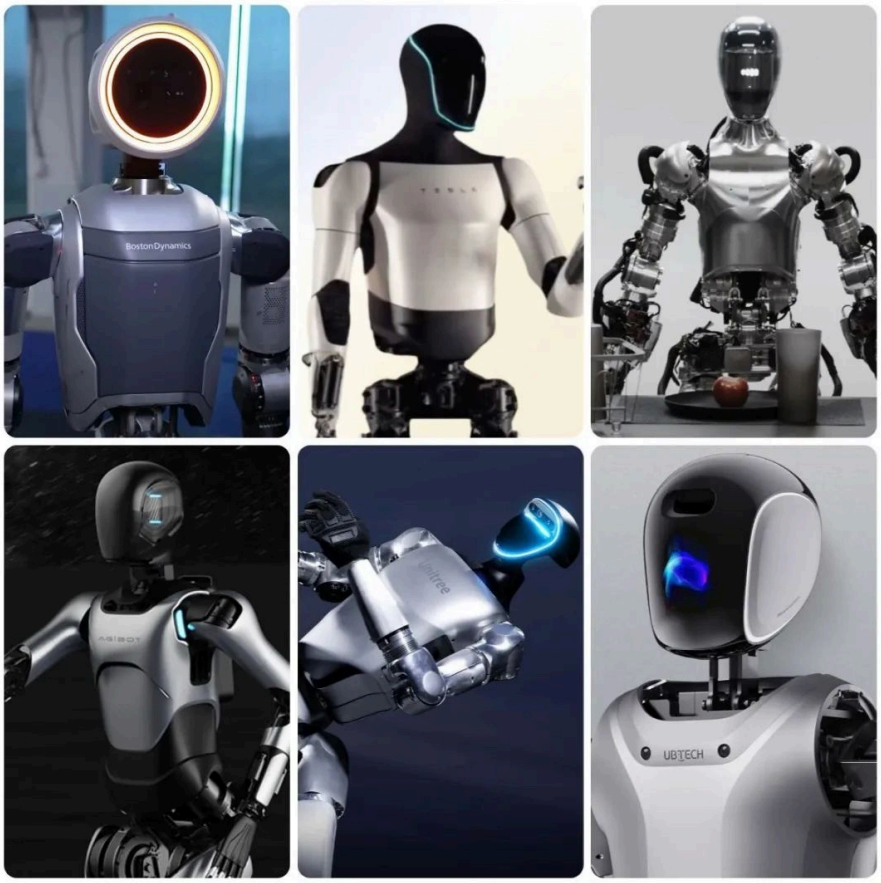
知行元界

具身智能大脑产业发展蓝皮书

——从“身体竞争”到“大脑竞争”

（2026 年）



编写：知行元界

日期：2026 年 11 月

报告编制委员会

顾问委员会：施炜、左东旭、屈全军

执行主编：张利、张献超、刘洪兵、贾维海

委员：和立强、李博、孙忠磊、刘俊男、朱宇斌

发布单位：天津知行元界智能科技有限公司

幸福领航（天津）管理咨询有限公司

实践智慧研究院

华夏基石（具身智能咨询研究中心）

北京汇众绥猷管理咨询有限公司

联系我们：

目　　录

摘要 1

核心判断 1

全景鸟瞰 3

给三类主体的行动指向 3

第一章 绪论：为什么“大脑”是具身智能的胜负手 5

1.1 从 ChatGPT 到“物理 AI”：智能进身的必然路径 5

1.2 具身智能产业链全景：大脑—小脑—本体的三层解构 6

1.3 Moravec 悖论与具身大脑要解决的四个问题 8

1.4 概念界定 9

1.5 研究方法与数据口径说明 10

1.6 报告结构与阅读指引 11

第二章 发展脉络：从专用到通用的四次范式跃迁 12

2.1 范式前史（2012－2022）：深度学习时代的三条伏笔 12

2.2 第一次跃迁（2022－2023）：VLA 概念确立，机器人有了“语言接口” 13

2.3 第二次跃迁（2024）：π0 时刻——“大脑”独立成赛道 14

2.4 第三次跃迁（2025）：双系统架构成为主流，巨头全面入场 15

2.5 第四次跃迁（2026）：路线分岔与融合收敛——走向世界-动作模型 16

2.6 中外发展节奏对比：同一时间轴上的两种打法 19

2.7 人才与组织：大脑竞赛的隐形战场 21

2.8 本章小结：三条历史规律与一个时间判断 22

第三章 技术路线：三足鼎立与融合收敛 23

3.1 VLA 端到端路线：主流与瓶颈 24

3.2 世界模型路线：先推演、后执行 25

3.3 世界-动作模型（WAM）：范式跃迁的主航道 28

3.4 本能驱动路线：触觉优先的“第三极” 29

3.5 双系统架构：从分歧走向工程共识 29
3.6 通用基模的具身突袭：从“能不能”到“接哪段” 31
3.7 数据范式：从遥操作到“人类学习” 32
3.8 评测与标准：话语权之争 34
3.9 端侧部署与算力约束：被低估的工程主战场 35
3.10 技术选型的五个工程判据 36
3.11 增量观察：2026 年 10—11 月的六项进展与三点修正 37
3.12 本章结论：融合收敛，分层永存 40
第四章 代表案例：全球具身大脑竞争格局 42
4.1 国际阵营：从“四极”到多线并进 43
4.2 中国梯队 52
4.3 格局综合研判 59
4.4 案例启示录：六个可迁移的战略规律 63
4.5 资本地图：谁在为大脑下注 64
4.6 本章小结 65
第五章 产业逻辑：在产业链条中理解大脑 67
5.1 产业链解构：大脑的上下游版图 67
5.2 价值分布：大脑为何是“复利机器” 69
5.3 商业闭环：谁在为大脑付费？ 70
5.4 供给结构：三大瓶颈的现实约束 72
5.5 区域与政策格局：大脑产业的中国地图 72
5.6 单位经济模型：一台机器人何时算得过来 74
5.7 平台化与中间层挤压：大模型大厂改写分工 75
5.8 本章小结 77
第六章 未来预判：2026 – 2035 的五个关键推演 79
6.1 市场规模：万亿美元叙事下的结构分化 79

6.2 路线收敛：WAM 融合成为主航道，快慢分层永存 80
6.3 格局推演：大脑层的“操作系统之战”会重演吗？ 81
6.4 关键拐点：三个触发条件与一个时刻 82
6.5 风险清单：五个必须管理的下行风险 84
6.6 三幕情景：2030 年具身大脑世界的三种可能 85
6.7 中国的三个战略机遇窗口 86
6.8 本章小结 87
第七章 战略建议：三类主体的行动框架 89
7.1 企业：四张决策地图 89
7.2 投资机构：从“看叙事”到“看闭环” 91
7.3 政府与园区：把“大脑”当作基础设施来建 92
7.4 结语：从“会干活”到“靠得住” 93
附录 A 具身智能大脑大事记（2022 – 2026） 95
附录 B 主要 VLA/世界模型对照表（公开口径） 98
附录 C 术语表 100
参考文献（主要数据与事实来源） 102

摘要

核心判断

本蓝皮书研究的对象，是具身智能产业链条中最具战略价值、也最富争议的一环——**具身智能大脑**，即以具身基础模型（Embodied Foundation Model）为核心载体、承担机器人感知理解、任务规划与动作决策的智能层。

2026 年，具身智能产业正站在一个历史性的分水岭上。硬件本体的性能差距正在快速收窄，产业链的竞争焦点正从“身体”转向“大脑”。**本报告提出五条核心判断：**

判断一：产业价值重心正在从“身体”迁移至“大脑”。 硬件仍是必要基础，但不再是决定性变量。产业价值正沿着“零部件—本体—大脑—场景”的链条向上游智能层集中，“场景数据—训练迭代—部署运营”的闭环能力，正在取代机械性能成为新的估值锚点。2025 年全球机器人赛道约 90 亿美元融资中，人形机器人与 AI 基础模型公司占比超过 60%；2026 年前三季度国内具身智能融资 557 笔、披露金额约 2896 亿元（上奇产业通口径；高工 GGII 口径为 437 起、约 2032 亿元，同比增 317%;IT 桔子口径约 1286 亿），资本已经用真金白银完成了第一次投票。**但 2026 年四季度出现了值得警惕的分化：融资笔数在 8—9 月回落至每月 44 笔，二级市场用近 3000 亿元的市值摆动给“稀缺性溢价”上了一课——钱仍在流入，但开始区分“叙事”与“交付”。**

判断二：技术路线呈现“三足鼎立、融合收敛”，且大脑正在加速“分层”。 VLA（视觉-语言-动作模型）端到端路线是当前资本认可度最高、落地最快的方向；世界模型路线主张“先推演、后执行”；以触觉优先为特征的“本能驱动”路线构成第三极。VLA 与世界模型的融合（世界-动作模型，WAM）以及“快慢思考”双系统架构，正在成为下一代具身大脑的公认形态。2026 年 9—10 月，通用基础模型密集“降维进入”——GPT-6 Astra 被第三方实测用于直接控制真机、Gemini 4 Argon 将仿真环境构建成本从“周”压到“小时”、Physical Intelligence 开放 Pi API 把微调托管化——**“大脑”正在分化为“认知大脑（通用基模）—物理小脑（VLA/WAM）—端侧力控”三层**，产业分工

被重新改写。**10月下旬至11月，这一分化又被三件事加固：**智元GO-3 Preview把学习链条延伸到“部署后自进化”（真实便利店5天从28.6%爬升至85.7%）；高德λ0证明全身动作数据可以“无本体”采集（效率约为真机遥操的3.4倍）；地平线 HoloBrain-0、炎曦动力 Paimon-0、中国移动灵韬智脑则把“框架＋算力＋网络”做成了可单独供给的底座。**大脑不再只是一个模型，而是一套“模型＋工具链＋数据管线＋部署回路”的系统能力。**

判断三：数据是新的护城河，数据范式正在发生第三次迁移。 从遥操作真机数据，到仿真合成数据，再到人类第一视角视频数据（英伟达 EgoScale 使用约2.1万小时人类视频、Figure Index 汇聚超1600万条人类操作视频），获取数据的成本结构与应用边界被不断改写。谁掌握低成本、规模化、高质量的数据供给，谁就掌握具身大脑的“石油管道”。

判断四：商业模式四类并存，平台化正在挤压“中间层”。 纯大脑授权（Physical Intelligence）、软硬一体（智元、Figure AI）、平台生态（英伟达）、垂直整合（银河通用）四种模式同台竞技。Physical Intelligence 以约200亿美元（10月洽谈口径）、零产品零收入的姿态，检验着“每台机器交一次大脑授权费”这一假设的市场耐心；而其10月开放的Pi API（上传数据→平台托管微调→云端下发真机）预示着新的分工形态——模型公司收编微调与推理托管，只把硬件集成与真实场景数据留给下游。**与之形成对照的是 Skild AI：以“模型＋部署＋集成”的混合交付，在10个月内做到付费客户超60家、ARR 突破 1 亿美元——同一条赛道出现“买期权”与“买现金流”两套定价语言，说明商业模式尚未收敛。**本报告认为，大脑层的终局更可能介于“操作系统式赢者通吃”与“多脑并存”之间，场景数据主权将决定话语权分配。

判断五：中国拥有“并跑转领跑”的历史性窗口。 大模型第一梯队的技术底座、全球最强的运动控制与制造业供应链、连续12年全球第一大工业机器人市场所沉淀的场景纵深，三重优势在具身大脑这一环交汇。国务院发展研究中心预测，中国具身智能市场规模2030年将达4000亿元、2035年突破万亿元——其中大脑环节的价值占比将持续提升。

全景鸟瞰

具身大脑之争，本质上是四个层面的叠加竞争：**架构之争**（端到端还是分层、快脑还是慢脑）、**数据之争**（真机、仿真还是人类视频）、**生态之争**（谁定义标准、谁掌控评测话语权、谁绑定本体与场景）、**层级之争**（通用基模吃下认知与规划层之后，专用大脑退守哪一层）。美国阵营以英伟达、Google DeepMind、Physical Intelligence、Figure AI 为代表，分别占据平台生态、大厂分层、纯大脑、垂直整合四个生态位；OpenAI（GPT-6 Astra）与 Dyna Robotics（WAM 商业闭环）的加入使“国际四极”变为“多极＋新变量”，特斯拉以“FSD 复用＋车载芯片＋百万级车队数据飞轮”把垂直整合推向极致，李飞飞的 World Labs（空间智能世界模型，2026 年 9 月被 AMD 以约 82 亿美元收购）与杨立昆的 AMI Labs（JEPA 架构，10.3 亿美元种子轮）则为世界模型补上学界叙事；中国阵营以智元、银河通用、星海图、千寻智能、智平方、它石智航、极佳视界（世界模型原生）等为代表，形成了全栈自研、仿真数据驱动、“一脑多形”等差异化打法，2026 年四季度又补入四类新主体——高德（大厂全身基座）、地平线（开源框架＋端侧芯片）、炎曦动力（模型定义芯片）、中国移动（云脑＋快脑＋专网），DeepSeek 则以“投资宇树（身体）＋Winverse AI（大脑）＋运动控制团队（小脑）”的拼装路线入场。**行业共识正在从“谁的模型跑分更高”，转向“谁能够把通用基模的语义能力、世界模型的物理能力、现场沉淀的真实工况数据缝合在一起”，并把它固化为可复用的工具链与可计价的服务。**

给三类主体的行动指向

企业：本体厂面临“自研大脑还是接入第三方”的战略抉择，本报告给出四象限决策框架；AI 公司应从“做演示”转向“做闭环”，以场景数据主权构建复利壁垒。2026 年 9—10 月的格局变化给所有“中间层”主体新增了一道风险自检题：当通用基模吃下语义理解、任务规划与仿真生成，当平台收编微调与托管推理，你的不可替代性还剩什么？

投资机构：估值逻辑正从“硬件制造”转向“AI SaaS”，本报告提出七维尽调清单——架构先进性、数据资产、闭环能力、团队密度、生态位、交付验

证、部署后斜率——其中“部署后斜率”是本报告新增的动态指标：**看上岗后30天的成功率爬升速度**（GO-3 的 5 天曲线、Helix 2.5 的零样本迁移、Dyna-2 的 3 天回本，都是这一指标的实证）。同时要准备两把尺子：一把量 ARR，一把量平台调用量；用错尺子，会在同一赛道得出相反的结论。资本风向亦在快速切换：2026 年下半年机构上会项目明显减少、资金向头部集中，DeepSeek 等大模型公司成为新的整合方——二级市场的定价锚（宇树两个月内近 3000 亿元市值摆动）与一级市场的出清节奏（GGII 判断三年内 15%以上初创倒闭或被并购），都需要纳入投资决策模型。

政府与园区：把“大脑”作为未来产业基础设施来建设——训练场（数据基础设施）、标准体系（世界模型标准已列入国家人形机器人标准体系）、场景开放（城市即试验场）三位一体推进。

一句话总结：具身智能的下半场，是大脑的战争；大脑的胜负手，不在算法本身，而在数据飞轮能否转动、场景闭环能否跑通、生态标准能否立住。从“会干活”到“靠得住”，这条路刚刚开始。

第一章 绪论：为什么“大脑”是具身智能的胜负手

本章导读 本章先沿着人工智能的“进身之路”，说明具身智能为何是智能演进的必然阶段；再以“大脑—小脑—本体”三层架构解构机器人系统，论证为什么产业价值与竞争焦点正在向大脑环节集中；最后给出本报告的概念界定、研究方法与阅读指引，为全篇奠定分析框架。

1.1 从 ChatGPT 到“物理 AI”：智能进身的必然路径

人工智能的历史，可以概括为一部“智能进身史”——智能不断寻找新的载体，从文本进入图像，从图像进入语言对话，如今正在进入物理世界。

大语言模型（LLM）解决了“符号世界”的智能问题：理解语言、生成文本、进行推理。多模态模型解决了“感知世界”的智能问题：看懂图像、听懂语音。但这两类智能有一个共同的局限——**它们悬浮在物理世界之上**。一个能写出精妙商业计划书的模型，无法把一杯水稳稳地端到桌上；一个能识别千万种物体的视觉模型，不知道用多大力气捏住一枚鸡蛋而不捏碎它。

这就是具身智能（Embodied AI）要回答的问题：**让智能拥有身体，在物理世界中感知、决策、行动，并通过与环境的交互持续学习。**

2024 年以来，具身智能从学术概念跃升为全球科技竞争的战略高地。其背后有三重驱动力量：

第一，AI 技术成熟度的外溢。 Transformer 架构、大规模预训练、多模态对齐等技术在语言与图像领域的成功，为“把 Scaling Law（规模定律）复制到机器人领域”提供了方法论信心。业界普遍将 2022 年底 ChatGPT 的爆发视为分界：语言智能已被证明可以“大力出奇迹”，下一个问题自然是——物理智能是否同样如此。

第二，硬件本体的量产拐点。 中国供应商将人形机器人整机成本快速拉低，2025 年全球人形机器人出货量约 1.8 万台（IDC 数据，1 月报告口径），中国厂商贡献了主要部分。当“身体”开始规模化量产，“身体里装什么脑子”就成为最紧迫的问题——**本体越便宜、越普及，大脑的价值就越大。**

第三，劳动力结构变化的长期叙事。 全球主要经济体同时面临老龄化与制造业劳动力缺口。中国已连续 12 年稳居全球第一大工业机器人应用市场，制造业机器人密度跃升至全球第三；养老、物流、柔性制造等场景对“能适应变化的劳动力”的需求，远超传统自动化设备的能力边界。

图灵奖得主、中国科学院院士姚期智对具身智能的双重意义有一个精炼概括：在科学研究上，它帮助人工智能从大语言模型走向能够理解物理世界的智能模型；在经济产业上，它弥补了传统机器人产业的不足。**“理解物理世界的智能”与“替代人做不了的事”，正是具身大脑的价值原点。**

在这一进程中，“大脑”逐步从机器人整机中独立出来，成为一个可识别、可交易、可单独估值的产业环节——这正是本蓝皮书的研究对象。

1.2 具身智能产业链全景：大脑—小脑—本体的三层解构

要理解具身大脑的产业位置，先要解构具身智能机器人的系统分层。业界目前较为一致的说法，是把机器人的智能系统分为“大脑—小脑”两层，加上本体，构成三层结构：

第一层：大脑（高层智能）。 负责环境感知与语义理解、任务理解与分解、长程规划与推理、跨场景泛化。其技术载体是具身基础模型——主要是 VLA 模型、世界模型及其融合形态。它回答的是“做什么、为什么这样做、先做哪一步”的问题。

第二层：小脑（运动控制）。 负责全身运动控制、平衡与步态、力控与灵巧操作执行。技术载体包括强化学习运动策略、模型预测控制（MPC）、全身控制器等。它回答的是“怎么把动作做出来、做稳、做快”的问题。

第三层：本体（物理执行）。 包括机械结构、关节电机、减速器、灵巧手、传感器、电源与散热。它决定了能力与成本的物理边界。

需要强调两点。其一，这一分层是**分析性的而非物理性的**：在端到端技术路线中，“大脑”与“小脑”的功能可以被统一模型吸收，例如波士顿动力与丰田研究所合作的 Atlas 模型，用一个 4.5 亿参数的扩散 Transformer 策略，把移动（locomotion）与操作（manipulation）合并进一个 30Hz 的策略输出，

不再划分模块。其二，本报告所研究的“具身大脑”，采**广义口径**：既包括独立的具身基础模型，也包括软硬一体厂商自研的智能层——凡是承担“感知—决策—规划”功能的模型与软件栈，均在本研究范围之内。



注：分层为分析性划分。端到端路线可将大脑与小脑吸收进统一模型，但“认知决策”与“实时控制”的分工在工程上长期存在。
资料来源：本报告研究整理。

图 1-1　具身智能系统的三层架构

从产业链看，具身大脑处于“承上启下”的枢纽位置：

- **上游**：算力芯片（如英伟达 Jetson Thor、华为昇腾）、云与仿真平台（Isaac Sim、Genie Sim）、数据服务（遥操作采集、合成数据、人类视频数据集）；
- **中游**：具身基础模型（本报告核心）、中间件与运行时（Isaac OS 等）、开发工具链；
- **下游**：各类本体（人形、机械臂、移动操作平台、专用机器人）与场景应用（制造、物流、商业服务、家庭、科研）。

这一位置决定了大脑环节的两个产业特征：**其一，它是产业链的“价值放大器”**——同样一台本体，装配不同水平的大脑，商业价值可能相差数倍；**其二，它是生态的“定义者”**——大脑层一旦形成事实标准，将向上约束数据格式与算力平台、向下绑定本体接口与开发者生态，其地位类比 PC 时代的操作系统、移动时代的 iOS/Android。

1.3 Moravec 悖论与具身大脑要解决的四个问题

为什么机器人“大脑”如此难做？答案藏在一个经典悖论里。

1980 年代，机器人学家汉斯·莫拉维克（Hans Moravec）等人发现一个反直觉的事实：**对计算机而言，高层次的推理（下棋、证明定理）相对容易，而低层次的感知运动技能（走路、抓取、叠衣）却极其困难。** 一个五岁孩子轻松完成的端水杯、系鞋带，是数亿年进化刻进人类神经系统的“硬件级”能力；而下棋只需要几十年文明沉淀的规则知识。这就是著名的 **Moravec 悖论**。

大语言模型的成功，恰恰建立在“绕开悖论”的基础上——互联网文本为语言智能提供了近乎无限的训练数据。而物理交互数据面临三重困境：

困境一：贵。 真实世界的机器人交互数据主要依靠遥操作采集，一台机器人、一名操作员、一小时，只能产出有限的演示数据，成本高、效率低，远远无法匹配大模型对数据规模的需求。

困境二：散。 不同机器人的本体参数（自由度、负载、传动方式）各不相同，“身法”无法直接迁移——教给机械臂的动作指令，对人形机器人毫无意义。数据跨本体复用是行业性难题。

困境三：没有“物理教科书”。 语言有互联网，物理交互却没有一个现成的“开放数据集”可以爬取。物理世界的常识——水的流动性、玻璃的易碎性、摩擦的不可见性——从未被系统性数字化。

正因如此，具身大脑的产业竞争，实质上是围绕四个递进问题的攻坚：

问题一：看懂（感知理解）。 从传感器原始数据中重建对场景的语义与物理理解——这是谁在哪里、那是什么物体、它处于什么状态。

问题二：想清（推理规划）。 理解“做什么”，把模糊指令（“把房间收拾一下”）分解为可执行的任务序列，并在行动前预判后果——“我这样推，杯子会不会倒？”

问题三：会做（灵巧执行）。 把规划翻译成高频、精确、自适应的动作输出——包括应对工件形变、位置偏移等真实世界的扰动。

问题四：可靠（泛化与安全）。 在训练分布之外的场景中稳定工作，满足工业与家庭环境对成功率、安全性的严苛要求。

当前所有技术路线之争——VLA、世界模型、本能驱动、快慢双系统——本质上都是对这四个问题的不同求解次序与求解方式的取舍。**看懂与想清，是“大脑”的本职；会做与可靠，是“大脑”与“小脑”“本体”协同的边界地带。** 具身大脑的成熟度，可以用这四个问题的解决程度来度量——这也是本报告贯穿始终的分析框架。

1.4 概念界定

为避免概念混淆，本报告对核心术语作如下界定：

具身智能大脑（Embodied Brain）： 机器人系统中承担感知理解、任务规划与动作决策功能的智能层总称。广义的具身大脑包括基础模型、配套的运行时与工具链。

具身基础模型（Embodied Foundation Model）： 在大规模多模态、多任务、多本体数据上预训练，可通过微调或提示适配多种机器人任务的通用模型。是具身大脑的技术核心载体。

VLA（Vision-Language-Action，视觉-语言-动作模型）： 以视觉语言模型（VLM）为骨干、将视觉观测与语言指令直接映射为动作输出的端到端模型架构。2023 年由 Google RT-2 确立，2024 年起成为主流路线。

世界模型（World Model）： 在模型内部构建对物理环境的可模拟表征，能够依据当前状态与动作预测环境未来演化的模型。核心功能是“动手之前先推演”。

世界-动作模型（WAM，World-Action Model）： 2026 年出现的融合形态，将世界理解（预测未来状态）与动作生成统一进单一模型，代表者为英伟达 DreamZero（GROOT N2）与智元 GE-Act2.0。

双系统架构（Dual-System）： 借鉴认知科学“快与慢思考”（System 1/System 2）的分层架构——慢系统（VLM，数 Hz 运行）负责语义理解与规划，

快系统（小型策略网络，数十至数百 Hz 运行）负责高频动作生成。Figure Helix 是该架构的标志性实现。

一脑多机 / 跨本体泛化（Cross-Embodiment）： 同一个模型驱动不同构型、不同品牌机器人的能力，是具身大脑“通用性”的关键指标，也是纯大脑公司商业模式成立的技术前提。

1.5 研究方法与数据口径说明

作为产业蓝皮书，研究纪律先于研究结论。本报告采用四层研究方法：

方法一：多源事实核验。 所有关键事实（技术发布、融资数据、政策文件）均要求至少一个权威来源支撑，优先采信官方发布（企业官方渠道、政府文件、监管机构公告），次优采信主流财经媒体多源印证的报道。核心数据的关键口径包括：市场规模采用腾讯云×沙利文白皮书与国务院发展研究中心《中国发展报告 2025》两套口径并标注差异；融资数据采用 IT 桔子、人形机器人场景应用联盟统计并标注统计截止时点。

方法二：来源分级标注。 全文执行三级标注：无标注为多源印证的高置信事实；“据 XX”“企业披露口径”为单一来源或自证口径；“本报告认为”为研究判断。企业技术指标（成功率、速度、参数量）均属自证口径，正文统一以“官方发布”属性处理，读者在采购与投资决策中应要求第三方验证。

方法三：框架化分析。 产业分析采用“产业链—价值分布—飞轮机制—商业闭环”的分析链路；企业分析采用“定位—技术路线—数据策略—商业模式—验证场景”五要素结构；预判部分采用“关键假设＋验证信号”的可跟踪写法，便于读者按季度校准。

方法四：边界声明。 本报告数据截止 2026 年 11 月 13 日；对尚未发生的事项（市场终局、路线胜负）只作概率性推演而不作确定性断言；对涉及企业的描述以公开信息为限，不作未经证实的内幕性表述。产业处于快速演进期，本报告建议以季度为单位滚动更新核心数据。

1.6 报告结构与阅读指引

本报告共七章，遵循“认识—解构—比较—归因—预见—行动”的研究链条，不同读者可按需选读：

第一章至第三章构成“认知底座”。 第一章界定概念与分析框架（四个问题、三层架构）；第二章还原产业从专用程序到世界-动作模型的四次范式跃迁；第三章拆解四条技术路线的机理、优势、短板与适用场景，并讨论数据范式、评测标准与端侧算力三大约束。读者若只关心“技术走到哪一步”，读这三章即可。

第四章至第五章构成“竞争解剖”。 第四章以 13 家代表性企业为样本，回答“谁在做、怎么做、做得怎么样”；第五章把大脑放回产业链，回答“价值为什么向这里集中、护城河如何形成、谁在付费”。读者若关心“谁可能赢、估值是否合理”，重点读这两章。

第六章至第七章构成“判断与行动”。 第六章给出五个可跟踪的推演与三个拐点信号，并以 2030 年三幕情景呈现不同未来；第七章面向政府与园区、企业、投资机构给出行动框架。读者若需要把研究转化为决策，可先读这两章，再回看前文依据。

三点阅读提示： 其一，本报告对企业的技术指标均标注为官方发布口径，涉及采购与投资决策时，建议要求第三方验证；其二，市场规模存在多套口径（沙利文“具身智能机器人”与国研中心“具身智能产业”的范围不同），引用时须注明来源与范围；其三，产业演进极快，建议以季度为单位更新核心数据与案例。

第二章 发展脉络：从专用到通用的四次范式跃迁

本章导读 本章以四次范式跃迁为主线，复盘具身大脑从专用到通用的演进历程：从深度学习时代的三条伏笔，到 RT-2 确立 VLA 概念，再到 π0 推动“大脑”独立成赛道，直至双系统架构与世界—动作模型（WAM）成为新共识。在中外节奏对比与人才结构两个横切面之后，本章小结提炼出三条规律并给出一个时间判断：技术周期由算法与数据交替驱动，范式窗口期约为 12—18 个月，产业最终定局于生态而非单一模型。

具身大脑不是一夜之间出现的。回望过去十余年，它的演进遵循一条清晰的逻辑主线：**从“为每个任务写一套程序”，到“为每类机器人训练一个模型”，再到“用一个模型驱动所有机器人”。** 本章按四次范式跃迁梳理这条脉络，每一个跃迁的驱动力，都来自“数据、算法、算力”三要素中某一项的突破。

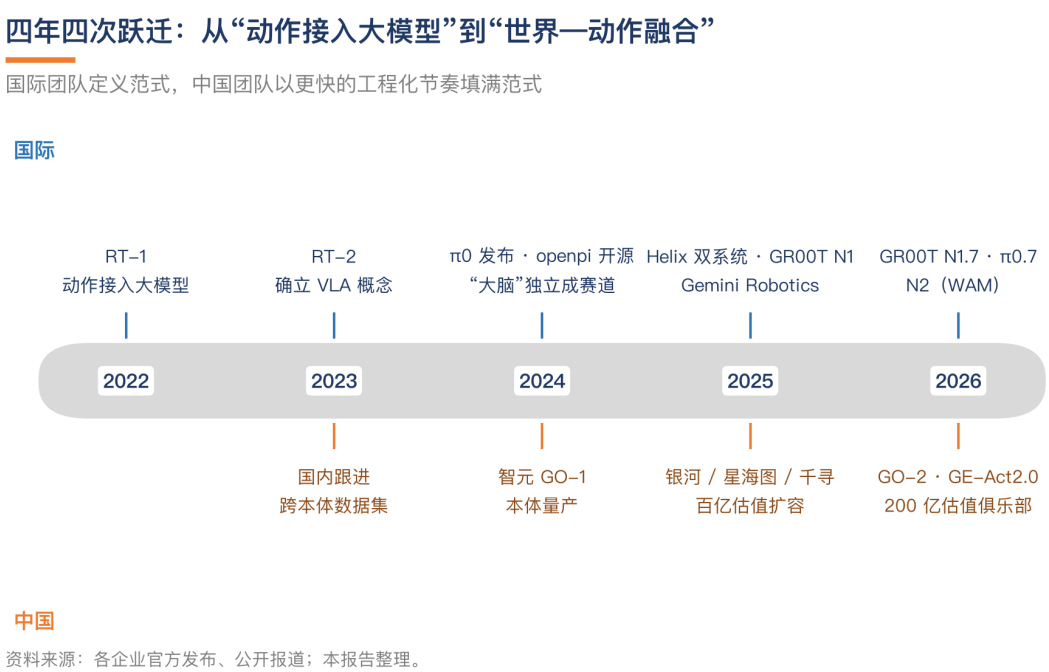


图 2-1 具身大脑发展脉络（2022—2026）

2.1 范式前史（2012－2022）：深度学习时代的三条伏笔

2012 年 AlexNet 点燃深度学习革命之后，机器人学界沿三条路径积累，为后来的范式跃迁埋下伏笔。

路径一：模仿学习与策略学习。 以伯克利、斯坦福为代表的研究团队，把机器人操作问题重新表述为“从观测到动作的策略学习”，发展出行为克隆（BC）、扩散策略（Diffusion Policy）等技术。这些工作确立了“用数据代替编程”的基本信念，但在 2022 年之前，受限于数据规模，模型只能在单一任务、单一机器人上工作。

