证券研究报告

东吴证券 SOOCHOW SECURITIES

物理AI：数据规模化叠加模型突破，物理AI产业迎来拐点

证券分析师：张良卫
执业证书编号：S0600516070001
联系邮箱：zhanglw@dwzq.com.cn
联系电话：021-60199793

证券分析师：周良玖
执业证书编号：S0600517110002
联系邮箱：zhoulj@dwzq.com.cn
联系电话：021-60199793

研究助理：姜标
执业证书编号：S0600126100009
联系邮箱：jiangb@dwzq.com.cn
联系电话：021-60199793

二零二六年十一月十四日

请务必阅读正文之后的免责声明部分

摘要

物理AI是AI应用的第三阶段演进。物理AI是继生成式AI、Agent之后AI的第三级演进，从输出信息、数字指令升级为输出实体动作，核心是掌握物理因果规律，面临Sim2Real虚实迁移挑战。已量产复制：AMR、工业巡检机器人；单场景量产：基地安保四足巡检机器人；探索阶段：人形机器人，尚未实现规模化部署。

Ego数据采集已实现标准化，将迎来规模化放量。数据采集包含遥操作、Ego、UMI、仿真数据四种数采方式，行业模型训练数据通常采用混合方案。EgoScale验证第一人称数据遵循Scaling Law，数据规模提升可稳定提升机器人性能。企业规划目标达百万-千万小时，但实际数据存量偏低，规划与落地存在缺口。

大语言模型将加速物理AI模型迭代。物理AI模型快速迭代，VLA模型学习视觉和语言指令到运动轨迹之间的映射，物理层面泛化能力较弱；世界模型学习环境如何随时间演化，重点是理解跨任务复用的物理规律；WAM模型是世界模型的演进，输入任务目标，输出是动作执行和未来世界状态的预测。大语言模型Chat-GPT6 Astra在任务执行过程中能够给出有价值的动作决策，超出行业预期，将加速物理AI模型发展，但仍存在执行速度较慢等瓶颈。

人形机器人是物理AI最合适的载体。物理AI的训练语料主要来自人类，比如Ego数据采集是人类的第一人称视角，用于训练的网络视频主要包含的是人类的行为动作，人形硬件天然适配模型和数据，训练效果较佳。

产业链公司梳理：风语筑：三维空间仿真数据资产，参控股搭建数采-仿真-训练闭环体系；**奥比中光：**全栈自研3D视觉，物理AI硬件感知入口；**福赛科技：**机器人零部件核心供应商；**五一视界：**仿真基础设施，51Sim高速增长，商业模式向训练片段订阅计费转型；**索辰科技：**CAE多物理场底座，“开物”产品矩阵；**宝通科技：**2026年7月和光轮智能成立合资公司，宝通科技持股70%，推进物理AI数据产品与工业机器人应用落地。

风险提示：物理AI模型能力进展不及预期；机器人商业化及规模放量不及预期；公开演示代表性不足；Sim2Real迁移效果不及预期。

目录



1、数据采集和处理：迎来规模化放量拐点

2、具身大脑：大语言模型将加速物理AI模型迭代

3、终端载体：人形机器人是物理AI最合适的载体

4、产业链公司梳理

5、风险提示

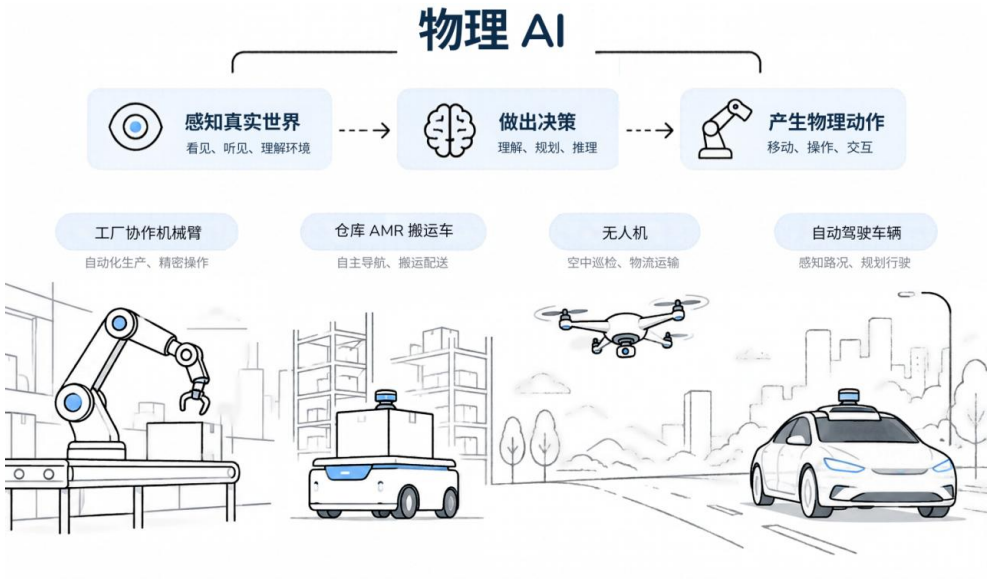
1、数据采集和处理：迎来规模化放量拐点

物理AI：让AI理解并应用物理规律实现实体执行

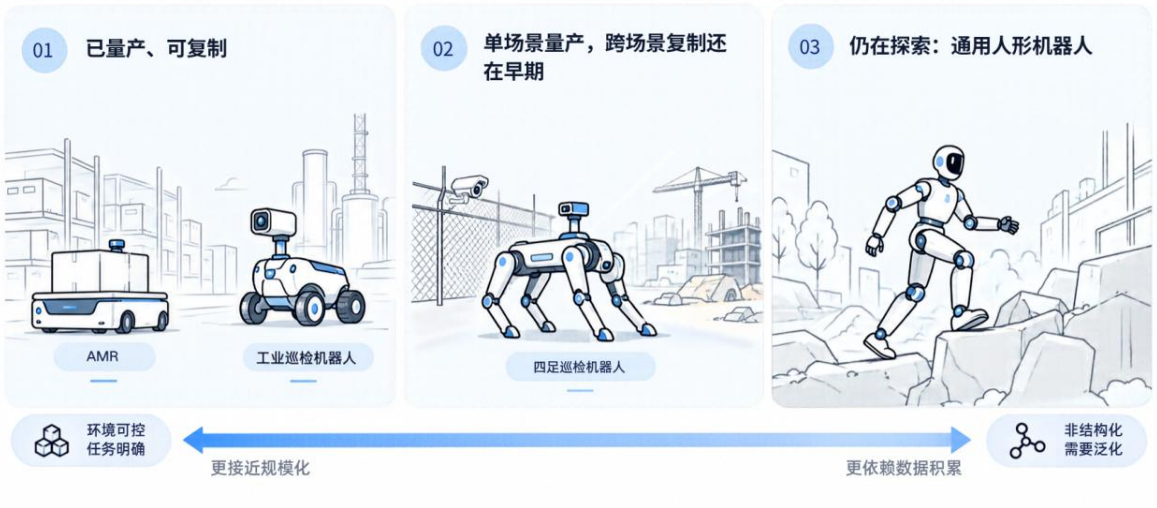
物理AI是指能够感知物理环境、依据物理规律推理并在现实世界自主执行动作、持续反馈学习的人工智能技术体系。物理AI的核心在于AI与物理世界的融合，其关键是让AI系统理解并应用重力、摩擦、材料特性等**物理规律**，实现从虚拟智能到实体执行的跨越。

物理AI覆盖场景广，各场景成熟度差别较大。按成熟度可分为以下三类：1）已经**量产且可复制**。比如自主移动机器人（AMR）、工业巡检机器人。这类场景环境相对可控、任务明确，边缘推理能力已经够用。全球范围看，Amazon等企业已经做到千台级部署；2）**单场景量产**，跨场景复制仍较早期。比如四足巡检机器人可以成熟应用在基地安保，但同一套系统去做建筑工地巡检，需要重新适配感知和规划模型；3）**探索阶段**。人形机器人较有代表性，目前还没看到人形机器人能在非结构化的环境里完成相对复杂的任务，并且规模化部署。

图：物理AI让AI理解并应用物理规律实现实体执行



图：物理AI覆盖场景广，各场景成熟度差别较大



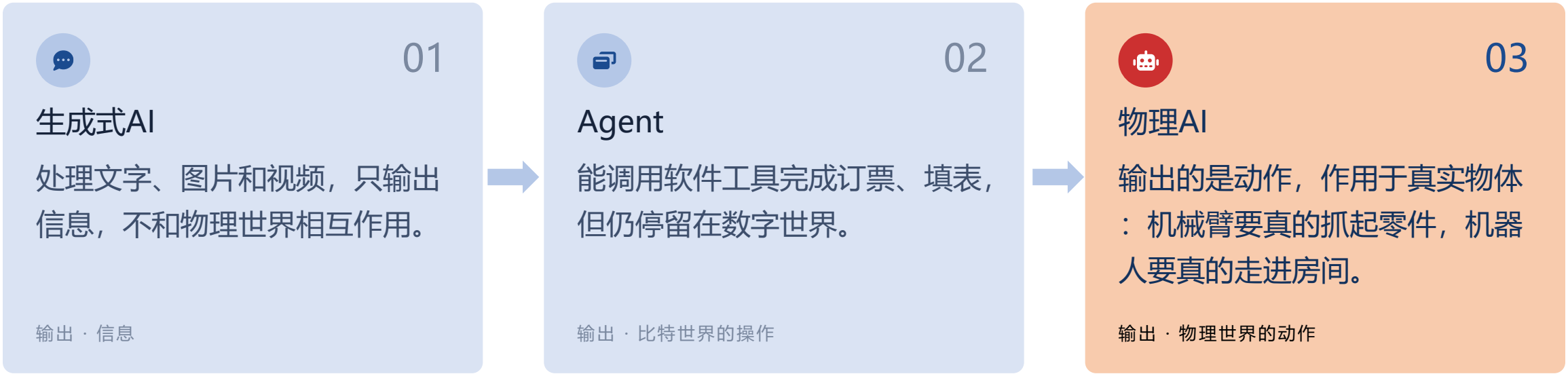
物理AI：AI演变路径从输出信息到输出动作

东吴证券 SOOCHOW SECURITIES

物理AI是生成式AI及Agent之后AI应用的第三级演进。1）AI应用的第一阶段是**生成式AI**，以文字、图片为处理对象，仅输出信息，不干预真实物理世界；2）第二阶段是**Agent**，可调用软件工具完成订票、填表等任务，但全部行为仍局限在比特数字世界；3）第三阶段迈向**物理AI**，模型输出不再只是文本或软件指令，而是可作用于实体的动作，比如驱动车辆转向、机械臂抓取、机器人移动等，真正实现和物理世界的交互。

传统AI处理统计相关性，物理AI以处理物理因果性为目标。传统AI给定输入，预测输出，处理的是输入和输出之间的**相关性**；而物理AI通过动作改变环境状态并接受反馈，必须处理**因果性**。数字世界表现良好的公司，进入物理世界会遭遇明显的Sim2Real鸿沟。

图：物理AI是生成式AI及Agent之后AI应用的第三级演进



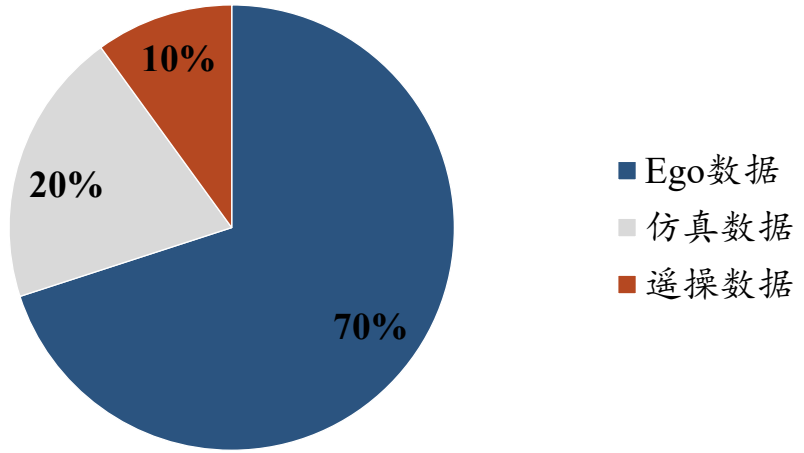
数据采集方式：四种方式分层互补

四种数据采集方式在质量、成本与规模化之间形成清晰分层，当前是互补关系。 1）遥操作数据质量最好，但设备与人力成本最高、规模化潜力一般；2）Ego与UMI以“无本体”换取规模，前者规模化能力、成本和效率较好，数据质量有限，后者较为均衡；3）仿真数据产能与成本最优，但质量较低，受Sim2Real制约。对应到训练环节，数据采集行业正走向“仿真+Ego打底、UMI补精度、遥操作补长尾”的组合。现阶段Ego是物理AI数据非常好的入口，若机器人部署数量提升（如10万台或100万台之后），大量真机数据会回流，数据飞轮转动，整个数据分布和逻辑会发生变化。

