

Agent商业化元年：模型退居二线

2026世界人工智能大会（即2026 WAIC）展馆里的热度分布很有意思。Kimi K3的发布吸引了不少技术从业者驻足月之暗面展台，全场Agent（智能体）展示密度更高，几乎每家展台均布局智能体方案，覆盖办公、企业服务、个人助理、行业解决方案全场景。

2025年WAIC中，智能体大多仅为演示demo，业内讨论集中在智能体与大模型差异、ToC与ToB智能体研发逻辑；时隔一年，大模型持续迭代，智能体正式进入商业化实战检验阶段。IDC最新《DAA 研究报告》数据印证行业转向：全球日活智能体（DAA）由 2025年的2860万个升至2026年的7940万个，2030年将达22.16亿个。该统计逻辑遵循百度创始人李彦宏提出的时代的度量衡指标：更关注每天有多少智能体真正进入业务流程、完成任务并创造价值。

Token调用成本直接决定企业AI开支，行业思考维度进一步细化，围绕智能体商业模式搭建、传统软件在智能体时代的定位等问题，本届WAIC上多家头部企业高管给出落地观点。



IT 系统现状，也能测算AI项目真实投资回报率 ROI。

2026 WAIC期间，安永大中华区人工智能与数据咨询服务主管合伙人陈剑光从企业咨询视角，解读行业争议话题：数字化成熟企业落地AI是否背负历史包袱。

陈剑光分析，认为存在转型包袱的观点有两层顾虑：老旧系统固化传统业务流程，改造投入成本偏高。但他持相反看法，“企业数字化做得比较好，恰恰更能发挥AI价值。因为系统稳定运行多年，数据质量相对比较高，是比较可信的”。

他进一步预判企业服务与用工形态变化：“未来客户有可能是人类，也有可能是个Agent”，企业内部将新增数字员工，业务流程需要重构，但底层基础系统无需替换，ERP（企业资源计划）、CRM（客户关系管理）依旧是 AI 时代核心底座，提供统一真实数据源。

“Agent决胜是谁更懂真实世界”

2026 WAIC上，各大科技巨头推出全品类智能体“全家桶”，依靠资源优势全面铺开赛道，垂直创业公司走出差异化路径。专注全球B2B赛道的百型智能深耕细分产业，坚持以交付完整业务结果为核心目标。

大会现场百型智能发布AI外贸员本体OntoZ，公司CEO韩美表示，率先跑出商业价值的主体，不一定是手握超大模型、充沛资源的巨头，更多是长期扎根垂直产业、深度吃透业务场景的创业公司。企业采购AI的核心诉求并非技术工具，而是可落地业务结果。

近两年市场智能体产品持续扩容，但她向北京商报记者透露，“企业很快发现，拥有越来越多Agent，并不等于拥有越来越多订单，企业真正需要的，不是单点工具，而是一套能够围绕业务结果持续运转的智能系统”。

韩美进一步点明行业普遍短板：“Agent最大的问题并不是不会执行，而是不理解企业。”针对行业竞争格局，她判断，企业选型不会单纯考量厂商规模，核心评判标准是厂商对自身业务的理解深度；行业长期壁垒不来自大模型技术，而是长期沉淀的真实商业场景落地反馈。

北京商报记者 魏蔚/文并摄

“Agent按任务收费难精确定义”

“智能体成败的关键在于场景。”2026 WAIC 期间，腾讯云副总裁吴运声发布ADP 4.0海外版时，点名智能体三大核心要素：场景、能力、知识。接受北京商报记者采访时，他重点谈及行业商业化难题，直言腾讯云不会采用按任务计费的商业化模式，核心原因是业务任务边界难以精准界定。

他进一步解释，稳定运行的智能体成本由调用模型、配套插件决定。针对企业普遍核算Token成本的趋势，吴运声表示Token价格持续下行，并分享可落地控本思路：拆分业务流程，模糊复杂工作交由智能体处理，标准化固定流程使用工作流承载，减少无效Token消耗。在他看来，企业精细化管控算力开销会涌现大量兼顾业务需求、压缩Token成本的方法。

“客户只为业务落地成果付费”

这是阿里云智能集团CTO李飞飞首次以现任身份出席WAIC。他在大会上提出，训练、推理技术最终目标是赋能企业业务、搭建

增长闭环，智能体嵌入业务流程是当前 AI 商业化核心突破口。

落地的前提是底层架构重构。他认为，智能体架构需搭建独立调度体系，实现智能体与底层模型解耦，由智能体自主判断任务类型，匹配适配度最优的模型，避免所有任务统一调用顶配大模型，推高企业无差别算力支出。

