

Agent 智能体行业研究报告

湖北社保科创基金管理中心 张泽锐

2026 年 7 月

摘要

2026 年，Agent 智能体正跨越技术可行的临界点，步入产业刚需的爆发期。国家“十五五”规划首次将其纳入顶层设计，“人工智能+”行动的全面铺开，意味着企业从要不要用转向怎么用的抢跑阶段。

剖析湖北产业链格局，其应用层繁荣、基础层薄弱的结构性特征十分突出，超 95% 企业集中于低价值应用端，而在算力、模型及平台中间件层明显缺位。对此，破局需要脱离单纯的第一梯队对标逻辑，转而立足本土工业土壤实施差异化布局：以本土汽车、光电子等优势场景为基座，开放真实产线与数据构建全国工业 Agent 场景验证中心，倒逼平台层补强；以十亿元量级专项基金实施订单 + 补贴的产业协同投资，定点培育具备平台化能力的链主企业补齐缺口；资源精准聚焦工业自动化、智能网联边缘、政务中间件及端侧决策四大赛道，依托本地场景支撑与跨区域可复制性，锁定商业化确定性最高的价值锚点。

核心结论：湖北最强大的底牌不是算力，而是千万台制造设备与千亿级工业实景。把握“智能体重塑千行百业”的本质，将应用高地转化为产业高地，湖北完全有能力走出一条以工业实景为基石的差异化特色之路。

一、产业现状

（一）产业概况

根据 2026 年 5 月国家网信办、国家发展改革委、工信部联合印发的《智能体规范应用与创新发展实施意见》，智能体是具备自主感知、记忆、决策、交互与执行能力的智能系统，是人工智能产品及服务的重要形态。从技术本质看，智能体与此前 AI 应用的核心区别不在于模型能力的高低，而在于运行范式：传统聊天机器人采用“请求—响应”的单轮交互模式，智能体则采用“感知—规划—执行—反馈”的多步闭环范式，能够自主拆解复杂目标为子任务序列，根据中间结果动态调整策略，在执行过程中主动调用外部工具，最终独立完成任务闭环。基于这一技术演进，当前围绕智能体已初步形成算力与数据服务、模型与中间件及智能体应用三大核心领域，覆盖产业链从基础支撑到价值落地的全链条。

OpenAI 将通用人工智能（Artificial General Intelligence, AGI）的演进路径划分为五个阶段，如表 1 所示。其中智能体处于第三阶段，是 AI 从被动应答迈向主动执行的分水岭。前两个阶段只解决理解问题，从智能体开始才真正解决行动问题，因此被普遍视为实现 AGI 的必经之路。

市场规模方面，根据艾媒咨询数据，全球 AI 智能体市场从 2025 年的 372 亿美元增长至 2030 年预计的 3,122 亿美元，年均复合增长率约 53%。中国通信工业协会数据中心委员会（CIDC）

表 1 通用人工智能（AGI）的五个阶段

级别	名称	描述
第一级	聊天机器人	具备基本自然语言对话能力，可回答简单问题并提供基础信息。
第二级	推理者	具备人类水平的逻辑推理能力，可分析并解决复杂难题。
第三级	智能体	能代表用户自主规划并执行任务，兼具思考与行动能力。
第四级	创新者	可在科学研究、技术创新等领域协助人类提出新理论或产品设计。
第五级	组织者	具备高度自主性与智能，可像人类组织一样进行复杂决策与管理。

资料来源：彭博社报道

数据显示，中国企业级 AI 智能体市场从 2024 年的 86 亿元增长至 2025 年的 212 亿元，预计 2029 年达 3,320 亿元，2024—2029 年年均复合增长率 107%。

智能体的产业链呈三层结构。上游基础层涵盖 AI 芯片、智算集群与数据服务，英伟达、华为昇腾等主导算力供给，截至 2025 年底全国已建成 42 个万卡级智算集群。中游模型与平台层汇聚大模型厂商（OpenAI、DeepSeek、智谱等）和编排工具（LangChain、Dify、Coze 等），是把模型能力转化为可用 Agent 的关键枢纽。下游应用层最为分散，企业服务、金融、制造、政务、医疗等垂直赛道各有深耕者。评估监控与安全治理体系贯穿全链，是 Agent 从能用到能放心用的最后门槛。

从发展脉络看，智能体在大语言模型持续迭代的基础上，与工具调用、任务规划等技术深度融合。早期主要应用于智能客服、代码生成等知识密集型场景；随着大模型推理能力跃升、API 调用成本以每年 5 至 10 倍的速度下降，以及通信协议、编排框架、Skills 生态等底层技术突破，应用场景迅速拓展至企业服务、智

能制造、金融科技、政务服务和医疗健康等领域。近年来，在国家“人工智能+”行动和十五五规划的推动下，智能体已正式上升为国家重点培育的产业方向。

总体来看，智能体产业正处于从技术可行向规模化落地跨越的关键窗口期。技术栈趋于收敛、成本快速下降、政策从鼓励转向考核，这些因素将推动产业快速进入爆发阶段。

表 2 中央层面智能体发展政策推进脉络

时间	政策	核心要点
2025.03	《政府工作报告》	“人工智能+”行动持续推进，首次强调规模化应用
2025.08	《关于深入实施“人工智能+”行动的意见》	国务院，首次设定智能体普及率量化目标（2027 年 70%、2030 年 90%）
2025.10	十五五规划建议	“智能体”首次写入五年规划，与多模态、具身智能、群体智能并列
2026.03	十五五规划纲要	第十二章第二节完整表述人工智能技术创新方向
2026.03	《政府工作报告》	首次提出“智能经济新形态”、“促进新一代智能终端和智能体加快推广”

（二）政策分析

1. 中央层面战略规划及政策体系

顶层设计层面，形成“规划引领、行动统领、报告持续推动”的支撑体系。“人工智能+”行动首次为智能体设定了明确的量化普及目标，到 2027 年，重点行业和社会各领域的普及率需超过七成，为产业发展锁定了清晰的时间表与考核标尺。“十五五”规划纲要则首次将“智能体”写入五年规划，全文“人工智能”出现频次较“十四五”时期的 6 次激增至 30 次，覆盖从基础研究、技术创新到应用治理的全链条。而政府工作报

告连续三年将“人工智能+”作为重点部署，措辞从2024年的开展行动、2025年的推进应用，演进至2026年的加快推广智能终端与智能体，并首次提出打造智能经济新形态，动词力度的逐年增强，清晰传递出国家战略由试点验证向规模化落地加速切换的明确信号。

