仅供内部参考，请勿外传

国金证券 SINOLINK SECURITIES

2026 年 10 月 24 日

计算机行业研究

买入（维持评级）

行业专题研究报告

证券研究报告

计算机组

分析师：刘高畅（执业 S1130525120005） liugaochang@gjzq.com.cn

分析师：郑元昊（执业 S1130525120004） zhengyuanhao@gjzq.com.cn

分析师：孙恺祈（执业 S1130526060005） sunkaiqi@gjzq.com.cn

物理 AI：此时恰好

行业观点

- **大模型能力正从数字世界向物理世界扩散，资本投入、模型升级与工具链完善共同推动本轮物理 AI 技术变革。**AMD 拟以约 82 亿美元收购 World Labs，验证世界模型与空间智能的战略价值。2026 年全球世界模型融资热度显著上升，AMI Labs、World Labs、Odyssey 等海外企业密集完成大额融资，国内 VAST 和极佳视界亦连续获得资金支持。模型御三家中，OpenAI 明确人形机器人研发方向，Google 已形成从视觉、动作到世界模型和端侧部署的长期技术积累，Anthropic 则通过能力研究与外部潜在接触补充机器人储备。NVIDIA 持续完善 Physical AI 工具链，推动物理 AI 由本体和零部件向模型、数据、仿真与硬件协同发展的产业体系扩展。
- **通用模型开始参与真实机器人操作，具身能力从技术演示逐步进入可验证阶段。**在未针对特定机器人本体进行大规模微调和后训练的情况下，通用模型已经能够根据语言指令和摄像头画面完成任务理解、空间定位和动作规划。Robocurve 测试中，GPT-6 Astra 在积木入碗任务中完成 19/20 次，但在圆形拼图任务中仅完成 2/20 次，说明模型已具备基础操作能力，精细接触和复杂任务仍是主要瓶颈。AA-Robotics 进一步通过统一机械臂、环境和任务比较不同模型，使通用模型的机器人能力开始具备可量化、可横向比较的验证方式。
- **行业 Benchmark 与分层架构逐步完善，物理 AI 的产业能见度和技术分工进一步清晰。**此前 BEHAVIOR、RoboChallenge 等评测受任务复杂、成本较高和企业参评意愿有限等因素影响，尚未形成广泛认可的机器人榜单；AA-Robotics 通过统一条件测试八类桌面任务，推动行业由企业演示走向标准化评测。Figure Helix 02 的 S0、S1、S2 架构则分别对应底层控制、动作策略和高层理解与规划。模型巨头预计先强化 S2 并逐步向 S1 渗透，抬升机器人智能的能力下限；与此同时，缺少数据和场景壁垒的独立模型企业可能面临竞争，真实交互数据、S0 低延迟控制、硬件适配和系统交付的重要性将进一步提高。
- **从交易结构看，物理 AI 具备逻辑顺、筹码轻、催化密和板块容量大的特点。**在科技板块轮动背景下，物理 AI 有望承接科技资金轮动。建议关注三条主线：一是机器人本体与 T 链供应链，包括特斯拉 Optimus、智元、宇树科技、上纬新材以及三花智控等；二是数据采集与空间感知，包括海天瑞声的数据集、奥比中光的 RGB-D 相机和禾赛 Kosmo 等三维数据产品；三是仿真、渲染与模型工具，包括摩尔线程、协创数据的具身智能仿真平台，索辰科技的 CAE 及物理 AI 平台、凡拓数创的 3D 渲染和 Sim2Real 方案等。

相关标的

- 1）本体：特斯拉、智元、上纬新材、宇树科技。
- 2）北美链：科森科技、斯菱智驱、岱美股份、世运电路、福赛科技、蓝思科技、浙江荣泰、三花智控、恒立液压、北特科技、拓普集团、科达利、万向钱潮等。
- 3）机器人大脑相关：摩尔线程、协创数据、索辰科技、仙工智能、宝通科技、五一视界、海天瑞声、道通科技、群核科技、智微智能、爱芯元智等。
- 4）其他核心机器人链：领益智造、奥比中光、绿的谐波、禾赛、速腾聚创、海康威视等。

风险提示

- 通用模型空间理解、任务规划及机器人控制能力进展不及预期的风险；机器人商业化、订单及量产节奏不及预期的风险；现有 Benchmark 无法充分代表真实场景能力的风险；仿真合成数据质量不及预期的风险；AI 芯片、算力和端侧部署进度不及预期的风险；行业竞争加剧、技术路线变化及相关公司估值波动的风险。

内容目录

一、技术变革：大模型能力向物理世界扩散，多重事件催化行情 4
1.1 高溢价交易验证世界模型的战略价值 4
1.2 模型御三家加速布局物理 AI 6
1.2.1 OpenAI 明确机器人研发方向，通用模型与本体结合预期强化 6
1.2.2 Google 多年投入形成跨层技术体系，模型与仿真协同构筑优势 8
1.2.3 Anthropic 以能力研究与外部接触补充机器人储备 11
1.3 英伟达持续强化 Physical AI 框架，NVIDIA 全栈布局降低开发门槛 11
二、产业进展：通用模型开始参与操作，公开评测提高行业能见度 13
2.1 通用模型涌现具身能力 13
2.2 机器人 Benchmark 的诞生 13
2.3 S0、S1、S2 构成具身智能的分层架构 15
2.4 巨头入局有望加速具身智能进入产业化阶段 16
三、交易分析：关注本体、数采及仿真龙头 17
3.1 板块位置与交易结构 17
3.2 三条投资主线 18
3.2.1 机器人本体与 T 链供应链 18
3.2.2 数据采集与空间感知 19
3.2.3 仿真、渲染与模型 20
四、相关标的 21
五、风险提示 21

图表目录

图表 1： Marble 可将图像、视频等输入转化为可探索的三维场景 4
图表 2： Marble 可将图像、视频等输入转化为可探索的三维场景 5
图表 3： HunyuanWorld1.0 架构 6
图表 4： HunyuanWorld1.0 可生成 360°沉浸式 3D 世界 6
图表 5： GPT-6 Astra 在 ARC-AGI-3 交互式环境评测中表现领先 7
图表 6： OpenAI 2019 年论文展示利用仿真训练的策略控制机械手完成魔方操作 7
图表 7： OpenAI 结合视觉状态估计与控制策略将机械手迁移至真机 8
图表 8： Google RT-1 架构 8
图表 9： Google RT-2 架构与训练 9
图表 10： Google RT-1-X 与 RT-2-X 架构图 9
图表 11： Google RT-1 与 RT-1-X 在小数据集场景下的成功率对比 9

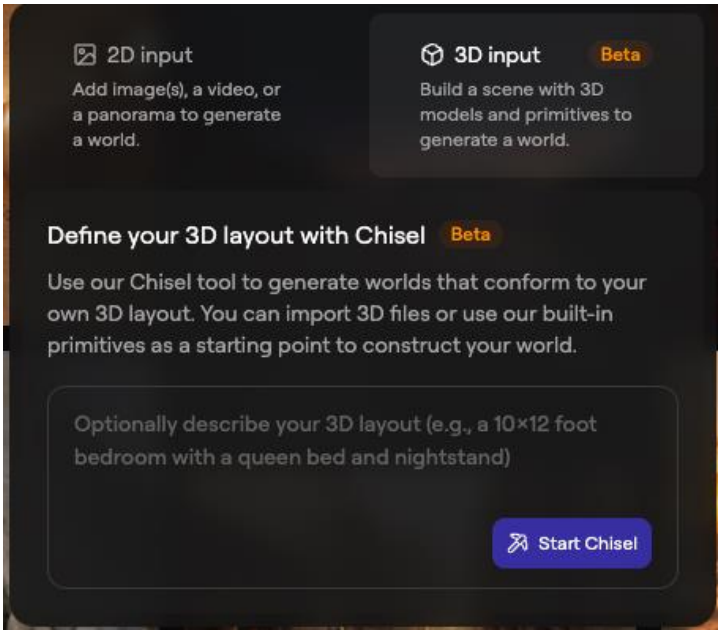
图表 12： Google-ER 能够识别物体、指向物体的各个部分 10
图表 13： Robotics 2 搭载 Inspire 灵巧手的全身操作任务成功率 10
图表 14： Robotics 2 搭载 SharpaWave 五指灵巧手的精细操作任务成功率 10
图表 15： Anthropic Embody 测评各大模型机器人控制能力 11
图表 16： NVIDIA Isaac GR00T N1 12
图表 17： 黄仁勋在柏林展示人形机器人参考平台 12
图表 18： Robocurve 测试发现 Fable 5.1 在输出较少 token 时完成了更多的工作 13
图表 19： Robocurve 测试结果 13
图表 20： Artificial Analysis 八类桌面任务测试 14
图表 21： Artificial Analysis 测试结果 14
图表 22： S0 基于人类数据的全身体制式控制 15
图表 23： Helix S1 视觉运动策略 15
图表 24： Helix S2 自主式长视距操控系统 16
图表 25： Helix S0-S2 架构 16
图表 26： 国证机器人产业指数 2026 年三季度以来走势 17
图表 27： 中证指数 CS 计算机指数 2026 年三季度以来走势 17
图表 28： 国证数字传媒指数 2026 年三季度以来走势 18
图表 29： Optimus 位于费蒙市第一条产线 18
图表 30： 上纬新材旗下启元机器人 19
图表 31： 奥比中光部分相机产品矩阵 19
图表 32： 禾赛 Kosmo 生成的数据能针对目标机器人的相机特性做视觉适配 20
图表 33： 索辰科技物理 AI 数据与训练平台 20
图表 34： 凡拓具身智能训练场方案核心架构 21