表：四种数据采集方式分层互补

技术	简介	核心优势
真机遥操	人类远程操控机器人完成任务，数据直接来自真机执行，全程需机器人参与，专注机器人特定动作；同步记录关节轨迹、力反馈、触觉与视觉画面。	数据质量高、成本高
Ego	通过头戴式第一人称视角设备，同步采集操作者视觉画面、IMU惯性数据、手部动作轨迹与空间位姿信息的多模态数据获取方式。	成本较低、规模化潜力大
UMI	通过手持夹爪接口由人类操作采集，无需机器人参与，聚焦可迁移的操作演示，但采集所用的末端夹爪、末端相机必须与机器人实际部署的硬件完全一致。	成本和质量较为均衡
仿真数据	在世界模型等虚拟环境中生成数据，但是无噪音，存在明显的Sim2Real问题。	产能上限极高，可生成失败场景数据

图：2026年具身模型训练数据以Ego为主（按数据时长测算）

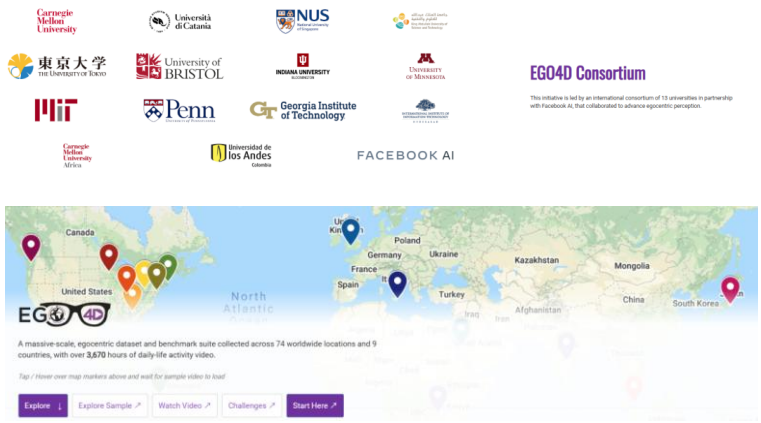


Ego数据：已实现标准化，需求端迎来规模化拐点

Ego4D数据集首次为第一人称数据建立了完整的行业标准。2022年2月，CMU、Meta和Google等顶尖机构联合发布Ego4D数据集，采用了统一的采集规范，使用消费级头戴相机，统一帧率和分辨率，在全球采集了超3000小时的日常视频，首次为第一人称数据建立了完整的行业标准，且定义了视线追踪、手部关键点等第一人称数据的多维度标注体系。

Ego数据需求端迎来规模化拐点。2026年5月，Genesis AI发布GENE-26.5，采用了20万小时高密度Ego数据；2026年8月，Dyna-2使用100万小时Ego数据预训练并验证Scaling Law趋势延续，单一场景要实现一定智能、通用泛化并保证成功率，需数十万至百万小时量级数据，若要实现跨场景通用，数据规模将大幅提升。数据需求方主要包括大模型公司、互联网大厂、机器人公司和灵巧手厂商。头部Ego数据供应商Maxinsights累计数据时长已超过200万小时，采集容量达到45万小时/月。

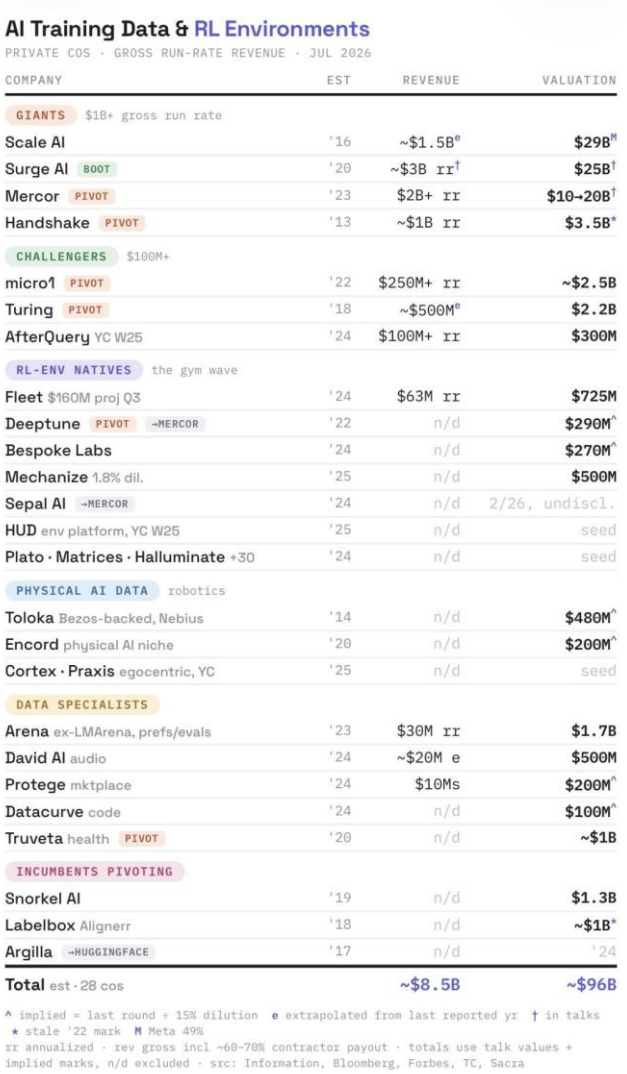
图：Ego4D实现Ego数据标准化



表：国内物理AI数据供应商数据时长可达百万小时级

公司	数据时长	主要客户	其他
觅蜂科技	100万小时高质量无本体数据，行业第一家拥有百万小时级、可对外售卖无本体数据的数据服务商	腾讯Robotics X、蚂蚁灵波	-
无问智科	累计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过2万小时	星动纪元、字节跳动、地平线等	2026年数亿元A轮融资，具身智能业务订单已达数亿元
光轮智能	累计交付超150万小时高质量人类数据，在优质场景的数据，复售率已经能够超过10倍	同时服务自动驾驶、具身智能和世界模型三大场景，客户包括英伟达、谷歌、字节跳动、阿里巴巴、银河通用、智元机器人等	2026Q1新增订单5.5亿元，超过2025年全年，估值突破150亿元

图：代表性AI训练公司

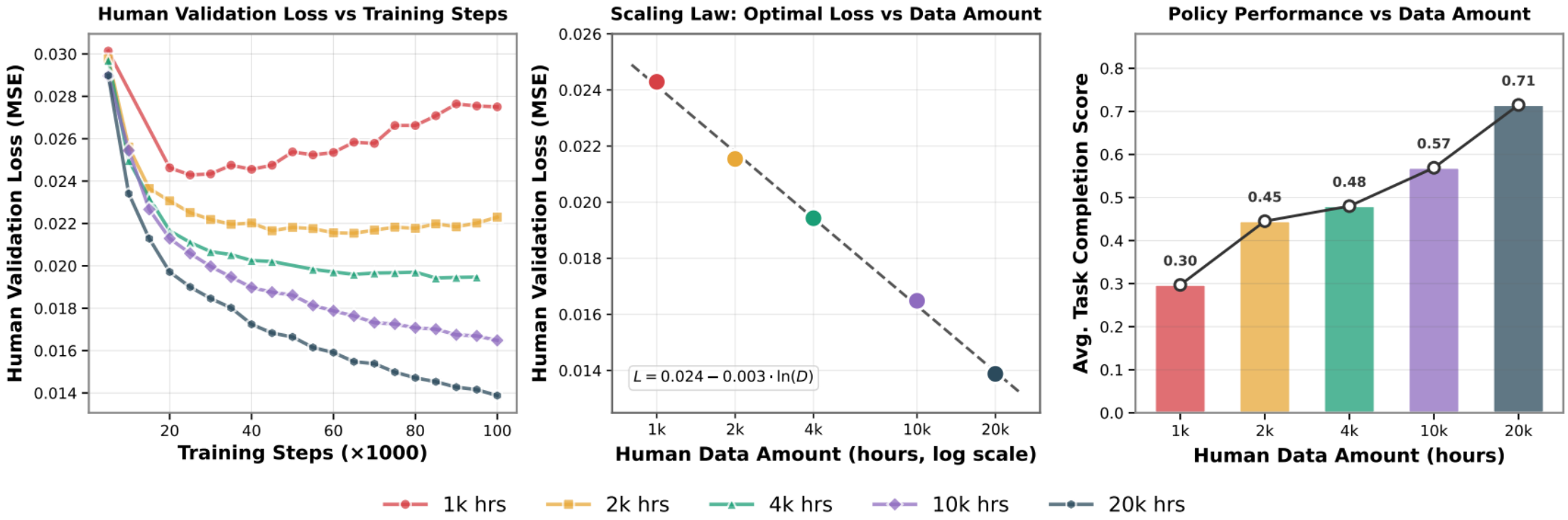


资料来源：深蓝具身智能，36Kr，东吴证券研究所

EgoScale：扩大数据规模可不断提升机器人通用能力

EgoScale证明扩大第一人称数据的规模可以不断提升机器人的通用能力。1）2026年2月NVIDIA发布EgoScale，证明了人类数据规模与机器人操控性能之间的对数线性关系，随着数据量的增加，模型的验证损失呈对数线性下降，而这种损失直接对应下游真实机器人的性能提升；2）EgoScale的研究表明，高质量的第一人称人类数据可以推动机器人的性能像大语言模型一样，通过增加数据量稳定提升。这给了行业一个明确的、可执行的路线图：只要不断扩大第一人称**数据规模**，就能不断提升机器人的**通用能力**。

图：扩充人类预训练数据量可稳定提升机器人性能



数据采量和存量：采量规划进入百万小时级，存量仍较少

行业数采规划已普遍抬升至百万乃至千万小时，但实际累计存量多在十万小时以内，规划与落地之间存在明显缺口。

规划端：京东计划两年积累1000万小时；觅蜂提出2026年实现千万小时级产能；亦数智能给出“2026年完成100万小时，未来三年迈向1000万小时”的路线。

存量端：千寻智能20万小时，鹿明与灵初各10万小时，多数企业仍在万小时级。同时，天奇股份以年产能1000万条卡位供给端。

表：部分企业数据采集存量和规划量差距明显

分类	公司采集类型	公司名称	目标数采量	当前累计存量
按数据类型	仿真	英伟达		暂未披露
		索辰科技		暂未披露
	动作捕捉	凌云光		1000小时
	UMI	鹿明机器人	100万小时	10万小时
		松灵机器人		1万小时
	Ego	简智机器人		100万小时
		京东科技	1000万小时	
按数采公司类型	灵巧手	灵心巧手		3000万条灵巧操作数据集
		戴盟机器人	100万小时	1万小时
	三方数采商	海天瑞声	100万小时	
		觅蜂科技	1000万小时	
		数据堂		每周5000小时
	机器人公司	灵初智能	100万小时	10万小时
		它石	100万小时	暂未披露
		星海图	1000万小时	
		亦数智能	2026年完成100万小时，未来三年迈向1000万小时	
	垂类公司	千寻智能	2026年100万小时	20万小时
		天奇股份		年产能1000万条

