

未来已来，注入硅基生命力

Best Results

6608.HK

2026年9月



免责声明

本文件是针对百融智能(以下简称“公司”)及其子公司和合并附属实体(以下统称“集团”)编写的, 仅供本次演示使用, 不用于任何其他商业或其他目的, 未经许可严禁截屏或转载。通过出席本报告或阅读本文件, 您(i)同意受这些条款的约束, 并(2)承认您将自行负责对市场 and 集团的市场地位进行评估, 您将自行进行分析, 并自行负责形成您对集团业务的潜在未来业绩的看法。本文件包含机密和专有信息。此文件内容未经独立核实。本文件中的某些事实性陈述和预测来自外部来源, 未经本公司、本集团任何成员或其任何董事、高级管理人员、雇员、代理人、附属公司、顾问或代表核实。本协议中所包含的信息或意见的公平性、准确性、完整性或正确性不作任何明示或默示的陈述、保证或承诺, 且不应加以依赖。本公司、集团成员、或其任何关联公司、或其任何董事、高级管理人员、雇员、顾问或代表均不对本文件中包含或提交的任何信息或与本文件有关的其他任何损失承担任何责任(由于疏忽或任何其他原因)。本文件中包含的信息, 其准确性无法保证, 在本文件发布之日提供, 可以在不另行通知的情况下进行更改, 不会更新或以其他方式修订以反映本文件发布之日之后可能发生的任何事态发展。

本文件包含前瞻性陈述, 反映了本公司在本文件中分别指出的日期对未来的当前信念和期望。这些陈述包含“预期”、“相信”、“打算”、“估计”、“期望”等词和类似含义的词。本文件中所包含的历史事实陈述以外的所有陈述, 包括但不限于有关中国经济、行业发展趋势和公司未来财务状况、经营结果、经营战略、未来经营的计划和管理目标(包括与公司产品和服务有关的发展计划和目标)的陈述均为前瞻性陈述。这些前瞻性陈述基于一些假设, 这些假设受制于已知和未知的风险、不确定性和公司无法控制的其他因素, 例如公司未来经营的政治、社会、法律和经济环境, 这些因素可能导致公司的实际结果和业绩与这些前瞻性陈述中表达或暗示的结果和业绩存在重大差异。这些前瞻性陈述仅反映本文件日期时公司管理层的观点, 并非未来业绩的保证。因此, 不应过分依赖这些说法。本公司、集团成员或其任何相关关联方, 或其任何董事、高级管理人员、员工、顾问或代表均不承担任何义务, 就该等日期或其他日期之后发生的新信息、事件或情况更新或以其他方式修改这些前瞻性陈述, 集团成员明确否认有任何义务或承诺传播任何更新或修订的任何前瞻性声明, 包括在此。本文件及与本文件有关的任何口头信息不构成或构成本公司或其任何股东、子公司或在任何司法管辖区的附属公司的证券出售或发行要约的一部分, 也不应被解释为购买或收购本公司或其任何股东、子公司或附属公司的证券的要约的招揽, 也不构成参与投资活动的诱惑。本文件的任何部分或其分发的的事实不得构成任何合同、承诺或投资决定的基础或被依赖。在香港, 本公司的股份不得向公众发售, 除非该等股份的发售已获香港联合交易所有限公司正式授权, 根据《公司(清盘及杂项条文)条例》(第32章)向香港公司注册处处长注册, 并已由香港公司注册处处长正式注册。未经注册的招股说明书不得在香港发行、发行或传阅, 但根据《证券及期货条例》(第571章)及其相关规则向专业投资者发行的情况除外。任何购买或认购证券的决定应完全基于本公司就有关发行而发出的招股说明书或其他发行通告所载的资料。本公司的任何证券已经或将不会根据美国《1933年证券法》(以下简称“证券法”)修正案进行注册, 除非根据《证券法》的豁免或在不受《证券法》注册要求约束的交易中, 本公司的证券不得在美国出售或出售。此处包含的信息并非(a)向您或公众提供本公司的股票或证券, 以现金或其他对价购买;或(b)向您或公众发出以现金或其他对价购买本公司任何证券的邀请。本文件不用于在美国或向“美国证券公司”发行, 也不构成出售或招揽购买或认购美国证券的要约《证券法》S条例中定义的“人”。该文件及其任何副本不得携带或传送到美利坚合众国及其领土或属地, 或直接或间接地分发给美利坚合众国及其领土或属地或向任何美国人。任何未能遵守此限制的行为都可能构成对美国证券法的违反。本文件不是在美国出售的证券报价。没有注册或豁免注册的证券不得在美国出售或出售。在美国进行的任何公开发行证券都将通过一份招股说明书进行, 该说明书可以从发行人处获得, 其中将包含有关公司和管理层的详细信息以及财务报表。通过出席本报告或阅读本文件, 您将被视为承认收到了有关本公司和集团的重大非公开信息, 您将被视为同意并同意(i)您应受本文件规定的限制的约束, 并对披露的任何信息(无论是书面的还是口头的)保持绝对保密, 您应赔偿并继续赔偿本公司、集团成员、其各自关联公司及其董事、高级管理人员、员工、顾问和代表免受因您直接或间接违反本免责声明中的任何条款而可能导致或招致的所有损失、损害赔偿、开支和成本;(ii)您在非根据适用法律和证券交易规则使用本文件中的信息时, 可能构成对适用法律、规则或法规的违反, 包括在内幕交易/交易方面;(iii)您将独自承担因任何该等违反而产生的任何后果。在其他司法管辖区分发 本文件中的信息可能会受到法律限制, 持有本文件的人应告知自己并遵守任何此类限制。本协议中所述的证券尚未并将不会根据美国、加拿大、澳大利亚、中华人民共和国(“中华人民共和国”)或日本的适用证券法进行登记, 并且, 在某些例外情况下, 不得在美国、加拿大、澳大利亚、中国或日本或向美国、澳大利亚、加拿大、中国或日本的任何国民、居民或公民提供或出售。本文件中的材料在英国传递给(i)在《2000年金融服务与市场法(2005年金融促进)令(经修订)》(以下简称“令”)第19条范围内的投资事项方面具有专业经验的人士;(ii)命令第49(2)(a)至(d)条所述的高净值法人团体、非法人团体和合伙企业以及高价值信托的受托人;或(iii)发送给可能以其他方式合法获得信息的人士(第(i)至(iii)所述的所有此类人士统称为“投资专业人士”)。本文件仅针对合格投资者和投资专业人士, 任何其他人士不应依赖或依据本文件或其任何内容行事。

在任何情况下, 本陈述的提交或本公司与任何收件人的任何进一步讨论均不应产生任何暗示, 表明本公司或集团的事务自该日期以来未发生任何变化。

本免责声明须根据香港法律解释。

“

我们是**聚焦AICC(智能联络)领域SOTA级垂类基础大模型开发与企业级Agent落地一体化企业**，真实业务沉淀高价值数据，驱动大模型迭代飞轮。

”

AICC是全球三大企业级AI落地场景之一

判断AI能在企业跑起来的3个标准：工作遵循文档化规则；碳基能判断输出是否正确；明确的碳基兜底路径；

百融聚焦场景



AI Coding 软件工程开发

No.1 重塑开发范式

硅基研发员工，颠覆传统效能，
大幅缩短开发周期。



AI Support¹ 智能联络交互

No.2 改造传统碳基模式

智能识别意图与精准应答，实现
成本中心到利润中心的跃升。



AI Search 企业智能搜索

No.3 重构信息入口

从被动检索到主动洞察，
大幅压缩信息获取与决策周期。

数据来源：Andreessen Horowitz (a16z), "Where Enterprises are Actually Adopting AI," April 2026.

