



AI MARKETING / FIRST-CLOSURE FIELD
GUIDE

AI营销场景优先级方法

如何选出第一个高价值闭环

不从工具清单寻找“最佳用例”：先用闭环完整性淘汰炫技，再让价值、风险、依赖与证据共同决定唯一首发场景。

生成候选池

从客户任务与经营瓶颈进入

穿过硬闸门

风险、数据与证据周期拥有否决权

识别真闭环

信号、判断、行动、验收、反馈

冻结首发场景

用 14 / 30 / 90 天做资本决议

出版说明

竹势 AI 营销智库出品

本报告面向企业董事会、经营层与首批 AI 试点负责人，提供一套可审议、可校准、可暂停的场景优先级方法。报告不提供行业平均分、虚构基准值或预设“最佳场景”榜单。评分用于暴露判断差异，否决门用于守住不可抵消的风险边界，依赖分析用于纠正单场景视角。

正文所列公司结果按公开披露口径呈现。公司自报、平台归因和供应商测量均明确标注，不能替代企业自身对照试验。所有第三方来源在正文以编号就近标注，并在文末以纯文本列示；文档不设置第三方可点击外链。

品牌首页：<https://shichangbu.ai/>

执行摘要

核心判断

第一章 首场景为何决定后续价值曲线

第二章 候选池：从客户任务、经营瓶颈与流程断点出发

第三章 五环闭环：信号—判断—行动—验收—反馈

第四章 八维评分：让价值、可行与证据进入同一张桌子

第五章 风险否决门：高分不得抵消红线

第六章 数据就绪：数据存在不等于闭环可用

第七章 证据设计：工时、现金与增量利润的边界

第八章 依赖图与场景组合排序

第九章 中国经营语境下的特殊约束

第十章 首个试点 Charter 与 14/30/90 天闸门

第十一章 复制、暂停与淘汰

第十二章 董事会治理与实施建议

管理者工具箱

附录与来源索引

BOARD BRIEF 02

执行摘要

竹势 AI 营销智库出品

FIRST-CLOSURE RULE / 首场景决议原则

第一个场景的任务不是展示 AI 有多聪明，而是以受控损失换取一条足以改变下一次资本决议的经营证据链。

从经营瓶颈建池

先闭环再评分

风险拥有否决权

证据决定复制

首个闭环应被视为一项经营证据工程，而非技术展示工程。

企业面对的真实难题并非缺少 AI 创意，而是候选场景过多、判断口径不一、技术演示与经营结果之间缺乏可追溯链路。2025 年麦肯锡全球调查显示，78% 的受访组织已在至少一个业务职能使用 AI，但从试点走向规模化影响仍是多数组织的未完成任务；营销与销售又是最常见的应用领域之一。[S1] 采用率上升说明供给与关注度上升，并不自动证明现金回报、组织吸收能力或风险控制已经成立。

本报告提出“场景筛选流水线 / 决策漏斗”：候选池生成、闭环资格审查、八维可解释评分、风险否决、依赖与组合重排、单一首发场景 Charter、14/30/90 天证据闸门、复制资格审查。该方法刻意把评分、否决和依赖分开。评分回答“相对值得做吗”，否决门回答“现在允许做吗”，依赖分析回答“即使值得做，是否必须先做别的”。三者混为一个总分，往往导致高价值想象掩盖生产风险，或导致容易上线的小工具挤掉真正能形成经营证据的闭环。

真正闭环必须具备五个环节：信号、判断、行动、验收、反馈。只生成文案、图片、建议或报表的项目，若没有进入真实业务动作、没有预先定义验收、没有把结果回写下一轮判断，最多属于能力验证，不应按经营闭环立项。端到端并不等于全自动。首发场景通常应保留关键人工复核，并把授权半径限制在可逆、可追踪、可暂停的范围。

任务级实验能够估计局部生产率和质量变化，却不能直接推出企业利润。Brynjolfsson、Li 与 Raymond 对 5,179 名客服人员的研究发现，生成式 AI 助手使每小时解决问题数平均提高约14%，低经验人员收益更大；该结果基于特定客服系统、工作流程和知识库，不能等同于减少编制、增加收入或形成营销现金回报。[S2] 首场景证据必须沿“活动量—流程表现—客户行为—财务结果”逐层建立，并明确哪些是领先指标、哪些是结果指标、哪些只能用于学习。

中国企业还需把平台接口、账号权限、个人信息、数据出境、广告与行业规则、集团审批、生成合成内容标识等约束前置。2025 年发布的《人工智能生成合成内容标识办法》明确显式与隐式标识要求；《生成式人工智能服务管理暂行办法》要求履行信息内容安全与个人信息保护义务。[S3] [S4] 因此，生成内容能否被合规发布、数据能否在当前目的下使用、外部动作能否留痕，直接决定首发资格。

约 10 个可证伪的核心判断

编号	核心判断	可证伪条件
1	采用率与价值实现不存在自动等号	若业务流程未重构、数据未接入、责任未改变，更多员工使用 AI 只会增加局部产出，未必改善收入、利润或现金。
2	第一个场景应优化“可信证据速度”	上线最快的项目可能长期无法观察客户行为或财务影响；应比较达到有效证据所需的完整周期。
3	闭环完整性优先于单点模型能力	模型再强，若输出无法触发动作、结果无法验收、反馈无法回流，经营价值仍不可复核。
4	工时节省不是现金回报	节省时间只有在转化为减少外采、避免新增编制、增加有效产能或提升收入时，才进入财务结果。
5	数据存在不等于数据可用	可用性包含合法目的、字段语义、时效、覆盖、主键、质量、访问权限和可回写能力。
6	授权半径与可逆性决定首发资格	高影响、不可逆、面向客户或资金的动作应更晚授权，并配置人工复核、额度和回滚。
7	单项高分必须接受依赖重排	高价值场景若依赖缺失的身份解析、事件采集或 CRM 状态，排序应后移或拆成能力建设。
8	首发只能有一个经营主闭环	多个并行“第一场景”会争夺数据、业务专家和管理注意力，降低对照质量。
9	复制资格高于试点成功	结果好但机制不可解释、依赖不可复用、风险靠明星团队人工兜底，不应直接扩张。
10	失败也应形成可购买的知识	明确暂停与退出条件，可把失败转化为数据、流程和控制能力资产；无退出设计的试点才是昂贵失败。

第一章 首场景为何决定后续价值曲线

竹势 AI 营销智库出品

01 经营瓶颈 客户与流程任务
02 闭环完整 信号到反馈
03 价值可证 口径与基线
04 风险可控 可逆与可停
05 首发 Charter 只放行一个
评分不能让缺少闭环、责任或回滚能力的场景穿过硬闸门。

管理层需要改变的第一项决策：停止按“最炫功能”或“最容易上线”选择首场景，改按最短可信证据链选择。

AI 项目早期常被两个偏差主导。其一是可见性偏差：图像、文案、数字人和对话演示容易让高管迅速感知能力，却未必触达经营瓶颈。其二是便利性偏差：数据容易取得、接口容易调用的场景被优先推进，即使该流程并非约束点。Goldratt 在《目标》中提出约束理论：系统产出受最关键约束决定，局部效率优化可能只增加在制品。把该思想用于 AI 营销，首场景应靠近限制增长或现金转换的瓶颈，而非选择最容易被自动化的步骤。

Drucker 对管理目标的强调也提供了边界：组织结果发生在客户与市场一侧，内部活动本身不是结果。AI 生成更多内容、完成更多分析，只能算活动。只有当这些活动改变了线索质量、响应速度、转化、留存、价格实现或营销成本结构，才开始接近经营结果。管理层应在项目批准前写清“若场景有效，哪个经营指标会以何种机制改变；若该指标没有变化，何时停止”。

案例 1 | P&G 与哈佛商学院“Cybernetic Teammate”现场实验（2024—2025）

主体与时间：哈佛商学院、沃顿等研究者与宝洁开展大规模现场实验，论文于2025年发布，参与者791人。业务情境：专业人员完成真实感较强的产品创新任务，比较个人、团队以及是否使用生成式 AI。具体动作：参与者使用 AI 支持构思与方案形成。作用机制：AI 扩展知识搜索与跨职能表达，使个人可获得类似团队的部分能力。公开结果：研究报告称，使用 AI 的参与者更可能提出进入前10%的方案，并改善跨职能知识整合；测量者为研究团队。[S5] 边界：实验针对限定任务与评分，不证明产品上市成功、消费者购买或利润增长。可复制部分是任务定义、盲评、对照和跨职能质量指标，而非“AI 可替代团队”的泛化结论。

