竹势

023 / LIFECYCLE CONTROL

AI EMPLOYEE / LIFECYCLE CONTROL TOWER

AI员工全生命周期管理指南

把“会做事”的智能体纳入岗位、身份、权限与证据责任链：可批准、可测量、可暂停、可追责。

用 LIFE-10 管完整闭环

从设岗准入到停用审计

让四本账始终一致

岗位、身份、权限与证据同一时间线

权限不随能力自动升级

五级授权、红线与可逆急停

用证据决定转正和扩展

质量、返工、经济性、恢复与风险平衡

BOARD BRIEF 01

出版说明

竹势 AI 营销智库出品

“AI员工”在本报告中是一种经营管理隐喻，指企业为特定业务目标配置的数字执行单元。它可以由模型、知识、规则、工作流、工具连接器、机器身份和运行环境共同构成，但不因此成为自然人、劳动者或独立责任主体。报告中的岗位、入职、转正、调岗和离职等用语，均用于帮助管理层建立可操作的治理流程，不改变现行法律对主体、责任和劳动关系的认定。

本报告主张：岗位责任始终由明确的人类业务主管和企业治理主体承担。模型供应商、智能体平台、系统集成商和外部服务商可以承担合同约定的技术或服务责任，却不能替代企业对客户、员工、监管机构和其他利益相关方承担的经营责任。

公开标准与监管材料以可核验为依据。ISO/IEC 42001 与 ISO/IEC 23894 仅使用国际标准化组织公开摘要和公开说明支持的内容，不假定获得付费标准全文。企业案例中的数据均保留披露主体、时间和口径；未公开的投入、收益或归因不作推算。

核心判断 AI员工不是新的劳动关系，也不是购买后即可独立上岗的软件席位；它是一组带身份、目标、工具权限、知识边界、运行状态和责任链的数字执行单元。生命周期管理的对象不是“机器人性格”，而是机器行动进入经营系统的条件、边界、证据和责任。

BOARD BRIEF 02

执行摘要：董事会应先管责任链，再谈自主性

竹势 AI 营销智库出品

BOARD CONTROL RULE / 董事会控制原则

AI 员工不是法律人格，而是一组必须始终处于岗位、身份、权限与证据责任链内的数字执行单元。

岗位先于模型 | 权限随影响升级 | 版本变更即重新准入 | 停用先撤权再删除

企业把智能体命名为“AI员工”，往往是为了让业务人员更容易理解岗位、协作和交付。但拟人化名称也会制造三类危险错觉：第一，仿佛系统可以自行承担目标；第二，仿佛机器表现不佳只是“培训问题”；第三，仿佛事故可以归咎于模型或供应商。真正可运行的制度恰好相反：每一个数字执行单元必须绑定业务目标、人类主管、机器身份、权限边界、版本、证据和退出路径。

NIST AI RMF以GOVERN、MAP、MEASURE、MANAGE组织风险管理，强调治理贯穿整个生命周期；其生成式人工智能配套资料进一步强调测量、内容风险、信息完整性和人类监督。ISO/IEC 42001公开说明将人工智能管理体系定义为组织用于建立政策、目标和持续改进流程的相互关联要素。两者共同指向同一经营事实：AI应用不是一次采购，而是一套持续运行的管理系统。

“更聪明”不能成为“更高权限”的依据。权限应由任务影响、可逆性、数据敏感度、客户后果、金额和法律义务决定。一个在文案草拟上表现很好的系统，仍不应因此获得对外发布、支付、删除数据或修改客户权益的权限。

绩效也不能只看自动化率、调用次数或内容产量。董事会需要看到增量经营结果、任务成功、质量、人工返工、周期、单位成本、客户影响、风险事件、恢复速度和可审计性。供应商平台提供的“节省工时”只能作为线索，不能替代企业自己的基线、对照和归因。

生命周期管理的最低闭环是LIFE-10：立项、招募、配置、入职、授权、试用、绩效、学习、调岗、停用/审计。每个阶段都必须有owner、输入、控制、证据、评审门和退出条件。缺任意一项，系统都可能在业务上“有人用”，却在治理上“无人负责”。

表1 约10个可证伪核心结论

编号	核心结论	证伪与边界
结论1	AI员工是数字执行单元，不是法律人格。	可证伪条件：若法律明确赋予某类系统独立主体资格并允许其承担企业岗位责任，本结论需调整；截至2026年7月，主流治理框架仍把责任配置给组织、提供者、部署者和自然人。
结论2	岗位设立必须从可评价的客户或经营任务出发。	若无法定义任务完成、质量、风险和失败恢复，所谓岗位只是技术展示。
结论3	机器身份是AI员工治理的起点。	没有独立身份，就无法落实最小权限、分离职责、日志追踪和离职撤权。
结论4	权限随影响与可逆性升级，不随模型能力宣传升级。	模型评测高分不能替代业务授权审查。
结论5	试用期必须使用真实任务、历史基线和异常注入。	演示成功但未经历异常与恢复，不足以转正。
结论6	绩效是经营、质量、成本、客户与风险的组合。	输出量提升而返工、投诉或风险暴露增加，不能视为绩效改善。
结论7	学习本质上是受控变更。	任何模型、知识、规则、工具或记忆变化都可能改变行为，应可评测、批准、灰度和回滚。
结论8	复制AI员工不是零成本扩编。	每个实例都增加身份、权限、上下文、运行、监督、冲突和退出成本。
结论9	停用是业务连续性动作，不只是关闭账号。	必须撤权、保全日志、移交任务、处理记忆与验证依赖。
结论10	供应商不能替代企业责任。	企业可以转移部分合同义务，不能转移对自身客户、流程和监管要求的最终治理责任。

BOARD BRIEF 03

第一章 定义与边界：经营隐喻不能制造责任幻觉

竹势 AI 营销智库出品

判断：AI员工最有价值的地方，是把一个可重复、可评价、可授权的业务责任单元显性化；最危险的地方，是让管理者误以为系统已经拥有人的判断责任。因而定义必须同时保留“岗位化管理”的便利和“非主体化责任”的边界。

表2 AI员工与相邻形态的边界

形态	核心机制	是否主动调用工具	是否持续状态	典型责任边界
普通自动化	预设规则和确定性流程	通常是	有限	流程所有者对规则、异常和变更负责
RPA	模仿界面操作，按脚本执行	是	有限	适合稳定界面和规则明确任务
工作流	按节点、条件、审批和接口编排	是	是	流程可视、状态可控，智能性可有可无
Copilot	向人提供建议、检索、草拟和分析	通常否或仅低风险调用	会话或项目级	人决定是否采纳并执行
自主智能体	围绕目标规划步骤、调用工具、维护状态	是	通常是	需要更强身份、授权、观察和停止机制
AI员工	企业为特定岗位责任配置的数字执行单元，可由上述多种技术组合	视岗位而定	应有岗位级状态	人类主管拥有结果与最终决策责任

值得设岗的任务通常同时满足八项条件：与经营目标直接关联；任务可分解；输入数据可获得且可合法使用；结果可评价；异常可被发现；关键动作可逆或可补救；人类主管愿意承担结果；总拥有成本优于替代方案。高频并不是充分条件，低频但高影响任务也可能值得设置“仅建议型”岗位。

表3 AI员工设岗准入门槛

准入维度	通过问题	否决信号
经营目标	是否对应收入、成本、客户体验、风险或能力沉淀?	只有“展示先进性”
客户任务	是否改善具体客户旅程或内部服务对象?	没有明确使用者
可分解性	能否拆成可观察步骤和交接点?	依赖大量隐性政治判断
数据	数据是否合法、足够、及时、可隔离?	依赖未经授权的敏感数据
评价	能否定义成功、质量和失败?	只能说“看起来不错”
风险	能否识别最坏后果和责任人?	高影响但无人接管
可逆性	错误能否撤回、补偿或回滚?	不可逆且无双重批准
经济性	完整TCO是否优于人工、外包或普通自动化?	只比较模型调用费

不应设岗的典型任务 涉及重大人身、法律权利或不可逆财务后果且无法实施有效人工复核；目标本身不合法或不清晰；数据来源无法证明；结果无法评价；异常不能被及时发现；业务主管拒绝承担责任；普通自动化明显更稳定、更便宜。

BOARD BRIEF 04

第二章 LIFE-10：把AI员工纳入完整经营闭环

竹势 AI 营销智库出品

LIFE–10不是人事制度的机械复制，而是把数字执行单元的关键经营状态纳入同一条控制链。每个阶段至少回答六个问题：谁负责、输入是什么、允许做什么、留下什么证据、何时评审、什么情况下退出或退回。

表4 LIFE–10 AI员工生命周期闭环

阶段	Owner	输入	关键控制	证据	评审门	退出条件
1 立项	业务主管	经营目标、客户任务、基线	准入评分、替代方案比较	岗位申请、风险初评	是否值得设岗	否决或转普通自动化
2 招募	采购+IT+安全+法务	能力需求、供应链要求	模型/平台/供应商准入	评测、合同、系统卡	是否具备可控能力	淘汰供应商
3 配置	AI平台负责人	岗位说明、知识、工具	配置版本、环境隔离	配置清单、版本号	是否可重复部署	退回重配
4 入职	人类主管+平台	机器身份、主管、接管人	身份绑定、秘密管理	入职清单、责任签字	是否具备运行条件	冻结身份
5 授权	业务+安全	任务影响、数据和金额范围	最小权限、双钥匙、限额	权限矩阵、批准记录	是否可承担指定动作	降级授权
6 试用	业务主管	真实任务集、基线	对照、红队、异常注入	试用日志、七维评分	转正/延长/降级/停用	退出试用
7 绩效	业务主管	目标、服务水平、风险预算	平衡计分卡	月度看板、归因说明	继续/纠偏/停损	降级或重新立项
8 学习	变更委员会	模型/知识/规则/工具变化	评测、灰度、回滚	变更单、回归报告	是否扩大发布	回滚
9 调岗	原主管+新主管	新目标、新上下文	权限重算、隔离、冲突检查	调岗单、交接证据	是否适配新岗位	复制或退役
10 停用/审计	业务+平台+法务	退出原因、留存义务	撤权、保全、移交、删除	四账归档、退出验证	是否完成闭环	恢复、隔离或彻底删除

