

ZENAVA

2025年企业级AI客服市场深度研究报告：

从争夺软件市场 到争夺劳动力市场的范式转移

天润融通AI研究团队出品

目录

CONTENTS

PART. 01

行业拐点：智能体时代全面到来

PART. 02

游戏规则重写：2025CX市场的技术与商业底层逻辑

PART. 03

案例研究：Sierra——狂飙式增长背后的“Agent OS帝国”

PART. 04

案例研究：Decagon——唯一敢说自己能“管住AI”的公司

PART. 05

案例研究：ASAPP——老牌独角兽的焦虑与破局

PART. 06

案例研究：Cognigy——为什么它值得NICE花10亿美金？

PART. 07

四大阵营对决：技术、商业、估值的三线竞争

PART. 08

下一轮竞争不是技术，是“治理、资本、商业模式”

PART. 09

真正的战场：全球劳动力的重构

PART. 01

行业拐点： 智能体时代全面到来

2025年已成为全球客户服务（Customer Experience, CX）行业由“数字化运营”迈向“智能化运营”的关键拐点。

根据我们对ASAPP、Sierra、Decagon与Cognigy四家行业标杆企业的深度追踪研究，具有代表性的技术供应商的持续追踪与对比研究，确认行业已完成从基于规则的聊天机器人（Chatbots）向具备自主推理与任务执行能力的生成式智能体（Agentic AI）的范式转移。这一转变不仅重塑了全球CX行业的技术架构，更彻底颠覆了企业软件的商业模式、估值逻辑与竞争格局。

从市场表现来看，行业已呈现出明显的分化趋势：

一方面，以Sierra为代表的新兴独角兽凭借“结果导向定价”和高自动化率实现爆发式的收入增长，在不到两年的时间内将估值推高至100亿美元，重新定义了SaaS企业的增长极限。

另一方面，传统AI厂商如ASAPP在二级市场面临估值压力，开始通过强化“人机协同（Human-in-the-loop）”架构寻求第二增长曲线。

与此同时，Decagon通过“智能体操作程序（AOP）”成功解决了企业对大模型不可控的恐惧，实现了900%的年均营收增长。而Cognigy则被巨头NICE以9.55亿美元收购，标志着行业整合大幕的开启，验证了成熟技术资产在战略并购中的确定性价值。

在这样的背景下，本报告将从战略定位、业务绩效、产品路线、技术经济学（包括Token成本结构）、资本路径等维度，对四家企业展开系统评估。目标是为企业管理者、投资机构与技术决策者提供对全球CX智能体市场发展方向的全面洞察与方法论框架。



PART. 02

游戏规则重写： 2025CX市场的技术 与商业底层逻辑

在深入分析具体公司之前，有必要首先建立理解2025年CX市场的基础框架。

随着生成式AI的快速成熟，行业关注点已从传统的自然语言理解（NLU）精度，转向更能体现业务价值的核心指标——包括“问题解决率（Resolution Rate）”、“端到端自动化能力（End-to-End Automation）”以及“基于Token的单位经济模型（Token-Based Unit Economics）”。

这种转变源于企业需求的结构性变化。过去的Chatbot多用于分担人工压力，评价体系围绕意图识别、命中率等技术指标展开；而在智能体时代，企业更关心AI是否能够真正闭环完成任务，减少人工接入，以及是否能够在多渠道、复杂流程中稳定实现自动化。

与此同时，Token成本快速成为管理层的重要决策变量。随着AI调用强度增长，模型选择、对话深度、RAG检索与工具调用所带来的Token消耗，直接决定了供应商和企业双方的单位经济结构，因此“成本可控性”与“规模化能力”成为新的竞争门槛。

基于上述变化，评估CX技术供应商已不再依赖单一技术指标，而必须从“解决效率、自动化深度、Token经济性”三条主线来审视其产品实力、运营能力与商业模式。本报告的后续分析也将围绕这一框架展开。

商业模式的代际更替： 从“席位（Seat）”到“结果（Outcome）”

传统客服软件（如Salesforce、Zendesk）长期以来依赖“按席位/按月（Per Seat/Month）”的订阅模式，其收入结构与企业的坐席规模高度绑定。本质上，这类软件被设计为“提升人工效率”的工具，价值评价体系围绕人力成本与生产率展开。

然而，在生成式AI驱动的智能体时代，AI不再是“辅助人工”，而是直接承担并替代部分人工任务。因此，商业模式也随之从“卖工具”转向“卖结果”。2025年的领先供应商普遍采用基于结果的定价（Outcome-Based Pricing），即按AI成功完成的工单数量计费。



其机制十分直接：企业仅为AI全流程解决的任务付费；若AI无法完成并转接人工，则该次交互不计费。这种“按成果付费”的模式本质上构成了一种新的商业契约，将交付风险从企业转移至供应商。

这种转变对行业的影响十分深远：Outcome模式显著降低了企业的采用成本与试错成本，但却对供应商的技术能力和单位经济提出了更高要求：必须通过高分流率（Deflection Rate）、高解决率（Resolution Rate）以及高效的Token成本控制，才能维持毛利空间和可持续的增长结构。

