中国具身智能数据产业分析报告2026

易观分析

2026年9月

本产品保密并受到版权法保护

Confidential and Protected by Copyright Laws

Analysys 易观分析

CONTENTS

1 中国具身数据市场发展背景与现状

2 中国具身智能数据产业生态与分析

3 中国具身数据行业典型企业分析

4 中国具身数据市场发展趋势预测

更多免费行业报告 并购家 www.ipoipo.net

01

中国具身数据市场发展背景与现状

中国具身智能产业发展阶段

具身智能产业2026年进入市场验证期，将逐步批量进入真实场景

- 中国具身智能产业经历了科研探索、产业起步、技术融合与场景示范阶段，在国家政策牵引、地方平台与产业集群协同推进下，多种形态的具身智能机器人已开启量产，预计2026年将进入市场验证阶段。
- 多个国家政策规划中，2030年被视为未来产业进入成熟期的分水岭，规划至2035年实现具身智能全面普及；此外，多位院士、专家在世界机器人大会及核心期刊指出，2030年是解决“泛化能力”和“低成本制造”两大瓶颈的关键年份；2030-2035年将是技术红利释放期，产业进入全面成熟与普及。

科研探索期

（~2010）

以高校和科研院所为主，聚焦基础理论与机器人样机研发，技术路线较为分散。

- 1990年，国防科技大学研制出中国首台两足步行机器人。
- 2000年中国独立研制首台具人外形、能模拟基本动作的人形机器人“先行者。

产业起步期

（2010-2021）

深度学习带动机器人感知能力提升，国内开始涌现一批机器人初创企业。

- 工业机械臂在汽车、3C电子等行业部署；扫地机器人、送餐机器人、配送机器人等开始进入商业和家庭场景。
- 优必选、宇树、云深处等专注人形机器人企业相继成立。

技术融合与场景示范期

（2022-2025）

大模型技术突破带动机器人发展，人形机器人创新体系初步建立，形成产业发展集聚区。

- 工信部出台《人形机器人创新发展指导意见》，计划到2025年，形成场景示范应用。
- 北京、上海、深圳等地建成具身智能产业园区。
- 部分头部机器人企业实现小批量交付。

市场验证期

（2026-2030）

政策与市场强力驱动，国家标准出台，产品开始批量进入真实场景，跑通PMF，并形成一批可复制的行业解决方案。

- 2026年为“十五五”规划第一年，具身智能为未来产业的重点培育方向。
- 2026年初，民政部、工信部等8部委部署机器人在家庭、社区、养老机构等场景的规模化应用。
- 2026年初，多家机器人企业宣布万台以上量产计划。
- 工信部计划到2027年，产业加速实现规模化发展，产品深度融入实体经济。

产业成熟期

（2030-2035）

具身智能泛化能力大幅增强，通用具身智能出现，具身产品普及。

- 国务院《十四五机器人产业发展规划》：到2035年，我国机器人产业综合实力达到国际领先水平，机器人成为经济社会发展的重要支撑。
- 工信部《人形机器人创新发展指导意见》：到2035年，人形机器人规模化应用广泛普及，在家庭服务、制造业、公共服务等领域形成成熟生态，成为新的经济增长引擎。

来源：网络公开信息，由易观分析整理

更多免费行业报告 并购家 www.ipoipo.net

中国具身智能产业核心优势

制造能力、产业链协同与市场驱动构建核心优势

制造业优势：

研发与制造能力推动成本下降，实现规模量产

- 制造业已形成高度成熟的模块化生产体系，能够在设计尚未完全固化的情况下迅速推进试制、迭代与小批量量产。这种“边设计、边制造、边优化”的能力，使机器人企业可以在实际交付中持续降本。
- 制造端对结构优化、材料替代、供应商整合的持续压缩能力，使机器人单机成本持续下降。

产业链协同优势：

供应链闭环，产业集群推动形成全产业链能力

- 具身智能产业链完整，在执行器、减速器、丝杠、电机、传感器、控制系统等核心硬件领域形成国产化闭环。机器人整机企业可以在本土完成大部分关键部件的采购与协同开发，大幅缩短研发周期并降低不确定性风险。
- 新能源汽车产业链与机器人产业链高度重合，可实现技术复用和加速发展。

市场驱动优势：

场景反哺技术，订单拉动迭代

- 机器人应用场景需求扩张，工业自动化、商业服务、医疗康复等领域对柔性自动化需求激增，促进市场扩张。订单开始直接拉动技术迭代，随着部分工业用人形机器人获得数亿元级订单并进入量产，带动了上游关节模组、电池、灵巧手、高精度传感器等部件的快速放量和技术突破。

智元机器人



智元机器人2025年全年出货量突破5000台，三大产品线（轮式人形、四足、人形）均超过1000台。

银河通用



银河通用与其产业园区中的核心零部件企业、工业制造企业联合开发，推动传统汽车零部件产业向机器人关键部件延伸。

宁德时代



宁德时代投入运行全球首条人形具身智能机器人规模化的新能源动力电池PACK生产线。

中国具身智能产业核心挑战

当前仍面临性能、数据、应用适配等多重发展瓶颈

单项能力已具备基础，但高性能本体部件制造仍面临制造难题	• 具身智能对力矩密度、功率密度、控制精度、轻量化与长续航的综合要求更高，需要在刚柔耦合传动、高力矩密度减速器、高功率密度电机、多模态传感与高能量密度动力系统等方面实现系统级优化。 • 当前单项能力已有基础，但围绕具身智能场景形成专用化、模块化、标准化的解决方案能力仍面临挑战，高端性能与长期可靠性将成为下一阶段竞争重点。
资本节奏与技术周期错配，导致部分企业机器人的展示大于真实场景应用能力	• 具身智能机器人的迭代逻辑更接近工业装备，需要通过长期场景运行与可靠性验证逐步优化，而部分资本预期仍强调短周期成果呈现，客观上加大了创业公司的展示压力，但真正决定商业化落地的稳定性、可维护性与成本控制能力仍需时间沉淀，如何在资本支持与工程节奏之间形成更为匹配的发展路径，是行业保持理性增长的重要前提。
高质量、大规模的动作与行为数据仍处于稀缺状态，泛化能力不足	• 机器人在单一任务中表现较好，但当环境变化或多任务叠加时成功率明显下降，反映出真实物理世界数据覆盖度不足。虽然仿真平台可在一定程度上弥补数据不足，但仿真与真实环境之间仍存在差距，未来竞争将不仅体现在算法架构本身，还体现在能否高效获取真实场景数据、构建持续学习机制并实现跨场景迁移能力。
人形机器人在技术路线、商业化模式、应用场景等方面尚未成熟，特别是在家用场景处于早期验证阶段	• 家庭场景对安全性、任务多样性和人机协同提出更高要求，但人形机器人在复杂开放环境中的泛化能力、大小脑协同能力、推理能力仍面临短板，短期内更有可能以局部功能切入，全面规模化应用仍需经历较长的技术与商业验证周期。

