文件帖 2026-10-29 07:14

来源：刨米机AI精选 | 先知股票数据源

--- 对话物理AI之具身数采-大衍科技高管访谈20261026.pdf ---
全文总结：
公司概况：大衍科技成立于2025年5月，聚焦数据、模型、仿真三条产品线，均围绕多元数据这一核心，解决当前具身智能的数据荒问题；首席科学家为萨迪克院士，首席技术官为原华为诺亚研究中心首席科学家董海威。
数据瓶颈：行业当前最大问题是数据荒，真机、仿真、异构数据均稀缺；对标大语言模型推导，通用模型或需上亿小时数据，行业高质量数据与需求相差约两个数量级。
技术路线：真机遥操作并非必经之路，仿真与异构数据未来可能性更多，业内提出真机、异构、仿真约1:100:100的比例；数据路线涵盖互联网视频、第一人称视角数据、UMI数据与真机遥操作，成本从几十元到上千元不等。
数据壁垒：核心不在单一模型，而在数据来源、数据清洗处理管线与交付能力三个环节；管线包含模型切片、数据融合、空间建模、骨骼映射等步骤。
触觉数据：触觉行业积累不足1%，是精细操作与工业、家庭服务场景的强需求；触觉频率约600Hz高于视觉约130Hz，多模态融合是难点；技术路线包括视触觉、光栅、电容电阻压阻及特种皮肤。
模型进展：底层模型只需赋予机器人基础行动与执行能力，力控与精细化操作仍需在细分领域做微调；世界模型三条路线中，杨立坤的物理规律理解路线意义与前景最大，但难度也最高。
行业格局：海外数据路线已趋于收敛，国内仍百花齐放；海外数据多由东南亚等低人力成本地区采集；具身数据需求明年预计大幅增长，降本空间主要在算力与存储环节。
会议实录：
今天我们请到大衍科技的CFO兼首席战略官刘延新刘总跟大家做交流分享。刘总此前的专业是机器视觉，也在具身行业有多年的从业经验，非常专业。另外，大衍科技也是一家全明星阵容的公司，首席科学家、首席技术官都是整个具身领域非常优秀的科学家代表。今天的会议分为两部分，先请刘总简单讲一下大衍科技的基础情况和团队信息，然后我们再进入行业交流环节。
刘总：感谢国金证券给我们这次机会，跟大家做一个分享。大衍科技其实是一家初创公司，去年5月份刚刚成立，主要聚焦在三个部分：第一个部分是做数据类的，第二个部分是
做模型类的，第三个部分是做仿真类的。这三条路线都围绕着统一的核心，即多元的数
据。不管是数据这部分，我们做的异构数据也好，还是做的触觉数据也好，其实都是在解决当前的数据荒问题。第二块是我们在有了一定数据体量之后，会做自己的模型。这个模型不光是传统意义上的世界模型，我们会把更多元的信息，比如视觉、触觉的信息，力觉的信息加入到模型

里面进行训练，这是我们第二条产品线。第三条产品线是做仿真的，仿真顾名思义就是解决接触力学的问题，用深度学习的模型或者世界模型，将以前物理引擎的部分慢慢替换掉，或者说跟物理引擎进行混用的状态，让sim-to-real这条路走得更顺畅一点。
刚才曹总也提到了，我们有一个比较豪华的明星阵容。我们的首席科学家是萨迪克院士，是加拿大皇家科学院、工程院院士，同时他在谷歌学术里面，在physicalAI领域、数字
孪生领域、沉浸式多媒体交互领域基本上都排在世界前五。我们的首席技术官是董海威教
授，原来是华为诺亚研究中心的首席科学家，之前也是任正非战略委员会成员之一。目前

两位已经全职回到国内，加入到我们公司，并且我们目前也已经把加拿大那边的博士团
队，大概五六个人左右全都挖回来了。
基于这样一个比较豪华的阵容，我们一直在拓展相关
业务。
市场那端，虽然我们成立大概在一年左右的时间，但我们跟基本上所有的头部具身公司、大脑公司，包括灵巧手公司都进行了一些战略合作。并且我们的产品，第一个是触觉数据手套，目前已经做到英伟达的Isaac平台里面了，我们也是国内第一家、首套把触觉手套做到这个
平台里面的公司。之前这家平台里面只有两家公司，第一家公司是Meta，应该是比较著名的，另外一家叫海斯，这两家公司做得比较早，我们算是第三家。同时我们自己的触觉世界模型目前也已经训完了，大概是两百多B的一个模型，这个模型目前也已经嵌入到英伟达的Isaac平台里面了，大家可以直接在访问平台里面去调用，主要是解决接触力学的问题，解决物理引擎底层的一些逻辑问题、一些漏洞。所以我们目前其实是布局在国内，但业务也会往海外拓展的一家公司。
问：从需求侧来看，现在数据带给具身的效果已经看到了没有？目前具身的卡点主要
是卡在数据、模型架构还是其他环节？
答：这两个问题其实是耦合在一起的，具身的意思是模型的体量提升、参数提升，甚至包括数据体量提升，然后堆积出来的一个泛化效果。所以具身其实是把所有数据、模型那一侧的所有技术都混合到一起了。从当前的研发上来看，首先大家能看到的是数据荒的问题，目前缺少不管是真机数据也好，还是仿真数据也好，甚至说是异构数据也好，目前还是属于一个稀缺的状态。当数据量不足到一定程度之后，大家才会去探索不同的模型结构，以及怎样有一个合理的模型架构，然后会有一个很好的效果。所以我觉得当前最大的问题还是要解决数据的问题。通过仿真去降低数据成本也好，通过异构把票样的数据进行数据采集也好，其实今年包括明年，我
觉得大部分的精力还是都放在数据这一个关键瓶颈上面的。
问：对于数据的量级，普遍说法是需要上亿小时。我们这边有没有一个说法或口径？
怎么看待这样一个缺口以及它的增长速度？

答：其实没有一个统一的定义能够回答，一个通用模型需要多少小时的数据。但是目前我们对标传统大语言模型，从它的token使用量推导出来了，目前据说是这个基础模型会需要上亿小时。但其实这两个数据的维度是不一样的，大语言模型可能只是一维的文字，而对于当前的具身世界模型来说，它其实是三维、四维，甚至我加入触觉、力觉，它有可能变成五维、六维这么一个数据量，所以它的数据体量是完全不一样的。上亿小时是不是够用，行业上也没有一个确定的答复。
但是目前从我们实际的开发端来看，像最早的单一动作，比如舞蹈表演、武术之类的，可以用几百小时训练一个比较好的状态；但如果要聚焦到工业商业领域的一些应用，需要很多泛化性能的话，那它需要大概几万小时或者10万小时。所以像国内很多大厂，包括国外的头部公司，他们对于异构数据的需求量，包括整体数据需求量，大概加起来会有一个千万小时的量级。那么这个量级是不是够用？目前没有人拿这么多数据去进行训练过。
今年年初有一个行业调研报告，大概是英伟达的需求量大概是1000万小时，但是整个行业里面所有的高质量数据可能只有50万小时，所以整个是差了两个数量级。站在当前这个视角来看，目前仍然要积累到几百万小时以后才能看出来，比如基础模型到底是不是能够达到好的泛化效果，所以短期还是在积累数据的状态。
问：刚刚说的千万小时的数据需求，主要集中在真机数据，还是其他类型也一起有？
答：目前来看，行业上有几个技术革新的地方。第一个地方就是大家摒弃了真机摇操作的数据，我用第一人称视角数据去进行训练，或者用UMI的数据去进行训练，这个模型表现仍然能够

达到预期的状态，比如任务成功率可以达到95%以上，这是第一个技术上的革新。第二个就是我们自己也在做同步的一些开发，包括仿真的数据、合成的数据。我用仿真环境里面虚拟的一些数据，去给一个真机的模型做训练，然后再做部署，它的任务成功率也仍然能够超过90%以上。所以在用我们自己的视角来看，真机摇操作不一定是一条必经之路，仿真和异构未来可能会有更多的可能性。

之前有一家行业机构调研，大概会在1:100:100这么一个状态，也就是真机数据大概可以占1%左右，那么剩下100倍的异构数据，加上100倍的仿真数据，去实现数据性能的泛化，以及模型能力的训练。所以我们觉得短期来看，对于大家最容易获得的数据，包括仿真数据，在未来的两到三年可能是比较合适的。真机因为受制于具身智能，它没有影子模式，不像自动驾驶一样，人在开的同时可以同步收集数据，将人类轨迹跟我预测的轨迹进行对比，它不会有这种强化学习的方案在里面。所以真机的摇操作数据一定是一个比较低效的状态，未来哪怕真机真的出到很多场景里去应用之后，也一定会是异构和仿真用得更多。

问：智能驾驶的数据对于具身是不是有一定参考性？

答：在特斯拉的数据里面，可能有很多也是废弃的数据，这里面实际有效的数据率也不知道是什么样的比例。特斯拉通过影子模式会收集很多数据，但并不能通过这么时长去简单做一个测算，它会根据自己的场景、OTA把所有数据进行分类，重复的数据会剔除，有价值的数据也会留存，也会打多重标签，所以它实际上不会用这么多数据去训练一个模型。但用这种方式去类比具身，其实是可行的，因为当前的技术底层架构、模型的底层架构基本上是一致的，哪怕我们演化到世界模型以后，底层仍然是继续在这套架构往上搭。从模型底层架构上来看，这两者其实是相通的。

但又有不同的地方，自动驾驶没有接触力学的部分，所以它不会涉及到触觉跟力学的这
些力控交互，而具身是需要这些数据的，这部分数据远比其他传统纯视觉的数据体量要大
了很多倍。所以从数据的复杂度来看，上亿小时的数据我觉得也是要的。

但同样还有一个观点，自动驾驶跟具身最不同的地方，就是影子模式的问题。自动驾驶可以有影子模式存在，去做人工的强化学习校验，但具身很难靠影子模式去收集数据。比如我铺了一台机器人在场景里面，把它的运动轨迹收集回来，并不意味着我能训练出一个很好的具身决策模型。因为自动驾驶能够参考人的轨迹，人的轨迹是一个比较好的标准；但具身的运动轨迹，比如它真的抓起了这个瓶子，那么它是用最优的方案去抓的吗？它的能耗是不是最低的？它的运动轨迹是不是最合适的？这是没有参考的。所以最近这两年大家都在转向去做第一人称视角的数据，把第一人称视角的数据做成未来一个标准，去考核机器人，这也是一条路径，基于这条路径再往下进化的话，那么需要的异构数据体量也可能会上亿小时。

问：力控是具身相较智能驾驶一个全新的增量。目前这块主要缺少的是触觉数据，触
觉数据要在什么时候、什么量级的时候，能够让具身模型做得比较好？

答：目前其实触觉的数据是非常少的，甚至于整个行业积累不到1%，甚至可以说很多家公司积累是0的状态。为什么会提到触觉跟力觉的数据呢？因为在精细操作当中，手指关节与物体的状态接触发生的那个时刻的物体滑移、柔性物体的形变，包括哪怕有一次抓取任务失败与恢复，以及这些长程任务，都需要有力觉的数据在里边，然后去跟传统的视觉、跟运动学做一个映射。但这里面不是说有一个固定小时门槛在那儿，因为随着任务的难度不一样，你的传统分辨率，包括接触物体的多样性不一样，那么它对于数据的需求量也是不一样的，所以没有一个固定的小时门槛。

但我们觉得，触觉的数据对于工业上的机器人来讲是一个强需求，甚至对于家庭服务业
的机器人也是一个强需求。比如我们工业上做网线插接的时候，要感受到插接力、应力释放，才能认定这个任务合格了、完成了，所以触觉的数据一定是必不可少的。但触觉现在又存在一个行业上比较瓶颈的问题，就是触觉是一个很高频的信号，目前触觉大概会做到600Hz，但视觉大