Sierra正是这一模式最具代表性的实践者，通过激进推进Outcome模式，重塑了企业级CX领域的商业预期。

核心技术指标基准 (2025行业标准)

下表总结了当前市场中领军企业的关键性能指标，作为评估四家目标公司的标尺：

关键指标 (KPI)	传统聊天机器人 (2020-2023)	生成式智能体 (2025)	领军者目标 (Sierra/Decagon)	商业影响分析
自动化分流率 (Deflection)	20%~30%	50%~70%	> 80%~90%	直接决定企业ROI；每提升10%可节省数百万美元人力成本。
客户满意度 (CSAT)	3.5/5.0	4.2/5.0	> 4.7 / 5.0	AI不再是“更差”的服务，而是“更好、更快”的服务。
上线周期 (Time-to-Value)	6~9个月	1~3个月	2~4周	极速部署成为核心竞争力，Decagon与Sierra均强调数周内上线。
单次交互 Token消耗	N/A (规则库)	1k~3k Tokens	动态优化	决定供应商的毛利率结构；需平衡模型智力与推理成本。

PART. 03

案例研究：

Sierra

——狂飙式增长背后的
“Agent OS帝国”

Can you help me track my order?





Sierra是2024–2025年周期中最具代表性的企业级AI新势力之一。公司由前Salesforce联席CEO、现任OpenAI董事会主席Bret Taylor与前Google AR/VR负责人Clay Bavor共同创立，具备强大的产品基因与行业资源。然而，真正使Sierra脱颖而出的并非创始团队本身，而是其在商业模式、产品形态与增长速度上的连续性突破，不仅快速确立了自己的市场地位，更证明了其作为“企业级AI操作系统”的巨大潜力。

公司定位：Agent OS与“AI即品牌”

Sierra将自身定位为企业级“代理操作系统（Agent OS）”，而非传统意义上的SaaS工具。其核心理念在于“AI即品牌（AI is the Brand）”：在客户与企业的主要接触点逐渐由人工转向智能体的背景下，AI的表现直接影响企业的品牌感知与服务体验。

在差异化战略上，Sierra并不以“降低成本”作为主要价值诉求，而是强调智能体在品牌体验中的战略角色。因此，其产品在高智能推理能力之外，更强调同理心表达（Empathy）、语气一致性和品牌风格对齐（Brand Alignment）。这一定位使Sierra能够服务WeightWatchers、Sonos、SiriusXM等对品牌体验要求较高的大型B2C企业。

在技术架构方面，Agent OS作为一个全栈平台，提供从低代码构建、实时调试到全渠道部署的一体化能力。平台能够与企业复杂的后端系统（如订单管理、CRM、计费系统）深度集成，使智能体不仅能够进行对话式交互，还能执行“升级订阅”“发起退货”“修改订单”等可验证的业务操作，从而实现真正的端到端任务闭环。



销售业绩与爆发式增长

Sierra的增长轨迹在企业服务领域极为罕见，其扩张速度甚至超过早期的Slack和Wiz，被视为新一代企业级AI厂商的增长样本。

根据TechCrunch与BEAMSTART报道，Sierra在2025年下半年即突破1亿美元的年度经常性收入（ARR），距离2024年2月正式发布仅7个季度，刷新了企业级软件的增长速度记录。

Sierra也并非仅服务于初创公司。数据显示，其客户中有15%的公司年收入超过100亿美元，50%的公司年收入超过10亿美元，显示其已成功进入预算充足、评估严格的 Fortune 500 客户群。Sierra的产品能力、合规性与可扩展性已在大型企业环境中得到验证。

典型客户案例亦体现了其跨行业适用性与可量化价值：

- **SiriusXM**: 服务3400万订阅用户，通过智能体提升客户留存与服务质量；
- **Sonos**: 覆盖1500万用户，显著缩短技术故障诊断的处理时长；
- **Casper**: 智能体达到74%的问题解决率，同时推动CSAT提升超过20%。

这些数据共同说明，Sierra已在消费品牌、数字内容、智能硬件等多个垂直领域验证其端到端自动化能力和商业价值。

估值与融资：百亿美金的资本神话

Sierra的融资轨迹清晰反映了资本市场对领先AI基础设施型企业的强烈偏好，也体现了其在技术路线、商业模式与增长潜力上的高度确定性。

融资概览如下：

- **2024年初 (Seed / Series A)**

完成1.1亿美元融资，投后估值约10亿美元，投资方包括红杉资本（Sequoia）与Benchmark等顶级机构；

- **2024年10月 (Series B)**

获得1.75亿美元融资，估值提升至45亿美元，由Greenoaks领投；

- **2025年9月 (Series C/F)**

再获3.5亿美元融资，估值达到100亿美元，由Greenoaks继续领投，Iconiq与Thrive Capital跟投。

从估值逻辑来看，Sierra在突破1亿美元ARR的同时获得100亿美元估值，对应约100倍的P/S（市销率）水平。如此高的定价不仅基于其当前的收入增长与技术领先地位，更反映了资本市场对其未来渗透企业服务链路、重塑万亿美元级BPO（业务流程外包）市场的长期预期。投资者普遍认为，随着智能体成为用户与企业交互的主要入口，Sierra有望成为企业与消费者交互的唯一入口（The Interface）。

