人形机器人系列专题之具身数据

物理AI核心资产，模型迭代孕育数据产业新蓝海

优于大市

核心观点

高质量物理交互数据是具身智能实现泛化的关键，具身Scaling Law初步显现。纵观数字AI大模型的发展历程，模型能力边界的突破始终伴随着数据、算力与算法的协同演进。数据是物理AI大模型发展的关键资源，是支撑模型训练、能力涌现与智能进化的底层驱动力。具身智能领域的Scaling Law已开始获得初步实证支持，如Dyna-2在Ego第一视角人类视频预训练数据的实验中发现，当数据规模从千小时扩展至百万小时，模型的表现呈现明显改善趋势，为具身Scaling Law提供初步证据，预计数据需求将持续扩大。

通用模型厂商加快入局具身智能，触觉数据稀缺性高。部分基础模型厂商数字AI业务ARR增速放缓，向物理世界的延伸有望成为其拓展应用边界的重要方向，具身数据采购需求迅速增长。以GPT-6 Astra为代表的多模态通用模型持续向物理世界拓展，在具身领域展现出较强的任务理解、规划与闭环纠错能力，进一步扩大了数据需求主体。当前具身数据积累以视觉信息为主，人类操作视频提供了丰富的行为与语义信息，触觉数据则需要在交互过程中同步记录接触状态，现有触觉数据的供给仍然相对有限。

具身数据采集路线多样，多源数据训练价值凸显。具身数据采集技术路线包括互联网视频数据、Ego第一人称视频数据、仿真合成数据和真实数据等类型，不同来源的数据具备互补价值。1）**互联网视频数据**规模庞大，是具身智能的数据底座，但缺乏力和触觉信息；2）**Ego第一人称视角数据**由采集者佩戴头戴式或颈挂式设备，在日常工作环境中自然完成任务，不依赖特定机器人本体，是规模化潜力较高、部署门槛较低的路线，未标注的Ego数据价格约几十元/小时，处理后的高质量数据几百元/小时，Ego设备则为千元级别；3）**UMI通用操作接口数据**由操作员手持UMI夹爪作为采集工具，采集末端操作信息，提供结构化动作信息，主要优势在于跨本体复用，不依赖特定机器人本体硬件，但动作空间与真实机器人有差距，数据价格几十到几百元/小时，核心装备万元级；4）**遥操作数据**提供真实机器人执行示范，是高价值的数据类型，主要用于模型后训练，数据价格为几百元至上千元/小时，触觉手套价格分化显著，国内厂商单套售价为万元级别；5）**仿真数据**支持高效率生成与场景扩展，但核心面临Sim-to-real Gap难题。

国内厂商积极布局数据采集与处理，具身数据有望实现百亿市场空间。具身智能数据产业链覆盖基础设施、数据采集、数据加工及应用等环节，具备软硬件协同、真实场景资源及持续数据交付能力的企业有望形成差异化优势。市场空间方面，展望明年，行业需求向亿小时级迈进，按数百元/小时价格测算，数据市场对应约百亿元市场空间潜力；长期看，借鉴特斯拉智能驾驶数据积累历程，在车辆持续在用、单车日均行驶1小时的假设下，测算到2024年FSD V13模型能力有明显突破时，特斯拉全球车队行驶时间测算约为60亿小时（测算值，仅供参考）。理论上，机器人的数据复杂度远高于智能驾驶汽车，若具身智能形成类似FSD V13重大模型能力突破，则数据需求大概率超越该时长，考虑到未来数据采集成本下降，远期数据市场规模也有望超千亿元，市场空间广阔。

风险提示：具身智能模型进展不及预期，具身机器人出货量不及预期，具身数据采购需求低于预期，数据技术路线变化，数据价格下降及行业竞争加剧。

行业研究·行业专题

汽车

优于大市·维持

证券分析师：唐旭霞
0755-81981814
tangxx@guosen.com.cn
S0980519080002

证券分析师：孙树林
0755-81982598
sunshulin@guosen.com.cn
S0980524070005

证券分析师：傅嘉良
0755-22940076
fujialiang@guosen.com.cn
S0980526100002

市场走势



资料来源：Wind、国信证券经济研究所整理

相关研究报告

《汽车行业2026年9月投资策略及中报总结——行业整体利润承压，Cybercab投入运营》——2026-09-15
《人形机器人本体系列之宇树科技招股书梳理》——2026-08-22
《汽车行业2026年8月投资策略-宇树机器人上市在即，关注汽车板块财报行情》——2026-08-20
《重卡行业专题——海外市场空间广阔，出海景气有望长期延续》——2026-08-20
《具身智能产业研究蓝皮书》——2026-08-02

内容目录

前言：数据成为具身智能核心资产与关键瓶颈 5
具身 Scaling Law 初现，数据成为物理 AI 迭代进化的重要资产 5
数据成为具身大模型迭代的核心瓶颈 7
基模厂商需求驱动具身数采加速扩容 12
具身数据：高质量触觉数据稀缺性高，采集与后处理成为核心环节 13
数据类型：视觉数据积累为主，高质量触觉数据稀缺性高 13
数据环节：数据采集与处理是具身数据关键环节 17
数据采集：技术路线多样，多源数据融合是大势所趋 19
互联网视频数据：数量庞大，构成具身数据底座 21
Ego-centric 数据：快速崛起的数据类型，规模化潜力大 22
仿真数据：采集效率及多样性高，但面临 Sim2Real 鸿沟 24
真实数据：价值高，采集成本高 26
市场格局：具身数据产业化进程加速 31
产业链：贯通上游硬件、中游采集、下游模型应用 31
商业模式：定制化服务构成主要商业形态 32
市场空间：以智能驾驶数据积累历程看具身数据市场空间 34
竞争格局：多元主体分工布局，数据质量与规模化交付构筑壁垒 35
风险提示 37

图表目录

图1： 数字 AI 大模型的 Scaling Law 5
图2： Dyna-2 在百万小时 Ego 数据中初步验证具身 Scaling Law 5
图3： 数据是具身智能技术发展的基础 6
图4： 数据、算法、硬件构成通用人形机器人三大基础 6
图5： 机器人数据飞轮 6
图6： 具身数据现存挑战 7
图7： 具身数据存量缺口 8
图8： 特斯拉智能驾驶数据测算 9
图9： 视觉与触觉数据互补程度高 10
图10： 视触融合比单模态的操作准确率高 10
图11： 多模态数据融合的整体过程 11
图12： GPT-6 Astra 展现机器人操作能力，与 VLA 协同提升复杂任务表现 13
图13： 触觉数据规模有限 15
图14： 直接加入触觉数据后可能反而导致模型效果变差 16
图15： 具身数据关键环节 17
图16： 相同数据量下场景多样性影响泛化效果 18
图17： 对采集的数据进行配比优化会影响下游任务效果 18
图18： 具身数据金字塔 19
图19： 具身数据采集技术路线对比 20
图20： 魔法原子的 Magic-W0 融合 EGO、UMI、真机与仿真四类操作数据进行预训练 21
图21： 枢途科技利用互联网视频数据构建具身数据集 22
图22： 枢途科技从基础 2D 视频中提取具身训练数据 22
图23： EgoMimic 研究显示 Ego 数据能够提高机器人操作能力 22
图24： EgoScale 研究显示 Ego 数据的增加能改善模型能力 22
图25： Ego-centric 数据采集过程 23
图26： 破晓智能全身 Ego 数据 24
图27： 破晓智能全身 Ego 数据采集及处理过程 24
图28： Sharpa Tacmap 通过触觉仿真训练实现灵巧手操作策略向真机的零样本迁移 25
图29： 光轮智能仿真数据 26
图30： 仿真数据存在 Sim-to-Real Gap 问题 26
图31： UMI 数据采集 27
图32： 智在无界 UMI 灵巧手 BeingBeyond U1 27
图33： 诺亦腾动捕数据采集方案 28
图34： GR00T N1 数据中有大量遥操作数据 29
图35： 遥操作数据采集过程 30
图36： π0.6 采用真机回流数据进行训练 31
图37： 具身智能数据产业链 32

图38： 觅蜂派 App 众包数据采集模式 34
图39： 觅蜂派 App 众包数据采集模式 34
图40： 特斯拉智能驾驶数据测算 35

表1： 具身机器人本体厂商数据情况梳理 8
表2： 具身数据来源多样 12
表3： 具身智能数据类型 14
表4： 不同原理的触觉传感器的对比分析 16
表5： 具身智能数据评价维度 19
表6： 主流具身数据模拟器对比 24
表7： 真实数据采集技术对比 26
表8： 国内具身数据主要玩家相关信息梳理 36

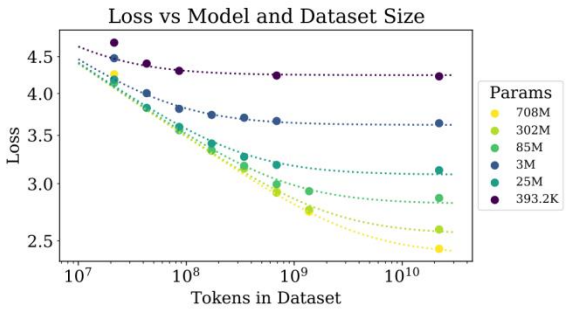
前言：数据成为具身智能核心资产与关键瓶颈

具身 Scaling Law 初现，数据成为物理 AI 迭代进化的重要资产

纵观数字 AI 大模型的发展历程，智能能力边界的持续拓展，始终伴随着数据、算力与算法的协同演进。大语言模型的发展证明了数据世界的人工智能存在规模定律(Scaling Law)，即随着数据规模、模型规模和计算资源的持续扩展，模型能力会不断提升。规模定律已经成为推动人工智能能力突破的重要底层范式，也让数据作为智能发展基础资源的价值愈发清晰。如今，随着人工智能从数字空间走向物理世界，通过积累更丰富的物理交互数据来提升机器人的感知、理解与行动能力，正成为具身智能发展的重要探索方向。

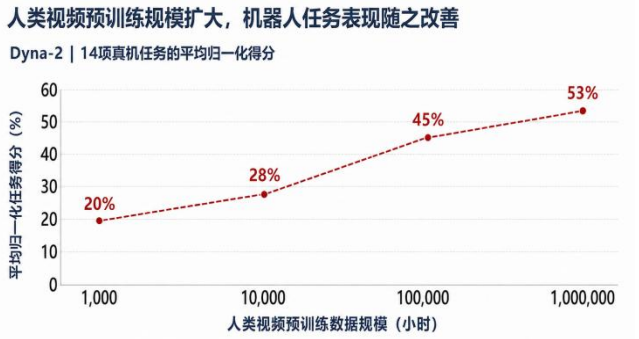
具身智能作为人工智能与多学科融合的前沿技术，其演进路径的探索主要建立在 Scaling Law 成立的假设之上，即数据驱动的通用学习范式在物理世界依然有效。尽管当前具身数据规模有限，但数据规模化的价值已开始获得初步具身领域的实证支持。Dyna Robotics 于 2026 年 8 月发布的 Dyna-2 研究报告显示，在第一视角人类操作视频预训练数据的实验中，当数据规模从千小时扩展至百万小时，模型在机器人数据上的离线预测表现，以及经过机器人数据后训练后的真机任务表现，均呈现改善趋势，为具身 Scaling Law 提供初步的证据，将持续拉动大规模数据需求。

图1：数字 AI 大模型的 Scaling Law



资料来源：Jared Kaplan 等，《Scaling Laws for Neural Language Models》，arXiv，2020，1-30，国信证券经济研究所整理

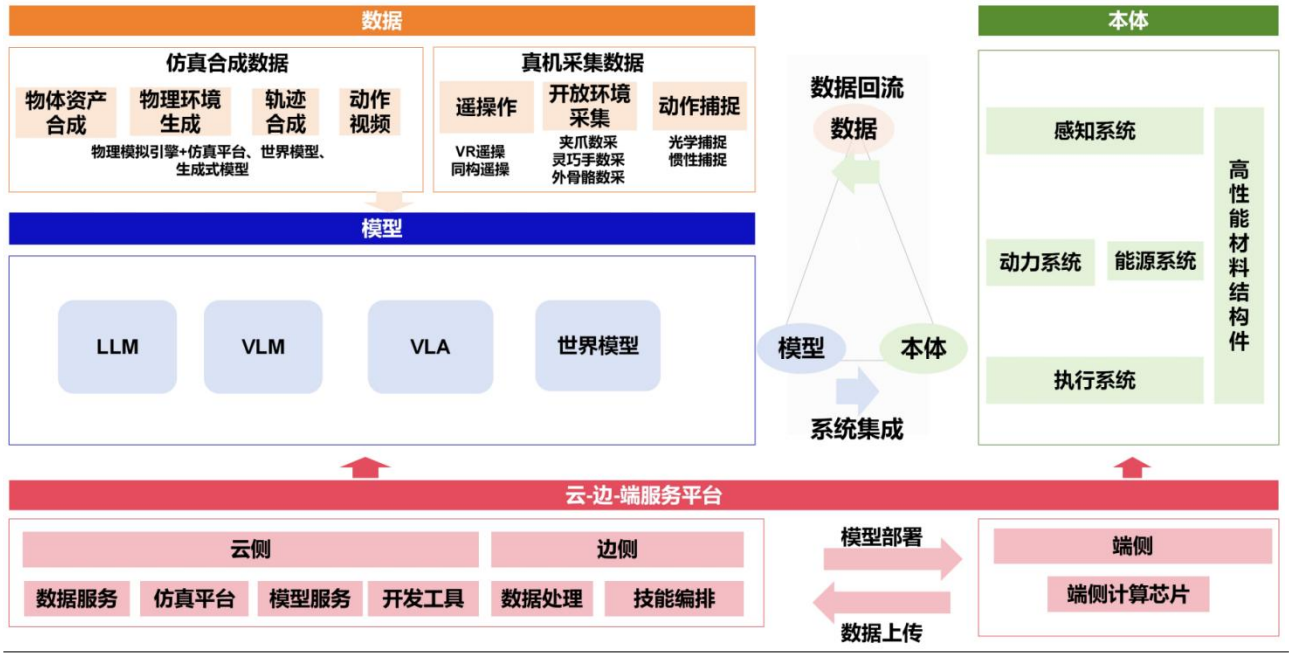
图2：Dyna-2 在百万小时 Ego 数据中初步验证具身 Scaling Law



资料来源：Dyna 官网，国信证券经济研究所整理

数据是物理 AI 大模型发展的关键资源，是支撑模型训练、能力涌现与智能进化的底层驱动力。物理 AI 需要机器人在物理环境的交互中形成自主进化的智能闭环：机器人本体被部署到真实场景中→在执行任务中收集真实行为数据→高质量数据被用于模型训练和迭代→更通用的模型被部署到更多的机器人上→收集更多、更高质量的数据，其中数据是闭环飞轮启动起点，支撑具身智能在现实环境中的有效运行、持续迭代。

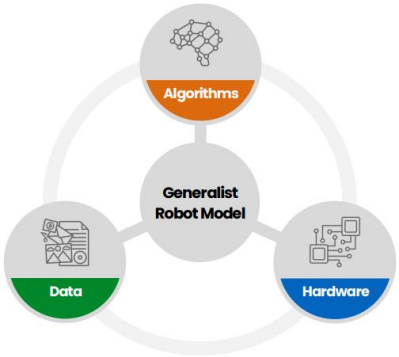
图3：数据是具身智能技术发展的基础



资料来源：中国信通院，国信证券经济研究所整理

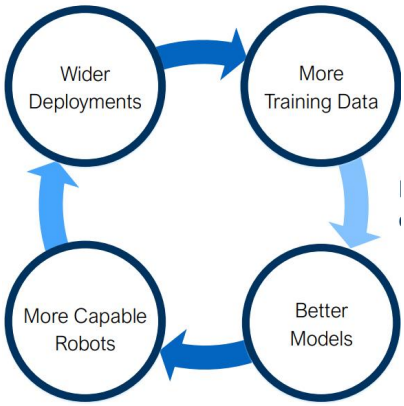
高质量、大规模数据是支撑具身智能模型发展的关键，数据预训练阶段的可扩展性初步得到验证。已有研究证明，大规模数据对具身基础模型的性能提升具有显著作用。清华大学 ManiBox 论文指出空间泛化所需数据量与空间体积呈现幂律关系，如果要在 38 倍的空间泛化内实现泛化，那么数据量需要扩展大约 38^0.35 倍。Generalist AI GEN-0 模型研究表明每增加 10 倍数据，模型性能也会成比例改进。