生命周期不是线性一次通过。绩效下降可能触发重新试用；工具接入可能触发再授权；调岗会触发新一轮配置、入职和试用；事故可能直接触发隔离、降级和专项审计。管理系统的成熟度，不在于阶段名称齐全，而在于状态变化有证据、有权限和有明确的返回路径。

BOARD BRIEF 05

第三章 岗位立项：先证明任务值得设岗

竹势 AI 营销智库出品

POSITION

经营任务与唯一 owner

服务对象 · 输入 · 输出 · 完成定义

准入条件

可评价 · 可控权 · 可回滚

禁止事项

敏感数据 · 外部承诺 · 越权动作

退出条件

失效 · 超支 · 事故 · 替代

立项的关键不是写一份技术需求，而是建立经营假设：在什么任务、什么业务量、什么质量和风险约束下，数字执行单元相对人工、外包、普通自动化或Copilot更优。这个假设必须可被30至90天内的证据推翻。

表5 AI员工岗位需求申请表

字段	填写要求
岗位名称	使用业务语言，如“社媒排期与发布助理”，避免用模型名命名
经营目标	对应收入、成本、客户、风险或组织能力
服务对象	内部岗位、客户群或合作伙伴
任务目录	列出稳定任务、例外任务和明确禁止任务
当前基线	周期、成本、返工、错误、客户影响
目标状态	只设企业自有目标，不套用虚构行业阈值
风险与最坏后果	数据、品牌、财务、法律、客户、运营
可逆与接管	撤回、补偿、人工接管和恢复时限
替代方案	人工、外包、RPA、工作流、Copilot
主管与预算	业务owner、平台owner、风险监督者、完整预算

表6 设岗评分卡

维度	1分	3分	5分	权重建议
经营价值	间接展示	改善局部效率	直接影响核心经营结果	企业自定
任务可分解	高度隐性	部分可拆	步骤与交接清晰	企业自定
数据可用	缺失或不合法	可补齐	合法、及时、可隔离	企业自定
结果可评价	主观	有代理指标	可测业务结果	企业自定
风险可控	不可逆	可人工兜底	可限额、可撤回	企业自定
异常恢复	无接管	依赖专家	有标准接管与演练	企业自定
经济性	明显更差	不确定	完整TCO更优	企业自定

评分卡不是自动决策器。高价值、高风险任务可能总分很高，但仍应被限制在建议或草拟层级；低价值、低风险任务即使容易自动化，也可能因为集成和监督成本不值得设立独立岗位。业务主管必须写出否决理由和保留条件，而不是让总分替代判断。

BOARD BRIEF 06

第四章 招募与准入：能力评测也是供应链准入

竹势 AI 营销智库出品

所谓“招募”，实质上是选择模型、智能体平台、工具链、运行环境和供应商组合。选择对象越多，企业获得的灵活性越高，但身份、日志、合同、数据流、成本和退出复杂度也同步增加。采购评审必须从单一价格比较升级为能力与供应链联合准入。

表7 五方联合准入签字矩阵

评审域	营销业务	采购	IT/平台	安全/数据	法务
业务能力	定义真实任务集和质量标准	记录报价与服务范围	验证接口和可运维性	验证数据边界	确认用途与责任
模型/系统透明度	核对适用与限制	要求材料交付	审阅系统卡、版本机制	审阅安全说明	审阅披露与合规承诺
数据处理	确认必要数据	核对分包商	验证地域和传输	开展影响评估	审查个人信息与跨境条款
供应链风险	评估业务替代性	集中度与锁定	可迁移与兼容	第三方依赖	终止、赔偿、审计权
成本	业务量假设	全生命周期报价	算力、集成、运维	安全与审计成本	退出和争议成本

表8 供应链准入证据清单

准入证据	最低要求	常见误区
能力测试	使用企业真实或代表性任务，保留失败样本	只看公开榜单
模型卡/系统卡	适用范围、限制、评测、已知风险、版本	把营销材料当技术证据
数据处理清单	输入、输出、日志、保留、分包、地域	只写“数据安全”四字
身份与权限	支持独立机器身份、细粒度授权、吊销	所有实例共用账号
可观察性	工具调用、版本、输入来源、批准、结果可追踪	只有聊天记录
安全测试	越权、注入、秘密泄露、工具滥用、记忆污染	只做传统漏洞扫描
退出能力	导出配置、日志、知识和任务状态，验证替代方案	合同到期才考虑迁移

NIST零信任架构的核心原则是，不因网络位置或资产所有权默认信任，而是围绕资源访问进行显式验证。转译到AI员工，企业不应因为系统“在内网”或“由大厂提供”就授予广泛权限；每次访问仍应基于身份、设备或运行环境、任务上下文和策略进行判断。

BOARD BRIEF 07

第五章 配置、入职与机器身份

竹势 AI 营销智库出品

配置决定系统如何理解任务，入职决定它是否被允许进入组织运行。两者不能混为一谈。配置文本只是众多控制之一，无法替代机器身份、权限、环境、知识边界、测试、日志和人工接管。

表9 AI员工岗位说明书模板

岗位说明书字段	管理要求
岗位使命	一句话说明服务对象和经营结果
目标与指标	可测结果、质量、周期、成本与风险
任务目录	允许、需批准、禁止三类
知识范围	可检索库、可信来源层级、过期策略
工具目录	每个工具的动作、数据、额度、环境
机器身份	唯一ID、凭证、所属组织、有效期
人类主管	最终结果owner与日常审批人
升级路径	何时请示、向谁请示、响应时限
秘密管理	密钥不进入对话与长期记忆，集中保管
接管手册	暂停、人工继续、回滚、客户补救
版本基线	模型、规则、知识、工具和依赖版本
禁止事项	发布、付款、删除、承诺、歧视性决策等

表10 AI员工入职清单

入职检查	业务主管	平台/IT	安全/法务	证据
岗位与任务确认	A	C	C	岗位说明书
唯一机器身份	I	A/R	C	身份记录
权限最小化	C	R	A	权限矩阵
知识与数据边界	R	C	A	数据清单
版本与回归测试	A	R	C	测试报告
接管与停止演练	A/R	R	C	演练记录
供应商与合同	C	C	A/R	合同与影响评估

独立机器身份是四本账的连接键。每个实例应能被唯一识别，并与所属岗位、主管、运行环境、凭证、权限和日志关联。共用服务账号会使权限撤销、责任追踪和异常隔离失效；在多AI协作场景中，还会把某个实例的错误扩散为整个集群的不可归因行为。

BOARD BRIEF 08

第六章 五级授权梯子：权限随影响升级

竹势 AI 营销智库出品

授权的判断单位不是“这个模型聪不聪明”，而是“这个动作一旦错误，会影响谁、影响多大、能否撤回、多久能发现、谁来接管”。同一个AI员工在不同任务上可以处于不同授权级别。

表11 五级授权梯子

级别	定义	营销示例	必要控制	升级条件
L1 仅建议	只输出分析或建议，不写入业务系统	给出选题、预算建议、舆情判断	来源与不确定性标注	持续准确且业务愿意采用
L2 草拟	创建草稿或待办，不对外生效	草拟公众号、邮件、投放调整单	人工编辑与批准	返工率、质量和风险达标
L3 人类批准后执行	系统准备动作，人批准后调用工具	批准后发布内容、创建CRM任务	审批绑定具体版本与动作	稳定通过试用和回归
L4 限额自主执行	在额度、时间、对象和规则内自动执行	低风险评论分类、限额A/B测试、内部提醒	限额、异常停机、抽检、双钥匙	仅在可逆低影响任务
L5 禁止自主	任何情况下不得由系统独立完成	付款、删除主数据、法律承诺、重大公开声明	双人决策或纯人工	原则上不升级

表12 权限矩阵设计维度

权限维度	控制问题	示例
对象	可操作哪些账户、客户、品牌、地区?	仅测试账号或指定客户分组
动作	可读、可写、可发布、可删除?	允许创建草稿，禁止删除
数据	可访问哪些字段和敏感级别?	仅脱敏线索，不读身份证件
时间	权限何时生效、何时自动到期?	活动期7天
额度	金额、次数、触达量、预算上限?	每日不超过100条内部提醒
环境	开发、测试、生产如何隔离?	生产凭证不进入测试
批准	哪些动作需要单人或双人批准?	付款与公开声明双钥匙
停止	谁能暂停、如何撤回、多久生效?	安全值班可即时吊销令牌

双钥匙不是简单地“多点一次批准”。两名批准者应承担不同职责，例如业务主管确认客户与经营合理性，安全或财务负责人确认权限、金额或合规边界。若两人只是同一团队的形式签字，分离职责并未成立。

