

2025

具身智能行业发展研究报告

系列报告之一：具身智能技术与行业应用简析

出品机构：甲子光年智库

智库院长：宋涛

报告撰写：翟惠宇

发布时间：2025年1月

「水木人工智能学堂」

水木AI知识荟 & 交流社群📢

📖 每日分享行业报告、行业资讯等！

🔗 链接海量AI行业精英！

🎉 不定时进行名校名企行活动！

🚀 足不出户，尽在水木AI知识荟！

🔥 扫码添加小编微信，免费进水木AI交流群

交流社群



去噪星球



去噪星球 每日仅需0.5元

公众号：水木人工智能学堂

目录

CONTENTS

Part 01 具身智能的发展背景

P02

Part 02 具身智能的发展现状

P17

Part 03 应用场景和代表厂商

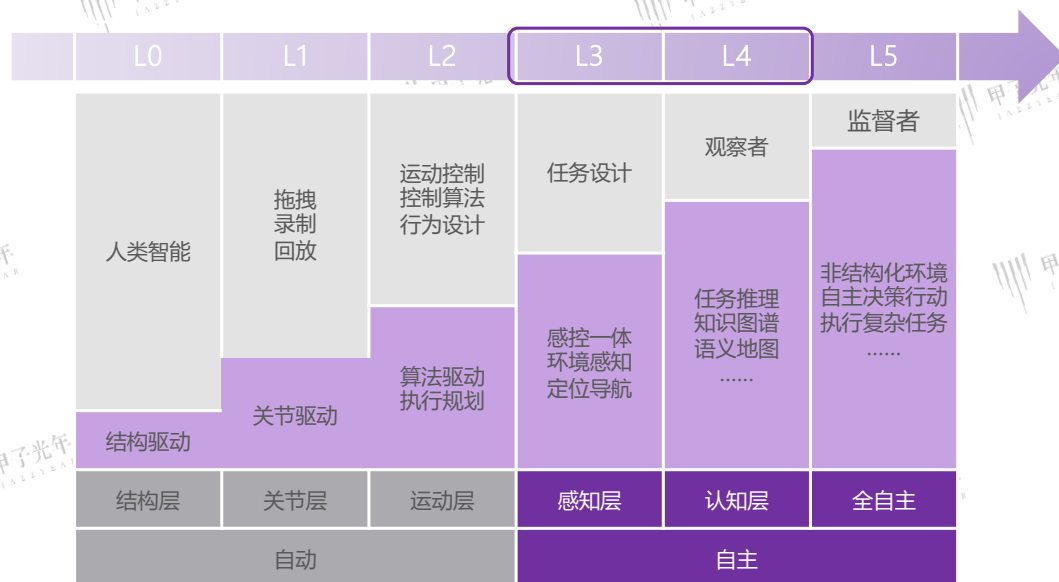
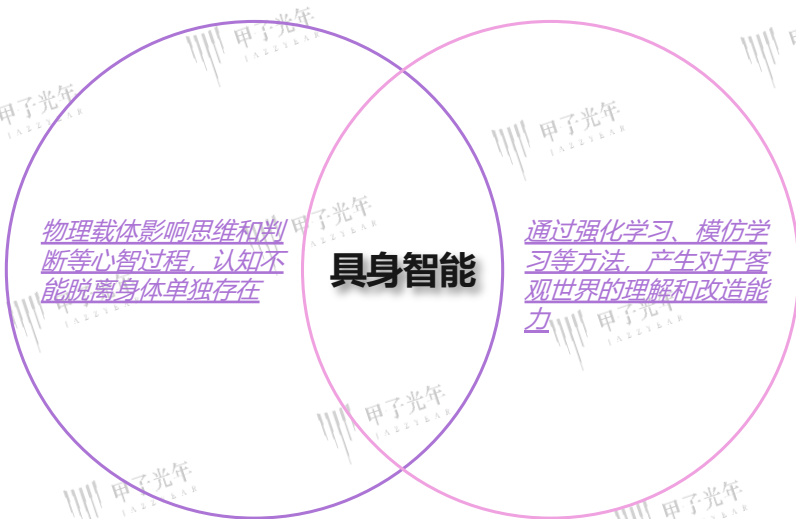
P20

Part 04 发展挑战与技术趋势

P36

具身智能指具备物理载体的智能体，强调智能体与物理环境的交互。

- “具身智能 (Embodied AI)” 指有物理载体的智能体，在与物理世界的交互过程中，通过感知、控制和自主学习来积累知识和技能，形成智能并影响物理世界的能力。
- 其关键在于“具身认知”，即该智能是通过身体和环境的交互过程中产生的，不能脱离实体，这是具身智能与离身智能的关键区别。
- 具身智能主要以各种智能机器人的形态出现，融合了机器人领域与计算机领域下的多个学科，是所有学科发展到相当成熟度之后才涌现出的能力。



以人类为主导

以机器为主导

多因素驱动我国成为具身智能技术与应用的沃土。

- 具身智能机器人是我国发展新质生产力的未来产业新赛道，集成了AI、先进制造、新材料等先进技术，将对社会产业变革和全球竞争格局产生颠覆性影响。
- 无处不在的应用潜力预示着庞大的市场需求、国家与地方长期的政策规划支持、核心零部件供应链完善形成集群效应、以及健康良好的人才和创业生态，多方面因素推动我国成为具身智能技术发展的沃土。
- 当前，国内的高校院所、科技大厂、初创企业均在布局以AI大模型和人形机器人为代表的前沿领域，我国在具身智能领域已经走在国际前列。

市场需求极为庞大

- 全球范围内工业机器人与服务机器人市场规模十分巨大。
- 当前，人口老龄化问题日益突出，老年人口数量不断攀升，同时人口出现了负增长态势。
- 年轻一代不再愿意在恶劣的环境中工作，枯燥、重复、低价值的劳动可交由机器人作业。
- 高盛预测，2035年全球人形机器人的总体可用市场（TAM）将达到380亿美元，预计出货量为140万台。

产业集群效应显著

- 制造业集群产业链完善，可大幅缩短研发周期，从1.5年降至0.7年。
- 靠近供应链与客户，能节省物流及定制时间成本，提升购买舒适度。
- 地方在土地、信贷、财政等方面大力支持。
- 得益于本地供应链规模量产，关键零部件成本快速降低。
- 在北京、上海、广东等地设立人形机器人创新中心。

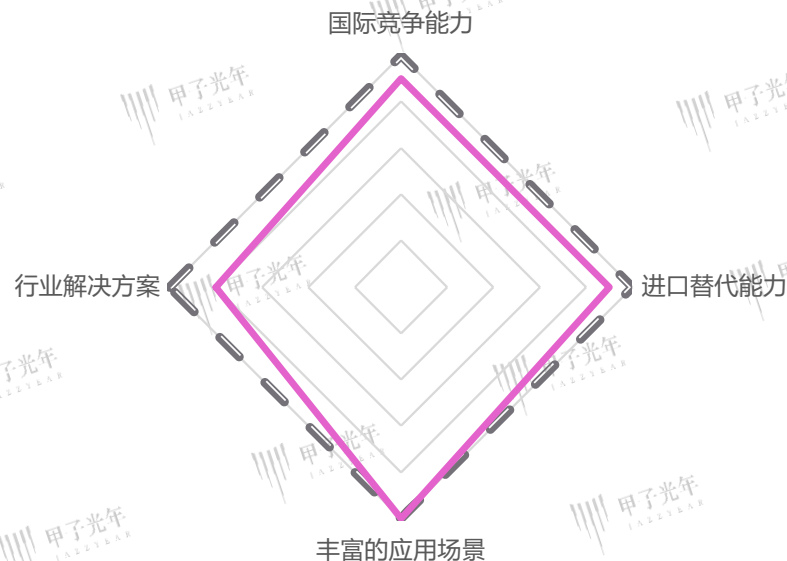
长期国地政策支持

- 国家和地方政策稳定推出并施行。
- 从顶层优化机器人产业结构，促进产业链协同。
- 聚焦底层技术、应用及创新驱动。
- 国家设立中期（2025年完成关键技术攻关）和远期（2027年创新能力显著提升，建成安全可靠本土产业链供应链体系）目标。

人才和创业生态良好

- 超过400所高校高效开设机器人工程专业。
- 机器人企业注册数量迅猛增长。
- 过去十年投资总额超千亿元。
- 多地设立机器人产业基金，如北京机器人产业发展投资基金（规模达100亿元）、上海人形机器人产业基金（20亿元）、吴中区机器人产业基金等。

图：我国在具身智能技术与应用已经走在国际前列



国际竞争能力：在技术和产品上可达到国际一流水平，未来能与国际一流厂商进行直接竞争
进口替代能力：基本做到关键产品和技术自主可控，在封锁时可以提供可用的技术和产品
丰富的应用场景：具身智能产品和技术在国内有广泛且长期的应用场景，未来将生成海量需求市场
行业解决方案：该行业有完整的产业链条，可为实体经济中各下游产业的需求与发展提供对应的具身智能机器人解决方案

政策持续加码，具身智能与人形机器人正走进科技发展的舞台中央。

- 各国均将以人形机器人为代表的具身智能上升到国家战略，通过跟进完善制度建设、提供资金补贴等方式推动技术发展。
- 国内相关政策主要围绕加快具身智能机器人重点场景应用、加强行业标准规范建设、设立专项财政与基金支持、加速人才引进与技术培育等方面，通过政策推动，在短期实现核心零部件的技术突破，在长期建设丰富产业应用和生态。

图：具身智能/人形机器人相关政策不断出台，多维度支持新技术落地

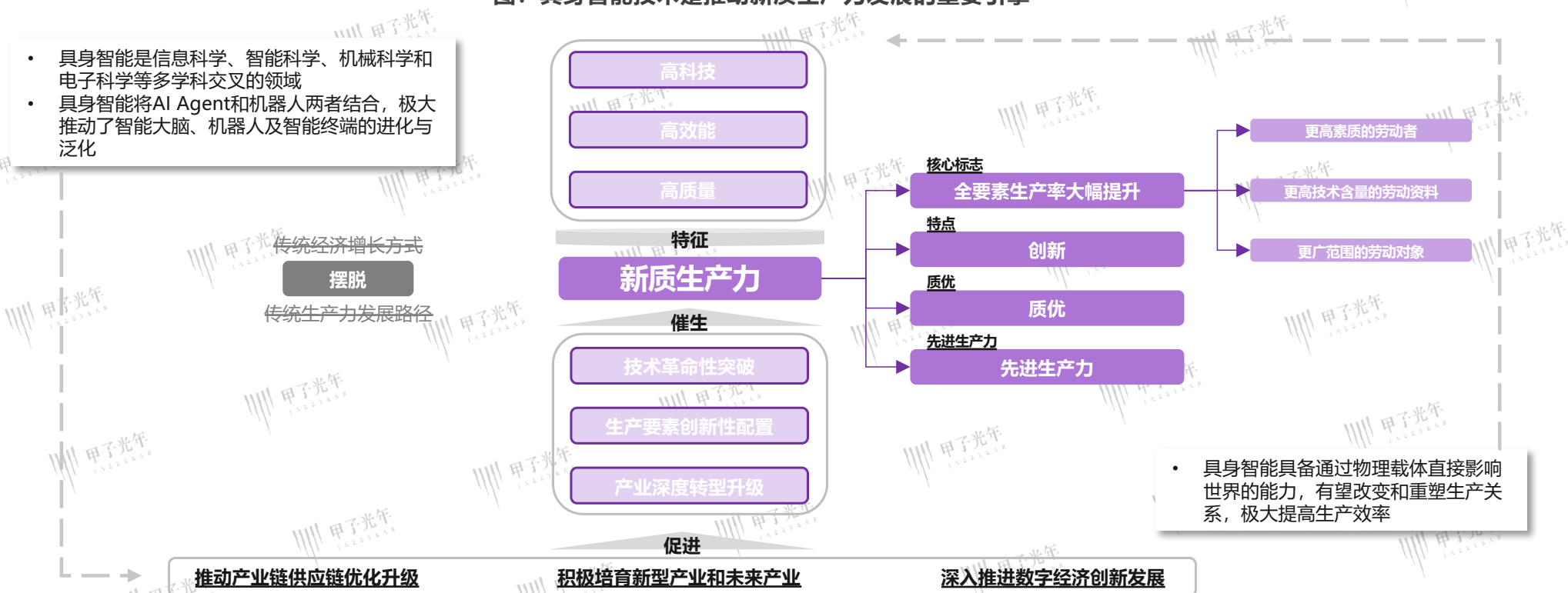
日期	部门	名称	主要内容
2023.1.8	工信部等17部门	《“机器人+”应用行动实施方案》	基金支持、试验中心、加速研发与推广
2023.5.31	深圳市委办公厅	《深圳市加快推动人工智能高质量发展高水平应用行动方案（2023-2024年）》	人才供给、加速研发、提升产业集聚水平
2023.6.15	上海市人民政府办公厅	《上海市推动制造业高质量发展三年行动计划（2023-2025年）》	核心部件攻关、加速应用推广、人才梯队
2023.6.28	北京市人民政府办公厅	《北京市机器人产业创新发展行动方案（2023-2025年）》	培育新产品、加紧布局、校企合作
2023.8.3	工信部等4部门	《新产业标准化领航工程实施方案（2023-2035年）》	研制人形机器人相关标准，包括术语、核心零部件、智能感知决策、运动控制、安全和应用等方面
2023.10.26	工业和信息化部	《人形机器人创新发展指导意见》	技术突破、重点产品、应用场景、人才引进
2024.1.18	工信部等七部门	《关于推动未来产业创新发展的实施意见》	加速创新和产业化、基金支持、应用场景、人才引进
2024.7.6	上海	《上海市人形机器人治理导则》	促进AI产业健康发展、确保人形机器人合乎人类道德和伦理价值