图4：数据、算法、硬件构成通用人形机器人三大基础



资料来源：Yuke Zhu，《Pathway to Generalist Robots: Scaling Law, Data Flywheel, and Humanlike Embodiment》，德克萨斯大学奥斯汀分校机器人感知与学习（RPL）实验室，2023:1-46，国信证券经济研究所整理

图5：机器人数据飞轮



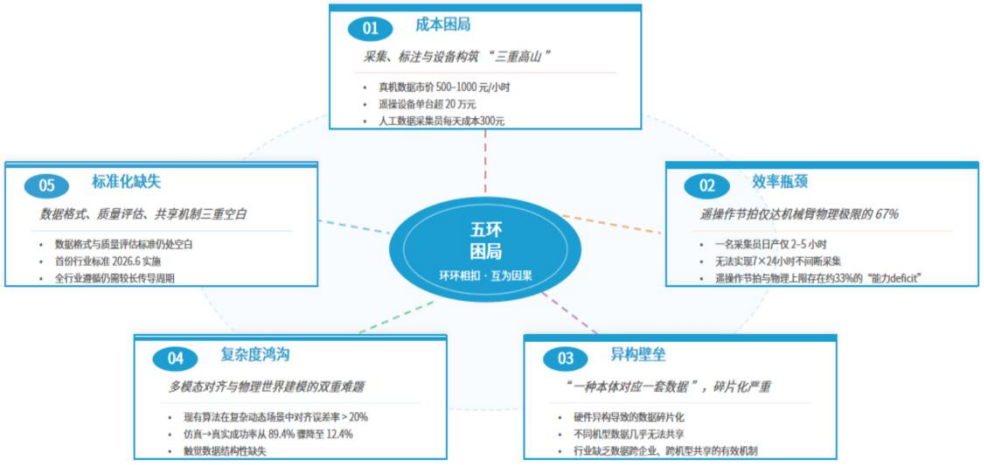
资料来源：Yuke Zhu，《Pathway to Generalist Robots: Scaling Law, Data Flywheel, and Humanlike Embodiment》，德克萨斯大学奥斯汀分校机器人感知与学习（RPL）实验室，2023:1-46，国信证券经济研究所整理

数据成为具身大模型迭代的核心瓶颈

数据是制约具身智能从“演示验证”走向“真实部署”的重要瓶颈。在具身智能的演进历程中，高质量数据被认为是跨越通用精细操作鸿沟的核心要素，随着硬件本体的逐步成熟，算法迭代的瓶颈逐渐转移至数据端。如何低成本、规模化地获取高质量的多模态数据，成为了未来影响具身智能大模型通用性和商业化落地的核心要素。当前行业尚未实现“数据-模型-本体-场景”的闭环，核心卡点在于真机数据采集尚未实现规模效应，合成数据质量不足以支撑可靠的行为学习和真实场景泛化。

根据亿欧智库，**当前中国具身智能数据面临多重制约，主要体现在数据供给、采集效率、数据兼容性、多模态处理及行业标准等方面。**①数据供给：具身智能模型对数据需求有较强的场景和本体依赖性，数据需要通过真实机器人操作、仿真环境生成等方式获取，并经过相应的标注与处理，导致数据采集和生产成本较高，规模化供给难度较大。②采集效率：数据采集流程复杂、效率有限，制约了数据积累速度及模型迭代进程。③数据兼容性：机器人硬件构型、传感器配置及控制系统存在差异，不同平台产生的数据在格式、维度和语义层面难以统一，跨平台迁移与复用面临较大障碍。④多模态处理：具身智能涉及视觉、触觉、力觉及本体状态等多种模态信息，数据之间具有较强的关联性，对采集设备、数据处理算法及专业人才提出了更高要求。⑤行业标准：行业尚未建立统一的数据采集、标注、处理及使用规范，不同研究机构和企业在数据生产与应用方面仍存在较大差异，同时也缺乏统一的模型能力评价体系和测试基准，不同技术路线难以开展客观、可量化的横向比较。

图6：具身数据现存挑战

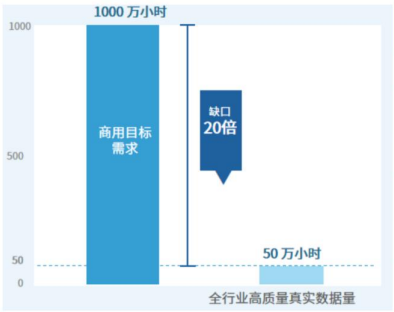


资料来源：亿欧智库，国信证券经济研究所整理

■ 数据需求量大、价格高

根据亿欧智库，具身智能要达到真正可用，至少需要 1000 万小时级别的真实场景交互数据，然而当前全行业高质量真实物理交互数据总量仅约 50 万小时，人类视频数据 100 万小时，数据缺口大。这一结构性稀缺源于具身数据的物理属性（无法从互联网直接获取，需要经过采集、仿真等高成本路径获取）使得数据成为当前产业发展的核心卡点。根据中国信息协会数据集和数据标注专委会等发布的《具身数据发展报告 2026》，具身智能真机有效数据的价格约 500~1000 元/小时，价格昂贵。

图7：具身数据存量缺口



模型或数据集 / 厂商	数据规模	数据类型
Dyna-2/Dyna Robotics	100万+小时	人类第一视角操作视频；部分带3D手部姿态及派生动作标签
GEN-1/Generalist	50万+小时	传感化的真实物理交互数据
Being-H0.8/智在无界	50万+小时	人类视频，以第一视角数据为核心
Psi-R2/灵初智能	10万小时级	人类手部操作数据，包含视觉、关节、语言、触觉等信息
LingBot-VLA 2.0/蚂蚁灵波	约6万小时	5万小时机器人轨迹＋1万小时人类第一视角视频
GigaBrain-0.7/极佳视界	3.73万小时	真机、UMI、人类第一视角、仿真、世界模型生成数据

资料来源：亿欧智库，Dyna，Generalist，智在无界，灵初智能，蚂蚁灵波，极佳视界，国信证券经济研究所整理

百万小时正成为多家厂商模型训练与数据采集规划的主要量级，面向通用物理智能的远期需求判断进一步达到千万至亿小时级。我们梳理了人形机器人本体厂商对具身数据的需求量情况，其中，星海图提出百万至千万小时的模型能力突破区间，银河通用计划到 2028 年将训练数据由 100 万小时提升至 1000 万小时，智元姚卯青则估计实现物理世界“GPT 时刻”可能需要约 1 亿小时数据。与上述目标相比，当前已披露的数据积累仍有较大提升空间：根据中国信通院，2026 年 10 月全国已建成启用 70 多家具身智能训练场，另有 40 余家在建或规划，相关数据平台已交付 150 万小时人类视频数据。自变量披露的操作视频规模为数万小时，Figure 披露的预训练与中期训练数据接近 100 万小时。**整体看，具身数据需求正向规模化扩张，数量积累与场景覆盖的双重缺口，将持续推动数据采集、处理及交付需求增长。**

表1：具身机器人本体厂商数据情况梳理

类别	公司	具体情况
数据需求	智元	➢ 智元合伙人、觅蜂科技董事长兼 CEO 姚卯青提出，实现机器人在物理世界中开箱即用、理解开放式指令及规划任务，所需数据或达亿小时级。
	星海图	➢ 联合北京亦庄成立“亦数智能”，通过亦数智能，星海图启动 100 万小时超高质量真实数据计划，目标在 2026 年完成百万小时数据积累，未来 3 年实现千万小时规模。
	跨维智能	➢ 公司在 EmbodiChain 技术说明中提出，规模化训练需要亿级乃至十亿级交互步数，并通过生成式仿真持续扩增符合物理规律的交互数据，提升数据生成效率。数量口径为交互步数。
训练规划与已有数据积累	银河通用	➢ 官网披露已构建百亿级具身智能数据基建“银河星数（AstraData）”；其中，GraspVLA 预训练完全使用合成数据，规模达到 10 亿帧“视觉—语言—动作”对。上述数据分别对应整体数据基建与特定模型的预训练语料。
	优必选	➢ 2026 年的数据收集目标是 100 万小时。优必选已在柳州、自贡、九江等地打造工业场景驱动的政产学研数据采集与测试中心，预计落地部署机器人将达 500 台以上。针对汽车生产、物流仓储、机加工智能制造产业等的模拟场景打造千万条规模的真机动作数据飞轮。满效开机下，单台机器人高质量数据条数（即经过数据采集、清洗、标注、对齐后的数据）可达 100 条/天，对应单台机器人年数据产量 400 小时以上。
	Figure	➢ 截至 2026 年 8 月 25 日，Index 平台收到超过 1,600 万条用户上传视频，应用覆盖 108 个国家；每秒处理约 30 分钟新上传视频。数据经过质量筛选、防欺诈审核、去重、再平衡及标注后，用于 Helix 训练。
	自变量	➢ WALL-OSS 采用数万小时以上的多源训练语料，包括自采机器人动作数据、开源动作数据及多模态视觉问答数据，分别支撑复杂任务学习、跨本体与跨环境泛化，以及视觉语言理解。该时长口径对应文章披露的整体训练语料。
采集部署、供给扩张	宇树	➢ 通过全身遥操作系统，今年年底之前宇树能部署几千台甚至一万台人形机器人。“每天采集 10 个小时的数据，最近的一两年甚至两三年之内，人形机器人的数据问题就可以基本解决。
	特斯拉	➢ 在 2026 年第三季度业绩会上，马斯克预计需要制造约 1 万—2 万台 Optimus，用于开展自我练习（self-play），并通过 Optimus Academy 中的多任务实践形成数据回流。该数量属于训练体系建设所需机器人规模的管理层判断。
	觅蜂	➢ 2026 年 8 月 31 日，第 20,000 套 MEgo 设备下线，累计生产超过 100 万小时高质量无本体数据，覆盖 22 大类场景、10,000 余个真实环境、50,000 余类物体以及 500 余种细分任务。百万小时数据中，裸手 EGO 约 50 万小时、EGO＋腕戴约 35 万小时、EGO+UMI 夹爪约 15 万小时。

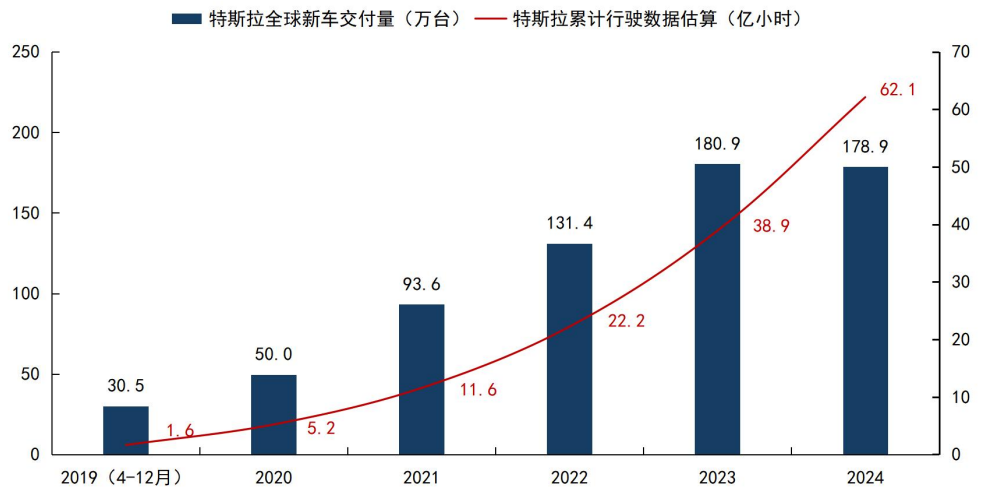
资料来源：觅蜂科技，星海图，跨维智能，银河通用，科创板日报，优必选，Figure，自变量，证券时报，特斯拉投资者关系，国信证券经济研究所整理

借鉴智能驾驶数据积累历程，具身数据有望实现百亿市场空间。当前，具身智能数据积累尚处早期，相较于数字 AI 可利用互联网沉淀的海量文本、图像及视频数据，具身智能训练更依赖真实物理交互中形成的感知—动作—反馈数据，采集涉

及硬件部署、环境搭建与任务执行，现阶段具身数据供给规模相对有限。随着应用场景拓展、任务复杂度提升及长尾问题涌现，高质量交互数据的需求成为具身智能的重要发展瓶颈。**展望明年，若行业需求向亿小时级迈进，按数百元/小时价格测算，数据市场对应约百亿元市场空间潜力。**相较当前行业普遍处于百万小时级的数据积累水平，未来需求扩张空间值得重点关注。

展望长期空间，我们参考智能驾驶行业的数据积累与能力演进路径，以特斯拉 FSD V13 实现显著能力跃升阶段的累计车队行驶时长为参照，估算具身智能数据市场的潜在空间。根据特斯拉，FSD V13 在安全性和舒适性上有重大改善，并具备从停车状态启动、驶离车位、倒车和泊车等能力。因此我们重点测算了特斯拉到 2024 年 FSD V13 模型能力有明显突破时累积的数据量。长期看，借鉴特斯拉智能驾驶数据积累历程，在车辆持续在用、单车日均行驶 1 小时的假设下，测算到 2024 年 FSD V13 模型能力明显突破时，特斯拉全球车队行驶时间测算约为 60 亿小时（测算值，仅供参考）。理论上，机器人的数据复杂度远高于智能驾驶汽车，若具身智能形成类似 FSD V13 重大模型能力突破，则数据需求大概率超越该时长，考虑到未来数据采集成本下降，远期数据市场规模也有望超千亿元，市场空间广阔。

图8：特斯拉智能驾驶数据测算



资料来源：特斯拉官网、特斯拉投资者关系官网、SEC 官网，国信证券经济研究所整理 注：按特斯拉交付量情况、单车日均行驶 1 小时、车辆持续在用估算累计行驶数据

■ 数据模态多，多模态融合仍存难点

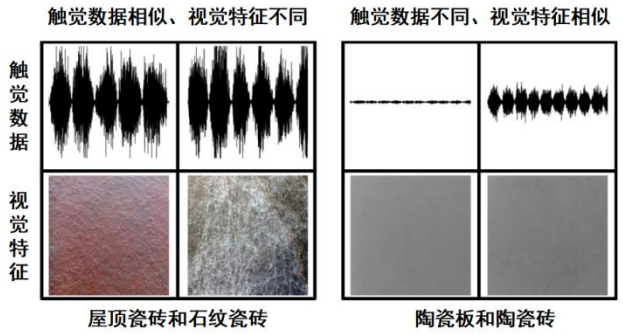
具身智能领域中的数据并非单一模态，而是由多种信息共同构成的复杂组合。典型的具身智能数据包括视觉(Vision，V)、语言(Language，L)、动作(Action，A)以及状态(State,S)等多模态信息的联合体。其中，视觉数据通常来自 RGB 或 RGB-D 相机，用于刻画智能体所处环境及目标对象的目标、操作指令或人类意图；动作数据刻画机器人在连外观与空间关系；语言数据用于表达高层次任务续或离散时间上的控制信号，如关节角度、关节返度或末端执行器位姿；状态数据则包含机器人自身的本体信息与环境状态，例如关节状态、力/力矩反馈以及触觉信息等。

将触觉融入感知体系成为人形机器人的重要技术方向。单一感知方式所能提供的信息往往十分有限，**视觉传感器**提供分辨率的图像**全局信息**，但难以准确识别物体表面材质、粗糙度、刚度，也受到遮挡、光照条件等诸多限制。**触觉传感器**能够在强遮挡、细粒度、对接触信息有需求的环境中帮助机器人更准确地感知和控

制目标物体，让机器人估计接触区域处的**局部信息**，但触觉只存在于机器人与物体直接接触的部位，单纯依靠触觉传感器难以对目标物体的位置、外观等全局信息进行全面描述，不足以了解物体的全局状态。

视觉和触觉感知拥有明显的互补特性，多模态融合成为智能感知的理想选择，有望满足人形机器人在复杂环境下的高级任务需求。不同物体可能拥有相似的触觉数据和不同的视觉特征，反之亦然。因此，视觉特征与触觉数据之间的互补性可以有效提高人形机器人的感知能力，满足高级任务需求，例如视觉+触觉的多模态信息进行有效融合后，机器人抓取的准确率有明显提升。**视觉与触觉传感器相互融合互补，实现多种传感器信息的综合感知和处理，提高机器人整体的感知能力。**

