

星海图如何在“无限游戏”中“结硬寨打呆仗”？

近日，北京亦庄，星海图首届全球开发者大会的会场外，双足人形机器人 Kengo 正做着动态展示，围观的人群里不时有人举起手机拍照。会场内，这家成立不到三年、估值已超 200 亿元的具身智能公司，一口气扔出了三枚“深水炸弹”：开源新一代 VLA 基础模型 G0.5，公布世界模型 Fast-WAM，与北京亦庄共建的数据公司正式揭牌。

会后，星海图创始人兼 CEO 高继扬接受了《中国电子报》等媒体采访。这位 90 后青年创业者，在具身智能赛道热度与争议并行的当下，用一场坦诚又精彩的对话，完整拆解了自己对这条漫长赛道的判断：数据到底该怎么采集？整机和智能之间是什么关系？商业化究竟该以怎样的节奏推进？他的回答构成了一个严密的逻辑闭环，也为行业提供了一份难得的、可供参照的思考框架。

本报记者 杨鹏岳

VLA 与 WAM 并非对立

数据质量决定智能的天花板

假设具身智能是一场长跑，数据就是最初的起跑线。高继扬首先将目光聚焦于此。

在他看来，具身智能与大语言模型最大的不同在于数据的可获得性。“大语言模型数据 99% 都是公开可得的，而具身智能数据 99% 都是私有的数据。”

语言大模型的能力涌现，是在互联网级别的数据量上发生的，而具身智能至今没有自己的“互联网”，真实世界的数据从来没有被系统性地采集过。正是基于这一判断，星海图成为行业内最早押注真实数据的公司之一。在此次大会上，星海图与亦庄共建的“亦数智能”正式揭牌，启动“100 万小时超高质量真实数据计划”，规划今年完成百万小时、未来三年迈向千万小时。

高继扬用一个类比，说明这一量级的意义：一个人从 0~18 岁，醒着的时间和物理世界交互的总时长大概是 12 万小时。100 万小时相当于约 8.3 个人的学习总时长，1000 万小时相当于约 83 个人。“我们相信在百万小时到千万小时之间的某一个数据量训练出来的具身基础模型，会带来突破性的改变。”他说道。

具体到数据采集方式，他详细拆解了四个维度：第一个维度是动作，做什么动作；第二个维度是对象，操作什么对象；第三个维度是场景，在什么场景里做事；第四个是本体维度。本质上，不同采集数据的方式其实都是为了能够更快更有效地铺满这四个维度。

高继扬指出，Human Centric Data (如 UMI 数据、Ego 数据) 和 Robot Centric Data (遥操作数据) 各有优势——前者跨本体能力强、采集效率高、成本低；后者则是模型最终部署到机器人本体上所需要的。关于仿真数据，他态度明确：“预训练全部都是真实数据，基本没用仿真数据。到 100 万小时之前，对于我们公司而言，看不到用仿真数据的必要性。”

他进一步强调，不应只关注数据单项成本，而应关注“智能总成本”——由数据成本、算力成本和研发团队成本三部分构成。在数据上量之前一定先把数据质量解决好，而最高质量的数据便是真实数据(包括真机数据和真人数据)。“如果你的数据不是那么有效，则产生的模型一定也没什么用。”

对于近期行业热议的 VLA 模型与世界模型 (WAM) 两条技术路线，高继扬明确表示两者并非对立。

“今年上半年有一些声音把 WAM 和 VLA 给对立起来了，其实我们一直不认为这两条路线是对立的，它们是同源共生，而且未来会越来越走向融合。”高继扬表示，训练 VLA 也好，训练 WAM 也好，回到其底层就是把一些多模态数据变成 Token，然后用多层的 Transformer 去做编码，再用某种方式做输出。

数据不仅是技术问题，也涉及安全边界。高继扬对此保持警觉：“很多数据从获取的那一刻起就有很多隐私问题，比如研发制造型企业里的核心资产之一就是工艺，老师傅们的活是怎么干的，这些数据需要握在企业自己手里。数据的安全和隐私问题肯定是未来几年特别重要的一件事。”