“类比SaaS行业付费逻辑，行业会走向智能体即服务模式，面向细分业务打造专属智能体，客户只为业务落地成果付费，而非算力资源消耗。”该模式成立根基在于推理基础设施形成规模效应，持续压低Token调用成本。智能体发展核心，是将算力、模型能力转化为可落地、具备商业价值的业务动作，李飞飞表示。WAIC现场，阿里云推出Agent Native-Cloud（智能体原生的云），同步上线Agent-Teams、AgenticComputer等企业级工具，覆盖基础设施、开发平台、云桌面，让智能体从外部接入工具，转变为企业原生内置能力。

“完整AI系统不只有模型”

PyTorch基金会执行董事Mark Collier自2010年参与OpenStack开源云计算项目，来华已有16年，现阶段工作重心聚焦Py-

Torch生态，基金会核心工作围绕模型训练、推理、智能体三大板块。

今年，智能体行业迎来爆发式转型，发展速度超出行业预期。“过去行业只关注模型生成Token的总量，但现在业内诞生了一个很好的概念——Token产出效率，不是所有生成的Token 都具备同等业务价值。”

“模型固然是核心，是智能体的‘大脑’，但现在行业开始重点搭建智能体调度适配体系，近期我也开始试用Hermes智能体框架。”Mark Collier表示，全球企业逐步摆脱单一模型依赖，或是自研垂直专用小模型，或是基于开源基座，通过强化学习搭配大小双模型适配细分业务。背后驱动力除控制调用成本外，还包含数据安全考量，规避企业核心业务经验、知识产权上传第三方API平台带来的泄露风险。

“大模型懂世界，但不懂你”

过去半年，金山办公CEO章庆元频繁被问及两个核心问题：大模型是否会吞噬传统软件？办公软件在AI时代还有没有活路？他的回答出人意料。

“在AI时代，软件应该会得到新生。但传

统软件不一定能活得下来。”该观点依托“大模型逐步成为基础设施服务”的行业洞察。章庆元梳理近一年三大行业变化：大模型稀缺性快速消退，选型核心标准转向性价比与Token 供给能力；开源产品、智能体批量涌现，行业竞争核心回归软件本体价值。

软件不可替代的价值如何体现？他做了通俗类比：“大模型是个博士生，它懂这个世界，但它不懂你。大模型的每次对话都是一次‘重生’，它没有记忆——所有的记忆和上下文，都存储在软件里。”

因此，“未来只有AI Native（AI 原生）的软件才能生存”，标准为大模型驱动、AI友好设计、结果导向型智能体。近期金山办公上线灵犀专业版与WPS Comate。他将灵犀独立拆分的逻辑类比相机行业迭代：在胶卷相机上改良终究存在局限，真正颠覆赛道的是全新产品；大模型提供通用智力，软件沉淀专属上下文，二者分工协同才是行业正确路径。

“未来企业客户有可能是Agent”

四大会计师事务所之一安永拥有特殊行业视角，每年完成数千家企业财报审计，同步提供数字化转型咨询，既能摸清企业现有

Market focus

AI终端新人胜旧人

手机行业的悲欢并不相通。7月16日，2026 WAIC开幕前夕，realme真我品牌负责人徐起发布致用户公开信，宣布真我将暂停中国市场运营，全部资源转向深耕海外性能、游戏赛道。实际上，这是一次早有预兆的退场——今年3月便有其收缩国内业务的市场传闻，品牌方当时曾出面辟谣，此后商城服务调整、售后并入OPPO体系、组织架构整合等动作接连落地，伏笔早已埋下。但在赛道另一端，是新贵进场、巨头出奇——7月的上海，WAIC如火如荼，AI终端成为其间的主战场。阶跃星辰于7月13日抢先发布号称全球首款的大模型原生智能体手机STEPX Neo，其也成为从大模型赛道跨界切入消费终端的先锋。传统终端厂商阵营中，中兴旗下努比亚在大会期间首发AI智能体手机，主打原生智能体；荣耀带来全球首款机器人手机（Robot Phone）的全功能实操首秀，AI终端战事硝烟渐浓。

存量周期下，传统手机赛道持续出清、硬件增长红利逐步消退、AI原生终端新势力加速集结，AI成为行业破局的核心变量，消费终端全面迭代的拐点未至将至。

遗憾离场

北京商报记者拨通真我官网客服电话，接线方已换成OPPO售后服务：“我是智能客服小O。”