所以，首场景不应承担“证明 AI 普遍有效”的宏大任务，而应承担更小、更严格的命题：在特定客户任务、特定流程和明确控制下，AI 是否使某个可观察结果改善，且风险在容忍范围内。

第二章 候选池：从客户任务、经营瓶颈与流程断点出发

竹势 AI 营销智库出品

客户任务 触发、阻力、未完成结果
流程瓶颈 等待、返工、判断与交接
经营暴露 收入、毛利、质量、风险
责任入口 owner、对象、动作、验收
候选从真实任务进入，不从模型功能或供应商清单进入。

候选池的输入应是待完成的客户任务、经营约束与流程失效，不是供应商功能菜单。

Christensen 的待办任务理论提醒管理者：客户“雇用”产品或服务完成进步，场景定义应从客户要完成什么开始。对营销而言，候选场景不应写成“用 AI 写公众号”，而应写成“将高意向客户常见异议转化为可检索内容，并在销售接触前降低解释成本”。前者是工具动作，后者包含客户任务、业务机制和可验收结果。

候选池建议通过三条路径生成：客户任务路径识别客户决策中的摩擦；经营瓶颈路径识别限制收入、毛利、现金或速度的约束；流程断点路径识别信息丢失、等待、返工、无人负责和反馈不回流。三条路径交叉出现的场景优先进入资格审查。

来源	提问	合格候选表述示例	常见误区
客户任务	客户在什么情境下需要完成何种进步？	新产品采购委员会需在两周内完成技术、财务和风险共识	把“写白皮书”当场景
经营瓶颈	哪个约束限制收入、毛利、现金或响应？	高意向线索48小时内未得到行业化回应，导致机会流失	优先自动化低价值重复劳动
流程断点	信息在哪一环丢失、等待、返工或失真？	广告创意表现未回写内容选题与下一轮素材	把多生成素材视为闭环

候选池需要控制粒度。过粗的“AI 社媒运营”无法评分，过细的“自动改一个标题”无法形成经营证据。合适粒度通常覆盖一个可独立验收的业务回路，拥有清晰起点、终点、责任人和反馈。

第三章 五环闭环：信号—判断—行动—验收—反馈

竹势 AI 营销智库出品



闭环资格审查先于评分；缺一环的候选不进入优先级排名。

Deming 的 PDSA 循环强调计划、执行、研究与调整。用于 AI 营销时，需要进一步拆解为五环：信号是触发所需的信息；判断是把信号转化为下一步决策；行动是真实改变客户、渠道、内容、预算或任务状态；验收是预先定义结果是否达到标准；反馈是把结果写回规则、数据、知识或下一轮动作。闭环不是“流程步骤很多”，而是输出能改变后续输入。

环节	必须回答	最低证据	淘汰信号
信号	由什么事件、状态或阈值触发？	时间戳、来源、对象ID、字段定义	仅靠临时人工口头发起
判断	依据何种规则形成建议或决策？	规则版本、输入引用、置信与例外	无法解释为何做出该选择
行动	谁或什么系统执行真实动作？	任务、发布、触达、更新或审批记录	输出停留在文档或聊天框
验收	成功、失败、异常怎样判定？	阈值、对照、抽样复核、结果字段	只凭主观满意度
反馈	结果如何影响下一轮？	回写路径、复盘节奏、规则更新记录	每轮重新开始、经验不沉淀

案例 2 | 腾讯广告技术与微信生态闭环（2024—2025）

主体与时间：腾讯，2024及2025财年。业务情境：营销服务需要在内容消费、广告定向、出价、投放和交易反馈之间提高效率。具体动作：腾讯披露持续增强 AI 广告定向，并扩展微信生态内闭环营销能力。作用机制：平台内行为与交易信号进入定向、竞价和投放优化，再由广告表现与交易结果反馈。公开结果：腾讯2025年营销服务收入同比增长19%至人民币1450亿元；公司称 AI 定向增强与微信闭环能力是增长驱动因素之一。[S6][S7] 测量者为腾讯，公司披露并非独立因果实验。边界：收入增长同时受广告需求、流量、广告负载与价格影响，不能把全部增长归因于 AI。可复制部分是事件信号、投放动作与交易结果的连接结构。

该案例显示，端到端价值更多来自信号、动作和反馈的系统连接，而非某一次生成质量。对非平台企业而言，首发闭环不必覆盖全链路，但必须形成一个能够回写的最小闭环。

第四章 八维评分：让价值、可行与证据进入同一张桌子

竹势 AI 营销智库出品

V 经营价值 收入 · 毛利 · 成本 · 速度
C 闭环完整 信号 · 行动 · 反馈
D 数据就绪 定义 · 质量 · 权限 · 时效
T 技术可行 连接 · 评测 · 降级
O 组织承接 owner · 岗位 · 例外
R 风险半径 可逆 · 可停 · 可追责
E 证据周期 领先证据 · 最终结果
X 复制潜力 机制 · 依赖 · 平台能力
分数用于暴露假设与分歧，不用于覆盖否决条件。

评分的目的不是制造精确幻觉，而是让假设、分歧和证据缺口公开化。

本报告采用八维评分：经营价值、流程稳定、数据就绪、系统连接、组织承接、评测能力、证据周期、可复制性。每维1—5分，分值必须附证据与置信度。建议权重由企业战略与风险偏好校准，禁止直接照搬固定行业权重。评分会前由业务、财务、技术、数据、风险分别独立打分，会中只讨论差异超过两分的维度。

维度	1分	3分	5分	管理问题
经营价值	与核心目标关系弱	可改善重要过程指标	直接作用于瓶颈并可连接财务结果	改善后能否进入收入、毛利、现金或避免成本？
流程稳定	流程依赖个人且频繁变化	主流程可描述，例外较多	SOP、输入输出与例外清晰	自动化的是稳定流程还是混乱流程？
数据就绪	字段缺失或用途不合法	可用但需清洗与补采	覆盖、质量、权限、主键与回写满足	当前闭环需要的数据今天能合法使用吗？
系统连接	无接口且动作靠复制粘贴	部分连接，关键动作人工完成	触发、执行、回写、日志完整	输出是否能进入真实系统？
组织承接	无负责人或无时间	有负责人但跨部门依赖未锁定	流程所有者、复核者与运维明确	谁对结果与异常负责？
评测能力	无基线、无验收	有过程指标与抽样	有对照、质量与结果指标	如何知道改善来自场景而非外部变化？
证据周期	超过90天且事件稀少	30—90天可形成有效信号	14—30天可出现稳定领先证据	最快上线后多久才有可信证据？
可复制性	高度依赖个人与一次性数据	同类团队可复用部分	规则、接口、数据契约和控制可复用	换团队后机制是否仍成立？

总分只用于同一批候选的相对比较，不用于跨企业、跨行业对标。建议同时记录置信度：A为直接数据或已运行证据，B为可核验流程与专家判断，C为未经验证假设。一个高分但大量C级证据的场景，应进入“补证据”而非“立即上线”。

案例 3 | BCG 顾问“Jagged Frontier”实验（2023）

主体与时间：哈佛、MIT、沃顿等研究者与 BCG，2023年；758名顾问参与。业务情境：咨询人员完成位于或超出模型能力边界的知识任务。具体动作：实验组使用 GPT-4，部分接受使用策略培训。作用机制：模型在其能力边界内提高速度与质量，但在边界外可能诱发过度依赖。公开结果：在适合 AI 的任务上，使用者完成更多任务、速度更快、质量更高；在模型不擅长的任务上，使用 AI 的参与者正确率下降。[S8] 测量者为研究团队。边界：实验任务并非企业完整经营流程。可复制部分是按任务类型评分、识别能力边界、设置例外与人工复核，而不是用一个总平均掩盖差异。

第五章 风险否决门：高分不得抵消红线

竹势 AI 营销智库出品

STOP 01

不可逆动作

权益、预算或外部承诺无法回滚

STOP 02

权限过宽

授权超过业务责任与监控能力

STOP 03

无法验收

质量、事实或合规没有可靠关口

STOP 04

无法暂停

异常不能隔离、降级或人工接管

高价值不是越过红线的通行证；首发场景必须保留人工关口和止损开关。

评分回答“值得做”，否决门回答“是否允许现在做”。

Porter 的战略取舍说明，战略的实质包含选择不做什么。首场景治理同样需要明确拒绝：未经授权的数据用途、不可逆的客户承诺、无复核的资金动作、无法回滚的批量发布、无法审计的外部执行，都不能通过“高价值”抵消。实物期权思想进一步说明，面对不确定性应以小额、分阶段、可退出投入换取学习权，而不是一次性承诺全部资本与权限。