资料来源：钛媒体，新华财经，企业官网，东吴证券研究所

Ego数采：方案多样化

Ego数采包括手机支架方案、双目 RGB-IMU头戴模组、4-6目SLAM头显、模块化套件以及定制/改型开发方案。1）手机支架方案：成本较低；2）双目RGB-IMU方案：整套头显价格在1000-1500美元之间，第一视角成像效果较好；3）SLAM头显：单机成本1000-2500美元；4）模块化套件：包括头戴采集单元、腕部相机和夹爪等，整套套件在1000-6500美元之间，可在同一坐标系下同步采集头部、手部和执行器数据；5）定制/改型开发：按需定制，项目周期风险高。

表：Ego数采一般包括五种方案

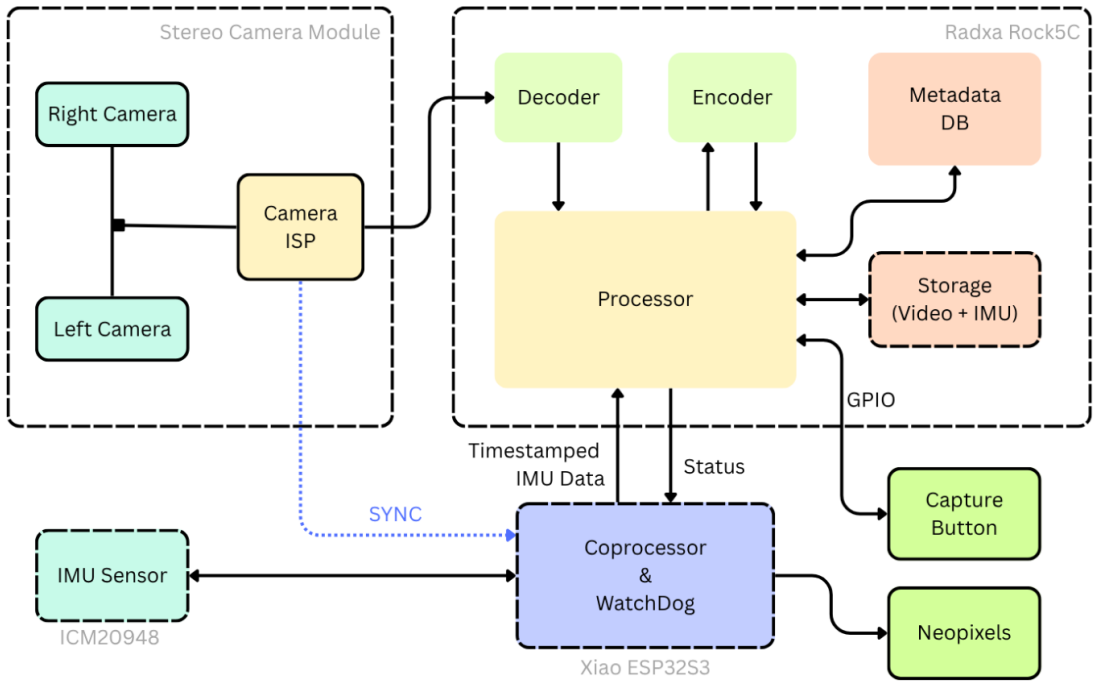
品类	典型配置	参考价格	优势	短板
手机支架方案	智能手机搭配头戴支架，采集 APP	手机+支架	易于规模化、成本低、使用者上手门槛低、支持云端上传	无硬件同步、卷帘快门、依靠应用程序处理、视场角有限
双目RGB-IMU头戴模组	2颗全局快门鱼眼相机（1080p-200 万像素）+六轴 IMU；USB 接口或本地板载SD卡存储	模组：150-400 美元；整套头显：1000-1500 美元	第一视角成像效果好，可通过双目获取深度，数据链路简单	全分辨率双目录制通常仅支持30fps；IMU频率200-500Hz，时间戳依赖主机生成
SLAM 头显（4-6 目）	4颗全局快门追踪单目相机+2路 RGB、IMU、板载 SoC（高通 XR2平台或瑞芯微 RK3576/RK3588），USB3/GMSL	单机：1000-2500 美元；企业定制配置需单独询价	端内输出6DoF位姿、覆盖范围广、支持手部追踪；常搭载红外LED，可对手持物体追踪	帧交错排布；追踪相机分辨率仅640×480；默认压缩录制；SDK封闭
模块化套件	头戴采集单元+腕部相机+UMI、夹爪、计算单元，共用统一时间基准	整套套件1000-6500 美元	在同一坐标系下同步采集头部、手部、末端执行器数据	每个模块都会增加重量与故障点；不同模块之间的同步效果参差不齐
定制/改型开发	在合作方平台上搭载自研传感器，或全新设计	一次性工程费用（NRE）：10000-100000 美元以上；最小起订量通常 10-5000	时序、架构、传感器、形态完全按需定制；可拥有设计知识产权	项目周期风险高、一次性研发成本高；认证与售后维护负担重

头戴式Ego设备：SBC和模组相机成本占比较高

东吴证券 SOOCHOW SECURITIES

目前头戴式Ego设备组件价值量集中在SBC和相机模组。1）以FPV Labs推出的开源低成本双目全局快门Ego采集方案Ego-OSCAR为例，其主要组件包括一块负责端侧编码与存储SBC、一套硬件同步双目相机、SD存储卡、散热器、一颗六轴IMU、移动电源以及一块用于时间戳打标、人机交互的实时微控制器等。2）在上述组件中，SBC的价值量占比最高，约为34.48%，其次是相机模组，成本占比约为31.97%，SD存储卡占比约为12.54%。

图：头戴式Ego设备主要由相机模组、SoC和IMU等组成



表：Ego-OSCAR成本占比较高的组件包括SBC和相机模组等

零部件	金额（元）	价值量占比	功能
Radxa Rock 5C（2GB）SBC	550	34.48%	采集、编码、本地存储
Dexcin USB 双目相机	510	31.97%	双目图像采集、全局快门
256GB SD存储卡	200	12.54%	支持约16–18小时录制
Radxa配套散热器	65	4.08%	散热管控
ICM-20948六轴IMU	65	4.08%	惯性传感
Xiao ESP32-S3	55	3.45%	人机交互、微控制器
10000 mAh移动电源	40	2.51%	整机供电
其他（支架、辅料等）	110	6.90%	接线、壳体、头戴支架、指示灯等
合计	1595	100.00%	

资料来源：Ego-OSCAR，东吴证券研究所

数据处理：竞争力取决于各环节的执行与衔接

数据最终能否转化为模型能力，取决于数据处理环节的执行质量与衔接效率。采集决定物理真实性，扩展缓解场景稀缺，标注打通动作与结果，筛选治理控制分布偏差，部署反馈则把线上失败案例回流为增量数据，形成“数据飞轮”。任一环节缺失，都会以模型缺乏物理认知、过拟合、跨传感器对齐失败或性能停滞的形式传导至最终效果；随着数据规模抬升，瓶颈将从“采不到”转向“用不好”。

表：数据处理过程包含多个环节

阶段	定义	主要产出	缺失风险
数据采集	通过遥操作、第一人称视角采集或自主运行，获取真实世界中的机器人操作数据与人类示范数据	原始多模态示范数据	模型缺乏对真实世界物理规律与环境的认知基础
数据扩展	借助仿真、生成式模型或合成增强，对有限的真实数据集进行扩增，以提升数据规模与多样性	规模更大、更多样的训练数据集	模型过拟合于少数可负担的采集场景
数据标注	为同步的多模态传感数据（视频、激光雷达、力觉、关节状态）标注空间、时间与动作层面的元数据	结构化、可直接用于模型训练的数据	模型无法将动作与结果关联，也无法跨传感器对齐数据
数据筛选与治理	对数据集进行筛选、去重与重新配比，使其覆盖模型部署时将会遇到的多样性场景与边缘情况	经过优化、具有代表性的训练集	模型在冗余或不均衡的数据上训练，实际部署时表现不佳
部署反馈	采集已部署机器人的运行数据与失败案例，并将其回流至数据采集、标注与治理环节	持续改进闭环（“数据飞轮”）	模型性能停滞不前，无法在迭代中累积提升

2、具身大脑：大语言模型将加速物理AI模型迭代

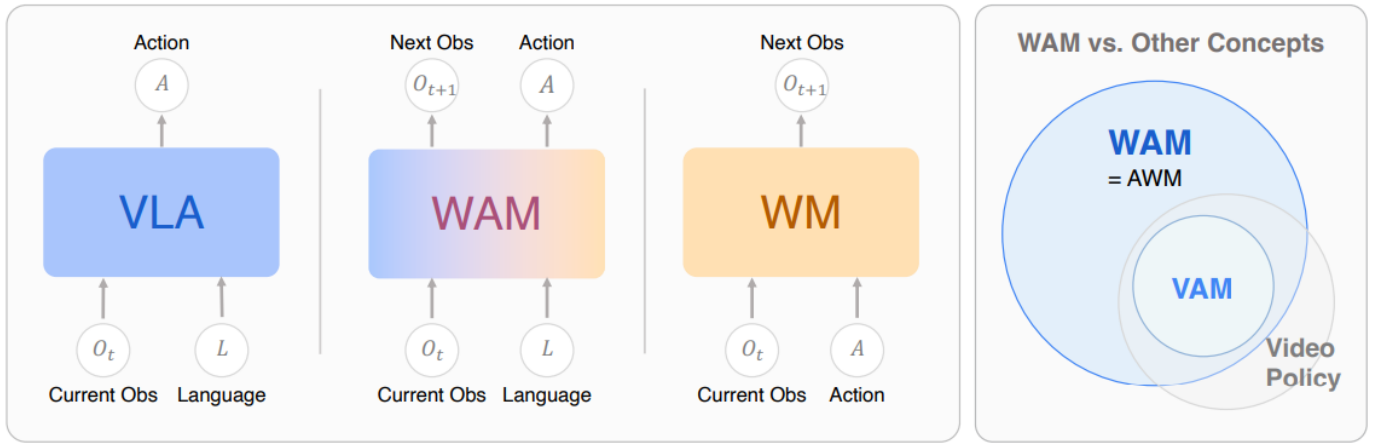
模型：从VLA到WAM再到大语言模型

1）视觉语言动作模型（Vision-Language-Action，VLA ）：主干网络建模的是语言与感知，而非世界动力学规律，目标是学习视觉和语言指令到运动轨迹之间的映射，**需要同一任务高度相似的示教样本**，但面对未见过的行为与环境时，**物理层面的泛化能力较弱**；

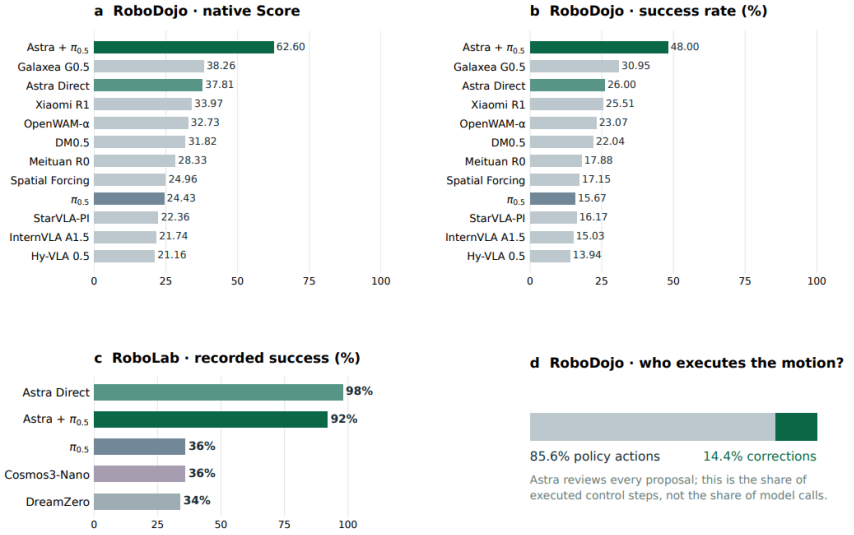
2）世界动作模型（ World Action Model，WAM ）：是**世界模型**的扩展，世界模型学习环境如何随时间演化，重点是理解**跨任务复用的物理规律**，因为物理规律的通用性高于语义，物体坠落、滑动的规律不会随物体种类改变，可复用到无相似示教样本的场景中，输入是动作指令，输出是**未来世界状态**；WAM是世界模型的演进，输入任务目标，输出是**动作执行和未来世界状态**的预测；

