



CHINA ENTERPRISE AI MARKETING /
MATURITY DIAGNOSTIC

中国企业AI营销

成熟度模型与基准报告

成熟度不是工具数量或永久标签，而是当前证据允许企业授予多大执行权、投入多少资本并承担何种风险。

识别真实成熟度

用八维依赖和五级证据门替代工具数量

诊断关键短板

封顶、否决和依赖规则暴露系统断点

建立可信基准

同群校准、缺失值与置信边界拒绝伪排名

形成升级处方

90 天把短板转为责任、证据与资本决议

BOARD BRIEF 01

出版说明

竹势 AI 营销智库出品

本报告面向企业董事长、CEO、CMO、CFO、CIO/CDO、营销运营负责人、数据与治理负责人以及 AI 转型业务负责人。报告建立一套可校准、可复核、可重复使用的八维五级成熟度模型，用于识别战略、组织、场景、流程、数据、技术、治理与绩效之间的结构性短板。

本报告不发布中国企业 AI 营销行业均值、分位数或排行榜。原因很明确：当前没有由竹势 AI 营销智库拥有或控制的代表性大样本问卷、客户访谈数据库或跨行业统一口径数据。公开材料中的采用率、效率数据与公司披露在样本、任务、行业、时间和指标定义上高度异质，只能作为外部锚点、案例口径或机制证据，不能被拼接为伪基准。

报告中的级别表示“当前证据状态”，不构成永久认证。自评必须附证据附件、验证责任人与有效期；关键证据缺失时，按保守原则降级或标注低置信。

- 1. 执行摘要与十项可证伪判断
- 2. 第1章 成熟度的经营定义与证据观
- 3. 第2章 八维依赖地图与系统判定规则
- 4. 第3章 五级证据阶梯
- 5. 第4章 战略维度：取舍、期权与决议日
- 6. 第5章 组织维度：责任、决策权与接口
- 7. 第6章 场景与流程：小闭环、标准运行与学习
- 8. 第7章 数据与技术：业务可用、可连接、可替换
- 9. 第8章 治理：按动作风险配置授权半径
- 10. 第9章 绩效：从活动到增量价值与风险
- 11. 第10章 基准：同群校准、缺失值与置信边界
- 12. 第11章 案例卷宗与专家判断
- 13. 第12章 升级处方、90天计划与董事会工具
- 14. 结语与来源索引

BOARD BRIEF 02

执行摘要与十项可证伪判断

竹势 AI 营销智库出品

BOARD DIAGNOSTIC RULE / 董事会诊断原则

成熟度是当前证据允许企业授予多大执行权、投入多少资本并承担何种风险的经营判断；它不是工具数量、内容产量或永久认证。

证据先于评分

依赖约束级别

同群优于总榜

升级可以撤销

董事会面对的真实问题并非“是否使用 AI”，而是企业是否已经形成可证明、可复制、可控风险的营销经营能力。高采用率可以与低成熟度同时存在：员工频繁使用通用工具，内容产量显著上升，但用例没有经营 owner、缺少结果基线、数据状态不可用、输出无法回写业务系统、关键动作没有日志和回滚，企业仍停留在局部劳动替代。任务级随机实验确实证明生成式 AI 可在特定边界内显著提高速度与质量，但同一研究也反复显示能力前沿不平整、人员收益不均、越界任务可能出现过度依赖。因此，成熟度必须评价完整经营系统，而不是评价工具热度。

本报告提出“八维五级成熟度模型”：八维分别为战略、组织、场景、流程、数据、技术、治理、绩效；五级分别为“无主试用、受控试点、闭环运营、可复制组合、适应性经营”。每一级都设置必须存在的证据、关键责任人、允许的执行半径、升级门和降级信号。整体级别不取八维简单平均，而采用“证据门 + 关键约束 + 依赖封顶 + 风险否决”的判定。

编号	可证伪核心判断
判断1	工具数量、员工使用率和内容产量均不能单独证明成熟；若没有经营 owner、基线与结果证据，整体最高只能判为“受控试点”。
判断2	八维是互补资产。数据、流程、组织与治理任何一项低于闭环门槛，场景规模化就会被封顶；单项领先只代表局部能力。
判断3	成熟企业优先构建端到端小闭环，再扩展场景组合；按热度堆叠单点用例通常增加接口成本和治理暴露。
判断4	探索项目可以暂不满足成熟业务的利润门槛，但必须有预算上限、可检验假设、停止条件和下一次决议日。
判断5	集中式 AI 中心适合建立标准、平台和高风险控制；业务嵌入式团队适合场景所有权和快速迭代。高成熟组织通常采用联邦式组合。
判断6	数据集中并不等于可用。字段定义、质量、权限、时效、血缘与业务状态缺一，AI 只能生成看似完整的答案。
判断7	私有部署不自动等于安全、低成本或高成熟。成熟技术架构必须可连接、可替换、可观察、可评测、可降级，并可核算总成本。
判断8	治理的目标是形成与风险匹配的授权半径，而非所有动作一律人工审批或追求零风险。没有日志、权限、暂停、回滚和人工责任的高自动化应被风险否决。
判断9	多触点归因不能单独证明增量，节省工时不能直接计入现金利润。绩效证据必须连接活动、行为、机会、增量结果、长期资产和风险。
判断10	在缺少代表性样本时，可信基准只能是同群校准与公开证据锚点。跨行业总榜会激励指标游戏，并掩盖监管强度、渠道结构和数据基础差异。

董事会一页结论

董事会问题	建议决议
现在处于什么级别？	以八维证据附件评分，应用封顶与否决规则后给出当前证据级别及置信度。
第一笔钱投哪里？	投向能在90天内形成端到端小闭环且最短板可修复的场景，不投向仅展示生成能力的孤立工具。
如何拨款？	按受控试点、闭环验证、复制扩展三道门分段拨付，每一门绑定证据、责任人和停止条件。
谁负责？	业务 owner 对经营结果负责；CMO 对场景组合负责；CIO/CDO 对数据与技术可运行性负责；法务治理对授权边界负责；CFO 对增量证据和资本效率负责。
何时停止？	连续两个验证周期未形成可接受的质量、风险或经营证据；关键数据长期不可用；责任无人承担；总成本超过预算上限且无能力沉淀。

第1章 成熟度的经营定义与证据观

竹势 AI 营销智库出品

E1 存在 能力是否具备
E2 运行 是否稳定受控
E3 结果 目标是否改善
E4 增量 改善能否归因

证据层级决定扩张额度，展示性材料不能兑换经营授权。

1.1 定义：稳定运行能力，而非技术展示

AI 营销成熟度，是企业围绕明确经营目标，将 AI 嵌入营销决策与执行，在责任清晰、数据可用、流程可控、技术可运行、风险可接受的条件下，持续产生可验证结果并把经验转化为可复制能力的程度。定义包含六个不可拆分的词：经营目标、稳定运行、可审计证据、责任主体、控制边界、持续学习。

Deming 在质量管理体系中强调系统与反馈回路，PDSA 不是一次验收，而是计划、执行、研究、调整的连续循环。Drucker 的目标管理强调目标必须转化为责任和绩效。两者在 AI 营销中的共同含义是：模型输出只是过程变量，只有当输入、动作、验收、反馈和责任形成闭环，企业才获得可管理的生产系统。

1.2 证据四层：事实、运行、结果、因果

证据层	回答的问题	典型证据	不能证明什么
E1 存在证据	能力是否存在？	合同、配置、系统截图、培训记录、用例清单	不能证明稳定使用和业务价值
E2 运行证据	是否按规则持续运行？	日志、任务记录、审批记录、异常与回滚记录、SLA	不能证明经营结果由 AI 造成
E3 结果证据	是否出现目标改善？	前后基线、质量指标、转化、成本、客户体验、风险损失	若无对照，不能排除季节、渠道与团队变化
E4 增量证据	改善是否可归因？	随机实验、准实验、地域/人群对照、增量测试、MMM与财务核对	不能自动外推到其他任务、行业和时间

Noy 与 Zhang 的受控实验显示，使用 ChatGPT 的专业写作任务平均用时下降约40%，质量提高约18%；Brynjolfsson、Li 与 Raymond 对客服人员的现场研究发现，该研究覆盖5,179名客服人员，按每小时解决问题数衡量，AI辅助使平均生产率提高14%，新手和低技能人员提高34%，对资深高技能人员影响很小；BCG 咨询顾问实验则显示，在技术能力前沿内，使用 GPT-4 的参与者完成更多任务、更快且质量更高，但在前沿之外更容易给出错误答案。这些研究支持“任务级能力真实存在”，却不支持“企业整体已成熟”。

