

字段	填写要求
场景与合格单位	明确输入、输出、质量门、业务量和责任人
基线	现有人时、外包、工具、周期、质量、转化与风险
九类账户成本	按固定、变动、共享、一次性拆分
五段 TCO	至少覆盖基准、压力、退出三种情景
价值证据	区分实测、对照、估算、未披露
依赖	数据、接口、权限、人员、流程、供应商
决议门	下一次加仓/持有/重构/退出所需证据
损失上限	预算上限、数据暴露、客户影响、恢复时间

5.4 案例：营销知识工作实验的单位选择

多项受控实验显示，生成式 AI 可在写作、分析和创意任务上缩短时间并改善部分质量，但能力边界依任务而变。若企业把“每份初稿成本”作为单位，会忽略事实核查、品牌调整和审批；若改为“通过质量门并进入下一业务环节的成果”，模型优势可能缩小，也可能因标准化而更稳定。管理启示：试点第一周最重要的工作不是部署工具，而是定义合格单位、基线和失败类型。

第六章 AI 营销投资组合与依赖关系

竹势 AI 营销智库出品



6.1 四维组合矩阵

维度	核心问题	评分证据	高分的含义
经营贡献	对收入、毛利、现金流、客户价值或风险有多大影响	基线、对照、业务 owner 承诺	结果重要且可归因
证据成熟度	任务、采用、质量、技术、经济性证据是否完整	实验、生产数据、稳定周期	不确定性已明显收敛
复用性	能力能否跨团队、渠道、品牌或场景复用	实际消费者、接口标准、维护机制	边际扩展成本下降
风险	错误、数据、合规、锁定和集中风险	风险评估、控制测试、退出演练	风险低或已被有效控制

组合评分不能机械相乘后自动决策。四维用于结构化讨论，最终仓位还需考虑相关性、依赖和集中度。多个场景可能共享同一模型、数据源或供应商，表面分散但实际具有共同故障点；一个底座项目自身没有直接收入，却可能解锁多个高价值场景。

6.2 组合仓位卡

仓位	适用状态	资金动作	管理要求
种子仓	高贡献假设、低证据、损失可控	小额探索资金	短周期、强基线、明确停止
验证仓	任务与采用证据初步成立	增加受控业务量	补齐质量、经济性和风险证据
核心仓	单位经济性稳定、复用与运营成熟	规模化运营承诺	SLA、容量、接管、季度调仓
底座仓	跨场景复用、直接收益难单列	共享能力资金	服务目录、消费者承诺、退役条件
观察仓	价值存在但依赖或证据不足	冻结扩张、保留最小能力	明确等待事件和复审日期
退出仓	需求消失、经济性失守或风险不可控	停止新增、执行迁移	知识转移、数据导出、合同收尾

6.3 依赖关系图的读法

依赖图以场景为右侧节点，以数据、集成、身份权限、评测、知识、运营团队和供应商为左侧能力节点。边表示“缺失即不能稳定运行”的关系。优先级由“解锁价值总量÷建设与风险成本”决定，而不是由提出需求的部门级别决定。依赖图还能识别集中风险：若十个场景都依赖同一不可替代模型或一位工程师，组合并未真正分散。

6.4 案例：摩根士丹利财富管理的共享知识底座

主体与时间：Morgan Stanley Wealth Management 自 2023 年起公开部署面向顾问的 AI 知识助手。业务情境：大量内部研究与政策资料，顾问需要快速获得受控答案。具体动作：建设检索、权限、评测和受控发布机制，并逐步扩大顾问使用。可核验结果：公司与技术供应商披露了高采用和知识检索效率，但未公开完整预算与净财务回报。结果主要由企业与供应商共同披露。归因限制：财富管理的知识治理、合规和高价值人工服务与一般营销内容不同。管理启示：共享知识、权限与评测属于底座仓，价值来自多个任务复用；若只要求单一助手项目承担全部底座成本，会低估组合价值。[S10]

第七章 四类证据决议门

竹势 AI 营销智库出品

加仓 证据允许扩大暴露
持有 等待明确事件与日期
重构 价值仍在，机制失败
退出 停止新增并安全迁移

决议	触发证据	预算动作	必须回答
加仓	合格单位成本改善、质量达标、采用稳定、风险受控	扩大业务量或范围	新增规模是否保持经济性
持有	价值存在但证据窗口未完成或外部依赖待定	维持最小投入	等待什么事件、何时复审
重构	价值假设仍成立，但流程、技术、供应商或责任设计失败	停止扩张，重做关键环节	哪项机制失败、损失如何封顶
退出	需求消失、单位经济性长期失守、风险不可接受或替代方案显著更优	停止新增、迁移与关停	数据、知识、客户和运营如何安全转移

7.1 证据层级

证据从弱到强可分为：演示可行、受控任务、真实流程试点、稳定生产运行、跨场景复用。演示证明“能做”，不能证明“值得做”；真实流程试点必须包括异常、人工复核、权限与数据；稳定生产至少覆盖一个有代表性的业务周期。每次拨款只要求与下一阶段相匹配的证据，避免早期过度证明，也避免用演示跳过生产证据。

7.2 战略必做项目的特殊处理

战略必做不等于无条件加仓。对于监管要求、关键平台迁移或竞争防御，退出业务目标可能不可接受，但仍可在路径、技术、供应商、范围和时间上重构。董事会应将“目标不可退出”和“当前方案不可退出”严格区分。

7.3 案例：IBM Watson for Oncology 的受限与退场启示

主体与时间：IBM 与多家医疗机构在 2010 年代推进 Watson for Oncology，后续业务收缩并出售相关资产。业务情境：将医学知识与患者信息结合，辅助肿瘤治疗建议。具体动作：投入知识工程、合作医院和产品化。可核验结果：部分研究报告一致性或可用性，但也出现适用性、数据本地化和建议质量争议；没有形成普遍可验证的经营成功。归因限制：医疗决策高度复杂，且不同国家指南、数据与临床流程差异大。管理启示：高价值和高战略叙事不能替代本地任务证据；当知识更新、流程适配和责任边界不成立时，应重构或退出，而不是因沉没成本继续加仓。[S11][S12]

