文件帖 2026-10-27 07:16

来源：刨米机AI精选 | 先知股票数据源

--- 物理AI的“卖铲子”环节：数据采集20261025.pdf ---
全文总结
本次电话会议为华源证券传媒互联网组关于物理AI数据采集方向的专题汇报。会议先由分析师介绍传媒板块的交易逻辑与物理AI的产业催化，再由分析师魏征围绕物理AI数据采集的产业逻辑、数据分类、商业模式以及重点公司进行系统梳理。以下为全文详细总结。
一、传媒板块的交易逻辑与物理AI的产业催化
传媒板块的交易充满相对偶然性。此前核心关注的几个逻辑：第一，Q4风险偏好相对来讲对AI应用这边逐渐起来，市场交易有可能在应用端进一步活跃；第二，整个传媒板块的资金配置，目前板块成交占比可能会超过0.4，其中游戏也超跌比较多，整个影视也是。从这两个交易角度来讲，市场更偏轮动；
产业方向上更聚焦大的多模态方向。多模态的进度，一方面AI视频从短剧到长剧到电影，是一个内容升维的过程，视频难度更长，从内容制作难度到产业壁垒都在强化，所以这里边确实会有一部分标的逐渐跑出；
另一方面，内容会进一步演绎到互动的内容形态当中去，这里包含字节的世界模型的进度。团队第一时间召开了字节世界模型的大会讨论其方向。从结论看，字节的世界模型年内一定会推出；第二，字节世界模型瞄准的应用方向会偏互动的影视和游戏这块内容。结合字节现在流量对于互动内容这块的扶持，对这个方向相对比较清晰；
从字节世界模型的维度看，团队更新了数据板块这边贡献的推荐。随着世界模型对物理、或者说对整个AI具备物理数据集这块采集、数采需求，这块是一个增量膨胀的阶段。因此前置推出了风语筑、宝通、凡拓这部分标的；最新又推荐了中广天泽、华数这部分新增的数据包，包括天翼数科。
二、海外市场对物理AI数据的强化支持
大家讨论的其实是字节系，但从整个海外市场的交互或者轮动的角度，我们看到OpenAI包括特斯拉这个阶段进一步强化了对整个数据、或者说物理AI这边数据的支持；
第一个是GPT-6Astra，在整个机器人操作和导航这方面已经开始挑战过去针对不同任务和机器人本体的传统模型。在这个情况下，GPT-6Astra作为独立模型的能力上线，而在整个桌面操作、移动平台、灵巧手这些具身智能方向也做了进一步强化。这也强调了对整个AI数据之一这边需求的强化，所以我们自己觉得这是一个连贯性的需求，就是从整个多模态这边也进入到了新的发展视角；

这个阶段进一步强化了对整个数据补给的需求。因为在大的模型基础上，原始的数据，大家觉得加意义不大，但是现在在整个动态数据集角度来讲，相对来讲是比较清晰
的：
第一方面是世界模型的数据集，比如高清多视角4D的视频数据为核心，这肯定比
较重要；
第二个是在整个具身智能这边的双手操作动作数据，有布局的公司也相对比较清晰。
所以在这个维度，再次强调对整个基于世界模型和物理AI，对素材、对仿真数据这块的贡献或布局的公司，相对来讲还是比较重要的。
三、传媒板块部分公司的布局情况
市场比较关心凡拓数创、宝利科技这边的布局基础。凡拓也好，还是宝通也好，本质
上来讲是进一步对整个物理AI数据集做了一定的布局；
凡拓的核心：早期就在工业物理AI这边做了比较深度的布局，现在包括素材、数据清
理、垂直模型和机端的部署都有一定的布局。公司领导在体外布局了素材的公司，这
家公司做了5~6个点的投入，阶段或成本也相对比较高；
目前已经具备的数采设备大概四五百万套，现在确实在累计地推动数采数据集的
标准数据产出；
另一方面，公司跟立讯精密合作多年，从最早的数字孪生产线已经布局，数字孪生当时帮立讯做了大概四个工厂的数字监测技术，已经在工业级别的数据采购和理解当中有所部署；
现在进一步延伸，以立讯为代表的工业互联网，进一步走向4.0，走向机器人部署的阶段。简单理解，机器进工厂其实不是一个简单的过程，它也需要对整个工厂的物理数据进行相对的理解，同时建立仿真数据平台，然后去做独立化的大脑、集成大脑训练，再到集成选型进行落地。实际上这是一个相对复杂的一体化工程路径，这一块市场可能会更加分明一些。所以在这一块，凡拓在工业化部署、
在数采阶段有一定的布局；
宝通科技也是一样：宝通过去是承载业务，之前的语数机器系统已经进入到机器人进
工厂这个领域当中去，同时又投资了光轮智能这样的公司。从逻辑理解，光轮智能还
是非常优秀的、中国最大的或者相对比较靠前的数据采集或仿真数据集售卖的公司。
这一块的垂直领域合作，也进一步推动了数采走向新的阶段；

所以传媒这边大部分公司，是围绕着过去的业务基础，从工业互联网角度阶段走向数
字孪生，走向世界模型的数据供给，然后包括进一步进入到法律平台部署的阶段。
四、物理AI与数据采集的产业逻辑（魏征）
和这一轮行情相比，之前机器人板块是围绕特斯拉链的核心零部件展开的行情，而这
一轮比较重要的变化是全球头部通用模型厂商正在布局物理AI。过去几年大模型主要

解决文本、推理、代码这些问题，随着通用模型在空间理解、物理推理和动作规划方
面的进展，模型开始逐步进入到物理的世界，也就是机器人；
机器人所需要的数据和之前我们所说的文本、图片、视频这种数据存在很大的差异。
所以我们认为，这一轮物理AI不仅只是模型和机器人本身的产业机会，其实也可能会
带来数据采集、数据加工、仿真训练基础设施的进展和行情；
国内方面，近期组织过字节的产业专家交流。从反馈看，字节的世界模型其实已经进入
比较关键的产品验证阶段，今年年内肯定会发布，大概就是11~12月份这个时间节
点。据产业专家交流，字节的VideoWorld2目前已经完成了预训练、互训练、强化
学习和评测等环节，整体处于ready的状态。从发布时间看，目前还没有完全确定，
但12月之前、今年一定会发布，可能会在年底和CDS360一起推出；
历史上，2024、2025年字节火山引擎的冬季Force原动力大会基本都在12月18~19号
举行，所以建议重点关注今年12月中旬的字节潜在产品催化；
接下来应重点关注字节链比较重要的公司，因为字节的世界模型需要很多3D的训练数
据。据产业专家交流，字节VideoWorld2的训练数据规模大概是1270万个小时，主
要来自五个方面：
互联网公开视频，大概8890万小时；
字节自己的内容生态，包括抖音、TikTok，大概210万小时；
学术公开数据集，大概80多万小时；
第三方采购以及仿真合成的数据，大约70多万小时；
专项采集的数据，超过20万小时。
从整体数据需求的结构看，第三方采购以及仿真合成大概70多万小时，以及专项采集
的20多万小时，在整体训练数据中占比相对有限。这恰恰体现了专业的物理AI数据和
普通互联网视频数据之间的差异。所以我们认为，未来对这块数据的需求会更加重视
专业化和结构化。据专家反馈，字节的专项采集数据涉及手工操作、机器人、飞行器
等场景，需要获取高清、多视角以及完整动作链条过程的数据，这部分数据相对比较
有异类和有限。所以字节仍然需要去第三方采购专业的数据，并且针对特定的物理场
景进行专项采集；
据专家交流，VideoWorld2初期主要面对互动影视、开放世界游戏、CG渲染、影视
特效以及XR沉浸式娱乐等场景进行布局。比如，未来互动影视的剧情可能不会完全固
定，而是根据用户的选择去实时生成后续内容；再比如，开放世界的游戏可以利用世

界模型去辅助生成3D场景、物体以及动态交互的场景。更长期来看，字节还是希望这
种世界模型的技术去和它的Pico等终端进行结合；
所以，字节世界模型的发展具有比较重要的产业意义：一方面，模型的训练需要大量
高质量真实的数据和仿真数据；另一方面，模型能力的提升以及国内首个世界模型的
发布，有望赋能影视制作、游戏以及XR的内容应用；
对于传媒的公司而言，物理AI可能会带来上游数据采集的机会；第二也可能去推动原
有的数字内容，向AI模型训练和新型交互应用延伸，比如AI游戏、AI影视、互动影游这些。
五、为什么重点看物理AI的数据采集
核心原因在于，现有的互联网几十年来积累的训练语料非常丰富，但物理AI的训练素
材缺乏规模化、缺乏高质量，以及缺乏可直接用于训练的真实或者说仿真的交互数
据；
从目前的产业技术路线看，比较值得关注的有三类数据：
第一类是Ego数据，即第一人称视角的数据。数据来源是通过头戴式或者胸带式
的设备，去记录一个人完成各种操作任务的全过程。优势在于能够利用人类自身的行为和经验去开展大规模采集，不需要所有数据都依赖真实的机器人去生成，所以具有扩大规模以及降低采集成本的潜力；
第二类是UMI数据，即通过一个通用的操作接口产生的数据。比如我们通过专门的手持操作设备去采集人类执行任务的姿态、位置、轨迹，以及和物体交互的信息，让它为机器人的学习提供信号；
第三类是机器人本体的数据，就是让一个真实的机器人放在一个场景里面去进行遥操作，或者自主去执行任务，然后再去采集机器人的状态、视觉、触觉、力觉的反馈等等这些数据。这类数据的成本相对比较高，采集数量也有限，因为它受限于机器人的数量、硬件成本、效率等等限制。

除了真实世界的数据之外，我们同样也看好第二种，就是仿真合成的数据。如果所有
训练数据都依靠人去收集、或者依靠机器人在真实场景中采集，整体成本非常高，同
时一些比较危险和复杂的场景也很难覆盖。所以未来更具规模化潜力的路径，是通过
建立仿真的平台，再结合3D建模、世界模型、物理仿真，在虚拟世界中生成更多的训
练场景，然后再把这些仿真数据喂给机器人，让机器人再回到真实场景中去进行测试
和验证。所以，未来物理AI的数据基础设施会逐步形成真实采集、数据加工、数据标
注以及仿真合成这样一个完整的闭环。
六、数据采集的商业模式与价值来源
传统的数据标注更多是劳动力密集型的服务，按照人员投入和项目交付去进行收费。

在大模型时代，数据的价值正在发生变化，特别是对于物理AI，它对高质量数据的稀
缺性更强。所以它不仅需要专业采集的设备和场景，还需要完成在空间中进行重建、
动作轨迹识别、多模态标注、数据质量校验等一系列比较复杂的加工处理，最后才有
可能成为一个真正可用于模型训练的数据；
所以未来数据采集这一块的产业价值来源于三个方面：
第一是规模化生产的能力，能够以较低的成本去组织专业的场景、人员以及设备，
实现数据大批量、大规模的持续供给；
第二是数据加工的能力，能够把普通的3D、视频进一步加工成具有空间坐标、动作
轨迹、物理属性的高质量训练数据；
第三是数据的复用和平台化能力，能够在合规授权的前提下，将同一个数据资产服
务于不同的模型、不同的任务和不同的客户，并进一步提供仿真训练、评测等增值服
务；
从产业机会看，未来可能会形成定制化的数据采集、标准化数据集的授权，以及数据
采集和仿真训练一体化服务的多种商业模式。对于传媒互联网板块来说，过去很多公
司积累了专业的视频拍摄、影视制作、3D数据资产、数字孪生以及虚拟场景建设等等
能力，这些能力其实和物理AI的数据采集和生产有一定的重合度、相关性以及业务交
集，所以未来一定要重点关注。
七、重点推荐公司
华源传媒互联网板块物理AI重点推5家公司：风语筑、中广天泽、宝通科技、凡拓数创
和天翼数科。

风语筑：过去长期从事企业展馆、数字空间和沉浸式内容制作，在3D场景构建、数据
资产制作以及渲染和空间交互方面有一定积累。今年公司投资并联合松延科技布局物
理AI相关业务，并成立合资公司风语科技，后面可能承接数据方面的一些业务。第二
个看点是Pico，公司之前已经基于自己的Pico技术生态推出了VR大空间项目，具备一定
的VR场景开发、沉浸式交互以及终端应用能力。Pico不仅仅只是一个内容展示终端，
未来也具备向第一人称视角的数据采集设备延伸的潜力。特别是现在物理AI数据对这
种Ego第一人称视角的需求比较大，公司在这一块会有一个潜力。
中广天泽：过去主要从事影视内容制作以及版权运营，长期以来积累了较多真实世界
的音视频内容资源，同时也具备专业视频采集、内容制作以及数据加工的能力。从公
司官网看，目前公司有3个AI数据产品：第一是世界模型的数据集，重点提供高清连续
多视角的4D视频数据、空间运动以及真实世界声学数据，能够帮助模型理解物体运动过

程、空间关系和物理规律；第二是AIGC的数据集，主要是依托公司原有的影视制作和音视频资源，为生成式AI提供高质量的训练素材；第三是具身智能的训练集，重点围绕人类的动作、真实的环境以及物体交互过程，提供相关多模态的训练数据。公司还有两个优势：第一是采标一体化形成的差异化优势——公司具有真实的场景数据，覆盖超过1万个室内外场景，涵盖工业生产、社区家居、商业消费、体育以及自然环境等领域，并且拥有数千人的垂直标注团队，同时具备前端真实场景采集以及后端的标注和结构化加工能力；第二是根据公司半年报，公司目前的数据集总规模超过10TB，客户包括基业新城、中国电信等，并且已经实现了相关AI算料产品的批量交付。

宝通科技：最值得关注的就是它有光轮智能的股权，双方合作围绕真实的工业场景去构建物理AI数据以及训练的基础设施。公司在工业运输、智能矿山这些领域拥有长期的业务积累，具备真实工业场景的资源；同时公司和光轮智能展开深度合作，围绕工业场景的数据采集、仿真合成、模型训练以及评测验证等环节进行了布局。

凡拓数创：过去主要从事数字创意、3D可视化和数字孪生相关业务，在3D场景制作、工业数字空间以及实时交互技术方面具备积累。根据公司此前的交流情况，凡拓在物理AI建立了四类数据布局，分别是空间数据、仿真交互数据、无本体数据和本体数据。在技术平台方面，公司也正在推进仿真平台的相关建设，探索为物理AI提供场景生成、仿真交互和模型训练能力。所以凡拓的优势不仅仅包括数据采集，还包括3D数据、3D数据资产、仿真平台，以及工业机器人训练解决方案这些。公司此前也披露，它和立讯精密进行了合作，去探索在真实流水线场景中的解决方案。

天翼数科：根据公司半年报，天翼数科已经积累了超过570万条多模态数据，其中3D数据超过200万条，涵盖高质量的纹理、铰链结构、可变形物体等等类型。公司部分高质量的3D数据和多模态具身机器人抓取的数据集，也完成了相关的数据资产登记。我们认为这一类数据和物理AI的训练需求也存在一定的匹配度。在此基础上，公司也在进一步布局空间智能模型以及相关的数据服务平台，希望能够将原有的数据资源转化为模型训练和下游应用的能力。

总的来看，我们认为物理AI正在推动机器人产业链的价值环节进一步向上游的数据和训练基础设施延伸。传媒板块中具备专业的内容生产、3D数据和仿真能力的公司，有望迎来新的业务发展机会。

会议实录：

首先有请我们同事魏征。今天跟大家汇报一下物理AI的情况。因为我觉得传媒板块的交易确实充满了相对的偶然性。前边我们核心关注的是几个逻辑：第一个是Q4风险偏好其实相对来讲对AI应用这边的风险偏好逐渐起来，那市场的交易有可能在应用端进一步活跃；第二个就是整个传媒板块这边资金的配置，整个现在的板块成交占比可能会超过0.4，这里边游戏其实也超跌得比较多了，整个影视也是。从这两个交易角度来讲，市场也更偏轮动。但产业的方向，我们之前就更聚焦了整个大的多模态这个方向。那这个多模态的进度，一方面我们觉得AI视频从短剧到长剧到电影，这是一个内容升维的过程。更长的这个视频难度，然后从内容制作难度到产业的壁垒其实都在强化，所以这里边确实会有一部分标的逐渐跑出。那另一方面，我们觉得这一块内容会进一步演绎到互动的内容形态当中去，那这里边包含了字节的世界模型的进度。我们也在第一时间召开了字节世界模型的大会，去讨论字节世界模型的方向。那确实从结论的角度来看，字节的世界模型第一是年内一定会推出；第二，整个字节世界模型瞄准的应用方向会偏互动的影视和游戏这块内容。那结合字节现在的流量对于互动内容这块的扶持，我们对这个方向相对比较清晰。那第三个从字节世界模型的维度看，我们更新了数据板块这边贡献的推荐。因为随着世界模型对物理、或者整个AI具备物理数据集这块的采集，包括数采这块数据的需求，这块是一个增量膨胀的阶段。所以我们前置地推出了像风语筑、宝通、凡拓这一部分的标的。那最新又推荐了像中广天泽、华数这一部分新增的数据包，包括天翼数科。大家讨论的其实是字节系，但是从整个海外市场的交互或者轮动的角度，我们看到了OpenAI包括特斯拉这个阶段进一步强化了对整个数据的、或者物理AI这边数据的支持。那第一个就是GPT-6Astra在整个机器人操作和导航这方面，已经开始挑战过去针对不同任务和机器人本体的传统模型。所以在这个情况下，我们觉得像GPT-6Astra作为独立模型能力上线，而在整个桌面操作、移动平台、灵巧手这个具身智能这边也做了进一步强化。这边其实它也强调了对整个AI数据之一这边需求的强化，所以我们自己觉得这是一个连贯性的需求，就是从整个多模态这边也进入到了新的发展视角。那这个阶段，我们觉得还进一步强化了对整个数据这块补给的需求，因为在大模型的基础上，原始的数据，大家觉得加意义不大，但是现在在整个动态数据集角度来讲，相对来讲是比较清晰的。那第一方面，就是世界模型的数据集，比如说高清多视角4D的视频数据

为核心，这肯定是比较重要的。那第二个，就是在整个具身智能这边的双手操作动作数
据，有布局的公司也相对比较清晰。所以我们在这个维度，再一次强调对整个基于世界模型
和物理AI，对素材、对仿真数据这块的贡献或者布局的公司，其实相对来讲还是比较重要
的。
然后我们再进一步去讨论相关的公司。那第一个，其实市场比较关心像凡拓数创、然后包括宝利科技这边布局的基础。那凡拓也好，还是宝通也好，其实本质上来讲，是进一步
对整个物理AI数据集做了一定的布局。那凡拓的核心，就是它早期其实就在工业物理AI这
边做了比较深度的布局。那现在包括素材，包括数据清理，包括垂直模型和机端的部署，
都有一定的布局。公司领导在体外其实布局了素材的公司，这家公司也做了5~6个点投入，
阶段或者成本也相对比较高一点。
那目前现在已经具备的数采设备，其实大概四五百万套
，所以现在也确实在累计地去推
动数采数据集的标准数据的产出。那另一方面，就是公司跟立讯精密合作多年，从最早的数
字孪生产线已经布局了。那数字孪生当时帮立讯做了大概四个工厂的数字监测技术，所以
已经在工业级别的数据采购和理解当中，有一定的部署。那现在进一步去延伸，以立讯为
代表的这种工业互联网，进一步走向4.0，走向机器人部署的阶段。那简单理解，就是机器
进工厂其实不是一个简单的过程，它也需要对整个工厂的物理数据进行相对的理解，同时
建立仿真数据平台，然后去做独立化的大脑、集成大脑训练，再到集成选型进行落地。所
以实际上这是一个相对比较复杂的一体化工程路径，这一块我们觉得市场可能会更加分明一些。所以在这一块，凡拓在工业化部署、在数采这阶段确实有一定的布局。
那宝通科技也是一样的。那宝通过去是承载业务，之前的语数机器系统已经进入到机
器人进工厂这个领域当中去，同时又投资了光轮智能这样的公司
。所以从逻辑理解，光轮智
能还是非常优秀的，这个数据采集或者仿真数据集售卖的最大、中国最大的，或者相对比
较靠前的公司。那这一块的垂直领域的合作，也进一步推动了数采进一步走向新的一个阶
段。所以，传媒这边大部分公司，是围绕着过去的业务基础，从工业互联网角度阶段走向
了数字孪生，走向了世界模型的数据供给，然后包括进一步进入到法律平台部署的阶段。
那我大概讲一下凡拓，以及宝通科技的基本情况。那后面也请我团队魏征进一步去讲整个物
理AI的阶段性的进度，以及我们对重点公司的方向梳理。好，那剩下的时间交给魏征，有请魏
征。

各位投资者下午好，我是华源传媒互联网组的分析师魏征。今天主要和大家交流一下近期市场关注度比较高的物理AI，以及我们在传媒互联网板块重点看好的一个子方向，就是物理AI的数据采集。我们近期也持续跟踪了字节的世界模型、具身智能数据采集，以及相关上市公司的布局，也组织了字节的产业专家交流。

那近期整个产业的催化其实是比较密集的。比如上周英伟达的GTC柏林大会、语数的物世界物理AI大会，以及特斯拉的业绩会等等，持续提升了市场对于整个物理AI的关注度。同时，国内字节的世界模型也在加快研发进展，后续可能会成为一个比较重要的催化。

我们认为这一轮行情和之前机器人板块围绕特斯拉链的核心零部件展开的行情不一样，这一轮比较重要的变化是全球头部通用模型厂商正在布局物理AI。那过去几年大模型主要是解决文本、推理、代码这样一些问题。其实随着通用模型在空间理解、物理推理和动作规划方面的进展，模型开始逐步进入到物理的世界，也就是说机器人。那它所需要的数据和之前我们所说的文本、然后图片、视频这种数据是存在很大差异的。所以我们认为，这一轮物理AI不仅仅只是模型和机器人本身的产业机会，其实也可能会带来数据采集、数据加工、仿真训练基础设施的一个进展和行情。

国内方面，我们近期组织过了字节的产业专家交流。从反馈来看，字节的世界模型其实已经进入了一个比较关键的产品验证阶段，今年年内肯定是会发布的，那大概就是11~12月份这个时间节点。那根据产业专家的交流情况，就是字节的VideoWorld2，目前已经完成了预训练、互训练、强化学习和评测等等环节，整体处于一个ready的状态。从发布时间来看，目前还没有完全确定，但是12月之前、今年一定是会发布的，可能会在年底和CDS360一起推出。

然后历史上，2024、2025年字节火山引擎的冬季Force原动力大会基本上都是在12月18~19号举行，所以说我们建议重点关注今年12月中旬的字节这边的一个潜在产品催化。所以说我们认为，接下来应该去重点地关注字节链的一些比较重要的公司，因为字节这边世界模型的话，它也是需要很多的这样一个3D的训练数据。

那根据产业专家的交流情况，字节的VideoWorld2的训练数据规模大概就是1270万个小时，主要来自五个方面：第一个就是互联网的公开视频，大概有8890万小时；那字节自己的内容生态，包括抖音、TikTok，大概是210万小时；然后第三一个就是学术公开的数

据集，大概是80多万小时；以及第四个，就是第三方采购以及仿真合成的数据，大约是70
多万小时；最后就是专项采集的数据，是超过了20万小时。
所以说我们从整体的字节数据需求的结构上来看，这个第三方采购以及仿真合成大概
70多万小时，以及专项采集的20多万小时，在整体的训练数据中占比相对有限。所以也恰
恰体现了这个专业的物理AI的数据，和普通互联网视频数据之间的一个差异。所以说我们
认为，未来对这块数据的需求会更加重视专业化和结构化。
那根据专家的反馈，字节的专项采集数据涉及到比如说手工操作、机器人、飞行器等等
场景，需要去获取高清以及多视角以及整个完整动作链条的过程的一个数据。那这部分数
据其实是相对比较有异类和有限的。所以字节仍然需要去第三方采购专业的数据，并且针
对特定的物理场景进行一个专项采集。那根据专家的交流情况，VideoWorld2初期主要是
面对互动影视、开放世界游戏、CG渲染、影视特效以及XR沉浸式娱乐等场景进行布局。
嗯，比如说，未来的互动影视的剧情可能不会完全固定，而是根据用户的选择去实时生成
后续内容。再一个就是对于游戏来说，可能开放世界的游戏可以去利用这个模型去辅助生
成一个3D场景、物体以及动态交互的场景。那更长期来看，其实字节还是希望这种世界模
型的技术去和它的Pico这些终端去进行结合。
所以说，字节的世界模型的发展是具有比较重要的产业意义的。一方面，模型的训练需
要大量高质量真实的数据和仿真数据；另一方面，模型能力的提升以及国内首个世界模型
的发布，有望去赋能影视制作、游戏以及XR的内容应用。所以说，对于我们传媒的公司而
言，物理AI可能会带来一个上游数据采集的机会。第二一个也可能去推动原有的数字内
容，向AI模型训练和新型交互应用的一个延伸，比如说AI游戏、AI影视、互动影游这些。
那为什么这一轮我们重点去关注物理AI的数据采集？我们认为核心原因还是在于，我们现有的互联网几十年来积累的训练语料是非常丰富的，但是物理AI的这样一个训练素
材，它其实是缺乏规模化、缺乏高质量，以及缺乏可直接用于训练的一个真实或者说仿真
的交互数据。
那从目前的一个产业技术路线来看，我们比较值得关注的其实有三类数据。第一类是Ego
数据，也就是说第一人称视角的数据。那它的数据来源就是通过头戴式或者胸带式的设
备，去记录一个人去完成各种操作任务的全过程。那它的优势是在于，能够利用人类自身

的行为和经验去开展大规模的采集，不需要所有的数据都依赖一个真实的机器人去生成，
所以说它具有扩大规模以及降低采集成本的一个潜力。
第二类是UMI数据，也就是通过一个通用的操作接口产生的数据。比如说，我们通过专
门的手持操作设备去采集人类执行任务的姿态、位置、轨迹，以及和物体交互的信息，让
它为机器人的学习、抓取和操作提供一个信号。
第三一类就是机器人本体的数据。那就是让一个真实的机器人放在一个场景里面去进
行遥操作，或者自主去执行任务，然后再去采集这个机器人的状态、视觉、触觉、力觉的
反馈等等这些数据。那这类数据，它的成本是相对比较高的，然后采集的数量也有限，因
为它受限于机器人的数量、硬件成本、效率等等这些限制。
那除了这些真实世界的数据之外，我们同样也看好第二种，就是仿真合成的数据。应
该说，如果所有的训练数据都依靠我们人去进行收集，或者依靠机器人在真实场景中去采
集，那整体的成本其实是非常高的。同时，一些比较危险和复杂的场景，其实也很难去进
行覆盖。所以说未来更具规模化潜力的路径，其实是通过建立这样一个仿真的平台，去再
结合一些3D建模、世界模型、物理仿真，在虚拟的世界中去生成更多的训练场景，然后再
把这些仿真数据喂给机器人，让机器人再回到真实的场景中去进行测试和验证。所以说，
我们认为未来物理AI的数据基础设施会逐步形成，真实采集、数据加工、数据标注以及仿
真合成一个完整的闭环。
那下一个问题就是说，为什么数据采集可能会形成一个新的商业模式？传统的数据标
注更多的是一个劳动力密集型的服务，按照人员投入和项目交付去进行收费。那在大模型
时代，其实数据的价值正在发生变化，特别是对于物理AI，物理AI对于高质量数据的稀缺
性更强。所以说它不仅需要专业采集的设备和场景，还需要完成在一个空间中进行重建、
动作轨迹识别、多模态标注、数据质量校验等等一系列比较复杂的加工处理，它最后才有
可能成为一个真正可以用于模型训练的数据。所以说我们认为，未来在数据采集这一块的产
业，它的价值是来源于三个方面的：
第一是规模化生产的能力，能够以较低的成本去组织专业的场景、人员以及设备，实现数据大批量、大规模的持续供给。第二个是数据加工的能力，就是说能够把一些普通的3D、视频
去进一步加工成具有空间坐标、动作轨迹、物理属性的高质量的训练数据。第三一个就是数据的复用和平台化能力，就是能够在合规授权的前提下，将同一个数据资产服务于不同的模型、

不同的任务和不同的客户，并进一步去提供一些仿真训练、评测等等增值的服务。

所以说，从产业机会来看，未来可能会形成定制化的数据采集、标准化数据集的授
权，以及数据采集和仿真训练一体化服务的多种商业模式。对于传媒互联网板块来说，过
去很多公司积累了专业的视频拍摄、影视制作、3D数据资产、数字孪生以及虚拟场景建设
等等能力。那这些能力其实和物理AI的数据采集和生产，它们是有一定的重合度和相关
性，以及业务的交集的，所以说我们觉得未来一定要重点去进行关注。
接下来，我们讲一下在华源传媒互联网板块物理AI的一个投资机会。那我们重点推5家
公司，分别是风语筑、中广天泽、宝通科技、凡拓数创和天翼数科。
那首先是风语筑。那这公司，它过去是长期从事企业展馆、数字空间和沉浸式内容制
作的一个公司，那它其实是在3D场景构建、数据资产制作以及渲染和空间交互方面会有一
定的积累。并且今年公司投资并联合松延科技去布局物理AI相关业务，并且他们成立了一
个合资公司叫风语科技，去后面可能会承接数据方面的一些业务。第二个看点呢，是
Pico。公司之前已经基于自己的Pico技术生态，推出了一个VR大空间项目，并且具备一定
的VR场景开发、沉浸式交互以及终端应用能力。这个Pico，不仅仅之前它只是一个内容展
示的终端，所以未来也具备去向第一视角的数据采集设备去延伸的一个潜力。特别是现在
物理AI数据对这种Ego第一人称视角的需求会比较大，那公司在这一块会有一个潜力。
第二个呢，是中广天泽。那中广天泽过去主要是从事影视内容制作以及版权运营。那
长期以来，实际上也积累了很多的真实世界的音视频内容资源，同时也具备专业视频采
集、内容制作以及数据加工的能力。那从公司的官网来看，目前公司会有3个AI数据产品。
第一个是世界模型的数据集，重点提供高清连续多视角的4D视频数据、空间运动以及真实
世界声学数据，能够去帮助模型理解物体运动过程、空间关系和物理规律。第二一个是公司有
AIGC的数据集，那主要是依托公司原有的影视制作和音视频资源，为生成式AI提供高质量的训练素材。第三一个是具身智能的训练集
，那重点围绕人类的动作、真实的环境以及物体交互过
程，提供相关多模态的训练数据。
那公司还有两个优势。第一个就是采标一体化形成的一个差异化优势
。那公司具有真实
的场景数据，覆盖超过1万个室内外的场景，那涵盖工业生产、社区家居、商业消费、体育

