



028 / WORK REDESIGN CLINIC

PERFORMANCE TRANSFER CLINIC / TASK
· PRACTICE · EVIDENCE

企业AI营销培训为何失效： 从学工具到改工作

把课程从知识供给升级为岗位任务、真实练习、经理反馈与绩效证据的工作系统改造。

找准六个转化断点

从目标错位到指标失真

让经理成为转化引擎

观察、复核、反馈与流程调整

重构 WORK 六段链

从经营结果倒推岗位行为

完成 30/60/90 天改造

任务基线、真实试点与证据评审

BOARD BRIEF 01

执行摘要：十个董事会结论

竹势 AI 营销智库出品

BOARD TRANSFER THESIS / 董事会转化判断

培训失效通常不是课程不够，而是企业把工作设计与绩效管理问题误当成知识传递问题。

核心判断：AI 营销培训失效，通常不是课程供给不足，而是企业把工作设计与绩效管理问题误当成知识传递问题。

表 1 | 十个核心结论

董事会结论	经营含义
1. 课程完成不是工作改变	结课、课时、满意度和证书只证明学习供给发生；知识掌握、工具采用、岗位行为、产出质量和经营结果是六类不同证据，不能互相替代。
2. 任何一段断裂都不能靠加课补救	目标任务不清、流程未改、数据不可用、权限受限、经理不检查或指标失真时，再多课程只会扩大“会说不会做”。
3. 应先选任务，再选能力与课程	优先改造高频、可观察、有基线、有反馈、风险可控且接近经营结果的任务；“了解模型功能”不是任务。
4. 培训对象不是员工个人，而是人—AI—流程系统	同一员工在有无模板、知识库、审批门、质量标准 and 经理反馈时，表现可能完全不同。
5. 真实项目是转化主战场	课堂演示适合建立共同语言，稳定迁移必须进入模拟任务、受控真实项目和正式工作流，并以真实产物验收。
6. 一线经理是最大杠杆也是最大约束	经理决定员工是否有时间试、是否允许暴露错误、是否在例会检查新行为、是否奖励复用与质量。经理不改，员工会回到旧方法。
7. 速度指标必须与质量、风险成对出现	只追求产出数量和周期，可能带来事实错误、品牌偏移、客户投诉、隐私泄露和返工上升。
8. L&D 不能单独为转化负责	业务主管拥有任务与结果，CMO拥有质量与客户影响，IT/数据拥有环境与权限，法务治理拥有风险门，L&D拥有学习设计与证据纪律。
9. 采购应按行为证据验收	供应商展示效果、讲师评分和课时不应成为主验收项；至少要交付任务卡、真实项目、经理脚本、通过标准、基线与复盘证据。
10. 规模化不是“更多人上课”，而是更多任务通过证据门	90 天应先锁定三类任务，完成基线、受控实践和流程修订，再决定扩展到哪些岗位、工具与权限。

表 2 | 董事会应改问的五个问题

错误问题	正确问题
员工学会了多少功能？	目标岗位任务是否以更快、更稳、更低风险的方式完成？
多少人拿到证书？	多少人连续多周期按新流程产出合格结果？
工具登录率是多少？	关键任务中的有效使用率、复用率和人工接管率是多少？
课程满意度高不高？	质量、周期、转化、收入或风险是否相对基线改善？
是否应该全员培训？	哪三个任务最值得先改，哪些前置条件尚未就绪？

第一章 现场热、回岗冷：培训失效的真实机制

竹势 AI 营销智库出品

01	02	03	04	05	06
目标错位	任务抽象	场景脱离	环境受阻	经理失联	指标失真

判断

“现场热”来自新奇性、即时反馈和低风险演示；“回岗冷”来自真实任务的复杂性、旧流程惯性、权限摩擦、质量责任和经理信号。培训现场优化的是体验，岗位现场优化的是风险与绩效，两套激励并不相同。

失效往往发生在回岗后的第一个工作周期，而不是课堂结束时。员工面对真实营销任务，需要同时处理品牌语境、事实核验、客户隐私、渠道规格、审批时点与结果责任；课堂中被刻意简化的变量重新出现，原本顺畅的操作会迅速变成高认知负荷决策。若岗位没有明确“哪些任务必须使用、哪些可以选择使用、哪些不得使用”，员工会把新能力放在低风险、低价值的边缘工作上，例如改写措辞或整理会议纪要，却不进入影响收入、成本与客户体验的核心链路。管理层看到的是采用不足，员工感受到的却是责任不对称：收益属于组织，错误由个人承担。

因此，现场热度不应被解释为转化已经发生。真正的转化证据至少要跨越一个完整业务周期：员工能够在正常工作压力下识别适用任务，调用批准的数据和工具，按质量门完成交付，并在异常时及时升级。营销岗位尤其需要观察“最后一公里”：内容是否真正进入审核与发布，投放建议是否进入实验与预算流程，线索摘要是否写回 CRM，复盘结论是否改变下一轮行动。若产物仍停留在个人对话窗口或临时文档，课程即使满意度很高，也只是建立了可能性。

机制

经典学习迁移研究将迁移结果视为学习者特征、训练设计与工作环境的共同产物。对 AI 营销而言，环境变量尤其关键：员工即使掌握操作，也可能因无可用数据、无品牌知识库、不能连接渠道、审批链过长、经理仍按旧产量考核而停止使用。Baldwin 与 Ford 的转化框架、Brinkerhoff 的成功案例方法以及 Gilbert 的绩效工程都指向同一结论：能力不足只是绩效缺口的一类原因。

表 3 | 培训失效六断点诊断

断点	典型表象	根因	不能用什么补救	管理动作
目标错位	课程围绕功能，业务围绕增长	未定义经营结果与目标任务	增加模块与课时	明确结果、任务、基线
任务抽象	会演示但不会处理真实 Brief	训练材料与工作情境不相似	再讲概念	用真实输入与边界练习
场景脱离	课堂产出漂亮，回岗无法复用	缺少模板、知识、系统连接	发更多课件	嵌入工作流与资产库
环境受阻	账号、数据、权限、审批不通	技术与治理前置条件缺失	责怪员工不积极	解决数据、权限和质量门
经理失联	员工试用后被要求恢复旧方法	主管未改变检查和奖励	增加自学任务	写入例会、复核与绩效
指标失真	高结课、高登录、低业务变化	把活动指标当价值指标	优化满意度	建立六层证据链

证据与边界

一项覆盖约 5,000 名客服人员的现场研究发现，生成式 AI 助手平均提高每小时解决问题数约 14%，收益主要集中在经验较少者；高经验者收益很小，部分质量指标甚至可能略降。该研究证明“嵌入工作任务的实时支持”可以带来生产率改善，但不能外推为任何课程、岗位或营销场景都会产生同等收益。[S1] 2026 年微软研究对 388 名财富 500 强零售企业员工的现场实验进一步显示，在相同工具访问条件下，围绕人机协作增加结构化支架会改变绩效，说明“怎么使用”与“是否有工具”同样重要。[S2]

表 4 | 课程的适用边界

何时基础课程仍有价值	何时必须停止加课
新工具上手、共同术语、基本风险意识	员工已会操作但真实任务仍无法完成
低风险功能探索和初始信心建立	数据、权限、流程或质量标准未就绪
新岗位的最低能力底线	经理不允许试错或仍按旧流程考核
法规要求的必要认知	问题是岗位职责与决策权不清

第二章 WORK 六段转化链

竹势 AI 营销智库出品



WORK 六段转化链把培训从课程工程改造成工作系统工程。顺序必须从经营结果向前倒推，而非从工具功能向后寻找应用。

表 5 | WORK 六段转化链

环节	定义	关键问题	最低证据	常见断裂
W—Work outcome	经营结果	要改善收入、毛利、周期、客户价值还是风险？	明确口径、基线、观察期	目标写成“提升 AI 能力”
O—Observable task	可观察任务	哪个岗位在何时对什么输入做什么动作？	任务频次、责任人、产物	只写“会用 AI”
R—Redesigned workflow	重构流程	人、AI、数据、审批如何交接？	新旧流程图、权限与例外	把 AI 叠加在旧流程上
K—Knowledge/practice	知识与练习	需要什么判断规则、示例、练习与反馈？	练习包、通过标准	只讲功能，不练判断
R—Reinforcement	经理强化	谁观察、复核、反馈、奖励与纠偏？	固定管理节奏	培训后无人追踪
K—KPI evidence	绩效证据	如何证明行为、质量与结果变化？	分层指标和对照/基线	只看结课与登录

表 6 | 六段断裂的补救规则

环节断裂	错误反应	正确补救
经营结果不清	增加战略课程	由业务负责人明确价值口径与时间窗
任务不可观察	增加案例数量	把职责拆成触发—输入—动作—产物—责任
流程不支持	再做工具教学	先改接口、审批、系统连接和模板
练习不足	发操作手册	增加高相似度任务、反馈与重复
经理不强化	做学习社群	把观察与复核写入周会和一对一
证据不足	做满意度调查	建立基线、样本、质量盲评和业务指标

管理含义

WORK 的价值不是给培训换一个名字，而是强制管理层承担前置责任。只要经营结果、目标任务和工作流未定义，L&D 就不应接受“先开一轮课再说”的委托；只要经理强化和证据口径未落实，项目就不应宣称进入规模化。