概是130Hz，运动学加上其他的信息频率大概又不一样。那么怎么把这些信息融合到一起，怎么能够让某一时刻的视觉对应到当时发生时刻那个触觉的反馈，这是数据融合比较麻烦的地方。现在整个行业都处于一个摸索前进的状态。

问：您说触觉的频率比视觉更高，也就是说人的反馈灵敏度在触觉上比视觉更高，可以这样理解吗？

答：是的。所以可以理解为未来机器人是不是有另外一种控制方式：上肢大关节的运动是靠视觉去完成的，比如我看到了那个杯子，伸手要去抓它，各个关节的运动轨迹是靠大脑这个VLA模型或者VLM模型去控制完成的；但真的到抓取的那一刻，我可能不依赖视觉，我依赖指尖的触觉，通过触觉这种高频的反馈去完成一个丝滑的动作。未来可能会有这种分布式架构的控制方式。

问：这样一个触觉模型，现在是独立的一套执行端的模型吗？它是不是可以简单理解为一个快系统、偏条件反射性的系统，而大脑是慢系统、规划型的系统？目前行业的技术路径是怎样的？

答：您之前理解得非常对，VLA其实相当于是人类大脑的一套控制方式，触觉那个高频的可能类似于我们的肌肉记忆，哪怕我看不到很多东西，我一样能摸出它的质感，然后把它抓起来完成一些动作，这就是触觉这条线。

当前发展到哪个阶段呢？目前有一个比较大的行业困境，就是目前没有一个统一的标准，大家用的数据传感器都不一样，收集回来的数据来源标准也不一样，结构里面的数据结构也是异构的。所以目前还没有形成一个统一收敛的技术规范，去实现这套类似于潜意识的处理方案。目前大家主流的方案还是基于世界模型的底层架构，去寻一个比如触觉专项的世界模型，用这种方式去做。这种方式有一个好处，也有一个不好的地方：训练完这个触觉模型，它既可以作为灵巧手的控制末端，即触觉的控制末端，也可以作为一个基础模型去做仿真器的接触力学的东西，这是它好的地方；但不好的地方就是这个模型过于庞大、过于臃肿，很难做到在末端做一个低能耗的、类似于端侧的部署。所以目前还是基于世界模型那套框架，大家在探索。

问：触觉的采集方式多种多样，具体落实到现在，有哪些触觉感知层面的设备？您更看好哪一种技术路线？

答：触觉有这么几种。第一种是视触觉，会有一个硬质的薄膜，背后有一个光学镜头，基于薄膜的形变去估计它的法向力，包括它的摩擦系数和纹理这些东西，这是第一种路线。第二种路线是基于光栅的方案，通过光栅的形变滑移去测算这个力的大小。第三种是比较低成本的方案，就是传统的方案，电容、电阻、压阻这类方案。

从工程化的角度来讲，最后一种是最容易实现的，成本最低的，也最容易做到批量化，但它的灵敏度也是最差的。相比于视触觉，视触觉可以把一个指节分成几千个触点，但这几千个触点是基于模型推理出来的，并不是真正实现采集的力控点有几千个；如果还原对应到压阻的话，那么一个指尖可能做到十几个、二十几个就已经算是很密集的了。所以从能力的上限上来讲，我觉得视触觉未来是行业的一个主流方案，但它要面临一些问题，就是它的易损、更换、售后维修，以及大批量使用之后的工程化问题、稳定性问题，尤其是它背后那个视触觉的推理模型的精准度问题，这是视触觉当前要面临的几个卡点。所以我觉得对于触觉采集来看，不管是这几种哪个技术路线，当前行业要确定的一个就是输出数据的结构要大家尽量实现统一，这个是目前大家需要去推动、需要做的。

问：这几种类型的设备，现在价格区间大概分别落在什么维度？
答：视触觉当然会贵一点，但目前也能做到百元级了；像压阻的那种就会更便宜了。还有一种新兴的方式，就是类似于特殊材料的特种皮肤，这种是行业上在延的阶段，但并没有实现工程化。压阻那种方式的量产价格可能只有几十元，所以这个价格对于整个机械本体上面来讲，其实不算是一个很贵的关键零部件；但它其实是一个易损件，可能用1000个小时、几百个小时就要更换一遍，所以从这个角度来看，它的成本还是比较高的。这是一个指尖的价格。目前很少有灵巧手公司会把它做到整个手掌都有，这其实也有一个问题：这类传感器背后一定要有一个控制单元，也就是一个小的算力平台，它的能耗也比较高，如何能够大面积铺开，并且能够用一个低功耗、低算力的处理器，这个其实也比较难。
问：大衍在触觉手套这块应该是世界领先的技术和产品。业内是怎么评价触觉产品的性能，用几个维度去评价？
答：触觉的评价指标有几个。第一个是分辨率，就是在同样的面积之上会有多少个触点，触点越多，分辨率越高，性能就越好。第二个指标是力的分辨率，比如有些单一触点可以实现0.01牛的分辨率，那就做得很好了。还有一个指标是单点的量程，有的一个单点可以做到0牛到最大40牛的量程，也有做到0到20牛的量程，这决定了最大力的上限。第三个就是频率，因为刚才讲到了，触觉频率大概600Hz，但不是生产中真的要做到600Hz，有些公司可能做到100Hz或者200Hz。这几个是数值上大家会关注的几个点。但还有一些是性能上大家会关注的点，就是对于力的感知，是一维的还是三维的，是不是能够感知法向力，这是性能上的评价指标。
问：目前行业里比较头部的几家，核心壁垒或卡点在哪里？最难解决的是什么？
答：目前法向力是比较难的，因为有些传感器天然的劣势带来的，就是没办法识别法向力。比如压阻的方案只有在向下压它的时候，才能感知到一些形变和力，如果想让它转成三维是比较难的，这是底层传感器的结构带来的差异。

问：9月份OpenAI发布了最新的通用模型，这个模型在具身领域表现很好，也有不少本体公司发表了论文。从你们研究的角度，有没有开始用这个模型？感受怎么样？
答：我目前测过这个模型，性能效果确实比较好。如果只是完成一些比较正常的、比如日常生活当中遇到的一些动作，它其实是OK的。它做得比较好的一件事是，它承担了语义理解、任务规划，包括常识推理这一整套的能力，让具身具有了解决实时控制和动作生成，包括去预测一些物理反馈的能力，基于这套能力，它可以实现当前这么好的效果。
但它还有一些需要再加强的地方，比如做一些日常当中遇到的简单动作是没问题的，但如果要让它做一些精细化的控制，还是做不到。这个时候它需要对很多地方做增强，比如扩展一些动作的对齐，对物理状态做恢复，甚至说触觉分配，然后去做一些微调。但我觉得这个底层模型不需要做这么详细的东西，这个底层模型只是让机器人具备一个基础的能力，最基础的行动和执行的能力；当我们在不同领域、不同细分赛道里面去投放、去应用的时候，像刚才提到的动作对齐、语义对齐，包括物理状态这些东西，是在单独的细分领域里面开发需要去做的。它的底层其实就是把大语言模型加上视觉的信息做一个对齐，这么一套方案，所以对于力控、对于精细化操作，它可以提供一些天然的劣势，包括现在即便是VLA做到直接末端的终端，能够实现5mm左右的力控精度也已经算比较高的了。
问：当前行业里具身比较通用的力控模型，用的是哪家的方案，还是自己在训？
答：目前大家其实都是自己在研发这么一个状态。通用的数据世界模型也好，或者说力控的模型也好，目前并没有很多开源出来的。虽然有些行业头部公司、百亿估值的触觉公司号称开源

了一些触觉的世界模型，但并没有人去用，也没有人去测评。比较核心的一个点就是世界模型、大语言模型有相关的benchmark指标参考，但力控目前还没有一个行业的指标在那儿，也没有做benchmark的参考依据。所以我觉得触觉还算是刚刚起步。

问：派0.5这个模型主要负责哪部分？

答：派0.5也是一个类似于偏大脑的模型，它跟OpenAI的模型应用方向其实类似，都是偏向于基础模型。现在有把它们拼在一起的方案，OpenAI的模型可能在语义上做得更好，派0.5可能在执行端做得更好，把这两个模型拼到一起可能会有一个好的效果。

问：之前车上面大家以特斯拉那一套端到端去做训练的标准，您觉得这一套端到端在具身上是不是很难落地？因为触力控的反馈跟视觉和决策层的处理目前还是分层的，所以后面具身是不是还是以分层的架构为主？

答：短期来看端到端是不太可能的。目前主流的方案都是分层架构的方案，简单理解就像拼图一样，我把不同的技能做成一个小模型，基于我想要应用的场景去把所有小模型拼到一起。比如打螺丝是一个小模型，涂胶是一个小模型，那么我在做工业场景里面产线装配应用的时候，就把这几个小模型拼到一起，用分层平台这种方式去做，这是行业比较主流的方案。

为什么做不到端到端呢？主要是因为：第一，自动驾驶的输出其实很简单，就是加速、减速、转向，而对于具身，比如一个人形机器人可能有60多个自由度，每个关节角度、力控都要输出的话，输出量太多了，太复杂了，跟自动驾驶完全不是一个量级。第二，自动驾驶的车辆上面有一个大电池，天然有很好的能源解决方案，但机器人体积没那么大，如果装相同算力大小的芯片，可能工作不了几个小时能耗就耗尽了。所以从工程化上来讲，端到端也是比较耗资源的一件事。

所以从这几个方面来看，短期来看分层的方式是一个比较好落地的方案。

问：具身数据有比较多的路线，包括从互联网上取视频数据、第一人称视角的、仿真的，还包括一些真实数据。这些数据是不是在不同阶段的模型训练中更加适合？比如哪些更适合用在预训练，哪些在后续或者部署阶段更合适？每条路线的成本情况能否拆解一下？

答：首先简单介绍一下当前几个路线。

第一个是互联网视频的路线，通过互联网的视频直接做转化，转化成一个可训练的数据。这套方案的边际效应是最大的，编辑成本最低，互联网上的数据可以随意拉取。但它有一个弊端，比如我在电影里面截取了一个人的动作视频，它会有遮挡、会有衣服，那么我提取骨架的时候会没有那么精准，所以这是精度上的一个缺失，但它的成本确实很低，目前一小时的互联网这类数据可能只能卖到大概30块钱、40块钱。

第二条技术路线是第一人称视角的数据。目前我们可以在互联网上看到，大家头戴一个头环，上面有几个相机，去采集人体第一视角的动作。第一人称视角的数据其实分成几个价值区间，比较差的比较多。最低的比如采集的裸数据，就是没有经过任何标注处理的数据，可能只有二三十块钱；但如果是标注好的高质量数据，也能卖到两三百块钱，所以这是数据标注带来的价值溢价。现在很多异构数据公司，其实核心的技术并不是这套采集设备，更多的是异构数据标注背后的一条数据生产管线。这条管线里面，第一个是大语言模型要去做切片，我采集了很长的一个小时的数据，要把一个完整的动作切出来，我们自己其实也做过尝试，用国内的天问去做切片

的话，会准个大概两三秒，这是第一个步骤；第二块要做的是数据融合，把多视角的数据拼到一起，把不同传感器类型的数据也拼到一起；第三块要做的是空间建模，做一个SLAM建模，把整个空间建出来，有助于做空间的理解，并且把人在这空间里面怎么运动的轨迹也提取出来；第四步

要把人体的所有骨骼全都补全出来，有了关节点、有了骨骼，把关节点的一些运动轨迹全都导出一个机器可读、可训练的一套语言，这样的数据才是最有效的数据。所以做完这一整套的异构数据，能卖到200块钱、300块钱甚至更多；如果只是单单像国内很多数采厂在做的，采一个裸数据，可能只有四五十块钱以内的价格。

第三条路线就是UMI的路线，UMI的价格也是从几十块钱到100块钱不等，主要是根据场景不一样，价格会有一些变化。如果只是一个普通的一张桌子，在上面做一些整理收纳、叠衣服这类的数据，可能就很便宜，大概十几块钱就够了；但如果是工业场景里面真实的工业场景，去做一些装配，那么这个数据溢价会更高，这是场景带来的价格的不同。

