

深度调研 × 架构设计 · 立项汇报 · 2026 年 9 月

AI 原生客户引擎

汽车客户全生命周期的数据、智能与 Agent 体系
——从 " 管理客户 " 到 " 赋能客户 " 的范式重构

调研规模 300+ 次独立检索 · 6 个广域方向 · 11 个深潜维度 · 467 条来源

覆盖范围 行业现状与开源生态 · 汽车标杆案例 · 七层架构设计 · 合规与路线图

口径声明 文中效果数字凡标注 " 厂商自报 / 供应商口径 " 者，均作上限表述

本演示为《AI 原生客户引擎：深度调研与架构设计报告》（5.4 万字 / 467 条参考文献）的汇报浓缩版

本汇报回答五个问题，对应报告十章证据

01	为什么传统 CDP 过时了？ 范式转移的证据链与 "AI 原生 " 的判定标准 报告第 1-2 章	P4-P9
02	汽车行业标杆已经做到了什么？ 主动服务 / 成交促进 / 用户运营的机制与口径甄别 报告第 3 章	P10-P12
03	客户引擎的架构是什么？ 七层参考架构、核心 / 外围数据、非结构化管线与 OneID 报告第 4-5 章	P13-P14
04	理解与 Agent 能力怎么建？ 客户记忆选型纪律、四类 Agent 角色、三大场景引擎 报告第 6-8 章	P15-P17
05	怎么落地、花多少、风险在哪？ 合规架构、开源选型、12 个月路线图与风险登记册 报告第 9-10 章	P18-P20

五个核心判断：不升级 CDP，新建 AI 原生客户引擎

- 1 CDP 品类正在消亡，您的试错不是个案而是行业基线。** mParticle 估值较峰值缩水 62.5% 卖身，独立 CDP 厂商六周内三连并购；Gartner 调查显示仅 22% 营销者高频使用已购 CDP。
- 2 方案是七层 AI 原生客户引擎：数据主权在自有湖仓，Agent 是数据的第一读者，合规是第一设计输入。** 从 L0 连接供给到 L6 评估反馈，合规、可观测、人机协同三横切贯穿。
- 3 一期主场景选主动服务，而非促进成交。** 它是唯一被端到端经济账验证的赛道：问界 / 特斯拉标杆机制公开、工程延迟基准现成、合规授权明确、可做对照实验；MVP 为维修进度主动推送。
- 4 客户记忆组件不预选型、先实测。** 公开基准中厂商自报与独立复测最高相差达 45 个百分点，且 " 删除合规 " 恰是开源框架实测最弱的一轴——须用自有语料建 30 题以上 golden set 再定选型。
- 5 身份主权自持，企微只是通道。** 以 UnionID、手机号等强标识主键落自有湖仓，SCRM 厂商可换、身份资产不丢；企微会话存档官方价 2026 年 1 月已翻倍，通道与数据资产必须解耦。

三条独立证据线汇聚： CDP 成了 " 别人家 AI 的燃料 "

① 资本线：三连并购

ActionIQ→Uniphore、Lytics 资本断供、mParticle 卖身，六周内相继发生；Census→Fivetran。

买方均非 Martech 公司—— CDP 的数据资产被 AI 公司收编。

-62.5%

mParticle 卖身价较 2021 年峰值估值的缩水幅度（\$800M→\$300M）

独立 CDP 厂商 2025 年退出事件（公开报道汇总）

厂商	结局	买方 / 方向	信号
mParticle	\$300M 卖身	Rokt（广告科技）	估值较峰值 -62.5%，仅略高于累计融资
ActionIQ / Lytics	被收购	Uniphore 等（非 Martech）	Lytics 2019 年后零融资，无力完成 AI 转型
Census	并入 Fivetran	数据管道公司	Composable 阵营同样被收编

② 使用线：建了不用

仅 22% 营销者对已购 CDP 高使用率（Gartner 2025 MQ 所引调查）；已部署 CDP 平均只用上 47% 的能力（Gartner 2023 营销技术调查，2022 年为 55%），逐年恶化。

22%

对已购 CDP 报告 " 高使用率 " 的营销者占比（Gartner，2025）

③ 内部人归因

收购方 Rokt CEO 点破：CDP 不直接参与投放、归因或竞价决策，离 " 价值创造层 " 太远，持续投入难自证。

与我们 " 数据不丰富、不知道怎么用 " 的试错互相印证。

47%

已部署 CDP 的平均能力利用率，2022 年为 55%（Gartner 2023 调查）

AI 原生的判定：Agent 是数据的第一读者，而非外挂的问答框

四条判定标准（全部满足才算 AI 原生）

① Agent 是数据第一读者

数据按 agent 的消费方式（语义层 + 工具 + 上下文 API）组织，而非按报表

② 决策与执行闭环在线

从洞察到触达动作在系统内完成，候选动作必须包含 " 不动作 "