以及自然环境等等领域。并且呢，公司拥有数千人的垂直标注团队，同时具备前端真实场
景采集以及后端的标注和结构化加工的能力。第二呢，就是根据公司半年报，公司目前的
一个数据集总规模超过了10TB，那其中公司的客户包括基业新城、中国电信等等，并且已
经实现了相关AI算料产品的一个批量交付。

第三一个是宝通科技。那公司最值得关注的就是它有光轮智能的股权，他们俩进行一
个合作，围绕真实的工业场景去构建物理AI数据以及训练的基础设施。那我们这家科技
呢，之前在工业运输、智能矿山这些领域拥有长期的业务积累，所以也是具备真实的工业
场景的资源。同时，公司和光轮智能展开了一个深度合作，他们俩围绕工业场景的数据采
集、仿真合成、模型训练以及评测验证等等环节进行了这样一个布局。
第四一个是凡拓数创。那凡拓数创过去主要从事数字创意、3D可视化和数字孪生相关
业务。那在3D的场景制作、工业数字空间以及实时交互技术方面具备积累
。并且根据公司此
前的一个交流情况，凡拓在物理AI建立了四类的数据布局，分别是空间数据、仿真交互数
据、无本体数据和本体数据。那在技术平台方面，公司也正在推进仿真平台的一个相关建
设，探索为物理AI提供场景生成、仿真交互和模型训练能力。所以说，凡拓它的一个优
势，不仅仅只是包括数据采集，还包括3D数据、3D数据资产、仿真平台，以及工业机器人
训练解决方案这些。那公司此前也披露，它和立讯精密也进行了这样一个合作，去探索在真
实的这种流水线上的一个场景中的一个解决方案。
第五家呢是天翼数科。那根据公司的半年报，天翼数科已经积累了超过570万条多模态
的数据。那其中3D的数据超过200万条，那涵盖了高质量的纹理、铰链结构和可变形物体
等等类型。那公司部分高质量的3D数据和多模态具身机器人抓取的数据集，也完成了相关
的数据资产登记。那我们认为，这一类的数据，它和物理AI的训练需求也存在一定的匹配
度。那在此基础上，公司也在进一步布局空间智能模型以及相关的数据服务平台，希望能够
将原有的数据资源转化为模型训练和下游应用的能力。
那总的来看，我们认为物理AI正在推动机器人产业链的价值环节，进一步向上游的数
据和训练基础设施去延伸。那我们传媒板块中具备专业的内容生产、3D数据和仿真能力的
公司，有望迎来一个新的业务发展机会
。那以上就是我们今天关于物理AI数据采集方向的一个

汇报。

--- 物理Ai系列：具身大脑专家交流20261025.pdf ---
全文总结
本次电话会议为物理AI系列之具身大脑专家交流，邀请在视觉技术、深度学习等领域有多年经验的专家，围绕物理AI与传统AI的差异、物理AI的三层落地架构与核心优势、商业化落地场景、具身智能与物理AI的关系、产业落地瓶颈及未来演进判断等话题展开分享，并在后段回应了投资者提问。
一、物理AI与传统AI的本质差异
传统大模型是纯数据驱动，基于统计概率去做拟合，本质上"见过数据就会、没见过就会错"，存在完全不符物理规则的极端工况失效问题，一般处理文本、图像这类信息；
物理AI最大的区别，是把物理机理、动力学约束、客观规律嵌入到模型里面，不是单纯靠数据去推结果，它所有的预测和控制输出都严格遵守现实世界的规则，具备可解释性、可泛化性、低数据依赖的特点，是能够对接真实物理世界的AI方式；
这源于黄仁勋提出的物理AI概念。前两周黄仁勋去柏林GDC与德国总理交流时，也讲到德国在物理AI这一块动作相对偏慢，但德国有一个比较好的先天优势——制造业基础比较强大，物理AI的基础元素仍然存在，在建设物理AI过程中会比较便利。
二、物理AI的三层落地架构
最底层是物理机理层，提供物理世界的边界条件、动力学参数，比如机械或机器人关键模组的摩擦力、弹性系数等，给定模型合理的物理边界；
中间层是AI拟合层，用神经网络去解决传统物理方程求解慢、非线性复杂建模困难的问题。有点像黑盒——传统方法有基本的逻辑和数学模型，去求解输入输出对应关系；而物理AI中间的拟合层也属于神经网络层，端到端输出的中间层是一个黑盒，没办法通过传统数学公式去衡量；
最上层是工程部署层，也可以认为是应用层，输出毫秒级推理的模型，包括仿真、控制、预测、数字孪生这些；
简单讲，物理AI就是三层：物理定边界、AI做拟合，最后实现智能化的控制。
三、物理AI的三大核心优势
第一，解决工业数据稀缺的问题。目前在具身智能应用过程中采集的数据，很难1:1复原工厂的数据；无论通过UMI设备、还是搭建类似工业制造场景去采集数据，相对于工业来讲还是太过理想——采集桌面都是白色背景，光线也非常稳定。物理AI可以在虚拟世界或者依托机理补全整个场景，通过变化不同材质、不同光线去做训练；

第二，极致提升仿真效率。传统动力学仿真基本是10分钟甚至小时级别的，整个过程周期非常长；而物理AI推理基本都是毫秒级的，能满足实时控制的需求；
第三，消除sim-to-real偏差，解决传统仿真与实物脱节、落地难的核心痛点，通过实现数字世界和物理世界精神对齐的方式。这也是现在具身智能遇到的最大瓶颈和难题——仿真环境里很好，从仿真导入到真机的时候，如何消除sim-to-real的gap。另外

一个，仿真里面之前的状况以及周边环境的状况，如果没有物理AI属性，那可能更多的只是一个聊胜于无的布局、一个简单场景类似而已。

四、物理AI的商业化落地场景

物理AI现在有三大载体，人形机器人、无人驾驶、无人机——从地上到偏作业的、再到天上。商业化主要集中在真正能变现、落地、终端用户愿意付费的高科技领域；

人形机器人和工业机器人可能是比较好的物理AI载体产品形态，因为这完全颠覆了现在工业机器人的作业逻辑。现在工业机器人基本上靠工程师和工厂规划去讨论整个产线布局、节拍甚至工艺生产流程，再由方案工程师、调试工程师完成产线交付，这个过程中人参与了很多前置工作。有了物理AI，工业机器人在产线里面想做的事情，可以提前在仿真里去构造出来；后面最难的是如何提高产线里的节拍和效率，现在基本依赖工程师经验，而有了物理AI引擎，可以和发那科、ABB的RobotStudio，以及罗克韦尔（Rockwell）联合去做，通过数字孪生技术更快地把整个产线节拍做快速提升。仿真里提升完以后，导入到真机里就能看到真实提升，这就是物理AI的商业落地价值；

第二是高端装备的结构化仿真，比如振动疲劳测试、芯片预测，可以替代传统部分仿真流程，像航空航天、飞机、汽车这些都包含这一类；

第三是无人驾驶，已经从乘用车扩展到商用车，比如无人矿卡这些无人设备的动态环境物理适配以及轨迹控制；

第四是基于物理退化规律的工业设备精准故障预测。设备在长时间服役过程中，能否通过大量数据、经过物理退化规律预测，提前知道什么时候会发生故障，提供更精准的预测。现在有些设备也能提供故障预警，比较麻烦的是厂方说要更换设备或零部件了，终端使用者会怀疑他的预测是否绝对正确；有了物理AI引擎，可以给予一个中立方去做判断。

五、具身智能与物理AI的关系

物理AI和具身智能不是完全从属的关系。具身智能可能是一个更大的范围，本质是让机器人像人一样具备感知环境、交互环境、自主执行任务的智能体；

但目前全大模型驱动的具身智能，最大的问题是不懂空间物理关系以及物品的物理属性。机器人与物理世界交互过程中，比如碰到杯子，杯子落下、重力加速度是什么，在普通环境里其实没有去做。现在看到的人形机器人更多是静态的pickandplace，没有去做像守门员守门、球体运行过来的抛物线运动轨迹提前预测这类复杂智能体工作；

所以物理AI其实是具身智能的底层基座、底座，负责帮助具身智能理解物理世界的属性关系，包括约束、动力学关系、姿态、环境理解。具身智能解决"想干什么"，物理AI解决"怎么能够比较好地干出来、不出错地干出来"。

六、具身智能产业真实落地的瓶颈

从今年工博会看，现在具身机器人基本上讲的都是解决某一个场景任务的能力，而不是模型能力。可以看到约90%的具身智能产品是人形机器人，10%是双手人形机器人，这背后是工博会面向工业制造，可以看出大家觉得具身智能可能优先在工业制造场景首先落地；

但展会里很多解决具体任务的场景，大家趋向于一致去做类似人形机器人搬包裹、纸箱或分拣小零部件的事情，我认为他们选的场景是错的。具身智能应该解决泛化型以及
柔性化的问题，而搬箱子、搬料箱、分拣零部件，在工业自动化背景里20年前就已经被工业机器人取代、实现了。这说明瓶颈还是模型能力不够，还不能做偏泛化型或高精度的
任务，因为具身机器人应该解决的是工业自动化最后一公里
——即非标自动化没能取代
的那些偏柔性体对象，比如汽车内外饰保险杠、发泡座椅、汽车零部件里的线束整理，这些活儿基本还是人工在干；
瓶颈一，环境泛化能力弱。目前实验室是比较规整的状况，机器人表现相对比较好。有些demo，比如星海图、自变量在WAIC上展示的纸盒折叠、把东西放进去再贴标签，这是长程任务的端到端，因为背景环境或作业对象相对比较规律
，他们demo的表现还不错；
瓶颈二，控制精度与稳定性不足。现在没有一家具身机器人用VLA能保证99.999%的稳定性或成功率，但工厂里其实很难容忍只有90%的成功率；另一个是控制精度不足，现在看到都是夹爪，没有精确抓取，每次抓的位置都不一样，这是目前VLA最大的问题。今
年工博会上有些厂商在尝试解决精度问题，比如北京某公司不生产整机，用一个天机的全（力）控臂做了一个线束（柔性线束）0.3mm的装配或插拔，在展会上看了几个回合基本都能准确插进去。他们不是单纯用VLA，而是也用了强化学习去控制；
瓶颈三，端到端落地成本太高，这也是具身智能产业没办法真正商业落地的最大瓶
颈。成本高可以分几个环节：数据成本高、训练成本高，还要搭建算力服务器和边缘设备，且现在人形机器人本身价格很高。通过模型去适配不同硬件的动力学参数、跨本体的问题仍然存在。尽管像蚂蚁、宁波某说我的大模型在家炒菜，用的是不同厂家的机器去做，但实际上有没有做后续训练、微调以及从零开始这些问题，展会上看不出来，跟技术人员聊，其实从从业者来讲没办法完全做到非常少工作量的框定；
所以每台机器人都需要微调，每台机器人针对需要干的场景都需要后训练，这个成本
太高。工业机器人虽然也要工程师调试、换了工件再重新调试，但现在具身智能去做一些场景和工艺，都是工程师把工艺拆解以后去遥操机器人、采集大量数据，而不是一
点点数据，背后付出的时间成本、算力成本、电力成本、维护成本看起来都比工业机
器人更贵。比如千寻在宁德时代做电池电芯的插拔，在现场训练了三个月，而这个动
作放在工业机器人现场布局后基本一周就结束了；三个月训练出来后才做到60%成功
率，后面还要再追加大量cornercase的数据。一台人形机器人基本上要六十几万，
而工业机器人现在便宜的三五万，再加调试人工三五天
，这就很难说未来具身智能产品
怎么去落地；
综上三个瓶颈：环境泛化能力弱、控制精度稳定性不足、端到端落地成本高。
七、具身智能未来产业演进的判断
从行业迭代方向来看，有些事情是明确的，比如具身智能和LLM、GPT相比，是不是也会存在GPT的拐点时刻，我们也在看到scalinglaw收敛的状况，所以大的方向应该是对

的。大家迭代方向比较明确：
一是，从大模型单点驱动，转向大模型加物理AI双驱动，这是目前人形机器人厂商比
较认可的方式——大模型是大模型，物理AI帮助大模型更好地去理解这个世界，是一
个双驱动架构；
二是，从实验室演示效果，转向可量产、高稳定、低微调的工程化能力，这是物理AI
帮助人形机器人或者具身智能量产落地的关键门槛；
三是，从感知智能、语言智能真正走向物理交互智能、行动智能，可能也是下一代AI
工业化的核心突破口。
专家表示以上观点比较狭隘，因为基本是基于工业自动化背景去讲的一些过程，在to C或
to G可能会有差异，大家可以一起讨论。
会议实录：
主持人：
这一场我们继续交流学习物理AI产业。今天邀请到的专家在相关领域具有多年丰富经验，对视觉技术、深度学习等都有非常深的认识。今天先请专家就物理AI跟传统AI的差异、物理AI的落地技术核心优势、落地场景，以及具身智能与物理AI的关系、具身智能当前落地的瓶颈、产业演进的判断等话题，给我们先进行一个阐述，后续再留一些时间给大家提问。要不先请专家开始。
专家：
好的。最近大家可以看到，物理AI包括之前黄仁勋提出要下场做人形机器人集群这一块，整个物理AI以及具身这一块的话题也变得非常热。这里面有几个观点或者说几个概念，还是希望跟大家阐述一下。
第一，物理AI和传统AI的本质差异。这一点我觉得还是得强调一下。传统大模型是纯数据驱动，基于统计概率去做拟合，本质上是"见过数据就会、没见过就会错"，存在完全不符物理规则的极端工况失效问题，一般处理文本、图像这类信息。物理AI最大的区别，是把物理机理、动力学约束、客观规律嵌入到模型里面，所以它不是单纯靠数据去推结果，它所有的预
测和控制输出都严格遵守现实世界的规则，具备可解释性、可泛化性、低数据依赖的特
点，是能够对接真实物理世界的AI方式。这也是黄仁勋提出物理AI概念之后的一个核心方向。前两周黄仁勋也去了柏林GDC，跟德国总理去聊，也讲到德国在物理AI这一块动作相对偏慢，但这里面他强调了一个德国比较好的先天优势——制造业基础比较强大，所有物理AI的基础和原子元素还是存在的，所以在建设物理AI过程中会变得比较便利。
第二个话题是物理AI的三层落地架构。这一块还是得跟大家梳理一下。物理AI从最底层来看，就是物理机理层，提供物理世界的一些边界条件、动力学参数，包括机械或者机器人
机身里面关键模组的摩擦力、弹性系数这些参数，给定模型合理的物理边界。中间层是AI拟合层，是用神经网络去解决传统物理方程求解慢、非线性复杂建模困难的问题。这里面有点像黑盒——过去传统方法有基本的logic和数学模型，去求解输入和输出对应的关系；但物理AI中间的拟合层也属于一个神经网络层，它端到端输出的中间层是一个黑盒，没办法通过传统数学
公式去衡量。最上层是工程部署层，我们也可以认为是应用层，这里面要输出毫秒级推理的
模型，包括仿真、控制、预测、数字孪生这些东西。所以简单讲，物理AI就是三层：物理定边
界、AI做拟合，最后实现智能化的控制。

物理AI的三大核心优势是什么呢？这里还是跟大家分享一下。
第一，物理AI可以解决工业数据稀缺的问题。我们现在在具身智能应用过程中发现，采集的这些数据很难1:1去复原工厂的数据，无论是通过UMI设备、还是搭建类似工业制造场景去采集数据，相对于工业来讲还是太过理想了，因为可以看到采集的桌面都是白色背景，光线也非常稳定。物理AI可以解决这部分问题，在虚拟世界或者依托机理补全整个场景，然后去变化不
同的材质、不同的光线去做一些训练。

第二，物理AI的优势是极致提升仿真效率。传统动力学仿真基本是10分钟甚至小时级别的，作为一个比较复杂的动力学仿真，整个过程周期会非常长；而物理AI推理基本都是毫秒级
的，能满足实时控制的需求。
第三，物理AI可以消除sim-to-real的偏差，解决传统仿真与实物脱节、落地难的核心
痛点，通过实现数字世界和物理世界精神对齐的方式。其实这也是现在具身智能遇到的最大
瓶颈和难题——仿真环境里很好，从仿真导入到真机的时候，如何消除sim-to-real的
gap。另外一个，仿真里面之前的状况以及周边环境的状况，如果没有物理AI属性，可能更多
的只是一个聊胜于无的布局、一个简单的场景类似而已。
第四个话题，前面都是讲技术的，后面讲物理AI核心商业化落地的场景有哪些。过程
当中我也跟大家聊到了人形机器人。现在物理AI的三大载体，大家都知道，人形机器人是一个、无人驾驶是一个，还有一个是无人机
——从地上到偏作业的，再到天上。这商业落地三
大载体有了以后，主要集中在真正能够变现、落地，终端用户愿意付费的高科技领域。我认为
人形机器人和工业机器人可能是一个比较好的物理AI载体产品形态
，因为这完全颠覆了现在
工业机器人的作业逻辑。现在工业机器人基本上都是工程师和工厂里面的规划去讨论整个产线的布局、节拍甚至工艺生产流程，然后由工程师、方案工程师以及后面的调试工程师去完成最后的产线交付，这个过程当中其实人参与了很多前置工作。如果有物理AI，工业机器人在产线里
面想要做的事情，是可以提前在仿真里面构造出来的
；后面最难的是工业机器人在产线里面
如何提高节拍和效率，现在基本依赖工程师经验，但这个经验是不是能够充分合理地最大化、对产线节拍提升的最大化？如果有一个基于物理AI的引擎，因为可以达到毫秒级（推理），可
以和发那科、ABB的RobotStudio去联合做，和罗克韦尔（Rockwell）去联合做，通过数
字孪生技术，把产线里面的整个节拍快速提升
。这个在仿真里去做，仿真里提升完了以后导
入到真机里面，就确实能看到真实的提升，这就是物理AI商业落地的价值。
另外一个就是高端装备的结构化仿真，比如振动疲劳测试、芯片预测，可以替代传统部分的仿真流程。这里面其实很多，高端装备的结构化仿真，像航空航天、飞机、汽车这些都包含这一类。
另外一个就是前面说到的无人驾驶，已经从乘用车扩展到商用车，比如无人矿卡这些无人设备的动态环境物理适配以及轨迹控制，这也是一个很好的商业落地场景。

第四个是基于物理退化规律的工业设备精准故障预测。
前面都在讲自动化布局以及无人
化、自动化与物理世界产生交互的过程，第四个是设备在物理长时间服役的过程中，有没有可能通过大量数据、经过物理退化规律的预测，提前知道工业设备什么时候会发生故障，提供更加精

准的预测。因为现在有些设备也提供故障预警，比较麻烦的是厂方说要更换设备或零部件了，终端使用者会有疑问——你这个预测绝对正确吗？现在有了物理AI引擎，可以给予一个中立方
去做判断。
后面讲讲具身智能和物理AI的关系。前面介绍了这么多，好像物理AI和具身智能很相似，那究竟物理AI是啥、具身智能是啥？我认为物理AI和具身智能不是一个完全从属的关
系，具身智能可能是一个更大的范围
，它的本质是让机器人像人一样具备感知环境、交互环
境、自主执行任务的智能体。但目前全大模型驱动的具身智能，最大的问题是不懂空间物理关
系以及物品的物理属性，甚至机器人在与物理世界产生交互的过程当中，比如碰到了杯子，杯子后面的落下、或者说重力加速度是什么，在普通环境里面这一块其实我们没有去做。因为现在看到的人形机器人更多是静态的pick and place，没有去做比如守门员守门、球体运行过来的抛物线运动轨迹提前预测这类比较复杂的智能体工作，现在还没看到在做。所以物理AI其实是具身智能的一个底层基座、底座，它负责帮助具身智能理解物理世界的一些属性关系，包括
约束、动力学关系、姿态、环境理解，这是它的一个基础。所以具身智能是解决"想干什
么"，物理AI是解决"怎么能够比较好地干出来、不出错地干出来"。
第六个话题，具身智能产业真实落地的瓶颈。作为从业者来讲，从今年工博会可以看到，现在具身机器人大家讲的不是模型能力，基本上都是讲解决某一个场景任务的能力。可以看到约90%的具身智能产品是人形机器人，10%是双手人形机器人
，这背后是工博会面向工业制
造，可以看大家觉得具身智能可能优先在工业制造场景首先落地
。但是我们现在看到很多展
会里解决某个具体任务，这是一件很好的事情，但我个人发现他们做的这些事情趋向于一致，去做了一些类似人形机器人搬包裹、纸箱或者分拣小零部件的场景，我觉得他们选的场景是错的，因为具身智能应该解决泛化型以及柔性化的问题
。搬箱子、搬料箱、分拣零部件，在工
业自动化背景里，20年前这些东西都已经被工业机器人取代、已经实现了。这里面说明了什么呢？说明了瓶颈还是在，就是模型的能力还不能去做一些偏泛化型、或者偏高精度的任
务。因为具身机器人应该解决的场景是工业自动化最后一公里
，什么是最后一公里？就是现
在工业自动化以及非标自动化已经取代的那些人、工厂里面剩余的那些工人，他们在产线上干什么，这个可能就是具身机器人应该去取代这些人去做的事情。因为非标自动化以及自动化没能取代人的，一般都是那些作业对象比较偏柔性体的，比如汽车内外饰的保险杠、发泡的座椅、包括汽车零部件里面的线束整理，这些活儿基本上都是人工在干。那大模型能不能适配具身机器人，实现像人这么精准、这么柔性地泛化去作业呢？这里面就是几个瓶颈：

一是环境泛化能力弱。目前实验室是比较规整的状况，机器人去做表现还是相对比较好的。有一些demo，比如星海图、自变量，他们在WAIC上展示的像纸盒的折叠、把这个东西放进去再贴一个标签，这是长程任务的端到端，因为背景环境或者说作业对象相对比较规律，他们demo

的表现还是不错的。
第二个是控制精度与稳定性不足。当然现在没有一家具身机器人用VLA能保证99.999%的稳定性或者成功率，没有一家能保证的；但如果在工业制造或者工厂里面去用，工厂里面其实很难容忍只有90%的成功率。另外一个就是控制精度不足，现在看到都是夹爪，都没有精确抓取，每次抓的位置都不一样，这其实里面就是现在VLA目前最大的问题。我也看到今年工博会有些厂商也在尝试解决精度问题，比如北京某家公司，他们不生产整机，不像宇树、乐聚、智元他们有自己的（人形）机器人，他们就用了一个天机的全（力）控臂，做了一个线束——就是柔性线束0.3mm的装配或者插拔，他们做得非常好，在展会上我们看了几个回合，基本上都能准确插进去。他们是怎么做呢？回过头来发现，技术上并不是单纯用VLA，他也用了强化学习去控制。
第三个是端到端落地成本太高。我觉得这也是具身智能产业没办法真正商业落地的最大的瓶颈。成本高可以分析几个环节：一个是数据成本高，另外训练成本高，然后还要搭建算力服务器和边缘设备，包括现在本身人形机器人的价格又很高。通过模型去适配不同硬件的动力学参数、跨本体的问题其实还是存在的。尽管像蚂蚁、宁波某企业说我的大模型在家炒菜，用的是不同厂家的机器去做，实际上有没有做后续训练、微调以及从零开始这些问题，在展会上看不出来，跟人家技术人员去聊，其实这个里面从我们从业者来讲没办法完全做到非常少的工作量。所以每台机器人都需要微调，每台机器人针对需要干的场景都需要后训练，这个成本太高。
那这么做的话，大家可能会有一个非常大的疑问。工业机器人前面说要工程师去调试，换了一个工件重新再去调试；其实现在看具身智能，具身智能去做一些场景和工艺，都是工程师
把工艺拆解以后去遥操机器人、采集数据，而且是大量数据，而不是一点点数据，背后付
出的时间成本、算力成本、电力成本、维护成本看起来都已经比工业机器人更贵。比如千
寻在宁德时代做电池电芯的插拔，他们在现场训练了三个月，而这个动作放在工业机器人
现场布局以后，基本上一周就结束了；三个月他们训练出来以后，才到60%成功率，后面还
要再追加大量cornercase的数据。这个确定的成本就不说了。然后一台人形机器人基本上
要六十几万，但工业机器人现在便宜了也三五万，再加上调试人工三五天，这就很难说未来
具身智能这个产品怎么去落地了。所以我提到三个瓶颈：环境泛化能力弱、控制精度稳定性不
足、端到端落地成本高。

第七个就是具身智能未来产业演进的判断。从行业接下来迭代方向来看，有些事情是明确的，比如具身智能和LLM、GPT去比的话，是不是也会存在GPT的那个拐点时刻？我们也在看到scalinglaw收敛的状况，所以大的方向应该是对的，大家去迭代的方向还是比较明确的：
一个是从大模型单点驱动，转向大模型加物理AI双驱动，这是所有人形机器人厂商目
前比较认可的方式——大模型是大模型，物理AI帮助大模型更好地去理解这个世界，是一
个双驱动架构。
第二个是从实验室演示效果，转向可量产、高稳定、低微调的工程化能力，我觉得这
是物理AI帮助人形机器人或者具身智能量产落地的关键门槛。

第三个，我认为具身智能从感知智能、语言智能真正走向物理交互智能、行动智能，
可能也是下一代AI工业化的一个核心突破口
。我这个观点可能比较狭隘，因为基本是基于我
的工业自动化背景去讲的一些过程，可能在to C或者to G会有一些差异，大家可以一起讨论。所以这几个话题，我就阐述一下我个人的一些理解和观点，欢迎大家提一些问题。
主持人：
好的，感谢专家的分享。下面看看大家有什么问题可以提一提。
（会议助理提示提问方式）
主持人：
有人有问题吗？
提问一：关于GPT-6与物理AI双驱动范式
参会者：专家您听得到吗？我有个问题请教一下。您刚刚讲到大模型和具身的技术范式，您觉得还是两个共同驱动。这次其实关注到，一个月前OpenAI的GPT-6出来，当时他们用GPT-6去结合具身机器人，感觉比之前提升了蛮多，您怎么看待这对整个行业的影响？
专家：这是OpenAI奥特曼他们那个零样本的，应该说是GPT-6，做了一个零样本的pick and place的成功率可以达到八十几还是九十几，后面又做了一个测试，就是让它去做一个新的、有难度的任务时，成功率会降得比较快。这个事情说明了什么呢？我通过大量的互联网数据、视频数据去训练的模型，可以让机器人去做一个简单的pickandplace任务，成功率
很高。这放在我们工业自动化里面可以这么理解：本身LLM解决的是物理世界物体是什么的问题，通过端到端，说我要去抓一个什么东西，通过头部相机、外部相机去做空间定位，基本上是能够抓起来的。因为目前具身机器人的手掌和物体的误差都非常大，不像工业机器人是非常精密的抓取，因为误差非常大，所以你放了一个物体，或者说放了一个没有经过训练的、但基模其实有大量数据的新的物体，它也能准确抓起来，我觉得这必然是能实现的。

但这个实现对我们后面有两方面影响：一方面是证明了通过大量数据可以看到scaling law的存在，这是事实；第二个事情，作为从业者来讲，虽然说它能够pickandplace，并
不能代表它未来做所有的事情、你给它下一个事情它都能完成
。因为我们现在已经在尝试一
些长程任务，工厂里面工人做的事情，在他独立工位不可能只做一个动作，基本是一个序列的动作；让具身智能机器人去实现一个序列的动作，过程是比较复杂的。现在行业内怎么做呢？以往我们希望是端到端，把整个长序列动作都给它看一遍，然后让它复现整个流程和动作，其实这里面会比较难，有很多事情，比如模型的记忆功能——我有没有记住上一个步骤干了还是没干、上一个步骤干了以后怎么触发下一个动作，这里面都是一系列的工程化问题，已经不是大模型的问题了。所以现在行业内我们看到长程任务端到端已经变成了叫"断到断"
，什么意思呢？比如
我去做一个事情，打开盖子，然后把旁边的一个物体放到盒子里面的仿真模型的洞里面，然后再在上面放一个说明书，再把盖子盖上，这么长的动作基本上都是通过工程师去分解：先训练我怎么100%或者90%成功地把盖子打开，然后再训练我怎么准确地抓取到手机模型，然后再训练把手机模型放入到盒子里面的仿真洞里面去，每一段全部分解下来，每一段用VLA去实现，然后再把这些东西做一个拼接，实现一个长程任务。
那如果说我们的终极目标是机器人说一句话就能理解我给的明确指令——当然我可以明确说
你先打开这个、再去跟那个，它就能直接去做。但目前我们看到，都还没有机器人自己去理
解，而是背后的工程师把他想要做的事情已经拆解、训练完了，然后提前才能达到这个效
果。所以这是目前我们看到的一个最大的问题。不知道这个答案能不能解答你的疑问。