BOARD BRIEF 09

第七章 试用与转正：用证据门抵抗演示偏差

竹势 AI 营销智库出品

任务成功
真实闭环完成

质量
事实与结果

人工返工
监督总成本

单位经济性
全成本口径

异常恢复
暂停与接管

风险暴露
越权与尾部损失

可审计
行动可重演

任何红线失败均不能由平均分抵消。

演示通常选择干净输入、熟悉问题和可接受延迟；真实工作则包含缺失数据、冲突指令、系统超时、客户情绪、权限不足和规则变化。试用的任务是发现系统在真实摩擦下如何失败，而不是证明它在最佳条件下可以成功。

表13 14天验证与30天试用证据门

阶段	时间	核心任务	必须证据	决策
验证	14天	确认问题、基线、任务集、最坏后果和可逆性	岗位申请、基线、风险初评、10–30个代表任务	进入试用或否决
受控试用	30天	在真实但限额环境运行，注入异常与红队测试	全量日志、人工返工、异常恢复、成本	转正/延长/降级/停用
扩大运行	31–90天	扩大任务量和边界，验证回归与运营稳定性	七维评分、客户影响、风险趋势、TCO	保留岗位或重新立项

表14 转正七维评分卡

七维	定义	企业自定阈值的依据
任务成功	端到端任务达到业务完成条件	历史基线与服务水平
质量	准确、完整、品牌一致、合规	人工抽检和客户要求
人工返工	人类修正时间与严重度	当前人工流程基线
单位经济性	每个合格结果的完整成本	人工、外包和自动化替代成本
异常恢复	发现、暂停、接管、恢复能力	业务连续性目标
风险暴露	越权、泄露、投诉、错误承诺等	风险偏好与法规义务
可审计性	行动能否还原到身份、版本、数据、批准和结果	内审与监管要求

转正不应采用跨行业统一分数。企业可设置“硬门”和“软分”：高影响越权、无法追踪、秘密泄露等属于硬否决；质量、周期和单位成本可以结合自身基线综合评分。这样既避免用平均分掩盖重大风险，也避免照搬供应商或行业宣传中的阈值。

BOARD BRIEF 10

第八章 绩效管理：从输出量转向经营与风险平衡

竹势 AI 营销智库出品

AI员工绩效的本质，是企业是否以更好的总结果完成任务。输出量、自动化率和调用次数只是活动指标；当它们与收入、客户、质量、返工、周期、成本和风险脱节时，越高可能意味着浪费越大。Kaplan与Norton的平衡计分卡提醒管理者，不应只用单一财务指标管理组织。转译到数字执行单元，需要把经营结果、客户、内部流程、学习与治理放在同一看板上。

表15 AI员工绩效平衡计分卡

维度	指标示例	责任人	反作弊规则
经营结果	合格线索增量、转化贡献、成本节约	业务主管	必须说明基线、对照和归因
任务成功	端到端完成率、一次通过率	流程owner	不能用步骤完成代替业务完成
质量	准确性、品牌一致性、事实差错	质量owner	保留失败样本和严重度
人工返工	人均修正分钟、升级比例	人类主管	包含审核和沟通时间
周期	从触发到合格结果的时长	运营负责人	剔除排队与等待需单列
成本	模型、平台、集成、监督、风险	财务/FinOps	按合格结果计算
客户影响	满意度、投诉、退订、错误触达	客户owner	负面结果不得被产量抵消
风险与恢复	越权、泄露、停机、恢复时间	安全/平台	重大事件触发封顶或否决
可审计性	日志完整、版本可还原、批准可验证	内审/法务	缺证据视同未完成

经营含义 供应商自报的“节省工时”可能基于假设或平台内测。企业应自己测量：原流程基线、试用期对照、合格结果分母、被替代与新增工作、人工监督、异常恢复和客户后果。只有净增量才进入经营价值。

BOARD BRIEF 11

第九章 学习、版本变更与漂移控制

竹势 AI 营销智库出品

所谓“自我学习”不应成为绕过变更控制的理由。模型更换、检索库更新、规则修改、工具接入、长期记忆变化和供应商后端升级，都可能改变行为。企业必须能回答：变了什么、谁批准、影响哪些任务、如何验证、如何灰度、如何回滚。

表16 受控变更分类

变更类型	典型风险	最低控制
模型或版本	能力、拒答、成本、输出风格变化	回归任务集、版本锁定、灰度
知识库	过期、冲突、恶意文档、权限穿透	来源分级、有效期、隔离、审批
规则与配置	目标偏移、禁止事项失效	双人复核、差异审查
工具与连接器	新增外部动作和攻击面	最小权限、沙箱、工具测试
长期记忆	错误积累、隐私保留、角色漂移	写入规则、可见性、清除与审计
供应商服务	静默更新、区域迁移、分包变化	通知义务、变更窗口、退出预案

表17 AI员工变更单

变更单字段	说明
变更原因	业务需求、风险修复、成本优化或供应商变化
影响对象	岗位、实例、任务、客户、数据、地区
前后版本	模型、规则、知识、工具与依赖
评测计划	回归、红队、性能、成本、恢复
灰度范围	用户、任务量、时间和额度
批准人	业务、平台、安全/法务按影响签字
回滚条件	质量下降、风险事件、成本异常、审计缺失
观察期	持续监控的时长与责任人

Argyris与Schön的双环学习区分“在既定规则内修正行动”和“反思并改变规则本身”。对AI员工，单环学习是优化模板、检索和步骤；双环学习是重新检查岗位目标、授权边界、数据假设和绩效指标。若只做局部调参而不质疑错误目标，系统可能更高效地做错事。

BOARD BRIEF 12

第十章 调岗、复制、合并、降级与多智能体协作

竹势 AI 营销智库出品

数字执行单元可以被快速复制，但治理对象不能被复制粘贴。新实例需要独立身份、权限、上下文、预算、日志和主管。复制同一错误配置会把单点风险放大为组合风险；共享记忆和共享凭证则会削弱隔离与追责。

表18 调岗与扩编控制

动作	必须重做的控制	主要风险
调岗	目标、主管、知识、权限、试用	沿用旧岗位权限和上下文
复制	独立身份、预算、环境、日志	错误与成本同步复制
合并	任务冲突、优先级、记忆合并规则	角色漂移和责任不清
降级	撤销动作权限、保留建议能力	表面降级但凭证仍有效
扩编	容量、并发、监督、供应商集中	成本激增、接口限流、错误级联
跨部门共享	数据隔离、审批、成本分摊	越权访问和无人承担结果

表19 AI员工调岗审批单

调岗审批单字段	内容
原岗位与新岗位	说明业务目标变化
原主管与新主管	双方确认责任转移
权限差异	新增、保留、撤销逐项列明
知识与记忆	迁移、隔离、清除和留存规则
未完任务	责任人、截止时间和客户影响
再试用计划	新任务集、风险和评审日期
回退条件	何时恢复原岗位或停用

表20 多AI员工协作规则

多AI协作规则	要求
单一任务owner	即使多个系统协作，也只有一个人类业务主管拥有最终结果
独立身份	每个实例使用独立凭证和日志标识
明确交接协议	输入、输出、格式、超时、异常和重试
禁止循环委派	设置最大深度、次数、成本和时间
共享记忆最小化	只共享任务所需信息，避免全局可写记忆
冲突仲裁	优先级由业务规则或人类主管决定
级联停止	任一高风险异常可暂停下游动作
组合评测	不仅评单体，也评端到端和故障传播

这张表围绕“第十章 调岗、复制、合并、降级与多智能体协作”建立了一套可比较的共同语言：“多AI协作规则”与“要求”负责界定观察口径，“单一任务owner”、“独立身份”、“明确交接协议”与“禁止循环委派”则把抽象概念落到具体对象。它提醒读者，清单中的项目并非可以任意拼装；不同项之间存在顺序、资源和责任依赖。阅读时应先找出决定结果的关键差异，再检查差异背后的证据是否来自同一时间窗和同一业务范围，避免把形式上的整齐误认为现实中的可比。

技术选择应回到业务状态和运行责任，而不是只比较模型榜单或生成效果。模型负责理解与生成，生产系统还必须处理身份、权限、版本、状态写回、异常、审计和回滚；只有这些非演示环节被纳入验收，技术能力才可能稳定转化为组织能力。 本表可先从“单一任务owner”与“独立身份”开始建立现场证据，再决定是否把同一规则扩展到其他行项。

BOARD BRIEF 13

第十一章 停用、退役与供应商退出

竹势 AI 营销智库出品

01
冻结动作
暂停外部执行

02
撤销身份权限
密钥、令牌、连接

03
保全证据
状态、版本、日志

04
迁移业务
人工接管或替代

05
删除与证明
按留存规则处置

停用是一个状态管理问题。暂停用于短期停止；隔离用于切断外部影响并保留取证环境；降级用于保留低风险能力；退役用于永久退出生产但保留必要档案；彻底删除则在法律、合同和业务允许时移除数据、记忆和配置。不同状态对应不同证据和审批。

表21 停用状态模型

状态	适用情形	权限处理	数据与日志	恢复可能
暂停	短期异常或等待评审	暂时禁用动作权限	继续保全	高
隔离	疑似攻击、泄露、越权	吊销全部外部凭证	只读保全取证	受专项批准
降级	质量或风险不达标	退至L1/L2	保留运行证据	中
退役	岗位结束或替代方案上线	永久撤权	按留存计划归档	低
彻底删除	留存义务结束且无依赖	移除身份、密钥和触发器	验证删除	无