3）大语言模型：Chat-GPT6 Astra是原生大语言模型，未经机器人数据训练即可跑通完整任务流程，2026年9月，Galbot Team经过系统测评表明Chat-GPT6 Astra在任务执行过程中能够给出有价值的**动作决策**，超出行业预期。

图：VLA通过语义推理与机器人状态生成智能体动作，WAM联合预测动作和未来世界状态



图：Astra作为通用具身策略可取得较好结果



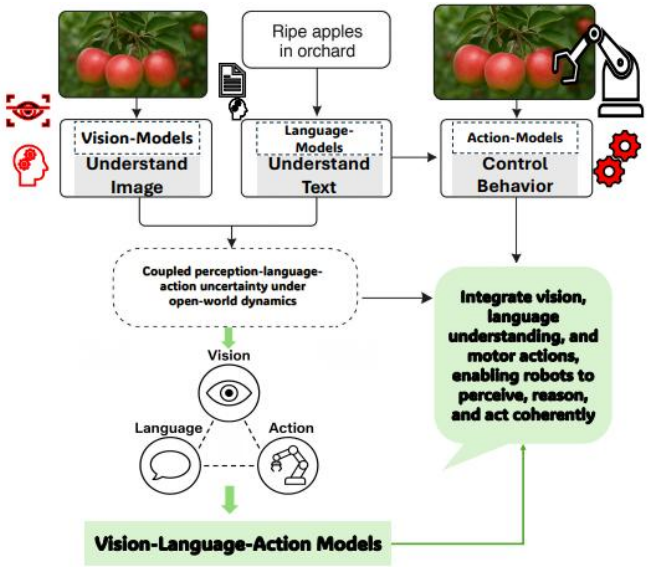
VLA：输入图像与语言指令，输出控制动作

独立模块：各功能模块独立。VLA模型出现之前，人工智能领域模型进展大多彼此割裂：视觉系统负责采集、解读与识别图像；语言系统完成文本理解与生成；动作系统则用于运动控制。这些独立模块各自在专属任务上表现尚可，但难以协同工作，**很难泛化至全新场景**。

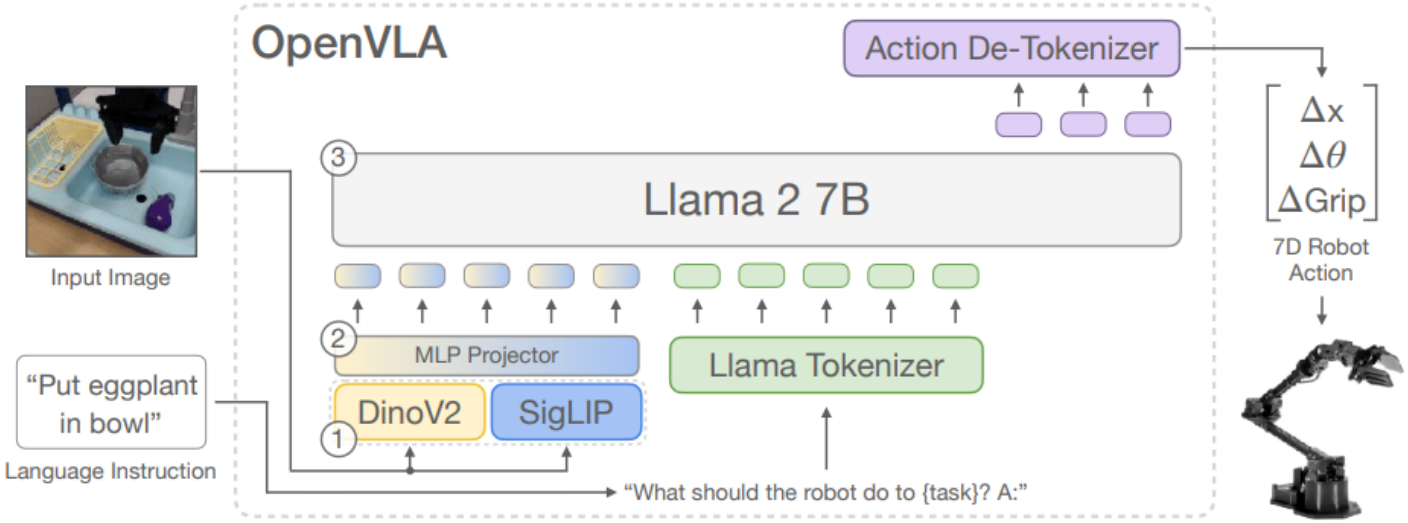
视觉语言模型（VLM）：融合视觉与语言，但无法输出动作。绝大多数人工智能系统仅专攻单一或两类模态，例如视觉–语言、视觉–动作、语言–动作，以VLM为例，该模型融合视觉和语言，但**模型无法基于多模态输入生成、执行连贯的动作**。

VLA：整合感知、自然语言理解与具身动作。VLA融合视觉输入、语言理解与运动控制能力，使具身智能体能够感知环境、解析指令，并动态输出适配动作，以谷歌DeepMind在2023年开发RT-2为标志性工作，其他还包括斯坦福大学等开发的开源OpenVLA、 初创公司Physical Intelligence开发的π_0/π_0.5/π_0.7等。这类模型学习的是响应式的“观测到动作”映射，并**不会显式建模物理世界在交互干预下如何演化**。

图：VLA模型整合感知、自然语言理解与具身动作



图：给定图像与语言指令， OpenVLA模型输出机器人控制动作



数据来源：康奈尔大学，OpenVLA，东吴证券研究所

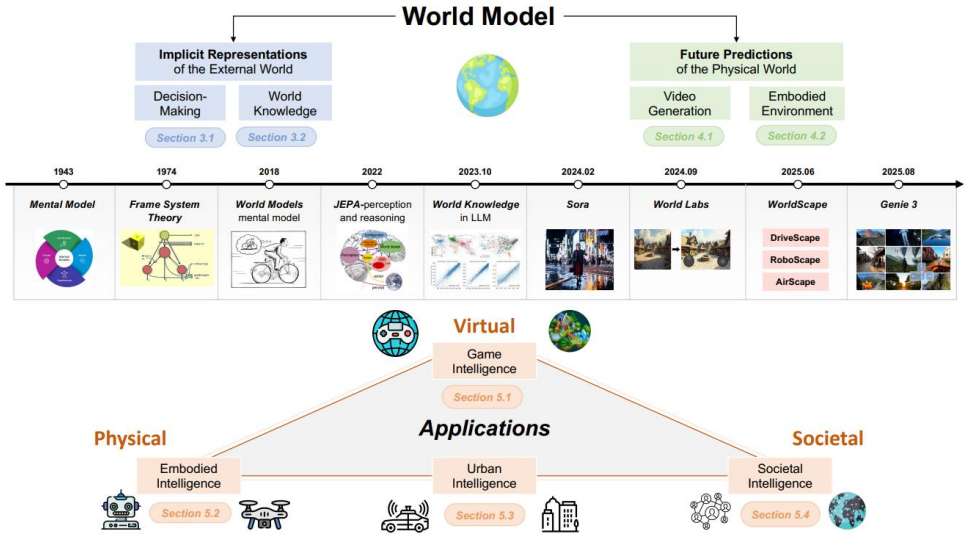
世界模型：理解外部世界的动力学规律，交互能力快速提升

世界模型：核心目标是理解外部世界的动力学规律。世界模型的目标是掌握世界的动力学规律，并能够确定性地**推演下一状态**，借此实现更长时程的演化外推，并支撑下游的决策与规划。主要可分为两类：**1）外部世界的隐式表征：**将外部现实映射为由隐变量构成的模型，以此形成对世界的隐式理解，大语言模型的出现大幅强化了该路线对世界知识的刻画能力；**2）物理世界的未来预测：**强调生成视频的真实感，要求视频能够复现物理世界的未来状态，前沿研究重心逐步转向构建具备真实交互能力的物理世界。

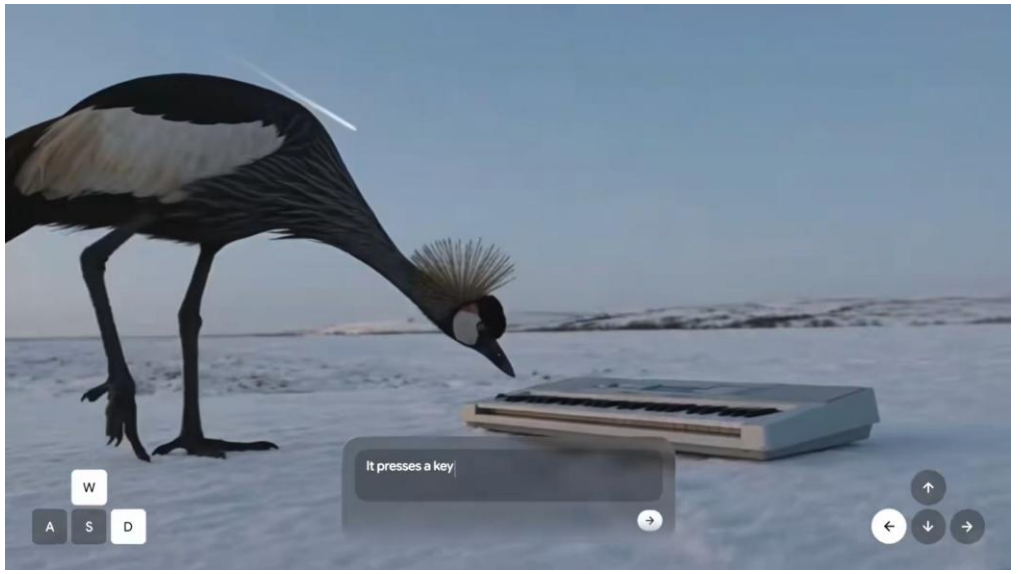
典型代表：谷歌DeepMind的DreamerV3、Genie系列、Meta的V-JEPA 2、蚂蚁灵波 LingBot-World等。

交互能力：谷歌DeepMind于2026年10月27日发布新一代世界模型**Genie 4**，相比Genie 3性能明显提升，面向**实时交互**式世界模拟，用户可通过键盘控制导航和相机，并可在模拟运行过程中输入自然语言改变天气、增加人物或物体，模型随后根据当前环境继续演变后续状态。

图：世界模型核心目标是理解外部世界的动力学规律



图：谷歌世界模型Genie 4支持实时探索动态场景



数据来源：清华大学，DeepMind，东吴证券研究所

世界模型：与大语言模型（LLM）的区别与协同

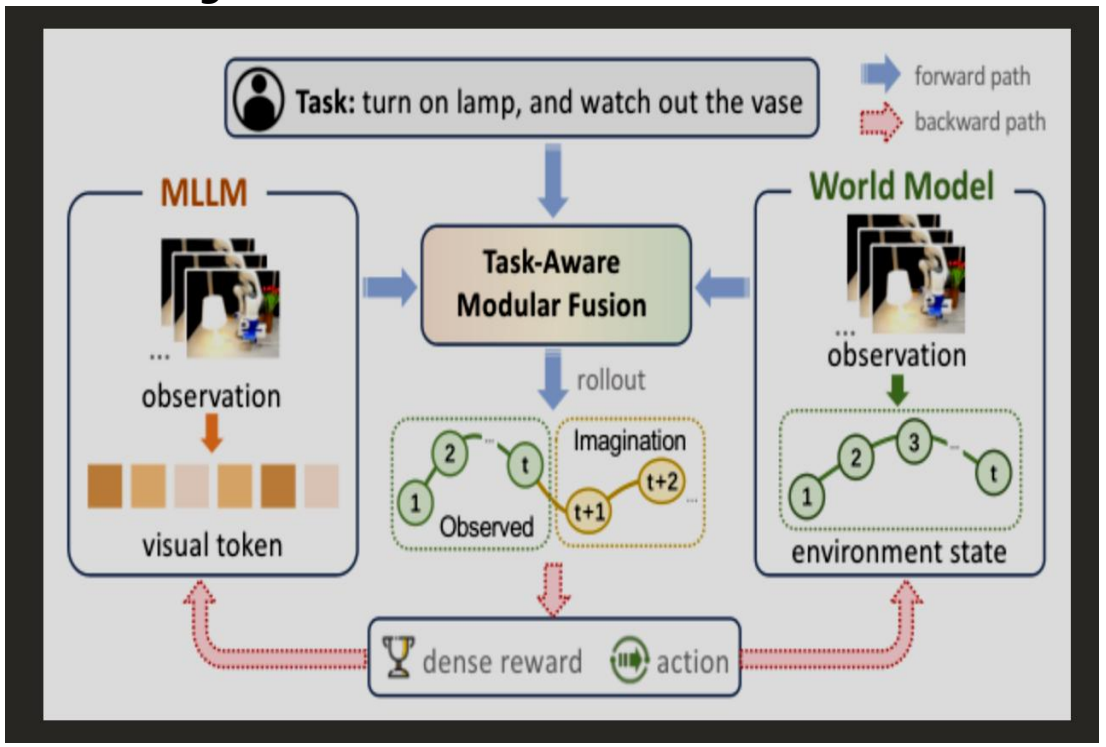
世界模型重点特征在物理因果规律。LLM依托数据学习语义关联，形成偏向描述性的符号认知，但缺少物理因果预测。世界模型则为状态预测器，依托数据学习物理因果规律，构建可推演的环境特征，擅长模拟动作带来的现实变化。