参会者：明白。其实就是现在大模型它也可能解决其中的一部分问题，对吧？
专家：对。
提问二：关于数据瓶颈
参会者：数据这块您怎么看的？不同场景的数据，包括工业场景或者人形场景的数据，现在是一个比较大的瓶颈吧，现在边际上有没有一些变化？
专家：这个就回到这个topic，就是讲大模型和物理AI、或者说通用智能模型的本质差异。因为前者是通过大量的互联网公开信息就可以获得的；但后者因为讲究的是具身落地和作业，里面很多场景，无论是toC还是toB，这些场景的数据很多都是偏私有化的。那为什么有物理AI以及数字孪生和仿真环境？它们就是为了规避如何获取这种私有化数据，既然太私有化，那我就在仿真里面自己去搭，但这个搭建的成本又很高。所以现在数据已经是一个比较……其实2026年是一个数据元年了，无论是政府还是企业还是高校，他们都搭建自己的数据的一些场景。但这个场景基本上都是人为复现的一些场景，无论是工业、家居还是生活的一些场景。但这个场景我认为前面也表达了，也是相对比较理想化、结构化的一些东西，因为它周边太干净了，采集的数据周边环境干扰非常小，noise非常少，干扰因子很少，相对是比较干净的。
但即使是这样，现在数据又分高质量数据和低质量数据。所谓高质量、低质量是怎么分的呢？为什么有的数据可以卖到600块钱一个小时，有的数据只能卖到200块钱一个小时？
因为低质量的数据，比如Ego（第一视角），它本身采集的成本就比较低，不需要像UMI、主从臂那样去买那么多硬件设备才能采集，它就是人以第一视角去作业去产生的数据。但这些数据用途又不一样，像Ego的数据，它只能给基模训练去用，意味着我训练一个认知万物的、类似LLM这样的大模型来讲，它是有帮助的；但是，你说我让它去看一个作业视频或者工人第一视角的作业过程，机器人就立刻能干这个活吗？我们发现是不能的。因为这里面还是需要知道真机机器人各个关节的URDF以及它的重心、电箱以后，它的各个关节和你的视频怎么去映射过去，然后我的机器人怎么去到达你的末端、你的作业对象，和作业对象之间发生一些物理关系。所以可以看到，数据它的类型不一样，有的数据可能是纯视频流，但有的数据会增加一些末端触觉进去；有的数据除了以上，还会加上重构的、和它整机机器人对应的关节的URDF的比例关系进去，或者它执行器的运动轨迹关系进去，这样的话，给到机器人去做实际应用场景出来的效果会更好。所以数据这一块基本上就是这么一个情况，我跟大家稍微拆解一下，无论是从技术还是从金额，跟大家分享一下。
提问三：关于大脑能力壁垒与厂商格局
参会者：专家老师您好，我这边后台收到一个文字提问，想跟您请教：GPT-6是否让行业看到一种趋势，就是大脑这端能力的壁垒在打破，或者说利好已有小脑、本体、运控能力的厂商？你如何看待通用大模型厂商和目前具身玩家的格局，二者的平衡和边界您可以介绍一下吗？
专家：这个问题有点长。前面是GPT-6对大脑模型厂商的技术壁垒是否会降低？我现在来看，我觉得做大脑和大模型的，可能在整个具身行业里面，它的生态位会比较高一点，我们讲究生态位。因为本身像本体，未来肯定是趋向于成熟，以中国的供应链来讲，哪一家其实都能造；但你大脑的能力能否去赋能你的本体，那这个相反可能会有一些变化。所以我觉得大脑不会把它的壁垒能力去降低。但这里面有一个问题，换了本体以后，它的执行成功率会大大降低。这个时候，OpenAI他们发现了这么一个问题，那可能会有个新的观点，就是是不是大模

型要和本体做双结合、双驱动以后，我的具身机器人在落地端落地的可能性会大大提高。
这有可能，因为目前从数据采集、训练以及跨本体的问题来看，我们看到还是需要一些本体的信息在的。

这也正好验证了奥特曼为什么要讲，2027年要自己把那些机器人造出来——这样的话，他的大脑对于他的机器人，它大脑更能去理解机器人本体的一些实际的物理参数。这其实也跟物理AI联系在一起了，更能理解它本身机器人的性能和能力范围，这样它的大脑去控这个本
体，才能做更好的匹配。就像一个长跑运动员，他能力很强、大脑很强，但如果只是一个小孩的身体，那其实他这个训练的模式什么他都能理解，但这个身体不能驱动他达到跑100米9秒以内，对吧？我觉得这个是（第一点）。
第二个，通用大模型和具身模型，这里面概念还是有点混乱的。什么叫通用大模型？
我们按照马斯克（提出的）第一性原理来讲，他讲的通用大模型是终局，就意味着未来具身机器人是不是能够像人一样达到通用智能的水平，意味着我这个机器人其实不是一个to B的产品了，应该更多像一个to C的产品，我打开这个机器人以后开箱即用，今天跟它说去做这么一件事，明天跟它说做那件事，我不需要再通过所谓的微调、训练、学习以后再怎么样，直接上了产线就能干活，那这个是通用智能。那具身智能和通用智能其实就相差一个载体，因为具身智能就是
embodiedAI，你的具身智能要嵌入到某一个载体里面去，帮你去实现这样的能力。这又回
到我前面回答的问题：具身智能肯定要和你的具身载体的物理属性匹配，你的大脑和你的本
体的能力要匹配，不然你大脑再强，你的本体是一个侏儒或者帕金森，那是没办法去完成
这样的响应的。 这是我以上个人的理解，仅供大家参考。
主持人：还有其他问题吗？下面有请。
提问四：关于数据采集产业链梳理
参会者：专家您好，可以帮我们做一下数据采集这个产业的产业链梳理吗？核心包括仿真数据平台、数据采集的硬件设备，还有零部件，以及你去数据采集的时候相关的服务，这几个核心环节能帮我们梳理一下吗？大概有哪些公司，各自的价值量占比情况是什么样的？
专家：这个问题其实也挺好的，我尝试着从从业者角度去梳理一下。因为具体的产值以及数据我可能没办法提供给大家，只能从结构和层面上帮大家去处理一下。就现在数据采集，对于人形机器人来讲，有几种方式：
第一，Ego（第一视角）这种方式，就是第一人称视角的无感采集，这也是成本最低的
一种采集方式。
第二，UMI的方式，就是我人带着头部一个摄像头，然后手是一个和本体基本上重构的假爪，我控制这个假爪去作业，相当于人直接在作业，但不是用我这只手，而是用一个带着UMI的假爪的形式去作业。不同厂家的UMI设备结构会有一些差异，有一些比如千寻，它不但有头部摄像头、外部摄像头，腰部还会再绑一个摄像头去采集，为什么？因为它的本体除了头部相机，腰

部还有一个（相机）。所以这会有一些差异。另外UMI设备有的可能会增加末端夹爪的触觉传感器，可能是用他实的或者其他一个触觉传感器贴在上面，去采集一些精密装配过程当中力的曲线的信息，这是一个UMI的设备。
另外就是主从臂。因为我们前面提到的这几种方式去采集，虽然比较直接，但它和真机还没有链接上。后面说到的这几个方式是驱动了真机去做采集。

真机采集有几种方式：一是Meta眼镜去遥操机器人采集数据，通过头戴Meta的VR眼镜和手持Meta的两个摇柄去做数据采集，这是一种方式。第二种是主从臂，就是做了一个主从结构的设备，穿戴在上面，类似外骨骼一样的东西，操控末端，远处的机器人就映射过去做相应的动作。第三种是Meta眼镜加IMU加IM手套，这是遥操、机型末端是灵巧手的作业的遥操数据采集，遥操设备是Meta眼镜或者苹果的Vision Pro眼镜，末端人穿上IM手套，手套上每一个关节都会有tracker，会精准捕捉灵巧手手指的动作，这样遥操机器人去做一些精密动作时就会比较方便。但这个成本会非常高，本身设备的成本就非常高。IM手套大概要8万一副，再加上Meta的眼镜，再加上tracker这些东西。

这里面相对前面的主从臂有个问题，就是我虽然灵巧手的姿态捕捉得比较完整，但每个手臂关节怎么动，它只能追踪到腕关节到末端，比如胳膊肘是向外走还是向内走，这一块机器人是没办法适配人去做的，外骨骼可能会好一点。所以这里面也会发现遥操的困难，我想让机器人的胳膊肘那个轴向外走、末端进来，还需要对遥操设备再做一些身体改造，比如在遥操员手部胳膊肘那边、或者肩膀这一侧，加一些IMU或者传感器，来把我整个人实际的动作映射到机器人上。基本上整个遥操、数据采集设备就分这几类，前面两大类属于不需要操控真机的数据，后面这几类趋向于叫真机数据。

后面的问题是数据成本以及产业链状况。这个产业链有统计过，大概全球的产业链大概是100多亿美元，这是数据的市场。但目前我们国内基本上都是以政府级、企业级和高校级来建立这样的数据采集。从去年底到今年上半年以后，会发现数据采集的标准以及格式大家都不一样，这一块数据面临最大的问题是如何被广泛应用，这是数据的第二个事情。另外一个，前面也提到，现在数据采集大家都是在企业里面、自己搭建的场景去做数据采集，那这个搭建的场景是否能够真实反映实际的家居场景以及工业制造场景里面的状况呢？我们认为其实不能完全替代。所以未来数据采集怎么去下到工厂里面？

这次黄仁勋和德国总理在柏林聊的时候，其实提到一个观点，德国有大量的工厂，这些大量工厂就是提供AI最好的基础和基础设施。所以未来数据采集能不能走到工厂里面，这是一个行业内要去做的事情。这里面去工厂采集数据，包括安全、数据的监管，是需要整个国家和行业去制定一些政策和策略去实现的。我认为只有这样，采集第一手的、最佳真实的数据，对未来具身的发展可能会有大量的帮助。这是我对数据资产这一块的观点和看法。

参会者：那我再问一下，比如像蜜蜂（宇树）它采集的这些数据拿去对外售卖的话，这个价格是多少钱每小时？

专家：这个看数据的质量。蜜蜂其实也是UMI的一套设备去采，所以它那个采集更多是供给它本身资源去做一些它模型的训练。他们基本上不是最便宜，差不多300块钱一个小时的视频数据。

参会者：行业普遍的是多少，您可以说一个区间或者说代表厂商的情况？

专家：基本上就是200到400之间，600的基本上都是指定采集。

参会者：指定采集是特殊场景吗？

专家：对，指定的一些特殊场景采集。

参会者：那他们获得数据的成本大概是多少呢？

专家：这个就是我前面说的，因为采集数据的方式不一样、设备不一样，相应的成本是不一样的。Ego（第一视角）数据是成本最低的，但这个数据只能给训练基模去用。UMI这种

方式的话，就看UMI的成本。其实他给你这个采集数据报价，比如报了310（元）或者410（元）一小时，他已经把人工成本以及设备成本算进去了。所以大家只要看那个数据的报价就行，基本上能够看到这个场景够不够复杂：如果场景比较复杂、动作序列比较多，数据卖出去就会比较贵；场景比较简单、单个任务的动作又比较简单，数据肯定会便宜。大家这么理解就比较清晰了。

参会者：您提到Ego的那个成本是最低的，那大概是多少呢？

专家：这个都是成万小时的去贡献给别人，而且是成万小时的各种场景去贡献给别人，这个我还真不知道价格。

我不知道现在从业者有哪些厂商已经买过、或者说哪些厂商Ego数据已经达到上万小时的数据包了，这个我还不知道。

参会者：那从设备的角度来讲，整个一套设备，Ego和UMI分别要多少钱一套？

专家：Ego的话，就头部一个摄像头就结束了，然后靠算法去维持、怎么去找你的骨骼，基本上千把块钱一套，比如3000块钱一套，基本上就能做这个事情。UMI的话，头部摄像头除外，外部还有两个摄像头，那两个小摄像头又是2000多，三个摄像头是6000多，再加上UMI的手爪，现在估计便宜了，早些年前一堆手爪的话要万把块钱。你说UMI整套设备，加上数据盒子这些东西，我估计3万块钱以内吧，整套。因为数据完了以后，还有很多细化工作，比如数据的清洗标注，现在有很多厂商在做，比如腾讯在做云数据那一块，它以众包的方式去做采集，提供云平台给到数据采集的厂商以及具身机器人使用的厂商，这就是大厂下场去做一些数据平台化的工作了。

参会者：硬件这边我还想请您帮我们再插得细一点，一个UMI里面所用到的摄像头，还有力传感器、IMU这些核心硬件，大概单个的价值量需要多少？

专家：如果用奥比中光的相机的话，单个1000多，实际上都1000多；有的是用鱼眼的模组的话，可能几百块钱就够了。IMU传感器的话，一个基本上是四五十块钱就能买到一个IMU。如果配了多个IMU，就乘以它的数量。还有一个就是供电和数据盒子。

参会者：一个IMU只要四五十块钱吗？

专家：对，便宜的四五十块钱一个IMU就可以买到。

参会者：那一套设备需要多少个IMU？

专家：就看你的设计。基本上腰部、头部、腕部、两个手，如果每个关节都要固定过去的话，可移动关节的连接部分都要放一个IMU。这么理解大家就很清晰了——从定义出发，比如手腕和胳膊肘这两个都是可以转动的，那你要在小臂上面放一个IMU；你需要捕捉多少个关节的位置和状态，关节和关节之间连接的骨骼处，你都需要放一个IMU。

参会者：那数据盒子是多少钱？

专家：数据盒子看你的存储量。现在本身内存也比较贵，因为它有的是放到云端，有的是存储到本机然后再上云，本机直接上云速度会比较慢，先存储到本机、然后再上传云可能会快一点，就看你的存储量。这个可以参考一下现在内存或者存储的价格，比如125G是多少钱，或1T的内存大概是多少钱。

参会者：还有触觉传感器这边呢，福莱新材、汉威他们做的那些？

专家：触觉传感器的话，UMI上用比较多的可能就是XYZ三向力的触觉片，那个触觉片会比较贵。触觉片的话，一个手掌需要……看你接触面积大小，正常的话一个手掌需要6片，两

个手掌需要8片，然后每片传感器的价值大概1800左右吧。
参会者：每片1800？
专家：对对对，今天1800，我们有买过。
参会者：那1800的话，这个是进口的吗？还是说国产的也有那么高的价格？
专家：国产的，他实（他山）就卖这么贵。
参会者：那这块我们国产的核心玩家，性能好的排个序是什么样？
专家：现在很多，这个我也不能说他们技术谁在前面谁在后面，就分几个，一个是高校系，一个是初创系。高校系是上海交大有一个叫……叫什么？出觉/千觉的一家厂商，叫千觉（千珏）。我记得一个千珏。还有好多我一会名字想不起来，会后如果说你们需要，就按这个去搜索一下，是触觉传感器。你们可以搜索一下触觉传感器的厂商。
参会者：那摄像头这边呢，这个格局怎么样？
专家：摄像头现在就几家。最早RealSense是英特尔的，快倒闭了，现在被康耐视（Cognex）美国康耐视收购了。国内的话就是奥比中光，还有图漾，他们也在做一些标准的具身摄像头，头部的、腕部的都做。
参会者：IMU呢？
专家：IMU的厂商也很多，具身领域之前有一家叫Z（某某），我们用的，然后其他的你可以搜一下。因为这个东西就是它配套的，设备集成完了以后，它可能就是一个隐藏的零部件，我们不太关注这个，因为我们买的都是整套设备。整套设备里面还没有提供整套设备的厂商，他自己用的什么零部件、厂商供应的，那也是他自己去选了好的。
提问五：关于数据采集行业未来比拼什么
参会者：最后还想请教一下，我们这个数据的话，后面因为现在玩家也很多，后面比拼的是什么？是靠人头吗？还是怎么样？
专家：我认为靠人肯定不行，靠人的话没有任何意义。（能听到吗？）我可以听到。就是数据这个事情，我也在思考，因为其实数据等于是人去帮助机器人去理解了整个作业的流程、分解了它。行业里面有个观点，如果是人通过这样的分解去做数据，为什么不按规则去叫机器人去执行呢？这个观点其实也是对的。所以我作为从业者的角度来讲，现在在想的一个事情是：我能不能有可能不需要遥操数据，让机器人去做一些事情？我在思考这个点。所以我觉得这件事能不能从一个（新）角度，以最低的成本，让你的模型具备一些物理功能感知的能力，我觉得是非常重要的，因为现在整个机器人的作业，都是背后的工程师把他要做的任务分解了，然后去遥操它去采集，采集了这些数据，再训练它的这个模型。
所以我现在在做的一件事情，可以跟大家分享一下，就是我现在在做的一个模型，叫物理功能认知大模型。未来我跟机器人说去拿矿泉水瓶并帮我拧开，那我只要说这个话，我根本不需要工程师遥操左手抓瓶身、右手拧瓶盖，这个动作其实我们小孩子去看这个矿泉水瓶的时候，看到瓶盖上面的螺纹，我们知道那个螺纹就是用来我的手接触它产生摩擦力，然后沿着瓶盖的……瓶盖为啥是圆的，是沿着它这个圆的切线方向去旋转。所以我未来想做的事情是，把物理世界这个东西是由什么部件构成的、这些部件之间的相互依赖关系——是绕轴旋转、还是抽拉、还是推拉——让机器人大脑、就是这个模型，让机器人对于这些物理功能有认

知了以后，我就直接调机器人的原子能力。因为人的作业能力，这些作业动作其实可以重
构的，人做事情无非就是推、拉、拽、压、按、拧、旋，这些动作就是组合起来，然后保

证你去完成了一件事情。好比炒菜有颠勺，这个动作无非就是上下运动。就这些动作就是原
子能力。你对物理世界功能有认知了以后，直接调这个原子能力就行了。这个事情我做出
来以后就不需要数据了，我的目的是直接要干掉数据。
好，谢谢专家，非常谢谢。我这边没
有问题了。
主持人：好的，那我们今天的时间也差不多了，感谢专家的分享，也感谢各位参会者，会
议到此结束。
（会议结束）

--- 物理AI加速发展，数采需求持续释放20261025.pdf ---
全文总结
一、会议主题
本次电话会主题为"物理AI加速发展，数采需求持续释放"。近期物理AI领域进展密集，特斯
拉、Optimus（Optimus）进展超预期，马斯克宣布弗里蒙特已启动早期生产规划，年产约100万
台的生产线大部分设备已到位；此外，AMD收购World Labs、引入空间智能模型研发团队，
OpenAI也重启机器人业务。随着技术持续迭代、产业投入不断加码，未来一到两年物理AI有
望迎来重要突破。在这一过程中，数据采集的重要性日益凸显，随着机器人大脑模型研发加
速、任务复杂度不断提高，模型训练对真实场景下高质量交互数据的需求持续增长，有望推动数
据采集需求进一步释放，为相关产业带来新的发展机遇。会议邀请了7位分析师，围绕行业最新
进展及产业链机会分享研究观点。
二、机械组（李哲）：标的排序从机器人本体到数据采集到端侧
李哲认为本次最大的变化是OpenAI切入机器人领域，且斯坦福用OpenAI最先进的模型去训
练机器人，把中间VLA的专用模型省略掉了，这更体现了模型的泛化能力，对传统人形机器人
公司的壁垒是一个削弱，往后二三线机器人公司的壁垒会受到侵蚀。但掌握核心的行业能耗龙
头企业，或者在一些垂直领域拥有统治权的设备企业，以及转型机器人的公司，因为其独特优势
会更有利。长期看，宇树科技、上纬新材，以及捷克科技、中控技术等比较有优势，排在第一梯
队。
第二梯队是数据采集。机器人后续实用化很重要的一点是数据采集和训练，目前多种技术
路线都在并行使用、且是配合关系；从机器人使用的需求量与现在能够提供的数量对比来看差距
非常大，所以可预见的一段时间内应该是多种技术路线共同发展，这里还会逐步扩展出很多投资
机会。现在在设备上、在服务场景上有优势的包括奥比中光、奥普特，以及与机器人训练密切相
关的灵巧手零部件，如兆威机电等，属于在数据采集等部件和服务方面比较有优势的设备企业。
第三块，物理AI除机器人外还有很多实现形式，比如各种可穿戴设备，明年可能会有很多
新产品推出，明年也会是一个比较大的年份
，这些公司也比较受益；而且现在大家对于AI基
建的强度预期有所减弱，其实有利于消费电子成本端的压力减轻，也是利好。这方面比较看好
OpenAI链的相关公司，包括博众精工、智立方，以及与眼镜核心设备密切相关的苏大维格、易瑞
科技等。整体看好排序是从机器人本体到数据采集到端侧
。推荐标的：本体龙头宇树科技、垂
直领域有核心技术的捷克科技和中控技术；数据方面奥比中光、奥普特和兆威机电；端侧的苏大

维格、博众精工和智立方。

三、汽车组（完颜尚文）：板块三维共振

当前物理AI、机器人板块有明显涨幅，周五下午板块集体爆发，物理AI、机器人迎来了市场级别的关注度。从产业端看确实有相应变化支撑新一轮行情，当前板块是三维共振状态：一是特斯拉量产加速、指引超预期；二是OpenAI全新入局，打开了板块新的叙事空间；三是数据这一要素开始得到市场关注。同时，产业有公司开始真实拿到来自上游模型厂商的订单。四季度有明确的催化因素，一是特斯拉12月4日的股东大会，二是特斯拉未来一到两个月内发布的Optimus（奥斯D）编程演示视频。往后看，OpenAI目标是在2029年IPO，也在为IPO寻找新的方向、新的估值逻辑，认为它会在提升这个方向上持续加大布局，能给整个板块贡献长期催化点。

特斯拉方面：马斯克时隔一年半再次给出特斯拉Optimus能够实现百万台的时间预期，最新指引是预计最快到2027年底实现周产两万台，也就是年产百万台的目标；而2024年、2025年4月份的指引当时讲的是预计在2029或2030年，这个时间提前了2~3年，是非常乐观的指引。技术上，目前机电系统大部分已经收敛，灵巧手还是当前的攻关中心，明年4月份AI5芯片会被提升到Optimus之上。

OpenAI方面：奥特曼在今年9月初确认将自研人形机器人，并明确表示真正的挑战是大脑而非本体；OpenAI已经开放了机器人领域20多个岗位，覆盖控制系统、执行器、电机设计、本体制造以及数据系统等多个环节，加快在软硬件方面的布局，并预计在2027年推出演示样机。从9月份发布的GPT Ultra上可以看到，在没有针对具身任务进行额外训练的情况下，它直接操控机械臂完成了抓取放置等操作，让大家看到了通过大模型向具身智能迁移的潜在路径。OpenAI的实际投入有望加速机器人大脑的能力突破，其整体路线是在向特斯拉靠拢的，从模型出发去走向本体，这也是行业当前正在形成共识的方向。

数据方面：数据线发育成了整个物理AI包括机器人板块最强的一个分支，本周主要发酵的是真实数据方向。产业链跟踪发现行业对智能数据的需求是实在的，核心原因还是训练具备通用开放能力的模型需要千万小时级的真实数据，而当前全球可用的高质量真实数据只有数十万到百万小时，仍远没有到能够支撑的临界点，当前具身模型端、数据端还面临量、质和异构三重的制约。

真实数据主要分两条采集路线：有本体路线和无本体路线。特斯拉在二季度业绩会上，马斯克明确表示特斯拉首批1~2万台Optimus主要用在自己场景里面来进行训练数据的采集，并采用了自博弈的机制，也就是引入了排查区的机制，因此非常重视真机这条路线；马斯克也明确表示不会把重心放在众包数据方面。另一方面，非夕AI采用了众包模式，把任务进行分包，具备真实场景的一方可以去承接任务、进行数据上传并拿报酬；现在智元也是类似模式，蜜蜂科技也是采用分包模式，可以在其平台上申请设备来进行数据采集。不管是真机模式还是无本体模式，大家对于数据量的需求是真实存在的，并且需求量在滚动提升。

另一条大的路线是模拟仿真路线，这是英伟达一直以来在布局和强调的路线，核心原因是真实数据的采集速度非常慢、整体成本也比较高——比如当前Ego数据的成本大概在100元人民币一小时以上，真机数据的价格还要在这个基础上再翻几倍；而仿真数据通过世界模型、合成数据加上大规模仿真运行生成，它的边际成本大概只有真机采集的1%。所以它可以低成本去覆盖各种常用场景。真机数据主要是去锚定数据质量，仿真数据是能够去放大规模的前提，模型能力迭代的需求会持续推动对数据需求的放量。当前能够卡位数据采集以及增量回升的厂商有望持续受益。

总结来看，当前整个板块呈现特斯拉量产加速、OpenAI全数局（全数入局）以及数据方向三线共振的状态，Q4特斯拉有两大催化，有望持续贡献新的催化因素。推荐方向：一是数据方向（全新的方向，本周得到市场级关注度），真实数据方面建议关注有本体的标的，如福莱新

材、华宇科技，以及机器人本体厂商（代表真实数据里有本体的方向），重点关注宇树科技和申昊科技（声平台）；仿真端建议关注当前具备物理仿真引擎能力的公司，主要是与物理世界阿尔特合作的公司。二是OpenAI产业链，认为当前比较类似2022~2023年的基点，刚开始布局、格局未定，核心是看公司有没有进行抢占的布局卡位，以及产品能力和客户关系能否复用，建议关注长盈精密、领益制造和奥普特。三是本体方面，认为当前格局已经出定，应优选高确定性并且具备潜在国品能力的公司，建议关注超华传动（对应"超华传动"）、万象前潮（对应"万象前潮"）和福赛科技。

四、海外组（孔融）：通用大模型解决大脑问题

海外头部公司能看到前期的一些订单，包括特斯拉、Optimus，对今明两年产业链预期的变化。但很多投资人在路演时更关注的问题是，如果机器人大脑的问题无法解决，产业的实际突破并没有来临。而在6月份看到了OpenAI GPT 6的发布，这是对整个机器人产业非常重要的一个时刻。通过一个通用的大模型，不经过机器人的专项训练，就可以通过视觉和推理直接当大
脑用，解决了机器人原有的大脑问题。

原来大家更多通过VLA去训练专用模型来提升智能化水
平，但还是有关于数据和泛化的问题；现在看到通过通用模型的能力提升，完全可以用作机器人的大脑，通过它的泛化能力提升机器人大脑的智能化水平，这是非常重要的变化。后续可以关注fine-tuning、长程任务的规划，以及自主去构建一些仿真的能力。这对后续的机器人机会是比较重要的时刻，其中的投资机会不只限于原有的机器人公司，也包括上游的、在数据上的公司，以及原有构建整体物理环境的素材类公司。后续围绕整个产业链及产业的变化，不只局限于
头部公司下的订单，更在于通用大模型的能力升级之后对整个产业的影响。

大家关注的素
材包括机器人公司，海外的以及国内的，国内如宇树等，海外如特斯拉的Optimus，这是产业链整体的一个机会。

五、计算机组（吕伟）：看好两个方向

机器人包括物理AI最近最大的变化其实是从GPT 6 Ultra这个模型之后开始的。过去顶尖的语言模型和世界模型或者多模态模型相对来说分开了两个路径，最近走向了融合；基模通过模型有了物理AI或者世界模型的能力，而且在不断增强，机器人的大脑，也就是融合后的语
言模型和世界模型，有了一个非常明确的拐点，这是特别重要的大的逻辑。

计算机组核心推荐两个方向：一是已经有明确的物理AI或相关收入落地、估值业绩基本
面非常好的偏白马组合，包括中控技术、道通科技；二是赛道比较独特、偏黑马的两个标
的，能科科技和江苏北人。道通科技整个估值大概在10倍左右，分红等各方面在科创团队处于前列，是物理AI的世界级代表企业，在无人机、汽车充电桩终端机型这些物理世界端口完成了AI token的变现，而且已经有了明确的商业订单，可以经设备端实现百亿级的token增量收入，比较明确。中控技术是TPT工业时序模型，上半年已经有非常明确的AI收入，预计全年AI收入可能会达到10个亿甚至更高；公司25年8月发布了流程工业的时序大模型TPT，已经在石化化工能源电力的十多个行业广泛应用，在沙特阿美这个MMR项目以及其他一些工业系统、AI领域都取得非常强的落地；公司整个基本面传统业务下滑收窄，从前两个季度来看都是不断环比优化，今年也是一个基本面的大的拐点。

两个黑马标的：能科科技是工业和军工AI里面一个非常独特的标的，在工业数字孪生里有自研的系列AI产品，聚焦具身智能训练和工业智能化，25个智能体已经在70个行业落地；去年年底的具身智能AI讯飞平台项目已经成为一个战略方向，公司整个重点的基本面也相对比较不错。

江苏北人在垂直的焊接集成领域，是现在最缺人的船舶焊接领域，自身累积绩效为4个数量级（对应"4个log号"），涵盖焊接工艺原理、复杂的工业视觉轨迹、机器人的集成和客户现场工程

化的经验；它通过深度科技等外部伙伴提供基础模型、多模态理解生成的能力，形成自有数据闭环加外部通用智能的双引擎。公司分三个阶段引进：第一阶段以L1到L2建立物理化标准，做AI焊接机器人；第二阶段L3到L5去做焊接机器人群控和自主工业能力；第三阶段把工艺能力沉淀为垂直行业大模型工艺平台。