表 3 各部委关于智能体的专项政策梳理

维度	时间	主管部门	政策	核心要点
技术创新及产业化	2025.09	国家发改委、国家能源局	《关于推进“人工智能+”能源高质量发展的实施意见》	八大类场景覆盖电网、油气、煤炭等全领域
	2025.09	交通运输部等七部门	《关于“人工智能+交通运输”的实施意见》	到2027年普及应用一批智能体、部署41个重点场景
	2025.10	国家卫健委等五部门	《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》	到2027年形成一批临床专病专科垂直大模型和智能体应用
	2025.10	中央网信办、国家发改委	《政务领域人工智能大模型部署应用指引》	四大类应用场景，明确“辅助型”定位
	2025.12	工信部等八部门	《“人工智能+制造”专项行动实施意见》	到2027年推出1000个高水平工业智能体
	2026.06	人社部等四部门	《关于加快推进“人工智能+人社”应用发展的实施意见》	到2027年培育一批人社行业大模型和智能体应用
标准规范	2026.05	网信办、发改委、工信部	《智能体规范应用与创新发展实施意见》	首次以智能体为核心主题的专项文件38条
	2026.06	国家市场监督管理总局	《人工智能智能体互联》7项国家标准	总体架构、身份码、身份管理、智能体描述等
风控	2026.04	国家网信办等五部门	《人工智能拟人化互动服务管理暂行办法》	对情感陪伴型智能体提出合规要求
基础设施	2025.08	国务院	《高质量数据集建设指引》	“数据要素×”典型场景
	2025.07	国家数据局	国家数据要素综合试验区	在十省市试点数据市场

2. 部委专项政策与制度支撑

各部委围绕 AI Agent 产业链密集出台专项政策，正加速从框架搭建向深度落地演进。标准层面，7 项《智能体互联》国家标准率先发布，为生态互通扫清底层障碍；产业化层面，工信部明确 2027 年推出 1000 个高水平工业智能体，并在医疗、政务等领域推动大规模应用示范，政策着力点从鼓励研发转向供需匹配；治理层面，针对拟人化互动服务出台专门办法，明确决策权限与分级治理框架，体现发展与监管并重的审慎思路；基础设施层面，高质量数据集建设与国家级数据要素试验区的协同推进，正缓解制约 Agent 落地的数据瓶颈。整体而言，密集的政策群为行业从试点转向规模化商业闭环提供了坚实的制度保障与需求托底。

3. 湖北省内政策梳理

湖北省紧跟国家战略部署，一是出台《关于加快推进人工智能产业发展的实施意见》，提出到 2027 年人工智能产业规模达到 3500 亿元、武汉市进入全国 AI 城市第一方阵的总体目标，将智能体明确列为关键技术攻坚方向和重点培育业态。二是发布《湖北省“人工智能 + 制造”专项行动实施方案》，到 2027 年推出 100 个以上高水平工业智能体、培育 1200 家以上智能工厂，建设省级智能体公共服务平台，鼓励中小企业开发轻量化智能体模块，发展“工业互联网平台 + 场景智能体”新模式。

在区域定位上，湖北已形成各有侧重的分工格局：武汉作为核心，东湖高新区聚焦智能体核心技术研发与 OPC 创新社区孵

表 4 湖北省主要城市智能体相关政策

城市	核心政策与关键举措	重点方向
武汉市	1. 算力券补贴：不低于 1000 万元/年。 2. 生态建设：“OPC 汉十条”，建立生态社区。 3. 产品奖励：武昌区率先出台智能体专项奖励。	省内核心增长极；产业规模 600 亿 +
襄阳市	1. 平台建设：启动“襄数智汇”智能体服务运营平台。 2. 项目落地：携手科大讯飞打造城市治理智能体示范区。	智能网联汽车与城市治理
宜昌市	1. 运营补贴：每年发放智能券（最高 30 万元/家）。 2. 重点项目：签约联想城市超级智能体。 3. 算力基建：加速推进智算“万卡”集群。	OPC 生态与算力基础设施

化，经开区侧重工业智能体与汽车制造深度融合；襄阳依托“襄十”算力圈和达安中心等自动驾驶测试资源，发展智能网联汽车与城市治理智能体应用；宜昌以点军区为核心，通过“算力基建 + 开源生态 + 场景驱动”四位一体布局，聚焦 AI 一人公司（OPC）生态和超级智能体落地。各地市围绕定位已密集出台配套支持政策，具体措施详见表 4。

4. 先进省市政策对比分析

如表 5 所示，在各省市竞相布局的格局下，政策导向正呈现出三大共性趋势：一是在供给端，以“算力 + 模型 + 语料”组合券形式降低开发门槛，如上海每年统筹超 10 亿元补贴、深圳推出专项“模型券”；二是在需求端，通过大赛、揭榜挂帅与示范项目创造增量市场，如北京丰台设立专项大赛、杭州以高额奖励吸引头部企业；三是在生态端，通过开源社区、行业立法及龙头企业牵引构建产业闭环，如合肥以科大讯飞为链主发布开源

表 5 先进省市智能体相关政策对比分析

省市	政策核心要点
北京市	“软八条”及 AI 智能体基础设施专项，通用智能体运营服务调用算力和模型成本最高支持 3000 万元；推动智能体自主互联协议和技能组件库开源建设；丰台区率先设立智能体创新创业大赛专项（最高 300 万元）。
上海市	发放算力券（6 亿元）、模型券（3 亿元）、语料券（1 亿元）；“揭榜挂帅”将科学智能 Agent 服务列入攻关方向；“模速空间”“模力社区”集聚效应明显。
深圳市	“模型券”直接覆盖智能体开发应用（最高 200 万元/年）；发布“人工智能 +”先进制造业行动计划；启动 OPC 创业生态引领地行动计划。
杭州市	十五五规划将智能体与芯片、大模型并列为全产业链核心环节；钱塘区 AI 新政“19 条”，智能体开发最高 300 万元/年；目标 2027 年智算规模超 100 EFLOPS。
南京市	率先启动 AI 产业立法，将 OPC 支持写入法律；发布全省首个城市级政务智能体推进方案，提出“一部门一智能体”。
合肥市	科大讯飞发布首个企业级 Agent 开源平台；省级出台“人工智能 + 万物”行动方案；设立 100 亿元未来产业基金。

平台、南京率先将 OPC 支持写入地方立法。

湖北未来可借鉴各地优秀经验优化政策体系：优先推动本地汽车与光电子链主企业开放核心产线场景，将链主企业在数字化改造中的真实痛点转化为智能体的采购订单，以给订单、给场景的确定收益代替单纯的“给补贴”；其次，依托光谷已形成的 OPC 社区生态，参考上海“模速空间”的运营机制，聚焦中小团队的轻量化开发与算力接入需求，形成有别于大厂生态的分布式创新承载力；核心策略应聚焦于订单供给 + 轻量开发的双规联动，以此形成具有湖北属性的产业壁垒。