1.3 高采用、低成熟的五种结构

表象	结构性断点	成熟度处理
大量员工日常使用	无统一业务目标与责任	战略、组织封顶
内容产量大增	没有发布验收与转化连接	流程、绩效封顶
自动化动作很多	无日志、暂停与回滚	治理风险否决
采购私有平台	数据状态不可用、接口不通	数据、技术不达标
展示多个成功案例	口径自报、无基线或不可复核	证据置信度下调

BOARD BRIEF 04

第2章 八维依赖地图与系统判定规则

竹势 AI 营销智库出品

经营闭环
价值 × 责任 × 控制

战略

组织

场景

流程

数据

技术

治理

绩效

整体级别受关键依赖和最短板约束；单维领先不代表系统成熟。

Goldratt 的约束理论指出，系统产出由关键约束决定，而非由各部分平均能力决定。组织互补性研究同样表明，技术、流程、技能与激励需要成套变化。基于此，本模型拒绝简单平均：一个治理失控、数据不可用的系统，不会因为战略表达漂亮或内容生成速度快而被判为高成熟。

2.1 八维依赖地图

维度	核心问题	主要上游依赖	主要下游影响
战略	为何做、做什么、不做什么？	董事会目标、资本约束	场景组合、预算、退出
组织	谁负责、谁决策、谁复核？	战略、岗位与治理	流程复制、异常处置
场景	在哪个业务闭环创造价值？	战略、数据可得性	流程、绩效
流程	如何输入、验收、异常、反馈？	场景、组织	数据状态、治理日志
数据	业务事实是否可定义、可访问、可追溯？	流程、法律权限	技术运行、绩效测量
技术	是否可连接、评测、替换、降级？	数据、架构标准	流程稳定、成本
治理	授权半径是否与风险匹配？	组织、法律、技术	可扩展边界、事故损失
绩效	价值与风险是否被验证？	战略、场景、数据	继续投资、停止、复制

2.2 系统判定：四类规则

规则	判定方法	设计理由
证据门	每级必须满足该级“必须证据”；口头陈述不计分	抑制乐观偏差与展示性材料
关键约束	整体级别不高于战略、组织、流程、数据、治理、绩效六项中的第二低值；场景与技术可局部领先	避免一个偶发低分过度惩罚，同时保留短板约束
依赖封顶	若无经营 owner、基线和结果证据，最高 L2；若数据或流程低于 L3，场景复制最高 L3	复制依赖稳定状态与责任闭环
风险否决	高自动化但无日志、权限、暂停、回滚、人工责任，整体不得超过 L2，并标记“高暴露”	自主执行扩大事故半径，不能以效率抵消控制缺失

2.3 行业差异：不改底线，调整证据强度

行业差异不应通过随意加权把监管要求稀释。高监管行业提高治理、数据和结果证据门；强品牌消费行业提高内容真实性、品牌一致性和客户体验门；B2B 长周期业务提高商机阶段、销售协同和长周期增量证据门；平台型业务提高实时性、规模稳定性与供应商边界门。核心底线不变：责任、日志、权限、暂停、回滚、结果基线均不能被行业权重抵消。

BOARD BRIEF 05

第3章 五级证据阶梯

竹势 AI 营销智库出品

L1

无主试用

只有零散能力

L2

受控试点

假设与边界明确

L3

闭环运营

结果可以复核

L4

可复制组合

契约与治理复用

L5

适应性经营

证据持续校准

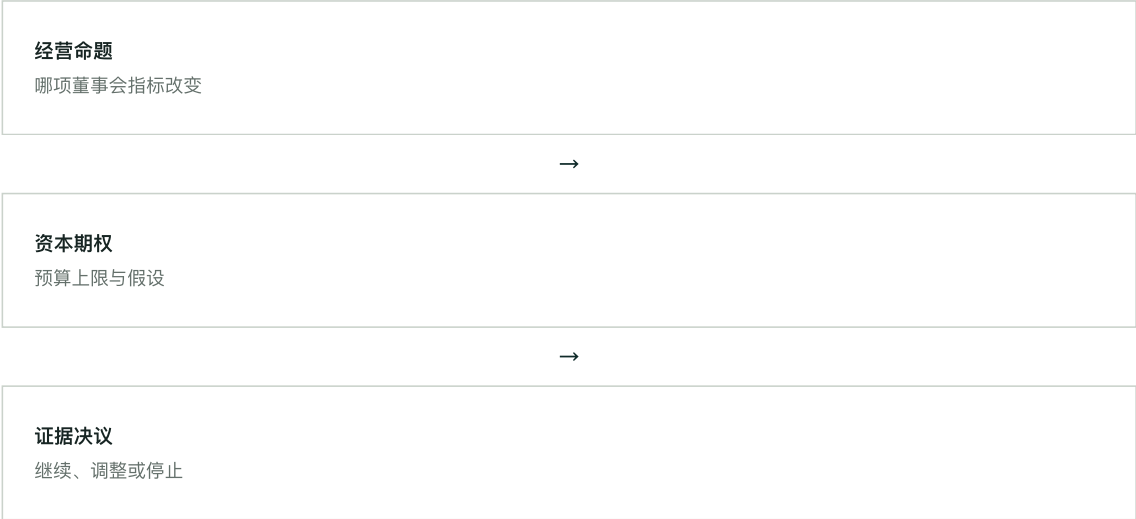
级别	经营目标	典型行为	必须证据	关键责任	允许执行半径	常见假象	升级门	停止/降级信号
L1 无主试用	局部人员尝试降低个人工作量	个人账号、零散内容、无法复盘	工具清单与零散样例	个人使用者	仅生成草稿，不接外部动作	把登录人数当成熟	指定业务 owner、选定场景、建立基线	发生外发事故、敏感数据进入无授权工具
L2 受控试点	在预算和风险边界内验证一个假设	明确范围、人工复核、周期性评估	试点章程、基线、日志、验收记录、停止条件	业务 owner + 项目负责人	低风险动作可自动，高影响动作逐项审批	试点热闹但无结果基线	连续两个周期达到质量、运行与结果门	证据缺失、成本失控、责任悬空
L3 闭环运营	一个端到端场景稳定运行并产生可验证结果	标准输入、异常处理、数据回写、绩效复盘	SOP、权限、日志、回滚、结果看板、财务核对	业务 owner + 跨职能责任组	规则内动作可自动，例外升级人工	把流程自动化率当价值	稳定性、增量证据和风险指标达标	客户体验恶化、异常积压、数据质量下降
L4 可复制组合	跨团队复制多个闭环并共享平台能力	联邦治理、标准接口、场景组合管理	复制包、接口契约、统一评测、组合预算、审计报告	高管赞助人 + 场景组合委员会	中风险动作在阈值内自主，持续抽检	规模扩张掩盖边际收益下降	复制成功率、总成本和组合回报达标	供应商锁定、系统耦合、风险集中
L5 适应性经营	AI 能力随市场变化持续校准并纳入经营系统	动态授权、持续实验、模型与供应商可替换、组织学习	持续评测、漂移监测、情景演练、董事会定期决议	董事会 + 经营层共同负责	基于实时风险与绩效动态调整半径	把“领先”当永久身份	持续保持证据有效性并修正边界	目标失真、指标游戏化、控制失效或外部规则变化

3.1 每一维的可审计描述

维度	L1-L5审计主线
战略	L1无目标；L2单场景假设、预算上限和决议日；L3场景与经营指标绑定；L4组合取舍和资本配置；L5动态调整战略与授权。
组织	L1个人自发；L2业务 owner 明确；L3跨职能 RACI 和异常责任；L4联邦式平台与业务接口；L5岗位、激励与能力持续重构。
场景	L1零散任务；L2单一可测场景；L3端到端小闭环；L4可复制场景组合；L5基于价值和风险动态编排。
流程	L1无标准；L2输入与人工验收；L3异常、回滚和反馈完整；L4跨团队接口契约；L5双环学习与持续优化。
数据	L1个人上传；L2授权数据集与基本质量检查；L3定义、质量、权限、时效、血缘可用；L4数据产品与共享契约；L5实时质量与用途治理。
技术	L1单工具；L2受控环境和基本评测；L3连接、观察、降级、成本可控；L4平台化、可替换、统一评测；L5动态路由、韧性与退出能力。
治理	L1无规则；L2分类审批与禁区；L3日志、权限、暂停、回滚、事故责任；L4分层授权和供应商治理；L5持续风险感知与演练。
绩效	L1活动量；L2基线与过程指标；L3结果与财务核对；L4增量测量和组合回报；L5长期资产、风险调整价值与动态资本配置。

