

探寻产业发展“新引擎”

# 走进“人形机器人学校”： 规模化“上课”为具身智能加注“数据燃料”

本报记者 贾丽

“欢迎来到北京人形机器人创新中心(以下简称‘北京人形’)机器人数据训练基地,这里是我们每天‘上学’的地方。”一台人形机器人“开口”,就让人瞬间置身于科幻世界。

在这座坐落于北京亦庄、占地近5000平方米的训练基地里,不同构型人形机器人正同时“上课”——学习内容并非“语数外”,而是给婴儿换尿布、为老人洗头穿衣、在厨房里清洗餐盘。操作者手把手地教导,机器人一遍遍地练习,一套动作要重复数百次,才能沉淀为一条有效的训练数据。

近日,《证券日报》记者走进这所“人形机器人学校”,探寻数据采集训练场如何为具身智能产业发展加注“燃料”。

## 采集海量训练数据

走进机器人数据训练基地的一个仿真训练试验场,记者看到,家居、商超、办公、工业、医药、康养等30多个典型场景被一一复刻。在“儿童房”,操作者正用机械臂给婴儿模型换尿布,动作虽慢,但精准无误;在“养老院”,机械臂正在为老人模型洗头、穿衣服;厨房里,机械臂小心翼翼清洗着餐盘;商超货架前,人形机器人一遍又一遍地练习分拣理货。

“操作机器人时,动作要自然、流畅。”现场工作人员向《证券日报》记者解释。记者随即体验操作一台机械臂,基地数据采集组组长郭明轩在一旁指导:“它的机械臂动作幅度很小,并且会延迟几秒夹紧,你得稍微使点劲。”一套动作虽然简单,但要经过几十次、上百次的采集,才能成为有效的“数据燃料”。

这些场景并非静态的“样板间”,而是可动态配置、组合重构的“数据工厂”。北京人形具身智能负责人孔超向《证券日报》记者表示,目前基地拥有超过120台设备,正全力为内部算法团队及外部合作伙伴(包括机器人企业和大模型公司等)“投喂”海量的训练数据。

在采集数据过程中,最忙碌的是采集操作人员。数据采集及动作捕捉、多模态同步、人工标注等环节,任何环节出现偏差都可能干扰模型。

据了解,机器人数据采集是指通过人类操作或自动化手段,收集机器人在物理环境中执行任务时产生的多模态行为数据,用于训练机器人AI模型,使其具备自主感知、决策和控制能力。在人形机器人产业中,数据采集既是决定机器人能否从实验室走向千行百业的底层基石,也是驱动技术迭代、商业化落地的重要引擎。

数据训练基地负责人蒋未来告诉记者:“三个月前,这里的数据合格率仅为50%。那时候我们面临各种各样的挑战,比如机器人手臂不小心碰到了不该碰的道具等。”如今,经过无数次人员培训、流程建立、问题追溯与质量标准优化,整体数据合格率已稳定在95%以上。

从“教机器人干活”到“让机器人自己学会干活”,数据训练基地正在为具身智能的规模化落地,铺下最底层也最为关键的一块基石。蒋未来表



图①机器人数据与训练基地 图②机器人正在分拣水果 图③操作者对拧瓶盖动作进行捕捉 图④数据采集员正对盘子分拣进行数据采集

示,随着具身智能机器人在更多真实开放场景中实践应用,该基地正迈向全球首个“百万小时数据”里程碑,助力人形机器人技术迭代和规模商用。

## 迈向全流程标准化

人形机器人要真正走进千行百业,需要的不是几百、几千条“精修”数据,而是海量、多样、高质量的数据“原料”。然而,数据采集之路绝非一帆风顺。

蒋未来坦言,目前数据采集面临三大核心难题:一是场景碎片化,真实环境千差万别,每个变量都在考验算法的泛化能力;二是人形机器人“方言”不通,不同构型机器人的传感器布局、关节自由度、控制接口各不相同,导致数据难以跨机型复用;三是数据质量参差不齐,采集、标注、质检任一环节出现偏差都可能产出“低质数据”,误导模型。