一、技术变革：大模型能力向物理世界扩散，多重事件催化行情

1.1 高溢价交易验证世界模型的战略价值

World Labs 的资本进展，为空间智能的战略价值提供了重要验证。公司由李飞飞与 Justin Johnson、Ben Mildenhall、Christoph Lassner 共同创立，核心团队覆盖机器学习、计算机视觉、三维生成与图形学。公司以空间智能为核心方向，聚焦物体、场景及交互关系建模；首款产品 Marble 支持从图像、视频、文本和三维布局创建具有空间一致性的场景。与主要处理文本和静态图像的任务相比，空间智能进一步要求模型理解位置、结构及环境变化，为 AI 从数字任务走向物理交互提供基础。

图表1：Marble 可将图像、视频等输入转化为可探索的三维场景



来源：Marble，国金证券研究所

世界模型正从早期研究走向产品应用。2018 年的《World Models》研究通过学习环境表征和变化规律辅助决策。World Labs 于 2024 年公开亮相时，媒体披露其已获约 2.3 亿美元融资；同年 12 月，公司展示可探索的持久三维场景，随后推进更大场景和实时交互生成。2026 年 2 月，公司宣布完成 10 亿美元融资，继续投入模型和产品研发。

AMD 拟以约 82 亿美元收购 World Labs，表明计算平台竞争正向世界模型能力延伸。2026 年 9 月 28 日，AMD 宣布签署收购 World Labs 的最终协议，交易以全股票方式进行，估值约 82 亿美元，预计于 2026 年底前完成，仍需满足监管审批等交割条件。我们认为，较大的交易规模反映出巨头对研究团队、空间生成能力及潜在工作负载的战略投入：世界模型既可能创造新的训练和推理需求，也可能影响机器人学习与仿真工具的开发方式。对计算平台而言，提前布局模型层有助于理解后续软硬件协同需求。交易估值体现战略价值，但在缺少同口径历史股权估值的情况下，不据此量化收购溢价率。

国内世界模型企业融资提速，一级市场对技术与数据能力的定价同步上移。极佳视界围绕世界模型、具身模型与机器人本体构建产品体系。据《直通 IPO》报道，公司于 2026 年 3 月至 6 月完成三轮融资，合计约 35 亿元。证券时报 4 月报道其估值突破百亿元，9 月的媒体报道则给出约 200 亿元的投前估值，反映出一级市场对世界模型能力的关注度提升。我们认为，融资节奏加快有望增强头部团队在算力、人才和数据侧的投入能力，推动模型与本体协同迭代；不同轮次的估值口径与交易条件存在差异，相关数据主要反映资本关注趋势。

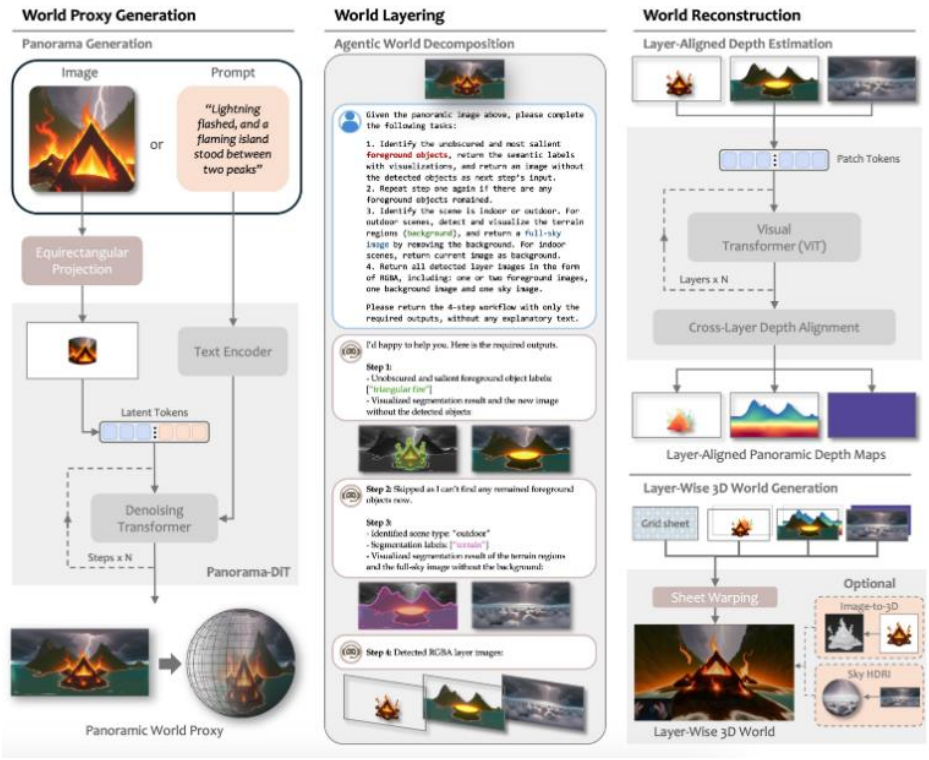
图表2：Marble 可将图像、视频等输入转化为可探索的三维场景

公司	技术方向	2026 年融资情况
AMI Labs	Yann LeCun 创立，基于视觉信息学习物理世界表征，侧重预测、推理与规划	3 月完成 10.3 亿美元种子轮
World Labs	空间智能及三维世界模型，可从图像、视频和文本生成持久化三维世界	2 月完成 10 亿美元融资
General Intuition	利用带动作标签的游戏视频训练世界模型和动作模型	6 月融资 3.2 亿美元，9 月再融资 2.2 亿美元，年内合计 5.4 亿美元
Odyssey	通用世界模拟，通过实时交互视频学习物理规律、因果关系和环境变化	6 月完成 3.1 亿美元 B 轮融资
Decart	实时交互世界模型及底层推理基础设施，覆盖游戏、仿真和机器人	完成 3 亿美元 B 轮融资
VAST（三启万物）	3D 原生基础模型及世界模型，面向游戏、工业设计和空间内容生成	7 月完成超 10 亿元 A3 轮；随后 B/B+轮合计约
极佳视界	世界生成模型与世界行动模型，向机器人及物理 AGI 延伸	3—6 月连续完成 10 亿元、15 亿元和 10 亿元融资，合计 35 亿元

来源：新加坡经济发展局、AMD、Coalition Capital、Odyssey、Radical Ventures、新华财经、中关村科学城、财新网、国金证券研究所

国内布局呈现机器人闭环、视频生成与三维世界构建多路线并行，空间智能的产业边界持续扩展。以爱诗科技旗下 PixVerse 为例，公司 2025 年 9 月宣布完成 6,000 万美元 B 轮融资，2026 年 3 月又宣布完成 C 轮融资并进入独角兽行列。腾讯的 HunyuanWorld 1.0 研究面向可探索、可交互的三维世界生成；商汤在 2026 年中期业绩材料中披露了与 Ace Robotics 联合开发的 Kairos 3.1 世界模型。与强调机器人数据和执行闭环的企业相比，视频生成企业首先面向内容创作及交互体验，收入来源与估值基础有所区别。后续产业分化的关键，在于视觉生成能力能否进一步转化为可控、可验证的物理预测，以及能否嵌入客户的训练与部署流程。

图表3：HunyuanWorld1.0 架构



来源：Github，国金证券研究所

图表4：HunyuanWorld1.0 可生成 360°沉浸式 3D 世界



来源：Github，国金证券研究所

1.2 模型御三家加速布局物理 AI

头部模型厂商向物理 AI 延伸，有望成为机器人研发范式变化的重要推动力。OpenAI 以内部研究、外部投资及硬件团队建设探索产品路径；Google 连接视觉、动作、世界模型与端侧部署，技术体系相对完整；Anthropic 的公开布局更突出语言、编程和智能体，机器人方向以能力研究及外部接触为切入点。三家公司的路径虽有差异，但均指向通用智能与物理交互的结合。我们认为，模型平台持续迭代与专用本体工程能力相结合，有望降低重复研发投入，并推动行业形成更清晰的软硬件分工。