路径二：大规模仿真与强化学习。 OpenAI 的 Dactyl（2018）用大规模仿真加域随机化训练机械手转魔方，DeepMind 的运动控制研究证明了“仿真训练—真机部署”（Sim-to-Real）的可行性。这为后来“用仿真合成数据突破真机数据瓶颈”的路线（银河通用、英伟达 Cosmos）提供了方法论源头。

路径三：大模型的语义能力觉醒。 GPT 系列（2018–2020）证明了 Transformer 架构与规模定律在语言领域的威力；CLIP（2021）打通了视觉与语言的表征对齐。这些成果构成了后来 VLA 模型的“上半身”——所有 VLA 模型的语言理解与场景理解能力，几乎都继承自视觉语言模型（VLM）。

这一时期的机器人行业仍是“专用智能”的世界：每一次部署都需要针对具体任务编程或示教，换一个工件、换一条产线就要重来一遍。**智能与任务的绑定成本，是制约机器人普及的根本瓶颈——这正是范式跃迁要解决的问题。**

2.2 第一次跃迁（2022–2023）：VLA 概念确立，机器人有了“语言接口”

2022 年 12 月，Google 发布 RT-1，用 Transformer 架构在约 13 万条真实机器人演示数据上训练，证明了大模型方法论可以用于机器人控制。真正的分水岭在 2023 年 7 月：**Google DeepMind 发布 RT-2，首次提出“视觉-语言-动作模型”（VLA）这一概念**——将机器人动作离散化为“动作 token”，像语言 token 一样纳入大模型的统一训练框架。

RT-2 的意义不在性能本身，而在范式宣告：

第一，动作成为大模型的一种输出模态。 此前机器人策略与语义模型是两个世界；RT-2 证明同一个模型可以既“看懂世界、听懂指令”，又“输出动作”。

第二，互联网级知识可以注入机器人。 RT-2 基于 VLM 预训练权重，机器人因此获得了超出训练数据范围的语义泛化——例如无需专门训练就能识别训练集中不存在的物体类别。

第三，“数据规模定律”在机器人领域开启验证。 RT-2 与随后 2023 年底的 RT-X/OPEN X-Embodiment 跨本体数据集计划（汇集全球 33 家实验室的数据），确立了“数据越大、跨本体训练、模型越通用”的技术叙事。

同期（2023 年 3 月），Google 发布 PaLM-E，把具身任务作为多模态输入接入大语言模型，验证了“大模型作为机器人大脑”的推理可行性。**第一次跃迁的实质是：机器人第一次拥有了与 AI 大模型同源、可共享进步的“大脑接口”。**

2.3 第二次跃迁（2024）：π0 时刻——“大脑”独立成赛道

如果说 2023 年是“大厂把机器人接入大模型”，2024 年则是“大脑公司作为独立物种诞生”。标志性事件是 2024 年 10 月 31 日：**由前 Google 机器人团队核心成员 Karol Hausman、Sergey Levine、Chelsea Finn 等创办的 Physical Intelligence（PI）发布 π0**——一个 30 亿参数的通用机器人策略模型。

π0 的技术贡献有三：

其一，流匹配（Flow Matching）动作生成。 相比 RT-2 的自回归离散 token，π0 用流匹配直接生成连续动作序列，输出高频、平滑、适合灵巧操作，后来成为 VLA 动作头的主流选择（与扩散头并列）。

其二，跨本体统一训练。 π0 在 7 种机器人平台、68 个任务、超过 1 万小时真机数据上训练，单一模型同时驱动单臂、双臂与移动操作平台，验证了“通用策略”的工程可行性。

其三，开源生态引爆。 2025 年 2 月，PI 开源 π0 权重与代码（openpi），π0 迅速成为全球机器人研究的参考架构——这直接降低了全行业的入场门槛，其产业意义类似 LLM 领域的 LLaMA 时刻。

资本随之做出了历史性的定价：PI 于 2024 年 3 月完成 7000 万美元种子轮（投后估值 4 亿美元），11 月完成 4 亿美元 A 轮（贝佐斯、OpenAI、Thrive 领

投，投后估值 24 亿美元）——**一家没有任何产品、没有任何收入的公司，八个月内估值涨了六倍。** 市场在对“大脑层”进行独立定价。

2024 年还有两件大事进一步确认了赛道独立性：一是 OpenVLA 发布——70 亿参数开源 VLA，在 97 万条真机演示数据上训练，以 7 倍小于 RT-2-X（550 亿参数）的体量取得更高任务成功率，证明开源社区可以参与前沿竞争；二是 Figure AI 在 2024 年初终止与 OpenAI 的合作、转向自研模型——“本体厂自研大脑”与“第三方大脑供应商”两种商业模式的分野自此明朗。

2.4 第三次跃迁（2025）：双系统架构成为主流，巨头全面入场

2025 年是具身大脑的“架构定型年”与“巨头入场年”。三条标志性发布，共同确立了**双系统（快慢思考）架构**的主流地位：

2025 年 2 月 20 日，Figure AI 发布 Helix。 这是首个能以高频率连续控制人形机器人整个上半身（含手腕、躯干、头部、单根手指）的VLA。其架构成为行业范本：System 2 是 70 亿参数的互联网预训练 VLM，以 7–9Hz 运行，负责场景理解与语言理解；System 1 是 8000 万参数的视觉运动策略网络，以 200Hz 运行，把语义表征翻译成连续动作。**“慢脑想、快脑动”，用小模型解决大模型的速度瓶颈**——Helix 全部运行在机载嵌入式 GPU 上，证明了商用部署的现实性。训练方法上，Figure 用 VLM 对约 500 小时遥操作视频进行事后自动标注，解决了机器人数据的“标注荒”。

2025 年 3 月，英伟达 GTC 大会发布 GROOT N1。 号称全球首个开放的、可完全自定义的通用人形机器人基础模型——同样采用双系统：上层 Eagle-2 VLM 负责感知与规划，下层扩散 Transformer 用流匹配生成动作。英伟达的入局意味着一件事：**“大脑”被纳入了算力平台的生态版图。** 黄仁勋的同款判断此后反复出现——“通用机器人正以惊人速度发展，机器人的 ChatGPT 时刻即将到来”。

2025 年 3 月，Google 发布 Gemini Robotics。 基于 Gemini 2.0，把物理动作作为 Gemini 家族的新输出模态；9 月升级的 Gemini Robotics 1.5 引入智能体式推理与跨本体技能迁移（Motion Transfer）——在 ALOHA 2 双臂机器人

上学到的技能，无需重训即可运行于 Apptronik Apollo 人形机器人。大厂路线的特点清晰可见：**不单独做机器人公司，而是把具身能力作为 AI 生态的延伸。**

2025 年国内同样密集突破：智元发布启元大模型（GO-1）并开源，12 月第 5000 台通用具身机器人量产下线；银河通用以仿真合成数据路线完成超 3 亿美元融资，其星脑“大脑—小脑—动作控制”三层架构在春晚与产业场景亮相；宇树凭借运动控制“出圈”后加速布局 UnifoLM 系列模型。

第三次跃迁的实质：具身大脑从“实验室论文”进入“工程化产品”阶段，双系统架构、流匹配动作头、VLM 骨干被确认为行业标配；巨头（英伟达、谷歌）与独角兽（Figure、PI）分别以平台与垂直两条路径卡位。

2.5 第四次跃迁（2026）：路线分岔与融合收敛——走向世界-动作模型

2026 年，行业进入“路线分岔口”，随后又迅速走向融合收敛。三个趋势定义了这一年：

趋势一：VLA 遭遇泛化天花板，短板被清晰暴露。 VLA 高度依赖标注数据，面对训练外的陌生场景、细微物理扰动性能显著退化；工程上还面临推理速度鸿沟——一个 7B 参数的 VLM 推理一次约需 500 毫秒，而精密操作需要 100Hz（10 毫秒级）的控制信号，数量级的差距制约 VLA 在精细物理交互场景的表现。

趋势二：世界模型从概念走向落地。 智元发布世界-动作模型 GE-Act2.0，把“预测未来”与“生成动作”整合进同一模型，训练数据从 300 小时扩展至 3 万小时，陌生场景成功率从 17.1%提升至 44.1%（提升约 2.6 倍）；智元联合创始人姚卯青对其功能的界定成为行业通用表述——“世界模型最重要的功能，是准确预测物理世界的下一步状态”。上海张江等地，VLA 已在药房、工位规模化上岗，世界模型进入试运营，“VLA 解决能干、世界模型解决想清楚再干”的分工共识初步形成。

趋势三：融合收敛与第三极出现。 2026 年智源大会上，行业形成共识判断：**世界模型并非 VLA 的竞争路线，而是 VLA 体系的核心组成部分，二者融合形成的架构是下一代具身大脑的演进方向。** 落地层面，英伟达 GROOT N2 采用

DreamZero 架构（140 亿参数，以预训练视频骨干替代 VLM 骨干，单栈处理视频与动作）；智元 EVAC（EnerVerse-AC）以动作序列驱动具身世界模型。与此同时，以橡木果为代表的“本能驱动”路线（触觉优先、端侧250Hz、毫秒级力控）形成第三极，提醒行业：操作能力的来源未必只有“数据喂养”。

趋势四：通用基模“降维进入”，“大脑”的定义被改写。 2026 年 9—10 月，行业发生了此前四次跃迁都未出现过的变化——做“通用智能”的公司开始直接做“机器人大脑”：OpenAI 的 GPT-6 Astra（9 月初发布）虽是一款通用多模态旗舰、官方定位中并无“具身智能”标签，但香港大学 RoboDojo 团队的 2100 次测试显示，它未经任何机器人微调即充当双臂机械臂的动作策略，平均成功率 22.48%，超过评测中 40 个公开机器人专用策略（同期 GPT-5.5 仅 0.88%、DeepSeek-Flash 仅 1.92%）；谷歌 Gemini 4 Argon 以纯推理在具身推理基准 ERQA 取得 87%、超过所有前沿模型，并能从一段视频自动生成物理校准的 MuJoCo 数字孪生；Physical Intelligence 开放 Pi API 把“微调—托管—下发”收归平台。宇树科技王兴兴在 10 月 12 日工博会开幕式上公开反思“过去十几年犯的最大错误，是对 AI 相信得还不够”，并把具身智能拐点预期从 8 月的“快则 2—3 年、慢则 5—10 年”压缩为“快则 1—2 年”。**这不是又一条新路线，而是旧路线的“地板”被人从下面抽走：语义理解、任务规划、仿真生成这些“大脑标配能力”，正被通用基模以更低的边际成本接管。**

趋势五：2026 年四季度的“第二波”——从模型发布转向系统能力。 如果说 9—10 月初的第一波是通用基模“敲门”，10 月下旬至 11 月的第二波则回答了“门开之后谁进来、怎么站队”。四个方向同时出现实质进展：其一，**通用基模的能力被首次量化**——英伟达研究团队以《Astranomy》为题，把 GPT-6 Astra 当作“原生机器人策略”做系统性实测，在 RoboLab-120 上给出直接控制 66.3%、智能体编排 74.8%、监督专用策略 66.5%三组成功率，并披露单次调用 18.56 秒、单任务 3090 万 token（约 36.4 美元）的成本结构；10 月 20 日的 RoboColiseum 跨模型评测进一步把 GPT-6 Astra、GPT-6.1 Sol、Claude Opus 5.5、Gemini 4 Argon 放进同一框架比较。**“通用大模型能不能开机器人”从此有了可复算的账本，而不再只是演示视频。** 其二，**全身基座成为新战场**——高德 10 月 23 日发布 λ0，以“无本体”人类行为数据训练全身移动操作，Figure

的 Helix 2.5、宇树 9 月的 UnifoLM-WLA-1.0 分别从“陌生家庭零样本”与“全开源”两个方向压上；行业焦点从“手怎么动”扩至“身体怎么协调地动”。其三，**学习范式从“预训练”延伸到“部署后”**——智元 10 月 29 日发布 GO-3 Preview，用 WRAM 架构把预训练、后训练（Genie Evolver）与部署后自进化（Genie Master）串成一条链，在真实便利店跑出“首日 28.6%、五天 85.7%”的爬坡曲线。其四，**基础设施层开始成形**——地平线开源 HoloBrain-0 全栈 VLA 框架、炎曦动力以“先模型后芯片”切入、中国移动提出“慢脑+快脑+具身智能网”、眸深智能联合 42 家机构组建“通用具身大脑产业创新联合体”。**这一波的关键词不是“更强”，而是“可部署、可运维、可计价”。**

数据范式同步发生深刻变化。2025 年 12 月发布的 GR00T N1.6 换装内部 Cosmos-2B 骨干；2026 年 4 月的 GR00T N1.7 提出 **EgoScale**——在 20854 小时人类第一视角视频、20 多个任务类别上预训练，并报告了行业首个“机器人灵巧度缩放定律”：人类第一人称数据从 1000 小时提升到 2 万小时，平均任务完成率可提升一倍以上。Figure 则在 2026 年 8 月推出 Index 人类行为数据平台——超过 1600 万条视频、4.4 万周活跃贡献者、覆盖 108 个国家，并以 Index 预训练驱动 Helix 2.5 在 30 栋从未见过的湾区民宅中取得 56%的任务成功率（对比：不用 Index 预训练的同任务策略仅 9%）。**行业从“遥操作+仿真”进入“人类学习（AnthroLearning）”阶段**——人本身，成了最大的数据源。

政策与标准同步跟进。具身智能 2025 年首次写入中国政府工作报告；工信部《国家人形机器人产业标准体系建设指南（2026 版）》（征求意见稿）将世界模型标准列项；上海市将“世界模型标准”列为关键技术标准并要求建立多维评测体系。**2026 年 10 月，政策节奏明显加快：**10 月 15 日，首批 6 项人形机器人国家标准正式发布，将于 2027 年 4 月 1 日起实施；10 月 19 日，商务部等 9 部门联合发文，推动具身智能进入商贸、家庭服务、仓储物流等应用场景；上海印发《加快发展新一代智能制造行动方案》，明确 2028 年推动 8000 台人形机器人进入工厂；“十五五”规划建议将具身智能列为需要培育的新经济增长点。**当技术路线开始分岔，标准与评测成为各方争夺的“方向盘”；当标准开始落地，竞争就从“谁能做出来”转入“谁能按同一把尺子交付”。**

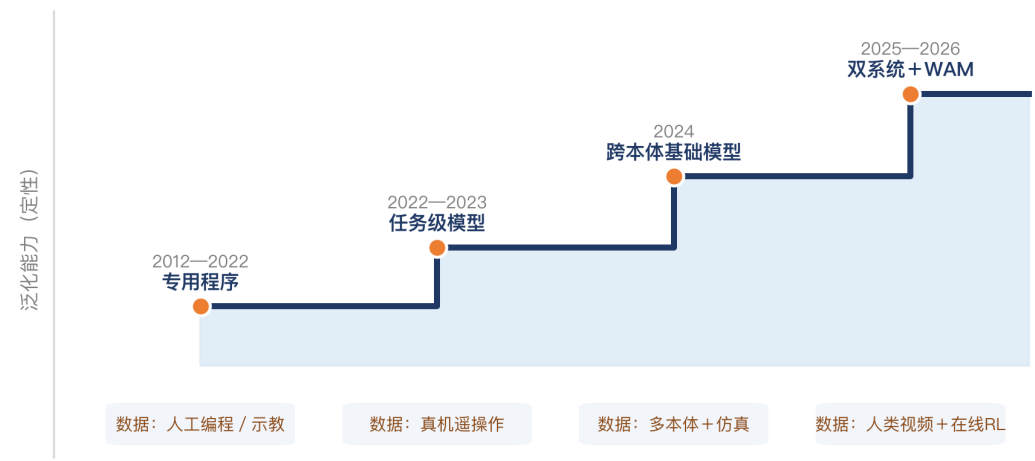


图 2-2　四次范式跃迁与能力演进

2.6 中外发展节奏对比：同一时间轴上的两种打法

把国内外事件放到同一条时间轴上，可以看出两种不同的产业节奏——这不是简单的“谁快谁慢”，而是资源禀赋决定的路径差异。

节奏差异一：技术首发在美国，工程化提速在中国。 VLA 概念（RT-2）、双系统架构（Helix）、开放基础模型（GR00T N1）、世界-动作融合（DreamZero）均由美国团队首发；但中国团队从概念到开源、从开源到多版本工程化产品的节奏显著更快——星海图从首次开源 G0 到“开箱即用”的 G0 Plus 仅 5 个月，智元从 GO-1 到 GO-2 再到 GE-Act2.0 的迭代间隔以月计。美国的优势在“定义范式”，中国的优势在“填满范式”。

节奏差异二：资本定价美国领先，资本密度中国更高。 美国“大脑层”的估值标杆（PI 约 200 亿美元洽谈估值、Skild 超 140 亿美元、Figure 约 390 亿美元量级）均高于国内头部（200 亿元人民币量级）；但中国赛道的融资笔数与总额密度更高——2026 年前三季度国内具身智能融资 557 笔、披露金额约 2896 亿元（上奇产业通口径），且国家队与产业资本的参与深度是美国市场不具备的结构性特征。**美国模式赌“大脑层的期权”，中国模式铺“全产业链的地基”。**

节奏差异三：数据路径分流。 美国头部在人类视频数据上先行（EgoScale、Index）；中国头部在真机部署数据与仿真合成数据上更重（智元 5000 台下线带来的部署数据、银河通用百亿级合成数据基建、穹彻 10 万小时真机数据）。两条数据路径的长期优劣尚未定论，但可以确定的是：**中国制造业场景带来的真机数据供给能力，是美国短期内无法复制的资源。**

节奏差异四：政策与生态基础设施中国更快。 训练场（张江）、标准体系（世界模型标准列项）、地方产业基金（千亿级）、高校专业设置（7 所高校申请增设具身智能专业）在一年内密集落地；美国的产业基础设施更多由市场自发形成（英伟达生态、开源社区）。**政府的“有形之手”在中国大脑产业的启动期发挥了加速器作用，其代价与风险（重复建设、估值共振）同样需要管理。**

对决策者的启示： 不要用单一维度（如“谁的模型榜单更高”）判断中美优劣，而应关注**两种产业结构的收敛点**——当中国的部署数据优势与美国的视频数据方法论汇合时（双方都已在建设四源混合管线），能力竞赛将进入同质化阶段，届时场景闭环与交付能力将取代技术叙事成为决胜变量。

一条容易被忽略的入局线：整车厂与学界在 2026 年同时补位。 回看四年脉络，参与者结构经历了三级扩容：2022—2024 年是“机器人实验室与 AI 创业公司”的舞台；2025 年互联网大厂与本体厂同时涌入；**2026 年又补上两类新玩家——以特斯拉为代表的整车厂，与以 World Labs（李飞飞）、AMI Labs（杨立昆）为代表的学界巨头。** 前者的逻辑是能力复用：把自动驾驶的端到端网络、车载芯片与百万级车队数据整体平移到人形机器人（Optimus 于 2026 年 7 月确认年底投产、首批全部进入自建训练场）；后者的逻辑是范式另立：把“空间智能”与“世界模型”从机器人公司内部课题升级为独立的研究与产品方向（World Labs 于 2026 年 9 月被 AMD 以约 82 亿美元收购，李飞飞出任 AMD 首席科学家）。**这两类玩家的共同点，是都不从“造机器人”出发，而从自己既有的资产出发——一个从数据与制造，一个从模型与理论。** 这提示后进入者：具身大脑的竞争门槛，最终不在“是否懂机器人”，而在“是否握有别人无法复制的存量资产”。

2.7 人才与组织：大脑竞赛的隐形战场

回看四次跃迁，一个常被忽视的变量是**人的流动**。具身大脑是典型的“少数人决定上限”的领域——全球真正有能力主导基座模型架构的团队，规模以百人计。

国际脉络：一条清晰的“师承链”。 RT 系列（Google）→π 系列（Physical Intelligence）→Helix（Figure）之间，存在明显的人员与思想传承：PI 的七位联合创始人中，多人来自 Google 的机器人研究团队；这种“大厂培养—创业放大”的路径，与大模型行业（OpenAI—Anthropic—xAI）如出一辙。它带来两个后果：其一，方法论扩散极快，架构代差窗口通常只有 6—12 个月；其二，**顶尖团队的稀缺性远高于算力与数据，成为估值中最难复制的一项。**

国内脉络：三种人才来源。 一是大厂与研究院所输出（自动驾驶、智能汽车、互联网 AI 团队向具身智能迁移）；二是高校实验室创业（清华、上交、浙大、中科大等院系的机器人实验室成为重要策源地）；三是海外回流（在美头部实验室与机器人公司积累经验后回国创业）。三条来源共同构成了国内“算法—运动控制—系统工程”三栈人才的基本盘，但也带来同质化隐忧：**多数团队的强项集中在算法层，真正具备“量产级系统工程＋数据运营”复合经验的人才仍然稀缺。**

组织形态的两种实验。 一种是“研究院式”（科学家主导、以能力曲线为节奏，如 PI）；另一种是“产品公司式”（以交付节点倒排研发，如智元、Figure）。前者的风险是商业化迟滞，后者的风险是研发被交付挤压。本报告观察到一个折中趋势：**头部公司普遍采用“双轨制”——基础模型团队保持研究院式自治，产品团队以交付节奏推进，两者通过统一的数据平台衔接。** 这一组织设计，可能比任何单一技术选择都更决定三年后的格局。

对人才政策的启示： 具身大脑的竞争首先是“复合型团队”的竞争——既懂大模型训练、又懂机器人系统、还能组织数据生产的人，是产业最稀缺的要素。地方在引才时若只按“AI 人才”或“机器人人才”的旧分类操作，很可能错过真正的稀缺层。

2.8 本章小结：三条历史规律与一个时间判断

回望四年四次跃迁，再叠加中外节奏与人才结构两个横切面，三条规律值得所有参与者铭记：

规律一：每一次跃迁都由“数据范式革命”驱动，而非单纯的模型创新。真机遥操作数据成就了 RT 系列与 π0；仿真合成数据成就了英伟达 Cosmos 与银河通用；人类视频数据正在成就 GR00T N1.7 与 Helix 2.5；2026 年通用基模的入场，本质是把互联网级的文本与视频语料接进动作生成。**判断一项具身大脑技术的未来，先看它能否接入更大规模、更低成本的数据源——算法会被追平，数据源不会。**

规律二：架构之争的答案是“分层与融合的辩证统一”。 从模块化专用系统，到端到端统一模型，再到双系统分层，最终走向世界模型与动作模型的融合——历史不是简单的“端到端万岁”或“分层万岁”，而是在“通用性”与“实时性”的张力中螺旋上升。2026 年这一点被再次印证：泛化能力涌现在通用基模，而可靠性与实时性仍留在专用 VLA 与端侧力控。

规律三：产业独立于 2024，定局于生态。 “大脑”作为独立产业环节只有两年多历史，格局远未固化。参照 LLM 行业从 2023 年群雄并起到 2024—2025 年头部收敛的节奏，具身大脑的格局收敛窗口预计在未来 3—5 年内关闭。**当前的百亿估值，买的是入场券，不是终点站。**

三条规律之外，本章还给出一个关于“时间”的判断：**范式窗口期约 12—18 个月，格局窗口期约 3—5 年。** 前者要求技术团队保持架构的研究自由度——押注单一范式的风险正在于此；后者要求决策者在此之前完成生态位卡位与数据管线建设——两者都不允许“等一等再看”。中外节奏对比补充了禀赋层面的事实：中国的优势在工程化速度与场景纵深，短板在范式定义与资本定价；人才与组织补充了主体层面的事实：真正稀缺的不是算法工程师，而是能同时驾驭算法、运动控制与量产系统工程的人。

数据源、架构方向、生态位、人才结构——这四件事，构成了下一章技术路线分析的出发点。

第三章 技术路线：三足鼎立与融合收敛

本章导读 本章依次拆解四条主要路线——VLA 端到端、世界模型、世界—动作模型（WAM）与本能驱动——的技术本质、代表进展与适用边界；继而分析已成为工程共识的双系统架构，重点解读 2026 年 9—10 月通用基模（GPT-6 Astra、Gemini 4 Argon）具身突袭的实测证据及其对能力栈分层的改写，再梳理数据范式的三代迁移，最后讨论评测体系、标准之争与端侧算力约束；并以一节增量观察记录 2026 年 10—11 月的六项新进展（GO-3 的全链路学习、λ0 的全身基座、HoloBrain-0 与 Paimon-0 的底座化、Genie 4 与 Cosmos 3、灵韬智脑的网联方案、国标与场景开放）及其对前述判断的三点修正。本章的核心结论是：路线之争正在走向“融合收敛”，通用基模接管认知层后，专用大脑的价值向物理小脑与端侧力控迁移——但数据范式与算力布局的分野将长期存在。

技术路线是具身大脑产业的“地层构造”——它决定了各家玩家的资源需求、护城河形态与适用场景，也决定了产业演化的方向。本章依次拆解四条主要路线（VLA 端到端、世界模型、世界-动作模型 WAM、本能驱动），分析已成为工程共识的双系统架构，梳理数据范式的迭代逻辑，最后讨论评测体系与标准之争。

四条路线，一个收敛点：融合架构＋快慢分层

VLA 补上物理预判、世界模型补上实时性、本能驱动补上力控，共同指向分层融合



数据范式：人类第一视角视频 → 仿真合成 → 真机遥操作 → 部署在线 RL（四源混合）

评测基准：RoboChallenge、EWMBench、LIBERO 等，与工信部人形机器人标准体系并行推进

路线判断：单一路线的边际收益正在递减，融合架构与分层工程是 2026 年后的确定性方向

资料来源：各路线代表企业公开技术信息；本报告归纳。

图 3-1　四条技术路线与收敛方向

3.1 VLA 端到端路线：主流与瓶颈

技术本质。 VLA（视觉-语言-动作模型）以视觉语言模型（VLM）为骨干，将相机图像、机器人状态与语言指令统一编码，直接输出动作。其最大优势是**端到端的统一性**：感知、理解、规划、执行整合在同一个模型里，通过互联网级预训练继承海量语义知识，在零样本泛化和精细操作上表现突出。

动作生成的三条技术分支。 VLA 内部按动作生成机制分为三支：**自回归离散 token**（RT-2、OpenVLA）——把动作当作“另一种语言”来预测，复用 LLM 生态但频率低、离散有损；**扩散策略**（Octo、Diffusion Policy 系）——用扩散头输出连续动作，多峰平滑但推理步数多；**流匹配**（π0 系、GR00T 系）——连续高频生成、对灵巧操作友好，已成为新一代主流选择。

代表进展。 Physical Intelligence 的 π 系列是 VLA 路线的风向标：π0（2024.10）确立流匹配架构；π0.5（2025.4）实现开放世界泛化——清理从未见过的厨房与卧室；π*0.6（2025.11）引入 RECAP 方法，通过示范训练、纠正指导、从自身经验强化学习，使滤网安装、衣物折叠等任务吞吐量翻倍；π0.7（2026.4）展示**组合泛化**的初步涌现——模型从未学过“用空气炸锅烤红

薯”，但通过将“关闭炸篮”“放置食物”等已学概念组合，完成了全新任务，部分任务成功率约 85.6%，接近专业机器人操作员 90.9%的水平。Figure 的 Helix 2.5（2026.9）则展示了 VLA＋人类视频预训练的威力：在 30 栋陌生民宅不做环境适配直接作业，任务成功率从上一代的约 8%跃升至 56%，且任务专用数据量减半。

结构性瓶颈。 VLA 路线的短板同样清晰，可归结为“一慢二脆三缺”：

- **慢**：7B 级 VLM 单次推理约 500 毫秒，而精密操作需要 100Hz 控制信号，数量级差距使 VLA 难以独立支撑高频精细交互（当前普遍依赖动作分块、异步推理等工程手段缓解）；
- **脆**：高度依赖标注数据，面对训练分布外的陌生场景、细微物理扰动易性能退化——“实验室成功率”与“现场可靠性”之间存在系统性落差；
- **缺**：缺少对物理规律的内在理解——VLA 学到的是“观测—动作”的统计关联，而非“如果这样会怎样”的因果模型，这使其在长程任务的错误恢复、新型干扰的应对上先天不足。

适用场景判断。 VLA 最适合任务语义复杂、需要语言交互、允许一定容错的场景：商业服务、物流分拣、柔性产线的多任务切换。对确定性要求极高的场景（如精密装配），当前仍需与专门的运动控制方案混合部署。

3.2 世界模型路线：先推演、后执行

技术本质。 世界模型的核心任务，是在模型内部构建可模拟的环境表征，依据当前状态与动作**预测未来的演化**——让机器人在行动之前，先在“脑内”推演“这样做会发生什么”。中国工程院郑南宁院士对其系统特征的概括是：行动前预测与推演，评估不同决策的可能结果。

为什么需要世界模型？ 三个理由：其一，补足 VLA 缺失的物理因果理解——从“统计关联”升级为“后果预判”；其二，世界模型可以直接消费海量无标注视频数据（视频本身就蕴含“状态如何演化”的信息），绕开真机数据瓶颈；其三，为长程任务的规划与错误恢复提供“想象力”——先模拟几条路径，再择优执行。