（三）核心技术与演进趋势

智能体的技术栈在 2025—2026 年完成了从“功能可用”到“工程可靠”的关键跨越。具体表现在三个层面：① 大模型基座

从文本理解走向多模态推理，测试时计算范式使推理能力显著提升；② 记忆系统从单次 RAG (Retrieval-Augmented Generation, 检索增强生成) 检索演进为 Agentic RAG 与记忆内化的三代技术体系；③ 工具交互从紧耦合 Function Call 走向 MCP、A2A 双协议栈的标准化架构。上述技术条件的同步成熟，使得 Agent 首次具备规模化落地的工程基础。接下来围绕 Agent 核心架构、模型基座演进、记忆系统、工具与协议、多智能体协作五个维度，梳理当前技术现状、关键瓶颈与演进趋势。

1. 核心架构

AI Agent 的技术架构沿袭了经典的智能体框架，即通过传感器感知环境、经效应器完成推理决策、再通过执行器作用于外部世界，记忆系统则贯穿全流程承担上下文保持与知识检索功能。如图 1 所示，在大语言模型 (Large Language Model, LLM) 驱动的 Agent 系统中，效应器的核心是 ReAct (Reasoning + Acting) 工作循环，即模型在“思考-行动-观察”的闭环中持续迭代完成任务规划与执行；执行器经历了从早期 Function Call 到 MCP 协议标准化工具接入、再到 A2A 协议实现跨 Agent 通信的演进路径；传感器则随着多模态模型的成熟，从纯文本输入扩展至音频、图像等多通道感知。

2. LLM 基础模型

LLM 是 Agent 的认知中枢，承担任务理解、步骤规划、工具选择与执行反馈的全流程决策功能。在 Agent 架构中，编排框

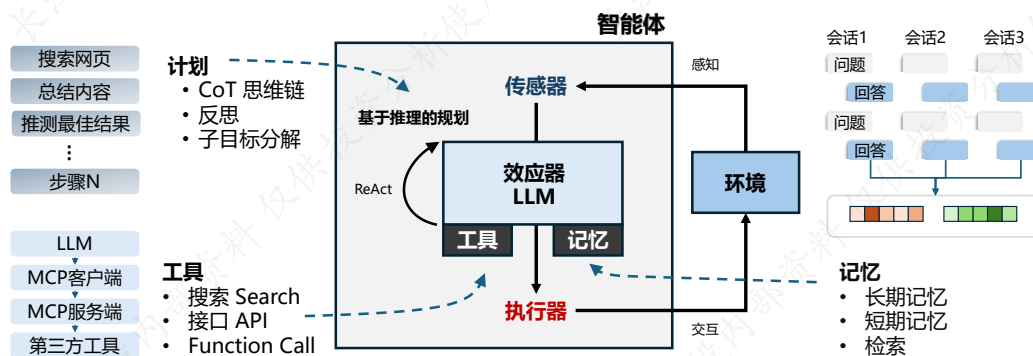


图 1 Agent 核心技术架构

架负责状态管理和调度，记忆系统负责信息存取，但每一个执行节点的判断和选择最终由 LLM 完成。因此，LLM 的能力边界直接构成 Agent 能力的上限约束。2024 年之前，主流模型在逐步推理中的表现高度依赖提示工程（Prompt Engineering）：人工设计推理模板引导模型逐步输出。其局限在于当步骤超过某个阈值后，模型偏离初始目标的概率快速上升。产业侧的应对是将 Agent 的执行步骤限制在较短的范围内。2025 年起，强化学习驱动的内生推理成为主流模型的通用能力，模型在输出结果前可以自主展开数万字的思考过程，这直接扩展了 Agent 的任务复杂度边界。当前主要局限在于，推理链延长到一定程度后准确率仍会出现下降，且模型在推理过程中可能编造未执行的操作步骤，目前缺乏对此类行为的统一评估方式。

推理成本是 Agent 规模化的前置条件。通用对话场景下一次推理调用即可完成交互，而 Agent 完成一项任务通常需要连续调用模型数十到数百次，单次成本的差异在累加后被显著放大。过去两年，两条技术路径共同降低了推理成本：架构层面，混合

专家模型（Mixture of Experts, MoE）将每次推理的计算量压缩到总参数规模的三十分之一至二十分之一；工程层面，模型量化和推理加速方案使中小模型的本地部署成为可行选项。当前同等能力的 Agent 任务推理成本较 2023 年已降低约一个数量级。

3. 记忆系统

记忆系统是 Agent 维持多步任务一致性的基础。Agent 的每个执行步骤都需要访问此前积累的信息——已完成哪些子步骤、各自的执行结果、曾出现的异常及处理方式。与通用对话场景相比，Agent 对记忆的访问频率和依赖深度更高，记忆系统的效率直接影响任务成功率。当前 Agent 记忆体系由三个技术环节构成：短期记忆、长期记忆与记忆演化。

短期记忆以模型的上下文窗口为载体，窗口容量从 2023 年的 32K token 扩展至 2026 年的百万级，容量约束已基本解除，摘要和卸载等配套策略在主流框架中趋于标准化。但容量问题解决后，利用效率构成新的约束：窗口越大，模型从中段信息的检索准确率越低，Agent 的多步执行持续向窗口内追加新内容，这一问题随任务增长而累积。

长期记忆负责跨会话的信息保留，其技术基础是 Embedding 向量化与语义检索，与 RAG 共享大部分底层技术组件。二者的功能定位不同：RAG 面向共享知识源（企业文档、产品手册等），长期记忆面向用户交互历史中的个性化信息（偏好、习惯、历史决策）。当前主流方案以 Mem0、ZEP 和腾讯云 Agent Memory 为

代表，在存储结构上分别采用图记忆、时间知识图谱和分层提取等路径，写入与检索的基础链路已有可用方案。

记忆演化是当前最不成熟、也最被低估的环节，包含三项操作。合并：将分散在多个会话中的语义相近的记忆片段整合为更完整的知识实体；反思：对任务执行结果做复盘，提取可复用的规则与经验；遗忘：对过时信息和低价值片段做淘汰。前两项已有工程实践，遗忘机制仍缺乏通用方法。

4. 工具与协议

Agent 的执行能力取决于能否调用外部工具：查询数据库、操作文件、访问 **API**。2023 年前后，工具接入的主流方式是 **Function Call**，即开发者为每个工具手动编写描述和参数定义，注册到模型后由模型生成结构化文本触发调用。这一模式的边际成本随工具数量线性增长，每新增一个工具，所有对接方都需要单独更新配置和适配逻辑。