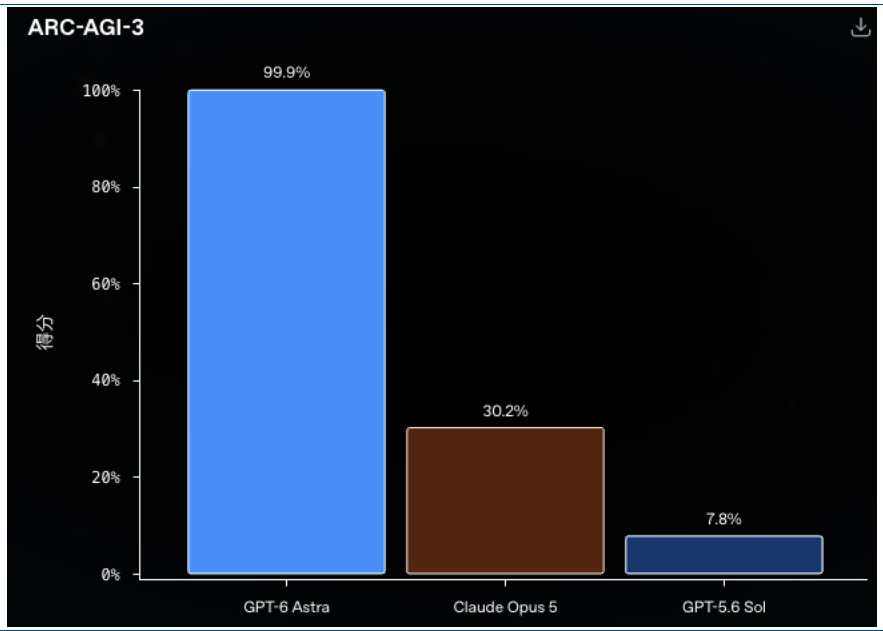
1.2.1 OpenAI 明确机器人研发方向，通用模型与本体结合预期强化

OpenAI 机器人研发方向进一步公开化，强化通用模型向实体终端延伸的预期。10 月 17 日的媒体报道援引 Sam Altman 在 Sources 播客中的表态，称 OpenAI 将研发人形机器人，重点在智能系统。其产业意义在于，模型能力的输出形式有望从数字环境中的文本、代码与工具调用，扩展至真实环境中的任务理解、动作选择和执行反馈。Altman 此前还谈及 2027 年可能出现演示，提供了后续产品观察窗口。我们认为，管理层表态首先催化

技术路线与组织投入预期，规模化应用仍取决于可靠性、安全性与成本的进一步验证。

2026 年 9 月 3 日，OpenAI 发布 GPT-6 Astra。该模型在 ARC-AGI-3 交互式环境评测中录得 99.9%的成绩，并在 96%的关卡中超过人类行动效率基准。尽管该评测不等同于机器人实机测试，但反映出通用模型在陌生环境中的理解、规划和交互决策能力进一步提升，为向具身智能迁移提供了能力基础。

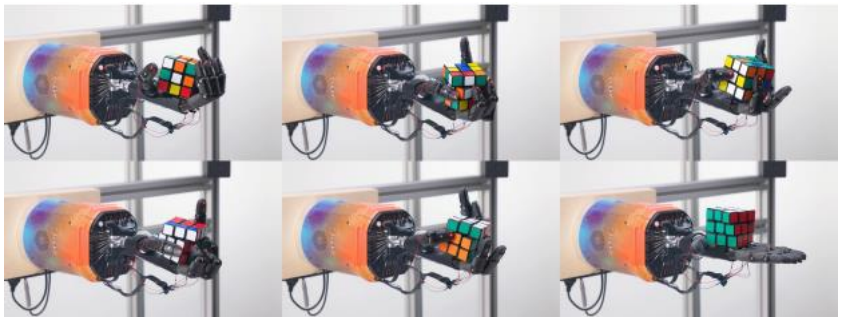
图表5：GPT-6 Astra 在 ARC-AGI-3 交互式环境评测中表现领先



来源：OpenAI，国金证券研究所

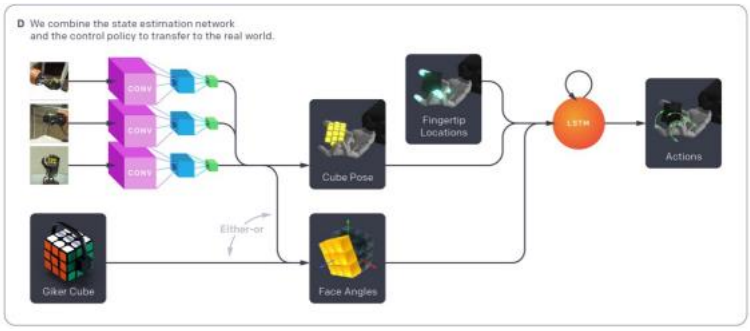
OpenAI 已在算法研究、产业投资与硬件组织层面形成多轮探索，为本轮研发提供历史积累。其 2019 年论文已经展示利用仿真训练的策略控制机械手完成魔方操作。2023 年，OpenAI Startup Fund 领投 1X 的 A2 轮融资；2024 年 2 月，Figure 宣布包括 OpenAI Startup Fund 在内的投资方参与其 B 轮融资，并与 OpenAI 签署人形机器人模型合作协议。人员方面，Caitlin Kalinowski 于 2024 年 11 月加入 OpenAI 推进机器人与消费硬件工作，并于 2026 年 3 月离职。上述节点反映公司长期关注机器人方向，当前合作与团队状态仍需结合最新公开披露分别判断。

图表6：OpenAI 2019 年论文展示利用仿真训练的策略控制机械手完成魔方操作



来源：Arxiv，国金证券研究所

图表7：OpenAI 结合视觉状态估计与控制策略将机械手迁移至真机



来源：Arxiv，国金证券研究所

通用能力复用有望改变机器人研发的成本结构，是本轮布局区别于早期专用策略研究的关键。多模态理解、代码生成与工具使用能力提升，使任务解释、环境识别和规划流程能够在更多场景中复用，减少逐项开发的投入。对本体企业而言，上层模型能力若能通过接口持续升级，既有硬件便有机会承接更多任务，研发重心也可进一步转向动作可靠性与场景交付。OpenAI 的产业影响因此不仅在于是否推出一款整机，更在于其模型迭代能否成为多个机器人平台共同受益的能力来源。

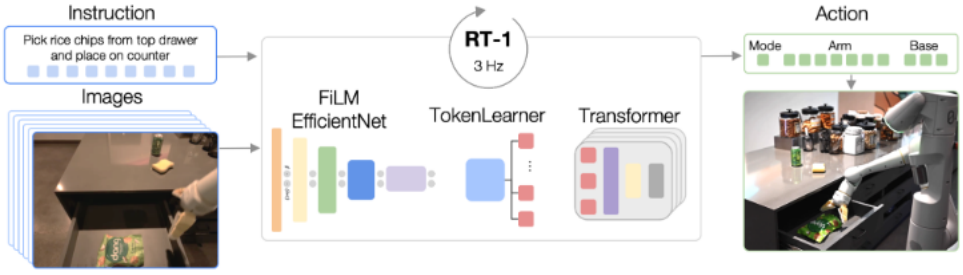
OpenAI 的产品组织方式将影响产业链受益环节，接口开放与部署验证是后续重点。若采取整机协同开发路径，核心零部件、制造和系统集成的重要性将上升；若通过模型接口、合作或授权向本体厂商输出能力，则更有利于扩大既有硬件平台的应用范围。当前阶段建议优先跟踪产品规格、对外接口、任务可靠性与批量部署信息，以产品和采购进展判断供应链弹性。管理层研发表态构成预期催化，向上市公司收入传导仍需经历客户认证与实际交付。

1.2.2 Google 多年投入形成跨层技术体系，模型与仿真协同构筑优势

Google 在机器人领域布局较早且投入具有连续性。RT-1 到 Gemini Robotics 2，其升级主线较为清晰：先用统一模型学习多种动作，再引入互联网知识和空间推理，随后加强长任务规划、跨本体迁移及全身控制。机器人模型的竞争重点，也由“能否完成一个动作”逐步转向“能否理解任务、适应变化并持续执行”。

2022 年 12 月，Google 发布 RT-1，将摄像头图像和语言指令作为输入，直接输出机器人动作。模型使用 13 台机器人在 17 个月内采集的约 13 万次操作数据训练，覆盖 700 余项任务。相比为每项任务单独开发策略，RT-1 的进步在于让一个模型学习多种操作，并适应新的物体和环境，为机器人模型规模化训练奠定基础。