还有一条路线，就是我们最早大家都开始做的，各种数采厂都开始做的就是真机摇操。真机摇操价格会很高，目前大概是在500至1000元不等，这个价格主要是因为硬件投入、人力投入、运营投入，高成本投入带来价格的溢价，当然真机的数据目前看也是价值最高的。

再回归到训练方法的问题，比如像互联网的数据、异构的数据，这类数据大多是用在预训练上面；那么UMI的数据，包括真机摇操的数据，更多可能是用在后续的微调上面，这是之前

主流的一套方案。但最近学术界里面也有一些不同的声音，包括我们自己也做了测试，拿纯仿真、纯合成的数据去训一个模型，它的任务成功率也在90%以上，虽然达不到百分之百，但这个完成率已经可以了。所以我觉得慢慢随着技术的发展，可能会有更多的训练方法出来，去颠覆传统的那套逻辑。

问：现在多模态模型对于视频这种模态的数据理解非常到位，数据标注这个环节的壁

垒在哪里？未来下游的数据需求方，直接拿一个类似模型就可以完成数据标注了，那数采

厂的数据标注壁垒会不会比较低？

答：就像我刚才说的，其实一个模型解决不了数据标注的问题。就像刚才那条数据生产管线里面强调的，大语言模型切片、数据融合、SLAM建模、骨骼映射这几步，每一步里面都会有一个单独的模型在做。所以短期来看，可能整个管线里面的精准度，比如SLAM建模的精准度、轨迹提取的精准度，这部分准确率可能是当前大家关注的一个指标。但再往下发展以后，比如不要纯视觉的数据了，要一些带力控的数据、带触觉的数据，那么这个时候数据融合那一端，可能它的模型技术比例就更高一点。再往后面发展，就是大家有批量的数据回来了之后，整个数据管线的生产效率，它的算力消耗、存储消耗，能做得越好，那么成本就会越低，市场就会越大。大概是这样的一个壁垒。

问：现在不同类型的数据，业内有没有一个使用的比例？

答：会有的。像银河通用王鹤老师他们那边做的，最早提出就是90%加10%，90%用合成、用异构，然后10%用真机的，目前这也算是一个主流的方案。但就像刚才说的，也会有很多国外的公司，纯用一种数据去训，也会有一个比较好的效果，也会有这种路线的公司在做。所以目前对于模型、数据怎么去用，并没有形成一个技术上的收敛，大家还是在尝试的阶段。

问：具身数据，比如整个人形机器人的数据，是不是和本体高度绑定？换一个机器人

本体之后，目前采集到的数据价值会不会降低？考虑到目前机器人本体的标准都还没有定

下来，比如一个机器人用多少关节都还没定，会不会未来无本体的采集会成为一个比较重

要的趋势？

答：会的。之前其实有一个卡点的问题，就是数据孤岛，每个数据采集厂用不同本体的机器

人采集回来的数据，只能给这一款机器人使用，不能实现跨本体的迁移。但最近涌现出来的UMI的数据也好，还是异构的数据也好，恰恰解决了这个问题：我不采集机器人的数据，我只采人的数据，由人的数据迁移给机器人，那么我可以迁移给语数的，可以迁移给智元的，可以是轮式的，可以是双足的。所以异构数据和仿真数据，包括UMI数据，最好的一个好处就是解决了跨本体迁移的问题。我觉得未来对这种数据的需求量会越来越多。包括前两天OpenAI发了一个需求，应该是明年会采集8000万小时的异构数据，目前海外的异构数据价格大概是80美金一小时，那也就是说OpenAI要投入大概6亿美金去做数据的积累。所以我觉得明年，包括今年年底，数据需求量会增长得非常夸张。

问：刚刚讲到异构数据后续的处理比较有门槛，那UMI这种数据后续的处理会存在比较
高的门槛吗？

答：UMI的数据，它的好处就是它的硬件结构还依然在，所以它只需要把数据录制出来就好了，不需要做太多后处理的工作，这是UMI最好的地方。它顶多是需要做一个运动、视觉跟其他信息的一些对齐就好了。所以我觉得UMI的数据标注这一环节并没有太大的技术壁垒。

问：海外数采的路线和国内数采的路线，现在是趋于收敛吗？比如都在往异构这个方
向靠，还是海外会有不一样的路线？

答：目前海外主流就是一个路线。我觉得目前学术圈已经形成了一个收敛的状态，但是国内相反，国内还是百花齐放的状态。

问：仿真数据这块，因为咱们也有做仿真。仿真数据比较重大的一个问题，就是可能
存在sim-to-real的gap，要解决这种gap可能需要哪些技术？目前的壁垒在哪里？

答：sim-to-real的核心问题是对质量、摩擦、光照，包括传感器的噪声、动力学的参数
进行随机化，然后看模型包括仿真器能应对得怎么样，这是一个比较通用的标准。目前仿真器里面有几个点要做：第一个是用真实数据去校准仿真器，这样才能有一个更好的输出效果；第二个是仿真训练后要用少量的真机数据去做一个微调，这样去做一个整体模型的性能提升；第三个是针对每次仿真失败的点，同样要用真机的数据去采集相关的数据，去做数据回流，不断提升仿真器的能力；第四个是要做到仿真器跟真机任务的统一和协同，就像我们在英伟达发布的那些demo视频里面看到的，现实当中真机怎么摆，仿真房间里面也怎么摆，通过把这两个数据打通，不断去提升仿真的效果。

从最底层的技术上来讲，如果要提升仿真器，有几条技术路线要走。仿真器最核心的就两块，第一块是物理引擎，第二块是渲染器。

物理引擎像英伟达用的就是PhysX，国内也有很
多其他的公司，比如风吟、索辰这类公司去做物理引擎；第二块是渲染器，渲染器就是让你看到的更真一点，所以无非是给两者进行一个提升。其实现在行业里面还会有一个第三种优化方案，就是用世界模型去跟物理引擎耦合。物理引擎通俗解释上来讲，是一个把很多物理公式拼到一起的大计算器，通过这种大计算器，也叫物理求解器，去做空间求解，得到下一个结果，这就是仿真、物理引擎要干的事儿。那么这个时候，我把一个物理公式拆解之后，让一个模型去把这个物理公式替换掉，有一个好处：原来物理公式里面不会考虑的因素，仿真器的模型可以考虑，比如温度、湿度这些所有参数，都可以用深度学习模型考虑进来，而传统的物理公式是忽略这些东西的；第二个好处就是用模型去做替换、去做耦合的话，可以基于GPU去做并发，效率可以提
升很大一个台阶。这是行业里面通常的三种优化方式。

问：世界模型目前看下来也有非常多不同的路线，不管是作为具身训练场还是作为大

脑，包括Sora那种渲染的、李飞飞的3D空间的，还有像杨立坤的路线。您觉得哪一条技术
路线未来会比较主流？
答：总结下来，世界模型的路线只有三种。第一种是做数据理解的；第二种是做物理空间理解的，就是李飞飞他们做的那一套东西；第三种是杨立坤的那条路线，他去做物理规律的理解，就是我要做物理定义。从这三者来看，杨立坤的那种其实是意义最大的，也是未来前景最大
的，但这个也是最难做的。
问：杨立坤他们那个路线，我感觉最难的是它主要是一个黑盒的状态，其实并不知道
最后训练出来的那个世界模型到底是不是真的能懂物理的规律。目前会有一些什么样的方
法在解决？

答：这个只能用第三方的测评，用真数据去测评的方式去评价。比如我同样模拟一个自由落体，模拟一个小球溅落，看最后的输出结果，它的落点、速度、加速度、轨迹是不是一致，只能靠这种方式去测评。但确实目前世界模型它就是一个黑盒，学术上也有一种争议的声音，就是你把物理求解器一个完全白盒的物理公式，转化成了一个黑盒的、不可解释的东西，那么里面的不可控的内容就太多了，是不是能做到真实的还原也不一定，所以行业内也会有一些反对的声音。
问：关于商业模式，客户怎么样去验收数据的价值？比如通过哪些指标，是数据的时
长还是有效性，包括它的有效率，即最终能用到的数据比例？
答：目前是这样的，拿到数据之后，大家验证标准最直接的一个标准就是我拿一个小模型去做一个预训练，然后看这个模型的预训练结果，跟真实数据训练出来的结果有没有什么差异，如果差异不大的话，就说明这些数据是可用的，这是最直观的一个方式。当然也有其他的方式，就是我会训练出来一些小的模型，比如去识别你的手眼动作是不是统一的，是不是都在同一个视频框架内的，也会有这类的小模型去做一个类似于筛选，或者叫清洗，也会有。但其实从目前行业上来看，拿一个模型去做训练，这是最直接的。
问：不同数据公司的壁垒构建，主要是从哪些维度构建的？是在于模型，还是在于采
集硬件，还是在清洗等各个环节？我们如何成为一个在终端客户会更认可的数据公司？
答：首先有几块。第一块是作为一家数据公司，你的数据来源，比如我要采集一个装配的数据，一定不是在数据采集场里面拿到的数据是有效的，一定是在真实场景里面工人采集的数据是最有效的，所以第一个要卡住数据来源，你跟更多真实的场景工厂合作越深，能采到的数据越多，这是第一个壁垒，就是从数据来源那一段。第二个就是数据回来之后，怎么去做数据
清洗和处理，这里面就会涉及到很多模型的东西，像刚才说的那一套数据生产管线就会纳入到这一个环节里面，你的模型能力越高，生产的效率越快，标注出来的数据越精准，这就是第二层，即数据处理那部分的能力，也就是你的壁垒。第三块其实是做交付那一段，你数据不光是要能标注好、标注准、生产出来，还要能够指导下游的客户，比如你应用这个数据的大脑公司，你要跟他一起去探索这个模型到底是怎么做的，你的数据怎么去用、在哪个环节去用是最有效的。所以第三个壁垒就是你能够摸清用户的训练方法和数据的使用方法，这三者才是一家数据公司壁垒真正要承担的一些角色。
问：从设备的角度，不同的设备目前的价格大概是什么样的？比如头戴式的、触觉手
套，还有各个维度不同的设备？
答：头戴式设备目前因为它就是摄像头，工业摄像头加上一些存储设备和时间同步的触发器，所以它的成本会比较低一点，目前大概能做到三四千块钱左右。UMI的设备会价格会贵一

点，因为它肯定会有一些传感器，包括它的数据可能会更多元一些，所以UMI的设备价格会高一点，有些厂商会卖到大概大几千、八九千左右，也会有一万多的。
触觉手套其实当前来看是最贵的一条路线，但这也是价格差异分化最明显的一条路线。国内有一些厂商大概会卖到9000块钱一套，即一副手套；同样也有一些公司会卖到5万多一套；海外那家Meta公司的触觉手套大概卖到8万多。所以触觉手套因为它的传感器的方案不一样，对应的成本也会不一样，价格差异还是蛮大的。
问：从数据端看，像80美金一小时这样一个数据价格，后面降本空间怎么看？通过哪些维度降本的空间会大一些？是在设备价格下降，还是在其他处理环节？
答：我们自己其实有一套成本拆解的逻辑。比如在硬件设备，一个触觉手套、一个头戴式设备，要能稳定采集1000小时的数据，把每一小时的硬件成本压缩到一美元，也就是6块多，这是硬件比较合理的价格区间。第二块就是所谓的数据生产、清洗的一部分成本，这里面包含了模型标注，即token的成本；第二块就是算力的成本。token目前因为之前把整个价格打下来了，目前我们能够做到大概7块钱一小时的token成本，那么算力和存储这部分大概会在14块钱左右。所以从这个成本结构上来看，算力和存储这一部分仍然有一个比较大的空间在，那么前面1小时1美元的成本基本上已经固化了，随着传感器的方案使用的不一样，可能这个成本会有一些降低。但其实目前大家的关注点还是在算力和存储上面，就是你的存储怎么去调用、你的模型怎么去压缩，然后还能保证你的输出效果，这部分可能是未来数据公司要做的事儿。
当然还有一部分数据成本，就是最前面采集的成本，这个成本我觉得地域差异会更大一点。像很多工厂里面的一些数据，可能天然就会贵一些，也可能大概要给工人补贴，比如50块钱、60块钱；比如家庭类的一些数据，去农村找一些干农活的人去采集这类的数据，可能就只要几块钱。这一部分其实是最不可控的一个成本。
问（电话尾号2716参会者）：现在数采领域，中美两地大概是一个什么样的状况？您刚才说真实的场景里面的数据是最好的，那现在比如这两边大家都是尽可能找一些真实的场景，然后让工人带着类似于摄像头这种方式去采数据吗？能不能简单介绍一下现在行业里面的现状？
答：首先是国内，国内就是数采厂，真实场景都有，但做真实场景数采的公司可能会少一点。海外是这样子的，海外的数据目前东南亚那端采集的会多一点，就是大家把设备给到这边低人力成本的地区，由他们去采集真实的数据，这种方式会多一些，所以海外肯定是这样。
问：刚才听您的介绍，其实我们现在不知道需要多少数据，只是现在说可能先抓紧把数据生产出来，但到底现在生产出来的数据，在未来的需求当中会占多大的重要比例，其实我们现在也搞不清楚，是吗？