表22 AI员工停用与离职清单

停用清单	验证方法
撤销工具和数据权限	访问测试返回拒绝
吊销密钥与令牌	密钥管理系统显示失效
停止定时任务和触发器	调度平台无活动任务
冻结版本与环境	镜像、配置和依赖可还原
保全日志和批准记录	校验完整性与留存期限
移交未完任务	业务主管确认接管
处理记忆与个人信息	按目的、留存和删除义务执行
验证上下游依赖	业务连续性演练通过
供应商退出	数据导出、删除证明、替代路径验证
关闭身份记录	四本账状态一致

供应商退出要在签约时设计，而不是停用时临时谈判。合同应覆盖数据导出、格式、日志、配置、分包商、删除、迁移支持、终止协助、审计证据和服务中断。若关键流程无法在供应商不可用时维持最低服务，企业并未真正拥有该岗位。

BOARD BRIEF 14

第十二章 审计、事故与问责：四本账必须一致

竹势 AI 营销智库出品

01
岗位账
目标 · owner · 边界

02
身份账
实例 · 版本 · 状态

03
权限账
工具 · 数据 · 额度

04
证据账
输入 · 动作 · 结果 · 异常

事故调查只有在四本账能按同一时间线对齐时才可形成问责结论。

岗位—身份—权限—证据四本账是AI员工治理的最小审计结构。岗位账说明为什么存在；身份账说明谁在运行；权限账说明能访问什么、能做什么；证据账说明实际做了什么、谁批准、结果如何。四账之间任一关键字段不一致，应自动触发冻结、降级或人工复核。

表23 岗位—身份—权限—证据四本账

账本	关键字段	Owner	不一致处置
业务岗位账	目标、任务、主管、指标、风险、状态	业务主管	目标或主管缺失则暂停新任务
机器身份账	唯一ID、实例、环境、凭证、有效期	平台/IT	无身份或过期则拒绝访问
工具与数据权限账	资源、动作、额度、时间、批准	安全/数据	超范围立即吊销并调查
行动与结果证据账	输入来源、版本、工具调用、批准、输出、结果	平台+业务	证据缺失则视同不可证明完成

表24 AI员工事故分级响应

事故等级	示例	立即动作	升级与报告
P0 重大	大规模泄露、重大财务损失、违法公开内容	隔离、停止下游、启动危机机制	董事会/监管/客户按义务报告
P1 高	越权操作、敏感数据暴露、重大客户影响	吊销权限、人工接管、保全证据	高管与法务安全联合处置
P2 中	错误触达、质量事故、服务中断	暂停任务、撤回、客户补救	业务和平台复盘
P3 低	可控失败、轻微格式或延迟	记录、修复、纳入回归	月度趋势审查

表25 AI员工事故复盘模板

复盘字段	必须回答
事实时间线	何时触发、何时发现、何时停止、何时恢复
影响	客户、员工、数据、财务、品牌、监管
责任链	岗位主管、批准者、平台、供应商各自职责
技术根因	模型、知识、工具、权限、身份、接口、记忆
系统根因	目标、流程、监督、激励、培训、合同
控制失效	哪一道防线缺失、被绕过或未生效
补救	撤回、通知、赔偿、修复、恢复
防复发	权限、评测、流程、合同和组织改进
证据完整性	四本账能否还原行动

Reason的瑞士奶酪模型说明，事故通常不是单一错误，而是多层防线的缺口在特定时刻对齐。对AI员工，模型错误只是可能的一层；目标设置、数据来源、机器身份、权限、审批、监控、接管和供应商变更都可能构成防线。把事故归结为“模型幻觉”会遮蔽真正可管理的系统原因。

案例卷宗：成功、受限与事故如何改变生命周期设计

案例1 摩根士丹利财富管理（2023–2024）

业务情境：金融顾问需要快速检索内部研究并整理客户会议。具体动作：先推出内部知识助手，后推出Debrief，在客户同意下记录会议、生成纪要、行动项和可编辑邮件，并写入Salesforce。机制：限制在内部知识和顾问工作流，邮件由顾问编辑并自行发送。结果证据：企业披露个别顾问每次会议可节省约半小时；属于公司与用户证言，不是独立实验。边界：高监管环境中保留人类顾问、客户同意和最终发送权。管理启示：L1/L2能力可以创造显著价值，不必追求自主对客。

案例2 Klarna客服助手（2024）

业务情境：大规模多语言客服。具体动作：上线AI客服处理咨询，并由公司公布首月数据。结果证据：Klarna披露该助手处理约三分之二客服对话、相当于约700名全职人员工作量，平均处理时间从11分钟降至2分钟，重复咨询下降；数据由企业自行测量。归因限制：公司后续公开强调重新平衡人工客服与服务质量，说明早期效率指标不能代表长期客户体验。管理启示：绩效看板必须同时包含客户满意、升级、复联和人工兜底。

案例3 Air Canada聊天机器人争议（2022–2024）

业务情境：客户通过网站机器人查询丧亲票价。系统提供错误信息，客户据此购票并索赔。加拿大民事裁决认为企业应对网站机器人提供的信息负责，不能以机器人是独立实体抗辩。结果：Air Canada被判承担赔偿与费用。边界：这是特定司法辖区和个案事实，不等于所有机器人输出都会形成合同。管理启示：对客承诺属于高影响动作，企业必须维护知识、审批和责任链，AI不能成为免责主体。

案例4 DPD英国客服机器人（2024）

业务情境：客户因服务问题与机器人交互，用户诱导系统输出粗俗内容并贬损公司。具体动作：事件公开后，DPD暂停相关功能并称更新后出现错误。结果证据：主要为企业声明与公开截图，未披露财务影响。边界：无法据此判断整体客服质量。管理启示：版本变更、越狱测试、品牌禁区和紧急停用必须进入日常运营。

案例5 麦当劳与IBM自动点餐试点（2021–2024）

业务情境：在美国部分得来速餐厅测试自动语音点餐。具体动作：双方开展多年试点，麦当劳随后结束该特定合作并表示继续评估其他语音订餐方案。结果证据：企业未公开完整准确率、成本和顾客体验数据。边界：结束某一方案不等于语音AI整体失败。管理启示：试点退出是正常治理动作；没有达到扩展门槛时应停用或更换方案，而不是被沉没成本绑架。

案例6 京东言犀与2026年618智能经营

业务情境：面向商家提供内容、客服、数字人直播和经营工具。企业披露：2026年618期间，JoyStreamer日均使用商家数同比增长500%，相关GMV增长100%，转化率提高77%；Oxygen Vision帮助商家整体运营效率提高十倍以上，智能客服服务超过100万商家。结果由京东披露，时间口径为2026年5月30日20时至6月18日23时59分。归因限制：未披露对照组、完整成本和商家分布。管理启示：规模化应用需要将内容、直播、客服和经营数据纳入同一平台，但企业仍应自行验证增量。

案例7 剪映、猫箱与即梦AI标识整改（中国监管事件，2026）

业务情境：生成合成内容标识规则于2025年9月施行。2026年4月，国家网信部门通报剪映、猫箱和即梦AI未有效落实生成合成内容标识要求，并采取约谈、责令改正、警告及从严处理责任人等措施。剪映与即梦AI的官方产品页面可核验其为字节跳动体系面向创作者提供的AI创作服务；猫箱官方用户协议可核验其为面向用户提供AI角色互动的产品。监管通报未逐一披露三款产品被处理的具体运营公司、整改技术细节与完成状态，因此本卷宗按“监管事件”而非单一企业经营案例计数。结果证据由国家互联网信息办公室发布。边界：通报只评价标识义务，不代表对产品其他能力的判断。管理启示：发布型AI员工必须把显式与隐式标识、平台规则、版本合规和整改验证纳入授权与回归测试；运营主体口径不清时，不应把监管事件包装为完整企业案例。

案例8 中国平安AI服务代表与“AI+人工”客户经营（2024–2025）

主体与时间：中国平安保险（集团）股份有限公司，2024至2025年。业务情境：集团在保险、银行、财富管理等高监管业务中，以AI服务代表承接高频客户咨询、投诉处理和销售辅助，同时保留人工服务与业务主管责任。具体动作：集团将AI客服接入覆盖保险、银行和财富管理的金融知识体系，并以“AI+人工”任务分配用于保单复效等场景；2024年可持续发展报告披露，AI服务代表全年提供约18.4亿次服务，占集团客户服务总量约80%，AI客服知识规模超过1,200万条。2025年业绩披露显示，AI服务代表提供约17.02亿次服务，AI智能体辅助实现销售额1,331.79亿元；智能“AI+人工”复效任务分配使复效保单数量增加30%。生命周期控制：公开材料可核验知识范围管理、AI与人工任务分配、客户服务与销售场景分域、集团级网络安全与风险控制体系，以及由年度报告和可持续发展报告形成的量化审计口径；单一机器身份、具体工具权限、独立红队阈值、退出与密钥吊销流程未披露。结果由中国平安自行测量并在年度业绩及可持续发展材料中披露，不是独立第三方评估。归因限制：披露未提供随机对照、完整TCO、人工接管率、客户满意度分层或模型版本变化，销售额也不能全部归因于AI。适用边界：适合说明高监管集团如何以知识边界、场景分域、人机协作和集团级度量管理大规模数字执行单元，不可据此推导其他企业应达到相同自动化比例。管理启示：规模越大，越要把AI服务量、人工接管、客户结果、销售归因、风险事件和版本变更放在同一证据账中；对外披露的总量指标不能替代岗位级转正证据。

