仅供内部参考，请勿外传

国海证券 SEALAND SECURITIES

证券研究报告
机械设备
2026年10月14日

人形机器人专题报告8
跨越与弥合：数据采集的必要性与重要性

评级：推荐(维持)

张钰莹(证券分析师)
S0350524100004
zhangyy03@ghzq.com.cn

仅供内部参考，请勿外传

最近一年走势



相对沪深300表现

表现	1M	3M	12M
机械设备	-8.2%	-14.7%	-3.9%
沪深300	-7.7%	-8.9%	-6.9%

相关报告

《人形机器人专题5：Sim to Real，具身大模型的问题、现状与投资机会（推荐）*工程机械*张钰莹》——2025-11-10

《人形机器人专题报告6：人形本体&灵巧手的电机进化“势”（推荐）*机械设备*张钰莹》——2026-01-03

《人形机器人专题报告7：“腕”与“踵”，商业化落地前夕，机器人的散热瓶颈（推荐）*机械设备*张钰莹》——2026-03-11

仅供内部参考，请勿外传

核心提要

本篇报告解决了以下核心问题：1、数据采集应用的理论前提：具身智能大模型真正落地应用前，为何、如何转动数据飞轮？2、数据采集应用的现状：从数据采集行业出发，各具身智能公司的大模型、数据采集工厂、数据采集硬件、平台及数据资产等的建设情况如何？

◆ **数据采集应用的理论前提：具身智能大模型真正落地应用前，为何、如何转动数据飞轮？**

1、从具身智能应用落地的目的出发，具身智能大模型需要打通表征与闭环两条主线，打破数据墙、表征墙与闭环墙三道墙，完成这一目的，可能需要百亿小时级别的高质量数据；典型的具身智能数据可以表示为视觉、语言、动作以及状态等，通过多模态数据跑通感知-决策-行动的闭环；

2、数据采集行业：从较宏观的角度来看，当前，数据采集行业逐步走向硬件设备成熟化+采集平台生态化+相关职业完善化，并催生了数据加工产业，主要有数据清洗、标注、治理、管理等；

3、数据采集方式：现有具身智能数据采集方式可归纳为四种核心范式：人类示教驱动、仿真环境驱动、真机跨平台数据聚合和世界模型生成，不同方式各有其优劣。

◆ **应用现状-布局、竞争：数据采集的行业现状与商业闭环**

1、整理了各本体厂大模型路径、进展以及数采工厂的建设情况；

2、 整理了各厂商在数据采集行业的布局，部分厂商布局全链路生态，包括硬件本体、数据集、大模型及场景落地；

3、数据采集硬件的不同形态，主要有手部数采设备、外骨骼类、柔性类和穿戴电信号类等。

◆ 相关标的：我们认为人形机器人处于技术突破、商业化落地初期，后续有望迎来量产空间，维持人形机器人行业“推荐”评级。建议关注：1）本体：宇树科技、优必选、极智嘉等；2）垂直场景本体应用：杰克科技、永创智能等；3）数据采集硬件：凌云光、汉威科技、福莱新材等。

◆ 风险提示：人形机器人行业进展不及预期风险；数采成本下降不及预期风险；数据应用安全性及合规风险；相关行业标准制定与行业发展速度不匹配风险；中美贸易摩擦风险；重点关注公司业绩不及预期风险；研究报告中使用的公开资料可能存在信息滞后或更新不及时的风险等。

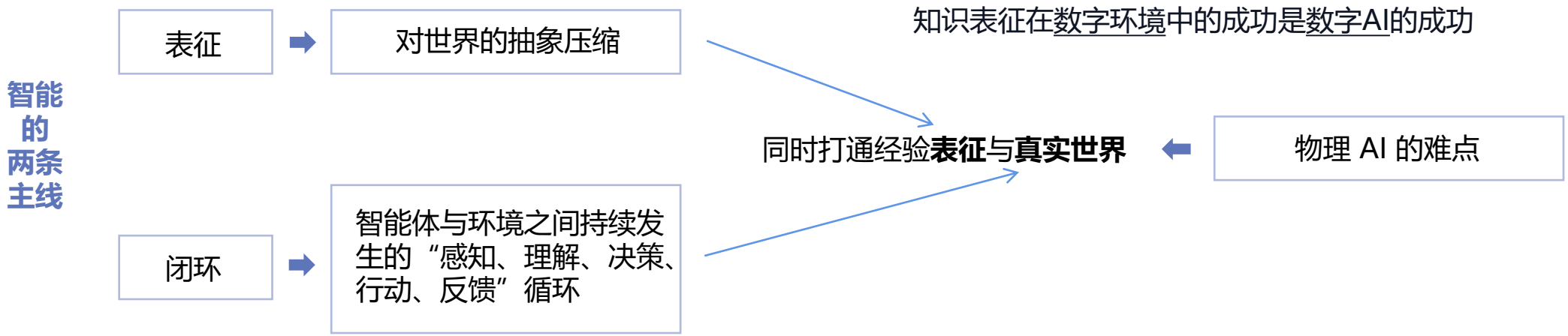
仅供内部参考，请勿外传

理论前提：具身智能大模型真正落地应用前，为何、如何转动数据飞轮？

仅供内部参考，请勿外传

模型：需要什么样的具身智能模型？

- 物理 AI 从“能演示”走向“能干活”，关键不在于单点模型能力，而在于能否形成可泛化、可优化、可进化的通用具身智能系统。
- 理解物理 AI，首先要回到智能的两条主线：表征与闭环。表征，是对世界的抽象压缩；闭环，是智能体与环境之间持续发生的“感知、理解、决策、行动、反馈”循环。数字 AI 的成功，本质上是知识表征在数字环境中的成功；而物理 AI 的难点在于，必须同时打通经验表征与真实世界闭环。



- 对应地，物理智能要实现规模化，必须突破三道墙：

数据墙	表征墙	闭环墙
真实交互数据缺乏且获取成本高	跨任务、跨场景、跨本体的统一物理表征尚未形成	真实试错代价昂贵、反馈缓慢，难以规模化

仅供内部参考，请勿外传

数据：需要什么样的数据?