否决门	红线问题	首发最低控制	触发后处理
合法与目的	数据取得、使用目的、内容与行业规则是否合法？	法律依据、目的限定、最小必要	暂停，完成法律与合规评估
授权半径	AI能否直接触达客户、资金、价格、账号或承诺？	角色权限、额度、白名单、审批	缩小动作范围或改为建议模式
可逆性	错误动作能否及时撤销、恢复与补救？	草稿、沙箱、回滚、版本记录	不可逆动作退出首发
人工复核	关键节点是否有具备能力和时间的人复核？	明确复核人、时限、抽样与升级	无承接人则不立项
隐私安全	是否涉及个人信息、敏感信息、跨境或第三方泄露？	脱敏、访问控制、日志、供应商条款	完成数据保护影响评估
品牌安全	是否可能造成虚假、歧视、侵权或品牌承诺偏差？	禁区、事实核验、素材权利、发布审批	降级为内部使用或退出
可观测性	动作、输入、输出与结果是否可追踪？	唯一ID、时间戳、日志与监控	先补可观测能力

案例 4 | Air Canada 聊天机器人错误承诺（2022—2024）

主体与时间：加拿大航空及不列颠哥伦比亚省民事解决法庭，争议发生于2022年，裁决于2024年。业务情境：客户通过官网聊天机器人询问丧亲票价，机器人提供与正式政策不一致的信息。具体动作：客户依据信息购票并事后申请退款。作用机制：面向客户的自动答复形成了可被依赖的价格与政策表述，但缺少一致知识源、复核和纠错。公开结果：法庭认定加拿大航空应对官网聊天机器人的误导信息负责并赔偿客户；测量者为司法裁决。[S9] 边界：案件金额有限，不能代表所有客服自动化。可复制部分是把价格、合同、资格与承诺类回答列为高授权动作，要求权威数据源、引用、升级与人工接管。

第六章 数据就绪：数据存在不等于闭环可用

竹势 AI 营销智库出品

定义 对象与状态一致
质量 完整、准确、可追踪
权限 目的、角色与期限
时效 满足决策窗口
回写 行动与结果可闭合
数据就绪面向当前闭环，而不是等待一座抽象的“完美数据平台”。

数据就绪度必须按当前场景与当前目的评估，不能用“公司有数据仓库”代替。

社会技术系统理论指出，技术、任务、人员与组织结构需要联合优化。数据就绪也不是技术部门单独完成的库存盘点，而是业务语义、合法目的、流程责任与系统能力的交集。一个客户表存在，并不表示可以用于自动触达；一个广告报表存在，并不表示创意资产拥有可连接的唯一标识；一个内容库存在，并不表示素材权利和版本可追溯。

检查项	关键问题	通过证据	常见假通过
目的与合法性	该字段能否用于当前判断和动作？	处理依据、告知/同意、用途记录	“数据是公司的所以都能用”
对象与主键	客户、内容、活动、线索能否跨系统对齐？	唯一ID、映射规则、重复处理	靠姓名或标题模糊匹配
覆盖与代表性	关键人群、渠道、时间是否缺失？	覆盖率与缺失分布	只看总体完整率
时效	数据到达速度是否满足动作窗口？	延迟分布、刷新频率	月报支持实时决策
质量与语义	字段定义、单位、状态是否一致？	数据字典、校验规则、责任人	同名字段含义不同
访问与回写	系统能否读取、写回并留痕？	接口、权限、失败重试、日志	只能导出Excel
权利与保留	素材、文本、客户信息能保存多久、如何删除？	权利证明、保留策略、删除机制	永久留存默认值

案例 5 | 阿里巴巴淘宝天猫商家 AI 工具（2023—2024）

主体与时间：淘宝天猫，2023年10月向部分商家推出，2024年初扩大开放。业务情境：商家需要生产营销素材、处理客服与分析经营数据。具体动作：平台推出10项 AI 工具，覆盖图像与营销内容生成、客服和经营分析。作用机制：平台拥有商品、流量、交易和商家身份等结构化数据，可把生成与分析嵌入经营界面。公开结果：阿里披露这些工具在2023年“双11”期间被使用超过15亿次；2024年“双11”期间，AI电商工具服务400万商家，生成超过1亿份营销图像、视频和文本，另有约29万商家通过“全站推”推动160万件商品销售增长。[S10][S11] 测量者为阿里巴巴，未披露独立对照、利润增量与商家分布。可复制部分是商品、内容、流量和交易对象的一致数据结构；使用次数不能等同于经营价值。

第七章 证据设计：工时、现金与增量利润的边界

竹势 AI 营销智库出品

DAY 0

基线

口径、对象、测量者

DAY 14

运行证据

质量、采用、异常

DAY 30

行为证据

周期、响应、机会

DAY 90

经营决议

复制、调整或停止

低频 B2B 场景可用领先证据管理试点，但不得把尚未实现的收入写成归因结果。

管理层需要同时看到领先指标、结果指标和财务转换条件。

任务级生产率证据通常停在“更快、更多、更好”。企业投资需要进一步回答：节省的时间被如何使用；是否减少外包、加班或新增岗位；增加的产能是否投向有需求的活动；客户行为是否改变；新增毛利是否超过模型、集成、治理和运营成本。若这些转换条件未发生，工时节省是潜在产能，不是现金收益。

证据层级	示例指标	证明什么	不能直接证明
活动量	生成数量、处理量、采用率	系统被使用并产生输出	质量、客户行为、利润
流程表现	周期、错误率、首次解决率、等待时间	流程更快或更稳定	现金已经回收
客户行为	点击、回复、预约、转化、留存	客户反应发生变化	变化完全由 AI 引起
财务结果	增量毛利、避免成本、现金回收期	经济结果及投入产出	长期可复制与风险可控
能力资产	数据契约、规则、接口、评测集	后续场景边际成本可能下降	本次场景已成功

证据周期应从“开始收集有效样本”算起，而非从项目启动日算起。最快上线的场景可能因为购买周期长、事件稀少或归因窗口长，90天仍无法判断。相反，一个需要两周集成但每天产生高频可验收事件的闭环，可能更快形成可信证据。

案例 6 | 生成式 AI 客服现场研究（2020—2021数据，2023发表）

主体与时间：研究者 Brynjolfsson、Li、Raymond；数据来自一家财富500强软件公司的5,179名客服人员。业务情境：客服人员处理技术问题。具体动作：系统基于历史成功对话向客服实时提供建议，人员可采用、修改或拒绝。作用机制：将高绩效员工的沟通与解决模式扩散给低经验员工。公开结果：每小时解决问题数平均提高约14%，低经验人员改善更明显；测量者为独立研究团队。[S2] 边界：结果依赖大量历史对话、明确工单结果和高频任务；没有证明减少工资支出或增加收入。可复制部分是高频事件、明确验收、人员保留最终判断与分层分析。

案例 7 | Klarna AI 客服助手（2024）

主体与时间：Klarna，2024年2月上线后首月披露。业务情境：多市场、高频客服咨询。具体动作：AI 助手处理订单、退款等客户对话。作用机制：连接账户与订单信息，提供24小时多语言处理。公开结果：Klarna自报首月完成230万次对话，约占客服聊天三分之二，等效700名全职人员工作量，重复咨询下降25%，平均解决时间从11分钟降至不足2分钟，客户满意度与人工相当。[S12] 测量者为Klarna及其技术合作方，属于公司自报。边界：等效工作量不等于裁减700名直接员工；公开材料未提供随机对照和完整成本。可复制部分是限定任务、高频反馈和质量/时效双指标。

第八章 依赖图与场景组合排序

竹势 AI 营销智库出品

共享基础 客户主数据 · 权限 · 评测 · 观察性 · 人工接管
先能力后场景 多场景共享且风险关键的依赖
互补组合 上下游共同闭合反馈回路
资源冲突 同一 owner、数据或系统窗口竞争
组合排序看边际学习与共享依赖，不把单项高分简单相加。