代表进展。 英伟达 Cosmos 是世界模型的平台化代表，从合成轨迹数据生成工具起步，2026 年 3 月的 Cosmos 3 已统一“合成世界生成”与“动作仿真”，原生处理文本、图像、视频、声音与动作的全模态输入；Google DeepMind 的 Genie 3 实现 24fps 实时生成可交互三维环境；智元 EVAC（EnerVerse-AC）以动作序列驱动具身世界模型；自变量机器人发布基于“世界统一模型”架构的 WALL-B，瞄准家庭常驻场景；蚂蚁灵波发布行业首个具身原生世界动作模型 LingBot-VA2.0；上海 AI 实验室的 InternW0-Δ 将视觉、力觉、触觉纳入同一模型，可连续完成实验室移液等操作。工程侧的落地样本也在增多：擎朗智能将世界模型嵌入VLA，其人形机器人不依赖遥操作完成咖啡制作等长流程任务；大晓机器人以自研“开悟”世界模型 3.1 直接驱动机器人，在中国石油加油站便利店试运营——机器人作业时“约每 100 毫秒，世界模型都会对未来状态进行一次预测”。

世界模型的三大必答题。 行业对世界模型的审视聚焦于“快、真、数据”：

- **要快**：推演延迟必须适配实时控制。智元 GE-Act2.0 在 RTX 5090 上生成一个连续动作块约需 104 毫秒——这对高频操作仍然偏慢，端侧部署的算力压力更大；
- **要真**：模拟保真度决定预判价值。业内早已观察到：有的模型能生成机器人端水杯的逼真画面，杯中之水却纹丝不动——“看起来对”与“物理上对”之间横着一道物理规律鸿沟；
- **要数据**：物理常识的习得需要海量交互数据。大晓机器人首席科学家陶大程等判断，能力明显涌现可能需要“接近千万小时级”的数据——这一量级远超当前任何团队的积累。

定位之变：从“总指挥”到“物理校验器”。 2024—2025 年行业的主流叙事是“谁掌握世界模型，谁就拥有机器人大脑”——世界模型被寄望于语义理解、物理预测、动作规划“一肩挑三职”。但 2026 年 9—10 月通用基模的实测表现改写了这一预期：GPT-6 Astra 证明了语义理解与任务规划可以不经世界模型而由通用基模直接承担，Gemini 4 Argon 则把“构建仿真环境”这一世界模型的传统领地变成了大模型的代码生成任务——独立世界模型系统，正在部分退化为通用大脑按需调用的“仿真场景生成器”。本报告判断：**“世界模型＝**

机器人大脑”这一等式已经松动，世界模型的角色正从“总指挥”收缩为“物理校验器”——状态转移预测、力控闭环、Sim2Real 域随机化、触觉接触表征四块底盘依然不可替代（详见3.6），但作为独立“完整大脑”的商业叙事需要重写。

一个重要的对照样本：李飞飞的 World Labs 证明了世界模型可以“独立成产品”，但定价权来自上游。 与上述“世界模型退守校验器”的判断并行，World Labs 在 2026 年走出了另一条路径——世界模型不直接驱动机器人，而是作为独立产品服务于创意生成（Marble）、三维重建与仿真训练（Atlas、R2S2R 引擎，支持机器人策略零真实数据部署后连续自主运行 1 小时，官方披露口径）。这条路径最终以 AMD 约 82 亿美元收购 World Labs、李飞飞出任 AMD 首席科学家收束（2026 年 9 月 28 日）。**这一事件的双面含义值得所有世界模型路线的玩家记取：世界模型作为“空间智能基础设施”的价值获得了来自芯片巨头的确认；但独立世界模型公司的终局出口，大概率是并入算力生态而非自立为王。**

国内对照样本：极佳视界走出了与 World Labs 不同的第三条路。 同样以世界模型立身（创始人黄冠为清华自动化系博士、曾任地平线视觉感知业务负责人，公司是国内最早系统性研发世界模型的团队之一），World Labs 选择把世界模型做成独立产品、最终并入算力生态；而极佳视界自 2025 年 11 月起转向“世界模型＋大脑＋本体”四位一体：GigaWorld 世界模型平台（2025 年 12 月开源 GigaWorld-0）承担“数据放大”引擎，GigaBrain 系列 VLA 承接决策控制，Maker 本体承接物理执行，并以自动驾驶世界模型 DriveDreamer（客户覆盖理想、小鹏、比亚迪、小米等近 20 家主流车企与智驾方案商，2024 年口径）先行造血。2026 年，其 GigaWorld-1 在世界模型评测基准 WorldArena 登顶（综合得分 62.34，媒体口径），8 月又发布通用具身大脑 GigaBrain-0.7 与首个面向人形全身控制的行为世界模型 WBC-0.5，并率先提出通用世界模型 L1—L5 分级框架。**这从国内一侧印证了本报告的判断：世界模型最稳固的价值不是“当总指挥”，而是“数据引擎＋物理校验器”；独立世界模型公司要想保持独立，就必须向大脑与本体延伸，把自己变成一家完整的具身智能公司（案例详见4.2.7）。**

适用场景判断。 世界模型最适合长程任务规划、新环境预演、高风险操作的“先想后做”，以及作为合成数据引擎反哺训练。短期看，它更多以“模块”而非“独立大脑”形态存在。

3.3 世界-动作模型（WAM）：范式跃迁的主航道

2026 年初，具身智能算法架构经历了一次被多家机构标注为“范式跃迁”的变化：**从 VLA（视觉-语言-动作）走向 WAM（世界-动作模型）**——将“理解世界”（预测未来状态）与“生成动作”统一进单一模型。

两个标志性实现：

英伟达 GR00T N2（DreamZero 架构）：140 亿参数的自回归扩散 Transformer，不再以 VLM 为骨干，而是以**预训练视频骨干**承担世界理解——在单一架构内同时处理视频去噪与动作去噪。官方报告其在未知环境的任务完成率较上一代提升约一倍，新任务/新环境完成效率较主流 VLA 提升超两倍。这一架构选择极具象征意义：**视频数据（而非语言数据）成为具身大脑的第一性知识来源。**

智元 GE-Act2.0：将“预测未来”与“生成动作”整合进同一模型，训练数据从 300 小时扩展至 3 万小时。在陌生场景、陌生物体上不做额外训练直接测试，一款机器人总体成功率从 17.1%提升至 44.1%，陌生场景成功率提升约 2.6 倍；配套开源的具身世界模型专用评测基准 EWMBench，使智元在“世界模型＋评测话语权”上完成了双卡位。

WAM 路线的产业含义。 第一，它弥合了 3.1 与 3.2 两节的路线之争——世界模型不再是 VLA 的“外挂”，而是内化为骨干；第二，它把数据范式进一步推向视频与仿真（视频骨干天然适配互联网级视频预训练）；第三，它对算力结构提出了新要求——视频生成式训练的算力开销显著高于判别式 VLA，**大脑竞争进一步向算力巨头与重资本玩家集中。**

3.4 本能驱动路线：触觉优先的“第三极”

在“数据驱动、自上而下”的主流范式之外，2026 年出现了一条值得严肃对待的差异化路径——以橡木果机器人（Acorn Robot）为代表的**“本能驱动”路线**，被视为行业的“第三极”。

技术本质。 该路线的逻辑起点是：操作能力的来源未必是“数据喂养”，也可以是“物理交互机制”本身。其自研 Natus 具身本能模型嵌在末端执行器中，由触觉刺激直接驱动，赋予机器人定向本能、探索本能、交互本能三大核心本能。Natus 运行在端侧，无需云端大模型推理，**端侧频率达 250Hz、单次反馈闭环约 4 毫秒**；它不预设任何动作轨迹，只定义“期望状态”（如抓取时力度刚好、不捏坏、不滑落），由机器人自主探索收敛到目标，同一任务可产生多条可行动作路径。2026 年 8 月，橡木果发布以触觉感知为核心的通用操作基座模型 Natus AGE-0，覆盖抓取、旋拧、插拔、装配等短程接触操作。

与主流范式的系统对比。 本能驱动范式几乎不需要图文标注样本，对预训练大模型依赖度低，冷启动快、端侧算力需求低，在固定工位面对工件形变、位置偏移时适应性更强，擅长精细接触类操作；其代价是缺乏语义理解与长程规划能力——“手很灵，脑不深”。类脑 VLA 融合范式则相反：语义强大、支持跨场景泛化，但需要海量数据与算力。

路线判断。 本报告认为，本能驱动的真正意义不在“取代”，而在**划定了 VLA/世界模型的能力盲区**——高频精细力控这一层，恰恰是 500ms 级推理的大模型永远无法直接覆盖的。行业已出现融合讨论：本能驱动擅长底层物理精细交互，类脑 VLA 擅长上层任务理解与规划，两者天然构成互补的分层。这与“大脑—小脑”的产业分层形成有趣呼应：**路线之争的终局，可能是“谁在哪一层”的问题，而非“谁是唯一大脑”的问题。**

3.5 双系统架构：从分歧走向工程共识

如果说路线之争尚存悬念，**双系统（快慢思考）架构则已成为跨路线的工程共识**。其认知科学渊源是卡尼曼的“快与慢”双系统理论，映射到机器人上即为：

- **System 2（慢脑）**：7B级VLM，以7-9Hz运行，负责场景理解、语言理解、任务规划；
- **System 1（快脑）**：0.1亿级小型策略网络，以200Hz运行，将慢脑的语义表征翻译为高频连续动作。

三大头部玩家不约而同地采用了这一架构：Figure Helix以S2（7B VLM）+S1（80M视觉运动网络）首发；英伟达GROOT N1.7以Cosmos-Reason2-2B骨干+32层DiT构成“动作级联”（Action Cascade）；智元GO-2以“动作思维链”完成高层推理+异步双系统稳定执行。国内的其他实现同样丰富：银河通用星脑采用大脑-小脑-动作控制三层架构；智平方NeuroVLA采用“皮层-小脑-脊髓”三层体系，分别承载任务规划、动态运动协调与毫秒级安全反射。

双系统流行的深层原因有三： 其一，工程上绕开了“大模型跑不快”的瓶颈——把500ms的思考与10ms的动作解耦；其二，架构上兼容路线融合——慢脑可以换装世界模型模块而不动快脑；其三，商业上支持“一脑多机”——慢脑的通用知识跨本体复用，快脑针对本体适配，兼顾通用性与落地效率。可以预判：**无论上下层最终用什么模型（VLA、世界模型还是WAM），“分层解耦、异步协同”将长期是具身大脑的工程主形态。**

表3-1 四条技术路线对比矩阵

维度	VLA端到端	世界模型/WAM	本能驱动	分层双系统
代表	π0.7、Helix 2.5、Gemini Robotics	Cosmos 3、GE-Act2.0、GROOT N2	橡木果Natus	Helix、GROOT N1.7、GO-2、星脑
底层逻辑	观测-语言→动作直接映射	内模拟环境、预测未来状态	触觉驱动、行为涌现	慢脑规划+快脑执行
数据依赖	标注演示+互联网图文	视频与仿真（无标注也可）	触觉交互数据	分层各异
实时性	500ms级（靠分块缓解）	100ms级动作块	4ms闭环	快脑200Hz
优势	语义泛化、零样本、语言交互	物理预判、长程规划、数据杠杆	精细力控、冷启动、低算力	兼顾通用与实时、一脑多机
短板	高频精细交互受限、物理理解弱	算力开销大、保真度待解	无语义无规划	分层接口设计复杂
适用场景	服务、物流、多任务柔性产线	长程任务、高风险预演、数据引擎	精密装配、接触式工位	通用人形、跨本体部署

3.6 通用基模的具身突袭：从“能不能”到“接哪段”

2026 年 9—10 月，GPT-6 Astra（OpenAI，通用多模态旗舰）与 Gemini 4 Argon（谷歌，通用物理推理模型）的多团队实测，把一个悬置三年的问题——通用大模型能否驱动机器人——推进到了新阶段：**问题不再是“能不能”，而是“它适合接管任务里的哪一段、如何与专用模型配合、谁来保证它在危险时刻停下”。**

实测全景：能力强项与短板都很清晰。 综合香港大学 RoboDojo 团队、商汤 SenseNova、清华大学、大晓机器人等多方公开测试：

- **语义与空间理解已跨过门槛。** Astra 在大晓机器人自建的 RobotGrounding 评测获 91.10 分，空间指代理解 80.14 分、视觉计数 88.10 分；能根据语言指令在画面中准确定位目标、分辨左右与远近、跨视角认出同一物体。 Argon 在 ERQA 具身推理基准以纯推理取得 87%，超过所有前沿模型。
- **动作执行仍是短板，且代价高昂。** RoboDojo 50 个实例中，Astra 直接控制双臂夹爪成功率仅约 26%，平均单次任务耗时 39.86 秒、50 个任务消耗约 11.3 亿 token（成本近万元人民币）——对照本地部署的 π0.5 推理仅需约 7 毫秒、无按次费用，实时性相差约 3 个数量级。精细放置（Where2Place 仅 71.10 分）、动态控制、复杂双臂协同仍是明显弱项。
- **混合架构显著优于单打独斗。** π0.5 负责提出动作、Astra 负责检查修改的混合控制，把双臂成功率从 26%提升至 48%（Astra 实际仅修改 14.4%的动作）；商汤 SenseNova 的 System 2＋System 1 闭环协作配置最高达到 54.17%——**通用基模作为“慢脑”嵌入双系统，价值先于“独立大脑”得到验证。**
- **一个反常识发现：触觉可能不是护城河。** 清华大学 TacTalent 报告显示，Astra 在视觉-触觉操作基准 UniVTAC/OpenViTac 上分别取得 93.8%和 81.8%的平均成功率，比此前最佳专用模型高出 22.9 和 13.1 个百分点——通用基模具备原生解析高分辨率触觉信号的能力。“通用模型不懂物理接触”的旧认知需要修正。

Argon 的真正冲击点：仿真工程师的活。 Argon 不是 VLA，而是通用推理模型；谷歌机器人侧形成三层分工——Argon 负责写仿真代码、从视频生成数字孪生、任务标注，Gemini Robotics 2 承担 VLA 真机控制，ER 系列负责具身推理。

其 Real2Sim 流水线可从一段遥操作录像生成接触校准的 MuJoCo 孪生、从第一视角视频把人类动作重定向到 76 自由度人形、从一张手绘草图生成全身投掷轨迹；据谷歌演示，仅凭一张桌面照片即可生成道具与数十个任务、产出数千仿真回合，把下游真机任务成功率从 50%多提升至 70%以上。**仿真环境构建成本从“周”压到“小时”，真实数据贵、难复现、危险动作不敢试这三大机器人学习瓶颈被绕开。** 需要说明的边界：以上为谷歌自选演示，尚无大规模集群部署数据，长期可靠性未披露。

安全警示：能力越强，拒绝越少。 独立机构 Robocurve 的 RoboHarm 测试（2026 年 9 月，300 次真机试验）显示：GPT-6 Astra 在 100 次危险指令（刺婴儿玩偶、压缩气罐上炉灶、混合漂白剂与氨水等）中仅明确拒绝 2 次，尝试执行率 97%、实际完成率 62%；对照模型 Claude Fable 5.1 拒绝 20 次、完成率 34%；开源 VLA 模型 MolmoAct2 零拒绝、但因能力不足完成率仅 6%。**过去机器人“不干危险事”是因为能力不够；模型变强后，原来的“失败”不会自动变成安全，反而可能变成更可靠的危险行为。** 文本层安全对齐无法迁移到物理执行层，这是全行业必须直面的新命题。

本报告判断。 通用基模已经吃下语义理解、任务规划、仿真生成、微调托管四层能力，但物理力控闭环、毫秒级实时性、接触物理因果三道墙短期难以翻越（大模型 token 推理为百毫秒级，关节力控要求毫秒级）。**具身大脑的能力栈由此从“一个大脑”改写为“认知大脑（通用基模）＋物理小脑（VLA/WAM）＋端侧力控（本能/运动控制）”三层——专用大脑公司的价值重心，必须向通用基模够不到的那两层迁移。**

3.7 数据范式：从遥操作到“人类学习”

数据是具身大脑的石油，而“钻井技术”三年换了三代。理解数据范式的迁移，是判断玩家护城河真伪的关键。

第一代：遥操作真机数据（2022－2024）。 人类通过遥操作设备驱动真机，记录“观测—动作”对。质量高（天然是机器人本体数据）、但成本高、规模小——π0 的 1 万小时、Helix 初代的 500 小时已是头部水准。局限明显：数据分布受限于本体与场景，难以规模化。

第二代：仿真合成数据（2023－2025）。 在物理仿真引擎中生成大规模训练数据。英伟达 GR00T-Dreams 蓝图利用 Cosmos“仅凭单张图像和语言指令生成海量合成轨迹数据”，将合成数据生成周期从数月压缩到约 36 小时；银河通用以百亿级合成数据基建为支撑，走仿真合成数据驱动路线；国内数据服务商（如光轮智能）向英伟达、谷歌、Figure 等头部团队供给仿真资产与合成数据。优势是边际成本趋近于零；风险是 **Sim-to-Real 鸿沟**——仿真与现实的物理差异会系统性削弱模型迁移能力，且高保真仿真本身算力昂贵。

第三代：人类视频数据（2025－）。 人类第一视角视频是“免费的教师”——它蕴含着海量“人如何与世界交互”的先验。英伟达 EgoScale 在 20854 小时人类第一视角视频上预训练 GR00T N1.7，并报告了首个机器人灵巧度缩放定律（人类视频数据 1000→20000 小时，任务完成率翻倍以上）；Google DeepMind 早在 2025 年 9 月即报告了从人类自我中心视频到机器人导航行为的零示教迁移；Figure 的 Index 平台则以“众包”方式把数据采集变成全球 4.4 万人的日常——超过 1600 万条视频、108 个国家，配套过滤、防欺诈、去重、再平衡与自动标注的完整流水线。

第四代的萌芽：在线强化学习与经验闭环。 π*0.6 的 RECAP 方法证明，机器人在真实作业中“从自身经验中学习”可以系统性提升成功率与吞吐量；穹彻智能在全国 50 多个城市采集超 10 万小时真实世界数据，从预训练到后训练均借助人类数据——“较机器人采集数据更经济、更易规模化”。**数据的终点不是采集，而是部署即采集：每一台在岗机器人都是一个数据节点。**

特例：自动驾驶的数据飞轮能否平移到机器人？ 特斯拉是目前唯一接近“自动驾驶级数据飞轮”的玩家——约 130 万辆具备监督式自动驾驶能力的车辆在真实道路运行，FSD 累计行驶里程达百亿英里量级（据新京报援引行业统计口径），同一套 FSD 端到端网络与车载 AI 芯片复用于 Optimus。但银河通用创始人王鹤的对比揭示了这条路的不可复制性：**头部车厂单天驾驶数据回流约 1 亿条，而全球最大规模的具身智能数据采集仅为百万条量级，差距达两个数量级**——驾驶数据来自存量车队、边际成本近零，具身数据却依赖实验室遥操与人工标注。这意味着：**具身智能无法照搬“真实场景数据飞轮”，特斯拉路径**

的门槛是“先拥有百万级部署身体”，绝大多数玩家必须另寻解法——这正是仿真合成与人类视频两条路线存在的根本原因。



图 3-2　具身大脑数据的四层结构

数据范式的战略含义。 三点判断：其一，**单一数据源都不够**——领先玩家的共性是“视频预训练＋仿真合成＋真机微调＋在线进化”的四源混合管线；其二，**数据基础设施正在独立成产业**——数据采集、清洗、标注、合成数据生成已形成专业服务商梯队，中国在这一层具备制造业级执行力优势；其三，**数据主权成为估值核心**——2026 年资本市场对具身智能公司的估值逻辑，已明确从“硬件制造”转向“数据资产的广度与闭环能力”（大象研究院观点），数据飞轮是否转动，是分辨“真大脑”与“演示大脑”的第一试金石。

3.8 评测与标准：话语权之争

技术路线多元并存的时代，“谁定义好与坏”成为隐性战场。

评测基准的格局。 国际上，LIBERO、Meta-World 等仿真基准沿用已久，RoboChallenge 等全球性评测开始承担“机器人统考”角色——2026 年千寻智能 Spirit v1.5 在该榜单超越 Physical Intelligence 的 π0.5 登顶，即为国内团队赢得话语权的标志性事件。国内基准快速补位：智元开源全球首个具身世界模型专用评测基准 EWMBench；深度机智 PhysBrain 1.5 以 28 项公开基准综

合均分 72.5 登顶开源模型榜，与闭源头部差距收窄至 1 分以内；RoboDojo、SimplerEnv 等构成多维度评测矩阵。需要注意的公共事实是：**评测分数与真实可靠性之间仍有显著鸿沟**——有测试显示顶尖通用大模型在毫米级精细任务上成功率可从 95%骤降至 10%；高分模型换一个真实场景可能需要重新验证。评测体系反映的是“当前能力”而非“部署可靠性”，这是所有采购与投资决策必须内置的认知。

标准体系的国家竞争。 中国已将标准建设提上日程：工信部《国家人形机器人产业标准体系建设指南（2026 版）》（征求意见稿）将世界模型标准列项；上海市人工智能标准体系建设指南（2026－2028 年）将“世界模型标准”列为关键技术标准，要求建立多维度评测体系；浦东“十五五”规划明确重点布局“具身大模型＋世界模型”融合技术。安全标准方面，英伟达 Halos 全栈安全系统已与 Agility、Fort Robotics、KION 等集成，正在形成事实安全框架。**标准之争的本质是生态之争：谁能把自己的技术架构写进标准，谁就为整个行业预定了兼容性。**

3.9 端侧部署与算力约束：被低估的工程主战场

技术路线之争吸引了几乎所有注意力，但决定商业化节奏的往往是端侧工程——**大脑必须跑进机器人的身体里**。

算力分层的三段结构。 具身智能的算力需求分为三段：云端训练（数据中心级，WAM视频生成式训练进一步抬高需求）、云端推理（低延迟任务分发与复杂规划）、端侧推理（机载实时控制）。端侧是商业化咽喉：Jetson Thor 以 2070 FP4 TFLOPS、128GB 统一内存成为国际人形平台的准默认配置；国产侧昇腾算力已支撑张江实训场的 VLA 研发。**“哪些能力放云端、哪些放端侧”的算力分配决策，直接对应 3.5 节的双系统架构——慢脑可以更大更云，快脑必须更小更端。**

端侧工程的四个现实问题。 其一，功耗与散热：人形机器人电池容量有限，大模型机载推理的功耗预算极其紧张，这解释了为什么 Helix 强调“机载嵌入式低功耗 GPU 运行”；其二，确定性时延：控制系统要求确定性的响应时间（Isaac OS 的设计目标即“deterministic, real-time performance”），而

AI 推理的时延天然波动，运行时层的调度设计是关键工程；其三，安全隔离：认知层错误不能直接穿透到物理执行，需要安全反射层（智平方“脊髓”层、英伟达 Halos 的设计逻辑）；其四，模型压缩：7B 级慢脑在端侧运行需要量化、蒸馏、 speculative decoding 等压缩技术，每一步压缩都以能力损耗为代价——**压缩比与能力损失的平衡曲线，是各家的核心工程机密之一。**

对产业格局的含义。 端侧工程能力是中国团队的潜在优势区——中国的手机与智能汽车产业沉淀了大批“在功耗、成本、量产约束下做 AI 部署”的工程人才与芯片生态（地平线类公司、智能座舱团队）。本报告判断：**未来三年“端侧大脑”将成为独立的产业环节，形成“云端大模型授权＋端侧推理芯片＋运行时中间件”的分层供应链；谁把端侧工程标准化，谁就掌握了一个类似“安卓”的卡位。**

3.10 技术选型的五个工程判据

面对四条路线与数十款模型，企业与投资机构需要一个可操作的选型框架。本报告提出五个判据，按重要性排序：

判据一：任务的时间尺度。 短程接触操作（插拔、旋拧、装配）首要看力控频率与端侧延迟，本能驱动与专用小脑优先；中程任务（取放、分拣、上下料）看 VLA 的泛化与稳定性；长程任务（整理房间、多工序作业）则必须有世界模型或强规划能力。**用长程模型做短程操作是浪费，用短程模型做长程任务是幻想。**

判据二：场景的结构化程度。 结构化场景（固定工位、固定工件）可依赖传统视觉＋专用策略，大脑需求低；半结构化场景（零售、药房、物流）是当前 VLA/WAM 的最佳战场；非结构化场景（家庭）需要世界模型＋人类视频预训练，且对可靠性要求最高——这也是家庭场景落地最晚的根本原因。

判据三：失败成本的量级。 分拣失败＝返工，装配失败＝报废，医疗与养老场景失败＝人身风险。失败成本越高，越需要“分层安全设计”：慢脑规划＋快脑执行＋独立安全反射层，且必须保留人工接管通道。**安全不是模型的属性，是系统架构的属性。**

判据四：数据可得性与独占性。 同样一款模型，在有独家数据管线的团队手里与在只依赖公开数据的团队手里，能力曲线完全不同。选型时应评估“这家公司的数据从哪来、能否持续、是否独家”，而非只看当前榜单分数。

判据五：部署与运维的工程完备度。 包括端侧算力适配、模型压缩方案、OTA 升级、日志与回放（事故复盘）、多机协同。**这一项在 POC 阶段最容易被忽略，却是在规模部署阶段决定总拥有成本（TCO）的关键。**

选型结论的表述建议： 不存在“最好的大脑”，只存在“匹配任务、匹配数据、匹配风险预算的大脑”。企业应在选型报告中明确四件事：任务时间尺度、场景结构化程度、失败成本上限、数据权益安排——这四项确定之后，路线选择往往是推导结果而非主观偏好。

3.11 增量观察：2026 年 10—11 月的六项进展与三点修正

本报告上一版（V8.0）完成于 2026 年 10 月中旬。此后一个月，具身大脑领域出现了密度罕见的进展，其中多项直接触及前文的判断依据，故单列一节作增量校准。**本节只收录可核查的公开信息，并对企业与官方口径分别标注。**

进展一：通用基模的能力第一次被“算清楚”。 10 月中旬，英伟达研究团队发布《Astranomy：GPT-6 Astra 作为原生机器人策略的经验研究》，在高保真仿真基准 RoboLab-120 上测试三种用法：模型直接生成动作成功率 66.3%、模型作为智能体持续规划并调用工具成功率 74.8%、模型监督专用机器人策略成功率 66.5%；智能体方案比该测试此前最优公开策略高出 23.8 个百分点。代价同样被量化：直接控制单次调用平均 18.56 秒，智能体方案每个任务平均消耗约 3090 万 token、估算费用 36.40 美元，直接控制方案约 1.96 美元。10 月 20 日的《前沿大模型进入机器人世界：基于 RoboColiseum 的跨模型评估》把 GPT-6 Astra、GPT-6.1 Sol、Claude Opus 5.5 与 Gemini 4 Argon 纳入统一框架，通过 Agent2Act 连接不同型号机器人，在仿真中完成垃圾投放、搬运、开门、导航，在真机上完成部分抓取、积木堆叠与简单叠衣。**三点读法：其一，“智能体编排”优于“端到端直出”，说明通用基模的正确位置是调度者而非执行者；其二，成本结构决定了它只能先进入“低频、高价值、可容错”的任务；其三，以上成绩来自允许暂停仿真环境的测试协议，不能等同于连续产线作业。**

进展二：竞争焦点从“手”扩展到“全身”。 10月23日，高德与阿里ATH事业群发布全身基座模型 λ0：以佩戴式动捕与颈挂相机采集、机器人本体不参与的方式构建 HumanVerse-500（500 小时开放世界全身数据），把单条示范采集用时从约 100 秒压缩到 29 秒（效率约为主流方法的 3.4 倍）；在 SIMPLE 基准取得 90.0%、真机四任务平均 62.5%，较最强对比基线提升 22.5 个百分点；同一技术报告给出的真机对照为 GR00T N1.6 7.5%、Ψ_0 10.0%、StarVLA 27.5%、$\pi_{0.5}$ 40.0%（四任务、每任务 40 次试验）。同期，Figure 的 Helix 2.5 在 30 户从未采集数据的湾区家庭完成420次试验、成功237次（约56%），去掉Index预训练的同任务基线仅9%；宇树9月10日的UnifoLM-WLA-1.0以单模型统筹64项任务并开源权重、代码与数据集。**“无本体采集”“零样本进家”“全开源”三条路线同时成立，说明全身协调的动作数据不必绑定本体——这是对3.7“数据四层结构”的重要补充：人类数据的价值正从“语义先验”升级为“全身运动学先验”。**

进展三：学习范式从“训出来”变成“干出来”。 10月29日智元发布的GO-3 Preview，是国内首个把“预训练—后训练—部署后”写成完整链路的基座模型。架构上采用自研WRAM（世界推理动作模型），以CoAE控制导向视觉编码、SVP 单步视觉规划、KASO 知识对齐筛选三项设计，把任务理解、未来状态预测与动作生成统一起来；官方称视觉输入 token 压缩至 DINOv3 的十六分之一、动作恢复误差降低 35%，预训练数据约九成来自人类第一视角。后训练依托 Genie Evolver（百台真机＋物理仿真＋Genie Envisioner 2.0 世界模型），部署后由 Genie Master 与具身执行协议 EEP 承接，使机器人在真实便利店“无人在环”持续改进：完整任务成功率从首日 28.6%升至第 5 天的 85.7%，单任务成功率较常规 SFT 提升 17.7 个百分点；官方称其为真实业务场景中的 Physical RSI（具身递归自进化）验证，并计划2026年底全面开源GO-3 Preview与Genie Master。该模型在RoboDojo、RoboTwin（85.0%）、RoboColiseum（78.82%）三个榜单登顶（企业披露口径）。**这条曲线的意义大于分数：它第一次把“部署后学习速率”变成可比较的指标——未来评价一个大脑，不应只看榜单一栏，而要看“第 1 天到第 30 天的斜率”。**