什么是高质量物理AI数据?

- 以终为始，能**让模型真正形成可用能力**，就是高质量数据。拆开来看，必须包含：硬件层面的同步与标定、语义标注规范、采集规范、数据分布合理、场景多样性、任务多样性，这些共同构成高质量数据。
- 数据对具身智能不是燃料，而是 “人的教育” 。好数据必须满足：真实开放场景、言传身教式长程示范、可规模化、可终身学习。不可规模化的数据，再精致也不是高质量数据。
- **真实性：**大多在实验室或人工场景采集，不够真实，需要泛服务、工业、家庭等真实场景；
- **高精度标注**：多模态时代一张图能有几千字标注，描述背景、前景、意图，但现在视频数据只有简单动作标注，环境和任务描述严重不足；
- **能让模型“涨点”的数据。**数据需要具备多样性、准确度、覆盖度等要素。对具身来说，未来还会出现更多质量维度。高质量数据加入，可以实现模型性能持续提升等。

AGI时代到底需要多少量级的物理AI数据?

从时长来看，十万小时、百万小时、一亿小时，行业里存在各种判断；部分行业观点为100亿小时是通用模型的基础门槛；当前已公开的开源数据集，基本以百万条/小时为量级。

表：部分已开源发布的数据集参数

厂商	宇树	智元	银河通用	星海图
数据集名称	高质量全身遥操作真机数据集 UnifoLM-WBT-Dataset	全域真实场景、全能硬件平台、全程质量把控数据集AgiBot World	灵巧手多样抓取仿真数据集 DexonomySim	发布百万小时超高质量真实数据计划
发布时间	2026年3月	2024年12月	2025年8月	2026年6月
数据集参数	340小时合计189万条动作轨迹	百万真机数据，100+场景、单条数据时长较长	超过950万条高质量抓取姿态	/

为什么以及如何构建数据生态?

1）多模态标准缺失：视觉、触觉、声音、肌电、脑电等数据，格式不同、坐标系不同、模态定义不同，可能花费大量时间做数据转换；2）数据筛选：a. 智能筛选高价值 / 长尾数据，把最有用的数据拿来训练；b. 需要失败数据做学习进化，失败轨迹、反事实数据、因果推理数据撑起强化学习与闭环仿真。

资料来源：觅蜂科技公众号，虎嗅网，澎湃新闻，国家数据局公众号，星海图公众号，智元官网，国海证券研究所

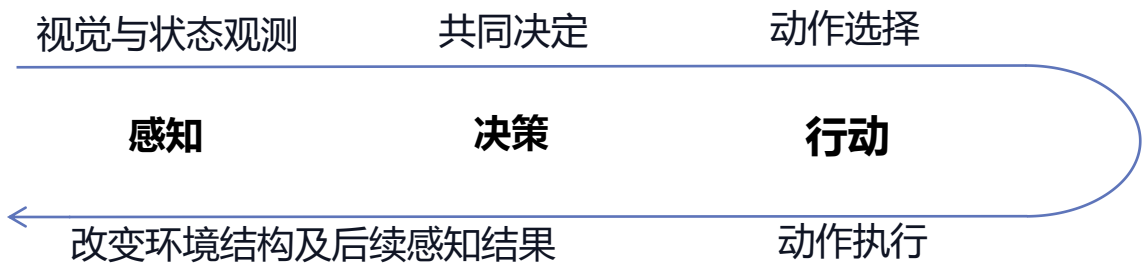
仅供内部参考，请勿外传

数据采集：多模态数据在大模型中的功能及配套硬件

- 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可以表示为视觉（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外观与空间关系；语言数据用于表达高层次任务目标、操作指令或人类意图；动作数据刻画机器人在连续或离散时间上的控制信号，如关节角度、关节速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/力矩反馈以及接触信息等。

数据维度	视觉	语言	动作	状态
所需硬件	RGB 或 RGB- D 相机	/	关节	传感器（力矩/六维力/电子皮肤等）
应用环节	刻画智能体所处环境及目标对象的外观与空间关系	表达/理解高层次任务目标、操作指令或人类意图	刻画机器人在连续或离散时间上的控制信号，如关节角度、关节速度或末端执行器位姿	状态数据包含机器人自身的本体信息与环境状态，例如关节状态、力/力矩反馈以及接触信息等

- 这类多模态数据在时间维度上通常呈现出显著的关联性：视觉与状态观测共同决定动作选择，而动作执行又反过来改变环境结构及后续感知结果，从而形成感知-决策-行动的闭环过程。因此，具身数据在采集过程中需要同时满足多模态同步、时序一致性以及物理可行性等约束，这也显著提高了数据采集与标注的复杂度。

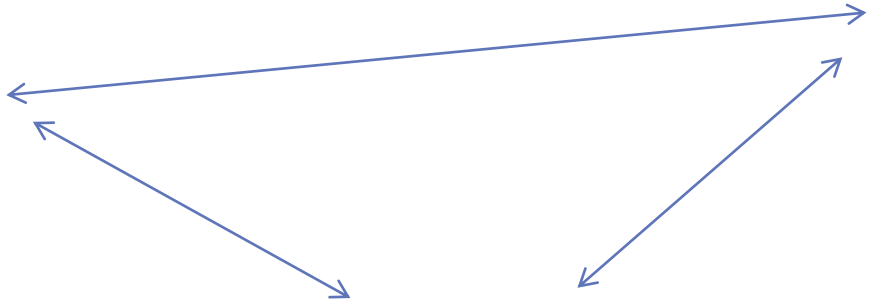


资料来源：丁贵广等《具身智能数据采集与处理综述》，国海证券研究所

仅供内部参考，请勿外传

数据采集产业：硬件成熟化+平台生态化+相关职业完善化

- 数据采集层承担从物理世界到数据世界的翻译功能。将人类真实场景操作数据转化为高质量、多模态的原始数据，提供具身智能所需数据信息。
- 数据采集层包含采集团队、硬件设备以及采集平台，三者属于协同关系，其中，采集团队保证采集的质量、效率和可重复性；硬件设备提供高保真、多模态、高精度实时同步的物理感知能力；采集平台统一调度、控制、记录、监控所有采集任务，最终产出带完整元数据的原始多模态数据包。



