

具身智能：范式演进、技术架构与产业发展路径

薛钦源¹, 万宏^{2*}

(1. 首都经济贸易大学财政税务学院, 北京 100070; 2. 中国石油天然气集团有限公司, 北京 100007)

摘要: 人工智能正从“离身计算”向“具身交互”经历深刻的范式革命。本文剖析具身智能的范式演进、内涵、技术架构、产业生态与发展路径。研究认为, 具身智能通过融合“认知大脑”与“物理身体”, 在“感知-决策-执行”的闭环交互中克服传统人工智能“知行分离”的局限, 是通往通用人工智能的关键路径。在技术功能上将具身智能解构为“感知-决策-执行-学习”四个层级, 在产业分工上剖析为“核心软硬件-系统集成-场景应用”三大环节。基于此, 进一步论证了具身智能对生产力与产业组织的重塑效应, 认为其将推动生产范式从静态自动化向动态自主优化演进, 并催生人机共生的新型组织形态。最后, 研判了产业发展面临的技术自主可控、场景落地难与生态碎片化三大瓶颈, 并提出“固本强基、场景驱动、生态共建”的战略路径, 为该领域的理论与政策制定提供学理参考。

关键词: 具身智能; 人工智能; 范式演进; 技术架构; 产业重塑

中图分类号: TP18; TP24; F426; G301 **文献标志码:** A

Embodied Intelligence: Paradigm Evolution, Technical Architecture, and Industrial Development Pathways

XUE Qinyuan¹, WAN Hong^{2*}

(1. School of Public Finance and Taxation, Capital University of Economics and Business, Beijing 100070, China; 2. China National Petroleum Corporation, Beijing 100007, China)

Abstract: Artificial intelligence is undergoing a profound paradigm revolution from “disembodied computing” to “embodied interaction.” This study dissects the paradigm evolution, core connotation, technical architecture, industrial ecosystem, and development pathways of embodied intelligence. Research indicates that by integrating the “cognitive brain” with “physical body” to establish closed-loop “perception–decision–execution” interactions, embodied intelligence overcomes the “knowledge–action decoupling” limitation of traditional AI, representing a critical pathway toward artificial general intelligence. Technologically, we deconstruct embodied intelligence into four functional layers (perception, decision–making, execution, and learning), while industrially we analyze three value–chain segments (core hardware/software, system integration, and scenario applications). Furthermore, we demonstrate embodied intelligence’s transformative effects on productivity and industrial organization, driving production paradigms from static automation toward dynamic autonomous optimization while catalyzing human–machine symbiotic organizational forms. Finally, we identify three critical bottlenecks: technological sovereignty constraints, application–scenario barriers, and ecosystem fragmentation, proposing a tripartite strategic pathway—“consolidating foundational capabilities, scenario–driven deployment, and collaborative ecosystem development”—to provide theoretical references for academic research and policy formulation.

Keywords: embodied intelligence; artificial intelligence; paradigm evolution; technical architecture; industrial transformation

基金项目: 中国社会科学院重大项目“中华民族工业文明的形成演进及驱动力研究”(2023YZD054); 首都经济贸易大学 2025 年度教改项目“适应新质生产力发展需求的交叉融合人才及培养模式研究”

1 引言

人工智能正从数字世界的计算智能，向深度融入物理世界的具身智能发生深刻的范式跃迁。不仅是技术路径的演进，更预示着以“知行合一”为特征、重塑生产力与产业组织形态的经济变革。作为国家重点布局的未来产业，对具身智能发展路径的认知与前瞻性研判已成为重要的理论与政策议题。然而，当前对具身智能的认知多局限于技术实现层面，对于其如何驱动产业生态演化、重塑价值创造逻辑的系统性经济学分析框架尚显不足。本文认为，具身智能的核心价值在于，通过融合“认知大脑”与“物理身体”，构建“感知-决策-执行”的实时闭环，从根本上克服了传统人工智能“知行分离”的局限，使其成为能够直接作用于物理世界的通用目的技术。此外，具身智能技术范式的特性决定了它将对经济系统产生深远影响。

为此，遵循“技术架构-产业生态-经济影响”的分析框架。首先剖析具身智能从“离身计算”到“具身交互”的范式演进与核心内涵，并构建其“感知-决策-执行-学习”四位一体的技术功能架构。以此为基础，进一步解构其“核心软硬件-系统集成-场景应用”的产业价值链。最后，论证其对微观企业生产函数和宏观产业生态的重塑效应，并针对产业发展面临的技术自主可控、场景落地与生态协同三大核心瓶颈，提出“固本强基、场景驱动、生态共建”的战略路径，以期为理解和分析具身智能这一未来产业的发展规律提供有力的学理支撑。

2 范式演进：从“离身计算”到“具身交互”

自20世纪中叶诞生以来，人工智能的发展已历经数次范式演进。从以符号主义为主流，将智能等同于逻辑推理的“经典人工智能”，到以连接主义为主导，依赖神经网络进行统计学习的“现代人工智能”，其主流探索路径始终聚焦于提升智能系统的计算与认知能力。特别是在深度学习浪潮的推动下，以大语言模型为代表的技术突破，使得人工智能在处理文本、图像等抽象信息时展现出卓越性能，构筑起强大的“数字大脑”^[1]。然而，无论是符号主义的逻辑推演，还是连接主义的统计学习，其本质均是将智能视为在抽象符号或数据空间内发生的“离身计算”，这种范式内在地造成了智能系统与物理世界的隔绝。