世界模型与LLM可以形成混合架构。大语言模型与世界模型二者可以协同搭建物理AI系统，LLM主要负责理解自然语言、拆解高层任务，世界模型负责模拟环境演化、预判动作带来的物理后果。大语言模型主要学习文本等数据中的规律，典型训练目标是预测**下一个Token**；世界模型则需要学习环境随时间变化的规律，预测**下一时刻的物理状态**。

表：世界模型与LLM的区别在于物理因果规律

对比维度	目标与学习范式	优势	局限	训练数据与表征形式
大语言模型（LLM）	作为token预测器，从文本、多模态数据学习语言统计与语义关联，目标产出通顺文本，建立符号语言层面、偏向描述性的世界认知	语义理解能力突出，适配自然语言交互，擅长开放式任务、知识调取、抽象推理和任务拆解	缺少物理落地支撑，容易产生幻觉，很难内生推演物理后果，长时序物理规划能力不足	训练素材以文本token序列为主；表征是隐式、统计型的语言嵌入
世界模型	作为状态预测器从交互、视觉传感数据学习物理时空因果规律，模拟动作带来的环境变化，搭建支持反事实推演的物理因果表征	掌握物理因果逻辑，可以在内部做场景预演规划输出结果符合现实物理规律，适配具身智能交互场景	语言理解、抽象推理能力偏弱，很难理解复杂语言指令、社会规则，不擅长语言表达	训练素材为视频、三维场景、传感器时序等连续数据；表征面向状态转移，侧重因果信息的高效压缩

图：BiTAgent，LLM与世界模型协同架构



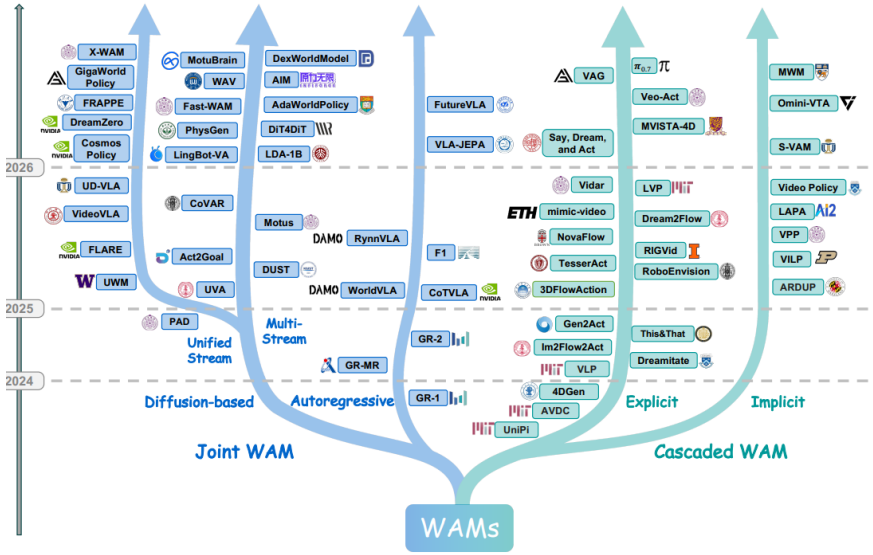
WAM：统一预测状态建模与动作生成

WAM：具身基础模型的一类，把**预测状态建模与动作生成统一**起来，目标是对未来状态与动作的联合分布建模，而非仅单独对动作建模，WAM演进过程中主要发展出联合式和级联式两类，两类架构路线是当前领域主流探索方向，它们属于不同结构方案，目前正处于**并行发展**阶段，并非严格的前后迭代替代关系。

1）联合式：将世界预测与动作生成深度耦合，世界预测、动作生成绑定在同一个模型内；典型的模型包括英伟达DreamZero、字节跳动GR-2、蚂蚁灵波LingBot-VA等。

2）级联式：将世界建模与动作执行大体解耦，基本分成两大模块，即先做世界模型预测未来状态，再单独交给策略模块输出动作，发展路线包括显式对齐、隐式对齐两种，典型模型包括谷歌Deep Mind等开发的UniPi、阿里WorldVLA等。

图：WAM主要可分为联合式和级联式两类



图：DreamZero可联合预测视频与动作



数据来源：OpenMOSS，英伟达，东吴证券研究所

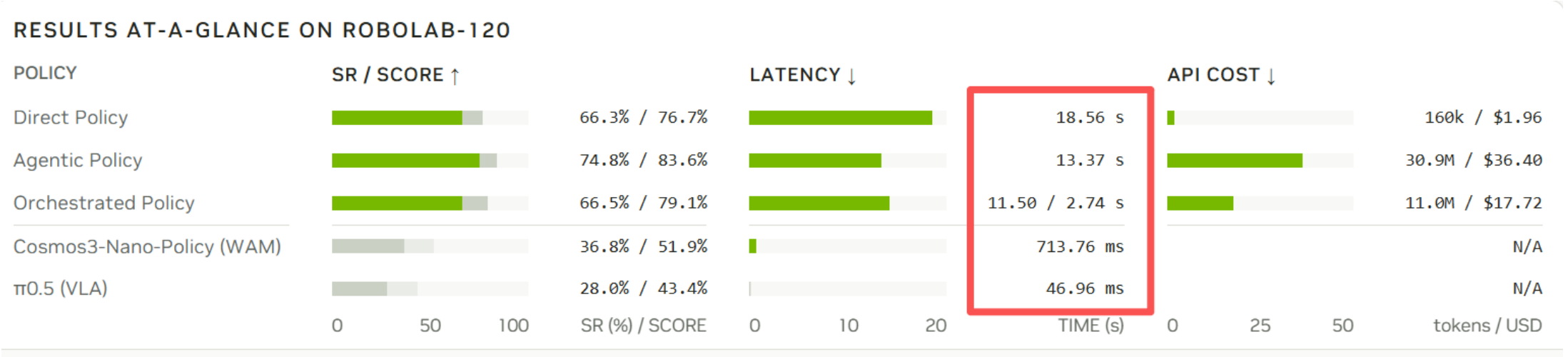
Astra：长程任务规划能力超预期，瓶颈在速度和效果判断

Astra用于物理AI仍存在速度和动作效果判断瓶颈。

1）控制**速度**不匹配：大模型控制机械的速度很慢，公开视频多为数倍甚至几十倍加速；大模型推理仅能做到几赫兹**频率**，英伟达Astranomy项目测得GPT-6 Astra在具身智能直接生成动作模式下的单次推理平均需18.56秒，而机器人底层伺服闭环需要200赫兹以上，二者完全不匹配。速度慢主要是模型本身的原因，比如Astra全参数1000B，运行速度较慢，但是并非所有参数都有效。

2）**动作效果判断**能力崩塌：大模型在任务泛化、理解、纠错、长程任务规划上具备优势，模型能够给出有价值的动作决策，却难以稳定生成、校验这些动作在物理世界里产生的实际效果。已训练策略与全身控制器有助于弥合这一差距，但想要实现高效协同，判断何时介入干预、何时把控制权交还给底层策略仍是Astra需要解决的问题。

图：直接用Astra生成动作模式下单次推理时长超10秒

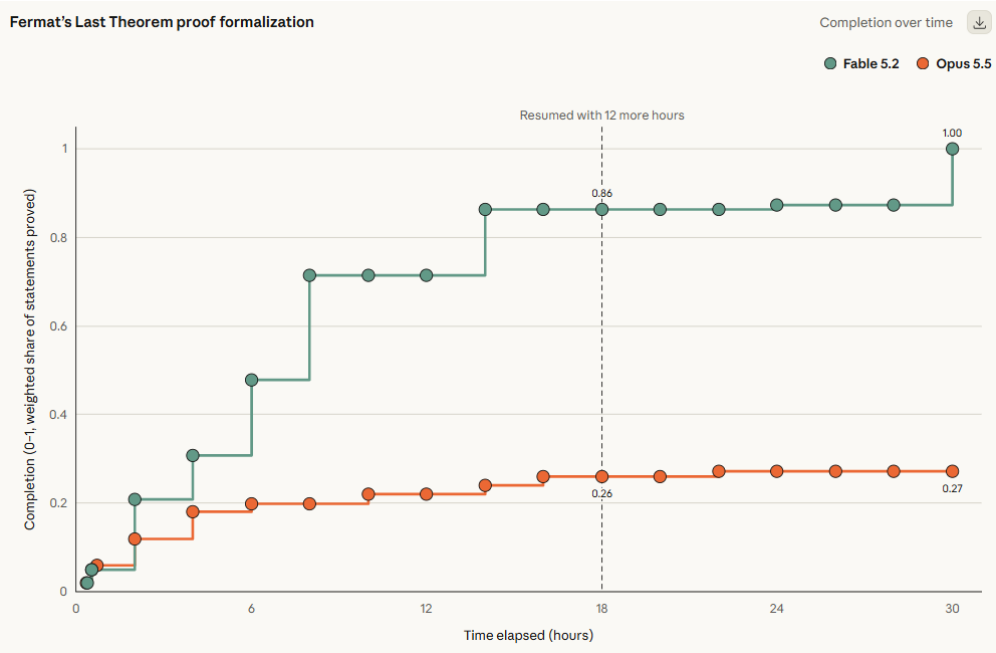


Fable 5.2：长任务处理能力和成本效率提升

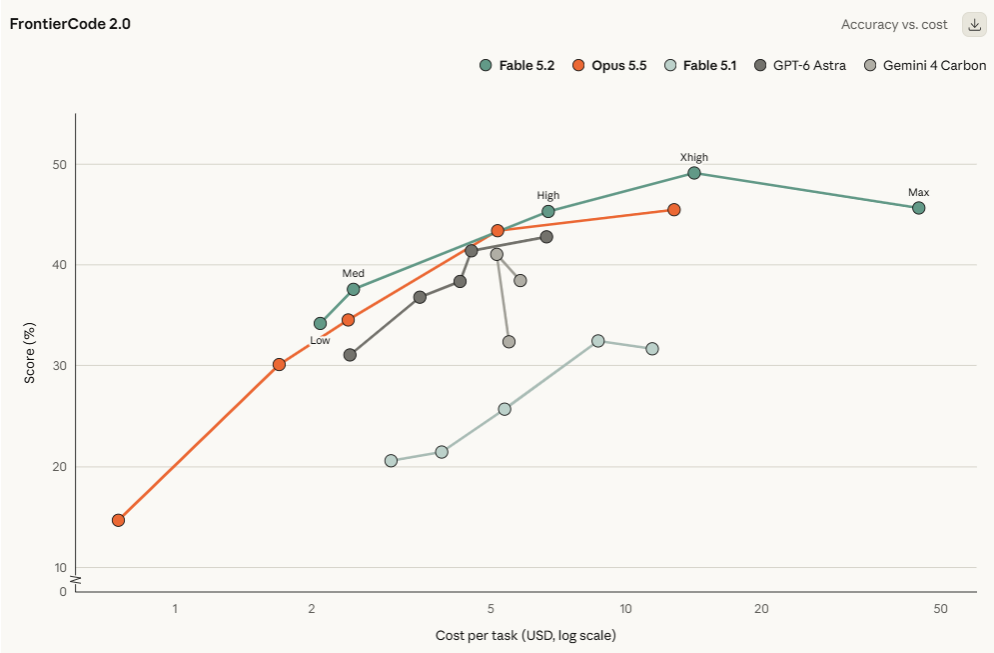
长任务处理能力提升。Fable 5.2的可用将近30个小时完成了费马大定理的形式化证明。而此前一个未发布的版本，配上专用的定制化工具调度框架，外加研究员引导，做完相同的工作，仍然需要11天，Fable 5.2的长任务处理能力明显提升。