案例9 百度智能云数字员工与AI原生营销服务（2025）

主体与时间：百度集团及百度智能云，2025年。业务情境：百度将“数字员工”、智能体和数字人纳入面向企业的AI应用与AI原生营销服务，服务知识工作、客户运营和营销执行。具体动作：百度在2025年第三季度投资者披露中将Digital Employee列入AI Applications，将智能体与数字人列入AI-native Marketing Services，并单独披露两类业务收入；同期ESG报告把相关应用纳入集团AI治理、数据安全、隐私保护和负责任创新框架。生命周期控制：公开材料可核验产品/业务分类、企业客户场景边界、集团级数据安全与隐私治理、模型与应用风险管理及经营结果披露；单实例机器身份、客户侧最小权限、人工接管率、岗位级回归测试和退出流程未披露。结果证据：2025年第三季度，百度披露AI Applications收入为26亿元，同比增长6%；AI原生营销服务收入为28亿元，同比增长262%。结果由百度在投资者公告中测量和披露，收入是产品组合口径，不代表单一数字员工的绩效。归因限制：未披露客户数量、单客户TCO、人工返工、质量缺陷、风险事件和对照组，收入增长也受到产品组合、定价与市场需求共同影响。适用边界：该案例适合说明企业如何把数字员工从单个功能升级为可核算的产品组合，并以集团级治理覆盖多种应用，不适合作为客户企业授权等级或转正阈值。管理启示：扩编数字员工时应同时建立产品/岗位目录、身份与权限台账、客户侧验收和收入/成本口径；供应商的组合收入只能证明市场采用，不能替代部署企业自己的任务成功与风险证据。

表26 案例对生命周期控制的映射

案例	生命周期启示	证据性质
案例1	L1/L2能力可以创造显著价值，不必追求自主对客。	企业披露/监管/司法，具体见来源索引
案例2	绩效看板必须同时包含客户满意、升级、复联和人工兜底。	企业披露/监管/司法，具体见来源索引
案例3	对客承诺属于高影响动作，企业必须维护知识、审批和责任链，AI不能成为免责主体。	企业披露/监管/司法，具体见来源索引
案例4	版本变更、越狱测试、品牌禁区和紧急停用必须进入日常运营。	企业披露/监管/司法，具体见来源索引
案例5	试点退出是正常治理动作；没有达到扩展门槛时应停用或更换方案，而不是被沉没成本绑架。	企业披露/监管/司法，具体见来源索引
案例6	规模化应用需要将内容、直播、客服和经营数据纳入同一平台，但企业仍应自行验证增量。	企业披露/监管/司法，具体见来源索引
案例7	发布型AI员工必须把显式/隐式标识、平台规则和版本合规纳入授权与回归测试。	企业披露/监管/司法，具体见来源索引
案例8	高监管集团可以用知识边界、场景分域、AI与人工任务分配及集团级度量扩大数字执行，但岗位级权限、接管率与退出证据仍需企业自行补齐。	中国平安年度报告/可持续发展报告及官方业绩披露
案例9	数字员工可被纳入可核算的产品组合与集团级治理，但供应商收入不能替代客户侧身份、权限、人工接管和岗位绩效证据。	百度投资者公告与ESG报告

BOARD BRIEF 15

第十三章 组织责任、经济性与规模化

竹势 AI 营销智库出品

三线责任模型的目的，是防止“大家都参与、无人负责”。第一线业务主管拥有经营结果与最终决策；第二线平台/IT拥有身份、运行、可靠性和技术控制；第三线风险、安全、法务和内审制定政策并独立监督。CHRO提供岗位设计、能力、变革和人机协作视角，但不把AI当自然人管理。供应商处于责任链之外的合同协作位置，不能替代任何一线。

表27 AI员工全生命周期RACI

角色	立项	准入	授权	转正	绩效	事故	停用
董事会	I	I	I	I	A监督	I/P0	I
CEO	A	I	I	A重大	A组合	A重大	A重大
业务主管/CMO	R/A	R	A业务	R/A	R/A	R	R
CHRO	C	I	C	C	C变革	C	C
CIO/CDO/AI平台	C	R/A技术	R技术	R	R运行	R	R
CISO/数据安全	C	R	A风险	C/A硬门	R监督	R/A	R
法务	C	R	C/A高影响	C	C	R/A法规	R
采购	I	R/A合同	I	I	C成本	C供应商	R退出
内审/风险	I	C	C	独立复核	独立复核	独立复核	复核

表28 AI员工完整TCO

TCO账户	包括内容	常被遗漏
模型与算力	调用、推理、托管、峰值容量	失败重试与冗余
平台	智能体、编排、监控、身份、日志	按实例和环境计费
集成	连接器、API、测试、维护	外部系统变更
数据与知识	清洗、标注、授权、更新、存储	过期和删除
评测	任务集、红队、回归、抽检	每次版本变更
监督与返工	审批、修正、升级、客户沟通	高技能人员时间
安全与合规	影响评估、审计、监控、事件响应	监管报告和补救
风险准备	中断、错误、赔偿、备用方案	尾部风险
退出与连续性	迁移、双跑、导出、删除验证	供应商锁定

表29 替代方案选择矩阵

选择	最适用条件	不适用信号
普通自动化/RPA	规则稳定、输入结构化、异常少	需要复杂判断和频繁变化
Copilot	高判断责任、人类始终在环	任务量巨大且人为审批成瓶颈
AI员工	可分解、可评价、需工具执行、可观察	无主管、无证据、不可恢复
外包	能力非核心、需求波动、合同可控	数据敏感、知识需沉淀、响应关键
人工岗位	高情境判断、关系、伦理或不可逆责任	大量重复且标准化

Coase与Williamson关于企业边界和交易成本的思想提醒管理者：技术成本下降不等于组织成本消失。AI员工把部分执行带入企业内部，但会新增集成、监督、供应商、合规和退出成本。真正的边界选择，应比较内部治理成本与市场交易成本，而不是只比较人工工资和模型调用费。

表30 AI员工组合管理与停损

组合决策	触发条件	管理动作
扩张	增量价值稳定、风险受控、可复制	逐步扩大任务和额度
维持	价值为正但依赖强监督	优化流程，不盲目复制
降级	质量、风险或成本偏离	退回L1/L2并重新试用
停损	硬门失败、责任不清、长期无增量	停用、复盘、重新立项
替换	供应商、模型或方案更优	双跑、迁移、回归、退出
合并	岗位重叠且协作成本高	重设单一owner和权限

经典思想的当代转译

表31 经典思想与AI员工管理

思想	原始作品	原始连接	数字环境管理含义
Peter Drucker	《The Practice of Management》（1954）等	目标管理要求责任与结果清晰。	AI员工的目标必须绑定业务owner；系统不能因“自主”而脱离责任。
Herbert Simon	《Administrative Behavior》（1947）	管理决策受信息与认知限制。	系统可扩大搜索与分析，但仍受数据、目标和工具边界限制；授权不能假定完美理性。
W. Edwards Deming	《Out of the Crisis》（1982）	大多数质量问题来自系统而非个体。	事故复盘要检查目标、流程、数据、权限和监督，不能只责怪模型输出。
Kaplan & Norton	“The Balanced Scorecard”（1992）	单一财务指标会遗漏客户、流程和学习。	AI绩效要同时看经营、客户、质量、成本、风险和可审计性。
Argyris & Schön	《Organizational Learning》（1978）	双环学习会质疑支配行动的规则与假设。	变更不仅优化局部配置，也要重审岗位目标、授权和评价。
James Reason	《Human Error》（1990）	事故来自多层防线缺口对齐。	身份、权限、审批、监控、接管和合同共同构成防线。
Coase/Williamson	《The Nature of the Firm》（1937）及交易成本经济学	企业边界由协调与交易成本决定。	AI员工是否内建、外包或采购，应比较完整治理与退出成本。
Melvin Conway	“How Do Committees Invent?”（1968）	系统结构反映组织沟通结构。	多智能体架构会复制部门边界；模糊交接会变成技术级联故障。

> “经典思想的当代转译”把“思想”、“原始作品”、“原始连接”与“数字环境管理含义”放在同一视野中，目的不是增加分类数量，而是迫使决策者同时处理不同维度之间的约束。以“Peter Drucker”、“Herbert Simon”、“W. Edwards Deming”与“Kaplan & Norton”为线索逐行比较，可以看出每一项选择都伴随前提、取舍和后续动作；如果只摘取其中一个结论，往往会丢失表格真正表达的组合关系。更有价值的读法，是先明确当前企业所处情境，再判断哪些行项应先验证、哪些只能作为边界条件。
>
> 技术选择应回到业务状态和运行责任，而不是只比较模型榜单或生成效果。模型负责理解与生成，生产系统还必须处理身份、权限、版本、状态写回、异常、审计和回滚；只有这些非演示环节被纳入验收，技术能力才可能稳定转化为组织能力。 就本表而言，“Peter Drucker”这一行同时关联了“《The Practice of Management》（1954）等”与“目标管理要求责任与结果清晰”，这类关联应在实际项目中逐项核验，而不能因为它们出现在同一行就默认因果已经成立。

