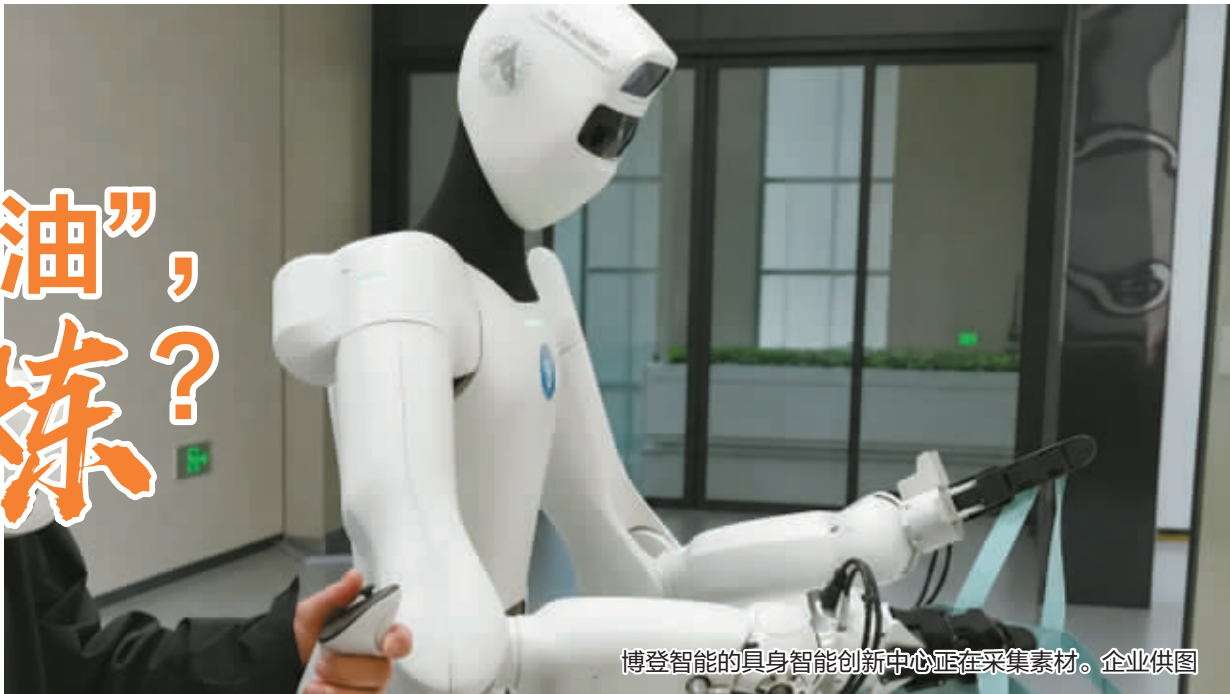


数据“新石油”， 宁波怎么炼？



1

千行百业，加“数”前行

在服装企业，设计师轻点鼠标，就能将2D草稿一键生成3D服装模型；在三甲医院，癌细胞的病理筛查，借助AI数秒就能完成初筛；在外贸公司，单证员打开系统，只需输入商品名称，就能查询出口到美国的最新关税……

是什么让AI如此聪明？答案在于数据。要想让AI识别图像、理解语言、进行决策，必须先“投喂”海量的高质量数据，才能“教”会它新技能。

如今，数据已成为继土地、劳动力、资本、技术之后的第五大生产要素，被誉为AI时代的“新石油”。

眼下，这场“新石油”的“炼化”行动，已悄然打响。谁能率先建成高质量数据集，让数据要素在合规前提下实现市场化流通，谁就能在未来的资源配置中抢占先机。

作为产业数据价值化改革先行城市，宁波正着力唤醒沉睡的数据，为千行百业筑牢智能底座。

如何把数据变成AI大模型的“燃料”？其中关键一步，在于高质量数据集的建设——不是简单粗暴地把数据“打包”，而是经过清洗、标注、结构化处理，把它变成训练AI的“教材”。

这不是理论推演，而是宁波正在布局的大事。

今年3月，宁波启动首批高质量数据集建设，23家重点企业率先“揭榜挂帅”，覆盖制造业、政务、医疗等多个领域。

作为国家数据局高质量数据集典型案例的拥有者，江丰生物正建设“多病种数字病理高质量数据集”。

宫颈癌等疾病的早期筛查，关乎患者的生命健康。江丰生物用高精度扫描仪将物理玻片上的生物信息转化为数字图像，再通过像素级标注，“教”AI识别癌变的信号。

这套数据涵盖宫颈细胞学、胃组织、肠组织及甲状腺病理四大高发癌

种，包含结构化数据37.91万条，图像数据规模达108TB，已累计完成27万例高质量病理数据的标注工作。它采用的标准化格式，可便捷接入医院影像归档与通信系统(PACS)，让数据直达临床一线。

在此基础上，江丰生物推出的AI辅助诊断系统，已在全国完成超150万例宫颈癌筛查，阳性检出率提升到5%以上，让数百例早期癌前病变患者得到及时干预。而基于它开发的AI辅助诊断产品，近两年销售额超5000万元，已出口全球50多个国家。