图表8：Google RT-1 架构

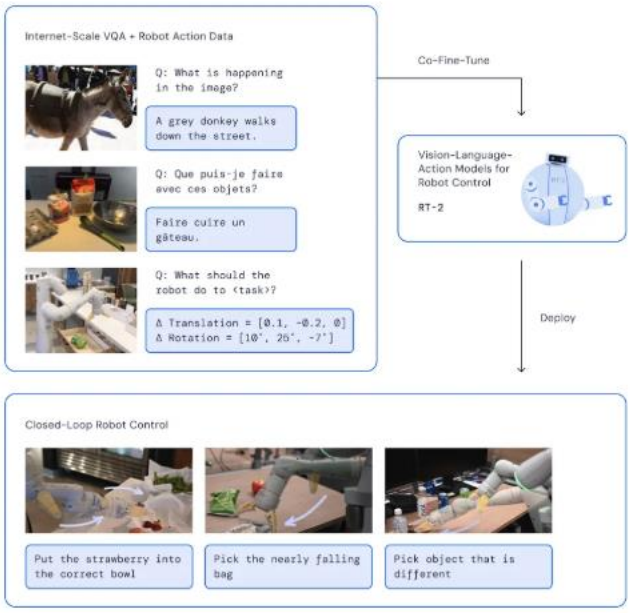


RT-1's architecture: The model takes a text instruction and set of images as inputs, encodes them as tokens via a pre-trained FiLM EfficientNet model and compresses them via TokenLearner. These are then fed into the Transformer, which outputs action tokens.

来源：Google Research，国金证券研究所

2023 年 7 月，Google 发布 RT-2，将大规模视觉语言预训练与机器人操作数据结合，形成视觉—语言—动作模型，即 VLA。相比主要从机器人示范中学习的 RT-1，RT-2 能够利用通用知识理解未专门训练过的指令，例如判断哪种物体可以临时充当锤子。在官方未见场景测试中，其表现由 RT-1 的 32%提升至 62%，验证了通用模型知识向机器人操作迁移的价值。

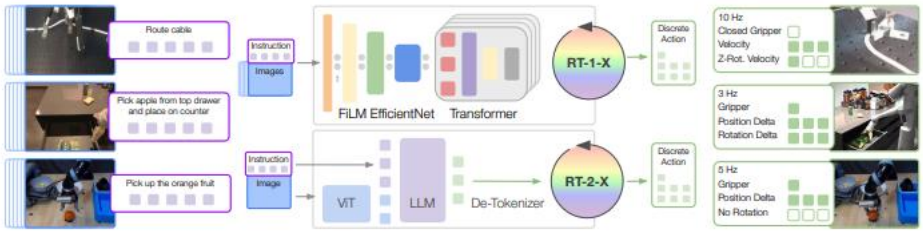
图表9：Google RT-2 架构与训练



来源：Google Research，国金证券研究所

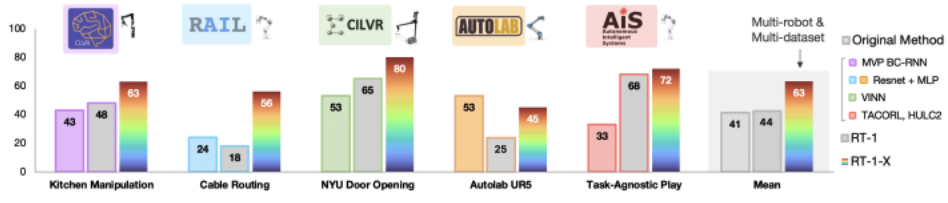
2023 年，Google 与多家机构推进 Open X-Embodiment 及 RT-X 研究，汇集 22 种机器人的数据。其核心变化是，不同机器人采集的经验可以共同训练模型，并改善其他平台的表现。这一路线为后续跨本体迁移提供基础，也有望降低每款新机器人从头采集数据和训练模型的成本。

图表10：Google RT-1-X 与 RT-2-X 架构图



来源：Arxiv，国金证券研究所

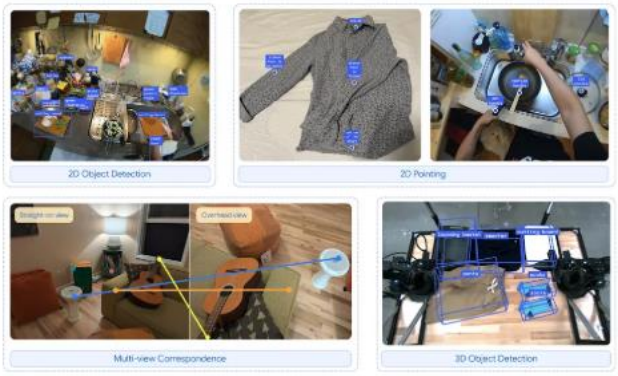
图表11：Google RT-1 与 RT-1-X 在小数据集场景下的成功率对比



来源：Arxiv，国金证券研究所

2025 年 3 月推出的 Gemini Robotics-ER 模型侧重具身推理，能够识别物体及其局部、寻找不同视角下的对应点，并检测三维空间中的物体。例如，看到杯子后，模型可以判断杯柄的抓取位置和接近路径。它可结合既有底层控制器，串联感知、状态估计、空间理解、规划和代码生成；在官方端到端测试设置中，成功率为 Gemini 2.0 的 2-3 倍。代码生成不足以解决任务时，还可通过上下文学习利用少量人类示范。其定位是增强机器人理解与规划，而非替代全部底层控制。

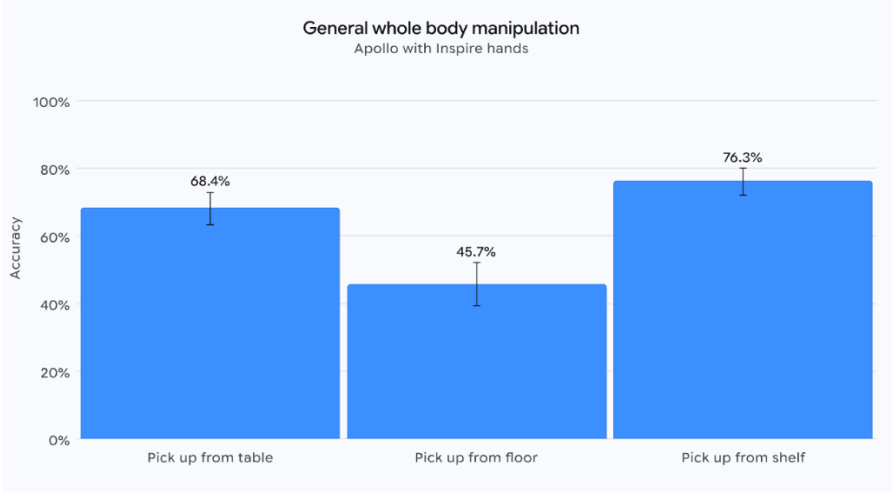
图表12：Google-ER能够识别物体、指向物体的各个部分



来源：Google Research，国金证券研究所

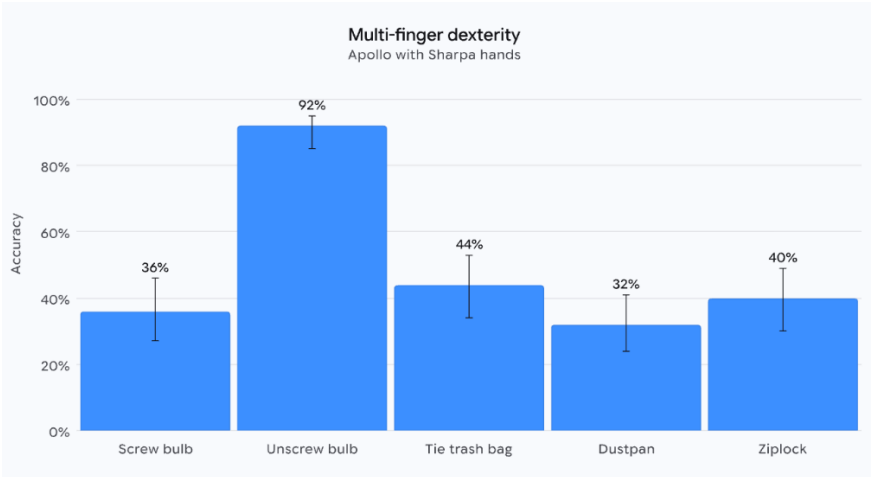
年2026，Google推出Robotics 2、ER 2和On-Device 2三类模型。Robotics 2将控制范围扩展至行走、下蹲、伸展与手部操作，使机器人能够移动到目标位置后完成任务；ER 2加强持续数分钟的任务规划、进度跟踪和失败纠正，并支持不同机器人协作；On-Device 2则继续降低新本体适配成本。相比上一代，升级重点是将移动、操作和长任务组织起来，但多指精细操作与执行速度仍有提升空间。