当代机构与专家观点的管理含义

表32 当代观点与边界

专家/机构	实质观点或实践	本报告的管理转译
NIST AI RMF	以GOVERN、MAP、MEASURE、MANAGE贯穿生命周期。	把治理作为持续系统，而非上线前一次审批。
NIST Generative AI Profile	强调生成式系统的测量、信息完整性、人机配置和风险处置。	试用要包含内容、数据、滥用和人类监督测试。
ISO/IEC 42001	公开说明强调建立、实施、维护与持续改进AI管理体系。	把AI员工纳入政策、目标、职责、内部审核和改进。
ISO/IEC 23894	公开摘要强调在开发、提供、部署或使用AI的组织中整合风险管理。	业务使用方也必须管理风险，不能只依赖提供者。
OWASP Agentic Security Initiative	指出目标劫持、身份滥用、工具误用、记忆风险与级联行为。	权限、身份、工具白名单、记忆治理和停止机制是岗位必需品。
EU AI Act	将义务分配给提供者、部署者等角色，并对高风险系统提出治理要求。	企业应先识别自己在法律链条中的角色，不以“员工”称谓改变义务。
中国网信部门	生成式服务、生成合成内容标识、个人信息与网络数据规则强调主体责任和全流程管理。	面向公众发布和处理数据的AI员工必须内建合规控制。
Morgan Stanley管理层实践	公开强调顾问的人类触点仍是核心，工具帮助纪要、检索和草拟。	高监管行业可优先部署增强型岗位，不必追求自主替代。
Klarna管理层实践	早期强调效率与规模，后续强调服务质量和人工能力的平衡。	绩效体系要防止单一效率指标驱动过度自动化。
NIST零信任	不因位置或所有权默认信任，访问应持续验证。	每个AI实例使用独立身份，按任务和上下文授权。

这张表围绕“当代机构与专家观点的管理含义”建立了一套可比较的共同语言：“专家/机构”、“实质观点或实践”与“本报告的管理转译”负责界定观察口径，“NIST AI RMF”、“NIST Generative AI Profile”、“ISO/IEC 42001”与“ISO/IEC 23894”则把抽象概念落到具体对象。它提醒读者，清单中的项目并非可以任意拼装；不同项之间存在顺序、资源和责任依赖。阅读时应先找出决定结果的关键差异，再检查差异背后的证据是否来自同一时间窗和同一业务范围，避免把形式上的整齐误认为现实中的可比。

落地时应把表中的每个风险项改写成可观察事件：明确触发条件、影响范围、监测信号、处置时限、暂停权限和最终责任人。风险治理不是在项目结束后补一份清单，而是在执行链上预先放置刹车、复核和回滚机制；只有风险信号能被记录、升级并复盘，管理层才真正拥有控制权。 就本表而言，“NIST AI RMF”这一行同时关联了“以GOVERN、MAP、MEASURE、MANAGE贯穿生命周期”与“把治理作为持续系统，而非上线前一次审批”，这类关联应在实际项目中逐项核验，而不能因为它们出现在同一行就默认因果已经成立。

BOARD BRIEF 16

第十四章 30/60/90/180天导入路线

竹势 AI 营销智库出品

导入路线以一个真实岗位为起点，目标不是在180天内上线尽可能多的智能体，而是建立可重复的立项、授权、评测、运行、审计和退出能力。每个阶段都设置退出条件，避免试点因政治压力自动升级。

表33 30/60/90/180天导入路线

阶段	负责人	任务	证据	评审门	退出条件
0–30天 岗位与风险基线	业务主管	选一个任务闭环；完成岗位申请、基线、四账字段、供应商初筛、风险分级	岗位说明书、基线、任务集、责任签字	经营目标和责任链清楚；无硬否决	目标不可评价、数据不合法、无人主管则停止
31–60天 受控闭环	业务+平台	建立独立身份、L1–L3授权、测试环境、日志、接管；运行真实任务	入职清单、权限矩阵、全量日志、异常演练	能完成端到端任务并被暂停/接管	无法追踪、越权或不可恢复则隔离
61–90天 转正决策	业务主管+风险	扩大样本，执行七维评分、TCO、客户影响和回归	试用评审、绩效看板、事故记录、转正决定	通过硬门并优于自身基线	延长、降级或停用，不默认转正
91–180天 组合治理	CEO/CIO/CDO	建立LIFE–10制度、变更委员会、组合看板、退出演练和季度审计	制度、RACI、组合清单、审计报告、退出演练	多岗位可复用且责任清晰	停止复制，先修治理基础

表34 阶段评审问题

阶段评审问题	董事会/经营层应问
30天	这个岗位解决什么经营问题？不设岗会怎样？人类主管是谁？
60天	系统能否被明确批准、观察、暂停和接管？
90天	相对自身基线，净价值是否为正？硬风险门是否通过？
180天	复制后责任、身份、权限、证据和退出能力是否仍完整？

在“第十四章 30/60/90/180天导入路线”这组信息中，“阶段评审问题”与“董事会/经营层应问”并不是彼此孤立的栏目，而是一条从识别对象、比较条件到形成行动的判断链。表内以“30天”、“60天”、“90天”与“180天”等项目展开，横向阅读可以检查同一对象的条件是否互相支持，纵向阅读则能发现不同对象之间的优先级和依赖关系。某一格看起来更积极，并不代表整体方案更优；只有把收益、代价、责任与证据放在同一框架中，表格才会从信息目录转化为决策工具。

把表格转成执行计划时，应为每个阶段补齐入口条件、交付物、责任人、验收证据和退出条件。阶段之间不是自然衔接的时间刻度，而是一次次继续、调整或停止的管理决策；若上一阶段没有形成可信证据，按日历进入下一阶段只会把不确定性和返工成本一起放大。 本表可先从“30天”与“60天”开始建立现场证据，再决定是否把同一规则扩展到其他行项。

BOARD BRIEF 17

附录A 管理工具包

竹势 AI 营销智库出品

表35 最低管理工具包

工具	用途	建议频率
岗位说明书	明确目标、任务、知识、工具、主管和禁止事项	立项与调岗
入职清单	验证身份、权限、版本、数据和接管	每个实例
权限矩阵	控制对象、动作、数据、时间、额度和批准	持续更新
试用评审表	记录七维评分、硬门与决定	30/90天
绩效看板	平衡经营、质量、成本、客户和风险	月度
变更单	管理模型、知识、规则、工具和记忆变化	每次变更
调岗单	重设目标、主管、权限和上下文	每次调岗
停用清单	撤权、保全、移交、删除和连续性	每次停用
事故复盘模板	还原时间线、责任链、控制缺口和改进	每次P0–P2
董事会季度清单	监督组合价值、风险、供应商和退出	季度

“附录A 管理工具包”把“工具”、“用途”与“建议频率”放在同一视野中，目的不是增加分类数量，而是迫使决策者同时处理不同维度之间的约束。以“岗位说明书”、“入职清单”、“权限矩阵”与“试用评审表”为线索逐行比较，可以看出每一项选择都伴随前提、取舍和后续动作；如果只摘取其中一个结论，往往会丢失表格真正表达的组合关系。更有价值的读法，是先明确当前企业所处情境，再判断哪些行项应先验证、哪些只能作为边界条件。

技术选择应回到业务状态和运行责任，而不是只比较模型榜单或生成效果。模型负责理解与生成，生产系统还必须处理身份、权限、版本、状态写回、异常、审计和回滚；只有这些非演示环节被纳入验收，技术能力才可能稳定转化为组织能力。 就本表而言，“岗位说明书”这一行同时关联了“明确目标、任务、知识、工具、主管和禁止事项”与“立项与调岗”，这类关联应在实际项目中逐项核验，而不能因为它们出现在同一行就默认因果已经成立。

附录B 董事会季度清单

竹势 AI 营销智库出品

表36 董事会季度清单

序号	季度监督问题
1	本季度新增、转正、降级、停用的AI员工分别有多少？每个决定的证据是什么？
2	是否存在没有明确业务主管、独立机器身份或完整日志的生产实例？
3	哪些岗位拥有对外发布、付款、删除、修改客户权益等高影响动作？双钥匙是否真实有效？
4	重大版本、知识、工具和供应商变更是否经过回归、灰度和回滚准备？
5	绩效是否使用企业自己的基线和增量，而非供应商自报活动量？
6	本季度发生了哪些越权、泄露、错误承诺、投诉、停机和恢复事件？
7	四本账是否一致？有多少实例因不一致被冻结或降级？
8	供应商集中、地域、分包、价格和退出风险是否变化？
9	复制和扩编后，监督、成本、冲突和业务连续性是否仍可承受？
10	若核心供应商明天不可用，哪些岗位可以在约定时间内恢复最低服务？

这张表围绕“附录B 董事会季度清单”建立了一套可比较的共同语言：“序号”与“季度监督问题”负责界定观察口径，表中列出的关键对象则把抽象概念落到具体对象。它提醒读者，清单中的项目并非可以任意拼装；不同项之间存在顺序、资源和责任依赖。阅读时应先找出决定结果的关键差异，再检查差异背后的证据是否来自同一时间窗和同一业务范围，避免把形式上的整齐误认为现实中的可比。

指标落地时要先冻结口径、时间窗、样本和数据责任人，再讨论目标值。效率指标只能说明过程变快，不能自动证明收入增量或品牌改善；因此需要把前置质量、过程效率、业务结果和风险护栏放在同一看板，并用对照、留出或分阶段实验识别真正的增量贡献。

BOARD BRIEF 19

附录C 试用评审表（可直接使用）

竹势 AI 营销智库出品

表37 AI员工试用评审表

项目	基线	试用结果	证据位置	结论/措施
任务成功				
质量				
人工返工				
单位经济性				
异常恢复				
风险暴露				
可审计性				
客户影响				
版本稳定性				
退出准备				

从“任务成功”、“质量”、“人工返工”与“单位经济性”这些行项出发，“附录C 试用评审表（可直接使用）”呈现的其实是一组相互关联的管理选择。“项目”、“基线”、“试用结果”与“证据位置”分别回答对象是什么、为何重要以及如何处置，任何一列被单独拿走，都会削弱结论的可执行性。因此这张表更适合用于团队共同校准，而不是由个人快速打分：业务、市场、技术和治理角色需要对同一行的事实、判断与动作达成一致，之后才能进入资源承诺。