第4章 战略维度：取舍、期权与决议日

竹势 AI 营销智库出品



Porter 的战略本质是取舍：选择独特价值与活动系统，也明确不做什么。AI 项目最常见的战略失真，是把“所有部门都试一试”当成战略，把供应商能力清单当成路线图。成熟战略必须把用例绑定经营问题、资源上限、替代方案和退出条件。

实物期权提供另一层修正：探索期的信息价值可能高于短期现金回报，因此不应以成熟业务的利润率过早扼杀试验；但期权的价值来自“有限下行、保留上行”，必须设置分段投入、到期决议和可放弃性。没有决议日的探索不是期权，而是拖延。

战略证据项	合格标准	审计问题
经营命题	明确收入、成本、现金周期、客户体验或风险中的主问题	若项目成功，哪一项董事会指标会改变？
场景取舍	列出优先场景与明确不做清单	为什么不先做更热门的用例？
预算与期权	阶段上限、释放条件、沉没成本控制	下一笔预算由什么证据触发？
退出条件	质量、风险、成本、数据、责任五类停止条件	什么情况下管理层会主动停止？
决议日	固定下一次继续/调整/停止日期	谁准备证据，谁有决定权？

中国平台型企业的案例说明，战略证据必须区分“能力覆盖”“采用规模”和“经营增量”。淘宝天猫在2024年11.11后披露，其AI电商工具包自当年10月以来服务约400万商家，生成超过1亿份图片、视频和文本；全站推广工具覆盖近29万商家、涉及超过160万件商品。该披露证明平台具备大规模供给与商家采用能力，也提供了部分商品销售增长的公司口径，但没有给出代表性对照、商家成本全口径或增量利润，因此不能把调用量、生成量

或覆盖商家数直接写成ROI。对董事会而言，这一案例限制了“先铺开、后找价值”的战略：规模数据可以触发下一阶段评估，不能替代场景取舍与资本回收证据。

战略维度的证据门应当形成一份“投资命题卡”：经营问题、目标客户或业务单元、当前基线、首个闭环、可承担最大损失、预算释放条件、供应商退出路径和下一次决议日必须同时存在。对于中国企业常见的平台依赖场景，还要说明流量规则、接口权限或平台政策变化是否会使原假设失效。探索期允许采用过程领先指标，但这些指标只能证明学习速度，不能提前确认财务价值。

当代管理研究对这一设计提供了两项约束。Raffaella Sadun长期研究管理实践、组织质量与技术生产率，核心结论是技术收益依赖管理互补性；本报告据此要求战略决议同时检查组织和流程能力。Ethan Mollick基于多项生成式AI实验强调能力边界呈不规则分布，适合一项任务的技术在相邻任务上可能失灵；本报告因此把“任务级评测”和“决议日重新评估”写入战略证据，而不把一次成功固化为长期承诺。

本章案例如何证明或限制判断：淘宝天猫说明大规模采用可以很快形成，但公开覆盖量无法替代可归因利润；P&G和BCG实验说明特定任务存在真实收益，却不能证明所有活动都应推进。战略成熟的可观察标志，是管理层能根据证据放弃热门但低闭环的项目，而不是项目清单持续增加。

第5章 组织维度：责任、决策权与接口

竹势 AI 营销智库出品

联邦式责任，不是模糊共治

中央团队守标准、平台与高风险底线；业务 owner 对真实流程与经营结果负责；法务、数据和财务分别持有明确否决或验收权。

Galbraith 的星型模型把战略、结构、流程、奖励与人员看作相互匹配的设计变量；社会技术系统理论强调技术与社会安排必须联合优化。AI 营销因此不能只交给技术团队，也不能只交给市场部个人。技术团队掌握平台、数据与控制，业务团队掌握客户、流程与结果；两者之间必须有可执行接口。

组织模式	适用情形	主要优势	主要风险	成熟做法
集中式 AI 中心	早期能力稀缺、高监管、需统一平台	标准快、风险集中管理	远离业务、排队慢	中心负责标准与平台，业务 owner 对结果负责
业务嵌入式	场景变化快、业务专业性强	反馈快、责任贴近结果	重复建设、治理不一致	共享底座与统一审计，业务可自主配置
联邦式	多业务单元、已形成多个闭环	兼顾规模与速度	接口复杂、权责模糊	中央制定不可妥协底线，业务持有场景 P&L

P&G 与研究团队在2024年5月至7月开展的新产品创新现场实验，为组织设计提供重要边界：AI 能让个人达到无 AI 团队的部分表现，并帮助商业与研发人员提出更平衡的方案；但实验任务有明确时间、范围和评审，不能据此推导跨职能团队可以被整体取消。管理含义是重新设计协作单元与接口，而不是简单裁撤角色。

组织成熟度首先检查三种权力是否落到具名角色：业务结果权、技术运行权和风险否决权。业务负责人决定场景目标和例外处理；CIO/CDO或平台负责人保证连接、观察和降级；法务、数据与治理负责人对个人信息、内容标识、客户权益和高影响动作拥有停止权。三者不能由一个“AI项目经理”笼统代替，也不能在事故发生后才临时分配责任。

P&G的预注册现场实验把组织问题具体化。同行评审论文口径为776名专业人员，部分外部报道采用791人，版本差异应保留并标注。实验显示，AI可以降低商业与研发专业边界，使个人方案更接近团队表现，但任务有共同目标、受控时间、明确评审和统一工具环境。它证明的是协作单元可以重新设计，不是证明跨职能责任可以被取消。复制到企业时，必须保留需求定义、事实核验、品牌判断和资源承诺的责任接口。

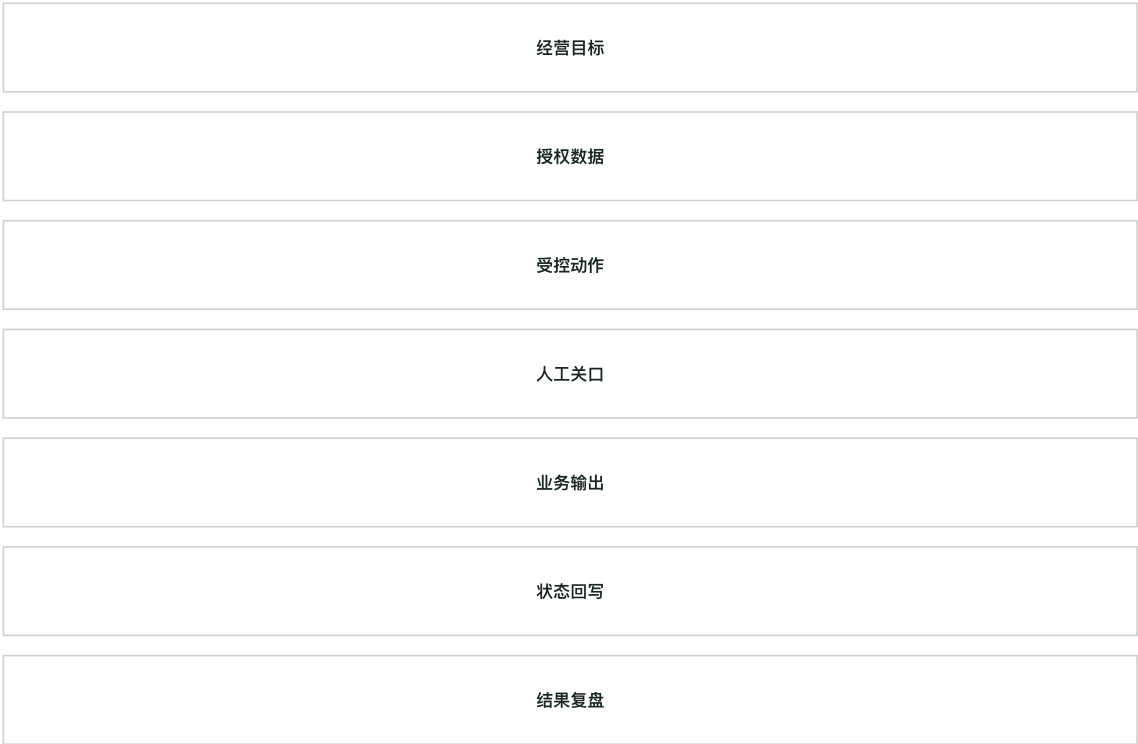
Karim Lakhani长期研究数字创新、知识生产与人机协作，并参与P&G实验。他的研究基础支持把AI视为改变知识分工的组织变量。本报告的修正：知识生成速度上升后，决策权、审核权和异常责任必须更清楚，否则信息吞吐增加只会把瓶颈推向审批与协调。Erik Brynjolfsson基于5,179名客服人员的现场数据发现，收益集中在新手和低技能人员；因此组织升级不能只培训“所有人会用”，还要按任务熟练度配置辅导、抽检和升级路径。