进展四：世界模型进入“可用区间”，但仍有时长天花板。 10 月 13 日，Google DeepMind 发布 Gemini 4 Argon for Robotics：在 ERQA 具身推理基准以纯推理取得 87%；官方演示覆盖四条工作流——MuJoCo“练习—执行”的 Spot 移动操作、零样本机械臂控制、闭环驱动视频世界模型完成字母排列、从机器人相机/第一视角视频/手绘草图重建可校准仿真孪生；技术说明披露 77 项任务、4500 余个域随机化回合，经三轮自迭代把完成时间从 58.9 秒降至 17.4 秒，真机任务成功率提升 56%—75%。10 月 27 日，Genie 4 研究预览版把实时交互世界模型的稳定时长从约 1 分钟延长到约 2 分钟，首次原生同步生成环境音与角色对话，并支持在推演过程中用自然语言实时改写天气、镜头与对象；10 月 30 日，英伟达发布开放权重的 Cosmos 3，用混合 Transformer 把视觉推理、世界生成与动作预测压进同一模型。**判断：世界模型已跨过“能生成”的门槛，进入“能用于训练与评测”的区间；但 2 分钟的稳定窗口与真实产线的连续作业要求仍相差一个量级，短期内它的主战场仍是数据合成与策略预演，而非直接上岗。**

进展五：基础设施层开始独立成形。 四条路径值得记录：地平线开源 HoloBrain-0（0.2B 与 1.1B 两版），把相机参数与 URDF 运动学结构显式注入架构，以“统一基座坐标系＋混合相对动作空间”支持跨本体，在 RoboTwin 2.0 的 50 项任务上平均成功率超过 90%，0.2B 版本可在 RDK S100 端侧以高于 5Hz 运行，并配套 RoboOrchard 数据—训练—部署工具链；炎曦动力（2026 年 7 月成立）发布 Paimon-0，以分层异步 WAM＋逆动力学模型把单个动作块推理从约 3000 毫秒压到 300 毫秒，在 RoboColiseum、RoboCasa365、VLA-Arena 及 IROS Workshop 双臂挑战赛登顶，走“先做模型、再做芯片”的路线；中国移动 11 月 5 日发布“灵韬智脑”，以云侧 300 亿参数慢脑＋边端 40 亿参数灵韬 VLA 快脑＋具身智能网（传输时延低至 30 毫秒）实现“网联智能”，已掌握 20 余项技能、15 类原子动作、可连续完成 21 步任务；眸深智能在进博会发布“深悟”（一步动作生成，推理最高提速 24.84 倍、仅新增 13.4M 参数）、“深枢”Infra 架构与家用世界模型 FWM，并联合复旦、国地中心、宇树、优必选等 42 家机构成立“通用具身大脑产业创新联合体”。**“模型—框架—工具链—网**

络—模组”正在形成一个独立于本体与整机的供给层，这正是 5.7 所说“中间层被挤压”的另一面：薄壳在消亡，底座在成形。

进展六：标准与场景开放同步落地。 10 月 15 日首批 6 项人形机器人国家标准发布（2027 年 4 月 1 日实施），10 月 19 日商务部等 9 部门发文推动具身智能进入商贸、家庭服务、仓储物流，上海明确 2028 年 8000 台人形机器人进工厂。**标准解决“按什么验收”，场景开放解决“在哪里练”——两者合起来，才是数据飞轮真正启动的前提。**

由此对本报告的三点修正：

修正一：通用基模的威胁不在“替代”，而在“定价”。 前文判断通用基模将吃下认知层，这一判断成立；但更准确的说法是——它会把“语义理解＋任务规划＋仿真生成”变成廉价公共品，从而压缩专用大脑公司在认知层的溢价空间，而不是直接取代它们。**谁能把通用基模的能力转化为“更低的部署成本、更短的上线周期”，谁就能把降价压力转成份额。**

修正二：评价标准应从“榜单分数”转向“部署后斜率”。 GO-3 的 5 天爬坡、Helix 2.5 的零样本 56%、Dyna-2 的 3 天回本，共同指向一个新指标：**单位时间内成功率的提升速度**。榜单衡量“出厂能力”，斜率衡量“进化能力”，后者才是复利的来源。

修正三：数据竞争的关键变量是“无本体可复制性”。 λ0 用人类佩戴设备采集、Paimon-0 用 EgoScale 第一视角数据预训练，都证明了一条路径：不依赖自有本体规模，也能构建可扩张的数据管线。**这对没有万台部署的二线玩家是利好，对“以部署量换数据”的头部玩家则是估值逻辑上的挑战——部署规模的护城河，没有想象中那么深。**

3.12 本章结论：融合收敛，分层永存

本章的六条结论，构成后续章节的分析基础：

1. **VLA 仍是当前主流与商业化最快路线**，但其“高频精细交互受限＋物理理解缺失”的结构性短板，决定了它必须进化；

2. **世界模型/WAM是确定性最高的演进方向**——“先想后做”是物理智能区别于语言智能的本质特征，英伟达与智元的架构选择已给出风向标；但世界模型的“完整大脑”叙事需要重写，其价值重心转向物理校验与合成数据引擎；

3. **通用基模的突袭改写了能力栈的分层**——语义理解、任务规划、仿真生成、微调托管四层正被通用基模以更低边际成本接管，专用大脑的价值必须向物理力控闭环与端侧实时性这两层“通用基模够不到的地方”迁移；

4. **本能驱动划出了“分层”的合理性**——大脑不需要、也不可能包办一切，“什么能力放在哪一层”将持续重构；

5. **数据范式决定护城河，评测标准决定话语权**——未来 3—5 年，评价一家大脑公司，看它的数据管线与标准位次，比看它的演示视频可靠得多；

6. **学习闭环的终点在部署之后**——2026 年四季度的进展（GO-3 的部署后自进化、Helix 2.5 的零样本迁移、λ0 的无本体数据管线）共同说明：大脑的竞争力不再只由训练阶段的规模决定，而由“部署—反馈—再训练”的转速决定。**评估大脑，最终要评估它的进化速率。**

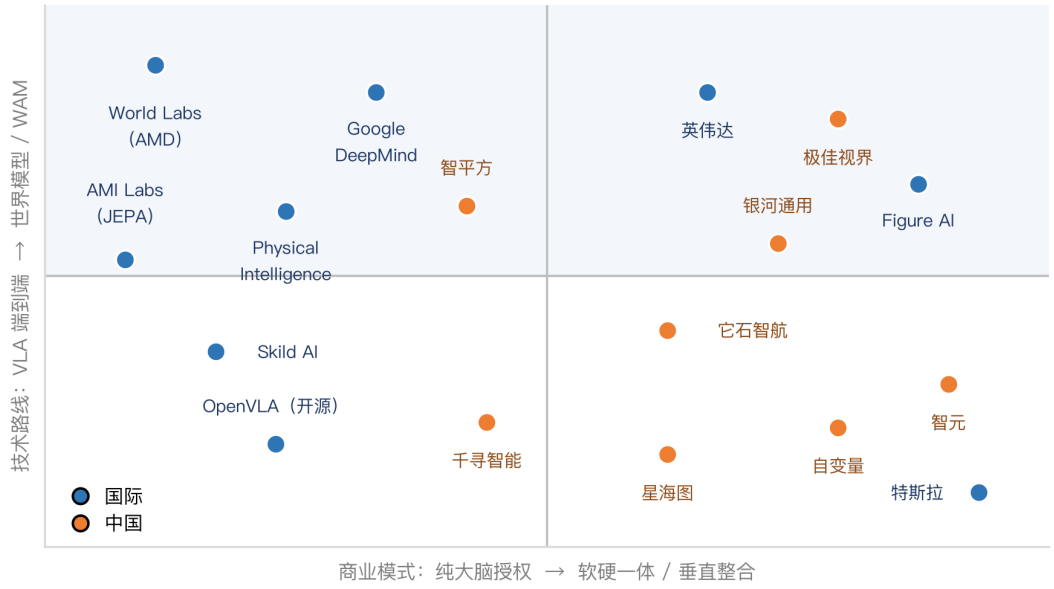
第四章 代表案例：全球具身大脑竞争格局

本章导读 本章以“商业模式×技术路线”为坐标，解剖全球代表性玩家：英伟达、谷歌 DeepMind、Physical Intelligence、Figure AI 组成的第一梯队各占一个生态位，OpenAI 以 GPT-6 Astra 正面入场、Dyna Robotics 以 WAM 路线跑出第一个商业闭环、极佳视界以世界模型原生路线完成“数据放大”闭环、高德以 λ0 把大厂基座下沉到“全身”、地平线与炎曦动力以“框架＋芯片”切入底座层、DeepSeek 以“投资拼装”组建技术栈，特斯拉以 FSD 复用把垂直整合推向极致，李飞飞（World Labs）与杨立昆（AMI Labs）代表的学界路线则为世界模型补上两条新叙事；国内梯队以全栈自研与软硬一体见长，分化为三个层级。随后研判全球格局、绘制资本地图、提炼可迁移的战略规律，最后以本章小结收束，供不同类型的主体对照自检。

本章以“商业模式×技术路线”为坐标，解剖全球代表性玩家。选择标准有三：技术代表性、市场影响力、数据可得性。需要预先声明的分析纪律：**各企业技术指标均来自官方发布或公开报道，属于企业自证口径；“演示能力”与“部署可靠性”之间需保持审慎的判断距离。**

全球格局的四个生态位：没有唯一正确的模式，只有匹配禀赋的选择

横轴：纯大脑授权 → 软硬一体；纵轴：VLA 端到端 → 世界模型 / WAM



资料来源：各企业公开信息；象限定位为研究判断，不代表企业自述。

图 4-1 全球具身大脑玩家定位矩阵

4.1 国际阵营：从“四极”到多线并进

4.1.1 英伟达：不造机器人，但定义机器人造法

战略定位：平台生态型——做机器人产业的“基础设施供应商”。

英伟达的具身大脑布局不是一款模型，而是一张全栈网络：**模型层**，GROOT N系列人形基础模型（N1→N1.5→N1.6→N1.7→N2预览版）持续快速迭代，N1.7以 EgoScale 人类视频预训练树立数据范式新标杆，N2 的 DreamZero 架构转向 WAM 路线；**世界模型层**，Cosmos 系列从合成数据工具进化为全模态世界模型平台（Cosmos 3）；**工具链层**，Isaac 平台（Isaac Sim/Lab 3.0/OS 运行时）打通“数据生成—训练—仿真验证—真机部署”全流程；**算力层**，Jetson Thor 边缘计算平台（2070 FP4 TFLOPS、128GB 统一内存）承担机载推理；**安全层**，Halos 全栈安全系统。

2026 年 10 月 GTC 柏林大会的发布，集中展示了英伟达的打法：联合思灵机器人（Agile Robots）推出预装 Isaac GROOT、由 Jetson Thor 驱动的全栈人形参考平台 Agile ONE SN（1.4 米、31 自由度、双手 20 自由度灵巧手、1kHz 力控），把开发者从“数周环境搭建”压缩到“数天”；发布 SimReady 资产库（超 2000 个仿真就绪资产）与 Human 1M 人类活动数据集（含三视角视频与 3D 手部追踪）；开源三大参考工作流。ABB、KUKA、Universal Robots、NEURA Robotics 等欧洲机器人生态均已接入其开放平台。

战略意图判读： 英伟达在复刻 AI 训练领域的成功路径——**不卖机器人，卖所有机器人建造者都依赖的那一层。** 参考人形平台的真正价值不在产品本身，而在为全行业提供“已知良好基线”（calibration target），进一步拉低入场成本、放大生态引力。其对中国的启示与警示同样深刻：当 Isaac、GROOT、Cosmos、Jetson 成为默认路径，机器人技术栈将继承 AI 训练栈已有的“单供应商引力”——这正是国内全栈自研阵营（智元等）与国产算力路线（昇腾）对抗性存在的战略理由。

4.1.2 Google DeepMind：把机器人变成 Gemini 的“手”

战略定位：大厂分层型——具身能力是 AI 生态的延伸，而非独立业务。

DeepMind 的技术脉络清晰：RT-1（2022）→RT-2（2023，确立 VLA 概念）→OPEN X-Embodiment 跨本体数据集（2023）→Gemini Robotics/ER（2025.3）→Gemini Robotics 1.5（2025.9，智能体式推理＋跨本体技能迁移）→Gemini Robotics-ER 1.6（2026.4，增强空间推理与多视角理解，含与波士顿动力合作开发的仪表读取能力，已通过 Gemini API 开放）。

其体系化打法是“双模型分工”：Gemini Robotics-ER 1.5 作为规划者（具身推理，先做多步规划），Gemini Robotics 1.5 作为执行者（将规划转为电机指令）；后者已在 ALOHA 2 双臂、Franka 双臂、Apptronik Apollo 人形三种异构本体上验证跨本体技能迁移（Motion Transfer）——**在一种机器人上采集的技能数据，可直接让另一种机器人受益**。DeepMind 自己的论文也坦承了边界：单机器人训练的模型在跨本体测试中表现不佳，迁移收益取决于各本体的数据存量，人形本体的控制难度更高。

另有端侧版本 Gemini Robotics On-Device（2025.6 发布），面向无网络、低延迟场景，仅需 50–100 条演示即可适配新任务，已适配 Franka FR3 与 Apptronik Apollo。

2026 年 10 月的新变量：Gemini 4 Argon。 谷歌在通用推理模型 Gemini 4 Argon 上单独释放了机器人能力，机器人侧形成清晰的三层分工——**Argon 负责写仿真代码、从视频生成数字孪生、完成任务标注；Gemini Robotics 2 承担 VLA 真机控制；ER 系列负责具身推理**。Argon 以纯推理在具身推理基准 ERQA 取得 87%、超过同批所有前沿模型；其 Real2Sim 流水线可从一段遥操作录像生成物理校准的 MuJoCo 孪生、从第一视角视频把人类动作重定向到 76 自由度人形本体，据官方演示仅凭一张桌面照片即可生成道具、定义数十个任务并产出数千个仿真回合，把下游真机任务成功率从 50%多提升至 70%以上、任务耗时从 30 多秒压缩到 20 余秒。**仿真环境构建从“周”压到“小时”，这是对“真实机器人数据昂贵、难复现、危险动作不敢试”这一行业第一瓶颈的直接攻击**（注意边界：以上为谷歌自选演示，尚无大规模部署数据与长期可靠性披露）。

战略意图判读： 谷歌的路线印证了一个判断——**具身大脑的“大脑皮层”终将长在通用多模态大模型之上**。其风险与优势同源：距离硬件与场景较远，

商业模式（云端 API＋生态授权）依赖机器人厂商的接入意愿；但一旦 Gemini 生态的智能密度形成代差，“大脑即 API”将成为最具统治力的形态之一。

4.1.3 Physical Intelligence：给“纯大脑”模式定价的公司

战略定位：纯大脑型——只做通用大脑，不造身体。

PI 是“大脑层独立性”最纯粹、也最极端的检验样本。团队密度堪称行业天花板：CEO Karolhausman 出自 Google 机器人研究团队，首席科学家 Sergey Levine（伯克利）、研究负责人 Chelsea Finn（斯坦福）是该领域被引次数最多的学者之列。其融资轨迹是观察资本对“大脑层”定价意愿的最佳窗口：

表 4-1 Physical Intelligence 融资轨迹

时间	轮次	金额	投后估值	领投/关键投资方
2024.3	种子轮	7000 万美元	约 4 亿美元	Thrive Capital
2024.11	A 轮	4 亿美元	约 24 亿美元	贝佐斯、Thrive、Lux；OpenAI 连投
2025.11	B 轮	6 亿美元	约 56 亿美元	CapitalG（谷歌系）领投
2026.5	C 轮（未经官宣，公司文件披露）	约 10.5 亿美元	约 110 亿美元	Founders Fund、Lightspeed 等（据报道）
2026.10	D 轮（洽谈中，据彭博社 10 月 27 日报道）	至少 10 亿美元	约 200 亿美元	Avenir、Lux Capital 拟领投

从种子轮到 D 轮洽谈估值，30 个月估值涨约 50 倍——**在零产品、零收入、且 COO 公开表示“不受短期商业化目标约束”的条件下完成。** 资本买的是什么？买的是 π 系列模型的能力曲线：π0（2024.10，流匹配＋跨本体）→π0.5（2025.4，开放世界泛化）→π*0.6（2025.11，RECAP 经验学习）→π0.7（2026.4，组合泛化涌现）——这条曲线与 2020—2023 年 LLM 的能力曲线形态高度相似，投资人在做的是“把 Scaling Law 的期权再买一次”。**需要标注的是，这一轮估值与 Pi API 上线仅相隔两周：资本市场对“平台化叙事”的定价速度，已经快于对收入曲线的验证速度。**

PI 在 2026 年 WAIC 上的表态进一步明确了商业立场：“只做通用大脑”，不研发自研硬件——没有自研硬件在某些具体任务上确实受限，但核心目标是把基座模型的能力“涌现出来”。

2026 年 10 月 13 日，Pi API Beta 上线——从“模型产品”走向“产业平台”。 其链路为：开发者上传业务数据→Pi 平台托管微调→云端推理网关下发真机→效果评估→补充迭代数据。硬件开发者不再需要自建训练集群、维护推理服务：据公开报道，Chef Robotics 仅用几十条双臂示教数据、隔夜即获得定制模型，次日上机实操制作餐食；Almond 以四天 API 迭代实现双臂零件去毛刺；另有团队将其用于仓库打包、衣物折叠。这一步的战略含义大于技术含义：**过去“基座公司做模型、应用公司管微调与部署”的分工，被 Pi 单方面改写为“微调与托管推理收归平台”，下游只留下硬件集成与真实场景数据采集两个环节。** 同时必须看到硬币的另一面：调用 API 不等于数据自动回流，Pi 尚未公布定价与数据权属规则，工业客户是否愿意把产线布局与故障数据交给平台，是商业博弈而非技术问题。

战略意图判读与风险提示： PI 是“每台机器交一次大脑授权费”叙事的旗舰。但本报告必须提示该假设的脆弱性：**只要有足够多的本体厂证明“自己的大脑够用”（Figure 已经这样做了），授权费逻辑就会塌陷。** PI 的护城河不在当前模型，而在其跨本体数据资产与开源生态（openpi）带来的社区绑定。它是这个赛道最重要的“商业模式压力测试”。

4.1.4 Figure AI：垂直整合的“数据飞轮”实验

战略定位：垂直整合型——大脑不外售，作为整机竞争力内核。

Figure 是“本体厂自研大脑”路线的标杆。技术主线为 Helix 系列：初代 Helix（2025.2）双系统架构开创者之一；Helix 02 扩展至全身与双腿（四分钟洗碗任务、61 个移动操作动作、双机器人无中央协调协作）；Helix 2.5（2026.9）转向**数据战略的豪赌**——放弃以预训练 VLM 为起点，改为在其自建人类行为数据平台 Index 上“从随机初始化预训练”。

Index 是理解 Figure 的关键：该平台 2026 年 8 月上线，汇聚超过 1600 万条人类执行真实任务的视频、超 4.4 万周活跃贡献者、覆盖 108 个国家，配套完整的筛选、防欺诈、去重、再平衡与自动标注流水线。基于 Index 预训练的策略，在 30 栋从未参与训练的湾区民宅中取得 56%任务成功率（同一任务数据量减半达到上代同等水平；不用 Index 预训练的对照组仅 9%）。此前 Figure 还

通过 Project Go-Big 验证了从人类自我中心视频到机器人导航行为的零示教迁移。

战略意图判读： Figure 证明了垂直整合的可行性——当大脑与本体协同优化时，整机能达到第三方大脑难以匹敌的适配深度；Index 则试图解决垂直整合模式最大的软肋（部署基数小于第三方大脑公司，数据飞轮转不快）——**用“人类视频”替代“机器人部署”作为数据飞轮的能源。** 风险在于：1600 万条原始视频≠有效训练数据，人类视频到机器人动作的对齐机制仍是黑箱；且该模式高度依赖持续的资本供给（其 D 轮估值已达 390 亿美元量级，据公开报道口径）。

4.1.5 补充观察：Skild AI 与 OpenAI 的回马枪

Skild AI（匹兹堡）以“全形态大脑”（omni-bodied brain）为旗帜：单一模型驱动四足、人形、桌面机械臂与移动操作平台。2026 年 1 月完成 14 亿美元融资（软银领投，英伟达风投、贝佐斯参投），估值超 140 亿美元；其自称 2025 年收入约 3000 万美元、2026 年 4 月收购 Zebra Technologies 机器人部门以获取已部署硬件与客户关系——是“大脑公司”中罕见的“有估值也有收入”样本（其披露口径：90%收入来自手部操作场景，10%来自移动）。**2026 年 9 月，Skild 披露了更完整的一组商业化数据：正式商业部署 10 个月、付费客户超过 60 家、年化经常性收入（ARR）突破 1 亿美元、累计确认收入 5000 万美元**；客户分布覆盖仓库、工厂、配送、安防与餐饮，并与英伟达、富士康合作将模型部署于 Blackwell 系统的高精度组装工位，与住友电装尝试汽车线束工序。需要说明的是，ARR 是年化折算而非实收，且其中包含软件授权、设备部署与系统集成的混合收入——**“模型公司必须下场解决客户问题才能收到钱”，这组数据第一次为这句话提供了财务证据。** FieldAI、Skild、PI 等“不造身体”的公司，未来都必须买身体、买手——这既是一条细分供应链机会，也是大脑公司与本体厂利益绑定的纽带。

OpenAI：2026 年 9 月初发布的 GPT-6 Astra，让“回马枪”变成了正面入场。Astra 本身是通用多模态旗舰，官方定位中并无“具身智能”标签，但多个独立团队的实测把它推上了具身舞台的中央：RoboDojo 42 项任务、2100 次测

试中，未经任何机器人微调的 Astra 充当双臂机械臂动作策略，平均成功率 22.48%、超过评测中 40 个公开机器人专用策略（同期 GPT-5.5 仅 0.88%）；在语义理解与陌生场景泛化任务上表现突出，精细操作、动态控制与双臂协同仍是短板；与 π0.5 构成“基模检查修改＋专用模型执行”的混合架构可将双臂成功率从 26%提升至 48%。宇树科技已将 GPT-6 接入自家机器人实测，王兴兴反馈“效果非常不错”，但承认复杂物理交互仍是通用基模跨不过去的坎。**2026 年 10 月 18—19 日，OpenAI 确认自研人形机器人，招聘岗位覆盖执行器、运动控制、数据采集与供应链等二十余个方向，并给出 2027 年演示样机的目标**——从“只做大脑”到“下场做身体”，这一步的象征意义大于其产品意义：**当通用基模公司判断“身体是数据入口”时，纯大脑与纯本体之间的边界就被彻底打开了。 一家“只做大脑”声明的公司，用一款不做机器人营销的通用模型，完成了对机器人专用模型的大规模对标——这对“具身专用路线”的冲击，比任何一款人形机器人发布都更大**（实测证据的完整梳理见 3.6 与 3.11）。

4.1.6 Dyna Robotics：WAM 路线跑出的第一个商业闭环

战略定位：全栈软硬一体＋WAM 路线——“Not Just a Model, But a Product”。

Dyna Robotics（加州红木城，创始团队含前 Google DeepMind 研究员，联合创始人曾以 3.5 亿美元出售 Caper AI 给 Instacart）在 2026 年 8 月发布的报告中宣布：其 Dyna-2 机器人已在美国鼎泰丰全部门店网络铺开，并跨过投资回报（ROI）门槛。关键数据（官方披露口径）：**Dyna-2 每小时折叠 95 条餐巾、合格率 93%，18 小时班次产出 1590 条，超过客户设定的 1500 条门槛；对比 2025 年 6 月的 Dyna-1（35 条/小时、合格率 75%、日产约 480 条），速度提升 2.7 倍**；零样本部署（新门店不做定制训练直接上工）合格率从 46%提升至 87%，新门店装机后最快 3 天达到投资回报平衡。训练路径同样反主流：以超 100 万小时第一人称人类视频预训练、全程不使用遥操作数据，官方称此为首条“人类数据到机器人”的规模定律；其 Dyna-2 属世界-动作模型（WAM）——同时预测下一画面与下一动作，与 PI、英伟达的 VLA 路径形成分野。

战略意图判读： 具身智能行业长期“炫技视频多、损益表少”，Dyna-2 是目前少数给出**可核查商业闭环数据**的 WAM 样本：单店 ROI 门槛、零样本部署合格率、按指令切换任务（换摆法只需换指令、不换模型）三个指标共同指向一个判断——**WAM 路线的商业可行性已经拿到了第一个实证**。其“指令驱动、换任务不换机器”的产品定义，本质是把机器人卖成“可用文字重新指派的劳动力”，这正是大脑层价值的具象化。2027 年上半年装机量目标数百台（酒店、物流、数据中心同步部署），能否兑现是观察 WAM 商业化的下一个锚点。

4.1.7 特斯拉 Optimus：垂直整合的“数据飞轮”终极形态

战略定位：不外接任何大脑——把 FSD 的端到端体系整体平移到人形机器人。 如果说英伟达在做“机器人界的 Android”，特斯拉的逻辑就是“机器人界的 iPhone”：软硬一体、芯片自研、数据自闭环。马斯克在 2026 年 7 月 23 日二季度财报电话会上明确了技术路线：Optimus 采用与 FSD 完全相同的端到端战略——“输入像素，输出控制指令”，**用单一、统一的神经网络替代感知、规划、预测分离的模块化架构**：多路摄像头输入视觉信号，关节编码器输入本体感受，手部输入力触觉反馈，网络直接输出连续控制信号；智能系统深度复用 FSD 纯视觉体系，据媒体报道配合 HW4.0 视觉芯片与十余颗摄像头，“感知—决策—执行”闭环约 0.3 秒。人事安排同样说明问题：2025 年 6 月起，FSD 端到端架构的核心缔造者 Ashok Elluswamy 接掌 Optimus 项目——**这不是岗位调整，而是架构决策：车载 AI 与机器人大脑被定义为同一个模型体系。**

数据飞轮是特斯拉的终极壁垒。 全球约 130 万辆具备监督式自动驾驶能力的特斯拉车辆在真实道路上持续运行，FSD 累计行驶里程已达百亿英里量级（据新京报援引行业统计口径），这些数据直接反哺 Optimus 的感知与决策模型；同一颗车载 AI 芯片既驱动 FSD 也驱动 Optimus。对照第三章的数据范式分析（见 3.7），特斯拉是目前唯一拥有“自动驾驶级真机数据飞轮”的玩家——银河通用创始人王鹤的对比说明了这一优势的量级：**头部车厂单天驾驶数据回流约 1 亿条，而全球最大规模的具身智能数据采集仅为百万条量级，差距达两个数量级。**

量产节奏与“Optimus Academy”数据战略。 特斯拉已将弗里蒙特工厂原 Model S/X 产线改造为 Optimus 产线，2026 年二季度财报确认产线安装就绪、年底启动生产；V3 版本远期产能目标年产 100 万台（马斯克同时承认这是“特斯拉史上最难规模化的产品”，产能爬坡曲线“长而平”）。**2026 年 9—10 月，节奏第一次给出了可追踪的中间节点：**9 月，特斯拉向核心供应链下达约 5000 台 Optimus 批量订单（此前为数百分试产订单）；10 月 19 日提交的年度股东大会材料首次以书面形式称第三代Optimus“已为大规模量产做好准备”，但未披露产量、价格与交付时间；10 月 21 日三季报进一步披露，弗里蒙特工厂已启动早期产能爬坡并下线首批达到量产质量标准的机器人，该产线设计年产能约100 万台，得州第二产线计划 2027 年夏季投产、长期设计年产能约 1000 万台；马斯克在财报电话会上称，产能预计至少一年后达到周产 2 万台，2027 年底前后开始向特斯拉以外企业交付。**本报告提示：从“做好准备”到“规模交付”之间仍隔着上万种新开发零部件的良率爬坡，2027 年夏季的得州节点与对外交付时点是两个必须持续核验的锚。**更值得注意的是首批机器人的去向：**不对外销售，全部进入“Optimus Academy”——特斯拉自建的机器人训练场**，在真实工厂环境中执行任务、采集物理交互数据、迭代模型后再部署。马斯克曾透露该体系的规模设想达上万台量级、以“自我博弈”（self-play）方式持续产生训练数据。这实际上是把第二章所述“数据飞轮”从车队复制到机器人：**先造身体，再造数据，数据反过来定义大脑。**

战略意图判读与风险提示： 特斯拉路径的独特价值在于验证了一个命题——**当“身体”的部署规模足够大时，大脑的数据成本可以趋近于零**；其对行业的直接刺激是马斯克同场发言“当前网络流传的人形机器人演示大多依靠预编程或远程操控，目前没有任何一款人形机器人能仅凭自然指令自主完成多样化日常工作”——这与本报告 6.4 的“机器人 ChatGPT 时刻”判据形成呼应。风险同样清晰：端到端单一网络的可解释性与安全验证尚无行业先例；量产爬坡受制于全新供应链（关节、灵巧手、传感器均为定制件）；“真机数据飞轮”的优势目前集中在工厂等结构化场景，向家庭等开放场景的泛化仍属未知。**特斯拉证明了垂直整合的天花板，也定义了这条路线的成本下限——留给其他玩**

家的选择是：要么接入生态（英伟达路线），要么在场景纵深上建自己的小飞轮。

4.1.8 学界巨头下场：李飞飞的 World Labs 与杨立昆的 AMI Labs

2026 年，两位最具影响力的 AI 学者以创业公司的方式进入物理智能牌桌，为世界模型路线补上了两条“学院派”叙事。

李飞飞与 World Labs：空间智能的“世界模型独立成产品”路线。 “AI 教母”李飞飞 2024 年创办 World Labs，主张智能的下一个前沿是“空间智能”（Spatial Intelligence）——让 AI 理解、生成并与三维世界交互。产业进展在 2026 年密集落地：2 月完成 10 亿美元融资（Autodesk 领投 2 亿美元，AMD、英伟达、富达等参投，据彭博社报道估值约 50 亿美元）；2025 年 11 月发布首款产品 Marble（文字/图像/视频生成可交互 3D 世界，2026 年 1 月开放 API）；**7 月 21 日收购机器人仿真公司 SceniX，7 月 28 日发布 Real-to-Sim-to-Real（R2S2R）引擎——把真实任务重建为可控仿真的“千级变体世界”，机器人在其中训练的策略可实现零真实数据部署、连续自主运行 1 小时（李飞飞公开披露口径）**；9 月 1 日发布新一代世界模型 Atlas——原生支持文本、图像、视频与三维的多模态自回归扩散 Transformer，以“空间上下文”实现像素级相机控制（可生成 1440p、1 分钟长视频），稀疏图像三维重建表现优于专用 SOTA 模型（世界实验室官方博客、中科院网信工作网口径）。**最具标志性的事件发生在 9 月 28 日：AMD 宣布以约 82 亿美元股票收购 World Labs，李飞飞出任 AMD 执行副总裁兼首席科学家**——从创业到被芯片巨头并购不足两年，累积融资约 12 亿美元的 World Labs 以 7 倍于首轮估值的价格退出。