第三章 六层证据：从结课到经营结果

竹势 AI 营销智库出品



表 7 | 学习转化六层证据

层级	回答的问题	可用指标	不可替代的下一层
1 课程完成	供给是否触达？	报名、出勤、结课、满意度	不能证明掌握
2 知识掌握	是否理解规则？	情境测验、风险辨识、判断题	不能证明采用
3 工具采用	是否在目标任务中使用？	有效使用率、连续周使用、任务覆盖	不能证明行为正确
4 岗位行为	是否按新动作执行？	流程遵从、人工复核、样例复用、升级及时性	不能证明质量
5 产出质量	结果是否更好且稳定？	事实准确、品牌一致、返工率、客户评分	不能证明经营价值
6 经营结果	是否改变业务或风险？	周期、成本、线索、转化、收入、投诉、事故	需要谨慎归因

表 8 | 禁止单独使用的指标

误用指标	为何危险	替代或配对指标
结课率	只说明完成流程	目标任务连续应用率
登录率	可能只是打开工具	有效任务占比、产物进入正式流程比例
生成数量	可能扩大低质产出	一次通过率、返工时长、复用率
节省时间自报	受记忆和期待偏差影响	系统时间戳、周期基线、抽样核验
满意度	易被讲师表现和新奇性影响	行为观察、盲评质量、业务指标
证书数	证明考试通过而非工作迁移	真实项目通过与经理签字

Kirkpatrick 四层模型有助于提醒企业从反应、学习走向行为与结果，但实践中常被简化成“满意度加测验”。本报告把采用、岗位行为和产出质量拆开，是因为 AI 工具可能被频繁使用却以错误方式使用，也可能提升速度却降低事实可靠性。Deming 的质量思想提醒管理者：多数绩效来自系统，若测量只奖励数量，系统会理性地产生更多低质量输出。Goodhart 与 Campbell 的规律则解释了为什么一旦结课率或登录率成为目标，它们会迅速失去作为价值代理的能力。

第四章 从岗位真实任务倒推能力

竹势 AI 营销智库出品

岗位任务一练习一证据卡

当前基线 · 理想动作 · 输入数据 · 工具权限 · 风险门 · 练习情境 · 反馈人 · 通过标准 · 业务指标

任务选择原则

适合优先改造的任务通常具备六个特征：频次足够高、输入相对可得、产物可观察、质量可判定、反馈周期较短、风险可被控制。创造性并不意味着不可训练，但应训练“判断与取舍”而非追求唯一标准答案。

任务倒推的最小单位不是岗位名称，而是可以被观察、重复和判定的工作事件。以“内容运营会使用 AI”为目标无法设计训练，也无法验收；把它改写为“在收到产品 Brief 后两小时内形成三套有来源标注的标题与正文方案，经品牌审核一次通过率达到既定门槛”，才可以继续定义输入、操作、风险门和证据。对于广告投放，目标不应是“会分析账户”，而应是“在固定数据窗口内识别异常、提出可验证假设、生成不越权的调整草案，并由授权人批准后进入实验”。任务颗粒度越接近真实交付，能力定义越不容易退化为功能清单。

并非所有高价值任务都应立即纳入训练。若任务的输入数据长期缺失、责任人不清、审批权冲突或质量标准无法达成一致，培训只会把结构性矛盾转移给员工。管理层应先判断缺口属于知识、技能、动机还是环境：知识缺口可以通过短课解决；技能缺口需要重复练习和反馈；动机缺口涉及奖励、风险与身份；环境缺口则需要改系统、流程或权限。只有当“更会做”确实能够改善结果时，训练投资才成立。对高度创造性的品牌策略任务，可以标准化问题定义、证据要求、评审语言和风险边界，但不应把创意判断压缩为唯一模板。

表 9 | 岗位任务—人机判断边界

营销任务	可观察动作	AI 角色	人的不可外包判断	主要证据
内容 Brief 生成	将目标、受众、主张、证据、渠道、禁区结构化	整理资料、生成初稿	确定商业命题与取舍	Brief 一次通过率、返工周期
长文/社媒改写	基于核验资料形成多平台版本	草拟与适配	事实、品牌、观点与发布责任	事实错误率、品牌一致性、发布周期
广告创意评审	按策略命题和标准评审变体	扩展方案、整理证据	创意方向、风险、预算	评审周期、胜出率、返工
投放账户体检	识别异常并形成建议草稿	拉数、诊断、草拟动作	预算调整和最终执行	发现时效、建议采纳、风险事件
CRM 跟进摘要	汇总客户状态与下一步动作	摘要、提醒、草拟话术	客户关系和承诺	跟进覆盖、响应时效、转化
活动复盘	连接目标、动作、数据与学习	整理数据、提炼模式	归因判断和下一步决策	复盘及时性、建议落地率

表 10 | 任务优先级评分尺

任务筛选维度	评分 1	评分 3	评分 5
经营价值	边缘便利	局部效率	直接影响收入/成本/风险
频次	季度以下	每月	每周或每日
可观察性	产物模糊	部分可评	动作与产物清晰
数据就绪	依赖缺失	可手工补齐	稳定可用
风险可控	高风险难隔离	可审批	低风险可回滚
反馈速度	数月	数周	当天或一周内

表 11 | 必须先改流程而非先培训的任务

不应先培训的情形	先做什么
岗位本身没有明确责任人	重写岗位任务与责任
输入数据长期缺失或口径冲突	治理数据源与业务定义
工具无企业许可或无法审计	完成技术与采购准入
输出无质量标准	建立事实、品牌、法律和客户标准
关键动作没有审批或例外机制	设计权限、升级与回滚
经理仍要求完全沿用旧流程	先改管理节奏和绩效口径

第五章 workflows、数据、权限与质量门

竹势 AI 营销智库出品

DATA 数据 可用、可溯源	ACCESS 权限 最小必要	QUALITY 质量 发布前标准	RISK 风险 升级与停止
-----------------------------	-----------------------------	-------------------------------	----------------------------

Hammer 的流程再造思想提醒企业，技术的价值不在于把旧动作自动化，而在于重新考虑工作为何存在、由谁完成、信息如何流动。AI 叠加在坏流程上，常见结果是更快地产生更多返工。联想公开材料描述其先盘点流程中可被智能化赋能的机会、按收益排序、小规模验证技术与业务价值，再推广并迭代；这是一种“流程—试点—推广”次序，而不是“课程—登录—规模化”次序。[S3]

工作流嵌入必须同时回答四个运营问题。第一，员工从哪里发起任务，是否需要在多个系统之间复制信息；第二，系统能否获得完成任务所需的最新数据与品牌上下文；第三，产出由谁复核、在什么条件下可以进入外部渠道；第四，结果和反馈写回哪里，能否被下一次任务复用。任何一项缺失，个人能力都会被等待、重复录入和责任模糊消耗。营销团队常见的“会做但不用”，并不是拒绝技术，而是新路径比旧路径多出若干登录、上传、复制和解释步骤，且最终仍要手工重做。

质量门应与风险等级成比例，而不是所有产出都走同一审批链。内部头脑风暴可以允许快速生成和低成本淘汰；面向客户的报价、医疗或金融内容、自动化触达和预算调整则必须提高数据来源、复核、授权与留痕要求。管理层需要避免两个极端：完全开放会扩大错误发布和数据泄露，完全封闭则迫使员工在批准体系之外寻找更顺手的工具。合理设计是把低风险练习环境、受控真实项目和正式生产权限连接成升级路径，使员工能通过证据获得更大授权。

表 12 | AI 营销 workflow 八要素

workflow 要素	设计问题	最低控制
触发	何时启动任务，谁发起？	触发条件和责任人
输入	哪些数据和资料可用？	权威源、版本、敏感级别
AI 动作	可生成、分析、建议还是执行？	工具白名单与动作边界
人工判断	哪些节点必须确认？	审批角色和时限
质量门	怎样判定合格？	检查表、抽样或双人复核
例外升级	何时停止或转人工？	红线、升级路径、回滚
状态记录	如何知道做到哪一步？	任务状态、日志、版本
结果回写	产物和学习进入哪里？	CRM、内容库、知识库或复盘库

表 13 | 训练中的数据分级

数据类别	允许方式	限制与训练要求
公开资料	可在批准工具中使用	仍需核验来源和版权
企业内部一般资料	仅限企业批准环境	权限最小化、版本控制
客户与潜在客户数据	按目的和授权使用	脱敏、访问控制、留痕
未公开财务/战略	原则上禁止进入外部工具	仅限批准隔离环境
个人敏感信息	高限制	合法基础、必要性、最小化
受版权约束素材	按许可使用	来源、授权范围、再发布检查

表 14 | 五道质量与风险门

质量门	检查内容	责任角色
事实门	数据、引用、产品能力、时间是否可核验	业务专家/编辑
品牌门	主张、语气、视觉与禁区是否一致	品牌负责人
法律门	广告、隐私、版权、行业监管是否合规	法务/合规
客户门	是否误导、冒犯或承诺越界	业务/客户负责人
执行门	是否允许发布、调预算、触达或写回系统	授权审批人