第八章 中央与业务分摊机制

竹势 AI 营销智库出品



8.1 分摊原则：可复用底座中央承担，使用与结果业务承担

中央平台通常承担统一身份权限、基础模型接入、日志、评测框架、数据标准、连接器目录和供应商治理；业务单元承担场景配置、内容与流程运营、采用、业务数据质量和经营结果。共享能力可采用分层分摊：基础公共能力由中央预算；可计量增量资源按使用量；专属需求由提出方承担；合规强制能力可按风险暴露或受益范围分摊。

8.2 先 showback，后 chargeback

AWS 的云财务管理指南把 showback 定义为向业务展示可归属成本，chargeback 则进一步形成内部计费。AI 能力早期数据标签、计量稳定性和可控性不足，立即收费会诱发规避使用、隐藏账户或局部自购。建议先用 1-2 个季度 showback，公开消费、质量和价值；当服务目录、用量归属、预算可控和争议处理成熟后，再对可变资源或专属服务 chargeback。[S13]

成本类型	建议分摊	原因
统一安全、日志、身份	中央或风险池	属于企业控制底线
共享数据与连接器	中央为主，消费者承诺	具有显著复用和网络效应
模型与算力变动成本	按合格任务或使用量	可计量且与需求相关
业务规则、内容与运营	业务单元	由业务可控并直接影响结果
专属集成与定制	提出方承担	避免公共预算补贴专属需求
探索试点	探索池 + 业务共同出资	保证学习价值与业务承诺

8.3 Ostrom 的共同资源治理修正

Ostrom 对共同资源的研究提示，共享能力不能只靠中央命令或完全自由使用。有效制度需要明确边界、参与者规则、监测、渐进式约束和争议解决。AI 营销共享底座应有服务目录、消费规则、容量上限、例外审批、消费者委员会和退役机制。完全免费会导致过度消费，过早全额收费则会破坏采用与学习。

第九章 供应商、锁定与退出

竹势 AI 营销智库出品

低首价 最低消费	
	→
专有流程 数据与配置	
	→
隐性人工 代运营与复核	
	→
退出负债 迁移与并行运行	

9.1 低首价的十类隐藏负债

隐藏项	采购必须追问
最低消费	未使用额度是否结转，增长后价格阶梯如何变化
数据迁移	导入、清洗、映射、导出由谁承担
专有工作流	流程定义、评测集、配置能否导出和重建
人工服务	标价是否依赖供应商长期代运营或人工审核
人工复核	供应商演示是否排除了客户侧审核成本
评测与监控	质量、漂移、日志和告警是否另收费
模型替换	能否切换模型，切换后质量与接口如何保障
数据权利	输入、输出、反馈数据是否用于训练，保留多久
终止条款	提前终止费、自动续约、价格调整与服务降级
退出支持	导出格式、时间、费用、过渡期和知识转移

9.2 同一工作负载比较法

供应商比较必须固定业务任务、数据规模、质量门、峰值业务量、人工复核规则和退出情景。报价表至少包含三年 TCO 区间、每合格单位成本、控制成本、合同弹性和集中风险。最低单价不等于最低经济成本：低质量导致的重试和人工修正可能远高于模型差价；专有工作流提高初期速度，也可能提高退出成本。

9.3 自建也会锁定

内部锁定来自专有架构、缺乏文档、关键人依赖、无法替换的数据管道和长期维护负担。自建决策应同样提交退出卡：关键组件替代路径、人员备份、开源或标准接口、数据导出、灾备与运维预算。采购部门不能把“自建”视为无合同风险，IT 也不能把“外购”视为无架构责任。

9.4 案例：Adobe Figma 收购终止与平台集中风险

主体与时间：Adobe 于 2022 年宣布收购 Figma，2023 年因监管阻力终止交易并支付约 10 亿美元终止费。该事件不是 AI 营销项目，但清楚展示了平台战略、监管和退出条款如何形成巨额或有成本。业务情境：设计与创意工作流高度集中。具体动作：签署大型并购协议并设置终止安排。可核验结果：交易终止和费用由公司公告披露。归因限制：并购与 SaaS 采购规

模不同。管理启示：企业在将创意、品牌资产与协作流程集中到单一平台时，应模拟监管、所有权变化、产品策略变化和退出成本；合同弹性具有真实经济价值。[S14]

第十章 风险分级、准备金与控制投资

竹势 AI 营销智库出品

已知可估损失

频发缺陷与人工接管

压力情景缓冲

容量、数据、合规与事故

运营接管与迁移

暂停、回滚和替代能力

准备金同时可以是现金额度、未承诺容量、人工能力或可替代架构。

10.1 风险分级

等级	典型场景	主要风险	最低控制
L1 低	内部摘要、非敏感选题辅助	低影响错误、效率损失	抽样评测、基础日志
L2 中	对外内容草稿、内部客户分层建议	品牌偏差、误导、数据使用	发布前审核、权限、版本记录
L3 高	客户自动触达、广告预算建议、个性化推荐	客户损失、歧视、合规、财务影响	持续评测、双人审批、回滚、接管
L4 关键	自动执行大额投放、敏感数据决策、高监管触达	重大财务、法律、声誉和系统风险	严格授权、限额、实时监控、演练、独立审查

10.2 准备金不是平均事故率乘以损失

Knight 区分可计算风险与真正不确定性。AI 场景的尾部事件常缺乏稳定频率，准备金不能只用历史平均。建议由三部分构成：已知可估损失准备、压力情景缓冲、运营接管与迁移资金。准备金可以是现金额度、未承诺预算、备用供应商、人工接管容量或技术冗余。

10.3 Deming 的质量成本：控制投入可降低总损失

Deming 的质量管理思想提醒管理层，检验不能替代过程设计，但预防、评测和改进能够减少返工与外部失败。AI 营销的控制投资应优先放在数据、规则、质量门、可观测性和反馈闭环，而不是在流程末端堆叠审批。低风险场景采用轻量抽检，高风险场景使用持续评测、限额和人工接管。