硬件设备

定义：包括机器人本体，视觉、力觉、触觉传感器硬件等

具体内容：
先对所有传感器进行参数校准
所有传感器同时工作
对原始数据进行无损或有损压缩

采集平台

定义：低延迟遥操作传输、多模态数据同步录制、实时监控仪表盘，以及标定工具和数据上传管道。

具体内容：
通过采集软件平台，将设计好的任务列表推送到各个采集终端；
采集各数据流并分文件记录；
提供实时可视化界面上传中央数据集群

相关职位

定义：包括遥操作专家、动捕演员、任务设计师、质量审核员等
核心：保证数据质量，如专家提供黄金标准示范，审核员实时剔除坏数据

具体内容：定义采集要完成的动作、任务清单；
操作员、专家等进行执行或记录；
质量审核员实时进行质量审核；
轻量级预标注等

仅供内部参考，请勿外传

数据加工产业：如何获得高质量数据

- 数据加工层的核心是将采集来的原始、多模态、非结构化数据转化为高质量、带标注、对齐好的训练就绪数据集。主要分为清洗、标注、治理、管理平台等。
- 具身数据采集成本高，可通过数据增强低成本扩充数据集，提升模型泛化性，常见的有视觉增强、物理参数随机化、运动增强和场景增强，增强后的数据保留原始标签并标记。
- 加工后的数据进入模型训练前，需要生成数据卡片，产出直接用于训练、可复现、可追溯的标准数据集。

原始数据	多模态、非结构化数据，量大，充满噪声和冗余
数据清洗	剔除异常帧、静置片段，修复缺失/乱序时间戳，多模态时空对齐，产出干净、连续的中间数据流
数据标注	选取最具代表性的样本进行精细标注——包括语义分割、动作切分和自然语言指令等，产出带丰富标签的高价值子集
数据治理	通过质量评估，去重去偏、合规脱敏、血缘追踪、访问控制等，产出高质量数据集
可以直接训练的高质量数据集	通过数据集构建与版本管理，产出可以直接训练高质量数据集

数据增强

通常在标注后、治理前，具身数据增强必须保持物理一致性。

·扩充数据集规模，降低采集成本
·提升模型泛化性，让模型适应不同光照、纹理、物理参数等变化
·弥补采集偏差，例如训练集中只有“红色方块”，通过颜色增强得到“蓝色方块”
·增强后的数据必须继承原始标签并标记

数据集构建/版本管理

将加工后的数据组织成规范化、可复现、可追溯的数据资产

·标准化输出：让数据格式、元数据、目的结构统一
·版本控制：记录每一次数据变更，支持回滚和对比
·可复现性：保证同一数据集版本在不同时间、不同机器上产生相同的训练效果
·数据卡片（Data Sheet）：提供详细文档，包括来源、传感器、任务分布、标注规则、伦理声明等

资料来源：亿欧智库，国海证券研究所

仅供内部参考，请勿外传

数据：具身智能大模型需要的数据类型

- NVIDIA GEAR团队负责人朱玉可认为，数据可以以三层金字塔的策略划分，底层是海量但被动的互联网人类视频数据，中层是可无限生成的合成数据，顶层是真实机器人数据，不应依赖单一数据源，而应以异质方式汇集三种数据，同时让世界模型扮演“数据海绵”的角色，吸收并整合所有类型的数据。

机器人本体的动作数据：较为珍贵，然而过去的方法普遍依赖于遥操作人工采集，效率太低而且价值不菲。来自真实世界传感器和硬件的数据没有现实差距，但每台机器人和每个人一样，一天只有24小时，可用时间可能远小于24小时

仿真合成数据：仿真引擎本身和真实世界有较大差距

UMI方法：大大降低了数据采集的难度，但仍然依赖于专门的数采人员进行采集。Egocentric数据相关的研究现在很热门，但面临着人和机器人因为结构不同导致的一系列retargeting问题



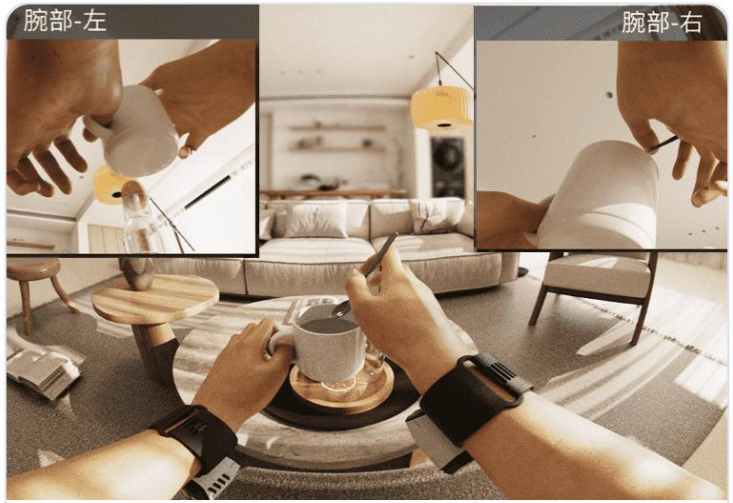
真机数据

真实环境多模态数据采集，覆盖环境感知、关节运动、触觉信息全维度



仿真数据

高保真物理与渲染管线生成大规模合成数据，低成本数据增广，加速模型迭代与泛化



人类演示数据

可穿戴设备与第一视角采集系统，构建可迁移的技能范式与行为基准

仅供内部参考，请勿外传

数据类型之间的效果比较

- 现有具身智能数据采集方式可归纳为 4 种核心范式：人类示教驱动、仿真环境驱动、真机跨平台数据聚合和世界模型生成，并在收集成本、数据质量以及规模方面存在较显著的差异。