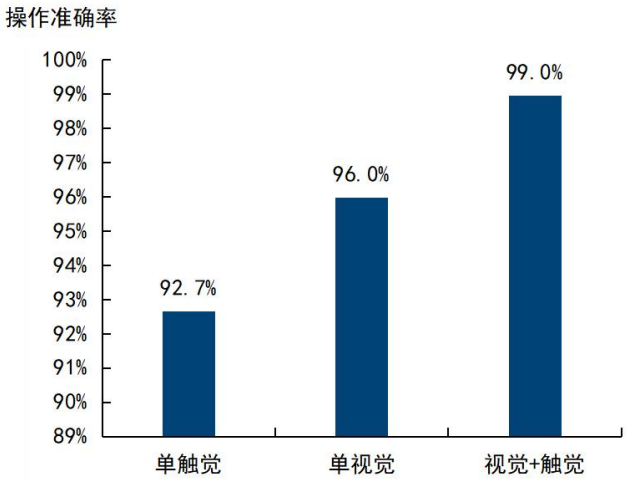
图9：视觉与触觉数据互补程度高



具体指标	视觉	触觉
视场角	大，全局性	小，局部性
信息丰富度	丰富	相对较少
数据收集难度	低	高
计算资源需求	高	低

资料来源：周天哲，《多模态融合赋能的物体识别与增强技术研究》，南京邮电大学，2023，1-76；Shan Luo 等，《Robotic Tactile Perception of Object Properties: A Review》，Mechatronics，2017，48:54-67，国信证券经济研究所整理

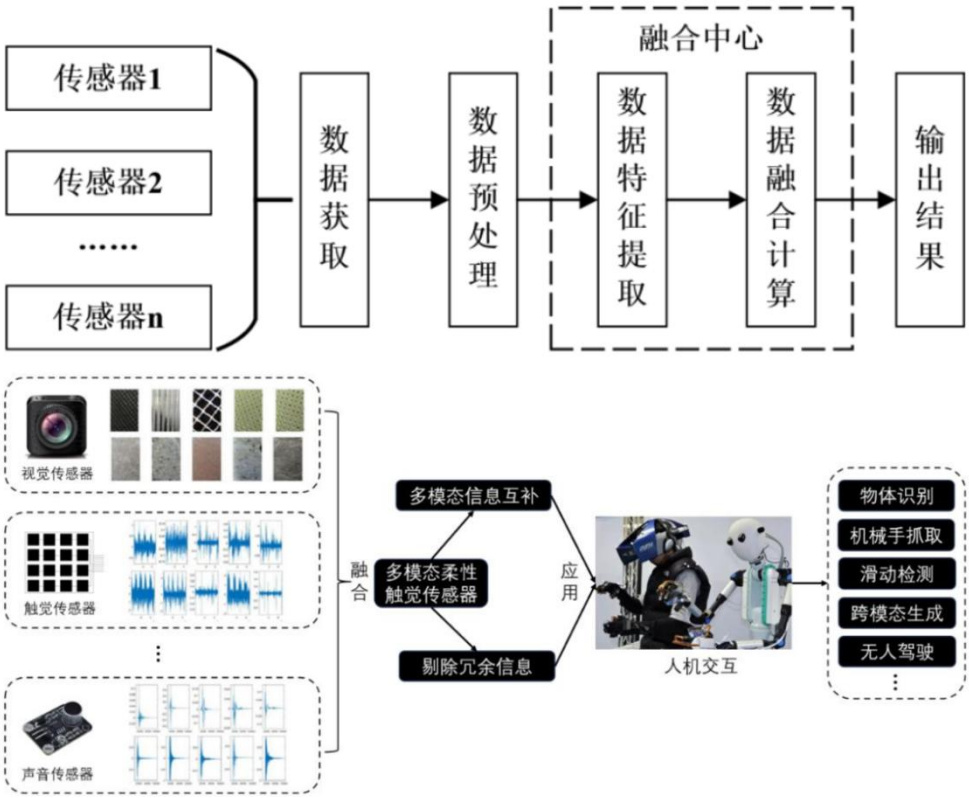
图10：视触融合比单模态的操作准确率高



资料来源：黄兆基等，《基于注意力机制和视触融合的机器人抓取滑动检测》，信息与控制，2024，53(2):191-198，国信证券经济研究所整理

多模态数据融合依然存在难点。不同的模态信息在形式和结构上存在不同，实现多模态的高效融合仍然难度较高。整体而言，多模态数据融合通常包括四个关键阶段，即数据获取、数据预处理、数据特征提取、数据融合计算。首先不同传感器获取原始数据，原始数据可能存在噪声或冗余，需要进行滤波、降维等预处理操作。随后，对预处理后的数据选择合适的融合方式进行融合，再进行决策，得到输出结果。

图11：多模态数据融合的整体过程



资料来源：尹一凡，《机器人触觉传感与多模态感知技术研究》，南昌大学，2024，1-71；杜宇坤，《基于视触融合的机器人抓取方法研究》，西南科技大学，2024，1-84，国信证券经济研究所整理

■ **数据标准未确立**

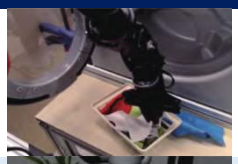
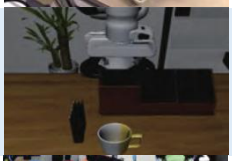
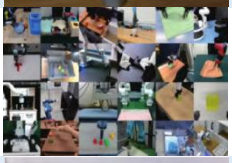
具身智能数据产业发展的核心难点在于统一的数据标准尚未确立。当前，遥操作、UMI、第一视角及动捕等采集路线并行，不同厂商在传感器配置、数据格式、时序同步、动作表征和标注规范等方面存在差异，数据质量评价与交付验收体系仍处于探索阶段。同时，不同机器人本体及模型架构对数据的需求各异，进一步增加了跨平台复用与效果验证的难度。标准体系的缺失推高了数据清洗、适配和交易成本，也使行业定价缺乏统一参照，加大了具身数据规模化生产和应用的难度。

■ **数据来源类型多，融合共享存在困难**

支撑具身智能预训练的有效数据形态尚未定型。大语言模型的成功本质上是用全世界的文本数据训练一个大语言模型，从而实现了数字 AI 领域的通用泛化，其数据来源相对较少。然而，具身机器人需要的数据不只是海量的互联网数据，还需要真实世界的动作数据、操作数据等，目前数据采集路线尚未收敛。

不同数据采集/仿真平台的数据集的融合共享存在困难。根据中国信通院，不同真机平台或采集设备在传感器类型、精度、采样率上差异显著，导致真实数据难以有效共享。不同仿真平台采集的数据在数据格式、坐标系、物理参数和任务映射标准等方面均不一致，仿真合成数据整合困难。从数据处理方法考虑，数据标注的精细度、视觉信息处理内容和采样频率可能均不一致。例如 π0.5 的训练数据包括对象检测和视觉问答标注信息。Gemini Robotics-ER 的训练数据包括抓取关键点、子任务标注和 3D 检测框等。

表2：具身数据来源多样

采集方式	示意图	采集成本	数据质量	数据	轨迹数	机器人种类数	场景数目	动作种类数目	是否有语言标注
人类示教驱动		高	高	RT-1 Datasets	130 k	1	15	700	是
				BridgeData	7.2 k	1	12	4	是
				BridgeData V2	60.1 k	1	24	82	是
仿真环境驱动		低	一般	LIBERO	6.5 k	1	5	130	是
				CALVIN	约 24 h	1	4	34	部分是
				Meta-World	n/a	1	1	50	否
跨平台数据聚合		一般	较高	OXE	1.4 M	22	311	217	部分是
				RoboTwin 2.0	100 k	5	50	50	是
世界模型视频生成		低	较低	DreamGen	n/a	n/a	n/a	n/a	是
				RoboEnvision	n/a	n/a	n/a	n/a	是

资料来源：丁贵广等，《具身智能数据采集与处理综述》，数据采集与处理，2026，41(2):332－346，国信证券经济研究所整理

基模厂商需求驱动具身数采加速扩容

GPT-6 Astra 等多模态通用模型能力持续提升，通用大模型能力边界持续向物理世界拓展，优势集中在语义理解、空间关系判断和任务规划，真实交互数据形成的动作能力仍具有重要价值。以 GPT-6 Astra 为代表的多模态通用大模型已开始探索机器人控制等应用，展现出较强的任务理解与闭环纠错能力。

- **RoboDojo 实验表明通用大模型 Astra 与 VLA 模型有互补价值。**RoboDojo 实验设置了两种 Astra 控制机器人的方案，包括 Astra Direct 直接生成机器人动作、Astra 与机器人 VLA 模型 π0.5 协同（π0.5 提供候选动作，Astra 结合视觉信息、机器人状态及执行历史，判断动作合理性并进行局部修正）。结果显示，Astra Direct 在 50 次任务中成功 13 次，成功率为 26%，48 个具有完整原生评分的任务平均得分为 37.81，在测评模型中排名第 3。**Astra+π0.5 的成功次数提高到 24 次，成功率达到 48%**，平均得分升至 62.60，排名上升至第 1，且 Astra 只修改了全部执行控制步骤中的 14.4%，说明 Astra 与 VLA π0.5 在语义推理与机器人动作先验上具有互补价值。

- **RoboLab 实验表明通用大模型 Astra 具备零样本泛化能力。**RoboLab 实验选取了 10 类任务、合计 50 次评测，GPT-6 Astra Direct 在 50 次实验中成功 49 次，成功率 98%；Astra+π0.5 完成 46 次，成功率 92%，成功率远超其他模型。同时，Astra 没有针对任务进行额外训练，模型根据当前画面、机器人状态和自然语言任务描述，持续判断下一步应该如何操作，展现出零样本泛化能力，模型能够识别目标、理解空间关系，将语言要求转化为动作，并在环境变化后继续调整策略。

基模厂商加快具身智能布局，数采市场有望进入加速扩容阶段。我们认为，基础模型在环境感知、语义理解及任务规划等“大脑”环节的能力不断提升，模型预训练、场景适配与部署迭代将共同拉动数据需求，具身数据有望成为物理智能发展的重要基础设施。随着部分基模厂商数字 AI 业务 ARR 增速放缓，多模态能力向

物理世界延伸，有望成为其拓展应用边界、培育新增量的重要方向，进而提升具身数据采购需求。展望 2027 年，若行业需求向亿小时级迈进，按数百元/小时的价格测算，数据市场对应约百亿元市场空间潜力。相较当前行业普遍处于百万小时级的数据积累水平，未来需求扩张空间值得重点关注。

图12：GPT-6 Astra 展现机器人操作能力，与 VLA 协同提升复杂任务表现



资料来源：Galbot Team，《Systematically Exploring the Capabilities of GPT-6 Astra as Embodied Policies》，arXiv，2026，1-46，国信证券经济研究所整理

具身数据：高质量触觉数据稀缺性高，采集与后处理成为核心环节

数据类型：视觉数据积累为主，高质量触觉数据稀缺性高

具身智能数据类型可按照感知模态、动作执行、认知决策及场景任务四个维度进行划分。

- **按感知模态划分：**视觉数据提供环境、物体及空间关系信息，构成具身模型的重要输入；触觉与力觉数据直接反映接触状态，补充视觉较难获取的物理交互信息；本体感觉数据描述机器人自身运动与姿态，语音、温度、气味等模态进一步拓展感知范围。
- **按动作执行划分：**数据覆盖移动底盘、双足步态及全身控制，并延伸至抓取、装配、放置、工具使用和双臂协作等操作任务，其中灵巧手数据支撑手指级精细操作，夹爪、吸盘等末端执行器数据对应具体工具的动作控制。
- **按认知决策划分：**指令与语义数据承载任务意图，VLA 语料连接视觉观察、语言指令与动作输出，世界模型数据支持物理结果预测，规划与导航数据服务于空间理解和路径决策。
- **按场景任务划分：**数据价值取决于多模态信息与实际作业需求的有效衔接。

行业场景定义机器人的应用环境，任务颗粒度进一步明确操作目标、动作流程及评价标准，并据此决定采集内容与数据组织方式。具身数据建设的核心在于围绕具体任务，将环境观察、本体状态、执行动作及交互结果组织为完整的训练样本，推动感知、决策与执行形成闭环，为模型训练及应用验证提供支撑。

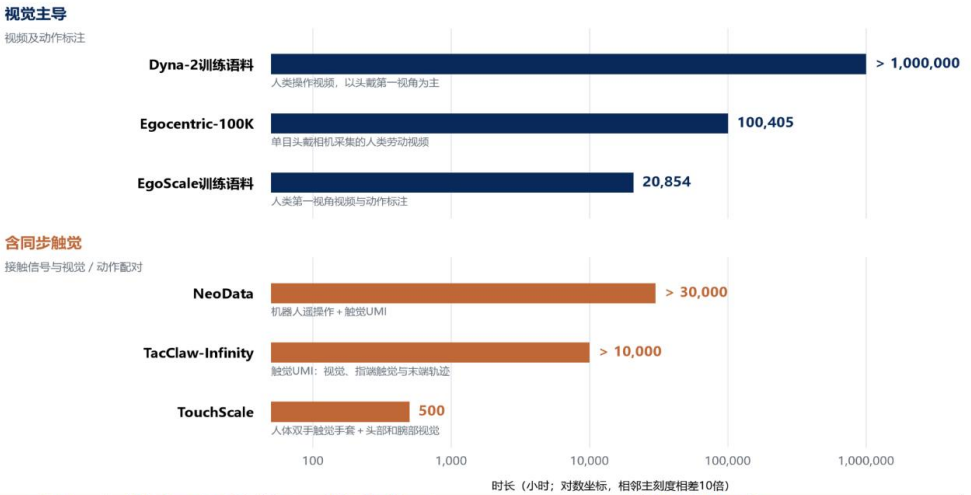
表3：具身智能数据类型

数据维度	细分类型	主要内容及作用
感知模态	视觉数据	通过相机采集环境、物体及空间关系信息，为机器人理解外部环境提供基础输入。
	触觉与力觉数据	直接反映接触状态及受力情况，补充视觉较难获取的物理交互信息，支持精细操作。
	本体感觉数据	描述机器人自身的姿态、关节及运动状态，为动作控制和身体协调提供依据。
	其他感知数据	包括语音、温度、气味等信息，拓展机器人对特定环境和交互需求的感知能力。
动作执行	整机运动数据	覆盖移动底盘、双足步态及全身控制，支撑机器人移动、平衡与身体协调。
	臂手操作数据	覆盖抓取、装配、放置、工具使用及双臂协作等任务，支撑操作技能学习。
	灵巧手数据	描述手指级动作及协同关系，为精细抓握、物体操控等能力提供训练支撑。
	末端执行器数据	记录夹爪、吸盘等末端工具的动作与工作状态，服务于具体工具的操作控制。
认知决策	指令与语义数据	以语言或符号表达任务意图，帮助模型理解操作目标与执行要求。
	VLA 训练数据	将视觉观察、语言指令与动作输出关联组织，支持端到端具身模型训练。
	世界模型数据	描述环境状态、动作及后续变化，支持模型学习物理规律并预测交互结果。
	规划与导航数据	提供空间理解、路径选择及行动规划信息，支撑机器人自主移动与任务决策。
场景任务	场景数据	按行业及应用环境组织数据，明确机器人的作业条件与应用边界。
	任务数据	按任务颗粒度定义操作目标、动作流程及评价标准，指导数据采集与训练验证。

资料来源：《具身数据发展报告 2026》，国信证券经济研究所整理

当前具身数据积累以视觉信息为主，触觉数据的供给相对有限。人类操作视频为具身模型提供了丰富的行为与语义信息，触觉数据则需要在交互过程中同步记录接触状态，现有同步视触觉数据的规模相较人类视频仍存在明显差距。**触觉数据通过数据传感器和接触物体之间的物理交互来测量接触物理属性。**触觉是人类在探索环境时不可替代的信息来源，能够将各种感觉信息（如压力、振动、疼痛和温度）传达给中枢神经系统，帮助人类感知周围环境。与视觉和听觉相比，人类的触觉在处理物体的材料特性和细节形状方面更胜一筹。与人类似，机器人的触摸传感（Touch Sensing）能够将外部刺激转换为可测量信号，实现机器与环境之间的良好交互。