③ 评估反馈内建

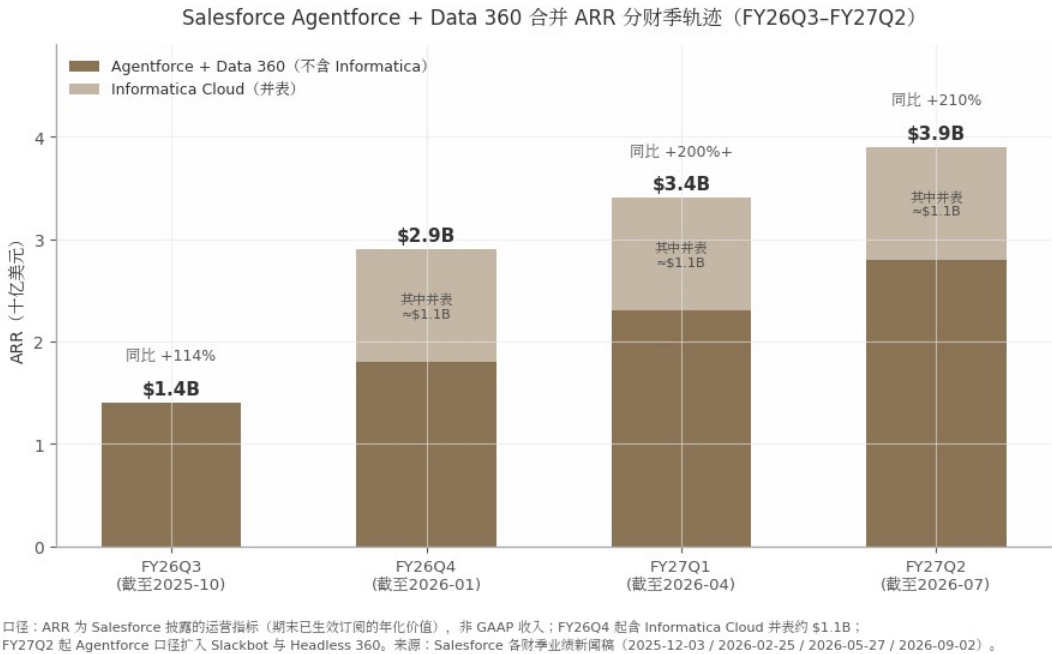
随机 holdout 对照与 uplift 度量是平台组件，归因 ≠ 增量

④ 统一画像是 " 引用 " 而非黄金记录

按用途选值 + 属性级 survivorship，取代单一真理源

维度	传统 CDP / CRM	AI 原生客户引擎
数据范围	结构化交易 / 行为事件为主	录音、聊天、舆情等非结构化全量入湖
非结构化	不支持或外挂附件	语音 / 会话 / 情绪理解为原生管线
时效性	T+1 批处理	事件驱动，供给按延迟分级至 sub-100ms
消费方	人（运营 / 分析师看报表）	Agent 优先，人通过 copilot 协同
价值闭环	止于分群与导出名单	决策 - 执行 - 评估在线闭环
度量方式	归因报表	holdout 对照 + uplift 增量度量

厂商生态剧变：数据底座 +Agent 合并收费，按结果计费成硬指标



读图： Salesforce 把数据云与 Agent 打包为同一收入线（ARR 口径含 Informatica 并表），"数据底座 +Agent" 一体化已是套件厂商的标准叙事。

定价革命：按结果计费（outcome-based）

厂商	单价口径	备注
Intercom Fin	\$0.99 / 解决	含 "假定解决" 争议
Zendesk	\$1.2-1.5 / 验证解决	2026-05 三层复核
HubSpot	\$0.50 / 解决对话	Breeze 客服 Agent

客服 AI 采购单价已透明到 \$0.5-2.5/ 解决（人工 \$5-12）。**自建的赢面不在单价，而在数据主权与跨域编排。** Gartner 预测 2028 年过半企业只为 "工作流结果" 付费。