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






仅供内部参考，请勿外传

人类数据采集方式 - 以英伟达为例

- EgoScale–第一视角人类视频
- 从第一人称视角的人类视频中学习，通过“预训练获取人类知识→对齐训练压缩知识→后训练表达知识”的三阶段方案，使得机器人仅需不到1%的真实机器人数据就能完成复杂操作任务

  通过跟踪人类手腕和手指在三维空间中的运动，即把人想象成一个机器人，头部运动就是动作空间，如此将第一人称视角视频转化为训练数据。

图：英伟达EgoScale演示示例



- SONIC–人类动捕数据
- 核心思路：首先将人类运动重定向到特定人形机器人的形态上，生成对应的动捕数据库，然后将运动跟踪作为强化学习的训练目标

- 优势：SONIC的关键在于动作的自然流畅度，这主要来自运动跟踪目标，让模型更好地模仿人类运动。这个模型可以接受遥操作、基础模型输出、甚至人类视频作为高级指令。

仅供内部参考，请勿外传

仿真数据采集方式 - 以英伟达为例

国海证券 SEALAND SECURITIES

- 通过机器人仿真，开发者可以在物理精准的数字孪生环境中训练、测试和验证机器人的行为——其中可包括仓库和工厂等场景的数字化表示，并在部署之前，通过机器人学习和可重复的测试场景进行验证。同样的方法也可以扩展到多机器人集群，使其能够基于实时生产数据、传感器输入以及推理能力，理解并与工业设施进行交互。
- 通过英伟达Isaac Sim平台，开发者可以从多种数据源中摄取数据，如计算机辅助设计 (CAD)、统一机器人描述格式 (URDF)，或通过 NVIDIA Omniverse NuRec 和 Isaac TeleOp 捕获的真实世界场景——并将其转换为 USD。随后，开发者可以通过指定材质、启用物理属性以及配置机器人和传感器模型来搭建仿真场景，然后将这些机器人与 NVIDIA Isaac Lab 结合用于机器人学习，并在 Isaac Sim 中完成仿真。

图：英伟达官网展示的Isaac Sim基础功能

使用 NVIDIA Omniverse NuRec 进行神经重建和渲染

将真实的传感器数据转化为交互式仿真使用 NuRec使用基于 3D Gaussian Splatting 的渲染来提高效率和准确性。

- 神经渲染入门
- 如何在交互式模拟中即时渲染真实场景
- 3DGUT：在 Gaussian Splatting 中启用扭曲的摄像头和二次光线
- 仅使用智能手机在 NVIDIA Isaac Sim 中重建场景
- 借助 NVIDIA Isaac Sim 和 World Labs Marble 更快地模拟机器人环境

逼真的物理模拟

对作为物理 AI 基础的对象和系统的物理行为进行建模。

Isaac Sim 可以模拟刚体和车辆动力学、多关节关节、SDF 对撞机等，实现逼真的物理模拟

- 物理仿真基础
- 传感器仿真入门指南
- NVIDIA® PhysX®
- 牛顿

可扩展的合成数据生成

利用合成数据加速 AI 模型训练。

通过随机化照明、反射、颜色以及场景和素材位置等属性来生成训练数据。

- 合成数据生成用例
- Omniverse Replicator 入门指南
- 使用合成数据扩展动作识别模型博客

仅供内部参考，请勿外传

应用现状 - 布局、竞争：数据采集的行业现状与商业闭环

仅供内部参考，请勿外传

大模型路径及进展：各厂商进展如何?

表：部分本体厂大模型、路线及数据路径等

公司	核心模型/平台	路线类型	数据路径	工具
英伟达	Cosmos 世界模型+Isaac 仿真框架+Isaac GR00T N系列模型	VLA + 世界模型	利用了大规模混合数据源：用于世界知识学习的互联网规模人体视频+用于物理对齐的真实遥操作数据+为安全扩展学习而生成的合成数据	搭配Isaac-Teleop遥操作硬件、SONIC等
谷歌DeepMind	Gemini Robotics 2 + ER 2 + On-Device 2	VLA	/	/
Figure	Helix系列+Index预训练驱动	VLA	多机器人、多任务、多操作员高质量操作数据集+机器人视觉数据+互联网规模数据	通过Index平台获取用户视频数据等
Physical Intelligence	π系列模型+RLT	VLA + 强化学习token改进精细操作	互联网规模预训练数据+开源机器人真机数据+高质量后训练真机数据等	机器人硬件本体等
宇树	UnifoLM-WMA-0 + UnifoLM-VLA-0	WM+VLA	高质量真机器人数据训练+继续预训练等	机器人操作数据+继续预训练等
智元	GO系列+Genie Sim仿真平台 + WAM（World Action Model，原生世界—动作模型）	VLA+WAM	GO系列：通过真机数据+互联网数据采集+基于已有真机数据生成隐式数据 Genie Sim：生成式仿真训练平台 WAM：遥操+无本体+Rollout回流+失败数据	通过genie studio等一站式开发平台，整合智元开源的百万级真机与仿真数据，提供自研及主流基座模型，内置完整仿真工具链，提供高保真仿真数据合成等
银河通用	AstraBrain +WAM-TTT (World-Action Model Test-Time Training)	VLA+WAM	合成仿真数据为主、真机数据为辅；人类示范数据预训练	/
星海图	G 0.5	VLA	/	通过大规模、多任务预训练，让模型形成可迁移、可组合的操作智能
千寻智能	Spirit AI系列	VLA	海量互联网视频数据+VLM及仿真数据+遥操作数据+真实任务数据，真实世界交互数据涵盖丰富的长尾复杂场景与多模态操作样本	2026年将积累100万小时真实世界交互数据

资料来源：各公司官网，腾讯科技公众号，国海证券研究所

仅供内部参考，请勿外传

商业模式：数采工厂+政府合作项目，让数据的飞轮先转起来

- 根据中国信通院，当前具身智能训练场在建设主体、建设内容和组织模式上呈现政府、企业双线推进的鲜明特征，逐步形成“政府引导+企业主导”“综合平台+行业垂直”双轮驱动的发展格局。