“数据饥饿”带来的“泛化性不足”,仍是阻碍人形机器人进入千行百业的“最后一公里”。“如果大家把具身智能机器人当做一辆汽车来看的话,数据就是汽车所需的燃料——汽油和电。它有助于机器人更好地实现运动控制自动化,具备更强的能力。”蒋未来补充说。

通过打造数据采集基地,各地可以将分散的场景集中再现,对多样的机器人进行统一调度,使采集、标注、质检全流程标准化。例如,作为具身智能“国家队”,北京人形已牵头制定国内首个具身智能数据行业规范《人工智能具身智能数据采集规范》,通过标准定义采集规范,引领行业发展。

目前,基地已组建起全国构型最丰富的机器人矩阵,拥有超过120台主流机器人设备,涵盖双足人形、轮式、机械臂、复合型等多种形态,还配备了头环式、夹爪式采集设备,动作捕捉、动作手套及遥操作驾驶舱等专业装备,具备真机遥操作、开放环境

采集、动作捕捉采集三大核心采集能力。

孔超表示,标准化的“数据工厂”流程,能让不同构型的机器人并行采集,实现高质量数据的规模化生产。截至目前,基地已为多家头部企业及科研机构交付超数万小时的高质量数据,服务客户覆盖物流、商超、办公、家居等多个领域。

## 资本竞逐“数据燃料”

当前,高质量数据已成为驱动人形机器人产业发展的战略资源。一场关于“数据燃料”的资本暗战正在全球范围内激烈上演。

人形机器人数据采集训练场、基地正在全国多地涌现。在上海张江,国家地方共建人形机器人创新中心搭建的异构人形机器人训练场里,来自10余家企业的超100台人形和类人形机器人正日夜进行训练。湖北人形机器人创新中心的数据采集场、青岛市具身智能机器人公共训练场也已落地。

此外,宇树、智元、星动纪元等品牌企业也已建设了内部或半公开的私有数据采集场,以持续夯实核心数据护城河。

“数据采集成本主要由三部分构成:资产折旧、人员成本、损耗率。”蒋未来透露,目前真机采集每小时的成本不低,且需要大量人力投入。但正是这种“重投入”,构筑了行业的核心壁垒。他进一步介绍,目前基地70%以上的产能用于服务行业客户,主要用于VLA模型训练和具身大脑研发,客户的数据需求已从去年的“几千小时”跃升至如今的“十万甚至几十万小时”量级。

中国机器人CR教育培训标准委员会委员魏国红对《证券日报》记者表示,数据训练场/基地在各地加速兴起,将助力人形机器人行业突破“小散重复”“后劲不足”等瓶颈,为各地打造全球标志性的人形机器人生态集群提

供有力支撑。

产业需求的指数级增长,也正在资本市场催生新的投资逻辑。从硅谷到中关村,从PE/VC到产业资本,纷纷将目光投向“数据基础设施”这一赛道。随着人形机器人在多种场景应用落地,“数据飞轮效应”加速启动,更多场景带来更多数据,更多数据吸引更多开发者与算法模型,进而推动数据需求的指数级增长。

放眼全球,这场围绕“数据燃料”的竞赛已然白热化。特斯拉、微软、亚马逊、Robotix等国际企业均在大力投入机器人数据采集基础设施建设,而中国企业在这场竞争中展现出独特优势。“我们在人力成本、场景多样性及产业配合度上都具备国际竞争力。”蒋未来表示,一旦“无本体”数据采集(具身智能领域中一种新兴的数据采集模式,通过轻量化、便携式设备直接记录操作行为)等新技术路线跑通,就能极大丰富数据规模,推动统一数据交易市场的形成。