中国集团型企业还要处理总部与业务单元之间的双重治理。总部适合维护共用平台、供应商清单、数据底线、评测集和事故响应；品牌、区域、电商、经销或事业群持有场景目标和客户结果。若总部要求所有内容、客户回复和预算动作逐项审批，系统会被控制成本拖垮；若业务单元可自行采购、接入数据和对外执行，又会形成重复建设与不可审计风险。成熟做法是中央规定不可妥协底线，业务在风险阈值内拥有配置权。

本章案例如何证明或限制判断：P&G证明AI能够跨越部分专业边界，但不消除任务定义与最终责任；NBER客服研究证明人员差异会显著改变收益。组织成熟不能以裁减岗位数衡量，而应检查责任是否落地、接口是否缩短、异常是否有人处理、经验是否沉淀为共享能力。

第6章 场景与流程：小闭环、标准运行与学习

竹势 AI 营销智库出品



缺少回写与复盘的生成任务，只是局部产能，不构成成熟闭环。

场景选择采用五因素矩阵：价值、频次、数据可得、风险、闭环性。Jobs-to-be-Done 帮助识别客户或内部角色真正要完成的任务；价值链视角要求把任务放回前后环节。高频、可测、风险可控且能回写结果的场景优先。

场景类型	价值	频次	数据	风险	闭环性	建议
内容草稿生成	中	高	易	低至中	低	适合作为入口，不应单独代表成熟
投放创意—实验—回流	高	高	中	中	高	优先闭环，要求平台数据和增量设计
线索评分—跟进—商机更新	高	高	中	中至高	高	B2B优先，需CRM状态与人工升级
客户自动回复	中至高	高	中	高	中	先限定意图和赔付边界
品牌危机自动处置	高	低	难	极高	低	仅做监测、建议与人工指挥

Deming 的 PDSA 适合稳定运营：明确标准、运行、检查偏差、调整。Argyris 与 Schön 的双环学习用于更深层修正：不仅调整动作，还要检验目标、规则和假设是否错误。创意探索不宜被过度标准化，成熟流程应区分两条轨道：探索轨允许多样性和低成本失败，运营轨强调可重复、合规和异常处理；两者通过评审门连接。

流程控制点	稳定运营轨	探索创意轨
输入	结构化字段、版本与责任人	允许模糊问题，但记录假设
验收	阈值、清单、自动检查与抽检	多方案评审、专家判断
异常	明确升级、暂停、回滚	失败可接受，但控制数据与外发
反馈	结果回写并更新规则	形成新假设与可进入运营的候选
发布	按风险自动或人工批准	未经评审不得直接规模发布

端到端小闭环至少包含六个可观察状态：经营需求进入、AI执行、人工或规则验收、外部动作或内部交付、结果回写、复盘调整。缺少其中任何一环，企业得到的都只是生产环节提速。以内容营销为例，生成稿件并不构成闭环；发布审批、合成内容标识、渠道版本、落地页或线索来源、销售跟进状态与效果复盘必须能够连接。

联想在2024年公开Lenovo Studio AI案例，为B2B内容流程提供了具名企业样本。其业务情境是产品资料、行业化销售材料和演示文档数量庞大，人工制作与外部代理造成较长交付周期，已有数字资产管理平台却只有少部分资产被销售使用。联想将企业内容、检索增强生成、版式编辑和人工发布审核连接为流程；官方披露超过11,500名销售人员和合作伙伴使用该平台，新产品资产制作速度最高达到原流程的8倍，并称内容制作最高可加快90%。这些结果由联想官方案例测量和披露，未提供随机对照、质量缺陷分布或完整成本，因此应解释为流程效率与采用证据，不能直接折算为增量收入。

联想案例的关键机制不是“生成更快”，而是把已有高质量资产、受众属性、编辑界面和人工批准连接起来。可复制条件包括可检索的合规内容库、产品事实版本控制、销售或营销人员能够修改、外发前具名审核以及对敏感信息的保护。边界同样清晰：材料更快交付不等于客户购买决策改善，企业仍需验证销售使用率、内容准确性、机会推进和客户反馈。

流程控制需要区分稳定任务与探索任务。稳定轨对输入字段、事实来源、发布权限和异常阈值做强约束；探索轨允许多方案、反常识假设和专业争论，但不得绕过数据禁区与外发控制。双环学习发生在复盘阶段：若连续优化文案仍不改变转化，管理层应检查受众、价值主张、渠道或产品条件，而不是继续增加生成次数。

本章案例如何证明或限制判断：联想证明当AI嵌入资产、编辑和审核流程时，制作周期可显著缩短；Noy与Zhang的444人随机实验和NBER客服现场数据证明明确任务中存在速度收益。三者都没有证明单点速度自然转化为营销增量，所以场景升级门必须包含下游状态和客户结果。

第7章 数据与技术：业务可用、可连接、可替换

竹势 AI 营销智库出品

数据可用 定义 · 质量 · 权限 · 时效 · 血缘 · 业务状态
技术可控 连接 · 替换 · 观察 · 评测 · 降级 · 总成本

私有部署或集中数据都不能绕过业务可用性与运行控制的验收。

数据成熟度不是把更多数据集中到湖仓，而是让业务对象具备定义、质量、权限、时效、血缘和状态。营销系统尤其容易出现“内容有了、客户状态没有”“平台报表有了、活动口径不一致”“线索记录有了、是否有效无人确认”的问题。

数据验收门	最低证据	失败后果
定义	指标字典、业务对象和状态机	同名指标含义不同，无法比较
质量	完整性、准确性、重复、异常规则与责任人	自动化放大脏数据
权限	用途、角色、最小必要访问与撤销	个人信息与商业敏感数据越界
时效	更新时间、延迟阈值、失效标记	系统基于过期状态行动
血缘	来源、变换、版本与输出去向	无法复核和追责
业务状态	线索、内容、活动、客户、预算的当前阶段	AI只生成文本，不能推动经营

技术成熟度评价“可连接、可替换、可观察、可评测、可降级、总成本可控”。NIST AI RMF 强调治理、映射、测量和管理；ISO/IEC 42001要求组织建立并持续改进 AI 管理体系。两者都指向生命周期控制，而非单次技术选型。私有部署只改变部分数据与基础设施边界，仍需要身份权限、密钥管理、日志、补丁、模型评测、供应链和运维能力。

技术验收门	证据	董事会关注
连接	API/连接器、失败重试、数据回写	能否进入真实流程
替换	模型与供应商抽象层、退出方案	是否被单一供应商锁定
观察	调用、成本、延迟、错误、质量日志	问题是否可发现
评测	离线集、在线抽检、业务指标	版本变化是否可控
降级	人工接管、规则引擎、备用模型、停机	故障是否扩大
总成本	许可、调用、集成、运维、审计、人员	单次便宜是否掩盖全生命周期成本

中国企业的断点往往发生在公域平台、私域触点和内部经营系统之间。广告与内容平台保留曝光、互动和转化信号，企业微信、门店或经销体系持有客户关系，CRM记录商机阶段，财务系统确认收入和退款。平台账户、手机号、OpenID、设备标识与企业客户编码不能当然互通；个人信息处理还需要目的、最小必要、授权与保存期限。数据成熟的证据不是“接入了多少平台”，而是每一个关键状态能否合法、稳定地关联，并说明未能关联造成的归因偏差。

联想Studio AI显示，成熟技术架构需要把检索、生成、编辑、敏感信息保护和人工审核组合起来。已有数字资产管理平台并未自动带来高利用率，说明数据集中化与业务可用之间存在检索、语义、权限和工作流鸿沟。技术验收应检查内容版本、引用来源、权限隔离、外发审批和回写记录，而不能只检查平台是否上线。

中国平安的公开披露提供了高监管、超大规模运行案例。平安2025年年度业绩材料称，AI服务代表在2025年提供约17.02亿次服务，占集团客服总量约80%；公司还披露AI智能体协助实现人民币1,331.79亿元销售。相关数字由平安公司统计并在年度材料中披露，证明其AI已进入大规模服务和销售流程，但披露没有提供随机对照，也未将销售额拆分为AI的增量贡献，不能将“协助实现的销售额”视为AI创造的收入。