表：部分本体厂数采工厂布局

公司	数采工厂	投用情况	投用规划
智元	上海浦东具身智能数据工厂	已投运	日均采集3-5万条
	南湖工业具身智能数据采集中心	/	全年实现60万小时数据采集
宇树	重庆人形机器人数据采集中心	2026年3月签约	/
优必选	江西九江人形机器人数据采集与训练中心项目	/	/
	广西防城港人形机器人数采与测试中心和人工智能科创教育示范项目	/	/
	四川自贡数投人形机器人数据采集中心项目设备采购项目	已投运	/
	广西具身智能数据采集及测试中心设备采购及安装项目	/	/
乐聚	长三角一体化示范区智能机器人训练中心	已启用	年产数据可超200万条
	青岛市人形机器人数据采集训练场	已启用	年产数据100万条
帕西尼	天津市帕西尼具身智能超级数据工程	已投用	/
	宿迁具身智能全模态超级数据采集工厂	/	/
	自贡具身智能全模态超级数据采集工厂	/	
	武汉具身智能全模态超级数据采集工厂	/	
	赣州具身智能全模态超级数据采集工厂	/	

仅供内部参考，请勿外传

数据采集硬件+平台化产品，各厂商加速全链路布局

- 当前，各具身智能公司正在持续构建高精度物理感知（传感器）- 海量全模态数据（数据集）- 泛化智能决策（大模型）- 多场景落地的具身智能核心生态闭环，以期使具身智能体快速理解深层复杂环境，有望提升任务泛化能力，助推具身智能领域跨越式发展。

图：帕西尼具身智能生态闭环



资料来源：帕西尼官网

仅供内部参考，请勿外传

智元-觅蜂：数采硬件+数据商城+仿真平台

- MEgo系列是觅蜂自主研发的无本体采集终端产品线，包含 MEgo Gripper（二指夹爪）和 MEgo View（头戴式视觉设备+腕部相机）两款独立产品，两者既可独立使用，也可协同配合，为具身智能训练提供多视角的高质量数据支撑。







双指、三维触觉夹爪	重量 g	续航 h	鱼眼镜头视角（FOV）	触觉阵列空间分辨率 mm	力分辨率 N
	587（Ultra 版） 540（标准版）	2.5	200	1.6	0.05

输出信息	3摄 RGB 图像	深度信息	IMU	空间定位轨迹	麦克风音频	触觉信息	夹爪角度	夹爪开口距离

MEgo View	头部相机	腕部相机	续航 h	鱼眼镜头视角（FOV）	触觉阵列空间分辨率 mm	力分辨率 N
	鱼眼相机*3+1双目相机*1组	广角相机（左右各 1 ）	2	200	/	/

输出信息	7摄 RGB 图像	深度信息	IMU	空间定位轨迹	麦克风音频

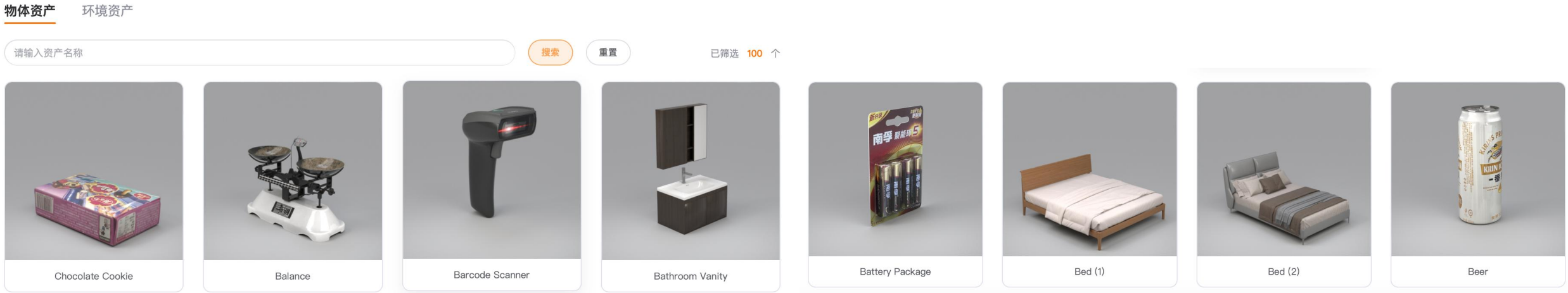
仅供内部参考，请勿外传

智元-觅蜂：数采硬件+数据商城+仿真平台

- 数据训练场景：截至2026年10月，平台构建起包括不同场景（家居/工业/餐饮/办公/商超）、数据类型（真机/仿真）和触觉信息（有/无），并建立具体的场景和动作指令，如厨房清洁等。



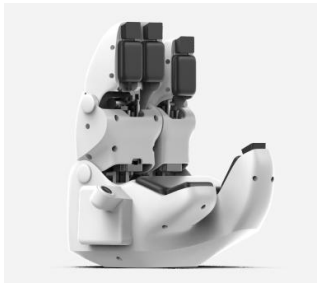
- 数据资产：截至2026年10月，已搭建8000+物体资产与仿真场景，9大类材质、180+类产品、12类原子技能，实现用户侧重建、专家轨迹获取、评测



仅供内部参考，请勿外传

帕西尼：数采硬件+数据商城+仿真平台







表：帕西尼部分数据采集产品及参数

产品	重量 g	自由度	多维阵列式触觉传感器	触觉传感器精度 N	触觉数据类型	相机	腕部定位方案	腕部定位精度
三指手套	650	4	指尖、指腹、掌心无盲区覆盖+帕西尼6D霍尔阵列触觉传感器	0.1	合力、分布力阵列	RGB*2	VR头显 + 手柄，支持其他定位方案	≤ 5 mm(VR头显+手柄)
三指执行末端（带电机驱动）	836	4	帕西尼6D霍尔阵列触觉传感器	0.1	合力、分布力阵列	RGB*2	/	/
五指数采手套专业版	325	17	指尖、掌心全域覆盖+触觉信号数1743个	0.1	合力、分布力阵列	RGB*1	VR 头显 + 手柄 (可支持多种定位方案)	≤5mm (VR头显 + 手柄)