更多免费行业报告 并购家 www.ipoipo.net

具身智能数据本质特征

具身数据的本质在于“物理世界交互的多模态记录”，其核心特征表现为物理交互性（力/触觉/形变）、多模态强耦合（时空对齐）与时序连续性（状态-动作轨迹）。与互联网文本/视频数据的根本差异在于，具身数据需绑定真实物理属性，这使其从训练的“副产品”转变为产业的“核心燃料”。当前产业正经历从“硬件能力驱动”向“数据规模与闭环驱动”的转型，但高质量、可跨本体复用、可评测的物理交互数据（尤其是失败样本与接触细节）稀缺，仍构成规模化落地的首要瓶颈。

物理交互性

必须包含力、触觉、形变、摩擦等真实物理属性，互联网已有数据不通用

多模态强耦合

视觉、本体姿态、触觉力觉、音频、语言指令需在同一时间轴上精确对齐



时序连续性

以 (状态, 动作) 序列构成轨迹，是行为克隆与模仿学习的核心载体

市场缺失可用数据，制约具身智能的整体行业发展

具身智能的硬件门槛正在被快速跨越——本体量产、成本下降、运动控制逐步成熟，产业已初步解决“机器人能不能动”的问题。然而，真正决定产业天花板的，已不再是硬件能力，而是“数据能不能喂饱模型”。当前数据瓶颈已从“总量不足”演化为“量、质、异构”三个维度的结构性制约：三者互为因果、彼此强化，具体分析如下：

量：总量断档与结构性稀缺

国际先进技术应用推进中心指出：要实现具身智能能力的“涌现”，至少需要百万小时级来自真实世界的物理互动数据，而目前行业积累的数量还不到需求的5%。

根因在于“高质量、大规模、低成本”不可能三角未破：真机数据质量高但单小时成本500-1000元；无本体数据便宜但需重定向、缺力觉；仿真数据无限但存在Sim2Real偏差；互联网视频规模大但无动作标签，不是不想采，而是采不起、采不动、采了也不对



质：Sim2Real鸿沟与失败样本缺失

斯坦福HAI《AI Index Report 2026》显示，机器人操控在仿真中成功率达89.4%，在真实家庭场景中骤降至12%，形成77个百分点鸿沟。

“质”的瓶颈体现在三方面：仿真与真实的物理偏差导致“仿真成功、真机失败”；行业数据集“过度干净”，系统性丢弃失败样本与长尾场景，模型泛化靠运气；数据质量缺乏评估标准，标注粒度不统一，不可审计、不可回归，导致“脏数据进、脏模型出”

异构：跨本体不通与标准缺失

数据被硬件锁死，资产无法沉淀。本体异构——各家关节数、坐标系、控制频率不同，一家本体采集的数据无法直接用于训练另一家本体的模型，换本体等于重采、重标、重训；格式孤岛——JSON、Parquet、ROS bag等格式并存，跨平台不可拼接、不可审计；数据孤岛——全国各训练场数据尚未连接，核心高价值数据因护城河与合规顾虑难以流通，后果是数据无法跨本体、跨平台流动，无法形成公共数据底座

更多免费行业报告 并购家 www.ipoipo.net

破局数据瓶颈：以“造治用通”协同驱动具身数据闭环进化

数据瓶颈的破局，需紧扣“量、质、异构”三维制约，构建“造、治、用、通”协同推进的系统工程：以“造数据”扩大规模供给，以“做对数据”筑牢质量根基，以“转飞轮”驱动闭环进化，以“通数据”打破本体壁垒。 唯有四管齐下，才能终结“采得少、训不好、用不了”的低水平循环，推动具身数据从碎片化的“企业私产”跃升为标准化的“行业基础设施”，如下：

造数据：本体厂与仿真商协同，解决“量”的供给

本体厂商依托量产机器人规模，在真实场景中持续采集物理交互数据，以真机数据锚定质量基线；仿真厂商构建高保真物理引擎与场景资产库，以合成数据覆盖长尾与极端场景，实现低成本放量；两条路线互补——真机数据保真、仿真数据扩边界，共同扩大高质量数据的总供给。

做对数据：标注与评测体系，解决“质”的保障

数据从“采集完成”到“可训练可用”，需经过标注、清洗、质量评估等环节。建立统一的标注规范（任务级/片段级/关键帧级/接触级）与质量认证标准，明确“什么是合格数据”，并通过自动化评测工具对数据可训练性进行量化评估，确保进入模型的数据是“干净、对齐、可回归”的。

转飞轮：训练场与数据闭环，解决“数据—模型—本体—场景”的循环

训练场作为数据与模型的“中转站”，将采集数据转化为模型能力，再将模型部署到本体上验证，执行中产生的新数据回流至数据集，形成“采集→训练→部署→回流”的闭环飞轮。训练场不仅提供规模化采集能力，更是数据资产持续增值的核心枢纽。

通数据：交易与标准体系，解决“异构”的壁垒

建立跨本体、跨平台的数据格式标准与接口规范，使不同来源的数据可拼接、可互认；通过数据交易平台与合规机制，在保护隐私与商业机密的前提下促进高价值数据流通；推动公共数据底座建设，让数据从“企业资产”变为“行业基础设施”。

Analysys 易观分析

02

中国具身智能数据产业生态与分析

更多免费行业报告 并购家 www.ipoipo.net

中国具身智能数据价值链产业图谱2026

Analysys 易观分析

行业客户（需求侧）

BYD · 国家电网 STATE GRID · NIO 蔚来 · CATL 宁德时代 · 小鹏 · LUXSHARE ICT · FOXCONN 富士康科技集团 · SF EXPRESS 顺丰速运 · CAINIAO 菜鸟 · YH 永辉超市 · 美团 · Haier

整机制造

UNITREE · 智元机器人 · GALBOT 银河通用

具身模型

NOEMATRIX · 星海图 GALAXEA AI · Robbyant 蚂蚁灵波科技

开发平台

BAAI 智源研究院 · X-HUMANOID 北京人形机器人创新中心 · 华为云 · 商汤

系统集成

中国移动 China Mobile · ZTE 中兴 · iSoftStone · Trust 百信

具身智能数据价值链核心+生态

全链路数据平台

觅蜂科技 MANIFORMER · 光轮智能 · 无问智科 WUWENAI · 库帕思 · X-HUMANOID 北京人形机器人创新中心 · 华为云 · 阿里云 · 腾讯云 · 百度智能云

真机遥操作

智元机器人 · UNITREE · UBTECH 优必选 · DAIMON ROBOTICS 戴盟机器人 · Spirit AI 千寻智能 · 智平方 AI ROBOTICS · ZERITH 零次方 · 灵心巧手 LINKERBOT · X SQUARE ROBOT 自变量机器人 · FOURIER 傅利叶 · PaXini 帕西尼感知科技 · SEER 仙工智能 · ROBOTERA 星动纪元 · 逐际动力 LIMX DYNAMICS