Token消耗与单位经济模型分析

由于Sierra采用激进的“基于结果定价”，其单位经济高度依赖对Token消耗的精细化管理，推理成本成为影响毛利率的关键变量。

从典型会话成本结构来看，根据Databricks等技术基准，一个中等复杂度的客服场景（如退换货流程）通常需要15–20轮交互：

- **输入侧Token：**包含RAG检索结果、对话历史、用户属性等上下文，每轮约需2k–4k Tokens；
- **输出侧Token：**AI回复约150–300 Tokens；
- **整体消耗：**单次完整解决（Resolution）会话的总Token需求约为30k–50k Tokens。

在主流模型（如GPT-4o）下，这对应约\$0.30–\$0.50的推理成本。若按照Sierra的市场定价（每次成功解决约\$2–\$5）测算，其毛利空间仍具吸引力。但该模型也带来显著风险：一旦AI未能在有限轮次内解决问题并最终转接人工，Sierra将承担全部Token成本且无法获得收入，分流率与解决率直接决定盈利能力。

为应对这一压力，Sierra极有可能采用了“模型路由（Model Routing）”架构，即对于简单的问候和意图识别使用低成本模型（如Llama 3 8B），仅在需要复杂推理（Reasoning）时调用昂贵的SOTA模型。此外，Sierra强调其“减少幻觉”的工程能力，也暗示了其在推理阶段可能使用了多次验证（Self-Correction）机制，这虽会增加Token消耗，但能提升一次性交付成功率，对Outcome模式而言是必要的投入。

而通过上述机制，Sierra在“结果付费”模式下实现了技术成本、交付确定性与商业回报之间的动态平衡。

PART. 04

案例研究：

Decagon

——唯一敢说自己
能“管住AI”的公司



如果说Sierra依托顶级资源与品牌势能“高举高打”，那么Decagon的崛起路线则完全不同——其核心竞争力来自极致的产品工程能力与强执行力，呈现出典型的“以工程驱动的降维打击”。

这家公司在2024-2025年间迅速进入行业视野，依靠对企业级AI落地痛点的精准把握，特别是在“可控性”“一致性”与“可维护性”等方面的突破，成功解决了企业在大模型应用中最普遍、最深层的担忧：**生成式AI不可控性**。

Decagon的产品设计哲学明显偏向工程化与确定性，其提出的AOP（Agent Operating Procedures）框架，将传统依赖Prompt的“概率型”智能体行为转化为“可编排、可测试、可追踪”的逻辑结构，让企业能够以SOP（标准作业程序）的方式管理AI行为，从而显著降低风险与维护成本。基于这一差异化能力，Decagon在互联网企业、SaaS公司以及对流程合规性要求极高的行业中快速打开局面，成为该细分领域增长速度最快的公司之一。

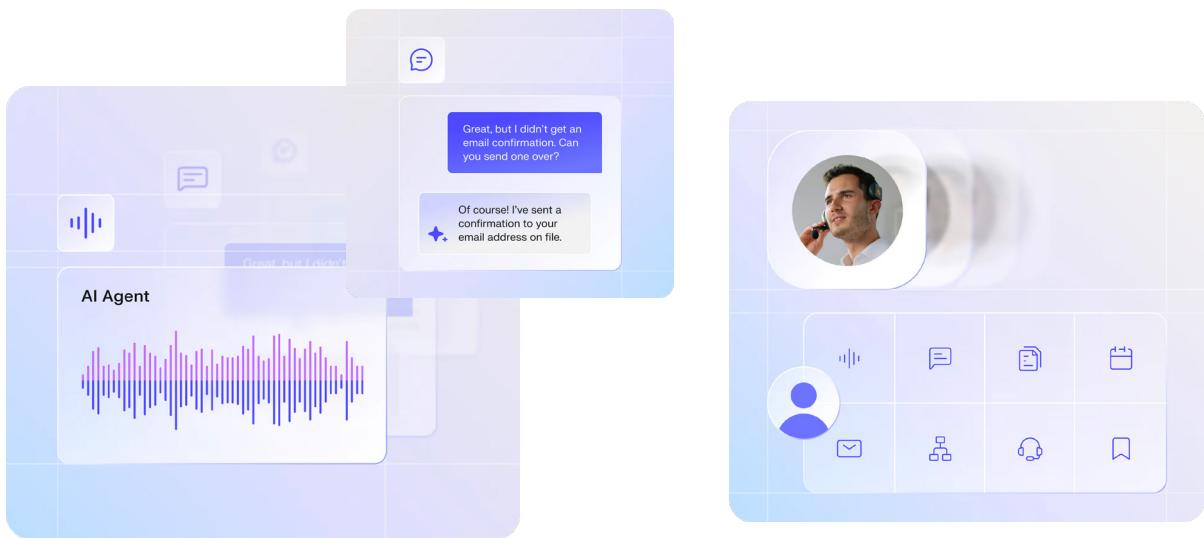
公司定位：AOP与确定性逻辑

Decagon的核心创新是其提出的“智能体操作程序（AOP）”框架。

在企业级应用场景中，客户不仅期望AI“足够聪明”，更希望其行为“可控、可预测、可复现”。传统的Prompt Engineering（提示工程）由于依赖概率型输出，往往存在脆弱性强、难以版本管理、跨团队维护成本高等问题，无法满足大型企业对稳定性和合规性的要求。