候选场景不应被当作彼此独立的项目；共享能力与资源冲突会改写排序。

单场景评分忽略四类关系：前置依赖、互补关系、资源冲突和共享能力。前置依赖如客户身份解析、内容资产ID、转化事件采集；互补关系如“创意生成”与“投放反馈回写”；资源冲突如多个场景同时占用同一数据工程师和业务复核人；共享能力如品牌知识库、评测集、审批、日志和连接器。

组合排序可采用“净启用价值”思路：场景本身价值，加上它为后续场景提供的共享能力价值，减去依赖缺口、资源冲突和风险成本。这里不要求把所有因素货币化，而要求显式标出关系。Stage–Gate 方法的管理意义在于每一阶段投入增加前都必须重新审议，而不是批准一次后自动走完全程。

依赖图应把“共享依赖”和“场景专属依赖”分开。共享依赖包括统一客户标识、事件口径、品牌规则、审批日志、评测集和通用连接器；它们一旦稳定，可以降低多个后续场景的边际实施成本。场景专属依赖只服务某一渠道、产品线或业务单元，过早建设容易形成局部最优。审议时可为每项依赖标注所有者、成熟状态、受益场景和失效后果，再判断它属于必须先完成的地基、可与试点并行验证的条件，还是应在证据成立后才投入的扩展。这样能够防止“先建大平台再找用途”，也能避免每个项目重复建设身份、数据、审批和评测能力。管理层需要批准的是最小共享底座及其服务边界，而非一次性批准覆盖全集团的基础设施。

互补场景的判断依据是两者能否共同完成同一经营反馈回路。例如，内容生成与投放回写具备互补性，因为后者为前者提供验收和学习信号；两个都生成内容的工具通常只是能力重叠。资源冲突则要按稀缺角色与时间窗口识别：同一名数据负责人、法务复核人或渠道运营者被多个试点同时占用，会让纸面高分失真。组合排序会议应先冻结关键资源容量，再比较候选场景对共享能力的贡献、对稀缺资源的消耗以及延误主业务的机会成本。若某项能力本身没有独立经营闭环，却是多个高价值场景的共同前提，可以采用“先能力后场景”，但必须设置明确的受益场景、验收标准和停止条件；没有这些约束的能力建设容易演变为长期平台工程。[S19][S21]

关系类型	识别问题	排序影响	管理动作
前置依赖	没有哪项能力就无法运行或评测？	缺失则后移或拆成前置项目	明确依赖完成标准与负责人
互补场景	与哪个场景组合才形成闭环？	可合并为一个主闭环的子模块	只保留一个经营主目标
资源冲突	是否争夺同一专家、数据、预算或渠道窗口？	冲突严重则错峰	锁定关键资源容量
共享能力	是否创建可被多个场景复用的资产？	适度上调启用价值	把资产交付写入Charter

案例 8 | 快手 AI 驱动在线营销服务（2025）

主体与时间：快手科技，2025年。业务情境：平台需要提升广告定向、素材理解、投放和转化效率。具体动作：公司将大模型用于智能投放、内容理解、推荐、AIGC短视频、数字人和数字员工等营销产品。作用机制：平台流量、内容和广告反馈共享底层数据与模型能力。公开结果：2025年第二季度在线营销服务收入人民币198亿元，同比增长12.8%；公司将增长部分归因于 AI 优化的智能营销产品。第三季度收入人民币201亿元，同比增长14.0%。[S13][S14] 测量者为快手，公司未披露独立因果分解。边界：平台收入还受宏观广告需求、客户结构和销售影响。可复制部分是先建设共享内容理解、事件反馈和投放基础，再扩展多个场景。

第九章 中国经营语境下的特殊约束

竹势 AI 营销智库出品

中国语境：把动作责任与标识义务写进流程

首发闭环需同时处理个人信息、生成合成内容标识、平台规则、内容事实与客户权益；合规节点必须进入 Charter，而不是上线后的补丁。

中国企业的首发资格必须同时通过平台、数据、集团治理与生成合成内容标识四类现实约束。

中国营销链路高度平台化。公众号、视频号、抖音、小红书、快手、淘宝天猫、企业微信、广告平台和本地生活平台各自拥有接口、账号、审核、内容与商业规则。场景在原型环境中成立，不代表可进入生产。接口权限、自动化限制、账号主体、内容审核、素材授权和结果回传都应在评分前核验。

集团企业还存在总部与业务单元、品牌与渠道、市场与法务、数据与信息安全之间的多层授权。首发场景若跨越多个法人、多个数据域或多个品牌，证据周期和风险半径都会扩大。更稳妥的做法是选择单一业务单元、单一人群、单一渠道窗口，建立可审计样板，再以共享控制扩展。

《人工智能生成合成内容标识办法》把文本、图片、音频、视频和虚拟场景纳入生成合成内容，并规定显式与隐式标识要求。[S3] 这会影响内容制作、文件元数据、平台发布和供应商交付。首发场景不能只检查生成质量，还要检查标识是否完整保留、平台是否识别、二次编辑是否破坏标识。

案例 9 | 中国生成合成内容标识制度（2025）

主体与时间：国家互联网信息办公室等四部门，2025年3月发布，2025年9月施行。业务情境：生成合成内容规模化传播带来识别、误导与责任问题。具体动作：制度建立显式与隐式标识要求，并要求服务提供者、内容传播平台和用户承担相应义务。作用机制：通过用户可感知标识与文件技术标识提高可识别性和责任追溯。公开结果：这是监管制度，不是企业经营案例，不存在ROI结果。[S3] 边界：具体行业广告、知识产权、个人信息和平台规则仍需另行适用。管理含义：生成内容发布类场景必须把标识保留作为验收项。

第十章 首个试点 Charter 与 14/30/90 天闸门

竹势 AI 营销智库出品

FREEZE 范围冻结 一个对象、一个动作、一个结果
14D 运行闸门 日志、质量、采用与异常
30D 学习闸门 机制、行为与领先证据
90D 资本闸门 复制、调整、暂停或淘汰
任务书必须写清责任、预算、对照、风险、例外和停止条件。

试点 Charter 是经营合同：定义命题、范围、证据、权限、预算、暂停和退出。

首个试点只能有一个经营主闭环，但可以包含必要子模块。Charter 至少包括：经营问题、目标人群与流程、五环定义、基线与对照、领先与结果指标、数据范围、系统连接、责任分工、风险控制、预算与资源、14/30/90天闸门、暂停和退出条件、复制假设。

时间闸门	核心问题	必须交付	通过条件示例	不通过动作
14天	闭环能否安全运行？	流程图、数据契约、最小连接、评测样本、权限与回滚演练	关键数据可读写；高风险动作需审批；异常可追踪	暂停生产动作，修复依赖或缩小范围
30天	是否出现稳定领先证据？	基线、对照/分组、质量抽样、周期与错误数据	至少两个流程指标稳定改善且质量不下降	调整规则、数据或工作设计；不得扩大权限
90天	是否形成经营结果与复制资格？	客户行为、财务转换、总成本、机制解释、复制清单	结果可复核；成本完整；风险可承受；依赖可复制	暂停、淘汰或限定为辅助工具

14天不承诺财务结果，重点是技术与治理可运行；30天不急于宣布规模化，重点是稳定领先证据；90天才审议经营结果与复制资格。对于长销售周期 B2B 场景，90天可能只能观察合格商机、响应与销售阶段推进，不能虚构收入归因。

长周期 B2B 场景的 Charter 应把证据链拆成可观察的阶段转换，而不能用90天内是否签约作为唯一成败标准。可审议的领先证据包括目标账户覆盖是否增加、有效触达是否发生、响应是否进入人工确认、需求是否被结构化记录、合格机会是否进入下一销售阶段、销售人员是否按规则完成跟进，以及丢单原因是否被回写。每个指标都要绑定定义、分母、数据来源、责任人和反证条件。例如，回复数量上升但目标账户质量下降，不能视为进展；销售阶段被批量改写却没有客户行为证据，也不能视为机会推进。管理层据此判断闭环是否改善了机会发现和推进机制，而不是提前声称收入由AI带来。

90天审议可采用“收入归因尚未成熟、机制证据已经形成”的双层结论。第一层检查流程因果链：系统输出是否被采用，采用后是否触发了规定动作，动作后是否出现客户侧可验证变化；第二层检查财务可转换性：这些变化在既有历史转化逻辑下是否有望形成订单，以及还需要多长观察窗口。若销售周期跨越多个季度，Charter 应保留后续跟踪队列，并把90天决策限定为继续观察、调整机制或停止投入。任何预测收入都应已与已确认收入分列，不能用管道金额、意向金额或平台归因替代财务结果。这样既避免因周期过长过早淘汰有效机制，也避免把乐观销售漏斗包装成投资回报。