10.4 案例：Air Canada 聊天机器人责任判决

主体与时间：加拿大不列颠哥伦比亚省民事解决法庭于 2024 年就 Air Canada 网站聊天机器人向客户提供错误票价政策信息作出裁决。业务情境：面向客户的自动服务。具体动作：公司使用聊天机器人回答票务问题。可核验结果：法庭认定航空公司对网站所提供信息负责，并判令赔偿。归因限制：案件金额有限，法律环境具有地区性。管理启示：供应商或自动化系统不能吸收企业对客户承诺的责任；对外自动触达应配置知识更新、引用、升级人工、日志和纠错机制，控制成本应进入场景 TCO。[S15]

第十一章 月观察、季调仓、年重置

竹势 AI 营销智库出品

MONTH 月观察 用量 · 质量 · 异常
QUARTER 季调仓 仓位 · 依赖 · 集中风险
YEAR 年重置 战略 · 合同 · 能力边界

11.1 三个节奏解决不同问题

节奏	关注对象	核心输出	不应做什么
月度成本观察	用量、异常、质量、采用、预算偏差	成本与价值异常清单、纠偏动作	频繁推翻长期架构
季度组合调仓	仓位、证据、依赖、集中风险	加仓/持有/重构/退出决议	只做项目进度汇报
年度预算重置	战略、能力边界、供应商、承诺容量	资金池、损失上限、长期合同和能力路线	把上一年预算机械延续

11.2 Beyond Budgeting 与双环学习

Beyond Budgeting 强调滚动预测、相对目标和分散决策，适合技术价格与需求快速变化的环境。AI 营销不能把全年用量、模型和场景锁死在一次预算中；但基础设施、人才和长期采购仍需要年度承诺。解决办法是“双层承诺”：年度确定总风险边界、共享能力和最低容量，季度根据证据调仓，月度通过异常管理调整运行。

Argyris 与 Schön 的双环学习进一步要求：复盘不仅问“执行是否符合计划”，还要问“原假设和评价规则是否正确”。当项目反复无法达到 ROI，可能不是执行差，而是单位定义、质量门、目标客户或流程边界错误。季度调仓会应允许修改假设，而非只要求团队继续优化同一方案。

11.3 案例：Google 对生成式搜索与广告的渐进式扩展

主体与时间：Google 自 2023 年起逐步扩展生成式搜索体验，并在 2024–2026 年持续调整 AI Overviews、广告呈现与 Gemini 相关产品。业务情境：搜索体验、广告收入和高计算成本之间的平衡。具体动作：分市场、分查询类型和分产品逐步上线，持续调整界面和商业化。可核验结果：Alphabet 财报披露搜索与广告业务持续增长，并持续增加技术基础设施投资，但没有公开单一 AI 搜索功能的完整单位经济性。归因限制：广告业务受宏观需求、产品组合和市场份额影响。管理启示：高流量核心业务应采用滚动扩展和容量管理，而不是一次性全量切换；年度基础设施承诺与季度产品调仓可以并存。

[S16][S17]

第十二章 30/90/180 天落地路线

竹势 AI 营销智库出品

0-30 建立可见性 账户 · 合同 · 成本卡 · 基线
31-90 形成组合决议 矩阵 · 依赖 · 证据门 · showback
91-180 接入治理系统 调仓 · 分摊 · 准备金 · 接管

阶段	关键交付	继续条件	调整/暂停/退出条件
0-30 天：建立可见性	九类账户、合同清单、前 5 个场景成本卡、基线、风险分级、资金池章程	成本可归属、单位可定义、owner 到位	无法获得基线或责任人，暂停采购扩张
31-90 天：形成组合决议	四维评分、依赖图、首轮证据门、showback、供应商退出清单	至少两个场景有真实流程证据，底座消费者明确	演示无法进入真实流程，重构范围或供应商
91-180 天：接入治理系统	季度调仓会、chargeback 试点、准备金、采购模板、绩效与运营接管	单位经济性稳定、质量达标、风险可控	采用低、成本失控、责任缺失则退出或降仓

12.1 前 30 天：先建立账户与基线

第一周完成合同、云、供应商、人力与外包盘点，并映射九类账户；第二周选择 3-5 个真实场景定义合格单位和基线；第三周估算五段 TCO 与风险等级；第四周成立三层资金池和初次组合评审。不要等待完美数据，先用区间和置信等级开始，但必须标出未知项及其决策影响。

12.2 90 天：让预算随证据释放

每个探索场景至少完成一次真实流程试点，包含数据、权限、异常、人工复核和业务结果。共享能力必须出现真实消费者，不接受只建平台不接场景。季度末形成第一版加仓、持有、重构、退出决议，并把未满足依赖转化为能力投资或范围收缩。

12.3 180 天：把组合治理接入日常经营

将场景成本卡接入财务 BP 月报，将供应商 TCO 与退出条款接入采购评审，将单位经济性和采用接入业务绩效，将风险分级与准备金接入风险委员会。成熟的标志不是报表更多，而是管理层能够在一个季度内停止低证据投入、把资金转向高价值场景，并保证退出过程不伤害客户、数据和运营。

完整案例组合：成功、受限与退出的共同教训

竹势 AI 营销智库出品

十个案例，一套资本纪律

企业自报、独立实验、财报、司法裁决和退出事件按各自证据等级使用；不把工时、平台收入或自动化率直接等同于增量 ROI。

案例	地区/行业	证据类型	主要启示
P&G 随机现场实验	国际/消费品	独立实证	先验证任务与质量，不能直接外推经营结果
腾讯营销服务	中国/互联网广告	上市公司披露	收入贡献与 AI 设备运营成本必须同看
招商银行企业级AI智能助手	中国/金融与高监管	企业2025年中期报告；结果由企业测量	共享能力投资需与采用、质量和工时兑现同时管理
百度智能营销与广告技术	中国/互联网广告	上市公司/平台披露	平台效果不能直接替代广告主 ROI
Klarna 客服自动化	欧洲/金融科技	公司披露与后续修正	自动化率不能替代质量与人工选择
Morgan Stanley 知识助手	美国/金融服务	企业与供应商披露	共享知识底座价值来自复用
IBM Watson for Oncology	国际/医疗	研究与业务退出	高战略价值不能替代本地证据
Air Canada 聊天机器人	加拿大/航空	司法裁决	企业仍对自动化承诺负责
Google 生成式搜索	美国/平台	财报与产品披露	核心业务采用滚动扩展与容量管理
Adobe-Figma 交易终止	国际/软件	公司与监管披露	退出条款和集中风险具有经济价值

