

机器人“喂不饱”，具身智能数据之困怎解

■中国城市报记者 邢 灿

在刚刚闭幕的2026世界人工智能大会(WAIC)上,具身智能、智能体、物理AI成为高频关键词。现场具身智能参展企业超过200家,超300台真机在工业、零售、商业服务等场景中同台亮相。

具身智能行业确实是在往前走。然而,和传统大模型发展初期一样,机器人也迎来了“数据荒”时刻。

国金证券近日发布的研究报告显示,要训练具备实际应用能力的具身智能模型,行业至少需要1000万小时的多模态交互数据,而当前全球积累的数据量尚不足需求的5%。

硬件日渐成熟,“喂养”机器人的核心燃料却严重短缺,多位业内专家认为,数据缺口背后,折射出整个产业在数据定义、采集路径、标准化建设与商业化落地层面的多重深层矛盾。

具身智能热潮下 为何无数据可用

上述报告指出,随着本体硬件逐步成熟,决定具身智能能力上限的关键变量,已从机械性能转向高质量数据。相较于算力和硬件,数据已成为物理AI时代最稀缺的生产要素。

与传统大模型相比,具身智能所需的数据有何不同?中国城市发展研究院投资部副主任袁帅告诉中国城市报记者,传统大模型依赖互联网上沉淀的文本、图像等数字信息即可完成训练,而具身智能的目标是让智能体在真实环境中完成感知、判断、决策到行动的完整闭环,因此它需要的并非孤立的单一模态信息,而是能反映物理世界真实交互逻辑的多维度数据集合。

袁帅介绍,此类数据需要同时涵盖智能体在操作过程中的视觉感知、动作轨迹、力觉触觉反馈、环境参数变化以及交互对象的状态变化等多重信息,并且所有维度的信息必须在时空维度上严格对齐,才能完整还原一次真实物理交互的全部逻辑,让智能体从中学习到符合现实规律的动作策略与问题解决方法。

广东省社会政策研究会副秘书长高承远向记者表达了相似观点:具身智能的核心在于“身体”与“环境”的实时交互,因此它真正需要的是多模态、高保真、带物理反馈的闭环数据。

与互联网大模型主要依赖文本不同,具身智能的数据必须包含物理世界的因果逻辑,比如一个抓取动作的力度、角度及物体形变等细微变化。

具身智能通常需要哪些常见数据?科方得咨询机构负责人张新原举例,一是机器人视角下的“视觉一动作”配对数据,也就是看到环境之后做出怎样的动作;二是物理动力学数据,包含重量、摩擦力、碰撞形变、物体姿态变化;三是长时序任务数据,比如完

成开门、取物、收纳等一整套连续操作;四是多场景泛化数据,覆盖不同光照、摆放位置、障碍物干扰,让机器人在真实环境里稳定执行任务。

数据供得上用得好 仍需跨越三道门槛

据专家介绍,目前,具身智能数据采集主要分为三条路线:一是真机遥控操作采集,指操作员操控配套主机械臂,远程控制实体机器人完成动作,同步记录全套机器人传感数据;二是无本体采集,即采集人员穿戴轻量化头戴相机、手持简易传感终端,在真实场景里完成日常动作,设备同步记录第一视角画面、手部运动轨迹、环境交互数据,无需实体机器人;三是仿真数据采集,指通过模拟虚拟物理世界,批量生成机器人操作数据。

当前,具身智能数据采集技术路线各有优劣。高承远举例,真机遥控操作采集精度高,但成本大、效率低;无本体采集和仿真数据采集成本低,可实现规模化,但存在“仿真到现实”的迁移鸿沟。

具身智能数据采集技术的应用呈现何种趋势?在高承远看来,当前具身智能数据采集正从“单一路径依赖”走向“混合增强”。头部玩家普遍采用“仿真预训练+真实数据微调+遥控操作兜底”的融合策略。同时,数据采集正从“人教机器”向“机器自主探索”演进,自监督学习和世界模型的引入,正在降低对人工标注的依赖。

张新原表达了相似观点。在他看来,纯人工采集成本高、产能有限,逐步转向“仿真生成为主、真机校准补充、人工标注校验”的混合模式。仿真负责大批量产出基础样本,真机用来修正仿真和现实之间的物理偏差,以此快速扩充可用数据集。