第十一章 复制、暂停与淘汰

竹势 AI 营销智库出品

复制机制

因果链在不同团队仍成立

抽离明星

移除冠军团队的隐性兜底

识别共享能力

平台资产与场景特例分开

保留停止权

成本、风险或证据失效即撤回

复制的是可解释机制与可承受控制，不是一次漂亮结果。

首发成功只证明“在当前条件下有效”，复制还需证明机制、依赖与控制可迁移。

Rogers 的创新扩散理论指出，扩散受相对优势、兼容性、复杂性、可试用性和可观察性影响。AI 场景复制同样不能只复制软件配置；需要复制业务规则、数据语义、角色能力、复核负荷和异常处理。若首发依赖明星团队、临时人工清洗、领导亲自协调或特殊渠道资源，结果属于“单团队冠军效应”。

决策	适用条件	必须保留的证据	禁止做法
复制	机制可解释；依赖可复用；质量与风险稳定；边际成本下降	版本、数据契约、评测集、复核负荷、异常率	只凭展示效果扩张
暂停	关键依赖暂缺；外部规则变化；质量波动；样本不足	暂停原因、恢复条件、保留资产	继续累积不可控动作
淘汰	经营假设被否认；风险不可接受；总成本长期高于价值	否认结论、沉淀资产、数据删除与权限回收	用沉没成本延长项目
限定使用	局部生产率成立但无法闭环	适用任务、人工责任、禁止范围	包装成全流程解决方案

高可靠组织研究强调对失败保持敏感、避免过度简化、尊重一线专业判断并维持恢复能力。用于 AI 营销，管理层应关注近失误、人工改写率、异常升级、回滚次数和未覆盖例外，而不只看平均准确率。

复制资格需要通过三种拆分检验。机制拆分检验回答结果由哪几个必要环节产生，移除其中一环后是否仍成立；团队拆分检验把明星负责人、临时加班、手工清洗和高层协调列为显性投入，判断普通团队能否在标准资源下运行；平台拆分检验确认哪些成果来自共享身份、数据、评测、审批和连接能力，哪些只依赖首发业务的特殊条件。只有关键步骤可说明、例外可处理、复核负荷可承受，才称为机制可复制。若复制方案仍要求原团队驻场兜底，实际复制的是专家服务，不是经营系统。

平台型共享能力也不能仅凭“可复用”标签获得扩张预算。管理层应观察后续场景是否真实调用同一数据对象、控制规则和评测资产，调用后是否减少交付周期、重复清洗和风险审查。若每个业务单元仍需重新定义口径、重做连接、重建规则，所谓平台只是统一界面。复制决策宜分为原样复制、参数化复制、重新设计和停止四类：原样复制适用于流程与风险高度一致的单元；参数化复制允许渠道、品牌或人群差异；重新设计用于机制相同但数据与授权结构不同的场景；停止适用于收益依赖特殊资源或风险无法收敛的情形。[S22][S24]

第十二章 董事会治理与实施建议

竹势 AI 营销智库出品

最终决议包

首发场景 + 经营假设 + 损失上限 + 责任 owner + 证据阶梯 + 风险开关 + 下一次决议日

董事会应批准的是证据纪律、风险边界和资本分段，不是某个模型品牌。

董事会与经营层应把首场景纳入投资组合治理。CFO负责财务转换条件与总成本，CMO负责客户任务与流程结果，CIO/CDO负责连接、数据和运行可靠性，法务与风险负责否决门，业务单元负责人承担结果，场景产品负责人承担闭环运行。责任不能落在“AI团队”这一模糊主体。

治理问题	董事会应要求的答案	拒绝接受的答案
为什么是它？	它直接作用于哪个经营约束，替代候选为何后移	因为同行都在做或演示效果好
如何证明？	基线、对照、样本、领先与结果指标、归因限制	上线后看数据
如何止损？	预算分段、权限、暂停、退出、数据删除与回滚	项目做完再复盘
谁负责？	流程所有者、复核人、技术运维、风险升级人	AI团队统一负责
何时复制？	机制、依赖、风险、边际成本四项均通过	首月指标好就全集团推广

当代专家观点可归纳为四组。Erik Brynjolfsson强调生产率提升往往依赖工作重构和互补投入；Karim Lakhani与研究同事通过现场实验提示 AI 会改变团队知识边界，但任务边界必须被识别；Ethan Mollick主张把 AI 当作协作者进行广泛试验，同时保持人类判断；Andrew Ng长期强调从具体工作流与可行项目切入；Daron Acemoglu提醒自动化收益取决于任务重组，过度自动化可能削弱劳动与组织价值；NIST 的 AI 风险管理框架强调治理、映射、测量和管理的持续循环；中国监管机构强调内容安全、个人信息保护与生成合成内容标识；咨询实践则反复指出规模化价值与流程重构、数据、治理和采用机制相关。^{[S1][S2][S3][S4][S5][S8][S15]} 这些观点支持本报告的共同结论：技术能力只能作为输入，经营系统决定结果。

第十二章补充 | 六类常见候选的审议演示

以下演示不是行业排名，也不预设任何企业应选择同一场景。它们展示同一套方法如何改变直觉排序。每个示例都必须用企业自己的基线、数据和风险条件重新评分。

一、社媒内容批量生成：容易上线，但常缺少客户行为闭环

许多企业把社媒内容生成列为第一场景，原因是需求可见、工具成熟、员工容易参与。若候选定义停在“每周生成更多帖子”，它通常在流程稳定和证据周期上得分较高，在经营价值、评测能力与反馈上得分较低。内容数量、制作周期和人工满意度可以很快测量，却难以回答内容是否触达正确人群、是否推动有效互动、是否带来线索或降低销售解释成本。

把该候选升级为闭环，需要将对象限定为某一客户任务和渠道窗口，例如“针对进入产品比较阶段的潜在客户，基于销售异议库生成证据型内容，经事实与品牌审核后发布；以内容触达后的资料下载、咨询预约和销售异议变化为验收，并把高频问题回写选题库”。此时，内容生成只是判断与行动的一部分，销售反馈和客户行为成为闭环。若企业没有内容资产ID、来源标记、落地页事件或CRM回写，数据依赖会把该场景排序后移。

二、广告素材变体：高频反馈有利于证据，但需防止平台归因替代企业判断

广告素材变体通常具备较短证据周期，因为曝光、点击、转化等事件频繁，且平台提供实验与归因工具。它可能成为优质首发场景，前提是企业能够控制受众、预算、投放时段和落地页差异，并将创意版本与结果唯一关联。若每个变体同时改变文案、视觉、受众和出价，结果无法解释；若只采用平台自动归因，也可能把平台优化、季节性和竞争变化误认为素材效果。

合格 Charter 应限制变量数量，保留人工审核，设置单日与总预算上限，规定不得自动修改价格、优惠承诺和敏感定向。领先指标可以是素材审批周期、可用变体率和学习速度；结果指标可以是相同投放条件下的转化成本或合格线索率；财务转换需进一步考虑毛利、退货、销售承接和客户质量。

三、线索分级与销售跟进：靠近收入，但数据语义和组织承接常成为真正瓶颈

线索分级看起来最接近收入，因此经常获得很高的经营价值评分。但若CRM阶段定义不一致、销售未及时更新结果、渠道来源字段缺失、历史成交样本偏差明显，模型只能放大旧流程的噪音。即使排序准确，销售团队没有响应时限、行业话术和负责人，客户行为也不会改变。

首发可先限定为“建议模式”：AI根据已批准字段生成优先级与理由，销售确认后触发下一步任务；系统记录建议、采纳、响应时间、阶段变化和最终结果。14天验证字段与流程，30天比较响应速度、覆盖率与销售采纳，90天观察合格商机推进。涉及个人信息、自动画像和差异化触达时，必须评估合法目的、透明度和歧视风险。

四、客服与私域自动回复：事件高频，但授权半径和品牌责任更大

客服与私域场景容易形成高频验收，能够测量首次响应、解决率、升级率和满意度；同时，它直接面对客户，错误信息可能形成价格、资格、产品效果或合同承诺。首发资格取决于知识源权威性、问题类型分层、引用能力、人工接管和错误可补救性。