便携/无本体采集

GenRobot · LUMOS ROBOTICS 鹿明 · 觅蜂科技 MANIFORMER · PsiBot 灵初智能 · DEXTELEOP · NOITOM 诺亦腾 · SIMPLE·AI · 青瞳视觉 CHINGMU · 艾欧智能

仿真与数据合成

光轮智能 · 无问智科 WUWENAI · 群核科技 MANYCORE · 大晓机器人 · 极佳 GigaAI · sudo · SIMPLEXDYNAMICS · 枢途科技 SYNAPATH · 跨维智能 DEXFORCE · 松应科技 ORCA · 深度机智 DeepCyber · FYSICS 飞诗科思智能科技

数据治理/标注

海天瑞声 · 云测数据 · 数据堂 Datatang · 景联文科技 JingLianWen Technology · 标贝科技 DataBaker Technology · 龙猫数据 LONGMAO DATA · BODENAI · 如祺出行 ON TIME · MindFlow · 整数智能 MOLAR INTELLIGENCE · 倍赛 BasicFinder

数据交易/流通平台

觅蜂科技 MANIFORMER · 库帕思 · 公象智能 · 百度智能云 · 北京国际大数据交易所 Beijing International Data Exchange · 上海数据交易所 SHANGHAI DATA EXCHANGE · 深圳数据交易所 SHENZHEN DATA EXCHANGE

训练场

X-HUMANOID 北京人形机器人创新中心 · 国地共建人形机器人创新中心 · 北京石景山人形机器人数据训练中心

多模态传感补维→采集硬件/传感器

PaXini · DAIMON ROBOTICS · XENSE 千觉机器人 · ENSURING TECHNOLOGY · 奥比中光 ORBBEC · LEANSTAR · MATRIX · 他山 TASHAN

算力与云服务商

阿里云 · 腾讯云 · HUAWEI 华为云 · 火山引擎 · 移动云 · 联通云 · 天翼云 State Cloud · 百度智能云 · DataCanvas 九章云极 · 商汤 · INFINIGENCE 无问芯穹

具身智能数据产业商业模式

具身智能数据行业呈现六类商业模式并存格局：数据集产品销售、定制化数采项目、数采硬件销售/租赁、仿真SaaS订阅、数据要素交易、全链路服务与训练场运营。当前仿真SaaS最为成熟，数据采集快速成长，数据交易与训练场仍处于极早期，如下：

维度	定制化数采项目	数采硬件销售/租赁	仿真SaaS订阅	定制化服务项目（治理/标注）	数据集产品销售/数据要素交易	全链路/训练场
计费方式	按小时/条/场景定向；高客单，交付周期1-8周	设备售价/租金 + 数据回购/授权费	平台授权年费 + 按调用量/Token/测试里程	按标注工时/条数/复杂度；人力密集型项目制	数据集：按套/时长（可复售）； 交易：佣金+授权费/入股	场地租用 + 设备共享 + 一体化服务采购
边际成本	高（依赖人力与本体，收入难复制）	中（硬件有制造成本，数据回流摊薄）	极低（合成数据，规模效应强）	高（人工/半自动标注，定制化程度高）	极低（标品复售）/ 趋零（平台交易）	中（重资产折旧，固定成本变可变）
成熟度	成熟（真机保真，当前主流）	成长（无本体轻量扩量）	最成熟（物理AI操作系统之争）	成熟（传统标注延伸，具身要求更高）	成长 / 极早期（交易规则与权属待完善）	起步（重资产，回报周期长）
核心壁垒	场景稀缺性与本体绑定	硬件专利与动捕精度	引擎精度与生态标准	标注规范与自动化工具	数据确权、合规与稀缺数据集	物理空间与全链路整合能力
典型企业	智元/觅蜂、宇树等	鹿明、帕西尼等	群核、光轮、无问等	海天瑞声、景联文等	百度、库帕思、北数所等	国资创新中心、云厂商等

更多免费行业报告 并购家 www.ipoipo.net

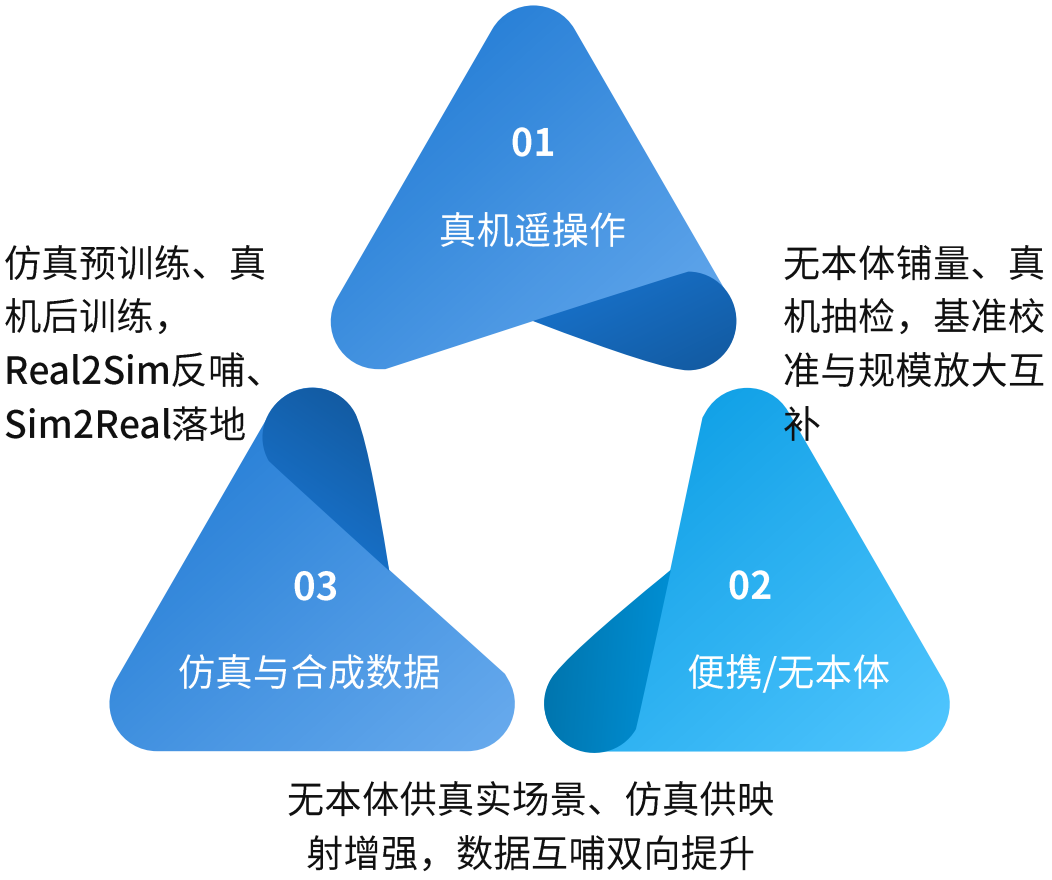
数据采集是具身数据价值链的源头，当前呈现三线并行格局

数据采集是具身数据价值链的“物理入口”——机器人从真实物理世界获取可学习信号的第一环。2026年，产业正从“单点真机遥操作”走向“真机基准 + 无本体规模 + 仿真增广”三维融合。其中，互联网视频蒸馏仅用于预训练阶段，不构成物理采集，故本报告重点围绕前三类路线展开分析，如下：