2.1 传统人工智能范式的内在局限

“离身计算”范式将智能封装于虚拟的数字世界，其核心局限在于“知行分离”，主要源于两大根本性难题。其一为“符号接地难题”。智能系统处理的符号，无论是逻辑表征还是数据向量，均缺乏与物理世界实体和感官经验的直接、稳定映射，导致其对世界的理解偏向于形式化而非实质性，难以真正稳定、高效把握符号背后的现实意义^[2]。其二为“认识论鸿沟”。智能的知识来源局限于预先给定的静态数据集，而非通过与环境的主动交互生成，使得其知识体系封闭且脆弱，缺乏对物理世界因果关系、动态变化及隐性常识的深刻理解。以上两大难题共同导致传统人工智能在面对复杂、多变、充满不确定性的物理任务时，常暴露出物理常识缺失、泛化能力脆弱和实时适应性差等问题。智能虽能“知”，却不能“行”，成为其迈向通用人工智能的关键瓶颈。

2.2 具身智能：作为应对方案的范式革命

为突破“离身计算”的禁锢，具身智能的发展应运而生，标志着人工智能研究从核心关注“计算与认知”，转向“感知-行动-认知”三位一体的范式革命。具身智能范式借鉴了认知科学、现象学与机器人学中的“具身认知”理论，其核心观点是：认知并非大脑孤立的运算过程，而是由身体的物理结构、感官系统和运动能力与环境的交互共同塑造和决定。身体不仅是行动的执行者，更是感知的媒介与认知形成的基础。因此，具身智能在研究范式上实现了两大根本性转变：一是从线性处理到闭环交互。具身智能摒弃了传统人工智能“感知-规划-执行”的分离式线性模式，构建了“感知-认知-行动”高度耦合的闭环系统，强调智能必须通过与真实世界的持续交互和即时反馈进行学习。二是从被动学习到主动探索。智能体不再是被动处理离线数据的“学习者”，而是成为主动探索环境、通过试错积累经验的“行动者”。智能在“交互-学习”的循环中自适应地“涌现”和演化^[3]，而非被预先编程。

2.3 认知大模型与物理身体的融合：迈向“知行合一”

需要强调的是，大语言模型等认知大模型的崛起，非但没有削弱具身化的必要性，反而为其提供了前所未有的强大“认知引擎”。二者的深度融合，为实现人工智能的“知行合一”提供了明确的技术路径。一方面，认知大模型充当具身智能的“大脑”，其强大的语言理

解、逻辑推理与世界知识，使其能够胜任高级任务规划、常识推理与决策中枢的角色，负责将抽象的人类指令分解为逻辑清晰、可执行的子任务序列^[4]。另一方面，物理身体则充当了“实践平台”，为大模型生成的知识和规划提供了在真实世界中验证、纠错和“接地”的场域。通过与物理世界的直接交互，智能体能够学习隐性的物理规律、获得真实的感官反馈，从而弥合抽象知识与物理实践之间的鸿沟^[5]。可以说，从符号主义的逻辑推演到连接主义的统计关联，再到具身智能的虚实融合，“身体”的回归正成为人工智能通往更高阶段智能的必然选择。

3 具身智能的概念与层次：身智一体的三阶进化

3.1 内涵界定：身智统一的智能系统

具身智能是一种将高级智能算法与物理实体深度融合的智能系统范式。它能够在真实物理环境中，通过“感知-认知-行动”的闭环交互，自主学习并完成复杂任务。其核心特征在于，智能不再是存于数字空间的纯粹计算程序，而是通过物理身体与真实世界进行直接、双向的交互，并在交互过程中实现感知、认知与行动的统一^[6]。

从系统论视角看，具身智能由两大相互依存、共同定义的核心要素构成：一是智能内核（大脑）。作为承载高级算法模型的计算中枢，负责处理多模态感知信息，进行环境理解、任务规划、决策生成与学习优化。以大模型为代表的认知技术正成为驱动其能力跃升的关键。二是物理本体（身体），作为智能内核与物理世界交互的唯一媒介，集成了传感器与执行器。物理本体的形态、材质与动力学特性，不仅是行动的工具，更从根本上决定了智能体感知世界的方式和交互的边界，构成了智能涌现的物理基础^[7]。二者共同构建了动态的“感知-认知-行动”循环反馈系统，行动结果又反馈为新的感知，使智能得以在与物理世界的持续交互中涌现和演化^[8]。

3.2 概念辨析：范式、载体与范畴

（1）具身智能与人形机器人：范式与载体的关系

将具身智能等同于人形机器人是一种常见的误解。二者的关系是“一般与特殊”。具身智能是一种研究范式，强调“智能如何通过身体与世界交互而涌现”的根本原理，其核心在于系统架构与学习机制，而非载体的特定形态。任何具备物理实体并能与环境自主交互的机器，无论其形态是轮式、足式、机械臂乃至无人机，均属具身智能

的范畴。人形机器人因其形态与人类相似，在为人类设计的环境中具备天然的通用性，是具身智能迈向通用场景的理想载体之一。然而，在产业实践中，物理本体的形态是功能需求、成本效益与环境适应性权衡的结果，面向特定任务的功能导向型异形机器人往往是更务实的选择。例如，仓储物流机器人采用轮式结构以追求效率与稳定。将具身智能等同于人形机器人，会窄化其理论内涵与产业图景。

（2）具身智能与智能体：物理世界中的特化

在人工智能领域，智能体（Agent）指任何能自主感知环境并采取行动以达成目标的实体，涵盖从软件到硬件的广阔范畴。因此，具身智能与智能体的关系是“子集与全集”，具身智能是智能体概念在物理世界中的“特化”或“实体化”，二者的核心区别在于交互对象的不同。虚拟智能体的交互对象是信息空间，遵循信息规则；而具身智能体的交互对象是物理空间，其感知和行动必须面对并遵循重力、摩擦力等物理定律。因此，具身智能体必须处理由传感器噪声、执行器误差和环境动态变化带来的巨大不确定性，对其系统的鲁棒性、实时性和安全性提出了远高于虚拟智能体的要求^[9]。