经典思想与当代修正

竹势 AI 营销智库出品

思想	原始连接	对 AI 营销预算的修正	管理含义
Drucker 资源配置	管理者决定资源投向机会	资源还包括数据授权和注意力	预算围绕经营假设而非工具
Coase/Williamson	交易成本与资产专用性	外购和自建都可能锁定	比较全生命周期与退出成本
March 探索/利用	学习与稳定产出的张力	探索资金不能按规模化回报考核	三层资金池使用不同证据门
实物期权	小额投资保留未来选择	试点价值包含减少不确定性	分阶段承诺、预设损失上限
Activity-Based Costing	按活动和成本动因归集	模型、复核和数据按合格单位归集	建立场景级真实成本
Ostrom 共同资源	共享资源需要边界与规则	共享 AI 底座不能全免费或全平均收费	先 showback 后 chargeback
Deming 质量成本	预防和过程改进降低总损失	评测与监控是控制投资	按风险配置质量机制
Beyond Budgeting	滚动预测与分散决策	快速变化与长期承诺并存	月观察、季调仓、年重置

具名当代专家与机构视角

竹势 AI 营销智库出品

观点只有在能改变一项预算决议时才进入驾驶舱

每项观点均对应精确原始著作、论文、标准或监管文件，并明确不可外推的边界。

专家/机构	背景与主张	适用边界	本报告判断
Rita Gunther McGrath	发现驱动规划要求把关键假设、里程碑和可承受损失写入阶段性承诺。	适用于高不确定性增长项目，不替代生产阶段的现金流与控制要求。	探索资金按证据释放，先购买学习权而非一次性承诺全部预算。[S27]
Amy C. Edmondson	团队心理安全有助于报告错误、提出异议和开展学习行为；可学习失败必须伴随清晰边界与复盘。	心理安全不等于降低绩效标准，也不能为无控制的客户风险开脱。	探索项目应预先定义损失上限、异常上报和复盘责任。[S32]
Erik Brynjolfsson、Daniel Rock、Chad Syverson	通用技术的生产率收益常滞后于无形资产、流程和组织互补投入。	宏观生产率解释不能替代单个营销场景的基线和归因。	数据、流程、培训和组织改造必须进入TCO与投资论证。[S33]
Ajay Agrawal、Joshua Gans、Avi Goldfarb	AI降低预测成本后，判断、行动、数据与责任等互补要素可能成为新的成本瓶颈。	“预测机器”是经济分析框架，不保证任何具体应用产生回报。	模型费用下降时仍需核算人工判断、流程执行、数据和错误成本。[S34]
Marco Iansiti、Karim R. Lakhani	AI规模化依赖数字化运营核心、统一数据与可复用流程，而非孤立试点堆叠。	平台化并非越集中越好，强监管、地域和业务差异仍需边界。	共享能力资金必须绑定服务目录、消费者、架构标准和运营责任。[S35]
Daron Acemoglu	AI的经济收益取决于可被可靠改进的任务范围；高估自动化覆盖会夸大总体收益。	宏观估算不是企业预算基准，且技术能力会持续变化。	组合评审应以可验证任务贡献为基础，并保留对过度自动化的下行情景。[S36]
FinOps Foundation	FinOps强调工程、财务与业务协作，以成本透明、单位经济性、预测和持续优化管理技术支出。	其主要实践来自云与技术财务管理，迁移到营销需补充质量、品牌与客户结果。	模型、算力和共享平台应形成可分摊的使用量、单位成本和价值指标。[S4][S5]
美国国家标准与技术研究院（NIST）	AI RMF 1.0以治理、映射、测量、管理组织风险工作，要求风险管理贯穿生命周期。	框架不规定统一控制等级或预算比例。	评测、日志、人工复核和事件响应应作为可验证的控制成本，并按风险分级。[S18]
国家金融监督管理总局	《银行保险机构数据安全管理办法》要求治理、分类分级、技术保护、个人信息保护及风险监测处置。	适用于银行保险机构，不能机械外推到所有行业和低风险内部任务。	高监管底座属于不可选投入，但仍应按数据级别、场景和损失暴露配置准备金。[S8]
中国人民银行及全国金融标准化技术委员会	金融科技规划与金融数据生命周期标准要求把数据治理、安全防护和运维保障纳入持续能力建设。	行业规划和推荐性标准不直接给出单个项目回报或会计处理。	金融场景的共享数据能力、删除销毁、审计与运维应纳入五段TCO。[S9][S37]

管理者决策清单

竹势 AI 营销智库出品

董事会

- ☐ 是否把 AI 营销定义为资本配置组合，而非年度工具清单？
- ☐ 总损失上限、集中风险与退出能力是否明确？
- ☐ 战略必做项目是否仍有路径重构和阶段证据门？
- ☐ 共享能力是否有真实消费者、维护责任和退役条件？

CFO 与财务 BP

- ☐ 九类管理账户能否穿透会计科目与合同？
- ☐ 五段 TCO 是否覆盖人工复核、控制、变更和退出？
- ☐ 单位经济性是否使用合格单位和保守归因？
- ☐ 季度调仓能否实际释放或转移预算？

CMO 与业务 owner

- ☐ 场景是否连接到经营结果和真实流程？
- ☐ 质量门、采用责任和运营接管是否明确？
- ☐ 释放工时是否能转为现金、产能或增长？
- ☐ 品牌、知识和客户资产是否有可治理的沉淀指标？

CIO/CDO 与平台负责人

- ☐ 共享数据、连接器、权限、评测是否按复用价值排序？
- ☐ 是否识别模型、供应商、关键人和数据源集中风险？
- ☐ 可观测性、容量、回滚和替换路径是否进入预算？
- ☐ 平台服务目录和成本动因是否可解释？