答：其实大家有一个共识，就是从大语言模型演化的过程来看，数据一定会对模型效果有比较好的提升，这是一定的；数据一定要积累到一个足够的体量，才能形成一个比较好的模型，这也是大家的共识。但具体这个数据是一个亿为单位的量级，还是说千万的量级，这个是目前没有定论的。
问：比如某一些场景，因为大家现在想解决机器人进去替代人类的一些问题，某一些场景大概需要多少个小时的时长作为训练的基础？有没有一些大致已形成、可以参考的经

验性数据？
答：有一些。比如单一动作的，如果只做一个打螺丝这类动作，大概会有2000至5000小时的数据，其实是做得会比较好的。
问：但打螺丝这个动作好像听上去比较简单？
答：它其实不简单。它其实要涉及到很多东西，比如你在汽车行业里面去打螺栓，可能螺栓比较大，比较好识别；比如电脑生产的时候，半导体上面板子上面的那些螺栓就很难识别。所以第一不光是要完成这个动作，还要对你在生产环境当中的空间有一个理解，这个理解是比较难的，也就是所谓的泛化性。

--- 具身智能数据采集：需求空间、盈利模式与硬件机遇20261027.pdf ---
具身智能数据采集：需求空间、盈利模式与硬件机遇
全文总结
这是一场关于具身智能数据采集行业的专家访谈，主要围绕行业变化主线、数据来源结构、真实采集与训练场的差异、数据链路能力、硬件成本与寿命、数据定价与商业模式、在手订单与市场空间等问题展开，核心信息如下。
一、行业主线：模型需求驱动，成本为辅
行业方式变化非常大。最早以动捕、光学为主，后来转向遥操（原文"姚涛"），到现在无本体采集非常火。更重要的不是成本，而是模型带来的需求变化。
源头是模型需求。模型对应的是机器人大脑，终端需求会牵动所有产业链发生变化。
2024年刚出来的派（π）模型实际是一个视觉语言模型，π0出现时基本用的是遥操数据做小范围的训练；到π0.5时，需要的数据量增多。后续发酵主要由于开源模型被大家广泛使用，并且都认同这种数据来源是一个确定性的数据来源，业务因此扩大。
乐字节在具身数据里线下占比非常高，时间点大概在2026年年初。原因还是VRA、π0.5这类模型开源之后，国内国外都按这种模型方式进行优化或改变，架构都按这个套路来做。架构里有预训练和后训练之分：预训练需要大量便宜的、成本低的、通用性的数据。当模型（如4月份做完的某大模型）以及其他模型验证这个方式是OK的、再适配之后，对整个市场就有很大的扰动。
数据来源结构：如果是针对具身模型的训练、世界模型的养料，或者从数据流通角度看，95%都是第一视角的数据，其余像5米（五米，原文）的数据占比较少。
关于特斯拉：特斯拉在采集方面一直没有对外做采集，一直是内部闭环的方式。
二、乐字节与五米（夹爪）数据：真实采集与训练场的差异
乐字节不耽误干活，头上戴一个装置就可以；五米（夹爪）不太一样，不能完全称为非本体采集方式，因为毕竟是要用人去操作一个夹爪，这一点比较tricky。
预训练时要看人类真实作业的场景、真实的操作情况。如果戴上夹爪去作业，作业的物体以及环节实际上都受限，所以市面上五米数据比较少的原因就在于此——不好采，而且采出来价格跟乐字节差不多，所以市场需求没有那么大。
数据采集应该是inhouse的。比如要训练某个工况，更希望流水线上的女工戴着乐字节去采，而不是把需求包给一个训练场。因为大家真正想要的是真实的数据——要采集真的擦桌子的数据，而不是假装在那里擦桌子的数据；乐字节并不影响双手操作。所以未来趋势是在真实产线上采，对客户订单也有这个要求：在真实场景、真正工作人员去做事情，模仿的是人类真正工作的情况，比如现场扫地阿姨戴着设备、打螺丝的工人去记录。
现在的订单也是在真实产线上，而不是弄一个训练场假设一个场景。
训练场是以前老
的遥操方式：老的遥操都是建立采集厂（原文"素材厂"），称之为任务式采集，针对单一

任务（比如叠衣服、叠完再恢复再叠、重复），而真实采集是在现实过程中戴上设备，比如收拾衣服、叠衣服，这里面会有一些其他动作，需要的是真实情况。

三、触觉数据与灵巧手硬件

目前还没有看到触觉数据的变化。触觉在落地上受限：触觉一般用在手掌、手上比较多，但手本身落地就很难——现在的手本身不成熟，可能几天就坏掉；其次手的操控
性还没有达到一定水平，所以现在落地以及在展示上，用的都是夹爪这种模式。所以触觉数据的核心点在于灵巧手这个硬件现在还不完善，触觉应用因此受限。

四、数据链路的构成与能力门槛

数据链路是公司能力之一。原始数据采集完要上传云端，然后做清洗、对齐、做持续同步、做标注，再做轨迹重建，以及一些结构化的变化。这条链路里每一个管道的处理都要达到一定级别，否则你可能一天只能出1000小时，满足不了大订单；某一环节卡住，整体数据就出不来。

还要根据客户订单做一些定制化要求，比如要一些中间态的图像、实现精度上的一些提升，所以数据链路里要重新做一些开发，数据链路还是比较重要。

创业公司要把这套能力建齐，如果人力充足，大概需要几个月
，比如3个月或几个月能
完成；但要把整体管路的日处理量做到万级，就需要对每一个管路做一些推敲。

五、硬件的构成、成本与寿命

五米（夹爪）的使用寿命比较短，只有几个月就坏了
，因为它不停进行夹爪作业，结构
件不太行。另一个问题是精度不够：长期使用后会有磨损等等，精度不够；结构件容易坏，校准也会有问题——原来归零是贴在一起归零的，做一段时间之后就会有误差。

头戴设备能好一些。但这个行业对电子产品的难点在于：一天8小时、甚至有时候一天20小时不停开机作业，对电子产品损伤比较大，比如发热问题；另外一线操作人员对设备的操作、保养做不到位，所以也有一定的损坏率。预计大概是半年到一年的周期。到期整套设备不能用了，要换新的，核心是里面摄像头等电子产品的问题。

六、价值量拆解与单小时成本结构

乐字节的有效时间：一个人戴8小时进行采集，经过数据链路后剩下的成品数据大概
在1~1.5小时左右（分场景，流水线这类重复、标准化的简单动作转换率比较高，需要的小时数也不多，比如只需要50小时，多了不要）。

硬件构成：一般乐字节设备有双目相机（用于看物体深度、测量尺寸长宽高，因为它需要这些参数）、RGB相机（可以看彩色，必须有的）；摄像头成本加起来基本上是整个设备成本的一半，剩下的是IMU、主板、电池、带存储卡（100G就够）这些，相对便宜。整套硬件差不多5000块钱以内。更高配主要体现在相机的帧率、分辨率、是否有快门等参数上。

单小时成本结构：一手采集过来的有效小时，价格一般在30~80元之间，折中约50元，这是一线采集的成本；数据链路的处理，因为有人工审核和系统耗费，大致在100~150
元，占比还是很高的。

销售价格：国外三四百；国内有波动，一手数据在150~300之间，二手数据折价在80~150。数据可以二次售卖，但二次售卖时依然要经过一些数据链路处理以满足客户
要求；这取决于数据通用性够不够，还要看买卖上的合同定制——比如买你的数据要求独享（即数据产权一起卖给他，不允许再售卖），如果是共享的就可以二次售卖，价格上会有一定幅度差异。

七、在手订单、客户与市场空间

在手订单大致在30~40万（小时），海外与国内差不多对半分。
海外客户主要是OpenAI，此外还有Meta、英伟达等。国内字节是公认舍得花钱的，之前
每个月2000万去买数据，2000万大概一个月也就10万小时；其他像美团等互联网公司也在买；京东在自己采。字节的量实际上跟国外也可以持平，跟模型的真实能力相关。
市场空间：头部几家按OpenAI加字节各100万小时估算，加起来就是千万小时级别。明年订单现在看不到，要等明年年初有预算才能看到，要看模型迭代是否需要加倍、几倍的增长。即使需求变成触觉手套，对公司来说只是硬件的更换，核心的数据清洗等数据
管路还是握在自己手里。
八、其他需求、存量设备与卡点
除乐字节和五米之外，其他类型数据需求基本没有。有一段时间手的数据有需求，是通过手套（动捕手套）进行采集的，量比较少，订单层面没有。
存量设备：按9月份量产2万套算，到现在大概有4~5万套。

数据链路的成本核心是人和系统（系统的开发与维护），里面会有一部分GPU、CPU和存储，再就是人力成本，主要在抽检、检查等环节。
卡点：第一，有没有那么多人的采集团队在同一时间采；第二，数据链路能不能达到
一天处理3万小时的能力。设备本身不愁卖、而且比较好卖，比如卖给对应的地方采集公司、原来的采集厂，以及想建立无本体采集能力的企业，都可以卖。相当于设备先行。
会议实录：
主持人：先从几个维度跟您讨论一些问题。采集这个行业我也看了，陆陆续续关注有一两年时间了，我发现它整个方式变化非常之大：我们刚开始看的时候，这个行业是以动捕或者光学为主，到后来变成了遥操（原文"姚涛"），到现在无本体采集肯定非常火。当然这中间的主线肯定是成本，但其实我觉得更重要的可能还是模型带来的需求变化。您怎么看这个问题？
专家：是的，您说的基本上是一个正确的思路，因为源头都是模型的需求。模型对应的是一个机器人的大脑，它的需求也就是终端需求，所以它会引动所有的产业链进行变化。因为原来的时候，像2024年刚出来的这种模型——以美国的π模型为例，实际上是一个视觉语言模型，2024年π0出现的时候，基本上用的是遥操数据做一个小范围的训练模式；直到出现π0.5的时候，它需要的数据就增多了。所以在后续的发酵，主要是由于这种开源模型大家都去使用，并且都认同这种数据来源是一个确定性的数据来源，所以它就把这种业务扩大了。
主持人：乐字节真正说在具身数据里面，线下它肯定占比非常之高，我毛估估觉得其他实际上肯定（很少）。它是什么时候开始在整个具身数据里面占据相当重要的地位的？大概是什么时间点？
专家：大概是在2026年年初，实际上是年初的事。主要还是由于它这个模型——原来VRA的、π0.5的这种开源，开源之后，国内国外都按照这种模型方式进行优化或者一些改变。所以国内国外的架构都按照这个套路来做，这个套路里有一个架构，就是预训练和后训练，目前大部分的模型是按照这个方式来做的。这里面预训练需要大量便宜的、成本低的、通用性的这种
数据。所以当模型（4月份做完的）以及其他模型验证这个方式是OK的，再适配的话，那对于这个市场就有一个很大的扰动了。
主持人：那么现在比如我们把具身的数据集看成一个数据集的话，整体的构成大概能占据多少？
专家：如果是针对具身模型的训练，还有世界模型这种养料而言，或者说是对于数据流通这方面来讲的话，95%都是第一视角的这种数据，其余的像这种五米的数据占得比较少了。