第六章 真实练习与在岗迁移

竹势 AI 营销智库出品

01 课堂演示	02 模拟任务	03 受控真实项目	04 正式工作流
每一级都必须交付工作样本、反馈记录与升级证据。			

Ericsson 的刻意练习强调明确目标、可重复任务和高质量反馈；Bjork 的“有益困难”说明适度增加检索、变式和间隔有利于长期保持；Sweller 的认知负荷理论提醒培训设计者不要同时让新手处理复杂工具、陌生业务、多个规则与高风险发布。对企业而言，最有效的路径不是把真实工作完全搬进第一堂课，而是逐级提高真实度。

真实练习不等于把未经准备的员工直接放入生产环境。有效设计应控制变量并逐步增加复杂度：先让员工在标准样例中识别好坏，再处理包含缺失字段、冲突事实和品牌禁区的变式任务，随后进入有教练旁观的真实项目，最后才独立承担正式交付。每次练习都要有可见反馈，反馈不能只评价文案“好不好”，而要指出事实、结构、品牌、客户、渠道和风险分别在哪一步偏离标准。没有分解反馈，员工只能凭印象反复尝试，难以形成可迁移的判断规则。

营销场景的练习库还应包含反例与异常，而非只展示成功答案。内容团队需要练习识别虚假引用、过度承诺、语气漂移和版权风险；投放团队要处理数据延迟、归因变化、预算冲突和小样本波动；CRM 团队要面对客户身份不确定、授权缺失、敏感信息和错误匹配。对高风险任务，沙盒中的通过标准应覆盖“知道何时停止并升级”，而不是只看是否完成。一个能主动报告不确定性、保留证据并请求复核的员工，往往比一个快速给出确定答案的人更适合获得生产权限。

表 15 | 真实工作嵌入度四级

嵌入等级	任务环境	证据	升级门
L1 课堂演示	统一样例、讲师操作	理解关键步骤和风险	通过情境测验
L2 模拟任务	仿真 Brief、脱敏数据、时间限制	独立完成并通过质量检查	连续两次达标
L3 受控真实项目	真实业务输入、受限权限、人工审批	产物进入真实项目但不自动外发	质量、周期与风险达标
L4 正式工作流	真实系统、标准权限、持续监控	连续多周期稳定产出	通过业务证据评审

表 16 | 刻意练习设计清单

练习设计要素	具体要求
情境相似	使用真实岗位输入、术语、时间压力和约束
变式	同一规则在不同产品、渠道、受众和异常情况下练习
即时反馈	错误发生后尽快指出事实、判断和流程问题
延迟检索	隔天或隔周重新完成关键判断，检验保持
样例对比	比较优秀、勉强、失败产物并解释差异
反思复盘	记录采用、修改、拒绝 AI 建议的理由
安全失败	在沙盒中允许错误，不把练习错误直接暴露给客户

表 17 | 五类高价值练习

练习类型	适用任务	通过标准示例
事实核验演练	报告、长文、社媒	关键事实全部有权威来源；无法核验则删除或标注
品牌一致性演练	内容、广告、客服	核心主张、语气和禁区符合品牌规范
异常升级演练	投放、客户触达	识别红线并在规定时间转人工
多版本选择演练	创意、标题、脚本	能说明选择依据而非只选“看起来最好”
复盘演练	活动、投放、内容	区分事实、推断和建议，避免过度归因

第七章 一线经理强化回路

竹势 AI 营销智库出品



Schein 的组织文化观点表明，领导者持续关注、测量和奖励什么，员工就会把什么理解为真正确规则。Edmondson 的心理安全研究则说明，员工只有在能够暴露不确定性和错误时，才会报告 AI 的问题并寻求帮助。经理若一边要求创新、一边惩罚所有试验性错误，员工会转向隐蔽使用或完全回避。

一线经理决定新行为在团队中是“被允许的实验”还是“额外的个人负担”。如果周会仍只追问最终数量，不检查过程证据；如果员工使用新方法后节省的时间被立即填入更多任务；如果出现一次可纠正错误就被公开否定，团队会迅速学会隐藏使用、回到旧流程或只在无关紧要的工作上尝试。经理强化不是口头鼓励，而是把目标任务、观察点、反馈时点、优秀样例和障碍升级写入固定管理节奏。

经理自身也必须接受工作改造。其新责任包括：判断任务是否适用，确认数据与权限已经就绪，对关键输出进行抽样盲评，区分人员能力问题与系统环境问题，并把重复出现的障碍提交给流程、IT 或治理负责人。对于营销经理，最重要的变化是从逐字改稿转向定义质量标准、解释业务取舍和管理例外。若经理仍以个人经验完成所有最终判断，却不把判断标准显性化，团队无法通过训练复制能力，AI 也只能制造更多等待审批的草稿。

表 18 | 经理强化回路

环节	经理动作	产物
会前目标	明确本周要改的任务、质量与风险边界	任务目标卡
在岗观察	查看真实输入、操作、人工修改和升级行为	观察记录
结果复核	按事实、品牌、客户与业务标准检查	复核结论
反馈纠偏	指出判断错误、流程障碍和下一次练习	反馈单
样例沉淀	把优秀和失败样例纳入团队资产	样例库
权限/流程调整	将反复出现的问题转为系统改进	变更请求

表 19 | 经理强化写入管理节奏

管理节奏	最小动作	不应做
每日/任务级	对高风险产物做快速审批	逐字替员工重做
每周例会	审查 2—3 个真实任务及证据	只问“用了没有”
双周复盘	比较基线、质量与返工	只表扬产出数量
月度评审	决定流程、权限和标准调整	把所有问题归为能力不足
季度人才盘点	识别判断力、复用力 and 治理能力	用证书数替代绩效

表 20 | 经理阻断诊断

经理常见阻断	员工反应	修复方式
要求新工具但不给时间	下班后私自试或放弃	配置练习时间和任务减负
只奖励快和多	牺牲事实与品牌质量	速度与质量成对考核
错误即公开责备	隐藏错误与风险	区分可接受试验和违规越界
经理自己不用新流程	团队判断项目不重要	经理亲自主持证据复盘
审批无限延迟	员工回到旧流程	设置时限与代理审批

第八章 绩效证据与学习转化仪表盘

竹势 AI 营销智库出品

LEARN 参与 / 掌握	BEHAVE 采用 / 行为	PERFORM 质量 / 周期	OPERATE 经营 / 风险
------------------	-------------------	--------------------	--------------------

评估应从“有没有发生”走向“是否稳定、是否更好、是否值得”。没有随机对照时，也可以采用基线前后比较、分阶段上线、同任务匹配组、盲评抽样和过程追踪，但必须明确其他变化可能影响结果。

六层指标之间不能相互替代。课程完成说明学习供给被触达；知识测验说明员工在特定条件下能够回忆规则；工具采用说明发生过使用；岗位行为说明目标动作进入真实流程；产出质量说明速度改善没有以错误、返工或品牌损害为代价；经营结果才回答收入、成本、客户价值或风险是否改变。管理层应同时观察层级之间的转化率。例如登录人数很高但正式任务覆盖率很低，问题可能在任务与环境；行为覆盖提高但一次通过率下降，说明质量门或练习不足；质量稳定但经营指标没有变化，则应重新检查场景价值或归因链。

速度指标尤其容易误导。初稿时间缩短并不等于周期缩短，因为复核、返工和等待可能把节省全部吞噬。内容场景应同时记录从 Brief 到可发布资产的端到端周期、事实错误率、品牌审核轮次和资产复用率；投放场景要观察建议采纳率、实验完成率、异常发现提前量与风险事件；CRM 场景则应跟踪记录完整性、跟进及时性、线索阶段推进和客户投诉。没有基线时先建立两到四周稳定观察，不宜用一次展示或个别明星员工的结果宣称组织价值。

表 21 | 学习转化仪表盘

指标层	领先指标	滞后指标	警戒指标
参与/掌握	情境测验、风险识别	关键规则保持率	高满意低掌握
采用/行为	目标任务使用率、连续周使用	流程遵从、样例复用	登录高、正式产物低
质量/周期	一次通过率、平均周期	返工率、客户评价	速度升、错误升
经营/风险	线索响应、投放异常发现	转化、收入、成本、事故	归因过度、风险外溢

表 22 | 指标定义示例

指标	口径示例	数据源	评审频率
有效任务采用率	使用批准 AI 流程并进入正式交付的目标任务/全部目标任务	工作流日志+业务系统	周
一次通过率	首次提交无需重大返工的产物/抽样产物	审核记录	周
周期改善	同类任务中位完成时长相对基线变化	时间戳	双周
事实错误率	抽样关键事实错误数/核验事实数	盲评	双周
复用率	调用批准模板、样例或知识资产的任务占比	资产库日志	月
业务结果	线索响应、转化、收入或成本变化	CRM/财务/投放	月/季度
风险结果	违规发布、隐私事件、客户投诉、紧急接管	事件系统	实时/月

表 23 | 因果评估强度阶梯

因果强度	方法	适用条件	限制
较弱	自报感受与案例故事	早期探索	期待偏差大
中等	前后基线与过程数据	稳定任务、无重大外部变化	难排除同期因素
中强	分阶段上线或匹配组	多团队、可错峰	组间差异仍存在
较强	随机对照或准实验	样本与伦理允许	成本高、外推有限

第九章 组织共责与决策权

竹势 AI 营销智库出品

转化责任不能外包给 L&D

CMO 定任务 · HR/L&D 设计练习 · 经理强化行为 · IT/数据保障环境 · 法务设置风险门

Argyris 的双环学习提醒组织，不仅要修正动作，还要质疑支配动作的目标、规则和假设。Kotter 的变革思想则强调领导联盟与短期胜利。AI 营销能力转化不能由 L&D 单点推动，因为许多障碍位于流程、系统、绩效与风险规则。