宁波企业不仅让AI“看懂”二维图像，还在“教”AI理解三维物理世界。

当前，具身智能正引发关注，但要让机器人“真刀真枪”干活，光靠文字和图像还不够，还得让它学习物理常识：握笔时摩擦力多大、走路如何避开障碍物……

为此，宁波博登智能为机器人建立

“培训学校”。工作人员通过Ego、UMI、动作捕捉、多机器人协同等主流数据采集模式，通过全自动的数据处理引擎，将人类的“肌肉记忆”转化为训练机器人的“教材”，“教”机器人完成打螺丝、擦桌子、货架取物等动作。

目前，博登智能已在宁波、湖州、马鞍山布局三大具身机器人创新中心，总面积超3万平方米，已实现年产50万小时机器人真机数据、百万小时级Ego场景数据的稳定产能，规模达到国内领先水平，让具身智能加快融入生产生活。

此外，吉利汽车研究院“吉利汽车星睿高质量数据集”，蓝卓数字科技“面向绿色石化的工业制造行业的多模态高质量数据集”，均入围国家高质量数据集建设先行先试试点。

作为制造业大市，宁波正将深耕细分场景的积累，转化为让产业加“数”前行的动能。

2

如何让数据流动起来？

建好数据集只是第一步。唯有让数据“供得出、流得动、用得好”，才能转化为生产力。

当前，宁波加快推进数据要素市场高质量发展，力争到2027年底，数据要素市场化配置改革工作走在全国前列，建成行业数据流通平台10个以上，数据企业达到2200家，创新应用场景100个。

为实现这些目标，宁波搭建让数据安全流通的“底座”。

2025年，宁波城市可信数据空间正式发布，通过城市公共数据与企业数据的深度融合，在普惠金融、医疗健康、智能制造等领域，持续释放数据要素乘数效应，现已上架1550个数据产品，建成8个行业数据流通专区。

我们熟悉的水利服务场景，背后就有它的功劳。该场景归集18类核心公共数据，覆盖水雨情、内涝积水、河道水位、水库库容等关键维度，基于公共数据授权运营，为城市防灾治理增效。市民能收到积水预警，在雨天避涝出行；保险公司则据此精准定损，将理赔效率

提升20%，综合运营成本节约15%。

除了公共数据外，宁波的一些龙头企业正“格局大开”，让数据从内部资产变成开放共享的行业资源，携手上下游产业链共同发展——

2024年，云聚智铱就推出全国首个纺织服装数据流通平台，依托浙江省服装产业大脑汇聚的超2万个数据目录，7大类数据资源，汇集30家数据供应商，形成120个数据产品，覆盖设计、生产、到市场销售要素等多个环节，助力中小微纺织服装企业的数字化转型。

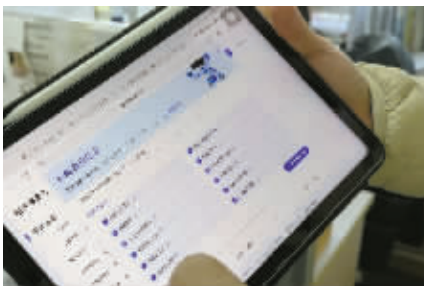
今年6月，均普智能旗下宁波具身智能机器人创新中心，联合博登智能、上海交通大学MINT实验室，正式开源全球首个面向真实机器人强化学习的大规模数据集RW-RL-Dataset首批数据，为全球具身智能团队提供可扩展、可复现、可协作的真实世界数据底座。

其首批开源内容，是超过1000小时的真实机器人交互数据，覆盖4类机器人系列、9大场景域、30余种任务模板及3类核心数据形态。

此后，不同的机器人研发团队，都能基于同一批真实轨迹开展实验，把分散在不同本体和场景中的宝贵经验，沉淀为行业的“公共知识”，为具身智能产业发展贡献“宁波力量”。

根据今年7月发布的《宁波市国家数据产业集聚区试点建设方案》，宁波还将采取如下行动，为繁荣数据产业生态提供有力支撑——

在数据要素市场建设方面，探索开展数据权益登记；重点在纺织服装、港航物流、金融服务、医疗健康、时空信息等特色优势产业，打造一批具有全国影响力的行业数据流通平台；有序推进数据资产化。在培育市场主体方面，建立宁波市数据企业（数商）培育库，推动数据与具身智能、智能制造、石化化工、港航物流、金融保险、医疗健康、智能网联、低空经济等产业融合创新。同时，加强数据基础设施建设，打造“算力+流通”双支撑底座，面向中小企业提供开发工具、数据资源、算力调度等普惠便捷的公共服务。



在宁波外贸公司中基惠通，AI能辅助报关单录入等多种场景。

记者 严瑾 摄

新闻多一点

第十六届智慧城市与智能经济博览会(以下简称智博会)将于2026年9月11日至13日在宁波市国际会展中心举办。本届智博会以“AI智变，慧聚未来”为主题，由宁波市人民政府、浙江省经济和信息化厅、中国信息通信研究院、中国电子信息行业联合会及三大运营商共同主办，采用“1+4”总体活动格局，即1场展览展示加“智领先声”“智聚产业链”“智出新才”“智趣生活”四大板块同期配套专题活动。

记者 严瑾 通讯员 徐波