主持人：其实我记得特斯拉应该更早，我印象里特斯拉可能在2024年下半年或者2025年的

时候，他们第一视角的视频数据应该都是公开承认过的主要训练模型数据来源，您有印象吗？
专家：特斯拉它一直看上去在采集方面没有对外去做这种采集，它一直是内部闭环的一种方式。
主持人：那么我还想请教一个问题，像刚刚咱们说差不多95%以上都是乐字节的数据，那么我们理解乐字节其实是不耽误干活的，我只要头上戴一个装置就可以了。但是五米有一些不太一样，我觉得五米不能完全称为非本体的采集方式，因为你毕竟是要用人去操作一个夹爪，这个是不是还是比较tricky？
专家：您说的也是很正确的。它在预训练的时候，实际上是要看人类真实作业的场景、真实的操作情况。所以如果说你戴上这种夹爪去作业，作业的物体以及环节实际上都是受限的。所以市面上五米的数据比较少的原因也是在于此，不好采；采了呢，它的价格跟乐字节的价格差不多，所以市场需求没有那么大，而且它采也不好采，所以它量比较少，这是主要原因。
主持人：那这里其实很容易带来一个疑问：数据的采集是不是应该是in house的？因为我想比如去训练某一个工况的时候，我更希望的是这个流水线上的女工戴着乐字节，而不是把这个需求比如包给你一个训练场，训练场的方式就是我模拟一套你的工况，然后再进行相应的操作。或者这么说，比如我想让他擦桌子，我只是想采集真的擦桌子的数据，而不是想采集他假装在那儿擦桌子的数据。因为大家真正想要的是真实的数据，而乐字节并不影响我双手的操作。所以未来是不是这个趋势，它会变成一个在真实产线上采的情况？
专家：对，客户的订单也有这个要求，就是你在真实场景、真正工作人员去做这个事情，他要模仿的是人类真正工作的情况。所以就要求是这种现场的、扫地阿姨戴着这些东西，或者是打螺丝的同志们去记录这些东西。
主持人：那所以你们现在的订单也是在真实产线上的，而不是说我去弄一个训练场、在训练场里面去假设一个场景？
专家：对，这个是以前的老的遥操的一种方式了。老的遥操的时候都是建立采集厂，采集厂里我们称之为任务式采集，任务式采集会对于单一任务，比如叠衣服，叠完之后再恢复、再叠，这样重复的方式进行采集。而真实采集呢，就是在现实过程中戴上这个东西，比如收拾衣服、去叠这个衣服，但这里面会有一些其他的动作在里面，所以他需要的是一个真实情况。
主持人：那你们其实是得去找真实的这种工厂产线上的操作工吗？

专家：对，所以现在行业里面就会分成地方性的采集商，是一个级别，就是他会有些地方的资源、有些人脉，能够进入一些工厂、一些学校，然后给这些工人带上去做采集——边干活边采集。所以就会形成这种地方性的采集公司，是比较多的。
主持人：然后我想问一下，乐字节类的订单肯定是碾压式的级别。那么除此之外，五米我们刚刚也讨论过。我观察到一个事情是π出来以后，其实国内外几家高校都用触觉的数据喂到这个π里面，发现任务的成功率提升还是挺多的。所以我们有没有看到触觉类采集的需求？在近期，比如半个月、一个月的时间，有没有一个比较大的需求增长？
专家：目前还没有看到触觉数据的变化。触觉主要是它在应用落地上受限。第一，触觉一般是在手掌、手上比较多，但手呢，本身落地就很难，就是它现在本身就不成熟，可能几天这一双手就坏掉了；第二，手的操控性现在还没有达到一定的水平。所以现在往往在落地，还有一些其他的展示上，都是用夹爪这种模式，在落地上都是夹爪。所以在这种触觉的应用上就会受限。

主持人：所以您觉得触觉数据核心的点其实是在于灵巧手这个硬件现在还不完善，对吧？
专家：对，是的。
主持人：然后我想请教一下，我也看了一些数据，尤其是硅谷那边采集订单的需求，单小时数据的价格肯定参差不齐，我看了非常多的纪要，方差分布还挺大的。但是从年初到现在，如果从同一个维度去观测的话，这个价格下降的趋势怎么样？
专家：海外看上去都是在比如三四百这个价格，没有看到太大的波动。可能是跟大单子有关系——比如这种10万的单子，国内去直接做出来，并且符合客户的要求，实际上是有一定难度的，特别是在上半年的时候是比较难做的。现在随着中国这种数据采集的整体供应链的形成，去采这些东西还是可以的，上半年是接不住的。
主持人：您指的这个链路、这个供应链指的是硬件还是什么？
专家：链路里面实际上就包含，比如上半年的时候是缺这种硬件设备，采集设备是很少的。因为采集设备有一定要求，比如要有这种RGB的彩色相机，又要有双目相机（看深度的时候），以及看颜色以及IMU这些东西。所以在年初的时候，他们也在探索用比如GoPro、用手机去采集，那质量会差一些，客户的要求有些是过不去的。所以当时设备实际上是缺的。另外，采集场景方面，它要求的也是一个广泛性、泛化性的采集场景，在地方上没有铺开，这种采集公司当时也没有说到处都是、这种事情都可以进行采集，所以他们也接不住，短时间内不能产出10万小时，说白了。另外，公司的数据链路，所谓的数据平台，就是数据出来之后要进行一系列数据处理，这一部分能力实际上是有待于提升的，也就是说这个数据链路也是公司的能力之一。
主持人：这部分数据链路的能力，您能展开讲讲吗？比如我们理解的是清洗、标注，然后最后算一个有效的数据——比如采8小时能出来2小时是有效的，这里核心能力点是在哪？
专家：核心能力点是说你能够对这一部分数据进行处理。因为每个点，比如原始采集完了之后要上传云端，上传云端之后要做清洗、对齐，做持续的同步，做标注，然后做它的轨迹和轨迹重建，以及它的一些结构化的一些变化。这条链路里面每一个管路的管道都要达到一定级别，否则你可能一天只能出1000个小时，也就是说不能满足大的订单。你可能能做这套东西，有平台，但是你的管线某一点卡住，就出不来整体的数据。另外还会根据客户的订单做一些定制的要求，比如他要一些中间态的图像，要实现精度上的一些提升，也就是说在数据链路里面要重新做一些开发。所以这里面数据链路还是比较重要。
主持人：比如一个创业公司从无到有，把这套能力建齐，因为现在设备应该相对来说不那么是门槛，设备只是有一定的交付周期，如果把这套能力建齐的话，您觉得大概需要多长时间？
专家：我觉得如果人力充足的话，大概几个月吧。你可以走流程，比如3个月或者几个月能够把它完成。但是你要把它整体的管路做成日处理量达到万级，那就需要对每一个管路做一些推敲了。
主持人：现在一般比如我们的硬件，以乐字节为例，硬件一般使用寿命大概多长时间？
专家：硬件的使用寿命，首先五米的这种夹爪寿命是比较短的，有几个月它就坏了，因为它在不停地进行夹爪的作业，而且它的结构件不太行。另一个就是后面精度不够了——长期使用完以后，比如有这些磨损等等，就精度不够了。一个是精度不够，另一个是它的结构件容易坏，就是它夹爪这部分容易坏。另外你长期夹的话，它的校准也是有问题：你原来归零是贴在一起归零的，但是你做一段时间之后就会有误差。头戴设备能好一些，但实际上做这个行业，对电子产品的难点是它要一天8小时、甚至有时候一天20小时不停地开机作业，所以这点对电子产品的损伤实际上是比较大的。比如发热的问题；再就是作业人员、一线操作人员，他对设备的操

作、保养做不到位，所以它也是有一定损坏率的。当然这里面也是因为比如第一代设备、第二代设备有迭代，也有它自己的问题。预计大概在半年到一年这么一个周期。
主持人：半年到一年，那就等于整套设备不能用了，得换新的了，对吧？核心是里面摄像头等电子产品的问题？
专家：对。

主持人：既然聊到这儿，咱们再聊聊价值量。比如现在这样一套硬件，我们大致给它算个账：它整体的硬件成本大概在什么水平？然后有效采集小时数，比如一天在多长时间，我们有过相应的测算数据吗？
专家：只能说乐字节的。有效时间，比如一个人戴着8小时进行采集，采集之后经过数
据链路，剩下的时间大概在1小时1.5小时左右，这个就是成品数据。
主持人：那等于一天也就是才1~1.5小时左右，如果工人正常上班的话。但是它是分场景的，有些场景是连续的，比如像流水线，它的转换率比较高？
专家：对，因为场景会影响你采集的效率。但像流水线这些频繁的、称之为简单动作的，因为它是重复的、标准化的，这些动作需要的小时数也不多，比如我们只要50小时，多了再不要，就是为保证（质量）。
主持人：现在一般来说，整套头戴式的设备里面主要含些什么呢？比如摄像头、相机、IMU等等，主要能大致给我们讲讲吗？
专家：一般乐字节设备有双目相机，主要去看物体的深度，能够测量尺寸长宽高，因为它需要有这个参数；第二个是RGB相机，可以看彩色，这也是必须有的。摄像头的成本加起来基本上是整个设备成本的一半，剩下的就是像IMU，然后再就是主板、电池、带存储卡这些，都是相对便宜一些，100G的就可以了。整体来说，整个这样硬件差不多5000块钱以内。
主持人：它还有什么比如更高维度、级别更高或者增配的说法吗？
专家：这里面主要是相机的一些参数，是不是一个高配版的相机，比如帧率、分辨率这些东西，是不是有快门。
主持人：那我们从另外角度拆一下。比如价格差不多三四百人民币一个小时，我们这个设备按一年去折旧的话，应该不会太贵，那核心成本主要是在数据链路的处理上，这个我们有算过吗？比如按三四百一小时的话，硬件占了多少、数据处理占了多少？
专家：毛利大概能占个30多吧。我可以给你个范围：采集过来的这种有效小时，价格一般在30~80元之间，你可以说折中就是比如50元，这是一线采集的成本。数据链路的处理
呢，因为这里面有一些人工的审核，还有一些系统的耗费，大致在100~150元。
主持人：那它占比还是很高的。
专家：对。卖的价格呢，国外的三四百；国内实际上也有波动值，他有卖这种一手的数据以及二手数据。一手数据在150~300之间，二手数据就折价了，在80~150。
主持人：也就是说你的数据可以二次售卖，但是二次售卖的时候你依然要经过一些数据链路的处理，满足客户的要求。就是你这个数据通用性够的话，就能够卖第二回？

专家：对，但还要再处理一下，这里面就是一个买卖上的合同定制。比如有些（客户）买你的数据的时候就要求你这个是独享的——所谓独享就是他独享，也就是说数据产权一起卖给他了，那你这个数据就不允许再售卖了。如果是共享的，那么你可以进行二次售卖，价格上会有一定幅度（差异）。
主持人：我们现在公司肯定算行业比较头部的，比如现在在手订单大概是一个什么水平？
专家：在手订单大致在30~40万（小时）。
主持人：这里海外跟国内怎么比呢？
专家：海外跟国内差不多对半分。
主持人：海外的话，你们应该主要还是OpenAI的，对吧？