Source：各政府网站，甲子光年智库。

具身智能具备新质生产力的关键内涵，是推动新质生产力建设的重要引擎。

- 新质生产力强调在科技创新的基础上大幅提升生产效率和创新能力，而具身智能技术作为人工智能和机器人技术的融合发展前沿，天然具备形成新质生产力的关键内涵，是推动新质生产力发展的重要引擎。
- 具身智能技术契合着解决未来社会矛盾的刚性需求，有望解决如工场劳动力短缺、社会老龄化等问题，把人类劳工不愿从事的枯燥、高危劳动中解放出来，并进一步推动全社会生产关系的改变与重塑。关注具身智能技术的发展对于我国实体经济高质量增长和国际产业竞争格局有着重大意义。

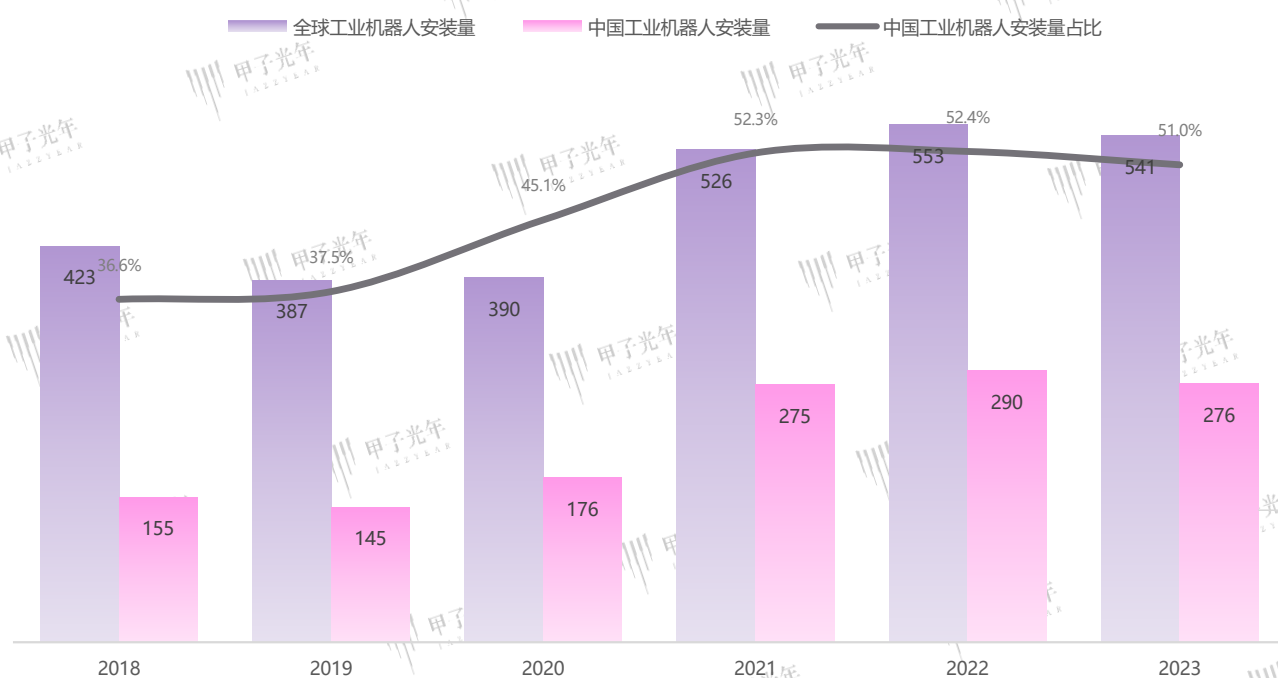
图：具身智能技术是推动新质生产力发展的重要引擎



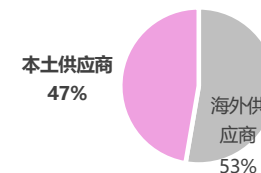
智能化技术将进一步放大机器人对工业领域人类劳动力的替代作用。

- 全球工业机器人整体仍然保持增长态势，其中中国工业机器人安装量一枝独秀，2023年安装量高达27.6万台，占全球安装总量的51%。
- 与此同时，来自国产品牌的工业机器人供应商也得到了快速发展，本土供应商已经近乎占领一半的国内市场。
- 人工智能与机器学习技术的进步，将进一步放大机器人对人类劳动力的替代作用，在灵活、柔性、非标的层面展现更大的应用价值。

图：全球vs中国工业机器人安装量（单位：千台）



中国工业机器人市场占比

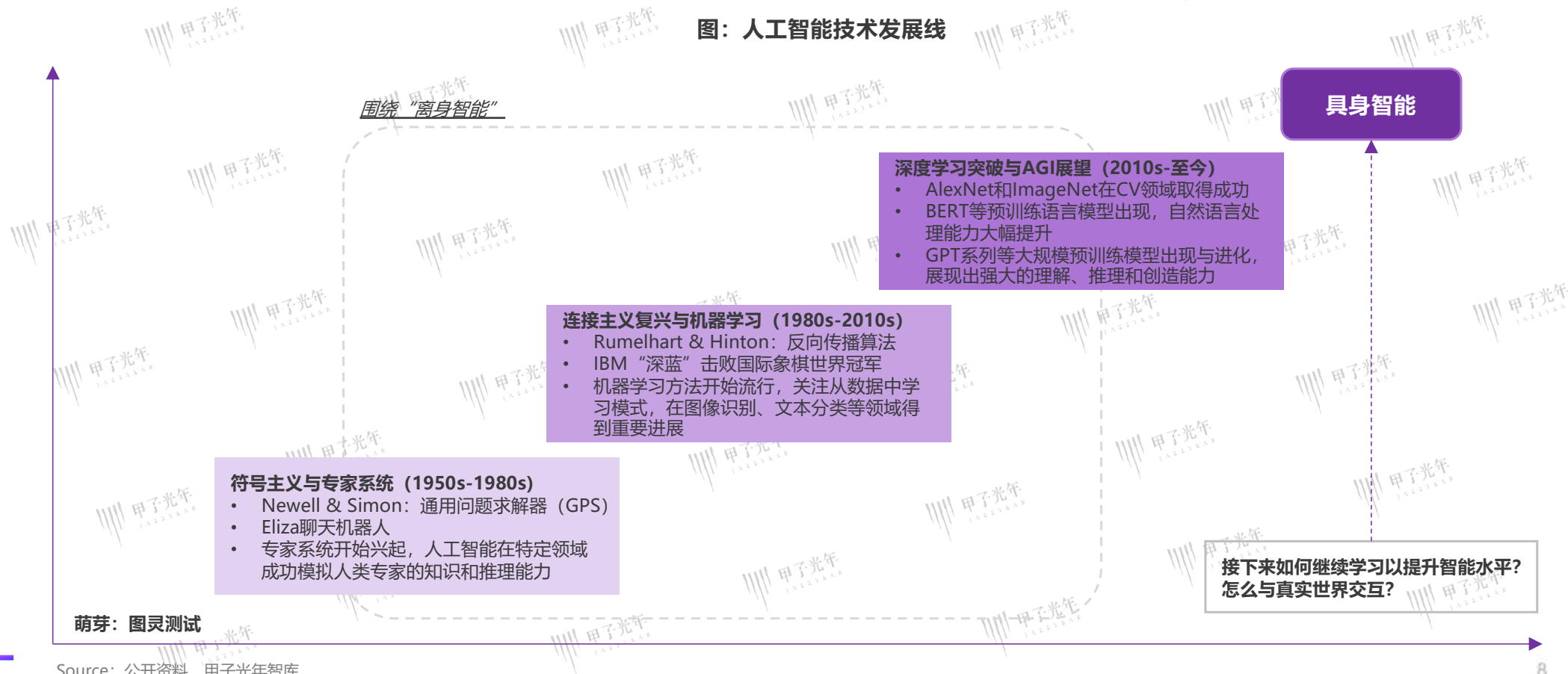


- 中国是工业机器人领域的全球**最大市场**
- 本土制造商在国内市场份额**快速提升**
- 供给侧改革稳步推进，高端制造业快速发展，国内制造业仍然有较大发展潜力，国内工业机器人市场规模有望**持续提升**
- 随着具身智能技术的加持，机器人的**自主决策能力**将快速提升，能够适应**更复杂、更多变**的任务

人工智能发展线：从理性主义到深度学习

- 在机器人领域的应用上，生成式AI技术正在兑现提升机器人关键能力的潜力，在环境感知、自主决策、学习与适应等多个方面均有表现。
- 基于网络数据训练的大模型似乎正在触及AI认知的边界，Scaling Law下的预训练大模型性能似乎快要达到极限，参数量和数据量的扩充已经难以给模型带来质的突破。AI想要进一步发展，**必须与物理世界建立更加紧密的联系**，具身智能将是AI继续向AGI进步的关键途径。

图：人工智能技术发展线



Source: 公开资料，甲子光年智库。

机器人发展线：从自动化工具到直接影响物理世界的智能体

- 从历史的角度来看，工业机器人的硬件进化不断发展，从自动化设备到移动机器人，从协作机器人到人形机器人，每一步都代表着技术的进步和创新。当下，已经进入了AI大模型与机器人在底层技术创新和应用融合的交叉时代。
- 技术的每一次进步都会带来机器人应用场景的突破，且一般会在精准、重复的工业场景作业流程中率先得到应用。未来，如何提升智能机器人在开放场景的可用性，是具身智能正在解决的问题。

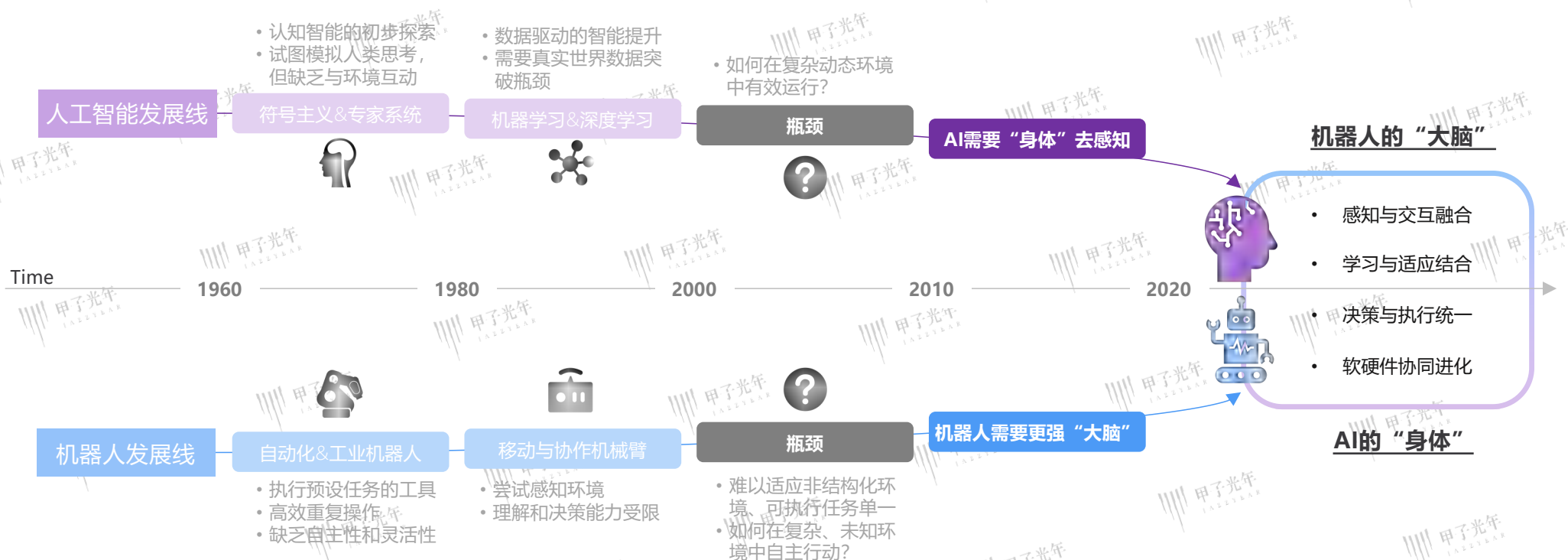
图：机器人技术发展线



具身智能：两条发展线的汇聚点，“AI的下一个浪潮是物理AI”