共责不等于人人负责。每一类目标任务仍需明确唯一的经营责任人，由其决定该任务是否值得改造并对业务结果负责；L&D 负责学习设计和证据方法，IT/数据负责批准环境、连接与可用性，法务与治理负责风险分级和升级规则，一线经理负责观察与强化，员工负责按标准执行并报告异常。若没有唯一的结果责任人，跨部门会议容易停留在课程内容、工具选型和合规条款，真正影响绩效的流程取舍无人拍板。

组织还应建立两条分开的升级路径：一条处理单次任务异常，例如事实冲突、客户敏感信息或错误发布风险；另一条处理系统性障碍，例如权限长期等待、数据口径不一致、审批链重复或质量标准冲突。前者保障当前交付，后者改变工作系统。只有把重复问题从个体辅导中提取出来，组织才会形成双环学习，而不是不断要求员工“更认真”。小团队可以由同一人兼任多个角色，但经营责任、技术责任和风险批准仍应在记录中区分。

表 24 | 组织共责模型

角色	拥有的责任	必须交付的证据
CEO/业务总负责人	价值方向、资源与冲突裁决	结果口径、试点授权、规模化决定
CMO/业务主管	目标任务、质量标准、客户与经营结果	任务清单、基线、质量评审
CHRO/L&D	学习设计、练习、教练与评估纪律	练习包、通过标准、转化数据
一线经理	在岗观察、反馈、奖励与例外升级	观察记录、样例、纠偏
IT/数据	工具准入、集成、权限、日志与支持	环境可用性、权限记录、故障数据
法务/合规/安全	风险分级、红线、审查和事件响应	政策、风险门、事件台账
员工/岗位人员	按规则执行、核验、报告异常并沉淀学习	真实产物、修改理由、异常报告

表 25 | 关键决策权 RACI

决策	A 最终负责	R 执行	C 咨询	I 知会
选择目标任务	业务主管	营销运营	L&D、IT、法务	CEO
批准工具与数据	IT/数据负责人	平台团队	法务、业务	L&D
定义质量标准	CMO/业务主管	领域专家	法务、品牌	L&D
设计练习	L&D 负责人	学习设计师/教练	业务经理	IT
决定正式上线	业务主管	营销运营	IT、法务、L&D	CEO
规模化拨款	CEO/经营层	项目办公室	CMO、CHRO、CIO	相关团队

第十章 供应商与课程采购治理

竹势 AI 营销智库出品

真实任务	工作样本	反馈强度	行为证据	经营边界
------	------	------	------	------

供应商天然会优化可展示的成果：课堂互动、短时生成效果、讲师评分和证书数量。采购方若以这些指标验收，就会得到“展示型课程”。Goodhart 与 Campbell 的警示在此直接适用：当代理指标被用于奖惩，它会被优化到失真。

采购文件应从“交付多少课时”改为“使哪些岗位任务达到何种证据门”。合格供应商需要先理解任务基线、系统环境、风险等级和经理角色，再设计内容与练习；若供应商只提供通用功能演示、预制成功样例和一次性证书，企业得到的只是短期体验。验收应包含至少一个受控真实项目、员工在无讲师代操作情况下的独立表现、经理复核记录、质量与周期数据，以及未达标时对流程、权限或课程的修正建议。

供应商自报效率数字只能作为场景假设，不能直接进入企业收益承诺。企业需要核验样本是谁、任务是否与本企业相似、比较基线是什么、是否计算复核返工、观察期多长，以及结果来自独立研究还是产品方案例。对于专项技术认证，证书可以证明最低知识或操作门槛，但不应自动转换为生产权限。生产授权必须依据岗位任务证据，并设置期限、抽查和撤回机制，防止一次考试替代持续胜任。

表 26 | 供应商六阶段治理

采购阶段	必答问题	淘汰信号
需求定义	目标岗位、任务、基线和工作约束是什么？	只提供通用课纲
方案评审	如何使用真实任务、练习、反馈与经理强化？	只承诺功能覆盖
安全评审	如何处理数据、账号、版权、日志和沙盒？	要求使用个人账号或公开工具处理敏感资料
试点验收	哪些行为与质量证据必须达标？	只交满意度和结课证书
规模化	如何复制到不同经理、岗位和地区？	依赖明星讲师个人表现
退出与沉淀	模板、样例、数据和方法如何归企业？	知识资产不可移交

表 27 | 供应商评分卡

评分维度	权重建议	证据
业务任务适配	20%	岗位任务分析与真实项目设计
学习迁移设计	20%	练习、反馈、间隔与在岗辅导
工作流与技术理解	15%	数据、权限、集成、日志方案
经理与组织机制	15%	经理脚本、复盘节奏、角色责任
测量与因果纪律	15%	基线、指标、样本和边界
风险与合规	10%	数据分级、版权、审批、事件响应
资产移交	5%	可编辑模板、样例库、运营手册

表 28 | 证书的合理边界

专项认证仍有价值的场景	边界
平台管理员、数据工程、安全与审计角色	认证证明最低知识，不证明业务绩效
高风险系统操作和法规要求	仍需在岗授权与持续复核
新版本关键能力更新	应与真实变更任务结合
外部职业资格或客户准入要求	不能替代企业内部质量标准

BOARD BRIEF 12

第十一章 风险、合规与影子 AI

竹势 AI 营销智库出品

RISK CLINIC / 影子 AI 警报

禁止不会自动产生安全；安全来自可用工具、真实练习、可观察日志和明确升级门。

NIST AI RMF 及其生成式 AI 配套文件强调在 AI 生命周期中明确治理、测量、管理和责任。ICO 的公开指导指出，单纯全面禁止员工使用生成式 AI 往往难以奏效，可能把使用推向地下；更可行的做法是提供合规工具、明确允许与禁止用途，并训练员工理解数据与输出限制。[S4][S5] 欧盟 AI 法案第 4 条自 2025 年 2 月起适用，要求提供者和部署者采取措施确保相关人员具备充分 AI 素养；这同样不是“上过一门课”即可机械满足，而要结合技术知识、经验、使用情境和受影响人群。[S6]

风险训练应以具体工作决定为中心，而不是要求员工记忆一长串政策。员工需要能够判断：当前输入是否包含客户、员工、商业秘密或未公开经营信息；使用的工具是否获得批准；产出是否需要事实来源、版权检查、专业审阅或人工批准；异常发生后应向谁报告、如何停止外发并保留证据。训练与系统控制要相互验证：若政策禁止上传某类数据，批准工具应提供分类提示、访问控制和日志；若系统无法落实，组织不能把全部责任转嫁给个人。

影子 AI 的根因通常不是员工缺少风险意识，而是正式路径无法完成工作。治理团队应定期比较批准工具与实际任务的适配度，观察员工是否因速度、语言、模型能力、文件限制或系统连接而绕开正式入口。发现违规使用时，既要控制当前风险，也要分析是否存在业务合理性。过度限制会压低公开报告率，使问题从可观察变成不可观察；过度宽松则会让低概率高损失事件累积。有效边界是提供可用的安全替代、分级权限、明确禁区和无惩罚的早期报告机制。

表 29 | AI 营销培训风险矩阵

风险	典型触发	训练与系统控制
数据泄露	把客户、合同、未公开计划输入非批准工具	数据分级、批准环境、脱敏、日志
事实幻觉	直接发布未经核验的生成内容	权威源检索、事实门、引用检查
品牌偏移	批量内容缺少品牌上下文	品牌资产库、禁区、抽样评审
版权风险	复制受保护文本、图片或风格	许可检查、来源记录、替代素材
越权执行	自动发布、触达、调预算	权限最小化、人工审批、回滚
偏见与不公平	客户分群、推荐或评分缺少检查	多样样本、偏差检查、人工复核
影子 AI	禁令与业务需求冲突	提供合规替代、清晰政策、低摩擦支持

表 30 | 风险分级与训练环境

风险等级	练习环境	上线条件
低：内部草稿、公开资料摘要	批准工具，可直接练习	抽样质量检查
中：品牌内容、销售话术、客户摘要	沙盒或受控真实项目	人工审批、日志、数据限制
高：公开发布、客户承诺、预算调整	仿真+严格受控环境	双人或专业审批、回滚
极高：敏感个人数据、受监管决策	原则上不作为普通培训任务	专项治理、法律评估、独立验证

第十二章 情境裁剪：规模、成熟度与监管

竹势 AI 营销智库出品

SIZE 组织规模	MATURITY AI 成熟度	RISK 任务风险	REGULATION 监管强度
--------------	--------------------	--------------	--------------------

表 31 | 情境裁剪矩阵

情境裁剪首先取决于任务风险，而不是企业规模。小企业虽然层级少、决策快，但往往缺少专职数据、法务和安全能力，因此更需要限制高风险自动执行，并使用少量批准工具和清晰人工复核。成熟企业拥有更多系统与专业岗位，却可能因接口、权限和跨部门责任产生更高协调成本。高监管行业不应放弃能力转化，而应从内部研究、知识检索、草稿辅助和合规检查等低风险任务开始，建立审计证据后再扩大权限。