专家：对，是的。
主持人：除了他之外，还有谁呢？
专家：还有像Meta，还有英伟达。（其他）再继续再看吧。
主持人：国内是不是单子就散一点、小一点？
专家：国内也有大的，国内就像字节是公认的舍得花钱的。他之前每个月2000万去买数据，2000万大概一个月也就10万小时。
还有一些其他的大的互联网公司也在买，美团等，京
东没有——京东在自己采。字节实际上跟国外的也可以持平，这就是跟模型的真实能力相关
的。
主持人：如果我们这么去估算的话，比如按字节这个数，我们也能排出它是100多万个小时，OpenAI也加上，这么多人。我们内部有过测算吗？比如今年肯定是刚刚开始的一年，今年整个市场的空间大概有多大？
专家：这只能是一个预估。头部这几家按照刚才讲的加总，比如OpenAI100万小时，
字节100万小时，加起来就是千万小时，就是这个水平。
主持人：现在能够看到明年的订单了吗？还是说要等明年年初有预算才能（看到）？
专家：这看不到，看不到。
主持人：你们现在OpenAI跟字节的（数据），你们就是交付过去了吗？
专家：对，正在交付。
主持人：有什么反馈吗？
专家：反馈还好吧，不能说是（特别好），就是持续交付、持续沟通的一个过程。

主持人：那其实这个市场空间还不小。我们国内除了咱之外，还能看到其他哪些我们觉得跟我们的体量也能够不相上下的？
专家：你是指我们这种平台性的采集公司，是吧？同行啊，像质检光轮（原文"检质光轮"）、路明（原文"陆名"）等。
主持人：比如国内这些人，你想假设一共1000个小时的话，咱们已经干了30~40万个小时了，对吧？
专家：对。
主持人：但这个可能就是放出来的、明年的量也不太好估？
专家：是的，明年就要看它模型的迭代，是否是需要加倍、几倍的这样的增长。
主持人：因为这个其实对你们来说是一样的，它如果变成其他，假设变成触觉手套的需求，对我们来说其实就是硬件的更换，但你核心的数据清洗等所谓数据管路，这个其实你们还是握在自己手里面对吧？
专家：对，是的。
主持人：我们会议秘书看一下线上有没有小伙伴有问题。（会议秘书提示提问方式）
主持人：专家，现在还没有其他人有问题，我再问一下：我们现在客户有跟你们沟通过，除了乐字节和五米之外的其他类型的数据需求吗？
专家：其他的基本上没有。另外要的是，只是说在某一段时间，比如有手的数据有一段是有需求的。手的数据怎么采呢？就是有这种手套，戴着进行采集的。
主持人：是动捕手套还是触觉手套？
专家：应该是动捕手套，这块因为比较少，没有进一步去研究。从订单来说是没有的。
主持人：像我们现在对应这个订单量，实际硬件比如大概存量有多少套？
专家：存量如果按照量产，9月份量产2万（套），到现在大概有4~5万套。
主持人：那比如整个负责数据链路的大概有多少人呢？
专家：这个可能不太好统计。数据链路核心的成本应该是人和你的系统，就是系统的
开发跟维护。在这里面它有一部分GPU和CPU，还有存储，然后再就是人力成本。人力成本主要

是在有一些比如抽检、检查这些环节。
主持人：比如存量有4~5万套，我们假设就用3万套，一天我们假设有效的就一个小时，那你比如说30万个小时，这不是很快就能采集完吗？

专家：这里面卡点是：第一个是有没有那么多人的采集团队去同一时间采；第二是你的数据链路能不能达到一天处理3万小时的能力。这里面是有几个卡点的。对我们来说，弄了这么一些设备，实际上因为现在是设备比较好卖，比如卖给对应的这种地方采集公司，以及原来的采集厂，然后他想建立这种无本体采集能力的，都是可以卖的，所以设备是不愁卖的，而且是比较比较好卖的。
主持人：我理解，就相当于设备先行。
专家：对，设备先行。
主持人：那我看看，应该没有人提问了。好的专家，今天非常感谢您的时间，跟着您学到很多，我回头细细想一想，如果有什么细节的问题搞不清楚，到时候可能再麻烦您。

--- AI的下一战：巨头集体押注“机器人大脑”，具身智能迎来“ChatGPT时刻”？20261028.pdf ---
全文总结
一、会议背景与整体观点
本次直播是在万得3C的邀请下进行的，上一次是在7月30日同样于万得3C平台做直播，当时汇报的机器人主题观点是看好三个结构性机会：数据（数采）、ToC应用和垂类应用。这两周整个人形机器人行情表现非常不错，特别是呈现全面开花的表现，也演绎了此前这三个主线的逻辑。本次直播一方面借近期市场关注点——大约前两周OpenAI创始人在参加一次播客后表示会投入去做机器人的硬件、会在明年发出样机，引发了机器人行情——首先汇报这几个月重点的变化，特别是大模型大公司、巨头们开始投入做具身智能，他们的战略方向是什么，以及在这个变化中可能迎来哪些新的投资机会。
二、机器人scalinglaw的真实存在
今年数据之所以出现爆发、巨头之所以要投入，是因为他们看到了在机器人和具身智能领域的投入是有回报的，机器人的scalinglaw是真实存在的。此前很长一段时间，模型训练数据增加，但训练效果没有特别显著的提升，所以那段时间不会有很大的投入；现在不一样了，已经有一些实际例子证明了机器人scaling law的存在，让大家看到了曙光，投入也进入加速阶段。
三、各巨头在具身智能方面的布局
OpenAI：9月份发布了最新的通用基座大模型GPT-6（Astro版本），发布之后可以看到它在操纵机械臂层面有突出的表现，证实了这个通用模型在机器人的空间理解、任务规划等方面表现出非常强大的能力，此后在一些机器人榜单上排名阶段性第一，各项任务中的成功率达到了30%多。虽然阶段性排名第一，但目前GPT-6 Astro在操纵机械臂时肯定还没有达到商业化状态：一是灵巧操作能力还不足，二是token消耗量特别大、成本非常高。但大家对这个表现的解读，会归结于未来具身智能实现的时间可能会被大大提前——此前产业界和学术界虽然有不一致的看法，但很多都觉得可能要5~10年才能看到具身机器人商业化大面积铺开；而现在由GPT-6 Astro的表现出发，大家可能会普遍把时间从5~10年提前到两到三年，未来很快可能看到机器人在各个行业、各个方向的落地。这个需求是完全存在的，不用疑虑，无论是ToB进工厂还是ToC进家庭做家务，需求都是真实存在的，只是目前机器人特别是大脑端的能力能不能达

到、覆盖掉这个需求。所以这是最深远的影响：这个行业可能会迎来一个加速式的发展。近
期宇树的创始人王兴兴上周在杭州开了物理AI大会，会上也发表论断，认为到GPT-7阶段可能就会击穿具身智能，这是近期行情表现好的原因——由于OpenAI的布局催化。

Anthropic：在具身智能方面的布局，我们认为目前还是在真机实验，以及一些研究人员和系统评测的投入当中，因为此前Claude推出了一个评测，覆盖了机械臂、四足以及人形机器人，也覆盖了像宇树机器人的一些操作。
谷歌：有DeepMind部门，此前发布了几代具身智能的机器人大脑模型，特别是在VLA层面的成果关注度非常高；另外谷歌在基础模型、机器人数据以及商业化端也在同步推进。
英伟达：在机器人方面定位依然是一个基础设施提供商，希望机器人时代大家也用它的卡，目前大家用得最好的机器人芯片还是Jetson（Orin）平台，此前也主要是用英伟达的机器人芯片。模型层面推出了GR00T系列模型，世界模型层面推出Cosmos系列；仿真也是一个代表性方向，无论是此前的运动控制层面，基本都用英伟达的仿真器。从仿真平台到各类模型和数据，英伟达都提供了全套，所以在未来机器人时代、物理AI时代，英伟达还是扮演基础设施提供商的角色。
国内两家代表性大模型公司：DeepSeek
最近也传出很多消息，梁文锋在接触具身智能公
司，也投了一家，此前在宇树上市时还参与了宇树的战略配售，所以DeepSeek可能从一个大模型公司往具身智能方向发展，投资跟合作已经证明了它的主动布局；Kimi此前在9月份也传出正在招聘相关人才，说明公司在这方面也有布局动作。以上是近期一些巨头的布局情况，其中最值得关注的就是OpenAI的动作，也正是OpenAI的动作引发了这一轮机器人行情。

四、模型进展证明机器人scalinglaw
近期无论是大厂还是机器人创业公司发布的模型进展，在很多层面都证明了机器人scalinglaw的存在。这将使得无论是具身智能公司还是大模型公司，未来都会进一步去投入数据，因为数据量的提升带来模型效果更好，可能在千万或者更高的小时数阶段跨越基点，使机器人真正商业化得以实现。
一是Generalist：它发布的模型从0到1.5，预训练数据从27万小时增加到50万小时以上，平均成功率由64%提升到99%。虽然更多是自己公司发布的数据情况，但在一定程度上证明了数据规模上一个台阶，任务成功率也会提升。
二是数据采集方法的变化。此前数据采集方式更多是摇操作，穿一套数采装备去摇操真机得到数据，成本非常高，通常用于机器人的后训练（微调）。预训练层面需要海量数据，此前行业分成两类：一类是仿真派，认为真机数据成本太高、效率太低，用真机数据做预训练遥遥无期，这也是此前很多判断具身智能商业化要10年甚至更久的原因。变化发生在从去年到今年，
预训练阶段开始用无本体数据，目前主要两类：一是Ego数据，二是UMI数据。

Ego数据：今年2月份英伟达发布了EgoScale论文，用大规模Ego数据（2万小时）进行预训练。所谓Ego数据，就是人类第一视角的视频数据，核心是在头上带一个摄像头，有时两个手腕处也带摄像头，把你如何干活拍下来。这篇论文证明：用2万小时Ego数据预训练，再用非常少量（50小时）的人类摇操作数据，加上4小时机器人自己运行的真机数据做对齐，就得到了非常好的效果。典型案例如叠衣服这类简单任务，在one-shot（只见过一次）的场景下取得了80%以上的成功率（88%的少样本迁移成功率）。后续延续的工作也非常有意义
：今年8月，一家有几个
从英伟达出来的人创立的创业公司发布了自己的模型二代，用了百万小时Ego数据
（百万小时人
类视频数据），效果非常好。随着数据量提升，在14项真机任务中的得分逐步提升：1000小时得分20%，到100万小时得分提升到53%。这就证明了机器人scalinglaw的存在。

所以从今年二、三季度以来，已经观察到机器人行业数据采集的爆发，这个爆发更多不是摇操作数采的爆发，而是Ego数据的爆发，也有一些UMI数据。Ego数据最核心就是几个摄像头（三个摄像头），现在又把触觉手套、触觉数据加进去，加进去之后已有模型厂商反馈模型效果有很好的提升。所以除了Ego数据设备、数采服务和数采设备之外，触觉数据也非常重要。

五、投资机会：数据金字塔与数采市场空间

投资机会与上次汇报观点大差不差，目前大家看到数据基建的爆发。本次结合最近调研、研究，对数据数采设备以及数采服务市场规模做了更好的摸排。

具身智能数据分为金字塔五层结构：

金字塔顶端是机器人真机数据，成本最高、效果最好，一小时可能卖到大几百块甚至上千块。我们认为这类数据后续应该都是具身公司自己去搞，因为训练模型时主要用于后训练、对齐，这块一定要用到真机数据。因为成本高，且这两年一直在谈数据的异构性，如果不是用自己本体获得的数据，效果肯定不是最好，所以最好是用自己本体的真机数据，未来可能由机器人本体公司自己掌握。当然现在也会有一些产业联盟搞数据集，把高质量数据放在数据集里面，谁去用就付费，形成行业里的高质量数据集。此前我国数据量比较少，近几年包括国家层面也构建了自己的数据集。

第二层是无本体数据，分成UMI数据和Ego数据两类。第三层是仿真数据，肯定也有意义，

跟自动驾驶一样：可以低成本、无限获得，在一些真实环境不太好模拟的极端数据上也可以用；但问题在于sim-to-real的gap，目前肯定无法跨越。有一篇论文提到，在仿真环境里成功率80%多，但迁移到真实环境里大概只剩20%多，这就是sim-to-real gap的问题，目前还是无法解决，所以仿真数据在一定程度上仍有其意义。再往下一层是通用数据，即互联网上的海量视频数据或文本数据，提供语义理解和推理先验，但这类数据缺乏真实的接触性信息和动作信息，不过它是海量的，比如训练大语言模型、多模态大模型用这类数据；但训练机器人，还是需要金字塔上层的这些数据。