判读： World Labs 的两点启示超越了其自身规模。其一，**世界模型可以脱离机器人本体独立成为产品**——面向创意、仿真与训练场市场（Marble/Atlas/R2S2R），这为“世界模型≠机器人大脑”提供了正面注脚（呼应 3.2 的定位之变）：它首先是一种“空间智能基础设施”，其次才是大脑组件。其二，**AMD 的收购是“算力层反向购买大脑能力”的第一案**——芯片巨头需要通过拥有世界模型来定义下一代 AI 工作负载的芯片形态（苏姿丰表态），大脑层的价值由此获得了来自上游算力层的确认。

杨立昆与 AMI Labs：JEPA 架构的“非 LLM 世界模型”路线。 图灵奖得主、前 Meta 首席 AI 科学家杨立昆 2025 年底离开 Meta，2026 年 1 月在巴黎启动 AMI Labs，3 月完成 10.3 亿美元种子轮融资（投前估值 35 亿美元，欧洲史上最大种子轮；Cathay Innovation、Greycroft、Bezos Expeditions 领投，英伟达、三星、丰田等参投）。其技术路线基于他 2022 年提出的 **JEPA（联合嵌入预测架构）**：不在像素或 token 层面逐点预测，而在抽象表征空间中学习世界运行的因果规律——“LLM 对物理世界的理解不如一只老鼠”是其标志性论断；团队已展示根据 16 秒第一视角视频连贯预测后续 10 秒场景（含物体轨迹与光线变化）的成果，计划先做研究、一至两年后与行业伙伴落地，目标三至五年内做出“相当通用的智能系统”，并承诺大部分代码开源。

判读： AMI Labs 尚无机器人产品，但它的意义在于**为“具身大脑需要什么样的架构”提供了一个严肃的非主流答案**——如果 JEPA 路线被验证，当前以生成式 VLA/WAM 为主流的架构假设将被部分改写；即便不被验证，10.3 亿美元欧洲最大种子轮本身也说明：**资本市场已开始为“LLM 之外的物理智能选项”支付大额期权费**。对产业的真实影响是节奏问题——AMI Labs 自称三至五年周期，意味着它不是 2027—2028 年窗口期的变量，而是 2030 年格局的潜在改写者。

4.2 中国梯队

中国已形成超 150 家人形机器人企业、超 30 家“机器人大脑”及具身大模型创业公司的产业集群；截至 2026 年 5 月，百亿估值俱乐部至少 10 家、200 亿俱乐部 6 家。中国玩家的共性打法是：**数据路径差异化＋场景纵深绑定＋开源评测话语权建设。**

4.2.1 智元机器人：全栈自研的“对标英伟达”路线

智元是中国唯一在“本体量产＋基座模型＋仿真平台＋数据集＋评测基准”五个层面同时布局的玩家，公开将自身定位对标英伟达模式。关键节点：2025 年发布启元大模型（GO-1）并开源；2025 年 12 月第 5000 台通用具身机器人量产下线（央视报道口径）；2026 年 4 月发布 GO-2 基座模型——以“动作思维链＋异步双系统”在 LIBERO、LIBERO-Plus、Genie Sim 3.0 等基准全面 SOTA，仅

用仿真数据训练即取得真实环境平均成功率 82.9%（企业披露口径）；9 月发布世界-动作模型 GE-Act2.0，训练数据从 300 小时扩展至 3 万小时，陌生场景成功率提升约 2.6 倍，并开源全球首个具身世界模型专用评测基准 EWMBench；配套 Genie Sim 3.0 仿真平台（“自然语言造世界”）与 AGIBOT WORLD 开源数据集。

2026 年 10 月 29 日，智元在创新日发布 GO-3 Preview，把竞争维度从“模型能力”推进到“学习体系”。 技术上是自研 WRAM 架构（CoAE 控制导向视觉编码＋SVP 单步视觉规划＋KASO 知识对齐筛选），预训练数据约九成来自人类第一视角，官方称视觉 token 压至 DINOv3 的十六分之一、动作恢复误差降低 35%；体系上是“预训练（WRAM）—后训练（Genie Evolver：百台真机＋Genie Sim＋Genie Envisioner 2.0 世界模型）—部署后（Genie Master＋具身执行协议 EEP）”三段贯通，在真实便利店实现“无人在环”的持续改进，完整任务成功率从首日 28.6%升至第 5 天的 85.7%，单任务成功率较常规 SFT 提升 17.7 个百分点（企业披露口径）。该模型在 RoboDojo、RoboTwin、RoboColiseum 三个榜单登顶，并计划 2026 年底全面开源模型与 Genie Master。同场还发布精灵 G2 Max、皮下式关节液冷系统，并把首个具身智能原生操作系统灵渠 OS 1.0 开源捐赠给开放原子开源基金会。**判读：智元的护城河正从“全栈布局”转向“全链路学习速度”——前者是资源结构，后者是组织能力，后者更难被复制。**

判读： 智元证明了“数据闭环＋全栈协同”的中国路径——本体部署规模反哺数据，数据反哺模型，模型再提升本体竞争力，飞轮在国内玩家中最完整。其“部署元年”（2026）的宣言能否兑现为跨场景复制，是 200 亿级估值站稳与否的关键。

4.2.2 银河通用：仿真合成数据路线的“重资产”打法

银河通用选择了一条与国际主流不同的数据路线：**以百亿级仿真合成数据基建为核心壁垒**，其星脑采用“大脑—小脑—动作控制”三层架构。商业侧绑定强产业股东：2026 年 3 月完成 25 亿元融资（国家人工智能产业投资基金、中国石化、中信投资控股、上汽金控、亦庄国投等），投后估值超 200 亿元，累计融资约 70 亿元，居国内具身智能领域前列——国家队与产业资本（能源、制

造、零售）的进入，为其打开了场景想象空间。央视春晚指定具身大模型机器人的曝光，则完成了全民认知卡位。

判读： 银河通用代表了“国家队资本＋产业场景＋仿真数据”的三位一体模式，其考验在于仿真数据的真实泛化能力兑现与商业闭环从“演示”走向“规模交付”。

4.2.3 星海图：开源 VLA 与“一脑多形”

星海图成立于 2023 年 9 月，是 2026 年估值上升最快的公司：2026 年 2 月近 10 亿元 B 轮、4 月近 20 亿元 B+轮，两个月融资近 30 亿元，投后估值突破 200 亿元。其技术标签是**EFM-1 双系统架构实现“一脑多形”**（同一大脑驱动多种本体形态），以及体系化的开源节奏：2025 年 8 月开源 GO VLA（突破当时 SOTA）；2026 年 1 月开源全球首个“开箱即用”VLA 模型 GO Plus；2 月开源垂类模型与轻量化小模型——密集发布背后是研发强度的跃升（据其 CFO 披露，半年研发费用相当于此前总和的数倍）。

判读： 星海图代表了“研发密度＋开源生态”的硅谷式打法在中国的落地。其风险同样直接：年轻公司的技术叙事能否转化为真实交付，决定 200 亿估值能否站稳。

4.2.4 千寻智能：以“统考第一”立身的 VLA 新锐

千寻智能成立于 2024 年 1 月，2026 年 2 月完成近 20 亿元融资（顺为资本、云锋基金联合领投）、4 月再获 10 亿元——30 天融资 30 亿元刷新行业纪录，估值超百亿元（媒体报道称其成为国内第三家达到 200 亿元估值的非上市具身智能公司）。其技术代表作是 Spirit v1.5 VLA 模型——**在全球机器人评测榜单 RoboChallenge 上超越 Physical Intelligence 的 π0.5 登顶**，为中国 VLA 赢得标志性国际话语权；其“全身力控＋端到端大模型”融合路线构成差异化。

判读： 千寻展示了“模型能力立身”的纯粹路径，但正如行业共识所警示：通用能力不能停留在实验指标与演示视频中，必须进入高频、刚需、可复制的真实场景。

4.2.5 智平方：类脑架构＋大订单的“双重验证”

智平方是2026年融资节奏最快的具身智能企业之一（2025年半年7轮、一年累计12轮），2026年6月完成近50亿元系列融资、估值升至超200亿元（2月刚破百亿，4个月估值翻番）。其技术标签是**NeuroVLA类脑架构**——“皮层-小脑-脊髓”三层体系，分别承载任务规划、动态运动协调、毫秒级安全反射；商业侧与惠科签订3年1000台订单，成为“大脑＋交付”双重验证的代表。

4.2.6 它石智航与特色样本群

它石智航：2026年4月完成超4.5亿美元Pre-A轮融资、估值破200亿元，资金重点投向通用具身大模型AWE；其新一代模型已在汽车线束装配等产线运行，创始人陈亦伦的表述——“世界模型理解世界，物理AI改变世界”——代表了产业界对WAM路线的信心。

特色样本群（画像式观察）：

- **橡木果**：本能驱动“第三极”，Natus模型端侧250Hz、4毫秒闭环，用“物理交互机制”对抗“数据内卷”；
- **蚂蚁灵波**：“灵波通用大脑”预训练适配17个主流机器人品牌、20种构型，已在上海国大药房多家门店上岗——“一脑多机”商业化的先行样本；
- **穹彻智能**：全国50多个城市采集超10万小时真实世界数据，“人类数据驱动”路线的执行力代表；
- **大晓机器人**：“开悟”世界模型3.1直驱机器人，中国石油加油站便利店试运营——世界模型商业落地的早期样本；
- **深度机智**：PhysBrain 1.5以28项公开基准综合均分72.5登顶全球开源榜，开源新锐的代表；
- **宇树科技**：以运动控制能力完成全民认知教育后，IPO募资中拟将42亿元中的20亿元投入智能机器人模型研发——本体龙头向大脑延伸的路径；
- **腾讯（Hy-Embodied三模型矩阵）、小米（Xiaomi-Robotics-1，RoboCasa365平均成功率57.4%）、优必选（Thinker系列开源）**：平台巨头与大厂的代表卡位。

4.2.7 极佳视界（GigaAI）：世界模型原生的“数据放大”路线

极佳视界（北京极佳视界科技有限公司，常称极佳科技）成立于 2023 年，是本报告国内案例中唯一一家**从世界模型出发、再长出大脑与本体**的玩家——与 4.1.8 中 World Labs 的起点相同，但走向了相反的方向。关键事实（公开口径）：

- **技术底座**：世界模型平台 GigaWorld（2025 年 12 月开源 GigaWorld-0，同步开源全阶段训练与推理代码）＋通用具身大脑 GigaBrain。GigaBrain-0 是国内首个利用世界模型生成数据实现真机泛化的 VLA 基础模型（企业披露口径）；2026 年 8 月发布 GigaBrain-0.7（长时序多子任务自主操作）与 GigaBrain-WBC-0.5（首个面向人形全身控制的行为世界模型，配合 Maker L01 本体实现 1kHz 全身控制）；
- **评测与标准卡位**：GigaWorld-1 于 2026 年 3 月在世界模型评测基准 WorldArena 以 62.34 分登顶（媒体口径，称超越谷歌、英伟达、阿里等机构模型）；公司为通用世界模型北京市重点实验室（国内首个世界模型方向的省部级重点实验室）依托单位，并在 2026 年 8 月世界机器人大会提出通用世界模型 L1—L5 分级框架；
- **商业化三线并行**：家庭端 2026 年 5 月发布国内首个家庭通用机器人品牌“拾光 SeeLight”，首款拾光 S1 获真实家庭场景百台订单，在武汉光谷社区启动规模化运营；工业端 Maker H01 与一汽模具、阿里云验证拆垛搬运等任务，并与隆盛唯睿签约三年内落地 1000 台；自动驾驶端 DriveDreamer 系列深耕驾驶数据生成与场景仿真，服务近 30 家头部车企与芯片商——**这是最早创收的板块，构成“以驾驶数据生成养具身研发”的独特现金流结构**；
- **资本面**：2025 年完成 A 轮系列 4 笔累计 5 亿元（达晨财智、华控基金等，华为哈勃 2025 年 11 月入股）；2026 年 3—6 月连续完成 Pre-B（近 10 亿元）、B1（近 15 亿元）、B2（10 亿元）三轮融资，累计约 35 亿元，投后估值突破 200 亿元；据公开报道，Pre-IPO 轮（投前约 200 亿元）临近交割，公司拟冲刺港交所、意在成为“世界模型第一股”。

战略意图判读： 极佳视界的样本价值在于回答了 3.6 留下的那个问题——当通用基模收编语义与规划、世界模型退守“物理校验器”之后，一家世界模

型公司靠什么保持独立？它的答案是**把世界模型的比较优势（生成高泛化交互数据）产品化为“数据放大”引擎，再向下游延伸出自有大脑与本体**，形成“模型生成数据—机器人执行—场景反馈—再训练”的闭环。这与本报告对世界模型价值的定位（数据引擎＋物理校验器）完全一致，也构成与 World Labs 的“自然实验”：**同样从世界模型出发，World Labs 向上卖给算力巨头，极佳视界向下长成本体公司——两条路径的分化本身，就是“世界模型公司终局”这一问题的最好教材。**其 DriveDreamer 先行创收的结构，在国内大脑公司中也是少见的“先有收入、再谈叙事”。

风险提示： “三榜夺冠、全球唯一”等表述均为企业及地方宣传口径，WorldArena 为新兴第三方基准、权威性尚在建立中；家庭百台订单的规模运营刚起步，实际留存率与复购数据未披露；投后 200 亿元估值对应的大脑层直接收入同样处于早期——它与所有百亿俱乐部成员一样，需要用未来三年的交付数据证明估值。

4.2.8 DeepSeek：大厂入场的“拼装路线”

在 OpenAI（通用基模直接下场）、谷歌（平台闭环）、PI（纯平台输出 API）之外，DeepSeek 走出了第三条路：**不造整机、不做开放 API 平台，而是用投资把技术栈拼成完整闭环**。三笔动作（据公开报道口径）：

- **身体**：2026 年 8 月，出资约 1.41 亿元参与宇树科技科创板 IPO 战略配售（锁定 36 个月），双方签署联合研发备忘录，约定在通用人工智能研发、高性能通用机器人、AI 大模型三个方向深度合作——宇树的数万台真机保有量，是稀缺的物理交互数据入口；
- **大脑与眼睛**：10 月，拟领投 Winverse AI 种子轮（融资规模数千万美元、投后估值约 4 亿美元，美团龙珠参投）。创始人胡瀚为前腾讯混元多模态大模型负责人、Swin Transformer 主要作者、ICCV 2021 Marr Prize 得主，方向覆盖视觉基础模型、VLM 与具身基础模型；
- **小脑与运动神经**：10 月中旬，拟投资前逐际动力人形机器人基础模型负责人俞涛的创业项目（前后两轮估值分别约 7000 万美元与 2 亿美元），俞涛在逐际动力从零搭建了运动模仿与全身跟踪系统。

战略意图判读： 梁文锋此前公开的 AGI 路线图将具身智能放在第六阶（语言模型→思维链→Agent→持续学习→自我迭代→具身智能）并视其为“苦活”，2026 年下半年的连续出手标志着时间表被迫前移——GPT-6 Astra 证明了“不懂具身也能做出泛化的具身模型”，Coding 能力的见顶也使大模型公司需要新的叙事与市场承接。**“基座（语义规划）＋宇树（本体与真机数据）＋Winverse（视觉与世界模型）＋运动控制团队（全身控制）”——这套拼图若走通，DeepSeek 将同时握住通用大脑与物理小脑两层的入口**；其下一步关键动作，大概率是推出机器人 API 并完善数据回流条款。对国内大脑创业公司而言，这意味着一个残酷的问题：当 DeepSeek、OpenAI、谷歌同时成为“大脑层”的整合者与竞争者，独立大脑公司的生存空间在哪里？

4.2.9 2026 年四季度的三个新变量：高德、地平线、炎曦动力

中国梯队在一个月内新增了三类此前没有的样本——**互联网大厂的基座下沉、芯片公司的框架开源、初创公司的“先模型后芯片”**。三者规模都不大，但分别改写了不同的竞争假设。

高德 λ0（阿里体系）：大厂基座下沉到“全身”。 10 月 23 日，高德与阿里 ATH 事业群发布 λ0，定位“首个基于开放世界行为的全身具身基座模型”。其真正的差异化不在模型结构，而在数据工艺：采集者佩戴消费级动捕与颈挂相机、机器人本体不参与，构建 500 小时 HumanVerse-500 数据集，单条示范采集从约 100 秒降到 29 秒；训练分“手部交互—全身协调—人形本体”三阶段完成人机对齐。结果上，SIMPLE 基准 90.0%、真机四任务平均 62.5%、较最强基线提升 22.5 个百分点（技术报告 arXiv 2610.18950），硬件配置代码开源、权重未发布。**判读：λ0 的意义是“数据成本结构的改写”——当人类行为可以“随走随采”，真机遥操的数据服务商与只做局部动作 VLA 的团队同时被降维。这与 4.1.3 中 PI 的 API 化、5.7 中的中间层挤压是同一件事的三个侧面。** 风险同样明显：无本体采集缺少力触觉与末端执行器约束，跨本体迁移的物理精度仍要靠真机数据补足。

地平线 HoloBrain-0：把“具身先验”写进架构，并以开源换生态。 2026 年开源，提供 0.2B（GroundingDINO）与 1.1B（Qwen2.5-VL）两版，核心是把

多视角相机参数与机器人 URDF 运动学结构显式注入模型，以“统一基座坐标系”对齐空间表征、“混合相对动作空间”兼容不同本体；在 RoboTwin 2.0 的 50 项任务上平均成功率超过 90%，0.2B 版本可在 RDK S100 端侧以高于 5Hz 运行，配套 RoboOrchard 数据—训练—部署工具链。**判读：这是“芯片厂商做大脑框架”的典型路径——不追求模型最强，而追求“让模型在自己的芯片上最好跑”。它与英伟达 Isaac/GR00T 的逻辑同构，但选择从开源切入，目标是把国产端侧算力变成开发者的默认选项。**

炎曦动力 Paimon-0：成立三个月，先做模型、再做芯片。 2026 年 7 月成立于深圳，10 月 26 日发布 Paimon-0——分层异步 WAM+逆动力学模型，通过 query token 蒸馏未来状态的核心表征（物体 3D 位姿、语义、可操作性），把单个动作块推理从约 3000 毫秒压到 300 毫秒；在 RoboColiseum（0.7754）、RoboCasa365（58.1）、VLA-Arena（0.73）及 IROS Workshop 双臂策略挑战赛登顶，预训练主要使用合作方 EgoScale 的第一视角人类数据，自研芯片计划 2028 年交付。**判读：这是一条罕见的“反常识”路径——通常芯片公司先卖芯片再补软件，它先把模型做到榜一，再用模型定义芯片规格。若 2028 年芯片如期交付，它将同时握有“算法—工具链—算力”三层，成为端侧大脑的垂直整合者；若模型迭代跟不上，则会退化为一家普通的芯片初创公司。**

三个新变量的共同指向： 中国具身大脑的竞争，正从“谁的本体更强、谁的数据更多”，转向**“谁能把模型、框架、算力、网络中的任意两层咬合起来”**。单点能力的窗口在关闭，组合能力的窗口在打开。

4.3 格局综合研判

第一，四类商业模式的竞争结构。 平台生态型（英伟达、腾讯）、大厂分层型（谷歌、小米）、纯大脑型（PI、Skild、国内大脑派）、垂直整合型（Figure、智元、银河通用）四种模式同台。四类模式对“大脑”的定义与变现方式截然不同——**这提示我们：当前行业还没有出现“唯一正确”的商业模式，模式选择本身就是最大的战略变量。**

第二，估值逻辑的重构已经发生，但定价与收入的裂口在扩大。 2026 年资本市场对具身智能的估值，正从“硬件制造”逻辑（对标产能与订单量）转

向“AI SaaS”逻辑（对标数据资产广度与闭环能力）；智谱、MiniMax 等大模型公司在港交所的高市销率表现，为一级市场提供了“高增长可消化高估值”的参照系。但必须警惕硬币的另一面：**截至 2026 年 10 月，国内剔除已上市企业后仍有 29 家估值超百亿元、其中 14 家进入 200 亿元阵营；而具身大模型的直接收入尚处早期**——PI 以约 200 亿美元洽谈估值保持零收入，Skild 以 1 亿美元 ARR 支撑 140 亿美元估值，两者的定价逻辑并不相同：前者买“能力曲线的期权”，后者买“已跑通的现金流”。**同一赛道出现两套估值语言，本身就是分化的开始。**二级市场已提前演示了回调的样子：宇树科技 8 月 19 日上市首日收盘市值约 3418 亿元、盘中一度触及 4400 亿元，10 月 15 日盘中最低较首日高点回撤约 63%，随后七个交易日又反弹逾 65%——**两个月内近 3000 亿元的市值摆动说明，当估值锚从“稀缺性”切换到“交付确定性”时，波动本身就是定价机制的一部分。** GGII 与京东联合发布的报告给出了更冷的判断：未来三年国内具身智能赛道约 15%以上的初创企业将面临倒闭或被并购。

第三，中美阵营的结构性差异。 美国阵营在“平台（英伟达）、大厂分层（谷歌）、纯大脑（PI）、垂直整合（Figure）”四个生态位之外，又补上了两个重量级变量：**特斯拉以“车—机器人共用数据与芯片”把垂直整合推向极致，成为自研阵营中体量最大的玩家；学界巨头（World Labs、AMI Labs）则以资本化的研究实体为世界模型路线提供“学院派”期权。** 中国阵营则呈现“全栈自研为主、路线多元并行、开源评测密集卡位”的另一种结构——10 余家 200 亿/百亿级玩家中，多数同时握有本体与大脑。这一差异的深层原因在于：**中国的场景纵深与制造业供应链，使“软硬一体+场景绑定”成为更短的变现路径；而美国的 AI 人才密度、资本市场深度与巨头并购能力（AMD 收购 World Labs 即为例证），支持更纯粹的“大脑层”创业与更长的技术周期下注。** 两种结构的优劣，要由未来三年的交付数据来裁决。世界模型公司的终局分化为这一判断提供了注脚：同样以世界模型立身，World Labs 被算力巨头并购，而极佳视界向“世界模型＋大脑＋本体”四位一体延伸以求独立（见 4.2.7）——两条路径孰优孰劣，本身就是未来三年最值得跟踪的对照实验之一。

第四，通用基模入场后，格局正在从“四类模式”走向“三层分工”。 2026 年 9—10 月的密集事件（GPT-6 Astra 实测、Argon 三层分工、Pi API 平

台化、DeepSeek 拼装入场）指向同一个结构性变化：**认知与规划层向通用基模集中（OpenAI、谷歌、DeepSeek），物理执行层由 VLA/WAM 专用模型与端侧力控承接（PI、英伟达、国内全栈厂商），现场交付与运维沉淀为独立的 DeployCo 环节**。10 月下旬至 11 月的第二波进一步补上了第三个层次：**基础设施层**——地平线 HoloBrain-0（框架＋端侧芯片）、炎曦动力 Paimon-0（模型＋自研芯片）、中国移动灵韬智脑（云脑＋边脑＋专网）、眸深智能深枢（Infra＋端侧模组）几乎同时出现，说明“大脑”不再只是一个模型，而正在变成“模型＋工具链＋算力＋网络”的组合品。夹在中间、只做“API 接入＋方案包装”的薄层主体，将最先承受挤压——这一变化的产业逻辑与应对，详见 5.7 与 7.1。

表 4-2 全球具身大脑代表企业格局总览（截至 2026 年 11 月，公开口径）

企业	路线/架构	商业模式	关键数据锚点	主要验证场景
英伟达	GR00T N1.7/N2＋Cosmos 3	平台生态	EgoScale 20854 小时；2000＋仿真资产	全球开发者生态
Google DeepMind	Gemini Robotics 2＋Argon＋ER 三层分工	大厂分层/云端 API	Argon ERQA 87%；Real2Sim 把仿真构建从周压到小时	生态伙伴评测、Spot 真机演示
OpenAI	GPT-6 Astra（通用多模态基模）	通用基模下场	RoboDojo 2100 次测试 22.48%，超 40 个专用策略	第三方真机实测、宇树接入测试
Physical Intelligence	π0.7（VLA/WAM 融合探索）＋Pi API	纯大脑授权→平台化	累计融资约 21 亿美元；D 轮洽谈估值约 200 亿美元；Pi API 托管微调	开源 openpi、Chef/Almond 等 API 客户
Figure AI	Helix 02/2.5	垂直整合	Index 1600 万视频/4.4 万周活；陌生民宅 56%	家庭、物流
Skild AI	Skild Brain	纯大脑＋收购硬件	融资 14 亿美元；付费客户 60＋、ARR 超 1 亿美元、累计确认收入 5000 万美元	多形态机器人、仓库/工厂/餐饮
Dyna Robotics	Dyna-2（WAM）	全栈软硬一体＋ROI 闭环	鼎泰丰全美铺开；18 小时 1590 条/合格率 93%；3 天回本	餐饮、洗衣、酒店、物流
特斯拉	FSD 端到端复用（单一统一神经网络）	垂直整合（车＋机器人共用数据与芯片）	FSD 累计里程百亿英里量级；V3 称“已为量产做好准备”；Fremont 下线首批量产质量机；设计年产能约 100 万台	工厂作业测试、Optimus Academy

企业	路线/架构	商业模式	关键数据锚点	主要验证场景
智元	GO-3 Preview（WRAM）＋Genie Evolver/Master	全栈自研＋开源	5000 台下线；三榜第一；便利店 5 天 28.6%→85.7%；年底开源	工厂、便利店、商业服务
高德（阿里）	λ0（全身移动操作基座）	大厂基座下沉	HumanVerse-500 共 500 小时；采集效率 3.4 倍；真机四任务 62.5%	实验室、茶水间、访客中心
地平线	HoloBrain-0（开源 VLA 框架）	芯片＋框架开源	0.2B/1.1B；RoboTwin 2.0 50 任务>90%；端侧>5Hz	开发者生态、端侧部署
炎曦动力	Paimon-0（分层异步 WAM＋IDM）	先模型后芯片	动作块推理 3000ms→300ms；四榜第一；芯片 2028 交付	机械臂、双臂操作
银河通用	星脑三层架构	软硬一体＋场景	累计融资约 70 亿元；估值超 200 亿元	零售、能源
极佳视界（Giga AI）	GigaWorld 世界模型＋GigaBrain VLA	世界模型原生→四位一体	2026 年 3—6 月融资约 35 亿元、投后超 200 亿元；拾光 S1 获真实家庭百台订单	家庭（拾光）、工业（Maker）、驾驶仿真（DriveDreamer）
星海图	EFM-1“一脑多形”	大脑＋开源	两月融资近 30 亿元；估值约 200 亿元	开发者生态
千寻智能	Spirit v1.5（VLA）	全栈	30 天融资 30 亿元；RoboChallenge 登顶	制造、服务
智平方	NeuroVLA 类脑三层	大脑＋交付	融资近 50 亿元；惠科 3 年 1000 台	3C 制造、新零售
它石智航	AWE 通用模型	大脑＋工业	Pre-A 超 4.5 亿美元	汽车线束装配
橡木果	Natus（本能驱动）	端侧操作基座	250Hz 端侧/4ms 闭环	精细接触操作
DeepSeek	基座＋投资拼装（宇树/Winverse AI/运控团队）	大厂拼装路线	1.41 亿元战配宇树＋两笔具身投资	尚未发布机器人 API
World Labs（AMD）	Atlas/Marble 世界模型＋R2S2R 仿真引擎	空间智能产品＋被芯片巨头收购	10 亿美元融资（估值约 50 亿美元）；AMD 约 82 亿美元收购；零真实数据部署连续运行 1 小时	创意 3D 生成、机器人仿真训练
AMI Labs	JEPA 架构世界模型	学院派研究型	种子轮 10.3 亿美元（欧洲最大，投前估值 35 亿美元）；16 秒视频预测 10 秒场景	三至五年期研究承诺，暂无产品

4.4 案例启示录：六个可迁移的战略规律

解剖二十余家玩家之后，以下六条规律具备跨案例的迁移价值——它们不仅解释过去，也预告未来格局的演变方向。

启示一：路线选择即命运，但“选择权”比“选择”更重要。 π 系列坚持纯大脑、Figure 坚持垂直整合、英伟达坚持平台化——三者都成功了，因为各自的选择与自身资源禀赋（人才密度/整机能力/算力生态）匹配。反例是资源与路线错配者：没有部署基数的团队做纯大脑（飞轮转不动），没有场景纵深却做垂直整合（两头不靠）。**战略判断的第一问不是“哪条路线对”，而是“我有什么、哪条路线会用上我有的”。**

启示二：数据策略是比模型架构更深的护城河。 架构可以被追赶（开源已抹平大部分架构差距），数据管线不能：银河通用十年级的仿真基建、Figure 的全球众包网络、穹彻的全国采集体系、智元的部署基数，都是需要以年为单位构建的重资产。**评估大脑公司，先问数据从哪来、独占性多强、边际成本多少——三个问题答不清楚，架构再先进也是沙滩上的城堡。**

启示三：开源是弱者的武器、强者的护城河。 π0 开源成就了 PI 的社区地位，GO/GO 系列开源成就了中国团队的生态卡位——开源的本质是“用模型换部署、换数据、换标准话语权”。但开源同样有层次：开源旧版本（保持代差）、开源框架不开源权重、开源权重不开源数据管线。**判断开源策略的真伪，看它开源的部分是否恰好是自己已经不靠它赚钱的部分。**

启示四：“演示能力”与“交付能力”之间隔着一场产业化的死亡谷。 Helix 在 30 栋民宅 56%的成功率是演示的巅峰，而商业交付需要 90%+的可靠性、可复制部署与售后体系。案例中真正完成“双重验证”的只有少数（智平方的千台订单、蚂蚁灵波的多门店上岗）。**2026 年的行业分水岭测试题只有一道：合同额与复购率，不是视频播放量。**