国内格局（甄别软文后）

- 销售易 NeoAgent 2.0（腾讯控股，接混元）与纷享销客 Agentic CRM 为主角；
- 神策有明确车企 CDP 方案（可对接评估，但量化 ROI 公开数据缺失）；
- 企微 SCRM 洗牌后头部化，与 Agent 融合度仍低；
- 本土 CRM 份额 2026 年预计过半（54.3%，转引口径）。

模型能力的三个事实，直接决定我们管线的技术选型

能力域	2025-2026 事实（带口径）	对本项目的含义
中文 ASR	开源第一梯队普通话 CER 2.9-4.6%（统一基准）；电话信道恶化至 9.8-14.7%（多源实测）；方言 2026 年大幅收敛	识别已商品化，差异化转向行业热词 + 角色分离 + 语义；电话录音是共同洼地，采集端双轨录音是最优解
全模态语音理解	上一代音频 LLM 普遍 30 秒硬限制（官方文档）；新一代宣称 10 小时 + 输入但无第三方保真实测；情绪识别学术上限 WA≈72%，厂商 88-97% 为不同任务口径	采用 "ASR→LLM 级联主路 + 全模态旁路" 混合架构；情绪只作辅助信号，不作决策依据
长上下文 vs RAG	头部模型上下文已达百万至两百万 token（2025-2026）；长上下文与 RAG 成本差最高可达 1250 倍（1M 窗口、无缓存条件）	客户记忆与画像不走 "全文塞上下文" 路线，走结构化供给 +RAG 的经济路线

对本项目的第一行动项

说话人分离在单声道录音上的最好成绩仍为 DER 11.2-13.3%（2025 跨语言基准）——意味着每 100 句客我对话仍有 11-13 句归属错误，意向归属可能被颠倒。**从采集源头拿双声道双轨录音接近零成本，后期算法分离昂贵且未解决。**因此：核查 Vios 平台录音输出是否双轨，是整个数据体系的第 1 号行动项。

开源盘点：可直接复用的很多，但有三个 " 许可 / 能力陷阱 "

类别	代表项目（ star 量级， 2026-09 检索时点）	许可	对本项目的借鉴价值
开源 CDP	RudderStack ~4.5k 、 Jitsu ~5k 、 LEO CDP	ELv2 / MIT	LEO CDP 是唯一 AI-first 同构蓝图（个人项目，只作参照系不作依赖）
Agent 编排	LangGraph 40k 、 Dify 151k 、 Coze Studio ~20k 、 FastGPT 25k	MIT / 附加条款 / Apache2.0	LangGraph 为生产默认（ ~400 家企业）； Coze Studio 是纯 Apache2.0 的许可最优解
Agent 记忆	Mem0 60k+ 、 Graphiti 20k 、 Letta 13k	Apache2.0 等	基准罗生门（见 P15 ），不预选型、先建 golden set 实测
RAG / 知识库	RAGFlow ~90k 、 LangChain 143k 、 Milvus 45k 、 Qdrant	Apache2.0 等	GraphRAG 索引 \$4-7/ 文档成本悬崖， LightRAG 作轻量替代
实体解析	Splink 2.2k 、 Senzing 、 Zingg	Apache2.0 / 商业	中文实体解析无现成开源实践，须自建比较器 + 强标识确定性匹配打底
语音栈	FunASR 16k 、 WhisperX 22.6k 、 emotion2vec	MIT 等	中文呼叫中心会话分析端到端 PoC 的首选开源栈

三个陷阱

① 许可成色： Dify 附加条款被 OSI 质疑非开源（集团 - 经销商多租户场景须法务认定）； RudderStack 已改 ELv2 。 ② star≠ 生产采用：框架选择可改变 30 个百分点表现，须按可靠性工程指标选。 ③ 厂商自测基准：记忆框架唯一独立复测与自报差 45 分，一切选型须自有数据复测。

失败率双锚的准确口径，与成功项目的五个共同条件

>40%

Gartner 预测（2025-06-25）：到 2027 年底超 40% 的 Agentic AI 项目被取消——分析师预测，无公开方法论。

95%

MIT NANDA（2025）：95% 企业 AI 项目 "无可测量 P&L 影响"——非同行评审、缺基线≠失败，但结构性警示成立。

67% vs 1/3

外购 vs 自建成功率（MIT 同研究口径）：外购约 67%、自建约三分之一——自建赢面只在数据主权与跨域编排。

成功项目的五个共同条件（多源一致）

- 1 窄场景起步，单点打穿再扩展
- 2 深度嵌入现有工作流，而非并列新系统
- 3 有明确业务 Owner 背指标
- 4 上线前定基线，对照组从第一天内建
- 5 workflow 重设计——McKinsey 口径：仅 21% 企业做到，但它是 EBIT 影响最强因子