1. a16z 原文落地场景名称是Support，主要指AI客服能力，本质上是AICC智能联络领域。

业务持续落地沉淀高价值私域数据，驱动大模型迭代飞轮 构筑百融核心竞争护城河。



百融自研三类领域专属大模型

覆盖语音交互、文档理解和营销决策，构筑企业级Agentic应用底座

BR-Voice

语音多模态大模型

聚焦实时语音交互闭环

BR-LLM-Proactive

主动引导式对话大模型

专注销售场景的意图识别和主动引导能力

BR-Vision-Doc

文档理解多模态大模型

聚焦多语种、多版式文档解析

领域专属大模型能力总览

语音 多模态大模型

聚焦实时语音交互闭环

CLAR CER

语音转写准确度，数值越低越好

0.92%

优于PAC/CFL/阿里SeACo方案

热词检索 F1

理解专属行业关键词的能力

97.03%

高于GLCLAP和阿里SeACo方案¹

TurnSense F1

AI判断客户对话状态能力，>95%为交互友好度高

96.35%

优于Easy/TEN/Smart专注
对话轮次检测方法

主动引导式 对话大模型

专注销售场景的
意图识别和主动引导能力

成交意向判断率

AI对客户购买意愿信号的理解力

90%

能力领先，4B小模型即达高分

意图识别准确率

AI识别用户潜在意图，100%最强

92%

8B小模型接近
外部14B+大模型效果上限

主动引导能力

AI主动说服客户进入下个销售环节能力

近人工销冠

真实业务数据
"主动引导"能力持续提升

文档理解 多模态大模型

多语种、多版式文档解析

CER

AI识别文字的准确性，数字越小越好

0.02

行业领先

TEDS

AI识别表格结构准确性，100%最强

95.80%²

拍照表格识别行业领先

ToC解析准确率

AI提取文档结构准确性，100%最强

96%

优于MinerU2.5方案³

两项国际挑战赛，验证百融领先语音垂类基础大模型能力

IEEE SLT 2026 SmartGlasses Challenge

智能眼镜语音交互挑战赛 | 多人语音识别理解

考察：穿戴场景、多人说话的真实环境，
能否实现从听清到听懂的端到端贯通

4个任务,全进 **Top 5**

全球范围

MLC-SLM 2026 Challenge

INTERSPEECH 2026挑战赛 | 多语种对话语音理解任务

考察：复杂长对话里能否听懂

Top 3 **Top 1**

全球范围

国内产业界

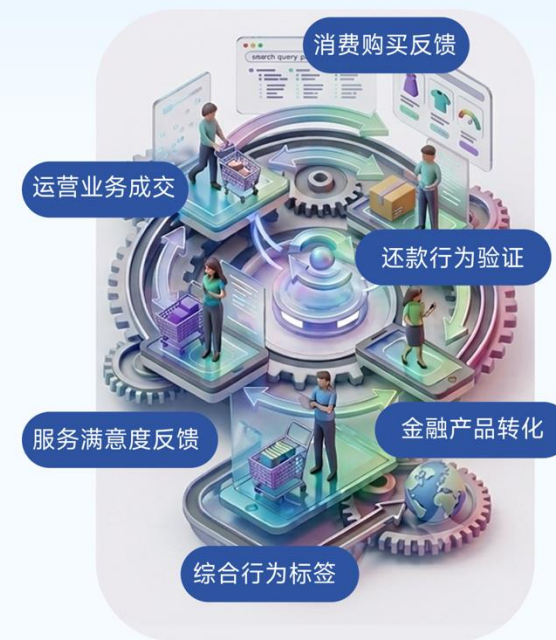
原生真实且不可合成获得的私域数据*， 构筑百融SOTA级垂类大模型的核心壁垒。



超 **10.5** 亿
合规个人数据



超 **1.5** 亿
合规企业数据

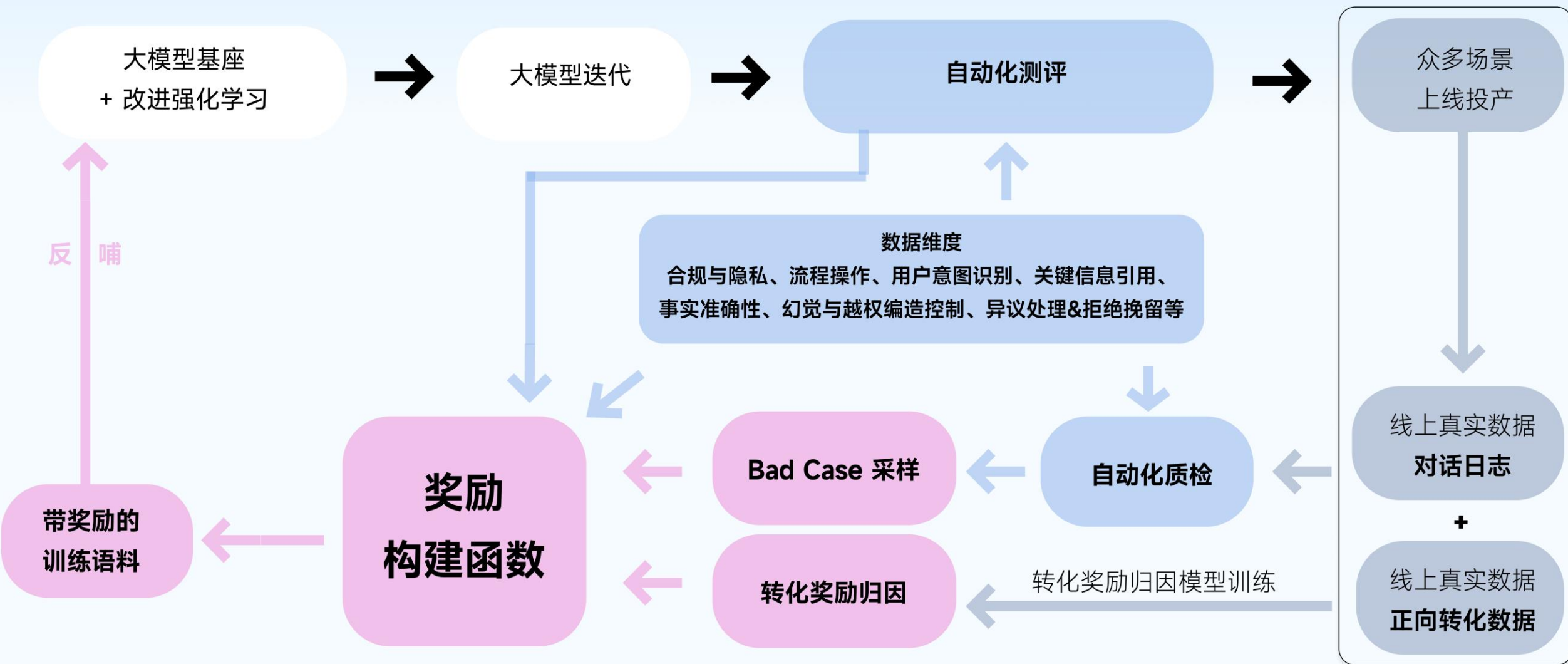


持续增加
的真实业务行为反馈数据

*数据采集与使用均符合相关法律法规。

真实业务数据构建奖励函数，反哺大模型强化训练闭环。

「自动化质检」与「奖励归因」后构建奖励训练函数，反向夯实大模型训练，每一次循环都在提升领域专属大模型的能力与壁垒。



百融领域专属大模型高性价比： 推理成本更低，公私有部署灵活，高效完成垂类任务。

百融模型性价比优势

0.4-30_B 90%

百融小参数模型¹
垂直任务上
领先通用模型

百融模型
输入Token成本
节省幅度²

1. 涵盖语音类模型、TTS模型、OCR模型等

2. BR-LLM-Proactive主动引导式对话大模型，相比市面如DeepSeek等高性价比大模型的对比

已合作行业客户

通用大模型公司



头部互联网



三大运营商



头部车企



头部金融机构



我们发现客户决定购买AI服务的两个考虑： 能否帮我提升收入？ 能否帮我降低成本？

硅基员工 (AI岗位外包)

降低成本：百融通过派驻硅基员工，按碳基员工的岗位收费，但工资远低于碳基员工；

APO (AI业务外包)

提升收入：百融的硅基员工承担整个业务流程，帮企业主完成结果，客户按结果付费；

调用API 构建专属Agent

降本提效：百融提供构建硅基员工平台，按客户调用API所产生的Token量收费；

我们认为：企业主最有机会为以上3种模式买单

我们通过打造AI结果云平台，为客户提供灵活买单模式。

智能体商店-百汇™



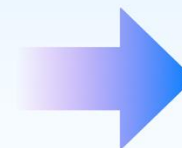
开箱即用、可复制、
可泛化的硅基员工



能力可用、逻辑不可
见的产权保护机制



可计量、可计费，
支持收益分配



AI岗位外包
or
AI业务外包

智能体平台-百工™



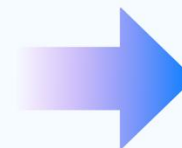
LLM Ops
兼容通用与领域大模型
基于场景智能推荐最优模型
LLM结果可观测、可计量