采购与风险

- ☐ 供应商比较是否固定同一工作负载和质量门？
- ☐ 最低消费、数据权利、价格调整和自动续约是否清楚？
- ☐ 退出支持、数据导出和并行迁移是否写入合同？
- ☐ 风险等级是否决定控制强度，而非一刀切？

附录 A：场景级成本卡模板

竹势 AI 营销智库出品

栏目	填写内容
基本信息	场景、业务 owner、运营 owner、资金池、仓位、评审日期
合格单位	单位定义、质量门、业务量、基线周期
成本	九类账户×五段 TCO，固定/变动/共享/专属
价值	成本节约、质量、增量结果、资产、学习、风险
证据	来源、测量者、样本、时间、置信等级、限制
依赖	数据、接口、权限、人员、供应商、控制
决议	加仓/持有/重构/退出，预算动作与责任人
退出	数据导出、工作流重建、客户连续性、知识转移

附录 B：组合仓位卡模板

竹势 AI 营销智库出品

字段	说明
经营贡献	收入、毛利、现金、客户价值或风险影响
证据成熟度	演示、任务、真实流程、稳定生产、复用
复用性	消费者数量、边际成本、标准化和维护机制
风险	数据、合规、品牌、财务、锁定、集中度
依赖与相关性	共享底座、共同故障点、跨场景阻塞
当前仓位与目标仓位	种子、验证、核心、底座、观察、退出
下一证据门	指标、样本、时间窗、决议日期
资金动作	新增、维持、冻结、转移、释放、准备金

附录 C：来源索引

竹势 AI 营销智库出品

[S1] Dell'Acqua, Fabrizio 等, The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise, Harvard Business School Working Paper, 2024; 对象为 791 名 P&G 专业人员及真实创新任务。

[S2] 腾讯控股, 2025 年年度业绩公告与年度报告, 2026-03/04; 营销服务收入、AI 广告推荐与业务表现披露。

[S3] 腾讯控股, 2026 年第一季度业绩公告, 2026-05-13; 营销服务收入、AI 推荐模型、设备折旧与运营成本披露。

[S4] FinOps Foundation, State of FinOps 2025 Report, 2025; 样本为 FinOps 社区调查, 管理 AI/ML 支出和单位经济性为上升重点。

[S5] FinOps Foundation, Cost Estimation of AI Workloads, 2025-09-15; 覆盖开发、试点、生产阶段的 AI 成本估算。

[S6] Klarna, AI Assistant Handles Two-Thirds of Customer Service Chats, 2024-02; 结果由公司测量。

[S7] Klarna 管理层关于恢复更多人工客户服务选择的公开访谈与公司沟通, 2025; 用于说明质量与成本权衡。

[S8] 国家金融监督管理总局, 《银行保险机构数据安全管理办法》, 2024-12-27发布, 2025-03-01施行, 部门规章/监管规范; 本报告仅用于银行保险机构数据治理、分类分级、技术保护、个人信息保护与风险处置的不可选控制边界, 不作为一般行业预算比例依据。

[S9] 中国人民银行, 《金融科技发展规划 (2022—2025年)》, 2021-12-31印发、2022-01-04公布, 行业发展规划; 本报告用于说明金融机构需把数据、技术、治理与人才作为持续能力建设, 不将规划目标解释为单一AI项目回报。

[S10] Morgan Stanley Wealth Management 与 OpenAI, AI @ Morgan Stanley Assistant 公开案例, 2023-2024; 采用与效率由企业和供应商披露, 未披露完整预算。

[S11] IBM Watson Health 与 Watson for Oncology 的企业公告、合作资料及后续资产处置披露。

[S12] 同行评审研究与医学机构对 Watson for Oncology 一致性、适用性和本地化限制的评估。

[S13] Amazon Web Services, Guidance for Cloud Financial Management on AWS; showback、chargeback 与成本归属方法。

[S14] Adobe, Termination of Figma Merger Agreement, 2023-12; 交易终止与约 10 亿美元终止费披露。

[S15] Moffatt v. Air Canada, British Columbia Civil Resolution Tribunal, 2024; 聊天机器人错误信息责任裁决。

- [S16] Alphabet, 2024–2025 年年度报告及季度业绩；搜索、广告与技术基础设施投资披露。
- [S17] Google, AI Overviews、Search Generative Experience 与广告产品的正式产品公告, 2023–2026。
- [S18] NIST, Artificial Intelligence Risk Management Framework 1.0, 2023；治理、映射、测量和管理框架。
- [S19] Drucker, Peter F., The Practice of Management, 1954；资源配置与管理责任。
- [S20] Coase, Ronald H., The Nature of the Firm, 1937；Williamson, Oliver E., Markets and Hierarchies, 1975；交易成本与资产专用性。
- [S21] March, James G., Exploration and Exploitation in Organizational Learning, Organization Science, 1991。
- [S22] Kaplan, Robert S. 与 Cooper, Robin, Cost & Effect, 1998；作业成本与成本动因。
- [S23] Ostrom, Elinor, Governing the Commons, 1990；共同资源治理。
- [S24] Deming, W. Edwards, Out of the Crisis, 1986；质量管理与过程改进。
- [S25] Hope, Jeremy 与 Fraser, Robin, Beyond Budgeting, 2003；滚动管理与适应性预算。
- [S26] Knight, Frank H., Risk, Uncertainty and Profit, 1921；风险与不确定性。
- [S27] McGrath, Rita Gunther；MacMillan, Ian C., Discovery–Driven Growth: A Breakthrough Process to Reduce Risk and Seize Opportunity, Harvard Business Press, 2009, 管理著作；本报告仅借鉴假设显性化、阶段承诺和里程碑方法, 不以其替代生产项目的财务纪律。
- [S28] Brynjolfsson, Erik；Li, Danielle；Raymond, Lindsey R., Generative AI at Work, Quarterly Journal of Economics, 2025（研究前身为NBER Working Paper 31161, 2023）, 现场研究；对象为客户支持人员, 本报告仅用于说明任务、经验与组织条件影响生产率, 不外推为所有营销岗位的统一收益。
- [S29] 百度集团, 2025 年年度及季度业绩材料；智能云、生成式 AI 与在线营销相关披露, 平台归因不等于广告主独立 ROI。
- [S30] European Union, Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 2024–06–13通过、2024–07–12公布, 法规；国家互联网信息办公室等七部门,《生成式人工智能服务管理暂行办法》, 2023–07–13公布、2023–08–15施行, 部门规章；本报告仅用于说明地区、服务对象与风险等级会改变合规成本, 不将其作为所有内部应用的统一控制模板。
- [S31] 招商银行股份有限公司,《2025年中期报告》, 2025–08–29, 上市公司中期报告；披露在零售金融、公司金融、风险管理、运营管理等领域上线AI智能助手并累计节省475万小时。结果由企业自行测量, 未披露完整预算、对照组、质量变化、客户增量或现金兑现比例。
- [S32] Edmondson, Amy C., Psychological Safety and Learning Behavior in Work Teams, Administrative Science Quarterly, 1999, 同行评审论文；本报告用于说明错误上报、异议与团队学习的组织条件, 不把心理安全解释为降低质量或风险标准。