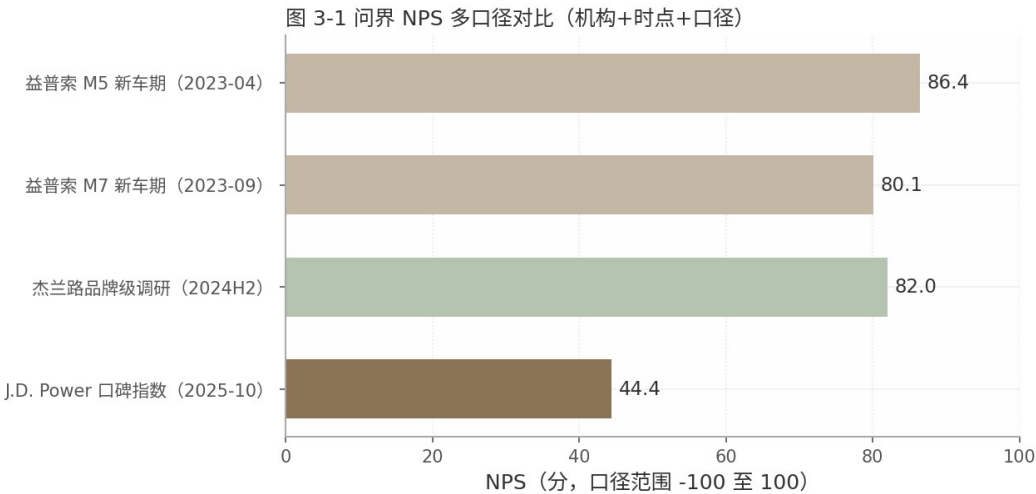
标本案例：Klarna 反转（须带时点引用）

- 2024-02 官方新闻稿：AI 助手首月 230 万会话、相当于 700 名全职客服、利润改善 \$40M
- 2025-05 CEO 承认过度裁员致质量下滑，回调人机混合
- 2025-11 财报口径 853 FTE / \$60M——成功叙事须带时点与反转

对我们的含义：Agent 体系按 "有界自主 + 审批门" 设计，45% 团队跑混合模式是行业常态；瓶颈在数据层而非 agent 层。

主动服务是已被标杆跑通的赛道：机制公开、经济账成立

标杆	机制（可复用点）	效果数字（口径）	证据等级
问界 CARE	106 项自研预警算法，远程诊断 / 故障预警 / 主动服务三平台，7×24 主动预见	累计主动服务 20 万次、省 37 万小时（厂商自报，2025-10）	机制 High / 数字厂商口径
特斯拉	远程诊断四层漏斗：90% 远程诊断、80% 不进店	2024 年 250 万人次远程诊断、均 25 分钟（厂商自报）	机制 High / 数字厂商口径
蔚来 / 福田等	服务卫士秒级诊断；福田 " 检测→督促返厂 + 工单推送 + 配件预准备 " 闭环	新能源重卡电池故障率 -67%（厂商口径）	机制 High / 数字厂商口径



由此锁定一期 MVP：维修进度主动推送

- 六节点（接车 / 诊断 / 配件到货 / 维修 / 质检 / 交车）主动推送，单一案例满意度 +25 个百分点；
- 需求侧第三方佐证：69.5% 用户想看到维保进度而 4S 不提供（凯睿赛驰调研）；
- 交付 / 零件延迟安抚已是行业标配，但补偿须感知客户外部损失情境（小鹏固定积分 SOP 撞国补窗口失效的教训）；
- 打扰风险低：服务类触达不受营销频控同等级约束，授权路径清晰。