- 人工智能和机器人技术的发展并非孤立进行，而是相互促进、共同演进的。人工智能为机器人赋予了“大脑”，使其具备感知、思考和决策能力；而机器人则为人工智能提供了“身体”，使其能够与真实世界进行交互，获取经验和知识。
- 具身智能的兴起是人工智能和机器人技术各自发展到一定阶段的必然结果，是两者深度融合的体现。人工智能的持续进步需要与物理世界的交互能力，而机器人的未来发展也离不开更高级别的智能化水平。

图：具身智能是AI与机器人两条线的交汇



目录

CONTENTS

Part 01 具身智能的发展背景

P02

Part 02 具身智能的发展现状

P11

Part 03 应用场景和代表厂商

P20

Part 04 发展挑战与技术趋势

P36

大模型与GenAI飞速发展，开启具身智能技术萌芽。

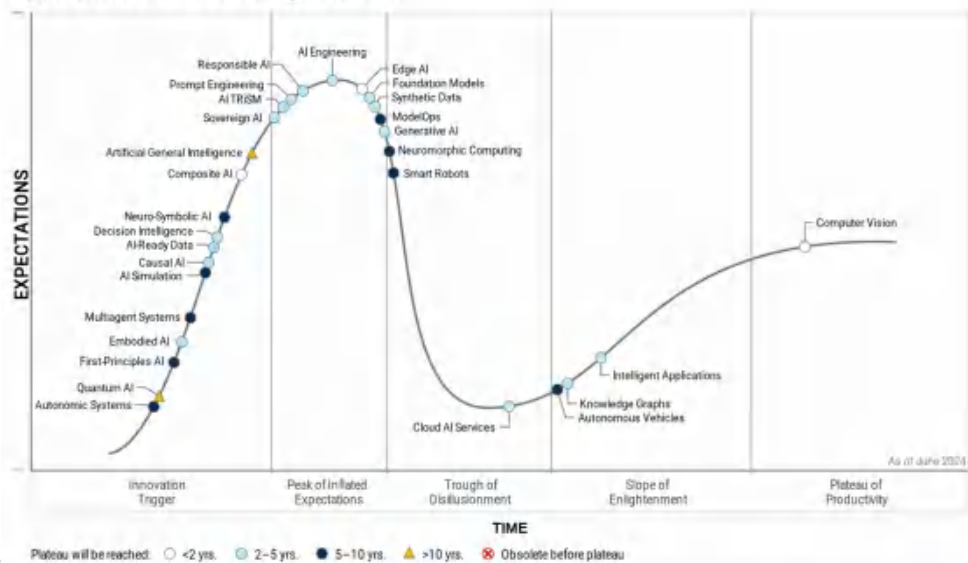
- 随着人工智能算法的发展，尤其是多模态大模型技术的突破性进展，将显著加速机器人产业的发展，提升机器人的智能水平，使得机器人能够自主进行判断和识别，执行复杂的多阶段语义推理任务。不仅提高了机器人的泛化能力，也快速推动了人形机器人通往量产的进程。
- 根据Gartner技术成熟度曲线，当前的具身智能仍处于技术萌芽期，虽然受大模型技术的推动成为科技产业的热点，但是初创公司技术和商业化路径还没有找到有效方案，在成本、技术等层面仍然需要攻坚克难。

图1：大模型与Agent技术是机器人产业突破临界点的关键



图2：具身智能尚处于技术发展的萌芽期

Hype Cycle for Artificial Intelligence, 2024



Gartner

智能机器人表现的提升依赖于软硬件多种技术的快速进步与降本。

- 以人形机器人为代表的具身智能机器人是链接虚拟数字世界和现实物理世界的最佳载体，是虚实融合的理想产物，其性能表现也高度依赖于软件与硬件的全方位进化。
- 从基础材料科学到人机交互技术，从多模态感知大模型到高精度机器人运动控制算法，软硬件多个层面的技术积累与快速进步，让具身智能在2023年后迅速让世界看到其应用潜力与应用可行性。

多模态感知&传感器

- 感知信息从低精度到高精度，从单模态到多模态
- 具身智能的实现需要做到视觉、听觉、温度、力度等多种模态的感知融合，方能全面地感知动态环境

感知结果&任务理解

- 大语言模型等单一模态模型，或图文大模型等双模态模型，并不能实际解决具身智能机器人需要实现的作业任务
- 多模态的感知数据需要多模态大模型进行深度融合，方能为后续的决策、规划和控制模块提供实时的精准时空融合感知
- 大模型与行业know-how、业务系统也需要实现深度融合，以更好地让机器人理解作业任务

自主决策与规划

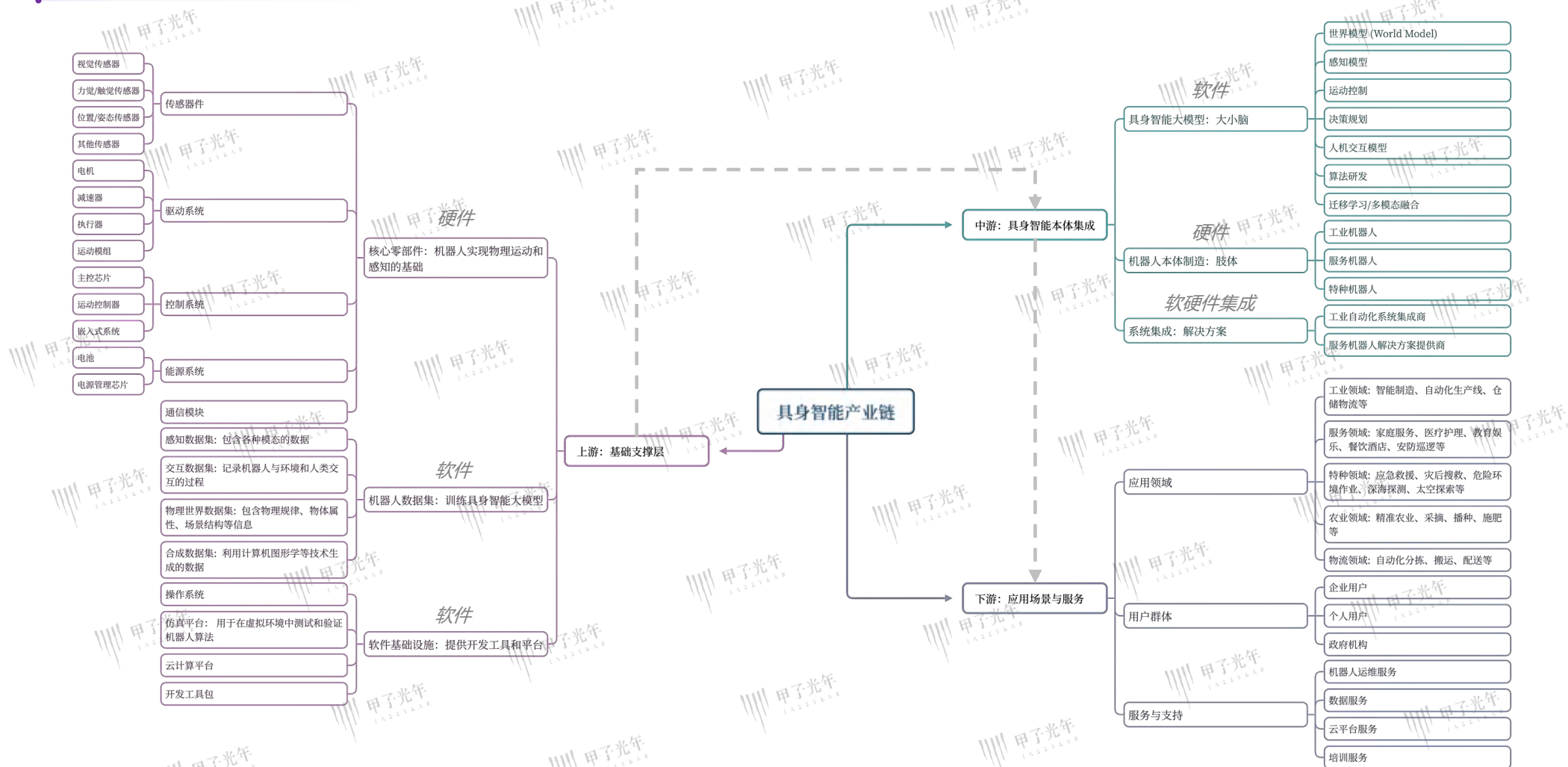
- 基于全面的环境感知与精准任务理解做出决策最优解
- 类似multi-agent技术发展趋势，智能机器人也需要实现单体智能到群体智能的进化

机器肢体&运动控制

- 一方面基于材料科学的进化，推动核心零部件的性能与成本表现
- 另一方面利用模仿学习、强化学习等方法，通过大数据集、合成数据等手段加速机器人控制算法的迭代
- 仿真技术的进步也为运动控制算法与机器人应用开发提供了一种有效手段



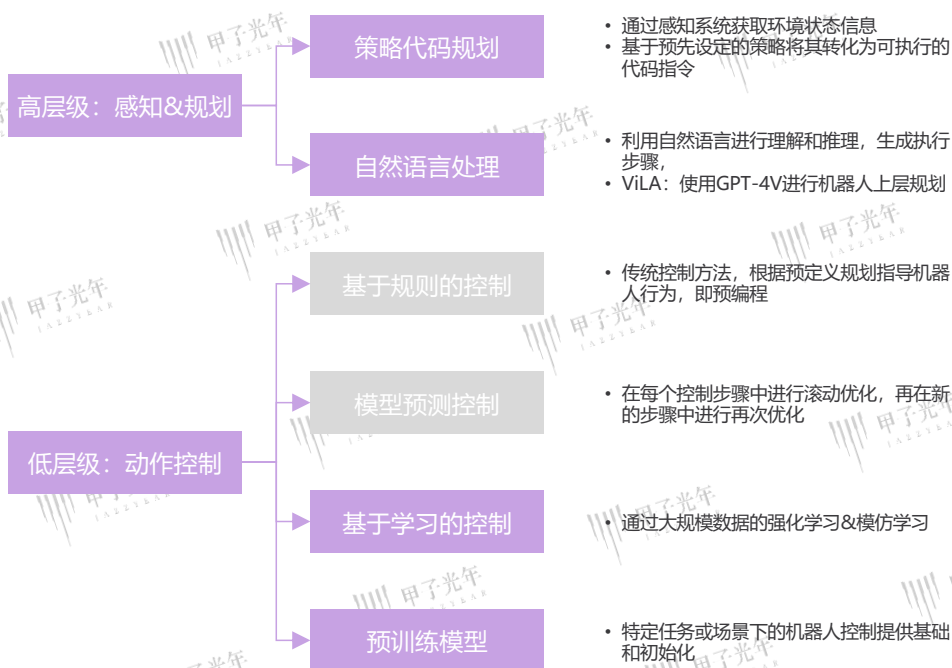
具身智能产业链：软硬件深度融合，复杂且充满活力的生态系统



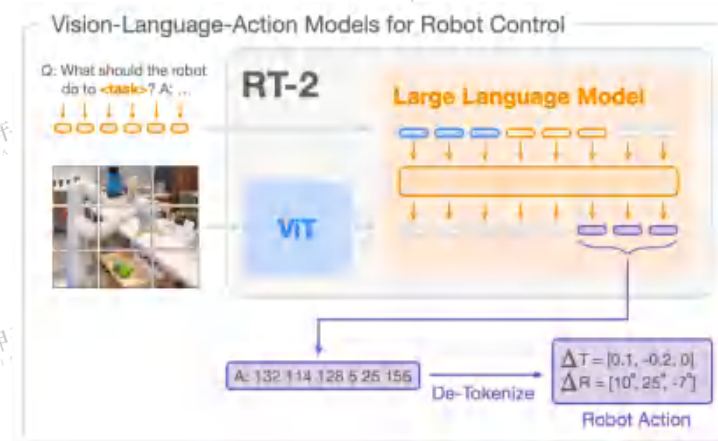
分层模型or一体化端到端，机器人面临与自动驾驶一样的选择题。

- 从“大脑-小脑-肢体”的架构来看，分层模型通过不同层次模型协作，利用底层硬件层和中间响应快的小模型弥补上层大语言模型的不足，推动机器人在实际场景的应用，并越来越多地采用基于学习的控制方法。
- VLA等端到端模型能够直接实现从人类指令到机械臂执行，即输入图像及文本指令，输出夹爪末端动作。
- 分层式路线一度是主流选择，因为端到端路线受数据制约难以达到性能要求；机器人数据正在逐步积累，端到端路线在未来可能成为主流。