整机与智能是有限游戏

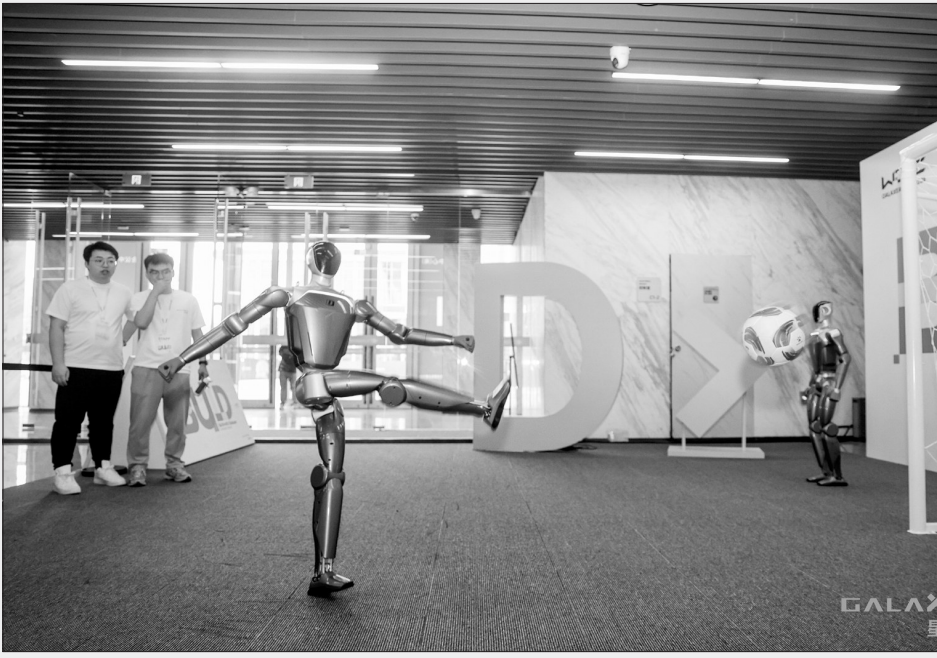
与无限游戏的统一

有了高质量的数据，接下来的问题是如何将数据转化为智能，又如何让智能在物理世界中发挥作用。这正是星海图这家企业试图回答的问题。

“整机和供应链是一个有限游戏，智能和应用才是无限游戏。”这正是高继扬对星海图战略最精炼的概括。他进一步解释：“整机是智能的基础。如果不玩好前面的有限游戏，是没有机会玩好后面无限游戏的。星海图花很多时间和精力去做自己的整机与供应链，目的还是为了做后面的智能和应用。”

从创业之初，星海图就确立了“整机+智能”这一战略方向。“具身智能未来的长期壁垒来自构建物理世界的数据闭环，而为了构建这个闭环，我们最重要的路径是整机+智能。”高继扬说道。

“过去两年外界对星海图有过不同判断，



星海图在开发者大会上公布了多项具身智能成果

有人说它是硬件公司，有人说它是数据公司，但其实星海图所做的一切，都围绕同一个核心——具身智能基础模型。”高继扬表示，硬件与数据都服务于模型，而模型最终要回答的是机器人如何在真实世界里创造生产力。

关于公司在当下的定位，高继扬表示：第一，星海图是一家具身大脑企业，以基础模型的预训练为核心，靠的是“学得越多越聪明”的规模效应，而非传统制造业“造得越多越便宜”的逻辑。第二，它是软硬结合最扎实的一家，从第一天就坚持自己做整机，动力单元已有 80% 自研或产业链共研。第三，它是国内最早、最坚定押注真机数据的公司，目前在国内布局最全、规模最大、技术最领先。

在星海图看来，具身智能不是一条简单的曲线，而是由本能智能、作业智能、进化智能逐级构筑的三个基石。

高继扬表示，本能智能直接作用于本体，让机器人像人一样天然学会驾驭身体，保持平衡、行走、奔跑；作业智能建立在本能智能之上，解决像人一样有序作业和操作的问题，语言是它的重要接口；进化智能则指向更深的问题——AI 能不能定义最优生产力形态。

“今天机器人的身体仍由人类工程师设计，而未来 AI 将围绕应用场景与任务，去定义和设计机器人自己的身体，让机器人从单一产品走向可复制、可迭代、可涌现的生产力系统。”在他看来，过去是人定义机器人，接下来则是 AI 定义机器人。

商业化第一阶段不会盲目冲量 无畏行业“泡沫”

技术能力的积累最终要接受商业化的

检验。高继扬对商业化节奏的判断，建立在对行业成熟度的冷静评估基础上。

“现在市面上几乎所有的销售都是以整机为中心的销售，如果基础模型能力尚不能支撑生产力场景的话，过早把机器人推到工厂或仓库，带来的结果只会是客户不满足、团队疲惫和供应链拉扯。”他坦率地指出，当前阶段，具身智能真正能够创造价值的市场只有两个——“科研开发者”和“表演娱乐”。第二阶段才是“面向生产力的场景方案订阅”。在生产力场景，目前还没有企业真正在其中有效作业，而这里的市场空间是科教研开发者市场和表演娱乐市场几万倍的规模。