上海具脑磐石科技有限公司联合创始人卢瀚宸告诉记者,随着具身智能技术架构与范式的迭代演化,有关数据采集的方式方法很大可能还会出现新的选项,比如脑电、肌电等方式。未来大概率将形成多种数据采集方式并存的混合数据范式,数据生产逐渐走向工业化,对于数据公司的要求会越来越综合。

受访专家认为,当前具身智能数据采集行业存在不少突出痛点,最核心的矛盾在于数据的供给规模和质量远远跟不上模型训练的需求。

一方面,高质量的真实交互数据获取成本极高。袁帅说,每一条有效数据的产生都需要人工操作配合真机运行,过程中涉及硬件损耗、人力投入、场景搭建等多项成本,导致数据生产的效率偏低,难以在短时间内积累到支撑模型实现能力涌现的量级。

曾在苏州从事过相关采集工作的李先生告诉记者:“当时,我是头上戴着采集设备,同样的动作要重复成千上万次。如果动作被系统判定不合

格,这相当于白做了。这些动作必须非常缓慢,胳膊要一直悬在空中慢慢移动,动作快了就采集不到数据。戴一整天下来,脑袋疼得厉害,胳膊也酸得不行。”

另一方面,数据的标准化程度依然不足。袁帅提到,不同采集主体生产的数据格式、标注规则不统一,不同企业之间的数据难以流通共享,形成了大量数据孤岛,进一步放大了数据供给的缺口。

此外,现有数据的泛化性依然有待提升,很多采集数据都是在特定环境、特定操作条件下获得的,一旦场景参数发生变化,数据的参考价值就会大幅下降,难以支撑模型在复杂多变的真实环境中实现稳定的操作能力,这也成为制约具身智能落地的关键瓶颈之一。

高承远告诉记者:“目前具身智能普遍缺乏极端场景下的训练数据,而正是这类数据,直接决定了其落地应用的可靠性。”

具身智能数据采集 会是一门好生意吗

关于具身智能数据的争夺战,已拉开帷幕。在2026世界人工智能大会(WAIC)上,京东首次系统展示了面向物理世界的具身智能数据布局,优必选、智元等头部企业也已落地规模化数据采集中心。

在袁帅看来,具身智能数据采集是整个具身智能产业生态中不可或缺的上游环节,其价值会随着下游应用场景的不断拓展而持续提升。这门生意的长期价值会随着行业整体规模的扩大而不断凸显,下游所有的具身智能应用开发都离不开高质量数据的支撑,因此数据采集赛道具备稳定的市

场需求基础。

张新原提醒,从商业角度看,具身智能数据采集更多绑定机器人厂商做定制化数据服务,是高壁垒、长周期的配套生意,短期很难靠卖数据集赚高额利润。中长期来看,单纯做数据采集没有长久竞争力,具备仿真引擎、自动标注、虚实闭环迭代能力的企业会占据优势。

高承远认为,这门生意的护城河不在硬件,而在数据资产壁垒和行业沉淀。谁能率先在特定场景(如工业装配、家庭服务)建立起高质量、可复用的数据闭环,谁就能形成网络效应。

高承远预测,行业未来走向会呈现两大特征:一是平台化,即出现跨本体、跨场景的数据中台和交易平台;二是服务化,数据公司将从“卖数据集”转向“卖数据能力”,为机器人厂商提供端到端的数据解决方案。

长期来看,具身智能数据采集不会是一门独立的“大生意”,而是会深度嵌入机器人产业生态,成为智能时代的“数据基建”。它的终局价值,在于让机器人从“实验室聪明”走向“现实世界可靠”。

袁帅分析,未来随着行业标准的逐步统一,数据的流通与共享机制会不断完善,数据生产的效率会持续提升、成本会逐步下降,数据采集与应用的联动会更加紧密,会有更多针对特定垂直场景的定制化数据服务出现。

同时,数据采集的技术路线也会不断迭代。袁帅预测,除了人工遥控操作之外,会逐步出现更多自动化、半自动化的数据采集方法,进一步释放数据供给的潜力,最终成为支撑具身智能从技术验证走向大规模落地的核心基础设施。



日前,在安徽省芜湖市鸠江区矽客机器人未来社区,工作人员通过真机遥控操作方式指导机器人开展技能示教,围绕咖啡馆、餐厅等服务场景进行任务训练和数据采集,为具身智能模型优化与场景应用验证提供支撑。 人民图片