AI 基础不同也决定训练密度。基础薄弱组织应先统一入口、数据底线和三类高频任务，避免同时引入过多工具；已有广泛个人使用的组织，应优先治理影子使用、统一质量标准并把优秀实践转成正式流程；已经建立工作流的组织，则要把重点转向经理授权、异常处理、跨系统证据和经营归因。单一课程包无法跨越这些阶段。可复制的是 WORK 六段转化链和证据门，具体课程时长、教练比例、审批层级与自动化深度必须随情境调整。

情境	优先做法	可合并角色	不可削减底线
小团队	选 1—2 个高频闭环，经理即教练	业务主管兼项目负责人	数据边界、质量门、基线
成长企业	围绕内容—线索—跟进建立共同流程	L&D 与营销运营协同	正式权限和复盘
成熟企业	分岗位、分区域、分系统试点	设转化办公室	架构、治理和因果评估
高监管行业	先做内部辅助和低风险任务	培训与合规联合设计	审计、审批、事件响应
低 AI 基础	增加共同语言与基础沙盒	内部种子教练	不得跳过真实任务
高 AI 基础	减少基础课，增加复杂任务和治理	专家社群自组织	防止影子工具与局部最优

表 32 | AI 能力转化成熟度

成熟度	主要目标	训练重点	规模化门
M0 分散试用	建立可见性与底线	政策、批准工具、低风险任务	无严重违规，目标任务明确
M1 受控试点	证明任务级价值	真实项目、质量门、经理反馈	质量与周期达到阈值
M2 流程嵌入	稳定行为和复用	流程、数据、权限、样例库	连续多周期稳定
M3 组合扩展	跨岗位复制	教练体系、度量、共责	业务结果与风险可接受
M4 持续优化	动态调整人机分工	实验、审计、能力更新	治理和价值闭环

第十三章 八个案例卷宗

竹势 AI 营销智库出品

CASE CLINIC / 八个转化现场

主体与时间 · 工作情境 · 介入动作 · 作用机制 · 结果证据 · 归因边界

案例 1 | 财富 500 强客服：实时嵌入优于脱离工作学习

2020—2021 年前后，约 5,000 名客服人员分批获得生成式 AI 对话助手。工具直接在客户服务工作中提供建议，而非独立课程。研究报告平均生产率提高约 14%，经验较少者提升更大；高经验者收益有限，部分质量表现可能略降。机制是把高绩效人员的知识以实时支持方式扩散。边界：客服任务结构化程度较高，不能等同于复杂品牌判断；研究不证明任何工具采购或短期课程都会有效。[S1]

案例 2 | P&G 与 HBS：AI 改变团队结构，不只是个人速度

2024—2025 年公开的 P&G 大型现场实验覆盖 791 名专业人员，比较个人、团队以及是否使用生成式 AI 的表现。公开论文显示，使用 AI 的个人能够达到传统两人团队相近的解决方案质量，AI 还帮助跨越职能边界并改善情绪体验。具体动作不是给员工做通识课，而是让其在真实产品创新任务中使用统一工具与任务设计。边界：实验任务和时间受控，不能直接外推长期组织绩效。[S7]

案例 3 | 财富 500 强零售商：同一工具，不同支架产生不同结果

微软研究在 2026 年披露对 388 名员工的现场实验：所有参与者均可访问相同 AI 工具，研究只改变协作支架。结果表明，结构化的人机协作方式会影响成效。对培训的意义是，工具访问和功能教学不是充分条件；任务分解、检查点和反馈结构本身需要被设计。边界：研究场景和指标需结合原论文解释，不应把供应商摘要当成普遍因果结论。[S2]

案例 4 | 联想：先盘点流程、验证价值，再培训推广

联想官方材料描述其以流程全景图识别可由 AI 赋能的机会，按预估收益排序，小规模验证技术和业务价值，收集反馈后再推广并迭代；同时以课程更新员工思维与工作方式。另有官方案例显示销售门户打通 24 个系统，面向数千名销售人员，销售智能体包含顾问式问答、AI 陪练、外呼计划、业务数据顾问和产品配置。作用机制是培训、系统入口、数据与流程共同存在。材料中的效率数字属于企业官方自报，不能独立归因于培训。[S3][S8]

案例 5 | 腾讯：把能力嵌入工程 workflow

腾讯 2026 年官方披露称，CodeBuddy 已支持超过 95% 的腾讯工程师，并使整体编码时间减少 40%；该结果属于企业自报。值得借鉴的不是数字本身，而是能力进入日常开发 workflow、形成团队标准与工具链，而非只做课程。对营销的转译是：内容、投放、CRM 等也需要统一入口、规范、质量检查和经理复核。边界：软件开发与营销任务不同，且官方披露缺少完整对照设计。[S9]

案例 6 | 海尔集团：AI 责任培训与管理体系结合

海尔集团 2025 年 ESG 报告要求参与 AI 产品设计、开发、部署和运营的员工完成 AI 伦理与治理基础课程，并结合管理制度推进负责任应用。其价值在于按角色与风险定义最低能力，而非把所有人训练成同一种熟练度。边界：该案例主要证明治理培训制度化，不提供营销绩效改善的因果证据。[S10]

案例 7 | 爱尔眼科：通过企业微信把助手带入日常工作

爱尔眼科与联想的公开案例称，其将本地部署的大模型接入自研 AierGPT，并通过企业微信向医疗和管理人员提供带个人知识库的助手，使 AI 进入日常工作场景。具体动作包括本地部署、知识管理、日常入口和场景连接。机制是降低切换成本并控制数据。来源为供应商案例，缺乏独立绩效评估，不能把“融入场景”等同于已经产生经营价值。[S11]

案例 8 | 全面禁用的反噬：影子 AI 风险

英国信息专员办公室公开指出，孤立的全面禁令通常难以奏效，可能推动员工在组织视野之外使用生成式 AI。失败机制是业务需求仍在，而批准工具、清晰政策和支持不足。正确做法是评估实际需要，提供合规工具，明确允许和禁止用途，并培训风险与限制。边界：部分极高风险任务仍可能需要禁用，但禁用必须配套技术控制、替代流程和持续监督。[S5]

表 33 | 八个案例的证据边界

案例	核心机制	证据性质	主要边界
客服现场研究	实时支持与知识扩散	学术现场研究	任务较结构化
P&G/HBS	真实任务与跨职能协作	大型现场实验	受控任务、长期性未知
零售商支架实验	同工具不同使用结构	企业现场实验	摘要信息有限
联想	流程盘点+系统+培训	企业官方披露	自报、难单独归因
腾讯	工作流嵌入与标准化	企业官方披露	自报、岗位差异
海尔	角色化责任培训	ESG 官方披露	无绩效因果
爱尔眼科	知识库+日常入口	供应商案例	缺独立评估
ICO 反例	禁令导致地下使用	监管指导	各行业风险不同

第十四章 30/60/90 天工作改造诊室

竹势 AI 营销智库出品

DAY 01–30 锁定任务与基线	DAY 31–60 真实项目练习	DAY 61–90 证据评审与改造
----------------------	---------------------	----------------------

表 34 | 90 天工作改造诊室

阶段	目标	关键动作	证据门
0—30 天	锁定三个高价值任务并建立基线	任务盘点、数据/权限检查、风险分级、经理任命、基线抽样	任务可观察；输入可得；质量标准明确
31—60 天	在真实项目中练习、教练与复盘	L2/L3 练习、在岗观察、每周复核、系统问题登记、样例沉淀	连续两轮达到质量和周期阈值
61—90 天	调整工作流、角色、指标并决定规模化	修订 SOP、权限、模板、绩效、供应商和支持机制	业务证据可接受；风险无失控；经理可复制

表 35 | 前 30 天交付清单

0—30 天交付物	责任人
三类目标任务清单与评分	业务主管/CMO
岗位任务—练习—证据卡	L&D+一线经理
当前流程与新流程草图	营销运营+IT
数据、权限、工具和风险清单	IT/数据+法务
基线样本与指标定义	业务分析/L&D
试点团队与经理承诺	业务负责人

表 36 | 真实项目周节奏

31—60 天每周节奏	内容
周一	明确本周任务、输入、质量与风险边界
周中	教练观察真实任务，记录人工修改与异常
周五	盲评样本、比较基线、复盘流程障碍
持续	优秀/失败样例入库；重复问题转系统变更

表 37 | 规模化证据门

61—90 天规模化决定	通过	暂停/修订
行为	多数目标任务连续按新流程执行	使用只发生在课堂或少数高手
质量	一次通过率和错误率达阈值	速度提高但返工、投诉或错误上升
环境	权限、数据和支持稳定	依赖临时账号、手工搬运或个人工具
经理	能独立观察、反馈和复盘	仍依赖外部讲师推动
经营	至少出现可信的周期、成本、客户或风险改善	无基线或无法区分同期因素
治理	无重大事件且升级有效	出现越权、泄露或无法追溯

BOARD BRIEF 16

结论与管理工具

竹势 AI 营销智库出品

FINAL TRANSFER PRINCIPLE

学习只有在工作流、管理节奏与绩效证据中被观察到，才算发生。

最终结论

竹势 AI 营销智库出品

企业 AI 营销能力转化的最小有效单元，不是一门课、一个账号或一张证书，而是“一个明确经营结果——一个可观察岗位任务——一条可运行工作流——一组高相似度练习——一个经理强化回路——一套分层绩效证据”。董事会应把培训预算视为工作改造投资，并要求与流程、技术、治理和经理责任共同立项。

