

“00后”当老师 “教”机器人“抓取世界”

记者探访山东省首个人形机器人数据采集训练场 年产数据百万条 助机器人“大脑”进化

在青岛崂山，一片独特的“训练场”正悄然运转。这里没有运动员的跑动呐喊，而是伺服电机低沉的嗡鸣与数据传输的无声洪流。31台人形机器人，在精密布置的工位上，日复一日地进行着抓取、行走、转身等基础动作。

这不是科幻电影的片场，而是山东省首个、也是目前唯一的市场化运营的专业人形机器人数据采集训练场。它的诞生，直指当前人形机器人产业发展的核心瓶颈——高质量、大规模、贴近真实场景的“真机数据”匮乏。近日，记者走进青岛市人形机器人数据采集训练场，探秘这里如何为人形机器人“大脑”的进化注入至关重要的“数据燃料”。



李慧茹

打造山东首个关键数据“供应商”

何为“首个”？“山东省内此前并非没有机器人测试环境，但多是公开、公益性质的体验或演示平台。”青岛市人形机器人数据采集训练场总经理李彦杰表示，“而我们，是第一个明确以专业化数据采集为核心目标，并基于完全市场化运营的实体训练场。”

“专业化”与“市场化”，体现在其清晰的商业逻辑和精准的产业卡位上。训练场并非实验室，而是面向全国机器人产业链提供关键数据服务。其客户画像明确：一是头部科技“大厂”，为其大模型训练和算法测试提供真机数据；二是高等科研机构，支撑其前沿学术研究；三是机器人行业的同行友商，助力攻克共性技术难题。

更为独特的是其技术基底。训练场部署了山东省内首个应用于工业复杂场景的“5G-A”专用网络。李彦杰解释：“5G-A可以理解成5G向6G演进的关键桥梁，它已经具备了部分6G的特性，尤其是超低时延和超大带宽。”在人机协同数据采集的瞬间，视觉信息、力觉感知、关节运动数据如同百川归海，对网络提出了近乎苛刻的要求。5G-A网络确保了海量数据能实时、稳定、低延迟地传输至处理平台。

“真实世界”中生产“真实数据”

训练场的核心运行逻辑，围绕着“真场景、真任务、真数据”展开。它并非虚拟仿真，而是依托于智慧家庭、工业制造等领域的真实场景进行1:1环境还原，部署的全部是可在实际中应用的机器人真机。

现场，一位数据采集师头戴VR设备，手握手柄，正在操控一台乐聚第四代人形机器人“夸父”完成物料分拣任务。采集员的每一个动作，都通过映射实时转化为机器人的动作。“我们的任务由简到繁。”李彦杰告诉记者，“从简单的站立、抓取、放下，到结合行走、转身的多步骤复合任务，最复杂的任务能分解成20多个步骤。”

在机器人执行动作的同时，一套精密的系统正在同步记录一切：多路高清摄像头捕捉的视觉信息、每一个关节电机的旋转编码数据……所有这些原始数



李慧茹通过VR映射操作机器人。陈小川 摄

据流在5G-A网络上汇聚奔涌。

“采集只是第一步，价值在于后期处理。”李彦杰强调。原始数据，如机器人的动作幅度、压力数值等，经过专业团队的清洗、标注，最终被打包成结构化的高质量数据集。一个看似简单的“抓取—放置”动作，在算法模型的训练和测试中，需要成千上万条这样的数据“投喂”。训练场的规模化能力正在于此——28个专业工位、31台真机，能够以工业化流水线的方式，持续产出标准化的高质量数据，年数据产生量可达百万条级别。

填平机器人“大脑”进化数据鸿沟

人形机器人可以分为三个部分，“本体”即机械结构，“小脑”控制运动，“大脑”则负责环境感知和行动决策。当前，全球人形机器人产业在“本体”和“小脑”方面已取得长足进步，机器人形体日臻完善，跑步、跳舞已非难事。然而，决定机器人能否真正融入复杂世界、完成实用任务的“大脑”，却仍是最大的短板。

“瓶颈的核心，就是数据的数量与质量。”李彦杰表示，“传统的实验室研发模式，数据产生效率低、成本高昂，且模拟环境与真实场景存在‘鸿沟’。许多在仿

真中表现优异的算法，一到真实世界就‘失灵’。”