资本的涌入也在加速技术迭代。“从去年至今,大家有目共睹,机器人的全身运控水平提升了。”蒋未来话锋一转,“但要真正进入‘打工’、到田间劳作,今年需要解决几个问题:完全自主的导航、双手的精细操作、环境认知智商。”

这是一个逐层突破的过程,每一层的突破都需要时间,但一旦实现突破,技术扩散的速度会很快。他认为,未来3年至5年内,行业或许会发生质变。未来,谁掌握了高质量数据,谁就掌握了具身智能时代的“燃料定价权”。

随着更多“数据工厂”投入运营,中国人形机器人产业正迎来从“机械执行”到“自主思考”的关键跨越。数据基地的扩容并非一件特别难的事,难的是持续产出高质量数据、构建标准化体系以及打通数据孤岛。国内产业链对“数据油田”的抢占,是在具身智能时代构筑全球竞争力的关键落子。

# 把“数据根基”扎深扎牢 让具身智能“枝繁叶茂”

贾丽

在北京亦庄的“数据工厂”里,我们见证了人形机器人产业真实而关键的一跃,产业链正从单纯的“造机器”,加速向更深层次的“养数据”转型。

人形机器人被称为“制造业皇冠顶端的明珠”,而高质量的数据,无疑是擦亮这颗明珠的关键。当前,全球具身智能行业的竞争已进入白热化阶段,谁能率先建成高质量、标准化、规模化的数据体系,为产业发展输送源源不断的“养料”,谁就能在未来竞争中牢牢掌握主动权。

然而,数据采集绝非简单的“场景堆砌”,而是一项兼具专业性与系统性的工程。笔者认为,面对场景碎片化、标准缺失、成本高昂等现实挑战,人形

机器人产业链各方在数据采集与训练的布局上,需要更精准、更长远。

其一,给数据“立规矩”,推动采集从“凭手感”走向标准化。

目前,行业内数据采集多依赖操作人员个人经验,具有一定的随意性与不稳定性,建议行业加快制定数据采集的工艺标准,将采集动作拆解为标准化的“工序”,引入“工业级”质量控制闭环。不仅要关注采集环节,更要前置介入场景设计与任务规划,并在后续环节加强清洗与标注的自动化水平。

其二,为数据“修公路”,以统一协议打通异构“孤岛”。

“方言不通”是目前制约数据规模化复用的最大堵点。不同构型机器人的传感器布局、控制接口各异,

导致数据难以在不同品牌间流通,造成了一定的资源浪费。建议由行业“国家队”牵头,联合产业链上下游企业,加速推广统一的具身智能数据采集标准与接口协议;同时探索建立数据格式转换的“通用翻译器”,使得同一套数据能够“迁移”到不同品牌机器人上进行训练。此外,各地应鼓励建设开源开放的公共数据集,在保护隐私与知识产权的前提下,促进数据要素的有序流动与共享,降低全行业的训练成本。

其三,帮数据“造场景”,以虚实共生破解长尾难题。

纵观行业发展,面对现实世界中无穷无尽的变量,单纯依靠真机采集数据,成本高企且效率有限,尤其是在工业危化、高空作业等高危、极端

场景中,真机采集更是难上加难。建议产业链各方加大对仿真引擎与生成式AI技术的研发投入,发展“合成数据”,通过构建高保真的物理仿真环境,在虚拟世界中批量生成海量、低成本的训练数据,重点覆盖现实中难以遇到的长尾、边缘场景。以“虚实结合”的方式,让机器人在虚拟世界中完成“通识教育”,再通过真机实践完成“专业深造”,大幅提升训练效率与泛化能力。

数据之争,本质上是未来智能生态的入场券之争。唯有把数据根基扎得够深、够牢,人形机器人这棵大树才能枝繁叶茂。从北京亦庄的“数据工厂”,到全国各地的数据基地,正为具身智能的发展蓄力赋能,推动中国人形机器人产业加速跑。