AOP的设计思想借鉴了传统客服中的SOP。企业可以使用自然语言定义业务规则（如：“若为VIP客户且退款金额小于50美元，则直接批准”），系统会将这些描述自动解析并“编译”为结构化、可测试、可追踪的执行逻辑。

换言之，AOP将“自然语言的灵活性+代码的精确性”结合在同一套框架中，使Decagon得以向企业提供更低的幻觉风险、更高的合规性与更强的可维护性，特别契合Notion、Rippling等对工程严谨性要求高、对流程可控性敏感的科技企业需求。



销售业绩与市场渗透

Decagon的增长效率极高，在短时间内完成了从互联网原生企业向大型传统企业的双向渗透，其商业化进展显著快于同周期的其他新创智能体公司。

根据Sacra数据，Decagon的ARR从2023年底的约60万美元提升至2024年底的约600万美元，同比增长约900%。至2025年中期，官方确认ARR已迈入千万美元级，并正向5000万美元规模逼近，增长曲线体现出强烈的产品拉动效应。

Decagon在科技与互联网领域获得了广泛采用，包括Notion、Duolingo、Bilt、Substack、Eventbrite等高速成长的独角兽企业。与此同时，2025年公司宣布与全球最大电信运营商之一 Deutsche Telekom（德国电信）达成商业试点并获得战略投资，这标志着其产品已具备支持电信级业务复杂度和稳定性要求的能力。

Decagon报告其智能体系统在多数客户场景中可实现接近70%的分流率（Deflection Rate），部分客户如Duolingo的分流率甚至超过80%。在成本端，ClassPass报告通过Decagon的系统将支持对话成本削减约95%，进一步验证了其在单位经济方面的显著优势。

融资与估值：一年三级跳

Decagon的融资节奏显著快于同周期AI创业公司，反映出其商业化进展与产品落地速度均超出投资机构预期。

融资轨迹如下：

- **2024年6月 (Series A)**

完成3500万美元融资，由Accel与a16z共同领投；

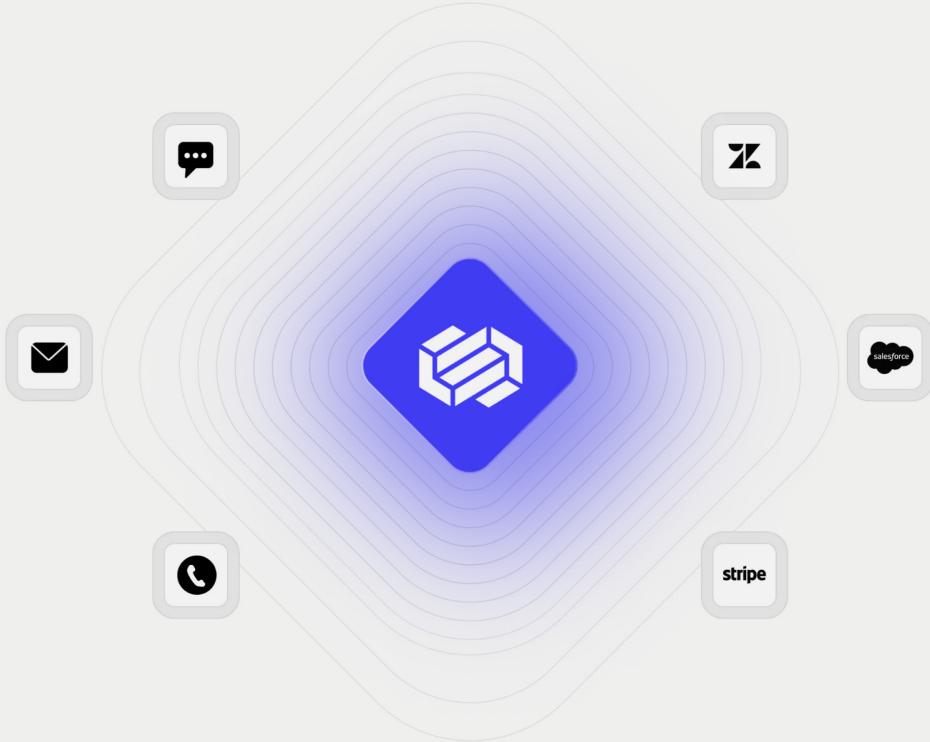
- **2024年10月 (Series B)**

获得6500万美元融资，投后估值约6.5亿美元，由Bain Capital Ventures领投；

- **2025年6月 (Series C)**

再获1.31亿美元融资，估值提升至约15亿美元，领投方再次为Accel和a16z。

市场预期方面，2025年9月业内消息显示，Decagon已在筹备新一轮融资，目标估值区间为40-50亿美元，以进一步强化其在与Sierra、Salesforce等大型竞争者对抗中的资源储备。若融资顺利落地，意味着Decagon的估值将在12个月内实现接近10倍的增长，成为该轮智能体创业公司中估值增长最快的企业之一。