图表13：Robotics 2搭载Inspire灵巧手的全身操作任务成功率



来源：Google DeepMind，国金证券研究所

图表14：Robotics 2搭载SharpaWave五指灵巧手的精细操作任务成功率



来源：Google DeepMind，国金证券研究所

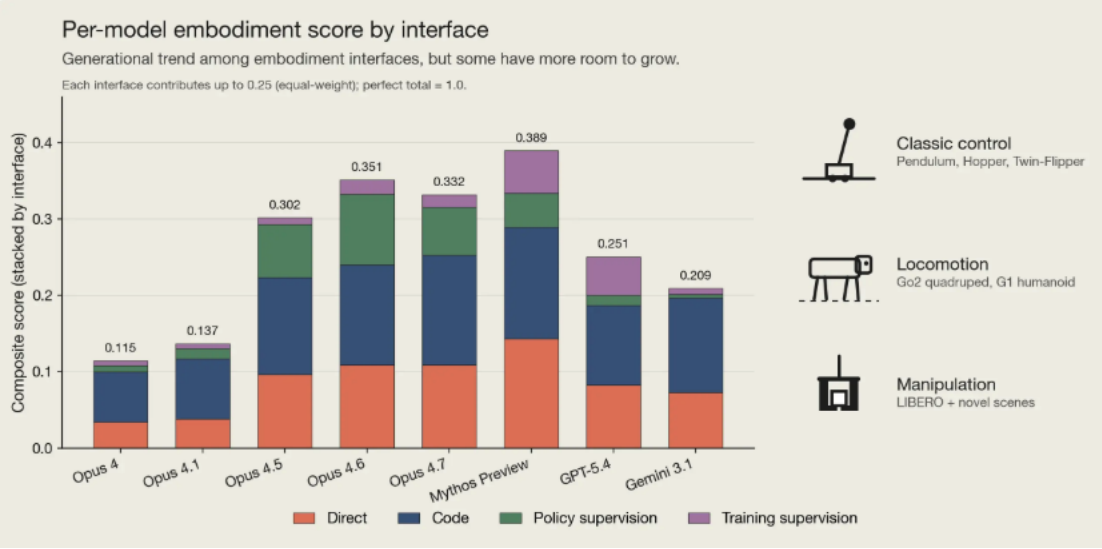
跨层协同而非单一模型参数，是 Google 机器人体系的核心观察点。高层推理负责理解任务与执行进度，动作模型将观测转化为控制目标，端侧部署则承接低延迟和稳定运行需求。上层能力迭代能否转化为下层策略泛化，合成数据能否改善实机表现，以及新本体适配成本能否下降，将共同决定体系优势的兑现程度。我们认为，具备模型、数据、仿真和部署工具积累的平台，更有条件缩短研发反馈周期，但本体适配仍是落地必经环节。对产业而言，平台能力的逐步完善有望降低开发者整合多个独立模块的成本，推动机器人研发从定制工程向工具化、平台化演进。

1.2.3 Anthropic 以能力研究与外部接触补充机器人储备

Anthropic 补充物理交互能力具备战略合理性，外部团队可能成为技术与数据积累的补充路径。外媒曾报道 Anthropic 与 Physical Intelligence 有收购接触；PI 否认已达成交易。语言、编程与工具使用为高层智能提供基础，而真实机器人研发还依赖交互数据、硬件工程和持续试验条件。相较完全自建，合作、投资或并购有望缩短能力储备周期，但目前相关信息仍属接触报道，尚不足以形成正式交易判断。

Anthropic 已有公开机器人研究，能力补足的重点在于将通用智能与控制、数据和部署环节连接。Anthropic 于 2026 年 7 月发表机器人能力研究，比较直接控制、程序控制以及调用已有策略等路径。相关研究为通用模型如何参与物理任务提供实验依据。相较 Google 较完整的多层产品体系，Anthropic 在本体控制、世界模拟及部署工具方面的公开布局仍有扩展空间，其后续进展值得从控制接口、真实测试和外部合作三个维度持续跟踪。

图表15：Anthropic Embody 测评各大模型机器人控制能力



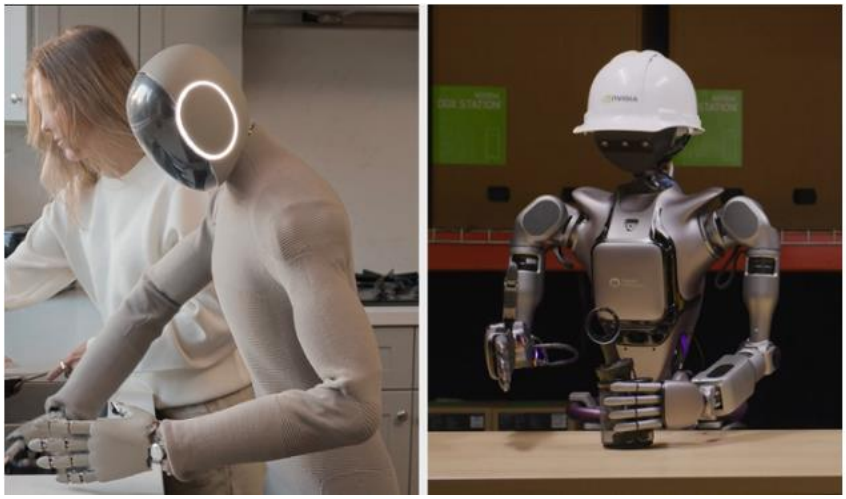
来源：Anthropic，国金证券研究所

巨头的竞争边界向真实交互延伸，将提高独立具身模型企业的差异化要求。对通用模型企业而言，外部团队的潜在价值在于补充真实数据、硬件工程和试验条件；对专用模型企业而言，可持续竞争力则更多取决于难以复用的数据资产、动作能力及部署经验。我们判断，行业分工可能从各自构建完整模型栈，逐步转向通用能力与专用能力协同。拥有真实任务积累、控制能力及客户场景的团队，有望在这一过程中保持更强议价能力。

1.3 英伟达持续强化 Physical AI 框架，NVIDIA 全栈布局降低开发门槛

NVIDIA 对 Physical AI 的持续投入，正将产业愿景转化为覆盖开发、训练与部署的基础设施。2024 年 3 月宣布 Project GR00T；2025 年 3 月发布 Isaac GR00T N1 并更新 Cosmos 相关工具；2026 年 10 月 21 日柏林 GTC 期间，进一步扩展世界模型训练、评估与部署工具。其中，2025 年 1 月 CES 发布 Cosmos 是重要节点，黄仁勋将世界基础模型与机器人发展的下一阶段联系起来，强调降低开发者获取数据和训练模型的门槛。模型、合成数据与仿真工具持续完善，有望推动机器人开发流程逐步标准化。

图表16：NVIDIA Isaac GR00T N1



来源：NVIDIA，国金证券研究所

柏林 GTC 强化采集、模型、仿真与部署协同，Physical AI 的产业内涵进一步清晰。黄仁勋在柏林主题演讲中介绍的人形机器人参考平台，连接了 Isaac Capture、GR00T、Isaac Lab 与 Isaac OS 等环节。GPU 提供训练与推理算力，Omniverse 承接数字环境构建，Isaac 服务机器人开发和仿真，Jetson 面向端侧部署，Cosmos 与 GR00T 分别补充世界建模与机器人智能。由此，Physical AI 的受益范围由人形机器人本体，扩展至自动驾驶、工业设备及其他具备感知、决策和执行需求的系统。

图表17：黄仁勋在柏林展示人形机器人参考平台



来源：NVIDIA，国金证券研究所

开发工具标准化有望扩大机器人研发需求，同时推动第三方厂商向专业化环节集中。现实采集、数字环境构建、模型训练、仿真评估及真实部署连接后，研发流程能够形成持续反馈，降低重复建设和试错成本。但基础工具普及也意味着通用功能更易获取，第三方软件的商业价值将更多来自复杂物理求解、真实场景数据及行业交付能力。对国内厂商而言，既有工业客户、场景资产和仿真积累能否转化为机器人训练所需的产品，是承接需求扩张的关键；与海外平台适配只是入口，客户付费与复用率才决定持续收入。

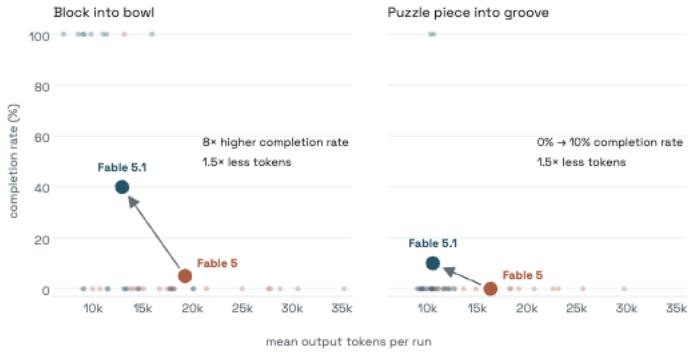
资本投入、模型扩散与工具链完善，共同构成本轮物理 AI 的催化闭环。World Labs 交易体现巨头对空间智能的战略定价，模型御三家推动通用能力向机器人迁移，NVIDIA 则补充训练、仿真与部署基础设施。我们认为，三类进展使机器人研究框架由本体和零部件，进一步扩展至数据采集、世界模型与训练软件。短期催化集中在技术和产品发布，中期验证重点则是模型能力能否转化为任务改善、客户导入与实际收入，产业链各环节的兑现节奏将存在差异。