六、电新组（李佳）：板块从技术收敛切换到量产与订单

核心观点是，板块正从技术收敛到量产和订单的关键切换期间，行业也正逐步走进订单和业绩兑现的阶段。

板块信息方面，过去两周人形机器人板块催化密集出现。特斯拉从样机正式跨入量产：10月19日，特斯拉在致股东材料里第一次以书面形式宣布，第三代类人机器人已经为大规模的量产做好准备；22日的三季报电话会上，公司又进一步披露了人形机器人将在加州工厂启动生产，首批达到量产质量标准的机器人正式下线，马斯克也重申了未来要冲到每周两万台的产能目标，这是整个行业从工程样机迈入量产周期的标志性事件。海外科技巨头军备赛全面打响：OpenAI关山亲自下场自研原型机器人、重点攻关智能系统，并计划于2027年推出演示样机，大规模招募机电、执行器人才；英伟达在柏林的GTC大会上把物理AI作为核心方向，升级了自己的通用机器人开发平台。以前这些巨头是提供大脑，现在都开始亲自下场，产业链的估值锚也从硬件组装往数据与模型迁移。国内政策和产业数据双双跟上：首批人形机器人国家标准正式发布，覆盖了本体构成、系统架构等关键指标，2027年4月开始正式实施；工信部披露今年上半年我国人形机器人整机产品产量超过4万台，全年将突破10万台。由此可以看到，产业逻辑正从技术收敛切换到量产与订单，资金也明显偏向有实际订单、已经进入头部供应链的标的。

推荐核心标的：鸣志电器，是全球的空心杯电机龙头，也是灵巧手最纯正的标的，目前公司全球空心杯电机份额稳居前三，在人形机器人灵巧手的电机上先发卡位优势非常明显，位居全球第一阵营；今年上半年公司机器人应用的营收同比增长60%左右，产品已经覆盖了灵巧手、上肢关节还有移动底盘等关键环节，并与国内百余家灵巧手和机器人企业建立了合作。汇川技术，是目前国内伺服系统的绝对龙头，聚焦工业级核心零部件和工业场景的半圆形（半自研）解决方案，能够同时服务多家本地客户而不与客户形成竞争；从公司现在的无框力矩电机及相关产品的出货数据看，公司目前已经位于行业前列，像髋关节模组、肩部关节模组都已经实现了批量化供应。伟创电气，是国内工控行业的老兵，也是具身智能收入放量最猛的标的之一，目前把工业控制的底层技术往机器人执行端复用，已经覆盖了空心杯电机、无框力矩电机、轴向磁通电机、关节模组等全套运动执行器解决方案；今年上半年公司披露的具身智能业务营收同比增幅接近350%，业绩增速非常亮眼，也已与国内外多家知名企业达成合作。科达利，是国内锂电结构件的龙头，机器人第二曲线在加速兑现，目前已经形成了减速器加丝杠加灵巧手三大产品方向系统布局；从产能看，公司今年6月份合资成立了泰国科盟，共建谐波减速器基地，达产后年产值大概在7亿左右，对应百万套产能，充分体现了公司拿到订单和份额的确定性。

七、电子组（贾国瑞）：看好SOC芯片

围绕SOC芯片，机器人中SOC的作用变化和发展趋势：从架构来讲，以往机器人的SOC基本上是一个大脑小脑分裂、然后外包的方案，目前逐步往集成化、SOC一体化的方式发展，现在主流方案基本上是一颗主控SOC加上一个AI处理器，或者一颗大的异构SOC，把AI计算、实时控制都集成到同一颗芯片。第一种方案基本保留主控芯片的架构、分配一个AI能力，后者的方案基本减少了整体的集成复杂度。第二个发展趋势是应用需求的变化：从支持运动或者基础感知，逐步转向要在机器人本地运行模型、包括生成一些动作策略。从以上的架构和功能端的改变看，对于SOC芯片的性能，一是整体算力门槛在逐步提升，二是对带宽的要求以及后续支撑迭代的要求在提升。

跟踪国内主流SOC厂商的产品进度可以看到，各家厂商都在快速推出新品、产品结构逐步丰富。以瑞芯微为例，以往主力是RK3588，目前包括RK122x，以及后续会集中推出的一些新产品，不管是从整体算力、CPU性能，还是NPU性能都在逐步增长。从其他厂商看，包括地平线、星宸等

公司也都推出了一些面向机器人的新品。从这些主流厂商导入客户的进度看，首先在小脑这边，很多公司都已经实现产品的量产，部分产品已经在国内头部的一些机器人厂商实现了搭载；从大脑的情况看，国产在逐步提升，以往英伟达、包括其他一些海外厂商占据主导，目前在逐步导入国产SOC的方案。除这些场景外，在信号采集这边也看到对国产有新增的需求。整体展望，国产SOC逐步从主控中的配套，往后升级到更完整的具身智能平台这样一个角色；目前一些新品
的发布远远快于客户放量的节奏，未来国内SOC厂商，包括瑞芯微这些龙头公司，不管是主控的延续应用，还是协处理器继续分配的需求，都非常乐观。核心观点是看好电子这边SOC公司后续有望受益于人形机器人的应用。
八、人形机器人研究院（王海洋）：具身智能达到大模型训练的重要拐点
上周日组织各首席联合做了机械板块的联袂推荐，之所以找到合适的买点，是因为上周末有非常重要的短期催化，以及看到了中长期催化的重要基本面变化。
短期催化：复盘特斯拉的飞车表演以及10月22日的业绩交流会，这两件事都是整个机器人板块最重磅的。特别是业绩交流会时，从产业链跟踪知道特斯拉很久没有讲机器人的相关内容，会重点强调机器人的大脑训练，以及对于明年量的指引和预期，这个时候是非常好的埋伏整个板块的短期交易情绪。
中长期逻辑：从现在来看，整个具身智能达到了一个大模型训练的重要拐点。过去一段
时间产业调研发现，越来越多头部大模型公司都在布局甚至全员投入具身智能。典型的新闻包括AMD这样的做芯片的公司去收购了李飞飞的World Labs这样一个非常重要的具身智能公司，价格也相对来说比较贵，为什么在这个时候布局，肯定代表了一个重要信号；以及上周一整个板块最重要的黑天鹅事件，就是OpenAI继今年六七月份大量招聘相关人员之后，在上周一正式宣布进军人形机器人，2027年发布人形机器人，并且2029~2030年会把整个机器人进入一些商业化的部署，这也是一个非常重要的催化剂；以及我们看到DeepSeek战略配售宇树科技，从产业里也非常容易捕捉到DeepSeek也在做一些跟具身智能相关的重要战略布局，可能在不久的将来也会有一些跟OpenAI一样比较大的动作。这么多公司在这个时间节点卷到具身智能大模型赛道，我们认
为肯定是大模型到了突破0~1的临界点
——前面可能是一个早期的理论模型探索，到这个时候
它的底层模型，无论是世界模型、VLA还是其他模型范式，其有效性得到大幅增强。很多人在说GPT 6 Ultra这个最新的模型对整个现在很多机器人的操作有一个非常大的泛化性提升，甚至已经超越了过往几年VLA的长时间训练，它可能目前在一些精细化的动作、准确度上还有些薄弱，

因为它毕竟刚开始做；王兴兴讲明年肯定要发布GPT 7，那么7这个模型可能会瞬间击穿整个具身智能底层目前困扰很多企业的大脑难题，这是行业内一些顶流大佬的判断，我们认为很有可能会成立。
这么多头部企业同时发力，肯定也是看到了一定的曙光。做AI是有一定时效性的，提前
很多年去部署人才、大模型，技术会迭代很快，AI就是"你方唱罢我登场"；这么多头部企业花这么多钱在这个时候部署具身大模型，一定是发展到了一个快车道的阶段——算力可能会提前买芯片、做AI基建部署，但真正去招团队、做模型、做数据采集训练，那一定是临近了快速发展的时点。所以我们反复强调，为什么这么多大厂在这个时间段有密集新闻出现，它代表的意义是：虽然现在没有看到整个AI大模型的出现，但应该在未来几个月甚至半年之内，我们会看到具身智能大脑的一些泛化性操作，会有一些非常惊艳的演示展示在世人面前。包括马斯克在业绩交
流会上也重点强调，在未来一到两个月之内，我们会看到特斯拉第三代机器人纯自主操作的一个非常重要的演示，会跟现在以往看到的所有机器人的演示大不一样，至少给了一个明确的信号以及时间节点，非常值得期待。当下各家公司的布局也非常重要：整个人形机器人板块，大家一直对量和产品商业化落地节奏产生诟病，原因就是大脑能力是一个木桶的短板，那么在当下，这

个底层逻辑的补强对整个板块一定是个巨大的利好，这个利好在未来的几个月一定会不断
地发酵和升级，大家会相信整个板块真正到了商业化落地的元年。
关于AMD的收购、OpenAI的动作，包括DeepSeek的新闻，都有一些公开资料；产业里还有更多，目前苹果等美国相关的M7公司，包括苹果、谷歌等等也都在做具身的事情，国内字节跳动、小米、华为最近也都发生了一些变化。
推荐股票方面，我们跟市场的观点略微有点不一样
。过往的机器人大家可能更加关注机
械、执行器也就是执行层这一块，无论是之前所谓台万的一些公司，以及核心供应的公司；最近大家对于是否需要Tier 2外组装、到底是由T公司自己完成还是供应链完成也有不同声音，包括供应商到底是国内还是国外等等，这些都是公司自己的角色和选择，根据市场竞争都会做动态调整，当下很难做明确判断。但我们想讲的是，到这个时候真正大脑开始发展的话，全世界搞具身智能肯定不只是特斯拉一家公司，它会有非常多的公司，就跟新能源车开始井喷式的发
展是一样的。那么我们真正需要看的是一个更宽泛的概念，不是说只看T链，而是看整个甚至是物理AI这样的概念。物理AI更明确的方向，就是现在的大模型训练；大模型训练因为标题比较虚，那么有一些落实的方向，关于数据它是一条线——巧妇难为无米之炊，大模型能力的提升短期对整个数据的需求是迅速增长的。关于数据，我们在今年7月1日国联民生的策略会上就邀请了国内顶尖的麦田智能CEO赵博士做过分享；7月初时整个二级市场对数据包括对Ego还非常陌生、不太关注，短短三个月左右的时间，市场对整个板块、这个概念的认知就发生了很大变化。所以

我们相信现在大家非常热议的数据这个概念，也将会是从现在开始到今年年底非常重要的一个新的成长概念的体现。相关标的其他首席已经做了比较全面的分享和跟踪。我们这边没有非常具体的标的，但认为整个做物理AI大模型训练，像芯片这一块，比如寒武纪（消D芯微）这样做专用算力芯片的，肯定是一个更加通用的技术能力零部件，包括相关技术和一些传感器，包括奥比中光、福莱新材这种做自旋力学（视触觉）的等等，这些芯片传感器肯定还是这一波物理AI的一个中心。
会议实录：
主持人：物理AI加速发展，数据需求持续释放。近期物理AI领域进展密集，特斯拉、Optimus进展超预期，马斯克宣布弗里蒙特已启动早期生产规划，年产约100万台的生产线大部分设备已到位。此外，AMD收购World Labs、引入空间智能模型研发团队，OpenAI也重启机器人业务。随着技术持续迭代、产业投入不断加码，未来一到两年物理AI有望迎来重要突破。在这一过程中，数据采集的重要性日益凸显，随着机器人大脑模型研发加速，任务复杂度不断提高，模型训练对真实场景下高质量交互数据的需求持续增长，有望推动数据采集需求进一步释放，为相关产业带来新的发展机遇。今天我们邀请到7位分析师老师，围绕行业最新进展及产业链的机会为大家分享研究观点。首先让我们有请机械组的李哲老师。
机械组李哲：各位投资者大家晚上好，我是机械组的李哲。应该说这个周末讨论度最高的应该就是物理AI这一块了，前面的逻辑我也不细讲了，可能大家现在主要关心的是标的选择。
首先就从标的上来说，我觉得这一次最大的变化有几个方面。其中最重要的是我们看到
OpenAI切入了机器人领域
，而且前一段时间，斯坦福用了OpenAI最先进的Ultra模型去训练机
器人，其实是把中间那个VLA的专用模型省略掉了。我觉得很重要的一点在于，它其实更体现
了这个模型的泛化能力，而且对传统的人形机器人公司，它的壁垒是一个削弱。往后，可
能二三线的机器人公司的壁垒会受到侵蚀。但是一方面，掌握核心的行业能耗龙头企业，或者在一些垂直领域里面拥有统治权的设备企业，以及转型机器人的公司，因为独特的优势，我认为会更有优势一点。所以从长期来看，我觉得像宇树科技、像上纬新材，然后包括像这个领域的杰克科技和中控技术等等，我觉得是比较有优势的。所以这个是我们认为排在第一梯队的机会。

那么第二梯队，其实最近大家比较关注数据采集。确实，对于机器人如果后续实用化的话，很重要的一点是数据的采集和训练。那么现在我们看到多种技术路线其实都在并行地用，而且是一个配合的关系。从机器人使用的需求量，和我们现在实际上能够提供的数量对比来说，是有非常大的差距的。所以我们认为在可预见的这一段时间，应该是多种技术路线共同发展。那么这里，我们认为逐步还会扩展出很多的投资机会。现在我们认为在设备上，还有在服务场景上有优势的，包括像奥比中光、像奥普特；也有与机器人的训练密切相关的，可能就是关于灵巧手，那么在灵巧手的零部件上，像兆威机电，这些我们认为是在数据采集等等这些部件和服务方面比较有优势的设备企业。

那么第三个，就是刚才讲的两块其实都是机器人相关。其实最近大家讨论的是物理AI，除了机器人以外，还有很多的实现形式，比如说各种可穿戴设备等等。其实最近我们认为，这些产品在明年可能会有很多的新产品推出，明年也会是一个比较大的年份，所以对这些公司来说，我们觉得也比较受益。而且现在大家对于AI基建的强度预期有所减弱，其实是有利于消费电子的、有利于消费电子成本端的压力减轻，所以也是利好他们的。所以这个方面，我们比较看好一个是OpenAI链的相关公司，包括像博众精工、像智立方，或者与眼镜的核心设备密切相关的苏大维格和易瑞科技这几个公司。整体来说，我们的看好排序是从机器人本体到数据采集到端侧。

推荐的几个标的：一个是本体里面的龙头宇树科技，还有核心的、在垂直领域有核心技术的，像杰克科技和中控技术；第二大是数据方面，包括像奥比中光、奥普特和兆威机电；第三是端侧的苏大维格、博众精工和智立方。我来汇报的大概是这些。

主持人：感谢李哲老师的精彩分享，接下来让我们有请汽车组的完颜尚文老师。

汽车组完颜尚文：各位领导晚上好，我来汇报一下汽车组的观点。我们可以看到，上周整个物理AI包括具身板块有非常明显的涨幅，特别是周五下午开始板块集体爆发，整个物理AI、机器人这边也在中国迎来了市场级别的关注度。从当前产业端来看，确实是有相应的变化来支撑这样新一轮的行情。我们认为，当前整个物理AI板块是三维共振的状态：第一点是特斯拉的量产加速、指引超预期；第二点是OpenAI的全新入局，打开了整个板块新的叙事空间；第三是数据这一个要素开始得到了市场的关注。同时，产业有公司也开始真实地拿到来自于上游模型厂商的订单。

往后看，我们认为四季度也有明确的催化因素，主要有两块：一个是特斯拉在12月4日的股东大会，另一块是特斯拉未来一到两个月内发布的Optimus编程的演示视频。再往后去看，我们认为OpenAI的目标是在2029年IPO，OpenAI也是在为IPO去寻找新的方向、寻找新的估值逻辑，那么我们认为它会在提升这个方向上持续加大布局，也能给整个板块贡献长期的催化点。

下面我们从各维度展开跟大家汇报。

首先是特斯拉这边。马斯克时隔一年半的时间，再次给了一个特斯拉Optimus能够实现百万台的时间预期。特斯拉的最新指引是预计最快到2027年底，能够实现周产两万台，也就是年产百万台的这样一个目标。而在2024年、2025年4月份的指引，当时讲的是预计在2029或2030年，所以这个时间提前了2~3年，是一个非常乐观的指引。从技术上来看，目前机电系统大部分已经收敛了，灵巧手还是当前的攻关中心，明年4月份AI5芯片会被提升到Optimus之上。

第二部分看一下OpenAI的布局。奥特曼在今年9月初确认将自研人形机器人，并且明确表示了真正的挑战是大脑而非本体，并且OpenAI已经开放了机器人领域大概有20多个岗位，覆盖了机器人的控制系统、执行器、电机设计、本体制造以及数据系统等多个环节，也是加快了在软硬件方面的布局，并且OpenAI预计会在2027年推出演示样机。我们可以从9月份发布的GPT Ultra上面看到，它展现出来在没有针对具身任务进行额外训练的情况下，直接去操控机械臂，完成了抓取、放置等等操作的能力，也让大家看到了通过大模型向具身智能迁移的潜在路径。我们认为OpenAI的实际投入是有望加速机器人大脑的能力突破的，并且我们认为OpenAI整体的路线

是在向特斯拉去靠拢的，从这个模型出发去走向本体，这也是行业当前正在形成共识的方
向。
第二点，我们可以看到数据这一条方向，也是发育成了整个物理AI包括机器人板块最强的一个分支。本周主要发酵的是真实数据这样一个方向。当前我们对产业链的跟踪可以看到，行业对于智能数据的需求是实在的，核心原因还是因为训练具备通用开放能力的模型还是需要千万小时级的真实数据，而当前全球可用的高质量真实数据只有数十万到百万小时，仍远没有到能够支撑的临界点。当前整个具身模型端、数据端还是面临着量、质和异构三重的制约。
我们看一下当前真实数据的情况。当前真实数据主要分为两个大的采集路线，一个是有本体的路线，一个是无本体的路线。行业内不同的公司现在也在采用不同的路线，比如说特斯拉，在二季度业绩会上马斯克也明确表示了特斯拉首批的1~2万台Optimus主要用在它的场景里面来进行训练数据的采集，并且特斯拉是采用了自博弈的这样一个机制，也就是引入了排查区的机制，因此我们可以看到特斯拉非常重视真机这样一条路线；有本体的特斯拉也明确表示自己不会把重心放在众包数据方面。另一方面，我们可以看到非夕AI现在采用了众包的这样一个模式，就是通过把任务进行分包，不同的具备真实场景的一方就可以去承接任务，然后来进行数据的上传，然后可以拿报酬，这是当前它的一个方式。包括现在的智元也是这样一个类似的模式，蜜蜂科技也是采用分包的模式，可以在它的平台上去申请设备，来进行数据的采集。不同的点可以看到，不管是真实真机模式还是无本体模式，其实大家对于数据的量的需求是真实存在的，并且需
求量是在滚动提升。
关于数据，另外一条大的路线就是模拟仿真的技术路线。这块也是英伟达一直以来在布局、在强调的路线，核心原因就是因为真实数据的采集速度还是非常慢的，并且整体成本也是比较高的。比如说当前Ego数据的成本，价格大概在100元人民币一小时以上；真机数据的价格还要在这个基础上再翻几倍。而仿真数据通过世界模型、合成数据，加上大规模进行仿真数据运行生成的

方式，它的边际成本大概只有真机采集的1%，所以它可以低成本去覆盖各种常用的场景。我们认为，真机数据主要是去锚定数据质量，仿真数据是能够去放大规模的一个前提，并且模
型能力迭代的需求会持续推动对于数据需求的放量。我们认为当前能够卡位数据采集以及增量回升的厂商是有望持续受益的。
总结一下，我们认为当前整个板块是呈现特斯拉量产加速、OpenAI全数入局以及数据方
向三线共振的状态，并且Q4特斯拉这边是有两大催化，我们去看的话，可能有望持续贡献新的催化因素。标的层面，我们主要推荐大家去关注三个方向：一个是数据方向，因为也是一个全新的方向，也是在本周得到了市场级的关注度。真实数据方面，我们建议关注有本体的标的，就是像福莱新材、华宇科技；还有建议大家去关注机器人的本体厂商，因为它其实是代表了真实数据里面有本体的这样一个方向，重点大家去关注宇树科技和申昊科技。仿真端，建议大家关注当前具备物理仿真引擎能力的公司，主要是包括与物理世界阿尔特合作的公司。第二条是OpenAI产业链，我们认为当前它比较类似于2022~2023年的基点，就是刚开始布局、然后格局未定，我们认为核心点是看公司有没有进行抢占的布局卡位，以及它的产品能力和客户关系能否进行复用，建议关注长盈精密、领益制造和奥普特。第三是本体方面，我们认为当前其实格局已经出定，还是应该优选高确定性、并且具备潜在国品能力的公司，建议关注超华传动、万象前潮和福赛科技。以上是我们组里的汇报，谢谢。
主持人：感谢完颜尚文老师的精彩分享，接下来让我们有请海外组的孔融老师。
海外组孔融：各位投资人晚上好，我是国联民生海外的孔融。今天跟各位汇报一下关于机器人板块的机会。我们在今年的6月其实也在电话会议跟各位汇报过关于机器人产业的观点。我们觉得从整个海外头部的公司能够看到前期的一些订单，包括特斯拉、包括Optimus，可能对于今年和明年从产业链端的一些预期的变化。但是很多的投资人在路演的时候更关注的问题是，如果机器人的大脑的问题无法解决，那么产业的实际上的突破并没有来临。但是，我们在6月份看

到了OpenAI GPT 6的发布，它其实是比较重要的，对于整个机器人产业非常重要的一个时刻。我们最近跟很多机器人产业的朋友交流的时候，大家非常兴奋，主要还是在于通过我们看到的一个通用的大模型，不经过机器人的专项训练，可以通过视觉和推理就能够直接当大脑用。这里面其实是比较重要的，去解决了机器人原有的大脑的问题。
原来可能大家更多地通过VLA去训
练专用的模型，通过这种方式来去提升它的智能化水平，但这里面还是有关于数据和泛化的问题；但是现在我们看到通过通用模型的能力提升，完全可以用作机器人的大脑，通过它的泛化能力提升机器人大脑的智能化水平，这里其实还是非常重要的一个变化。

我们觉得在后续可以关注：一个是对于fine-tuning，以及长程任务的规划，以及自主地去
构建一些仿真的能力。这对于后续的机器人的机会，我觉得是一个比较重要的时刻。这里面的投资机会，不只是仅限于原有的这些机器人的公司，也包括上游的、包括在数据上的一些公司，包括原有构建整体物理环境的其中的一些数据类的公司。这对于后续从整体机器人的大脑的视觉的理解，再到后续的执行任务，依然是重要的。我们也能够看到，在这段时间可以把原有的机器人通过常识和任务理解，来最终去执行任务。所以我们觉得后续围绕整个产业链以及产业的变化，也不只局限于头部的公司下的一些订单，更在于通用的模型能力升级之后对于整个产业的影响，包括这里面大家关注的数据包括机器人公司，海外的、以及国内的，大家关注的包括像宇树等等，海外的可能大家关注的像特斯拉的Optimus，这是一个产业链的整体的机会。以上是我们简单的观点汇报。
主持人：感谢孔融老师的精彩分享，接下来有请计算机组吕伟老师。
计算机组吕伟：各位投资者大家好，我是国联民生计算机吕伟，给大家汇报一下我们计算机关于物理AI机器人的最新观点。从我们的角度来看，机器人包括物理AI最近最大的变化其实是从GPT 6 Ultra这个模型之后开始的。因为过去来说，顶尖的语言模型和世界模型或者多模态模型相对来说分开了两个路径，那么最近在GPT 6之后走向了融合，基模通过模型有了物理AI或者世界模型的能力，而且在不断增强。我们也看到了整个机器人的大脑，也就是说这个融合后的语言模型和世界模型，有了一个非常明确的拐点，所以这是我们认为特别重要的整个大的逻辑。
从标题上来说，物理AI我们最近核心推荐的是两个方向：一个是已经有明确的物理AI或者相
关的收入落地的、整个估值业绩基本面非常好的一个偏白马的组合，一个是中控技术，还有道通科技；另外一个就是相对来说这个赛道比较独特、偏黑马的两个标的，能科科技和江苏北人。道通科技整个估值大概在10倍左右，它的分红等各方面在科创团队处于前列。公司是物理AI的世界级代表企业，在无人机、汽车充电桩终端机型这些物理世界端口方面完成了AI token的变现，而且已经有了明确的商业订单，可以经设备端实现百亿级的token的增量收入，比较明确。中控技术是TPT的工业时序模型，已经在上半年有非常明确的AI收入，预计全年的AI收入可能会达到10个亿甚至更高；从8月份开始，公司在25年8月份已经发布了流程工业的时序大模型TPT，已经在石化化工能源电力的十多个行业广泛应用。公司在最近的沙特阿美这个MMR项目，包括其他的一些工业系统、AI领域，都取得非常强的落地。另外，公司整个基本面、传统业务的下滑收窄，从前两个季度来看都是不断地环比优化，所以公司今年也是一个基本面的大的拐点。
两个黑马标的：一个是能科科技，它本身是工业和军工AI里面一个非常独特的标的，在工
业数字孪生里面有自研的系列AI产品，聚焦具身智能训练和工业智能化，25个智能体已经在70个行业落地了；它在去年年底的具身智能AI讯飞平台项目已经成为一个战略方向，公司的整个重点

的基本面也相对来说比较不错。那么最后一个新的标的是江苏北人，公司在垂直的焊接集成领域，是现在最缺人的船舶焊接领域，自身累积绩效为4个数量级，涵盖焊接工艺原理、复杂的工业视觉轨迹、机器人的集成和客户现场工程化的经验；它通过深度科技等外部伙伴提供基础模型、多模态理解生成的能力，形成自有数据闭环加外部通用智能的双引擎。公司分三个阶段引进：第一阶段以L1到L2建立物理化标准、AI焊接机器人；第二个阶段L3到L5去做焊接机器人群控和自主工业能力；第三阶段把工艺能力沉淀为垂直行业大模型工艺平台。我们主要的看好方向和4个标的就是这些，谢谢大家。
主持人：感谢吕伟老师的精彩分享，接下来让我们有请电新组李佳老师。

电 新 组 李 佳 ：各位领导晚上好，我是国联民生电新组的李佳，代表我们组更新一下人形机器人板块的最新观点。核心观点是，板 块 正 从 技 术 收 敛 到 量 产 和 订 单 的 关 键 切 换 期 间 ，行 业
也 正 逐 步 走 进 订 单 和 业 绩 兑 现 的 阶 段
。下面我将从板块动态到我们推荐的具体标的给领导们更新一下观点。
首先跟进一下板块信息发生了什么。过去两周，能够看到人形机器人板块的催化密集出现。首先大家最关心的肯定是特斯拉从样机正式跨入量产：10月19日，特斯拉在致股东材料里面第一次以书面的形式宣布，第三代类人机器人已经为大规模的量产做好准备；22日的三季报电话会上，公司又进一步披露了人形机器人将在加州工厂启动生产，首批达到量产质量标准的机器人正式下线，马斯克也重申了未来要冲到每周两万台的产能目标，这 是 整 个 行 业 从 工 程 样 机 迈 入 量
产 周 期 的 标 志 性 事 件 。第二，海外的科技巨头军备赛也全面打响：首先OpenAI关山亲自下场自研原型机器人、重点攻关智能系统，并计划于2027年推出演示样机，大规模招募机电、执行器人才；英伟达在柏林的GTC大会上把物理AI作为核心方向，升级了自己的通用机器人开发平台。以前这些巨头是提供大脑，现在都开始亲自下场，产业链的估值锚也从硬件组装往数据与模型迁移。第三，从国内去看，政策和产业数据也是双双跟上：首批人形机器人国家标准正式发布，覆盖了本体构成、系统架构等关键指标，2027年4月开始正式实施；工信部也披露了今年上半年我国人形机器人整机产品的产量超过4万台，全年将突破10万台。由此可以看到，产 业 逻 辑 正 从 技 术 收 敛 切 换 到 量 产 与 订 单 ， 资 金 也 是 明 显 偏 向 有 实 际 订 单 、 已 经 进 入 头 部 供 应 链 的 标 的 。
接下来给领导们更新一下我们推荐的几个核心标的的具体逻辑。首先是鸣 志 电 器 ，鸣志是全球的空心杯电机龙头，也是灵巧手最纯正的标的，目前公司全球空心杯电机份额稳居前三，并且公司在人形机器人灵巧手的电机上先发卡位优势非常明显，是位居全球第一阵营的。今年上半年，公司机器人应用的营收业务同比增幅在60%左右，产品已经覆盖了灵巧手、包括上肢关节还有移动底盘等关键环节，并与国内百余家灵巧手和机器人企业都建立了合作。第二家是汇 川 技 术 ，汇川是目前国内伺服系统的绝对龙头，聚焦的是工业级的核心零部件和工业场景的半自研解