技术路线	真机遥操作	便携/无本体采集	仿真与数据合成
定位	提供最高质量的物理真实数据，是模型后训练（SFT）与最终落地的“金标准”	脱离完整机器人本体，通过可穿戴设备（手套、外骨骼、动捕）采集人体动作，实现长尾场景与规模化数据获取	在虚拟物理引擎中生成数据，解决真实世界难以覆盖的极端/多样化场景，实现数据增广与评测闭环
技术特征	依赖真实机器人本体（如人形、双臂），通过VR、动捕、力反馈手柄等方式由人类远程操控，记录本体状态、视觉、力觉等全量数据	人穿戴设备直接记录动作，成本降至真机的1/3左右；结合众包模式，走向“伴随化采集”（普通员工边工作边采）	物理引擎（刚/柔/流体）+域随机化+合成数据生成，遵循“99%合成+1%真机校准”的Sim2Real逻辑，常延伸至“研发-仿真-训练-验证-部署”全链路
数据质量	最高，具备真实的物理交互、摩擦力、碰撞动态，无需复杂的跨本体映射	中高，场景真实，但存在“人体→机器人”的跨本体重定向鸿沟（视角、关节结构、动作空间差异），需算法弥补	可控，物理规律可设定，但存在Sim2Real迁移鸿沟，需真机数据反哺校准
成本与规模	成本最高（500-1000元/小时级），受限于本体数量、场地与人力，边际成本难降，规模扩展性弱	成本适中（200-400元/小时），设备国产化（如简智订单破万）叠加众包网络，产能易进入百万小时级，规模性强	前期物理引擎研发与场景资产构建投入高，一旦引擎搭建完成，边际成本趋零，数据产量近乎无限，泛化能力极强
适用场景	复杂精细操作（灵巧手）、最终场景验证、模型微调基准	大批量日常操作预采集、长尾场景补充、动作轨迹快速录制	大规模预训练增广、极端/危险场景模拟、算法评测与验证

三大路线协同发展与下一步趋势

三条路线并非孤立竞争，而是围绕“数据质量—规模—成本”三角，形成前后递进、相互校准、动态平衡的协同体系，围绕数据金字塔各司其职：真机锚定质量、仿真放大规模、无本体补充真实。



下一步趋势

趋势①：虚实融合闭环是产业共识的终局范式，但成本曲线的动态变化将持续重塑各路线占比——协同是常态，比例是变量，如下：

- 当前：仿真边际成本趋零，但前期研发高；无本体规模化后单小时成本已低于仿真（部分场景）。
- 趋势：随着无本体设备国产化+众包网络成熟，真实数据成本持续下探，部分场景下“真实众包”可能比“仿真合成”更便宜。
- 影响：协同比例将动态调整——仿真侧重“真机难以覆盖的边界场景”，无本体承接“可规模化采集的常规场景”，真机专注“高价值垂直场景”。

趋势②：硬件入口化：设备即数据入口，形成“卖设备+持续买数据”闭环

趋势③：标准化与标品化：数据从定制走向资产

*仿真需要购买数字资产、建模、调用 GPU 渲染，而随着无本体设备规模化运营，**单小时真实数据的成本已经低于仿真**。*

——觅蜂科技创始人/CEO 姚卯青

更多免费行业报告 并购家 www.ipoipo.net

仿真与数据合成发展最为成熟，开启“物理AI操作系统之争”

仿真与数据合成



仿真与数据合成是当前六类商业模式中成熟度最高的板块，开启“物理AI操作系统之争”：

- 头部企业已形成“仿真引擎+场景资产库+合成数据管线”全栈能力
- 商业模式以SaaS订阅为主（年费+按调用量+一体机硬件），边际成本趋零，单条真机演示可生成百倍合成数据
- 合成数据成本可降至真机1/10以内，是当前突破“百万小时级”瓶颈的关键引擎

发展驱动力

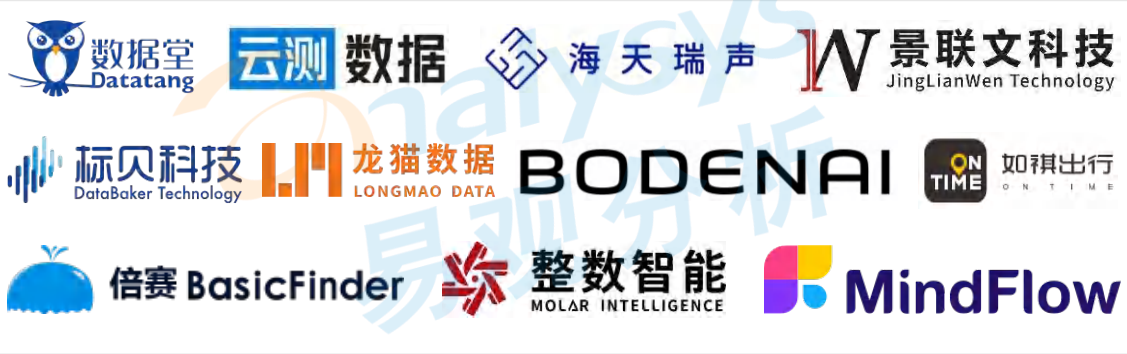
- 突破量的瓶颈：真机采集成本高、速度慢，仿真以近乎零边际成本覆盖长尾与极端场景
- 物理引擎成熟：高保真物理仿真（摩擦、形变、接触响应）精度持续提升，缩小Sim2Real鸿沟
- 世界模型兴起：无问智科等将“世界模型×数据工厂×世界模拟器”结合，仿真从数据生成器升级为模型训练的“数字孪生底座”
- 标准共建：头部参与国际开源物理引擎（如Newton）标准共建，生态壁垒加速形成

下一步趋势

- 从“卖仿真软件”到“卖物理AI底座”：头部企业正从工具提供商升级为"数据—仿真—评测—部署"全栈基础设施
- 虚实深度融合：真机微调+仿真预训练的联合方案（Sim2Real）成为标配，闭环反哺仿真精度
- 参与标准定义：通过开源引擎、数据接口、评测基准（如RoboFinals）构建CUDA级生态锁定
- 算力绑定：仿真对GPU需求巨大，云厂商（华为云、阿里云、腾讯云、百度智能云）通过"仿真+算力"一体化切入