技术架构与消息吞吐量

规模效应方面，Decagon的平台能够支撑数千万级终端用户的并发访问，其系统架构在设计之初便围绕高并发处理能力与稳定性进行了优化，具备支撑大型互联网与电信企业需求的工程基础。

技术细节上，不同于ASAPP强调在人机协同环节保留人工介入，Decagon更依赖AOP业务规则进行结构化定义，通过“预设逻辑”替代大量实时推理。

这一机制导致平台在初始化阶段需要消耗更多 Token，用于将自然语言规则解析、编译为可执行的向量化或图谱化结构；但在实际运行过程中，由于多数决策路径已具备确定性，系统可以避免在每次对话中重复触发高成本的思维链（Chain-of-Thought）推理，从而整体提升推理效率并优化单位经济性。

PART. 05

案例研究：

ASAPP

——老牌独角兽的焦虑与破局





ASAPP成立于2014年，是上一代企业级AI客服技术的代表性厂商之一，以实时语音识别、情感分析与座席辅助（Agent Assist）能力闻名，在航空、通信等高价值行业深耕多年。凭借稳定的交付能力、深度定制化能力以及长期积累的行业数据，ASAPP在早期智能客服市场中建立了稳固的领先地位。

然而，随着生成式AI的快速普及与智能体（Agentic AI）架构的崛起，ASAPP原有以模型微调、自动化建议与座席增强为核心的技术路线面临结构性挑战。为了应对这一代际变化，公司正在推进一场从“增强人工效率”向“部分替代人工执行”的战略转型：既强化自研大模型与新一代智能体能力，也继续保持其在语音渠道、人机协同（Human-in-the-loop）与合规性管理方面的传统优势。

对ASAPP来说，这既是一段竞争压力显著上升的时期，也是一段通过技术升级与组织调整寻求第二增长曲线的关键阶段。

公司定位：GenerativeAgent与Human-in-the-loop

ASAPP将自身定位为提供“AI原生（AI Native）”客户体验的平台，其与新兴智能体公司的最大差异在于始终强调“人”的价值与作用。

ASAPP并不追求完全替代人工坐席，而是构建了成熟的HILA体系：当系统对某一判断或执行动作的置信度不足时，会自动通过坐席界面请求人工介入。人工的处理结果不仅用于即时解决问题，还会被自动记录并沉淀为高质量标注数据，用于后续模型优化。这种“实时使用、实时学习（Learn-while-Serving）”的机制，是ASAPP在十年行业实践中形成的核心竞争壁垒之一。

在渠道能力方面，ASAPP具有明显的“语音优先（Voice-First）”优势。由于其长期深耕航空、通信等高语音占比行业，公司在实时语音转录、情感分析、语音理解与语音合成等方向形成了深厚的技术积累，使其在处理电话渠道（Voice Channel）等高复杂度场景中具备更高的稳定性与成熟度。

估值倒挂与二级市场压力

ASAPP的财务表现呈现出典型的“AI转型期”企业特征：收入保持稳健增长，但估值水平承受一定压力。

估值变化方面，公司在2021年5月完成Series C融资时的投后估值约为16亿美元。然而，截至2025年，根据二级市场平台（PM Insights、Notice.co）的交易数据，公司股权的隐含估值区间约为11.7亿至12.3亿美元，相较高点出现约25%–58%的折价。估值回调在一定程度上反映了市场对老牌独角兽增长动能减弱的担忧。

从市场定价逻辑来看，与Sierra等新兴智能体厂商动辄50–100倍P/S的估值形成对比，ASAPP更多被投资者视为偏传统的软件资产，估值倍数亦更趋向成熟SaaS企业的水平。

收入表现方面，ASAPP在收入规模上仍保持较强竞争力：2024年收入预计在7500万至1.2亿美元之间；2025年预测收入约为1.8亿美元。尽管绝对规模不小，但其约50%的同比增长率明显低于Sierra、Decagon等处于高速扩张周期的新兴智能体企业。