决方案，使得公司能够同时服务多家本地客户而不与客户形成竞争；从公司现在的无框力矩电机及相关产品的出货数据去看，公司目前已经位于行业前列，并且像髋关节模组、肩部关节模组都已经实现了批量化供应。第三家是伟 创 电 气 ，伟创是国内工控行业的老兵，也是具身智能收入放量最猛的标的之一，伟创目前把自己的工业控制的底层技术往机器人执行端复用，目前公司已经覆盖了包括空心杯电机、无框力矩电机、轴向磁通电机、包括像关节模组等全套运动执行器解决方案；今年上半年，公司披露的具身智能业务营收同比增幅接近350%，业绩增速非常亮眼；从公司实际上绑定的客户情况来看，也已经跟国内外多家知名企业达成了合作关系。最后一家是科 达 利 ，科达利是国内锂电结构件的龙头，并且能够看到公司的机器人第二曲线在加速兑现，目前公司已经形成了减速器加丝杠加灵巧手三大产品方向的系统布局；从产能情况来看，公司在今年6月份合资成立了泰国科盟，共建谐波减速器的基地，达产之后年产值大概在7亿左右，对应的百万套产能，这也充分体现了公司拿到订单和份额的确定性。以上是我们电新组对于人形机器人板块的最新观点，谢谢。
主 持 人 ：感谢李佳老师的精彩分享，接下来有请电子组贾国瑞老师。
电 子 组 贾 国 瑞 ：好的，各位现场投资人大家好，我是电子组的贾国瑞，今天主要汇报一下关于人形机器人电子组这边的观点，主要是围绕SOC芯片这么一个情况。首先汇报一下我们看到机器人中SOC目前的一个作用变化和发展的趋势。首先从架构来讲，以往的机器人的SOC基本上就是一个大脑小脑分裂、然后外包的方案；目前是逐步往集成化、SOC一体化的方式去发展，那么现在主流方案基本上就是一颗主控SOC加上一个AI处理器，或者说是采用一颗大的异构SOC，把整体的AI计算、实时控制都集成到同一颗芯片。那么第一种方案基本上是保留主控芯片的架构，然后分配一个AI能力；后者的方案基本上是减少了整体的集成的复杂度。那么第二个发展趋势，就是我们看到SOC应用需求的变化，就是以往是从支持运动或者基础感知这一个需求，目前是逐步

转向到要在机器人本地去运行模型，包括生成一些动作策略。那么从以上的架构和功能端的改变来看，对于SOC芯片的性能：第一个就是整体算力门槛是在逐步提升的，第二个就是对于带宽的要求，以及后续支撑一个迭代的要求，也在提升。
从我们目前国内一些主流的SOC厂商的产品进度跟踪下来可以看到，各家主控厂商都在进行新品的快速推出，包括产品结构的逐步丰富。我们举例瑞芯微这个情况来看，以往的主力是这个RK3588，目前来看，包括RK122x，包括后续会集中推出的这些新的一些产品，不管是从整体的算力，还是CPU的性能，包括NPU的性能都在逐步增长。那么从其他的一些厂商来看，包括像地平线、星宸等等这样的公司，也都推出了一些面向机器人的新品。从目前这些主流厂商导入客户的进度来看，首先我们可以看到在小脑这边，很多的公司都已经实现产品的量产，部分的产品已经

在国内头部的一些机器人厂商实现了一些搭载。那么从大脑的情况来看，国产在逐步地提升，以往就是英伟达、包括其他一些海外厂商占据主导的情况，目前在逐步地导入国产的一些SOC的方案。那么除这些场景外，在信号采集这边，我们也看到对国产有一些新增的需求。
所以整体上我们后面去展望的话，我们觉得国产的SOC逐步就是从主控中的配套，往后面
去升级到更完整的具身智能平台这么一个角色
。那么可以看到目前一些新品发布是远远要快
于客户放量这么一个节奏的，所以往未来去展望，我们觉得国内的SOC厂商，包括像瑞芯微的这些龙头公司，不管是主控的延续的一些应用，包括协处理器继续分配的这些需求，都是非常乐观的。所以整体上我们核心观点就是，电子这边还是看好SOC公司，后续是有望受益于人形机器人的应用。以上就是我们电子组关于机器人这么一个观点，谢谢各位。
主持人：感谢贾国瑞老师的精彩分享。接下来让我们有请国联民生人形机器人研究院负责人汪海洋老师。
人形机器人研究院汪海洋：尊敬的各位投资者，大家晚上好，我是国联民生人形机器人研究院的汪海洋。上周日的时候，我们也组织了各位首席联合做了一个机械板块的联袂推荐。为什么我们在上周日的时候找到了一个非常合适的买点呢？因为其实在上周末的时候，有一个非常重要的一些短期催化，以及我们看到了一些中长期催化的基本面的重要变化。
先说短期催化。我们复盘一下，就是特斯拉的飞车表演，以及10月22号的业绩交流会，其实这两个事情都是整个机器人板块最重磅的。特斯拉公司有一些比较重要的变化，特别是业绩交流会的时候，我们也是从产业链的跟踪知道它很久没有讲机器人的相关内容，会重点强调一些机器人的大脑训练，以及对于明年量的指引和预期。那么在这个时候来说，其实是一个非常好的埋伏整个板块的短期交易的情绪。
那么中长期逻辑来说，包括我们这周起持续强调、以及未来几周会不断跟各位投资者反复强调的就是，从现在来看，整个具身智能达到了一个大模型训练的重要拐点
。为什么这么说
呢？就是我们在过去的一段时间产业调研发现，越来越多的头部大模型公司都在布局，甚至全员投入整个具身智能。那么典型的几个重要的新闻，包括AMD这样一个做芯片的公司，去收购了李飞飞的World Labs这样一个非常重要的具身智能公司，而且这个价格也相对来说还是比较贵的。那么为什么在这个时候去布局，那肯定也代表了一个重要的信号。以及上周一整个板块最重要的一个黑天鹅事件，就是OpenAI继今年六七月份大量招聘相关人员之后，在上周一正式宣布他们要进军人形机器人，2027年发布人形机器人，并且2029~2030年的话，会把整个机器人进入一些商业化的部署，这也是一个非常重要的催化剂。以及我们看到DeepSeek去战略配售宇树科技，从产业里面也非常容易去捕捉到，其实是有非常大的动作的，就是DeepSeek其实也在做一些跟具身智能相关的重要战略布局，可能在不久的将来也会有一些跟OpenAI一样的比较大的动作。

那么这么多公司在这个时间节点卷到具身智能大模型这个赛道的话，我们认为肯定是这个大模型突破了一个0~1的临界点。前面的时间可能是一个早期的理论模型的探索，然后到这个时候，它的整个底层的模型，无论是世界模型、VLA还是其他的一些模型范式，它的有效性得到了一个大幅的增强。那么很多人在说它的GPT 6 Ultra那个最新的模型，就是它可以对整个现在的

很多机器人的一些机操有一个非常大的泛化性的提升，甚至已经超越了在过往几年VLA的一个长时间的训练；它可能目前还在一些精细化的动作、准确度上还有一些薄弱，因为它毕竟这个模型刚开始做。那么就像王兴兴讲的，明年肯定是要发布GPT 7的，那么7的这个模型的话，可能会瞬间击穿整个具身智能底层的一些目前困扰很多企业的大脑上的一些难题。那么这种就是行业内一些顶流大佬的判断，我们认为很有可能会是一个真实的。

那么这么多的头部企业同时在这个时候发力，相信肯定也是看到了一定的曙光，虽然它肯定有一定的挑战。我们知道搞AI是有一定的时效性的，就是我提前很多年去部署整个人才、大模型的话，那肯定它的技术会迭代得很快，都知道AI就是一个星期两个星期就你方唱罢我登场。那么这么多的头部企业花这么多的钱在这个时候部署整个具身的大模型，它一定是发展到了一个快车道的阶段。它不会说是提前很多年就去开始培育，那么算力的话，它可能会提前去买一些芯片、去做一些AI基建的部署，但是真正去招团队、去做整个模型、去做数据采集的训练，那一定是临近了这样快速发展的时点。所以我们反复强调，为什么这么多大厂在这个时间段有新闻密集地出现，它代表的意义是什么？它的意义是，我们现在虽然没有看到整个AI大模型的出现，但应该是在未来的几个月甚至半年之内，我们会看到整个具身智能大脑的一些泛化性的操作，会有一些非常惊艳的演示展示在世人面前。包括我们回过头来讲，马斯克在业绩交流会上的时候，也重点强调了在未来一到两个月之内，我们会看到特斯拉的第三代机器人纯自主操作的一个非常重要的演示，会跟现在我们所有人看到的一些以往的机器人的演示会大不一样。那么至少他也给了一个明确的信号以及时间节点，这也是非常值得期待的。也有可能就是一两个月之后，我们再次看到特斯拉三代机器人的一些重要的演示，我们会觉得这个机器人大脑能力有一个重要的突破。

那么在当下，各家公司的布局也是非常重要的。我们知道整个人形机器人板块，大家一直对量和整个产品商业化落地、To C等等的节奏产生诟病，原因就是大脑能力是一个木桶的短板，那么在当下来说，这个底层逻辑的补强对整个板块一定是个巨大的利好。这个利好在未来的几个月一定会不断地去发酵和升级，大家会相信整个板块真正到了一个商业化落地的元年。以上是我们根据上周的一些内容，在本周多花了一些时间，进一步强调了大脑模型的边际变化。

关于AMD的收购、关于OpenAI的一些动作，包括DeepSeek的一些新闻，其实都有一些公开的资料；产业里还有更多的，目前苹果等美国的一些相关的M7公司，包括苹果、包括谷歌等等，也都在做一些具身的事情；包括国内的字节跳动、包括小米、包括华为，其实有一些变化大家也都是清楚的，最近也都发生了一些变化，我们时间关系就不一一赘述了。

关于推荐的股票，其实我们跟市场的观点也略微有点不一样。我们知道过往的机器人，大家可能更加关注机械、关注整个执行器，也就是执行层、机械这一块，无论是之前的所谓台万的一些公司，以及核心供应的公司；包括最近大家对于是否需要Tier 2外组装、到底是由T公司自己来完成还是供应链完成，也有不一样的声音，包括供应商到底是国内还是国外的等等。这些其实都是一些公司自己的角色和选择，根据市场竞争，都会做一些动态调整，在当下其实我们也很难做一个明确的判断。但我们想讲的是，到这个时候真正的大脑开始发展的话，那么全世界搞具身智能肯定不只是特斯拉一家公司，它会有非常多的公司，就跟新能源车开始井喷式的发展是一样的。那么我们真正需要看的是一个更宽泛的概念，不是只看T链，而是看整个甚至是物理AI这样的一个概念。

物理AI更明确的方向，就是我们讲现在的大模型训练。大模型训练因为标题比较虚，那么有一些落实的方向，关于数据它是一条线的。就当下来说，巧妇难为无米之炊，大模型的能力的提升，肯定短期对于整个数据的需求是迅速增长的。关于整个数据，我们其实在今年的7月1号，在我们国联民生的策略会上，就邀请了国内顶尖的麦田智能的CEO赵博士给我们做过一个分享。在7月初的时候，整个二级市场对数据、包括对Ego还是非常陌生的，当时不太关注；但我们可以看到，在短短过去三个月左右的时间，市场对于整个板块、这个概念的认知就发生了很大的变化。所以我们相信，现在大家都非常热议的数据这个概念，也将会是从现在开始到今年年底非常重要

的一个新的成长概念的体现。相关标的，其他首席其实已经做了非常全面的分享，后面也会做一些比较详细的分享和跟踪。我们这边倒不是说有一些非常具体的标的，但包括整个做物理AI大模型训练，我们认为整个像芯片这一块，比如寒武纪这样做专用算力芯片的，肯定是一个更加通用的技术能力零部件，包括这个能力技术，包括一些传感器，就是包括奥比中光、福莱新材这种做视触觉的等等，我觉得这些芯片传感器的话，肯定还是这一波物理AI的一个中心。

--- 物理AI、具身智能与人形机器人20261025.pdf ---
全文总结
本次电话会由国泰海通证券机械组肖琴曦主讲，主题是物理AI、具身智能和人形机器人三个概念的辨析，三者内涵与相互关系，以及涉及的投资机会和选股思路。肖琴曦指出，这三个概念是最近一周大家反复聊到的，不少领导对它们的内涵和关系有较多询问，因此本次会议先把三个概念定义清楚。
物理AI是黄仁勋在去年GTC大会上提出并"喊出来"的概念，核心是让AI能够更好地理解物理世界，并且能够在物理世界里采取行动、执行动作
。它涉及对物理规律、真实世界的空间
关系、所作用行动物体的物理状态及其背后物理因果关系的理解，基于这些理解进行预测、决策和行动。物理AI的产业范围包含所有AI能够赋能的物理世界，范畴较大，包括自动驾驶（英伟达在自动驾驶上也有布局）、工业智能系统（各种工业自动化以及工业自动化设备），也包括人形机器人。物理AI的产业范畴比较宽。
具身智能也是去年新出的概念，核心讲的是智能体能够通过身体、传感器和执行器与环
境进行交互并完成任务。肖琴曦理解，具身智能的内涵更偏算法和"软"的部分，更多强调一个智能体怎样感知环境、理解任务，然后形成动作编排，这些基本上都是软件、算法层面和模型层面的东西。从狭义理解看，它更偏软的部分，但不仅包含人形机器人对环境的感知、任务理解和编排，也包含移动机器人、机械臂这类智能体的环境感知和行动能力提升。从产业范围看，具身智能包含人形、移动机器人、机械臂这些智能体的算法和软件能力提升。
人形机器人，大家理解更偏硬件层面，即本体，团队认为它是物理AI未来最重要、也是最强的硬件载体。物理AI的产业范畴包含了人形机器人、自动驾驶和一些智能工业系统，其中智能工业系统从产业空间上讲可能是最大的
；从人形来讲，也许到2030年以后它也会成长为
一个非常大的产业，也会是物理AI最重要的载体。目前产业发展状态下，人形机器人的软硬耦合程度还是比较高的，人形本体公司基本上也把具身智能一起做了，所以这两个功能现在都是由人形机器人厂商、创新企业承载的。三个概念的关系是：物理AI包含了具身智能和人形机器
人；具身智能更偏软件和算法层面，包括仿真平台、训练方法等；人形机器人更偏硬件。
过去3年产业一直在探索能否把大脑和小脑分开，形成大小脑分离的产业分工。英伟达、谷歌DeepMind都把自己的定位放在大脑层面，硬件自己不做，跟第三方人形机器人硬件公司合作，希望未来做出通用的人形通用模型、通用大脑。国内也有代表性公司在探索，比如蚂蚁灵波（对应上文"蚂蚁林波"）就把自己定位为做人形机器人的通用大脑，自己不做硬件，基本跟乐聚、以及北京、江苏的几家做硬件的公司合作。

最近OpenAI发布了一款新模型（对应上文"extra"），让大小脑分离的产业分工可能
性进一步增强。此前人形机器人的模型架构已基本形成混合架构或分层架构，即一个快系统、一个慢系统，最早是Figure提出来的，去年Helix 01用的是三层架构，去年推出的Helix 02版本形成了双层架构；两个月前谷歌也发布了一款新的具身智能模型，同样采取双层架构。两个月前开电话会时就讲到，谷歌采取分层架构意味着像谷歌这样的头部公司都已认可双层架

构，具身智能领域双层架构已得到产业一致认可，这次OpenAI推出的新模型进一步强化了这一共识。OpenAI在谷歌双层架构基础上再往前走了一步：谷歌把大脑和小脑的功能做了比较强的划分，大脑只做环境感知和任务、状态预测，小脑是一个VLA模型，任务的行为动作编排及执行都在小脑（VLA模型）里完成，这是谷歌双层架构及以前所有双层架构采取的大小脑分离架构；而OpenAI的新模型把小脑VLA的部分功能放到了大脑模型里去完成，没有完全替代VLA模型的功能，但把VLA里任务编排、任务轨迹编排的动作完成了。由于这一变化，大家认为未来大小脑分离、做产业化分工的可能性进一步增强。

但OpenAI这个新模型目前还是非常早期的阶段，任务完成的准确度比较低，看到的数据大概只有40%左右的准确度。如果要完全实现大小脑分离，用具身的大脑模型去赋能所有其他硬件本体、实现跨本体的泛化功能，后面还要做很多模型的训练和升级，去提高任务完成的准确度，才有可能实现商业化落地。

所以这次发布，团队认为还到不了一个可以商业化全面落地的阶段，只是让看到了这个可能性。未来最核心的工作就是提高任务完成的准确度，核心是要用更多的物理世界的数据去训练模型。

由此引出数据采集——最近一周大家讨论非常多的方向。大模型厂商未来要往具身走，也要采集真机的带有物理量的数据去训练模型，才能提升模型准确度，才有可能往商业化落地走，这也是大元模型厂商做具身模型未来要投入的方向。团队也推了长盈精密（对应上文"长云精密"）这家公司，因为其子公司给OpenAI提供了数据采集的力控方案。今年整个数据采集的进步也是比较大的，现在产业里基本已经形成了一个数据金字塔，用来训练具身模型的数据类型基本分为四类：

一是Ego（第一人称）数据，即带一个摄像头进行轨迹采集，是目前数据量占比最大的一类，采集成本相对便宜，目前1小时的价格基本在100~400块钱之间。二是UMI的夹爪数据，除夹爪外也配摄像头，除采集运动轨迹数据外还会有重力加速度数据，可能还会加一些触觉数据在上面；但目前产业里大多数夹爪都没有加触觉数据，因为力的数据、位置数据这种数据数量还远远不够，如果把触觉数据加进去，数据对齐的工作还是比较大的，所以现在触觉数据的渗透率非常低，还没有到全面加触觉数据的阶段。UMI能采集的数据类型比Ego多一些，单价也会比Ego数据略贵一点。再往上走是增肌遥操（遥操作）的数据，以及真机实采的数据。金字塔最顶尖是真机实采这类数据，典型代表是特斯拉：像低端的Ego和UMI数据它不会自己去采，而是通过第三方采购；它也有仿真的合成数据对模型进行预训练；真机实采数据方面，它今年大概会做4.5~2万台的Optimus放到它的场景里边，进行真机实采的真实数据采集。如果往下推，可能到明年，特斯拉会成为人形机器人领域真机数据积累量最多的一家本体厂商。这类数据成本比较高，现在对外价格基本会去到1000~2000块钱一个小时，目前数据量也比较少、比较稀缺。

模型厂商对这四类数据都是需要的，但会用不同数据做不同阶段的训练：Ego和UMI可能用来进行模型的预训练，真机实采这种高质量数据会用来做模型的后训练，去提升模型的鲁棒性。 因为中国人力资源成本相对比较便宜，今年有一些数据商、新成立的创业公司来做数据采集工作，更多则是本体厂商自己搭建团队来做数据采集。

团队也做了真实世界数据未来市场空间的测算。中国信通院在今年6月30号发布了具身智能训练场研究报告，对截至今年6月30日全球的真机真实世界数据量做了统计：截至今年6月30

日，全球真实世界数据大概不到１００万小时。
国内从去年到今年建了比较多的数据采集场，
尤其以地方政府主导的方式在做。按照目前中国的情况，估计今年全年（２０２６年）全球真机真
实世界的数据大概在３００万小时的水平
，今年上半年不到100万小时，下半年速度可能会加速
一点。未来各家本体厂商，至少国内前10家左右今年也都定了数据采集目标，头部基本以百万小时作为全年采集量目标，所以明年整个数据量估计可能会去到千万小时的水平
；未来也有可
能整个真实世界数据会去到1亿个小时。但真实世界数据到底需要多少个小时，行业里目前没有定论，因为现在数据量还太少，首先得先把数据采集起来、对模型进行训练，后续再来看对数据需求量的大小。第一阶段的任务就是要到千万小时水平，后面也许会往１亿小时的水平去
走。
按照1亿小时的数据量进行测算，对Ego数据、触觉数据、真机实采数据以及其他模态数据分别做了测算：如果全球数据需求量去到１亿个小时，对应Ｅｇｏ数据的理论市场空间大概是
８０~９０亿元；触觉数据的理论市场空间在４６亿元；真机实采数据的市场空间会去到接近３００
亿元——真机实采数据在各类数据里占比最大的核心原因是成本相对较高，它的小时数可能跟Ego的小时数差不多，但因为单小时价格比较高，所以整体理论市场空间比较大。其他模态，比如重力、力矩等，理论市场空间大概接近２００亿元。

会议实录：
各位投资人晚上好，我是国泰海通机械组的肖琴曦。今天晚上我们这个电话会议的主题是物理AI、具身智能和人形机器人。这三个概念可能是最近这一周大家反复聊到的一些概念，所以有领导可能对这三个概念的内涵还有各自的关系有比较多的询问。因此我们今天晚上主要围绕物理AI、具身智能和人形机器人的内涵以及它们之间的关系，包括涉及的一些投资机会，以及我们的选股思路，来给各位领导做一下汇报。
首先，我们还是得把这三个概念定义清楚，因为这三个概念其实还是蛮容易混淆的。
从物理ＡＩ这个概念来讲，实际上是黄仁勋在去年的ＧＴＣ大会上面喊出来的这么一个概
念。所以它的核心其实是为了让ＡＩ能够更好地去理解物理世界，并且能够在物理世界里边
去采取一些行动、执行一些动作。所以整个物理AI其实涉及到了对物理的一些规律，还有包括真实世界的一些空间关系，以及我们作用的行动的一些物理的物体的一些状态，然后包括它背后的一些物理的因果关系的理解，还有就是基于这些理解，它要进行一些预测和决策，还有包括一些行动。所以其实物理AI的产业范围，包含了所有的AI能够赋能的物理世界，所以这个范畴相对来讲还是比较大的，包含了自动驾驶——因为大家知道英伟达在自动驾驶上面也有一些布局；第二个是包含了工业的智能系统，各种工业自动化以及一些工业的自动化设备；然后也包括了人形机器人。所以其实整个物理ＡＩ的概念，产业的范畴是比较宽的。
第二个就是具身智能这个概念，这个概念其实也是去年新出的一个概念。实际上它的核心讲的是，智能体能够通过身体、传感器和执行器跟这个环境进行交互，并且完成任务。所以我
理解下来的话，具身智能的内涵其实更偏算法和软的部分，所以它更多的是强调一个智能体怎么样去感知环境，去理解任务，然后形成一个动作的编排，所以这些基本上都是一个软件和算法的一个层面、模型层面的一些东西。所以具身智能从狭义的理解上面来看，我认为它其实是更偏软的部分。当然这个软的部分，其实不仅仅是包含对人形机器人对这个环境的感知和任务理解和编排，它也包含了对于移动机器人，还有包括一些机械臂这类智能体，它的对环境的感知和它的行

动能力的提升，这个都属于具身智能的概念。所以从具身智能的产业范围来看，实际上它就包含了对于人形、对于移动机器人、对于机械臂这些智能体它的算法和软件能力提升的这么一个范畴，这个是具身智能的内涵。

那么第三个就是人形机器人。人形机器人可能大家理解会更偏硬件的层面，就是这个本体，它是一个我们觉得是物理AI未来最重要的、也是未来最强的一个硬件载体。刚才提到物理AI的产业范畴就包含了人形机器人、自动驾驶和一些智能的工业系统，这里边我们觉得智能工业系统从产业空间上来讲可能是最大的；那从人形来讲，也许到30年以后，它也会成长成为一个非常非常大的一个产业，也会是物理AI最重要的一个载体。所以人形大家会理解成就是更偏硬件形态一些。当然现在目前产业发展的状态来看，实际上人形的软硬耦合程度还是比较高的，人形本体公司基本上也把具身智能一起做了，所以从目前这个情况来看，软硬耦合程度比较高。具身智能和人形这两个功能，其实现在都是由人形机器人厂商、创新企业所承载的，这是人形机器人这么一个概念。

所以从这三个概念的内涵来看，它的产业范围，我们理解就是物理AI是包含了具身智能和人形机器人的；具身智能会更偏于软件和算法层面的一些东西，包括这个仿真平台、训练的方法这些；而人形机器人更偏硬件。

当然刚才我们也讲到了，其实现在人形机器人的厂商，它很多也会把具身和人形这两个业务都涵盖下来。所以我们之前在路演的时候一直在讲，人形目前这个产品形态是一个软硬耦合程度相当高的产品形态。

那实际上产业在过去的3年其实也一直在探索，有没有一种可能性，能够把我们刚才讲的具身智能和人形机器人，就是大脑和小脑分开，那么形成大小脑分离的这么一个产业分工。之前比如说像英伟达，还有包括像谷歌的DeepMind，其实他们都是把自己的定位放到了大脑的层面，然后硬件实际上他们自己是没有做的，都是跟第三方的人形机器人硬件公司去做合作的。所以他们其实一直在探索，希望能够未来能够做一个通用的人形通用模型、通用大脑。国内其实也有一些公司在做这个探索，比较有代表性的，国内的像蚂蚁灵波，他就是把自己的业务定位为做人形机器人的通用大脑，他自己是没有做硬件的，他基本上是跟像乐聚，还有北京、江苏的几家做硬件的公司在合作，所以他自己不做硬件。所以过去这几年一直在做这么一个探索。

最近我们看到，OpenAI发布了一款新的模型，这个模型发布了之后，我们就看到了这种大小脑分离的产业分工的可能性，我们觉得概率是增加了，因为他这个模型做的变化跟之前的具身模型的变化，就是它往VLA的层面其实跨了一步。以前我们讲，其实人形机器人的模型架构基本上是已经形成了混合架构或者叫做分层架构，就是一个快系统，一个慢系统。这个最早是Figure提出来的，在去年Helix 01的时候，他就是用了一个三层的架构；去年他推出的Helix 02这个版本，实际上就形成了一个双层架构。然后在两个月之前，谷歌也发布了一款新的具身智能的模型，谷歌这一款新的具身智能模型也采取了一个双层架构。所以其实在两个月之前，我们开电话会议的时候就讲到，谷歌采取了一个分层的架构，其实意味着产业里边像谷歌这样的头部公司都已经认可了这种双层的架构。我们认为在具身智能的领域，双层架构已经得到了产业的一致认可。那么这次OpenAI也推出了这个模型，其实就是更加强化了双层架构的产业的一致认可。

那么OpenAI在谷歌的双层架构上面再往前走了一步，就是谷歌还是把大脑和小脑的功能做了一个比较强的划分。大脑只去做环境的感知和任务的预测、状态预测这么一个工作，那么它的小脑其实是一个VLA的模型，任务的这个行为的动作编排以及执行，实际上都是在小脑、

就VLA这个模型里面去完成的，这个是谷歌的双层架构，包括以前所有的这些双层架构，都是采取这么一个大小脑分离的架构。那OpenAI这个模型，实际上就是把小脑的VLA的部分功能放到了大脑的模型里面去完成，它没有完全替代VLA这个模型的功能，但是它把VLA里边任务编排、任务轨迹编排的动作给完成了。所以由于它有这样的一个变化，大家会认为未来的话，也许大小脑的分离、做一个产业化的分工的可能性是进一步增强了。

那么目前来看，OpenAI这个模型还是一个非常非常早期的阶段，因为目前它这个模型的任务完成的准确度还是比较低的，我看到数据应该是只有40%左右的准确度。所以如果是要完全实现大小脑分离，它可以用具身的大脑模型去赋能所有的其他所有的硬件本体，也就是我们讲到的，实现一个跨本体的泛化的功能的话，那它后面还要做很多模型的训练和升级，去提高任务完成的准确度，才有可能去实现商业化的落地。所以其实他这次发布，我们认为还到不了一个可以商业化全面落地的阶段，只是让我们看到了这么一个可能性。那未来其实还是有很多工作要做，最为核心的一个工作就是我刚刚讲的，要提高任务完成的准确度。那么怎么提高呢？核心就是要去用更多的物理世界的数据去训练它的模型。

所以这就来到我们今天要说的第二个事情，就是关于数据采集，也是最近这一周大家讨论得非常多的。因为大模型的厂商，这种大元模型的厂商，他们未来要往具身走，他也要去采集真机的这些带有物理量的数据，去对它的模型进行训练，它才能够提升模型的准确度，才有可能往下一阶段商业化落地的阶段去走。所以这个也是未来我们觉得如果大元模型厂商要来做具身模型，未来要去投入的这个方向。所以最近这一周，大家对于数据采集也非常关注，我们也推了长盈精密这个公司，也是因为他子公司给OpenAI提供了数据采集的力控方案。

数据采集，其实今年早些时候我们在路演的时候就讲，今年整个数据采集的进步也是比较大的。现在产业里边基本上已经形成了一个数据的一个金字塔，也就是现在用来训练具身模型的数据类型基本上是分为四类数据：

一个是Ego的数据，也就是带一个摄像头来进行轨迹的采集，这个是数据量目前占比最大的一个数据类型，它相对来讲采集的成本也会比较便宜一点。现在1小时的Ego数据，每小时的价格基本上就是在100~400块钱之间的一个价格区间。

另外就是第二类数据，就是这种UMI的夹爪的数据。UMI的数据，它除了UMI这个夹爪，它也会配摄像头，所以它除了采集运动轨迹的数据之外，它还会有一些其他的，比如说重力加速度的数据，还有包括像一些可能还会加一些触觉的数据在上面。但目前来看，其实产业里边大多数的夹爪都没有加触觉的数据，因为实际上力的数据、位置的数据，这种数据的数量还远远不够，所以大多数如果把触觉的数据加进去的话，其实数据的对齐工作还是比较大的。所以现在其实触觉的数据的渗透率非常非常低，还没有到全面加触觉数据的这么一个阶段。所以这个是UMI的数据，它的数据类型能够采集到的数据类型就比Ego要多一些，那么所以它的单价也会比Ego的数据会略贵一点点。

然后再往上走的话，就是这个增肌遥操的数据，还有就是用真机实采的数据。所以基本上现在业内就形成了这么一个数据的金字塔，在金字塔的最顶尖就是真机实采的这类数据。比较有典型的代表就是特斯拉，特斯拉像低端的Ego和UMI的数据，它是不会自己去采的，它是通过向第三方采购买过来；另外它也有仿真的合成的数据来对它的模型进行预训练；那么真机实采的数据，就是他今年大概会做4.5~2万台的Optimus放到他的场景里边去，进行真机实采的真实数据的采集。所以如果是这个开始往下推的话，那可能到明年的话，特斯拉就会成为人

形机器人领域真机数据积累量最多的一家本体厂商。所以这类数据它的成本是比较高的，现在对外的价格基本上也会去到1000~2000块钱一个小时这么一个价格，那这一类的数据目前的数据量也是比较少的、也比较稀缺。所以这个是在数据的类型上面形成了这么一个数据的金字塔。