数据治理/标注是“做对数据”的核心保障，当前处于从传统AI标注向具身专业化转型阶段

数据治理/标注



数据治理/标注是“做对数据”的核心保障，当前处于从传统AI标注向具身专业化转型阶段：

- 代表企业从语音/视觉标注延伸至具身数据的多模态对齐
- 商业模式为定制化服务项目制，按标注工时/条数/复杂度计费，人力密集型
- 需处理视觉、本体、触觉、语言等多模态的时间对齐、坐标系转换与质量评分，复杂度远超传统标注

发展驱动力

- “质”成为瓶颈：量的缺口逐步缓解后，数据质量（对齐、标注粒度、可训练性）成为模型性能的关键
- 自动化工具升级：“AI预标注+人工审核”模式降低人力成本，提升效率
- 标注规范沉淀：行业逐步定义任务级/片段级/关键帧级/接触级等标注粒度标准，提升可复用性

下一步趋势

- 半自动与合成标注融合：利用仿真数据的精确标注反向辅助真机数据标注，形成协同
- 从项目制走向平台化：输出自动化治理工具（时间对齐、轨迹重建、质量评分）以SaaS赋能客户
- 参与标准制定：标注规范的行业标准话语权成为新壁垒，头部通过定义规范获取规则优势

更多免费行业报告 并购家 www.ipoipo.net

数据交易是“通数据”的底座，当前处于极早期阶段

数据交易/流通平台



数据交易是“通数据”的底座，当前处于极早期阶段：

- 交易平台初现（觅蜂数据商城、库帕思、百度数据平台，及北京/上海/深圳数据交易所）
- 商业模式为数据集产品销售 + 数据要素交易（佣金+授权费+合规服务费）
- 核心高价值数据仍闭源，交易规模小、标准缺失

发展驱动力

- 数据要素化政策：国家推动数据作为生产要素流通，数据确权、评估、交易制度逐步完善
- 平台网络效应：数据积累到一定规模后，标准化交易可降低全行业采集成本
- 合规机制成熟：隐私脱敏、权属界定技术（如联邦学习、可信执行环境）为流通提供前提

下一步趋势

- 从私产走向要素：标准化数据集逐步标品化，高复售率数据成为“数字资产”
- 合规先行：交易所与平台通过确权、脱敏、合规评估建立可信机制，合规能力成为准入门槛
- 使用权与所有权分离：“授权使用+收益分成”模式替代一次性买断，平衡开放与保护
- 国资主导公共底座：国资平台承担公共数据底座与标准制定，推动互联互通

训练场是“转飞轮”的物理载体，当前处于起步阶段

70+

座训练场已建成启用，另有一座 40+ 座在建或规划中

23+	2,500+	28 万㎡	13+
覆盖省份	已部署机器人	总面积	部署 100+ 机器人的中心

已投用训练场区域分布（座）



省级分布：浙江 12 · 江苏 11 · 广东 10 · 北京 8 · 上海 8 · 四川 5 · 安徽 5

发展驱动力

- 降低行业试错门槛：共享训练场向中小企业提供设备与场地，降低数据采集的固定资产投入
- 闭环数据飞轮：采集-训练-部署-回流的闭环使数据资产持续增值，形成复利效应
- 政策新基建：国家及地方大力建设公共训练场，作为具身智能“算力+数据”基础设施
- 规模效应：集中化采集与训练显著提升效率，京东等通过社区化采集快速放量

下一步趋势

- 从单点服务向生态平台：头部（国资中心、云厂商）从单一采集走向全链路运营，构建数据飞轮
- “训练场+服务商”双重属性：兼具数据采集与平台服务，形成规模壁垒
- 标准化与互联互通：跨训练场数据标准逐步统一，推动形成全国公共数据底座
- 云化与算力绑定：云厂商通过“训练场+算力+存储”一体化切入，提供底层支撑

更多免费行业报告 并购家 www.ipoipo.net

全链路数据平台是具身智能数据价值链的终极整合形态

Analysys 易观分析

全链路数据平台



全链路数据平台是当前具身智能数据行业生态位最高、整合度最强的环节。如图所示，参与者呈现“垂直数据商+云大厂+创新中心”的多元格局：

- 垂直头部数据商（觅蜂科技、光轮智能、无问智科）：具备从采集（真机/无本体/仿真）、治理标注到数据流通的全栈能力，是前期“三线融合”的核心推动者
- 云厂（华为云、阿里云、腾讯云、百度智能云）：提供底层算力、存储与云服务底座，正向下游延伸，试图将平台变为云服务的“数据入口”
- 产业生态方（北京人形机器人创新中心）：推动公共数据底座与标准制定

整体上，行业不再满足于交付静态数据文件，头部玩家正构建“采-存-算-标-训-评”一体化平台，试图成为具身智能的“数据操作系统”。

发展驱动力

- 打破不可能三角：三线融合动态重塑成本曲线，解决数据匮乏
- 飞轮闭环：打通“数据回流—模型训练—部署验证—新数据采集”，实现持续迭代
- 硬件入口+标品化：卖设备锁定源头，标品高复售摊薄成本

下一步趋势

- 生态平台化：垂直数据商（觅蜂/光轮）与云大厂深度绑定，云厂商提供算力底座，数据商提供“数据+算法”闭环，共同抢占“底层基础设施”入口
- 标准定义者胜出：平台竞争终局是“标准竞争”，参与开源物理引擎、定义数据格式与评测基准的平台，将获得极高转换成本，实现“通数据”的行业级打通
- 公私数据协同：构建“公共数据底座+企业垂直定制”双层结构——公共平台（如北京人形机器人创新中心）推动通用场景标品化，高价值垂直场景仍由真机全链路深度定制

Analysys 易观分析

03

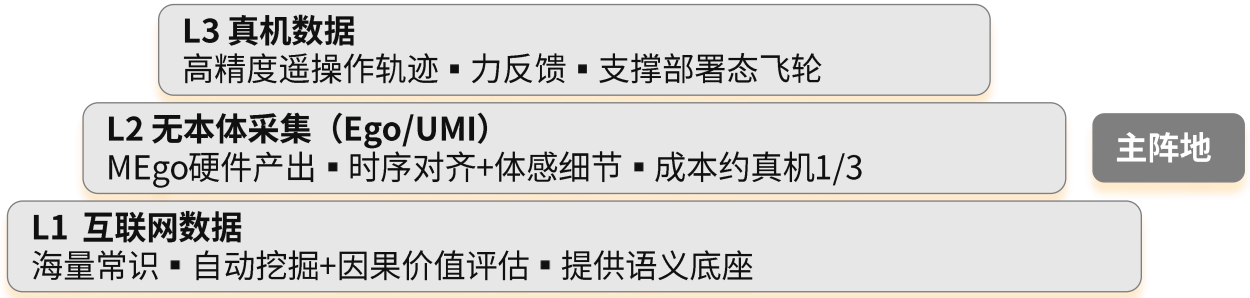
中国具身智能数据环节典型企业分析

更多免费行业报告 并购家 www.ipoipo.net

典型企业分析 - 觅蜂科技

觅蜂科技2026 年 2 月由智元机器人内部孵化拆分成立，是智元「一分为四」战略（灵巧手→临界点、数据→觅蜂、租赁→擎天租、四足→智元酷拓）中的数据业务主体。创始人姚卯青判断“模型无强壁垒，数据才是壁垒”，智元据此将数据业务独立运营，面向全行业提供数据基础设施。成立半年内完成三轮数亿元融资，红杉中国、国方创投、中国电信先后领投，覆盖市场化基金、上海国资与央企资本。