平安案例对数据与技术成熟的管理含义，是规模运行必须建立统一客户状态、业务规则、身份权限和人工接管。金融产品条款、适当性、承保、理赔和续期状态都具有时点与版本属性；如果系统只读取静态文档而不能读取当前业务状态，就可能在回答正确知识的同时作出错误行动。可复制条件包括受监管数据域、分级授权、规则与模型协同、全链路日志和人工服务通道。财务结果、客户体验与合规事件仍需分别验证。

技术产品视角中，Stuart Russell关于可控AI的研究强调系统目标必须与人的真实偏好和控制权保持一致；本报告把这一理论转化为可执行要求：高影响系统必须支持停止、回滚、人工接管和供应商退出。Fei-Fei Li以人为本的AI主张提醒企业，技术效率不能以牺牲客户尊严、员工判断或弱势群体权益为代价；因此数据与技术验收还应包含可访问性、偏差监测和申诉路径。

本章案例如何证明或限制判断：联想说明检索与工作流使存量资产变得可用，平安说明AI可以进入大规模业务状态；两者均为企业自报，缺少完整因果识别。它们支持“技术必须连接真实系统”，同时限制了把用户数、服务量或协助销售额直接当作成熟度与ROI的做法。

第8章 治理：按动作风险配置授权半径

竹势 AI 营销智库出品

R1 内部可逆 自动执行 / 持续抽检
R2 有限外部影响 阈值内执行 / 批量审批
R3 客户与预算 人工批准 / 双人复核
R4 权益或不可逆 AI 只建议 / 禁止自主

中国现行规则要求企业在生成式 AI、个人信息、数据安全、网络内容和生成合成内容标识等方面承担相应责任。《生成式人工智能服务管理暂行办法》明确服务提供者的信息内容、个人信息与安全义务；《人工智能生成合成内容标识办法》自2025年9月1日起施行，规定显式与隐式标识要求。NIST AI RMF 与 ISO/IEC 42001提供组织层面的风险管理与持续改进框架。

治理不应追求所有动作零风险。高可靠组织研究强调对失败保持敏感、尊重专业、保留韧性。成熟治理通过动作风险分级配置授权半径，让低风险、高频、可回滚动作自动运行，让高影响、不可逆、涉及权益或监管的动作由具名人员决策。

风险级	营销动作示例	允许半径	最低控制
R1 低影响可逆	内部摘要、标签、草稿、排期建议	可自动执行	日志、版本、抽检
R2 有限外部影响	常规社媒草稿发布、低金额预算建议	规则阈值内自动或批量审批	权限、品牌检查、暂停、回滚
R3 高影响可修复	客户回复、线索评分、预算调整、个性化推荐	人工批准或双人复核	证据引用、人工责任、事故流程
R4 权益/监管/不可逆	价格承诺、赔付、敏感人群决策、危机声明	AI仅建议，不得自主执行	法务/业务负责人批准、完整留痕、演练

治理红线	判定
无具名责任人	不得对外自主执行
无日志或日志不可检索	不得进入 L3
无暂停和回滚	高自动化触发风险否决
未授权个人信息或敏感数据	立即停止并启动事件处置
供应商责任边界不清	不得扩展到高影响场景
生成合成内容标识义务未落实	不得公开发布相关内容

Air Canada责任案给出治理的最低线：企业不能把自动系统描述成独立于企业的主体，也不能以“机器人自行生成”排除责任。对中国企业，自动回复、价格承诺、售后政策、促销规则和金融或医疗相关说明同样需要统一知识源、版本有效期、证据引用与具名责任人。机器人回答与网页、合同或线下政策不一致时，企业承担的并非技术准确率问题，而是客户权益与承诺管理问题。

中国的生成合成内容标识要求增加了内容链路的证据门。企业需要确定自己在服务提供、内容制作、传播平台使用和委托供应商中的角色，把显式标识、隐式标识、文件元数据、平台提示和发布记录纳入验收。集团多品牌或代理商体系中，品牌方不能仅在合同中笼统要求“合规”，而应约定谁添加标识、谁保留日志、谁处理平台退回、谁对未经授权的人像声音和个人信息负责。

个人信息与供应商责任是中国经营场景的另一条硬边界。营销系统常把客户画像、聊天记录、手机号、交易、位置或未成年人信息交由模型、云平台、内容工具和外包机构处理。治理证据应包括处理目的、字段清单、最小必要、访问角色、跨主体委托条款、删除与撤回机制、事件通知以及供应商停服或更换时的数据处置。私有环境若权限混乱、日志缺失或补丁滞后，风险仍然存在。

Arati Prabhakar在美国科技政策与工程管理领域长期推动创新、安全和权利并行；其政策环境与中国不同，但“治理不能等到规模化后补做”具有通用管理意义。本报告进一步要求治理动作与中国法律角色和业务责任对应。Stuart Russell的可控AI理论说明，保留人类控制不应依赖操作者临时发现错误，而应通过系统化暂停、回滚和授权设计实现。

本章案例如何证明或限制判断：Air Canada证明责任随企业服务而存在；中国标识制度明确了生成合成内容的流程义务；平安的大规模高监管应用说明治理并不必然阻止规模，但企业披露的服务量不能替代事故、投诉、误导和人工升级率。治理成熟的标准是风险被发现、限制和恢复的能力，而不是“尚未发生公开事故”。

第9章 绩效：从活动到增量价值与风险

竹势 AI 营销智库出品

风险调整经营结果 增量利润 · 客户价值 · 长期资产 · 事故损失
机会与行为 商机 · 转化 · 留存 · 销售采用
活动与运行 内容 · 触达 · 周期 · 质量 · 成本

归因分配功劳，增量测量判断是否创造了新的结果。

Kaplan 与 Norton 的平衡计分卡提醒管理者同时观察财务、客户、内部流程与学习成长；营销生产率研究强调营销投入与财务结果之间需要可解释连接。Goodhart 定律警告：当指标成为目标，指标就可能失去测量价值。AI 营销尤其容易把产量、速度、调用量和节省工时变成绩效终点。

绩效证据树	示例	财务处理
活动	内容数、实验数、触达数	只作投入与运行信息
行为	点击、停留、回复、销售采用	需要排除机器人流量和口径变化
机会	有效线索、商机阶段、复购意向	需与CRM状态和人工确认连接
增量结果	增量收入、转化、留存、成本避免	优先随机/准实验；否则披露假设
长期资产	品牌搜索、第一方数据、知识产权、流程能力	不能全部当期确认收益
风险	投诉、错误承诺、偏差、泄露、停机、监管事件	纳入风险调整回报和停止条件

多触点归因分配“已发生转化”的功劳，但不能单独证明营销活动创造了增量。节省工时只有在人员成本实际减少、产能用于可计价工作或避免新增招聘时，才可能转化为现金效益；否则应记录为能力释放，不应直接计入利润。成熟绩效体系将财务核对、客户体验和风险损失一起纳入。

绩效证明需要把“谁测量”写入证据附件。公司财报、业务系统、供应商案例、平台归因和独立实验的可信边界不同。淘宝天猫披露的商家覆盖、内容生成量、流量分析次数与商品销售增长属于平台自报；联想的8倍速度与11,500名使用者属于企业案例口径；平安的服务量和协助销售额属于集团年度披露。三者都具有经营相关性，但没有统一对照设计，不能被合并为行业平均收益，也不能彼此比较高低。

Klarna的连续披露进一步说明单一效率指标会反噬经营判断。2024年公司强调AI客服处理量、解决速度与成本收益，随后管理层公开承认过度关注成本并恢复更多人工服务选项。该变化不否定初期效率，而是揭示客户关系、复杂问题和品牌体验未被充分计入目标函数。Sebastian Siemiatkowski的经营表述来自公司实践，具有管理参考价值，也带有企业战略传播属性；本报告将其作为停止条件设计的证据，不作为独立因果证明。

绩效树必须防止三个会计误区。第一，节省工时若未减少支出、避免招聘或转移到可计价产出，只能记为能力释放。第二，AI参与的销售额不等于AI增量收入，需要对照、转化漏斗与财务核对。第三，内容或客服量增加可能带来投诉、退货、品牌稀释、人工返工和供应商成本，必须纳入风险调整回报。CFO应要求每项收益同时列出计算公式、数据源、测量责任、时间窗口和反事实。

本章案例如何证明或限制判断：Klarna证明效率与体验可能方向相反；淘宝天猫证明采用与运行规模可被量化，却不能直接证明商家增量利润；平安证明AI可深度参与服务与销售，但“协助实现”不等同因果贡献。绩效成熟的升级门不是数字更大，而是指标链能够被复核且不会因目标化而被游戏。