[S33] Brynjolfsson, Erik; Rock, Daniel; Syverson, Chad, Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics, NBER Working Paper 24001, 2017, 学术论文；本报告用于说明通用技术收益依赖无形资产与组织互补投入，不用于估算单一企业ROI。

[S34] Agrawal, Ajay; Gans, Joshua; Goldfarb, Avi, Prediction Machines: The Simple Economics of Artificial Intelligence, Harvard Business Review Press, 2018, 经济管理著作；本报告使用“预测成本下降会提高判断、行动和数据等互补要素价值”的分析框架，不把模型能力等同于自动经营结果。

[S35] Iansiti, Marco; Lakhani, Karim R., Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World, Harvard Business Review Press, 2020, 管理著作；本报告用于共享数字核心、运营模式和规模化边界，不主张所有能力无条件集中。

[S36] Acemoglu, Daron, The Simple Macroeconomics of AI, Economic Policy, 2025（早期版本发表于2024），学术论文；本报告仅用于提醒AI收益受可靠改进任务范围约束，不将宏观估算作为企业预算基准。

[S37] 全国金融标准化技术委员会，《金融数据安全 数据生命周期安全规范》（JR/T 0223—2021），2021-04-08发布并实施，推荐性金融行业标准，中国人民银行主管；本报告用于金融数据采集、传输、存储、使用、删除、销毁及运维保障的生命周期成本边界。

附录 D：九类账户的深度控制规则

竹势 AI 营销智库出品

D.1 软件账户：许可并不等于能力被采用

软件账户需要同时管理“购买权利”和“实际使用”。席位制产品常见的失真是：部门为了获得折扣一次性购买大量许可，后续仅少数员工活跃；平台制产品则可能基础价低，但高级权限、审计、连接器和测试环境另行收费。建议把软件账户拆成基础许可、增量模块、环境、支持与闲置五个子项。月度观察不只看登录人数，还要看关键任务采用率、产出进入真实流程的比例和功能重叠。对重复能力，应设置整合期限，而不是因为合同尚未到期继续分散运行。

采购折扣必须与弹性共同评价。多年预付可降低名义单价，却会减少技术替换和需求收缩时的选择权。只有当工作负载稳定、质量经过生产验证、退出路径明确且最低容量确实需要时，才适合长期承诺。探索期宜使用可取消或小容量合同，即使单价较高，其期权价值也可能更大。

D.2 模型账户：把重试、代理循环和质量失败纳入成本

模型账户不能仅按输入输出量归集。一次业务任务可能触发检索、规划、多轮调用、工具执行、错误恢复和最终复核；表面单次价格很低，完整任务成本却由循环次数与失败率决定。应记录每合格任务的调用链长度、缓存命中、模型路由、重试原因、人工修正和失败退出。对高频任务，优先优化上下文、检索质量、结构化输出和停止条件，而不是简单换成更便宜的模型。

模型质量也具有非线性经济后果。对于内部低风险摘要，轻微质量差异可能只增加少量复核；对于客户报价、合规内容或广告预算建议，少数严重错误可能抵消大量日常节约。因此模型选择应通过任务级评测集、严重缺陷率和压力情景，而不是只比较平均分。

D.3 算力账户：容量承诺要和需求驱动因素绑定

自托管或专用算力可带来数据控制、性能和稳定性，但也引入利用率、折旧、运维和升级风险。容量计划应由实际需求驱动因素建立，例如每日合格任务量、峰值并发、上下文规模、响应时限和可接受排队时间。不能以“未来都要用 AI”为由提前购买难以利用的容量。对于波动较大的营销战役，弹性资源可能更经济；对于稳定高频且有严格数据要求的任务，承诺容量可能更合适。

D.4 数据账户：低质量数据会在每次调用中重复收费

数据成本具有前置与持续两种形态。前置包括清理、映射、标注、版权和授权；持续成本包括更新、质量监控、删除请求、保留、版本和主数据维护。企业常把一次性数据项目视为完成，忽略业务规则与产品信息持续变化。应为每个数据域设立 owner、质量 SLA、更新频率和退役规则。知识库中大量过期、冲突或不可授权内容，会通过错误检索、重试和人工核查反复放大成本。

D.5 内容账户：产量不是资产，经过治理的复用才是资产

内容账户既包括原创与素材采购，也包括版权、翻译、剪辑、渠道适配、审校和下架。生成速度提高后，企业可能产出更多低价值内容并增加审核负担。内容资产只有满足可检索、权利清楚、版本明确、效果可追踪和可再次调用，才具有资产属性。建议用“合格发布资产成本、跨渠道复用率、有效生命周期、返工率、严重品牌缺陷率”共同评价。

D.6 集成账户：连接成功只是开始，异常处理决定运行成本

连接器开发的初始费用通常易见，长期成本来自认证变化、字段变化、接口限流、错误重放、数据一致性和第三方平台政策。每个接口必须有服务 owner、监控、故障降级、测试环境和变更预算。若供应商声称“已连接”，采购应验证真实读写范围、权限粒度、失败状态、回滚和审计，而不是只看演示。