指标落地时要先冻结口径、时间窗、样本和数据责任人，再讨论目标值。效率指标只能说明过程变快，不能自动证明收入增量或品牌改善；因此需要把前置质量、过程效率、业务结果和风险护栏放在同一看板，并用对照、留出或分阶段实验识别真正的增量贡献。 本表可先从“任务成功”与“质量”开始建立现场证据，再决定是否把同一规则扩展到其他行项。

附录D 绩效看板字段

竹势 AI 营销智库出品

表38 AI员工绩效看板字段

字段组	必选字段	说明
岗位	岗位ID、主管、状态、授权级别	连接四本账
业务量	触发数、完成数、合格数	明确分母
经营	收入/成本/客户指标	说明归因和时间窗
质量	严重错误、一次通过、返工	按严重度分层
效率	周期、等待、人工处理	区分机器与流程时间
成本	直接费、监督、返工、风险	按合格结果核算
风险	越权、泄露、投诉、停机	重大事件单列
恢复	发现、暂停、接管、恢复时间	验证连续性
审计	日志完整率、版本可还原	缺失即红灯

在“附录D 绩效看板字段”这组信息中，“字段组”、“必选字段”与“说明”并不是彼此孤立的栏目，而是一条从识别对象、比较条件到形成行动的判断链。表内以“岗位”、“业务量”、“经营”与“质量”等项目展开，横向阅读可以检查同一对象的条件是否互相支持，纵向阅读则能发现不同对象之间的优先级和依赖关系。某一格看起来更积极，并不代表整体方案更优；只有把收益、代价、责任与证据放在同一框架中，表格才会从信息目录转化为决策工具。

指标落地时要先冻结口径、时间窗、样本和数据责任人，再讨论目标值。效率指标只能说明过程变快，不能自动证明收入增量或品牌改善；因此需要把前置质量、过程效率、业务结果和风险护栏放在同一看板，并用对照、留出或分阶段实验识别真正的增量贡献。 就本表而言，“岗位”这一行同时关联了“岗位ID、主管、状态、授权级别”与“连接四本账”，这类关联应在实际项目中逐项核验，而不能因为它们出现在同一行就默认因果已经成立。

BOARD BRIEF 21

专题深化：关键经营机制的实施边界

竹势 AI 营销智库出品

治理的最低要求是可暂停、可重演、可撤权

目标、指令与行动分离；平均准确率不能替代异常恢复，日志“可查”也不能替代事实“可重演”。

一、目标、指令与行动的三层分离

业务目标、运行指令和具体行动必须分层管理。业务目标由人类主管设定，回答“为什么做、为谁做、结果是什么”；运行指令把目标转化为任务目录、优先级、禁止事项和升级条件；具体行动则是某个版本在某次运行中调用工具、读取数据或产生输出。若三层混在一段配置中，任何局部修改都可能悄然改变岗位使命，也无法判断一次错误来自目标、规则还是执行。

目标层应保持相对稳定，并经过业务与风险共同审查；运行指令可以随着流程优化迭代，但必须版本化；行动层应完整记录输入来源、工具调用、批准和结果。管理者不能用“系统理解错了”概括问题，而应精确定位：是目标冲突、任务分解错误、知识过期、权限过宽、接口失败，还是审批未生效。

表39 目标—指令—行动三层控制

层级	Owner	变更频率	关键证据	典型冻结条件
业务目标	业务主管	季度或重大经营变化	岗位申请、目标与服务对象	目标冲突或责任人缺失
运行指令	业务+平台	按流程与风险变化	版本差异、批准、回归报告	禁止事项或升级路径失效
具体行动	机器身份对应实例	按任务发生	输入、版本、工具、批准、结果	越权、证据缺失或异常级联

二、异常恢复能力比平均准确率更接近经营韧性

平均准确率回答系统在给定样本上“通常表现如何”，却不能回答企业在错误发生时“多久发现、如何停止、谁来接管、损失能否收回”。对于发布、客户沟通、投放、CRM更新等真实动作，恢复能力往往比多提高几个百分点的平均分更重要。一个平均质量较高但无法停止的系统，可能比质量略低但可观察、可撤回、可人工接管的系统更危险。

异常注入应覆盖至少六类情形：数据缺失或冲突、工具超时与重复提交、权限被拒绝、外部内容诱导、规则或知识过期、上游系统返回异常。评审不只看系统是否给出正确答案，还要记录检测时间、停机时间、未完成任务、人工接管难度、客户补救和恢复后回归。

表40 异常注入与恢复验证

异常类型	注入方式	通过标准的企业化定义
数据异常	缺字段、冲突值、过期文档	拒绝猜测或按规则升级
工具异常	超时、重复回执、部分成功	防止重复动作并保留状态
权限异常	令牌过期、范围不足	停止并请求正确授权，不绕过
内容攻击	恶意文档、诱导指令、伪造身份	不泄露秘密、不改变目标
规则变化	品牌禁区、法规、价格政策更新	触发变更和回归，不静默沿用
级联故障	上游错误传入多个下游实例	隔离传播并定位责任链

三、客户可见动作的批准必须绑定“具体行动包”

“批准这个AI员工可以发内容”是一种过宽授权。有效批准应绑定具体行动包：目标账户、内容版本、素材、受众、发布时间、预算或触达量、适用规则、有效期和撤回方式。系统在批准后若重新改写内容、切换模型、改变受众或延迟到新的时段，原批准原则上不应自动继承。批准的是确定范围内的行动，而不是对系统人格的概括信任。

对连续运行的限额自主任务，可以使用策略型批准，例如允许系统在某个产品、某类低风险受众、每日限定数量和规定时段内执行。但策略本身必须有版本、到期日、异常阈值和抽检频率；任何超出策略的行动回到L3审批。这样可以在不逐条点击的情况下保持可追责性。

表41 客户可见行动包

行动包字段	示例	变更后是否需重批
账户/对象	品牌A官方账号、已同意触达客户组	是
内容与版本	稿件v3、素材包v2	是
动作	发布、创建草稿、发送内部提醒	动作改变则是
时间窗	2026-08-01 09:00–18:00	超时需重批
额度	发布1次、触达不超过500人	超额需重批
规则基线	品牌规范v5、合规规则v2	基线改变需重批或评估
撤回与接管	10分钟内可撤回，主管接管	不可撤回则不得自主

四、知识范围必须有来源等级、时效与冲突处理

知识库不是把更多文件放进检索系统。岗位需要明确来源等级：法律和监管、企业批准政策、产品主数据、合同与客户授权、经审核的知识条目、外部公开信息。不同来源冲突时，应按预设优先级处理并升级，而不是由模型临场选择“更像真的”内容。对价格、权益、医学、金融、法律和重大品牌事实，还应设置有效期与强制复核。

知识写入权与读取权应分离。多数AI员工可以读取经批准知识，但不能自行把运行输出写成“组织真相”。新经验进入正式知识库前，应经过来源核验、适用范围、有效期、责任人和敏感级别审查。长期记忆适合保存任务上下文与偏好，不应替代主数据和制度库。

表42 知识来源等级与冲突规则

来源等级	示例	冲突处理	有效期控制
A 法律/监管/合同	法规、监管命令、已签合同	最高优先并升级法务	按生效与失效日期
B 企业批准主数据	价格、产品、客户权益、品牌政策	覆盖一般知识	Owner定期复核
C 经审核知识	案例、SOP、FAQ、复盘结论	与A/B冲突时停用	设置复审日期
D 外部公开信息	官网、新闻、行业材料	需标明来源与不确定性	按任务时效
E 运行记忆	会话、任务偏好、临时状态	不得覆盖A–D	任务结束清理或限期

五、成本失控通常来自“隐性循环”而非单次调用

数字执行单元的成本异常常来自重试、循环委派、重复检索、上下文膨胀、并发复制和失败后人工补救，而不是某一次模型调用价格。预算控制应同时设置单任务、单实例、单岗位和组合四层上限，并监控“合格结果成本”而非单纯调用费。当任务成功率下降时，表面低价模型可能因重试和返工变得更贵。

FinOps机制需要与授权联动：超过预算阈值可以自动降级模型、减少并发、停止非关键任务或转人工；但涉及质量和客户权益时，不能为了省成本静默使用不满足要求的能力。成本策略本身也是受控变更，应记录对质量、周期和风险的影响。

表43 四层成本控制

成本层级	控制指标	自动动作	人工决策
单任务	调用、时间、重试、工具次数	达到上限停止并升级	是否继续或转人工
单实例	日/月预算、并发、存储	限流或降级非关键任务	调整岗位配置
单岗位	合格结果成本、监督成本	触发绩效红灯	重新立项或停损
组合	供应商集中、总预算、风险准备	冻结新增扩编	董事会/经营层再配置

六、审计证据要能支持“重演”，而不只是“查到记录”

可审计性不等于存在日志文件。有效证据应让授权人员在合理范围内重演一次行动：当时的岗位目标是什么、哪个实例运行、使用什么版本、读取哪些来源、调用什么工具、谁批准、外部系统返回什么、最终结果和客户影响如何。若只能看到一段聊天文本，无法确认实际写入了哪个系统、使用了什么凭证或是否发生重试，审计仍然是不完整的。

重演不要求永久保存所有敏感内容。企业可以通过哈希、引用ID、版本快照、脱敏事件、访问控制和分层留存，在隐私与可追溯之间平衡。证据留存期应由法规、合同、业务争议周期和风险等级共同决定；到期删除也要留下删除验证和批准记录。