资料来源：帕西尼官网，国海证券研究所

仅供内部参考，请勿外传

帕西尼：数采硬件+数据商城+仿真平台

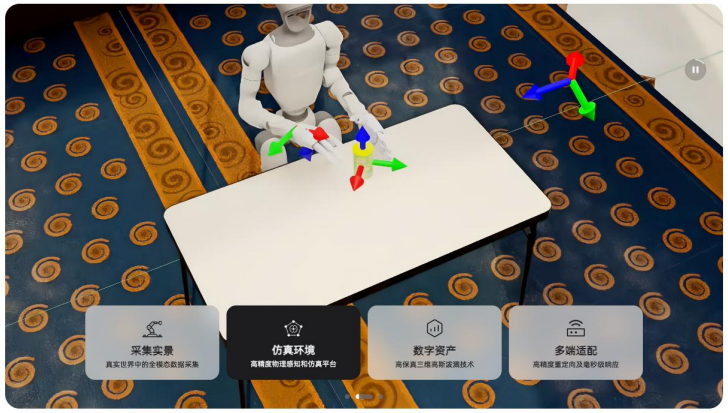
- OmniSharing DB 全模态数据集
- 帕西尼推出的全模态数据集，包含多维度触觉、多视角视觉、语音、文本、自体感知、空间轨迹的全模态数据，解决具身智能体不同场景下的快速泛化难题。

数据集效果展示

全方位展示验证数据集的实际应用效果和技术价值



采集数据



仿真环境

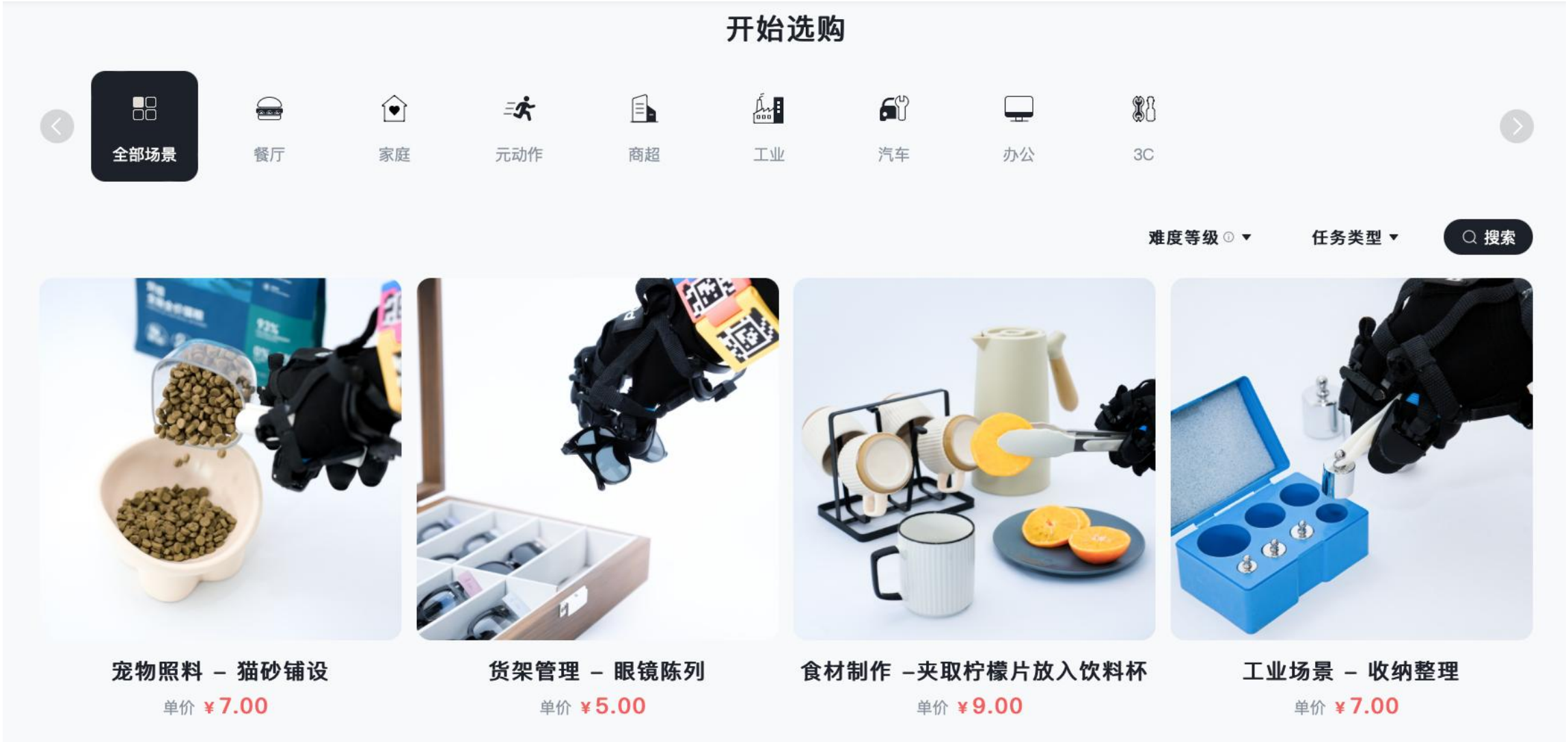


数字资产

仅供内部参考，请勿外传

帕西尼：数采硬件+数据商城+仿真平台

- 搭建全模态具身智能数据云商城，提供海量全模态、高自由度、高质量的真实物理交互数据



资料来源：帕西尼官网

仅供内部参考，请勿外传

数据采集硬件供应 - 手部硬件类：他山科技

◆ TS-V视触融合训练平台

集成视觉触觉多模态传感技术，实现机器人在复杂场景下对任何物体的稳定抓取。系统支持实时感知接触状态与施力状况，自适应控制夹爪动作，抓取更加稳健、柔顺、泛化。



◆ TS-VT视触融合训练平台（数据采集）

提供专业的多维触觉数据记录、标注、清洗与导出功能，适配多款灵巧手及夹爪等末端执行器。用户可轻松采集三维力、接近觉、触点变化等数据，大大降低了数据采集的成本，为训练触觉感知算法与评估机器人交互性能提供可靠基础。