图13：触觉数据规模有限



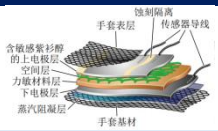
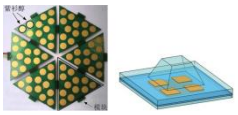
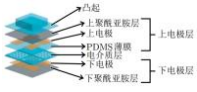
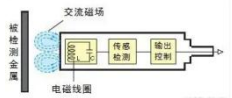
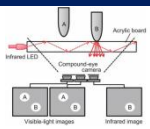
类型	数据集 / 训练语料	披露规模	数据内容及口径
视觉主导	DYNA Dyna-2训练语料	超过100万h	人类操作视频，以头戴第一视角为主，覆盖烹饪、整理、折叠、组装；企业训练语料规模
视觉主导	Build AI Egocentric-100K	100,405小时	头戴相机采集的人类劳动视频，作者数据仓库提供申请访问
视觉主导	NVIDIA EgoScale训练语料	20,854小时	人类第一视角操作视频，配套手部姿态及动作标注
含同步触觉	NeoData	超过30,000小时	真机遥操作与触觉UMI采集，视觉、触觉及动作同步；公开子集5,000小时
含同步触觉	TacClaw-Infinity	超过10,000小时	触觉UMI采集，包含视觉、指端触觉及末端轨迹
含同步触觉	TouchScale	500小时	人体双手触觉手套与头部、腕部视觉同步采集

资料来源：DYNA 官网，Build AI 官方数据，NVIDIA Research 官网，NeoteAI Team 等，《N0 -Foundation: Towards the Age of Tactile Intelligence》，arXiv，2026，1-35；Fucai Zhu 等，《TacClaw-Infinity: A Large-Scale Vision-Tactile-Language-Action Dataset for Dexterous Manipulation》，arXiv，2026，1-9；Dayou Li 等，《TouchScale: 500 Hours of Human Vision and Touch for Visual-Tactile Learning》，arXiv，2026，1-22，国信证券经济研究所整理

高质量触觉数据的采集与融合是当下核心难点。

- **数据采集方面，**采集触觉数据需要将触觉传感器集成至机器人末端、手持采集设备或可穿戴手套，并围绕不同物体、材质和操作方式组织交互。**在硬件方面，触觉传感器的技术路线尚未收敛，也加大了当前触觉数据的采集难度。**触觉传感器可分为压阻式、电容式、压电式、霍尔效应和视触觉等多种类型。传统基于电信号的触觉传感器（如压阻式、压电式、电容式）需要大量的电子仪器与线路来传输和处理高带宽的传感数据，随着传感器空间分辨率的提高，单位面积内的传感单元数量增多，会出现复杂的布线和串扰等一系列问题，信号收集电路的设计与制作往往比较复杂，而且由于信息转导模式的限制，空间分辨率普遍较低，存在迟滞大、易受电磁干扰等问题。而视触觉传感器直接将接触区域的信息以图像的形式进行编码传输，不需要在传感器上布置复杂的线路以处理电磁信号，极大增加传感器的空间分辨率，并且高维的图像信息适合与基于数据驱动的机器学习方法相结合，以深度挖掘其中的丰富接触信息。

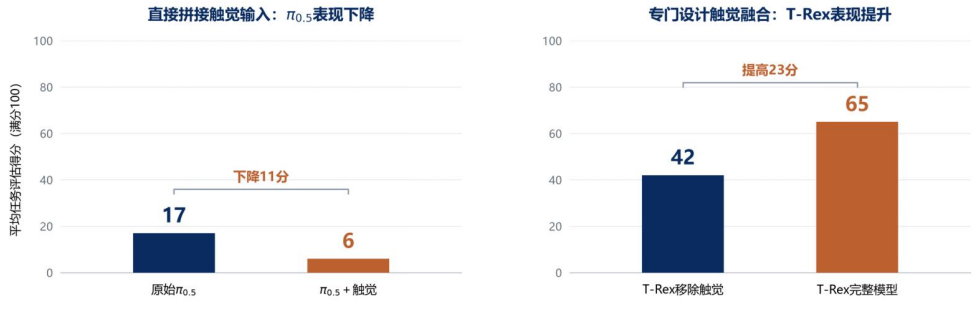
表4：不同原理的触觉传感器的对比分析

	压阻式	压电式	电容式	霍尔效应	光学式
示意图					
技术原理	压阻效应是压阻材料受到应力时，载流子数量变化，进而改变内部电阻率，采集电路通过将电阻的变化转化为便于测量的电信号来检测传感器的受力大小	压电效应是当晶体受到法向力时，内部产生电极化现象，撤掉外力时，晶体又恢复到不带电状态。压电式触觉传感器通常被连接到电荷放大器和测量电路上，采集外力相关电信号实现触觉检测	外力作用时，两极板间的相对位置间距和面积发生变化，从而导致电容变化，通过检测电容变化量来获取受力信息	霍尔效应指通电流的半导体薄片中，由于垂直磁场中的洛伦兹力作用，半导体中的电荷向不同侧边偏转，形成电场。当永磁体与霍尔传感器之间的距离发生变化时，传感器的电场强度也相应改变，从而计算形变程度	传感器与接触物体发生触碰挤压时，传感皮肤会产生相应的变形，由视觉相机捕获变形并在设备上成像、计算、分析、解释与变形相关的物理信息
优点	灵敏度较高，检测范围广，结构简单，成本低	线性度好，响应灵敏，动态响应较好	灵敏度高，响应速度快，结构简单，空间分辨率高、功耗低	精度高，响应速度快，具有线性特性	高空间分辨率，高精度，抗电磁干扰
缺点	温度敏感，有滞后现象，重复性、线性度较差	空间分辨率差，测量电路繁杂，易受噪声、温度影响	测量电路复杂，易受噪音、电磁干扰	温度漂移现象，强磁环境下存在干扰问题，成本较高	光照要求高，尺寸较大，可能存在发热问题，成本较高

资料来源：高工机器人，邓刘刘等，《智能机器人用触觉传感器应用现状》，现代制造工程，2018，(2)：18-23；朱盛鼎等，《触觉传感器与电子皮肤研究进展》，电子机械工程，2022，38(4)：4-9；尹宝凡等，《柔性触觉传感器智能感知技术与应用研究进展》，传感技术学报，2025，38(1)：1-20；邹宜桢，《基于视觉的触觉传感器及其在机器人灵巧操作中的应用研究》，浙江大学，2023，1-98，国信证券经济研究所整理

- **<u>数据处理方面，</u>**触觉数据需经过标定、噪声处理和时间对齐，建立接触信号与视觉、动作及任务过程的对应关系，触觉信息的融合方式同样影响训练效果。例如，2026 年 6 月由斯坦福大学、加州大学伯克利分校、英伟达等机构研究者联合发布李飞飞参与署名的 T-Rex 论文认为，将单时刻指端力信号直接拼接至 π0.5 的状态输入后，其在 12 项接触密集型任务中的平均评估得分由 17 降至 6；而针对 T-Rex 开展的六项任务消融实验中，完整模型得分为 65，移除触觉后降至 42，凸显了触觉表征与融合机制的重要性。我们认为，触觉数据的难点不仅在于数据采集，还在于触觉与其他模态数据的融合方式。

图14：直接加入触觉数据后可能反而导致模型效果变差



模型	这是什么方法、具体怎么做	是否使用触觉	平均得分
ViTacFormer	视觉与触觉结合的模仿学习方法。根据示范学习一段连续动作，并通过预测后续触觉帮助学习两种模态的联系。	有	3
RDP	全称Reactive Diffusion Policy。采用“较慢的动作规划＋较快的触觉修正”：先规划动作，再根据接触反馈及时调整。	有	6
Tactile-VLA	将触觉接入视觉—语言—动作模型，让机器人同时考虑操作位置与受力。本文复现时采用指端力 / 力矩作为触觉输入。	有	15
EgoScale	先从大量人类第一视角视频中学习手部操作，再将这些能力迁移到机器人。	无	35
π0.5	原始VLA基线，利用图像、任务指令和机器人自身状态生成动作。	无	17
π0.5＋tactile	在上一行基础上，将当前一步的10个手指力 / 力矩信号，直接拼接到原有状态输入中。平均得分从17降到6，你正文中的负向案例就来自这两行。	有，直接拼接	6
T-Rex	T-Rex完整方案，包含专门的触觉编码、训练流程和高频动作修正。	有，专门设计融合方式	65

无触觉得分高

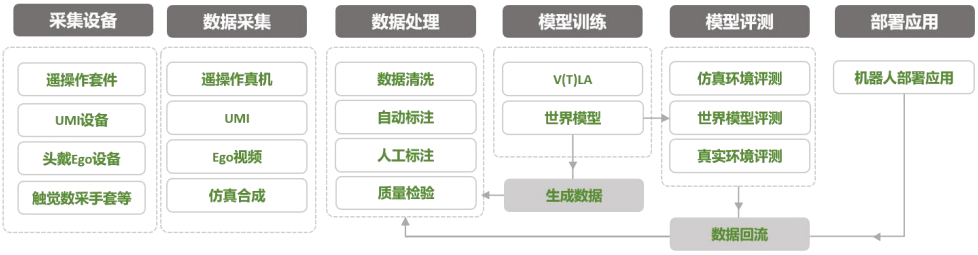
资料来源：Dantong Niu 等，《T-Rex: Tactile-Reactive Dexterous Manipulation》，arXiv，2026，1-30，国信证券经济研究所整理

数据环节：数据采集与处理是具身数据关键环节

具身数据的关键环节包括上游采集设备、数据采集、数据处理、模型训练、评测应用等。当前，具身数据量不足、数据采集方式尚未收敛，数据采集和处理成为目前的卡点。

- 数据采集设备：核心难点是产品的系统级设计，全套标定参数以及多模态信息采集传感器之间的毫秒级时间戳对齐，以保证采集设备的可靠性，以及采集回的数据稳定性、可复现性以及可用性。
- 数据采集环节：主要包括采集任务定义、场景搭建/租赁、人员培训、数据采集，在此过程中，需要进行项目计划制定、进度跟踪、人员调度、风险预案制定等。核心竞争力在于对模型数据需求的理解及采集任务定义、场景资源的获取以及组织运营能力管控能力。
- 数据处理环节：包括对采集来的数据进行清洗、标注、质量检验（QA）。无本体数据采集门槛降低之后，数据处理难题随之放大，海量、多模态、高对齐精度要求都是对人工标注的挑战。
- 模型：是数据的应用端，数据的质量需要模型端来评价，验证反馈越快，数据飞轮转的越快，具身智能性能提升的速度越快。

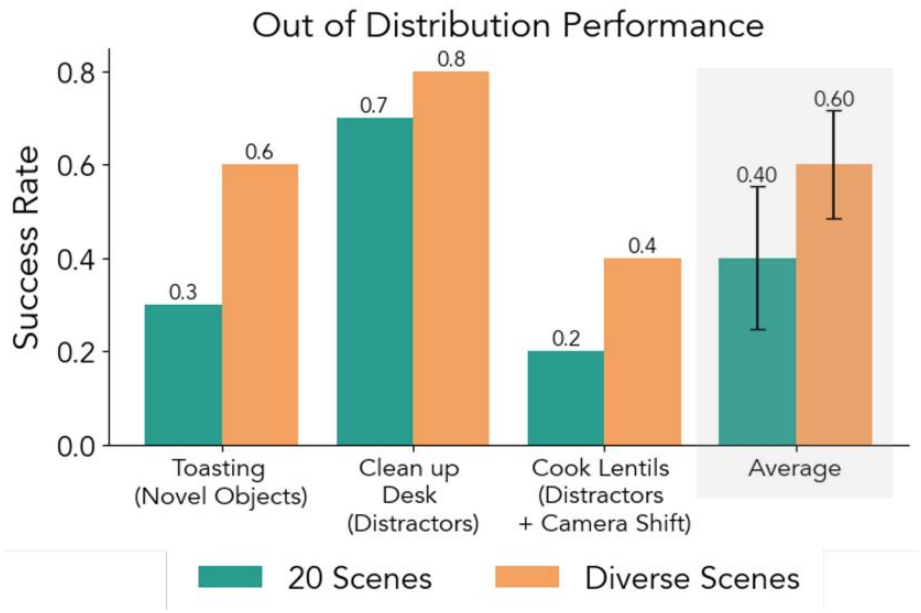
图15：具身数据关键环节



资料来源：高工人形机器人，国信证券经济研究所整理

数据采集是具身数据的起点，也是形成高质量训练数据的核心环节。采集阶段决定了原始数据的供给规模、场景覆盖与信息完整性，为后续清洗、标注、数据集构建及模型训练提供基础。DROID 研究发现，在示教数量相同的条件下，多样化场景数据带来更好的分布外任务表现，进一步体现了采集端对数据训练价值的影响。我们认为，持续获取丰富、完整的真实交互记录，是具身数据产业发展的基础，采集能力构成高质量数据供给的首要支撑。

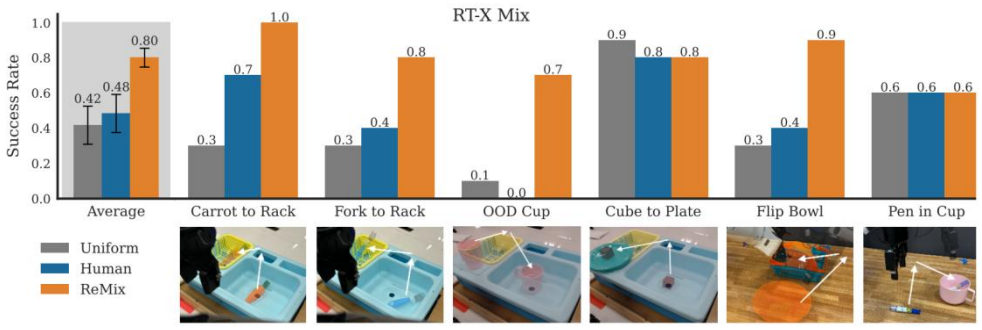
图16：相同数据量下场景多样性影响泛化效果



资料来源：Alexander Khazatsky 等，《DROID: A Large-Scale In-The-Wild Robot Manipulation Dataset》，Proceedings of Robotics: Science and Systems（RSS Foundation），2024:1-13，国信证券经济研究所整理

数据处理则是连接原始数据采集与模型训练的核心环节，也是将数据记录转化为有效训练样本的关键。采集获得的视频、传感器信号和动作轨迹，需要经过标准化、时空对齐、动作适配、语义标注及质量筛选，建立环境观察、任务目标、执行动作与交互结果之间的准确对应。处理质量直接影响训练样本的一致性、完整性和可用性，进而影响模型的学习效果。ReMix 研究显示，采集数据之后对数据进行配比优化能够改善下游机器人任务表现，进一步体现出数据组织方式对训练价值的影响。我们认为，随着 EGO、UMI、遥操作和动捕等采集路线共同发展，**数据处理能力将成为连接数据采集与模型应用的重要壁垒。**

图17：对采集的数据进行配比优化会影响下游任务效果

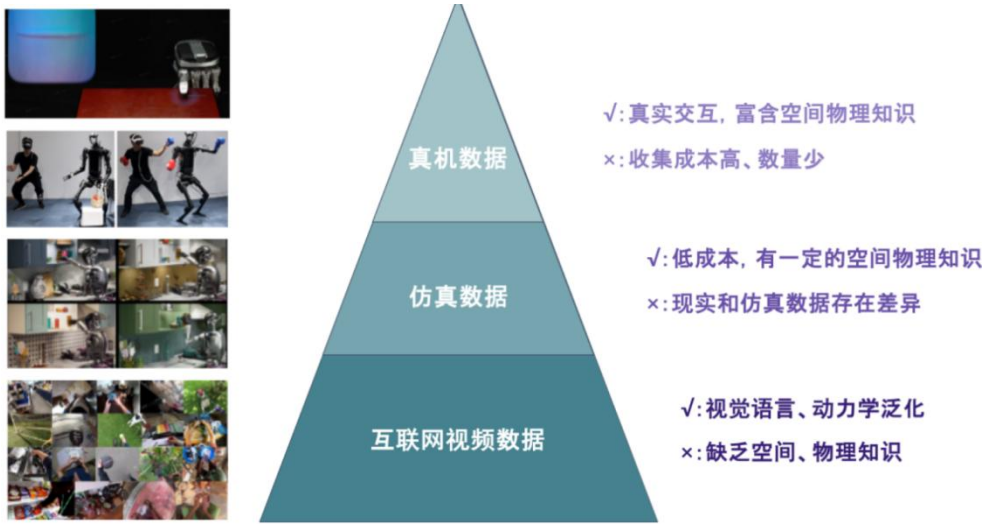


资料来源：Joey Hejna 等，《ReMix: Optimizing Data Mixtures for Large Scale Imitation Learning》，Proceedings of Machine Learning Research（PMLR），2025，270:145-164，国信证券经济研究所整理 注：灰色为按原始样本量形成的配比，蓝色为人工选择配比，橙色为 ReMix 优化配比