MCP (Model Context Protocol) 改变了工具与 **Agent** 之间的连接方式。引入客户端-服务器架构，工具提供方以服务端形式一次性完成接口描述和能力暴露，**Agent** 框架作为客户端通过统一协议动态发现和调用，不再需要为每个工具做单独适配。工具接入的边际成本由此从线性增长压缩至接近常数。

A2A (Agent-to-Agent Protocol) 解决了不同智能体之间“语言不通”的协作壁垒。如果说 **MCP** 是 **Agent** 连接外部工具的“统一接口”，**A2A** 则是 **Agent** 与 **Agent** 对话的“通用翻译”。

它基于 HTTP 与 JSON 协议，使不同来源的 Agent 无需额外适配即可相互发现、委托任务并同步状态。多智能体协同的边际成本由此从“逐对定制”降至“标准化调用”。

（四）产业链分析

1. 产业链上中下游

AI Agent 产业链呈“算力与数据驱动基础、平台与框架衔接转化、应用场景实现价值”的三层漏斗结构，同时配套监测评估与安全治理体系贯穿全程。价值链条的当前重心正从模型生产向工程化落地转移。

上游：基础层（算力、数据与基础设施）。这一环节是制约 Agent 规模化部署的硬性天花板。截至 2025 年底，中国智能算力超 1,590 EFLOPS，但高端 GPU 受出口管制影响，华为昇腾、寒武纪等国产芯片正加速替代，CUDA 生态的迁移成本与开发周期仍是国产化进程中的核心风险。算力之外，高质量数据的供给是当前更为严峻的隐性瓶颈。Agent 的规划、记忆与工具调用高度依赖垂直领域（特别是工业、政务）的微调数据与操作日志，谁掌握高质量、可合规流通的行业数据集，谁就占据了产业链上游的主动权。

中游：模型与平台层。这一层承担着“将通用大模型的智力转化为可执行任务”的关键枢纽职能，是当前产业链竞争最为激烈的领域。行业内正形成平台即底座的竞争格局：头部大模型厂商（如字节 Coze、阿里百炼、百度千帆等）以低代码构建工具构

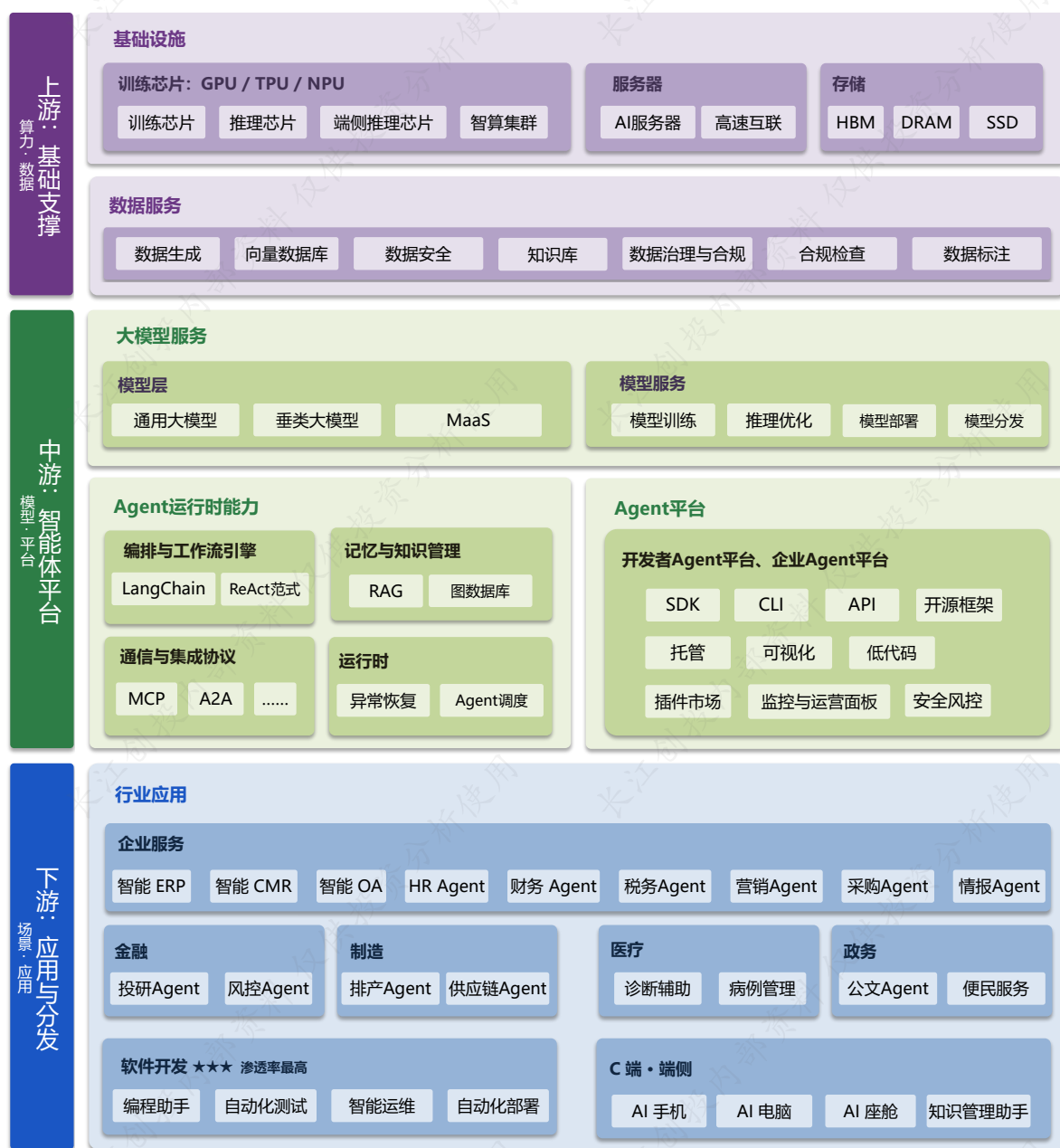


图 2 AI Agent 产业链全景图

建生态锁定；开源框架（LangGraph、Dify 等）则抢占开发者心智。其中，MCP 与 A2A 协议的标准化进程至关重要，为产业构建起一套跨工具调用与跨智能体协作的通用连接协议。中游环节的核心价值在于能否有效降低 Agent 的开发门槛、解决多步任