D.7 培训账户：从课程人数转向关键任务胜任度

培训不应以参加人数、课时和满意度为主要产出。真正目标是让角色在真实任务中达到质量、时效和风险标准。预算应覆盖管理者如何设定任务、员工如何复核和接管、运营团队如何处理异常、财务如何看单位成本。新流程上线后需要现场辅导和绩效规则调整，否则课程结束即回到旧流程。

D.8 运维账户：没有运营接管，试点只是临时演示

运维包括监控、事件、知识更新、容量、版本、支持和人工接管。试点团队往往由最熟悉技术的人临时值守，规模化后这些隐性劳动无法持续。进入核心仓前必须指定运营 owner、服务时段、事件等级、恢复目标、人工接管容量和变更窗口。若没有持续预算和人员，项目应停留在验证仓，不能以“上线”名义进入生产。

D.9 顾问账户：以能力转移和可独立运营作为验收

顾问能够加快诊断、架构、实施和治理，但也可能形成长期依赖。合同应区分一次性设计、实施交付、代运营和持续优化，避免低价项目通过后续人工服务回收。验收除文档外，应包括内部团队能否独立运行、关键配置和评测是否可访问、问题处理是否有知识转移、供应商退出后是否可以维持最低服务。

附录 E：组合相关性、集中风险与共享底座的计算方法

竹势 AI 营销智库出品

E.1 项目收益不是独立随机变量

经典投资组合理论假设可以用收益、波动和相关性描述资产，但 AI 营销项目存在三项重要差异。第一，收益往往由同一底座驱动，项目之间高度相关；第二，损失分布可能有厚尾，例如错误自动触达在短时间内扩大；第三，收益可被组织采用和流程变更共同决定，历史序列很短。因而本报告借用“分散、相关和仓位”的思想，却不机械使用均值方差优化。

组合评审至少建立四张相关性表：共同供应商、共同模型、共同数据源、共同关键人员。任何单点同时支撑多个核心仓，都应纳入集中风险。缓解方式包括替代模型、标准接口、离线降级、人员备份、数据快照、容量冗余和合同保障。并非所有集中都应消除；集中可降低集成与治理成本，但必须与故障影响和替换时间匹配。

E.2 共享底座的“解锁价值”

底座价值不能简单等于消费者项目收益之和，否则容易重复计算。建议分三层报告：直接服务价值，指底座本身减少的重复建设和运维成本；解锁价值，指因底座存在才可能进入验证或核心仓的场景贡献；选择权价值，指未来新增场景的边际进入成本下降。财务决议以直接价值和已承诺解锁价值为主，选择权价值作为战略补充，不强行货币化。

底座能力	直接价值	解锁价值	主要风险
统一身份权限	减少重复接入和审计	支持跨部门受控使用	权限设计错误形成系统性暴露
企业知识与检索	减少重复查找和内容整理	支持内容、客服、销售多个场景	过期和冲突知识扩大错误
评测与日志平台	减少各项目自建控制	允许高风险场景进入生产	指标失真或日志不完整
模型网关与路由	集中合同、缓存和成本控制	快速替换模型并扩展任务	形成新的平台单点
连接器目录	减少重复接口开发	加快端到端闭环	第三方变化带来连锁故障

E.3 示例组合：一家多品牌消费企业

假设企业同时推进商品详情页生成、广告创意变体、会员触达、舆情摘要和活动复盘。商品详情页与广告创意共享品牌知识、素材权利和评测；会员触达依赖客户数据、同意管理和人工接管；舆情摘要风险较低，可作为早期验证场景；活动复盘需要渠道数据连接。若只按单个项目 ROI 排序，企业可能先扩张广告创意，却没有品牌评测和素材权利底座。更稳健的组合是：用舆情摘要验证运营流程；把品牌知识与评测列为底座仓；广告创意进入验证仓；会员触达在同意与接管能力完成前保持观察仓；活动复盘与渠道连接器共同立项。

附录 F：四类决议门的会议操作手册

竹势 AI 营销智库出品

F.1 加仓会议

加仓不是对团队努力的奖励，而是对证据的资本回应。材料必须包括基线、合格单位、至少一个完整业务周期的单位经济性、质量尾部、采用、风险事件、容量和运营接管。会议先检查是否存在选择性样本或把人工支持排除；再判断新增规模是否改变成本结构，例如进入更高价格阶梯、需要全天候运维或提高监管等级。加仓决议必须写明新增范围、预算上限、下一证据门和可逆性。

F.2 持有会议

持有适用于价值假设仍合理，但关键证据尚未到达的情形。持有必须有明确等待事件，例如完成一个销售周期、获得接口授权、监管规则落地或供应商发布替代能力。没有事件和日期的持有等于拖延。持有期间只保留获取下一证据所需的最小能力，停止功能扩张和长期合同。

F.3 重构会议

重构首先定位失败机制：任务定义错误、数据不足、质量不稳定、流程未改变、员工不采用、供应商不匹配、控制成本过高或责任不清。每次只重构少数关键变量，保持可比较性。若价值假设本身已经失效，不应以重构名义继续消耗；若战略目标必须实现，则应比较替代路径，而不是默认沿用现有技术。

F.4 退出会议

退出应在立项时设计，在决议时执行。退出清单包括停止新增数据和新用户、冻结配置、导出数据与评测集、迁移 workflows、通知相关角色、安排客户连续性、终止合同、释放容量、保留审计记录和完成知识转移。退出后的复盘重点是更新组合规则和供应商条款，而不是寻找责任替罪羊。

附录 G：供应商比较与退出演练模板

竹势 AI 营销智库出品

评估域	必须提供的证据	红旗
价格与容量	完整价格阶梯、最低消费、超额、折扣条件	只给低用量样例或拒绝压力情景
质量	企业任务评测、严重缺陷、漂移和版本变化	只展示通用基准或精选案例
人工依赖	供应商与客户侧人工工时、接管率	把代运营隐藏为产品能力
数据权利	处理目的、保留、训练、导出、删除	权利模糊或导出不可验证
集成与标准	接口、日志、身份、配置和工作流可移植性	关键流程只能在专有界面维护
服务连续性	SLA、故障通报、灾备、降级与支持	无明确恢复目标
合同弹性	期限、续约、涨价、缩容、终止和所有权变化	自动续约、单方涨价、退出费不透明
退出演练	在限定时间导出并由替代方案读取	只有承诺，没有实际演练

