

企业级客服 Agent 选型评估表

2026-10-29 | 评分表 1/2 | 适用：IT 决策者、运营负责人联合评审

企业/项目：_____ 场景：_____ 评审人及日期：_____

先用同一批真实任务验证各方案，再填写 0-4 分。0=未提供或验证失败；1=仅口头承诺/演示；2=有限场景试通；3=本企业场景验收通过且证据完整；4=真实试点持续达标，异常与恢复已验证。

编号	评估项	权重	否决项	必须展示或留存的证据	方案 A 0-4分	方案 B 0-4分	方案 C 0-4分
01	业务与对话 / 真实任务与上下文	10%	否	同一测试集覆盖改口、多意图和连续追问；记录任务是否完成，核对不能重复收集已知信息。			
02	业务与对话 / 渠道与语言适配	8%	否	在实际接入渠道与目标语言测试；如需跨渠道关联，验证身份映射与授权，不能默认已打通。			
03	知识与质量 / 知识检索与依据	12%	否	提供答案依据、知识版本和更新流程；测试过期、冲突资料与权限范围内的检索。			
04	知识与质量 / 澄清、拒答与纠错	10%	否	测试资料缺失、模糊问题和不当指令；确认会澄清或转人工，不生成未经确认的业务承诺。			
05	系统与执行 / CRM与工单衔接	8%	否	在测试环境演示字段映射、事件触发、失败反馈和状态回写；区分API可用与业务流程已接通。			
06	系统与执行 / 动作授权与执行可靠性	12%	是	验证最小权限、必要确认、幂等与重复请求处理；失败不能误报成功，保留操作结果及责任记录。			

建议评分方法：总分 = $\sum$ （权重 × 得分 ÷ 4），满分 100。权重为编辑建议，可调整但须合计 100；未完成评分前不要将空白视为 0 或输出确定排名。

企业级客服 Agent 选型评估表

2026-10-29 | 评分表 2/2 | 适用：IT 决策者、运营负责人联合评审

四项建议否决项：06 动作授权、07 数据权限与审计、08 异常恢复、09 人工接管。建议均须达到 3 分并留存证据；阈值由企业按场景确认。总分不能抵消硬项未过。

编号	评估项	权重	否决项	必须展示或留存的证据	方案 A 0-4分	方案 B 0-4分	方案 C 0-4分
07	治理与安全 / 数据隔离、权限与审计	10%	是	提供数据流向、访问和训练用途说明；演示越权拦截、敏感信息处理、日志及保留删除策略。			
08	治理与安全 / 异常停止、降级与恢复	8%	是	实测接口超时、模型不可用与知识异常；验证可停止自动动作并降级，业务回退与恢复需预先定义。			
09	运营与协作 / 人工接管与上下文交接	8%	是	测试客户主动找人工、AI无法完成及人工离线；确认入口可用、记录可读，并交代后续安排。			
10	运营与协作 / 运营指标与持续复测	6%	否	约定解决、重开和转人工口径；保留失败样本，测试知识或模型变更后能否复测。			
11	成本与退出 / 全成本与计费口径	4%	否	拆分许可、调用、集成、维护、质检与人工兜底；明确失败、重复请求和所谓解决的计费规则。			
12	成本与退出 / 数据可迁移与退出安排	4%	否	核对知识、日志、客户数据可导出范围、格式及退出费用；明确交接与供应商变更支持。			

评审结论	方案 A	方案 B	方案 C
加权总分（满分100）	_____	_____	_____
四项否决项（通过/待验证/未通过）	_____	_____	_____
待补证据与责任人	_____	_____	_____

企业级客服 Agent 选型评估表

2026-10-29 | 试点验收记录 | 适用：IT 决策者、运营负责人联合评审

每个任务填写一条记录。复制本页可用于更多样本；Excel 版提供三方案评分及可继续填写的记录区。

样本编号	<div></div>
业务场景与意图	<div></div>
授权范围	<div></div>
预期结果	<div></div>
实际结果与系统回写	<div></div>
人工介入原因	<div></div>
约定窗口内重开	<div></div>
耗时与全成本	<div></div>
证据链接	<div></div>
验收人及日期	<div></div>

指标先约定：自主完成必须有明确的任务结果或业务回写，并在约定观察窗口内检查重开；把全部请求、适用范围内请求、转人工及失败记录分开。单次成功案例不能替代真实流量试点。