内部动荡与战略调整

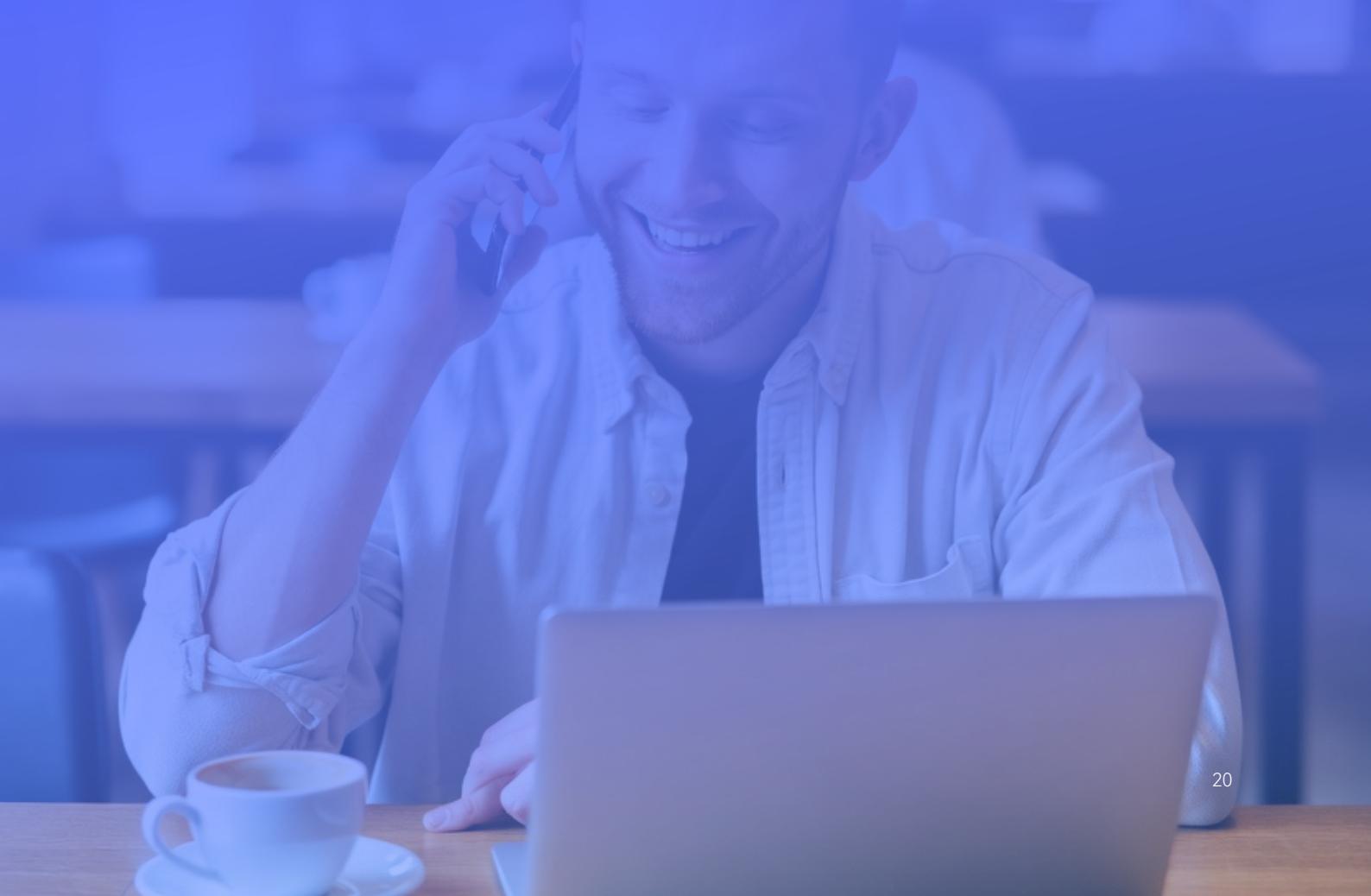
2025年9月ASAPP任命Priya Vijayarajendran出任新任CEO，这一举措通常意味着公司从创始人领衔的早期阶段进入由职业经理人主导的成熟运营阶段。此类调整往往旨在强化组织效率、提升财务纪律，并为潜在的资本路径（如上市、并购或战略合作）做好准备。

市场信息显示，ASAPP在2024–2025年间进行了多轮人员优化，并同步收紧薪酬政策（包括取消年度普调）。这些措施旨在改善现金流表现、提升单位经济模型的可持续性。在企业服务软件行业从“增长优先”向“盈利优先”转变的背景下，这类调整已成为行业普遍趋势。

PART. 06

案例研究：

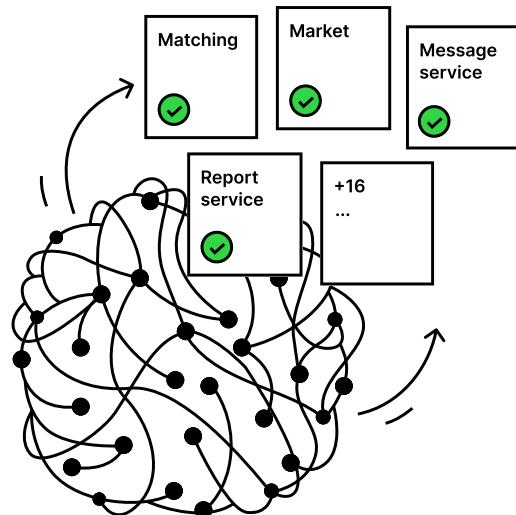
Cognigy ——为什么它值得NICE 花10亿美金？



codigy

Cognigy的发展轨迹代表了本轮企业级AI浪潮中成熟技术公司的另一种典型结局——依托稳健的产品力和行业积累，最终被大型科技巨头以战略并购的方式吸收。作为欧洲领先的企业级对话式AI与流程编排平台，Cognigy在自动化、全渠道集成、合规能力以及全球化交付体系方面具备长期优势，其产品在航空、制造、零售等传统行业中拥有广泛应用。