务编排的稳定性，并缩短从 Demo 到上线的时间周期。

表 6 AI Agent 产业链各环节代表企业

层级	子环节	覆盖内容	代表企业/产品
算力与数据	AI 芯片	训练芯片、推理芯片、端侧推理芯片	英伟达、AMD、华为昇腾、寒武纪、海光信息
	智算集群与云服务	万卡/十万卡集群、算力调度	阿里云、腾讯云、华为云、火山引擎、百度智能云、中科曙光
	训练与推理框架	分布式训练、推理优化引擎	PyTorch、TensorFlow、昇思 MindSpore、飞桨 PaddlePaddle
	数据服务	采集、清洗、标注、合成数据	星环科技、数据堂、海天瑞声
	向量数据库	向量存储与检索、图存储	ChromaDB、Pinecone、Milvus、Qdrant
大模型	海外通用大模型	API 调用、私有化部署	OpenAI（GPT）、Anthropic（Claude）、Google（Gemini）、Meta（Llama）
	国内通用大模型	API 调用、私有化部署	智谱（GLM）、科大讯飞（星火）、阿里（通义千问）、百度（文心）、DeepSeek
智能体平台与中间件	行业垂类模型	金融/医疗/制造等行业专用模型	恒生电子、卫宁健康等
	MaaS	模型推理、微调、部署平台	智谱、百度智能云、阿里云百炼、火山引擎、华为云
	编排与工作流引擎	Agent 任务规划、多步执行管理	LangChain、LangGraph、CrewAI、AutoGen
	通信协议	Agent—工具、Agent—Agent	MCP（Anthropic）、A2A（Google）、AGNTCY（Cisco 等）
	记忆与知识管理	会话记忆、长期记忆、知识检索融合	Mem0、EverMind、LangMem
	技能与工具生态	公共技能库、工具连接器、插件市场	Skills（Anthropic）、Coze 插件市场、百炼插件中心
	Agent 开发平台	低代码/无代码/API Agent 构建	字节 Coze、阿里百炼、Dify、OpenAI Agents SDK、Google ADK
	评估与监控	Agent 行为追踪、质量评估、运行时监控	MLflow、LangSmith、Datadog LLM Observability
	安全与治理	权限控制、数据脱敏、人机协同、合规审计	Credo AI、Arize、华为云 ModelSec

续表 6

层级	子环节	覆盖内容	代表企业/产品
智能体应用	企业服务	ERP、CRM、OA、IT 服务智能化	用友、金蝶、金山办公、Salesforce (Agentforce)、ServiceNow
	生产制造	工业、供应链 Agent	汉得信息、赛意信息、能科科技、鼎捷数智
	金融	风控、投研、客服 Agent	京北方、恒生电子、宇信科技、中科金财
	政务	公文处理、政策匹配、便民服务	新点软件、科大讯飞、南威软件
	医疗	辅助诊断、病历管理、药物研发 Agent	卫宁健康、医渡科技、鹰瞳科技
	端侧/C端 Agent	AI 手机、AI PC、智能座舱	华为、小米、OPPO、联想、蔚来、小鹏
	开发者 Agent	个人自主 Agent、AI Coding、开发者工具	通义灵码、智谱 CodeGeeX、Cursor、GitHub Copilot、Claude Code

资料来源：公开资料整理

下游：应用层。这一层直接兑现产业链的经济回报，当前正处在从“人机协同”向“单任务自主执行”过渡的关键节点。渗透率最高的赛道是 AI 代码生成 (AI Coding) 与企业级 SaaS (ERP、CRM 和 OA 的智能化改造)；工业、金融、政务等复杂场景仍处于深度渗透阶段。目前下游面临的核心商业矛盾是“推理成本高企”与“客户付费意愿”之间的博弈：Agent 的多步推理与工具调用导致单次任务消耗 Token 量是普通对话的 10 至 100 倍，如何在客户可接受的成本预算内保证执行准确率，是决定应用层能否大规模商业化闭环的终极考验。

全链支撑：评估监测与安全治理。随着 Agent 从“对话”升级为“具备执行权限的数字员工”，系统可靠性、合规性与可

追溯性成为产业化的“压舱石”。当前已形成技术层（MLflow、LangSmith、Datadog 等运行时监控与链路追踪）与制度层（权限控制、数据脱敏、人工干预机制及合规审计）的双轮支撑体系。在业务场景尤其是涉及工业控制与金融交易时，治理能力直接决定了 Agent “能做什么、不能做什么”，是当前判断一家 Agent 企业是否具备大规模交付能力的核心分水岭。

2. 产业生态分布

第一梯队（北京、上海、深圳）占据产业链“制高点”，掌握平台与基础设施的话语权。北京集结了字节扣子、百度千帆、智谱 AutoGLM 等头部平台型企业，依托其大模型底座与大规模流量分发，主导着 Agent 开发框架与分发入口；上海凭借“模速空间”“模力社区”等高水平集聚区，在风险投资与模型生态上形成了强大的磁吸效应；深圳则以华为昇腾为根，构建从底层算力到开发工具链的国产化闭环。第一梯队的共同特征是以平台即生态的逻辑构建护城河，实际上锁定了 Agent 产业的定价权与上游利润。

第二梯队（杭州、南京、合肥、武汉）各具差异化禀赋，大多处于垂直深潜阶段，尚未形成平台级生态。杭州依托阿里百炼的云原生 Agent 平台，试图将 Agent 能力内嵌入庞大的企业级 SaaS 生态；南京以制度创新为破局口，率先启动 AI 产业立法，试图通过规则先行吸引创新要素；合肥以科大讯飞为链主牵引产业链，侧重开源与教育场景；武汉凭借“模态空间”孵化器与

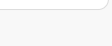
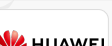
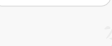
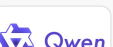
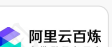
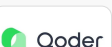
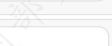
城市	算力与数据	大模型	平台与中间件	智能体应用
北京 一梯队	 寒武纪  火山引擎  百度云  飞桨  数据堂  海天瑞声	 GLM  Kimi-Dev  文心一言	 智谱·AI  百度云  火山引擎  Coze  文心智能体平台	 用友  金山  致远互联  i  文心快码  CodeGeeX  京北方  中科曙光  Lenovo  mi  亿渡科技
深圳 一梯队	 Ascend  腾讯云  MindSpore  华为云	 Tencent Hunyuan  ModelScope	 HUAWEI  腾讯元器  ModelScope	 金蝶  HUAWEI
上海 一梯队	 zilliz  MINIMAX  卫宁健康	 MINIMAX  卫宁健康		 HAND  鼎捷数智  NIO  商汤
杭州	 阿里云	 Qwen  阿里云百炼	 阿里云百炼	 HUNDSUN  Qoder  AionClaw  海康威视
苏州			 Dify	 Epoint
合肥		 讯飞星火		 科大讯飞
武汉		 紫东太初	 模态跃迁	 Thingo