那么不同的数据，其实模型厂商也都是需要的，当然他会用不同的数据来做不同阶段的模型的训练。像Ego和UMI可能会用来进行模型的预训练，真机实采的这种高质量数据会来做模型的后训练，去提升模型的鲁棒性。

所以这个是模型的数据的类型，以及不同数据采集的价格、目前的一个成本。那因为中国人力资源成本相对来讲比较便宜，所以其实中国今年是有一些数据商、新成立的创业公司来做数据采集的工作，那更多的话就是本体厂商他们也会自己搭建团队来做数据的采集，所以是这样一个模式。

然后我们也做了一下真实世界数据未来市场空间的测算。我们也看到中国信通院在今年的6月30号发布了一个具身智能训练场研究报告，它对今年截至6月30日全球的真机真实世界数据的数据量做了一个统计。那截至今年的6月30日，全球的真实世界数据大概就是不到100万小时的数据水平。现在国内实际上从去年到今年建了比较多的数据采集场，尤其是以地方政府主导的这种方式在做。所以按照目前中国的这么一个情况，我们估计今年全年，按照这么一个资本开支的状况，今年2026年全年全球真机的真实世界的数据，我们估计是在300万小时这么一个水平，今年上半年大概不到100万小时，下半年我们认为可能速度会加速一点，那么全年预计在300万小时左右这么一个水平。那么未来，我们看到各家的本体厂商，至少国内前10家左右，今年也都定了数据采集的目标，基本上头部的都是以百万小时的全年采集量作为一个目标，所以明年我们估计整个数据量会去到千万小时的这么一个水平。那么未来，我们觉得也有可能整个真实世界的数据会去到1亿个小时。

但是到底真实世界的数据需要多少个小时，其实现在行业里面是没有定论的，因为现在数据量还太少了，所以首先大家得先把数据采集起来，然后对模型进行训练，后续再来看对于数据需求量的这个大小。所以第一阶段的任务就是要到千万小时这么一个水平，后面也许会往1亿小时这么一个水平去走。所以我们认为这个曲线，现在产业里也许说不太清楚，但至少1亿小时这个数据——我讲的是真实的这个数据，不是合成仿真数据——这个数据量我觉得未来可能是会达到的。

那所以按照1亿小时这么一个数据量来进行测算，不同的数据它未来的空间，我们对Ego的数据、触觉的数据、真机实采的数据以及其他模态的数据进行了一个测算。如果是全球的数据需求量去到1亿个小时，对应到的Ego的数据的理论市场空间大概是在80~90个亿左右；触觉的数据的理论市场空间会在46个亿；真机实采的数据的市场空间会去到接近300亿。所以真机实采的这个数据，它的空间看起来是各类数据里边占比最大的，核心还是因为它的成本相对来讲是比较高的，它的小时数可能跟Ego的小时数差不多，但是由于它的单小时的价格比较高，所以整体的理论市场空间就比较大，大概接近300个亿这么一个水平。其他的一些模态，比如说重力、力矩，这些模态的理论市场空间大概是在接近200亿的这么一个水平。整体这几类加起来，就是这样一个测算结果。

--- 机器人盛世的起点20261025.pdf ---

机器人盛世的起点
全文总结
这是长城证券机械团队首席分析师邓明娜在电话会上做的汇报，核心观点是"大脑要成熟，真机先行"——大脑成熟的前提一定是海量真机数据，海量真机数据的前提一定是海量真机本体。汇报围绕数据采集、物理AI、机器人本体三条主线展开，最后给出对应标的。
一、行情与产业背景：OpenAI的动作印证了"真机先行"的判断
9月3日，OpenAI发布了GPT-6X轴（原文）模型，之后发表论文表明，GPT-6的X模型作为机器人的基础大模型或作为机器人的大脑，能够显著提升机器人的能力。原因是OpenAI做了一个关键动作：买了100万小时的乐字节（原文"1eagle"）和英米（原文）的数据来投喂给模型，才达到这个效果。11月18日（上周日），OpenAI首席执行官奥特曼说他们要做机器人，并且说机器人长什么样其实并不重要——可以是人形，也可以是狗型，甚至其他形态；他们做机器人的真正核心目的是研发机器人大脑。OpenAI本来就是做大脑的，为什么还要做本体去研发大脑？这正是因为大脑要成熟，真机先行：大脑成熟的前提一定是海量真机数据，海量真机数据的前提一定是海量真机本体。这进一步印证了团队的观点，也因此引出了当前行业炒得很火的两个方向——真机数据对应数据采集，真机本体对应物理AI，人形机器人则可以说是物理AI的终极载体。行情的演绎顺序大概是：第一步数据采集，第二步物理AI，第三步机器人本体。
二、数据采集：必须依赖真机数据，是最底层、最稀缺的资产
很多朋友会问：不是有互联网上的视频数据、图片数据，也有世界模型产生的数据吗？为什么还要真机数据？
世界模型无法替代真机数据：世界模型确实可以产生很多虚拟数据，但很多复杂的物理方程没有精确的解析解。比如玩《黑神话》时，悟空在水里运动、做动作、发技能，水流很不逼真；看一些仿真动画片，水流也很难真实还原，因为流体力学本来就没有精确解析解。软物体挤压、捏脸时脸部的挤压形状也未必真实，因为很多分子之间相互作用；手指抓杯子时每根手指的摩擦力是多少，不知道水下（原文）时粉尘多、光传播受多少影响，也不知道。也就是说，不可能把真实世界完整1:1搬进电脑，这会产生系统性的物理偏差，即使增加无数次的虚拟数据、无数次的数据增广，也无法消除这个自带的系统性物理偏差。逻辑上就必须需要真实数据来反馈修正模型。

互联网视频数据用处不大：互联网视频、图片虽然是人拍的，但缺少机器人可以执行的动作维度（action）。比如你看视频里一个人从桌上拿水杯、上车或做某个动作，你看到了，但不知道怎么去下达动作指令。因为二维视频没有深度信息——视频里手往前伸，但具体伸了多少厘米、往上抬了多少角度、走到哪里、抬多少角度用多少力，都不知道。而且拿空杯子、满杯子、湿杯子、滑的杯子，需要的力也不同；视频里没有这些数据。看似表面光滑、以为摩擦系数很低，实际摩擦系数可能很高，用力也不一样。所以互联网视频数据能用，但用处不大，因为没有力矩、没有数据、没有位置信息，很难通过后端方法把数据用好。
结论是：虚拟世界数据和互联网生成数据都不够好，一定要（采）真机数据，真机数据是最底层、最稀缺的底层资产。数据采集设备就是生产数据的产值（原文）。
关于需要多少数据：跟踪下来，跟产业广泛交流，这个数字是海量的、且很难以量化。对比2023年GPT训练模型用的海量数据，只要确定这条路线是对的、一定要走，就可以大力往上做。GPT是靠海量文本数据大力出奇迹跑出来的，机器人大脑也需要海量数据大力出奇迹，需要持续通过海量真机数据推动模型性能不断提升。像字节（原文"童AI"）之前的模型之所以有用，是因

为拿了100万小时的数据投喂；现在跟国内某大厂（原文"国小售"）行业里也有10亿美金级的数据采集资本开支在做，典型的大模型公司正在往这方面投资。不止字节，像谷歌、Meta，国内的大语言模型公司都要加强去做，也就都需要这些数据、需要采集设备。
从数据体量看，大家收到的新闻里，全球维度今年大概是1000万小时的数据体量，到三
年后可能有百亿级别的有效数据体量，翻1000倍，并且都是有效数据
。原始数据到有效数据
是有转换率的：8小时的数据，可能能用的只有1~1.5小时，中间需要清洗，很多奇怪的动作数据也用不到。按10亿有效数据对应100亿维度原始数据
，一天采集8小时（不是24小时，因为人
要休息，而且采集对象不是专门找的数据采集员，而是工厂里正在干活的工人，采集公司给一些劳务报酬，很多是兼职），一年365天，对应的是350万台级别的销量。目前一台一体机（原文"一米一国"）的成套设备对应约一万块钱一台，那么这是350亿级的市场空间，空间非常大。如果给35亿利润（按10%净利率）、给20倍估值，就是上千亿级别的市值空间，容量非常大。
还要提醒一点：采集设备是有耗材属性的。相机天天开着容易发热、容易损坏，比如夹爪天天夹、一天夹8小时、夹一年容易坏。了解下来，目前这些一体机的生命周期在半年到一年左右，就要换新的。所以这并非一次性（二阶）需求，而是会成为耗材台数（持续复购）。结论：空间非常大，这条路是对的，设备一定要上。

从设备构成看，一体机的构成比较简单也比较集中：头上戴一个头套，上面有5个摄像头；手上右边再带一个数据采集的摄像头，合计7个摄像头；再配一个IMU（惯性测量单元）。对投资而言就比较好选票了。既然方向一定要走，就选这个方向目前放量最大的方向。了解下来，冷齐科技（原文"冷齐"，应为"隆岐"一类，此处按原文"冷齐"更接近"灵巧"疑为"零一"系，
按原文保留音译差错勘正为"凌云"极可能不符，故按原文音译"冷齐科技"处理）做采集设
备是目前这个方向量最大的公司，下游客户是蜜蜂（原文，应为"密蜂"，即某数据采集平台），蜜蜂的客户像OpenAI、谷歌都是客户，国内像字节等也都是客户。蜜蜂是国内非常顶级的数据采集平台，今年要做1000万小时的体量，量很大。冷齐科技给它做采集设备，已放量，8月、9月每个月都是1万台级别的出货量，价格1万块钱一台，每个月就是1亿元级别的收入。明年经营（订单）翻倍，明年就是24万台、24亿的订单金额，量非常大，爆发速度也非常快。所以冷齐科技适合走成数据采集方向的"中军王牌"——所谓中军王牌，就是在这个方向里量最大、业绩增量（产量、收入、利润）最大。设备细节上，核心量和价值量的维度是摄像头：一台设备7个，每个摄像头一两千块钱，价值量最大；IMU（惯性测量单元）30块钱一个。摄像头对应的是奥比中光（原文"二比中端"）和奥普特，做摄像头的第一视觉；华依科技（原文"华谊科技"）做IMU这个关键单元。
三、物理AI：数据采集的下一层，是"AI+执行器+传感器"的结合
物理AI非常火，团队的定义是：AI真正产生作用、真正作用于物理世界。之前用AI在网上、在电脑里产生生产力——让它做PPT、查资料做整理，这是OK的；但怎么让它到物理世界里产生生产力？这就是物理AI。物理AI＝AI＋执行器＋传感器，它可以是人形，也可以是其他形态，总而言之是物理AI，用于人类物理世界去解决真实问题、真正产生收益、真正有量、真正有用。
物理AI也可以反哺数据与模型。第一，不同机器人形态不一样，需要不同的数据。本体A采集的数据怎么放到本体B上用？轮式加夹爪的数据怎么用到双足人形加五指灵巧手上？因为机器人的身高、体重、形态不同，电（机）不同（线性、谐波或行星、丝杠、行星滚柱丝杠或滚柱丝杠、电机转速多少），都是问题。所以同一套数据、同一个大脑，换个身体效果可能就会变。最好的方式是不差钱就用自己生产的真机本体采集自己的真实应用数据，或者客户真实应用

场景的交互数据来训练自己的模型，最后用在自己的一线身上。物理AI在各应用场景里用的模型也可以反哺数据采集、进入传统模型，这就是大家讲的后训练。AI（原文"经济出"）对应的是海量文本数据；给机器人（清洗和效率，原文）的定义是：物理AI使用加上海量真实应用场景的数据，这就是机器人的GPT时刻。训练场和真实场景是有差距的，不可能完全模拟真实应用场景，确实需要实战，需要真实场景的交互数据来反向（训练）机器人——一边干活一边记录失败，再用积累的数据去改进模型。比如数据采集用的是经币（原文）和一对（原文），采集的数据类似于学生学习：给机器做一个正确的示范，让机器去学习；而真实应用场景的数据，是让机器人自己去做、反复做，比如做炒肉这道菜做了无数次，30分、40分、50分、60分、80分、90分，给它打分，分数高的就让它朝更高维度走，形成整套流程和结果。模型有基础模型和分层模型（垂类模型）：基础模型相当于读到高中，对世界基本了解；垂类模型相当于读大学选专业，选电子、工程等专业再去学习。这些都需要真实应用场景的数据来反哺。机器人要真正去干活，从中期产品里失败了再后训练、再让模型更好地干活。机器人也和人一样，要边干边学。

团队的判断是：谁能在应用场景真正落地，谁就领先，谁就会拉开巨大的竞争差距。

因为你掌握了后续的把控权，是裁判员，你先用了、数据你先拿到了，谁就更有优势，且这个优势非常高。这个阶段，失败也是一种资产——我就想要失败数据，因为失败数据才能让飞轮进一步运转，知道自己哪里错了，比一直埋头做更重要；飞轮转一圈，模型强一分、数据多一截。而且应用场景跑通以后是有闭环的，数据能复现；如果没有真实应用场景，后训练的数据都是一次性消耗品。所以物理AI无论对数据的作用、对增量资产收入利润，还是在系统领域里的金融定位，都是很有意义的方向。物理AI是AI价值加深的着力点（原文"解个体"），而人形机器人就是物理AI的终极应用载体。所以真正有用的，就是在物理AI里找"中军王牌"，找谁真的放量了、能达到可能的放量，能解决人的问题，飞轮的数据都是有应用品（原文）的数据，可能已经开始转了，这个壁垒会非常强。

目前物理AI方面，体量最大的是奥普特，用AI视觉替代（人工）资源（原文），用在苹果产线上。公司在物理AI这个方向已经从头做了九年时间。从2020年（原文"220"）模型加入以来，公司找的第三方开源模型，已经给到自己的模型；现在已经做了四个模型给（客户），公司自己的（全）模型是自己的。从2020年到现在做了四年多，并且还有1000万级别的有效数据资产，这个体量在它这个领域里应该是唯一的，唯一且绝对明显。壁垒的本质是时间的复利：奥普特做了九年才走到模型迭代、觉得（历经）四年才成熟，其他竞争者要走这条路，也要把这么长的时间走一遍，所以壁垒很高。体量上，今年上半年1.8个亿以上的苹果给它的收入体量，今年全年三个（季度）以上，明年是6~10个亿级别的体量，今年是换（苹果大）年。所以奥普特是目前跟踪下来体量最大的物理AI公司，被定义为物理AI方向的"中军王牌"。

四、OpenAI链：找订单最多、卡位最好的标的

欧（OpenAI）链的逻辑一样：找这个方向里谁的订单最多、谁卡位能力最好。选出来的标的是博众精工（原文"博众京东"）——OpenAI有"甜点圈"（原文，应指AI端侧智能眼镜），博众精工就是做这个设备的，今年是5~9亿的订单体量（原文"5~91"），里面还配十几个（零件）订单体量，应该是OpenAI链里跟踪到的最大订单体量的公司，相信博众精工会成为OpenAI链里的中

军王牌。也可以找卡位好的：螳螂（原文"圳可"）是对AI做一线研发的；奥普特也在OpenAI
的供应商名录里，OpenAI做机器人，机器人肯定有眼睛、有视觉，奥普特很有可能成为OpenAI的"一线眼睛"，做OpenAI的视觉；科达利也在给OpenAI合作做执行器。所以OpenAI链里，主线是博众精工，卡位好的包括奥普特、科达利。
五、特斯拉链：值得持续关注，产量维度已明确
特斯拉链（T面）关注度也比较高。特斯拉本周三（原文"这周三"）开了（发布会），表示两个月内要发布最新一代机器人——奥奇摩斯（Optimus）第三代，这也是催化。从产量角度，马斯克首次明确定量：明年年底到周度生产2万台级别，对应年化100万台级别。所以今年
也可以整理起来了。选票方向上，对于大体量资金的领导，可以关注恒立液压和科达利这一力度组合，定义为机器人里的"绝代双骄"——核心逻辑是两个千亿级公司，在应用（原文"用行"）这个事情上确定放量，并且主业也非常不错；大公司资源多、人才多、平台大，又都在努力干，为什么干不过别人？就是这个逻辑。恒立液压和科达利相信会在机器人里面有不错的产业地位。
其他公司：丝杠里面比较领先的是北特科技、恒立、永泰（原文"北科科技文泰"）、浙江荣
泰；其他（部件）里面是科达利和斯菱（原文"斯林"）；电机（原文"直播间"）里面是恒博
和福赛；像谐波（原文"西东"）一样的是恒立。这个事情今年也可以持续关注，目前也挺能放量：今年年底大概周度生产2500台，现在大概是周度1000到1500~2000级别，都比较复合生产，也很不错。
六、国产本体：本质上还是数据采集的产值，同时是物理AI的终极载体
最后是国产（本体、自主机器人）。本质上还是回到最初的观点：大脑要成熟，真机先行；大脑成熟的前提是海量真机数据，海量真机数据的前提是海量真机本体。所以真机本体
本质上可以作为数据采集的产值，同时也是物理AI的终极载体。对应的核心标的：生成能源（原文，应指某本体公司）、集成机器人（原文）、资源（原文，应指某公司）、语速
（原文，应指某公司），包括范围路线也都是可以（关注）。
总而言之，数据采集方向、物理AI方向、机器人本体方向都是值得关注的方向。投资逻
辑就是：先数据采集，再物理AI，然后机器人本体。
会议实录：
各位投资人晚上好，欢迎大晚上参会，我是长城机械的首席分析师（工程师，原文）邓明娜。我们知道机器人板块其实已经行了一年了，但我们也一直在紧紧跟踪，其实这个产业一直是有进展的。那么我们先来看看机器人这个板块发生了什么。

首先，9月3日，OpenAI发布了GPT-6X轴的模型。之后，它的（论文）也发布了论文，论文里面表明，GPT-6的X模型如果作为机器人的基础大模型，或者说作为机器人的大脑，是
能够显著提升机器人的能力的。那么这是为什么呢？其实OpenAI做了一个关键的动作，就
是买了100万小时的乐字节（原文"1eagle"）和英米（原文）的数据来投喂给模型，才达
到这个效果。那么随后，在11月18日，也就是上周日，OpenAI的首席执行官奥特曼说他们要
做机器人，并且说机器人长什么样其实并不重要——你可以是人形，也可以是狗型，甚至

可以是其他的形态。他们做机器人的真正核心目的是研发机器人大脑，这个太有意思了。我们知道OpenAI本来就是做大模型的，本来就是做大脑的，那么为什么你要去做一个机器人的本体来研发机器人大脑呢？其实关注我们的朋友可能会知道，我们一直有一个观点：大脑要成熟，真机先行。大脑成熟的前提一定是海量的真机数据，海量真机数据的前提一定是海量真机本体。OpenAI的行动也进一步印证了我们的这个观点。这背后其实就已经出来了真机数据和真机本体，也是最近行业炒得很火的方向：真机数据对应的是数据采集，不论是数据采集平台也好，还是数据采集设备也好，总而言之是真机数据；真机本体对应的就是另外一个方向——物理AI。而人形机器人可以说是物理AI的终极载体。那么我们觉得这个行情的演绎顺序是：第一步，先数据采集，你的采集平台也好、采集设备也好；第二步，物理AI；第三步，机器人本体。

关于数据采集，其实很多朋友会问我们：不是有很多数据吗？有互联网上的视频数据、图片数据，也有世界模型产生的数据，为什么（还要真机数据）呢？其实我们知道，世界模型有VR（原文）、有世界模型，世界模型确实可以产生很多虚拟数据，但是很多复杂的物理方程是没有精确的解析解的。举个例子，如果我们玩《黑神话》的话，我们拿着悟空去水里面运动也好、做动作也好、发技能也好，我们知道这个水流是很不逼真的，可以说是假的；我们再看一些仿真的动画片，会发现水流其实很难模拟。为什么呢？因为流体力学本来就没有精确的解析解。比方说你看到的水流、你折挂（原文）进的水流，或者说软物体挤压、形状扭曲，比如拍一个3D的视频拍人的脸，脸蛋的挤压形状其实并不逼真，因为很多分子之间是有相互作用的。再比方说，你的拇指、手去抓一个杯子，摩擦力是多少——每个手指头的摩擦力是多少，不知道；或者说粉尘多的时候，光折射、光线传播受本身的影响有多大，你也不知道。所以这些东西都是没有精确解析解的。也就是说，你不可能把真实世界完整地1:1搬进电脑里面去，这就产生了系统性的物理偏差。这个系统性的物理偏差，即使你增加无数次的虚拟数据、无数多的数据增广，都无法消除本身自带的系统性物理偏差。这个逻辑的存在，你就必须需要真实的真机数据，一定需要真实数据来反馈修正你的模型。所以，这就是我们核心观点的由来：大脑成熟的前提一定是海量真机数据。

当然，可能有的朋友会进一步质疑：除了世界模型的虚拟数据以外，不是还有互联网上的视频数据吗？不是还有视频、还有图片吗？这不都是人拍的吗？都是电影也好、都是人拍的吗？OK，是有用，但是互联网上的视频数据缺少机器人可以用动作维度执行的action。比方说，你看一个视频、互联网上的视频，一个人从桌上拿了个水杯，或者他上了一辆车，或者他做了某个动作，你看到了，其实你也看到了，但你不知道怎么去下达动作指令。为什么呢？因为二维视频是没有深度信息的，就是视频里面的位置和真实世界的位置不一样。比方说视频里面你的手往前伸了，你看到它往前伸了，但具体伸了多少厘米你不知道；往上抬了多少角度，你看到抬了，但你不知道；你把这个水杯拿到嘴边，走到哪里、抬多少角度、用多少力，不知道。所以网上的视频数据是二维的、没有深度信息，也不可能（直接使用）。并且，你拿空杯子、满杯子、湿的杯子、滑的杯子，需要的力也不一样，视频里面没有这个数据。比方说你看到一个杯子，表面很光滑，你以为它的摩擦系数很低，其实它的摩擦系数很高，可能它用的力也不一样。所以互联网上的视频数据能用，但用处不大，因为视频数据里没有力矩、没有数据、没有位置这

些信息，你也很难通过后端的方法把数据用好
。所以这样呢，你也需要有真机数据。那么结
论是什么？虚拟数据（原文"数据访问数据"）和互联网上的生成数据都不够好，结论就是
你一定要（采）真机数据，真机数据一定是最底层、最稀缺的底层资产。数据采集设备就
是生产数据的产值（原文）。
到这个问题维度，就有很多朋友会问：你到底需要多少数据？这个数据的量是多少小时、空
间有多少量？其实我们跟踪下来、跟产业里面广泛交流，这个数字真的是海量的，并且是很难
以量化的。其实我们对比2023年的GPT——GPT用多少数据去训练模型，也是海量。那么，其实
我们只要确定这条路线是对的、这条路一定要走，你就可以大力地往上去干。GPT怎么跑出来
的？也是海量文本数据大力出奇迹。机器人大脑其实也需要海量的数据来大力出奇迹，需
要你持续地通过海量的真机数据来推动模型性能的不断提升。
这个东西一定要走，那么数
据也一定不会少。像我们跟踪的，童AI（原文）之前的模型为什么这么有用？其实它拿了100万
小时的数据投喂，就产生了这么好的效果。现在，我们跟国内某大厂（原文"国小售"）行业里，
也有10亿美金级的数据采集资本开支在做。所以这个行业里的大头部大模型公司，典型的公司正
在往这方面去投资。其实不止它，像谷歌、Meta，像国内的大语言模型公司，都要加强去做这条
路线。那么你势必也都需要这些数据，都需要数据（采集）设备，这是海量海量的数据需求。
当然我们也知道，很多朋友还是想（知道）每一个数字。那我们就说，其实大家都可以收到
一些新闻，比方说我们今天聊的是蜜蜂（原文）的新闻里面的一些内容：全球维度，今年大概
是1000万小时的数据体量，那么到三年后可能有百亿级别的有效数据体量，翻1000倍，并
且都是有效数据。如果我们大家参加周末的电话会不少，关于数据采集的，那么我们知道大家
应该都有一个共同的结论：原始数据到有效数据是有转换率的，不是说8小时就都能用。你8

小时里面的数据，可能你能用的就是1~1.5个小时，中间你需要清洗，中间很多比如说一些基本
的数据、一些很奇怪的动作，你也用不到。所以这还有一个量级的差距。那么（按）群众（原
文）的（说法），因为这个有1亿的说法、有10亿的说法、有100亿的说法，我们就按照10亿的
有效数据对应100亿维度的原始数据
，按照一年365天、一天采集8小时来算——为什么8小时不
是24小时一直采？因为别人要休息的。并且，你并不是专门找一个人来给你专门采，而是别人正
在那个地方干活，比如打饭阿姨正在给你打饭，她就（顺便）带给你设备，你这个采集的公司就
是给她一些劳务报酬，她不是专门的数据采集员，很多是兼职的，主要是工厂里面的工人，别人
要正式干活。我们按照一天采8个小时来算，一年365天，对应的是350万台级别的销量。目前
我们知道，一台一体机（原文"一米一国"）的成套设备，对应的是1万块钱一台，那么这是350
亿级的市场空间，空间是很大的。如果变成爆发式的变大，如果我们给（净利率）1/10，就是
35亿利润，给20倍，就是上千亿级别的市值空间。所以这个空间的容量非常非常大。
我们还想再提醒一点，也是很多朋友可能忽略的一点，觉得这个是一次性（二级，原文）
的。其实这个生产设备是有耗材属性的：你现在就算有人，工人或者说别人也只是兼职来弄一
下，设备也容易发热，毕竟相机天天开着，容易坏。比方说你的夹爪天天夹，一天夹8小时，夹
一年容易坏。所以我们了解下来，目前这些一体机的生命周期是在半年到一年左右，就要换新
的。所以这并不是一次性的，而是会成为耗材台数（持续复购）。所以结论是什么？空间非常
大，并且这条路是对的，设备一定要上。
目前我们看下来，一体机里面是一些什么东西呢？比方说一个头套，头上戴一个头套，上面

有5个摄像头；然后手上配套，右手边再带一个数据采集的摄像头，合计7个摄像头；然后再配一个IMU（惯性测量单元），就是这个构成。所以比较简单，也比较集中。这样对我们投资而言，就比较好去选票了。既然这个方向是一定要走的路，那么我们就选这个方向目前放量最大的方向。目前我们了解下来，冷齐科技（原文）做采集设备，是这个方向量最大的公司，下游客
户是蜜蜂（原文），蜜蜂做本体大工（原文），蜜蜂的客户像OpenAI、像谷歌，都是客户；国内像字节等，也都是客户。所以蜜蜂是国内非常非常顶级的数据采集平台
。今年，我们看新闻
的话，也知道他们要做1000万小时的体量，这个量是很大的。冷齐科技就是给它做采集设备的，也已经放量了，8月、9月每个月都是1万台级别的出货量，价格1万块钱一台，那么每个月就是一个亿元级别的收入体量。明年的订单是翻倍的，明年就是24万台、24亿的订单金额，这个量非常大，并且爆发的速度非常快。所以，我们觉得冷齐科技适合走成数据采集方向的中军王牌。
什么叫做中军王牌？就是在这个方向里面，量是最大的——量的最大增长、产量增长、收入增长、利润（增长），这个叫做中军王牌。所以我们相信冷齐科技会走成（这个）产业方向的设备上的中军王牌。

我们再去研究这个设备、这个细节的话，前面讲到，这个东西里面核心的量和价值量的维度是摄像头：一台（设备）里面有7个摄像头，每个摄像头一两千块钱一个，所以这个价值量最大。然后IMU（惯性测量单元）是30块钱一个。那么摄像头对应的就是奥比中光（原文）和奥普特，这是做摄像头第一视觉的；华依科技（原文）是做IMU这个关键单元的。所以这是数据采集的一个分析。
除了数据采集，还有物理AI。物理AI也是数据采集的下一层。
我们知道物理AI现在非
常火，各种电话会、专家会都非常火。那么什么是物理AI？因为我们的定义是：AI真正产生作用、真正作用于物理世界。比方说你之前用AI，在网上买一个Cloud的code也好，或者你的豆包也好，或者网（原文"网巴黎"）也好，这个东西你在网上、在电脑里面产生生产力——你让它给你打个PPT、查一些资料做整理，这个是OK的。那么怎么让它到物理世界里面去产生生产力呢？这就是物理AI。那么总结就是：物理AI = AI + 执行器 + 传感器的结合，它可以不是人
形，它可以是各种形态，总而言之是物理AI，用到人类的物理世界中去解决人类的真实问
题，真正产生收益、真正有量、真正有用，这个就是物理AI。
物理AI也可以反哺我们的数
据、我们的模型。第一个点是什么呢？比方说你需要（不同）的数据。我们回到数据的维度来看物理AI：因为不同的机器人样子不一样，机器人和机器人需要不同的数据。比方说你在本体A上面去采集数据，本体A肯定是轮式加夹爪；那么本体B是双足人形加五指灵巧手。那么我用本体A采集的数据，怎么放到本体B上面去用呢？因为本来是轮式加夹爪，怎么用到双足人形和五指灵巧手上面去？这个数据是可以复用的吗？因为机器人的身高、体重、形态，比方说有什么电机、线性（执行器）是什么样（行星、谐波），是丝杠——是行星滚柱丝杠还是滚柱丝杠，电机转速多少，这些都是问题。换句话说，同一套数据、同一个大脑，但你换个身体，效果可能就会
变。那你最好是——如果你不差钱，最好的方式是用自己生产的真机本体机器人采集自己的真实应用数据，或者客户真实应用场景的交互数据来训练自己的模型，最后用在自己的一线身上，这是最好的。所以物理AI在各应用场景里面去用它（相关）的模型，也可以去反哺你的数据采集、进入传统的模型。所以这就是大家讲的后训练。那么AI对应的是什么？是海量文本数据。我们给机器人（清洗和效率，原文）的定义是：物理AI使用加上海量真实应用场景的数据，这就
是机器人的GPT时刻。当然，大脑肯定是（要）成熟，大脑成熟是前提之一。因为你的训练场