3.3 演进层次：从算法赋能到身智协同进化

具身智能并非静态的技术概念，而是动态发展的技术体系。根据智能内核与物理本体的耦合程度，可将其演进路径划分为三个层次，亦对应着不同的产业发展阶段与价值创造模式。

（1）初级层次：算法赋能型

初级层次具身智能的核心特征是“强人工智能大脑+传统机械体”，即软件定义硬件。产业上表现为将先进的人工智能算法模块（如视觉识别）集成于现有的、相对成熟的机械硬件（如工业机器人、汽车）之上，实现智能化改造。在此，智能是主导性的创新要素，物理本体主要扮演被动执行者的角色。例如，为传统工业机器人加装视觉识别能力，或为汽车搭载高级驾驶辅助系统。该层次能以较低的硬件改造成本，快速提升存量设备的自动化水平。

（2）中级层次：形态优化型

中级层次具身智能的特征转变为“特化躯体+适应性智能”，体现了硬件与软件协同设计的思想。物理本体的设计开始扮演主动角色，通过对机器形态、材料、结构的创新，使其本身具备一定的“物理智能”，从而

分担部分原需复杂算法完成的计算任务。例如，采用软体材料的手爪能被动适应不同物体的形状，从而简化抓取算法；精巧设计的被动行走机器人，仅需极简控制就能实现稳定步态。在此，硬件创新与软件开发协同并进，物理本体的形态本身即蕴含了部分智能。

(3) 高级层次：身智协同进化型

高级层次的具身智能是理想形态，其核心特征是共同进化的身智系统。在此层次，智能内核与物理本体不再是主次分明或简单相加的关系，而是在设计、学习和演化过程中形成不可分割、高度协同的有机整体。智能算法根据物理反馈不断优化自身，物理结构则根据智能需求迭代升级。其未来图景是智能体能根据任务自主优化其运动模式乃至形态，云端“大脑”可连接并协调多

种不同形态的物理本体，实现群体智能与经验共享，最终标志着具身智能从生产工具演变为具备自主学习和创造能力的“智能物种”^[10]。

4 具身智能技术功能架构：感知 - 决策 - 执行与学习闭环

具身智能作为融合了先进人工智能与复杂物理实体的系统工程，其内在技术架构遵循动态的闭环控制模型。从技术功能架构视角看，该模型通过“感知-决策-执行”的连续循环与物理世界进行实时交互，并通过一个贯穿始终的学习与迭代机制实现能力的自我演化和持续提升。基于此框架，可将具身智能的技术体系解构为四个相互依存、层层递进的功能层级，如图1所示。

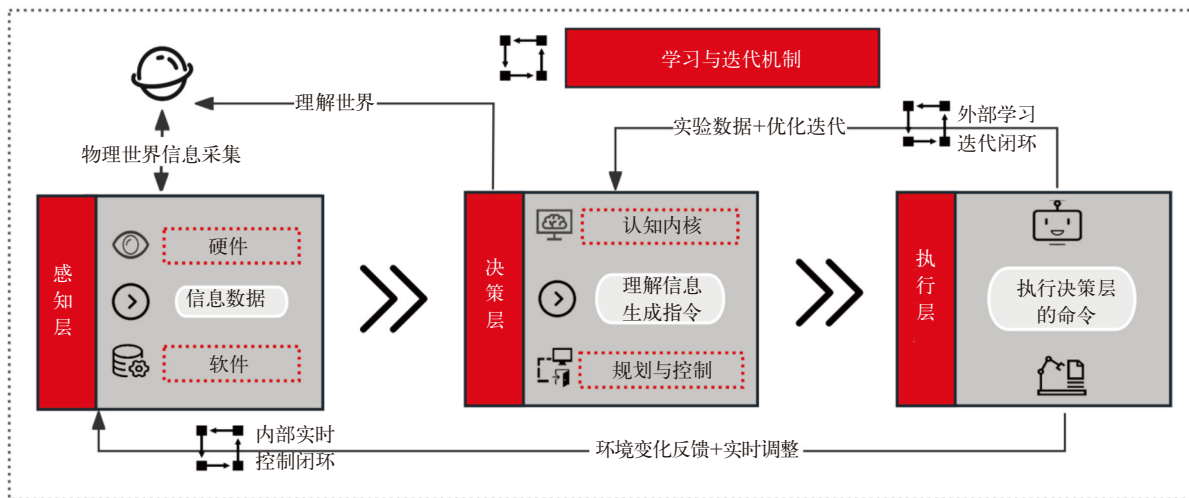


图1 具身智能技术功能架构示意图

4.1 感知层：物理世界的数字化映射

感知层是具身智能体与物理世界建立信息联系的起点，其核心功能是将环境及自身的状态物理信息，转化为可供智能内核处理的结构化数字信号，完成从物理现实到数字表征的映射^[11]。作为整个系统的“感觉器官”，感知层所提供信息的广度、精度与实时性，直接决定了系统决策与行动的质量上限，是智能涌现的根本信息基础。该层级是硬件与软件的深度融合体。硬件层面，它依赖于多模态传感器矩阵，如提供三维空间几何信息的激光雷达与深度相机、捕捉纹理与色彩的视觉传感器，以及赋予精细操控能力的力矩传感器和电子皮肤等触觉传感器，共同构建对物理世界的全方位度量能力。软件