图 3 中国智能体相关重点企业分布格局

密集高校人才供给，已集聚 Agent 相关企业超 800 家、产业规模突破 600 亿元。但需要注意的是，武汉当前的产业繁荣高度依赖应用层开发，无平台企业意味着没有生态入口，企业利润极度依赖外部大模型厂商的 API 调用成本。

第三梯队（成都、西安、长沙等）则以承接一二线城市的垂直场景落地、数据标注及部分代码外包为主，尚未形成完整的产业闭环，在当前价值链中处于弱势地位。

3. 专利布局情况

国内智能体专利申请在过去两年连续翻倍增长，2024 年 4,282 件，2025 年攀升至 1.04 万件，同比增速分别为 91.8% 和 143.3%，截至 2026 年 6 月累计存量 2.5 万件。发明专利占比超过 95%，行业整体处于技术密集的快速扩张阶段。

从技术分布看，当前专利高度集中于信息检索和商业管理两个方向，分别对应 RAG 与知识库检索、企业流程自动化等通用基础能力，申请量的集中意味着技术门槛仍在应用层。多智能体协作与通信协议、可解释性推理等更接近核心技术环节的方向，专利积累明显不足。

（五）市场规模预测

1. 全球市场

全球 AI Agent 市场正处于从“概念验证”到“规模化落地”的加速期。据艾媒咨询统计，全球市场规模预计从 2025 年的 372 亿美元增长至 2030 年的 3,122 亿美元，年均复合增长率约 53%。区域格局上，北美凭借基础模型先发优势与成熟的企业级 SaaS 生态占据主导地位；亚太地区在政策驱动与产业数字化需求的共振下成为增长最快的区域市场；欧洲则在流程自动化与 AI 合规治理领域形成差异化竞争优势。

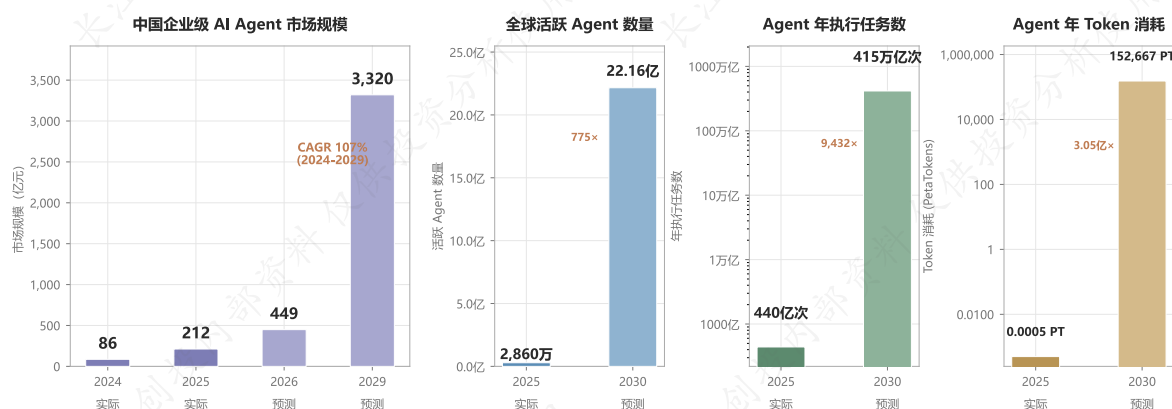


图 4 全球及中国 AI Agent 市场核心指标（市场规模、活跃 Agent 数、年任务数、Token 消耗）

2. 中国市场

中国市场展现出比全球更为陡峭的增长斜率。根据中国通信工业协会数据中心委员会与科智咨询联合发布的白皮书，中国企业级 AI 智能体市场规模从 2024 年的 86 亿元增长至 2025 年的 212 亿元，预计 2029 年将突破 3,320 亿元，2024—2029 年复合年增长率高达 107%。不过需要指出的是，各机构对“Agent 市场规模”的统计口径存在显著差异。从表 7 可见，狭义口径（仅计算平台与软件）与广义口径（包含全产业链基础设施）的预测结果差距可达数倍。当前行业处于标准尚未统一的快速发展期，但剔除口径影响后，行业整体处于“爆发式增长轨道”的趋势已形成普遍共识。

3. 核心驱动因素

本轮市场爆发具备清晰的供需逻辑与政策托底。在成本端，大模型推理 Token 价格以年均 5 至 10 倍的速度下降，直接降低了 Agent 的单次执行成本；在能力端，Agent 正从辅助对话迈向

表 7 主要机构对智能体市场的预测对比

机构	区域	基准年数值	预测年数值	CAGR	口径范围
IDC	全球	2024: \$5.3B	2030: \$47.1B	~44%	狭义（平台 + 软件）
Markets and Markets	全球	2024: \$5.1B	2030: \$47.1B	~45%	与 IDC 基本一致
艾媒咨询	全球	2025: \$37.2B	2030: \$312.2B	~53%	广义（全产业链）
CIDC	中国	2024: ¥86 亿	2029: ¥3,320 亿	~107%	企业级 Agent
中商产业研究院	中国	2025: ¥69 亿	2030: ¥287 亿	~33%	狭义（平台 + 软件）
艾媒咨询	中国	2025: ¥804 亿	2030: ¥6,968 亿	~54%	广义（全产业链）

资料来源：各机构公开报告整理

自主执行。如图 4 所示，据 IDC 预测，全球活跃 Agent 数量将从 2025 年的 2,860 万个增长至 2030 年的 22.16 亿个，年执行任务数将从 440 亿次跃升至 415 万亿次；在政策端，“人工智能 +”行动设定的 2027 年 70% 普及率目标，锁定了政企大规模部署的刚性需求下限。三者共振，构成了 Agent 产业高速增长的核心支撑。

（六）竞争格局与重点企业分析

1. 全球竞争格局

全球 Agent 产业呈现“北美引领产品化、中国聚焦场景规模化、欧洲深耕合规治理”的三极分化。北美凭借基础模型先发优势与企业级 SaaS 存量生态，率先跑通“存量客户升级 + Agent 增值付费”的商业模式；中国则以丰富的应用场景、自上而下的政策推力及活跃的开源社区为支点，在应用层和部分平台层快速形成竞争力，但在基础模型能力与协议标准制定上仍处跟跑；欧洲基于其严格的监管框架，在企业流程自动化与 AI 合规治理领域建立了差异化壁垒。

2. 国内竞争格局

国内 Agent 产业当前呈现“平台层高度集中、应用层群雄割据、开源生态生机勃勃”的格局。平台层已被少数头部云厂商及大模型公司垄断，依靠算力入口与亿级流量锁定开发者生态；应用层则按垂直赛道分散竞争，企业软件、工业制造、金融 IT、政务等领域各有深耕多年的主力厂商，行业 Know-How 与存量客户关系构成主要护城河。值得关注的是，国内开源社区与 OPC 生态是区别于海外闭环格局的独特变量，在长尾场景和边缘应用中提供了极具活力的多样性补充。