表 38 | 岗位任务—练习—证据卡

字段	填写要求
目标任务	用“在何种触发下，对什么输入，完成什么产物”描述
当前基线	周期、质量、返工、风险或业务结果
理想动作	人和 AI 的步骤与判断边界
输入数据	来源、版本、敏感级别
工具权限	允许工具、动作、审批人
风险门	事实、品牌、法律、客户、执行
练习情境	模拟/真实程度、异常与时间限制
反馈人	经理、领域专家、法务或教练
通过标准	行为、质量、周期与风险阈值
业务指标	线索、转化、收入、成本或风险

表 39 | 高管管理清单

董事会/CEO	CMO/业务主管	CHRO/L&D	IT/数据	法务/治理
只批准有结果口径、责任人与证据门的项目	选择任务、质量标准 和经营指标	设计练习、教练和测量	保障工具、数据、权限与日志	定义风险分级、红线与事件响应
季度审查价值、风险和扩展边界	主持真实样例复盘	拒绝以课时和满意度宣称成功	报告环境摩擦和影子使用	抽查高风险任务与供应商

表 40 | 经理每周六问

一线经理检查问题	证据
本周哪两个真实任务用了新流程？	任务记录
员工接受、修改或拒绝了哪些 AI 建议，为什么？	修改理由
速度改善是否伴随质量下降？	质量与周期对照
反复问题是能力问题还是系统问题？	问题分类
哪个优秀样例应进入资产库？	样例条目
需要调整什么权限、模板或审批？	变更请求

经典思想与专家视角的当代转译

竹势 AI 营销智库出品

表 41 | 十位专家与经典思想

思想/专家	原始贡献	对 AI 营销能力转化的转译	边界
Baldwin & Ford	学习迁移由学习者、设计与环境共同决定	不能把回岗失败只归因于员工	经典框架早于生成式 AI
Robert Brinkerhoff	以成功案例追踪培训如何产生结果	研究少数成功与失败链路，改系统而非只算平均	需防止选择性叙事
Thomas Gilbert	绩效缺口常来自信息、工具、激励和环境	先诊断环境，再决定是否培训	模型需结合组织现实
Peter Drucker	以贡献与结果定义知识工作	从经营结果倒推任务	创造性工作难完全量化
Michael Hammer	技术应推动流程重构而非旧流程自动化	先重构交接、审批和数据流	激进再造可能造成组织损伤
K. Anders Ericsson	刻意练习依赖明确目标、重复和反馈	用真实任务和专家反馈替代演示	复杂职业的“专家”形成周期长
Amy Edmondson	心理安全支持报告问题与学习	允许员工暴露 AI 错误并及时升级	心理安全不等于降低标准
W. Edwards Deming	绩效主要由系统和质量管理决定	速度与质量配对，修系统而非责怪个人	仍需个人问责
Chris Argyris	双环学习质疑目标和规则本身	反复失败时改绩效规则和岗位假设	需要管理层开放性
Donald Kirkpatrick	评估从反应走向学习、行为与结果	不得停在满意度和考试	四层间并非自动因果链

BOARD BRIEF 19

公开来源索引

竹势 AI 营销智库出品

表 42 | 公开来源索引

编号	机构/作者	标题	日期	样本或口径
[S1]	Erik Brynjolfsson、 Danielle Li、Lindsey R. Raymond	Generative AI at Work	2023/2025 修 订	约 5,000 名客服人员，现 场分阶段引入；生产率与质 量口径
[S2]	Microsoft Research；相 关研究团队	Scaffolding Human–AI Collaboration: A Field Experiment	2026	财富 500 强零售企业 388 名员工，同工具、不同协作 支架
[S3]	联想集团	智能化转型：价值导向的体系 化推进；品牌纪事实践	2023—2024	流程盘点、收益排序、小规 模验证、课程与推广闭环； 企业官方披露
[S4]	美国国家标准与技术研究 院 NIST	AI Risk Management Framework 1.0；Generative AI Profile	2023—2024	跨行业自愿风险管理框架与 生成式 AI 风险措施
[S5]	英国信息专员办公室 ICO	Staff use of AI: is a blanket ban ever the answer?；AI policy guidance	2024—2025	员工使用、影子 AI、数据 保护与组织政策指导
[S6]	European Union	Regulation (EU) 2024/1689, Article 4 AI literacy	2024；2025– 02 起适用	提供者与部署者采取措施确 保相关人员具备充分 AI 素 养
[S7]	Fabrizio Dell’Acqua 等； Harvard Business School / P&G	The Cybernetic Teammate	2025	791 名专业人员，产品创新 任务现场实验
[S8]	联想集团	销售门户、销售智能体与数字 化转型案例	2024/2026	24 个系统、数千销售、AI 陪练与业务场景；企业自报
[S9]	腾讯	Productivity Agent Suite / CodeBuddy internal transformation	2026	覆盖工程师与编码时间口 径；企业官方自报
[S10]	海尔集团	Environmental, Social and Governance Report 2025	2026	AI 产品相关人员基础课 程、伦理与治理要求
[S11]	联想集团、爱尔眼科	AierGPT 企业知识助手案例	2024—2025	本地部署、企业微信入口、 个人知识库；供应商案例
[S12]	Donald Kirkpatrick、 James Kirkpatrick	Evaluating Training Programs / Four Levels	1959 起；后续 版本	反应、学习、行为、结果评 估框架

[S13]	Timothy Baldwin、Kevin Ford	Transfer of Training: A Review and Directions for Future Research	1988	学习者、训练设计、工作环境与迁移
[S14]	Robert Brinkerhoff	The Success Case Method	2003	通过成功与失败案例追踪培训价值链
[S15]	Thomas Gilbert	Human Competence	1978	绩效工程与环境支持
[S16]	K. Anders Ericsson 等	The Role of Deliberate Practice in the Acquisition of Expert Performance	1993	刻意练习、反馈与专家表现
[S17]	Robert Bjork、Elizabeth Bjork	Making Things Hard on Yourself, But in a Good Way	2011	有益困难、检索、间隔与迁移
[S18]	John Sweller	Cognitive Load Theory	1988 起	工作记忆负荷与教学设计
[S19]	Amy Edmondson	Psychological Safety and Learning Behavior in Work Teams	1999	团队心理安全与学习行为
[S20]	W. Edwards Deming	Out of the Crisis	1982/1986	系统、质量、测量与管理责任

BOARD BRIEF 20

附录：实施模板

竹势 AI 营销智库出品

表 44 | 四周转化追踪表

周次	目标任务	基线	本周新行为	质量结果	障碍类别	下一步
W1						
W2						
W3						
W4						

表 45 | 董事会项目验收清单

问题	是/否	证据/备注
是否定义经营结果和基线？		
是否明确可观察岗位任务？		
是否完成数据与权限检查？		
是否有事实、品牌、法律与执行质量门？		
是否安排真实项目和反馈人？		
一线经理是否有固定强化节奏？		
是否同时跟踪速度、质量与风险？		
供应商是否按行为证据验收？		
是否有异常升级、人工接管和回滚？		
是否通过 90 天规模化证据门？		

附录：关键经营场景深化

竹势 AI 营销智库出品

一、内容生产：从“更快出稿”转向“更快形成可发布、可归因的内容资产”

内容团队最容易把 AI 能力转化误解为“单位时间生成更多文本”。这一口径忽略了营销内容的完整成本：选题判断、证据搜集、品牌校准、跨渠道适配、法务审核、素材协调、发布、数据回收和复盘。若只训练初稿生成，员工可能把节省的写作时间重新消耗在事实核验、重复修改和跨部门解释上，甚至因产出数量增加而使审核成为新瓶颈。因此，目标任务应被定义为“从批准 Brief 到可发布内容包”，而不是“写一篇文章”。基线至少记录端到端周期、重大返工次数、关键事实错误、品牌偏移、审核等待和发布后表现。练习也应覆盖证据不足、产品能力变更、敏感表达、不同渠道长度限制以及突发热点等异常情境。经理强化的重点不是逐字润色，而是要求员工说明：哪些材料来自权威源，哪些判断由人保留，哪些建议被拒绝，以及为什么。只有当这些判断过程可复核，团队才是在形成组织能力，而不是依赖个人偶然技巧。

对内容任务的经营验收应采用成对指标。速度必须与一次通过率、事实准确率和返工时长配对；产量必须与有效发布率、资产复用率和客户行动指标配对；采用率必须与批准知识资产调用率和高风险升级及时率配对。对于品牌叙事、危机回应和高管署名内容，AI 可以承担资料整理、结构建议和版本比较，但最终立场、承诺、语气和发布责任不可外包。对于标准化产品介绍、活动信息适配和内部摘要，可在经过样例验证后提高自动化程度。差异不在于“创意或非创意”这个粗分类，而在于错误代价、判断可编码程度、反馈速度和可回滚性。

