

“小模型+智能体”跑出大效益 制造业AI探路最后一公里

证券时报记者 康殷

证券时报记者从业内了解到,“人工智能+制造”正成为相关上市公司提质增效的新抓手。人工智能(AI)视觉质检、设备预测性维护等单点应用为企业带来降

本收益,AI仿真、数字孪生和智能体开始进入研发设计、柔性生产、供应链协同等更深环节。TCL华星、美的集团、国星光电等企业的经验显示,工业AI的核心不在于模型“更大”,而在于其是否理解工艺、嵌入现场并形成可复制的组织能力。

1 小模型跑出高回报

在TCL华星的产线上,面板工艺波动以微米计,一两微米的偏差就可能影响产品表现。过去某些异常靠人工班次检查,一个班12小时,可能七八个小时后才被发现,等暴露时已形成批量异常货。AI实时监控后,异常能更快锁定,小问题在扩大前就被拦截。

TCL华星AI巡检项目的推动人、数字化转型中心中心长余必海给了一组对比:过去华星某工厂一年小异常事件有400多起,现在只有几十起,减少接近一个数量级。目前LCD良率部分突破98%,行业平均水平是95%—96%;实行AI巡检后,华星正冲刺99%的良率目标。

在面板生产现场,过去人工判片、检查、修补等工作,如今靠图像识别小模型

和现场智能体,一个班只需少数人维护AI模型。

上述只是TCL华星的AI应用案例。据了解,TCL华星在制造领域已开发图像识别等小模型约两万个、智能体一万多个。制造环节的人效、良率、产能、异常事件等指标,因AI应用出现显著改善。

这类小模型能跑出回报,是技术特性与工业场景双向选择的结果。资深企业管理专家、高级咨询师董鹏分析,视觉质检、异常预警属于边界清晰的问题,定义明确,历史数据充足,价值反馈也极快,企业算得清投入产出比。事实上,这些场景无需颠覆性改造,训练周期短、部署成本可控,企业不需要承担太大的试错风险。

2 大小模型协同工作

制造领域真正创造大价值的,目前主要还是小模型,而不是大模型。小模型在确定性较强的场景中表现稳定;大模型泛化性虽强,但存在“幻觉”,不宜用于高风险环节。换言之,特定场景下AI不是越聪明越好,而是要足够可靠。

由中国信通院、清华大学人工智能研究院、罗兰贝格及华为联合编撰的《工业与AI融合应用指南》指出,大模型虽然凭借庞大的参数体系在自然语言理解、内容生成上表现优异,但其“幻觉”风险、可解释性不足以及高昂的算力成本,使其在面對设备异常预测、工艺参数毫秒级优化这类“失之毫厘谬以千里”的生产制造场景时,显得力不从心。

反观工业小模型,它们多为感知类、判别式模型,专注于特定场景,经过长期输入与输出关系的打磨,专业度更高。IDC的统计数据也印证了这一点,目前工业AI小模型的应用占比高达70%。

天使投资人、资深人工智能专家郭涛认为,“小模型干现场、大模型做参谋、智能体当传令官”的分层架构匹配工业制造的严苛约束,将成为长期主流路线。

作为TCL战略孵化的工业AI企业,格创东智提出“大小模型协同”体系,其中大模型承担企业级认知与决策能力,通过理解企业知识体系、业务流程以及多源数据,大模型能够完成复杂语义理解、跨系统知识推理以及智能决策生成。而小模型则深度嵌入具体工业场景,通过工业机

理与生产数据训练,形成针对特定业务环节的专业模型,例如设备预测维护模型、质量缺陷识别模型、工艺优化模型以及能耗分析模型等。

工业AI小模型单点应用效果被验证后,越来越多制造企业把AI推向研发设计、柔性排产和能耗管理等更深环节。

国星光电依托针对LED集中度稳定性训练的AI模型,对产品进行智能化分档和测试异常及时预警,实现试点产品光电参数一致性(集中度水平)提升10%以上,异常响应时间缩短约98%,提升产品良率,降低制造成本。