二、产业进展：通用模型开始参与操作，公开评测提高行业能见度

2.1 通用模型涌现具身能力

通用模型已能借助工具接口参与机械臂操作。 Robocurve 的测试将模型接入 YAM 机械臂，通过相机和机器人状态获取观测，再调用末端位姿接口；逆运动学由机器人系统处理。Fable 5.1 等现成模型因此能够完成识别、规划和动作调用。其意义在于减少高层任务的专项开发，但系统仍依赖标定、控制器和安全约束。

图表18：Robocurve 测试发现 Fable 5.1 在输出较少 token 时完成了更多的工作



来源：Anthropic Robocurve，国金证券研究所

抓放取得进展，精细接触仍是短板。Robocurve 报告中，积木入碗任务里 GPT-6 Astra 完成 19/20 次，Fable 5.1 完成 8/20 次；圆形拼图嵌入任务中，两者均仅完成 2/20 次。两款模型在积木任务的平均每次试验耗时分别为 2.5 分钟、6.8 分钟，拼图任务为 3.4 分钟、5.9 分钟。每模型每任务仅测试 20 次，且存在设备台次和测试时间差异，结果尚不能代表工业可靠性。

图表19：Robocurve 测试结果

Task	Model	Mean stage	Completions	Rate	Output tokens/run	Est. cost/run	Minutes/run
Block into bowl	Fable 5	1.30	1 / 20	5%	19.2k	$2.69	8.2
Block into bowl	Fable 5.1	2.40	8 / 20	40%	12.9k	$2.12	6.8
Block into bowl	GPT-6 Astra	3.95	19 / 20	95%	2.1k	$0.94	2.5
Puzzle into groove	Fable 5	1.50	0 / 20	0%	16.3k	$2.63	7.9
Puzzle into groove	Fable 5.1	2.35	2 / 20	10%	10.5k	$2.18	5.9
Puzzle into groove	GPT-6 Astra	2.00	2 / 20	10%	2.7k	$1.36	3.4

来源：Anthropic Robocurve，国金证券研究所

总体来看，本轮模型升级的核心变化并非单项任务成功率的提升，而是通用模型开始涌现具身能力。在未针对特定机器人本体进行大规模微调和后训练的情况下，模型已经能够根据语言指令和摄像头画面完成任务理解、空间定位、轨迹规划，甚至生成可交互的仿真环境。这意味着机器人智能正在由依赖专用模型和单一本体训练，逐步转向由通用模型提供基础能力。

2.2 机器人 Benchmark 的诞生

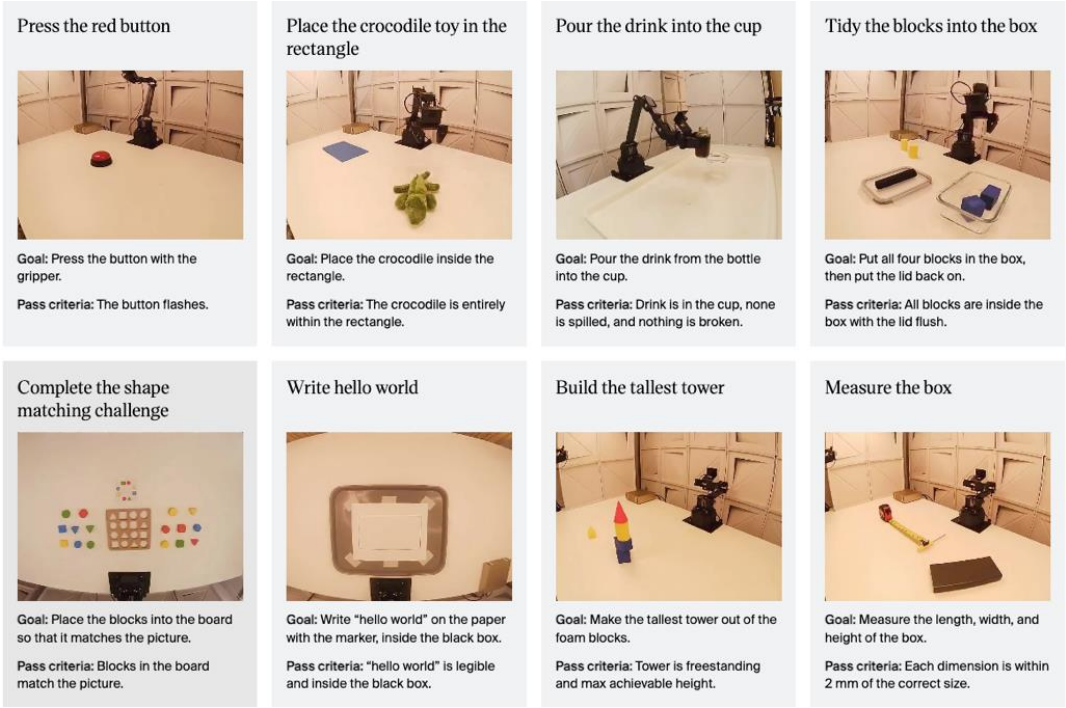
大语言模型已经形成较成熟的公开榜单和标准化测试，模型能力可以通过统一题目持续比较。具身智能的评测则更加复杂，不同机器人在硬件结构、传感器、控制接口和任务环境上存在明显差异，企业发布的演示视频也往往缺少统一测试条件，难以直接横向比较。

BEHAVIOR Challenge 等赛事较早尝试建立具身智能评测体系，在模拟家居环境中测试机器人的任务理解、长程规划和操作能力。但由于任务复杂、评测成本较高，公开参赛团队及可比结果仍然有限；同时，模型排名可能影响外界对公司技术能力和估值的判断，企

业公开参评的意愿相对谨慎，具身智能尚未形成类似大语言模型榜单的广泛影响力。

Artificial Analysis 推出的 AA-Robotics v0.1 进一步降低了横向比较的难度。该评测在统一机械臂、桌面环境和任务条件下，让模型根据摄像头画面自主编写控制代码，零样本完成按按钮、倒水、收纳积木、形状匹配、写字、搭塔和物体测量等 8 项任务，综合考察视觉理解、空间推理、任务规划和动作控制能力。

图表20：Artificial Analysis 八类桌面任务测试



来源：Artificial Analysis，国金证券研究所

图表21：Artificial Analysis 测试结果

● Full score (100) ● Partial (1 to 99) ● Zero (0)

Task	GPT-6 Astra	GPT-6.1 Sol	Gemini 4 Argon	Claude Opus 5.5	Claude Fable 5.1
Press the red button	●	●	●	●	●
Place the crocodile toy in the rectangle	●	●	●	●	●
Pour the drink into the cup	●	●	●	●	●
Tidy the blocks into the box	●	●	●	●	●
Complete the shape matching challenge	●	●	●	●	●
Write hello world	●	●	●	●	●
Build the tallest tower	●	●	●	●	●
Measure the box	●	●	●	●	●
AA-Robotics Score	69.4	64.6	55.3	48.8	46.3

来源：Artificial Analysis，国金证券研究所

从结果看，GPT-6 Astra 以 69.4 分排名第一，GPT-6.1 Sol、Gemini 4 Argon、Claude Opus 5.5 和 Claude Fable 5.1 分别获得 64.6、55.3、48.8 和 46.3 分。五款模型均能完成按按钮和放置玩具等简单任务，但在倒水、搭建积木和精确测量等任务上明显分化，其中没有模型完整完成尺寸测量，说明通用模型已经具备基础机器人操作能力，但精细控制和复杂接触仍是主要短板。