二、广告投放：培训不能越过预算权限与实验制度

投放团队的能力转化常出现一种危险错觉：学员能读懂平台建议、生成优化方案，就被视为已经具备自动调预算或改定向的能力。实际上，投放结果同时受竞价环境、季节、创意、人群、落地页、库存、价格和归因口径影响。训练必须首先区分诊断、建议、草稿、审批和执行五种权力。初级阶段可以让 AI 拉取数据、检查异常、生成假设和形成变更草稿；预算、出价、受众排除和暂停账户等动作应根据金额、可逆性和客户影响设置审批。员工通过标准不应是“建议看起来专业”，而应是能够说明假设、证据、预期影响、观察窗口、停止规则和替代解释。

真实项目应采用小预算、可回滚和单变量优先的实验纪律。经理每周复核的不只是 ROI 数字，而是实验是否按计划执行、归因口径是否一致、是否存在选择性汇报，以及负面结果是否被记录。若团队只沉淀成功案例，系统会逐渐形成幸存者偏差。供应商若承诺通过短期课程直接提升投放回报，应要求其说明样本、账户基础、同期活动、平台变化和归因窗口；无法说明时，该承诺只能视为营销主张。培训可以提升分析与执行纪律，但无法消除市场波动，也不能代替产品竞争力、创意质量和渠道策略。

表 46 | 投放任务的权限与训练证据

权限层	允许动作	训练证据	升级条件
观察	查看账户、报表和异常	能解释数据口径与限制	数据缺失或口径冲突
诊断	提出原因假设	区分事实、推断与待验证项	涉及重大异常或合规风险
建议	形成优化选项	写明收益、风险和观察窗	预算或客户影响超过阈值
草稿	生成待审批变更包	可回滚、变更项清晰	跨账户、跨地区或敏感人群
执行	在批准范围内操作	日志、双人复核或自动校验	达到金额、风险或异常门槛

三、CRM 与线索跟进：采用率高不代表客户关系改善

CRM 场景往往能快速获得高登录率，因为员工可以用 AI 摘要客户记录、生成邮件和提醒下一步。但营销和销售的真正问题不是“是否生成话术”，而是线索是否被及时、恰当地推进。训练应把目标任务定义为一个客户状态变化：例如，高价值线索在规定时间内完成资格判断并进入下一阶段；沉默客户根据授权和历史互动获得合适触达；活动线索被正确标记来源并反馈给市场团队。AI 可以压缩信息整理和草拟时间，但客户意图、商业承诺、价格权限、异议处理和关系判断仍需要人负责。

质量标准应包括摘要是否遗漏关键承诺、是否混淆联系人、是否使用过期价格、是否产生未经授权的触达、是否把推断写成事实，以及销售是否理解并确认下一步。经理强化要进入商机例会：抽查若干线索，比较 AI 摘要与原始记录，检查销售对建议的修改理由，并识别系统性问题。若错误来自 CRM 字段缺失或销售长期不录入，继续训练摘要功能不会解决根因；管理层必须先修正数据纪律和流程责任。经营结果评估应连接响应时效、有效跟进覆盖率、阶段推进率、失联率、客户投诉与转化，而不是仅看生成邮件数量。

表 47 | CRM 场景的分层指标

CRM 任务	领先指标	质量指标	经营指标	风险指标
线索资格判断	规定时间内完成率	字段完整、一致性	MQL→SQL 转化	错误淘汰/误升级
跟进摘要	正式记录覆盖率	关键承诺遗漏率	销售准备时间	隐私或错客户
触达草稿	采用与人工修改率	品牌与事实准确	回复率/预约率	未经授权触达
商机复盘	按期复盘率	归因清晰度	阶段停留时间	选择性汇报

四、管理者能力：经理也需要被重新设计工作

许多项目把经理定义为“支持者”，却没有改变经理自己的工作量、信息入口和绩效责任。结果是经理口头鼓励员工使用新方法，但没有时间审查样例，也不知道如何区分合理试错与违规。经理能力转化至少包括四类任务：选择适合 AI 的工作、检查员工的人机判断、识别系统障碍、根据证据调整权限与指标。培训经理时，不应让其重复学习所有操作，而应使用真实团队案例练习决策：什么可以授权，什么必须审批；速度提高但质量下降时如何处理；高绩效员工拒绝使用时如何判断是合理边界还是行为惯性；员工通过个人工具绕开限制时如何同时解决需求与风险。

经理的通过标准应是能主持一次证据复盘，并把问题分类为知识、练习、流程、数据、权限、工具、激励或角色问题。只有知识和练习类问题适合继续训练；其余问题要进入运营、IT、治理或组织决策。这样可以避免把所有失败都变成“员工能力不足”。对经理的绩效也应增加团队复用率、问题关闭周期、风险报告及时性和样例资产沉淀，而不是单纯要求团队达到登录目标。高层要明确，发现并上报系统问题不是项目失败，而是试点阶段的核心产出。

表 48 | 经理证据复盘判断题

经理判断题	错误答案	合格答案
员工产出更快但事实错误增加	要求更仔细或再培训	暂停扩展，检查质量门、来源和奖励指标
员工不用批准工具而用个人账号	直接处罚全部员工	控制风险，同时调查批准工具是否满足任务需求
只有少数高手获得明显收益	要求高手分享技巧后全员复制	分析任务差异、基础能力与高手隐性判断
审批成为瓶颈	取消审批	按风险分级、授权额度和抽样复核重构
业务指标没变但周期变短	宣布成功	确认节省时间是否被重新配置并观察更长结果窗口

五、L&D 的角色转型：从课程运营者到绩效系统设计伙伴

L&D 最重要的转型不是增加 AI 课程目录，而是改变接单方式。面对“给市场部做一次全员培训”的需求，应先要求业务方提供目标任务、当前基线、典型输入、质量问题、系统环境和经理责任。若业务方无法提供，L&D 应组织任务诊断工作坊，而不是直接寻找讲师。课程设计要从知识点清单变成任务链：学员在什么情境下做什么决定，使用哪些资料 and 工具，何时转人工，产物由谁评价，失败后如何反馈。学习平台的数据也应与工作流、内容库、CRM 或审核记录连接，否则只能看到学习活动，无法看到工作迁移。

L&D 还应保护评估纪律。项目发起方常希望尽快宣称成功，供应商希望展示高满意度，员工希望获得证书，经理希望少增加管理负担。L&D 必须坚持分层证据，不把自报节省时间直接折算成财务收益，不把相关性写成因果，不隐瞒速度与质量的权衡。其价值不是证明每个项目都成功，而是帮助企业更快识别何处有效、何处受阻、何处不应扩展。对于失败试点，应形成停止或修订建议，并保留基线、样例和问题数据，使失败成为组织学习资产。

六、证据评审会：把培训复盘升级为经营审议

90 天项目结束时，不应举行以学员展示为主的结业汇报，而应召开证据评审会。参会者至少包括业务负责人、目标岗位经理、L&D、IT/数据、法务治理和财务或业务分析。会议按六层证据顺序审查：供给是否触达、知识是否掌握、目标任务是否采用、新行为是否稳定、质量是否改善、经营或风险是否出现可信变化。任何一层出现断裂，都要明确归因假设和下一步责任。比如，员工掌握且愿意使用，但正式采用率低，可能是权限、系统入口或经理信号问题；采用率高但质量下降，则应修订知识源、质量门、奖励指标或任务边界，而非继续扩大覆盖。

评审结论应只有四种：规模化、限定规模化、修订后复试、停止。规模化意味着相同任务在更多团队复制，不能被解释为立即向全员开放所有工具；限定规模化适用于低风险任务达标而高风险任务仍需受控；修订后复试用于环境或流程障碍可解决的情形；停止适用于价值不足、风险过高、数据条件不成立或其他方案更优。每项决定都应写明适用任务、岗位、工具、权限、数据、质量阈值、观察期和退出条件。

表 49 | 证据评审四种结论

评审结果	适用条件	后续动作
规模化	行为、质量、环境、经理、经营与治理均过门	复制到相似任务和团队，保持监控
限定规模化	部分任务达标，风险或环境存在边界	限定岗位、权限、数据和场景
修订后复试	价值假设仍成立但系统或设计有缺陷	明确责任、期限和重新基线
停止	价值不足、风险不可接受或替代方案更优	关闭权限、保留学习、重新配置资源

七、不同成熟度组织的投资组合

低成熟度组织不应同时采购多种平台和大规模课程。其首要投资是批准工具、最低治理、一个统一入口、三个任务卡和一名有时间的经理。中等成熟度组织的主要瓶颈通常从上手转向数据、知识资产、流程连接和跨部门审批，应减少通用课程，增加岗位项目、教练和系统改造。高成熟度组织则要防止局部团队形成各自标准和重复建设，需要统一任务分类、风险分级、评估口径和可复用资产，同时允许业务单元在边界内实验。高监管组织应把培训记录、权限记录、模型或工具版本、人工审批和异常处置纳入审计证据。

投资分配也应改变。一个以工作转化为目标的项目，课程制作可能只占总成本的一小部分；更多资源应投入任务分析、真实项目、数据准备、系统连接、质量评审、经理时间、教练支持和评估。若预算只覆盖授课而不覆盖这些环节，管理层应降低项目目标，只把它定义为**基础认知**，而不能要求经营结果。反之，若项目以经营结果立项，就必须把相关流程与技术成本纳入总投资，并设置阶段拨款和停止规则。