技术路线 ▪ 三层数据金字塔



数据闭环 ▪ Real2Sim2Real



产品体系与产品能力

硬件	**MEgo Gripper采集夹爪 ▪ MEgo View头戴感知** 毫米级轨迹重建 ▪ 亚毫秒全局同步 ▪ 超300°全景+腕部特写 ▪ 无线快换
工具链	**MEgo Engine自动化处理管线** 自研VIO位姿算法 ▪ 空间重建 ▪ 长时序自动切分 ▪ 效率较人工提升10x
数据产品	**L1/L2/L3分层标准化数据集** 22大类场景 ▪ 1万余真实环境 ▪ 累计产出100万+小时
平台	**一站式物理AI数据服务平台** 面向全行业开放 ▪ 与中国电信共建云网一体化采集网络

运营情况

- 规模：MEgo 设备累计下线 2 万台，累计产出超 100 万小时数据，覆盖 22 大类场景、1 万余个真实环境
- 产能目标：2026 年千万小时级产能，2030 年百亿小时，当前累计 100 万小时 vs 年内千万小时目标
- 成本：无本体路线成本约为真机 1/3（行业参照约 300–400 元/小时），硬件自研 + 量产带来 BOM 下降空间
- 数据资产属性：场景覆盖广，是复售的基础

易观分析

- 优势：智元体系数据工程know-how直接继承，起点即壁垒；中国电信领投带来云网与合规上云资源，具备排他性；MEgo硬件2万台下线，成本与铺量可控；三层数据金字塔覆盖全类型，不依赖单一路线
- 挑战：智元体系外客户与复购有待验证；千万小时产能需9倍放量，执行风险高

典型企业分析 - 光轮智能

光轮智能于2023年1月成立，创始人谢晨博士（前英伟达/Cruise/蔚来仿真负责人），定位“物理AI的数据与评测基础设施”，聚焦仿真合成+人类第一视角（Ego）双主阵地。2026年3月完成10亿元A++/A+++轮，成为全球首个具身数据独角兽；5月蚂蚁集团领投，估值突破20亿美元（约150亿元），两个月翻倍；6月再获10亿元战略融资，估值与融资密度均居赛道首位。

技术路线 ▪ 双主阵地

仿真合成 ▪ SimFoundry
全球唯一自研高精度GPU物理求解器
求解-测量-生成三位一体

人类第一视角 ▪ EgoSuite
7国500+环境 ▪ 1万+任务
周产2万+小时示范数据

标准话语权 ▪ Newton仿真技术指导委员会唯一中国企业
EgoVerse人类数据委员会 ▪ 主导或参编20项国家及行业标准

数据闭环 ▪ Real2Sim2Real 四模块

EgoSuite 人类示范采集 → SimFoundry 仿真训练生成 → RoboFinals 仿真评测失败定位 → RoboStack 真机部署与回流

产品体系与产品能力

硬件	**多模态采集设备矩阵：VR / 外骨骼 / UMI夹爪** RGB-D+手部与全身位姿+触觉 ▪ NVIDIA AR/VR技术栈 ▪ Jetson Orin NX端侧推理
工具链	**SimFoundry全栈仿真平台** 自研求解器 ▪ 与NVIDIA共定SimReady标准 ▪ 支持PhysX/MuJoCo/Newton/ GENESIS
数据产品	**EgoSuite数据集 ▪ EgoSuite-Open100k 开源10万小时** 7类场景 ▪ 129类场景 ▪ 15,000采集任务 ▪ 全量脱敏授权 ▪ 部分数据集复售超10倍
平台	**RoboFinals评测+RoboStack部署编排** 评测基于NVIDIA Isaac La...

运营情况

- 规模：全球 7 国 500+ 环境、1 万+ 任务、周产 2 万+ 小时、累计交付 30 万+ 小时
- 复售倍数：部分数据集已向十余家客户复售，复售倍数超 10 倍
- 收入与订单：2026 Q1 新增订单 5.5 亿元，单季收入超 2025 全年
- 成本结构：仿真路线边际成本约为真机 1%，标品复售进一步优化毛利结构

易观分析

- 优势：标准制定者姿态，话语权即壁垒；Q1订单5.5亿元，数据资产化已被商业验证；自研GPU物理求解器，全栈不绑第三方引擎
- 挑战：真机数据非主阵地，L3高精度后训练环节可能被卡位；既卖数据又办评测，“运动员兼裁判”的中立性存疑

更多免费行业报告 并购家 www.ipoipo.net

典型企业分析 - 无问智科

无问智科成立于2022年，由刘盛翔（前百度Apollo测试负责人）携百度/清华团队创立，智驾“测试验证”基因。以“世界模型×数据工厂×世界模拟器”跑通Real2Sim2Real，推“无垠”基座与“无穹”模拟器。2024营收数千万元；2026年4月/7月/9月连融三轮，洪泰领投A轮，老股东超额追投；具身订单数亿元、客户覆盖/复购均超80%，闭环叙事+量化验证兼备。

技术路线 ▪ 双轮驱动

真实世界端 ▪ 无感野采
自研多模态设备捕获真实人类行为
Raw→Model-Ready效率提升500%-1000%

仿真世界端 ▪ 4D世界生成
SimReady 资产构建3天→5分钟（300x）
柔性物体相关性90% ▪ 接触误差＜1%

数据闭环 ▪ Real2Sim2Real

无感野采 数十万小时积累 → **无垠自动化** 清洗 ▪ 对齐 ▪ 重建 → **无穹仿真** 百万SimReady → **三级评测** 仿真 ▪ 硬件 ▪ 世界模型

量化验证 ▪ WAIC公开披露
在线强化学习使仿真成功率 9.7% → 79.8%｜迁移真实机器人后任务成功率 +50%｜篮子收纳 +20%
客户复购率超 80% ・ 订单同比增长超 3000%｜德清训练场 3000㎡ 室内 + 505 亩室外