层面，则通过先进的感知算法，如多传感器数据融合、即时定位与地图构建等，对原始数据进行处理、解释与融合，旨在从嘈杂、高维的信号中提取出关于环境的稳定、准确的语义信息。当前，感知层的核心挑战在于如何在保证高性能的同时，实现低功耗、小尺寸与低成本统一，并确保在非结构化与动态环境中的感知鲁棒性。

4.2 决策层：认知到规划的智能中枢

决策层是具身智能的“大脑”，负责对感知信息进行高级处理，完成从环境理解、任务规划到行为决策的完整认知过程，并最终生成对执行层的具体控制指令^[12]。该层级是智能体“智慧”的核心体现，其先进性决定了智能体行为的自主性、适应性与前瞻性。决策层内部可

进一步解构为两大核心模块：一是认知内核。主要由大型语言或多模态基础模型驱动，负责高级别的语义理解、常识推理与长期任务规划。例如，它能将“把桌上的苹果放到冰箱里”这类模糊的人类指令，自主分解为诸如“走向桌子→识别苹果→规划抓取→导航至冰箱→打开冰箱门→放入苹果→关闭冰箱门”等逻辑清晰、物理可执行的子任务序列。认知内核的引入，使得具身智能超越了传统自动化机器的范畴，赋予了其理解复杂指令和应对开放式任务的能力。二是运动规划与控制。作为连接“思考”与“行动”的关键桥梁，该模块负责将认知内核生成的高级任务指令，转化为具体、平滑且安全的物理动作序列。涉及路径规划、运动学与动力学求解、轨迹优化等算法，并需综合考虑本体的物理约束、环境障碍物及任务的实时性要求。当前，决策层呈现出典型的“软件定义”特征，其核心竞争力体现在“模型-算法-芯片”的技术体系上。同时，该层级正经历从基于规则的传统控制论，向数据驱动的端到端学习范式的演进。

4.3 执行层：智能意图的物理实现

执行层是具身智能的“骨骼与肌肉”，作为智能意图的物理载体，负责精确执行决策层的指令，通过与环境的物理交互完成最终任务。该层级的性能直接定义了智能体物理能力的边界，如力量、速度、精度与灵活性，是高端装备制造能力的集中体现。执行层的核心由各类执行器及其集成的机械本体构成，关键零部件包括提供动力的伺服电机、实现高传动比与扭矩放大的精密减速器等，这些部件共同构成了机器的关节、灵巧手、移动底盘等功能结构。近年来，软体机器人学与新材料科学的发展正不断扩展执行层的能力边界，使得智能体能更好地适应非结构化环境和与脆弱物体的交互。例如，形状记忆合金、介电弹性体、气动人工肌肉等新材料与新构型的引入，使机器能够更好地适应非结构化环境和脆弱物体的操作。当前，该层的发展不仅依赖于机械、材料与电气工程的深度融合，也受限于精密加工工艺与先进制造水平。

4.4 学习与迭代层：系统进化的赋能引擎

学习与迭代层并非一个独立的物理模块，而是贯穿于“感知-决策-执行”全流程的赋能机制。学习与迭代赋予系统纠正偏差、适应变化及从经验中学习的能力，是具身智能区别于传统开环自动化系统的本质特征，也是其实现能力持续进化的核心引擎。该层级通过两个不同

时间尺度的反馈循环驱动系统进化。一是内部实时控制闭环。一个高频、低层次的内部循环。执行器的动作会即时改变自身与环境的状态，感知层捕捉到这些变化后形成反馈信号，决策层的控制模块据此进行毫秒级的实时调整，以抑制扰动、补偿误差，确保动作的稳定与精确。此闭环是经典控制论在系统中的应用，是保证系统鲁棒性的基础。二是外部经验学习闭环，一个基于任务成败、能耗、效率等结果的更高层次反馈。具身智能体在与环境交互的过程中，任务的成功与否、能耗、效率等结果构成了更高层次的反馈。系统通过强化学习、模仿学习等机制，利用交互经验数据，在云端或本地对决策模型进行离线或在线的优化迭代，从而实现能力的持续进化。当前，学习闭环的完善催生了对仿真平台、数据服务与云端算力的巨大需求。通过“仿真训练-现实部署-数据反馈-模型优化”的数据飞轮效应，智能体能够实现经验共享与群体智能，从而指数级地加速技术的成熟与应用落地。

5 产业生态与价值链：三元协同的价值创造体系

若技术功能架构揭示了具身智能的内在机理，那么产业生态则描绘了其价值实现的全景图。从产业经济视角来看，具身智能的产业生态可解构为由上、中、下游构成的完整价值链。该链条上游是技术基石与价值源点，中游是系统集成与价值转化，下游则是场景应用与价值实现。三者相互依存、协同演进，共同驱动产业发展，如图2所示。

5.1 上游：技术基石与价值源点

上游环节为整个产业生态提供核心的硬件、基础软件与关键算法，是具身智能能力边界的技术决定者与价值创造的源头。此环节具有技术密集、资本密集、高进入壁垒和高附加值的特征，其技术水平与供给自主性直接决定了整个价值链的性能上限与安全韧性。上游可进一步划分为两大板块。

1) 硬件核心零部件板块。作为实现“具身”的物理基础，是高端制造能力的集中体现。此板块主要包括三类。一是感知单元，如CMOS图像传感器、激光雷达、力矩传感器等，作为智能体的“五官”，决定了其获取物理世界信息的精度与维度；二是执行单元，如高性能伺服电机、精密减速器、控制器等，作为智能体的“肌肉与骨骼”，