表 50 | 不同成熟度的投入重心

成熟度	课程投入	工作流/技术投入	经理/教练投入	评估重点
低	基础认知与安全底线	批准环境、统一入口	种子经理	可用性与低风险行为
中	岗位化练习	数据、知识库、审批、系统连接	在岗教练	质量、周期与复用
高	专家更新与新场景	跨系统编排、监控、治理	经理社区与内部教练	组合价值、风险与跨团队复制
高监管	角色化合规与异常演练	隔离、审计、双人审批	合规联合教练	可追溯、事件与控制有效性

八、董事会季度审议议程

董事会或经营层无需审查课程细节，但应审查能力转化是否成为可治理的投资组合。季度议程可分为五部分。第一，价值：哪些目标任务相对基线改善，收益口径是否与财务和业务一致；第二，采用与质量：使用是否进入正式流程，速度、一次通过率和客户影响是否同步；第三，风险：是否出现数据、版权、错误发布、越权或影子使用，控制是否有效；第四，组织：经理强化是否执行，岗位和决策权是否需要调整；第五，组合：哪些场景扩展、暂停或退出。审议材料应呈现分布和异常，而不是只给平均数，因为平均改善可能掩盖某些团队的显著反噬。

董事会还应防止“AI 能力”成为模糊资本化叙事。只有可复用的工作流、知识资产、质量标准、数据连接、经理机制和稳定行为，才构成可沉淀的组织能力；一次成功演示、明星员工技巧或短期活动不能等同于能力资产。对外披露或内部投资回报估算必须区分已实现收益、可验证的过程改善和条件性潜力。若财务收益主要来自员工自报节省时间，还要回答这些时间是否真正减少成本、增加产出、改善客户体验，还是只是被其他工作吸收。

表 51 | 董事会季度审议表

季度议题	董事会证据	关键追问
价值	基线、改善幅度、样本与观察期	收益是否已实现，还是潜在时间节省？
采用	目标任务覆盖与连续使用	是否进入正式流程，还是偶发试用？
质量	一次通过、错误、投诉与返工	速度提升是否以质量为代价？
风险	事件、接管、越权与影子使用	控制是否真正被使用并有效？
组织	经理执行、岗位变化、资产复用	能力是否依赖少数个人？
组合	扩展、暂停、退出清单	下一笔钱投在哪个任务，为什么？

九、岗位认证：从一次考试改为持续授权

企业可以保留岗位认证，但认证对象应从“知道多少”转为“在规定边界内能否稳定完成任务”。一个可信的认证至少包含知识底线、模拟任务、受控真实项目、经理观察和周期性复核。知识底线用于判断数据、事实、版权、品牌和权限风险；模拟任务检验在异常情境下的判断；受控真实项目证明员工能处理本企业的输入、工具和审批；经理观察证明行为能够在工作节奏中保持。认证通过后获得的不是抽象荣誉，而是与任务相匹配的权限，例如可独立生成内部草稿、可提交公开内容审批、可在限额内形成投放变更包。权限应随任务风险、员工表现和工具变化动态调整。

一次认证不能成为永久授权。模型、平台规则、产品信息、组织政策和监管要求都会变化，员工也可能长期不使用某项能力。企业应设置再认证触发条件：工具或工作流重大更新、岗位调动、连续出现质量问题、发生风险事件、长期未执行目标任务，或高风险权限到期。再认证不必重复完整课程，可以用变更说明、关键情境测试和抽样任务完成。这样既避免证书崇拜，也降低不必要的学习负担。认证数据应服务于权限和支持决策，而不是成为简单排名。

表 52 | 岗位认证的证据结构

认证组件	证明什么	不证明什么
知识底线	理解规则与风险	真实工作中的稳定执行
模拟任务	能处理典型与异常情境	能应对全部业务复杂性
受控真实项目	能在企业环境产出合格结果	长期经营收益
经理观察	行为进入工作节奏	系统不存在障碍
周期复核	能力与权限仍匹配	未来永不需要更新

十、员工沟通、公平与心理契约

AI 能力转化会改变任务分配、评价标准和岗位身份，若沟通只强调效率，员工容易把项目理解为隐性裁员或监控升级。管理层应明确三件事：第一，哪些重复动作将被减少，哪些人类判断变得更重要；第二，哪些使用数据仅用于改进系统与支持员工，哪些可能进入绩效评价；第三，员工在工具错误、数据风险和不合理目标面前有哪些申诉与升级渠道。未经说明就使用细粒度操作日志评价个人，可能损害信任并诱发规避行为。日志首先应用于流程诊断、风险追溯和支持资源配置，只有在规则透明、口径稳定且员工可解释的前提下，才适合进入绩效。

公平还意味着不把已有数字熟练度当成唯一潜力。现场研究显示，经验较少者有时从嵌入式 AI 支持中获得更大收益，但这并不意味着所有新手都会自然受益。不同员工在业务知识、语言、无障碍需求、学习时间和设备条件上存在差异。企业应提供多种练习方式、可访问的材料、合理的学习时间和同伴支持，并用真实任务表现而非自我展示能力判断进步。对于因风险边界而不能使用某些工具的岗位，不应以较低登录率惩罚。对高经验员工，也不应强迫在所有任务使用 AI；其专业价值可能体现在识别工具不适用、纠正错误和沉淀规则。

表 53 | 员工沟通与心理契约清单

沟通主题	管理层应说明	员工权利/责任
岗位变化	哪些任务减少、增加或升级	理解新责任并提出实际障碍
数据使用	哪些日志被收集及用途	查看规则、报告误用
绩效评价	哪些指标进入评价、如何申诉	提供真实证据，不规避风险控制
错误处理	可接受试验与违规红线	及时上报、停止和升级
发展机会	认证、教练与新角色路径	参与练习并沉淀经验

十一、经营收益的财务化边界

将培训收益财务化时，最常见的错误是把“自报每周节省两小时”直接乘以工资并宣称成本节省。时间被节省不等于现金成本减少：员工可能把时间投入其他工作，企业也未减少编制、外包或加班。更稳健的做法是把收益分为三层。第一层是已验证过程改善，如任务周期、返工和等待下降；第二层是已实现经营收益，如外包费用减少、活动提前上线、线索响应提升或风险损失减少；第三层是条件性产能，如在相同人员下可承担更多任务。三层应分别披露，不能混合成单一 ROI。

成本也必须完整。除了课程和许可，还包括经理时间、员工练习时间、数据清理、系统集成、知识资产维护、审核、法务、模型或调用费用、供应商支持和机会成本。某些任务在初期可能因增加核验和审批而变慢，这是建立安全工作流的合理成本；管理层应观察学习曲线，而不是在第一周要求立即降本。反之，若长期需要大量人工修复，说明任务边界、知识源或工具并不适合，应停止扩展。财务评估的目标不是证明 AI 必然正收益，而是支持更好的继续、修订和退出决策。

表 54 | 收益财务化分层

收益类别	可接受证据	财务处理
过程改善	系统时间戳、返工和等待记录	作为运营改善，不直接等同现金
已实现收益	发票、预算、收入、成本或损失变化	可进入财务收益，说明归因
条件性产能	单位人员任务量与质量稳定	作为产能潜力，需说明是否被利用
风险避免	事件概率与损失情景、实际事故下降	谨慎估计，避免虚假精确
学习资产	可复用流程、样例、知识库	视为能力资产，不随意货币化

十二、最小可行的企业 AI 营销能力转化标准

一个项目只有同时满足以下条件，才应被称为企业 AI 营销能力转化：存在明确的经营结果和基线；目标被拆成可观察岗位任务；批准工具、数据、权限、知识源与质量门已经就绪；学员在模拟与受控真实项目中反复练习并获得及时反馈；一线经理在例会、复核和绩效中持续强化新行为；采用、行为、质量、周期、经营与风险证据被分层记录；项目能够根据证据扩展、限定、修订或停止。缺少其中任何一项，项目仍可作为认知教育、工具入门或探索活动，但不应对外或向董事会宣称已经完成工作改造。

这一标准也给出了预算纪律。认知教育可以按覆盖率和掌握度验收；任务级试点必须按真实产物、质量和周期验收；流程级改造必须加入数据、权限、系统和经理责任；经营级项目则必须有基线、观察窗口和归因边界。不同层级的目标、成本和证据不能混用。企业不应要求一场通识课承担经营转型责任，也不应允许一个高成本转型项目只提交课程满意度。最成熟的做法不是追求所有员工同时达到同一水平，而是让不同岗位在适合的任务、权限和风险边界内形成稳定、可复用、可审计的人机协作能力。

PUBLISHER

关于竹势 AI 营销智库

竹势 AI 营销智库出品

竹势 AI 营销智库聚焦 AI 市场部与 AI 营销，面向企业创始人、CEO、CMO 及市场增长团队，构建“6+1”知识架构。六大核心模块组成“道、法、术、器、技、例”六脉体系：“道”负责认知刷新，研究 AI 时代的营销本质与战略判断；“法”关注体系建设，沉淀可持续、可复制的方法论；“术”提供落地路径，把策略转化为具体行动；“器”评测工具与平台，帮助企业选择合适的能力组合；“技”分享可以立即应用的实战技巧；“例”通过标杆案例验证方法与成效。“+1”专题围绕 AI 营销的重要趋势与关键经营问题，推出系统、深入、可下载的专题白皮书。

竹势智库通过专业研究、实践经验与管理框架连接认知、决策和执行，帮助企业看清方向、减少试错，逐步建设能够创造真实经营价值的 AI 市场部。