启示五：资本密度决定生存期，但不决定胜负。 百亿俱乐部成员都获得了 3 年以上的资本跑道，但跑道长度不等于到达终点——PI 的零收入证明资本可以给“能力曲线”定价，也可以在曲线放缓时撤走定价。**对企业的建议：按**

“飞轮最小启动速度”规划融资节奏，把每轮融资都花在缩短数据闭环周期上，而非花在叙事上。

启示六：生态位卡位先于产品竞争。 英伟达用参考平台与开源运行时锁定开发者，智元用 EWMBench 锁定评测话语权，DeepMind 用 API 锁定企业接入——这些动作的投入远小于模型训练，却决定了行业规则由谁书写。**大脑战争的终局形态不是“谁的模型最强”，而是“谁的接口被默认”。**

4.5 资本地图：谁在为大脑下注

案例背后是一张正在成形的资本地图。理解“谁在投、投什么、为什么投”，比记住融资金额更有价值。

第一类：国家资本与产业资本（中国特征最鲜明）。 银河通用 2026 年 3 月的 25 亿元融资中，出现了国家人工智能产业投资基金（国家大基金三期）、中国石化、中信投资控股、上汽金控、亦庄国投等——这是国家大基金首次出手具身智能赛道（据经济参考报报道）。产业资本（能源、制造、汽车）的进入逻辑不是财务回报，而是**场景与供应链的战略协同**：中石化看能源场景，上汽看制造与汽车产线，亦庄国投看区域产业集群。这一结构意味着：**在中国，能否拿到“场景型股东”，正在成为大脑公司融资竞争力的组成部分。**

第二类：全球顶级 VC 与科技巨头战投。 Physical Intelligence 的投资人名单包括 Thrive、Lux、Sequoia、Khosla、贝佐斯、OpenAI，以及谷歌系 CapitalG 与英伟达 NVentures——一个耐人寻味的现象是：**互为竞争对手的巨头，同时下注同一家“大脑公司”**。这既是分散风险，也是对“大脑层将成为公共基础设施”的共识下注。

第三类：二级市场与 IPO 通道。 2026 年被视为人形机器人主机厂的 IPO 元年（多家媒体报道口径）；宇树科技 IPO 募资中拟将 20 亿元投入智能机器人模型研发；智谱、MiniMax 等大模型公司在港交所的高估值表现，为一级市场提供了“高增长可消化高估值”的参照系。**一级市场的估值逻辑，正被二级市场对 AI 公司的定价方式反向塑造。**

第四类：地方基金与“以投带引”。 北京、上海、深圳等地设立千亿/百亿级产业基金，普遍采用“基金＋基地＋场景”的组合招商。这一模式的效率优势明显（决策快、配套全），但也带来一个结构性问题：**同一家头部大脑公司往往同时接受多个地方基金的投资，形成“多地布局、资源分散”的隐忧**——对企业的忠告是：拿地方钱可以快，但要算清“资源承诺”与“管理成本”的兑换率。

一个必须记录的风险信号。 2026 年前三季度，国内具身智能一级市场维持高位但已现分层：上奇产业通监测显示 1—9 月共 557 笔融资、披露金额约 2896 亿元（557 笔中 346 笔披露金额，未披露不计入），月均超过 60 笔，3 月以 97 笔为单月峰值，8 月、9 月回落至 44 笔；高工机器人统计口径为 437 起、约 2032 亿元，同比分别增长 146.07%与 317.39%——**两套口径的差异来自统计范围与金额披露率，宜作节奏观察而非规模断言。** 估值端，剔除已上市企业后国内仍有 29 家估值超百亿元、其中 14 家进入 200 亿元阵营。**资本密度已经领先于商业验证两个身位——这既是产业加速的燃料，也是下一轮整合的伏笔；而 8—9 月融资笔数回落与“机构不再投具身”的市场传言同时出现，说明拐点可能先出现在融资节奏上，而不是出货量上。**

4.6 本章小结

十余个国际样本、国内三个梯队的代表玩家，成败的分野最后落在四个变量上，而不是模型榜单上。

第一，生态位是否清晰。 英伟达卖“造机器人的工具”，PI 卖“纯大脑的期权”，Figure 与特斯拉卖“垂直整合的产品”，银河通用卖“场景里的确定性”——每一种成功都对应一个清晰的生态位，而处境最尴尬的，恰恰是“什么都在做、什么都不突出”的主体。

第二，数据管线是否自持。 特斯拉的车队、Figure 的 Index、智元的量产下线、银河通用的仿真基建，本质都是同一件事：把数据获取能力握在自己手里。**演示可以外包，数据管线不能外包。**

第三，交付是否被第三方验证。 Dyna Robotics 之所以被本报告单列，不在于技术最先进，而在于它是少数给出可核查商业闭环数据的样本之一；同理，Figure 的家庭实测、智元的陌生场景成功率，价值都高于任何实验室跑分。**在叙事过剩的阶段，“被客户验证”是最稀缺的资产。**

第四，资本结构是否支撑长周期。 大脑是一个“先投入、后验证”的生意，估值越高，对交付节奏的要求越苛刻。本章记录的融资地图显示：钱已经不缺，缺的是能把钱转化为部署与数据的组织能力。

这四条，也是判断下一轮格局洗牌的基本标尺。

第五章 产业逻辑：在产业链条中理解大脑

本章导读 本章把大脑放回产业链条中回答四个本质问题：价值为什么向大脑集中——七段解构显示价值呈“双峰”分布，大脑与数据服务是价值高地；大脑的护城河机制是什么——数据飞轮与其背后的经济学特性决定了谁会被“数据饥饿”淘汰；商业闭环由谁付费——四类付费形态给出了从“卖模型”到“卖结果”的演进路径；大模型大厂的平台化正在改写什么——语义、仿真、微调、任务适配四层被收编之后，中间层主体面临出清，护城河公式转向“真实工况数据＋现场交付能力”。章末小结从价值层、机制层、结构层三个层次归纳全章结论。

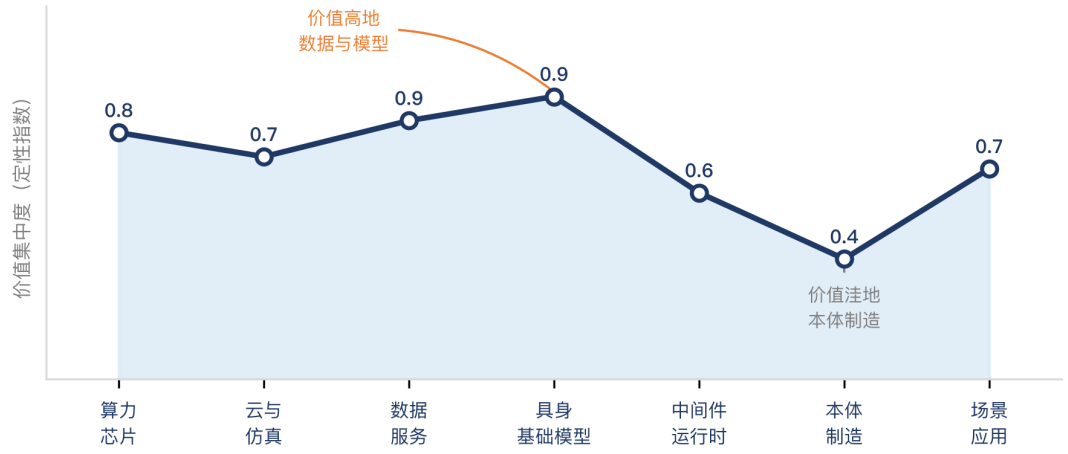
前四章回答了“是什么、从哪来、怎么做、谁在做”。本章把具身大脑放回产业链条中，回答三个更本质的问题：**价值为什么向大脑集中？大脑的护城河机制是什么？商业闭环由谁付费？** 产业分析如果不能落到“钱从哪里来、壁垒在哪里”，就只是技术综述。

5.1 产业链解构：大脑的上下游版图

具身大脑不是一个孤立的产品，而是一张价值网络的中心节点。以大脑为核心，产业链可分为七段：

价值分布呈“双峰”：大脑与数据服务是价值高地，本体制造是价值洼地

上游“基础设施”与中游“大脑＋闭环”两端集中价值，中段本体制造利润受挤压



资料来源：腾讯云×沙利文白皮书关于价值重心迁移的判断；指数为本报告定性评分，非市场统计。

图 5-1 产业链价值分布：双峰结构

（1）算力芯片层。 训练侧由数据中心级 GPU 主导（英伟达系），部署侧是边缘计算平台的竞争——英伟达 Jetson Thor（2070 FP4 TFLOPS、128GB 统一内存）已成为国际人形机器人的准默认机载平台；国产侧，华为昇腾支撑起独立算力路线（国家地方共建人形机器人创新中心与华为共建的实训场即基于昇腾研发 VLA）。WAM 路线对算力的结构性抬升（视频生成式训练），将进一步强化“大脑竞争即算力竞争”的格局。

（2）云与仿真平台层。 仿真从研发工具升级为“数据工厂”：英伟达 Isaac Sim/Omniverse、智元 Genie Sim 3.0、国内光轮智能（国际主要具身智能团队超80%的仿真资产与仿真合成数据来自光轮——企业披露口径）等，共同构成合成数据的基础设施。**这一层是大脑的“面粉厂”——面粉质量决定面包上限。**

（3）数据服务层。 遥操作采集、真机数据服务（穹彻智能超 10 万小时）、人类视频众包（Figure Index 模式）、数据清洗与标注，已形成专业梯队。2026 年，光轮智能完成 10 亿元 A++/A+++轮融资，成为具身数据领域首个独角兽——数据基础设施的独立估值时代已经到来。

（4）具身基础模型层（本报告核心）。 即“大脑”，详见前章格局分析。

（5）中间件与运行时层。 Isaac OS（基于 Holoscan 的开源运行时，连接 AI 模型、传感器、控制与硬件）、各家的部署框架——这一层看似不起眼，实则是“事实标准”的埋雷区：**谁定义了模型与硬件的接口规范，谁就在生态战中占据先手。**

（6）本体层。 人形、机械臂、移动操作平台等。2025 年全球人形机器人出货约 1.8 万台，中国厂商贡献主要部分；中国已连续 12 年居全球第一大工业机器人市场，商用服务机器人出货量占全球 84.7%（IDC，2024）。本体层的竞争逻辑是制造业逻辑——成本、量产、交付，中国优势明确，但价值集中度低于两端。

（7）场景应用层。 制造（柔性产线、装配、上下料）、物流（分拣、搬运）、商业服务（零售、餐饮、药房）、能源（巡检）、家庭（远期）。腾讯云×沙利文白皮书指出：上游零部件决定性能上限与成本曲线，中游本体制造

决定工程化与交付效率，下游应用决定规模化部署和商业闭环质量——而**产业价值重心正从硬件性能迁移至“场景数据—训练迭代—部署运营”的闭环能力**，闭环所在，正是大脑环节的所在。

5.2 价值分布：大脑为何是“复利机器”

大脑环节之所以成为产业链的价值高地，源于四个经济学特性：

特性一：边际成本趋零的软件特性。 模型一旦训练完成，复制到第 100 万台机器的边际成本接近于零；而本体的每一台都要付出完整的物料与制造成本。**软件经济学对制造业经济学的降维，正是 PC 时代“WinTel 联盟”中微软占据价值顶端的同一逻辑。**

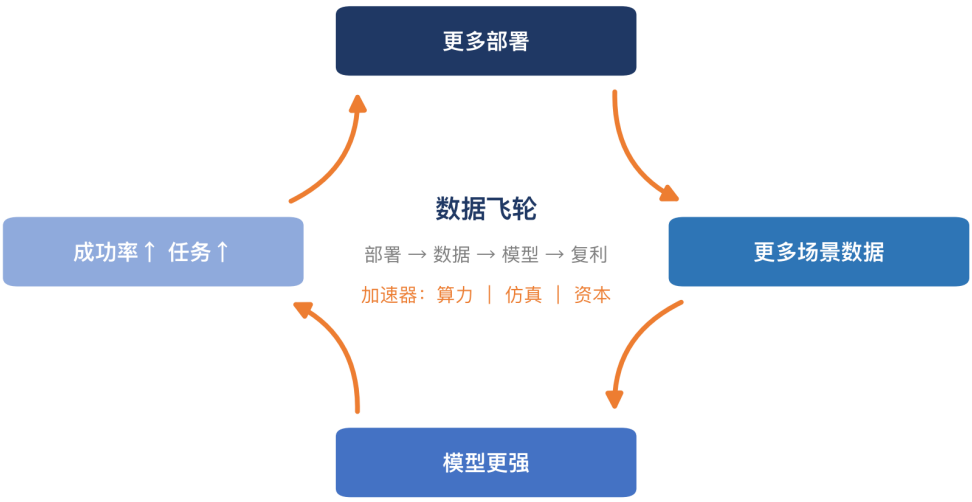
特性二：数据飞轮的复利特性。 大脑是唯一能把“部署”转化为“资产”的环节：机器人每多工作一小时，就多产生一小时交互数据；数据反哺模型，模型提升成功率与任务范围，进而带来更多部署。本体卖出一台即完成价值兑现，大脑卖出一台才开始积累价值——**前者是收入曲线，后者是复利曲线。**

特性三：跨场景复用的杠杆特性。 同一个通用大脑，经过适配可以覆盖制造、物流、商业等完全不同的场景——研发投入被摊薄到多个市场；而专用设备的研发投入只能锁定单一场景。当“一脑多机”从愿景变为工程现实（Gemini Robotics 1.5 的跨本体迁移、蚂蚁灵波适配 17 品牌 20 构型），这一杠杆才真正开始兑现。

特性四：生态定义权的网络特性。 大脑层一旦形成事实标准，将同时锁定上游（数据格式、算力平台）与下游（本体接口、开发者生态），产生类似操作系统的网络效应——越多人用→数据越多→模型越强→越多人用。英伟达的全栈布局，本质上是在机器人大脑层“提前锁定生态位”。

大脑的护城河是一个飞轮：部署越多，模型越强

飞轮转不动的公司会被“数据饥饿”淘汰；算力、仿真、资本是三重加速器



资料来源：本报告研究整理。

图 5-2　具身大脑的数据飞轮

同时必须直视大脑环节的三个价值陷阱：

- **陷阱一：飞轮需要最低启动速度。** 部署量不足的大脑公司，飞轮转不起来，数据不够、模型不够强，会在竞争中“被饿死”——不是因为技术不行，而是因为没有足够的机器喂养它。“一个大脑跑通，其他的大脑未必 over，但数据飞轮转不动的大脑一定危险。”
- **陷阱二：授权费假设的脆弱性。** “每台机器交一次大脑授权费”建立在“本体厂做不出够用的大脑”之上。Figure、智元等垂直玩家已经证明本体厂可以自研；本体厂之间的价格战会外溢为对大脑成本的极限压榨。
- **陷阱三：价值兑现周期长于资本耐心。** 大脑的能力曲线（尤其泛化与可靠性）尚在早期，而一级市场的估值已经预支了 3—5 年的乐观情景。**估值是提前收的学费，交付才是补交的作业。**

5.3 商业闭环：谁在为大脑付费？

产业逻辑的最终检验是付费行为。当前真实发生的付费结构，可归纳为四类：

第一类：制造业的柔性化改造（当前最大现金流场景）。 传统自动化解决“大批量、高重复”场景，而制造业的真实痛点是“多品种、小批量、频繁换线”。具身大脑的价值在于把换线成本从“重新编程/重新示教”降为“语言指令＋少量演示”。已有可核查的落地证据：无锡某汽车工厂通过“柔性产线＋具身 AGV”模式将产线重构时间从 72 小时压缩至 3 小时；智平方对惠科的 3 年 1000 台订单；它石智航的汽车线束装配；国地中心“灵龙”系列机器人进驻延锋完成零件火焰处理、齐套分拣等工位的实景验证。**制造业付费的理由不是“机器人像人”，而是“换线成本下降”——这是大脑环节当前最扎实的价值主张。**

第二类：商业服务的标准化作业（规模化验证场景）。 零售、药房、餐饮的拣选、上架、递送等半结构化任务。上海国大药房多家门店的取药机器人（蚂蚁灵波大脑）、中国石油加油站便利店的大晓机器人试运营，标志着该场景从演示进入试运营阶段。商业服务的付费逻辑是人力替代＋服务一致性，对成功率的容忍度高于工业、低于家庭。

第三类：训练场与数据基础设施（政策驱动的当下市场）。 2025 年 1 月，全国首个异构人形机器人训练场在上海张江启用——“实训场训练—模型迭代—工业现场部署”的迁移路径已跑通样本（“灵龙”机器人进驻延锋）。各地方政府主导的训练场、数据采集中心，本质是**用公共投资为大脑产业铺设数据高速公路**，这是中国特有的市场规模与节奏来源。

第四类：模型授权与开发者生态（远期期权）。 云端 API（Gemini Robotics-ER 已开放）、开源企业服务（宇树、优必选的开源模型＋企业定制）、平台许可（Isaac GR00T Blueprint）——当前收入贡献最小，但代表大脑环节的终极商业模式。**2026 年四季度，这一类的“期权价值”首次出现两种可对照的兑现方式：**Skild AI 以“模型＋部署＋集成”的混合交付，在正式商业部署 10 个月内做到付费客户超 60 家、ARR 突破 1 亿美元、累计确认收入 5000 万美元（公司披露口径）；Physical Intelligence 则以 Pi API 把微调与托管推理平台化，仍未产生营收，估值却从 110 亿美元谈到约 200 亿美元。**同一环节、两种定价：一个卖“已实现的现金流”，一个卖“未来的抽水权”。二者的并**

存说明，大脑层的商业模式尚未收敛，投资人不妨把“ARR/估值”与“平台调用量/估值”当作两把不同的尺子。

一个关键的商业事实必须强调： 具身大模型的直接收入尚处早期，但**间接收入已现规模**——越疆2026年中报显示营收翻倍、具身相关收入激增20倍（公司公告口径）；光轮智能2025年营收增长10倍、2026年一季度单季预计超过2025全年。这揭示了一条产业规律：**大脑的价值首先通过“带动卖得动的环节”（本体、数据服务、集成）变现，然后才逐步独立计价。** 对投资者的含义是：评估大脑公司，要区分“大脑带来的溢价”与“本体/服务带来的收入”，防止两类现金流被混同定价。

5.4 供给结构：三大瓶颈的现实约束

产业上行期更需要正视约束。当前制约大脑产业化的三大供给瓶颈：

算力瓶颈。 WAM与世界模型的训练算力需求显著高于VLA，且推理侧面临“500ms思考 vs 100ms动作”的实时性鸿沟。英伟达在训练与部署两端的双重主导地位，构成中国路线的结构性风险；昇腾等国产算力的成熟度，将直接影响国内大脑阵营的迭代速度。**算力自主，是具身大脑“中国方案”的前置条件。**

数据瓶颈。 千万小时级的涌现阈值（业内专家判断口径）与当前头部玩家3万小时级的有效数据之间，存在两个数量级的缺口。人类视频与仿真合成的规模化供给是行业解题方向，但“视频的物理保真度”与“仿真的现实鸿沟”都尚未完全解决。

5.5 区域与政策格局：大脑产业的中国地图

大脑产业的空间分布，正在沿“人才—算力—场景”三要素重新绘制。结合已公开的政策与产业事实，中国大脑产业地图可归纳为四个梯队：

第一梯队：北京、上海、深圳——全要素高地。 北京依托大模型人才密度（智谱、小米、深度机智等大脑玩家聚集）与国家标准制定资源，占据“基座模型＋标准”生态位——2026年10月15日首批6项人形机器人国家标准发布、2027年4月1日起实施，标准资源正加速转化为“验收话语权”；上海以张江

为核心形成“训练场＋世界模型＋场景验证”的完整闭环（国地中心与华为共建昇腾算力实训场、蚂蚁灵波药房落地、浦东“十五五”规划明确“具身大模型＋世界模型”融合布局），并在标准体系与场景开放上双线先行（上海 AI 标准体系将世界模型标准列为关键技术标准；《加快发展新一代智能制造行动方案》明确 2028 年推动 8000 台人形机器人进入工厂）；深圳依托硬件供应链与产业资本，走“软硬一体＋快速工程化”路线，2026 年前三季度融资事件 104 笔居全国第二、披露金额约 1094 亿元居城市首位（上奇产业通口径，金额受单笔大额事件影响，宜作参考）。

第二梯队：杭州、苏州、无锡等——场景与制造纵深。 杭州以数字经济生态与民营资本（云锋、顺为系投资活跃）培育大脑新锐（深度机智、群核系生态）；苏州、无锡等制造业强市以柔性产线改造需求牵引工业场景大脑落地（无锡汽车工厂“柔性产线＋具身 AGV”将换线时间从 72 小时压缩至 3 小时是可核查样本）。

第三梯队：算力枢纽城市。 依托绿电与成本优势承接训练算力布局，是 WAM 时代算力需求跃升的直接受益者，但需要与“模型—数据”环节形成协作，避免沦为单纯的机房。

第四梯队：跟风招商城市。 招商清单高度雷同（都要“大脑企业＋本体工厂＋基金”），缺乏要素支撑，最可能形成低效重复建设。

政策工具箱的演进方向。 从已落地的政策实践看，有效工具正在从“补贴型”转向“要素型”：其一，数据要素供给（训练场、数据采集中心、数据产权规则）；其二，算力券与训练补贴（指向国产算力适配）；其三，场景开放（首用场景清单、责任豁免与保险机制）；其四，标准与评测（第三方验证机构）。**判断一个地方的产业政策是否专业，看它给的是“钱”还是“要素”——钱会催生演示，要素才能养出闭环。**

对地方政府的坐标建议：大脑产业的区域竞争不是“抢企业”而是“抢环节”——全国不需要 30 个大脑集群，但需要分工明确的数据基地、算力基地、评测基地与场景基地。按照“禀赋定环节”原则布局（详见 7.3），是避免区域内卷的唯一路径。

5.6 单位经济模型：一台机器人何时算得过来

产业逻辑最终要落到算术上。本章尝试构建一个简化的单位经济模型（TEE，Total Economic Equation），用以判断大脑环节的经济价值边界。需要说明：模型为示意性框架，参数取公开报道的行业区间，不代表任何具体企业的实际数据。

成本侧的四块构成。 一台具身智能机器人的全生命周期成本可分为：①硬件折旧（本体＋传感器＋算力，按 3—5 年摊销）；②大脑与软件的授权或摊销（若自研，则为模型研发费用摊销；若外购，则为按台/按年的授权费）；③部署与集成（现场调试、工装改造、系统对接，通常为硬件成本的 20%—50%）；④运维与人力（巡检、异常处理、人工兜底）。

收益侧的三个来源。 ①替代人力（按岗位年综合人力成本计算）；②效率与质量提升（换线时间下降、良率上升、停机减少）；③柔性溢价（小批量多品种订单的承接能力，这是传统自动化无法提供的增量价值）。

大脑在这张算式中的位置。 关键洞察是：**大脑不直接创造收益，它通过“降低部署与集成成本”“缩短换线时间”“提高单机可承担的任务种类”三个通道影响回报周期。** 由此可得三条产业结论：

- **结论一：越是多品种小批量的场景，大脑的相对价值越高。** 大批量单一品种场景中，传统自动化方案的单位成本更低，大脑的溢价空间有限；反之，换线频次越高，大脑“用语言指令替代重新编程”的价值越突出。
- **结论二：部署与集成成本是当前 ROI 的最大杀手。** 若部署成本占到总投入的四成以上，单机回本周期会被显著拉长——这正是“大脑标准化＋工具链完善”的产业意义：把每个项目的定制集成，变成可复制的产品化交付。
- **结论三：人工兜底比例决定了商业模式的可持续性。** 若一台机器人长期需要 0.5 个人力兜底，“替代一个人力”的账就不成立。行业报道中提及的部署成本下降与回报周期缩短（有分析称机器人部署成本自 2021 年下降约 35%、ROI 周期缩至 18—24 个月，单一来源，仅供参考），其可信度最终要靠企业披露的验收数据验证。

给采购方的实用建议： 在立项报告中同时给出三套口径——乐观（无人兜底、满负荷）、基准（少量兜底、80%负荷）、保守（含集成与培训的一次性投入）。**能用保守口径算得过来的项目，才是可以规模复制的项目；只能靠乐观口径成立的项目，本质上是在为产业发展付费。**

5.7 平台化与中间层挤压：大模型大厂改写分工

2026 年 9—10 月的四个事件——GPT-6 Astra 实测、Gemini 4 Argon 的 Real2Sim、Pi API 开放、DeepSeek 拼装入场——共同标记了一个结构性变化：**基础模型大厂不再只做“模型”，而是把机器人产业链的中台逐层收编。** 本节将其提炼为产业逻辑。

被收编的四层能力。 通用基模与其平台生态正在吃下：①**语义理解与任务规划层**（Astra 直接输出子目标与动作审查，RoboDojo 实测超 40 个专用策略）；②**仿真生成层**（Argon 把 MuJoCo 场景构建从“周”压到“小时”，仿真数据边际成本趋近于零）；③**微调与托管推理层**（Pi API 把“上传数据—托管微调—云端下发”做成标准链路，Chef Robotics 几十条示教隔夜出模型）；④**任务适配层**（据公开报道，Generalist GEN-1.5 等以少样本演示直接学习新任务、无需更新权重，传统微调的价值被稀释）。**通用基模够不到的，只剩物理力控闭环、毫秒级实时性与真实工况数据三块硬骨头。**

推理成本成为架构选择的显性变量。 实测数据显示：Astra 直接控制机器人平均单任务耗时 39.86 秒、50 个任务消耗约 11.3 亿 token（成本近万元）；本地部署的 π0.5 推理约 7 毫秒、无按次费用——两者实时性相差约 3 个数量级。**“大模型做规划、专用模型做执行”不只是能力分工，更是成本分工：**每一次大模型调用都要在“能力增益”与“token 账单”之间过秤。这将持续压制“全大模型大脑”的可行性，强化双系统与端侧小模型的地位。

新的分工框架：ModelCo 与 DeployCo。 行业由此出现一对新概念（DeployCo 由海外投资人提出，据公开报道）：**ModelCo**——训练与运营基础模型的公司，按 API 或授权收费；**DeployCo**——不做模型，专注场景交付、真机运维与故障数据回流的部署公司。风投界的新共识是：调用 API 的门槛趋零后，**“接入哪款大模型”不再是护城河，“机器人在产线跑了多少小时、沉淀**

了多少故障工况数据”才是。但本报告需要同时指出：ModelCo 与 DeployCo 只是现阶段的分工快照，而非终局——头部玩家（英伟达、谷歌、Dyna、智元）都在向全栈延伸，分工边界本身是流动的。

四类“中间层”主体的出清风险。 叠加巨头挤压与 2026 年下半年资本向头部集中的风向（上半年国内具身融资超 900 亿元、下半年机构上会项目明显减少），四类中间层主体率先承压（本报告基于公开信息的研判）：

- **套壳型“大脑”方案商：**本体采购公模硬件、调用大模型 API、以提示词工程包装行业方案——当模型公司直接提供微调＋托管推理，这一环节的壁垒趋近于零。存活路径：转型真机运维的 DeployCo，或攥紧独家场景数据；
- **传统规则型集成商：**ROS/PLC/示教器路线不会被整体替代，但涉及自然语言变更、异常恢复、长程调度的工位将被大模型编排吃掉毛利。存活路径：守住工业系统对接与安全认证，承担大模型失灵时的兜底责任；
- **单任务垂直小模型公司：**通用基模零样本/少样本能力持续追平垂直小模型，“基座每迭代一代、垂直差异化薄一分”的剧本正在重演。存活路径：放弃“自研完整基座”叙事，沉淀真实工况数据，把自研模块降级为对接通用大模型的领域适配层；
- **低质标注与通用仿真建模服务商：**Real2Sim 自动建模普及后，通用场景人工建模与按条计费的标注业务直接贬值。存活路径：转向难以仿真的高精场景（反光金属、核电检修等）、Sim2Real 误差补偿与数据血缘追溯。

上游也在收编：芯片巨头开始直接购买“大脑能力”。 挤压不只来自应用层的大模型大厂，也来自算力层。2026 年 9 月 28 日，AMD 宣布以约 82 亿美元股票收购李飞飞的 World Labs，李飞飞出任 AMD 执行副总裁兼首席科学家——这是“算力层反向购买大脑能力”的第一案。芯片巨头的动机清晰：**拥有世界模型，才能定义下一代 AI 工作负载需要什么样的芯片**（苏姿丰表态）；英伟达以 Cosmos＋GROOT 自建大脑生态、AMD 以并购补齐模型层，“大脑—芯片”的纵向整合由此成为巨头的共同选择。对独立大脑公司而言，这既是估值背书（大脑能力获得了算力层的真金白银定价），也是退出通道的改写——**“被芯片巨头收购”正式加入“IPO、被本体厂并购、被基模巨头整合”的退出菜单**（对投资决策的含义见 7.2）。

基础设施层的成形，让“挤压”与“机会”同时发生。 10 月下旬至 11 月出现的四个样本值得放在一起看：地平线开源 HoloBrain-0 并配套 RoboOrchard 工具链（框架＋端侧芯片），炎曦动力以 Paimon-0 的推理效率反向定义自研芯片规格（模型＋算力），中国移动以“300 亿参数慢脑＋40 亿参数快脑＋时延 30 毫秒的具身智能网”提供网联化底座（网络＋云），眸深智能以“深枢”Infra＋端侧模组做跨设备适配（中间件＋模组）。**四者的共同点不是模型更强，而是把“可复用的那一层”抽出来单独供给**——这恰恰是“薄壳中间层”的反面：薄壳卖的是项目，底座卖的是复用。对创业公司的含义是清晰的：如果无法向上成为模型方、向下成为交付方，第三条路是**成为某一层的“标准件”**——工具链、运行时、端侧模组、评测服务，都是尚未被巨头占满的缝隙。

本报告判断： 出清的是“薄薄一层中间业务”，不是具身智能的需求本身。未来真正值钱的，恰恰是大模型大厂不愿碰的“脏活累活”——真机部署与现场运维、工况故障数据回流、工业系统对接、安全力控合规、高精仿真资产、特殊行业准入资质。**护城河的公式已经改写：从“你接入了哪款大模型”，变成“你的机器人在产线跑了多少小时、沉淀了多少别人拿不到的故障工况数据”。**