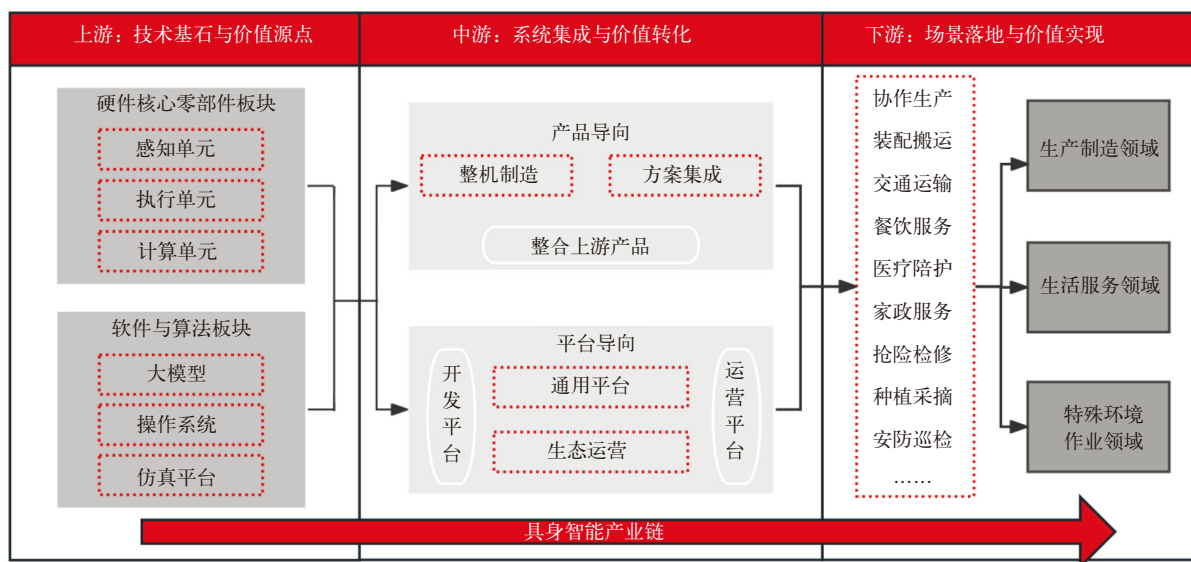


图2 具身智能产业链示意图

决定了其物理交互的能力边界，包括力量、速度与精度；三是计算单元，以 GPU、NPU 等人工智能芯片为核心，作为智能体“大脑”的算力基座，决定了其端侧实时决策的效率与能力。

2) 软件与算法基础平台。作为智能实现的逻辑核心，构成了产业生态的“操作系统”和开发底座，决定了产业的创新效率与生态开放性。主要包括：一是基础大模型，以大型语言模型和多模态大模型为代表，为具身智能体提供通用的世界知识、逻辑推理和任务规划能力，是其“大脑”实现智能涌现的核心引擎；二是操作系统与中间件，如机器人操作系统（ROS/ROS2），为异构硬件和复杂应用提供标准化的通信、驱动和开发工具，是降低开发门槛、构建开放生态的关键；三是仿真与开发平台，提供高保真度的物理仿真环境，是加速算法训练、测试与验证的“数字孪生”引擎。

5.2 中游：系统集成与价值转化

中游环节是连接上游技术供给与下游市场需求的产业枢纽，其核心职能是将上游分散的硬件、软件与算法进行系统集成、工程化与产品化，形成面向特定场景的具身智能整机或可供二次开发的通用平台。中游在价值链中扮演着“价值转化器”与“产品定义者”的关键角色，当前呈现出两种并行的发展路径。

1) 产品导向，典型代表是整机制造与方案集成。该模式以交付“交钥匙”式的功能型产品或解决方案为

核心，参与者主要是各类机器人制造商（工业、协作、服务机器人等）和系统集成商。通过整合上游零部件与自身在结构设计、控制工程、行业工艺等方面的需求，为下游客户提供解决特定痛点的具身智能产品。例如，面向汽车产线的焊接机器人系统或服务于智慧物流的自主移动机器人集群。此模式的竞争核心在于系统工程能力、供应链管理、规模化生产及对垂直行业需求的深度理解。

2) 平台导向，典型代表是通用平台与生态运营。该模式是具身智能时代涌现的新兴业态，其目标是打造可供二次开发的通用软硬件平台。以人形机器人公司为例，参与者提供具备基础通用能力的机器人本体和开放的开发接口（API/SDK），赋能第三方开发者或下游客户进行上层应用创新。此类平台型企业致力于成为具身智能时代的“Android”或“iOS”，其核心目标是通过构建开发者社群和应用商店来催生网络效应，从而掌握产业生态的主导权。

5.3 下游：场景应用与价值实现

下游环节是具身智能产业价值的最终实现场域，涵盖了所有引入具身智能产品与服务的终端行业。下游需求的广度与深度，决定了整个产业的市场规模与长期发展潜力。当前，具身智能的应用场景正从结构化、确定性的工业环境，向非结构化、动态化的社会生活环境加速渗透，主要可划分为三大领域。

1) 生产制造领域。生产制造是当前具身智能应用最成熟、价值最明确的领域。其核心价值在于提质、增效、降本，推动制造业向柔性化、智能化转型。在汽车、电子、新能源等行业，协作机器人、移动复合机器人等已在装配、检测、物料搬运等环节广泛应用，成为发展新质生产力的关键支撑。