徐起将此次调整称为“按下暂停键”。根据公开信内容，真我未来将把全部资源集中到海外市场，聚焦性能与游戏赛道深耕；针对国内存量用户，真我现有机型的销售、售后质保将全部由OPPO官方团队承接，待下一代ColorOS系统发布后，真我手机用户可直接升级至该系统，用户相关权益将得到完整保障。

战略调整背后，是真我在国内市场的困局。第三方监测数据显示，2026年一季度国内手机市场，OPPO系整体份额为16.7%，其中真我品牌仅贡献1.2%；到2026年第23周，真我国内市场份额进一步滑落至0.7%，逐步走向边缘。

真我官网的品牌介绍，仍停留在2021年的高光。当年9月Counterpoint发布全球

手机市场报告显示，realme出货量位列全球第六，成为全球第四大中国手机品牌。彼时凭借性价比与设计等卖点，真我在全球快速积累超3亿用户，一度成为行业公认黑马。

数年国内市场征程，从异军突起到战略收缩，最终以一场转身暂别本土。

新人胜旧人

细数手机行业，真我并非年内唯一一家官宣收缩退场的品牌。

2月27日，魅族科技发布战略转型公告，宣布暂停国内手机新产品自研硬件项目，整体发展方向从硬件主导转向AI驱动的软件业务，以Flyme开放生态为基座对外赋能，保留海外手机业务、AI眼镜等产品线独立市场化运作。4个多月后，7月15日，效力魅族近20年的中国区CMO万志强官宣离职，投身创业项目。

有人离开，也有人涌入。本届WAIC，AI智能体终端成为手机行业新的主战场，跨界

新贵与传统厂商集结，押注下一代终端。

大模型厂商率先完成布局。7月13日，阶跃星辰抢先发布全球首款大模型原生智能体手机STEPX Neo，搭载智能体原生操作系统Step AOS，内置个人智能体阶跃Amoo，可跨应用自主完成日常生活及办公任务；传统终端厂商也在加速技术落地。中兴旗下努比亚在大会期间首发AI智能体手机NaviX Ultra，主打原生智能体架构，搭载已完成网信办备案的豆包端侧大模型，通过系统级GUI Agent架构实现跨应用自主执行多步骤任务；荣耀则带来全球首款机器人手机Robot Phone，内嵌云台结构并搭载智能体系统。

“市场饱和、同质化严重及成本压力导致的盈利空间收窄，最终使得传统玩家如realme、魅族等选择收缩。而新势力如阶跃星辰、努比亚携AI的进场，则反映出终端价值从终端本身转向端侧的、原生的智能。”苏商银行特约研究员武泽伟对北京商报记者表示，AI智能体手机可能成为差异化突破口，长远来看或为行业催生新的增长点。

行业洗牌

手机行业的压力是客观存在的。

根据Counterpoint Research智能手机市场监测初步统计，受存储芯片供应短缺持续冲击，2026年二季度全球智能手机出货量同比下滑11%，创下2013年以来二季度的最低水平。由于存储供应商将产能优先分配给AI数据中心而非消费电子领域，DRAM与NAND价格在整个季度持续大幅上涨，整机物料成本的攀升多轮传导至消费端，其中对价格敏感度最高的入门级、中端机型受冲击最为显著，也进一步压缩了主打性价比路线品牌的生存空间。

全球市场格局方面，头部效应持续加剧。三星凭借垂直供应链优势与高端产品表现，以24%的市场份额重返全球第一，是头部品牌中同比增速最高的厂商；苹果依托高端机型的稳定需求，出货量同比增长3%，市场份额升至20%的历史新高，同时也是二季度唯一未上调手机售价的主流品

牌。而产品布局更集中于中低端价位段的小米、OPPO、vivo均录得两位数同比跌幅，在成本上涨与需求疲软等多重压力下处境承压。机构预计，2026年全球智能手机全年出货量仍将同比下降约14%，存储短缺问题或将持续至2027年，接下来主流厂商将普遍转向优先保障盈利能力，而非单纯追求出货规模。

行业承压的另一面，是AI终端的合规化进程正在加速落地。近日，网信部门公布新增7款手机端生成式人工智能服务备案信息，包括苹果、华为、OPPO、vivo、小米、三星、努比亚七大终端品牌完成合规备案，AI手机规模化商用落地有了“通行证”。

科方得智库研究负责人张新原对北京商报记者表示，若新品牌能证明“AI重构用户体验”可行，将倒逼传统巨头加速转型；但反过来，若AI仅是营销噱头，行业可能先经历一轮“试错期”再走向收敛。短期看，技术创新窗口已开，但开发者生态、云服务配套与规模效应，仍是这条新赛道最终的胜负手。

北京商报记者 陶凤 王天逸