关于无本体数据：这一波真正爆发的是无本体数据。UMI数据从前年开始越来越多，它是一个手持的、有末端操作器（一般是夹爪），带着摄像头以及一个姿态跟踪模块，去执行动作，比如洗碗、拿东西，会把运动拍下来，另外有轨迹信号，有些末端还会加触觉信号；它可以适用于任何本体，通过逆向往定向转换成机器人的轨迹，适用于训练任何本体。Ego数据就是几个摄像头加触觉手套（有的也用机电手套）去做数据收集，是目前正在爆发的数据。

未来需要多少数据、数据配比如何，是两个问题：

数据量问题：参考行业内大厂或头部公司的指引，今年一些公司如智源旗下的蜜蜂目标是750万到1000万小时，京东也有比较高的目标、今年想增加500万小时，还有其他一些公司有百万级别的目标。所以今年加起来，中国市场上可能有2000万小时以上甚至2500万小时的数据需求量。未来参考智源旗下的逆风给出的量级是百亿级别小时数（存量概念）。如果假设到2030年达成百亿小时数据目标，按每年提升做假设；如果提前，比如未来两到三年（如2029年）实现，那么当年测算的需求量又会增加。

数据配比问题：目前并没有明确，我们也去拜访调研过很多机器人模型公司、机器人本体公司。有一家非常头部的、首席科学家层面的人谈到，在他的层面当然希望都用最好的金字塔顶尖数据，但不可能，因为无本体数据其实是效率和成本最终折中下来的选择——如果不考虑效率、不考虑成本，大家都想用最好的真机数据，但不可能。所以目前无本体数据的数据配比至少在50%以上，甚至可能60%以及更高的配比，用无本体数据去做模型训练。基于此也做了关于无本体数据的测算：如果未来要达到百亿小时、年新增几十亿小时的量级，我们认为数据服务当年的体量有可能达到千亿市场规模。

六、产业链与设备端机会

目前做数据采集的公司越来越多，7、8月份在世界人工智能大会和世界机器人大会上调研时，看到好多家都推出了数采方案以及无本体数据的解决方案。目前确实是发展初期的行业，但

也确实在爆发，玩家涌进来很多。
数采服务市场：全球有McKinsey AI？——Scale AI这类给大模型公司提供数据服务的领先公司，估值在200多亿美元级别；全球另一家有Physical Intelligence（PI）？——此处以原文为准：数采服务市场全球有Scale AI，估值200多亿美元级别，另有一家公司估值大致在

十几亿美元水平，都算目前产业头部的公司。目前一级市场估值水平并没有特别高，跟大模型领域公司还有数十倍差距。
国内做得比较靠前的有简智、路明、光轮等几家，各自有各自专攻的方向，目前算国内排名比较靠前的几家数采服务商。
设备端：目前有一个特点，数据服务商也会做一些自己的设备，做设备的有些也会往下游服务延伸。设备端玩家更多，因为Ego数据最核心就是3个摄像头加触觉手套：做摄像头的其实都可以做，或者说对摄像头要求不高，采购一些也可以做出来；触觉手套，所有此前可以做触觉传感器的公司都可以去做。这些公司其实都发布了相应产品。
我们的观点是：如果未来这个市场确实要爆发，一定会考验设备端的标准化能力和量
产能力。在设备端、Ego摄像头这个层面，我们还是会看好奥比中光，虽然它是7月份发布的数采产品，但这家公司实打实会有产品的标准化能力以及量产能力。触觉手套方面，目前也有很多家公司发布，可以去看后续跟一些下游大模型厂商以及机器人厂商合作的情况。上市公司领域，包括像福莱新材、日盈电子、汉威科技等此前做传感器的公司，其实都有相应的产品。
七、ToC与垂类应用方向
ToC方向：建议继续关注上纬新材，它在9月22号发布了两款机器人新产品Q1和T1。最近提到累计下线，其实才刚过一个月，累计下线就已经5000台。明年产品放量非常值得期待，因为定价也不高，Q1是1.99万，而且公司在初期就开始配套开发了很多可玩性，包括3D打印外装，以及内置一些教育模块去学习等等，对一些中高端家庭还是非常有吸引力。如果明年如期卖到数十万甚至几十万的量，我们认为它在具身智能领域里出货量将会排名第一。
垂类应用：近期重点推荐杰克科技，今天正好杰克在开它机器人的产品发布会，我们会继续看好这家公司。看好的核心逻辑在于，我们看好这个时间点一些垂类应用方向会开始起
量。在7、8月份的世界人工智能大会和世界机器人大会上，看到一些垂类应用方向的进步，当时观察到的场景是物流分拣：像星动纪元、自变量等公司，其实都已经达到准确率超过99%，效率已经接近或者超越人工，完全可以到了可以替人的阶段。物流分拣包括物流行业，也被认为是物理AI率先会落地的行业。刚提到的这几家公司都是一些非上市公司的创业公司。
落到上市公司层面，建议关注杭叉集团（原文"航沙集团"应为杭叉集团）：公司昨天开了三季报业绩会，重点提及了他们在叉车单车智能方向的布局，以及未来在物流场景里做更多智能化布局的事情，公司质地也非常好，是叉车龙头公司，估值也比较低，这是一个可关注的标的。

另一个可关注的标的是中邮科技：这家公司其实是一个应用方，作为中国邮政旗下的硬科技平台，可以在应用端利用现在已经开始突破的技术，首先给中国邮政体系内做到降本增效、获得收益，另外也可以拓展到中国邮政体系外的市场。
除了物流场景之外，找到的另外一个场景就是杰克科技。杰克作为一家做缝纫机的公司，质地也非常好，发布的机器人目前在跟模板机结合、跟订扣专机以及所眼（锁眼）专机相结合方面，准确率已经完全没问题，效率也在逐步提升，可以实现在纺织服装这个行业里面积其
替人的目标。纺织服装行业目前的工人人数非常多，全球将近1亿纺织工人，随着特别是中国目前想找一个几千人的纺织工厂已经很难，缺人缺劳动力还是现实存在的问题，所以中国市场有实在的机器替人需求。行业在往东南亚转移，对于全球龙头的服装品牌来说，选择合作代工厂时也希望看到他们装备能力的提升，所以在这种情况下，杰克新一代的AI产品和机器人产品，也会助

力这些服装代工厂去绑定他们的下游品牌客户。所以我们非常看好它后续AI缝纫机跟机器人
的放量。
会议实录：
主讲人：在万得3C的邀请下，上一次是在7月30号，当时也是在万得3C这个平台上跟大家
做了一次直播。当时我们汇报的机器人主题和机器人的观点，就是看好三个结构性机会，分别是
数据（数采）、ToC应用和垂类应用
。我们看到这两周整个机器人行情表现非常不错，特别
是今天有一个全面开花的表现，也演绎了我们此前这三个大主线的逻辑。
今天一方面也是借近期市场的关注点：大约前两周，OpenAI的创始人在参加一次播客之后，
表示自己会投入去做机器人的硬件，会在明年发出样机，由此引发了机器人的行情。本次直播会
先跟大家汇报这几个月重点的变化，特别是大模型的这些大公司、巨头们开始去投入做具身智
能，他们的战略方向是什么，以及在这个变化当中，我们可能会迎来哪些新的投资机会。
首先跟大家汇报巨头的布局以及机器人的scalinglaw。
今年大家看到数据的爆发、巨
头要投入，其实是他们看到了在机器人和具身智能这个领域的投入是有所回报的，机器人的
scalinglaw是真实存在的。因为此前可能有很长一段时间，模型训练数据增加，但训练效果
好像没有特别显著的提升，所以在那段时间就不会有特别大的投入。但是现在不一样了，目前已
经发生了一些实际例子，证明了机器人scaling law的存在，让大家对这个投入看到了曙光，投
入也就进入了加速阶段。
首先看巨头的布局跟战略准备。

OpenAI：9月份OpenAI发布了最新的通用基座大模型GPT-6的Astro版本，发布之后，大家就
看到它在操纵机械臂层面有突出的表现，证实了这个通用模型在机器人的空间理解、任务规划等
方面表现出了非常强大的能力。此后，在一些机器人榜单上，它也是排名阶段性第一，在各项任
务中的成功率达到了30%多。当然，虽然阶段性排名第一，但目前GPT-6 Astro在操纵机械臂的时
候，肯定还没有达到一个商业化的状态：首先是它的灵巧操作能力还不足，另外就是token消耗
量特别大，成本非常高。所以大家对GPT-6 Astro在操纵机械臂、操纵机器人方面表现的解读，
其实会更多归结于未来具身智能实现的时间可能会被大大提前
。因为此前产业界跟学术界虽
然有不一致的看法，但很多都会觉得是不是要5~10年才能看到具身机器人商业化大面积铺开；但
现在不一样了，由GPT-6 Astro这个表现出发，大家可能会普遍把这个时间从5~10年提前到两到
三年。那么可能在未来很快的时间里，我们看到机器人在各个行业、各个方向上的一些落地。因
为这个需求是完全存在的，完全不用疑虑：无论是ToB进工厂，还是ToC进家庭做家务，
这个需求其实都是真实存在的
，只是说目前机器人特别是大脑端的能力能不能达到、能不能覆
盖掉这个需求。所以这是他最深远的影响，就是整个行业可能会迎来一个加速式的发展
。
近期，宇树的创始人王兴兴上周在杭州开了物理AI大会，会上也发表了他的论断，认为到了
GPT-7这个阶段，可能就会击穿具身智能。这是近期为什么有这么好的行情表现，就是由于
OpenAI布局的催化。
再看其他大厂的布局。说到大厂，不得不提Anthropic，它在具身智能方面的布局，我们
认为目前其实还是在真机实验，以及一些研究人员和系统评测的投入当中，因为此前Claude推出
了一个评测，覆盖了机械臂、四足以及人形机器人，也覆盖掉了像宇树机器人的一些操作，这是
另一个大家所关注的技术布局。
谷歌：有DeepMind部门，此前发布了几代具身智能的机器人大脑模型，特别是在VLA层面，
是一个大家关注度非常高的成果。在其他方面，谷歌在基础模型、机器人数据以及商业化端也是
同步推进的。

英伟达：它的布局此前我们有一期专题也写过报告汇报过，就是借英伟达开GTC大会的时候也汇报过。英伟达在机器人方面的定位依然是一个基础设施提供商，在机器人时代，英伟达希望大家也用他的卡。目前大家用得最好的机器人芯片，还是Jetson（Orin）这个平台，此前也主要是用英伟达的机器人芯片。模型层面，它推出了GR00T系列的模型；世界模型层面，也推出了Cosmos系列。整体仿真也是一个代表性方向，因为无论是此前的运动控制层面，基本都用英伟达的仿真器。从仿真平台到各类模型和数据，英伟达都提供了一个全套。所以在未来机器人时代、物理AI时代，英伟达还是扮演一个基础设施提供商的角色。

国内的话，重点看两家代表性的大模型公司：DeepSeek和Kimi。DeepSeek最近其实也有很多消息传出来，梁文锋在接触具身智能的公司，也投了一家；此前在宇树上市的时候，他也参与了宇树的战略配售。所以DeepSeek可能从一个大模型公司往具身智能方向去发展，因为投资跟合作已经证明了它的主动布局。Kimi此前在9月份也传出正在招聘相关的人才，说明公司其实在这方面也有布局的动作。