2) 生活服务领域。生活服务是最具增长潜力的蓝海市场，核心价值在于弥补劳动力缺口和提升人类生活品质。在商业服务（如智能餐饮配送机器）、医疗康养（如手术、陪护机器人）、家庭服务（如通用保姆机器人）等场景，具身智能展现出巨大的社会与经济潜力，通用人形机器人更有望成为继个人电脑、智能手机后的下一代消费级终端。

3) 特殊环境作业领域。特殊环境作业场景通常具有高危险性、高重复性或人类难以进入的特点，是具身智能替代人类劳动的刚需市场，其核心价值在于保障生命安全与拓展人类作业边界。例如，交通运输中的自动驾驶汽车，安防巡检中的机器人与无人机，以及能源、矿采、农业等领域的特种机器人，正逐步将人类从繁重、危险的劳动中解放出来。

6 经济与组织影响：生产范式与产业结构的双重重塑

作为能够直接作用于物理世界的通用目的技术范式，具身智能的影响超越了传统软件人工智能的边界，正以一种新型生产要素的形态，从根本上改变生产要素的组合方式与价值创造逻辑，其影响将在微观的企业层面和宏观的产业层面渐次展开，引发生产力与产业组织的系统性范式变革。

6.1 微观层面：企业生产函数与组织能力的变革

在企业内部，具身智能的渗透正驱动生产方式从“自动化”走向“自主化”，并催生出与之相适应的新型组织形态与能力，进而引发生产方式与组织管理的系统性变革。

(1) 从静态优化到动态演进：生产函数的内生性增长

传统自动化旨在优化给定的、静态的生产流程，通过机器的“程序化执行”实现了泰勒制下基于静态效率的优化，而具身智能则通过“感知-决策-执行”的实时闭环与持续学习能力，推动生产方式从“自动化”向“自

主化”，进而将生产函数从外生变量转变为能够内生性增长的动态系统^[13]。首先，具身智能赋予了生产系统在执行中学习与自我优化的能力。传统工业机器人严格遵循预设程序，面对环境扰动与任务变更时适应性较差。具身智能不再是严格遵循预设程序的被动执行者，而能通过持续与物理世界的交互，实时分析多模态数据，并利用强化学习等机制迭代自身行为策略。这种“干中学”的能力，使生产系统从线性流程转变为能够自我完善的“智能有机体”，将知识积累与经验优化内化为生产过程本身，实现了全要素生产率的内生性、持续性提升。其次，数据驱动的闭环决策显著提升了资源配置效率。遍布现场的智能体构成物理世界的“神经末梢”，与云端人工智能中枢共同组建了数字孪生系统，使生产管理从事后补救式的“开环控制”，转向基于实时数据与预测分析的“闭环治理”，从而在全局层面优化资源利用，提高资源的配置效率。

(2) 从科层管理到人机共生：组织形态与劳动分工的重构

具身智能的引入正解构传统的科层式管理架构，催生出更扁平、敏捷的组织形态和人机共生的新型协作关系。一方面，组织结构趋向扁平化与网络化。具身智能系统承担了大量重复性的操作与监控任务，并能自主处理多数现场异常，使得传统依赖层层汇报的垂直管理模式变得冗余，其监督与信息传递职能被智能系统部分替代，有效降低了内部交易成本。决策权下沉至由人类专家与具身智能体共同组成的、以任务为导向的“混编团队”，促使组织向更灵活的网络化结构演进。另一方面，劳动分工被重新定义，形成基于比较优势的“人机共生”新范式。具身智能并非旨在完全替代人类，而是在更高层次上形成了认知与体力的优势互补^[14]。人类的价值从直接的物理操作，转向处理高度不确定性、进行创造性设计及对智能系统进行价值对齐与伦理监督等复杂任务；而机器人则负责高精度、高强度的重复性作业。“人类赋智、机器赋力”的共生关系，催生了对兼具工程技术与人工智能素养的新型“灰领”劳动力的巨大需求，对现有教育与职业培训体系构成了深刻挑战。

6.2 宏观层面：产业价值链与商业模式的生态级重构

企业层面的变革汇聚在一起，将通过重塑生产要素的组合效率与价值创造逻辑，最终重构整个制造业的产

业生态与商业竞争格局。

(1) 驱动柔性生产：消解“规模-范围”的二元困境

传统制造业长期面临“规模经济”与“范围经济”的二元困境。表现为大规模标准化生产难以满足个性化需求，而小批量定制化生产则成本高昂。以人形机器人为代表的通用具身智能体，凭借其高度的灵活性与软件定义下的任务泛化能力，为打破这一困境、实现“大规模定制”新范式提供了可能。其核心在于，通用具身智能体作为一种“柔性资产”，通过加载不同算法即可在不同工序间快速切换，极大降低了产线转换成本，使得生产系统能够以接近大规模生产的效率，响应碎片化、多样化的市场需求^[15]。例如，消费者需求通过数字平台直达工厂，智能制造系统实时解构订单、调度生产，并与上下游供应商的智能系统协同，实现物料的准时化供给与产品的快速交付。具身智能驱动的柔性生产将优化整个供应链，从根本上改变了以产定销的传统生产逻辑。

(2) 重塑价值分配：催生数据驱动的新兴商业模式

具身智能将重塑产业价值链的利润分布，并催生以数据和服务为核心的新型商业模式^[16]。首先，价值链的核心正从硬件制造向“软件、算法与数据”迁移。在具身智能时代，物理本体的价值固然重要，但系统的“智慧”程度，即基础大模型、操作系统与上层应用软件，逐渐成为决定其能力边界与附加值的关键。因此，将引致产业利润池的重新分配，上游提供“机器人脑”的基础模型与操作系统开发商，有望占据价值链的制高点。其次，机器人即服务（Robot as a Service, RaaS）模式兴起，重塑商业逻辑。RaaS模式将具身智能体从一次性售卖的固定资产，转变为可按需订阅的运营服务。不仅通过降低应用门槛，加速了技术向中小企业的渗透，也为服务商带来了持续的现金流和宝贵的一线应用数据。这些数据反过来成为训练和优化人工智能模型的“燃料”，形成强大的“数据飞轮”效应，驱动整个产业生态的加速迭代与正向循环。