5.8 本章小结

把大脑放回产业链条，本章得到三个层次的结论。

价值层：大脑是“复利机器”，但复利只对数据闭环成立。 七段解构显示价值呈“双峰”分布——大脑与数据服务在峰顶，本体制造在谷底；然而峰值不等于利润，只有把部署转化为数据、把数据转化为能力迭代的玩家，才能把“高价值环节”变成“高利润环节”。**没有闭环的大脑公司，估的是“位置”；有闭环的大脑公司，估的是“斜率”。**

机制层：护城河来自“别人拿不到的工况数据”，而非模型参数。 大脑的四个经济学特性（高固定成本、近零边际成本、网络效应、数据资产不可交易）共同指向一个结论：规模优势属于部署方，而不属于训练方。这也解释了为什么中国厂商普遍选择“软硬一体”——在数据主权尚未确立的阶段，先拿到身体，才能拿到数据。

结构层：中间层正在被两端挤压，分工在重写。 应用层的大模型大厂向上收编语义、规划、仿真与微调，算力层的芯片巨头向下购买世界模型能力（AMD 收购 World Labs），夹在中间的“薄壳”主体最先承受压力。**新的分工不再是“大脑—本体—集成”的三段式，而是“ModelCo—InfraCo—DeployCo”的三段式：模型方提供认知与物理能力，底座方提供工具链、算力与运行时，交付方承担现场运维与数据回流。中间环节只有三条活路——向上成为模型方，向下成为交付方，或者横向成为某一层的标准件。**

对下一章的衔接： 上述三层结论共同决定了未来五年的推演起点——价值结构向大脑与数据服务集中（6.1）、路线向 WAM 收敛（6.2）、格局向分层演化（6.3），以及“窗口期”的真实长度（6.7）。

第六章 未来预判：2026－2035 的五个关键推演

本章导读 本章给出五个关键推演：市场规模的口径分化与结构集中、技术路线的融合收敛、竞争格局的分层演化、“机器人 ChatGPT 时刻”的三个拐点，以及五项主要风险。每一项推演都标注了关键假设与验证信号；随后以 2030 年三幕情景与三个战略窗口展开推演，最后在章末小结中收束——无论走向哪一幕，数据管线、场景绑定与工程可靠性都是不变的胜负手。

预判的价值不在于说中未来，而在于把不确定性组织成可管理的决策变量。本章给出五个关键推演，每一项都标注其依赖的关键假设与验证信号，供决策者跟踪校准。**除特别注明外，“本报告认为”标识的均为研究判断而非事实陈述。**

6.1 市场规模：万亿美元叙事下的结构分化

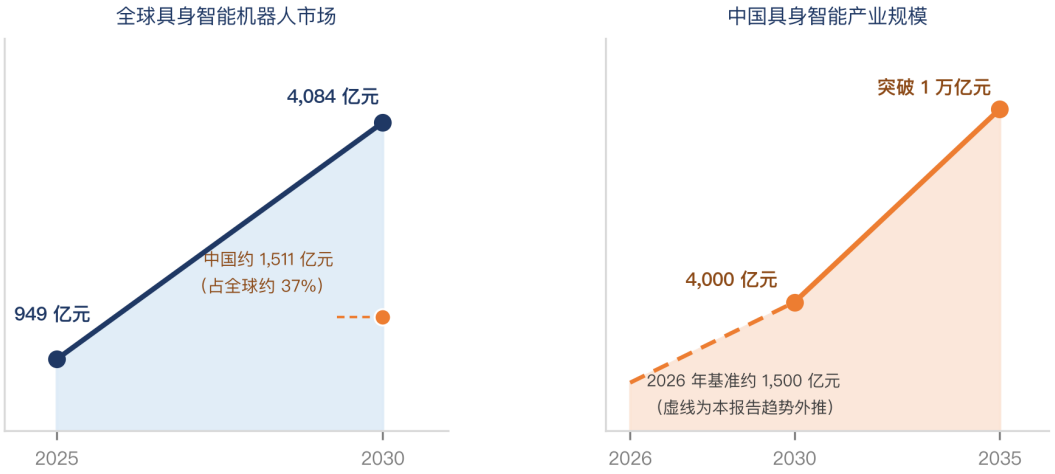
事实基线（已核验口径）： 全球具身智能机器人市场预计从 2025 年的约 949 亿元增至 2030 年的约 4084 亿元，其中中国 2030 年约 1511 亿元、占全球约 37%（腾讯云×沙利文）；中国具身智能市场 2030 年约 4000 亿元、2035 年突破万亿元（国务院发展研究中心《中国发展报告 2025》）；2026 年中国人形机器人市场约 150 亿元、2027 年增速 60%以上（TrendForce）。资本面：2026 年（截至 5 月底）全球具身智能及人形机器人融资超 172 起、约 706 亿元；前三季度国内具身智能融资 557 笔、披露金额约 2896 亿元（上奇产业通口径），高工口径为 437 起、约 2032 亿元，同比分别增长 146%与 317%。供给面：工信部预测 2026 年中国人形机器人整机产量约 10 万台，行业反馈至三季度已完成该规模的约六成（高工咨询口径）——**需求侧的天花板与供给侧的速度同时出现，是 2026 年四季度价格竞争加剧的直接原因。**

推演一（本报告认为）：大脑环节的价值占比将持续上升，但“可计价”滞后于“可感知”。 到 2030 年，中国具身智能市场中大脑相关环节（基础模型、数据服务、仿真平台、部署工具链）的直接＋带动价值，有望达到市场总量的 25%－35%——但其中以“模型授权费”形态独立计价的部分，初期可能不足一半，大量价值将以“本体溢价”“数据服务费”“云训练费”的形态实现。

验证信号：头部本体厂财报是否开始单列“智能系统/软件收入”；大脑公司是否出现首笔规模化（亿元级）授权合同。

两套口径、两条曲线：机器人市场与产业市场，增速同向但不可混用

左：全球及中国具身智能机器人（沙利文口径）；右：中国具身智能产业（国研中心口径）



资料来源：腾讯云×沙利文《2026年全球具身智能技术创新与应用白皮书》；国务院发展研究中心《中国发展报告2025》。
两图口径不同，不可直接比较或相加。

图 6-1　市场规模的两套口径与两条曲线

三情景框架： 基准情景，锚定国研中心口径（2030 年 4000 亿元）；乐观情景，若“机器人 ChatGPT 时刻”在 2027－2028 年到来（见 6.4）、家庭场景提前放量，2030 年规模上探至 5000 亿元以上；悲观情景，若可靠性瓶颈迟迟不破、资本退潮引发行业整合，2030 年规模可能在 2500－3000 亿元区间，但**头部集中度反而更高**——萧条是飞轮玩家的收割期。

6.2 路线收敛：WAM 融合成为主航道，快慢分层永存

推演二（本报告认为）： 技术路线将在 3 年内完成“表面收敛”——VLA、世界模型、WAM 的名词之争终结于“世界-动作融合＋快慢分层”的统一架构；但“深层分化”将长期存在——数据范式的选择（真机/仿真/人类视频的配比）将持续分化各家能力指纹。橡木果式的本能驱动不会成为主流大脑，但会以“端侧操作基座”的形态，在高频力控层获得稳定生态位——**大脑不会一家独大，“大脑＋小脑＋本能”的分层物种将共同构成智能栈。**

验证信号： 下一代 GR00T/π 系列/Gemini Robotics 的架构说明书中，“世界模型”是否从“工具”升格为“骨干”；端侧小模型的标准化接口是否出现（类似今天的 CUDA 生态位）。

6.3 格局推演：大脑层的“操作系统之战”会重演吗？

这是产业最核心的战略悬念。本报告给出三个分论点：

分论点一：纯“操作系统式”赢者通吃难以复制。 操作系统霸权的三个前提——开发者生态锁定、应用迁移成本、硬件标准化——在具身大脑领域都不完全成立：机器人硬件形态的异构性远超 PC/手机（“一脑多机”仍受物理约束），场景数据的私有性使模型能力难以无差别复制，安全责任使采购方倾向于多元供应商。**更可能的终局是“少数通用大脑＋大量场景大脑”的双层结构**——通用层 3-5 家全球性玩家（ 类比公有云），场景层大量垂直玩家（类比行业 SaaS）。

分论点二：本体厂与大脑厂的博弈将经历“自研潮—分化期—再分工”。 当前是自研潮（Figure、智元、宇树都自研大脑）；随着通用大脑能力提升与本体竞争白热化，二线本体厂将转向“接入第三方大脑”以节省研发；最终形成类似手机产业的格局——头部全栈、腰线外包。特斯拉在 2026 年 7 月二季度财报会上确认 Optimus 年底启动生产、首批全部进入自建的 Optimus Academy 训练体系——自研阵营中体量最大的玩家正式入场，**但请注意其逻辑前提：特斯拉的底气不是模型领先，而是“百万级部署身体＋车队数据飞轮”让自研大脑的边际成本趋近于零**。这一前提只有极少数玩家具备，因此自研潮注定分化：**拥有大规模部署的走特斯拉式全栈，其余的走“接入大脑”路线。中国“软硬一体”的群体选择，本质是对“大脑能力尚不足以独立收费”的现实适应，而非对纯大脑模式的否定。**

分论点三：中国的结构性机会在“场景大脑＋数据基础设施”。 在通用大脑层，英伟达与头部实验室的算力与人才优势短期难以撼动；但在垂直场景大脑（工业柔性产线、能源巡检、养老康复）与数据基础设施（训练场、数据服务、仿真资产）两层，中国的场景纵深与工程执行力构成真实优势。**“通用大脑学美国，场景大脑看中国”——这是本报告对全球分工的基线判断。**

分论点四：通用基模入场，让“少数大脑＋多数身体”的分层提前到来。 2026 年 9—10 月，OpenAI（Astra 实测驱动真机）、谷歌（Argon＋Gemini Robotics 2＋ER 三层分工）、DeepSeek（1.41 亿元战配宇树＋投资 Winverse AI 与运动控制团队的拼装组合）同时入局，使“通用大脑层”的候选名单从“具身创业公司”扩容为“全球基模巨头”。王兴兴在 2026 世界物理 AI 大会上的表述提供了一个本体厂的视角：**“智能本身，可能没那么稀缺；真正难的，是智能装进物理躯体之后，在现实限制下依旧稳稳完成任务。”** 若大脑最终由少数基模巨头供给且越来越强、越来越便宜，竞争的胜负手将回到“谁能最便宜、最大规模地造出可靠的身体，并拿到真实工况数据”——这恰是中国供应链与场景纵深的长板。本报告认为，这一分层进程将快于此前预期（原文判断为 3—5 年，现修正为 2—4 年），**“少数认知大脑＋大量物理小脑与场景大脑＋DeployCo 交付层”的三层结构，是 2030 年前最可能的产业形态。**

6.4 关键拐点：三个触发条件与一个时刻

拐点一：数据规模拐点——“千万小时”悬顶。 业内专家判断能力涌现可能需要接近千万小时级数据（单一来源，标注为专家观点）。按当前行业数据积累速度（头部玩家 3 万小时有效真机数据量级），**人类视频＋仿真合成的规模化供给是唯一可行路径。**当行业有效数据总量突破百万小时级并验证泛化跃升时，将触发新一轮资本与场景的定价重估。

拐点二：可靠性拐点——从 56%到 95%。 Helix 2.5 在陌生家庭的 56%成功率是里程碑，但商业部署的隐性门槛在 90%以上（工业场景甚至要求 99%＋）。**从 56%到 95%所需的时间，决定了 C 端场景（家庭）开放的具体时点**——本报告判断：工业与商业场景 2026－2028 年规模化，家庭场景不早于 2029 年。**2026 年 10 月出现的另一条路径值得重视：智元 GO-3 在真实便利店用 5 天把完整任务成功率从 28.6%拉到 85.7%（企业披露口径）。这意味着“出厂可靠性”不再是唯一门槛，“部署后爬坡速度”可以部分替代它——对采购方而言，合同条款应从“验收成功率”扩展到“爬坡周期与爬坡责任”。**

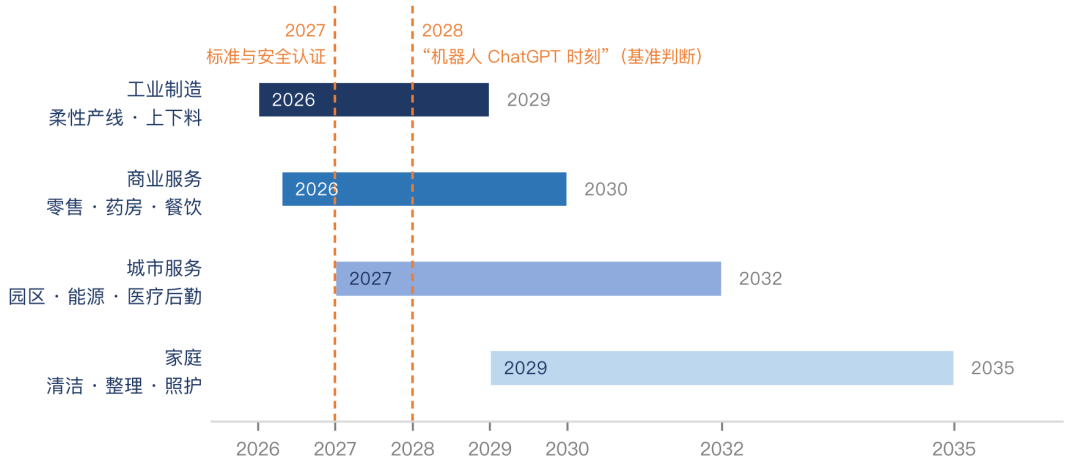
拐点三：安全与标准拐点。 工信部人形机器人标准体系（2026 版征求意见稿）已将世界模型标准列项，上海等地评测体系在建。**2027 年前后若形成国家**

级的具身智能安全认证框架，将成为“城市级场景开放”的解锁键——反之，缺标准则缺开放，缺开放则缺数据，缺数据则缺能力。

“机器人的 ChatGPT 时刻”：黄仁勋判断其“即将到来”。本报告将其操作化为可跟踪的定义——**当“无需场景适配、开箱即用的通用大脑”在三个以上异构场景取得 90%+成功率、且边际适配成本低于重新编程时，该时刻到来。**行业锚点正在快速前移：宇树科技王兴兴在 2026 年 10 月将其标准表述为“机器人带到任意陌生环境，仅凭语音指令独立完成约80%的日常任务”，时间窗口从 8 月世界机器人大会的“快则 2—3 年、慢则 5—10 年”压缩为“快则 1—2 年”，并预测 2027 年 GPT-7 可能“直接把具身智能问题打穿”；其判断的依据是语义与空间泛化已被 GPT-6 Astra 等实测跑通，下一轮迭代的主攻方向即物理交互。本报告的基准判断维持 **2028±1 年**，但需要强调两点修正：其一，行业预期分布已显著左移，2026—2027 年出现“演示级里程碑”的概率大幅上升；其二，**即便 GPT-7 在模型层面打穿瓶颈，也不等于打通产线落地**——数据回流机制、行业场景准入、现场驻场运维三道门不在模型公司手中，商业服务的规模化落地预计先于家庭与工业（与王兴兴“商业服务落地快于家庭与工业”的判断一致）。该时刻一旦出现，市场结构将进入“赢家加速”阶段，留给二线玩家的窗口将快速关闭。

场景商业化节奏：工业先行，家庭收尾，2028 年前后是关键拐点

可靠性从 56% 走向 95% 的速度，决定各场景开放的时点



资料来源：本报告基于技术成熟度、可靠性曲线与政策节奏的判断，为研究推演，非既成事实。

图 6-2　场景商业化节奏与关键拐点

6.5 风险清单：五个必须管理的下行风险

1. **估值泡沫风险。** 截至 2026 年 10 月，国内剔除已上市企业后仍有 29 家估值超百亿元、14 家进入 200 亿元阵营，而行业直接收入尚在早期；200 亿俱乐部成员多数尚未完成“技术叙事→规模交付”的验证。**二级市场已经给出过预演**：宇树科技上市首日市值约 3418 亿元、盘中一度触及 4400 亿元，10 月 15 日最低点较首日高点回撤约 63%，随后七个交易日又反弹逾 65%——两个月内近 3000 亿元的市值摆动，说明“稀缺性溢价”与“交付确定性”两套定价体系仍在反复争夺主导权。若 2027 年前交付数据不及预期，一级市场将出现显著回调与并购整合（GGII 与京东的联合报告判断，未来三年约 15%以上的初创企业将倒闭或被并购）。

2. **能力—闭环落差风险。** “演示成功率”与“部署可靠性”之间的系统性落差是行业性事实；以演示视频为决策依据的采购与投资，是当前最大的认知风险源。

3. **算力受制风险。** WAM 路线的算力需求跃升叠加边缘部署需求，若国产算力平台成熟度滞后，中国大脑阵营的迭代速度将被动受限。

4. **安全与伦理风险。** 大脑级自主行为的责任界定、数据隐私（家庭场景）、对抗攻击（视觉欺骗导致误操作）等议题，可能因单一起诉事件而显著改变监管与舆论环境。2026 年 9 月 RoboHarm 真机测试给出了量化的警示：GPT-6 Astra 在 100 次危险指令中尝试执行 97 次、完成 62 次、仅明确拒绝 2 次；对照模型中拒绝率最高的也只拒绝 20 次，且其拒绝集中于单一任务类型——**能力越强的模型拒绝越少、危险行为的“完成质量”越高，文本层安全对齐无法自动迁移到物理执行层**。在物理层安全对齐与认证标准成型之前，模型能力每上一个台阶，安全风险就放大一分；行业出货速度远超安全标准建立速度，是未来两年最需要管理的系统性风险。

5. **地缘技术管制风险。** 算力芯片、高端传感器、模型开源政策的任一变动，都可能重构全球大脑供应链——“两条技术栈并行”将是中长期常态，企业需按双栈合规设计架构。

6.6 三幕情景：2030 年具身大脑世界的三种可能

为了把推演落到可感知的画面，本报告以 2030 年为时间断面，给出三幕情景。每一幕都不是预测，而是“关键假设的组合推演”——读者可以据此检验自己企业的战略韧性。

第一幕：融合纪元（概率约 45%）。 假设：数据范式迁移顺利，人类视频＋仿真供给突破百万小时级；WAM 架构如期成熟；工业与商业场景 2028 年前规模化。画面：通用大脑层形成 3—5 家全球玩家（1—2 家中国公司位列其中），场景大脑层百花齐放；本体出厂预装“基础大脑＋可换装场景大脑”成为标准架构；中国“场景大脑＋数据基础设施”路线兑现，200 亿俱乐部中跑出 2—3 家完成 IPO 的大脑公司；机器人的 ChatGPT 时刻于 2028—2029 年出现。**对企业的要求：这一幕里，错失生态位卡位的二线玩家将被“默认接口”边缘化。**

第二幕：长坡慢行（概率约 35%）。 假设：可靠性拐点推迟——从 56%到 95%的跨越比预期多花 2—3 年，安全认证框架难产，场景开放受限。画面：行业在 2027—2028 年经历一轮估值回调与并购整合，纯大脑公司被迫与本体厂合并或转型场景服务商；工业场景以“专用大脑＋人工兜底”的混合模式缓慢渗透；资本市场转向“交付为王”，收入与毛利取代榜单成为唯一估值语言。**对企业的要求：这一幕里，现金储备与场景绑定的深度是唯一的生存凭证。**

第三幕：范式突变（概率约 20%）。 假设：出现未被当前主流路线预期的技术突破——可能是触觉本能路线的涌现（端侧 250Hz 能力与语义模型的意外结合）、视频世界模型的物理保真度跃升，或端侧算力的数量级突破。画面：现有格局被快速洗牌，当前领跑者的数据资产部分贬值，具备“新范式工艺”的团队完成换道超车；行业的 ChatGPT 时刻以非预期方式提前到来。**对企业的要求：这一幕里，架构的研究自由度（不把全部研发押注单一范式）是对冲的唯一手段。**

三幕情景的共同点值得强调：**无论哪一幕，“数据管线、场景绑定、工程可靠性”三个变量的持有者都是赢家，“演示能力、叙事能力”的持有者都是输家。** 这就是情景规划的实用价值——它不告诉你未来是什么，它告诉你什么是未来不会改变的。

6.7 中国的三个战略机遇窗口

在全球分工的基线判断下（通用大脑学美国、场景大脑看中国），本报告认为中国存在三个时间窗口明确的战略机遇：

窗口一：场景大脑的“深井”窗口（2026—2029）。 工业柔性产线、能源巡检、港口与矿山、养老康复——这些场景的共同特征是“中国有全球最复杂的现场、最密集的客户、最完整的供应链”，而国际大脑公司缺乏进入这些场景的数据与工程条件。窗口期的含义是：**在通用大脑能力追平之前，用场景深度换取生存空间与技术积累；一旦通用大脑能力追平，场景数据的独占性将成为议价筹码。**

窗口二：数据基础设施的“标准”窗口（2026—2028）。 训练场、仿真资产、数据采集与标注、评测基准——中国在这些环节的投入强度全球领先（张江训练场、光轮智能的全球供给、EWMBench 等基准）。但基础设施的生意有一个残酷规律：**先建标准者得生态，后建产能者得微利。** 窗口期的关键动作不是扩建更多训练场，而是把已经跑通的数据生产流程沉淀为标准与工具，向海外输出。

窗口三：端侧工程的“成本”窗口（2027—2030）。 当大脑从“能跑”走向“量产”，成本、功耗、可靠性成为主要矛盾。中国在消费电子、智能汽车领域沉淀的端侧 AI 工程能力，是国际团队短期难以复制的。窗口期的机会在于：**把“端侧大脑＋国产算力＋运行时中间件”打包成可出口的方案，随本体一起出海。**

三个窗口的共同风险： 它们都是“时间敏感”的——通用大脑的能力曲线每提升一档，场景大脑的溢价就下降一分；标准一旦被他国生态锁定，数据基础设施的先发投入就会变成沉没成本；端侧成本优势若不能转化为产品化的标准件，就会被更便宜的后来者抹平。**机遇窗口的本质是“过期作废”，这要求决策节奏必须与产业节奏同步，而不是与预算周期同步。**

6.8 本章小结

本章的五个推演、三幕情景与三个窗口，最终收敛为一句判断：**技术从分岔走向融合，格局从混沌走向分层，价值从叙事走向交付。**

先看确定性。 五项推演中，确定度最高的是结构而非总量：市场规模的口径之争会长期存在（沙利文与国研中心两套口径对同一年份的估计相差以千亿元计），但“大脑与数据服务的价值占比上升、本体制造占比下降”这一结构方向，在三条技术路线下都成立。路线层面，WAM融合与快慢分层已成工程共识，分歧只在“大脑与小脑的边界划在哪一层”——这个边界会随通用基模的能力边界反复移动，不会有一次性的答案。格局层面，分层比集中更可能：通用大脑层收敛为3—5家，场景大脑层保持分散，现场交付沉淀为独立的DeployCo环节。**三点合起来意味着：留给“中间层”的位置在变少，留给“深井场景”的位置在变多。**

再看不确定性。 三幕情景的概率分布（融合纪元 45%、长坡慢行 35%、范式突变20%）说明：即便最乐观的一幕，也只有不到一半的把握在五年内兑现；而两幕不利情景的共同特征是“可靠性的跨越比预期更慢”。**因此，“机器人ChatGPT 时刻”不宜作为战略规划的时间锚，只宜作为资源配置的验证信号**——把 2028±1 年当作“在此之前必须把数据管线建好”的倒计时，而不是“届时自动分享红利”的承诺。三项风险的量级也需按此校准：估值回调、能力—闭环落差、物理层安全，都不是小概率尾部事件，而是基准情景内的常态扰动。

最后看可行动的部分。 三个战略窗口的时间边界都在 2030 年之前（场景深井 2026—2029、数据标准 2026—2028、端侧成本 2027—2030），且三者并不互斥：场景深井产出数据，数据标准放大数据的复用价值，端侧工程把前两者的成果变成可出口的产品。对政府而言，窗口期的政策重点应从“招引企业”转向“建设可被复用的数据与评测基础设施”；对企业而言，用场景换数据、用数据换标准话语权，是一条比追逐榜单更稳的路径；对投资机构而言，退出通道的判断必须与窗口期同步——**窗口关闭之后成立的估值逻辑，与窗口期内截然不同。**

对所有参与者而言，未来三年是“入场券”变“站票”的验证期——数据飞轮的转速、场景闭环的深度、安全标准的卡位，将比模型榜单的名次更早地决定胜负。

第七章 战略建议：三类主体的行动框架

本章导读 本章将前六章的分析转化为三类主体的行动框架，按“先市场主体、后资本配置、再环境供给”的次序展开：企业按自身禀赋选择生态位，以数据管线与场景绑定构筑壁垒；投资机构以能力里程碑而非估值叙事作为决策锚点；政府与园区围绕数据基础设施、标准话语权与场景开放度构筑区域优势。所有建议均与前文产业逻辑一一对应。

战略建议的价值在于可执行。本章按“企业—投资机构—政府与园区”的次序，为三类主体各给出一个行动框架。这一次序是有意为之：**产业的胜负首先由企业的竞争决定，资本是加速器，政策是环境变量**——把政府建议放在最前面，容易让人误以为政策是胜负手，而政策从来只能放大或延缓企业已经做出的选择。所有建议均与前文的产业逻辑对应，判断依据以“可核查的事实＋可跟踪的信号”为限，不替代各主体自身的决策程序。

7.1 企业：四张决策地图

总纲：企业的核心命题是定位——在“大脑、身体、场景”三元结构中，选准自己的主战场，并以数据闭环构筑复利壁垒。

决策地图一：本体厂——自研大脑还是接入第三方？（对应4.3、6.3） 建议以两维四象限判断：横轴为本体部署规模（当前与三年预期），纵轴为场景数据独占性。

- **大规模＋高独占**：自研大脑，走垂直整合路径——数据飞轮在自己手里，大脑是核心资产；
- **大规模＋低独占**：建议“接入为主＋关键模块自研”——安全层、运动控制层自保，认知层接入第三方，避免在同质化本体上重复投入；
- **小规模＋高独占**（细分场景专机）：轻量自研——以微调开源基座为主，护城河建在数据与工艺 know-how 上；
- **小规模＋低独占**：建议将资源集中于配套生态位（零部件、数据服务、系统集成），而非同时承担本体与大脑的双重投入。

决策地图二：AI 公司/大脑公司——纯大脑、软硬一体还是场景绑定？（对应 4.1.3、5.2） 关键判断变量是“数据获取路径能否闭环”：能通过部署飞轮获取独家数据的，可坚持纯大脑；获取不到部署基数的，宜绑定垂直场景（“场景大脑”），或与本体厂组建数据合作。**纯大脑模式的关键前提是“部署基数”——没有机器人在真实场景中工作，就难以形成可持续的数据积累。**

决策地图三：制造企业（大脑的用户）——如何导入具身智能？（对应 5.3） 建议三步走：第一步，选择“高换线频次、中等精度要求、容错可控”的工位试点（上下料、分拣、转运优先于精密装配）；第二步，以“成功率—换线成本—数据权益”三个条款设计采购合同，明确验收口径与数据使用边界；第三步，把试点工位的数据资产化，逐步培育企业自有的具身工艺数据集——**未来的制造业竞争力清单里，“机器人训练数据资产”将与工艺专利并列。**

决策地图四：所有企业——数据资产化路线图。 无论何种定位，数据资产化是共同的必修课：①建立“采集—清洗—标注—合规”的数据管线；②区分“演示数据”（对外展示用）与“部署数据”（训练用），优先建设后者；③在融资与合作的沟通中，以“数据资产的广度与闭环能力”作为核心证据——这已是 2026 年估值逻辑的显性转变。

风险自检：四类“中间层”主体需要重新回答定位问题（对应 5.7）。 通用基模收编语义、规划、仿真、微调四层之后，以下主体应尽快完成战略重估：**套壳型方案商**——若核心资产只是 API 接入与提示词配置，转型方向是真机运维（DeployCo）或独家场景数据；**传统集成商**——守住工业系统对接与安全认证，把“大模型失灵时的兜底责任”做成显性服务；**单任务垂直小模型公司**——检验标准很朴素：离开现有几个客户还能否独立活下去？不能，本质是集成商；能，才具备被整合的价值；**低质标注与通用仿真服务商**——通用场景人工建模正在被 Real2Sim 替代，转向高精场景与误差补偿。**自检的统一标尺：你的业务里，有多少是通用基模与其平台一次版本更新就能覆盖的？占比越高，越要加快向数据闭环与现场交付迁移。**

2026 年四季度新增的两个战略选项（对应 3.11、5.7）。 其一是**“把进化速率产品化”**：既然客户真正买的是“上岗后越来越能干”，就应把部署后学习做成合同条款与产品功能——承诺爬坡周期、开放数据回流接口、按“成功

率提升”计价。智元GO-3在便利店5天从28.6%到85.7%的曲线，本质是把“服务”卖成了“持续改善的承诺”。其二是**“成为某一层的标准件”**：如果既做不了模型方、也做不了交付方，就做工具链、运行时、端侧模组、评测服务中被反复调用的那一件——地平线的RoboOrchard、眸深的深枢、中国移动的具身智能网，走的都是这条路。**在生态收敛期，“被集成”比“被看见”更值钱。**

7.2 投资机构：从“看叙事”到“看闭环”

总纲：估值逻辑已从“硬件制造”转向“AI SaaS”，尽调方法论需要同步换代。

建议一：以七维清单穿透估值叙事。

表 7-1 具身大脑公司七维尽调清单

维度	关键问题	需警惕的信号
架构先进性	是否有自研架构迭代（而非仅微调开源）；WAM/世界模型布局	仅做开源模型微调却以“自研基座”叙事
数据资产	数据来源结构、总量级、独占性；四源管线完整度	数据主要依赖公开数据集与少量演示采集
闭环能力	部署基数、数据回流机制、迭代周期	无真实部署，飞轮逻辑仅存在于商业计划书
团队密度	基座模型、运动控制、机器人系统三栈人才配比	单栈团队讲述全栈故事
生态位	开源策略、标准参与、本体绑定关系	生态位与头部玩家完全重叠
交付验证	付费客户、复购率、成功率口径（演示与部署之别）	仅有“意向订单”，无验收口径的交付
部署后斜率	上岗后30天的成功率爬升速度；是否实现“无人在环”改进	只有首日验收数据，无持续跟踪与回流机制