路线1：分层模型路线，不同层级由多个神经网络分别训练优化再联合



路线2：端到端模型路线，从目标输入到指令输出



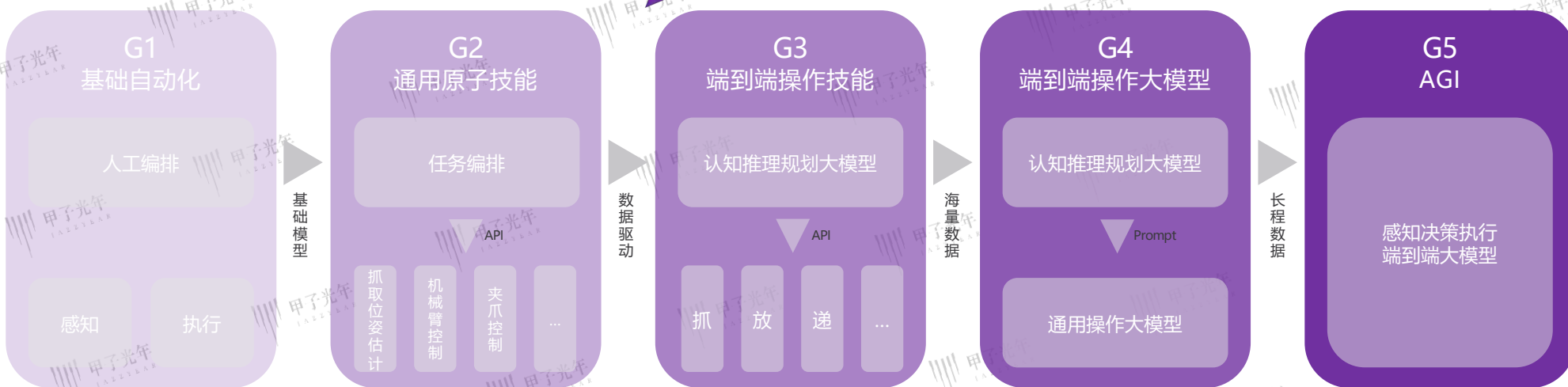
- 以RT-2为例的VLA模型，大规模数据下能够实现泛化能力
 - Input: 图像及文本指令
 - Output: 输出是夹爪末端动作
 - 直接端到端地实现从人类指令到机械臂执行
- 端到端模型的难点在于数据采集难：类似于过往自动驾驶技术的发展，在没有足够的车辆上路之前，极难实现数据的飞轮效应
- 现行资源下的计算效率低下

通往One-Model端到端大模型是个循序渐进的过程。

- 数据规模、模型泛化性、响应速率等问题是当前端到端具身大模型需要逐步攻克的弊端，需要在数据采集和模型训练上寻求新的突破方法。
- 根据智元机器人定义的路线图，当前具身智能大脑已经具备认知、推理、规划的能力，且而小脑层面机器人技能任务的训练也是通过深度学习以驱动的。随着场景、数据的增多，多个特定小模型会逐渐泛化为通用操作大模型，并最终实现与上层模型的融合。

图：具身智能的发展阶段，从模块化走向One Model端到端

当前行业前沿所处的阶段



具身智能阶段

采集方式不断进化，仿真数据有望驱动机器人实现智能跃迁。

- 当前的具身智能是通过深度学习技术（模仿学习、强化学习等）从大量数据中学习并实现的，巨量的、优质的、多样的数据，是具身智能机器人能够实现各种各样自主操作的基础。
- 数据采集方式呈现多元化，人工示教（遥操、动捕）、机器人自主探索、仿真合成数据等方式各有优劣，为具身智能训练数据集添砖加瓦。
- 自动驾驶的数据采集方式具备借鉴意义：在规模化部署以后，机器人在影子模式下自主探索，通过与环境互动收集多元数据，在云端形成数据闭环，再反哺模型的调整与训练，这也是具身感知、学习的具体表现。

训练具身智能的常见数据采集方法及优劣势

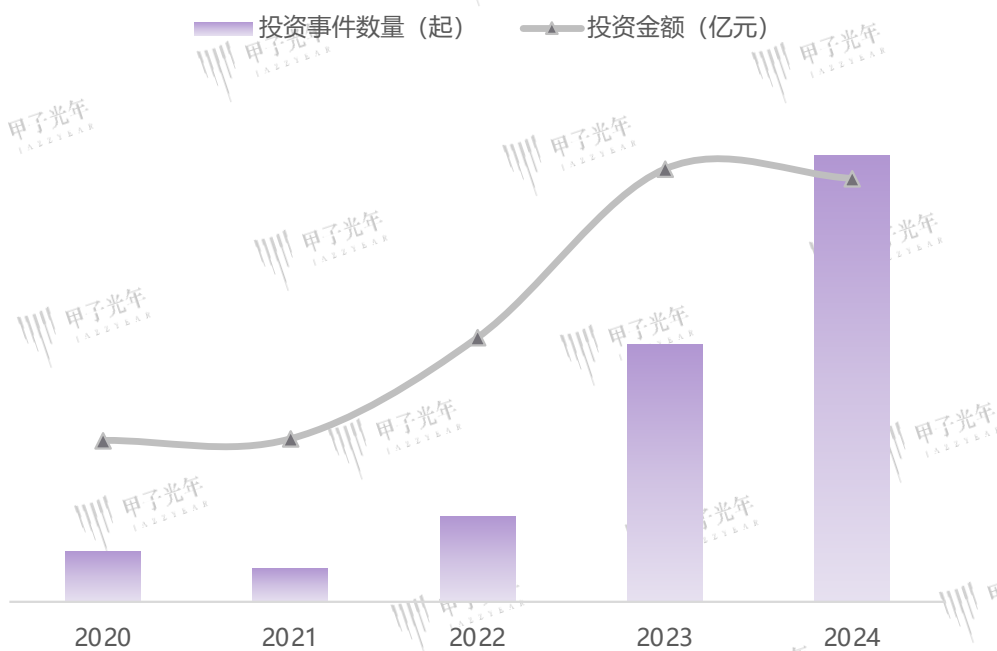
	方法	优势	劣势
起步阶段（过去几年） 人工示教与轨迹记录	- 操作人员直接通过遥控器、手柄、力反馈设备或其他界面控制机器人执行任务 - 同时记录机器人的关节角度、末端执行器位姿、传感器数据	- 直观易懂、快速采集 - 捕捉人类经验	- 成本高、耗时长 - 场景难以全覆盖
发展阶段（当前主攻方向） 仿真环境生成数据	- 在虚拟环境中（如世界模型）模拟机器人的工作状态，生成大量的感知数据和动作数据 - 可以控制环境的各种参数，利用大模型的泛化性，生成各种各样的数据	- 低成本、高效率 - 可控性强 - 生成的数据自带标注	- Sim-to-real gap - 仿真环境本身难以构建 - 物理现象难以模拟
规模化阶段（未来规模化部署后） 机器人自主探索与交互	- 允许机器人在环境中自由探索，通过与环境互动来收集数据。 - 真·具身感知：观察其对环境对自身状态的影响。	- 无需人工过多干预 - 能够覆盖更多场景 - 经过不断试错学习	- 数据质量难题 - 需要优秀的奖励函数设计 - 需要很多实体在现实层面部署
开源数据集&预训练模型	- 利用已有的、公开可用的数据集进行训练 - 或使用在其他任务上预训练好的模型作为基础，进行迁移学习或微调	- 降低采集成本 - 解决自身采集的数据稀疏问题	- 缺乏标准，数据不一定适配 - 标注质量参差不齐 - 预训练模型迁移学习表现不佳

Source: 甲子光年智库整理。

2024年迎来具身智能投资热潮，本体与具身模型受到资本青睐。

- 如果说OpenAI的ChatGPT引爆了2023年对大语言模型的投资热潮，那么Tesla入局人形机器人和黄仁勋的那句“AI的下一个浪潮是具身智能”，则彻底带火了具身智能与人形机器人领域的投资热潮，成为2024年科技产业投资的最大热点。
- 当前具身智能机器人可用性不足的原因还是“智能化”水平不足，算法环节无法达到要求。从投资角度看，头部人形机器人的本体研发集成厂商已经经历多轮融资，估值较高，行业整体投资热点正从人形机器人本体向具身智能模型和其他上游零部件迁移。

图1：国内人形机器人投资事件数量与金额变化（2020-2024）



Source: IT桔子（统计口径或另有出入，仅选取公开融资信息），公开资料，甲子光年智库整理。

图2：从本体制造到具身大模型，投资热点有所迁移

国外市场率先转向

公司名称	融资时间	融资金额	投后估值
1X Technologies	2024 年 1 月	1 亿美元	未提及
Figure AI	2024 年 2 月	6.75 亿美元	26 亿美元
Skild AI	2024 年 7 月	3 亿美元	15 亿美元
Physical Intelligence	2024 年 11 月	4 亿美元	24 亿美元

国内资本于2024年下半年跟上

公司名称	融资时间	融资金额
千寻智能	2024 年 8 月	近 2 亿元
穹彻智能	2024 年 9 月	累计数亿元
星云智慧	2024 年 10 月	2000 万元
星海图	2024 年 11 月	超 2 亿元
自变量机器人	2024 年 11 月	亿元级

*仅选取部分初创公司，大厂未包含在内。

技术发展一脉相承，车企加速部署人形机器人。

- 自动驾驶是具身智能的一个重要场景，都具备“感知-决策-规划-控制”的算法架构，具身智能和自动驾驶在技术实现路径上是一脉相承的，且算法与零部件可实现高度复用，这是Tesla给业界带来的启示。
- 当下，智能辅助驾驶已经跑通商业化路径，进入卷交付、卷规模、卷性价比的阶段，这为具身智能未来的技术打磨路径和商业化提供了一些参考。
- 同时，汽车行业从主机厂到供应商，从投资人到创业者，都在从自动驾驶转向追求更多场景的具身智能：主机厂或自研或投资具身智能公司，而人形机器人企业也在寻求机器人“进场打工”的机会。

图1：自研或投资，国内车企与头部供应商加速人居具身智能产业

车企	基因	产品或业务布局
广汽集团	传统主机厂	第三代轮组人形机器人GoMate
奇瑞汽车	传统主机厂	联合Aimoga研发Mornine
比亚迪	传统主机厂	自研、投资、合作
上汽集团	传统主机厂	对外投资
北汽集团	传统主机厂	对外投资
长安汽车	传统主机厂	预计2027年前发布人形机器人
小鹏	造车新势力	AI人形机器人Iron
小米	造车新势力	自研Cyberone
华为	智驾供应商	成立具身智能产业创新中心
宁德时代	动力电池供应商	自研、投资，机器人电池
地平线	智驾供应商	计算芯片与开发者套件

图2：是供给也是需求，车企与humanoid的双向选择

机器人厂商	车企	当前应用场景
自研自用		
Tesla	Tesla	分拣电池、搬运等其他任务
小鹏	小鹏	抓取、组装、推车

车企提供场景

Figure AI	宝马	搬运拿取零部件，放置与组装
优必选	比亚迪、吉利、东风、一汽、蔚来等	搬运、质检、分拣、组装
宇树科技	吉利、蔚来	搬运
傅里叶	上汽	零部件组装与操作
Apptronic	奔驰	取物、搬运、装备

- 机器人进厂多为战略合作的试点开展，主要目的仍然是数据采集与训练
- 车企制造流程高度标准化和流程化，能够为机器人提供高度确定性的封闭场景
- 人形机器人可以在该场景内通过重复性高的动作学习与泛化，丰富制造场景任务库

目录

CONTENTS

Part 01 具身智能的发展背景

P02

Part 02 具身智能的发展现状

P17

Part 03 应用场景和代表厂商

P20

Part 04 发展挑战与技术趋势

P36

具身智能行业厂商图谱

源于不同技术背景的科技企业聚焦具身智能，其技术基因、产品形态、性能表现及应用场景的差异性正在帮助它们融入工业、服务、特种应用等各类应用场景。在AI、机器人及多元领域参与者的共同作用下，我国具身智能机器人行业正经历快速成长与扩张。