成本效率提升明显。在Fable 5.2 Medium模式下，Fable 5.2完成高难度任务所需的token数量少于Fable 5.1。在FrontierCode 2.0的评测中：Fable 5.2与Opus 5.5的水平大致相当，但单任务成本仅约后者一半；Fable 5.2 Xhigh模式在所有已发布参评模型中取得最高分。

图：Fable 5.2长任务处理能力提升



图：相近成本水平下Fable 5.2优于GPT-6 Astra编程测试最高分



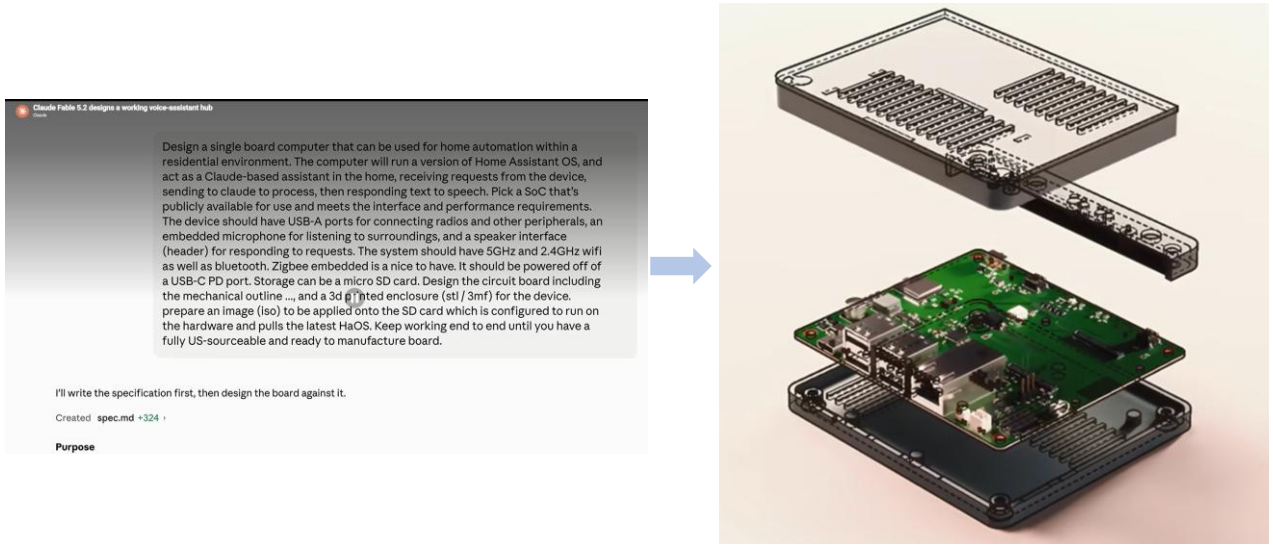
Fable 5.2：工程领域能力逐渐强化，机器人控制能力有优势

通用AI在硬件设计和机器人控制等工程领域能力逐渐延伸。Claude Fable 5.2不仅在编程、物理和数学等领域表现强劲，在硬件设计和机器人控制等工程领域的表现也明显提升：

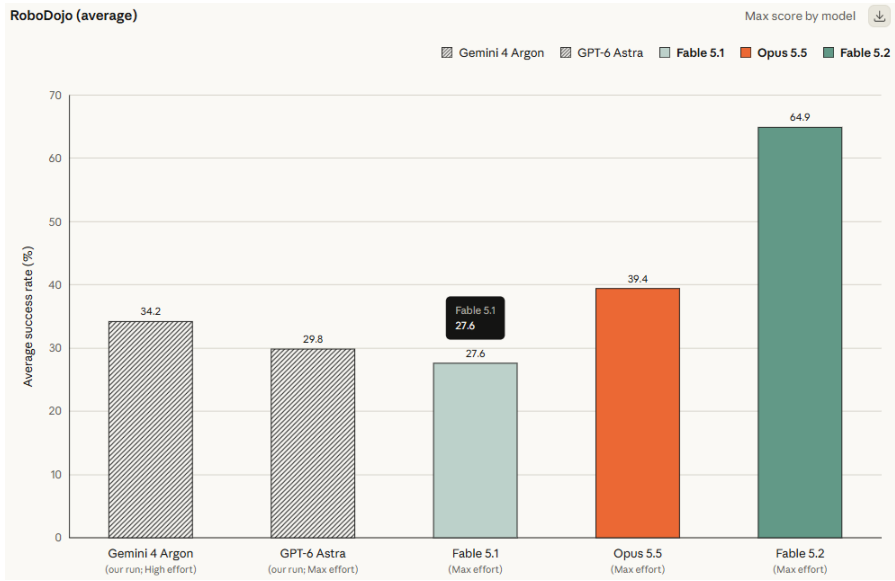
1）硬件设计：使用单段提示词，令Fable 5.2输出三套可投产电子设备方案， Fable 5.2可自主输出规格、选型元器件、完成四层PCB布线、CAD外壳建模与固件开发；

2）机器人控制：Claude Fable 5.2 Max模式在RoboDojo Sim测试中取得64.9%的平均成功率，达同样测试条件下GPT-6 Astra成功率（29.8%）的2倍以上，并领先上一代Fable 5.1（27.6%）与Opus 5.5（39.4%）。

图：Fable 5.2硬件设计能力逐渐强化



图：Fable 5.2在双臂仿真机器人的控制能力评估中表现较强

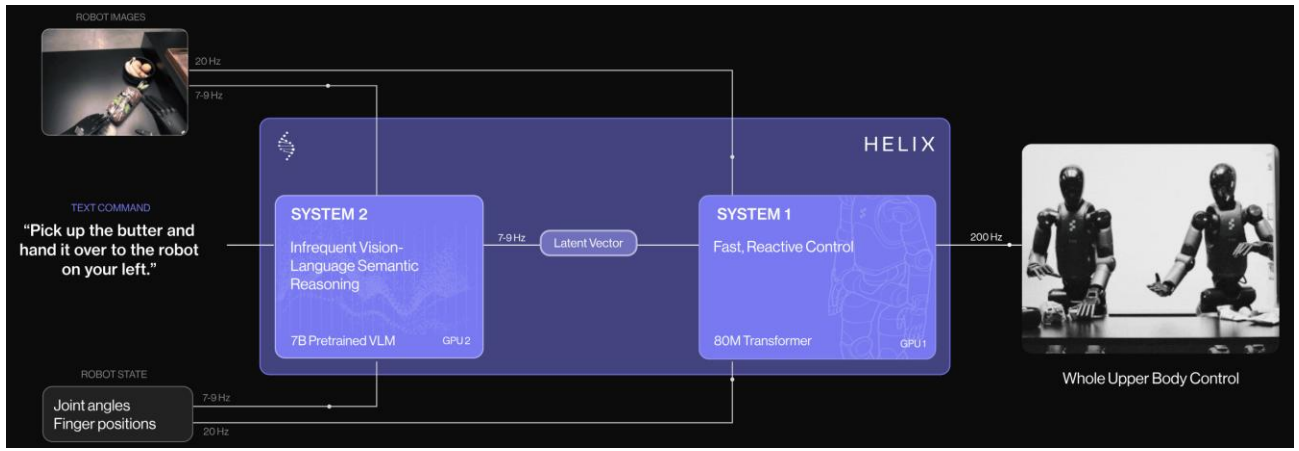


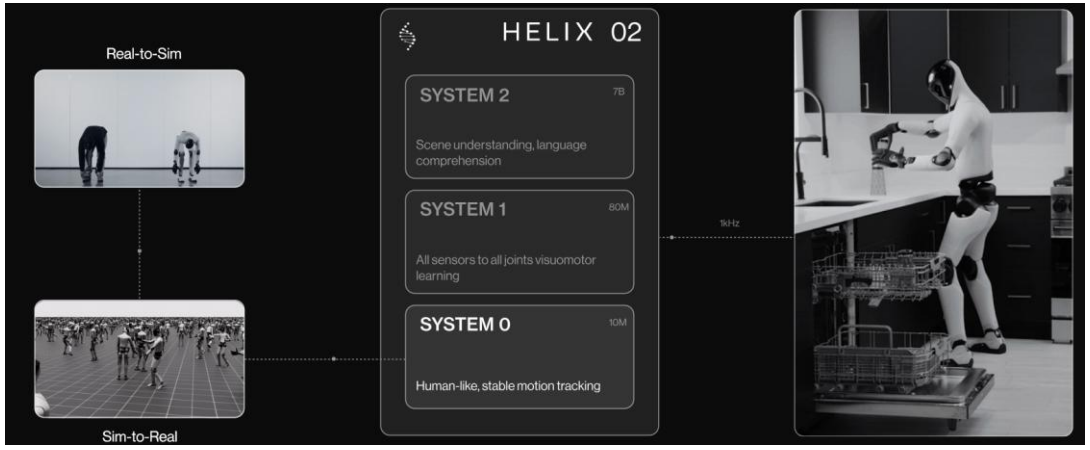
模型部署：分层结构是潜在主流方向

VLM的主干网络具有较强的通用性，但运行速度不够快；而机器人的视觉运动控制策略虽然响应迅速，却缺乏足够的通用性。Figure AI的Helix模型尝试采用System 1与System 2结合架构的VLA模型，专为实现高频率、灵巧且覆盖整个类人机器人上半身的运动控制而设计，Helix通过两个相互协作、经过端到端联合训练的系统解决了速度不匹配问题：解耦式架构使两个系统能够按照各自最适合的时间尺度独立运行。

System 2：理解高层任务目标，进行语义推理和场景分析。部署在机器人本体上的VLM模型，运行频率为7-9 Hz，负责场景理解和语言指令解析，使机器人能够在不同物体和环境背景下实现泛化能力；**System 1**：快速执行动作，并根据实时环境变化进行动态调整。高速响应的视觉运动控制策略，以200 Hz频率运行，将System 2生成的潜在语义表征转化为精确连续的机器人动作。后续Helix 02进一步扩展了该结构，引入以1 kHz频率运行的**System 0**，负责机器人的实际运动执行，处理全身范围内的平衡控制、接触交互以及身体各部位之间的协调。

图：不同层级的模型共同控制机器人的动作执行





3、终端载体：人形机器人是物理AI最合适的载体

载体：人形机器人是物理AI理想载体，灵巧手是关键部件

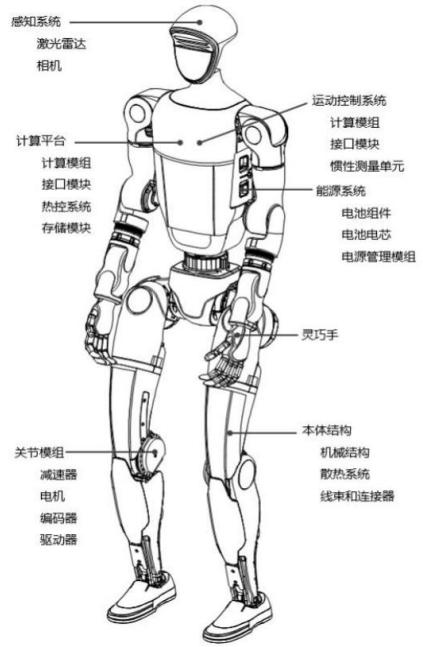
人形机器人是物理AI最合适的载体。1）物理AI的训练语料主要来自人类，比如Ego数据采集是人类的第一人称视角，用于训练的网络视频主要包含的是人类的行为动作；2）人形硬件天然适配模型和数据，训练效果最佳；3）人形轮式机器人、双足机器人（如宇树）的硬件成熟度均在持续提升。

人形机器人包括感知、计算、控制、能源和执行等模块，灵巧手是关键部件。典型人形机器人包括感知系统、运动控制系统、计算平台、本体结构、关节模组和灵巧手等模块，灵巧手技术路线多样，是人形机器人大规模量产的关键难点。