产品体系与产品能力

硬件	**自研多模态数据采集设备矩阵** RGB-D+手部与全身位姿+触觉 ▪ NVIDIA AR/VR技术栈 ▪ Jetson Orin NX端侧推理
工具链	**无垠物理AI数据基座 ▪ 无穹世界模拟器** 采集管理 ▪ 标注管理 ▪ 合成泛化+世界重建 ▪ 可微分物理疫情 ▪ 训练基础设施
数据产品	**多源异构数据集（真实+合成） ▪ 百万级SimReady资产库** 覆盖物流仓储/工业制造/家庭服务/商业零售
平台	**可量化评测体系（三级）** 仿真预验证 - 物理测试 - 产线部署 ｜ 与拓斯达共建

运营情况

- 规模：累计沉淀数十万小时高质量真实物理交互数据，月度数据生产能力超 2 万小时
- 订单与增长：具身智能业务订单达数亿元，订单同比增长超 3000%，收入预计同比增速 1000%
- 客户质量指标：头部客户覆盖度超 80%，客户复购率超 80%
- 成本与产能：长三角（德清）训练场室内 3000 m^2、室外 505 亩、开放场景 937 km^2，日产数据上千小时；SimReady 构建从 3 天→5 分钟大幅降低仿真侧成本

易观分析

- 优势：效果验证强——RL提升9.7%→79.8%，真实任务+50%；双轮闭环（野采数十万小时+百万级SimReady），SimReady构建效率300倍提升；复购超80%、订单同比+3000%
- 挑战：规模与生态位势落后——无自有本体/硬件生态，数据规模数十万小时相对较弱

典型企业分析 - 群核科技

群核科技成立于2011年，创始人黄晓煌，酷家乐起家，2026年4月成为“杭州六小龙第一股”。纯仿真合成路线，卡位“空间数据”——5亿+结构化3D场景、4.8亿带物理参数3D模型，用十年数据管线把渲染空间转成物理正确三维表示，SpatialVerse 供具身仿真数据。财报仍靠酷家乐订阅（约97%收入，毛利率83.8%，2026H1营收4.05亿），是“家装SaaS现金流 + 空间智能期权”的典型样本。

技术路线 ▪ 存量资产驱动的纯仿真合成

5亿+3D空间场景	4.8亿3D模型	1800万套设计方案
十余年家装设计沉淀	约6000万可公开调用	获取成本已摊销

转换管线 ▪ 资产→具身可用数据

SpatialLM 视频重建场景布局 → SimReady参数化 质量 ▪ 摩擦 ▪ 弹性赋予 → IsaacSim渲染 OpenUSD多传感器RTX ⇢ 真机部署反馈 环节缺失

产品体系与产品能力

硬件	无采集硬件入口
工具链	SpatialVerse空间智能平台 物理正确合成数据集 ▪ 多模态鲁棒扩增 ▪ 支持亿次迭代训练 ▪ 自研启真渲染引擎
数据产品	InteriorGS / InteriorNet 室内认知数据集 80+室内环境 ▪ 全球最大室内场景深度学习数据集之一
平台	Aholo空间智能开放平台+空间大模型家族 SpatialLM ▪ SpatialGen ▪ Lux3D ▪ LuxReal ▪ 日均24亿token调用

运营情况

- 数据资产：5 亿+ 3D 空间场景、4.8 亿 3D 模型、1800 万套空间设计方案，存量资产，且获取成本已在过去十余年摊销完毕，边际成本极低
- 具身业务收入：SpatialVerse 2024 年客户 8 家、收入 340 万元；2025 年客户 16–18 家、收入 520 万元，占全年总营收不足 1%；2026 上半年订单金额 680 万元（超 2025 全年），占比仍低，海量资产尚未能转化为具身侧订单

易观分析

- 优势：十年沉淀5亿+场景、4.8亿模型，成本已摊销；83.8%毛利率+自研引擎+NVIDIA集成，仿真底座扎实；已上市，资本通道占先
- 挑战：具身收入年收入百万级、占比不足1%，故事与收入落差大；家居场景强但工业/物流弱，恰是付费主战场错位

更多免费行业报告 并购家 www.ipoipo.net

典型企业分析 - 鹿明机器人

鹿明机器人成立于2024年，喻超（前追觅/CyberDog）创立，曹俊亮（交大博士）管硬件，研发占70%+、十余位博士。UMI无本体采集绝对主阵地，FastUMI Pro做“人采—跨本体复用”的低价数据闭环。融资累计近10亿，2026年5月A1+A2共4.93亿由三菱电机连续领投，恒生/德马等产业方入局，拿产业资本换真实产线（质检/装配/分拣），走“无本体扩量+工业场景回流”路线。

技术路线 ▪ 单路线重注

FastUMI Pro
首款量产全功能数采系统

FastUMI Ego
即插即用 ▪ 无需建图

FastUMI Go
背包版 ▪ 计划投放1万台

核心指标：单条擦剂50s→10s（5x效率） ▪ 成本降至传统遥操作1/5（计入本体/100-1/200）
8道工业级质量评估：数据有效率70% → 95%+ ▪ 与本体解耦，适配市面90%+机械臂

数据闭环 ▪ 本体 + 数据双主线飞轮

数据产品输出
商城92个任务SKU

→

整机场景落地
进入三菱质检环节

→

真机数据回流
反哺Prime RO

产品体系与产品能力

硬件	**FastUMI 数采三形态 + 自研本体 LUS / MOS / 鹿小明 / LUD / NIX** 自研关节模组与视触觉模组
工具链	**8 道工序工业级数据质量评估体系** 有效率 70% → 95%+，仅交付可复现轨迹｜未形成对外售卖的工具产品
数据产品	**Lumos 数据商城（全球首个具身数据超市）** 92 个任务 SKU ・ 10 万+ 条轨迹 ・ 37+ 家机构采购｜1 万小时积累 vs 100 万小时目标
平台	**Lumos NexCore（2026-08 发布）+ 具身大脑 Prime R0** Prime R0 在 MolmoSpaces 零样本榜单位列 Top 1

运营情况

- 规模：已初步完成 1 万小时高质量真机数据积累与基座模型训练；数据商城 10 万+ 条轨迹；目标 2026 年建成100 万小时产能
- 成本：单条采集成本降至传统 1/5，计入本体成本则 1/100–1/200；数据有效率 95%+（行业普遍 70%）
- 产能与利用率：FastUMI Go 计划投放 1 万台；百万小时产能目标 vs 当前 1 万小时积累，存在 100 倍放量压力

易观分析

- 优势：优势：UMI采集单条50s→10s、成本1/5、适配90%机械臂，硬件效率硬指标领先；具备本体+数采+具身大脑，飞轮结构独特
- 挑战：技术路线延续存疑；成立不足两年铺开四条业务线，战线过宽；1万小时积累对100万小时目标差100倍，规模叙事弱