Agent Builder
低代码可视化编排，企业级知识库/
专业语音套件，插件/MCP/沙箱，
智能体评测、优化、一键发布



Agent Runtime
稳定运行时封装，效果可视可衡量
智能体在线自迭代，
内置计价与分成功能

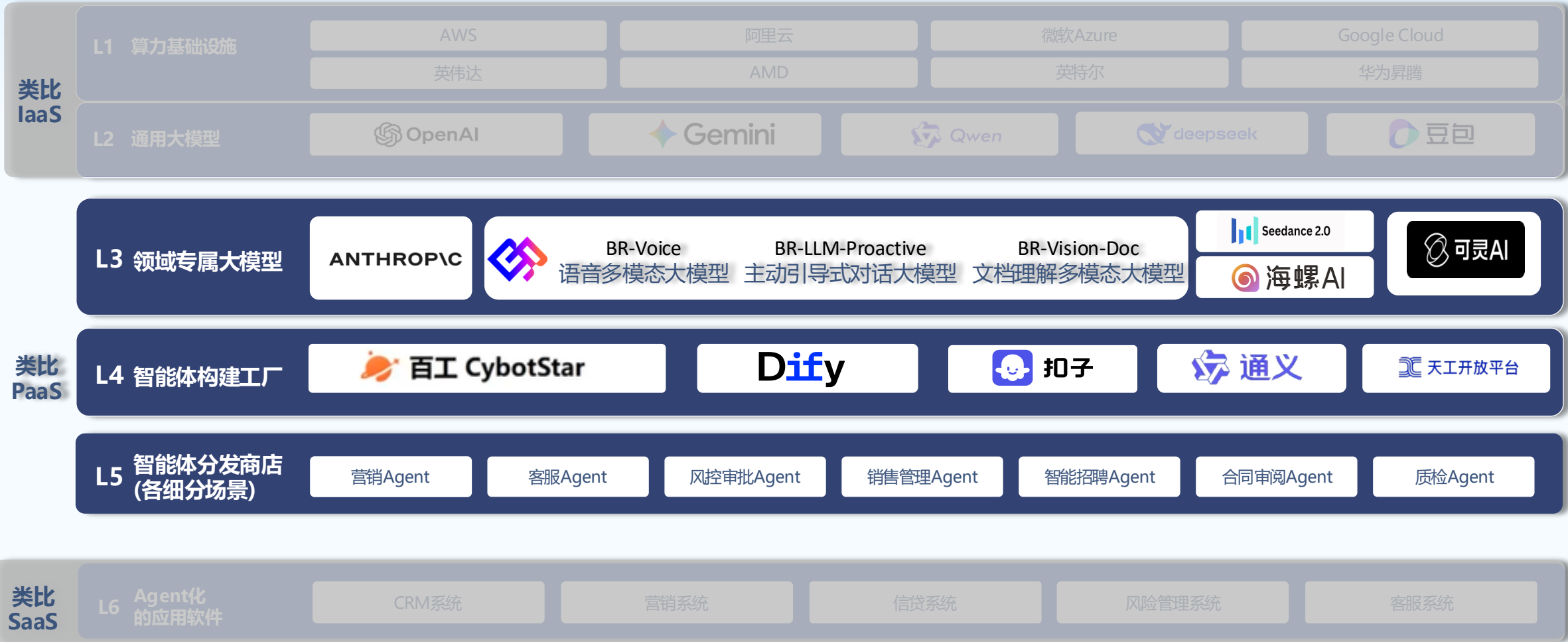


调用API构建
专属Agent

AI Infra-百基™

算力基础设施 + **领域专属模型** + 推理引擎

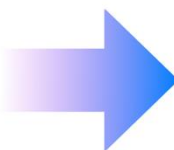
百融AI结果云平台架构的行业类比



**核心业务聚焦：输出AICC领域硅基客服，
AI化改造「昂贵低效的传统碳基客服模式」。**



AI化
改造



中国Contact Center人力成本巨大，企业主对AI化改造需求迫切。

1.9万 亿元

全场景人力成本测算

核心假设

(单位：人民币)

分类	估算从业人数 (万)	人均年薪 (万)	估算总年薪 (亿)
根据AI工具可替代1) 电话营销；2) 分歧调解；3) 信贷；4) 客服与运营等坐席，进行测算：			
电话营销	1100 ¹	10.3	11,300
分歧调解	200 ²	7.9	1,580
客服与运营	1,000 ³	6.0	6,000

1.经专家访谈调研和电话客服行研报告，预估电话营销行业从业人数与年薪；

2.经专家访谈研究，预估分歧调解从业人数；

3.经行业专家访谈调研，预估银行、电商、运营商、物流、航司、OTA、教育、医疗等行业客服与运营从业人数。

中期来看，百融将聚焦以下新场景，
赋能企业与外部客户的智能交互管理。

物流

航司

银行及券商

运营商

其他场景

分歧调解、
车险、教育及医疗电销等



百融硅基客服不只是完成对话，而是要帮助企业客户实现
营销触达、意图洞察、销售转化、客户服务

基于以上场景，百融中期可AI化改造的人力成本测算¹



约**1,496**亿元
中期可AI化改造的人力成本

1. 百融AICC业务中期聚焦的新场景中，可被AI化改造的碳基坐席成本池
2. 碳基客服预估数：基于BPO行业专家访谈、百融AICC业务线访谈的估算人数
3. 估算总年薪 = 该场景碳基客服预估数 x 中期(3-5年)预估AI改造率 x 该场景的碳基客服年薪

核心假设

智能联络硅基客服基于不同AI改造率，优先聚焦以下场景：

(单位：人民币)

场景分类	碳基客服 预估数(万) ²	中期可AI化改造 碳基客服总年薪(亿) ³
物流	15	115
航司	6	49
运营商	25	126
银行及券商	103	794
分歧调解	30	216
其他(教育及医疗电销、车险等)	28	197

数据来源：
A.碳基客服数：中证协官方披露数据、中银协报告、各股份制银行招聘信息、三大运营商招聘信息、行业专家访谈；
B.替代率估算参考：
-Anthropic《Labor Market Impacts Of AI》- Observed Exposure(Customer Service Representatives) ~ 70.1%；
-麦肯锡《Global Banking Annual Review 2025》- The Net Effect On Banks' Aggregate Cost Base 15%-20%；
-招行25年股东大会数据参考-AI处理业务对应的工时和人类员工实际投入工时的比值~1:9；
-BPO专家访谈；
C.平均人力成本参考：各大招聘网站岗位薪酬数据、BPO行业专家访谈；

基于新一代AICC技术的百融硅基客服， 不止会对话，更能在企业服务规范下稳定完成工作。

新一代AICC

服务场景 | 呼出场景，复杂呼入场景

技术路径 | 语音多模态，主动对话大模型

适合情景 | 开放提问，长多轮对话，自然打断

深度绑定 | 调度知识库，接入业务系统

VS

传统AICC

服务场景 | 优先落地在呼出场景

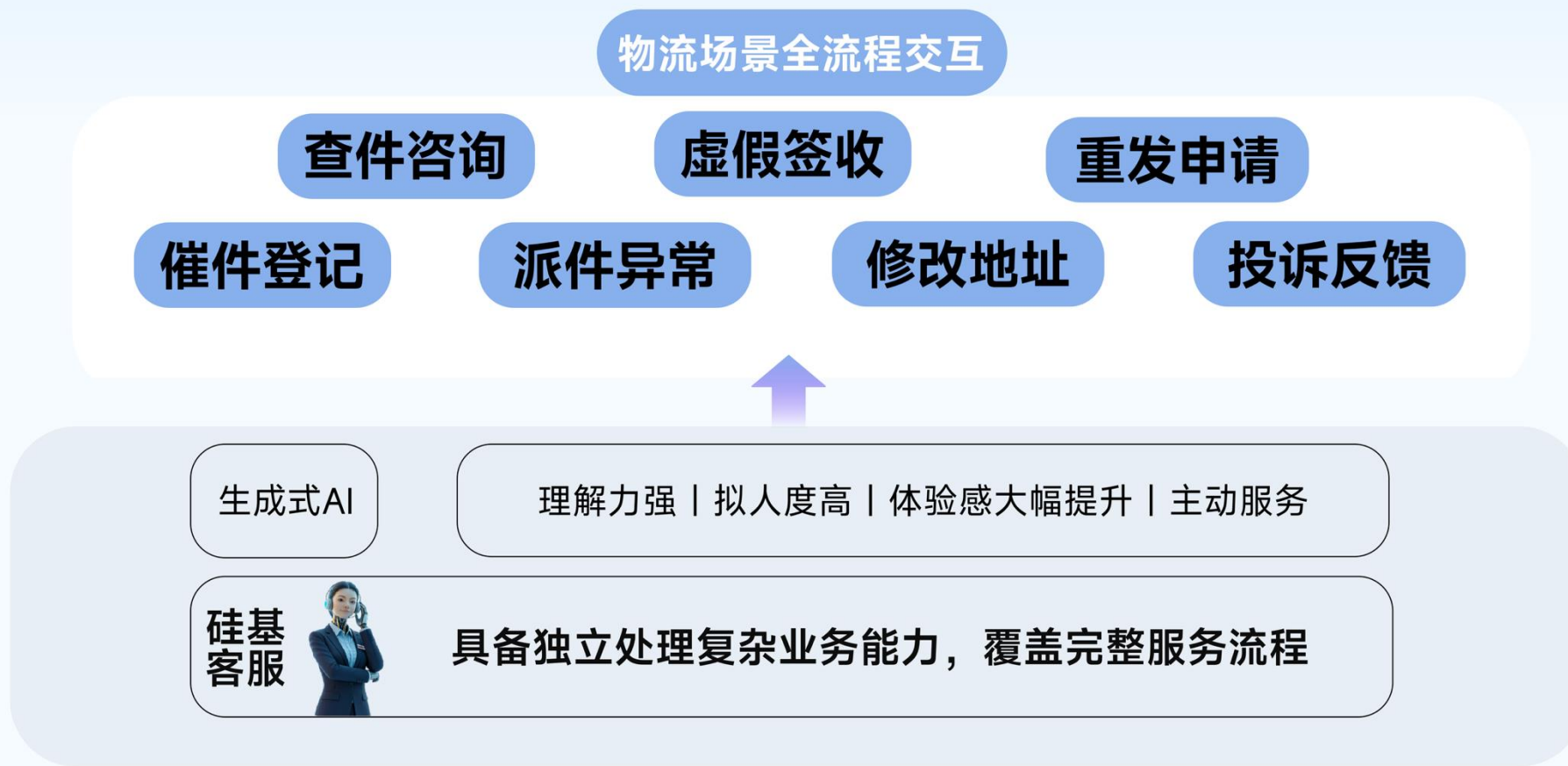
技术路径 | NLP，关键词匹配，预设话术

适合情景 | 脚本外呼，通知提醒，菜单应答

局限情景 | 开放提问，多轮复杂对话，自然打断

从“预设话术，机械应答”到“开放式理解的人机交互，与业务落地闭环”