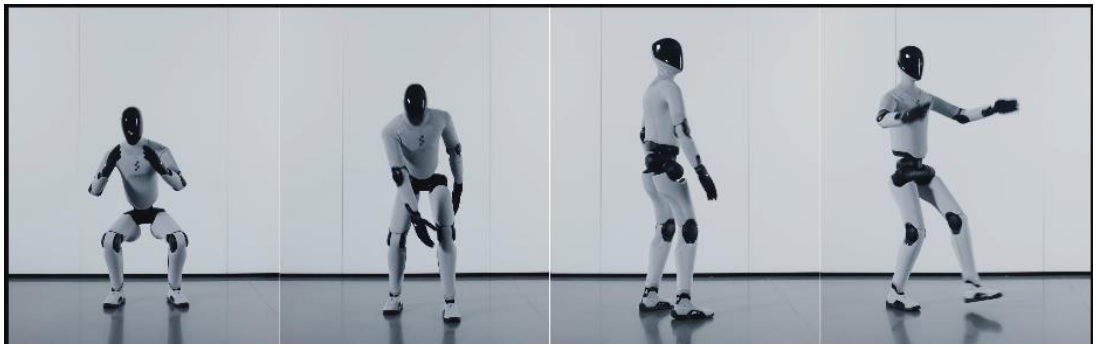
Benchmark 的出现意味着具身智能正在从依赖演示视频和公司叙事，逐步进入统一测试、横向比较和持续迭代的阶段。随着评测体系完善，模型能力进步将更容易被跟踪，具身智能产业的技术能见度和基本面可验证性也有望进一步提升。

2.3 S0、S1、S2 构成具身智能的分层架构

随着机器人任务复杂度提升，单一模型难以同时兼顾高层推理、快速反应和底层控制，分层架构逐渐成为行业共识。Figure Helix 02 是其中的代表性系统，其将机器人智能划分为 S0、S1 和 S2 三层，使不同模型按照各自适合的频率运行，形成从任务理解到电机执行的完整链路。

S0 是底层控制和执行层，负责将位置、姿态和关节目标转化为连续的执行器指令，并处理全身平衡、接触和动作协调。Helix 02 的 S0 以 1,000Hz 运行，保证机器人动作的实时性和稳定性。该层对运动控制、硬件适配和工程经验要求较高，也是国内机器人本体及零部件企业积累相对深厚的环节。

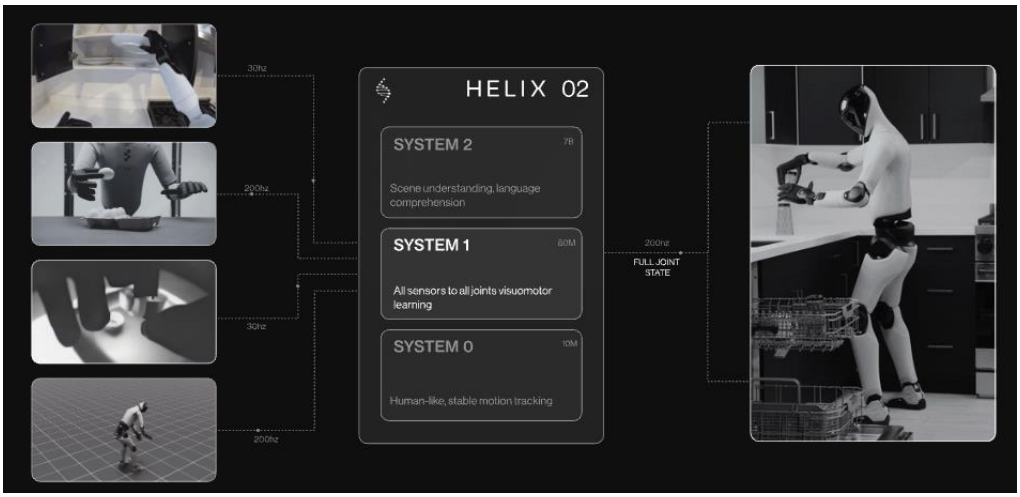
图表22：S0 基于人类数据的全身体制式控制



来源：Figure Helix 02，国金证券研究所

S1 是 VLA 或动作策略层，根据摄像头、触觉传感器和机器人状态生成下一步动作。Helix 02 的 S1 以 200Hz 输出全身关节目标，再由 S0 完成底层执行。不同模型的接口并不完全相同，部分 VLA 直接输出关节或末端动作，部分模型则生成轨迹或未来视觉状态，再由控制器转换为机器人指令，因此不能简单将 S1 统一理解为输出带轨迹的图像。

图表23：Helix S1 视觉运动策略



来源：Figure Helix 02，国金证券研究所

S2 是高层语义理解和任务规划层，负责理解语言指令、识别场景并拆解任务。在 Helix 02 中，S2 输出语义目标，引导 S1 生成动作；在其他技术路线中，S2 也可以引入世界模型，预测动作对环境造成的变化，并据此调整后续规划。

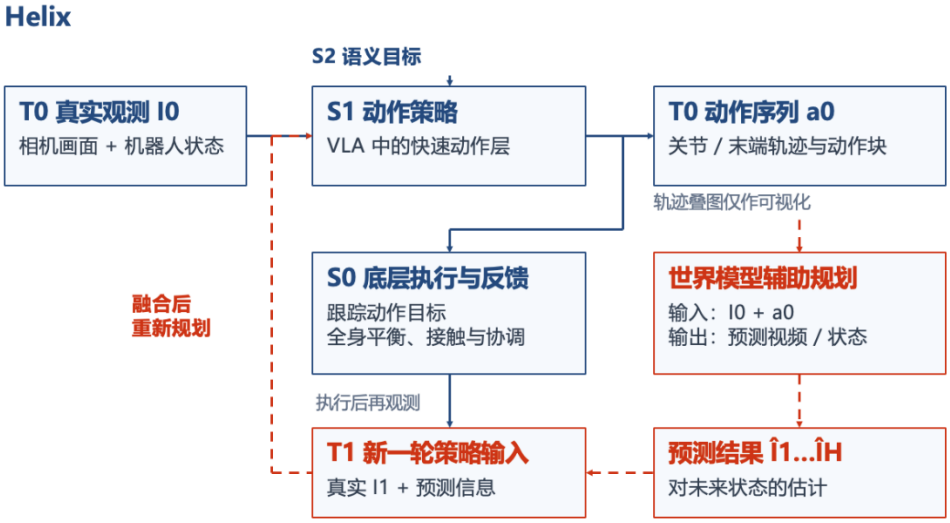
图表24：Helix S2 自主式长视距操控系统



来源：Figure Helix 02，国金证券研究所

三层系统共同构成闭环：S2 确定任务目标，S1 根据实时感知生成动作，S0 完成稳定执行，传感器再将环境变化反馈给上层模型并重新规划。该架构兼顾通用模型的推理能力与底层控制的实时性；随着 S2 进一步融合世界模型，机器人将不仅能够理解当前环境，还可以预测动作执行后的状态变化，为长程规划和闭环纠错提供基础。

图表25：Helix S0-S2 架构



来源：figure，国金证券研究所

2.4 巨头入局有望加速具身智能进入产业化阶段

模型巨头进入物理 AI，首先将增强 S2 层的任务理解、场景推理和长程规划能力。随着多模态模型进一步融合空间理解和世界模型，机器人能够更准确地理解语言指令，并预测动作可能造成的环境变化。未来机器人公司即使缺少自研基础模型，也有望通过接入通用模型获得基础的高层智能。

在此基础上，模型巨头的能力还可能继续向 S1 动作策略层渗透，将视觉、语言和机器人状态直接转化为动作或轨迹。通用 VLA、模型接口和开发工具逐步开放后，机器人企业训练基础智能的门槛有望下降，整机的任务理解和操作能力下限将被整体抬高。

但巨头入局并不意味着所有具身模型公司都会受益。如果 S2 能力逐渐由通用模型统一提供，S1 也被基础模型部分吸收，缺少数据、场景或硬件壁垒的独立 VLA 和世界模型公司，其技术稀缺性可能下降。行业竞争重点将进一步转向 S0 底层控制、真实任务数据、硬件适配、系统集成和安全执行等更接近实际部署的环节。

总体来看，模型巨头入局有利于加快机器人智能能力普及，缩短产品研发和迭代周期，对机器人产业整体构成利好；但产业价值可能向数据、硬件和场景闭环重新分配，不同类型

具身智能公司的受益程度将出现明显分化。

三、交易分析：关注本体、数采及仿真龙头

3.1 板块位置与交易结构

从 2026 年三季度至今的交易结构看，物理 AI 仍处于产业催化和主题交易驱动阶段，市场定价尚未形成高度一致的共识。板块覆盖机器人本体、核心零部件、数据采集、仿真渲染和模型软件等多个环节，不同公司的业务相关度和基本面差异较大，资金更容易向产业逻辑清晰、流动性较好的龙头集中。

从指数表现看，机器人、计算机和数字传媒板块均经历明显调整，前期获利盘得到较充分消化。从股价位置和阶段回撤幅度看，上述板块的整体筹码结构均相对较轻，其中机器人板块仍处于相对低位。

在科技板块内部轮动较快、市场交易热度仍然较高的环境下，资金倾向于寻找位置较低、产业逻辑清晰且催化较为密集的新方向。物理 AI 同时具备世界模型获得资本验证、模型巨头进入机器人以及产品持续迭代等催化，机器人本体、数据采集和仿真训练等环节有望承接科技资金扩散。