Analysys 易观分析

04

中国具身数据产业发展趋势

更多免费行业报告 并购家 www.ipoipo.net

技术发展路径

具身智能技术发展路径与数据能力建设沿三条主线演进：数据生产从单一走向融合，数据闭环从线性走向虚实共生，基础设施从软件工具走向软硬一体，如下：

三线融合

单一路线无法破解“高质量、大规模、低成本”的不可能三角。三线融合的本质，是根据任务难度与成本约束动态配置路线占比——仿真守边界、无本体扩规模、真机保质量，成本曲线变化将持续重塑配比，但多源协同作为生产范式本身不会逆转

物理引擎×生成式 AI

引擎的物理正确性是仿真数据能否迁移到真机的底线；生成式AI的自动化生成解决人工建模的规模瓶颈。两者结合，才使仿真从“人工搭建的有限场景”走向“算法生成的无限任务分布”，引擎精度决定数据下限，生成多样性决定数据上限。

无本体采集

以“人持设备采操作”绕开真机遥操作的高成本，在真实物理环境中获取人类操作数据，同时规避本体硬件依赖。定位介于真机与仿真之间：比仿真更接近真实物理交互，比真机遥操作成本低一个数量级

Real2Sim2Real

“真实采集→仿真重建→批量扩增→真机验证”的全链路闭环，正在取代“采集—标注—交付”的线性流程。真实数据提供物理基准，仿真负责批量扩增与长尾覆盖，真机验证反馈修正仿真误差，形成持续进化的飞轮，世界模型既是数据生产工具，也是质检与评测平台

少样本 / 自进化

模型训练范式正从“海量外购数据依赖”转向“少样本学习 + 强化学习自进化”，模型可通过自身部署反馈持续进化，而非无限依赖外部数据供给

仿真 × 算力

仿真正在从“可选的软件工具”演变为“软硬一体的基础设施”。GPU物理求解器、云平台、分布式采集网络、数据治理管线的协同，构成具身智能数据生产的基础底座，企业不再自建仿真能力，而是按需调用平台服务

数据采集 | 数据闭环 | 模型训练 | 基础设施

产业发展趋势

具身智能数据产业正经历三重结构性变迁：能力层竞争维度从“数据规模”转向“数据价值”，产业层权力格局从“分散供给”走向“分化重构”，终局资产形态从“数据商品”迈向“数据要素”。数据竞争的本质，从卖数据转向定义数据价值的标准。

能力层趋势（当下基石，企业竞争方式）

趋势①硬件入口化——入口壁垒取决于软件粘性

采集硬件成为数据主权前哨，“卖设备即买数据”模式初步成型。但硬件本身可被追平，真正的入口壁垒在于软件粘性——数据格式绑定、云端平台依赖、评测基准锁定

趋势②标准化标品化——从数据集销售走向能力交付

标注维度精细化、格式规范化、评测基准统一化，使数据从“原材料”变为“即用型能力”。竞争单位从“小时数”转向“有效信息密度”，高复售标品成为可持续商业模式

趋势③任务难度定价——从成本定价走向价值定价

任务难度决定稀缺性，稀缺性决定价格——高难任务（长时序、强泛化、复杂操作）溢价凸显，低难任务价格趋零。任务难度与数据价格的结构性挂钩，建立市场价值闭环

产业层趋势（中期演变，产业结构与权力分配）

趋势④买方自建与中立性分化——接口权力的争夺

头部本体厂商与基座模型公司正将数据生产内化，自建采集体系、数据管线与评测标准，以掌控接口权力。中立第三方（评测机构、开源社区、公共训练场）在服务长尾市场中寻找空间。买方自建决定接口权力归属，中立方的生存空间取决于头部玩家的自建边界

趋势⑤产能军备与结构性过剩——数据资产加速折旧

行业正经历产能军备竞赛，但结构性过剩已然显现：原始数据贬值、高难数据稀缺并存。生产权向有本体/场景的下游迁移，外购需求收缩；少样本/自进化范式进一步压缩外购空间，加速存量数据资产折旧。静态数据集价值窗口收窄，“有效信息密度 × 闭环增益”成为新价值标尺

终局：趋势⑥数据要素化 + 平台整合——超百亿平台型企业涌现

实景数据从“商品”走向“要素”，确权、定价、入表、流通制度逐步完善，实景实训专项与数据交易所提供落地路径，平台整合加速，超百亿平台型企业涌现，

更多免费行业报告 并购家 www.ipoipo.net

发展机会与空间

进攻性机会（抢标准）	整合性机会（绑场景）	平台性机会（建基础设施）	要素性机会（吃制度红利）
机会①第三方评测认证 当前“有能力≠有产品”的空白真实存在——既有评测机构既卖数据又办考场，中立性代价明显；部分厂商的评测体系停在内控，未产品化。谁能做出“不卖数据、只办评测”的独立第三方，谁就能直接吃掉既有玩家的中立性溢价	**机会②垂类场景联合工厂** “资产绝对量没有意义，场景匹配度才是兑现开关”。场景方有数据但不懂AI，数据商懂AI但进不去场景。设计出“数据归属清晰、收益分配明确”的合作框架，是率先跑通垂类场景联合工厂模式的关键 **机会③多源数据采集服务（真机 / 无本体 / 众包）** 数据缺口高，真实物理交互数据最稀缺，真机遥操、无本体穿戴、场景众包、工业一线操作，先有采集网络就有议价权。抓手：轻量化设备+场景准入+标注流水线	**机会④仿真与合成数据基础设施** “引擎保真 × 生成式扩量”是规模供给唯一可放大路径核心是物理正确性 + 任务分布覆盖 + 与真机闭环对齐 **机会⑤数据工程与“炼油厂”能力** 行业从“有数据”转向“有效信息密度”，格式孤岛和标注水数据是最大痛点。跨本体动作映射、多模态对齐、失败样本库、自动质检，抓手是标准化格式、可复售标品、任务难度分级 **机会⑥闭环训练场 / 公共实训底座** 中小企业不愿重资产自建，公共训练场把固定成本变可变成本，是区域新基建。政府+本体厂商+数据商共建，提供“采集—训练—验证—回流”全链路 **机会⑦中立平台与工具链** 头部买方自建挤压纯外包，但中小模型厂商和长尾本体仍需中立层。数据超市、跨本体训练平台、评测社区、开源物理引擎，抓手是开源评测 + 付费增值	**机会⑧数据要素流通与资产化（交易所 / 确权 / 入表）** 数据从“商品”走向“要素”，各地数据交易所与开源社区正在坐实制度基础设施。数据产权登记、质量评价、资产入表、场内交易。抓手：高价值垂直场景数据集、一登双证、合规流通

Analysys 易观分析

易观分析

中国数智化领域专业的科技和市场分析机构

- 网址：www.analysys.cn
- 客户热线：4006-010-231
- 服务邮箱：ygfx@analysys.com.cn



关注易观分析
获取更多报告

本产品保密并受到版权法保护

Confidential and Protected by Copyright Laws

更多免费行业报告 并购家 www.ipoipo.net