◆ 触觉模拟仿真平台

实现了从数据实时可视化、实时三维力到触觉模态建模训练的完整闭环，精准还原触觉传感器的真实触感信息，缩小Sim2Real的差异，是推动触觉传感落地的重要工具。



仅供内部参考，请勿外传

数据采集硬件供应 - 外骨骼类：灵巧智能

- Dexcap外骨骼数据采集系统，提供丰富且精确的数据采集，超越低自由度设备的数据局限性，双臂各9自由度、双手各21自由度和腰部5自由度，并在精准捕捉、支持从算法到硬件的全链路开发、便携式穿戴设计等具体应用上进行全面优化。



产品参数
重量：4.2kg
末端位置误差：X轴 ≤ 0.5mm，Y轴 ≤ 0.5mm，Z轴 ≤ 0.1mm

表：灵巧智能外骨骼核心参数

1000Hz通信频率	实现超低时延 数据输出与动作反馈
高精度动作捕捉	65个磁绝对式编码器实现精准动作捕捉
65+全域超高自由度	从手指微操到腰部扭转 捕捉连续流畅复杂的人体运动
模块化组合设计	人体上肢DexCap/双臂UL/双手GL可根据不同应用场景按需配置灵活扩展
8-12h续航	便携式穿戴人体工学设计，保证研究工作的连续性

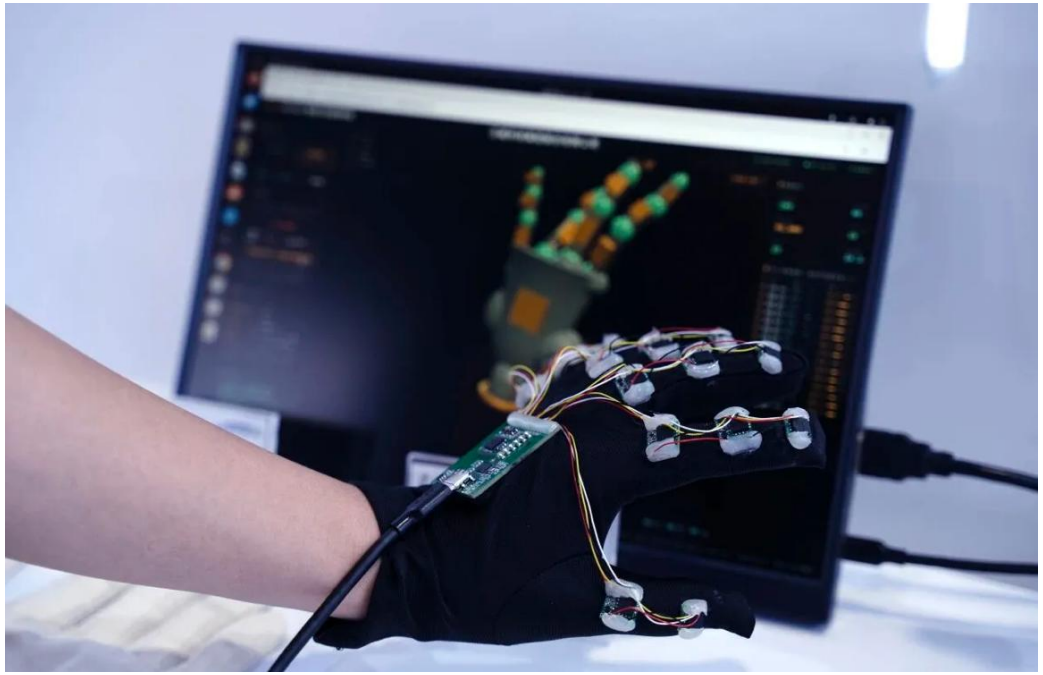
资料来源：灵巧智能官网，国海证券研究所

仅供内部参考，请勿外传

数据采集硬件供应 - 柔性类：汉威科技

- 汉威科技布局用于机器人灵巧手的视觉+触觉多模态数采终端、用于数据采集与训练的国内首款弹性动捕服等。

视觉+触觉多模态数采终端



机器人数据手套，整合柔性压力传感器、弯曲传感器、惯性传感器，可以实时采集手指尖等部位的压力分布情况和手部抓握、反转、按压等动作，为机器人灵巧手数据采集、动作训练提供多维感知数据。

用于数据采集与训练的国内首款弹性动捕服



该产品于2025年9月推出，产品全身关键关节如肩关节、肘关节、膝关节等部位分布了近30个高精度弹性可拉伸传感器，采用高分子弹性材料，拉伸率超过50%，分辨率达到0.05%，配备数据处理软件等，用于动作示范、训练学习和远程操控。

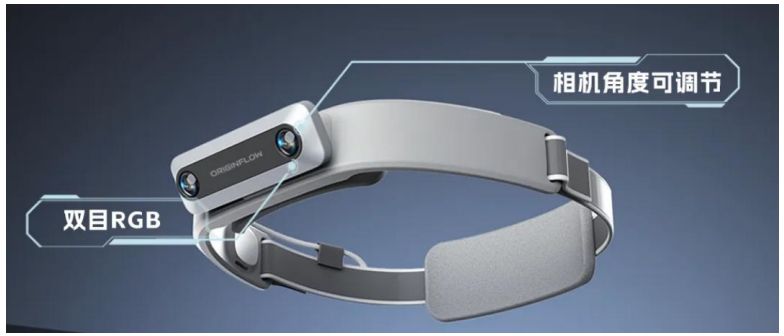
资料来源：汉威科技官网，汉威科技公众号

仅供内部参考，请勿外传

数据采集硬件供应 - 穿戴电信号类：Originflow

- 公司推出的便携式多模态数据采集套件，从人的真实操作中，捕捉第一视角视觉、运动状态和神经肌电等多模态信息，将人类在物理世界中积累的经验，转化为机器可以学习、复用和迁移的数据。