和你的真实场景其实是有差距的，你不可能完全模拟你的真实应用场景，你确实需要去实战，需要真实场景的交互数据来反向（训练）你的机器人——你需要一边干活一边记录失败，再用积累的数据去改进你的模型。打个比方来说，你前面数据采集用的是经币（原文）和一对（原文），那么你采集的数据就类似于你的学（习）——比如释放（做）一道菜一定要学，给机器做一个正确的示范，让机器去学习。那么真实应用场景的数据，就是你真实用起来是怎么一回事：你让这个机器人自己去做，比方说让它去做炒肉，就让它去做，它做了无数次，比方说30分、40分、50分、60分、80分、90分，你给它打分；那么分数高的，你就让它朝着更高的维度去走，它的整套流程以及这个结果。并且，这个模型有基础模型和分层模型，或者说相应的垂类模型。其实大家可以这么理解：基础模型，如果你对比人来说，就相当于你读了一个高中，你读到高中了，
你对这个世界也基本上了解了，这个叫基础模型；那么这个垂类模型呢，比方说你要去做
一个具体的事情，就相当于你读大学了，你选一个专业，电子专业也好、工程也好，你选
一个专业，然后你再去学习，这是实际里的学习模式。
这些东西都需要真实应用场景的数据
来反哺。也就是说，机器人要真正去干活，从中期产品里面失败了，再到后训练，再让这个模型更好地干活。所以机器人和人一样，也要边干边学。那么，我们目前的观点是这样的：谁能
够在应用场景真正落地，谁就领先，谁就会拉开巨大的竞争差距。
为什么呢？因为你核心
掌握了后续的把控权，这个裁判员是你；你先用了，数据你先拿到了，那么谁就更有优势，并且这个优势非常高。那么在这个阶段，其实失败也是一种资产，因为我就是想要你的失败数据。为什么呢？因为你的失败数据才能够让我的飞轮进一步去运转——我知道自己哪里错了，比我一直埋头做更重要；我飞轮运转一圈，我的模型强一分，我的数据多一截。并且你在应用场景跑通以后，你是有闭环的：你的数据是能复现的。如果没有这个真实应用场景，那么你的后训练的数据都是一次性消耗品。所以，物理AI无论是数据上面的作用也好，还是增量资产的收入利润
也好，还是在这个系统领域里面的金融定位也好，都是很有意义的一个方向。
物理AI是AI
价值加深的着力点（原文"解个体"），那么人形机器人也是一样，也是物理AI的终极应用载体。所以真正有用、真正解决人类的问题的，我们还是（要在）物理AI里面找中军王牌，找谁真的放量了、真的达到了可能的放量，能解决人的问题，飞轮的数据都是有应用品（原文）的数据，可能已经开始转了，这个壁垒会非常非常强。
目前物理AI方面，体量最大的是奥普特，是用在苹果上面、用AI视觉去替代（人工）资源（原文），因为公司在这个方面已经从头开始做了九年时间。2020年（原文"220"）模型加入以来，公司找的第三方开源模型，然后已经给到自己的模型。相当于现在，今天已经做了四个模型给客户，就是公司自己的模型；当然（基础）模型是第三方的，但这个（全）模型是自己的。从2020年到现在做了四年多，并且还有1000万级别的有效数据资产
，这个体量在它这个（领域）
里应该是唯一的，唯一且绝对明显。我们知道壁垒的本质是什么？壁垒的本质是时间的复利。奥普特做了九年才（走）到模型迭代，历经四年才成熟；那么其他的经营者要走这条路的话，我走了这么长时间，你一定要走一遍，所以这个壁垒很高。那么从体量上看，今年上半年1.8个亿以上的苹果给它的收入体量，今年全年三个（季度）以上，明年是6~10个亿级别的体量，今年是换（苹果）大年。所以奥普特是我们目前跟下来体量最大的物理AI公司，我们把
它定义成为物理AI方向的中军王牌。

那么，除了数据采集和物理AI，我们再讲讲最近大家非常非常感兴趣的方向——欧（ＯｐｅｎＡＩ）链。OpenAI链呢，我们的逻辑还是一样的：去找这个方向里面谁的订单最多、谁的卡位能力最好，这样选出来我们的标的。那么，博众精工（原文"博众京东"）——大家知道OpenAI有那个甜点圈（原文，应指AI端侧智能眼镜），博众精工就是做那个甜点圈（设备）的。今年是5~9亿的订单体量（原文"5~91"），里面还配十几个（零件）的订单体量。这应该是OpenAI链里面我们跟踪到的最大订单体量的公司。所以我们相信博众精工会成为OpenAI链里面的中军王牌。那么也可以去找卡位好的：我们知道螳螂（原文"圳可"）是对AI做一线研发的，奥普特也在ＯｐｅｎＡＩ的供应商名录里面。我们知道OpenAI是做机器人，那么机器人肯定是要有一个眼睛的，肯定是有视觉的，那么目前在OpenAI里面的就是奥普特，奥普特很有可能会成为OpenAI的一线的眼睛，做OpenAI的视觉。那么像科达利也在给OpenAI去合作这个执行器。所以这是OpenAI链：订单最多的是博众精工，卡位好的像奥普特、科达利。

那么，当然还有关注度也算是比较高的，就是T面（特斯拉链），我们的老铁（原文"老贴面"）。我们知道T面这块是很多很多的。比方说特斯拉这周三（原文"这座三"）、这个周末时间，也开了这个（发布会），说是两个月内要发布最新一代机器人——奥奇摩斯第三代，这也是催化。那么我相信，特斯拉链也是一个非常值得关注的方向。从产量的角度，马斯克也明确说了，首次明确说了这么大的体量：明年年底到周度生产２万台级别，这是什么概念？一年５０００多（周）年化，就是１００万台级别。所以１００万台级别明年年底出来了。所以今年，我觉得也可以整理起来了。那么从选票的方向来说，对于大（体量）资金、大体量的领导，我们觉得你可以去关注一下恒力液压（原文"恒力液压"）和科达利这个力度组合，我们把它定义成机器人里面的绝代双骄。绝代双骄核心是什么概念呢？就是两个千亿（市值）公司，在应用这个事情上确定放量了，并且主业也非常不错，在主业上面也在做很多事情。那么，（逻辑是）我是千亿（百亿，原文）公司，你百亿公司，我们俩都在努力干；我的资源、我应该比你更有钱，我应该人才比你更多、平台比你更大、资源比你更多，那么我也在努力干，我为什么干不过你呢？核心就是这个逻辑。所以我们相信恒力液压和科达利一定会在机器人里面会有不错的（增长）、会有不错的产业地位。那么其他公司呢，比方说像丝杠里面，比较领先的是北特科技、恒立、永泰（原文"北科科技文泰"）、浙江荣泰；其他（部件）里面，就是科达利和斯菱（原文"斯林"）；电机（原文"直播间"）里面，就是恒博和福赛；像谐波（原文"西东"）一样的就是恒立。所以这个事情，我们觉得今年也是可以持续关注的，并且目前也是挺能放量的：在今年年底的话，大概是在周度生产2500台，现在的话应该是周度1000到1500~2000级别，都比较复合（产能）生产，所以也是很不错的。

那么最后就是国产的——我们讲的是国产的本体、国产的机器人（原文"经济"）。那么本质上这是什么东西呢？本质上我们还是回到我们最初的观点：大脑要成熟，真机先行。大脑成熟的前提是海量真机数据，海量真机数据的前提是海量真机本体。所以真机本体本质上就可以作为数据采集的产值，同时也是物理AI的终极载体。那么对应的核心标的：生成能源（原文）、集成机器人（原文）、资源（原文）、语速（原文），包括范围路线也都是可以（关注）。这是国产（方向）。

所以总而言之，我们觉得数据采集方向、物理ＡＩ方向、机器人本体方向，都是我们值得

关注的方向。那么投资逻辑就是：先数据采集，再物理AI，然后机器人本体。以上就是我
们长城机械给各位领导的汇报，如果有关注的问题，也欢迎联系我们，我们后面再做一个详细的汇报和分享。谢谢大家。

--- 具身模型-OpenAI入局，未来格局将如何演绎20261025.pdf ---
全文总结
一、会议背景与主题
本次电话会是系列电话会的第一章节，主题是具身模型。本周起每天晚间9点半将介绍物理AI的不同环节：今天讲模型，明天讲数据，后续讲传感器、灵巧手加双臂，核心是围绕整个行业最新的变化和增量环节。团队认为，整个行业最新迎来了比较大的变化，此前二级市场比较平淡，从上周开始市场终于注意到，整个行业在过去一年有比较大的进展，主要围绕模型层面和数据层面，这两者的变化可能比供应链层面和硬件层面更大
。团队此前梳理过机器人行业的
四大环节：模型、数据、偏硬件、应用场景，其中硬件相对比较成熟，不是行业短板，最大
的短板其实是模型，而限制模型发展的就是数据
，所以这几点是最值得关注边际变化的环节
——如果模型环节的突破往前进一步，就代表行业最大的瓶颈被解开一步，整个行业就能再往前迈进一步。
二、行业最大变化：OpenAI入局
最近行业最大的变化就是模型层面，从今年6月份开始，OpenAI正式开始布局机器人，并公开招聘硬件人才；在今年9月的GPT 6S（对应上文"GP 76S"）当中已经加入了几十万小时的机器人数据，在测评当中已经展现出非常强的自身控制能力，在很多包括Robot Dojo等常规测评中已经得到了很高的分数，超越了全球其他的开源模型。短短几个月时间，大语言模型公司转型
去做机器人模型，就已经能呈现出这么强大的能力。
随着它未来从大语言模型、相当于PS2
环节、偏大脑的模型层级，往偏本体S1和运控的层级去做，并且加入很多软硬件协调的东西，它很有可能在未来模型竞争格局中担当非常重要的地位，去引领整个行业的技术迭代与技术收敛。
行业一直等待的时间就是机器人的GPTmoment
，也就是当模型收敛能够应用于很多机器人本
体、让机器人本体能够在真实场景中发挥应用能力的时刻，OpenAI的入局加速了这个时刻的到来。
三、具身模型的发展阶段与现状
机器人模型分成两个时代：AI之前更多是传统控制，偏向数学、动力学的控制方式，机器人需要时时刻刻计算自己身体的重心、下一步的输出轨迹和落脚点，非常机械式、动作缓慢，也没有大数据模型的应用。大模型出来之后，行业经历两个阶段：第一阶段是模型公司直接用大模型去控制机器人，但控制能力非常差，很难跟物理世界及硬件结合，效果很差，只能做一步动作；随后多模态模型发展，多模态模型与机器人本体结合开始应用，目前市面上大部分模型都处于
这个阶段，包括谷歌、OpenAI加Pi 0.5、国内的语数智源等，都是上面接一个多模态模型、再接一个机器人本体模型构成架构。这一步被列为阶段四，下一步可能会有新的事件模型或其他模

型架构，大家期待有更有效率的新架构提升机器人的学习能力，因为过去大模型控制方式效率偏低、需要很大算力、很多数据、很长时间思考，但阶段四还没发展到终点，现在大家还是沿着多模态模型这条主线往前演进。
四、全球玩家格局
整体看，具身智能领域全球主流玩家的构成相对偏大厂、大模型公司或云厂商，创业型公司

占比较小。北美代表性公司：谷歌做了很多年具身智能（早期RT系列到现在的Gemini加机器人系列）；英伟达一直不遗余力推进机器人模型，希望做机器人基座、赋能下游机械公司，推出GR00T系列（从N1迭代到N1.7），还有Cosmos模型用于生成机器人数据，以及专门用于机器人的世界模型等，是偏基础设施的模型公司，开源给更多机器人公司使用；特斯拉技术方案保密，但模型应该非常强；OpenAI是最近最大的变化
，模型当中未来会加更多机器人控制，下一步也会
做机器人硬件；还有Physical Intelligence（对应上文"派"），是具有代表性的机器人模型创业公司，模型处在第一梯队、非常领先，最新迭代到Pi 0.7，在模型架构上非常具有创新性和引领性。
中国分两类：一是偏大厂，如小米、阿里、腾讯等，大厂不会错过未来这么大的产业趋势，都认为这是非常明确的方向、未来有很大的想象空间，但现阶段硬件技术、算法技术还没有收敛，所以没有把它当成商业化项目去做，而是当成偏预研的项目和研究部门，基本都会有一个robots lab做早期机器人数据采集、硬件预研和模型开发，官网也能看到一些公开发表的机器人模型，做得还比较不错。二是具身智能创业公司，代表是智源、星海图、银河通用、天行智能这几家，知名度最高，估值基本都在上百亿到几百亿——智源最新已经融到七百亿，银河通用也差不多三四百亿，星海图也是两三百亿，基本都是两百亿往上，在所有的具身公司当中估值排在第一梯队；创始人大部分来自自动驾驶行业、大厂或高校研究机器人相关的人才，主要聚焦在机器人大脑以及与硬件结合的环节，是国内做得最优秀的创业公司之一。整体看，模型的玩家或竞争格局跟本体还是有比较大的区别，还是有很多大厂来做这件事情
，因为这是更加前沿性的
方向。
五、代表性公司的模型架构
Physical Intelligence的模型基本是开源的，有很详细的记录文章，特别是在Pi 0.5之前权重都是开源的，可以调度它的模型放在自己的真机上做测试，这是比较少见的会把自己的权重完全开源出来的公司；星海图也做了开源的工作，信息量最大、最值得研究。它从2024年开始做模型，第一代是Pi 0模型，给行业树立了标杆，成为很多行业的主流方向；后续两年基本每隔几个月迭代一款，到Pi 0.5增加了陌生环境，Pi 0.6增加了强化学习，Pi 0.7增加了世界模型和很多新模块，提升泛化能力。它在北美地位相对靠前，也是百亿美元估值的公司。

最基础的是Pi0结构，是一个VLA结构，大脑部分是一个分层架构，相当于大脑一个
S2、小脑一个S1，就是一个快慢脑——大脑是慢脑，小脑是快脑。
首先通过大脑做环境识
别感知、任务推理，再把下一步任务输出给小脑，小脑再通过函数的计算最后输出一系列动作。展现出来的形态就是机器人面对世界的变化，比如给它一个动作指令"帮我倒一杯水"，它结合视觉用机械臂把桌上的水拿来递给你。详细看，大脑是一个偏开源的模型，加了一个视觉网络、再加一个Gemma（谷歌开源的模型，2.6B参数），小脑部分基于本体当前状态，再用流匹配的架构——通过噪声分布去噪，完成动作去噪，再输出一系列动作，即接下来一段时间要输出的动作序列。当时Pi 0模型大概有三种异构本体的控制方式，包括7自由度的机械臂、18自由度的移动机械臂、14自由度的双臂；训练数据是混合的预训练数据，包括Pi自己采集的数据集、互联网开源数据等。
Pi0.7的主要架构跟Pi0差不多，中间还是VLA，基于一个VLM，更为详细：左边是预训练过的VLM模型，有更大参数的5个开源模型来做推理跟计算，右边部分还是expert加上流匹配架构的模型，通过去噪输出动作序列，两个模型的参数都有增加。其次这个模型比之前丰富很多，加了很多新东西，包括observation memory（历史记忆）；动作指令方面，比如"Clean the kitchen"（清扫厨房），之前可能是简单任务、单任务，而清洗厨房是一个非常复杂的任务、比较抽象、里面包含很多子任务，接下来通过high level policy把组合式任务拆解出来，生成子任务集，输出下一步要做的任务；此外还有世界模型，相当于在脑海当中预测做了下一步动作会导致什么样的结果，如果想象中是坏的结果，它就会规避这个动作，再去生成一个更加正确的动

作。同时模型架构当中还加了一个叫meta data的打分系统，对输出的动作、对数据做打分，得分比较高的就会输出，可以提升模型输出的准确率，也可以提升数据的使用效率。这是Pi 0.7相比Pi 0比较大的边际变化。由此可见，VLM、VLA模型可以跟世界模型做结合，双方并不是完全不交叉的关系，而是可以在具体的模型架构上做结合，只要能把模型能力、准确率提升，其实模型没有那么严格的区分。

Pi 0.7还有一个变化是数据更丰富，加入了示范数据、自主数据、互联网数据和第一人称视角数据。第一人称视角从这个模型的变化当中可以看得出来，它确实是2026年最大的增量、最多的来源方向，在之前2024年的模型当中比较少见，但从2025年底到2026年，可以看到越来越多的模型当中加入了第一人称视角数据，也因为这类数据比较低成本、可获得。

谷歌一直基于VLA架构做模型，最新从之前纯端到端、从输入到输出的模型架构，也变成了快慢脑的架构，慢脑是自己的Gemini，快脑也设计了一个机器人的小模型；最早是RT-2，最新就是Gemini 1.5，今年又迭代了一把，但跟1.5整体都比较像，也是快脑加慢脑加快脑的架构，可以应用在机器人本体上，有一些很复杂的能力输出，而且可以实现跨本体的迁移，所以谷歌也是未来的主要玩家之一。

英伟达推出很多模型，首先有GR00T N1，专门针对机器人的VLA模型，从架构上看也是分层架构、快慢脑的结构，S2就是一个VLM大脑，S1是一个扩散（Diffusion Transformer）的动作模型，基本上跟前面的模型架构非常相似；N1.7也是类似的升级，加入了一些Cosmos模型的元素，整体复杂程度有所提升，架构可能跟前面类似，同时训练当中加入了更多数据，加入了Ego等数据集，有2万多个小时的第一人称视角数据去覆盖很多场景，也就是模型更复杂，Ego数据变成了很核心的增量。英伟达除了GR00T之外，也有世界模型系列。世界模型跟VLA有一个区别：VLA模型更像大模型，是一个输入到输出、输出下一步token、把token直接形成动作，所以容易犯错，面对新场景时可能没有那么好的表现，因为它没有自我学习的能力，更多是基于自己学到的东西直接做输出；而世界模型本质上偏视频生成，去对未来做预测，预测下一步的环境、下一步的动作，先生成一段虚拟的动作序列，再把它转换成实际的动作，所以它可以对未来做想象，在输出时也可以规避掉一些错误的结果，输出更准确的动作序列，面对开放的环境或新环境也有更好的学习和适应能力。英伟达在世界模型方面有两个：一个是Cosmos，不是专门针对机器人的，可以处理视频、语言、动作、音频等，是全方位世界模型；另一个是基于Cosmos之上、专门针对机器人的世界模型。从特点看，它的模型参数量比VLA更大，泛化能力更强，有更好的跨本体迁移能力，也有更好的小样本、零样本泛化能力，一些新任务也可以完成且有不错表现；缺点是参数比较大、动作更慢、算力消耗更大，相当于要无时无刻生成新的视频来做动作输出。所以这两个模型各有优劣。

阿里也是偏VLA架构，基于自己通义千问模型的VLA架构；字节也是一样，现在主流还是VLA架构，但它在其他业务上也有世界模型，可以展望在机器人上未来也会有世界模型。智源有自己的G0、G1、G2，都基于VLM模型，加了一个plan planner（规划器）在中间做衔接，类似一个中间层，再去到动作模型实现输出。银河通用偏世界模型，最近发布的是VAM，是视觉动作的模型，跟其他公司不太一样，数据上也包含更多仿真数据，路线确实跟其他公司不一样。星海图更多也是VLA结构，基于通用大模型加一个动作。腾讯是世界模型跟VLA两条路线目前都在做。海外还有语数，也是两条路线都做，理念是"我不做选择"，就follow行业有什么先做布局，既有VLA、世界模型，也有世界模型；它整体做的时间比较晚，大概从去年年底到今年年初才陆陆续续开始做，所以能力上相比行业头部在时间上要慢个一年，上市之后再追赶、努力补这个方

向。整体看，大模型公司开源确实对于偏硬件能力特别突出的公司来说会更利好。

美国GeneralistAI（对应上文"GeneralAR"）
也是有模型工厂的公司，但模型架构比
较保密，公开信息里没有找到；它在很多现实生活任务上的得分，在数据增加的情况下有很大增强——一是模型在迭代，二是数据量基本每次翻倍增长来提升能力。国内像智谱，在最新论文当中也加了数据。

六、技术路线与测评
从路线上讲，现在主要有三类：第一是分层架构，即一个VLM再加一个动作模型，是现
在最主流的架构，像Pi、阿里、腾讯、语数等都是这个路线，可以发挥VLM的能力上限，用大脑做任务推理，再用动作模型生成与硬件衔接的输出，但有自己的缺陷，现在还处在发展状态中，也是绝大多数资源投入的方向。第二是完全一体化的自回归模型，把所有信息放在一个平面
里跑回归，这个路线基本已被证实是相对比较淘汰的方向，因为它带来的信息缺失太大，对世界的感知和推理能力比较弱。第三是世界模型，对未来进行预测、输出下一步动作，做发散
式的模型，现在玩家还比较少，除银河通用之外，今年有很多创业公司在做，美国能数出几十家、中国也能数出几十家，每家公司都以非常夸张的估值融资。从技术看，这么多新公司出现说明它是一个很新的方向，里面有很多idea和创意；世界模型其实没有一个统一的架构，每家公司的架构都不一样，它是一个很笼统的名称和代称，只要有想法、创意和相应的学术成果，都能往这个方向转换，所以这个方向更偏前沿，可能还是要先把第一条路往前走一走，再看世界模型未来有没有更好的突破去颠覆。现在模型架构没有办法得出完全结论，大家都还是以VLA这条
路线往前迈进。
行业还存在一个问题是测评没有那么透明，这块也有很多榜单在提升测评。一般从四个方向测评：推理的泛化能力、数据的使用效率、任务的执行质量和推理能力，但更多考虑的相对比较显性的是第一个——泛化能力，即考察模型在未见过的场景当中的任务成功率，这是现在对
模型最核心的要求，大部分测评也是针对这一点设计。行业中有主流测评，包括仿真基准、真机实采、行业标准和赛事测评，最近比较火的是Robot Dojo，测评OpenAI模型的得分、引领其他公司，是相对比较权威的来源；另一个比较权威的来源是behavior challenge，是李飞飞在学术当中设立的标准，可能是最严苛的考验，有100多个任务、分布在几十个场景当中，挑战最大，也是行业更多参考的数据维度。从以上梳理可以看到，整个行业还是相对比较偏早期和发散
式，整个架构是没有收敛的状态，这种时候确实需要有一些新的变化给行业指引方向。

七、OpenAI入局的意义与影响
OpenAI在2020年之前其实就已经做过机器人，后面发现机器人数据积累太慢、效果太差，因此转向了大语言模型；时隔6年多重新开启机器人业务，一方面说明模型经过这几年发展到了很高的高度，到了真正可以做机器人业务的阶段；另一方面机器人行业过去几年发展非常快，2022

年全球没有几家机器人公司、也没有几台机器人本体销售，到今年全球每年都有10万台的机器人销售，成本也从当时一台几百万降到最便宜的机械臂可能两三万、语数的可能10万，成本降了一大步，整个行业做数据采集的成本、壁垒和生态都截然不同。所以OpenAI在这个时候重新做机器人，是因为它认为行业到了一个临界点，很多关键要素都已经具备，它来做有望给行业带来往前迈进的一大步。现在的VLA模型三层架构是S2、S1和harness，GPT6Ultra主要做的就是S2这
个部分，即VLM部分，它的能力本身能够提升整个机械的控制；OpenAI未来要做的不只是S2，

还要做硬件业务、做软硬件一体，会招聘大量机械人才，在2027年推出软硬件结合的机器人产品，即从S2延伸至S1、延伸至harness环节，从而把整个机器人的模型能力有一个全方位的提升，大脑跟小脑的能力都会有一个全方位的提升，非常值得期待。

从最新的GPT得分看，它6月份刚开始做、开始买数据，9月推出的GPT 6当中就能涌现出非常强的能力。大家用两个方式测评：一是直接用extra去控制，二是用extra加一个Pi 0.5模型去做控制，最终是extra加Pi0.5模型到了榜单第一名，其次是extra直接控制，其次就是其他的开源模型，短短几个月加了N十万小时机器人数据，OpenAI的模型就已经刷爆了机器人榜单。所以说明大语言模型公司进军具身智能领域，对这些公司来说可能是一个降维打击，因为它有很深的模型能力、很多大模型人才、很多数据积累，以及海量算力资源，会给行业带来巨大革新。

在具体操作上，除模型架构本身在迭代，硬件动作也越来越多。从今年6月份看到很多招聘，行业当中也能听到OpenAI已经从苹果、包括特斯拉等头部公司招聘大量硬件和数据相关的人才，最新岗位上也还是有很多软硬件、运控、电机等各方面人才的招聘，说明它还在招兵买马、做产品预研，未来一定会做很大量的机器人硬件工作。这个硬件也是利好，因为它的人才要么从中国招聘、要么从苹果、特斯拉招聘，对中国供应链非常熟悉，而且现在能够提供低成本、高性价比产品的还是中国公司，所以未来它一定不可避免地会来中国采购零部件，可以期待未来一年之内会陆续听到它来中国审厂、下订单、送样等消息，所以这一条线是后续要加大关注和跟踪的重要方向。目前硬件还处在预研阶段，比较早期，只有很少量的消费电子公司在配合它做预研工作；数据采集工作其实早就开启，已经采购大量数据，未来也会加大数据采购。总体影响：一是明显加速整个行业的产业阶段，优先让机器人的GPT moment早日到来；二是会重塑行业格局，特别是一些纯模型的公司会更慌，这类公司可能更多在美国，因为美国有大量公司不做本体、只做模型，受到的冲击最大，有可能被这种最强大脑的公司完全超越；但对中国的公司来说冲击没那么大，因为中国很少有纯软件的公司，更多是软硬结合或纯硬的公司，其实相对利好。因为OpenAI最强的还是大脑层面，如果它把这一层开源，偏硬件的公司也可以基于开源大脑去做自己的S1、harness，或者followOpenAI的路线去迭代自己的模型，还有自己的本体。所以大脑并不会赢者通吃，它在硬件构型上应该有更好的效果，但对其他本体来说还是各有所长，每家公司会找到自己的战略定位，去做更好的软硬结合。另外一些软件能力没那么强、硬件能力比较强的公司明显受益，因为它们本身硬件好、性价比高，但过去因模型能力缺失无法展现实力，加上大模型之后有望发挥更好的功能、服务现实世界。第三是数据方向，明天会详细展开，数据环节一定是最陡峭的，因为大语言模型公司做具身智能一定还是沿用大量数据、大量算力的思路，所以数据一定是投入最大的方向。

八、标的梳理与总结

OpenAI入局对纯软的公司是天大利空，但对偏硬的公司是利好。梳理出的方向：一是本体，包括语数（硬件最强公司之一，性价比高、运控表现非常好，唯一不足是代码，如果放下去看代脑对它肯定是好消息），还有上文提到的公司、优必选等偏软硬结合的公司；二是数据采集，包括数据、数据采集设备等；三是OpenAI链，现在还是聚焦在一些消费电子公司，因为OpenAI未来也要做硬件，这些公司有原有业务的迁移，跟消费电子公司的业务接触目前会更多，但未来也会有越来越多的机器人公司加入。这些是团队这一轮行情最看好的标的跟细分方向。

在第二部分的补充中，柯总指出：OpenAI的入局，市场还没有意识到这一事件的深刻影
响。目前看，大语言模型入局机器人实际上是在加速机器人具身模型的落地，一定程度上天平已经在往大模型这一阵营倾斜，但并不代表其他阵营没有机会
。团队对技术性资料研究的初
步结论是，未来大语言模型再加所谓的具身模型，最后再加物理AI或世界AI，可能是一个
融合的角色，就像现在很多国内外机器人模型公司在做分层研究和尝试一样。即使GPT 6
Ultra，也离不开在具身领域的公司贡献和已经做出来的基础，它只是站在巨人的肩膀上、过来人的肩膀上再进一步。总体而言这是一件好事情，因为机器人能不能放量取决于机器人的具
身模型，现在有了重磅级玩家进来，一定对未来机器人放量是促进和加速。第二点，OpenAI入局实际上是补齐它的数据链，明天会用几个PPT再汇报数据，但整体而言，如果要进行高质量的分级数据采集，一定离不开数据的系统和相关的硬件，包括再往后两天要汇报的传感器以
及机械臂和灵巧手，这恰恰是中国制造业所擅长的，也是中国创业企业和上市公司的机会
所在。第三点，特斯拉依然是在这个位置值得关注的，因为最终人形机器人必然要经历放量，而特斯拉不仅已经开始为它的供应链出货，并且供应链已经在缩圈，头部的这些上市公司团队继续看好，相关上市公司也在PPT最后一页做了罗列。
会议实录：
这是我们系列电话会的第一章节，主题是具身模型。从今天开始，每天晚上9点半都会去介绍整个物理AI的不同环节。今天讲模型，明天会继续讲数据，后面还有传感器、灵巧手加双臂，核心是围绕整个行业最新的变化和增量的环节。