人形机器人企业

- ✓ 同源技术积累
- ✓ 软硬件研发能力
- ✓ 人形/四足仿生机器人、服务机器人，以及上游零部件等跨界玩家



工业具身智能企业

- ✓ 设计、运控技术同源
- ✓ 工业场景需求洞察
- ✓ 机械臂、协作机器人



科技大厂 & AI企业

- ✓ AI算法技术优势
- ✓ 资金和生态优势
- ✓ 互联网/科技企业



汽车企业 & Tier 1

- ✓ 智驾技术架构
- ✓ 离散制造场景



具身智能创新中心

- ✓ 地方性产业链企业共建
- ✓ 地方性机构、基金支持



Source: 公开资料，仅例举了行业内部分代表性企业且不分先后，甲子光年智库整理。

载体形态多样，场景需求决定具身智能的落地形态。

- 具身智能的物理载体形态呈现多样性，根据其移动特性或方式，可以分为固定底座机器人、轮式（履带式）机器人、足式机器人、仿生机器人等，其中近期火爆的人形机器人属双足人形。
- 具身智能具体形态的选择上需要洞察场景实际的需求，并不存在“最佳形态”的单一解，但人形机器人具备多场景的“泛用性”，在具身智能的验证、以及特定场景中的最终应用有既定优势。
- 在实际应用场景中，不同形态的机器人可以进行有机地组合以满足具体的场景需求，如在轮式机器人上配备机械臂或灵巧手，配备轮式底座的机械臂既具备了轮式机器人移动迅速、能效高的特点，也兼备机械臂和灵巧手精细操作的优势。

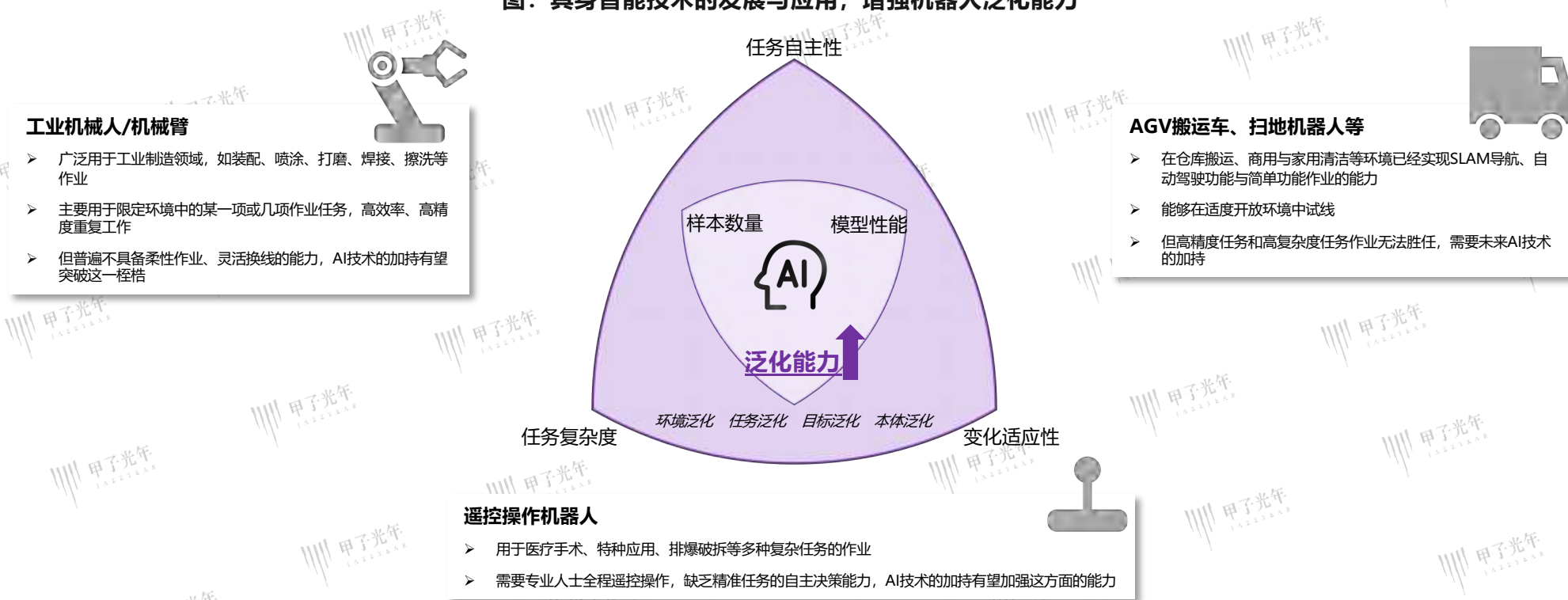
图：具身智能的形态分类



AI助力机器人泛化性能提升，突破能力三角制约，进入更多场景。

- 基于人工预先编程的自动化设备，机器人已经被广泛用于工业制造领域并且已经相当成熟，但受限于预编程（或反复示教）的时间成本、机器人的智能化水平等原因的限制，机器人的应用仍然受限于“任务自主性”“任务复杂度”“变化适应性”组成的能力三角形。
- 大模型技术的快速发展，为具身智能技术的进步与应用提供了历史性的突破机遇，在提升机器人智能化水平的过程中，机器人将更全面、精准、敏捷的进行环境感知、任务理解和准确执行，不断提升机器人在复杂环境中处理复杂任务的能力，泛化水平将不断提升。

图：具身智能技术的发展与应用，增强机器人泛化能力



从专用到通用，从ToB到ToC，具身智能机器人应用场景持续拓宽。

- 作为人工智能与机器人技术的最前沿，具身智能的应用场景会极度多元化，极具想象空间。
- 在工业制造场景，机器人的任务执行和流程和任务本身有高度规则性（可通过编程和训练来高效完成重复性高和强度大的任务）。其中，人类不愿意做的脏活、累活、危险任务，机器人具备极高的劳动力替代性。而服务场景更加开放，相关任务更加多样化，更加不可预测，需要机器人具备更强的自主决策能力与泛化适应性。



应用场景从制造业导入，向商业和家庭服务场景逐步渗透，远期有望应用于航天航空等极限环境。

• 趋势1：从ToB到ToC

- 随着人形机器人在媒体和日常生活中的曝光度增加，消费者对人形机器人的接受度和期待值也在提高。
- 核心零部件的技术进步使得生产成本降低，使得产品价格更加亲民，从而推动人形机器人进入家庭和个人消费市场。
- 消费者对于智能化服务的需求将日益增长，如家庭服务机器人、教育机器人、陪伴机器人等。

• 趋势2：从专用到通用

- 核心零部件的技术和人工智能、机器学习和自适应控制算法的发展使得机器人能够更好地理解和适应多变的环境。
- 随着社会对机器人应用的需求增加，单一功能的机器人已无法满足市场的多样化需求，需要机器人具备更广泛的应用能力。

具身智能率先落地的场景：工业制造

- 工业制造流程的特点使得该场景在具身智能技术落地应用方面占据先发优势，如柔性生产需求迫切、工作环境结构化程度高、成本效益优势突出等需求特点，刺激着工业制造场景客户对工业具身智能的应用更加期待。
- 工业具身智能机器人能够有效提升工厂生产任务的灵活性与适应性，并在作业过程中实现自主学习，不断增强其复杂任务执行能力与操作精度。
- 具身智能的训练需要数据，而质量更高的数据需要从真实场景中来，机器人落地最快的场景仍然是场景相对封闭的工业制造场景，越早实现规模化落地，就有望越早实现模型的能力提升。

工业制造场景率先应用具身智能的原因

对灵活生产需求性高

传统机械臂

- 交付周期长、灵活性差、维护成本高

环境结构化程度高

生产环境

- 固定且相对结构化：产线布局、作业流程更规律
- 机器人更容易感知与作业，落地性更强

积累数据

- 结构化场景下，在安全底线上，更易采集机器人感知与操作数据，方便后续能力持续提升

成本效益优势突出

降本增效

- 提高生产效率与标准化
- 减少人力成本，缩短生产周期

商业化动力强

- 对核心零部件的要求不高，满足工况使用即可
- 工厂客户成本可控，效益可见，更愿意投入

具身智能机器人率先在工业制造场景创造价值

生产排单 灵活性↑ 适应性↑

- 快速响应生产线切换和参数调整
- 适应当下多品类、小批次的生产

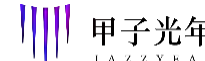
增强执行能力与精度

- 具备更精准的感知、更强的认知与规划能力
- 提升效率的同时，处理复杂任务的能力不断提升

智能决策 & 自主学习

- 多模态大模型驱动，能够自主拆解复杂任务并决策实施
- 持续收集和分析数据，并优化自身性能，不断提升智能化水平

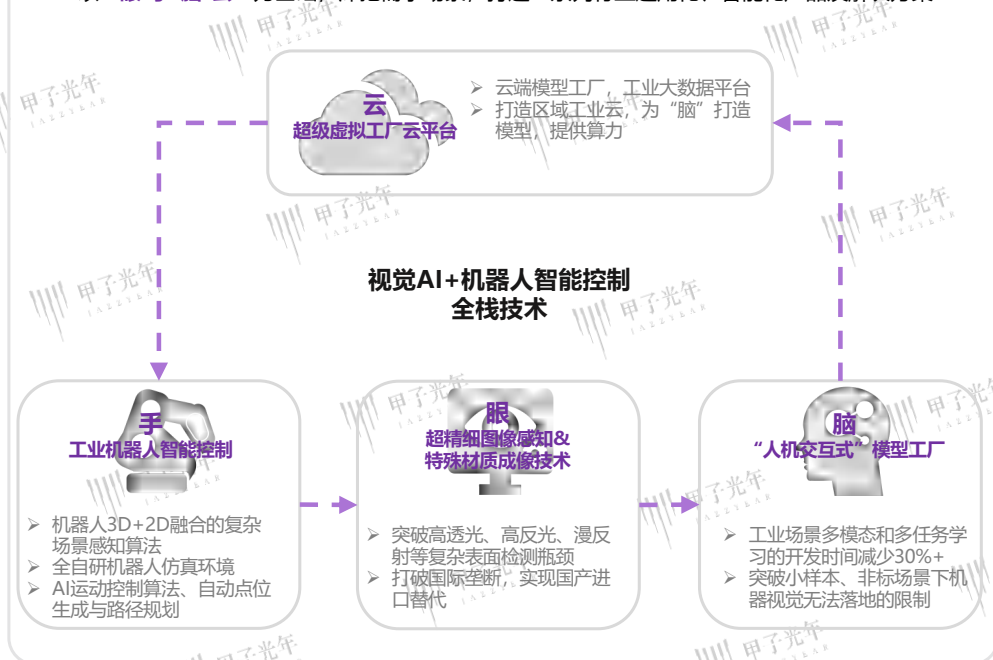
具身智能工业场景应用代表厂商：微亿智造



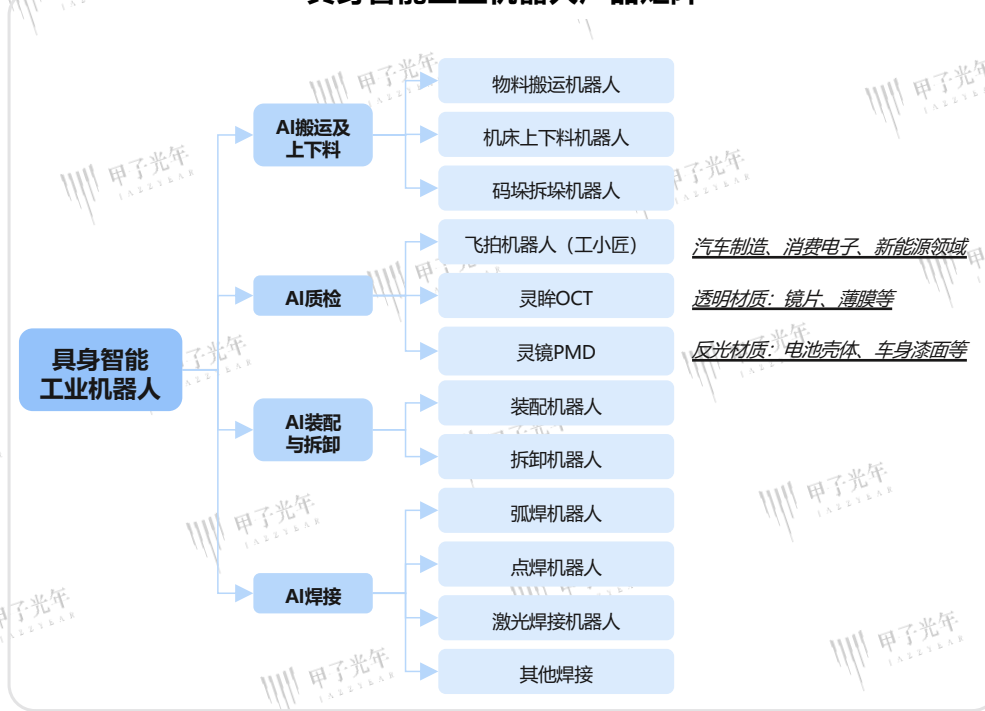
- 微亿智造致力于以工业人工智能及大数据技术，助力工业企业快速实现数智化转型升级。
- 长期布局“工业AI算法+超精细视觉感知模组+工业机器人智能控制”的全栈技术，以“**眼-手-脑-云**”的实施架构为基础，打造一系列将感知算法与驱控算法相结合的“软硬一体”智能化产品，广泛适用于质检、打磨、抓取、搬运及焊接等各类工作，实时感知并适应复杂变化的工业环境。
- 已组建面向3C、汽车、新能源等行业大客户的销售及解决方案团队，建立完善的服务体系，并与京东科技、中国移动、浪潮、中国联通等合作，共同推出服务于全国中小型制造企业的平台化解决方案。