物流场景：硅基客服大幅提升物流工单处理效率。



物流场景：精准查找定位运单，减少人工介入。

我有一个快递显示放到快递柜，但我没有收到取件码。



好的，您有一个放到蜂巢柜没收到取件码的快递，请问运单号是多少呀？我帮您重新发一下取件码。



我看下，运单号是43418587。



我记一下，43418587，这是八位数字，请问后面还有吗？



09181481。



好的，您的运单号是434185870981481。



太快了，我没听清，你慢一点。



好的不好意思，我慢一点说，您说的运单号是434185870981481。



有两位有问题，第一个是倒数第三位应该是5不是4，还有个倒数第四位应该是2不是1。



好的，我重新拼一下，435185870982581，一共15位，对吗？我帮您查一下。



协助查件，重新发送取件码



提供单号位数不准确



准确识别，引导用户补齐单号



客户要求放慢播报速度



语气和情绪识别，按照用户要求实时响应调整语速



根据两段客户语音自动合并完整的运单号，在长数字中定位片段进行修改

物流场景：硅基客服部分工作效率、成本优于碳基客服。

单位：人民币

硅基

2分钟

平均处理话务时长¹

近碳基水平

满意度

1.8万通

月处理通话峰值²

远低于碳基

硅基客服平均工资

VS

碳基

4分钟

平均处理话务时长

98%

满意度

0.5万通

月处理通话峰值

8-10万/年

碳基客服全口径³平均工资

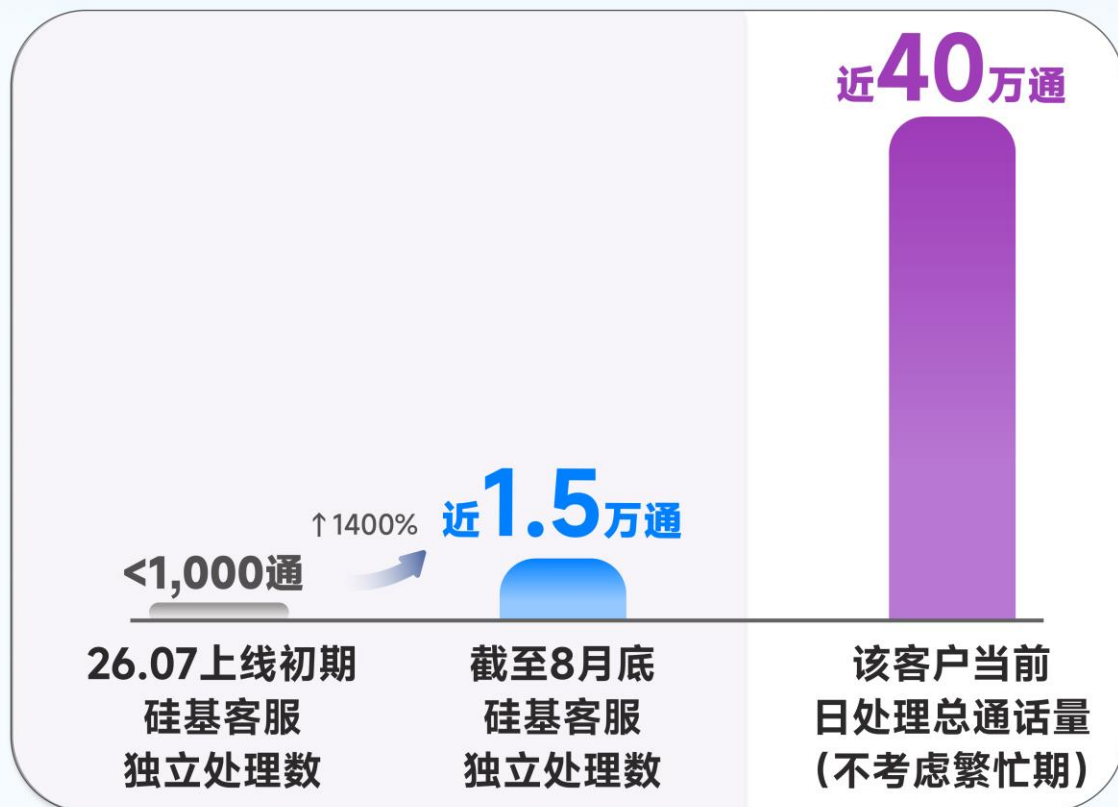
1. 含通话时长及挂断电话后话务处理时长，硅基客服话务处理时长远低于碳基客服。

2. 碳基员工：月工作26天，每日上1班（8小时），每班接通峰值为200通。硅基客服：月工作30天，每日上3班（每班8小时），每班接通峰值为200通。

3. 企业实际承担成本，含工资 + 社保 + 福利 + 场地分摊，数据参考各大招聘网站。

物流场景：硅基客服日处理量快速攀升，承接复杂呼入查询

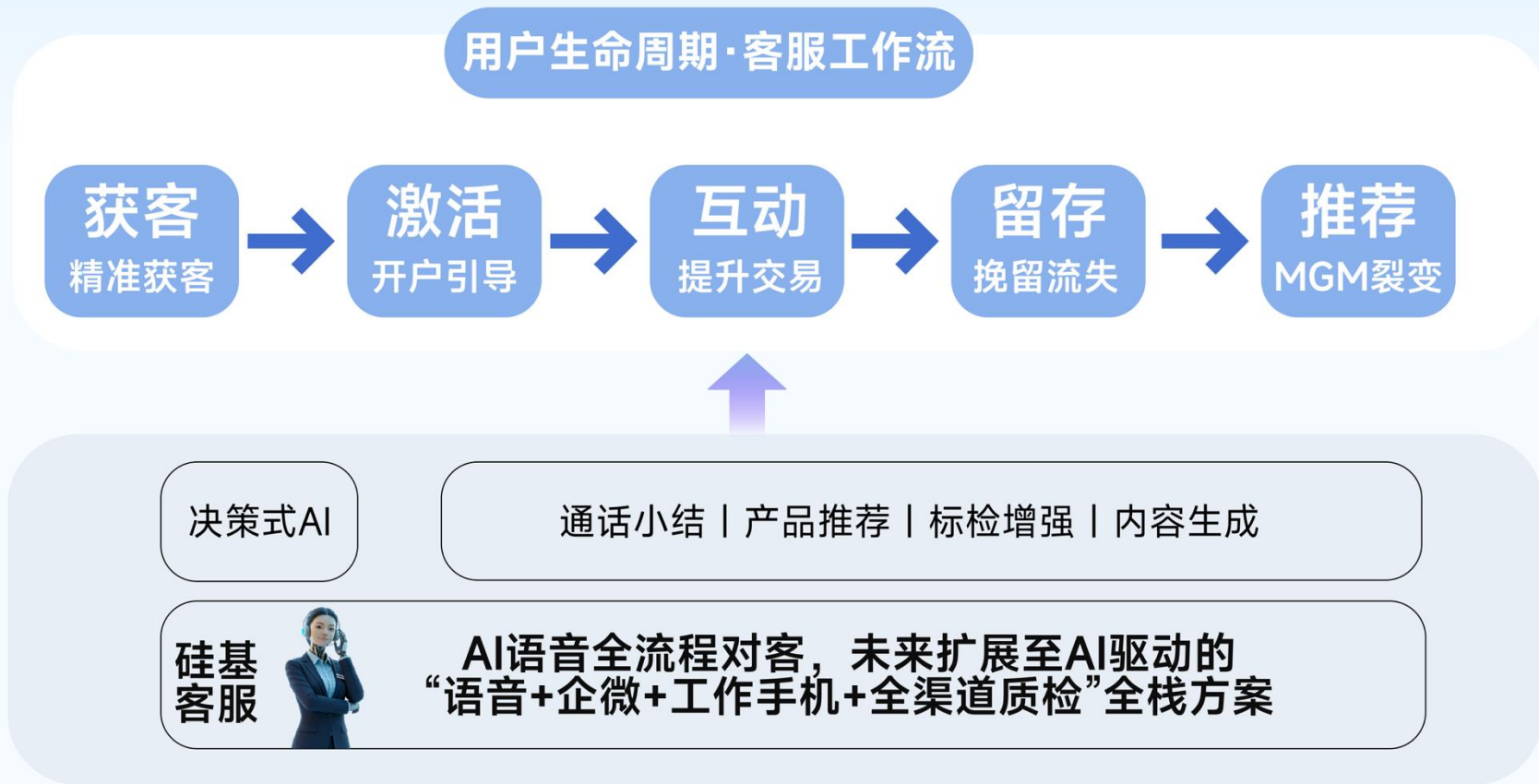
某客户 | 硅基客服日处理通话数



硅基客服与碳基用户 最长通话时长

超 **20** 分钟
并成功独立处理问题

券商及银行场景：硅基客服服务用户全生命周期，承接转化任务。



券商场景：硅基客服实时调整话术，高效引导营销转化。



您好，这边是XX证券，来电是想提醒您，目前有个老客户专享活动，约定年化收益6%，5000元起投，期限...

不需要，最近没钱，理财现在都亏，不买。



用户强势打断，典型的防御心理、市场信心不足、直接拒绝。



AI打断，实施调整沟通策略。

AI识别客户防御心理，实时大模型微调话术和语气，及时的情绪识别到话术回复（反应时间在200ms内）。



特别理解您！正因为现在市场不好，我才不建议您去看高风险的产品。我们这两天针对“求稳”的老客户，刚上线了一款“纯债+大额存单底层”的半年期闭合新品，非常适合用来做您资金的“避风港”。

你们上回推荐那个，到期根本没拿到宣传的收益，都是骗人的。



开始动摇，但依然挑剔，客户翻旧账，质疑信誉。



情绪安抚+专业Know-how解释，顺势提出加微链接。



您提的这个情况太关键了，我先代表团队跟您道个歉。但也正是吸取了之前的教训，行里这次推出的新品，100%锁定了AAA级央国企的稳健债。直白点说，就是用“国家队”的信用在帮咱们降噪。它过去的兑付率在99.2%以上。如果我把这款产品的历史业绩曲线微信发您瞅一眼，您现在方便加个微信吗？