3. 省内重点企业分析

从产业层级结构看，湖北当前呈现出“应用层初步集聚，平台层与基础层系统性缺位”的典型特征。在算力与数据、大模型服务两个上游层级，省内尚无具备规模竞争力的企业；平台与中间件层仅靠“模态跃迁”一家覆盖，团队规模与融资体量尚处早期。应用层则形成了梯队化的初步沉淀：格创东智（TCL 孵化、C 轮数亿元、营收超 10 亿元）依托工业场景已服务超 600 家湖北企业；数命科技以 17 人团队实现 2025 年营收 2 亿元，在 C 端 Agent 商业化验证速度上表现突出；另有“猴子说话”“浮木科技”等一批新生力量在端侧和政务赛道崭露头角。

（七）湖北省产业基础与资源禀赋评估

在各地竞逐 Agent 产业的背景下，湖北的核心禀赋并非源于算力或资本的“绝对规模”，而在于其“产业场景密度 + 高校人

表 8 湖北省 Agent 产业重点企业

产业链层级	公司	方向	核心产品/业务	竞争力
算力与数据	—	—	—	本层暂无优势企业
大模型服务	—	—	—	本层暂无优势企业
平台与中间件	模态跃迁	Agent 基座	多智能体协同框架,覆盖水利水文、高校教育等场景	40 人团队 (硕博占比超 80%), 累计融资数千万元, 获湖北省科技进步奖一等奖
应用层	猴子说话	端侧 Agent	办公智能体 Simen, 可在个人电脑本地运行	产品未发布已获近 4000 万元投资, 10 人团队约 80% 代码由 AI 生成
应用层	格创东智	工业 Agent	章鱼智脑 Agentic AI 平台	TCL 孵化, C 轮融资数亿元, 营收超 10 亿元, 已服务超 600 家湖北企业
应用层	浮木科技	政务 Agent	“浮木 AI” 社会治理可信增强引擎	已服务全国十余省份, 近三年营收复合增长率近 100%
应用层	数命科技	Agent 应用	AIWIN 任务经济操作系统、AI 寄养平台	17 人团队 2025 年营收 2 亿元, C 端 Agent 商业化验证速度最快
应用层	归心认证	Agent 应用	多模态 AI 人才认证与实战考核引擎	华科大 00 后学生创业, 获 4000 万元融资
应用层	灵境万维	Agent 应用	数字内容创作 AI 工具链	成立半年获数百万元天使投资
应用层	火星登陆	空间智能	为具身智能提供空间计算基座	2025 年成立, 种子轮及天使轮融资近千万元
应用层	PsyGo	C 端 Agent	通用智能体平台、多智能体协同能力	开箱即用、支持私有化部署; 竞品对标: OpenClaw、字节 Coze

资料来源: 公开资料整理

才密度 + 综合成本洼地”所形成的叠加优势, 使其天然具备成为 Agent 产业创新创业试验田的潜力。

从产业场景与人才供给看, 湖北拥有全国独一无二的工业“语料库”。汽车制造与光电子信息两大万亿级产业集群, 为工业 Agent 提供了涵盖产线质检、供应链调度、设备运维等丰富且复杂的真实验证场景; 同时, 武汉大学与华中科技大学长期在 NLP、AI+ 制造、AI+ 安全等领域的技术积淀, 成为持续输出优

质创业团队的核心引擎。

从算力基建与要素成本看，武汉作为全国八大算力枢纽之一，且是全国唯一“一跳直达”其余七大枢纽的城市节点，其网络延迟优势在 Agent 推理与跨地域任务调度中具备实质性的商业价值。武汉人工智能计算中心与超算中心合计 1.8 万 P 的算力储备，“江城模境”大模型公共服务平台汇聚的 1400 余个开源模型，叠加要素成本仅为北京三分之一、上海二分之一的优势，为处于早期烧钱阶段的 Agent 创业企业提供了宝贵的经营缓冲期。光谷 30 个 OPC 生态社区与 400 余家 AI 类 OPC 市场主体，已初步显现出轻量化开发、低成本试错的集群氛围。

综合来看，湖北在工业场景、人才密度与综合成本上的先发优势，已使其具备了孕育优质 Agent 企业的根基。但若无法在短期内破除应用端拥堵、平台端缺位的结构性瓶颈，并补齐本土资本的接力断层，湖北 Agent 产业的发展极易在达到现有规模天花板后陷入增长停滞，难以实现对领跑地区的实质性超越。

二、面临的机遇和挑战

（一）发展机遇

从宏观到微观，三重力量正同时推高 Agent 产业的天花板。一是政策端进入考核倒逼阶段。“人工智能+”行动首次设定 2027 年 70%、2030 年 90% 的智能体普及率量化目标，“十五五”规划首次将智能体写入五年规划，政府工作报告连续三年措辞升级，企业采购 Agent 已从鼓励探索变为经营刚需。Gartner 预测，

集成任务型 AI Agent 的企业应用占比将从 2025 年的不足 5% 跃升至 2026 年的 40%。二是技术端迎来跨越式成本下降。2025 年 DeepSeek R1 以 OpenAI o1 十分之一的成本实现同等推理能力，国产模型推理价格已降至海外同级的十分之一量级；MCP 协议快速成为行业事实标准，A2A 协议打通了 Agent 间的通信屏障。IDC 预计，全球活跃 Agent 数量将从 2025 年的 2,860 万个激增至 2030 年的 22.16 亿个。三是湖北要素禀赋构筑了独特的场景与成本优势。汽车制造与光电子信息两大万亿级产业提供了工业 Agent 最天然的验证土壤；数据要素综合试验区牌照赋予湖北在行业数据开放上的先行先试权；同时，武汉的要素成本仅约为北上杭深的三分之一至二分之一，光谷已形成 30 个 OPC 生态社区及 800 余家企业、600 亿元规模的初步集聚效应。

（二）面临挑战

在巨大机遇的背后，湖北乃至全国 Agent 产业面临四个必须正视的结构性挑战：

1. 产业落地仍受制于工程化临界点。当前 Agent 从演示级到生产级的跨越尚未完成。伯克利大学在 2025 年的研究显示，约 68% 的生产级 Agent 执行步骤被限制在 10 步以内，而多 Agent 协作中的错误级联效应更为严峻，单 Agent 5% 的错误率在多体协同中可膨胀至 86%。此外，Gartner 筛选的数千款标榜“智能体”的产品中，真正具备自主决策与执行能力的仅约 130 款，“智能体洗白”现象普遍存在。