7 具身智能产业发展的制约与路径选择

尽管具身智能展现出重塑经济社会的巨大潜力，但其从前沿探索迈向成熟的产业化应用，当前仍面临一系列制约因素，构成了技术、应用与生态层面的发展挑战。

为此，需要构建前瞻性的战略布局，以“固本强基、场景驱动、生态共建”三位一体的路径，推动产业实现高质量发展。

7.1 核心制约因素：三大系统性瓶颈

(1) 产业链自主可控挑战：核心基础能力供给不足

具身智能系统的性能上限，根本上由上游核心技术与零部件的供给水平决定。当前我国在该领域仍面临自主可控的挑战，构成了产业发展的首要制约。进一步看，在硬件层面，关键零部件对外依赖度较高。具身智能“身体”的高度精密性依赖于高端执行器与传感器。在高功率密度伺服电机、高精度减速器等执行部件，以及高端六维力/力矩传感器、高分辨率工业相机等感知部件上，部分核心市场仍由少数国际企业主导。硬件层面的依赖不仅推高了产业成本，也对产业链供应链安全构成潜在威胁。在软件层面，底层技术生态根基尚不牢固。具身智能“大脑”的效能依赖于算力、操作系统与基础模型。在面向端侧实时控制的高能效、低功耗专用人工智能芯片，以及具备工业级稳定性的国产机器人操作系统方面，仍存在短板。此外，专为具身交互优化的、深度融合物理规律的多模态具身大模型仍处于探索初期。

(2) 场景“落地难”困境：物理世界的不确定性挑战

从数字世界迈向物理世界是具身智能的本质跃迁，但也因此直面物理世界的复杂性与不确定性挑战。一方面，存在数据匮乏与泛化脆弱的矛盾。具身智能的学习高度依赖在真实物理环境中交互所产生的大规模、多模态数据。相较于互联网上取之不竭的文本与图像，物理交互数据的采集成本高昂、标注困难且场景覆盖有限，导致当前模型训练严重依赖仿真环境。加之，现阶段“仿真到现实”的迁移鸿沟依然显著，模型在虚拟环境中表现优异，但在充满噪声、动态扰动的真实场景中，其感知与决策的鲁棒性往往大幅下降，泛化能力不足成为规模化部署的关键障碍。另一方面，决策风险与安全伦理的挑战。在物理世界，错误的决策可能直接导致设备损坏、人员伤亡等严重后果。如何确保智能体在开放环境中的决策可解释、可预测、可控制，是其从实验室走向规模化应用必须跨越的伦理与技术门槛。目前，针对具身智能的安全验证、风险评估与伦理规范体系尚未完全建立，制约了其在高可靠性要求场景的应用。

(3) 产业“碎片化”制约：创新生态协同效率不高

作为高度交叉融合的新兴领域，具身智能的成熟依

赖于开放协同的创新生态，而当前我国产业生态仍处于培育初期，呈现“碎片化”特征，协同效率不高。一方面，产业链上下游协同不足。人工智能算法公司、机器人本体商、核心零部件供应商与终端用户之间尚未形成紧密的“共生”关系。算法研发与硬件设计脱节、应用需求未能有效传导至技术供给端等问题，导致部分研发成果与市场需求错配，难以转化为具备商业竞争力的产品。此外，统一的行业标准与接口规范的缺乏，加剧了产业碎片化，形成了阻碍技术高效流动的“创新孤岛”。另一方面，复合型人才供给结构性短缺。具身智能的发展亟需大量兼具人工智能、机器人学、机械和材料等多学科背景的复合型人才^[17]。现有以专业分割为主的教育体系难以满足产业对跨学科人才的迫切需求，能够引领技术方向的战略科学家与具备复杂系统工程能力的领军工程师更为稀缺，已成为制约核心技术攻关与产业化进程的关键瓶颈。

7.2 战略路径：“三位一体”的系统性布局

针对上述挑战，需采取系统性、前瞻性的战略布局，以“固本强基”应对技术瓶颈，以“场景驱动”破解应用困境，以“生态共建”弥合产业断点，形成三位一体的发展路径。

(1) 固本强基：构建自主可控的核心技术与零部件体系

夯实产业技术基础，是应对外部不确定性、掌握发展主动权的根本前提。应采用使命导向的创新战略，集中力量攻克核心瓶颈。硬件层面，实施“强链补链”工程，设立国家级科技重大专项，组织“产学研用”联合体协同攻关减速器、伺服系统、高端传感器等关键短板，并通过首台（套）示范应用、税收优惠等政策工具激励企业研发，加速实现高端零部件的性能突破与自主供给。软件层面，启动“铸魂开源”计划，支持研发面向具身智能的低功耗、高性能人工智能芯片与计算架构；鼓励构建国家级开源机器人操作系统与开发平台，提供标准化接口与丰富的开发工具链以凝聚开发者生态；同时，支持建设国家级、高质量的具身交互多模态数据集并开放共享，为算法模型迭代提供高价值“数据燃料”。