数据采集：技术路线多样，多源数据融合是大势所趋

具身智能领域数据包括海量互联网数据、仿真合成数据和真实数据等类型。在构建具身基础模型时，往往会以互联网多模态数据预训练的视觉语言模型为基座，让模型先具备通识知识，再结合仿真合成数据或真机数据训练，使其看懂物理环境、理解行为逻辑、学会动作技能并适应现实世界。仿真合成数据包括通过物理模拟引擎结合仿真平台、世界模型或其他合成技术合成的 3D 物体资产、3D 场景、本体轨迹或动作视频等，采集效率较高，但数据质量不足以支撑可靠的行为学习和真实场景泛化。真机数据是通过部署在真实环境中的机器人收集的数据，这类数据能较好地反映机器人实际行为，质量高，但采集效率低，且成本需要考虑机器人和大量人工费用。

图18：具身数据金字塔



资料来源：中国人工智能学会，国信证券经济研究所整理

从评价维度看，在具身智能数据采集中，质量、成本、规模、通用性不可兼得，通常从数据质量、规模化、成本效率、跨机型复用等维度对具身数据进行评价。

表5：具身智能数据评价维度

评价维度	定义	核心衡量要素
数据质量	对 VLA 模型训练的直接价值	动作标签精度（关节角度、位姿等）、多模态同步（视觉、语言、动作、力触觉等）、物理真实性（摩擦、形变、光照变化等）、任务多样性（长尾场景等）、噪声水平（标注错误、时间戳出错位等）等
规模化	单位时间内可采集的数据量及扩展上限	单实例采集速度、并行扩展能力、人力依赖度、时空约束、数据生命周期等
成本效率	单条有效数据的综合成本	硬件、人力、后处理成本、边际成本、失败率成本等
跨机型复用	数据是否可迁移到不同构型的机器人的可用性与迁移损耗	动作空间标签、观测空间一致性、形态差异容忍、策略泛化性等

资料来源：亿欧智库，国信证券经济研究所整理

我们判断，不同的数据采集方式在数据精度、规模化以及多样性方面各有千秋，未来具身数据的采集路线不会收敛到单独一种路线，而是多种来源的具身数据在不同环节可能都有应用潜力。伴随大模型对数据需求维度的几何级膨胀，单一技术路线可能难以满足“规模、成本、精度、泛化”的要求。行业未来预计将迈向

多源融合采集时代，以 EGO 数据注入通用物理常识，以仿真合成数据海量覆盖长尾边界，以 UMI 采集分布式扩充真实交互动作，最后依托高精度遥操作进行垂直场景的专家级微调。

- 真机遥操作能够直接获取机器人在真实环境中的操作示范，为模型学习具体任务提供数据基础，但硬件成本和采集规模、速度上略显劣势；
- UMI 和 EGO 数据可降低采集过程对机器人本体部署的依赖，可用于预训练阶段，用于赋予机器人通用知识和基础能力，拓展数据获取的场景范围；
- 动捕数据兼具真实数据和合成数据的优点，正在补充操作和全身运动控制任务的数据需求；
- 仿真与数据合成则通过场景变化和轨迹扩展，提高数据生产效率及样本多样性，但存在“仿真到现实迁移鸿沟”（Sim-to-Real Gap）。

图19：具身数据采集技术路线对比



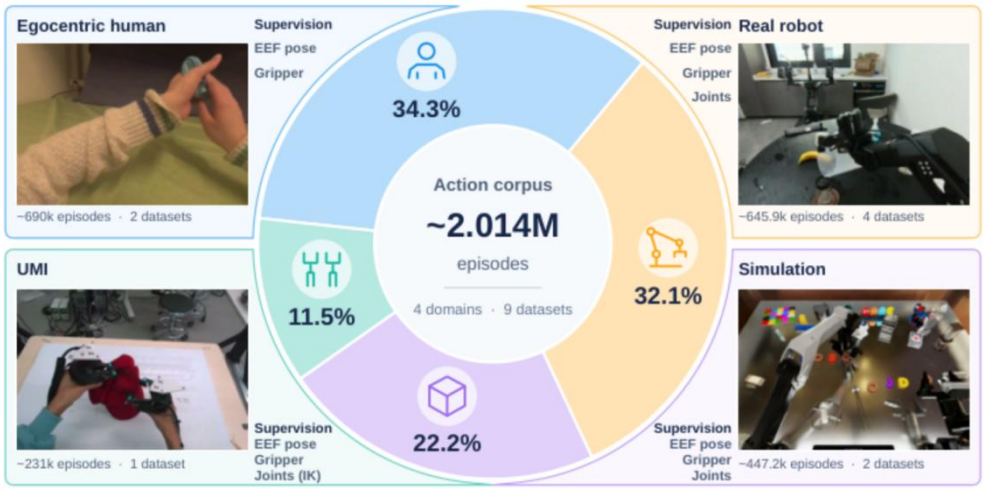
采集方式	技术原理	成本经济性	产能与规模化	数据质量与保真度	核心技术瓶颈	适用任务
遥操	操作员通过主从机械臂、VR/AR手柄或动捕设备远程操控真机，机器人本体与多模态感知同步记录。	数据价格约500—1000元/h；设备单套超20万元，机器人本体几十万元，触觉手套价格分化明显，几千至几万元均有	遥操作员单日有效产出仅2—3小时，常需1台设备配2人；节拍约为机械臂极限的67%，难以实现7×24小时连续产出	行业公认的“黄金标准”，数据天然处于真实机器人动作空间	人工操作效率低、人员成本高，采集节拍受操作者熟练度与设备运行约束	高精度精密装配、接触密集型长程任务
仿真	在物理引擎与仿真平台中程序化生成机器人轨迹与感知数据，配合域随机化、演示扩增、数学孪生等技术批量造数。	仿真数据生成成本约为真机采集的1/100；主要成本来自仿真平台开发、算力及场景构建	产能上限高、边际成本低，可通过程序化生成与并行仿真批量造数	仿真成功率89.4%，但真实家庭泛化测试约12%；头部方案拟合误差可控制在5%以内	Sim-to-Real Gap是核心挑战；柔性物体仿真、摩擦建模与形变标定仍是技术难点	海量标准化预训练、危险极端场景，须配合真实数据
无本体	不使用机器人本体，由人手持或穿戴便携终端采集末端轨迹与接触信息，再通过逆运动学映射到机器人动作空间。	未标注的Ego数据价格几十元/h，处理后的高质量数据几百元/h，Ego设备千元级别；UMI数据价格几十到几百元/h，UMI核心装备万元级	规模化潜力高，部署门槛低，无需配置完整机器人本体，具备进入家庭、工厂等真实环境规模化采集的潜力	可采集人类操作轨迹与接触信息，但人类动作与机器人运动学约束并不完全一致	难点在于后处理，需解决SLAM轨迹恢复、手部姿态估计、逆运动学求解与动作重定向；弱纹理环境下SLAM容易失效	大规模冷启动预训练、日常长尾场景补充
互联网数据	把分散、异构、多来源的数据集转换为统一格式（RLDS已成为事实标准）与统一动作空间，形成可比可训练的联合数据集。	单小时售价几十元，数据处理与训练对接成本可达采集本身的3—5倍	可利用已有数据集与公开资源快速扩充数据规模，但受数据格式兼容性、数据效果等限制	价值在于跨本体多样性带来的能力涌现，但部分数据存在RGB单模态、腕部相机或深度信息缺失等问题	存在遮挡、衣物干扰，骨骼提取精度不足	跨本体通用策略预训练、横向对比评测

资料来源：亿欧智库，《具身数据发展报告 2026》，国信证券经济研究所整理

部分厂商已开始采用多源数据融合作为具身模型预训练的重要路径。魔法原子的 Magic-W0 在预训练阶段融合 Ego、UMI、真机、仿真四类操作数据，覆盖 9 个数据集、约 201.4 万条有效操作轨迹。其中，EGO、UMI、真机与仿真数据的轨迹数量占比分别为 34.3%、11.5%、32.1%和 22.2%。模型通过统一的状态—动作表示对齐

不同来源的数据，联合学习动作生成与物理世界变化，展示了多种数据采集路线共同支撑具身模型训练的实践。

图20：魔法原子的 Magic-W0 融合 EGO、UMI、真机与仿真四类操作数据进行预训练



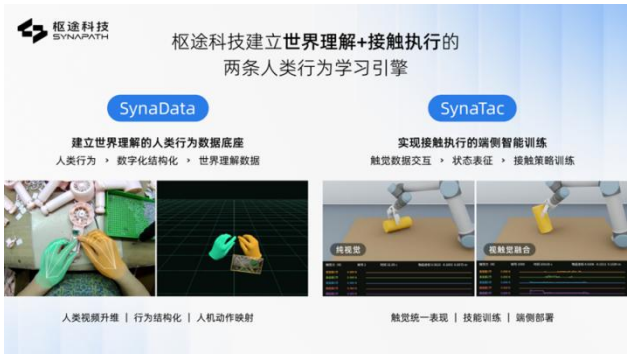
资料来源：Xuhua Chen 等，《Magic-W0: A Structured World-Action Foundation Model for Physical Intelligence》，arXiv，2026，1-29，国信证券经济研究所整理

互联网视频数据：数量庞大，构成具身数据底座

互联网视频数据规模庞大，是具身智能的数据底座。互联网沉淀的人类视频已经超过百亿小时，获取及生产成本较低，并且来源于真实物理场景，被认为是具身模型预训练的重要燃料。但这些数据没有力和触觉信息，只能支撑模型“看懂世界”，无法支撑机器人完成真实物理世界的精准交互。

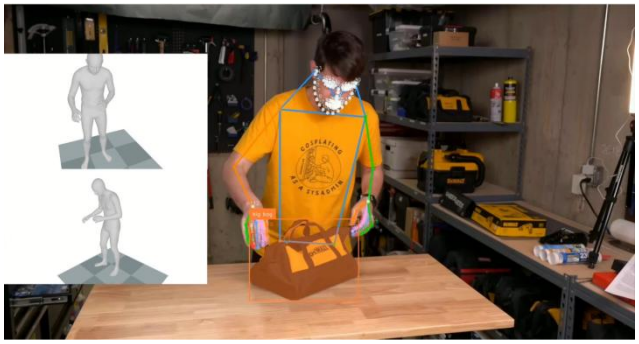
互联网视频在场景、物体和技能等维度的多样高，部分厂商进行视频蒸馏。互联网视频数据的多样性高，微软亚洲研究院的 VITRA 研究认为，互联网上的人类视频蕴含了丰富的场景语义和物体知识，通过将“人手”视为机器人末端，利用三维视觉技术构建数据管线，可将海量视频转化为结构化的 VLA 数据，单小时售价几十元。当前，有部分厂商从互联网视频提取多模态具身数据来充分利用互联网数据。例如，枢途科技从互联网采集海量 RGB 视频，并基于视频数据升维、跨域 retargeting 等技术突破，枢途 SynaData 数据管线解决方案将视频转化为多模态、高精度的具身训练数据，为具身智能的大规模训练提供可持续的高质量数据来源，枢途的综合数采成本降为行业平均水平的千分之五。

图21：枢途科技利用互联网视频数据构建具身数据集



资料来源：枢途科技，国信证券经济研究所整理

图22：枢途科技从基础 2D 视频中提取具身训练数据



资料来源：枢途科技，国信证券经济研究所整理

Ego-centric 数据：快速崛起的数据类型，规模化潜力大

第一人称视角（Ego-centric）数据是一种新兴的视频数据类型，目前已成为具身智能的重要数据来源。互联网数据主要是第三人称的视角，而 Ego 数据则是第一人称视角，核心思路是采集者佩戴头戴式或颈挂式设备，在日常工作环境中自然完成任务，不依赖特定机器人本体，被认为是规模化潜力较高、部署门槛较低的路线，在模型启动训练中发挥了关键的“冷启动”作用。例如，特斯拉采用 Ego 数据路线，由数十名员工正模仿着日常生活中的各种动作，在 8 小时轮班中重复数百次，由固定在头盔和沉重背包上的 5 台摄像机全程记录。

- 2024 年的 EgoMimic 研究显示，在 EGO、三维手部轨迹与机器人遥操作数据联合训练的条件下，机器人连续拾取入碗任务得分由 68 分提升至 128 分，增幅约 88%；叠衣任务成功率由 73%提升至 88%，提高 15 个百分点。同时，在拾取入碗任务中，人类操作示范的每小时采集数量约为机器人遥操作的 10.4 倍，体现出该类数据在训练效果与采集效率上的双重价值。
- 2026 年英伟达的 EgoScale 研究显示，预训练阶段 Ego 数据由 1,000 小时增加至 20,000 小时能够显著降低模型误差，并且经过下游任务训练后，真实机器人的平均任务完成度评分由 0.30 提升至 0.71，相对提高约 137%，从而验证了 Ego 数据扩张对模型学习效果及真机执行能力的双重增益。

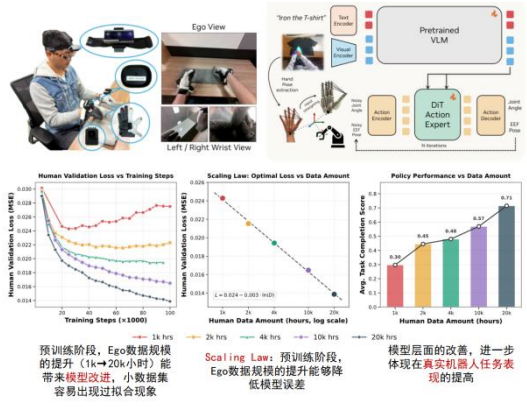
图23：EgoMimic 研究显示 Ego 数据能够提高机器人操作能力



方法序号	方法	使用的训练数据	方法含义	拾取入碗得分	叠衣得分	叠衣成功率	抓取袋子提手成功率	装袋得分	装袋成功率
1	ACT	机器人遥操作数据	学习人遥控机器人完成任务的示范，作为基准方法	39	82	55%	54%	82	22%
2	MimicPlay	人类示范+机器人遥操作数据	分层学习：人类数据帮助学习动作规划，机器人数据训练底层执行策略	71	78	50%	40%	53	8%
3	EgoMimic（无Ego）	机器人遥操作数据	使用EgoMimic架构，去掉人类数据，作为同架构对照组	68	104	73%	60%	92	28%
4	EgoMimic（含Ego）	EGO+三维手部轨迹+机器人遥操作数据	对齐人类与机器人数据，联合训练操作策略	128	114	88%	70%	110	30%
4相对于3的提升				+88.2%	+9.6%	+15pct	+10pct	+19.6%	+2pct

资料来源：Simar Kareer 等，《EgoMimic: Scaling Imitation Learning via Egocentric Video》，2025 IEEE International Conference on Robotics and Automation（IEEE），

图24：EgoScale 研究显示 Ego 数据的增加能改善模型能力



资料来源：Ruijie Zheng 等，《EgoScale: Scaling Dexterous Manipulation with Diverse Egocentric Human Data》，arXiv，2026，

2025:13226-13233，国信证券经济研究所整理

1-22，国信证券经济研究所整理

采集流程：①基于头戴式 RGB 第一人称视频流；②结合 SLAM 轨迹恢复、三维手部姿态估计、逆运动学与动作重定向；③构建统一相对末端执行器动作空间；④全域采用头戴主相机+腕部辅相机双视角，保证手部交互可见性。

图25：Ego-centric 数据采集过程



资料来源：Bowen Yang 等，《AoE: Always-on Egocentric Human Video Collection for Embodied AI》，Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops（IEEE），2026:3893-3902，国信证券经济研究所整理

优势：①核心数据丰富：记录的数据包括了第一人称视角中双手的操作、丰富的手-物交互细节、环境数据、身体关键点数据，同时包含了人类的决策逻辑、视觉注意力切换。采集员只需佩戴头环、头戴式相机，使用双手操作即可。**②采集方便，规模化潜力大：**由于是可穿戴设备，很容易进入家庭、商业等场景中进行数据采集。**③采集成本低：**未标注的 Ego 数据价格约几十元/h，处理后的高质量数据几百元/h，Ego 设备则为千元级别。根据中国信通院，**Ego 数据单台套售价在数万元左右。**

劣势：①缺乏力和触觉反馈：Ego 作为一种视频数据，缺乏精细力和触觉反馈数据。**②缺乏外部视角（如 Exo 数据+3D 环境）补足：**Ego 数据通常靠视觉进行定位，需要通过外部视角了解人与物体的整体关系，因此需要进行动作识别、轨迹估计和本体映射等后处理。例如英伟达 EgoScale 人类第一视角数据需结合 SLAM 轨迹恢复、三维手部姿态估计、逆运动学与动作重定向等进行数据后处理，以构建统一的末端执行器动作空间。