2. 本地资本接力链断裂，生态锁定效应薄弱。北京、上海、深圳已形成从天使轮到 Pre-IPO 的完整投资链条，而湖北以本地国资和早期风投为主，单笔融资多停留在千万元量级。Manus 的迁出轨迹极具警示意义：该团队在光谷完成产品从 0 到 1 的研发后，因本地缺乏承接 A/B 轮中后期大规模融资的生态与资本，B 轮 7500 万美元融资最终流向硅谷，总部亦随之迁至新加坡。优质项目出省出海已成为湖北当前最隐蔽的产业流失风险。

3. 产业链结构呈现“倒金字塔”风险。全省 800 余家 Agent 企业中，超过 95% 集中于低门槛的应用层，算力与数据服务层、大模型服务层均无优势企业，平台与中间件层仅模态跃迁一家覆盖。这种高度依赖外部 API 和框架的格局，导致本地应用企业在算力成本与模型调用上极度缺乏议价权，极易陷入规模做大但利润微薄的困境。

4. 监管出清与合规门槛正在重塑赛道。Agent 自主决策带来了全新的责任认定与安全边界难题。随着《人工智能拟人化互动服务管理暂行办法》正式施行，主流平台集中下线了大量 C 端智能体功能。短期来看，这一合规阵痛将冲击部分 C 端应用，但长期看，它将加速行业出清，推动产业重心向具备稳健治理能力的企业端场景倾斜。

三、下一步思路建议

（一）出台支持政策建议

国家层面，一是建议加快制定智能体产业发展战略与专项规

划，将万卡智算集群、高质量数据集、开源社区等纳入国家新型基础设施布局。二是期待加快《人工智能法》立法进程，明确智能体自主决策的责任归属与权限边界，推行基于风险的智能体安全评估与分级认证。三是建议在科技重大专项中增设智能体技术方向，重点支持推理模型架构、多智能体协作协议、Agentic RAG、端侧推理芯片等基础共性技术，并建设国家级智能体评测基准与公共测试平台。四是布局“智能体科学与工程”一级学科，联合头部企业建设产教融合实训基地。

省级层面，一是加快出台《湖北省智能体产业创新发展三年行动方案》，明确量化目标与区域分工，同步推动省级产业促进条例立法。二是依托武汉人工智能计算中心等国有算力基础设施，设立 Agent 推理专项补贴。三是加强产业基金设立的政策和资金支持力度，联动社会资本。四是加强人才培养与产学研融合，探索通过设立 Agent 技术专项职称评审、对 Agent 领域领军人才给予认定和专项奖励。

（二）产业发展建议

湖北在 AI Agent 产业上的竞争策略，应明确放弃与北上广深在“通用算力与通用模型”上的全面对标，转而走一条基于自身制造业底色的差异化道路。核心策略可围绕“战略锚点、赛道切口、空间落位、要素保障”四个维度展开。

战略锚点上，聚焦两大核心地标。第一，依托汽车与光电子两大支柱产业，确立“中部 AI+ 智能制造应用创新高地”的定

位，在工业 Agent 与制造 Agent 领域确立全国影响力；第二，借助数据要素综合试验区的先行先试红利，打造“数据要素驱动 AI 产业示范区”，以数据流通和行业数据集建设作为先发优势。

产业重心上，聚焦四个垂直领域。智能制造 Agent（结合汽车制造和光电子信息产业基础）、智能网联汽车 Agent（发挥“光芯屏端网”产业集群的硬件优势）、医疗健康 Agent（利用武汉丰富的医疗资源和高校科研积累）、智慧政务 Agent（抓住政务数字化转型的政策窗口）。这四个方向既有产业基础支撑，又能与第一梯队形成差异化竞争。

区域布局上，构建“一核引领、多点支撑”的空间格局。以光谷为核心打造智能体产业集聚区，集中布局研发孵化、算力服务和创新平台；以武汉经开区为制造应用示范区，推动 Agent 在汽车制造和工业场景的先行先试；在宜昌、襄阳等地根据产业特色布局智慧物流和农业 Agent 应用试点。

生态构建上，构建“数据 + 算力 + 人才”的产业闭环。依托湖北数据要素试验区先行先试的优势，率先建立行业级高质量数据集（尤其是汽车与光电子领域），降低 Agent 微调成本；整合省市算力资源，以“算力券”等形式定向补贴中小型 AI 企业，形成算力洼地；依托武汉密集的高校资源，设立“AI Agent 产教融合实训基地”，使人才供给与本地产业需求精准对接。

（三）集团工作思路

作为湖北制造业生态的核心参与方，集团在 Agent 产业中的

角色应区别于单纯的市场化 VC。集团的战略价值不在于跟投的力度，而在于能否将汽车、光电子等制造板块的真实产线需求，转化为 Agent 企业“落地即获订单”的市场入口。以此为基点，集团的工作思路应从“资本配置”与“场景供给”两个维度穿透：以产业基金作为撬动外部资源的支点，以并购整合作为资本退出的闭环，在投得准、留得住、退得出之间形成一套完整的产业投资闭环。具体而言，战略与执行围绕以下三个层面展开。

1. 战略定位：做“生态构建者”而非“财务投资者”

集团的 AI Agent 投资不应停留在散点式财务跟投，而应立足于自身制造业板块的“长板效应”，构建场景反哺投资、投资验证场景的闭环。将汽车、光电子、高端装备等制造板块转化为 Agent 企业的真实试验田与付费首发市场，以实际的产线痛点倒逼 Agent 技术落地，将早期资本投入转化为可验证的产业订单，实现产业增值与资本回报的深度绑定。

2. 资本赋能：以基金撬动外部资源，补齐早期标的短板

建议设立总规模 10 亿元的 AI Agent 产业专项基金，采用“母基金 + 直投”的双轮驱动模式。一方面联合省级引导基金发挥财政资金的杠杆效应；另一方面，主动与具备大模型产业背景的外部一线 VC 建立联合投资与战略合作关系，将外部机构的项目研判能力与全国项目资源池纳入本地投资决策的参考体系，以此弥补湖北早期创新标的供给不足的短板。对于平台层与中间件等战略环节，可采取集团直投或主导并购的方式，以投带

引，将产业链关键节点锚定在省内。

3. 退出闭环：以“并购整合”为主要导向的资本循环

在当前 IPO 审核周期显著拉长的背景下，集团应将并购整合作为优先级最高的退出路径。通过系统梳理集团体系内链主企业（东风、华工科技等）的技术采购与并购需求，引导其出具明确的 Agent 技术产品采购清单；通过定期举办产业协同对接会，推动被投企业的技术能力优先嵌入湖北制造数字化改造项目中，以“业务协同拉动股权整合”，实现投资退出与集团实体产业升级的双重闭环。