# 资本密集押注脑机接口赛道

本报记者 吴奕萱

受政策、技术、市场等多重因素影响,我国脑机接口产业活跃度持续提升,近日发布的《中国脑机接口商业化前景报告》(以下简称《报告》)显示,2025年以来,脑机接口企业融资进入空前活跃期,2026年前三个月,脑机接口企业融资总金额已超2025年全年。

多位受访专家对《证券日报》记者表示,当前我国脑机接口产业正处于从技术验证向商业化落地加速迈进的关键节点。融资额的爆发式增长,意味着产业已进入资本加持的提速阶段,行业整体将逐渐从概念炒作转向务实的产品落地与临床应用。

IT桔子数据显示,截至4月7日,年内脑机接口产业融资事件已有21起,融资总金额超40亿元。

其中,今年1月份,浙江强脑科技有限公司(以下简称“强脑科技”)完成20亿元的B+轮融资,成为国内脑机接口赛道最大单笔融资;3月份,格式塔(成都)科技有限公司完成1.5亿元天使轮融资,刷新该领域天使轮融资最高纪录。

从投资方类型来看,本轮融资热情的参与主体呈现出明显的多元化特质。除了中金资本、红杉中国等一线机构外,也不乏上海国投、浙江创新投资等国资平台的身影。多家互联网巨头和产业资本也加速入场。

二级市场方面,据《证券日报》记者不完全统计,今年以来,已有超过10只脑机接口概念股获得机构调研。其中,河南翔宇医疗设备股份有限公司、江苏爱朋医疗科技股份有限公司等5家上市公司获100家以上机构调研。

部分独角兽企业加速奔赴资本市场。今年2月份,证监会网站显示,博睿康技术(上海)股份有限公司(以下简称“博睿康”)在上海市证监局办理辅导备案登记,拟首次公开发行股票并上市。

中国投资协会上市公司投资专业委员会副会长支培元对《证券日报》记者表示:“资本的密集涌入为脑机接口产业发展注入了强劲

动能。一方面,脑机接口技术研发周期长、投入高,雄厚的资金支持为企业研发创新提供了有力保障;另一方面,资本也将吸引更多人才和资源向脑机接口领域集聚,进而完善产业链配套,提升产业整体竞争力。”

在资本的加持下,我国脑机接口产业有望迎来商业化落地的“黄金窗口期”。据《报告》预测,2026年脑机接口市场规模有望突破50亿元并保持高速增长,到2030年这一数字将达到150亿元。

分技术路径来看,我国非侵入式脑机接口技术的产品相对成熟,已在康复医疗、消费电子、教育等领域实现商业化落地。以强脑科技为例,该公司深耕非侵入式脑机接口技术,成功研发出“智能仿生手”“智能仿生腿”“脑机智能安睡仪”“正念舒压系统”等多款智能产品,部分产品已在电商平台上架销售。

侵入式脑机接口技术则进入“准商业化”阶段,并在临床试验方面取得显著突破。例如,今年3月份,国家药监局批准了博睿康植入式脑机接口手部运动功能代偿系统创新产品注册申请,实现脑机接口医疗器械全球首发上市,标志着国际首个侵入式脑机接口医疗器械进入临床应用阶段。

现阶段,企业正积极推进脑机接口产品从实验室走向大规模量产。今年1月份,上海脑虎科技有限公司的“超级工厂”项目在江西省赣江新区正式动工,该工程将承担国内首款、全球第二套“全植入、全无线、全功能”脑机接口系统的量产任务,目标实现万套级稳定交付,预计于2026年下半年建成并投入运营。

脑机接口研究院院长高承远在接受《证券日报》记者采访时表示:“随着各大技术路径的协同推进,未来3年至5年,我国脑机接口产业有望呈现‘医疗康复需求先行、消费级逐步渗透’的商业化落地趋势。此外,随着产业链持续整合,行业‘头部效应’将愈发明显。相关企业需聚焦核心器件的自主研发、临床数据积累与量产能力提升,立足真实场景实现商业闭环。”