全身模态 Ego 数据补足全身视角。针对 Ego 数据的缺点，部分厂商开发全身 Ego 数据，为现有的 Ego 数据补足全身模态。Ego 数据目前主要记录双手的操作，却难以完整捕捉全身运动，也感知不到手接触物体时的压力变化，而这些被简化掉

的交互信息，往往才是完成任务的核心。破晓智能联合天珑移动发布全身全模态数据采集系统 SeePhys Hub。整套系统由一个四目头环、两个腕部手环、三个腰腿部动作追踪器和一副触觉手套组成。可同步采集第一视角视觉（360°环视覆盖）、手部全模态数据（掌内视角视觉、手部位姿追踪、高密度触觉）与全身运动信息（位姿与受力）。

图26：破晓智能全身 Ego 数据



资料来源：破晓智能，国信证券经济研究所整理

图27：破晓智能全身 Ego 数据采集及处理过程



资料来源：破晓智能，国信证券经济研究所整理

仿真数据：采集效率及多样性高，但面临 Sim2Real 鸿沟

仿真数据主要通过物理引擎构建虚拟机器人与操作场景，在无需真实硬件参与的情况下自动生成大规模交互数据。与人类视频数据不同，仿真数据通常依赖程序化策略、规划器或强化学习算法在仿真环境中自主执行任务，从而自动生成包含视觉观测、状态信息与动作序列的交互数据。

表6：主流具身数据模拟器对比

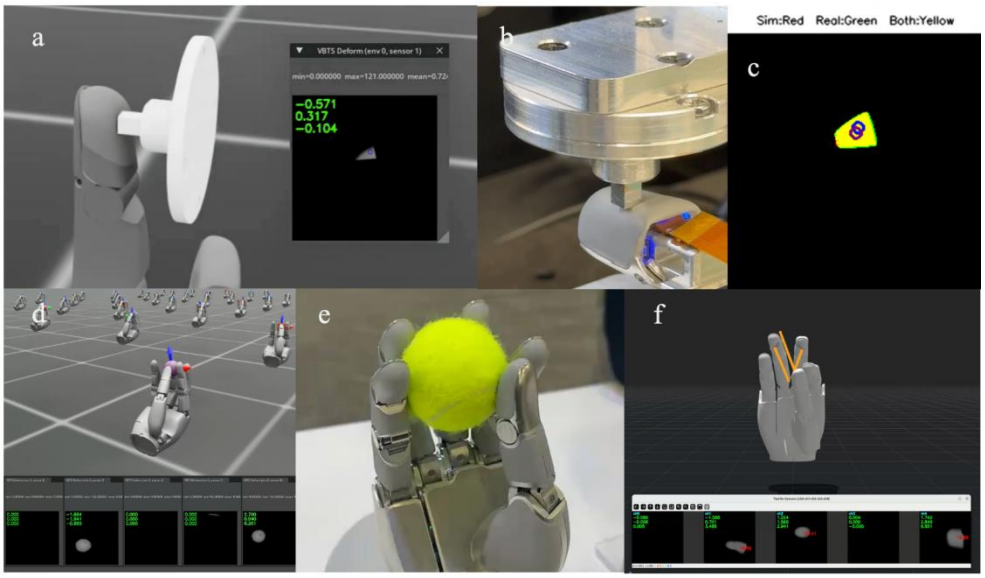
模拟器	底层物理引擎	并行能力	渲染速度
Isaac Gym	PhysX（GPU）	极高（10k－100k 并行）	中
Isaac Sim	PhysX 5	中（有限并行）	高
MuJoCo（传统版）	解析动力学引擎（连续可导）	中（CPU 单核极快）	中
MuJoCo 3.x（新版）	解析动力学+GPU 加速	高（GPU 加速+多线程）	中
PyBullet	Bullet（PGS 接触求解）	中（CPU 并行有限）	中
SAPIEN 3.0	PhysX（深度定制）	中/高（GPU 加速）	高
Genesis 1.2	统一求解器（ABD）	极高（>30k 并行）	高

资料来源：中国人工智能学会，国信证券经济研究所整理

仿真数据已在灵巧操作策略训练中展现应用价值。Sharpa 联合港科大及 NVIDIA 提出 Tacmap 触觉仿真框架，通过统一仿真与现实中的接触形变表示，支持操作技能从虚拟环境迁移至实体机器人。Tacmap 完全利用仿真交互数据训练强化学习控

制策略，并将其直接部署至 SharpaWave 灵巧手，在无需真机微调的情况下完成球形物体的掌内连续旋转，验证了仿真训练支撑接触密集型操作技能学习的可行性。

图28：Sharpa Tacmap 通过触觉仿真训练实现灵巧手操作策略向真机的零样本迁移

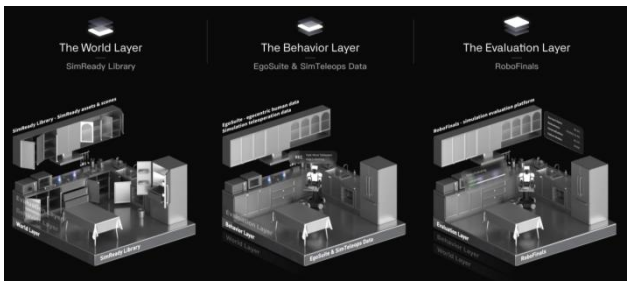


资料来源：Lei Su 等，《Tacmap: Bridging the Tactile Sim-to-Real Gap via Geometry-Consistent Penetration Depth Map》，arXiv，2026，1-8，国信证券经济研究所整理

优势：①成本低、可扩展性高：仿真数据减少了人工参与，仅需承担算力开销，能以较低的边际成本生成大量多样化的训练数据，并支持高度并行的数据采集流程，在数据规模和多样性方面优于真机数据。**②数据生成速度快、效率高：**例如，银河通用研发针对端到端 VLA 模型预训练的全仿真合成数据生产管线，在一周内就能生成十亿级机器人操作数据集（包含视频-语言-动作三个模态）。**③可生成长尾场景，失败数据成本低：**仿真数据能够生成人类未经历但符合物理规律的极端 Corner Case 场景，机器人可以在仿真中反复摔倒、反复失败，而且不会造成真实损失。**④数据多样性高：**仿真数据迅速切换场景、物体、环境参数，具备较强的多样性。

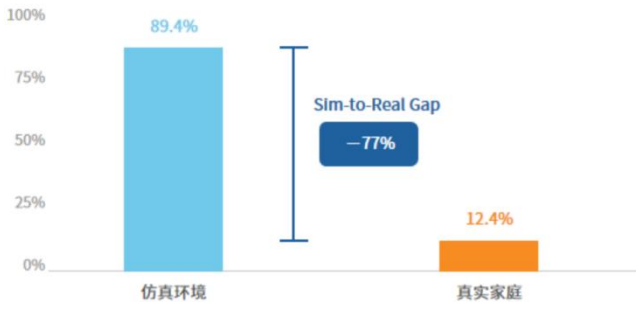
劣势：Sim-to-Real Gap。2026 年 4 月，斯坦福 HAI 发布的《AI Index Report 2026》中提到，机器人操控在仿真环境中的成功率高达 89.4%，但在真实家庭场景中骤降至 12%，存在 77 个百分点的“仿真到现实迁移鸿沟”（Sim-to-Real Gap）。Sim-to-real gap 主要瓶颈并非来自视觉逼真度不足，而是源于动作动力学和接触建模的不匹配：①物理模拟无法完美复制真实世界的复杂性和不确定性，即使先进的模拟器也会简化某些物理过程，忽略微小但可能重要的细节，也很难真实还原物理世界中的摩擦力、材料软硬等特性。②传感器模拟通常无法捕捉真实传感器的所有噪声特性和限制。③机器人模型可能与实际硬件存在差异，且即便仿真数据高度接近真实传感器数据，动作执行过程中的微小误差仍会在长时序任务中不断累积，最终导致策略在真实机器人上的性能显著退化。

图29：光轮智能仿真数据



资料来源：光轮智能，国信证券经济研究所整理

图30：仿真数据存在 Sim-to-Real Gap 问题



资料来源：亿欧智库，国信证券经济研究所整理

真实数据：价值高，采集成本高

真实数据采集是利用物理世界的设备采集真实数据的采集方式，分为无本体采集和有本体采集。不同真实数据采集技术在数据获取成本、规模化程度和数据价值上各不相同，形成多元协同互补的数据供给格局。

- **无本体采集：主要用于模型预训练阶段学习大规模行为先验。**通过人类操作员手持末端执行器或穿戴采集设备，在真实场景或生产工作中执行任务，以采集人类运动视频、轨迹、手部动作，以及人类通过自身判断对末端执行器的操作信息等。这类数据从操作员及其周围环境中得到，经常需要多种设备组合及识别算法处理，并从人体向具身智能本体映射，以覆盖多模态动作信息和训练具身智能模型。
- **真机采集：**依托具身智能本体获得数据，例如遥操作、真机回流数据，主要包含本体视觉、触觉、运动轨迹、关节角度、末端状态等多模态信息。这类数据从本体及其周围环境中得到，因此与本体结构一致性高，而且能够在预训练中为模型提供真实执行信息，同时自主作业回流数据进一步为模型补充各类长尾情况与失败纠错经验。

表7：真实数据采集技术对比

路线	数据采集方法	成本	规模化程度	真机映射难度	数据用途
无本体采集	末端操作数据采集（UMI）	低	高	高	支持多场景下末端精细交互操作任务
	动作捕捉数据采集	中等	中等	中	支持全身或手部复杂运动交互
真机采集	遥操作数据采集	高	低	低	无域偏移问题，与本体运动、力控、结构一致
	自主作业回流数据（真机回流）	高	低	低	具备真实环境信息、真实扰动、失败案例等高价值数据

资料来源：中国信通院，国信证券经济研究所整理

■ UMI：成本与质量的平衡，可实现跨本体复用

UMI(通用操作接口，Universal Manipulation Interface)数据本质上是末端操作数据采集，在数据质量、采集效率和成本方面取得较好平衡。通常由操作员手持或穿戴 UMI 夹爪作为采集工具，采集末端操作信息。UMI 夹爪利用与机器人末端执行器形态相近的夹爪作为操作接口，通过传感器实时采集操作员的夹持开合、位姿变化及运动轨迹数据；数据采集手套由操作员穿戴于手部，实时采集手部动作视频、运动轨迹和力触觉信息等多模态数据；腕部相机针对操作员的手部操作动作形成视频。夹爪内配有传感器，能够确定末端在空间中的未知位姿，算法层

面只需完成逆运动学求解，即根据末端位姿，计算机械臂的旋转角度。

采集流程：①设备准备与简单校准夹爪开合度等数据。②操作员手持夹爪操作物体，设备实时记录 RGB 图像、深度图、设备位姿、夹爪开合度等数据。③数据后处理包括轨迹平滑、时间对齐和质量检查，数据采完即可使用。

优势：①可采集力、触觉等物理交互数据：末端操作数据采集能够获取操作员手部运动轨迹、操作细节以及力、触觉等多模态真实物理交互数据。例如，智在无界发布的 UMI 灵巧手 BeingBeyond U1 可采集视觉、动作、力、触觉数据，能够实现动态追踪最大误差小于 1mm，平均误差小于 0.1mm。**②采集方便、成本可控：**UMI 采集无需机器人本体，结构轻量、部署灵活，更容易进入真实的家庭场景，无需在数据工厂中完成。根据亿欧智库，灵生科技发布的 LivUMI 方案以万元级成本实现高精度具身数据采集。UMI 数据价格几十到几百元/h，UMI 核心装备万元级。**③跨本体复用：**UMI 路线不依赖特定机器人本体硬件，一套采集数据可适配多种机器人形态，相当于为不同本体的机器人构建了一个通用的类人操作接口。

劣势：①效率偏低：UMI 仍然需要人类手持设备进行采集数据。**②动作空间与真实机器人有差距：**UMI 的动作空间并不完全等同于真实机器人部署时的动作空间，且 UMI 大多为夹爪，人形机器人末端主要为灵巧手，夹爪类 UMI 数据在灵巧手的训练中可复用性不高，部分厂商如智在无界 UMI 灵巧手则针对这一劣势进行迭代，推出 UMI 类灵巧手，更好适配灵巧手的数据需求。

图31：UMI 数据采集



资料来源：Cheng Chi 等，《Universal Manipulation Interface: In-The-Wild Robot Teaching Without In-The-Wild Robots》，Proceedings of Robotics: Science and Systems (RSS Foundation)，2024:1-14，国信证券经济研究所整理

图32：智在无界 UMI 灵巧手 BeingBeyond U1



资料来源：智在无界，国信证券经济研究所整理

■ 动捕数据

动作捕捉（Motion Capture）指的是将人的动作和姿态数字化的过程，具体来看是通过计算机视觉、传感器、磁追踪、外骨骼等不同技术手段把人的动作姿态捕获下来、转化为数字序列的过程。相比于遥操作，动捕数据示教侧重于使用各种动捕设备（如**数据手套**、动作捕捉服等）记录操作员的动作，并为机器人作示教，而不是精确对应机器人运动，在采集的数据可用性上存在差异。

动作捕捉数据采集主要包括光学捕捉和惯性捕捉两种方式，用于采集人体行为数据，多用于全身复杂运动或手部灵巧操作技能学习。

- **光学捕捉：**这类数据需要经骨骼解算和动作重定向后，才能映射到机器人本体，映射过程中可能出现动作失真、约束不匹配和信息损失等问题。同时，光学动捕方案对场地要求严格，需布置多视角高精度相机，成本在数十万到

百万不等，具体取决于相机数量、动作捕捉场地空间和精度要求等。

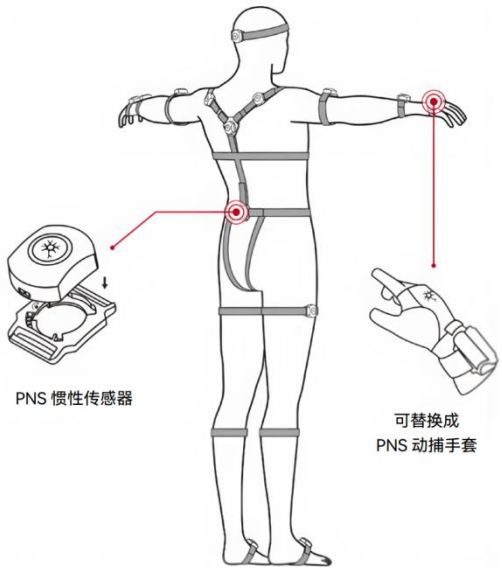
- **惯性动捕：**环境适应性强、便携性强，但惯性传感器本身不能进行空间绝对定位，存在漂移问题，成本在数万到数十万元不等。

优势：①能够获取结构化、多关节协同的人体运动信息，降低数据采集对机器人本体的依赖。在纯人体采集模式下，示范者可以直接完成动作，无需每条数据都由机器人实际执行，有利于减少机器人占用及真机试错，并在机器人尚不具备相应能力时提前积累训练素材。**②数据可复用。**经过统一表示的人体动作库可以重复用于不同模型的训练，提高数据复用价值。

劣势：①Embodiment Gap：人体动作与机器人可执行动作之间存在差异，需要承担动作映射及物理可行性适配成本。人与机器人在肢体长度、关节自由度、活动范围、质量分布和驱动能力等方面存在差异，人体能够完成的动作并不一定能由目标机器人稳定复现，动捕轨迹通常需要经过动作重定向、可行性筛选及控制策略训练，才能转化为机器人可执行的运动。**②Functional Retargeting：机器人只是在模仿动作的形状，却没有理解动作要完成什么。**动捕数据主要是把人的动作映射为机器人动作，系统只是建立了运动学上的对应关系，并没有真正建立操作本身在语义上的关联，因此动作捕捉数据可能出现关节角度超限、力矩不够、平衡失败等问题。**③单纯位姿数据对真实交互信息的覆盖不足，且高精度采集仍存在设备与工程成本。**动捕数据主要记录人体位置和姿态，不直接获得目标机器人执行任务时的接触力、驱动响应及任务结果；对于抓取、插接等依赖接触状态的操作，还需结合物体信息、力觉或触觉数据，以及真机执行验证。**④依赖专用设备和场地。**部分专业动捕方案需要专用设备、场地布设和标定，数据还需经过清洗与处理。

图33：诺亦腾动捕数据采集方案

PN S 惯性传感器	17 个
PN S 备用传感器	1 个
PN S 动捕手套	1 双
PN S 传感器充电盒	1 个
PN S 数据收发器	1 个
充电盒电源线	1 套
PN S 全身绑带	1 套
Type-C 充电线	2 条
防爆手提箱	1 个
全向天线	1 条
Axis Studio 软件	1 套