建议把问题划分为三层：低风险事实查询可在受控知识库内自动回答；需要账户数据的事项在身份核验后处理；价格、退款、合同、医疗、金融、法律与重大投诉默认转人工。人工复核并非固定不变，可在积累稳定证据后按问题类型逐步放宽。平均准确率不能替代高影响错误率，企业应单独统计错误承诺、隐私泄露、投诉升级和近失误。

五、营销分析与周报：风险低、可逆性强，但容易成为“漂亮报告”

自动分析和周报通常具有较低风险，适合作为能力建设，却常被误判为经营闭环。报告是否按时生成、图表是否完整，只证明信息加工效率。若管理会议不使用结论、没有触发预算、内容、渠道或销售动作，报告不会改变经营结果。

要形成闭环，必须把每条洞察绑定行动、负责人、截止日期和预期指标，并在下一周期检查执行与结果。例如，异常发现触发素材暂停、落地页修复或销售跟进；系统记录谁批准了什么动作及结果。若企业当前连基础指标口径都不一致，首发任务应先建设指标契约，而不是生成更复杂的自动解释。

六、自动投放与预算调整：潜在价值高，但通常不适合成为最早授权的场景

自动投放能直接改变资金使用和客户触达，理论价值高，风险半径也大。首发阶段更适合“诊断—草稿—人工审批—执行包—结果回写”的闭环，由AI持续巡检账户、提出有依据的调整建议，并在沙箱或草稿中显示预算、出价、定向和素材变化。只有当建议质量、异常率、回滚和监控经过多个周期验证，才考虑有限额度的自动执行。

授权应按动作拆分，而非按系统整体授予。暂停低表现素材通常比提高预算更可逆；调低单日额度比改变长期出价策略风险更小；白名单账户比全集团账户更适合试点。管理层应把每类动作的最大金额、最大人群、最大持续时间和复核时限写入权限策略。

候选类型	直觉优势	常见缺口	可能的首发改造
社媒内容生成	上线快、团队有感	缺客户行为与反馈	限定客户任务，绑定CTA、来源与销售反馈
广告素材变体	高频事件、易实验	变量混杂、平台归因偏差	限制变量、统一受众预算、记录版本
线索分级	靠近收入	CRM语义与销售承接不足	建议模式、记录采纳与阶段推进
自动客服	高频、时效可见	错误承诺与隐私风险	问题分层、权威知识源、人工升级
分析周报	风险低、可逆	不触发真实动作	洞察绑定行动、负责人和复盘
自动投放	直接影响经营	资金与品牌风险高	先做诊断草稿与有限审批执行

当代专家观点：主张、证据基础与适用边界

专家观点用于帮助管理层提出更好的问题，不能替代企业自身证据。以下观点覆盖劳动经济学、组织实验、产品实践、风险治理和监管。

专家/机构	证据或经验基础	与本题相关的主张	本报告的使用边界
Erik Brynjolfsson	企业现场数据与数字经济研究	AI生产率依赖任务设计、经验扩散和互补投入	任务生产率不自动转换为利润
Danielle Li、Lindsey Raymond	5,179名客服现场研究	不同技能群体收益不同，应分层评估	特定客服场景不可泛化全部营销
Karim Lakhani	P&G、BCG等现场实验	AI可能改变团队与专业知识边界	仍需识别锯齿状能力边界
Ethan Mollick	教育、组织实验与广泛实践	应通过实际使用发现能力，同时保持人类责任	鼓励试验不等于放弃控制
Andrew Ng	AI产品与企业项目实践	从清晰、可行、有数据的具体 workflow 切入	可行性高的项目仍需经营价值审查
Daron Acemoglu	任务经济学与自动化研究	价值取决于创造新任务和改善工作，不只是替代劳动	宏观判断需转化为企业流程证据
NIST	AI RMF及跨行业风险治理	治理、映射、测量、管理应持续循环	框架不替企业决定具体风险容忍度
国家网信部门	中国生成式AI与内容治理制度	内容安全、个人信息与标识是生产条件	仍需叠加广告、行业 and 平台规则
Michael Porter	竞争战略与活动系统	价值来自相互强化的活动与取舍	不能把技术功能清单当战略
Robert Cooper	新产品Stage–Gate研究	不确定项目应分阶段增加资源承诺	闸门需加入AI特有数据与授权风险

评分校准会议：如何避免权重政治与数字幻觉

评分会议最常见的问题不是数学错误，而是部门用分数维护立场。业务团队容易高估价值并低估例外，技术团队容易高估可行性或把接口完成视为成功，风险团队可能用最坏情形否定全部试验，财务团队可能在证据尚未形成时要求精确ROI。校准机制应把这些角色差异转化为可审议信息。

第一步，所有评分人依据同一张候选定义卡独立打分，避免会议中的锚定。第二步，评分必须附依据和置信度；没有数据的判断可以保留，但应标为C。第三步，只讨论差异超过两分或置信度冲突的项目，讨论目标是确定需要补什么证据，而非强行达成一致。第四步，权重在看到候选分数之前由决策委员会批准，防止为偏好的场景倒推权重。第五步，做敏感性分析：在合理权重范围内，排序是否稳定；若轻微调整就改变第一名，说明结论脆弱，应先补证据。

评分结果建议同时展示三张图：价值—就绪二维图显示相对位置；风险否决表显示不可进入生产的原因；依赖图显示前置与共享能力。单一雷达图容易产生“面积大即更好”的错觉，也无法表达红线和依赖。

财务审议：把潜在产能转换为可验证经济结果

CFO在首场景中不应只计算模型调用费，也不能在试点前要求一个看似精确的三年净现值。更适合的做法是建立成本与收益转换树。成本至少包括软件与模型、集成、数据清洗、评测、人工复核、培训、变更管理、风险控制、运维和机会成本。收益分为避免成本、释放产能、增量毛利、风险损失降低和能力期权。

避免成本要求存在真实可避免支出，例如减少外采、加班或已批准但尚未招聘的岗位；释放产能要求明确被释放时间将投入何种高价值任务，并测量新增产出；增量毛利要求客户行为变化、收入确认和单位经济性；风险损失降低需要有基线事件和合理概率，不能用极端事故金额制造收益；能力期权指数数据、接口、评测和治理资产降低后续场景成本，但应单独列示，不能与本次现金收益混算。

释放产能能否进入现金或毛利口径，取决于管理动作是否随之发生。若人员合同、外包采购、加班安排和招聘计划均未改变，节省工时只代表闲置或可重新配置的容量；若被释放时间投入更多销售会谈、实验迭代或高毛利客户服务，还需记录新增活动及其结果，才能讨论增量贡献。现金转换通常来自已取消的外采、未发生的新增招聘、减少的付费返工或更快回款；毛利转换还要求收入增加且履约成本没有同步上升。对暂时无法转化的能力，可作为期权单列：它可能降低未来场景的启动成本、缩短响应时间或提高风险承受能力，但期权价值需要后续真实调用才能兑现。财务审议因此应同时呈现已实现现金、可验证经营贡献、待验证产能和能力期权，禁止相互替代。

同一项效率改善在不同约束下会产生相反财务结果。需求不足时，更多内容产能可能增加审核、投放和品牌管理成本；销售跟进能力已饱和时，新增线索会降低响应质量；产能瓶颈恰好限制收入且下游能够承接时，效率改善才更可能转化为毛利。CFO与业务负责人需要共同确认约束位置、下游容量和成本行为，再决定是否把效率写入商业案例。约束发生变化后，原有收益假设必须重算，不能把试点时期的单位改善线性外推到全量规模。[S19]

收益类型	确认条件	可用证据	不能接受的替代
避免成本	预算或合同实际减少	取消外包、减少加班、避免已批准招聘	把节省小时乘平均工资
释放产能	时间被转投并产生新增有效工作	任务日志、产出与结果	员工表示更轻松
增量毛利	收入变化可比较且扣除变动成本	实验/准实验、订单与毛利	点击率上升直接等同利润
风险降低	有历史基线、控制改变与事件概率	错误率、投诉、损失事件	用最坏事故金额乘主观概率
能力期权	资产可被后续场景复用	接口、数据契约、评测集、控制模块	把未来所有可能收益计入本次ROI