整个行业最新还是迎来了比较大的变化，二级市场此前整体都比较平淡，从上周开始市场终于注意到，整个行业在这一年的时间其实是有比较大的进展，我觉得这主要是围绕在模型层面跟数据层面，可能是比供应链层面和硬件层面更大的两个变化。我们之前也梳理过整个机器人的四大环节：模型、数据、偏硬件，还有应用场景。其实整个硬件是相对比较成熟的，也就是说它现在不是整个行业的短板，最大的短板其实是模型，然后去限制模型发展的就是数据，所以这
一点其实是最值得有边际变化的环节。如果模型环节的突破往前进了一步，也就代表着这个行业最大的瓶颈解开了一步，整个行业就能够再往前迈进一步。所以接下来，我觉得更多的是要关注具身模型在未来发展的迭代。
最近行业最大的一个变化其实就是模型。从今年6月份开始，OpenAI正式开始去布局机器人，并且公开去招聘一些硬件人才；在今年9月的GPT 6S当中，其实已经是加入了几十万小时的机器人数据，而且在测评当中已经能够展现出来非常强的自身控制能力，在很多包括像Robot Dojo等常规测评已经得到了很高的分数，超越了像国内以及全球其他的一些开源模型。也就是大概短短几个月的时间，这个大语言模型公司去转型做机器人模型，就已经能够呈现出这么强大的一个能力。所以我们确实是可以展望，随着它未来从大语言模型、相当于就是PS2的这个环节、偏大脑的模型层级，去往偏本体S1和运控的层级去做，并且去加入很多软硬件协调的一些东西，那它很有可能是能够在未来的模型竞争格局当中担当非常重要的一个地位，去引领整个行业的技术迭代跟技术收敛。这个行业一直等待的一个时间就是机器人的GPTmoment，也就是当它的
模型收敛能够去应用于很多机器人本体，让我们机器人本体能够发挥出来在真实场景当中去应用的能力的这个时刻。那么OpenAI的入局，我觉得是加速了这个时刻的到来，并且带来这个时刻的可能也是GPT本身。所以这个其实是行业最大的一个变化，也给整个行业带来了很多希望。
我们今天会从两个方向去讲：第一个就是具身模型现在是处在一个什么样的状态，除了OpenAI，其他公司现在做得怎么样了，全球中国跟美国两大阵营分别在模型领域是处在一个什么

样的生态位，然后有哪些关键的瓶颈，有哪些测评的指标去测试这些模型的优劣，以及大家目前的排序大概是怎么样的；其次第二个部分，就是大概去看OpenAI的入局会给这个行业带来什么样的变化，在投资上带来什么样的投资机会。
第一个部分先关于机器人模型。整个机器人模型其实分成两个时代，第一个是AI之前，第二个是AI之后。AI之前更多是一些传统的控制，像后来的一些公司，它其实偏数学、偏动力学的控制方式，机器人需要时时刻刻去计算自己身体的重心，去计算自己的下一步输出轨迹，需要去计算下一步的落脚点，然后落脚之后再去计算。所以它整个是非常机械式的发展，整个动作非常缓慢，也没有任何大数据模型和大数据的应用在身体上面，所以那其实是偏早期的一个年代。整个行业的变化其实是在大模型出来之后，大模型出来之后，整个行业也经历了两个阶段：第一个阶

段是先有模型公司直接用大模型去控制机器人，然后就会发现它的整个控制能力还是非常差的，很难去跟物理世界以及跟硬件去做相应的结合，效果很差，可能只能做一步动作、做不了多步动作；然后在这一两年的时间，随着多模态模型的发展，多模态模型跟机器人本体的结合开始去应用，最典型的代表就是现在行业里我们见到的市面上大部分的模型其实都是处在这样一个阶段，包括像谷歌的、包括像OpenAI加Pi 0.5这个最近最火的结构，然后到国内像语数智源，其实大家都处在这样一个阶段，就是上面去接一个多模态模型，然后再接一个机器人本体的硬件模型，构成一个架构。现在这个是阶段四，下一步可能会有新的模型，可能是事件模型，也可能是别的模型，就是在模型上，其实大家还是期待有一个新的模型架构、更有效率的方式去提升机器人的学习能力，因为过去那种大模型的控制方式效率还是偏低的，还是需要很大的算力、很多的数据、很长时间的思考来做这种工作，但这个可能是下一步、是后话。但现在整个阶段四还没有发展到终点，大家还在沿着这条路再往前迈进，也有很多发散的新架构、新算法出来，所以现在大家还是沿着这个阶段四，也就是多模态模型这个主线再往前演进。
这一页大概是现在整个具身智能领域全球主流的一些玩家，可以看到整个玩家的构成还是相对比较偏大厂、或者比较偏大模型公司或者云厂商这样的企业，创业型公司占比还是比较小的。比较有代表性的，首先看北美偏大厂的公司：谷 歌 已经做了很多年的具身智能，从早期的RT系列，再到现在的Gemini加机器人系列，都是这个行业当中非常代表性的公司；英 伟 达 一直在不遗余力地推进机器人模型，希望做整个机器人的基建、去赋能下游机械公司，推出的GR00T系列从N1已经迭代到现在N1.7，还有Cosmos模型用于生成机器人数据，以及专门运用于机器人的世界模型等不同模型，它其实是一个偏基础设施的模型公司，开源给更多的机器公司供他们使用；另外还有特 斯 拉 ，它的技术方案比较保密，我们都不知道，但可以知道的是它的模型应该是非常强的，因为在硅谷也有公司对它的评价；然后就是O p e n A I ，应该是最近最大的变化，它的模型当中未来也会加更多的机器人控制，并且在下一步也会去做机器人硬件；还有一家公司是P h y s i c a l I n t e l l i g e n c e （ P I ）
，它是一个比较有代表性的机器人模型公司，也是一个创业型
公司，它的模型现在也是处在第一梯队、非常领先的位置，最新迭代到Pi 0.7，在整个模型架构上面非常具有创新性和引领性。这几家大概是北美相对比较有代表性的公司。
中国的话大概分为两种，第一种是偏大厂，第二种是偏创业型公司。大厂像小米、阿里、腾讯等等，基本上大厂不会去错过未来这么大的产业趋势，他们都认为这是一个非常明确的方向，未来有很大的想象空间，很多业务可以做，而且是一个很大的市场，所以现在这个阶段整个行业其实不明朗，很多硬件技术、算法技术还没有收敛，所以他们没有把它当成一个商业化项目去做，而是把它当成一个偏预研的项目和研究部门，基本上都会有一个robots lab、机器人的实验室，去做一些早期的机器人数据采集、整体硬件的预研，还有机器人模型的开发，他们也都在官网上可以看得到有一些公开发表的机器人模型，也都做得比较不错。其次就是一些具身智能的创

业型公司，代表性的是像智源、星海图、银河通用、天行智能这几家，大家知名度最高，它们现在的估值基本上都是上百亿、几百亿的，智源最新已经融到七百亿，银河通用也是差不多三四百亿，星海图也是两三百亿，基本都是两百亿往上这么一个估值，在所有的具身公司当中都属于估值排在第一梯队的公司。它们大部分的创始人都是来自自动驾驶行业、大厂或者高校的研究机器人相关的人才，然后去创立的公司，主要聚焦在机器人大脑以及跟硬件结合的环节上面，应该是

国内做得最优秀的创业公司之一。所以整个模型的玩家或者竞争格局，可能跟本体还是有比较大的区别的，还是有很多大厂来做这件事情，我觉得也是更加前沿性的一个方向。
后面我们会快速过一下有一些代表性的公司，它们的模型现在处在一个什么样的状态，大家的模型架构分别是什么样的。首先就是PhysicalIntelligence，我把它放在最前面，是因为这家公司的模型基本上是开源的，有很详细的记录文章，特别是在Pi 0.5之前，它的权重都是开源的，你完全可以去调度它的模型放在自己的真机上去做测试，这个是比较少见的会把自己的权重完全开源出来的公司。其他像星海图也是做了开源的工作，所以它的信息量是最大的，也是最值得我们去探讨跟研究的。它从2024年就开始做这个模型，第一代是Pi 0模型，当时这个模型推出来其实也是给这个行业树立了一个标杆，变成了很多行业的主流方向；在后续的两年时间之内，基本上每隔几个月会迭代一款模型，到Pi 0.5增加了陌生环境，到Pi 0.6又增加了强化学习，到Pi 0.7也增加了世界模型，增加了很多新的模块去提升整个模型的泛化能力。它在北美的地位也是相对比较靠前的，也是百亿美元估值的一家公司。
这个是它最经典的一个模型架构，一个最基础的Pi 0结构，它的结构其实就是一个VLA结构，大脑部分是一个分层的架构，相当于就是大脑一个S2，然后小脑是一个S1，就是一个快慢脑，大脑是一个慢脑，小脑是一个快脑。它首先会先通过大脑去做环境的识别感知、任务推理，然后再把下一步的任务输出给你的小脑，然后小脑再去进行一个函数的计算，最后再去输出一系列动作；最后展现出来的形态，就是这个机器人面对一个世界的变化，比如说你给它一个动作指令，比如帮我倒一杯水，然后它面前有很多的水，你再加上语言指令，它经过计算就输出出来一些动作，然后这个机器人就可以结合自己的视觉，用自己的机械臂帮你去把桌上的水拿起来，再递给你，大概就是这样一个结果。详细去看的话，大脑其实是一个偏开源的模型，加了一个视觉网络，再加一个Gemma（Gemma是谷歌开源的一个模型），是2.6B的参数；小脑部分是基于Q7，就是本体当前的状态，然后再用一个流匹配的架构，也就是说，它通过一个噪声分布去去噪，完成了这个动作去噪，然后再去输出系列动作。可以看得到，绿色部分下面就是noise，上面其实就是输出了一系列动作，A_t、A_t+1、A_t+M等等，就是接下来一段时间你要去输出的动作序列，然后最终测试它能够去控制的本体——当时这个Pi 0模型大概有三种异构本体的控制方式，包括7自由度的机械臂，然后18自由度这种移动的机械臂，以及14自由度的双臂，有不同的对于本体的控制能力。它当时所训练的数据也是一种混合的预训练数据，包括Pi自己采集的数据集、互联网开源的数据，以及网上可以开源得到的数据，通过这种方式去做计算、去做回归，然后去实现了这样的一个模型输出。这个我觉得是最基础的，很多公司可能大概就是这样一个模型架构。
Pi0.7的主要架构还是跟Pi0差不多的，中间还是一个VLA，VLA去加一个基于一个VLM，这个会更为详细。左边是一个也是预训练过的VLM模型，然后有更大参数的5个开源模型来做推理跟计算；右边部分还是这个expert一样流匹配架构的一个模型，通过去噪去输出动作序列，这两个模型的参数都有增加。其次，这个模型比之前要丰富很多，加了很多新的东西，包括observation memory，就是它历史的记忆；第二就是你动作指令，比如"Clean the kitchen"就是去清扫厨房，可以看到，之前可能是一个简单任务、单任务，而这个清洗厨房其实是一个非常复杂的任务，比较抽象，里面包含很多子任务的一个指令，接下来它就是通过一个计算方式，就是下面的high level policy，去把组合式的任务拆解出来，去生成一些子任务集，输出下一步要去做的任务是什么；并且它还有一个世界模型，世界模型就相当于我会在脑海当中去预测，我做了下一步的动作之后，会导致出来一个什么样的结果，那如果它想象中是一个坏的结果的话，它就会规避掉这个动作，然后它就会再去生成一个更加正确的动作，然后如果这个计算完成的话，它就会输出下一步要做的动作。同时在这个模型架构当中，它加了一个东西叫meta data，就是一个打分系统，相当于它会针对它输出来的动作去做一个打分、对数据去做打分，然后得分比较高的就会去输出，所以它其实是可以去提升整个模型输出的一个准确率，第二是可以去提升数据的使用效率。所以这是Pi 0.7模型相比Pi 0模型一些比较大的边际变化。从这个当中其实也可以看得到，VLM、VLA模型可以跟世界模型去做一个结合，就是双方并不是一个完全不交叉的关系，而是可以在具体的模型架构上去做结合，只要能够把这个模型的能力提升、把它的准确

率提升，其实这个模型它没有那么严格的区分。所以这大概是Pi 0.7。
Pi 0.7还有一个变化就是它的数据也更丰富，它加入了示范数据、自主数据、互联网数据和第一人称视角数据。第一人称视角其实从这个模型的变化当中是可以看得出来，它确实是2026年最大的增量、最多的来源方向，在之前2024年的模型当中是比较少见的，但是从2025年底到2026年，其实可以看到越来越多的模型当中是加入了第一人称视角的数据，也就是因为这个数据是比较低成本、可获得的。
后面的话就更快地过一下。谷歌一直在基于VLA这个架构去做模型，最新从之前它是一个纯端到端，就是从输入到输出的一个模型架构，现在也变成了一个快慢脑的架构，它的慢脑就是自己的Gemini，它的快脑也是设计了一个机器人的小模型。它最早期的就是RT-2，最新的话其实就是Gemini 1.5，但是在今年它又迭代了一把，但跟1.5整体都还是比较像的，也是一个快脑再加慢脑再加一个快脑这样的架构。它这个是可以应用在机器人本体上的，有一些很复杂的能力的输出，而且可以实现跨本体的迁移，所以谷歌我觉得也是未来的主要玩家之一。
其次就是英伟达，它推出很多个模型。首先它有一个GR00T N1，专门去针对机器人的VLA模型，从这个架构当中也可以看得出来，也是一个分层架构、一个快慢脑的结构，S2就是一个VLM大脑，S1就是一个扩散（Diffusion Transformer）的动作模型，基本上跟刚刚的模型架构整体还是非常相似的。然后在N1.7当中也是类似，做了一定的升级，加入了一些它Cosmos这个模型的元素，整体的复杂程度都有所提升，整体架构可能跟前面类似，同时它在模型训练当中也是加入了更多的数据，加入了Ego skill数据集，有2万多个小时的第一人称视角数据去覆盖了很多的场景，也就是这个模型更复杂，以及Ego数据变成了很核心的一个增量。英伟达除了GR00T之外，也有世界模型系列。世界模型跟VLA有一个区别，VLA模型更像我们大模型，就是一个输入到输出、然后输出下一步token，然后把这个token就直接去形成一个动作，所以它容易犯错，或者他面对这种新的场景的时候，可能没有那么好的一个表现，就是因为它没有这种自我学习的能力，更多的是基于自己学习到的东西去直接做输出；而这个事件模型不太一样，世界模型它的本质其实是偏这种视频生成的，它是去对未来去做预测，去预测下一步的环境、去预测下一步的动作，然后去先生成一段虚拟的动作序列，然后再去把它转换成一个实际的动作，所以它可以去对未来去做想象，那么它在输出的时候其实也是可以去规避掉一些错误的结果，然后去输出一些更准确的动作序列，所以它面对一些开放的环境或者新的环境，也有更好的一个学习跟适应的能力。英伟达在世界模型其实有两个，一个是Cosmos，Cosmos不是专门针对机器人的，它是一个多元化的，可以去处理视频，然后把语言、动作、音频等等做一个全方位的世界模型；像GR00T zero是基于Cosmos之上，专门针对机器人的一个世界模型。从这个模型的特点当中可以看得到，第一就是它的模型参数量要比VLA更大一些，第二就是它的泛化能力要更强，有更好的跨本体迁移能力，有更好的小样本快速、零样本泛化能力，就是有一些新的任务它也可以去完成，而且有不错的表现，其次就是小样本快速适应新本体。总体来看，就是有更好的泛化能力、场景泛化能力和跨本体迁移能力，但它的缺点就是参数要比较大，以及它的动作是更慢的，算力消耗也要更大，相当于它要无时无刻去生成这种新的视频来做动作的输出，所以这两个模型各有优劣。
下面就是一些细节，论文当中的一些细节，它怎么去实现这个模型的动作输出。其次就是阿里，它也是偏VLA这个架构，也是有一个基于自己通义模型的VLA架构；字节也是一样的，它现在也是变为VLA，当然它可能在其他业务上面也有世界模型，所以也可以去展望，在机器人上未来也会有世界模型，当然它现在主流还是VLA这个架构，也是一个VLM加一个动作这样的构成。然后智源也是有自己的G0，还有G1、G2，也都是基于VLM模型，然后再加了一个plan planner、加了一个规划器，然后在中间去做了一个衔接，这个类似一个中间层，然后再去到这个动作的模型，然后去实现输出。然后银河通用不太一样，银河通用是偏世界模型的，它最近发布的是一个VAM，是一个视觉动作的模型，所以它跟其他公司可能不太一样，它的数据上面也是会包含更多的仿真数据，它这个路线确实是跟其他公司不太一样的。然后星海图，它更多的也是一个VLA的结构，就是基于一个通用的大模型去加一个动作。然后腾讯是世界模型跟VLA两条目前都在

做。然后国外还有一些，还有语 数 ，语数也是两条路线都做，对于语数来说，它的理念是我不做选择，它就follow，行业有什么我先去做布局，所以它既有VLA模型，也有世界模型，它整体做的时间还是比较晚的，大概从去年年底到今年年初才能陆陆续续开始去做，所以它这个能力上面，我觉得可能相比行业的头部，在时间上面要慢个一年时间，那在上市之后再去除追赶跟努力补这个方向。整体上面，我觉得这种大模型公司去开源，确实对于这种偏硬件能力特别突出的公司来说是会更加利好。

其次像美国的G e n e r a l i s t A I ，它也是一个有模型工厂的公司，但是它的模型架构是比较保密的，并没有从公开信息里面找到它的模型架构，然后它也是有很多现实生活当中的任务的得分，在数据增加的情况之下有了很大的增强。第一就是它的模型在迭代，第二就是它的数据量基本上每次就是翻倍增长的趋势去提升能力。然后国内可能像智 谱 ，它在一个最新的论文当中是加了这个数据。

最后去做一个总结。从路线上面来讲，现在主要有三类：第 一 个 就 是 分 层 架 构 ， 就 是 一 个 V L M 再 加 一 个 动 作 的 模 型 ， 是 现 在 最 主 流 的 架 构 ，像行业当中我们看到的像Pi，像阿里、腾讯、语数等等都是这样一个路线，它可以去发挥出来VLM的能力上限，用大脑去做任务推理，然后再去用动作模型生成跟硬件去做衔接、去做输出，但是它有自己的缺陷，现在还处在一个还在发展的状态当中，也是现在绝大多数资源在去投入的方向。第 二 条 就 是 一 个 完 全 一 体 化 的 自 回 归 模 型 ， 把 所 有 信 息 都 放 在 一 个 平 面 里 跑 回 归 这个路线，基本上现在是被证实是相对比较淘汰的一个方向，因为它带来信息的缺失太大，而且对于世界的感知、这种推理能力还是比较弱的。第 三 种 就 是 世 界 模 型 ， 对 未 来 去 进 行 预 测 、 去 输 出 下 一 步 动 作 ， 去 做 发 散 式 的 模 型 ，现在的整个玩家都还是比较少的，除了银河通用之外，今年其实是有很多的创业公司在做，在美国你可能能数出几十家，中国也能数出几十家，而且每家公司都是以非常夸张的估值去做融资。从技术来看，为什么会有这么多新公司出来去做、而且是新的公司，也说明它是一个很新的方向，里面可能有很多的idea和创意；世界模型其实没有一个统一的架构，每家公司的架构都不一样，它是一个很笼统的名称和代称，所以只要你有这个想法跟创意，然后有相应的自己的学术成果，他都能去往这个方向去做转换。所以这个方向我觉得是更偏前沿的，可能还是要先把第一条路往前走一走，然后再去看世界模型未来有没有更好的突破能够去颠覆掉。所以现在这个模型架构可能没有办法说得出一个完全的结论，就是怎样是一个最主流的模型，现在大家都还是以VLA这条路线再往前去迈进。

然后这个行业还存在一个问题，就是它的测评可能也没有那么透明，这块也是越来越多的榜单去提升这方面的测评。一般来说会从四个方向去做测评：推理的泛化能力、数据的使用效率、任务的执行质量和这种推理能力，但我们可能更多考虑的相对比较显性的其实是第一个，就是泛化能力，也就是考察模型在未见过的场景当中它的任务成功率，第一就是任务成功率，第二就是在未看见到的场景当中的成功率，这其实是我们现在对于模型最核心的一个要求，大部分测评其实也是针对这一点来做相应的设计。现在整个行业当中有主流的这些模型测评，包括有一些仿真的基准，有一些真机实采的，有一些行业的标准和赛事测评，最近比较火的就是Robot Dojo这个测评，就是测评OpenAI的模型的得分，然后引领其他公司，是这个测评相对比较权威的一个来源；另外一个比较权威来源是behavior challenge，这个是李飞飞在学术当中设立的一个标准，可能是最严苛的考验，它有100多个任务，然后分布在几十个场景当中，所以它的挑战是最大的，这个可能是这个行业大家去更多参考的数据维度。所以从以上的梳理，其实是可以看得到，整 个 行 业 还 是 相 对 比 较 偏 早 期 和 发 散 式 ， 整 个 架 构 是 没 有 收 敛 的 一 个 状 态 ，这种时候确实需要有一些新的变化给这个行业去指引方向。

那么这个时候英伟达或者OpenAI的入局，给行业带来了一些非常重要的变化。OpenAI在2020年之前其实就已经做过机器人，后面发现机器人的数据积累太慢、在机器人上面效果太差，因此

转向了大语言模型；现在大概又时隔了6年多的时间，他重新开启机器人业务。一方面是说明这个模型经过这几年的发展来到了一个很高的高度，到了真正可以去做机器人业务的阶段；其次就是这个机器人行业在过去几年的时间也发展得非常快，从2022年可能全球没有几家机器人公司、也没有几台机器人本体销售，到今年全球大概每年都有10万台的机器人销售，而且成本从当时可能一台要几百万，现在已经降到最便宜的机械臂你买个可能两三万，像语数的可能10万，成本也相比当时是降了一大步，所以我们去做很多机器人的数据采集，以及它的成本跟它的壁垒，包括整个行业的生态，都是截然不同的。所以OpenAI在这个时候开始重新去做机器人，是因为它认为行业到了一个临界点，到了可以去做支撑的时候，很多关键要素都已经具备了——模型的发展、包括现在本体、包括很多数据的来源，都比较具备，所以他来做，其实是有望给行业带来很大的往前迈进的一步。

现在这个VLA模型经过我们刚刚的梳理，三层架构是S2、S1跟harness，其实G P T 6 U l t r a 它 主 要 去 做 的 就 是 S 2 这 个 部 分 ， 也 就 是 V L M 这 个 部 分 ，它的能力本身就是能够去提升整个机械的控制的。OpenAI未来要去做的不只是S2，它要去做的还有硬件业务，要做软硬件一体，然后会招聘大量的机械的人才，在2027年去推出自己的软硬件结合的机器人产品，也就是它会从S2去延伸至S1，延伸至harness这个环节，从而去把整个机器人的模型的能力有一个全方位的提升，就是大脑跟小脑的能力都会有一个全方位的提升，我觉得是非常值得期待的。然后从最新的GPT的得分当中，其实也可以看得出来，它6月份刚开始做、开始去买数据，在9月份推出的GPT 6当中就能够涌现出来非常强的能力。大家用两个方式去测评：第一个就是直接去控制，第二个就是用extra去加一个Pi 0.5模型去做控制，最后发现是extra加Pi 0.5模型到了榜单第一名，其次是extra直接去控制，其次就是其他的开源模型。也就是通过短短几个月的时间、加了几十万小时的机器人数据，OpenAI的这个模型就已经刷爆了机器人的榜单。所以这其实就说明，这种大语言模型公司去进军具身智能领域，对这些公司来说可能是一个降维打击，因为它有很深的模型能力，也有很多大模型人才、很多数据的积累，以及海量的算力资源，还有更多的能力去做投入，所以会给这个行业带来一个巨大的革新。

在具体的一些操作上面，除了模型架构本身肯定在迭代，它的硬件的动作我们看到也是越来越多的。从今年6月份看到很多的招聘，并且从行业当中也可以听得到，OpenAI已经从苹果、包括特斯拉很多同行业头部公司去招聘大量的硬件和数据相关的人才，最新的岗位上面也可以看得到，还是有很多的软硬件、包括运控、电机各方面人才的招聘，说明它还在招兵买马，然后再去做一些产品的预研，未来也一定会做很大量的机器人硬件的工作。我觉得这个硬件肯定也是一个利好，因为它的人才要么是从中国招聘的、要么是从苹果、特斯拉去招聘的，其实大家对于中国的工业还是非常熟悉的，而且现在能够提供这种低成本、高性价比产品的还是由中国公司来提供，所以未来它一定不可避免地也会来中国去采购零部件。所以我们觉得是可以去期待，在未来一年的时间之内，我们也会陆陆续续听到它来中国去审厂、包括下订单、然后去送样等等这样的消息。所以未来OpenAI一定是要去重视的，一定是最强的、边际变化最大的一个方向，所以这条线我们觉得后续要去加大关注跟踪，应该也会看好这个环节。

除了硬件呢，但是现在硬件还处在预研阶段，还是比较早期的，只有很少量的消费电子公司再去配合它做一些预研工作；然后它的数据采集工作其实早就开启了，已经采购了大量数据，并且在未来也会加大数据的采购。所以它对行业的影响，我们觉得整个行业的产业阶段那肯定是加速的，一定会去优先让机器人的GPT moment早日到来；再其次，它一定也会去重塑行业的格局，而且特别是一些纯模型的公司，我觉得是更慌的，但这种公司它可能更多的是在于美国，因为美国有大量的公司它是不做本体的、只做模型，然后这类公司我觉得是最大受到冲击的，有可能就是由这种最强大脑的公司完全超越。但是对于中国的公司来说，其实冲击没有那么大，因为中国很少有这种纯软件的公司，更多还是去做这种软硬结合的公司，或者是纯硬的公司。其实对中国的公司来说，它是相对利好的，因为OpenAI它最强的还是那个大脑层面，如果它把这个层面去开源，那么偏硬件的公司也可以去基于一个开源大脑去做自己的S1，去做自己的harness，去做很好的软硬结合，或者它也可以去follow OpenAI的路线去迭代自己的模型，而且还有自己的本体。所以我觉得大脑它并不会赢者通吃，它更多的在一个硬件构型上面应该有更好的效果，但是

对于其他的本体来说，应该还是各有所长，就是每家公司会找到自己的一个战略定位，然后去有更好的软硬件的结合，去发挥自己的特点。另外的话，就是一些可能软件能力没有那么强、而硬件能力比较强的公司，这些公司其实是明显受益的，因为他们本身这个机器人就是硬件好、性价比高，但是由于模型能力的缺失，导致他们其实在历史当中是无法去展现很多的实践能力，那么加了这种大的模型之后，是有望能发挥出更好的功能，然后去服务我们现实世界。第三个明确的方向就是数据，这个我们明天也会详细去展开，就是数据这个环节一定是最陡峭的，因为大语言模型公司去做具身智能，一定还是沿用现在的思路，就是很大量的数据、大量的算力，所以数据一定是投入最大的一个方向，这个方向明天我们还会去仔细展开，具体去讲利好的一些细分的环节和标的。

最后就做一个总结和一些标的梳理。OpenAI入局，我觉得对于纯软的公司是大利空的，但是对于偏硬的公司还是利好的。

所以我们梳理出来第一个方向首先是本体，就包括语数，它是硬件我觉得最强的公司之一，而且性价比很高，并且整个运控表现也非常好，唯一不足的就是它的代码，那么如果放下去看代脑对它来说肯定是个好消息；其次就是上文提到的公司，还有优必选等，这些也都是一些偏软硬结合的公司。第二个方向就是数据采集，包括数据、数据采集设备等等。第三个就是OpenAI链，这个的话现在还是聚焦在一些消费电子公司，因为它未来也会去做一些硬件，然后它有更多的做原有业务的迁移，所以这部分其实跟消费电子公司的业务接触目前会更多，但我觉得未来也会有越来越多的机器人公司去加入。所以我们大概梳理出来这些方向，也是我们这一轮行情最看好的一些标的跟细分方向。

以上是我这边的一个汇报。柯总看一下有没有什么要补充的。

感谢各位领导收听。刚才舒杰也是用了将近40页的PPT，把整个具身模型当下全球的进展做了一个罗列。我们想说，OpenAI的入局，应该说市场还依然没有意识到这个事件的深刻影响。从目前来看，大语言模型入局机器人，实际上是在加速机器人具身模型的落地，一定程度上天平已经在往大模型这一阵营去倾斜了，但并不代表其他的阵营没有机会。而且从我们截至目前对于这些技术性和相关资料研究得出的一个初步结论就是，未来大语言模型再加我们所谓的具身模型，最后再加物理AI或者世界AI，它可能是一个融合的角色，就像现在很多国内外的机器人模型公司，在进行了一些分层的研究和相关的尝试一样。那么即使GPT 6 Ultra，它也离不开在具身领域的这些公司已经做出的贡献和已经做出来的基础，它只是说站在巨人的肩膀上，或者过来人的肩膀上，再进一步。我们认为总体而言这是一个好事情，因为机器人能不能放量取决于机器人的具身模型，那么具身模型现在看有了重磅级的玩家进来，一定是对于整个未来机器人的放量是一个促进和加速的作用。

第二点，刚才提到OpenAI本身要入局的话，实际上是补齐他的数据链，那么这一块明天数据也会再用几个PPT再进行汇报。但整体而言，如果要进行高质量的分级数据采集，一定离不开数据的系统和相关的硬件，包括再往后两天我们要汇报的传感器，以及机械臂和灵巧手，这恰恰是中国的制造业所擅长的，以及中国的创业企业的机会和上市公司的机会之所在。

第三点就是特斯拉，其实依然是我们认为在这个位置值得关注的，因为最终人形机器人必然要经

历放量，而特斯拉不仅已经开始为它的供应链进行出货，并且我们也看到它的供应链已经在缩圈了，那么头部的这些上市公司我们也是继续看好的，相关的上市公司也在我们PPT最后一页进行了罗列，有感兴趣的领导可以会后跟我们取得联系，可以向舒杰说。