行吧，那你发我看看。



大模型动态商机捕捉+挂断自动触发。



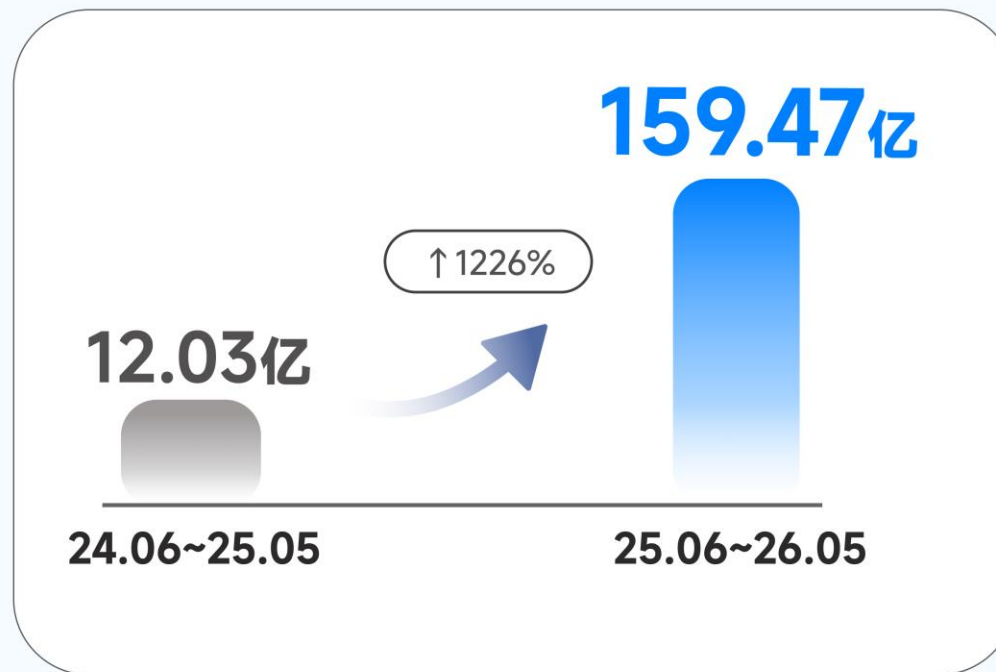
太好了，李先生，不耽误您时间。我这边已经把带您专属标签的微信号通过短信同步发到您手机了，您直接点击就能添加。另外，我刚刚在系统里帮您把这款产品的专属申购额度锁定了24小时（非强制）。等您加了微信看完资料，咱们随时微信上聊，祝您生活愉快！

券商场景：硅基客服在部分环节效率优于碳基客服¹。

单位：人民币



百融硅基客服 促成的交易额²



1. 对比数据为券商在百融AI语音过程中，同样用户的人工AB测试结果。

2. 客户通过百融硅基员工触达顾客并成交的金额，按实际接通数 x 接通转化率（参考真实项目数据）x 人均入金金额（参考真实项目数据）x 人均年换手率（参考行业水平）测算。

券商客户数快速增长，已合作券商持续增购硅基客服¹。

30[↑]

已跑通的业务场景，
如休眠激活、
断点挽回、入金激励等

210[↑]