整套产品由智能计算主机、头戴式摄像机以及智能肌电手环组成，可以同步采集第一视角视觉、头部运动、手部运动和肌电信号，让一次发生在真实世界中的操作，以更加完整和同步的方式被记录下来。







仅供内部参考，请勿外传

相关标的

- 我们认为人形机器人处于技术突破、商业化落地初期，后续有望迎来量产空间，维持人形机器人行业“推荐”评级。**建议关注：**1）本体：宇树科技、优必选、极智嘉；2）垂直场景本体应用：杰克科技、永创智能；3）数据采集硬件：凌云光、汉威科技、福莱新材等。

表：相关标的梳理（截至2026年10月12日）

标的	代码	总市值（亿元*）	净利润（亿元）		
			2026E	2027E	2028E
宇树科技-W	688836.SH	1,703.97	6.62	10.32	14.43
优必选	9880.HK	340.80	-3.20	1.80	8.00
极智嘉-W	2590.HK	90.94	-0.03	2.91	6.75
杰克科技	603337.SH	167.21	9.90	11.75	14.39
永创智能	603901.SH	52.48	2.70	3.25	3.96
凌云光	688400.SH	184.02	5.21	3.86	5.07
汉威科技	300007.SZ	91.98	1.83	2.20	2.57
福莱新材	605488.SH	80.22	1.15	1.70	2.43

*港股标的市值单位为亿港元

仅供内部参考，请勿外传

风险提示

- **人形机器人行业进展不及预期风险：**人形机器人行业硬件技术及软件技术（大脑及大模型、小脑及运控技术等）发展低于预期，将对人形机器人技术成熟的速度、后续商业化落地和放量节奏产生不利影响；
- **数采成本下降不及预期风险：**在行业发展初期阶段，数据采集成本下降存在不及预期的风险；
- **数据应用安全性及合规风险：**数据采集如果落地家庭端，涉及隐私保护、数据安全合规风险；数据集市场化交易中的资产权属与交易风险等；
- **相关行业标准制定与行业发展速度不匹配风险：**存在数据采集行业标准制定与完善无法匹配行业未来可能的较高发展速度的风险；
- **中美贸易摩擦风险：**若美国对人形机器人中国配套供应链厂商加强技术封锁和关税限制，将对机器人国产产业链中的企业技术革新造成不利影响，增加关税负担和经营成本等；
- **重点关注公司业绩不及预期风险：**数据采集行业处于起步探索阶段，存在相关公司短期内业绩兑现不及预期或实际业务贡献有限的风险；
- **研究报告中使用的公开资料可能存在信息滞后或更新不及时的风险：**人形机器人行业技术方案多样，迭代速度快，研究报告中使用的公开资料可能存在信息滞后或更新不及时的风险。

仅供内部参考，请勿外传

研究小组介绍

机械小组介绍

张钰莹，机械行业首席分析师，致力于前瞻研究及深度研究，擅长自上而下与自下而上相结合挖掘机械板块投资机会。
闫智，机械行业分析师，现阶段聚焦燃气轮机、PCB设备、金刚石等AI设备成长方向。

分析师承诺

张钰莹, 本报告中的分析师均具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立，客观的出具本报告。本报告清晰准确的反映了分析师本人的研究观点。分析师本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收取到任何形式的补偿。

国海证券投资评级标准

行业投资评级

推荐：行业基本面向好，行业指数领先沪深300指数；
中性：行业基本面稳定，行业指数跟随沪深300指数；
回避：行业基本面向淡，行业指数落后沪深300指数。

股票投资评级

买入：相对沪深300 指数涨幅20%以上；
增持：相对沪深300 指数涨幅介于10%～20%之间；
中性：相对沪深300 指数涨幅介于-10%～10%之间；
卖出：相对沪深300 指数跌幅10%以上。

仅供内部参考，请勿外传

免责声明和风险提示

免责声明

本报告的风险等级定级为R3，仅供符合国海证券股份有限公司（简称“本公司”）投资者适当性管理要求的客户（简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。客户及/或投资者应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司的完整报告为准，本公司接受客户的后续问询。

本公司具有中国证监会许可的证券投资咨询业务资格。本报告中的信息均来源于公开资料及合法获得的相关内部外部报告资料，本公司对这些信息的准确性及完整性不作任何保证，不保证其中的信息已做最新变更，也不保证相关的建议不会发生任何变更。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。报告中的内容和意见仅供参考，在任何情况下，本报告中所表达的意见并不构成对所述证券买卖的出价和征价。本公司及其本公司员工对使用本报告及其内容所引发的任何直接或间接损失概不负责。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露义务。

风险提示

市场有风险，投资需谨慎。投资者不应将本报告为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向本公司或其他专业人士咨询并谨慎决策。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，请公众投资者慎重使用未经授权刊载或者转发的本公司的证券研究报告。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息。本报告不构成本公司向该机构之客户提供的投资建议。

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。本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

郑重声明

本报告版权归国海证券所有。未经本公司的明确书面特别授权或协议约定，除法律规定的情况外，任何人不得对本报告的任何内容进行发布、复制、编辑、改编、转载、播放、展示或以其他任何方式非法使用本报告的部分或者全部内容，否则均构成对本公司版权的侵害，本公司有权依法追究其法律责任。

仅供内部参考，请勿外传

国海证券 SEALAND SECURITIES

国海证券·研究所·机械研究团队

心怀家国，洞悉四海

国海研究上海

上海市黄浦区绿地外滩中心C1栋国海证券大厦

邮编：200023

电话：021-61981300

国海研究深圳

深圳市福田区竹子林四路光大银行大厦28F

邮编：518041

电话：0755-83706353

国海研究北京

北京市海淀区西直门外大街168号腾达大厦25F

邮编：100044

电话：010-88576597