组织承接：首场景是一项工作设计变更

AI进入流程后，人的工作不会只是“少做一些”。判断点、责任、技能、节奏和绩效都会变化。若员工仍按旧KPI被考核，往往会抵制系统、重复做一遍或只在演示时使用。首场景应明确哪些任务由AI准备、哪些由人判断、哪些必须升级、哪些旧步骤被删除。保留旧流程再叠加AI，通常造成增负而非增效。

流程所有者必须拥有调整SOP和协调跨部门资源的权力。复核人需要被分配真实容量，不能假设审核“顺手完成”。数据负责人要对语义和质量负责，技术负责人对运行与回滚负责，风险负责人对红线和例外负责。场景产品负责人负责把这些角色连接起来，但不替代业务负责人承担经营结果。

培训也应围绕工作任务而非通用功能。员工需要理解何时相信、何时核验、如何报告异常、如何留下可学习反馈。评估系统时，应同时观察采用率、绕行率、人工改写率、复核等待、异常升级和员工对责任边界的理解。

失败模式与纠偏动作

高分错觉的可观察信号包括：总分很高却无法画出五环、关键数据仍靠人工补录、风险门写着“后续确认”、排序对微小权重变化极其敏感。纠偏动作是暂停讨论总分，回到证据缺口、否决门和依赖图；必要时把候选降级为能力验证，不得直接进入生产。权重政治常表现为部门在看到候选结果后要求修改权重，或把本部门可控指标赋予过高比重。纠偏应由跨职能委员会在评分前批准权重与调整窗口，并保留敏感性分析；排序不稳定

时，追加证据比继续争论权重更有效。

试点永久化的信号是阶段闸门不断延期、人工兜底长期存在、预算以“小额续费”滚动追加，却始终没有复制或退出决策。纠偏动作是为每次延期写明尚缺证据、责任人和最终截止点；超过截止点仍不能判断，应停止或重新定义场景。过早自动执行通常出现于建议质量尚未稳定、异常样本未覆盖、回滚未演练时就扩大权限。可观察信号包括人工复核率被管理层视为负担而被压低、异常升级下降但客户投诉或人工返工上升、系统动作缺少可追踪理由。纠偏应退回草稿或建议模式，缩小账户、人群、金额和持续时间，恢复双人复核与回滚演练，直到控制证据与经营证据同时达标。

失败模式	早期信号	根因	纠偏
展示成功、生产失败	演示流畅，真实数据频繁报错	数据与例外未进入设计	用真实样本重做14天闸门
内容更多、业务不变	产量上升，客户指标无变化	缺少CTA、分发、销售与反馈	重定义为客户任务闭环
人工复核成为瓶颈	待审队列增长、上线延迟	权限过宽或审核无容量	按风险分层、缩小自动化范围
模型表现波动	同类输入结果不稳定	评测集、规则与版本管理不足	建立黄金样本、版本与回归测试
数据泄露或越权近失误	出现不该访问的字段或动作	角色权限与目的控制不足	立即暂停、回收权限、完成影响评估
冠军团队效应	试点团队好，第二团队失败	依赖隐性经验与领导协调	把经验转成SOP、数据契约与培训
ROI被工时替代	报告只写节省小时	没有财务转换条件	由CFO确认避免成本或增量毛利
多场景争抢资源	数据与业务专家频繁切换	没有唯一首发与组合治理	冻结其他生产试点，保留候选池

董事会审议脚本：用十二个问题压缩决策噪音

正式立项会不应从供应商演示开始。建议由业务流程所有者先用十分钟陈述经营约束和当前基线，再由数据、技术、财务、风险依次回答证据与边界。以下十二个问题可以作为会议议程，也可以作为投票前的最低信息要求。

审议问题	合格回答要求
1. 当前损失是什么	用等待、流失、返工、错误、成本或机会损失描述，避免只说效率低。
2. 谁的任务被改善	明确客户、销售、运营、渠道伙伴或管理者，不接受笼统“市场部”。
3. 哪个环节是系统约束	说明该环节改善后为何会改变整体结果，而非把问题转移到下一环。
4. 为什么现在做	说明业务窗口、数据条件、组织容量与不行动成本，避免用技术热度作理由。
5. 最小闭环是什么	明确五环和边界，列出暂不自动化的步骤。
6. 最强替代方案是什么	比较流程简化、规则自动化、培训、外包或传统软件，AI不是默认答案。
7. 哪些假设最脆弱	指出价值、数据、行为改变和归因中最可能被否证的假设。
8. 哪个错误最不可接受	从客户伤害、资金、合规、品牌和运营连续性评估。
9. 14天能证明什么	只承诺运行、数据、控制和评测可行性，不提前宣称经营成功。
10. 30天能否稳定复现	要求多个周期、不同样本和异常条件下仍保持质量。
11. 90天怎样进入财务	写清从流程指标到客户行为再到毛利或避免成本的转换。
12. 失败后留下什么	确保数据契约、评测集、接口、控制和否证结论可复用。

投票建议采用四种决议：批准进入14天、补证据后再审、降级为内部辅助、否决。不得用“原则同意、细节后补”绕过否决门。批准文件中应同时记录少数意见和未解决假设，以便30天闸门重新审议。

最小可行证据链：一个可复核示例

假设一家B2B企业发现，高意向官网线索进入CRM后，销售平均需要36小时才给出行业化回应，且大量回应缺少客户所在行业的案例与风险说明。候选场景定义为：当目标行业线索提交表单后，系统读取经授权的公司与需求字段，检索已批准的行业证据和案例，生成带引用的跟进建议与邮件草稿；销售在四小时内审核发送；系统记录采纳、改写、发送、回复、预约与商机阶段，并把新异议回写知识库。

该场景的信号是表单提交与线索评分，判断是行业、需求与证据匹配，行动是生成草稿并创建限时审核任务，验收是四小时响应率、事实错误率、客户回复和预约，反馈是销售改写与客户异议回写。首发不允许自动发送，不读取与目的无关的个人字段，不生成价格承诺。14天检查CRM字段、知识引用、任务与日志；30天比较试点组和相似历史或同期对照的响应时间、采纳率、错误率；90天观察预约率、合格商机推进和销售节省时间的实际去向。

即使30天显示响应时间显著下降，也不能立即宣称收入增长。若销售把节省时间用于更多有效跟进，并且合格商机推进提高，才具备财务转换路径。若预约率不变，可能说明响应速度不是约束，或内容缺乏说服力；此时应回到客户任务和瓶颈，而不是继续增加模型复杂度。若只有一名明星销售能高质量审核，组织承接评分仍然不足，不具备复制资格。

方法适用边界

本方法适合需要进入真实营销流程、涉及跨职能资源或具有外部风险的AI场景。对于纯个人探索、一次性创意脑暴或低影响内部辅助，完整治理可能过重，可以采用简化版本；一旦输出触达客户、影响预算、处理个人信息、改变价格或形成正式对外内容，就应恢复完整否决门和证据闸门。

评分也不能替代战略判断。若企业战略选择进入新市场、重建品牌或改变渠道模式，某些基础能力的短期财务证据可能较弱，但仍具有战略必要性。此时应明确它属于战略能力投资，不应伪装成短期ROI项目；同时仍需设置阶段产出、风险边界和退出条件。反之，短期指标优秀的场景若加深对单一平台或供应商的锁定，也可能不符合长期战略。

对于极低频、高损失事件，例如重大舆情、危机沟通或合规审查，常规A/B实验和90天样本可能不可行。企业可以采用历史回放、红队演练、专家盲评、情景模拟和近失误指标，但必须说明证据强度较低，避免把演练表现等同真实事件结果。对于高监管行业，首发更适合内部检索、证据整理、草稿与复核支持，生产授权应随证据逐步扩大。

管理者工具箱

工具 1 | 候选场景定义卡

使用方法：每个候选只填写一页，由业务流程所有者负责。若无法用一句话写清客户任务、瓶颈、动作与验收，说明粒度仍不合格。

字段	填写要求
场景名称	用“对象+触发+动作+结果”命名
客户任务/经营瓶颈	描述客户或内部流程要完成的进步及当前损失
起点与终点	明确触发事件、最终状态和时间窗口
当前基线	处理量、周期、质量、客户行为与成本
AI作用	识别、判断、生成、推荐、执行或监控
非AI互补投入	数据、流程、培训、接口、权限、内容资产
五环摘要	信号、判断、行动、验收、反馈
关键风险	数据、授权、可逆、品牌、合规
最短证据周期	从有效样本开始到可审议结果
流程所有者	对业务结果负责的人