第10章 基准：同群校准、缺失值与置信边界

竹势 AI 营销智库出品

监管强度
企业规模
业务模式
渠道结构
数据基础
投资阶段

先与自身历史比较，再与条件相近的同群比较；公开锚点只证明机制可能，不生成伪行业排名。

“基准”在本报告中指一套可重复的同群比较规则，不指全国行业排名。企业先与自身历史基线比较，再与条件相近的同群比较，最后才引用公开证据锚点。公开锚点用于判断某种机制是否可能、某项控制是否必要，不能推导本企业应达到同样数字。

同群变量	分组建议	原因
监管强度	低/中/高	决定数据、内容、客户权益和审计门槛
企业规模	中小/大型/集团多业务	影响平台化、权限和复制复杂度
业务模式	B2C/B2B/平台/渠道/服务	决定周期、客单、触点和结果证据
渠道结构	公域投放/私域/线下/经销/混合	决定数据可得性与归因
数据基础	分散/可连接/数据产品化	决定闭环速度
AI投资阶段	试用/试点/规模化/再治理	避免把不同阶段混排

数据状态	评分处理	披露方式
有直接证据	按证据等级评分	列证据日期、责任人、有效期
有间接证据	暂按低一档或标中置信	说明代理变量和偏差方向
证据缺失	不以零替代事实；该项记“未证实”	整体置信度下降，关键门缺失则封顶
口径不一致	不强行换算	并列展示并说明不可比
过期证据	超过有效期后降级	要求重新验证

建议输出“级别 + 置信度 + 关键缺口”，例如：L2 受控试点（中置信），原因是流程和治理证据完整，但增量结果仅有前后对比，数据血缘未完成。不得只输出一个漂亮总分。

中国企业构建同群时，应增加四类经常被全球通用框架忽略的变量。其一是渠道权力：企业是否高度依赖单一电商、短视频、搜索或社交平台，接口、流量规则和归因数据是否由平台控制。其二是公域与私域断点：内容曝光、直播互动、企业微信、门店成交、经销商回款和CRM商机是否能合法关联。其三是集团复杂度：多法人、多品牌、多事业群、区域公司和代理商是否共享数据与审核规则。其四是内容与个人信息责任：生成合成标识、肖像声音授权、个人信息委托处理及供应商再委托是否影响可执行半径。

同群校准不能把这些差异折成一个“行业修正系数”。例如，同样是L3闭环运营，直营电商企业可能要求订单与内容实验直接连接；经销体系企业可能只能先证明线索分发、经销商确认和回款匹配；金融企业还需证明客户适当性、人工复核和审计留痕。证据门相同地要求责任与可复核，但结果链和时间窗口不同。

基准表至少披露六项元数据：证据截止日期、业务范围、同群变量、指标定义、缺失值处理、置信等级。企业自评若缺少附件，应标记“管理层陈述，未验证”；供应商演示应标记“能力存在证据”；平台或公司公开案例应标记“外部锚点，不可外推”。缺失值不得用同群均值填充，因为在没有代表性样本时，这种填充会制造不存在的精确性。

Goodhart定律在成熟度管理中表现为“为了升等级而生产证据”：人为增加运行次数、把审批拆成日志、选择容易改善的指标、隐瞒失败场景，都会让评分失真。因此每次校准应抽查原始业务记录、失败样本和停止决议，并允许成熟度下降。级别下降并不等于管理失败；当法律、平台政策、供应商版本或业务模式变化时，主动降级和重新验证反而是高成熟行为。

本章案例如何证明或限制判断：淘宝天猫、联想和平安分别代表平台生态、全球B2B科技企业和高监管金融集团，公开指标口径完全不同。它们可以作为机制锚点，不能组成中国企业排行榜。P&G 776人与外部报道791人的版本差异也说明，即使同一研究，样本版本和发布日期仍需注明；基准可信度来自披露差异，而不是消除差异。

第11章 案例卷宗与专家判断

竹势 AI 营销智库出品

案例只在说明完整机制时成立

主体与时间 · 业务情境 · 具体动作 · 作用机制 · 公开结果 · 测量主体 · 归因限制 · 适用边界

案例1：P&G 新产品创新现场实验（2024）

主体与时间：宝洁公司专业人员；实验在2024年5月至7月开展，研究后于2025年形成工作论文，2026年发表于 Organization Science。

业务情境：商业与研发人员围绕真实的新产品创新挑战工作，随机分配为个人或双人团队，并有无生成式 AI 条件。

具体动作：参与者在限定任务、时间和评审机制下使用 AI 形成产品方案。

公开结果：研究报告显示，个人使用 AI 的表现可匹配无 AI 团队的部分表现；AI 还使不同职能提出的方案更平衡。样本版本存在776与后续报道791的差异，本报告采用同行评审论文与论文摘要中的776作为主口径，并提示版本差异。

归因限制：研究是预注册随机现场实验，内部效度较强；但任务集中于产品创新，不能外推为跨职能团队可整体取消。

可复制条件：明确任务边界、评审标准、数据权限和团队接口；把 AI 作为协作结构变量而非单纯个人工具。

案例2：BCG 咨询顾问“锯齿状前沿”实验（2023）

主体与时间：758名BCG顾问参与预注册实验。

业务情境：18项现实、复杂、知识密集任务，部分在 GPT-4 能力前沿内，另有前沿外任务。

具体动作：随机分配无 AI、使用 GPT-4、使用 GPT-4 加操作培训。

公开结果：前沿内使用 AI 的顾问平均完成更多任务、速度更快、质量更高；能力较低者收益更大。前沿外任务出现更高错误风险。

归因限制：实验任务和当时模型版本限定明显，不能直接转化为企业 ROI。

可复制条件：建立任务级评测、识别前沿、对前沿外任务保持人工验证。

案例3：Klarna 客服自动化与人类服务回补（2024—2026）

主体与时间：Klarna 于2024年2月发布 AI 客服助手成果，2025年公开强化人类服务选项，2026年继续在投资者材料中强调 AI 驱动的人效。

业务情境：全球支付与购物服务，高频、多语言客服。

具体动作：AI 助手处理大量聊天，企业同时调整人员与服务模式。

公开结果：公司自报首月完成230万次对话、约占客服聊天三分之二、解决时间由11分钟降至不足2分钟，并估算利润改善；这些是公司口径。其后管理层承认对成本关注过度、需要保留人类连接，说明效率与体验存在边界。

归因限制：公司自报指标缺少独立实验与完整成本口径；后续财务改善不能单独归因于客服 AI。

可复制条件：按意图和客户价值分层；保留人工通道；将客户体验、重复咨询、投诉和升级率纳入停止条件。

案例4：Air Canada 聊天机器人责任案（2024）

主体与时间：加拿大不列颠哥伦比亚民事解决法庭，Moffatt v. Air Canada, 2024 BCCRT 149。

业务情境：客户询问丧亲票价退费规则，聊天机器人提供与网页政策不一致的信息。

具体动作：客户依赖机器人答复购票并申请退费；企业拒绝后发生争议。

公开结果：法庭认定企业应对网站聊天机器人提供的信息负责，并判赔相关差额、利息与费用。

归因限制：这是法律责任案例，不证明所有聊天机器人都会失败。

可复制条件：统一知识库、政策版本控制、引用证据、对承诺性答复设置人工批准与赔付责任。

案例5：生成式 AI 客服现场研究（美国，2020—2021数据）

主体与时间：Brynjolfsson、Li 与 Raymond 对一家财富500强软件企业客服系统开展现场研究。

业务情境：客服人员处理真实客户问题，AI 提供实时建议。

具体动作：分阶段部署生成式 AI 辅助工具并比较人员表现。

公开结果：NBER工作论文主口径覆盖5,179名客服人员；按每小时解决问题数衡量，AI辅助使平均生产率提高14%，新手和低技能人员提高34%，对资深、高技能人员影响很小。研究同时观察到客户对客服人员的语言更积极、要求转人工主管的情况下降。