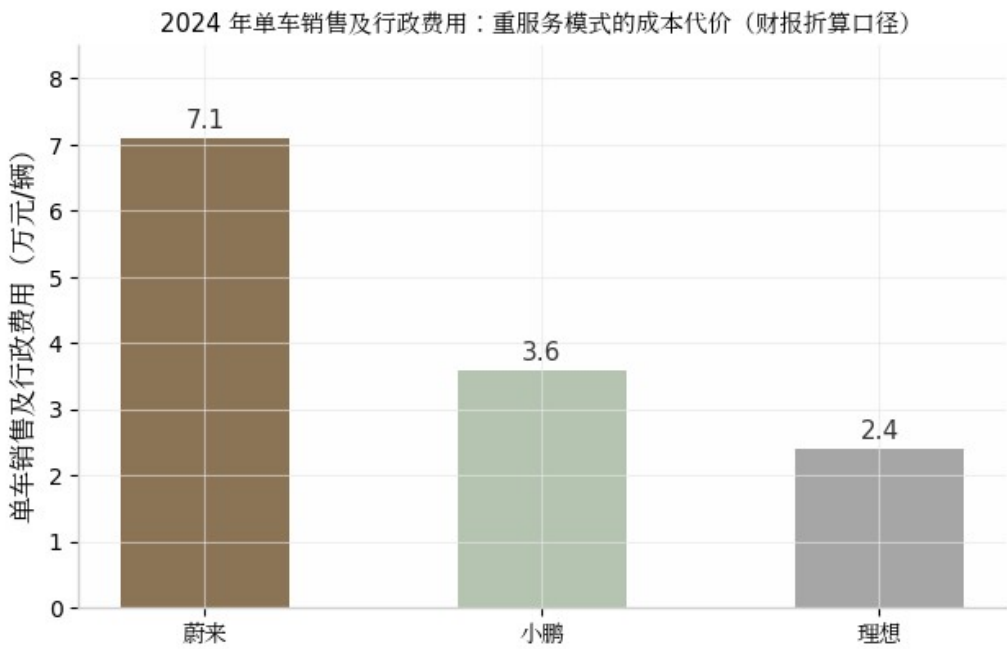
成交侧标杆打法已收敛，但一切效果数字须用买方基准校准

标杆打法（供应商口径，作上限看待）

- 红旗 / 瓴羊 半年成交数据冷启动 → 两店 A/B 测试 → Top10% 线索 3.5 倍转化、成交 +11%（供应商自报）——方法论细节完整，是一期范本
- 长安福特 中央 DCC 双呼模式，进店率 +10%（供应商口径）
- 吉利银河 IM 私聊→意向 62%（窄口径）
- 北汽福田 600 万保客 + 车联网数据驱动增换购，线索成本仅 80 元（供应商口径）
- 奕派 生命周期服务地图，置换转化 +18%（厂商口径）

买方基准（甄别一切厂商数字的锚）

- 垂媒线索有效性 25-28%、线索成交率 32-35%、单台成交成本约 2200 元（经销商访谈 + 财新，Medium-High）
- 行业线索成交转化率公开口径互相打架（4-8% / 6.8% / 12.7%）——无口径定义的数字不引用、不决策
- AI 外呼宣传接通率 70-85%，实测均值 56.4%—— 供应商数字普遍约 30% 水分



蔚来涟漪模式的反面账：转介绍占销量 40-69%（口径混用、非审计）的同时，服务无忧单户年亏 4000 元（李斌自认）、单车 SG&A 数倍于理想。

含义：用户运营要做，但关怀强度须有单位经济约束——这正是 "频控 +holdout" 内建的理由。

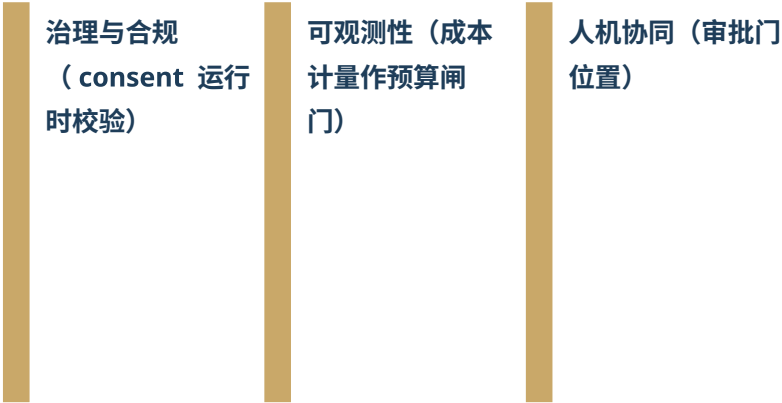
本项目最大的风险不是技术：客户数据归属是商务谈判问题

案例（时点）	事实	对我们的含义
奔驰澳洲（2023-08 判决）	38 家经销商诉奔驰代理制败诉，标的 6.5 亿澳元；法官认定经销商 " 事实上胜诉、法律上败诉 "	代理制 / 直营的本质动机是主机厂回收客户关系与全量数据；数据条款须先于系统建设谈判
大众（2025-12）	终止欧洲电动车代理直销，C 端回归经销商	代理制浪潮退却——数据归属无标准答案，须以合同逐条约定
smart D2C（现行）	佣金 = 销售佣金 + 交付佣金两段制，区域独家 " 少商多店 "，数据与定价归车企	最清晰的分账样本：渠道让渡数据的前提是获得经营对价