↑ 600%

上线某头部券商半年
硅基客服增至

13^{家¹}

↑ 333% YoY

已签约&
合同流程中券商

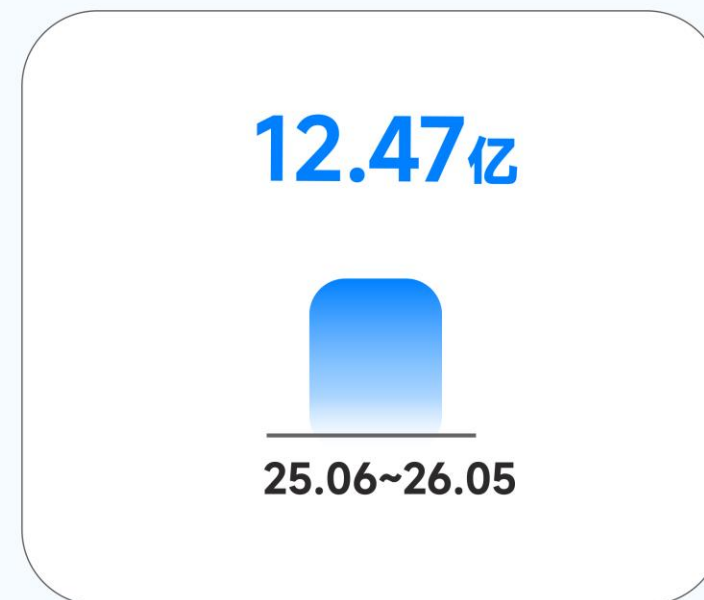
1. 截至2026年8月底。

银行场景：硅基客服在部分环节效率已优于碳基客服¹。

单位：人民币



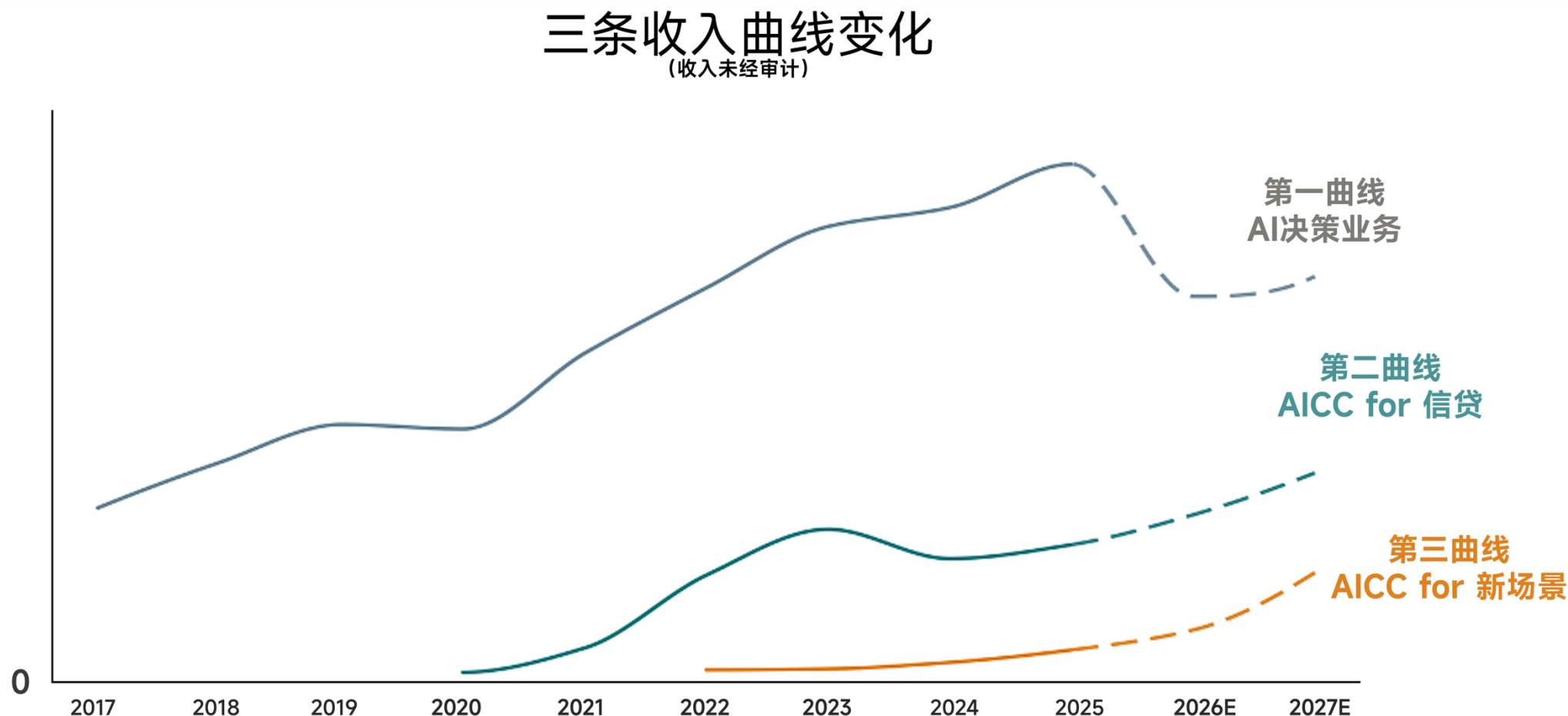
百融硅基客服促成的交易额²



1. 对比数据为券商在百融AI语音过程中，同样用户的人工AB测试结果。

2. 客户通过百融硅基员工触达顾客并实际促成的GMV金额。

业务规划：积极在新场景落地AICC技术



第一曲线:2014年基于决策式AI，为金融机构提供智能分析及决策能力，压舱石业务提供现金流。26年处于业务调整期，预计27年回稳。

第二曲线:2017年百融开始为金融机构进行信贷产品的营销推广，25年后响应政策调整方向，聚焦「24%利率以下产品」的营销推广。

第三曲线:集团核心战略业务，2026年聚焦为券商、物流、运营商、航空等新场景提供AICC能力，业务持续增长。

除了助力企业主与外部客户交互， 还更好赋能员工与客户的交互。

百销·销售管理 | 企业数据流物理入口

将碎片化对客沟通、客户真实反馈
提炼沉淀成宝贵、可复用的企业资产

门店经营动态

真实对客记录

拜访客户交流

客户沟通反馈



零售门店管理

1. 经营动态随时可获
2. 督查效率提升90%+
3. 自动生成经营洞察报告辅助决策

一线销售管理

1. 规范销售话术标准
2. 避免违规过渡承诺
3. 金牌销售经验沉淀为成单技巧

AI赋能零售门店管理案例： 信息收集效率与决策速度双层跃升

过去：滞后且缓慢

AI赋能后：即时且迅速

门店经营情况汇报

👤 人工记录层层上报，信息传递慢且失真。

👤 智能录音转写，每日晨会结束，总部即时获取**所有门店纪要和AI洞察经营报告**。

督查工作时长反馈

👤 400+家门店实地调研需200天，整理汇总还需额外1-2个月。

👤 AI自动化整理与分析，**半小时**完成全域数据处理，生成标准分析报告。

管理者决策速度

👤 时效性低，决策需依赖层层汇报的信息。

👤 基于实时、全面的一线门店数据支撑快速决策。

某餐饮连锁集团（400+分店）案例

AI赋能一线销售案例： 数字化沉淀销冠经验提升成单效率

过去：不透明且混乱

销售沟通动态反馈

👤 销售人员分散，销售拜访、沟通管理难度大。

客户需求理解

👤 依赖主观汇报，客户需求反馈失真。

销冠经验复用

👤 成功销售话术、经验难以结构化总结传播。

AI赋能后：实时且清晰

👤 拜访报告**实时生成**，随时把握一线动态。

👤 客户反馈投喂AI分析，及时生成结构化客户需求体系。

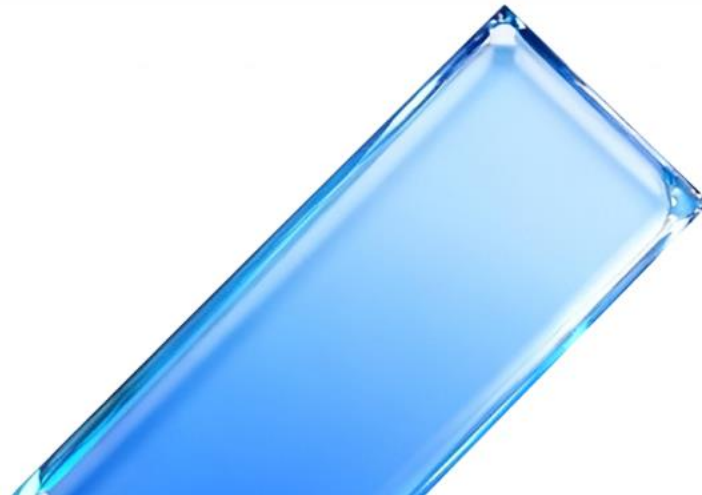
👤 辅助**优化销售话术**，提升销售能力和组织管理效能。

某制造业案例



百融智能
resultscloud.com

财务表现



26H1 收入组成

单位：人民币，YoY%

9.14 亿

↓ 43%

AI决策业务

基于决策式AI，为金融机构提供审批与风控硅基员工，自动完成从数据分析到决策输出的全流程，替代传统人工审批岗。按API调用量或年费收费。

3.72 亿

↓ 26%

其他业务

1.基于部分信贷产品的智能营销及运营收入。
2.赋能保险经纪人提升销售转化，按促成保费收取佣金分成。

3.77 亿

↓ 62%

AICC智能联络业务

基于 AICC 技术，面向多行业部署硅基客服员工（企业级 Agent），承接客服、营销、查询及客户运营等工作。按结果付费，涵盖工作量、派驻员工数及规模分润等。

1.66 亿

↑ 52%

AICC for 新场景

1,126 万

↑ 195%

AICC for 信贷场景

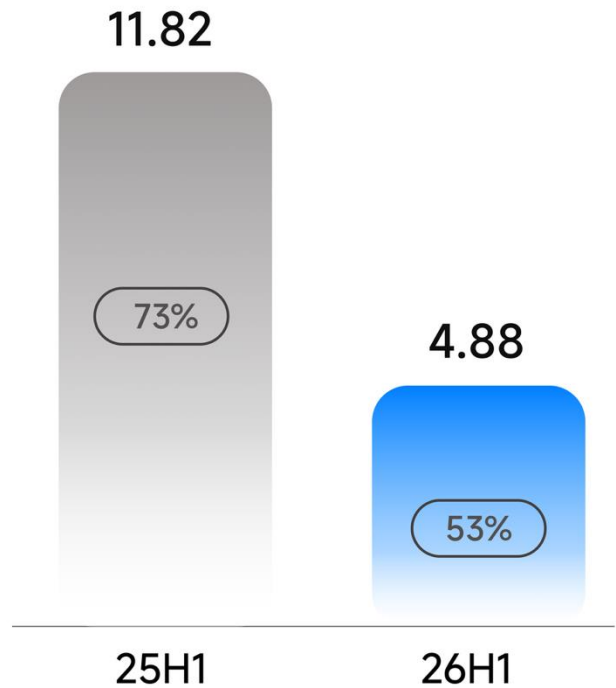
1.55 亿

↑ 47%

26H1 毛利及毛利率

单位：人民币亿元，YoY%

本期变动原因：毛利率相对更高的老业务收入及占比下降所致。



毛利YoY

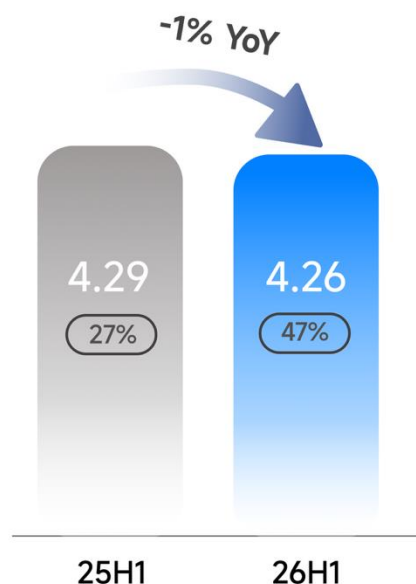