然而，与依靠高速增长和资本推动崛起的新兴智能体公司不同，Cognigy更像是一家进入“成熟期”的企业技术提供商。其收入结构稳定、客户覆盖度高、技术栈以流程编排和集成能力为核心，而非以智能体推理能力为主。这使其在并购市场中成为巨头补齐产品线、增强AI能力的理想标的，最终被NICE以战略价格收购，也体现了其技术资产和企业客户基础的长期价值。



并购事件：NICE9.55亿美元收购Cognigy

2025年7月28日，全球云联络中心（CCaaS）龙头NICE（NASDAQ:NICE）宣布以战略收购方式正式拿下Cognigy。本次并购总价约为9.55亿美元，包括现金、股票以及约5000万美元的延期支付与保留金。

从战略逻辑来看，NICE旨在通过Cognigy的核心能力——尤其是其在流程编排（Orchestration）、全渠道自动化以及欧洲市场的深度渗透——进一步强化CXone Mpower平台的竞争力，以对抗Genesys、Salesforce等全球竞争者。对于Cognigy而言，这笔交易是一次极为成功的资本退出，不仅为其长期投资方（如Insight Partners、Eurazeo）带来可观回报，也标志着其技术路线和企业级能力在全球市场获得了高度认可。

财务数据与估值倍数分析

Cognigy的收购为企业级AI客服市场提供了一个相对理性的估值锚点。在收入规模方面，市场普遍预计Cognigy在2025年的ARR约为6500万美元，并有望在2026年实现约80%的增长，ARR达到1.15亿至1.2亿美元区间。

从估值倍数来看，本次约9.55亿美元的交易对价对应2025年预期收入的估值倍数约为**14.7倍**，对应2026年收入的倍数约为**8倍**。这一估值区间显著低于新兴智能体厂商动辄50-100倍的P/S水平，但处于大型企业在战略并购中常见的倍数区间。

NICE支付的价格反映了其以“战略协同”“产品补全”与“确定性现金流”为核心的定价逻辑，而非对指数级增长空间的押注。这也说明，尽管顶级智能体企业（如Sierra）可以获得超高估值溢价，但对于成熟技术平台而言，市场更倾向于依据收入质量、可预期增长与整合价值进行定价，并非所有AI公司都能够享受类似Sierra的估值倍数。

业务规模与Token吞吐

Cognigy是典型的“重型”企业级自动化平台，强调高稳定性、强流程编排能力以及大型企业场景下的高吞吐量支持。在交互量方面，Cognigy服务的客户规模庞大，例如Lufthansa（汉莎航空）每年需处理超过1600万次对话，Henkel每年处理约500万次对话。平台能够在高并发、多流程、多语言环境中保持稳定运行，凸显其在传统行业中的技术成熟度。

在Token消耗与成本结构方面，Cognigy采用“编排层（Orchestration Layer）”模式，而非直接承担推理成本的“智能体执行层”。其平台允许企业自带模型（Bring Your Own Model, BYOM），因此即便Cognigy负责处理数十亿级别的交互量，底层推理Token的成本仍由企业直接向Azure、OpenAI等云服务支付。Cognigy自身的收入主要来自“流量费”“会话费”或“编排费”等软件层收费。

这种模式的优势在于规避了Token成本波动带来的毛利风险，使其收入结构更加稳定；但相应地，也限制了其通过Token利差获得额外利润的空间，导致其商业模式更接近传统企业软件，而非智能体公司常见的Outcome-Based或Usage-Based利润结构。

PART. 07

四大阵营对决： 技术、商业、估值的三线竞争

为了更直观地展示四家公司的竞争态势，我们从财务、产品、技术和市场四个维度进行了详细对比。

评财务与估值概览 (2025 数据)

维度	SIERRA	Decagon	ASAPP	codigy
成立年份	2023	2023	2014	2016
总部	旧金山	旧金山	纽约	德国
最新估值 (2025)	\$100亿	\$15亿(寻求\$40亿+)	~\$12-16亿 (二级市场)	\$9.55 亿(收购价)
ARR (预估)	>\$1.0亿	~\$3000万-\$5000万	~\$1.2 -1.8亿	~\$6500万
P/S 估值倍数	~100x	~50x-80x	~8x-10x	~14.7x
融资总额	\$6.35亿	\$2.31亿	~\$5亿	~\$1.65亿
状态	爆发期(独立)	成长期(独立)	成熟期(转型中)	被NICE收购