落地要求：给渠道的 " 数据共享对价 " 清单（先于系统建设）

- ① 渠道经营看板——门店级线索转化、试驾执行率、服务履约的可视化回流，让经销商从数据上交者变为数据受益者
- ② 分级名单机制——线索评级结果与门店共享，换取全链路跟进数据回传
- ③ 倒闭风险隔离——30+ 车企倒闭致数百万车主服务断供（威马 / 极越 / 哪吒 / 高合），架构须内建离线降级与异地容灾，售后数据托管预案
- ④ 行业背景——八大上市经销商集团 2025 年七家合计亏 60 亿，渠道议价能力窗口期短暂

七层参考架构：数据主权在自有湖仓，按 "Agent 第一读者 " 倒推



三条设计原则

- 数据主权在自有湖仓——模型与通道可换，资产不丢
- 按 Agent 第一读者倒推—— MCP 工具清单设计优先于编排选型
- " 赋能关系 " 落到对象模型——客户 / 车辆 / 顾问 / consent 为一等公民

可靠性硬约束

单步 95% × 20 步 ≈ 36% 端到端——拆短链路 + 确定性校验 + 节点级重试，而非换模型

边界

车联原始遥测不灌入客户链路（成本论证见报告 4.3）

非结构化管线：洞察上限在采集层就已锁死



数据源	关键工程事实（口径）	决策
Vios 录音	双轨录音接近零成本解决角色归属；单声道算法分离 DER 11.2-13.3% 未解决（2025 基准）	第 1 号行动项：核查 Vios 输出是否双轨
企微会话存档	官方 API（RSA+AES、增量拉取、延迟 5-15 分钟）；2026-01 涨价 100%（900/1800 元 / 席 / 年）；专区 " 可用不可见 "	一期开通（开始留存历史）、二期接入；个人号历史聊天无合规路径
会话抽取	schema-first+LLM+ 校验器 F1 0.866 vs 规则 0.258（学术）；汽车七字段：意向车型 / 预算 / 用车场景 / 决策角色 / 决策周期 / 竞品 / 核心异议	抽取即回填画像；字段级 F1 取代转写 CER 作业业务指标
论坛 / 舆情 / VOC	车质网 2025 年投诉 227,803 宗（+31.6%）；智能工牌已有 21 品牌 8500 门店产品化先例	四期再接入；工牌 / 展厅视频须显性告知 + 可拒绝、禁用人脸识别
OneID 双主键：手机号 + UnionID 确定性锚定打底，概率匹配扩展；微信资产全部绑同一开放平台主体，自建应用持有 UnionID 明文。开源实测示警：Splink 百万条 9.24 秒但召回偏弱、真实脏数据 precision 60-76.5%（MERAI 论文）——中文实体解析无现成开源方案，比较器须自研并自有数据复测。		

客户记忆是 " 伪成熟 " 组件：唯一可信路径是自建 golden set 实测

证据口径	数字	条件限定	可用性
厂商自报（ Mem0 官方， 2026-04 ）	LongMemEval 94.4	托管平台新算法、无公开 harness	仅作上限参考
独立复测（ KTH 等同口径横评）	Mem0 49.0–57.5 ； RAG≈Mem0 且 TCO 仅 1/8.4	特定基准 +judge+ 版本，互不兼容	方向性参考
删除 / 遗忘合规实测（ DeepLethe ）	开源 Graphiti 仅 4.4–7.0% ； 所有框架最高约 68%	测的是删除控制面，与召回能力正交	对 PIPL 删除权是硬预警

结论：没有任何已发表数字可直接拍板选型。用自有语料建 30 题以上 golden set （含删除、时点查询、矛盾更新三类题） + 开源 harness 实测；PIPL 被遗忘权要求自建 mutation-time 删除层（实测可升至 91.7–93.2%）。

画像与标签：车企独有飞轮

- 10 大类 100+ 标签起步；反面教训：2000+ 标签案例仅 15% 高频使用
- 销售顾问人工标注 → 算法训练样本是行业独特杠杆（企微汽车行业版实战）
- " 画像 = 记忆的阈值化沉淀 " ：同类偏好 5 场景重复才写入，防画像抖动
- 准确率验证四法：留出集 / 人工抽检 / 覆盖度 - 准确度 - 稳定性 /A/B 实验