# 算力租赁升温 上市公司抢占市场先机

本报记者 丁莹

日前,全球AI(人工智能)大模型聚合平台OpenRouter发布的最新测算数据显示,3月30日至4月5日期间,全球AI大模型词元总调用量达27万亿,环比增长18.9%。中国AI大模型词元调用量连续5周增长,且连续5周超过美国。

激增的词元调用量,直接带动了AI算力需求。算力租赁成为缓解算力供给缺口的重要模式。接受《证券日报》记者采访的业内人士表示,包括算力租赁、算力设备回收与再利用等在内的多元化算力服务,能快速匹配应用端企业的算力需求,也推动DaaS(Device as a Service,设备即服务)市场迎来黄金增长期。

算力租赁作为算力服务市场的重要组成部分,市场规模正快速扩容。根据市场研究机构IDC预测,2026年全球算力租赁市场规模将突破800亿美元,年复合增长率超25%。

从中国市场来看,行业正从“规模扩张”迈向“价值深耕”。中研普华产业研究院最新发布的《2025-2030年算力租赁行业市场调研与投资潜力预测报告》显示,行业正从单一的基础设施租赁向一体化服务转型。

DaaS作为一种为企业提供设备租赁、技术服务、设备管理和回收再利用的成熟服务模式,在AI技术快速发展的时代浪潮下,迎来了巨大的增量市场。

凌雄科技集团有限公司(以下简称“凌雄科技”)董事会主席、小熊U租创始人胡祥雄近日在接受《证券日报》记者采访时表示:“随着AI大模型、智能体加速应用,AI算力需求呈指数级增长,尤其是对缺乏足够资金和技术能力自建数据中心的中小企业而言,算力租赁、算力设备回收再利用等模式可大幅降低其前期投资成本,降低中小企业获取、使用算力的门槛。公司长期布局DaaS,构建了完整的服务生态,拥有庞大的长期服务客户群体,当下迎来新一轮市场机遇。”

日前,工业和信息化部印发《关

于开展普惠算力赋能中小企业发展专项行动的通知》,围绕算力资源配置提升行动、算力服务普惠供给行动等重点任务作出一系列安排,包括鼓励设立中小企业专属算力池;探索“算力银行”“算力超市”等创新业务,支持中小企业存入闲置算力资源等。

面对算力服务市场的爆发式增长,多家上市公司在算力租赁、算力设备回收与再利用等领域持续探索,抢占市场先机。

协创数据科技股份有限公司(以下简称“协创数据”)在算力服务领域构建了包括算力租赁与云服务、交付实施与运维服务等在内的多元化收入模型。该公司算力服务业务以自有算力资产运营为主,主要通过采购人工智能服务器、自建算力集群并向客户提供服务,在多地部署并运营算力中心,为客户提供以高端算力为基础的云计算租赁业务。协创数据的服务器再制造业务,包含服务器的回收、拆解、研发、生产、销售以及维修服务。

针对中小企业群体难以低成本获得算力资源这一核心痛点,凌雄科技于2025年构建起全面的智算生态链,采用“回收+再制造+租赁”闭环运营服务模式,通过回收GPU、服务器等设备,在设备回收业务和设备租赁业务的深度协同下,将其转化为可租赁的算力,为中小企业提供高性价比的算力设备租赁服务。

奥瑞德光电股份有限公司构建了“算力弹性供给+智能运维管理+全生命周期服务”的一站式算力综合服务体系。客户可根据业务特性选择本地化部署或云端资源调用,允许用户灵活、高效、低成本地获取所需的计算能力,精准适配人工智能模型训练、实时推理等智能算力需求场景。

中国金融智库特邀研究员余丰慧在接受《证券日报》记者采访时表示:“企业需要深入研究不断变化的市场和企业客户需求,持续创新服务模式,使广大中小企业能够真正以低成本租赁和使用AI服务器等算力设备以及云算力服务,进一步提升算力资源利用效率。”