工具 2 | 闭环完整性检查表

检查项	是/否	证据位置	缺口处理
触发信号可被系统识别			
判断规则或评测标准明确			
动作进入真实流程			
结果可被唯一对象记录			
验收阈值预先定义			
失败与异常有升级路径			
反馈会改变下一轮输入或规则			
人工与AI责任边界明确			

工具 3 | 多维评分表及校准说明

评分前先通过闭环资格与风险否决。每分必须附依据和置信度A/B/C。不同评分人独立完成，再讨论分歧。

维度	权重（自定）	评分1—5	置信度A/B/C	依据	主要缺口
经营价值					
流程稳定					
数据就绪					
系统连接					
组织承接					
评测能力					
证据周期					
可复制性					

工具 4 | 风险否诀门

门类	通过/否决/待补证	控制与证据	负责人
合法目的			
授权半径			
动作可逆性			
人工复核			
隐私与安全			
品牌与内容安全			
外部平台规则			
可观测与审计			

工具 5 | 数据就绪度检查

数据对象	目的合法	主键	覆盖	时效	质量	读取	回写	责任人

工具 6 | 证据周期与指标表

业务假设	领先指标	结果指标	财务转换条件	基线	对照/比较	观察窗口	归因限制

工具 7 | 场景依赖图与组合排序

候选场景	前置依赖	互补场景	共享能力	资源冲突	风险成本	净启用判断

工具 8 | 首个试点 Charter

模块	批准内容
经营命题	
范围与不做事项	
目标对象与流程	
五环定义	
基线与对照	
领先/结果/财务指标	
数据与系统	
角色与RACI	
风险控制与授权	
预算与资源	
14/30/90天闸门	
暂停与退出	
复制假设	

工具 9 | 14/30/90 天阶段闸门

闸门	通过证据	未通过原因	决策：继续/修复/缩小/暂停/淘汰	批准人
14天				
30天				
90天				

工具 10 | 复制 / 暂停 / 淘汰决策清单

问题	是/否	证据	决策影响
结果可由独立数据复核			
改善机制可解释			
关键依赖已标准化			
复核负荷可承受			
异常与近失误在阈值内			
总成本包含集成与治理			
换团队后流程仍可运行			
数据权利与标识持续满足			
边际复制成本下降			
退出与权限回收可执行			

附录 A | 经典管理思想与本方法的连接

竹势 AI 营销智库出品

思想	原始作者/作品	连接	当代修正
目标与结果	Peter Drucker, 《管理的实践》	内部活动必须连接外部客户与经营结果	增加数据、风险与可观测要求
待办任务	Clayton Christensen 等	从客户要完成的进步定义候选	企业场景还需纳入流程与治理
PDSA	W. Edwards Deming	行动结果必须进入学习循环	AI闭环需明确数据回写与版本
约束理论	Eliyahu Goldratt, 《目标》	优先处理系统瓶颈而非局部效率	同时评估客户、数据和组织约束
战略取舍	Michael Porter	明确哪些高风险动作不做	用否决门固化不可抵消边界
实物期权	金融与战略投资理论	分阶段小额投入购买学习权	以14/30/90天闸门控制资本与权限
Stage–Gate	Robert Cooper	每阶段增加投入前重新审议	加入证据质量、风险与复制资格
社会技术系统	Tavistock传统	联合优化技术、任务、人员与结构	把组织承接作为独立评分维度

附录 B | 完整案例一览与证据等级

竹势 AI 营销智库出品

案例	地区/类型	结果测量者	证据用途	主要限制
P&G Cybernetic Teammate	国际/消费/实证	独立研究团队	团队与任务质量	非市场投放与财务结果
腾讯广告技术	中国/平台/企业	公司披露	闭环与经营连接	多重增长因素
BCG Jagged Frontier	国际/B2B/实证	独立研究团队	任务边界与风险	实验任务
Air Canada chatbot	国际/高监管风险	司法裁决	授权与责任	单一案件
淘宝天猫商家AI工具	中国/消费平台	公司披露	数据与平台嵌入	缺少独立对照
客服现场研究	国际/B2B/实证	独立研究团队	生产率与技能扩散	不等于现金回报
Klarna AI助手	国际/金融科技	公司自报	高频闭环指标	公司口径
快手在线营销	中国/平台/企业	公司披露	共享能力与组合	因果分解不足
中国内容标识制度	中国/监管	监管制度	发布资格与控制	非企业案例

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，旨在帮助管理团队用受控方式验证 Agentic Marketing 的价值，并在规模化之前建立必要的证据、权限与问责机制。

shichangbu.ai

专题中心 · 返回顶部 ↑

来源索引（纯文本）

竹势 AI 营销智库出品

[S1] McKinsey & Company, 《The State of AI: How Organizations Are Rewiring to Capture Value》, 2025-03-12; 全球调查, 1,491名受访者, 指标为组织AI使用与价值实践。

[S2] Erik Brynjolfsson, Danielle Li, Lindsey R. Raymond, 《Generative AI at Work》, NBER Working Paper 31161, 2023; 5,179名客服人员, 采用分阶段上线数据, 生产率为每小时解决问题数。

[S3] 国家互联网信息办公室等四部门, 《人工智能生成合成内容标识办法》, 2025-03-14发布, 2025-09-01施行; 适用于文本、图片、音频、视频、虚拟场景等生成合成内容标识。

[S4] 国家互联网信息办公室等, 《生成式人工智能服务管理暂行办法》, 2023-07-13发布, 2023-08-15施行; 涉及内容安全、个人信息保护与服务规范。

[S5] Fabrizio Dell’Acqua 等, 《The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise》, 2025; P&G现场实验, 791名参与者。

[S6] Tencent Holdings Limited, 《2025 Annual Report》, 2026-04; 公司披露AI广告定向与业务经营信息。

[S7] Tencent Holdings Limited, 《2025 Annual and Fourth Quarter Results》, 2026-03-18; 营销服务收入及公司归因口径。

[S8] Fabrizio Dell’Acqua 等, 《Navigating the Jagged Technological Frontier》, Harvard Business School Working Paper, 2023; 758名BCG顾问现场实验。

[S9] Civil Resolution Tribunal of British Columbia, Moffatt v. Air Canada, 2024; 官网聊天机器人错误信息与企业责任裁决。

[S10] Alibaba Group, 《Taobao and Tmall Upgrades Consumer Shopping Experience with AI》, 2024-06-13; 10项商家AI工具及使用次数, 公司口径。

[S11] Alibaba Group, 《Taobao and Tmall’s 11.11 Shopping Festival Records Robust Growth》, 2024-11-12; 商家覆盖、素材生成与全站推结果, 公司口径。

[S12] Klarna, 《AI Assistant Handles Two-Thirds of Customer Service Chats in Its First Month》, 2024-02-27; 公司自报230万对话、时效、重复咨询与等效工作量。

[S13] Kuaishou Technology, 《Second Quarter and Interim 2025 Results》, 2025-08-21; 在线营销服务收入与AI产品公司归因。

[S14] Kuaishou Technology, 《Third Quarter 2025 Results》, 2025-11-19; 在线营销服务收入、智能投放与转化口径。

[S15] NIST, 《Artificial Intelligence Risk Management Framework (AI RMF 1.0)》, 2023; 治理、映射、测量、管理四类功能。

[S16] Peter F. Drucker, 《The Practice of Management》, 1954; 目标、客户与管理结果。

[S17] Clayton M. Christensen 等, 《Competing Against Luck》, 2016; Jobs to Be Done理论。

[S18] W. Edwards Deming, 《The New Economics》, 1993; PDSA学习循环与系统管理。

[S19] Eliyahu M. Goldratt, Jeff Cox, 《The Goal》, 1984; 约束理论。

[S20] Michael E. Porter, 《What Is Strategy?》, Harvard Business Review, 1996；取舍与活动系统。

[S21] Robert G. Cooper, 《Winning at New Products》, 相关版本；Stage–Gate产品创新治理。

[S22] Everett M. Rogers, 《Diffusion of Innovations》, 第5版, 2003；创新扩散属性。

[S23] Eric Trist, Fred Emery 等, 社会技术系统研究；技术、任务、人员与组织联合优化。

[S24] Karl E. Weick, Kathleen M. Sutcliffe, 《Managing the Unexpected》, 相关版本；高可靠组织。

版权与使用：本报告由竹势 AI 营销智库发布。引用公司自报数据时，应保留测量者、口径和限制。