群体智能：放大与反哺（对应您的要求）

- 个体→群体三路径：显式相似 / PU learning / 图嵌入
- 群体→个体反哺：群体差异对比、联邦学习冷启动、模式检测
- 预测建模：LLM embedding+ 轻量分类器（流失 89% ）优于零样本打分
- 警示：整车超低频购买不适用 BG/NBD 传统模型；lookalike 须防敏感属性推断

四类 Agent 角色 + 有界自主：可靠性是工程问题，不是模型问题

Agent 角色	职责	读写权限	HITL 级别	评估指标
数据丰富 Agent	从对话 / 工单回填画像与记忆（参照微软 D365 Data Enrichment Agent：建议 / 自动两档 + 回滚 + 审计）	读全部、写画像建议	低置信人工审	字段级 F1 / 采纳率
场景触达 Agent	主动服务 / 关怀 / 成交的触达执行	读画像、写触达记录	审批门分级	增量转化 / 退订率
顾问 Copilot	企微侧边栏：一句话操作 CRM、语音录单、金牌话术	读写 CRM	人在环中	跟进及时率
质检 Agent	对其他 Agent 的输出做质检（D365 QA Agent 已 GA）	只读 + 评分	抽检复核	质检覆盖率

可靠性工程：单步预算反推法

- 悖论：单步 95% 可靠 → 20 步端到端仅约 36%
- 反推：20 步链路要 90% 端到端，单步须 ≥ 99.5%
- 应对：拆短链路 + 确定性校验 + 节点级重试 + pass^k 看板——**而不是换更强的模型**
- 编排默认 LangGraph（1.0 GA 2025-10，节点级断点续跑是分水岭）
- 写路径五层控制：零信任身份 / OAuth 2.1 / 读写分离 scope / 工具级审计 / 写操作人工批准

主动式 Agent：干预时机是硬科学

- Pare-Bench（143 任务）：前沿模型干预时机得分仅 42%，**"提议率越低越好"**
- 清华 Proactive Agent：把 "用户可拒绝" 作为一等信号
- 架构：Observe-Execute 分离 + 事件驱动触发 + 审批门
- 触达合规硬约束：同意 / 录音留存 / 退订通道；21:00-8:00 禁呼
- "AI SDR 全自动叙事已死"：45% 团队跑混合模式，瓶颈在数据层

同一座数据湖，三组取数视图：场景 × 数据需求总表（节选）

场景	触发点示例	所需数据（从湖中取什么）	批次
主动服务	维修进度六节点推送；交付 / 零件延迟安抚	工单节点时间戳、订单状态、配件物流、历史会话情绪 ⚠	P0 （一期）
主动服务	车联网故障预警→主动服务工单	故障码 / 里程 / 电量信号（脱敏聚合）、服务历史	P1 （三期）
主动关怀	保养 / 续保 90-60-30 三级 SOP	保养周期、保单到期日、历史响应偏好 ⚠	P1 （三期先行）
主动关怀	生日 / 年度用车报告等个性化关怀	画像偏好、用车年报聚合数据	P2
促进成交	意向评分→试驾邀约；沉默线索分层唤醒	会话七字段抽取、行为事件、历史跟进记录	P2 （三期随后）
共用底座	全局频控中心、consent 状态校验	consent_ref、频控计数、OneID 主档	一期准入件

落地顺序即合规成本梯度：⚠ 字段对应三条合规链路（外部损失情境推断 / 会话存档 / 位置足迹），可直接作 PIA 底稿；完整版为 21 行字段级清单（报告 8.4 节表 8-1）。**频控不预设数字：**分通道分人群实测确定，唯一硬约束是渠道方限制。

决策仲裁原则：候选动作必须包含 " 不动作 " ；自动化决策按 PIPL 第 24 条内置说明机制 + 非个性化选项 + 人工通道；差评转化概率、情绪三维模型（频率 × 强度 × 关联性）驱动黄 / 红 / 总监三级 SOP。