美的集团在制造工厂部署AI智能体,其中荆州工厂由“工厂大脑”统一调度,14个智能体覆盖38个核心生产场景,排产响应速度提升约90%,多场景平均提效超80%;泰国空调工厂交付前置时间缩短43%,客户投诉率下降32%。

TCL科技高级副总裁、TCL华星首席执行官赵军透露,自研的星智大模型已深度融入研发、制造、运营全流程,制造环节的ADC自动缺陷分类和ADR自动缺陷修补技术实现“检测—判定—修复”闭环,缺陷判断准确率提升至95%;研发领域助力问题解析与材料研发效率提升20%—30%。

董鹏认为,未来一到两年,研发仿真、工艺参数优化和智能排产最有望规模化破局,企业应选择工艺成熟的主力产线作为“磨刀石”,先跑通一条线再横向复制,避免一开始就铺向全厂。



图片来源:AI生成

3 数据质量和可靠性决定胜负

■ 小模型已产生显著的降本价值,而从AI工具走向“AI原生工厂”不会一蹴而就,设备、材料、工艺、产线固有的物理极限都制约着改造速度。

与此同时,工业AI从单点应用走向全链渗透,还存在堵点卡点。

赵军将AI落地卡点归纳为四方面:模型可靠性、工业数据飞轮建设、智能体工程能力、领域专家协同机制。“工业领域容错率非常低,一次偏差就可能引发产线停机,模型必须具备确定性输出、可量化置信度、可追溯推理链路以及完善的兜底机制。”他表示,半导体显示领域高价值数据稀缺,必须建立采集、清洗、标注、训练、验证的完整闭环,才能驱动大模型从“辅助工具”向“自主智能体”升级,这更需要领域专家持续参与问题定义与结果校验。

郭涛建议,工业现场容错率接近于零,必须建立“模型建议、人工确认、系统

执行”的三段式防护体系,坚持“人在回路”原则,AI输出只能作为参考,不能直接操控设备。他认为评估工业AI成色,真正值得跟踪的是良率、能耗和周转三项硬指标,以及数据治理成熟度。

值得注意的是,今年1月工业和信息化部等八部门印发《“人工智能+制造”专项行动实施意见》,提出到2027年推动3—5个通用大模型在制造业深度应用,推出1000个高水平工业智能体,打造100个工业领域高质量数据集,推广500个典型应用场景。

行业整体仍处早期,绝大多数企业还处于“量变”阶段,智能体数量很多,但还没有真正引发质变。业内共识是,小模型已产生显著的降本价值,而从AI工具走向“AI原生工厂”不会一蹴而就,设备、材料、工艺、产线固有的物理极限都制约着改造速度。

随着“人工智能+制造”专项行动纵深推进,能否率先打通数据治理、系统对接与组织协同,将决定谁能率先跑出工业AI的第二增长曲线。



抢抓暑期旺季机遇 在线旅游平台加速推进AI能力落地

证券时报记者 曹晨

当前正值暑期旅游旺季,多家在线旅游平台持续加码AI能力落地。

“暑期出行需求激增,平台加码AI,既能缓解旺季咨询压力、满足游客个性化行程需求,也契合AI+消费政策导向,提升转化效率。未来AI下单将逐步普及,功能从攻略推荐向直接交易演进,但仍需人工兜底,形成人机协同模式。”中国旅行社协会副秘书长龙飞对证券时报记者表示。

盘古智库高级研究员江瀚对记者表示,暑期旺季流量高峰倒逼平台依靠AI降本增效,用户咨询与行程规划需求指数级上升,传统人工客服边际成本高昂,AI可大幅压缩履约成本。同时,旅游行业产品同质化严重,AI成为差异化破局抓手。平台依靠对话搜索、智能总结抢占用户决策入口,把多页面跳转的复杂操作整合至单次对话,重塑用户使用习惯。

8月10日,阿里旗下飞猪推出AI产品飞猪帮帮,升级面向消费者的AI服务能力。飞猪方面介绍,飞猪帮帮不局限于旅行问答与行程规划,定位为一站式智能助理,覆盖行前、行中、行后全流程服务,主打“一句话就出发”。