总体来看，物理 AI 当前的交易优势主要来自相对位置较低、筹码压力较小和产业催化密集，但行情持续性仍需机器人订单、产品交付和收入兑现进一步验证。。

图表26：国证机器人产业指数 2026 年三季度以来走势



来源：国证指数，国金证券研究所

图表27：中证指数 CS 计算机指数 2026 年三季度以来走势



来源：中证指数，国金证券研究所

图表28：国证数字传媒指数 2026 年三季度以来走势



来源：国证指数，国金证券研究所

3.2 三条投资主线

3.2.1 机器人本体与 T 链供应链

特斯拉 Optimus 已在弗里蒙特工厂启动高产量生产线的前期生产，公司预计仍需一年左右解决机电设计、可靠性及 AI 软件等问题，目标一年后达到周产能 2 万台，并于 2027 年底开始向外部企业客户交付。当前技术难点主要集中于手部设计，需要在耐用性与灵敏度之间取得平衡；AI 方面则计划通过自我博弈和仿真训练持续积累数据，并预计于 2027 年 4 月左右集成 AI 5 芯片，推动算力提升一个数量级。

图表29：Optimus 位于费蒙市第一条产线



来源：Tesla，国金证券研究所

机器人本体端重点关注宇树科技和上纬新材：宇树已经形成覆盖不同尺寸和应用场景的人形机器人产品矩阵，上纬新材则开始布局消费级具身智能机器人，

2026 年 9 月 20 日上纬新材旗下消费级具身智能品牌启元机器人在上海举行“我和我的个人机器人”全球产品发布会，启元 Q1、启元 T1 正式开售，10 月 1 日起按订单顺序陆续发货。产业链方面，可以关注三花智控、拓普集团等 T 链代表公司，其在机电执行器、机器人部件和规模制造方面具备一定基础，后续弹性主要取决于下游机器人放量节奏。

图表30：上纬新材旗下启元机器人



来源：盖世研究院，国金证券研究所

3.2.2 数据采集与空间感知

随着具身模型持续迭代，高质量空间和交互数据的重要性进一步提升。奥比中光的深度相机能够为机器人提供 RGB-D 数据和三维空间感知，禾赛则通过 Kosmo 等产品将真实环境转化为高质量三维数据。该方向的核心逻辑是，数据采集设备有望由机器人硬件配套件，逐步升级为模型训练和世界模型构建的数据入口。

图表31：奥比中光部分相机产品矩阵



来源：奥比中光官网，国金证券研究所

图表32：禾赛 Kosmo 生成的数据能针对目标机器人的相机特性做视觉适配

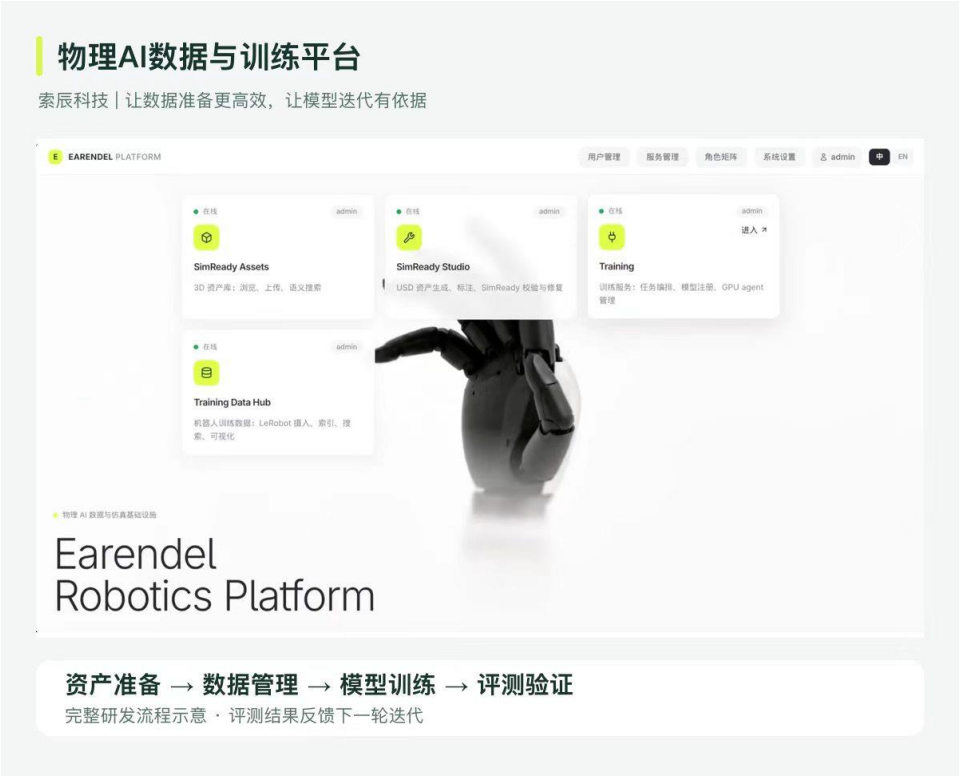


来源：禾赛官网，国金证券研究所

3.2.3 仿真、渲染与模型

真实机器人数据采集成本较高，仿真环境和合成数据能够提高训练效率，并降低模型在真实设备上的试错成本。索辰科技依托 CAE 仿真能力布局“开物”物理 AI 平台、合成数据和世界物理模型；凡拓数创则依托 3D 渲染和数字孪生能力，向机器人仿真训练、合成数据及 Sim2Real 闭环延伸。相关业务仍处于产业化早期，后续需要重点跟踪产品交付和收入兑现情况。

图表33：索辰科技物理 AI 数据与训练平台



来源：索辰科技公众号，国金证券研究所

图表34：凡拓具身智能训练场方案核心架构



来源：凡拓数创官网，国金证券研究所

四、相关标的

1）本体：特斯拉、智元、上纬新材、宇树科技。

2）北美链：科森科技、斯菱智驱、岱美股份、世运电路、福赛科技、蓝思科技、浙江荣泰、三花智控、恒立液压、北特科技、拓普集团、科达利、万向钱潮等。

3）机器人大脑相关：摩尔线程、协创数据、索辰科技、仙工智能、宝通科技、五一视界、海天瑞声、道通科技、群核科技、智微智能、爱芯元智等。

4）其他核心机器人链：领益智造、奥比中光、绿的谐波、禾赛、速腾聚创、海康威视等。

五、风险提示

■ 通用模型具身能力进展不及预期的风险。

当前通用模型在空间理解、任务规划和机器人控制等方面已展现一定能力，但相关测试多集中于标准化环境和有限任务。若模型在复杂场景、长程任务、精细操作和跨本体泛化方面的提升速度低于预期，可能影响物理 AI 产业的发展节奏。

■ 机器人商业化及规模放量不及预期的风险。

当前人形机器人及具身智能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、续航和安全性未能满足实际需求，或下游客户付费意愿不足，可能导致机器人量产、订单和收入兑现节奏低于预期。

■ Benchmark 及公开演示代表性不足的风险。

现有机器人评测体系仍处于发展初期，测试任务、硬件平台和评分标准尚未完全统一，标准化环境中的测试结果以及企业公开演示，可能无法充分反映模型在真实复杂环境中的实际能力。

■ 仿真合成数据及 Sim2Real 进展不及预期的风险。

仿真训练和合成数据能够降低真实数据采集成本，但虚拟环境与真实世界之间仍存在差异。若仿真精度、数据质量或跨场景迁移效果不及预期，可能限制机器人模型的训练效率和真实环境泛化能力。

■ AI 芯片、算力和端侧部署进展不及预期的风险。

具身模型训练及机器人实时推理对算力、存储、通信带宽和功耗提出较高要求，若芯片研发、算力供给或端侧部署进度低于预期，可能限制模型迭代速度和机器人规模化应用。

- 行业竞争加剧及相关公司估值波动的风险。

随着科技巨头和创业公司持续进入物理 AI 领域，模型、机器人本体及零部件环节的竞争可能加剧，并带来产品降价、技术路线调整和供应链份额变化。部分相关公司业务仍处于产业化早期，若基本面兑现不及预期，可能面临估值回落和股价波动风险。

行业投资评级的说明：

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；

增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；

中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；

减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。

特别声明：

国金证券股份有限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。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海
电话：021-80234211
邮箱：researchsh@gjzq.com.cn
邮编：201204
地址：上海浦东新区芳甸路 1088 号紫竹国际大厦 5 楼

北京
电话：010-85950438
邮箱：researchbj@gjzq.com.cn
邮编：100005
地址：北京市东城区建内大街 26 号新闻大厦 8 层南侧

深圳
电话：0755-86695353
邮箱：researchsz@gjzq.com.cn
邮编：518000
地址：深圳市福田区金田路 2028 号皇岗商务中心 18 楼 1806

【小程序】
国金证券研究服务

【公众号】
国金证券研究