建议至少每年对核心供应商执行一次桌面退出演练，对关键场景执行一次技术演练。桌面演练验证合同、联系人、资产清单和责任；技术演练选择一小段真实数据和工作流，验证导出、重建、质量和切换时间。退出能力并不表示企业准备频繁更换，而是形成议价权、降低集中风险并防止关键知识流失。

附录 H：董事会情景推演

竹势 AI 营销智库出品

情景一：模型单价下降 60%，总成本却上升

原因可能是业务量增加、上下文变长、代理循环增多、更多任务进入生产，或低价诱发无价值消费。管理层不应要求简单压低单价，而应检查每合格单位模型成本、任务价值密度和用量预算。若单位成本下降且经营贡献增加，总支出上升可能合理；若任务质量和价值没有增长，则需要限额、路由和停止低价值任务。

情景二：试点节省大量工时，但财务结果未变化

释放工时可能没有转化为减少外包、避免招聘、增加有效产能或加快收入。业务 owner 需要在加仓前提交工时兑现计划：取消哪些工作、减少何种采购、把人员转向哪些高价值活动、如何测量新增产出。无法兑现的时间仍可作为员工体验或韧性价值，但不能全部计作现金回报。

情景三：共享平台投入超支，业务采用不足

先区分平台能力是否被真实场景需要、服务体验是否阻碍采用、业务是否承担采用责任。若消费者承诺从未存在，应冻结扩张并将平台降为最小底座；若采用受成本归属或接入复杂度影响，先简化服务目录和 showback；若业务选择外部工具，应比较重复建设与例外价值。平台不能用未来想象继续获得预算。

情景四：高价值自动触达场景发生严重错误

立即执行限额、暂停自动执行和人工接管，同时保留日志、识别受影响客户和完成纠正。季度决议不应只问是否退出，而要判断错误来自知识、权限、流程、模型还是监控。若控制可重构且业务价值仍高，可降仓后重新验证；若责任、数据或合规无法控制，应退出。准备金应覆盖客户补救、人工接管和迁移。

情景五：供应商提出更低单价换取三年承诺

董事会应比较折扣现值与失去的选择权。需要压力测试需求下降、技术替换、质量变化、价格阶梯、最低消费和退出费。若工作负载已稳定、替代成本低、合同允许缩容和数据可移植，长期承诺可能合理；若仍处于验证仓，折扣通常不足以补偿锁定风险。

附录 I：管理账户与法定会计、绩效体系的接口

竹势 AI 营销智库出品

I.1 管理账户不替代会计判断

九类账户服务经营决策，并不直接决定费用化或资本化。相同的集成、数据整理或软件配置，在不同合同、控制权、可辨认资产和会计政策下可能有不同处理。财务团队应保留从管理账户到总账科目、成本中心、项目号和合同的映射，但不得为了会计便利牺牲经营可见性。董事会需要看到的是全生命周期承诺、现金流时间、可取消性和或有负债；会计确认则由适用准则和专业判断决定。

尤其应区分四类概念：费用预算反映当期资源消耗；资本性承诺反映跨期、难取消或形成长期能力的合同与投入；共享能力投资反映可被多个场景消费的底座；或有风险反映事故、迁移、最低消费和合同变化可能造成的额外现金流。四者可能交叉，但在决议材料中必须分别展示。

I.2 与绩效管理连接，避免局部最优

若业务团队只对收入负责，可能无节制消费共享资源；若平台团队只对单位技术成本负责，可能通过降低质量或限制使用实现表面优化；若采购只对折扣负责，可能增加长期锁定。绩效应采用共同结果与角色责任相结合的方式：业务 owner 对采用、质量和经营结果负责；平台 owner 对服务成本、稳定性、复用和替换能力负责；财务 BP 对归属、预测和调仓执行负责；采购对合同弹性、数据权利和退出条款负责；风险团队对控制设计与剩余风险透明负责。

共同指标不宜过多。建议组合层只保留六项：经质量调整的单位成本、保守归因的经营贡献、关键任务采用率、严重缺陷率、共享能力复用率、退出准备度。各团队可在下层增加诊断指标，但不能用大量活动指标遮蔽这六项结果。

I.3 预算争议的处理顺序

发生分摊争议时，先判断成本是否可控：业务无法控制的统一安全要求不应完全按使用量收费；平台无法控制的业务内容返工不应由中央吸收。其次判断是否可计量：计量不稳定时先 showback，不急于内部结算。再次判断是否可替代：业务有合理替代方案时，可通过架构例外和全成本比较，而不是行政禁止。最后判断行为影响：任何分摊规则都要观察其是否诱发隐藏账户、规避平台、过度消费或抑制合理探索。

I.4 决议记录的最低标准

每次组合决议应形成一页记录：决议对象、当前仓位、关键证据、不同意见、预算动作、损失上限、责任人、下一证据门和复审日期。记录的目的不是增加文书，而是防止市场环境变化后无法还原当时假设。退出项目同样应记录释放的资金、保留的资产、未解决风险和对组合规则的修正，使失败转化为可复用学习。

实施时应坚持一个底线：任何新增预算都必须能追溯到场景、证据门、责任人和退出安排；任何已承诺预算也都应允许在证据变化后被重构、转移或终止。

PUBLISHER

关于竹势 AI 营销智库

竹势 AI 营销智库出品

竹势 AI 营销智库关注企业增长、营销组织变革、智能体运营与 AI 治理，帮助企业老板、CEO、CMO 和数字化管理者把复杂技术议题转化为可讨论、可比较、可执行的经营决策。

本报告的判断框架、责任边界和实施工具，旨在帮助管理团队把 AI 营销支出转化为可比较的经营承诺，并在加仓前建立必要的证据、单位经济性、风险边界与退出责任。