基于这一判断，高继扬将具身智能的商业模式变化划分为三个阶段。

第一阶段是当前所处的“整机销售”阶段：年化自然增长率在 30% 至 100%。第二阶段是“方案订阅”：当具身智能基础模型能够在生产力场景中打开应用和规模交付，机器人整机会成为收费入口，但真正的毛利来源是方案的订阅，年化增长率将达 3 倍至 10 倍。第三阶段是“Token 销售”：随着智能从“单一场景有用”走向“多场景、多任务通用”，收费来源从方案转向按智能消耗与任务复杂度计量的 Token 销售，每年将呈现 10 倍乃至 100 倍的增长。

高继扬的判断是——具身智能行业前期是研发制造型行业，中后期则是人工智能的规模化服务，且增速会越来越快。

“星海图不会在整机销售阶段追求绝对的第一，而是更加看重由智能驱动的第二阶段商业化。”在他看来，整机销售阶段中盲目去冲一个绝对的量，给企业带来的更多是负债，而非资产。“我们会按自己的节奏走，等到智能真正开始驱动第二阶段的商业模式了，那时候才是拼第一的时候。”

关于行业泡沫，高继扬的态度很清醒：任

何行业的发展都会伴随着泡沫的产生。人们对一个事物的发展无法做出精准预测，所以有时给出了过高的阶段性估计。赢家往往在泡沫膨胀阶段和破灭阶段都可以获益。

在他看来，泡沫膨胀期并非全然负面——通过这个阶段获取优质资源，大家一起把这些资源整合起来把事做成。与此同时，也时刻准备着市场下行那一天的到来——“这件事一定会发生”。星海图的应对之道是四个字——“务实创新”。

“解决问题的创新才是创新，天马行空对企业来说没有什么价值。”他多次提到自己喜欢曾国藩的“结硬寨，打呆仗”：“做很多事情都是这样，没什么捷径。知行合一，正面推进，一步一个脚印往前走。”

在融资与支出方面，他也提出了自己的看法：“融资和支出是两条路径。融资上，你改变不了资本市场的客观周期，只能在上行时多融钱。支出上，具身智能的驱动力是 AI Scaling Law。传统 VC 按年匀速花钱钱的模型在 AI 这块不奏效，因为 Scaling Law 意味着今年花 1000 万元，明年就要花 1 亿元。”

具身智能的“GPT 时刻” 有潜力改写 GDP 结构

当所有的基础设施、技术能力和商业模式齐备，具身智能将以怎样的方式进入社会？具身智能的“GPT 时刻”何时到来？高继扬给出的答案，或许与多数人的想象不同。

“这个时刻可能不会像大语言模型那样明显。大语言模型是每个人在手机上就能感受到的东西，而具身智能可能在一个地方落地了，大家都不知道。”高继扬认为，具身智能的落地过程是随着能力边界扩大逐个解锁场景的。“可能若干年之后回头一看，原

来机器人已经无处不在了，但我们已经记不得是从哪一年开始的。”

高继扬描绘了落地的渐进路径：第一步，现在是厘米级别精度，能应用到工业领域的上料、搬运，物流领域的分拣、共包、分包，以及电商零售领域的拣选和打包等场景。随着操作精度的上升，到毫米级别精度的时候，制造业里面更多场景——装配、线束、插拔、检测等——将会全部被打开。再往未来看，制造业只是一个开始，农业、建筑业等都是具身智能可以去改变的行业。

“从制造业、服务业迈向农业、建筑业的过程，实际上是机器人从结构化场景(人类改造过的场景)走向非结构化场景(人类没有改造过的、正在进行的场景)的过程，需要双足形态才能进入。”高继扬表示，随着在非结构化场景中有动态适应能力的基础模型的出现，再往未来看，最终目标是自我复制的生产力——即机器人自己制造机器人。到了那时候，就是具身智能的终极形态。

“具身智能是未来的中心性产业，所有行业都可以和具身智能产生上下游业务往来。未来真正的潜力是改写 GDP 结构，而不是简单地用于表演或者改变制造业的某一个具体环节，这都不是它的真正未来。”高继扬表示，“当整机、智能、数据、商业化这四个齿轮开始咬合并加速转动，具身智能对社会生产方式的改造才真正开始——而这场改造，可能早已静默地发生在我们身边。”

回到行业整体格局，高继扬对中国具身智能产业的前景充满信心：“过去行业一直很关注中国在硬件、零部件供应链上面的优势，其实从今年开始，我国在数据供应链上的优势也会凸显出来。数据工程链的优势叠加整机供应链的优势，会决定在未来两到三年的时间里，我国整个具身基础模型的能力将会超过美国，拿到世界第一的位置。”