归因限制：单一企业与客服任务，结果不能外推到创意、战略或自动执行。

可复制条件：高质量知识库、实时工作流嵌入、人员分层培训和客户结果监测。

案例6：专业写作随机实验（2023）

主体与时间：Noy 与 Zhang, 444名具有大学教育背景的专业人员。

业务情境：撰写新闻稿、短报告、分析计划等中等复杂写作任务。

具体动作：随机获得 ChatGPT 使用权限。

公开结果：平均完成时间下降约40%，质量提高约18%，能力差距缩小。

归因限制：短时、明确写作任务；不包含品牌长期影响、法律审核和发布结果。

可复制条件：将生成能力嵌入验收流程，不把写作速度直接计入利润。

治理证据单元：中国生成合成内容标识制度落地（2025）

主体与时间：国家互联网信息办公室等部门发布《人工智能生成合成内容标识办法》，2025年9月1日起施行。

业务情境：文本、图片、音频、视频、虚拟场景等生成合成内容进入公开传播。

具体动作：要求服务提供者按规定添加显式和隐式标识，传播平台核验文件元数据并提示用户。

公开结果：这是合规制度而非经营效果案例；企业未落实标识可能形成发布与供应商治理风险。

归因限制：适用范围需结合企业角色、服务形态和具体场景判断。

可复制条件：把标识要求写入内容 workflow、素材元数据、发布验收和供应商合同。

案例7：淘宝天猫商家AI工具规模应用（2024）

主体与时间：阿里巴巴集团旗下淘宝天猫，2024年11.11大促后由集团官方披露商家AI工具运行情况。

业务情境：大量商家需要在大促周期完成商品素材、流量分析和广告推广，平台试图以AI工具降低制作和运营门槛。

具体动作：平台提供图片、视频和文本生成、流量分析及全站推广工具，并在商家经营链路中大规模部署。

公开结果：阿里巴巴官方称，AI电商工具包自2024年10月以来服务约400万商家，生成超过1亿份营销图片、视频和文本；超过80万商家完成200多万次流量分析；近29万商家通过全站推广工具使超过160万件商品获得销售增长。数字由平台测量并自报，未披露代表性商家样本、完整成本、增量对照或利润口径。

归因限制：采用量、内容生成量和平台归因的销售增长不能直接证明商家增量利润；大促流量、补贴、商品力和其他投放共同影响结果。

适用边界与可复制条件：适用于平台型、高频电商运营；企业应分别记录采用、运行、商品结果和利润证据，并保留平台规则变化、广告成本和商家差异的置信边界。

案例8：联想Studio AI营销与销售内容流程（2024）

主体与时间：联想集团，2024年5月由联想官方发布内部应用案例。

业务情境：B2B产品资料、行业化销售内容和演示文档数量庞大；人工制作或委托代理导致交付周期长，已有数字资产管理平台中的大量内容未被充分使用。

具体动作：联想建设Studio AI，将企业内容资产、检索增强生成、受众与行业信息、可视化编辑、敏感信息保护和人工发布审核连接起来，为销售与合作伙伴生成定制资料。

公开结果与测量主体：联想官方称超过11,500名销售人员和合作伙伴使用该平台，新产品资产制作速度最高为原流程的8倍，部分官方页面概括为内容制作最高加快90%。结果由联想自测并披露，财务增量结果未披露。

归因限制：没有随机对照、缺陷率分布、客户转化或完整总成本，速度提升不能直接折算为销售收入。

适用边界与可复制条件：适用于拥有大量可信内容资产和多行业销售需求的B2B企业；需要版本化事实库、权限、编辑能力、敏感信息保护和外发前人工批准。

案例9：中国平安AI客服与销售流程规模运行（2025）

主体与时间：中国平安保险（集团）股份有限公司，2025年度经营数据于2026年年度业绩与年报材料中披露。

业务情境：金融保险客户服务和销售流程规模大、产品规则复杂、权益与合规要求高，需要在效率、服务体验和风险控制之间平衡。

具体动作：平安将AI服务代表和AI智能体嵌入客服、销售、续期及相关业务流程，并以集团级技术、数据与风险体系支撑运行。

公开结果与测量主体：平安公司披露，2025年AI服务代表提供约17.02亿次服务，占集团客服总量约80%；AI智能体协助实现人民币1,331.79亿元销售。相关数字由公司业务系统统计并在年度材料中披露。

归因限制：“服务占比”证明运行规模，“协助实现销售额”证明流程参与，但没有随机对照和AI增量贡献拆分，不能将该销售额视为AI创造的收入；相关财务增量未单独披露。

适用边界与可复制条件：适用于具备统一客户身份、业务状态、规则体系、权限日志和人工服务通道的高监管集团；复制时必须保留适当性、人工复核、投诉申诉、暂停与事故责任。

11.3 当代专家与经营者的实质观点

人物	经验/研究基础与主张	适用边界	本报告判断
Erik Brynjolfsson	基于客服现场研究，AI 对经验较少人员的帮助更大。	适用于有结构化知识与反馈的任务；不能推导所有岗位都同样受益。	组织应按任务和人员分层设计培训与质量门。
Raffaella Sadun	组织实践与管理质量决定技术是否转化为生产率。	技术效果依赖组织互补性。	成熟度必须同时评价组织和流程。
Karim Lakhani	AI 改变知识生产与协作结构。	实验边界明确，不等于取消团队。	重新设计协作单元与接口。
Ethan Mollick	AI 能力边界呈锯齿状，用户需要持续实验与判断。	模型版本变化快。	建立任务级评测与人工复核。
Sebastian Siemiatkowski	Klarna 先强调效率，后承认需要人类连接和质量平衡。	公司自述含战略传播成分。	客户体验必须成为自动化停止条件。
Arati Prabhakar	作为美国科技政策领导者强调安全、权利与创新并行。	政策环境与中国不同。	董事会不能把治理留到规模化之后。
Fei-Fei Li	以人为本的 AI 强调技术服务人的目标和尊严。	原则需转化为流程控制。	对客户与员工影响纳入绩效。
Stuart Russell	可控 AI 需要系统目标与人类偏好保持一致。	理论层面不能替代现场控制。	高影响动作保留暂停、回滚和人工责任。

专家观点的使用原则是把研究或经营主张嵌入具体证据门，而不以声望替代事实。Brynjolfsson、Sadun、Lakhani和Mollick主要提供学术实证与组织管理视角；Siemiatkowski提供企业经营中的效率与体验修正；Prabhakar提供科技政策与监管治理视角；Fei-Fei Li和Russell分别从以人为本与可控性约束技术设计。每项观点均需与企业证据、法律义务和本企业现场结果交叉验证。

BOARD BRIEF 14

第12章 升级处方、90天计划与董事会工具

竹势 AI 营销智库出品

01-10

定位约束

owner、同群、基线

11-30

受控运行

日志、质量、异常

31-60

验证结果

对照、客户、成本

61-90

复制决议

继续、调整、停止

12.1 短板诊断与升级处方

最短板	典型症状	90天处方	升级证据
战略	项目多、无人说清经营问题	砍到1—2个场景；设预算上限、停止条件、决议日	董事会批准的试点章程
组织	市场、技术、法务互相等待	指定业务 owner；建立RACI与每周异常会	责任矩阵与决策记录
场景	只做内容生成，无结果回流	补齐发布、线索/实验、回写、复盘	端到端运行日志
流程	依赖个人技巧，异常无处置	定义输入、验收、例外、暂停、回滚	SOP与异常演练
数据	字段缺失、口径冲突、权限不清	建立最小数据产品与质量责任	数据字典、质量报告、权限记录
技术	单一供应商、成本不透明	增加观察、评测、降级和替换层	评测集、成本看板、退出测试
治理	统一人工审批或完全放开	按R1—R4配置授权半径	风险分级、日志和演练
绩效	只有产量与节省工时	建立基线、对照、财务核对和风险指标	增量测试或明确置信边界

升级顺序由关键约束决定。若内容生成质量尚可但无法连接线索和订单，继续增加内容工具不会提高系统产出；若数据与流程完整但所有动作逐项等待总部批准，组织和治理接口是瓶颈；若自动执行规模大而日志、暂停和人工责任缺失，应先降级执行半径。Goldratt的管理含义在这里非常具体：预算投向解除约束，而不是投向得分最高或展示效果最强的维度。

中国企业升级还需设计“平台变化情景”。电商、短视频、社交和广告平台可能调整接口、流量机制、内容标识、投放规则或数据返回口径。90天计划中应至少进行一次接口失效或供应商停服演练，验证人工接管、数据导出、替代工具和客户沟通。集团企业则应选择一个业务单元形成复制包，再用第二业务单元检验接口契约，而不是一次性全集团上线。

案例对升级顺序的启示是：联想先解决内容资产可用和审核流程，再扩大使用者；淘宝天猫先证明商家覆盖和运营规模，但商家仍需建立自己的利润证据；平安在高监管流程中实现大规模运行，但对外披露仍不足以证明每个场景的增量贡献；Klarna的经验要求把人工服务回补写入停止和修正条件。Stage–Gate的每一道门都应允许“继续、调整、暂停、停止”四种决议。