据介绍,现阶段,飞猪帮帮可完成订单退改、酒店升房、值机选座、预订接送机、协助填写入境卡、开票报销等十余项功能,支持并行处理多项任务。数据显示,飞猪帮帮输出结果可用率较上一代产品提升超70%,同类任务平均处理耗时缩短近10%。

飞猪首席执行官庄卓然表示,“依托飞猪帮帮及各类AI实践沉淀下来的技术能力、实践经验与行业认知,将成为飞猪向AI原生平台演进的核心资产,具体涵盖模型微调与适配能力、智能Agent系统搭建与评测体系,以及数据资产治理能力”。

除飞猪平台外,多家头部平台近期持

续加码AI相关应用。8月5日,携程商旅官微发布消息,携程商旅智能行程规划AI Agent“程星途”已完整接入用友商旅云,推动差旅管理从“人操作系统”向“人与AI协同”跃迁。

今年4月,携程商旅CEO宋涛在首届全球商旅伙伴大会现场发布了“携程商旅AI生态”全景图,这一生态全面覆盖差旅出行、差旅管理、采购管理等全流程场景。该AI生态包含差旅问答、行程预订、差旅审批、差旅洞察、合规风控、财务结算、资源采购共7个Agent。

7月27日,美团宣布旗下本地生活AI原生助手“小团”完成全面升级。小团2.0不仅能帮用户搜信息、做比较,还能结合实时信息,协助完成下单、打车、订位等各类本地生活服务相关操作,从此前的“问小团”全面升级为“让小团帮忙”。

此前,美团发布《2026五一·消费洞察报告》显示,AI工具也成为不少游客规划

行程的重要帮手。“五一”假期期间,美团AI管家“小团”查询量比春节增长89%,平均消息完读率达61%,户外、文化、体验类目的地是用户问询最集中的方向。

6月8日,同程旅行宣布即将全面接入微信AI智能体生态,并成为首批接入微信AI生态的在线旅游平台。据悉,同程旅行的机票、火车票、酒店、门票、度假等核心产品的旅行预订服务,将同时接入微信AI智能体生态。同时,同程旅行旗下艺龙酒店科技、万达酒店及度假村的超3200家酒店,也将全面接入微信AI智能体。

AI下单的发展前景备受市场关注。江瀚表示,长远来看,AI下单将从“辅助工具”演进为完整“交易闭环”。“现阶段AI多用于攻略参考,未来有望直接完成锁价、下单。但支付、退改签场景存在信任壁垒,平台需解决AI答复准确性、责任划分难题,才能打消用户支付顾虑。”

多家公司宣布半年度分红 海康威视分红超50亿元

证券时报记者 臧晓松

8月11日晚间,上市公司集中发布2026年半年度分红实施方案或预案,多家企业分红总额超亿元。

行动教育(605098)拟每10股派发现金红利12元(含税),合计拟派现1.43亿元。上半年,公司各项经营指标稳步向好,实现营业收入4.28亿元,同比增长24.44%;实现归母净利润1.72亿元,同比增长30.67%。经营活动产生的现金流量净额2.22亿元,现金流充裕。

新天然气(603393)拟向全体股东每10股派发现金红利8元(含税),预计合计派现3.39亿元。今年上半年,公司实现营业收入19.04亿元,比上年同期下降6.58%;实现归属于上市公司股东的净利润3.86亿元,比上年同期下降37.95%。

江河集团(601886)拟向全体股东每股派发现金红利0.25元(含税),预计合计派现2.83亿元。今年上半年,公司实现营业收入约为108.24亿元,同比增长15.90%,营收规模创下历年同期最好水平。公司实现净利润为4.41亿元,同比增长34.34%,同样创下历年同期最佳水平。

甬金股份(603995)计划向全体股东每10股派发现金红利3元(含税),预计合计派现1.09亿元。上半年,公司实现营业收入227.50亿元,同比增长13.05%;实现归母净利润3.6亿元,同比增长17.90%。