资料来源：诺亦腾官网，国信证券经济研究所整理

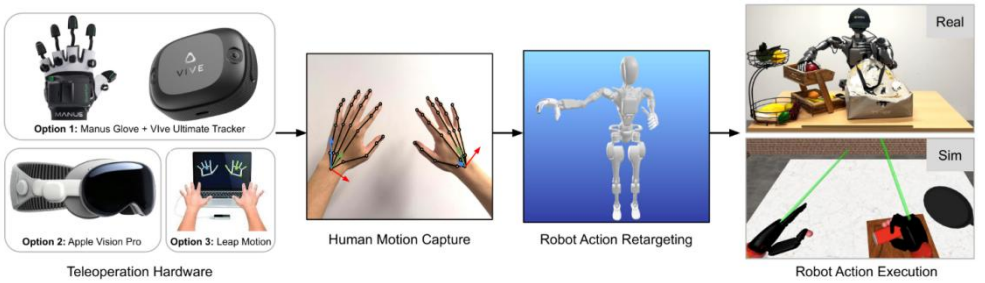
■ 遥操作：数据质量高、成本高

遥操作数据是由数采人员通过 VR 设备、外骨骼等主端控制设备，控制具身智能本体执行真实任务，同步记录本体产生的视觉、力觉、触觉等多模态数据的采集方式，与机器人本体结构一致性强。例如，由人操作两只机械臂，再远程控制最

终执行任务的两只机械臂，因此会形成主臂（Leader）和从臂（Follower）共四只机械臂的配置。

智元于 2024 年底发布 AgiBot World，并于 2025 年 3 月推出包含超过百万条人工遥操作轨迹的 Beta 版本，推动真实机器人数据开放共享。AgiBot World 数据集已进入国际主流模型的训练体系，英伟达第一代具身基础模型 GR00T N1 中，AgiBot World 数据占其预训练真机数据约 81.5%。

图34：GR00T N1 数据中有大量遥操作数据



资料来源：NVIDIA，《GR00T N1：An Open Foundation Model for Generalist Humanoid Robots》，arXiv，2025，1-36，国信证券经济研究所整理

采集流程：①场景准备：布置操作环境，放置目标物体，调整相机角度，确保光照充足。②设备校准：穿戴动捕服、校准传感器，验证关节映射，测试通信延迟。③任务演示：操作员执行任务，系统记录 RGB 图像、深度图、关节角度、末端位姿、力/力矩等数据，自动标注成功/失败。

优势：包含机器人真实本体与真实世界之间的交互，能够在模型能力不足时高效获取复杂技能示范。通过操作员直接控制机器人执行任务，遥操作可以同步获取视觉观察、机器人状态及动作指令，为模型提供明确的学习目标，降低从零探索的难度，尤其适用于精细装配、双臂协同和多步骤操作等任务。操作员还可以围绕目标技能设计采集流程，对关键动作进行重复示范，使数据供给与训练需求保持较高匹配度。在新任务导入阶段，即使机器人尚不具备自主完成能力，也能够通过人类示范积累有效训练样本。因此，遥操作在**模型冷启动、新技能拓展及高难度任务学习**中具有较高价值。

劣势：采集成本高、效率低。遥操作数据采集投入成本较高，需要配置本体、操作人员、采集场地和任务物料等，同时采集效率也依赖真人遥控。根据亿欧智库，遥操作设备购入价格＞20 万元/套，根据中国信通院，真机遥操作单小时有效数据成本在规模化运营后约 275 元（含设备折旧成本、人工成本、场景成本等），小规模采集时单小时成本可达数千元。根据智元机器人合伙人、觅蜂科技 CEO 姚卯青，真机遥操作 8 小时工作大概能平均产出 2 到 3 小时的有效数据，数据有效率接近 1/4，推高采集成本。

图35：遥操作数据采集过程



资料来源：亿欧智库，国信证券经济研究所整理

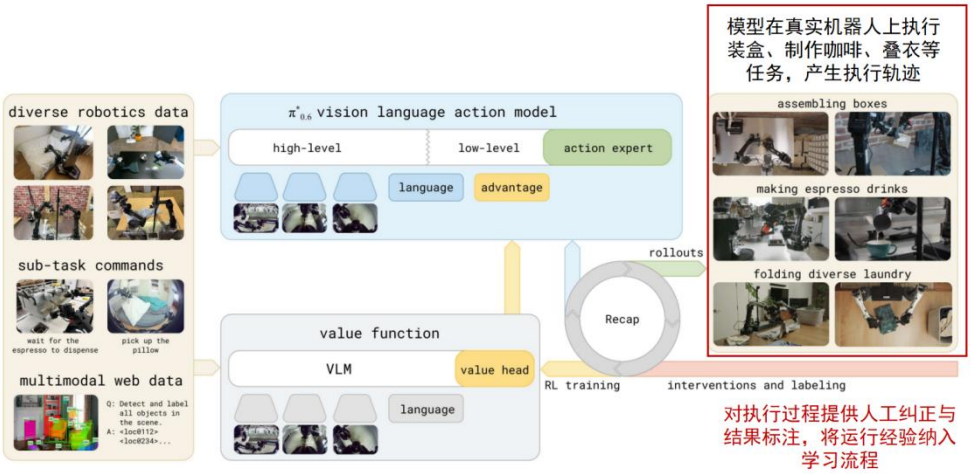
■ 真机回流数据

真实场景部署后的自主作业回流数据，是对模型持续迭代优化有益的高价值真机数据。这类数据自动采集本体在真实任务执行过程中的成功样本、失败案例、人工接管记录和环境长尾问题等信息。但受限于现阶段机器人的部署规模，短期内难以快速形成大规模高质量供给，更适合作为小规模高价值场景数据，用于对模型开展评估、验证和持续优化。

优势：学习价值高。用真机自主执行任务的回流数据不断优化面向具体任务的作业能力。真机自主执行任务回流的执行成功、执行失败或任务恢复等多种情况的数据，相比人工采集数据更具备长尾效应，因此对于模型来说具有高学习价值。Physical Intelligence 推出的π0.7 模型中，除了人类视频、开源机器人数据、遥操作真机数据外，引入π0.6 强化学习训练后自主执行任务的高价值回流数据，为模型制造出从高质量动作到低质量动作的数据分布，提高推理策略的执行质量。

劣势：①数据采集难度高。根据中国信通院，假设实现机器人“ChatGPT”时刻的模型参数量级为 55B，按大语言模型的缩放定律，训练所需数据量约 1.1 万亿 token。若全部由真机回流提供，按控制频率 8Hz、每步约 3 个预测 token、每天有效交互 2 小时、每年工作 300 天测算，约需 21200 台机器人连续工作一年，产出千万小时级有效数据。而当前全球可用的高质量真实数据仅数十万至百万小时级，未达到支撑具身基础模型智能涌现的门槛。**②对初始模型能力、评价机制和闭环工程要求较高。**自主采集的任务范围受到现有策略能力约束，模型能力不足时，可能反复执行简单动作或产生大量重复失败，难以稳定获取复杂技能经验。失败记录本身并不提供正确动作，必须结合成功判定、奖励信号或人工纠正，才能有效支持策略优化，回流数据能否转化为模型能力，取决于评价与训练机制的完整性；直接模仿未经筛选的自主轨迹，反而可能固化原有偏差。

图36：π0.6采用真机回流数据进行训练



资料来源：Physical Intelligence，《π*0.6: a VLA That Learns From Experience》，arXiv，2025，1-18，国信证券经济研究所整理

市场格局：具身数据产业化进程加速

产业链：贯通上游硬件、中游采集、下游模型应用

具身智能数据产业链覆盖基础设施、数据采集、数据加工及应用等环节，多元主体共同构建产业生态。从产业链结构来看，上游算力与云服务商、采集硬件及传感器厂商、训练场等为数据生产提供计算资源、感知设备和场景支持；中游通过真机遥操作、便携式及无本体采集、仿真与数据合成等方式获取数据，并经由治理、标注及流通环节形成可供模型训练使用的数据供给；下游对接整机制造、具身模型、开发平台和系统集成，最终服务于汽车、工业制造、物流及零售等行业客户，不同行业的作业环境、操作对象和任务要求存在差异，数据服务需要围绕具体任务组织采集、加工与验证。具备持续场景接入、数据筛选加工及模型效果验证能力的企业，有望积累更具应用价值的数据资源，增强客户合作的持续性。

图37：具身智能数据产业链



资料来源：易观分析，国信证券经济研究所整理

商业模式：定制化服务构成主要商业形态

具身智能数据商业模式以定制化服务为主，数据授权、工具订阅及场景运营等模式同步探索。具身智能模型训练涉及机器人本体、操作任务与应用环境的深度适配，客户对数据类型、采集方式和交付标准的需求存在较大差异，推动数据供给以项目制、定制化服务为主要形态。围绕数据采集、加工、生成、应用与评测，行业逐步形成多元化商业模式，其盈利能力与规模化潜力主要取决于数据复用程度、生产自动化水平、场景资源及交付成本。

■ **数采工厂**

数采工厂是机器人的“培训基地”。数采工厂的核心功能主要是：①真实场景搭建，覆盖工业、服务、家庭、医疗等多种场景，为机器人提供贴近实际部署环境的学习条件；②专业设备部署，包括机器人本体、遥操作终端、动捕系统、多模态传感器等采集硬件；③人在环质量把控，由具备产业经验的专业采集员进行长序列连续采集，为机器人学习注入人类直觉和物理常识。

政府引导和企业自建成为两大模式。根据中国信通院，从建设主体看，具身智能训练场主要有政府引导、企业自建两类模式。**①政府引导模式下，**政府提供场地、资金或政策支持，由具身智能企业、创新中心、科研院所或具身智能产业链上下游企业等联合建设运营。如上海人形机器人麒麟训练场由政府指导，国家地方共建人形机器人创新中心牵头建设。**②企业自建模式下，**企业基于自身研发需求，或场景数字化转型需求，自主投资建设，部分企业也会逐步对外开放能力，向行业供给数据资源。如上海浦东智元数据采集工厂由智元机器人建设，构建家居、餐饮、工业、商超和办公等真实场景，除供给自身研发数据外，也对外开源 Agibot World 系列开放场景真实数据集。

数采工厂需要重资产投入，成本主要为数采系统、人力运营、场地基建、算力存储。根据《中国具身智能数据采集产业发展蓝皮书》，帕西尼对数据工厂的成本

结构进行了拆解：

- **核心专业采集系统：**机器人本体单台造价几十万元。一个部署 100 台机器人的训练场，仅本体投入就达上千万元。
- **人力与运营成本：**高质量的数据采集需要庞大且具备产业经验的专业采集团队来支撑。采集员需要理解产业场景、掌握设备操作、进行人在环的质量把控，这一岗位的门槛高于普通数据标注员，人力成本较高。
- **场地与基建投入：**以天津帕西尼具身智能超级数据工厂为例，其空间达到 1.2 万平方米，其他训练场基本也在几千平方米的占地空间。
- **算力与存储底座：**具身智能的数据产出涵盖触觉、视觉、语音、文本、关节角度与空间轨迹等，需要匹配数百 PB（1 PB=1024 TB）级的庞大存储容量与极高的实时算力。

■ **众包采集**

众包采集将零散的场景、分散的人类操作经验整合起来，提高数据多样性。与传统集中式训练场不同，众包模式让数据采集走出工厂，走入真实家庭与工作场景。众包采集模式的经济意义在于将数据采集从"重资产集中式"拓展为"轻资产分布式"，有望降低单条数据的采集成本，并显著提升数据的场景多样性。但挑战在于质量控制，分布式采集难以保证数据质量的一致性，需要配套的自动化质量检验工具链。

2026 年 9 月 23 日，觅蜂科技发布全球首个全品类高质量物理 AI 数据众包服务平台觅蜂派，面向社会开放物理 AI 数据采集任务，用户通过觅蜂派 App 领取数据采集任务，佩戴 MEgo 设备在目前覆盖 22 大类场景、5000 多个任务和 50000 多个真实环境中完成指定操作，即可按照审核后的有效数据时长获得相应报酬，具体场景包括物流、仓储、住宿服务、教育、养老照护、餐饮、零售、公共服务、办公、汽车服务、医疗、建筑施工、交通出行、农业生产、文旅等。根据觅蜂派 App，众包的数据设备需要向觅蜂派申请，设备租金为 39 元/天，而单次任务报酬则在 20~50 元，任务时长为 2~60 分钟，用户完成 2 项任务后即可覆盖当天设备租金。

图38：觅蜂派 App 众包数据采集模式



资料来源：觅蜂派 App，国信证券经济研究所整理

图39：觅蜂派 App 众包数据采集模式



资料来源：觅蜂派 App，国信证券经济研究所整理

市场空间：以智能驾驶数据积累历程看具身数据市场空间

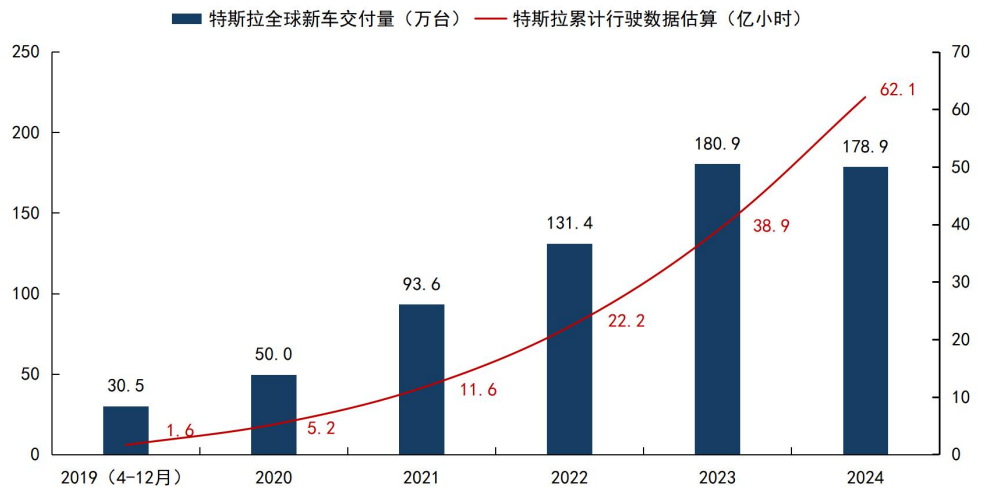
借鉴智能驾驶数据积累历程，具身数据有望实现百亿市场空间。当前，具身智能数据积累尚处早期，相较于数字 AI 可利用互联网沉淀的海量文本、图像及视频数据，具身智能训练更依赖真实物理交互中形成的感知—动作—反馈数据，采集涉及硬件部署、环境搭建与任务执行，现阶段具身数据供给规模相对有限。随着应用场景拓展、任务复杂度提升及长尾问题涌现，高质量交互数据的需求成为具身智能的重要发展瓶颈。**展望明年，若行业需求向亿小时级迈进，按数百元/小时价格测算，数据市场对应约百亿元市场空间潜力。**相较当前行业普遍处于百万小时级的数据积累水平，未来需求扩张空间值得重点关注。

展望长期空间，我们参考智能驾驶行业的数据积累与能力演进路径，以特斯拉 FSD V13 实现显著能力跃升阶段的累计车队行驶时长为参照，估算具身智能数据市场的潜在空间。根据特斯拉，FSD V13 在安全性和舒适性上有重大改善，并具备从停车状态启动、驶离车位、倒车和泊车等能力。因此我们重点测算了特斯拉到 2024 年 FSD V13 模型能力有明显突破时累积的数据量。长期看，借鉴特斯拉智能驾驶数据积累历程，在车辆持续在用、单车日均行驶 1 小时的假设下，测算到 2024

年 FSD V13 模型能力明显突破时，特斯拉全球车队行驶时间测算约为 60 亿小时（测算值，仅供参考）。理论上，机器人的数据复杂度远高于智能驾驶汽车，若具身智能形成类似 FSD V13 重大模型能力突破，则数据需求大概率超越该时长，考虑到未来数据采集成本下降，远期数据市场规模也有望超千亿元，市场空间广阔。