12.2 90天验证计划

阶段	时间	核心动作	证据门	停止条件
门0：立项	第1—10天	确定经营命题、owner、同群、基线、风险分级、预算上限	试点章程与证据清单	无owner、无数据合法路径、目标不可测
门1：受控运行	第11—30天	构建最小闭环；人工复核；记录日志、质量、成本、异常	连续两周稳定运行	错误率超阈值、客户或员工不可接受
门2：结果验证	第31—60天	开展对照或准实验；核对客户体验、机会与成本	结果证据与归因说明	无改善且无可修复机制
门3：复制决议	第61—90天	评估数据、流程、治理和总成本；形成复制包	董事会继续/调整/停止决议	边际收益下降、风险或锁定超上限

90天验证不追求完成全部平台建设，而是形成一组足以支持资本决议的证据。第1—10天固定口径和责任；第11—30天验证稳定运行与异常处理；第31—60天建立对照、准实验或至少可解释的前后基线；第61—90天验证复制、总成本、供应商退出和治理韧性。任何阶段若出现未授权个人信息处理、对客错误承诺、无法暂停的高影响动作或责任无人承担，应立即停止相关自动执行，而不是等待阶段结束。

12.3 角色行动清单

角色	90天必须完成
CEO	选择一个经营命题；任命业务 owner；确定下一次决议日。
CMO	管理场景组合；建立品牌、客户体验和增长验收门。
CFO	确认基线、现金口径、能力释放与利润的区别；实施分段拨款。
CIO/CDO	保证数据可用、系统可连接、可观察、可降级和可替换。
业务负责人	对真实流程、采用、异常和经营结果负责。
数据负责人	维护定义、质量、权限、时效、血缘和业务状态。
法务治理	完成风险分级、标识、个人信息、合同与事故责任设计。

12.4 董事会检查表与决议模板

- ☐ 是否存在具名经营 owner，且对结果而非工具上线负责？
- ☐ 是否有可复核基线、结果指标、风险指标与证据有效期？
- ☐ 八维中哪一维是当前关键约束？修复它是否比新增工具更有价值？
- ☐ 数据是否具备定义、质量、权限、时效、血缘和业务状态？
- ☐ 高自动化动作是否都有日志、暂停、回滚与人工责任？
- ☐ 供应商自报结果是否被清楚标注，是否存在独立验证？
- ☐ 同群比较是否控制监管、规模、业务模式、渠道、数据基础和投资阶段？
- ☐ 下一笔预算由什么证据触发？最大损失是多少？
- ☐ 停止条件是否具体到质量、风险、成本、数据和责任？
- ☐ 下一次董事会决议日是否已经写入日历？

董事会决议：本次批准/暂缓/停止 _____ 场景；批准预算上限 _____；经营 owner _____；风险级别 _____；必须提交的证据 _____；停止条件 _____；下一次决议日 _____。

下一次董事会决议日应当写明四类输入：八维证据变化、同群与外部锚点变化、法律和平台规则变化、场景组合的边际回报。决议材料不得只展示成功样本，还要提交失败记录、人工接管、投诉、数据缺失、供应商异常和停止动作。只有当管理层看到系统如何失败，才可能判断系统是否具备可复制的韧性。

BOARD BRIEF 15

结语：成熟度是一项可撤销、可复核的经营判断

竹势 AI 营销智库出品

最终输出不是一个漂亮总分

当前级别 + 证据置信度 + 关键约束 + 90天处方 + 下一次决议日

企业不需要追求八个维度整齐地“平均升一级”。真正有效的升级，是识别当前阻断价值闭环的关键约束，用90天验证修复是否有效，然后决定继续投资、调整设计或停止。成熟度的价值不在于获得一个分数，而在于把模糊的 AI 热潮转化为董事会可以审议的证据、责任、风险与资本选择。

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 ↑

纯文本来源索引

竹势 AI 营销智库出品

[S1] Noy, Shakked; Zhang, Whitney. Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. Science, 2023. 444名专业人员；写作任务随机实验；平均用时下降约40%，质量提高约18%。

[S2] Brynjolfsson, Erik; Li, Danielle; Raymond, Lindsey. Generative AI at Work. NBER Working Paper 31161, 2023. 5,179名客服人员；按每小时解决问题数衡量，平均生产率提高14%，新手和低技能人员提高34%，资深高技能人员影响很小。

[S3] Dell’Acqua, Fabrizio 等. Navigating the Jagged Technological Frontier. 2023. 758名BCG顾问预注册实验；18项现实咨询任务；能力前沿内速度、数量与质量改善，前沿外存在错误风险。

[S4] Dell’Acqua, Fabrizio 等. The Cybernetic Teammate: A Field Experiment on Generative AI and Teamwork. Organization Science, 2026；早期工作论文2025。宝洁专业人员随机现场实验，论文摘要主口径776人。

[S5] Klarna. Klarna AI Assistant Handles Two–Thirds of Customer Service Chats in Its First Month. 2024–02–27. 公司自报：230万次对话、约三分之二聊天、解决时间与重复咨询口径。

[S6] Klarna Group plc. 2025年度及2026年投资者材料。公司披露AI驱动的人效与经营数据；不能单独归因于客服AI。

[S7] Moffatt v. Air Canada, 2024 BCCRT 149. British Columbia Civil Resolution Tribunal, 2024–02–14. 企业对网站聊天机器人误导性信息承担责任。

[S8] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023；Generative AI Profile, 2024. 治理、映射、测量、管理及生成式AI风险行动。

[S9] International Organization for Standardization. ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system. AI管理体系要求与持续改进。

[S10] 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, 2023–07–13，2023–08–15施行。服务提供者内容、安全与个人信息义务。

[S11] 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。显式标识、隐式标识与传播责任。

[S12] 中华人民共和国个人信息保护法, 2021；中华人民共和国数据安全法, 2021；中华人民共和国网络安全法及相关配套规则。

[S13] W. Edwards Deming. Out of the Crisis, 1982；The New Economics, 1993. 系统观、变异与PDSA。

[S14] Eliyahu M. Goldratt; Jeff Cox. The Goal, 1984. 约束理论及系统瓶颈。

[S15] Michael E. Porter. What Is Strategy? Harvard Business Review, 1996. 战略取舍与活动系统。

[S16] Jay R. Galbraith. Designing Organizations, 1995及后续版本. 星型模型与信息处理视角。

[S17] Eric Trist; Fred Emery 等. 社会技术系统研究，20世纪50年代起。技术与社会系统联合优化。

[S18] Robert S. Kaplan; David P. Norton. The Balanced Scorecard, 1992/1996. 财务、客户、内部流程、学习成长的平衡。

[S19] Charles Goodhart. Problems of Monetary Management, 1975. 指标成为控制目标后会被扭曲的规律。

[S20] Robert G. Cooper. Winning at New Products及Stage–Gate相关研究. 分阶段评审、资源释放与停止。

[S21] Chris Argyris; Donald Schön. Organizational Learning, 1978. 单环与双环学习。

[S22] Karl E. Weick; Kathleen M. Sutcliffe. Managing the Unexpected. 高可靠组织、对失败敏感与韧性。

[S23] Peter F. Drucker. The Practice of Management, 1954. 目标、责任与绩效管理。

[S24] David A. Aaker 等营销资产研究； Rust, Ambler, Carpenter, Kumar, Srivastava. Measuring Marketing Productivity, Journal of Marketing, 2004. 营销投入与财务结果连接。

[S25] Alibaba Group. Taobao and Tmall’s 11.11 Shopping Festival Records Robust Growth, 2024–11–12. 平台官方口径：AI电商工具包服务约400万商家，生成超过1亿份营销内容；超过80万商家完成200多万次流量分析；全站推广覆盖近29万商家和超过160万件商品。采用与平台归因结果不可直接外推为商家ROI。

[S26] Lenovo. Lenovo Harnesses Generative AI to Accelerate and Optimize Creative Content, 2024–05–07；Lenovo Powers Lenovo — Studio AI官方案例。公司自报：超过11,500名销售人员和合作伙伴使用；新产品资产制作最高8倍提速，部分官方页面表述为最高90%更快；财务增量未披露。

[S27] Ping An Insurance (Group) Company of China, Ltd. 2025 Annual Results and Annual Report, 2026–03. 公司披露：2025年AI服务代表约17.02亿次服务，占客服总量约80%；AI智能体协助实现人民币1,331.79亿元销售。公司自报，未拆分AI增量贡献。

[S28] 国家互联网信息办公室等. 人工智能生成合成内容标识办法配套说明与实施要求，2025. 用于生成合成内容显式、隐式标识及传播责任的流程设计。