核心技术架构

➢ 以“**眼-手-脑-云**”为基础，深挖需求场景，打造一系列行业通用化、智能化产品及解决方案

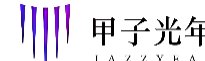


具身智能工业机器人产品矩阵



Source: 微亿智造。

应用案例：具身智能工业机器人“创Tron”



- 客户为保定市某电气公司，专注于电力系统用大型配电柜、储能柜及其配套设备的供应。该公司拥有SMT线、涂敷线和机箱装配线等生产线，其中手插件装配每片PCB板耗时180至300秒，其他产品生产节拍各异且常需更换型号。
- SMT产线单日处理超过100种产品，全年超2000种，且每年新增500多种新品类，导致产线频繁换型。引入具身智能工业机器人后，显著降低了调试成本，提升了生产线的灵活性与适应性，成为离散制造业新质生产力的最佳实践。

案例解析：“创TRON”助力打造离散制造智能化柔性适配生产线



客户痛点

- 依赖人工操作：人工精度低，速度慢，一致性无法保障
- 产线换型频繁：物料品种多，传统机械臂无法自主学习，调试成本过高



解决思路

- 引入具身智能工业机器人解决方案，降低人工依赖，提升机械臂柔性作业能力
- 快速学习接插件和操作对象，可生成复杂的接插任务
- 使用视觉伺服引导机械臂完成高速、高精度装配工作



应用效果

- 提效：取代人工作业流程，提升线体生产效率，提升装配工作精度
- 降本：大幅降低产线调试时间，降低换型调试成本
- 安全：实时监测周围环境，人机协作更加安全
- 迭代：持续积累操作数据，适应不断变化的生产环境与任务

更适合工业体质的具身智能机器人：“创TRON”



- ✓ 更好地理解复杂的工业环境
- ✓ 快速调整动态变化的生产任务
- ✓ 提升产线的灵活性和适应性

复杂环境感知：具身视觉模块，实时捕捉环境变化

像素级3D定位、复杂任务多模态感知

强任务理解：无需示教与编程，快速理解和拆分任务

大模型驱动的机器人生成式翻译器

实时动态规划：动态识别跟踪与避障，实时生成最优轨迹

基于三阶约束轨迹规划、高维度实时路径规划

高速精准执行：实时控制频率1KHZ，快速响应规划轨迹

超闭环控制、高速高精度触发、“驱控感一体”视觉伺服技术

自然交互

柔性作业

持续迭代

降本提效

安全稳定

具身智能工业场景应用代表厂商：配天机器人



配天机器人



甲子光年
JIAZYGUANGNIAN

- 配天机器人是一家专注于工业机器人、核心零部件及行业自动化解决方案的提供商，是京城机电旗下的国家级高新技术企业。
- 公司始终致力于机器人技术的自主研发及高端装备技术瓶颈的突破，同时前瞻性布局打造以具身智能、机器人行为大模型为核心的“AI+机器人”。
- 植根于工业制造场景，配天在“AI+机器人”已经成功研发免示教焊接软件模块，基于绎零机器人运动控制引擎，通过视觉检测和感知技术实现对焊接任务的快速识别和自主调整，无需人工示教即可投入使用，适应工厂灵活的排产需求。

配天机器人：工业机器人全家福



- 配天专注工业机器人及自动化领域13年；
- 控制器、伺服系统等智能核心部件自主研发；
- 工业机器人产品44款，负载范围3-280kg。其中：
 - 六轴工业机器人负载范围：3-280Kg
 - SCARA工业机器人负载范围：3-20Kg

覆盖行业：

3C

汽车

新能源

化工

玻璃制造

食品加工

橡胶和塑料

金属加工

工程机械

.....

场景应用：

焊接

打磨

搬运码垛

喷涂

上下料

分拣

切割

装配

.....

配天机器人：工业自动化解决方案

行业范围

制药化工

船舶行业

建筑行业

包装行业

半导体行业

新能源材料

AI加持：以免示教焊接任务为例，从专机专用到单机多任务并行，并在低制造、部署、运营成本下，大规模部署在结构化、半结构化、非结构化环境中，并逐步向多样化任务延伸。

产线应用

搬运上下料工作站&产线

机器人焊接工作站&产线

智能仓储物流

拆码垛、成品处理、包装产线

AOI产品缺陷检测

定制化非标产线

典型案例

无源滤波器自动化智能柔性装配生产线

医药行业粉针剂类全自动混粉投料系统

装配式建筑智能化产线

金融行业智能分拣码垛工作站

植物纤维餐具检品机

应用案例：船厂BK板免示教焊接

- 某船舶制造厂商专注于碳钢船体结构的生产，其生产线包含焊接工序。该工序中，需人工焊接BK立板与BK基板，且生产模式呈现小批量、多品种的特点。传统焊接机器人因其柔性不足，难以适应此类生产需求。
- 配天免示教焊接软件模块通过集成3D视觉系统，能够自主识别工件信息，无需预先进行人工标记。该模块可自动提取焊缝特征，并基于此选择合适的焊接路径规划和工艺参数，最终驱动机器人完成焊接作业，有效满足该产线的柔性化生产需求。

案例解析：AI赋能船厂BK板免示教焊接



场景需求

- 主要生产碳钢板，常规厚度12-25mm，需人工将BK板**按照划线标识**进行装配与焊接
- 依靠人工上下料、装配与手工焊枪焊接
- 多品种小批量，且产品间均存在差异



解决思路

基于绎零控制引擎的免示教焊接：满足非结构化场景下的非标产品生产



免模型专家算法驱动的离线编程：通过3D相机自主识别工件类型、板厚和焊缝类型，自动提取焊缝，结合免模型专家算法自适应路径规划和工艺匹配；

✓ **焊缝跟踪与实施修正系统：**极高的视觉定位精度，实时的电弧跟踪修正；

✓ **面向多种工艺场景的解决方案：**船舶制造、轨道交通、航空航天、建筑钢构、电力制造等板材焊接、管材焊接，覆盖多种接头类型与多种焊缝类型。



应用效果



一键运行：

工件放置（非结构化）

任务配置

设备连接

速度设定

点云扫描

焊缝提取

路径规划

ARL生成

一键空运行

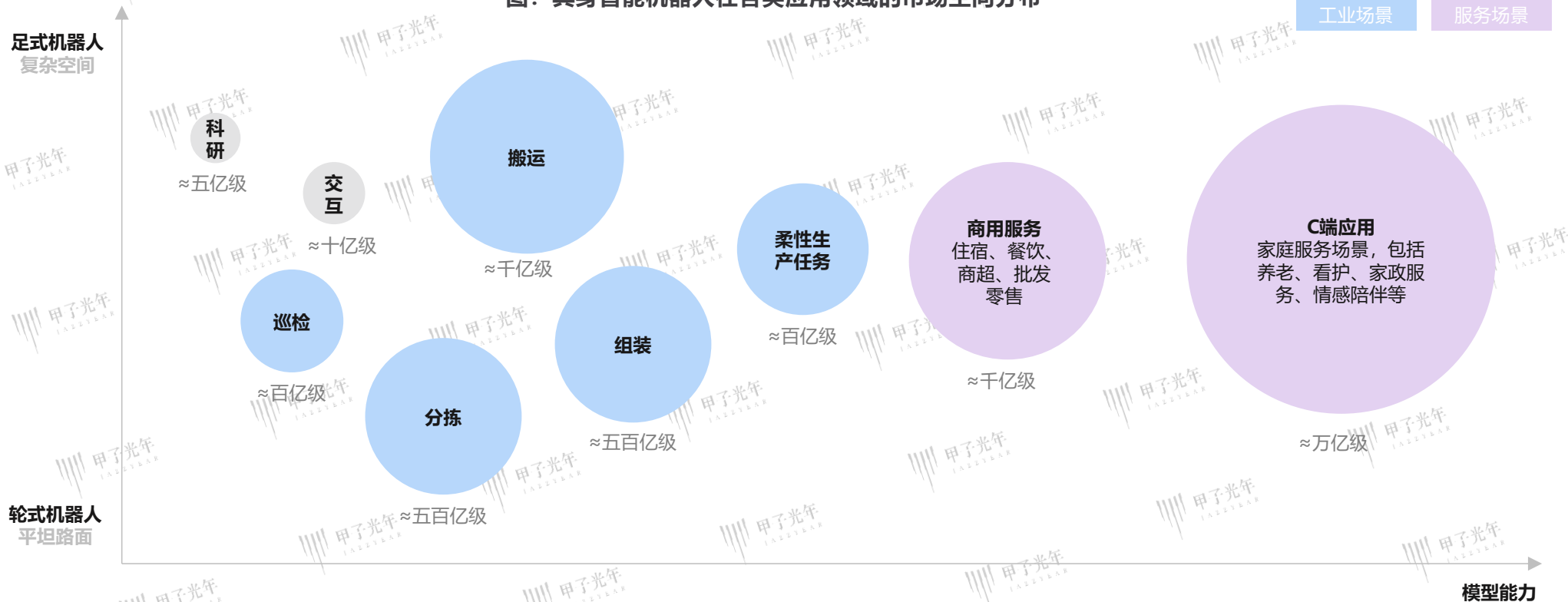
一键实焊

全自动流程

从工业制造的确定性走向家庭服务的灵活性，具身智能价值极具想象力。

- 机器人在C端的应用最具想象力，但短期来看，任务相对聚焦，对泛化能力要求不高工业制造场景下的任务正在更快进入商业化阶段。
- 在工业制造场景实现商业化落地之后，海量机器人的具身数据叠加算力技术的进步，机器人的能力将循序渐进逐步解锁，并向商用服务、家庭服务等更开放的场景进行延伸。

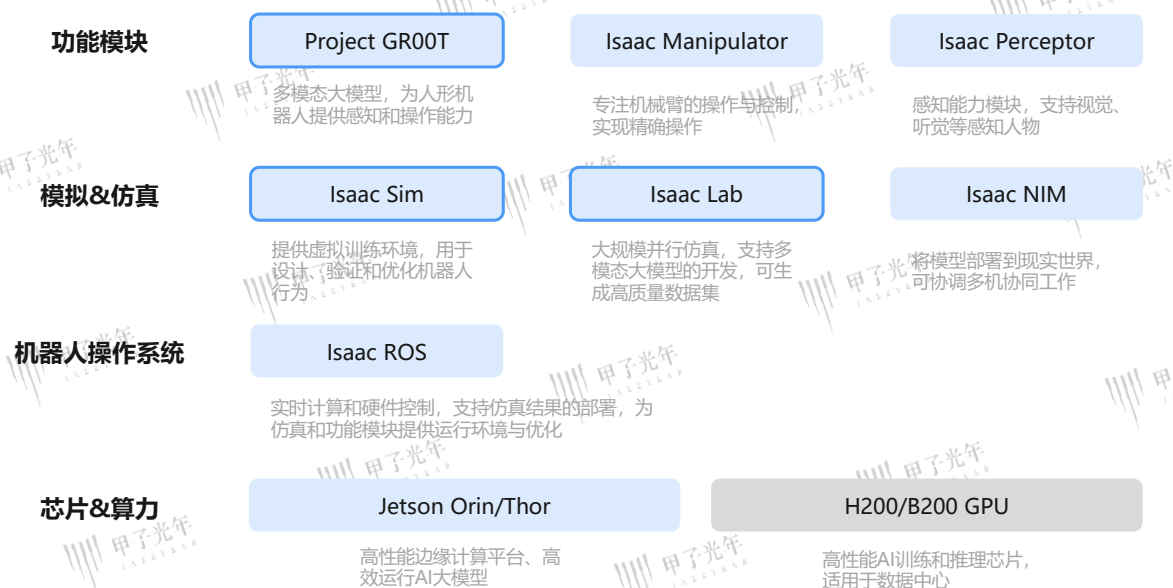
图：具身智能机器人在各类应用领域的市场空间分布



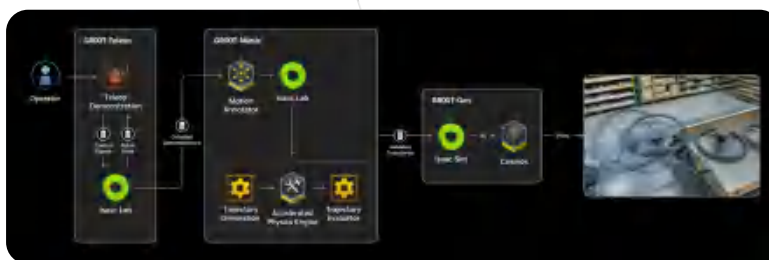
具身智能代表厂商：英伟达

- 大模型训练与推理的需求背景下，英伟达凭借AI训练芯片成功坐上了人工智能算力领域的铁王座，是AI计算领域的绝对领导者。
- 黄仁勋曾表示，AI的下一个浪潮是具身智能。当前，英伟达已经从多个方面部署具身智能，包括计算平台的升级、多模态大模型的研发、软件开发工具包的发布和对外具身智能机器人公司的投资，正在形成一个完整的具身智能底层技术生态体系。
- 在具身智能领域，英伟达维持了其“**底层算力驱动者与研发生态构建者**”的生态定位，专注于提供基础设施、算力支持和工具框架，同时通过硬件与软件的紧密结合，为开发者和企业赋能。