(2) 场景驱动：以真实应用为牵引加速技术迭代与数据闭环

具身智能的技术成熟与价值实现，必须在与物理世界的真实交互中完成。应以应用场景为“试验场”和“孵

化器”，牵引技术迭代，形成数据驱动的良好循环。一方面，打造国家级应用示范标杆。聚焦智能制造、智慧物流、医疗康养和特种作业等关键领域，选择产业基础好、市场需求强的区域或行业，打造黑灯工厂、全流程无人化仓储等标杆工程。政府可通过定向采购、应用补贴等方式激励用户端试用国产创新产品，为其提供宝贵的真实场景数据和迭代机会。另一方面，构建“仿真+现实”双循环测试验证体系。大力发展高保真物理仿真平台，为算法提供低成本、高效率的大规模训练环境；同时，建设开放的、多场景的物理测试场，为产品提供标准化的性能评测、安全认证与可靠性验证服务。通过“仿真训练-现实测试-数据反馈-模型优化”的闭环，加速弥合仿真到现实的鸿沟，提升智能体的鲁棒性与泛化能力。

(3) 生态共建：培育开放协同的产业集群与治理框架

具身智能产业的壮大离不开开放、协同、共赢的生态系统。需从产业组织、人才培养和制度建设三方面着手，为产业长远发展提供支撑。第一，培育世界级产业集群。依托京津冀、长三角、粤港澳大湾区等地的优势，引导产业链上下游企业、科研机构与金融资本集聚，形成特色鲜明、优势互补的世界级产业集群。鼓励成立产业联盟，制定共通的技术标准与行业规范，促进信息共享与协同创新，提升产业整体竞争力^[18]。第二，完善复合型人才培养体系。在高等教育阶段，鼓励高校设立“人工智能+X”交叉学科，推动人工智能、机器人、机械和材料等院系深度融合，开展项目式教学；深化产教融合，鼓励企业与高校共建联合实验室，精准对接产业需求。第三，前瞻布局法律与伦理治理框架。政府、行业与研究机构应协同合作，前瞻性地研究制定具身智能相关的法律法规与伦理准则，清晰界定产品安全标准、数据隐私边界、事故责任划分等关键问题，构建清晰、稳定且包容的治理环境，确保技术发展安全、可控、向善。■

参考文献

- [1] 石英. 知识社会学视域下的人工智能：原理、功能及影响 [J]. 人文杂志, 2025 (5): 1-14.
- [2] 霍书全. 人工智能符号接地问题研究的意义和挑战 [J]. 上海师范大学学报 (哲学社会科学版), 2019, 48 (3): 98-107.

- [3] GUPTA A, SAVARESE S, GANGULI S, et al. intelligence via learning and evolution[J]. Nature Communications, 2021, 12 (1): 1-12.
- [4] 任磊, 董家宝, 曾宪超, 等. 数字族谱: 驱动工业具身智能世界模型[J]. 中国科学: 信息科学, 2025, 55 (7): 1748-1765.
- [5] JIN D, ZHANG L. Embodied intelligence weaves a better future[J]. Nature Machine Intelligence, 2020, 2 (11): 663-664.
- [6] 李开阳. 弱具身、具身认知与具身 AI[J]. 自然辩证法研究, 2025, 41 (3): 81-87.
- [7] 刘永谋, 白英慧. 具身智能“身体”实现的四个环节[J]. 福建论坛(人文社会科学版), 2025 (4): 26-33.
- [8] 袁依格, 何卓丰, 李威, 等. 具身智能驱动的智能制造应用发展研究[J]. 中国工程科学, 2025, 27 (3): 67-82.
- [9] 张伟男, 刘挺. 具身智能的研究与应用[J]. 智能系统学报, 2025, 20 (1): 255-262.
- [10] LIU H, GUO D, CANGELOSI A. Embodied intelligence: A synergy of morphology, action, perception and learning[J]. ACM Computing Surveys, 2025, 57 (7): 1-36.
- [11] 马愈卓, 张永飞, 贾伟, 等. 面向计算机视觉的数据生成与应用研究进展[J]. 中国图象图形学报, 2025, 30 (6): 1872-1952.
- [12] 曾凯, 王耀南, 谭浩然, 等. AI 大模型驱动的具身智能人形机器人技术与展望[J]. 中国科学: 信息科学, 2025, 55 (5): 967-992.
- [13] 史丹, 万宏, 薛钦源, 等. 能源央企加快形成新质生产力研究[J]. 北京石油管理干部学院学报, 2025, 32 (1): 15-23.
- [14] 薛钦源, 史丹, 史可寒. 新质生产力的形成逻辑、新质特征和理论要素[J]. 当代财经, 2024 (7): 3-16.
- [15] FAN H, LIU X, FUH J Y H, et al. Embodied intelligence in manufacturing: leveraging large language models for autonomous industrial robotics[J]. Journal of Intelligent Manufacturing, 2025, 36 (2): 1141-1157.
- [16] BROERING A, NIEDERMEIER C, OLARU I, et al. Toward embodied intelligence: smart things on the rise[J]. Computer, 2021, 54 (7): 57-68.
- [17] 韦妙, 汪雨梦. “人—技术”关系视域下生成式人工智能赋能职业教育教法改革探析[J]. 职业技术教育, 2025, 46 (14): 6-14.
- [18] 史丹, 何辉, 薛钦源. 数据分类分级制度与数据要素市场化: 作用机制、现实困境和推进策略[J]. 福建论坛(人文社会科学版), 2024 (4): 58-76.

作者简介: 薛钦源(1993—), 男, 博士, 副教授, 主要研究方向为现代化产业体系、新质生产力。

万宏(通信作者)(1972—), 男, 硕士, 正高级经济师, 主要研究方向为能源政策、新质生产力。

收稿日期: 2025-07-23