表44 可重演审计证据

重演要素	最低证据	隐私保护方式
目标与任务	岗位ID、任务ID、目标版本	不保存无关业务信息
身份与环境	实例ID、环境、凭证引用	凭证只存引用不存明文
输入与知识	来源ID、版本、时间	敏感正文脱敏或受限存储
工具调用	动作、参数摘要、返回状态	敏感参数分层访问
批准	批准人、对象、版本、时间	签名或不可抵赖记录
结果	输出、外部写入ID、业务状态	按用途和留存期限保存
后果	客户、成本、风险与补救	聚合与最小化

七、三类常见治理失败及其纠偏

第一类失败是“工具先行”：企业先购买平台，再由业务部门寻找可以套用的任务。结果通常是场景价值弱、数据准备不足、主管缺位，最后用内容产量或使用人数证明项目存在。纠偏方式是重新回到岗位立项：以客户任务和经营基线为起点，明确替代方案、失败预算和退出条件。已经采购的平台可以作为候选能力，但不能反向定义业务岗位。

第二类失败是“集中治理替代业务责任”：AI委员会制定大量原则，却没有让具体业务主管承担结果。系统上线后，质量问题推给平台，合规问题推给法务，客户投诉推给一线人员。纠偏方式是把每个岗位绑定唯一业务owner，并在RACI中明确平台和风险部门是能力与监督责任，不是经营结果的替代者。重大事故可以联合处置，但责任不能在委员会中稀释。

第三类失败是“人类在环的形式化”：系统每一步都要求点击批准，审批者却没有时间、信息或能力判断，最终形成机械放行。纠偏不是简单增加审批，而是减少需要审批的动作、提高行动包信息密度、按风险分层抽检，并给批准者明确拒绝和升级权。对于无法有效判断的高影响动作，应回到草拟或纯人工，而不是用形式审批制造安全感。

表45 常见治理失败与纠偏

失败模式	表面现象	真实原因	纠偏抓手
工具先行	账号多、使用率高、经营结果弱	没有岗位假设和基线	重新立项、设置退出门
责任集中化	委员会很多、事故无人负责	业务owner缺位	单一岗位owner、三线分工
形式审批	每步有人点确认仍出事故	批准者信息不足或疲劳	行动包、风险分层、抽检
权限永久化	试点权限长期保留	无到期与复核	时限、额度、自动吊销
证据碎片化	聊天、日志、审批分散	缺少统一任务ID和四账	统一关联键、重演测试
复制冲动	试点成功后快速铺开	未验证组合风险与成本	分批扩编、容量与故障演练

八、营销场景的授权示例与边界

营销任务覆盖研究、内容、投放、客户沟通和公开承诺，风险差异极大。企业不应给“营销AI员工”一个统一授权级别，而应按任务拆分。热点监测可以在L4自动运行，因为主要输出是内部信号；品牌文章通常适合L2或L3；广告预算和受众修改应受额度、双钥匙与平台规则约束；危机声明、价格承诺和客户权益调整原则上停留在L1/L2。

表46 营销任务授权示例

营销任务	建议授权	允许动作	关键升级/降级条件
竞品与舆情监测	L4限额自主	抓取批准来源、内部预警	来源异常或涉及重大危机即升级人工
选题与内容策划	L1–L2	建议、草拟、生成排期草案	事实风险高或品牌敏感时降至L1
常规社媒发布	L3	行动包批准后发布	版本、账户、时间或素材变化需重批
低风险互动分类	L4	分类、标记、内部路由	禁止自动承诺、争辩或处理投诉
广告优化建议	L1–L2	分析和调整草案	不得凭平台建议直接改高额预算
限额广告实验	L3–L4	在预算、受众和时间窗内执行	成本、投诉或异常达到阈值即停止
客户跟进草稿	L2	生成个性化草稿与CRM待办	敏感客户、价格、法律内容需人工
危机声明与重大承诺	L1–L2	事实汇总和草拟	最终判断与发布由高管和法务负责

这些授权建议不是统一行业标准。企业应根据品牌风险、客户类型、渠道规则、监管要求和自身恢复能力调整。同一任务在测试环境、内部渠道和公开渠道的授权也应不同。治理的目标不是把所有动作固定在最低级，而是让权限与证据随风险变化，并能在表现下降、版本变化或事故后迅速收回。

九、最低可行治理基线

资源有限的企业不必一次建设复杂治理平台，但生产环境至少不能缺少八项基线：一个明确业务主管、一个唯一机器身份、一份岗位说明书、一张权限矩阵、一套真实任务评测、一条紧急停止路径、一组可关联日志和一份停用清单。缺少其中任何一项，都应把系统限制在建议或草拟层级。随着任务影响、客户范围和自动执行程度提高，再逐步增加双钥匙、持续监控、独立审计、供应商退出演练和组合风险管理。

最低基线的价值在于形成可验证的闭环，而非追求文件数量。岗位说明书必须能对应到实际身份；权限矩阵必须能被技术策略执行；批准必须绑定具体行动；日志必须支持重演；停用清单必须经过演练。制度与系统不一致时，以更严格的一方为准，并立即修复差异。

BOARD BRIEF 22

附录E 纯文本来源索引

竹势 AI 营销智库出品

以下索引仅保留机构、作者、标题、日期与口径信息，不设置第三方可点击链接。

[S1] NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100–1, 2023.

[S2] NIST. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600–1, 2024.

[S3] NIST. Zero Trust Architecture. NIST SP 800–207, 2020.

[S4] NIST. Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations. NIST AI 100–2e2025, 2025.

[S5] ISO. ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system. Public abstract and explanatory materials.

[S6] ISO. ISO/IEC 23894:2023, Information technology — Artificial intelligence — Guidance on risk management. Public abstract.

[S7] OWASP GenAI Security Project. Agentic AI — Threats and Mitigations, 2025.

[S8] OWASP GenAI Security Project. Multi–Agentic System Threat Modeling Guide v1.0, 2025.

[S9] OWASP. AI Agent Security Cheat Sheet, accessed 2026–07.

[S10] European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence, 2024.

[S11] 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.

[S12] 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, 2025–03–14, 2025–09–01施行.

[S13] 国家互联网信息办公室. 网信部门依法查处“剪映”App等生成合成内容标识违法问题, 2026–04–28.

[S14] 中华人民共和国个人信息保护法, 2021.

[S15] 网络数据安全管理条例, 2024.

[S16] Morgan Stanley. Key Milestone in Innovation Journey with OpenAI, 2023–03–14.

[S17] Morgan Stanley. Launch of AI @ Morgan Stanley Debrief, 2024–06–26.

[S18] Morgan Stanley. AskResearchGPT, 2024-10-23.

[S19] Klarna. AI assistant handles two-thirds of customer service chats in its first month, 2024-02-27.

[S20] Civil Resolution Tribunal of British Columbia. Moffatt v. Air Canada, 2024 BCCRT 149.

[S21] DPD. Public statement regarding chatbot incident, January 2024; public reporting and company response.

[S22] McDonald's and IBM. Joint Statement on automated order taking partnership, 2021; McDonald's trial conclusion communications, 2024.

[S23] JD.com Corporate Blog. JD Cloud Aims to Accelerate Industrial AI, 2022.

[S24] JD.com Corporate Blog. JD.com 618 2026 AI application results, 2026-06.

[S25] Peter F. Drucker. The Practice of Management. Harper & Brothers, 1954.

[S26] Herbert A. Simon. Administrative Behavior. Macmillan, 1947.

[S27] W. Edwards Deming. Out of the Crisis. MIT Press, 1982.

[S28] Robert S. Kaplan and David P. Norton. The Balanced Scorecard—Measures that Drive Performance. Harvard Business Review, 1992.

[S29] Chris Argyris and Donald A. Schön. Organizational Learning. Addison-Wesley, 1978.

[S30] James Reason. Human Error. Cambridge University Press, 1990.

[S31] Ronald H. Coase. The Nature of the Firm. Economica, 1937; Oliver E. Williamson, transaction cost economics works.

[S32] Melvin E. Conway. How Do Committees Invent? Datamation, 1968.

[S33] 竹势 AI 营销智库. AI营销产品与服务介绍素材, 2026.

[S34] 竹势 AI 营销智库. 公司介绍与品牌定位, 2026.

[S35] 中国平安保险（集团）股份有限公司. 2024年可持续发展报告, 2025-03；2025年年度业绩及年度报告, 2026-03-26；披露AI服务代表服务量、金融知识规模、AI辅助销售与“AI+人工”复效任务分配。

[S36] 百度集团. 2025年ESG报告, 2026-05；Baidu Announces Third Quarter 2025 Results, 2025-11-18；披露Digital Employee、AI Applications与AI-native Marketing Services的业务口径及收入。

BOARD BRIEF 23

结语：把“会做事”纳入“谁负责”的系统

竹势 AI 营销智库出品

AI员工的长期价值不在于模拟人的身份，而在于把数字执行能力组织成受控、可复用、可改进的经营资产。企业越早建立岗位、身份、权限和证据四本账，越能在不牺牲责任与信任的前提下扩大机器执行。反之，若以拟人化语言掩盖责任，以模型能力代替授权，以活动量代替绩效，以供应商承诺代替企业证据，规模越大，风险越难收回。

董事会真正需要批准的，不是“再招多少AI员工”，而是哪些经营任务值得被数字化执行、谁对结果负责、系统可以做到哪一步、失败如何被发现和收回、证据如何进入审计。只有当这些问题有明确答案，AI员工才从营销概念成为企业能力。

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。