图：Apple开发HALO系统控制宇树G1机器人在新场景中活动



图：人形机器人由感知、控制和执行等多个模块组成



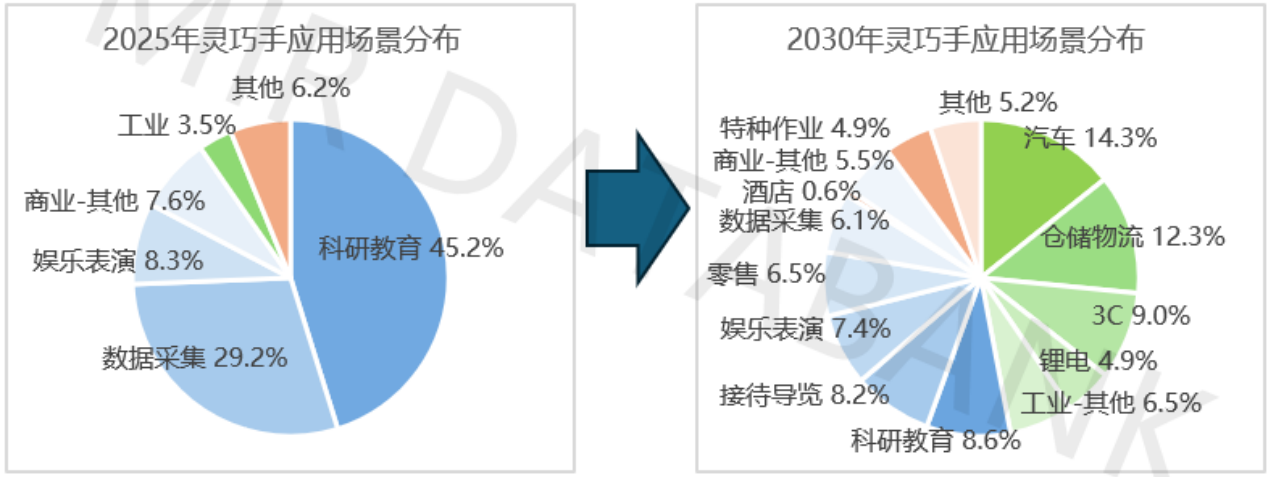
数据来源：Apple，宇树科技，东吴证券研究所

灵巧手：技术路线分化，逐渐从研发验证走向商用落地

灵巧手是人形机器人核心模块之一，现有四条主流路线。1）分类：灵巧手的传动结构有四条路线：连杆传动、腱绳传动、直驱、混合传动。混合传动是融合两种及以上传动结构，针对拇指、四指、近端与远端关节的差异化作业需求分区匹配不同方案，当前主流组合为“腱绳+直驱”。2）特点：灵巧手各技术路线各有特点，但单一路线较难兼顾高精度、高负载、轻量化与高柔顺等需求。**混合传动**依托分区传动设计，在精度、负载、轻量化与柔顺性之间实现综合平衡，有效解决纯直驱方案散热问题突出、纯腱绳方案精度不足的问题，是未来灵巧手硬件演进的核心方向。

灵巧手逐渐从研发验证走向商用落地。据MIR数据，按 2025 年出货量（台）计，中国灵巧手市场中科研教育和数据采集合计占比74%，产品仍集中在实验室和展台，真实场景应用中尚未规模化落地。随着汽车线束插接、3C精密装配等场景以及灵巧手所用的力学传感器小批量订单在2026年集中落地，增长主引擎正从科研采购加速转向行业级应用。

图：2025年灵巧手应用场景包括科研教育和数据采集等，后续逐渐丰富



表：灵巧手技术尚未完全收敛

技术路线	优势	劣势
连杆传动	· 刚性传动，负载能力强 · 结构简洁，制造成本较低 · 控制算法成熟，量产良率高	· 灵活性较差，关节活动范围受限 · 手部自重大，不利于轻量化设计 · 精细化操作能力有限
腱绳传动	· 驱动器可外置，手部自重轻 · 被动柔顺性好，可自适应抓取不规则物体 · 仿生度高，关节自由度高，运动接近人手	· 腱绳易磨损，寿命较短，维护成本高 · 控制算法复杂，传动精度受限 · 电机外置，散热设计难度较大
直驱	· 传动精准，响应迅速 · 刚性传动，力反馈特性好，抗冲击能力强 · 结构紧凑，可实现各关节独立精准控制	· 动力单元内置于手指，结构笨重，体积偏大 · 齿轮啮合易产生噪声 · 内部结构复杂，维护难度大
混合传动	· 兼顾高负载与高灵活性，综合性能均衡 · 可分区优化设计，关键部位保精度，其余部位做轻量化或自适应处理	· 系统复杂度高，设计、控制与制造难度大幅提升 · 多技术融合，硬件成本显著上升

MIR DATABANK

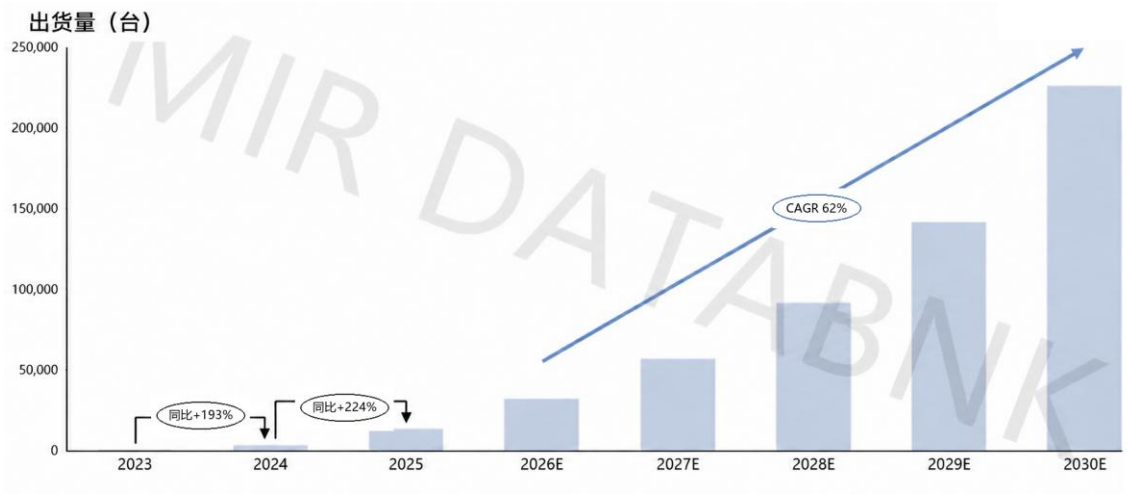
数据来源：MIR，东吴证券研究所

灵巧手：市场规模快速提升，部分企业初步具备规模化交付能力

市场规模将快速提升。据MIR预测，2026年全年灵巧手出货台数将突破3万台，小批量行业级落地订单开始成为市场增长的新支撑。

部分企业初步具备规模化交付能力。从精密硬件结构、传感感知、柔顺力控算法、场景化软件方案、工程交付服务等多维度综合评估，国内灵巧手行业已形成一批具备规模化交付能力的代表性企业。其中，因时机器人、强脑科技、灵心巧手、Sharpa、中科硅纪等企业在软硬件一体化能力、技术路线完整度与产品落地上表现突出。

图：中国灵巧手市场规模将快速提升



表：灵巧手厂商梳理

序号	英文简称	灵巧手厂商综合榜单	图标
1	Inspire-Robots	因时机器人	
2	BrainCo	强脑科技	
3	LinkerBot	灵心巧手	
4	Sharpa	Sharpa	
5	CasiaHand	中科硅纪	
6	DexRobot	灵巧智能	
7	WUJI	舞肌	
8	Xynova	曦诺未来	
9	AGILINK	临界点	
10	DH-Robotics	大寰机器人	

*排名不分先后

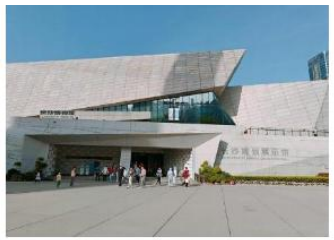
4、产业链公司梳理

风语筑：三维空间数据铸就壁垒，积极布局物理AI

空间数字化沉淀的可仿真三维资产，使公司成为物理AI稀缺的数据供给方。公司具有丰富的数字展示、文旅/城市空间数字化项目积累，拥有大量带物理属性、可仿真的真实三维空间数据资产，适配物理AI训练需求；公司具备空间采集、三维重建、数字孪生能力，和字节即梦AI合作落地非遗、文旅AIGC项目，是物理AI训练稀缺的数据供给方。

参股/合资/共建深化物理AI全产业链布局。2026年6月公司拟以自有资金不超过2,000万元投资松应科技（自研ORCA OS物理AI操作系统），随后又于7月与松应科技合资设立上海风应科技，聚焦物理仿真引擎、Sim2Real仿真训练及空间合成数据生成。联手宇树科技等企业打造杭州市具身智能展示与应用推广中心，推动物理AI技术从研发验证向深度交互场景落地转化。

图：公司具有丰富的数字化空间项目经验


长沙规划展示馆

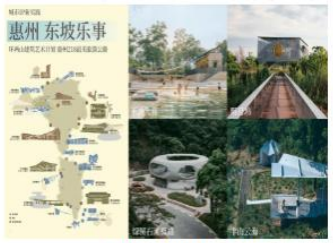


重庆规划展示馆（新馆）


晋城城市展览馆


中国共产党历史展览馆


杭州市具身智能
展示与应用推广中心



惠州东坡十六乐事
艺术计划

表：公司通过参股/合资/共建深化物理AI全产业链布局

布局	角色	意义
松应科技	以自有资金不超过2,000万元投资	布局物理AI操作系统
风应科技	和松应科技合资设立，控股70%	推动数采-仿真-训练-评测闭环
杭州市具身智能展示与应用推广中心	和宇树科技等企业联手共建	推动物理AI技术从研发验证向深度交互场景落地转化

奥比中光：3D视觉龙头，卡位物理AI感知入口

“全栈式+全领域”技术体系，卡位物理AI感知入口。公司可提供全套自主知识产权3D视觉感知产品，全面布局主流3D视觉技术，灵活适配不同应用领域。 2026年8月，公司于WRC大会上展示EGO、UMI、WristCam、Hub等产品，形成覆盖真实交互全关键视角的数据采集矩阵。

芯片、光学、算法全栈自研，构筑端侧物理AI硬件的适配能力。公司自主完成多代深度引擎与dToF/iToF感光芯片开发，覆盖结构光、双目、ToF三条路线，具备百万级量产能力，可为物理AI提供轻量化、高集成的三维感知方案。

生态接入与下游拓展同步推进，3D感知需求随物理AI放量持续扩容。 产品已融入NVIDIA Isaac/Jetson、苹果MacOS、微软、英特尔等平台，海外拓展提速；下游聚焦智能机器人、三维扫描、工业测量与消费级 3D 打印，构成长期增长主线。

图：奥比中光通过自主研发不断迭代产品



表：3D 视觉产品矩阵

技术路线	代表系列	特点与场景
结构光	Astra	近距离高精度，三维扫描/刷脸支付
双目	Gemini	低成本、室外适应性好，机器人导航
iToF/dToF	Femto	远距离、抗强光，物流/工业测量