图：Nvidia在具身智能领域的布局广泛，并于2025年发布了用于合成数据生成的Isaac GR00T Blueprint以加速智能机器人研发



NVIDIA Isaac GR00T Blueprint For Synthetic Motion Generation



- GR00T-Teleop工作流，借助Apple Vision Pro捕捉人类动作
- GR00T-Mimic工作流，将捕捉到的人类示范扩展成更大的合成运动数据集
- GR00T-Gen工作流，基于Omniverse和Cosmos世界模型会通过域随机化和3D升维，指数级扩增这个数据集

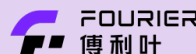
*物理AI模型的开发成本很高并且需要大量真实数据和测试。Cosmos世界基础模型可以让开发者能够轻松生成大量基于物理学的合成数据以用于训练和评估其现有的物理AI模型

- 与14家人形机器人厂商达成合作
- 其中6家为中国企业



UNITREE 宇树

GALBOT



ROBOTERA 星动纪元



具身智能代表厂商：Google DeepMind

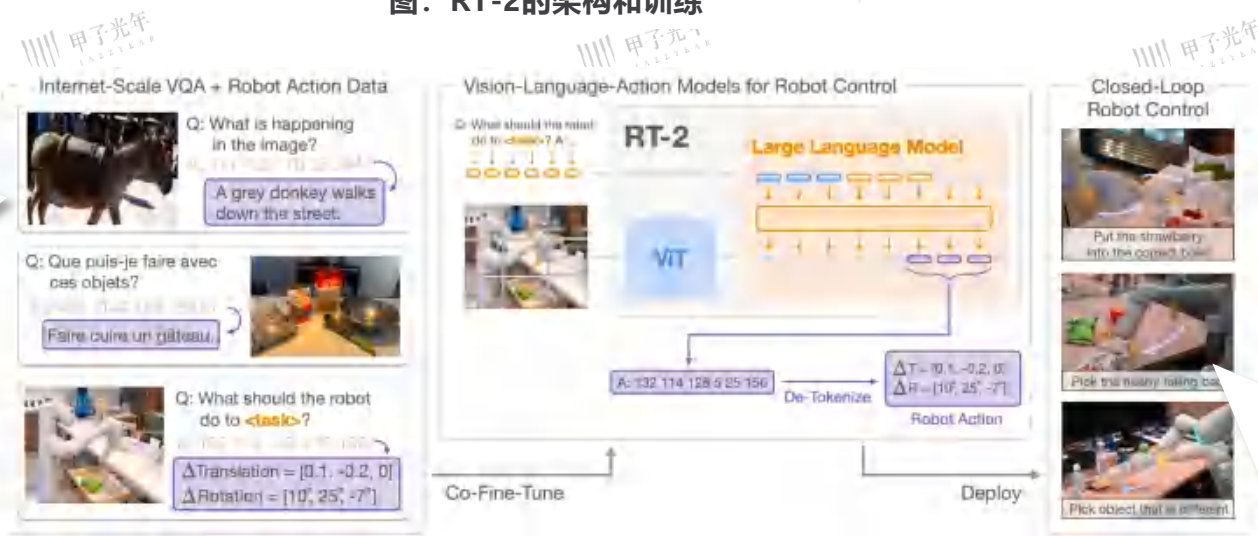


甲子光年
JAZZYKAR

- 来自Google DeepMind团队的RT-2模型能够从机器人数据和网络数据中学习，在保留web-scale能力的同时，将知识转化为机器人控制的通用指令，实现了从视觉语言模型到机器人动作输出的直接转化，提升了机器人控制的泛化能力和语义推理能力。
- Deepmind基于PaLI-X和PaLM-E作为RT-2的等预训练视觉语言模型，将其调整为VLA（视觉语言动作模型）。
- 采用RT-2类似的训练方式，使用Open X-Embodiment数据集进行训练，得到了RT-2-X模型，相对于RT-2有了更好的泛化表现。

图：RT-2的架构和训练

- 大型预训练模型在多种任务中展现强大能力，但机器人获取类似能力面临数据规模和模型应用的挑战
- 自然语言和视觉语言模型难以有效整合到机器人控制中



- 对一个预训练的VLM模型在机器人和网络数据上进行共同微调
- 生成的模型接收机器人摄像头图像并直接预测机器人要执行的动作
- 结合VLM预训练模型和机器人数据，RT-2实现了强大的机器人控制策略
- 泛化性能显著提升

局限一：

- 虽然加入预训练VLM可提升对语义和视觉概念的泛化能力，但机器人并不能产生对新动作的创造能力
- RT-2的物理技能局限于机器人数据中所见的技能分布
- 数据集在技能维度上的多样性是未来机器人执行更多动作的关键突破点

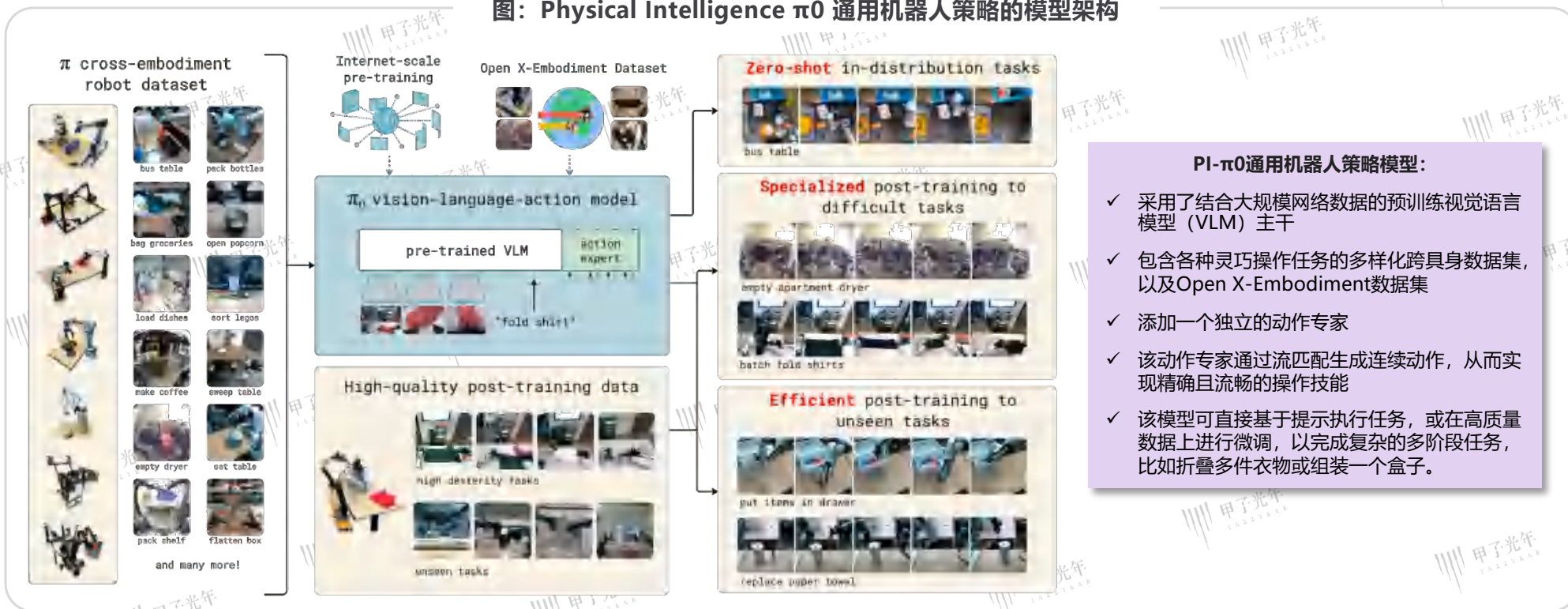
局限二：

- 模型的计算成本很高，随着高频控制的场景的增多，实时推理会成为主要瓶颈
- 目前可用于RT-2的视觉语言模型数量较少
- 对模型量化和蒸馏技术的探索显得尤为重要，需要让此类模型以更高的速率运行，或在更低成本的硬件上运行

具身智能代表厂商：Physical Intelligence

- Physical Intelligence是一家2024年3月成立于美国旧金山的具身智能机器人初创公司，致力于将通用人工智能引入物理世界，开发大规模的人工智能模型和算法，为机器人提供动力，目前估值已达24亿美元。
- PI于2024年10月发布了通用机器人基础模型 $\pi 0$ ，先在高度多样化的机器人数据上进行预训练，调整为更强大的VLA，然后针对复杂任务进行微调。
- 基于强大的预训练模型与多源数据集， $\pi 0$ 能够实现零样本学习的任务处理能力，以及经过高质量后训练数据微调之后诞生的复杂任务执行能力。

图：Physical Intelligence $\pi 0$ 通用机器人策略的模型架构

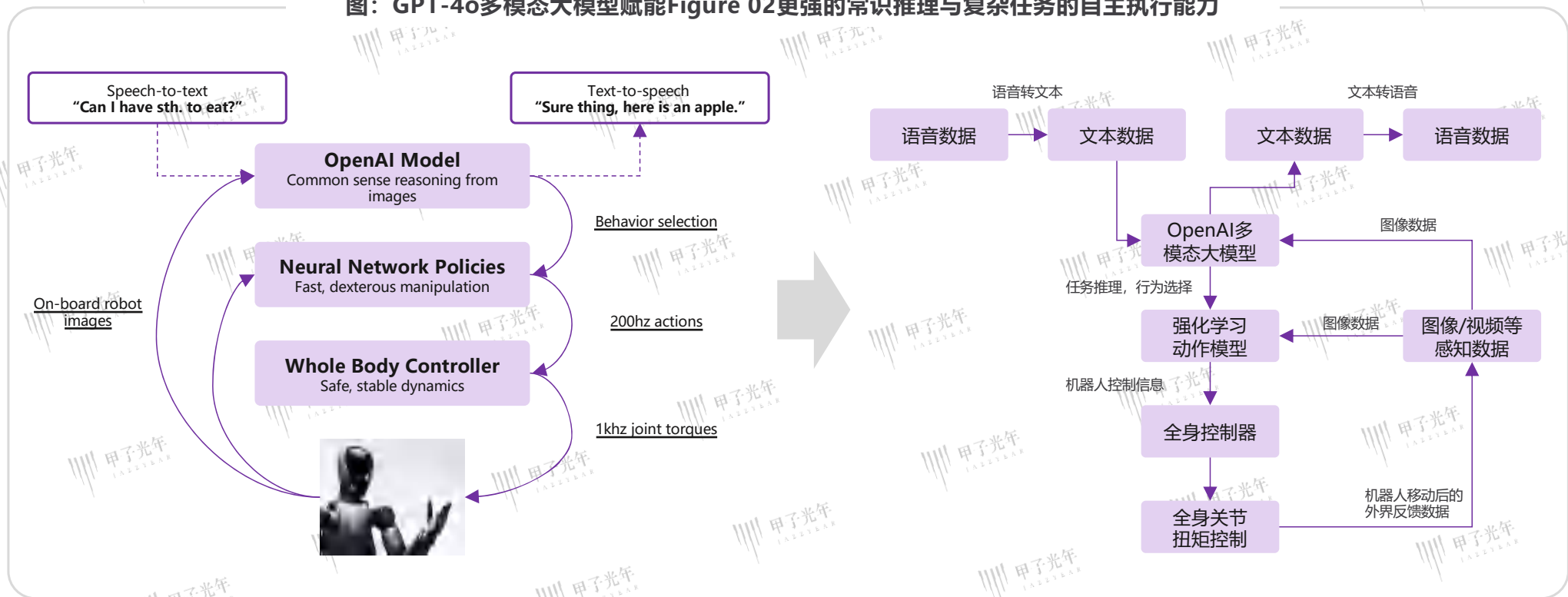


Source: Physical Intelligence, $\pi 0$: A Vision-Language-Action Flow Model for General Robot Control.