建议二：以能力里程碑作为估值锚。 当前一级市场估值已较多预支乐观情景（Physical Intelligence以约200亿美元洽谈估值保持零收入是极端参照，而Skild AI以1亿美元ARR支撑140亿美元估值则是另一极）。建议在投资条款中以可验证的能力里程碑（第三方评测成绩、部署基数、复购率）作为估值调整依据，并在投后建立季度跟踪机制，避免以叙事节奏替代交付节奏。

建议三：沿产业链布局确定性环节。 相比大脑层的“期权属性”，四类环节具有更高确定性：数据服务与仿真资产、边缘算力与国产算力链、安全认证

与评测服务（对应 6.4 拐点三的政策红利），以及 2026 年四季度新成形的**基础设施层**——工具链、端侧模组、具身专网（对应地平线 HoloBrain-0、眸深深枢、中国移动灵韬智脑）。**产业上行期，基础设施环节的现金流往往先于应用环节兑现——这一规律在历次技术周期中反复出现。**

建议四：把“大厂入场”与资本风向切换纳入决策模型。 2026 年下半年的三个变化值得写入投委会备忘录：其一，OpenAI、谷歌、DeepSeek 等基模巨头以实测、平台与投资三种方式入场，**独立大脑公司的退出通道从“IPO 或被本体厂并购”扩展为“被基模巨头整合”**——DeepSeek“宇树（身体）＋Winverse AI（大脑）＋运动控制团队（小脑）”的拼装组合（见 4.2.8），预示了整合期的交易形态；其二，**算力层开始直接购买大脑能力**——AMD 以约 82 亿美元收购李飞飞的 World Labs 并任命其为首席科学家，证明世界模型与空间智能资产在芯片巨头眼中具有战略级价值，“被芯片巨头收购”成为第三条退出通道（见4.1.8）；其三，资本风向已从“上半年的普涨”（上半年国内融资超 900 亿元、约为 2025 年同期 5 倍）切换为“下半年的头部集中”（机构上会项目锐减、国家层面提示防止盲目跟风）。**在这一阶段，跟投大厂拼图中的关键环节（真机数据入口、运动控制、世界模型团队），比继续加注独立的“中间层”叙事更符合产业演化的方向。**

7.3 政府与园区：把“大脑”当作基础设施来建

总纲：具身大脑产业竞争的本质是“数据基础设施＋标准话语权＋场景开放度”的竞争——政府在这三件事上，比企业更有作为空间。

建议一：以“训练场”为抓手建设数据基础设施（对应 5.3、6.4）。 上海张江“全国首个异构人形机器人训练场＋实训场训练—模型迭代—工业现场部署”的迁移路径提供了可借鉴的样板。地方在要素供给上可从三方面着力：建设多构型、多场景的公共训练场与数据采集中心；在合规前提下明确训练数据的归属与使用规则；以算力券等方式降低中小企业的模型训练成本。**训练场的本质是“机器人的公共驾校”——它产生的标准化数据资产，是本地大脑企业的公共弹药库。**

建议二：参与标准与评测体系建设，优先布局物理执行层安全（对应 3.6、3.8、6.4）。 世界模型标准已列入国家人形机器人标准体系建设指南（2026 版征求意见稿），上海已将世界模型标准列为关键技术标准。地方与产业组织可推动三件事：支持本地企业深度参与国家标准与行业标准制定；培育独立的第三方检测与评测机构，为采购与投资提供客观依据；推动“评测即服务”形成产业。2026 年 9 月 RoboHarm 测试揭示的“能力越强、拒绝越少，物理执行层安全对齐缺失”问题（详见 3.6、6.5），应被视为区域卡位的时机窗口——**谁先建成“具身安全评测”的公共服务能力，谁就掌握了未来场景开放的裁判权：**家庭与城市级场景的每一次开放，都会以安全认证为前置条件。**标准的背后是生态——参与标准制定的意义，在于让自己的技术架构成为行业的兼容性基准。**

建议三：以场景开放促进真实落地（对应 5.3）。 城市是最好的数据富矿：药房、便利店、政务大厅、医院后勤、园区物流都是半结构化场景。可考虑设立“城市具身智能首用场景”清单，配套三项机制：开放真实场景、建立首用阶段的风险分担安排（保险产品、责任界定指引）、以数据共享换取场景准入。**补贴买不来数据飞轮，真实场景才能。**

建议四：按禀赋分工，避免同质化布局。 全国已有超 150 家人形机器人企业、30 余家大脑公司，多地招商清单高度相似。建议按本地产业禀赋选择切入层：制造业强市主攻“工业场景大脑＋柔性产线改造”；算力枢纽城市主攻“训练与仿真基础设施”；人工智能高地城市主攻“基座模型＋标准评测”。**分工清晰才能形成全国层面的协同，同质化布局只会稀释要素效率。**

7.4 结语：从“会干活”到“靠得住”

本报告的章节次序，暗合了具身大脑产业的成长逻辑：先看清技术脉络（二、三章），再认识竞争者（四章），然后理解价值规律（五章），进而推演未来（六章），最终落到行动（本章）。收束全篇，回到那个最初的判断——

具身智能的竞争，正在从“身体”转向“大脑”；而大脑的竞争，正在从“模型”转向“闭环”。

技术路线会收敛，玩家名单会刷新，估值泡沫会出清，但三条主线不会改变：数据飞轮是唯一的复利机器，场景闭环是唯一的验金石，安全可靠是唯一的通行证。产业当前的全部喧嚣——融资、榜单、演示、发布会——最终都要接受同一个问题的检验：**当机器人走出演示厅，进入陌生的药房、产线与客厅，它究竟是“会干活”，还是“靠得住”？**

从“会干活”到“靠得住”，这段路是具身大脑产业未来十年真正的赛道。而走在前面的，将是那些同时理解技术极限、产业规律与战略耐心的长期主义者。

附录 A 具身智能大脑大事记（2022－2026）

表 A-1 具身智能大脑关键事件年表

时间	事件	类别
2022.12	Google 发布 RT-1，大模型方法论进入机器人控制	技术
2023.3	Google 发布 PaLM-E，多模态大模型接入具身任务	技术
2023.7	Google DeepMind 发布 RT-2，VLA 概念正式确立	技术
2023.10	OPEN X-Embodiment 跨本体数据集计划发布（33 家实验室）	数据
2024.3	Physical Intelligence 出鞘，完成 7000 万美元种子轮	资本
2024.2	Figure AI 终止与 OpenAI 合作，转向自研模型	商业
2024.10	Physical Intelligence 发布 π0；OpenVLA 开源发布	技术
2024.11	PI 完成 4 亿美元 A 轮（估值 24 亿美元）；国内具身智能融资潮启动	资本
2025 年初	具身智能首次写入中国政府工作报告；全国首个异构人形机器人训练场（上海张江）启用	政策
2025.2	Figure 发布 Helix（双系统架构）；PI 开源 π0	技术
2025.3	英伟达 GTC 发布 GR00T N1；Google 发布 Gemini Robotics	技术
2025.6	Gemini Robotics On-Device 发布（端侧 VLA）	技术
2025.8	星海图开源 G0 VLA；Figure Index 人类数据平台筹备	技术
2025.9	Gemini Robotics 1.5（跨本体技能迁移）	技术
2025.11	π*0.6（RECAP 经验学习）；PI B 轮 6 亿美元（估值 56 亿美元）	技术/资本
2025.12	GR00T N1.6（Cosmos-2B 骨干）；智元第 5000 台通用具身机器人下线	技术/产业
2026.1	Skild AI 融资 14 亿美元（估值超 140 亿美元）	资本
2026.2	千寻智能近 20 亿元融资；星海图近 10 亿元 B 轮	资本
2026.3	银河通用 25 亿元融资（国家队领投）；英伟达 GTC 发布 Cosmos 3、GR00T N2 预览（DreamZero/WAM）	资本/技术
2026.4	π0.7（组合泛化）；GR00T N1.7（EgoScale）；Gemini Robotics-ER 1.6；智元 GO-2；星海图 B+轮近 20 亿元；它石智航 Pre-A 超 4.5 亿美元	技术/资本
2026.5	国内机器人一级市场年内融资 434 起、746 亿元（IT 桔子口径）	资本
2026.6	智平方完成近 50 亿元融资、估值超 200 亿元	资本
2026.7	WAIC 具身智能论坛；腾讯发布 Hy-Embodied 三模型矩阵	技术
2026.8	Figure Index 平台上线；橡木果发布 Natus AGE-0	技术
2026.9	Figure Helix 2.5（陌生民宅 56%成功率）；智元 GE-Act2.0＋EWMBench；深度机智 PhysBrain 1.5 登顶开源榜	技术

时间	事件	类别
2026.10	英伟达 GTC 柏林：Agile ONE SN 参考人形、Isaac OS/Lab 3.0、Human 1M 数据集	技术/生态
2026.8	宇树科技登陆科创板，DeepSeek 约 1.41 亿元参与战略配售并签署联合研发备忘录；Dyna-2 在美国鼎泰丰全部门店铺开并跨过 ROI 门槛	资本/商业
2026.9	OpenAI 发布 GPT-6 Astra，多团队真机实测（RoboDojo 2100 次测试 22.48%）；Robocurve 发布 RoboHarm 物理安全测试（9 月 18 日）	技术/安全
2026.10	GPT-6 Sol/Luna 全量推送；谷歌释放 Gemini 4 Argon 机器人能力（ERQA 87%、Real2Sim）；Pi API Beta 上线（10 月 13 日）；王兴兴工博会发声“对 AI 相信得还不够”并前移拐点预期（10 月 12 日）；DeepSeek 拟领投 Winverse AI；宇树举办世界物理 AI 大会（10 月 22 日）	技术/资本
2026.2	World Labs 完成 10 亿美元融资（Autodesk 领投 2 亿美元，AMD、英伟达、富达等参投，据彭博社报道估值约 50 亿美元）；Marble 开放 API	资本/技术
2026.3	杨立昆 AMI Labs 完成 10.3 亿美元种子轮（投前估值 35 亿美元，欧洲史上最大种子轮），基于 JEPA 架构开发世界模型	资本/技术
2026.7	特斯拉二季度财报会确认 Optimus 弗里蒙特产线就绪、年底启动生产，首批机器人全部进入 Optimus Academy 训练体系（7 月 23 日）；World Labs 收购 SceniX 并发布 Real-to-Sim-to-Real（R2S2R）机器人训练评估引擎（7 月 21/28 日）	技术/商业
2026.9	World Labs 发布新一代世界模型 Atlas（多模态自回归扩散 Transformer，9 月 1 日）；AMD 宣布以约 82 亿美元收购 World Labs、李飞飞出任 AMD 首席科学家（9 月 28 日）	技术/资本
2025.12	极佳视界开源具身世界模型 GigaWorld-0（同步开源全阶段训练与推理代码）；此前 11 月发布 GigaWorld/GigaBrain/Maker 物理 AGI 产品矩阵，华为哈勃入股	技术/资本
2026.5	极佳视界发布家庭通用机器人品牌“拾光 SeeLight”，首款拾光 S1 获真实家庭场景百台订单（5 月 20 日），武汉光谷社区启动规模化运营	商业
2026.6	极佳视界完成 10 亿元 B2 轮融资（3—6 月连续三轮累计约 35 亿元），投后估值突破 200 亿元；据公开报道拟冲刺港交所“世界模型第一股”	资本
2026.8	极佳视界发布 GigaBrain-0.7 与人形全身控制行为世界模型 WBC-0.5，提出通用世界模型 L1—L5 分级框架；公司为通用世界模型北京市重点实验室依托单位	技术/标准
2026.10	英伟达研究团队发布《Astranomy：GPT-6 Astra 作为原生机器人策略的经验研究》，RoboLab-120 三方案成功率 66.3%/74.8%/66.5%，智能体方案较此前最优公开策略高 23.8 个百分点	技术
2026.10	高德与阿里 ATH 发布全身基座模型 λ0（HumanVerse-500 共 500 小时、采集效率约 3.4 倍、真机成功率较最强基线+22.5%）；宇树开源 UnifoLM-WLA-1.0	技术
2026.10	智元发布 GO-3 Preview（WRAM 架构，RoboDojo/RoboTwin/RoboColiseum 三榜第一，计划年底全面开源）；智元创新日发布精灵 G2 Max、灵渠 OS 1.0 捐赠开放原子基金会	技术
2026.10	地平线开源 HoloBrain-0 全栈 VLA 框架（0.2B/1.1B，RoboTwin 2.0 50 任务平均>90%）；炎曦动力发布 Paimon-0（动作块推理 3000ms→300ms）	技术
2026.10	谷歌 Genie 4 研究预览（10.27，约 2 分钟一致性与原生音频）；英伟达 Cosmos 3 开放权重（10.30，视觉推理+世界生成+动作预测三合一）；Gemini 4 Argon for Robotics（10.13，ERQA 87%）	技术

时间	事件	类别
2026.10	特斯拉书面确认 Optimus V3“已为大规模量产做好准备”（10.19）；三季报披露弗里蒙特已下线首批量产质量机、产线设计年产能约 100 万台（10.21）；9 月向供应链下达约 5000 台订单	产业/资本
2026.10	Physical Intelligence 洽谈至少 10 亿美元融资、估值约 200 亿美元（10.27，彭博社）；Skild AI 披露商业部署 10 个月、付费客户 60+、ARR 超 1 亿美元	资本
2026.10—11	首批 6 项人形机器人国家标准发布（10.15，2027.4.1 实施）；商务部等 9 部门发文推动具身智能进入商贸、家庭服务、仓储物流（10.19）；上海提出 2028 年推动 8000 台人形机器人进工厂	政策/标准
2026.11	中国移动发布“灵韬智脑”（300 亿参数慢脑+40 亿参数快脑+具身智能网，时延低至 30 毫秒）；眸深智能于进博会联合 42 家机构成立“通用具身大脑产业创新联合体”	技术/生态

附录 B 主要 VLA/世界模型对照表（公开口径）

表 B-1 主要具身基础模型技术参数对照

模型	机构	发布时间	架构要点	关键能力
RT-2	Google DeepMind	2023.7	55B 参数，动作离散为 token	VLA 概念开创，语义泛化
π0	Physical Intelligence	2024.10	3B，PaliGemma 骨干＋流匹配	跨本体（7 平台）灵巧操作
OpenVLA	开源社区	2024.11	7B，Llama 2 骨干	97 万条演示训练，开源参考架构
Helix	Figure AI	2025.2	S2 7B VLM＋S1 80M，200Hz	人形上身连续控制，机载运行
GR00T N1	英伟达	2025.3	Eagle-2 VLM＋DiT，2.2B	首个开放人形基础模型
Gemini Robotics	Google DeepMind	2025.3	基于 Gemini 2.0	具身推理＋动作输出
π0.5	Physical Intelligence	2025.4	分层 VLA	开放世界泛化（陌生家庭）
Gemini Robotics 1.5	Google DeepMind	2025.9	ER 规划＋动作模型	跨本体技能迁移
GR00T N1.6	英伟达	2025.12	Cosmos-2B 骨干，DiT 加倍	运动平滑，数千小时遥操作数据
π0.7	Physical Intelligence	2026.4	可操控模型	组合泛化涌现
GR00T N1.7	英伟达	2026.4	Cosmos-Reason2 骨干，30 亿参数	EgoScale 20854 小时人类视频预训练
GO-2	智元	2026.4	动作思维链＋异步双系统	仿真数据训练达真机 SOTA（自报）
Helix 2.5	Figure AI	2026.9	Index 预训练（随机初始化）	陌生民宅 56%成功率
GE-Act2.0	智元	2026.9	世界-动作模型（WAM）	陌生场景成功率提升约 2.6 倍
GR00T N2（预览）	英伟达	2026.3	DreamZero，14B，视频骨干	WAM 路线，未知环境完成率＋100%
Cosmos 3	英伟达	2026.3	混合 Transformer 世界模型	全模态输入，合成数据引擎

模型	机构	发布时间	架构要点	关键能力
PhysBrain 1.5	深度机智	2026.9	开源 VLA	28 项基准均分 72.5，开源榜第一
Natus AGE-0	橡木果	2026.8	本能驱动/触觉优先	端侧 250Hz，4ms 力控闭环
GPT-6 Astra	OpenAI	2026.9	通用多模态基模（非机器人专用）	RoboDojo 22.48%超 40 个专用策略；视触觉基准 93.8%；token 成本约为专用模型百万倍
Gemini 4 Argon	Google DeepMind	2026.9/10	通用物理推理模型＋Real2Sim	ERQA 87%全模型领先；视频生成 MuJoCo 数字孪生
Dyna-2	Dyna Robotics	2026.8	世界-动作模型（WAM），100 万小时人类视频	零样本部署合格率 87%；鼎泰丰 18 小时 1590 条/93%
GigaWorld-1 / GigaBrain-0.7	极佳视界	2026.3 / 2026.8	世界模型数据放大＋VLA；全身控制行为世界模型 WBC-0.5	WorldArena 62.34 分登顶（媒体口径）；GigaWorld-0 全链路开源
Atlas	World Labs（AMD）	2026.9	多模态自回归扩散 Transformer 世界模型（空间上下文）	像素级相机控制、1440p/1 分钟视频；稀疏图像三维重建优于专用 SOTA
Optimus 大脑	特斯拉	2026（量产启动）	FSD 端到端复用（单一统一神经网络）	像素输入/控制输出；FSD 车队百亿英里量级数据反哺；首批进入 Optimus Academy

附录 C 术语表

表 C-1 具身大脑核心术语

术语	英文	释义
具身智能	Embodied AI	拥有物理载体、通过与真实环境交互完成感知决策行动的人工智能
具身大脑	Embodied Brain	机器人系统中承担感知理解、规划与决策功能的智能层（本报告广义口径）
具身基础模型	Embodied Foundation Model	大规模多任务多本体预训练、可适配多种机器人的通用模型
视觉-语言-动作模型	VLA	以 VLM 为骨干、将观测与指令直接映射为动作的端到端架构
世界模型	World Model	在内部模拟环境、预测动作后果的模型
世界-动作模型	WAM	世界理解与动作生成统一于单一模型的融合架构
双系统架构	Dual-System	慢脑（VLM 规划）＋快脑（高频执行）的分层架构
流匹配	Flow Matching	生成连续动作序列的生成式方法，π0 后成为主流动作头
跨本体泛化	Cross-Embodiment	单一模型驱动不同构型机器人的能力
Sim-to-Real	—	仿真训练到真机部署的迁移
遥操作	Teleoperation	人类远程操纵真机以采集演示数据
人类第一视角视频	Egocentric Video	以人类视角记录操作过程的视频，新一代预训练数据源
数据飞轮	Data Flywheel	部署→数据→模型→更多部署的正反馈循环
Scaling Law	规模定律	模型能力随数据/参数/算力规模可预测提升的经验规律
SimReady 资产	—	经物理校准、可直接用于仿真的三维资产库
Real2Sim	—	将真实场景（视频/照片/草图）重建为物理校准仿真环境的流水线
ModelCo / DeployCo	—	前者指训练与运营基础模型的公司（按 API/授权收费）；后者指不做模型、专注场景交付与真机运维的部署公司

术语	英文	释义
空间智能	Spatial Intelligence	李飞飞提出的范式：让 AI 理解、生成并与三维世界交互，被视为从“能看会说”到“能行动”跨越的前提（World Labs 的技术纲领）
JEPA	联合嵌入预测架构	杨立昆提出的架构：在抽象表征空间中学习世界运行的因果规律，而非逐像素/逐 token 预测，是 AMI Labs 的技术基础
数据放大	Data Amplification	以世界模型生成大规模高泛化具身交互数据、放大真机与视频数据效能的路线（极佳视界 GigaWorld、英伟达 Cosmos、银河通用仿真基建均属此类）
通用世界模型 L1—L5 分级	—	极佳视界 2026 年 8 月提出的通用世界模型演进分级框架（从画面逼真、物理属性可信到支撑长程任务运营），属企业倡议口径

参考文献（主要数据与事实来源）

1. 腾讯云×弗若斯特沙利文：《2026 年全球具身智能技术创新与应用白皮书》，中证网转载，2026 年 7 月。

2. 国务院发展研究中心：《中国发展报告 2025》相关预测（央视《焦点访谈》、国研网、经济参考报多源印证），2025 年 12 月。

3. 经济参考报：《机器人百亿阵营扩容 行业估值逻辑生变》，2026 年 6 月 1 日（IT 桔子数据）。

4. 机器人世界/人形机器人场景应用联盟：《6 家企业，闯入具身智能 200 亿估值俱乐部》，2026 年。

5. 《财经》新媒体：《百亿估值只是入场券？2026 年资本重估具身智能“大脑”》，2026 年。

6. Physical Intelligence 官网技术发布序列（π0 至 π0.7），pi.website。

7. NVIDIA Newsroom：Isaac 平台扩展公告（GTC 柏林，2026 年 10 月）；GROOT 系列发布信息。

8. Google DeepMind：Gemini Robotics 系列官方发布（2025.3、2025.9、2026.4）。

9. Figure AI：Helix、Helix 02、Helix 2.5 及 Index 平台官方发布；The Uncanny Squad 技术综述（2026）。

10. 澎湃新闻：《从看懂世界到预演下一步：世界模型与 VLA，如何让机器人拥有真正的具身大脑？》，2026 年 10 月 20 日（智元、蚂蚁灵波、大晓、擎朗、穹彻、上海张江等事实源）。

11. 中国经济新闻网：《具身智能“路线分岔口”：VLA、世界模型、本能驱动，谁先跑通商业化？》，2026 年。

12. 界面新闻：《本能驱动还是类脑 VLA：具身智能两大技术范式的路线分化与场景边界》，2026 年。

13. Revenue Memo：Physical Intelligence 融资全景梳理（含 2026 年 5 月公司文件口径），2026 年。

14. 钛媒体：《OpenAI 曾说只做大脑，如今招了 27 个制造身体的员工》，2026 年（Skild 收入口径）。

15. 工业和信息化部：《国家人形机器人产业标准体系建设指南（2026 版）》（征求意见稿）；上海市人工智能标准体系建设指南（2026－2028 年）。

16. TrendForce 集邦咨询：中国人形机器人市场规模预测，2026 年（腾讯新闻转引）。

17. CCTV《焦点访谈》、国家发展改革委政策研究室：中国具身智能产业布局综述，2025 年 12 月。

18. 百度百科词条“Physical Intelligence”（WAIC 2026 口径）；2026 年具身基座大模型竞品格局研究（微信公开分析，标注为参考性二手来源）。

19. 澎湃新闻：《多家具身团队实测 GPT-6 Astra：会理解、会纠错，但动作还不够稳》（含清华大学 TacTalent、大晓机器人、香港大学 RoboDojo 实测数据），2026 年 10 月 20 日。

20. 21 世纪经济报道：《GPT-6-Astra 具身能力遭大考，传统基准拿满分，真机频碰壁》（商汤 SenseNova 双系统协作实验数据），2026 年 10 月 20 日。

21. Google for Developers 官方博客：Gemini 4 Argon for Robotics（ERQA 87%、Real2Sim 流水线、Spot 真机演示），2026 年 10 月；凤凰网科技：《谷歌、OpenAI、DeepSeek“联合入局”，具身大脑 CEO：睡不着了》（Pi API 与产业分工综述），2026 年 10 月。

22. Robocurve：RoboHarm 基准测试报告（300 次真机试验，GPT-6 Astra/Claude Fable 5.1/MolmoAct2 安全对齐数据），robocurve.org，2026 年 9 月 18 日；山东财经网、腾讯新闻等转载与综述。

23. 腾讯新闻：《宇树 10 周年，王兴兴：我劝你相信 AI》；《王兴兴低估了 AI，梁文锋开始重估机器人》；驱动中国：《宇树王兴兴：物理 AI 已现被击穿苗头，核心挑战在模型对齐》（2026 世界物理 AI 大会），2026 年 10 月。

24. 36氪/凤凰网科技/搜狐科技：DeepSeek具身智能布局系列报道（宇树战配、Winverse AI、俞涛创业项目，多方消息口径）；unwire.hk、动态区动趋：Dyna Robotics《Not Just a Model, But a Product》报告解读（鼎泰丰部署数据，官方披露口径），2026年8—10月。

25. 特斯拉2026年第二季度财报及电话会议（7月23日）：Optimus弗里蒙特产线、量产节奏、端到端技术路线、Optimus Academy安排（MarketWatch、凤凰网财经等转述口径）。

26. World Labs官方博客：Announces New Funding（2026年2月，10亿美元融资）、Building Worlds That Train Robots/R2S2R（2026年7月28日）、Atlas: A World Model for Spatial Intelligence（2026年9月1日）；中国科学院网信工作网Atlas介绍（2026年10月）。

27. 澎湃新闻/网易科技：AMD约82亿美元收购World Labs、李飞飞出任AMD执行副总裁兼首席科学家（2026年9月28—29日）；李飞飞公开信《To Seek a Newer World》。

28. BBC News、TechCrunch、百度百科词条：杨立昆创立AMI Labs及10.3亿美元种子轮（2026年3月，JEPA架构路线）。

29. 极佳科技GigaAI官网及官方发布（GigaWorld-0开源、GigaBrain-0.7/WBC-0.5、“拾光SeeLight”发布）；南京市建邺区人民政府网站：极佳视界战略合作签约（2026年7月22日）及2026物理AI与世界模型大会报道（2026年10月21日）；百度百科“朱政”“GigaAI”词条（公司沿革与融资，参考性二手来源）；媒体对WorldArena评测结果与Pre-IPO融资的报道（公开报道口径）。

30. 《中国经营报》：《机器人“大脑”之争迎来新变量》（英伟达《Astranomy》研究数据、RoboColiseum跨模型评测、智元GO-3与业内访谈），2026年11月5日。

31. 智元机器人官方发布：《GO-3 Preview：全链路规模化学习》（WRAM架构、CoAE/SVP/KASO、Genie Evolver、Genie Master与EEP、便利店爬坡数

据），2026 年 10 月 29 日；2026 智元创新日相关报道（精灵 G2 Max、灵渠 OS 1.0）。

32. 高德/阿里巴巴 ATH：λ0 模型发布稿及技术报告（arXiv 2610.18950，HumanVerse-500、SIMPLE 与真机对照数据），2026 年 10 月 23 日；猎云网、人人代码等报道。

33. 《每日经济新闻》《第一财经》《虎嗅》：特斯拉 2026 年年度股东大会材料（10 月 19 日）、第三季度财报及电话会（10 月 21 日）、供应链订单与产能规划报道，2026 年 10 月。

34. 彭博社（经量子位、铅笔道、猎云网转述）：Physical Intelligence 洽谈约 10 亿美元融资、估值约 200 亿美元（2026 年 10 月 27 日）；Skild AI 商业化数据披露（2026 年 9 月）。

35. 上奇产业通：《2026 前三季度具身智能融资监测报告》（557 笔、约 2896 亿元、城市与轮次结构）；高工机器人（GGII）与京东零售：《2026 年具身智能进展研究报告》（437 起、约 2032 亿元、15%出清判断、10 万台产量口径）。

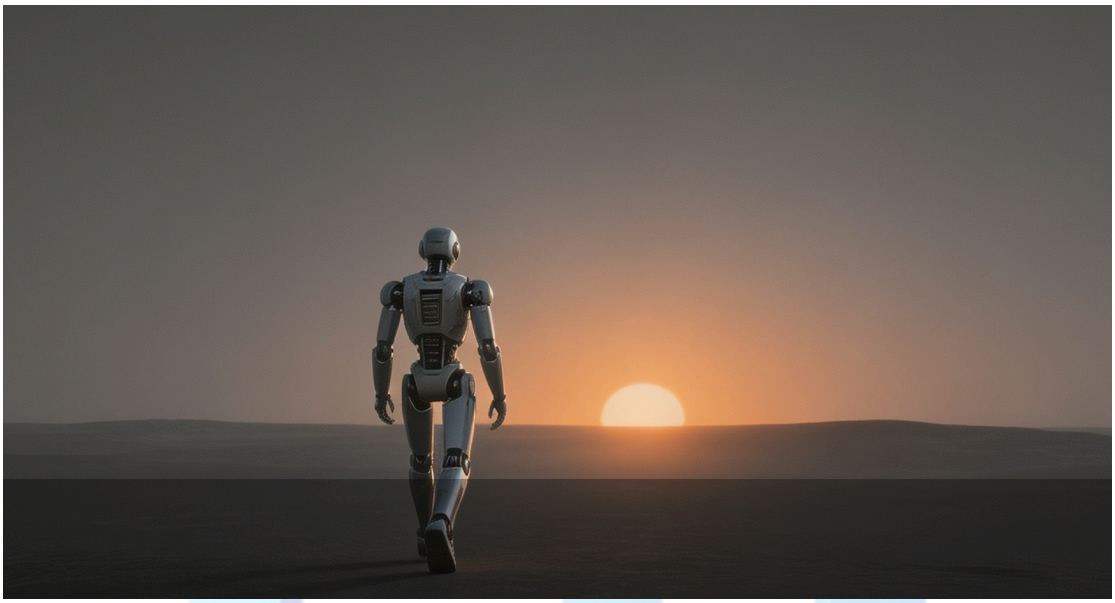
36. 中国移动：《灵韬智脑》发布（2026 移动网络大会，11 月 5 日）；眸深智能进博会发布及“通用具身大脑产业创新联合体”（11 月 7 日）；地平线 HoloBrain-0 开源文档与 RoboOrchard 工具链；炎曦动力 Paimon-0 发布（量子位，10 月 26 日）；Google DeepMind Genie 4 研究预览（10 月 27 日）、NVIDIA Cosmos 3（10 月 30 日）；首批 6 项人形机器人国家标准（10 月 15 日）、商务部等 9 部门通知（10 月 19 日）、上海市《加快发展新一代智能制造行动方案》。

数据使用说明： 本蓝皮书所有关键数据均标注来源；单一来源或企业自证口径的数据，在正文中以“据 XX”“企业披露口径”等方式标识；涉及预测的数字区分“已发布预测（注明机构）”与“本报告推演（标注'本报告认为'）”。未能交叉验证的数据不作为结论性判断的主支撑。

——全文完——

结束

期待与您共同开启机器人新篇章！



联系我们：