



专访光轮智能联合创始人兼总裁杨海波： 数据，机器人产业的新基建



在这一轮堪比“淘金热”的科技产业浪潮中，市场大量关注投向机器人本体研发与具身大模型算法突破，而作为底层基础设施的数据环节，更有可能抢先兑现产业刚需，成为率先跑通具身智能商业化的核心领域。

2026中关村论坛年会期间，北京商报记者与光轮智能联合创始人兼总裁杨海波展开深度对话，围绕具身智能的数据困境、行业实际需求、技术从实验室走向市场的核心要素等议题进行交流。

记者了解到，光轮智能服务智元机器人、银河通用等头部机器人团队，字节跳动、阿里等科技大厂，以及比亚迪、吉利等制造业巨头，是服务头部科技企业及机器人厂商的数据与仿真基础设施服务商。企业将自身定位为具身智能赛道的“送水人”与“卖铲人”，在机器人本体尚未大规模普及的行业早期，凭借贴合产业需求的数据服务实现了技术和商业层面的同步增长。

访谈中杨海波表示，数据是驱动具身智能跨越技术鸿沟、落地工厂与家庭场景的核心基础设施，而物理交互数据的标准化采集与规模化生成，正是具身智能产业化落地的首要前提。

Q: 具身智能行业目前数据稀缺，是数据本身不足，还是数据采集环节存在难度？

A: 具身智能的数据现状与大语言模型、自动驾驶行业有本质区别，数据稀缺并非单纯源于数据量不足，而是数据类型、采集逻辑存在根本差异。大模型依托互联网沉淀的数十年文本数据，自动驾驶有大量上路车辆持续产生道路采集数据，这两类数据采集成本相对可控，主要成本集中在筛选、清洗与标注环节；而具身智能没有现成的线上文本或道路采集数据可用，必须依靠实时实景的物理交互数据，还需额外采集力反馈等物理交互信息，本身数据缺口极大。此外，自动驾驶面向结构化道路场景，具身智能则要应对千变万化的非结构化场景，数据规模化、标准化难度远高于前两者，数据供给成为行业核心刚需痛点。

Q: 光轮智能2023年成立，当时大模型、自动驾驶方案等赛道热度极高，公司为何选择聚焦数据这一细分？

A: 我们始终认为，AI的技术发展是以数据为核心。当时我们判断，物理世界的具身智能需要海量3D实景数据，这是行业的空白切入点，因此没有直接投身大模型研发。初期我们从服务四轮机器人，也就是自动驾驶这类相对简单的场景切入提供数据服务，2024年下半年，行业对具身智能数据的需求大幅提升，我们顺势聚焦该赛道，精准匹配产业实际需求。

Q: 公司布局了仿真合成数据、人类行为数据等多种路径，未来会有所侧重吗？具身智能产业最需要什么数据？

A: 具身智能数据必须同时满足高质量、大规模、多样性三大核心要求。高质量指数据不只是单纯的视频素材，还要包含完整的物理交互信息，保

障模型训练的有效性；大规模是因为具身智能处于训练早期，数据量需求至少是自动驾驶的1000倍，少量高质量数据无法满足训练要求；多样性要求场景、任务、执行主体足够多元，甚至失败、有瑕疵的数据都具备极高价值。分享一个案例，我们曾有客户以1.5倍价格收购这类“不那么成功”的案例数据。我们采用实景采集与仿真生成并行的技术路线，实景数据既能满足客户需求，又能为仿真场景扩容，实现效率最大化。

Q: 对比机器人本体的炫酷动作，如何让外界理解机器人数据的价值？

A: 我们可以把数据服务比作机器人的教育体系，机器人的硬件是身体，大模型是大脑，数据就是大脑学习的知识。我们既为机器人提供训练数据以提升能力，也提供评测数据为机器人“出题考试”，帮助其发现不足并持续迭代。未来机器人走进千家万户，或许需要像汽车驾照一样的国家级准入考试，我们就是搭建这套培养与考核体系的核心服务商。

Q: 具身智能数据的成本主要集中在哪些环节？

A: 具身智能数据的核心成本分为四部分：一是场景成本，包括真实场景运营、仿真场景搭建的费用；二是人力成本，涵盖数据采集、标注的人员相关费用；三是设备成本，涉及实景采集、仿真操作的硬件投入；四是数据处理成本，包括数据清洗、自动化标注、存储、质检的平台研发与运营费用。

Q: 公司近期完成了10亿元大额融资，投资方最认可公司的哪些核心优势？

A: 投资方的认可主要集中在三点：一是赛道价值，各方一致认为具身智能数据是决定行业发展的核心要素，赛道前景明确；二是行业地位，我们是业

内最早布局具身智能数据、行业认知最领先、技术实力最强、服务头部客户最稳定的企业，交付能力与技术水平处于行业前列；三是生态合作优势，这次融资的一个特点是，我们获得新希望集团、奥克斯等场景方的产业投资，当然，投资的背后是我们的生态合作，我们负责挖掘场景数据价值，场景方提供真实落地环境，双方价值高度互补。后续我们会继续扩大与产业方的场景合作，进一步深挖场景数据价值。

Q: 公司选择在北京海淀中关村布局，发展过程中获得了哪些支持与利好？

A: 选择落地中关村核心是看重这里独有的人才、政策、生态三大优势。一是人才优势，公司的办公室几乎正对清华南门，能够吸纳顶尖高校学生参与研发，区域智能领域人才密度领先；二是创业环境优势，海淀区政府的政策扶持、资本配套完善，创新创业氛围浓厚；三是生态优势，海淀已形成从算法、算力到数据的全链条协同创新生态，产业链完整，适配具身智能这类体系化创新领域的发展。

Q: 公司目前营收增速如何？相比机器人本体企业，是否能更早实现商业化与盈亏平衡？

A: 我们正处于高速增长阶段，2025年全年收入是2024年的10倍，预计2026年一季度收入就会超过2025年全年。我认为2026年是具身智能数据规模化的元年，我们从商业端也看到了百倍于去年的需求出现。我们作为数据服务商，相比本体、模型厂商能更早实现商业化收入、更早达成盈亏平衡。我们是具身智能赛道的送水人，就像淘金热潮里的卖铲人、送水人，是为整个行业筑牢地基的基础设施提供方。

北京商报记者 陶凤 王天逸/文
高嘉骏/摄