↓ 59%

26H1 成本及费用（占收入比）

单位：人民币亿元, YoY%

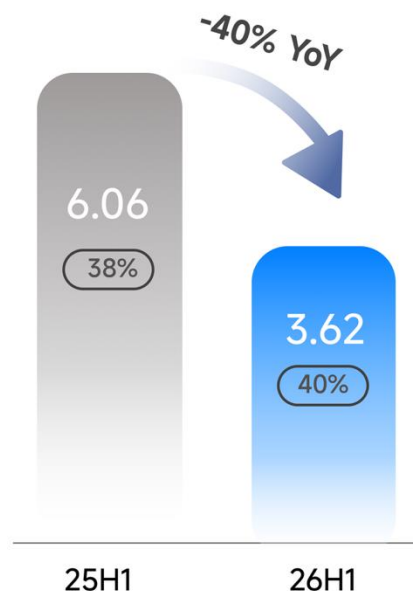
销售成本

成熟业务收缩，数据及运营成本大幅下降，但被原 BaaS 保险云收入增长带动的经纪佣金上升所抵消，报告期内微降。



销售费用

成熟业务收缩，减少相关投放获客。



研发费用

为夯实AI人才基础并增强核心竞争力，集团加大AI研发人员投入。



管理费用

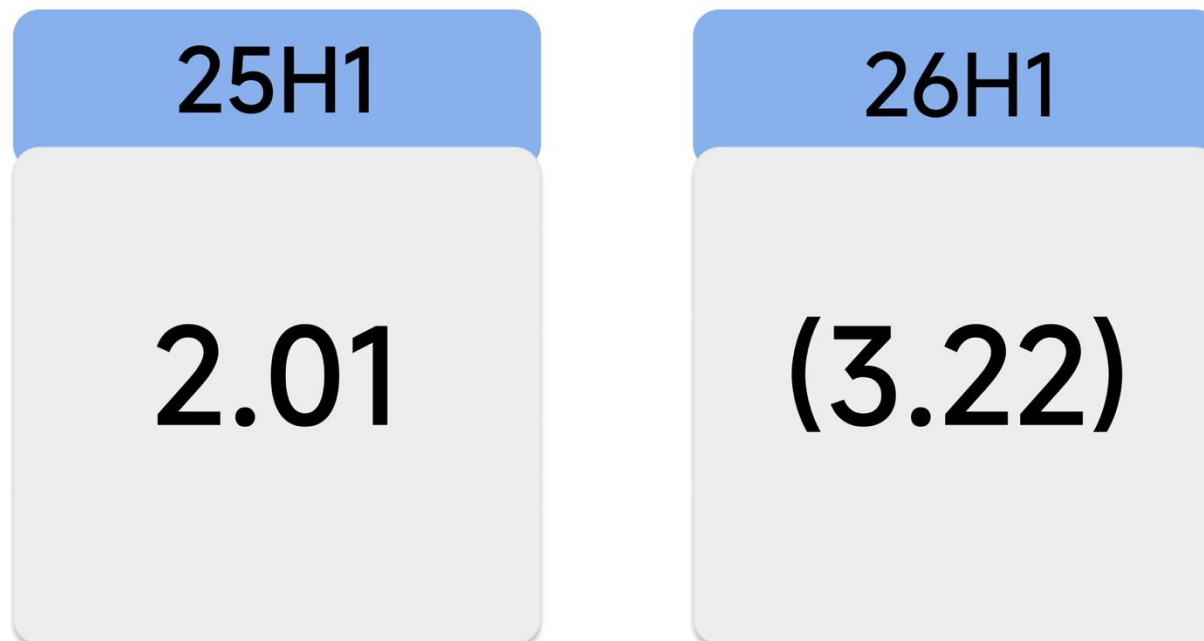
因咨询及其他专业服务费减少所致。



期间溢利

单位：人民币亿元

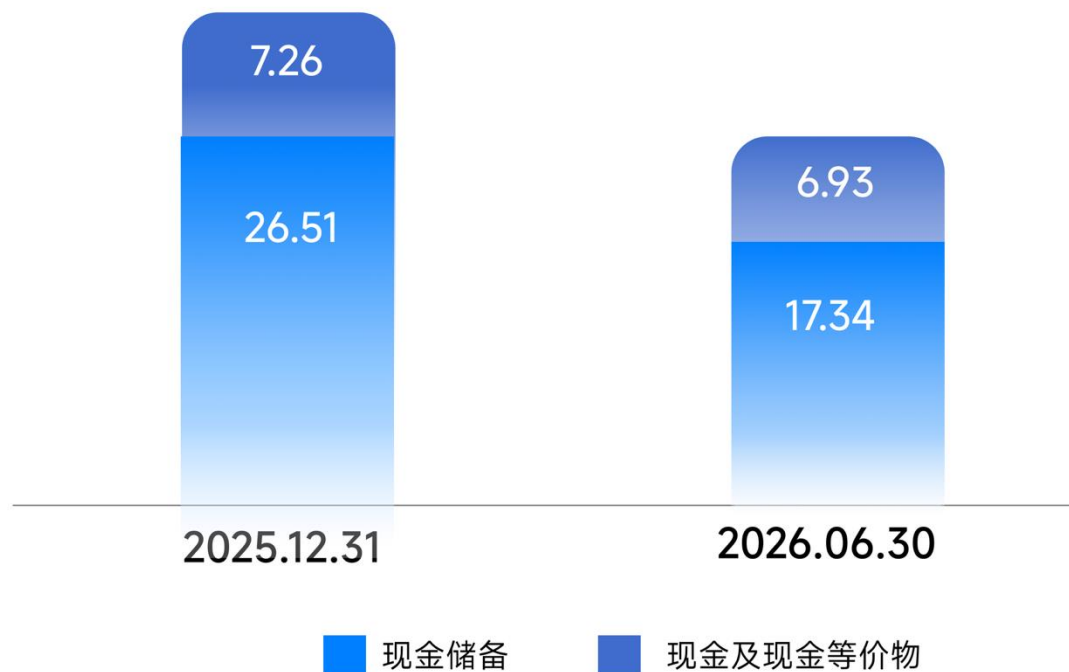
本期变动原因：毛利率相对更高的老业务收入及收入占比下降带动毛利下滑，毛利下滑幅度大于费用下滑。



现金及现金等价物、现金储备

单位：人民币亿元

上半年现金主要用途：经营净流出 3.80 亿元、股权投资 3.11 亿元、股份回购 1.41 亿港元。集团无有息负债。



*现金储备包括受限制现金、定期存款及理财投资



百融智能

resultscloud.com

补充信息一：

SOTA级领域专属大模型介绍

01 语音多模态理解大模型BR-Voice：实时语音交互

极快的响应速度，交互过程中能感知到用户的副语言信息(如性别、年龄、情感等)，对话过程支持用户打断、对话轮次的检测，模型支持多个可选音色、音色克隆、合成语音的表现力控制。

BR-Voice 技术指标

语音识别ASR任务(CER字错误率指标)

Model	AISHELL2 公开评测集	WenetSpeech 公开评测集	KeSpeech 公开评测集
BR-Voice	2.30%	4.80%	3.92%

副语言信息理解任务(准确率指标)

Model	性别识别 (AISHELL3公开评测集)	年龄识别 (StepAudio公开评测集)	情绪识别 (ESD 公开评测集)
BR-Voice	99.28%	95.86%	90.29%

BR-Voice 技术优势

响应速度快：端到端响应可达500ms

语音感知能力：支持语音识别、性别识别、年龄识别、情绪识别、场景识别等

全双工交互：交互过程中支持用户语音打断、对话轮次的检测判断

高表现力语音生成：可选择多种音色、支持音色克隆、表现力的控制(如情感、风格等)

工具调用能力：支持工具调用能力，如：网页搜索、知识检索等

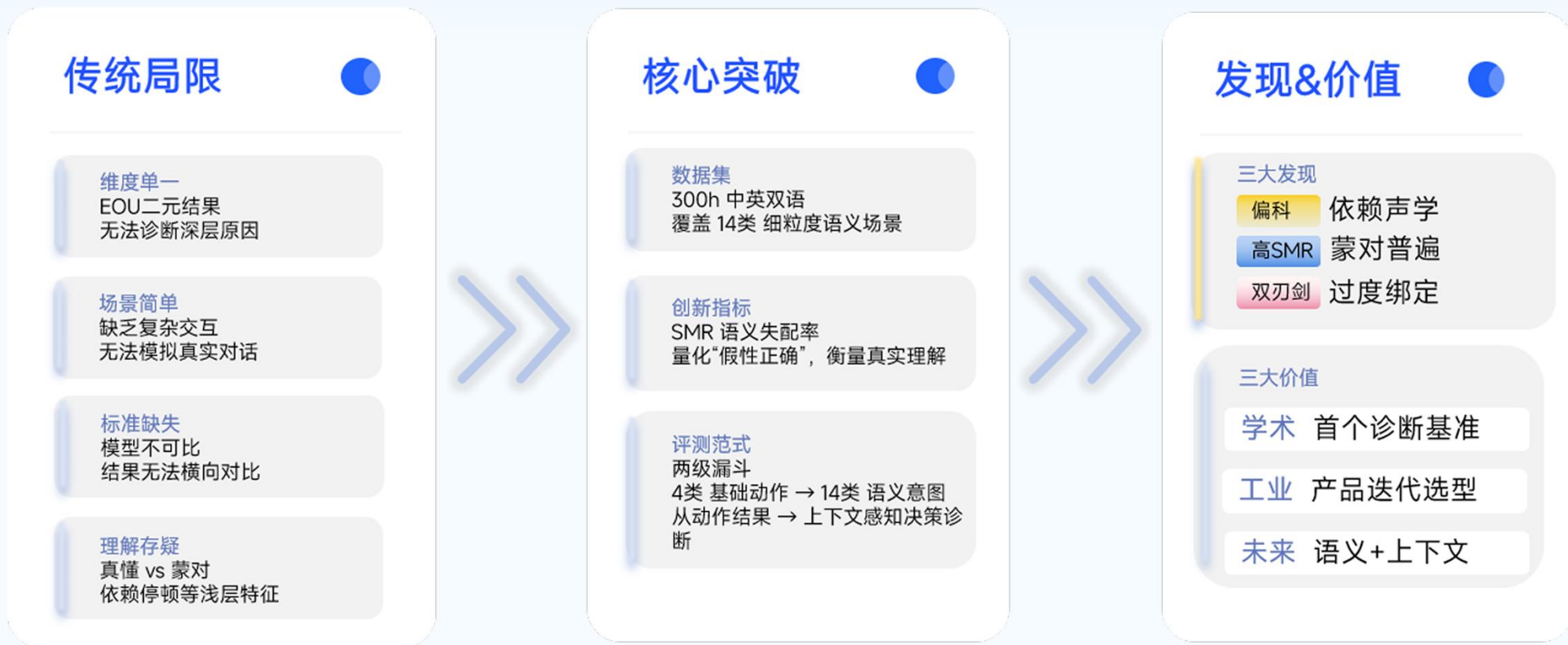
01 语音多模态理解大模型：端到端模型，比传统3段式模型更快更准



01 百融定义全双工语音交互评测标准

BENCH-CODETT收录于INTERSPEECH 2026主会，建立行业首个诊断式评测基准。

传统语音评测只看“用户声音停没停”（二元EOU），模型只要学会抓停顿、语气就能“蒙对”，不代表它真的听懂了对话；BENCH-CODETT建立行业首个诊断式评测基准，引入SMR（语义失配率）创新指标，不再只看“动作对不对”，而是看“理解真不真”。



核心使命：推动人机交互从“声学驱动”向“语义+上下文联合驱动”演进

适用范围：顾客联络服务-人工与智能客户服务协同

- ✓ 呼入任务分配
- ✓ 呼出任务分配
- ✓ 客服智能切换
- ✓ 协同服务改进

牵头单位：中国标准化研究院
起草单位：百融云创，30+ 家
行业标杆企业
主要起草人：百融AICC业务线
负责人...

『填补行业空白 树立人机协同服务统一标尺』



02 主动引导式对话大模型BR-LLM-Proactive聚焦引导交互进入下个营销环节

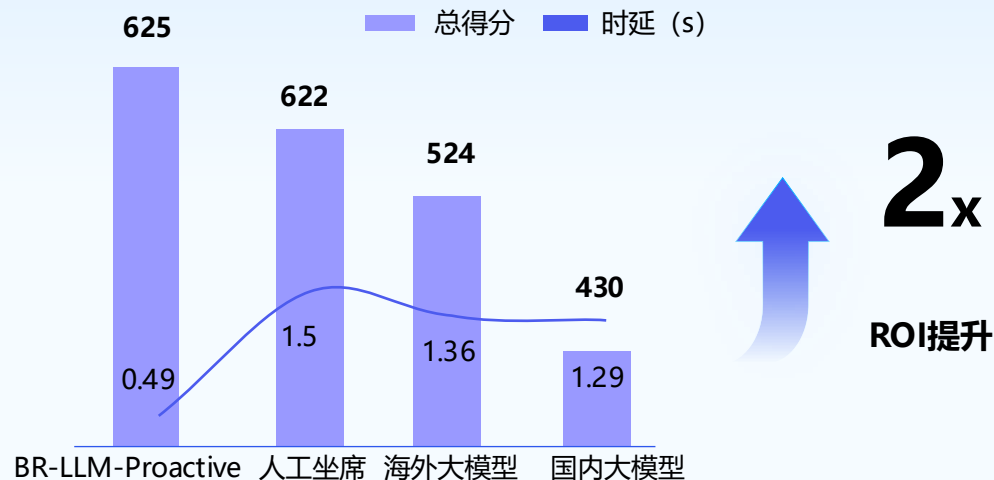
从“被动响应”到“主动引导”的范式升级



被动对话
错失良机



主动对话
引导客户



注：评测脚本来源于2000组对话，由营销业务专家按 0-2 分对其进行评分。

- 【海量规模】真实场景抽取：基于海量线上对话，精准锚定成交关键节点
- 【百万沉淀】对标金牌模范：深度学习百万级金牌销售风格，打造标准化话术模板
- 【千人团队】专业筛选迭代：汇聚1000+业务专家智慧，持续优化并筛选最优解