沃华医药(002107)拟每10股派发现金红利1.17元(含税),合计派现6753.35万元。湖北宜化(000422)拟每10股派发现金红利0.50元(含税),预计合计派现5434.75万元。长城科技(603897)拟每10股派发现金红利2.00元(含税),预计合计派现4128.71万元。东百集团(600693)拟向全体股东每股派发现金红利0.04元(含税),预计派现3479.38万元。富祥股份(300497)拟向全体股东每10股派发现金红利0.50元(含税),合计派现2693.24万元。

此外,海康威视(002415)发布《2026年半年度权益分派实施公告》,将向全体股东每10股派发现金红利5.50元(含税),累计分红总额达50.41亿元。公告显示,本次权益分派股权登记日为8月18日,除息日、现金红利发放日为8月19日。海康威视此前披露的2026年半年度报告显示,报告期内,公司实现营业总收入468.23亿元,比上年同期增长11.97%;实现归属于上市公司股东的净利润78.96亿元,比上年同期增长39.57%。

宇树科技中签结果出炉

证券时报记者 胡华雄

8月11日晚间,通用机器人行业龙头企业宇树科技披露网下初步配售及网上中签结果。

发行人和保荐人(主承销商)于8月11日(T+1日)上午主持了宇树科技首次公开发行股票网上发行摇号抽签仪式。摇号抽签仪式按照公开、公平、公正的原则进行,摇号过程及结果已经上海市东方公证处公证。

公告显示,宇树科技本次网下中签号码共有19414个,每个中签号码只能认购500股宇树科技A股股票。在此之前,宇树科技披露的公告显示,其网上发行有效申购户数创出科创板有效申购户数历史新高,同时网上中签率创下科创板历史新低,宇树科技也因此成为科创板史上最难中签的新股。

值得注意的是,由于宇树科技发行价为150.80元/股,网上投资者如果中签,需根据中签结果缴纳认购款,中一签(500股)宇树科技新股需缴款7.54万元。网下初步配售结果方面,A类投资者获配数量约1933.38万股,获配数量占网下发行总量的比例为85.36%;B类投资者获配数量约331.64万股,获配数量占网下发行总量的比例为14.64%。公告资料显示,A类投资者包括公募基金、社保基金、养老金、年金基金、银行理财产品、保险资金、保险资产管理产品、合格境外投资者资金,B类投资者则为除A类外的其他投资机构。

强瑞技术 拟定增募资不超10.5亿元

证券时报记者 聂英好

强瑞技术(301128)8月11日公告,拟定增募资不超过10.5亿元,用于AI服务器精密液冷部件产品及服务器自动化组装、检测设备研发与制造项目,高端半导体设备精密零部件生产制造项目,强瑞技术龙华区总部基地建设项目及补充流动资金。

同日晚间,强瑞技术还发布2026年半年度报告,上半年实现营业收入13.74亿元,同比增长63.68%;实现归母净利润8611万元,同比增长56.83%。

今年4月,强瑞技术公告宣布,拟投资18亿元建设人工智能产业研发智造总部,项目核心聚焦内容包括AI服务器液冷产品及生产设备、工业机器人、半导体精密设备等领域。

在6月接受机构投资者调研时,强瑞技术介绍了公司液冷服务器inner manifold中央分水器等液冷散热产品的最新进展。强瑞技术inner manifold中央分水器产品目前主要应用于Rubin服务器的柜内计算托盘内,面临的市场空间较大。

“随着Rubin服务器机柜出货量的逐渐提升,现阶段强瑞技术所获取的订单数量正逐步较快放大。强瑞技术该款产品正在对接多个目标客户,目前Rubin服务器机柜尚未达到大批量出货的节点,后续获取的相关产品订单数量预计将较大幅度增加。”强瑞技术表示。

在液冷服务器组装和检测设备方面,强瑞技术表示,公司所生产的液冷服务器组装测试线最终应用于北美客户的AI服务器)已交付两条线,目前正在生产第三条线,计划年度交付10—13条线;与此同时,强瑞技术正在配合客户开发下一代的液冷服务器检测设备,技术对接和业务进展较为顺利,已开始取得部分配套的检测治具订单。