这座训练场的根本价值，正是为了填平这道鸿沟，解决“最后一公里”的数据支撑问题。它提供了一个低成本、高效率、高保真度的数据生产环境，让算法能够在无限接近真实世界的条件下进行迭代和验证。无论是汽车零部件分拣，还是未来更复杂的服务场景，都需要海量真实数据来“训练”机器人的大脑，使其理解物理世界的复杂性与不确定性。

谈及近期热门的“灵巧手”技术，李彦杰表示，训练场同时部署了夹抓手和灵巧手两种末端执行器进行数据采集。“灵巧手更拟人，是未来趋势，但目前成本过高（高端产品单只可达十几万元），且在频繁的物理交互中易损坏，因此当前产业落地仍以经济可靠的夹抓方案为主。”他表示，“但决定人形机器人应用天花板的，还是‘大脑’的智能水平。手是工具，认知和决策才是核心。”

新职业诞生与国产化之路

从2025年1月国内诞生第一家人形机器人数据采集训练场，至今短短一年，国内人形机器人训练场已近40家。这

在产业链上激起连锁反应，一个全新的职业正在兴起。“我们称之为‘数据采集师’或‘数据集采标注师’。”李彦杰表示，这个负责操控机器人、设计采集任务、处理数据的新工种，可能会被正式确立，成为智能制造领域的新兴岗位。

在自主可控的背景下，训练场内机器人的国产化率同样令人瞩目。以主力机型“夸父”为例，其元器件国产化率已达95%以上。这确保了数据采集的底层硬件安全，也为中国机器人算法的迭代提供了坚实的基础。关于人形机器人何时能广泛服务于日常生活，李彦杰保持了审慎的乐观：“工业场景相对结构化，正在快速突破，例如汽车生产、电力巡检、高危作业等。但家庭、商超等非结构化场景复杂度呈指数级增长，仍需漫长的时间和数据积累。我们现在所做的，就是为未来打下坚实的数据基石。”

在这个训练场里，每一次机械臂的抬起，每一次数据的传输，都是朝着未来智能世界的一次微小而坚实的迈进。它不再只是一个测试空间，而是成为了驱动人形机器人产业从“炫技”走向“实用”、从“实验室”走向“应用场”的关键基础设施。在这里，数据不再是副产品，而是核心产品；智能的进化，正通过最朴实无华的重复动作与海量数据记录，一步步照进现实。

给机器人当“老师”很有成就感

训练场内的数据采集师以“00后”为主，李慧茹就是其中一员。李慧茹去年刚毕业，她坦言，选择这份工作“有点意外”，从未想过会和机器人打交道。但真正沉浸其中后，她感觉“蛮有趣的，很有意义”。她的工作远非按下按钮那么简单。她需要头戴VR设备，手握手柄，成为机器人的“影子导师”。在一个模拟超市货架的工位上，她每日的任务是通过操控机器人完成整理货架、抓取商品等动作，并在此过程中生成标准、有效的动作数据。

成为熟练的“机器人教师”并非一蹴而就。李慧茹清晰记得初学时的笨拙感：“手，不像自己的手。”这种疏离感源于精密的操控映射——VR头盔上的摄像头是机器人的“眼睛”，会实时捕捉她的动作，系统识别后驱动机器人做出完全一致的动作。熟练后，一种奇妙的“人机合一”感便会产生：“你知道机器人会跟着你走，我这边怎么动，它就怎么动。”

这份工作有着严格的“教学标准”。每个采集任务都有预设的“标准时长”，可以提前，但不可以往后拖。超时即意味着数据失效。正是在这种对时间和精度的严苛要求下，海量结构化的高质量数据源源不断生产出来，用以“喂养”和“训练”机器人的“大脑”。

当被问及家人朋友如何看待这份新兴职业时，李慧茹笑了：“跟他们说，他们感觉科幻、高级。”但对她而言，更深的驱动力来自于参与创造未来的切实感。“也许未来，超市里工作的机器人就用到了我今天采集的数据。”当遇到复杂任务并努力克服后，她能获得一份独有的“小小的成就感”。李慧茹与同事们正站在人机交互的最前沿，他们通过肢体语言为机器人注入最初的行为范本，在虚拟与现实的边界上充当智能进化关键阶段的“基础数据教师”。在这个崭新行业里，年轻人不仅是见证者，更是直接参与塑造未来的“奠基人”。

青岛晚报/观海新闻/掌上青岛记者 陈小川



扫码看视频