合规不是事后审查，而是与湖仓分层同级的架构输入

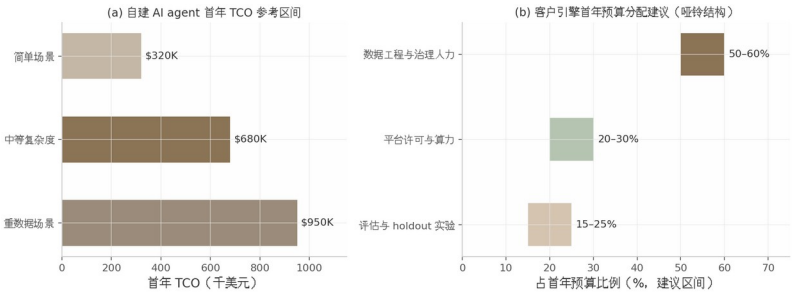
法规要件（条文 / 判例）	关键要求	架构原语（落点）
PIPL 敏感个人信息单独同意（§28-30；北互 2023 判例：捆绑无效）	逐项单独同意，一揽子同意无效	consent 中心：字段级 consent_ref + retention_till，运行时校验（L1 横切）
PIPL 自动化决策（§24）	说明机制 + 非个性化选项 + 人工通道	NBA/ 线索评分的决策解释接口与人工复核通道（L4）
删除权 / 被遗忘权（PIPL §47）	删除请求须可执行、可验证	crosswalk 可撤销 + mutation-time 删除层（L1/L3）
汽车数据规定 + 2026 出境指引（27 类 51 项）	座舱数据默认不收集、车内处理、远程访问即出境	车联遥测脱敏聚合后入湖；出境字段级目录管控
采集红线判例（阿维塔 2025-11 / 宝马人脸案 4617 条 / 群控案 2020 判赔 260 万）	录音显性告知 + 可拒绝；展厅禁用人脸识别；个人微信抓取违法	采集端告知机制为 L0 准入件；企微官方存档为唯一合规通道

大模型接入三件套：脱敏网关（云端只处理逻辑） + 私有化 / 专区部署 + 合同禁止供应商训练；供应商评估用信通院生成式 AI 个保 8 部分标准。PIA 五类触发情形本项目几乎全中，报告留存 ≥ 3 年。

最新自律动态（2026-09-15）：《智能网联汽车数据安全合规与发展倡议》发布——恰逢本项目启动时点，合规姿态本身即是渠道与用户信任资产。未决灰区（画像拒绝权缺位、员工同意有效性口径分化、远程访问出境执行尺度）已在报告第 9 章如实标注。

选型建议：开源优先但守住许可底线，预算的大头是数据工程人力

层	主推	替代 / 备注
流与湖仓	Kafka+Flink → Iceberg	Serverless Kafka 起步（千元级 / 月）
OLAP 供给	Doris / StarRocks	实时客户 360 有货拉拉 / 淘宝实证
实体解析	Splink + 自研中文比较器	100 万条 9.24 秒，可扩亿级；中文须自研
Agent 编排	LangGraph	生产默认；节点级断点续跑
内部 Agent 平台	Coze Studio （Apache2.0）	Dify 附加条款须法务认定
语音栈	FunASR + CAM++	中文 ASR/ 分离 / 情感一站式



首年 TCO 参考区间 \$320K-950K（咨询机构估算口径）；**数据工程人力占预算 ≥ 50%**——成本是哑铃形，控范围比选组件更重要。

Build-vs-Buy 决策框架：外购成功率 67% vs 自建约三分之一（MIT 口径，方向性证据）——客服单点环节优先外购（单价已透明至 \$0.5-2.5/ 解决）；**数据底座、OneID、画像 / 记忆、场景编排必须自建**，因为自建赢面在数据主权与跨域编排。Composable 隐性成本：3-5 名专职数据工程师（年化 \$45-100 万）。预算纪律：按 " 试点 ×3-4 倍 " 报生产化（60% 项目超支 30-50%）。CARIAD 累计 € 12-14B 的回撤案例与本项目差 3-4 个数量级——**我们的风险在组织与数据，不在金额。**

12 个月四阶段：红旗范式起步，客户面场景最后上



30 天立即行动（均不依赖平台采购决策）

- ① Vios 录音双轨核查（采集团队）
- ② 渠道数据归属谈判启动，带上 " 共享对价 " 清单（商务 + 法务）
- ③ 企微会话存档开通评估（法务 +DPO ）
- ④ 一期两店 A/B 试点设计 + 基线确定（运营 + 数据）
- ⑤ consent 中心需求冻结（数据团队）
- ⑥ 记忆组件 golden set 出题（30+ 题）

风险登记册 Top3（完整 Top5 见报告 10.4）

- ① **渠道数据归属**——非技术风险，谈判先行
 - ② **卖方口径主导的度量环境**——买方基准（线索有效性 25-28%、成交成本约 2200 元）做 sanity check
 - ③ **企微平台条款单向可变**——身份与画像资产落自有湖仓
- 首年预算参考 230-680 万元，人力 ≥ 50% ；先修数据与协议，再谈平台与模型。