具身智能代表厂商：Figure AI

- “Figure is giving artificial intelligence a body.” 公司专注于研发和生产 AI 人形机器人，希望通过先进的AI扩展人类的能力。
- OpenAI大模型赋能的人形机器人，GPT系列多模态模型赋予了Figure 02对多种类信息输入的感知与理解能力，并使机器人通过语言与动作实现与周围环境、个体的交互。
- 当前的Figure 02已经能完成叠衣服、餐桌清理、购物袋包装等相对复杂任务，并在汽车工厂中执行零件装配等劳动密集型任务。

图：GPT-4o多模态大模型赋能Figure 02更强的常识推理与复杂任务的自主执行能力

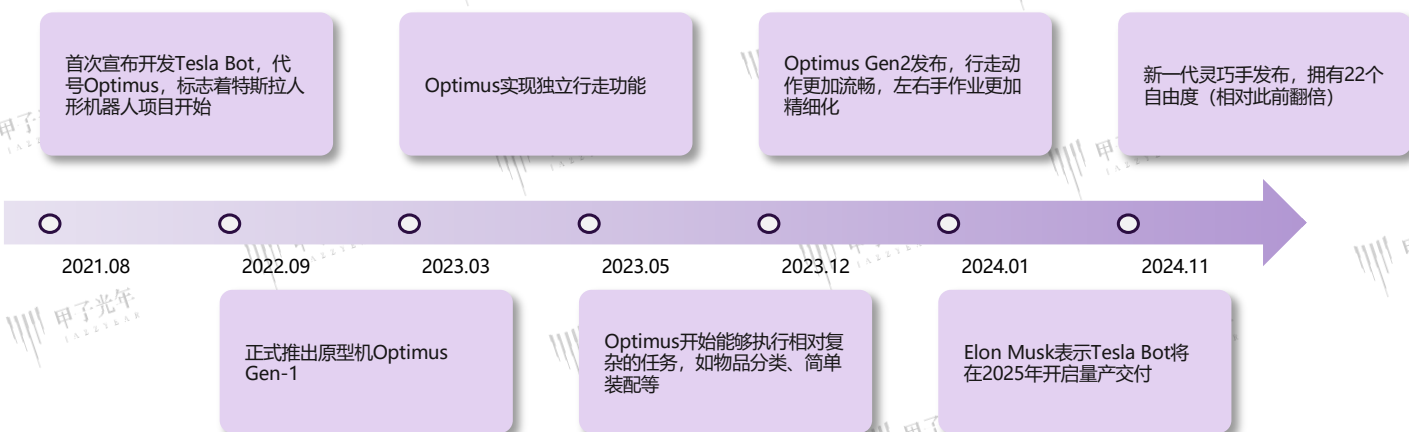


Source: Figure官网, 公开资料, 甲子光年智库整理。

具身智能代表厂商：特斯拉

- 特斯拉通过自身的品牌影响力与对“低成本量产”的目标，催化了人形机器人行业的发展，将这个原本专业且小众的领域快速走入大众视野。
- 得益于自动驾驶领域的数据和算法优势，特斯拉顺理成章地率先转向具身智能人形机器人是顺利成章的。
- 同时，特斯拉通过工艺技术创新和极致产品设计，成功地自研了电动汽车的核心零部件并极好地控制了成本，而这些电动汽车的核心零部件也可以大量地复用在人形机器人之上，例如视觉传感器、关节、电机、电池、热管理系统等。

特斯拉Optimus发展历程及企业研发优势



- ✓ 软件层面：Optimus在软件层面与特斯拉自动驾驶领域有高度重合的环节，自动驾驶技术、数据和模型可以平滑迁移
- ✓ 硬件层面：作为全球领先的电动汽车厂商，汽车身上的传感器与视觉系统、电池与能源技术、电机控制技术都能够作用于Optimus上，并且得益于大规模的生产，各类硬件成本高度可控，直接影响人形机器人的成本价格
- ✓ 资源层面：强大的资金优势、数据与算力储备、以及特斯拉的品牌影响力，为后续机器人量产与销售提供保障



目录

CONTENTS

Part 01 具身智能的发展背景

P02

Part 02 具身智能的发展现状

P17

Part 03 应用场景和代表厂商

P20

Part 04 发展挑战与技术趋势

P36

具身智能作为新兴技术，仍然面临训练数据与模型能力等多重挑战。

- 通过多学科融合发展，具身智能已经展现出了赋能人类经济生活各方面的能力与潜力，但产业链各环节的发展与应用水平不一，仍然给具身智能的整体发展带来了多维度的挑战。
- 机器人的智能化水平仍受到现有方法与能力的制约，其感知能力、执行能力、学习能力、自适应能力、硬件性能、验证方法等受限于技术水平与产业链现状，虽然展现出了潜力，但距离实际落地应用仍有较远的距离。

1 数据获取难题

- 高质量机器人数据集缺失：收集真实世界的机器人数据耗时且昂贵，但仿真数据暂未解决sim-to-real gap，3D数据集建设也仍旧缓慢
- 行业中已经出现不少开源的机器人数据集，但数据质量参差不齐，且缺乏数据采集基准，无法实现跨场景、跨任务应用的通用机器人训练

2 弱解释性问题

- 现有数据驱动的具身智能体暂时无法真正地理解知识、行为与环境之间的因果关系，难以在真实环境中可靠、稳健的运行
- 在商用/家用服务场景下，机器人的“失效成本”很高，强调安全底线，一旦发生事故将可能造成严重后果，因此相对封闭的工业场景更适合率先落地

3 模型能力待提升

- 当前的多模态大模型仍然处于快速发展中，语言、视觉、触觉等多模态融合感知能力尚浅，无法支撑机器人在开放场景的运行
- 在复杂环境与长周期的任务执行能力不足，现有的智能体与计算能力无法实现知识的有效转移与泛化，任务规划器无法实现通用场景的适应性

4 技术路线不确定

5 缺乏验证方法

6 算力水平制约

7 伦理与安全问题

数据赋能，联盟与开源数据集驱动具身智能机器人产业增长

- 高质量的数据是训练高性能具身智能机器人的基础，联盟与开源数据集的建设将有力推动相关技术的进步和应用落地，加速整个行业的增长。
- 通过行业联盟、跨界合作等方式，共同构建高质量、大规模的具身智能数据集，解决数据稀疏和碎片化的问题。
- 开源数据集能够降低研发成本，加速技术迭代，吸引更多开发者参与，促进创新生态的繁荣。

全球：数据生态蓄势待发

全球范围内已涌现出多种类型的具身智能数据集

- **行为模仿数据集**：记录人类或其他智能体完成任务的过程，如来自Google的Open X-Embodiment（跨平台、大规模机器人轨迹数据，加速通用策略研究）、RoboNet、DROID等
- **强化学习环境与数据集**：提供机器人与环境交互的数据，例如Habitat、RoboSuite等
- **多模态感知数据集**：包含视觉、触觉、听觉等多种传感器信息，用于提升机器人的感知能力，如Saycan

开源数据集的价值正在显现

- 开源数据集对于加速具身智能机器人研发进程、降低开发门槛、促进技术普及
- X-Embodiment的开源，促进了通用机器人策略的研究，降低了跨平台开发的门槛

跨界合作，多方积极构建具身智能数据集

- **科研机构与高校**：如Stanford、Carnegie Mellon、MIT等
- **科技巨头企业**：如Google、Meta、NVIDIA等
- **初创AI与机器人企业**

中国：本土数据需求亟待突破

国内在数据集建设方面仍面临严峻挑战

- **数据量相对不足**：与全球领先的数据集规模相比仍有差距，如Open X-Embodiment
- **数据质量参差不齐**：缺乏统一的标准和规范
- **数据孤岛现象**：数据分散在不同机构和企业，难以互联互通，开源的项目相对较少
- **数据安全与隐私问题**

有部分企业与机构开始着手开放数据集建设，攻克数据卡点

- **智元机器人AgiBot World**：基于全域真实场景、全能硬件平台、全程质量把控的百万真机数据集
- **国地共建具身智能机器人创新中心RoboMIND**：首个通用的具身智能机器人训练开源数据集
- **UNITREE G1人形机器人操作数据集**
- **鹏程实验室ARIO**

依托广大市场与应用场景，中国开源数据集有极大的发展潜力

- 开源数据集打破行业数据瓶颈，提升数据质量
- 在先进制造及其他有中国特色应用场景积累数据集

世界模型正在打造具身智能技术进步的“演武场”与“弹药库”。

- 世界模型能够辅助具身大模型进行训练，提供的训练型的仿真：其生成的视频交给具身大模型，具身大模型通过它的规划执行接下来的动作，接下来的动作交互产生新的场景、新的视角，再通过世界模型继续生成新的数据，进行闭环仿真的测试，成为具身智能机器人的“演武场”。
- 另一方面，以NVIDIA新发布的Cosmos世界基础模型为例，WFMs能够为自动驾驶、具身智能机器人等多种AI模型提供符合物理规律的逼真数据，大幅度降低数据采集与标注的成本，充足的“弹药”有望在未来极大地加速具身智能的学习速度。
- 在世界基础模型技术的发展推动下，具身智能的浪潮似乎有望来得更快一些。

特点	World Labs	Deepmind-Genie 2	Nvidia-Cosmos
侧重点	从单张图像生成 3D 世界，强调空间智能	生成可交互、可玩的 3D 环境，用于训练智能体	构建物理感知 AI 模型，生成合成数据，强调物理准确性
技术方法	深度学习模型，3D 重建	自回归潜在扩散模型，Transformer，CFG	世界基础模型平台，不同规格模型
核心能力	2D 到 3D 的转换，遵循几何物理规则	生成可控的互动环境，模拟物理效果，长时间记忆	生成物理上准确的合成数据，理解物理因果关系
应用领域	3D 内容生成、VR/AR、机器人导航	游戏开发、AI 智能体训练、虚拟环境模拟	自动驾驶、机器人、物理模拟

未来的应用场景

- **机器人导航：**提供环境地图和物体识别；
- 建筑设计、室内设计：从平面图生成 3D 模型；
- 游戏开发：生成游戏场景、角色和道具；
- 虚拟现实/增强现实：创建沉浸式体验；
- 在线广告：生成商品 3D 模型，用于展示和虚拟试用；
- 文化遗产保护：对古代建筑和文物进行 3D 重建。

- **机器人训练：**在虚拟环境中训练机器人执行各种任务；
- 游戏AI：训练更智能、更逼真的游戏角色；
- 虚拟社交：创建虚拟社交环境，用于虚拟会议、虚拟活动等；
- 教育和培训：创建交互式学习环境，例如虚拟实验室、虚拟手术室等；
- 电影和动画制作：生成虚拟场景和特效。

- **合成数据生成：**为各种 AI 模型提供训练数据，降低数据采集和标注的成本；
- **自动驾驶：**训练自动驾驶系统，提高其在复杂交通环境中的适应性和安全性；
- **机器人：**训练机器人在复杂环境中执行各种任务，例如工业自动化、物流配送等；
- 物理模拟：用于科学研究和工程设计，例如模拟流体流动、材料形变等；
- 灾难预测和模拟：模拟自然灾害或人为事故，用于应急预案制定和救援演练。

THANKS

谢谢

北京甲子光年科技服务有限公司是一家科技智库，包含智库、媒体、社群、企业服务版块，立足于中国科技创新前沿阵地，动态跟踪头部科技企业发展和传统产业技术升级案例，致力于推动人工智能、大数据、物联网、云计算、AR/VR交互技术、信息安全、金融科技、大健康等科技创新在产业之中的应用与落地。



关注甲子光年公众号



扫码联系商务合作

分析师

翟惠宇微信
zhaihy1203

智库院长

宋涛微信
stgg_6406