技术架构与产品差异化

维度	ASAPP	SIERRA	Decagon	codigy
核心哲学	Human-in-the-loop (HILA) 人机无缝协作	AI is the Brand 高度拟人化与同理心	AOP (流程即代码) 确定性与可控性	Enterprise Orchestration 流程编排与集成
定价模式	混合/席位费 (Seat-based)	基于结果定价的颠覆性创新	对话量/解决量混合	席位/流量费
集成深度	深度定制 (AWS生态强相关)	Agent OS 全栈低代码平台	Code-Level 开发者友好的代码集成	广泛的预置连接器 (SAP、Salesforce等)
安全策略	物理隔离、PII屏蔽	品牌护栏、实时监察	逻辑锁死通过AOP杜绝不可控路径	企业级合规认证 (GDPR等)
目标客群	航空、电信 (B2C巨头)	消费品牌 (DTC, Retail)	科技、SaaS、互联网	全球大企业 (欧洲制造/航空)



Token 消耗与成本结构深度解析

Token消耗是AI客服公司最重要的可变成本 (COGS)。基于各公司的业务模式，我们构建了以下成本分析模型：

维度	Token消耗模型	成本承担方	风险分析
 SIERRA	极高 (High Volume)每次解决需消耗大量Token用于推理、自我纠错和工具调用。	Sierra承担(包含在“结果付费”中)	风险最高。必须依靠极高的模型路由效率 (用小模型解决简单问题) 来维持利润。如果遇到恶意攻击或极其复杂的用户，成本可能倒挂。
 Decagon	中高(Medium-High)AOP需要预处理规则，但在运行时路径相对确定。	客户/Decagon混合	相对可控。AOP的确定性减少了无效的“思维链”推理，提高了Token利用率 (Token Efficiency)。
 ASAPP	中等 (Medium) HILA 模式下，人类介入会截断AI的无效长对话。	客户/ASAPP混合	成本优化较好。ASAPP拥有大量垂直领域专有小模型，推理成本低于通用大模型。
 codigy	透传(Pass-through)作为编排器，通常仅消耗路由Token。	客户承担	风险最低。Cognigy赚取软件费用，客户直接支付Azure/OpenAI账单，因此对模型价格波动不敏感。

PART. 08

下一轮竞争不是技术， 是“治理、资本、商业模式”

随着智能体在企业级场景中从实验性应用迈向大规模部署，行业竞争将从“模型能力”转向“可控性、生态整合与商业模式”的系统性竞争。

首先，**智能体治理（Agent Governance）** 将成为下一轮竞争的核心战场。Decagon以AOP建立的确定性逻辑体系，使市场进一步意识到：“让AI生成内容并不难，难的是确保其可控、可验证、可追踪。”未来12-24个月，企业将从依赖“黑盒式智能体”，转向选择具备版本控制、回归测试、模拟仿真等工程化治理能力的平台。ASAPP和Decagon在该领域布局较早，有望占据治理标准化浪潮的先发优势。

其次，**行业并购整合将进一步加速。** Cognigy被NICE收购只是序幕。随着Salesforce（Agentforce）、Genesys、Five9等传统巨头在智能体时代面临技术短板，它们对关键技术资产的收购意愿显著上升。拥有差异化技术壁垒的中型厂商——如Sierra的Agent OS、ASAPP多年积累的HILA数据体系——均可能成为潜在收购目标。特别是ASAPP，鉴于其成熟的企业客户基础与估值压力，被Cisco、Avaya、AWS等大型厂商整合的概率正在提升。

最后，“**结果定价（Outcome-Based Pricing）**”将从创新模式迈向行业常态。Sierra以Outcome模式打开市场后，企业迅速感受到“零风险付费”的优势，这将迫使ASAPP、Cognigy等老牌厂商逐步从“卖软件”转向“卖结果”或“卖劳动力”。这一变化将深刻重塑收入结构和估值逻辑，使行业从传统以ARR为核心的SaaS模型，转向以GMV或“成功解决的任务量”为基准的服务型收入模式。

总体来看，智能体时代的竞争将从“技术领先”演变为“可控性+生态整合能力+商业模式创新”的综合竞争，行业格局也将在治理标准确立与并购整合加速的推动下进一步重塑。

PART. 09

真正的战场： 全球劳动力的重构



2025年的AI客服市场呈现出明显的分化格局：新兴力量快速崛起，传统厂商积极转型，而各类技术路线在特定场景中各自占据优势。

对于强调品牌体验与服务创新的企业，Sierra以其“结果付费”模式与高完成率的智能体能力，正在成为下一代CX基础设施的代表。对于注重工程严谨性、流程可控性与可维护性的技术型企业，Decagon的AOP架构提供了介于灵活性与确定性之间的最佳组合。

在流程高度复杂、受合规约束较强且仍需保留人机深度协同的行业（如航空、银行、电信），ASAPP依旧是最稳健的合作伙伴，其HILA模式在可控性与效率之间实现了长期验证的平衡。而对于寻求全球化交付能力、全栈集成与长期稳定性的跨国集团而言，并入NICE后的Cognigy则成为最具有安全性和可持续性的选项。

综合来看，行业竞争的焦点正在从“软件市场份额之争”转向对**全球数万亿美元劳动力市场（Labor Market）重新配置权的竞争**。智能体并非替代某个工具，而是在重塑企业的运营方式与人力结构，决定未来企业服务市场的深层格局。

ZENAVA



官网: www.zenava.ai

邮箱: marketing@zenava.ai

服务支持: 1010-9099