展望长期空间，我们参考智能驾驶行业的数据积累与能力演进路径，以特斯拉 FSD V13 实现显著能力跃升阶段的累计车队行驶时长为参照，估算具身智能数据市场的潜在空间。根据特斯拉，FSD V13 在安全性和舒适性上有重大改善，并具备从停车状态启动、驶离车位、倒车和泊车等能力。因此我们重点测算了特斯拉从 2019 年影子模式到 FSD V13 版本（2024 年）更新时的数据量。借鉴特斯拉智能驾驶数据积累历程，在车辆持续在用、单车日均行驶 1 小时的假设下，自 2019 年影子模式累计至 2024 年 FSD V13 模型能力突破期间，特斯拉全球车队行驶时间测算约为 60 亿小时（测算值，仅供参考）。理论上，机器人的数据复杂度远高于智能驾驶汽车，若具身智能形成类似 FSD V13 重大模型能力突破，则数据需求大概率超越该时长，考虑到未来数据采集成本下降，远期数据市场规模也有望超千亿元，市场空间广阔。

图40：特斯拉智能驾驶数据测算



资料来源：特斯拉官网、特斯拉投资者关系官网、SEC 官网，国信证券经济研究所整理 注：按特斯拉交付量情况、单车日均行驶 1 小时、车辆持续在用估算累计行驶数据

竞争格局：多元主体分工布局，数据质量与规模化交付构筑壁垒

具身智能数采行业呈现多类主体参与、沿产业链分工竞争的格局。上游以视觉、触觉、力觉等设备厂商为主，竞争优势主要体现为传感精度、多模态同步能力及采集设备成本；中游涵盖专业数据服务商、训练场运营商及仿真平台厂商，围绕场景覆盖、数据质量、交付效率和客户适配能力展开竞争；下游机器人及应用厂商依托本体、模型和真实作业场景，推进数据采集、训练与应用反馈闭环。我们认为，随着具身模型训练需求增长，规模化采集能力与高质量数据供给能力将共同构成竞争壁垒，具备软硬件协同、真实场景资源及持续数据交付能力的企业有望形成差异化优势，行业价值分配也将进一步向能够提升模型训练效果的数据产品与服务环节集中。

表8：国内具身数据主要玩家相关信息梳理

细分领域	公司	市值	营收	归母	研发	具身数据相关进展
视觉感知	思特威	420.3	90.3	10.0	6.2	2026 年 8 月与奥比中光签约深化合作，拟将 SmartGS 全局快门 CIS 集成至奥比中光无本体数采平台，覆盖 EGO、UMI、WristCam 等硬件产品。
视觉感知	禾赛科技	131.4	30.3	4.4	8.0	JT 系列激光雷达已进入机器人应用，2026 年 7 月官网披露原力灵机 Apex 搭载两颗 JT，可实现传感器与执行器微秒级同步，Apex 影子模式支持真实作业数据采集。
视觉感知	速腾聚创	54.4	19.4	-1.5	6.5	AC1 融合激光雷达、相机及 IMU，将深度、图像语义和运动姿态进行硬件时空融合；配套 AI-Ready 生态提供标定、数据融合及 SLAM 等开源工具。
触觉感知	汉威科技	115.6	24.1	1.6	2.3	控股子公司苏州能斯达。智能数据手套采用弹性拉伸传感器，应用于指关节训练数据采集、机器人遥操作和装配示教。柔性可穿戴动捕服覆盖 28 路传感器，用于人体运动采集及机器人示教。公司同时布局柔性触觉传感产品。
触觉感知	恒辉安防	43.9	12.3	1.0	0.7	传感封装及机器人柔性织物产品。基础集成型传感封装定制数据集采手套已有少量商品化订单；机器人柔性织物防护件实现小批量销售，柔性复合皮肤持续试样迭代。
触觉感知	南山智尚	55.1	17.9	1.3	0.7	2025 年 9 月发布织物触觉智能手套，融合柔性传感与感驱一体技术，面向灵巧手遥操作、机器人数据采集及 VR 交互。
触觉感知	福莱新材	140.4	28.6	0.8	1.2	柔性触觉传感产线完成调试优化，产品实现批量交付，围绕具身智能及人机交互的触觉数据需求持续迭代。福莱新材推出以触觉采集为主的数据手套系列，融合手部姿态、腕部空间信息、第一视角视觉与任务标签，并强调多模态数据的时间对齐。
触觉感知	晶华新材	47.9	21.7	0.9	0.9	控股子公司北京晶智感布局柔性触觉传感器及数采硬件。数采手套、自研智能夹爪等新品推进落地，指尖传感器取得头部客户量产订单，全掌触觉方案完成客户验证并进入小批量导入。
触觉感知	日盈电子	63.7	10.4	-0.5	0.5	推进机器人柔性触觉传感电子皮肤及相关线束布局。
触觉感知	兆威机电	225.2	17.2	2.5	1.7	参股企业途见科技发布柔性触觉数据采集手套，提供灵巧手触觉感知、机器人全身触觉及数据采集触觉系统，整合柔性材料、传感器、电路与算法。
IMU	华依科技	41.7	5.1	-0.6	0.5	2026 年 7 月与觅蜂科技合作,自研动捕专用 IMU 将全面搭载于觅蜂 MEgo 系列全品类物理 AI 训练数据采集设备。华依高性能 IMU 产品以高稳定、低噪声和多节点同步优势，适配单手套 10 颗、双手套 20 颗的分布式阵列采集方案，为下一代数采手套提供关键感知支撑。
动作捕捉	利亚德	145.7	70.6	3.1	4.0	全资子公司虚拟动点在 InfoComm China 2026 展出遥操作数采及异构机器人数据采集平台，接入睿尔曼、OpenarmX 本体，提供采集、清洗、标注流程。同步展示光学定位与无标记识别融合采集方案。
动作捕捉	凌云光	175.5	29.1	1.6	4.8	与睿尔曼等机器人厂商开展具身数据采集合作，发布 FZMotion 具身数采方案，以光学动捕同步获取人体、本体和环境运动信息，支持无本体示教、机器人姿态轨迹真值采集，以及多采集席位配置。
动作捕捉	奥飞娱乐	117.1	24.7	0.7	1.8	间接持股诺亦腾机器人，诺亦腾提供定制数据服务及 Adam-U 多方联合数据采集操作平台，关联动捕、遥操作与人类动作数据采集。
EGO / UMI	奥比中光	409.9	9.4	1.3	2.0	提供无本体数据采集硬件平台，覆盖头戴 EGO、腕部 WristCam、手部 UMI 等形态。UMI F-3 面向精细操作学习，同步采集图像、IMU 和夹爪开合数据，配套 EgoViewer 软件及多模态同步能力。
EGO / UMI	德马科技	56.3	16.9	1.0	1.1	全资上海德马具身智能机器人，发布 OmniEgo 头戴式采集终端，承担硬件制造、工程化及供应链。募投文件规划工业智能数据采集终端生产建设项目，总投资 6299.46 万元，建设期 3 年，设计年产 1.5 万台。通过与北京麦擎的合资公司开展数据服务。
EGO / UMI	均普智能	126.3	25.6	0.1	0.7	2026 年 7 月发布 Primus Forge 全链路数据平台及 Primus Ego 头戴式第一视角设备，覆盖采集、标注、训练、测评和部署。五目 270° 采集硬件记录工人实操，并结合工业产线开展应用验证。
EGO / UMI	美格智能	103.7	37.5	1.4	2.3	发布具身智能数据采集终端解决方案，提供动作捕捉可穿戴便携终端，以及支持 4—6 路 GMSL 相机的多视角视觉终端。依托 AI 模组，实现多传感器时间同步、多路原始视频采集及端侧处理。
EGO / UMI	龙旗科技	176.3	421.2	5.9	27.2	具身智能数采设备 JDM/ODM 业务，提供轻量化视觉采集、多模态采集和定制联合开发。多模态方案支持视觉、力觉、关节及控制信号同步，并提供自有量产产线真实工况验证。
遥操作与采集末端	长盈精密	356.0	188.2	6.0	13.9	参股公司天机智能发布 Marvin 遥操作数采系统，采集真实环境中的人类操作数据，同步记录机器人运动、力传感等多通道信息，支持数据存储、可视化分析和操作回放。
遥操作与采集末端	山东威达	44.4	19.2	2.5	1.0	参股知行机器人，发布优采数据采集手，将视觉、位姿、力觉与触觉同步采集，并配套自研数据采集软件。设备采用二指夹持结构，统一输出频率 30Hz，整机 600g；触觉压力采集范围 0—160N，力觉识别精度 0.1N。
数据采集与训练场	天奇股份	69.3	27.6	0.5	1.3	主导建设的江苏省具身智能机器人工业数据采集与实训中心已形成真机、UMI、EGO 三类采集路线，覆盖 36 个一级场景、163 个二级场景，累计沉淀近 2 万小时多模态数据。2026 年 9 月交付极石汽车阿布扎比机器人应用及数据采集方案。
数据采集与训练场	开普云	43.9	4.2	-0.1	0.6	无锡具身智能训练场已落地，采用专业运营团队、标准化流程和模块化作业方式开展多机器人连续动态采集，具备年采集 50 万小时真机交互数据的生产能力，并建立采集、加工、训练、验证流程。
数据采集与训练场	多伦科技	50.7	4.0	-0.3	0.5	与众擎共同设立江苏擎安，建设机器人场景数据采集中心及创新中心，提供数据采集、算法训练、模型验证等服务。
数据采集与训练场	东方精工	156.1	52.3	7.3	1.0	子公司东方合智承接数据采集及训练业务。2026 年 5 月佛山具身智能机器人数据训练中心揭幕，以智能制造、商业及家庭场景为牵引，结合遥操作、数字仿真及真机部署，构建高质量真机数据集。
数据采集与训练场	云从科技	93.7	5.0	-5.6	2.1	参与国地中心项目，推动跨区域、跨行业数据归集与模型训练，并探索真实与虚拟环境融合的训练平台。
数据采集与训练场	世纪恒通	40.4	10.6	-0.4	0.4	正在交付部分具身智能 EGO 数据采集与标注业务，开展穿戴式动捕采集，在模拟水厂、药店等场景同步采集视觉、触觉及运动轨迹，配合人机协同标注形成训练数据。
数据加工与资产管理	海天瑞声	74.1	3.8	0.1	0.6	已建立具身场景搭建、入户采集能力及供应商储备。多模态数据生产平台新增具身数采标注子系统，覆盖上传、分配、标注、质检及导出全链路体系。
仿真合成与三维数据	索辰科技	132.5	4.7	0.3	1.1	开物产品体系已包括行者合成数据及营造·万象世界物理模型，具备模型构建、数据生成、智能体训练和应用开发能力，推进具身智能等物理 AI 场景的产品验证和客户导入。
仿真合成与三维数据	尚品宅配	20.1	35.5	-2.5	1.7	2026 年 9 月联合汤元科技发布 WorldSimReady-Home 仿真数据集，在自有平台、Hugging Face 及魔搭同步免费开源，面向家庭服务机器人的仿真训练、算法评测及策略迭代。公司依托家居空间与设计资产布局具身数据服务，与上纬新材启元品牌合资合作家庭场景。
仿真合成与	群核科技	44.0	8.2	-4.3	2.9	SpatialVerse 空间智能平台生成物理正确的合成三维场景与虚拟数据集，用于智能机器人、具身 AI

三维数据						等训练；2025 年推出 SpatialTwin，将仿真能力延伸至工业数字孪生、模拟验证与具身机器人训练。
仿真合成与三维数据	天娱数科	124.4	20.8	-0.5	0.4	高质量纹理 3D 铰接数据、多模态 VLA 具身机器人抓取数据等数据集已在北京国际大数据交易所完成数据资产登记；3D 铰接数据集入选北京市高质量数据集典型案例。
仿真合成与三维数据	风语筑	71.5	16.6	-0.2	0.6	与松应科技合作设立控股风应科技，面向物理 AI 搭建仿真环境、合成数据与训练基础设施；设立浙江风语筑具身智能科技，布局空间数据及场景能力。
仿真合成与三维数据	宝通科技	91.5	34.6	2.1	1.4	与光轮智能合作建设面向工业机器人的持续学习基础设施，结合工业、矿山等真实场景，推动数据、仿真与部署反馈闭环。
仿真合成与三维数据	宇信科技	93.8	36.2	4.3	3.6	投资光轮、智平方、灵心巧手等，与光轮智能设立合资产业创新平台，覆盖场景定义、产品研发、数据运营、仿真评测、系统集成及部署反馈。
仿真合成与三维数据	阿尔特	49.5	9.8	-2.5	0.5	阿尔特、宇信科技、光轮智能三方合资合作；阿尔特提供真实工业场景及工程化支撑，输出汽车研发、精益智造及机器人产品工程化经验，提供真实工业场景。
数据平台与基础设施	五一视界	190.8	3.5	-1.8	0.8	AperData 由 AperEgo 头戴设备、AperWristCAM 腕部终端、AperFinger 夹爪手柄及 AperOS 构成，覆盖采集、时间对齐、标准化、解算、质检标注、增广及标准数据集导出。
数据平台与基础设施	仙工智能	57.1	4.4	-0.5	0.8	成立杭州开物创新具身智能科技有限公司，连接真机采集、多源接入、治理、模型训练、仿真评测及能力回灌，建设机器人大脑的数据基础设施。
数据平台与基础设施	当虹科技	38.4	3.2	-0.9	0.9	BlackEye Vision 面向机器人远距离遥操作，RobRTC 支持低时延视频传输；方案压缩与传输视频、深度及激光雷达点云，并通过远程操控节点采集真实交互数据。
数据平台与基础设施	协创数据	714.8	122.4	11.6	4.2	OmniBot 贯通具身智能数据采集、仿真、训练、评估及部署，将分散工具整合为可持续迭代的完整闭环；以 FCloud 统一算力与模型能力为底座，配套数据预处理、模型评估和训练监控工具。
数据平台与基础设施	华数传媒	122.7	91.8	4.7	6.0	组建浙江具身智能数据运营公司，围绕基础设施建设、数据产品运营及高质量数据集开展布局。
机器人本体与模型	宇树科技	2617.1	17.0	2.8	1.4	发布 G1-D 数采训练全栈方案，是集数据采集、处理、标注、审核及数据资产管理的一站式全流程数采平台，集项目管理、任务分配、进度追踪与状态分析于一体，通过模板化配置，一键生成数据采集任务，全流程状态实时跟踪，协作更顺畅，采集更高效。
机器人本体与模型	优必选	273.4	20.0	-7.0	5.1	Walker S2 具有仿真自动数采功能，可执行抓取、分拣、嵌装及动作回放，同步记录机器人状态、控制动作、头部和腕部四路 RGB 图像，以及任务、时间戳等信息，保存为 LeRobotDataset 训练数据集。
工业场景与数据回流	领益智造	862.6	514.3	22.9	23.8	建设具身智能数据资产平台，覆盖真机、仿真与互联网数据的采集、沉淀和应用；已搭建多个真实工作场景用于数据采集和真机训练，在上下料搬运、AOI 外观检测等环节部署工业具身机器人。
家庭场景与数据回流	上纬新材	649.4	18.0	0.4	0.5	旗下启元机器人与尚品宅配合作建设“机器人友好家”，2026 年 9 月成立合资公司。合作规划覆盖居家场景构建、数据采集、仿真建模及具身数据 DaaS；现阶段合资公司聚焦机器人二次开发与销售，并规划家庭场景搭建和数据沉淀。

资料来源：wind，各公司公告，各公司官网，国信证券经济研究所整理 注：市值数据截至 2026 年 10 月 30 日，营收、归母、研发费用均为 2025 年数值，单位为亿元人民币

风险提示

具身智能模型进展不及预期，具身机器人出货量不及预期，具身数据采购需求低于预期，具身数据技术路线变化，数据价格下降及行业竞争加剧。

免责声明

分析师声明

作者保证报告所采用的数据均来自合规渠道；分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

国信证券投资评级

投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后6到12个月内的相对市场表现，也即报告发布日后的6到12个月内公司股价(或行业指数)相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数(HSI.HI)作为基准；美国市场以标普500指数(SPX.GI)或纳斯达克指数（IXIC.GI）为基准。	股票投资评级	优于大市	股价表现优于市场代表性指数10%以上
		中性	股价表现介于市场代表性指数±10%之间
		弱于大市	股价表现弱于市场代表性指数10%以上
		无评级	股价与市场代表性指数相比无明确观点
	行业投资评级	优于大市	行业指数表现优于市场代表性指数10%以上
		中性	行业指数表现介于市场代表性指数±10%之间
		弱于大市	行业指数表现弱于市场代表性指数10%以上

重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司

关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中意见或建议不一致的投资决策。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，我公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

国信证券经济研究所

深圳
深圳市福田区福华一路 125 号国信金融大厦 36 层
邮编：518046　总机：0755-82130833

上海
上海浦东民生路 1199 弄证大五道口广场 1 号楼 12 层
邮编：200135

北京
北京西城区金融大街兴盛街 6 号国信证券 9 层
邮编：100032