数据来源：奥比中光，东吴证券研究所

福赛科技：聚焦机器人轻量化，深化海外产能布局

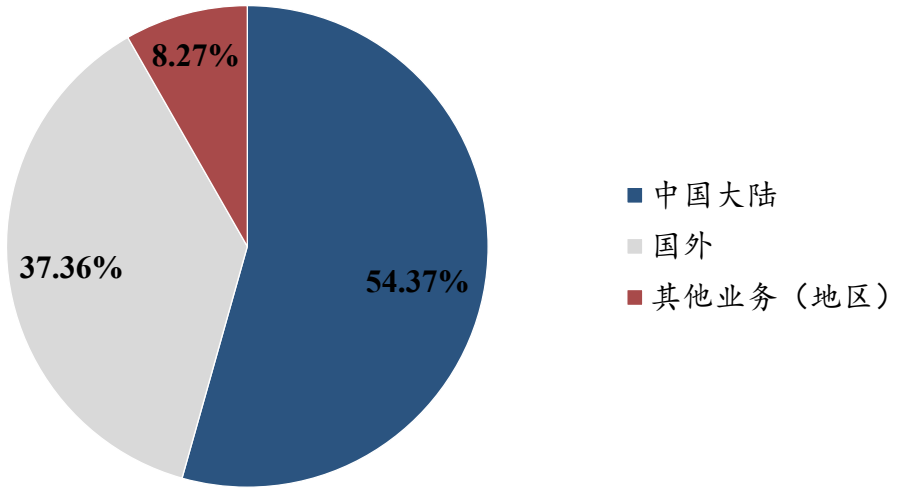
汽车内饰产品矩阵不断丰富，前瞻布局机器人轻量化。公司主要从事汽车内饰开发和销售，逐渐形成以功能件为主、装饰件和新能源三电嵌件为重要发展方向的业务布局。公司已与众多国内外汽车整车厂商或知名零部件供应商建立合作关系。公司前瞻布局机器人轻量化业务，已完成部分机器人零部件产品研发、小批量试制及送样实验工作。

海外收入占比较高，加速海外产能布局。2026H1墨西哥三座工厂约占公司总营收的1/3，公司拟投资3.8亿元建设泰国生产基地，进一步强化全球化布局，增强公司中长期竞争优势和增长前景。

图：公司汽车内饰产品矩阵不断丰富



图：2026H1公司国外收入占比较高



数据来源：福赛科技，Wind，东吴证券研究所

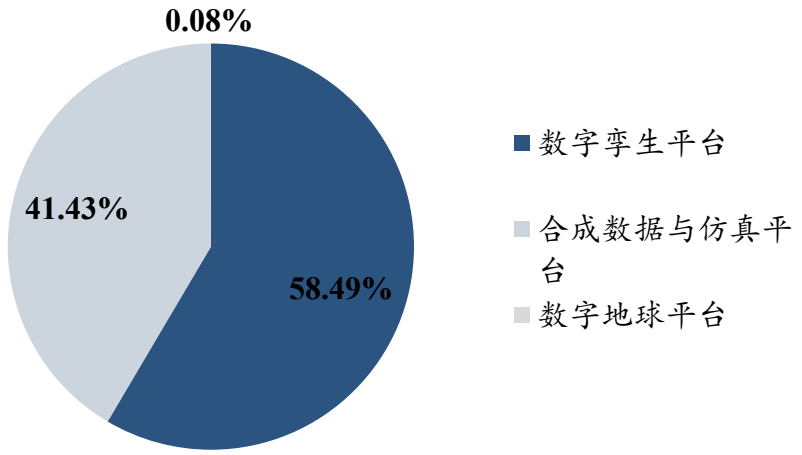
五一视界：51Sim放量，物理AI基础设施成型

51Aes、51Sim及51Earth三项业务铸就物理AI核心基础设施企业。公司依托51Aes数字孪生、51Sim合成数据与仿真、51Earth数字地球三大核心产品，业务涵盖自动驾驶、具身智能及空天探索等领域，应用场景逐步由数字孪生构建扩展至智能驾驶、具身智能数据与训练以及物理AI工厂。

51Sim是公司物理AI核心引擎，服务智能驾驶与具身智能企业，行业市占率第一。公司51Sim合成数据与仿真平台为具身智能领域提供高逼真仿真训练环境与海量合成数据生成能力，支撑AI系统在虚拟环境中的训练与验证。据沙利文报告，按相关业务收入计，2025年51Sim在中国端到端高阶智驾仿真及数据平台市占率53.5%，排名第一。

卡位物理AI基础设施，打开远期成长空间。公司依托AperData、51Claw、AperOne等产品切入物理AI的数据采集处理、Real2Sim2Real训练闭环及现场部署运营环节，有望打开公司远期成长空间。

图：2026H1五一视界收入来源包括数字孪生和合成数据与仿真



表：短期-中期-长期三阶段业务演进

阶段	核心产品	定位
短期	51Sim	按2025年相关业务收入计，高阶智驾仿真与数据平台，市占率 53.5%
中期	AperData+AperOne	具身数据采集与仿真验证闭环
长期	物理AI工厂	按训练片段计费的订阅/消耗式服务

数据来源：五一视界，Wind，东吴证券研究所

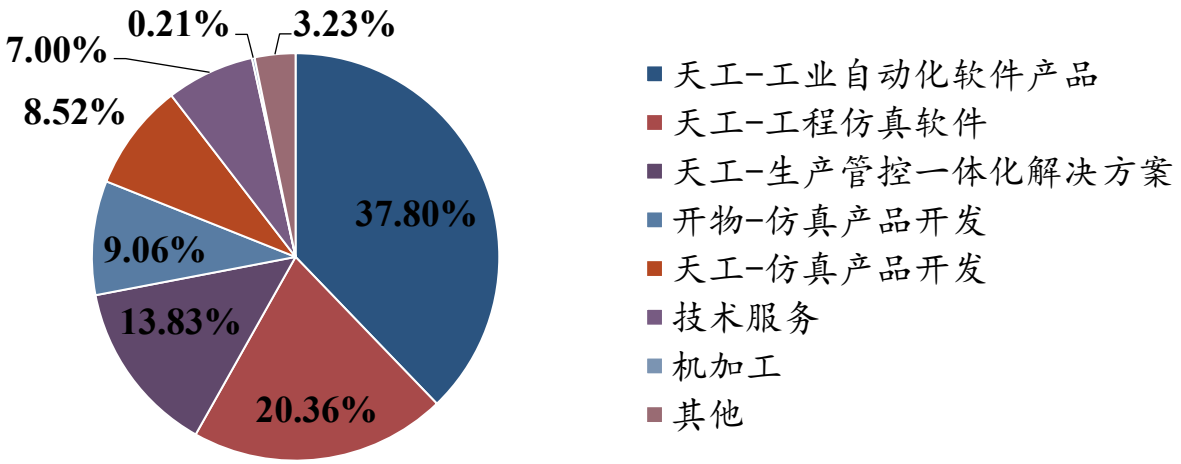
索辰科技：具备物理AI全栈能力，CAE求解器铸就差异化

以CAE求解底座优势切入物理AI，构建差异化技术壁垒。公司具备深厚CAE工程仿真、多物理场求解技术积累，打造物理机理+AI算法+工程能力的核心壁垒；依托原有仿真能力向物理AI延伸，和纯数据侧厂商形成路线差异，是其物理AI业务的核心特点。

“开物”物理AI产品线成型，核心产品矩阵持续迭代落地。公司“开物”系列产品作为物理AI第二增长曲线，2026H1收入658.85万元，新增发布“营造·万象”世界物理模型、大圣智能体、行者合成数据，完善产品体系，面向工业能源、具身智能、低空经济推进场景验证与客户导入。

资源倾斜+算力生态合作，落地“算法+算力”双轮生态。公司变更募投项目，将资金转向物理合成数据基座与场景示范项目，集中资源押注物理AI；同时与海光、中科星图等头部厂商签署战略合作，打通算力、航天测控等产业资源，共建物理AI生态。

图：2026H1索辰科技营收主要来自天工和开物系列产品



表：公司“开物”产品矩阵

产品	定位	面向场景
营造·万象	世界物理模型	工业能源、低空经济
大圣	智能体	工业/具身任务规划
行者	合成数据	训练数据生成
智能工坊	物理AI开发平台	场景验证与交付

数据来源：索辰科技，Wind，东吴证券研究所

宝通科技：布局工业物理AI，场景与渠道构筑核心壁垒

东吴证券 SOOCHOW SECURITIES

合资平台宝光是物理AI业务的核心承载主体。公司与光轮智能依托该平台完成能力分工协同，公司输出场景、工程、真机数据与客户资源，合作方提供高保真仿真引擎、合成数据以及机器人评测工具链，所有虚实闭环的研发、项目交付与商业化探索均在此平台内统筹推进。

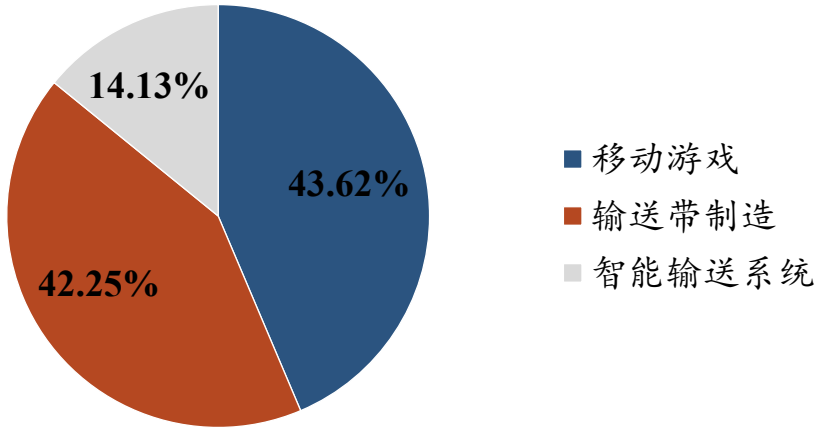
场景资源是业务落地的核心壁垒。依托输送带运维主业积淀，持续积累高质量工况真值数据，以清扫、巡检、矿业四足机器人作为数据采集终端，实现真机数据与虚拟样本双向回流迭代，匹配工业客户安全生产、减少停机损失的刚性需求。

渠道优势有望推动物理AI业务快速复制推广。国内依托成熟工程交付体系，推动仿真评测、机器人一体化方案落地标杆项目；海外借助长期矿业客户资源，已实现头部矿企的矿业四足机器人验证，为数据与仿真服务出海打下基础。

图：宝光将结合宝通和光轮优势，推进物理AI数据等产品落地



图：2026H1输送带制造与智能输送系统收入占总收入的56.38%



数据来源：宝通科技，Wind，东吴证券研究所

5、风险提示

风险提示

物理AI模型能力进展不及预期风险。当前通用具身模型在空间理解、任务规划、机器人控制等测试多集中于标准化环境与有限任务，若模型在复杂场景、长程任务、精细操作及跨本体泛化方面提升速度低于预期，将影响物理AI产业整体发展节奏。

机器人商业化及规模放量不及预期风险。物理AI产品仍处于研发验证和小批量导入阶段，若产品可靠性、成本、续航、安全性未能满足落地需求，或是下游客户资本开支意愿偏弱，将导致量产、订单兑现节奏低于预期。

公开演示代表性不足风险。物理AI评测体系仍处于发展初期，测试任务、硬件平台与评分标准尚未统一，标准化环境下的测试结果与企业公开演示，难以充分反映模型在真实复杂环境中的实际表现。

Sim2Real迁移效果不及预期风险。虚拟仿真环境与真实物理世界天然存在差异，合成数据难以完整还原接触、摩擦、材料形变、传感器噪声等长尾工况，若仿真精度、跨域迁移效果不及预期，将约束模型训练效率与真机泛化能力。

免责声明

东吴证券股份有限公司(以下简称“本公司”)经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供符合中国证监会《证券期货投资者适当性管理办法》规定的专业投资者使用，且接收人须为本公司认定并建立服务关系的专业机构投资者使用，本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。本报告不面向普通投资者发布或传播。任何非专业机构投资者收到本报告后，应立即退回并销毁，不得依据本报告内容做出任何投资决策。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中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投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，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。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后6至12个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证50指数），具体如下：

公司投资评级：

买入：预期未来6个月个股涨跌幅相对基准在15%以上；

增持：预期未来6个月个股涨跌幅相对基准介于5%与15%之间；

中性：预期未来 6个月个股涨跌幅相对基准介于-5%与5%之间；

减持：预期未来 6个月个股涨跌幅相对基准介于-15%与-5%之间；

卖出：预期未来 6个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

增持： 预期未来6个月内，行业指数相对强于基准5%以上；

中性： 预期未来6个月内，行业指数相对基准-5%与5%；

减持： 预期未来6个月内，行业指数相对弱于基准5%以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所

苏州工业园区星阳街5号

邮政编码：215021

传真：（0512）62938527

公司网址： http://www.dwzq.com.cn

东吴证券 SOOCHOW SECURITIES

东吴证券 财富家园