02 BR-LLM-Proactive：500通AI语音对话测评，百融分数领先



BR-LLM-ProActive 回复效果得分最高、延时最低、成本最低

模型	BR-LLM-Proactive	海外某大模型1	国内某大模型1	国内某开源大模型1	国内某开源大模型2	人工座席
0分	16	64	27	40	90	33
1分	343	348	338	380	390	312
2分	141	88	135	80	20	155
总分	625	524	608	540	430	622
平均延迟/秒	0.49	1.36	0.59	13.61	1.29	~1.5
百万token开销/元	0.07	35.65	0.8	4	4	-
500通对话成本/元 (~150万token)	0.105	53.475	1.2	6	6	1900

1. 评测对象是6组电话座席员工。前5组由不同的大模型支撑的AI硅基员工组成，最后一组由随机抽取的碳基电话员工组成；
2. 随机抽取了某企业3000位真实客户，然后将其随机分配给以上6组员工，每组员工负责与500位各不相同的客户进行电话交流，形成3000通电话录音；
3. 选取一组资深碳基电话座席人员组成评审组，请他们对5组通话录音进行效果评分（0分代表答案不好，1分代表答案普通，2分代表答案优秀），然后对每组通话对应的500个得分进行加总，从而算出每组通话的总分；
4. 百融AI硅基员工座席的总分最高，碳基员工座席的总分次高。但若只看最高分（2分），人工座席最多；
5. 此测试说明：AI的能力距离最优秀的碳基员工还有差距，但AI可以大规模复制出中上水平的硅基员工。

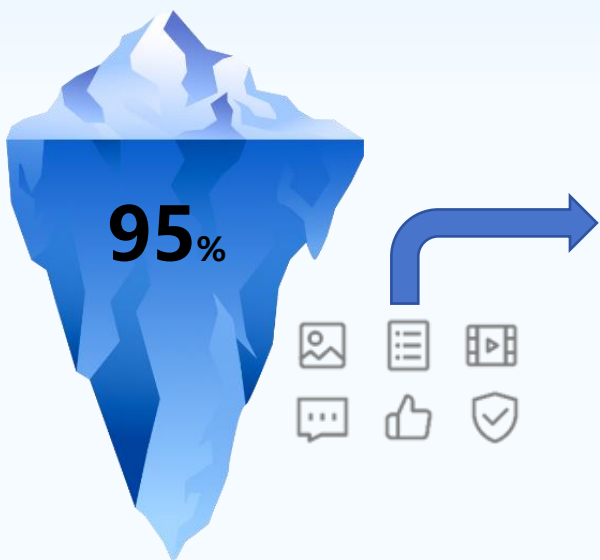
02 BR-LLM-Proactive：某联客服客户自测数据



在2026年6月某联客服POC测试反馈中，由客户自行筛选几百通多轮对话真实呼入业务数据，并自行完成模型效果评测
百融BR-LLM-Proactive表现远超客户预期

指标名称	客户目标值	BR-LLM-Proactive测试值
严重错误率	0.00%	0.00%
幻觉率	≤3.00%	0.85%
问题解决率	≥80.00%	94.00%
知识回答准确率	≥95.00%	98.29%
交互自然度达标率	≥95.00%	98.00%
情感识别成功率	≥90.00%	100.00%
多轮对话保持率	≥80.00%	96.00%

03 文档理解多模态大模型BR-Vision-Doc：聚焦多语种、多版式文档解析



落地效果：真实业务场景深度优化

核心技术：应用“视觉压缩+理解”联合优化及领域自适应技术

性能提升：解析准确率 **>98%**（特定场景），解析速度提升 **80%**

标注体系：硅碳协同高效标注

人机协作模式：采用“AI预标注+专家审核”机制，专家定义文档结构与业务语义标签

标注效能倍增：效率提升 **65%**，一致性（IAA）从0.75上升至达成 **0.92**

数据底座：海量高质量训练样本

多模态全覆盖：深挖合同、报告、扫描件等高价值文档

亿级规模支撑：累计 **>1.2 亿页** 真实业务文档样本

03 BR-Vision-Doc：复杂文档解析测评



复杂文档上解析准确率对比

模型	BR-Vision-Doc	国内大模型	海外模型1	海外模型2
扫描件	99.86%	98.78%	99.3%	89.15%
翻拍件	97.64%	85.97%	90.11%	76.31%

注：评测集由真实业务场景中的银行流水表、财务报表、医学检查单、简历等构成，包括扫描件和照片翻拍。



百融智能
resultscloud.com

补充信息二：

全球领先AICC领域企业参考



01 AI语音基础模型 - ElevenLabs

我们判断：语音领域多维度的SOTA是其业务持续落地增长的先发条件。

\$110 亿

最新估值 ^[1]

2026.02 Series D · Sequoia 领投

\$5 亿+

最新 ARR ^[2]

2026.04, 2025 年底 \$3.5 亿

~22×

ARR 估值倍数 ^[1,2]

\$110 亿 / \$5 亿 ARR

语音基础模型 SOTA 四大锚点

1

STT 全球第一 ^[3]

Scribe v2 在独立第三方 AA-WER 基准以 2.3% 误字率夺得第一，胜 Google Gemini 3 Pro (2.9%)、OpenAI Whisper Large V3、Mistral Voxtral Small

2

TTS 多语种覆盖最广 ^[4]

Eleven v3 支持 70+ 语言，业内最广；Audio Tags 情感标记、Text-to-Dialogue 多人对话、复杂文本错误率下降 68%

3

超低延迟实时 ^[5]

Flash v2.5 TTS 模型推理延迟 ~75ms；Scribe v2 Realtime STT 延迟 <150ms，支撑实时 Voice Agent 场景

4

学术界认可 SOTA ^[6]

arXiv、PMC 多篇论文将 ElevenLabs 作为语音克隆与合成的 SOTA 基准对比参考

Source: ^[1] TechCrunch (Series D, 2026.02) · ^[2] ElevenLabs 官方 blog (\$500M ARR, 2026.05) · ^[3] Artificial Analysis AA-WER v2.0 基准 · ^[4] ElevenLabs Magazine - Eleven v3 指南 · ^[5] ElevenLabs Docs - Latency · ^[6] arXiv 2507.10985 / PMC PMC12459763 学术论文

02 按结果付费模式的企业级语音Agent平台 - Sierra

我们判断：按结果付费模式已被高速增长 ARR 验证期商业可行性，资本市场也越来越认可此模式。

\$158 亿

最新估值 [1]

2026.05 Series E · Tiger Global / GV 领投

\$2 亿

最新 ARR [2]

2026.05 Sacra 估算 / 公司披露

8 个季度

从 0 到 \$150M ARR [4]

2024.02 launch → 2026.02 达 \$150M

业务模式（按结果付费）

- "完成才付费"：[5]

AI Agent 解决问题才付费，升级到人工免费，成功解决的 conversation 为主要计价单位，与客户 ROI 强绑定

- 持续增长模式：

SaaS 式年度合同 + Usage 浮动叠加结构

- 混合计价结构：[2]

简单交互按量、复杂交互按已解决，年度合同含最低使用量保证

场景落地

- 客服与挽留：[6]

电商、订阅、金融等场景的工单处理、流失挽回、续订转化

- 一线企业客户：[4]

40% 财富 50 强已签约，含 SoFi、Brex、Cigna、Rivian、SiriusXM、ADT 等

- 高自动解决率：[6]

部署成熟客户实现 70-90% 全自动解决率，无需人工介入

高速增长验证按结果付费模式可行性

- 极速放量：[4]

7 个季度达 \$100M ARR，再 2 季度翻至 \$200M；

- 估值跳跃：[1]

8 个月估值 \$100 亿 → \$158 亿，+58%；累计融资约 \$16 亿

- 105× ARR 估值倍数：[3]

Sierra 的 105× ARR 倍数为"按结果付费"模式建立独立资本估值范式

03 AI智能客服领跑者 - Decagon

我们判断：~18 个月达 8 位数 ARR，验证 AI 智能客服赛道的庞大需求空间。

\$45 亿

最新估值 [1]

2026.01 Series D · Coatue / Index 领投

~\$3500 万

估算 ARR [2]

2025.10, Sacra 第三方估算

18 个月

0 → 8 位数 ARR (2024年) [3]

Decagon 官方披露增速里程碑

业务模式

• 持续增长模式：

SaaS 式年度合同 + Usage 浮动叠加结构

• 深度集成交付： [5]

Forward-Deployed Engineer 模式，为每个企业定制业务流与系统集成

• 按工单计费： [4]

可选 per-conversation 或 per-resolution，无公开费率，企业级议价

场景落地

• AI Concierge： [3]

全渠道 AI 客服 Agent，SaaS、教育、出行、订阅、零售等高并发场景，覆盖语音、对话、邮件、SMS 多触点

• 一线企业客户： [1]

Notion、Duolingo、Rippling、Hertz、Eventbrite、Bilt 等一线企业

• 深度优化效果： [6]

ClassPass 6 个月客服成本下降 95%、Duolingo 80% 工单 deflection

高速增长与资本市场加持

• 极速放量： [2]

2024 年底 \$10M ARR → 2025.10 达 \$35M ARR，~3.5× YoY 增长

• 客户基数翻倍： [2]

12 个月客户数 4× 增长，2025 年新增 100+ 企业客户

• 一线 VC 加持： [1]

Coatue、Index 领投，a16z、Accel、Bain Capital、Ribbit 跟投



百融智能
resultscloud.com

感谢您的聆听!