以上是近期一些巨头的布局情况，最值得关注的就是OpenAI的动作，也是由于OpenAI的动作引发了这一轮机器人的行情。

那么接下来，我们想说最近无论是一些大厂也好，还是机器人的创业公司也好，他们发布的一些模型进展，其实在很多层面上都证明了机器人scaling law的存在。机器人scalinglaw的存在，就会使得无论是具身智能公司还是大模型公司，未来都会进一步去投入数据，因为这证明了数据量的提升会让模型效果更好，可能在一定的阶段，千万或者更高的小时数，就会跨越基点，使得机器人真正的商业化得以实现。

这里面我们想提到两家公司。一是Generalist，它已经发布了自己的模型，从0到1.5，预训练数据从27万小时增加到50万小时以上，平均成功率由64%提升到99%。当然，它这个更多是自己公司发布出来的数据情况，但在一定程度上证明了数据规模上一个台阶，任务的成功率也会提升。

另外就是数据采集方法的变化。此前整个数采的方式，更多是摇操作，就是穿一套数采装备，然后去摇操一个真机，得到这个数据，成本非常高，通常也是用于机器人的后训练，也就是微调，用机器人的真机数据来操作。但是在预训练层面，其实是需要海量数据的。这个海量数据此前行业大致分成两类：一类是仿真派的代表，认为真机数据成本太高、效率太低，如果用真机数据去做预训练，可能是遥遥无期的。这也是为什么此前很多判断具身智能商业化时代的到来要到10年甚至之后，就是觉得预训练所需要的数据量庞大，用真机数据其实是遥不可及，成本又高、投入又大，效率也非常低。

那么变化发生在哪儿呢？变化从去年到今年，就是在预训练阶段用无本体数据。无本体数据目前主要是两类：一类是Ego数据，一类是UMI数据。

先来看Ego数据。今年2月份英伟达发布了EgoScale这篇论文，用大规模Ego数据进行预训练，用了两万小时。所谓Ego数据，就是人类第一视角的视频数据，核心就是你头上戴一个摄像头，有时候两个手腕处也带一个摄像头，把你如何干活拍下来，用这类数据用作机器人模型的预训练。在今年2月份这篇论文上，已经证明了用2万小时的Ego数据进行预训练，再用非常少量的50个小时人类摇操作数据，再加上4个小时机器人自己运行的真机数据，用少量数据去做对齐，就得到了非常好的效果。一个典型的案例是，在一些任务中，比如叠衣服这种比较简单的任务，在one-shot（只见过一次）的场景下，取得了80%以上的成功率，也就是88%的少样本迁移成功率，这是今年2月这篇论文提到的、用大量的Ego数据所得到的训练机器人的效果。

后续还有一个延续的工作，我们觉得非常有意义，在这里也讲一下。今年8月，这家公司也是有几个从英伟达出来的人创立的创业公司，已经发布了自己模型的二代，用了百万小时的Ego数据，也就是百万小时的人类视频数据，得到了非常好的效果。随着数据量的提升，可以看到它在14项真机任务中的得分逐步提升：1000小时的话得分在20%，到100万小时得分提升到了53%。这就证明了机器人scalinglaw的存在。

所以其实在今年，特别是从二季度、三季度以来，我们已经观察到机器人行业数据采集的爆发。这个数采，更多层面不是摇操作数采的爆发，而是Ego的爆发，也有一些UMI数据，主要还是Ego数据的爆发。Ego数据最核心其实就是几个摄像头、三个摄像头，现在已经把触觉手套、触觉数据也加进去。触觉数据加进去之后，已经有一些模型厂商反馈，加了触觉数据之后模型效果也有很好的提升。所以我们觉得后续除了Ego的数据设备、数采设备跟数采服务之外，触觉数据也是非常重要的。

这是在上一期汇报的基础上，想谈一下最近发生的变化，就是大厂的布局以及机器人的scaling law。那么在这种情况下，投资机会有哪些呢？其实跟我们上次汇报的观点是大差不差的。目前大家看到了数据基建的爆发，这次也想跟大家汇报一下，我们结合最近的调研、研究工作，对于数采设备以及数采服务这个市场有多大，去更好地摸排了一下。

还是老生常谈，把基础情况再跟大家汇报一下。可能有些投资者是最近新接触这个行业，我们把具身智能的数据分成金字塔的五层结构。

金字塔的顶端，就是成本最高、效果最好的机器人真机数据，这个数据一小时比如要卖到大几百块、上千块，这都是有的。这类数据我们认为后续应该都是具身公司自己去搞，因为这个数据会用在训练模型上，目前主要还是在后训练，就是对齐，这块一定要用到真机数据。因为这个数据成本比较高，而且这两年一直在谈的一个问题就是数据的异构性——如果不是用自己本体获得的数据，效果肯定不是最好，所以最好是用自己本体的真机数据。所以这个未来可能是机器人本体公司自己的。当然现在会有一些产业联盟搞一些数据集，可以把这种高质量数据放在数据集里面，谁去用就去付费，获得行业里的一个高质量数据集。目前也是因为此前我们国家的数据量比较少，但近几年，包括我们的国家层面等等，也构建了我们自己的数据集。这是第一层。

第二层就是无本体数据，分成UMI数据和Ego数据两类，这两类待会儿会展开再提一下。再往下就是仿真数据，仿真数据肯定也是有意义的，就跟自动驾驶一样：首先仿真数据可以低成本、无限制地获得，这是它的特点；另外，在一些真实环境不太好模拟到的极端数据上，用仿真数据也可以。但仿真数据的问题就在于sim-to-real的gap，目前肯定是无法跨越的。我们看到有一篇论文里面其实提到，在仿真环境里成功率在80%多，但如果迁移到真实环境里面，大概就只剩下20%多，这就是sim-to-real gap的问题，目前现实情况还是无法解决。所以仿真数据在一定程度上还是有其意义的。再往下一层就是通用数据，就是互联网上的海量视频数据或者文本数据，去提供一些语义理解跟推理先验。但这类数据因为缺乏真实的接触性信息和动作信息，不过它是海量的，比如训练我们的大语言模型、多模态大模型，是海量这一类数据；但去训练我们的机器人，还是需要金字塔上层的这些数据存在。

接下来把各类数据大致的情况具体做一个介绍。这里还是想介绍一下无本体数据，因为其实这一波真正爆发的是无本体数据。UMI数据，从前年开始看到行业里这种UMI数据已经越来越多了，它其实就是一个手持的、有一个末端的操作器，一般就是一个夹爪，然后带着摄像头，以及一个姿态跟踪模块，去执行动作，比如去洗个碗、拿个东西，它就会首先把你这个运动拍下来，另外一个就是有轨迹的信号，有些可能末端也会加一个触觉信号，然后它可以适用于任何本体，它去做一个逆向的重定向，就可以转换成机器人的轨迹，适用于去训练任何本体。Ego数据的话，刚刚也提到了，其实就是目前几个摄像头，再加上触觉手套，或者有的也用机电手套这种模式，去做数据的收集。这一类数据其实就是目前正在爆发的数据。

这个数据介绍完之后，还有两个问题：未来到底需要多少数据？数据配比怎么样？关于数据量的问题，我们也去参考了行业里的大厂，或者行业里的一些头部公司给出的指

引。今年其实有一些公司，比如智源旗下的蜜蜂，它的目标是750万到1000万；京东也有一个比较高的目标，今年想要增加500万；然后还有其他一些公司会有百万级别的目标。所以今年加起来，我们觉得今年中国市场上都能有2000万以上，或者2500万这么一个数据的需求量。

未来需要多少呢？也是参考智源旗下的逆风给出的未来的量级，就是百亿级别小时数，当然这是一个存量的概念。所以我们根据这个假设，如果到2030年达成百亿小时的数据目标，按照每年提升去做假设。如果说这个会提前，比如未来两到三年，比如2029年，那么有可能我们测算的表格有点变，那么当年算下来的需求量就又会增加。这是关于数据量的问题，未来可能是百亿级别的数据量。

关于数据配比的问题，目前确实并没有明确，我们也去拜访调研过很多机器人的模型公司、机器人的本体公司。我印象中有一家非常头部的、首席科学家层面的人谈到过，在他的层面当然希望都是最好的、用金字塔顶尖的那个数据，但是不可能，因为像无本体数据，其实是效率跟成本最终折中下来的一个选择。如果不考虑效率、不考虑成本，那大家都想用最好的真机数据，但是不可能。所以目前来说，这些无本体数据的数据配比至少是一半以上，甚至可能
60%以及更高的配比，用无本体数据去做模型的训练。

所以我们在这里也做了一个测算，其实是关于无本体数据的测算，得到的数据服务的体量会更大一些。这里面的结论是：如果未来要达到百亿小时、年新增在几十亿小时的量级，我们认为数据服务当年的体量有可能达到千亿市场规模。

关于数据产业链的情况，目前做的公司也越来越多了。像7月份、8月份，我们再去世界人工智能大会和世界机器人大会上去调研的时候，也看到好多家都推出了他们的数采方案，以及无本体数据的一些解决方案。所以目前确实是发展初期的一个行业，但确实也在爆发，玩家也涌进来很多。

当前在数采服务这个市场上，全球有ScaleAI，这家公司估值大致在几亿美元这个体量；如果去对标全球给大模型公司提供数据服务的领先公司，那么可能还有很多倍的差距，Scale AI的估值水平在200多亿美元这个级别。目前还有另外一家公司，估值大致在十几亿美元的水平，都算是目前产业的一些头部公司。目前其实一级市场的估值水平并没有特别高，跟大模型领域的公司还是有数十倍的差距。国内的话，像做得比较靠前的简智、路明、光轮这几家，目前还是排名比较靠前，各自有各自专攻的方向，算是在国内目前排名比较靠前的几家数采服务商。

设备上，目前有一个特点，就是数采服务商也会做一些自己的设备，做设备的有些也会往下游服务去延伸。但设备上的玩家更多，因为刚刚也提到，如果是Ego数据的话，最核心就是3个摄像头，然后再加触觉手套。所以做摄像头的其实都可以做，或者说对摄像头的要求不高，采购一些，然后也可以做出来。触觉手套的话，所有之前可以做触觉传感器的公司，都可以去做触觉手套。所以这些公司其实都发了相应的产品。我们当前的观点是，如果未来这个市场确实要
爆发，那么一定会考验设备端的标准化能力跟量产能力。

如果在这个设备端、Ego层面，摄
像头这个层面的话，我们还是会看好奥比中光。虽然它是7月份发布的数采产品，但这家公司是实打实会有产品的标准化能力以及量产能力的。触觉手套方面，目前也有很多家公司发布，可以去看后续跟一些下游大模型厂商以及机器人厂商合作的情况。上市公司领域，包括像福莱新材、日盈电子、汉威科技等等此前做传感器的公司，其实都有相应的产品。这是数采相关的。

接下来还有两个方向也想跟大家提及一下。

一个是ToC。To C的话，我们建议继续关注上纬新材，它在9月22号发布了两款机器人的新产品，Q1和T1。最近也提到他们的累计下线，其实才刚过了一个月的时间，累计下线就已经5000台了。对于明年这个产品的放量，我们觉得还是非常值得期待的，因为这个定价也不高，Q1的话就1.99万，而且公司其实在初期就开始配套开发了很多的可玩性，包括3D打印的外装，包括内置一些教育模块去学习等等，对于一些中高端家庭来说，还是非常有吸引力的一类产品。如

果明年如期卖到数十万甚至几十万的量，我们认为在具身智能领域里面，它的出货量将会
排名第一。

另一个就是垂类应用，我们近期重点在推荐杰克科技
，今天这个时间正好杰克正在开它
机器人的产品发布会，我们会继续看好这家公司。我们看好的核心逻辑，其实在于我们看好这个时间点，一些垂类应用方向会开始起量。
我们去看垂类应用这个方向，其实也是在7、8月份的世界人工智能大会和世界机器人大会
上，看到了一些垂类应用方向的进步。当时观察到的场景就是物流分拣：像星动纪元、自变量等这些公司，其实都已经达到了准确率超过99%，效率已经接近或者超越人工，完全可以到了一个可以替人的阶段。物流分拣这个，包括物流这个行业，也会被认为是物理AI率先会落地的行业。刚提到的这几家公司，其实都是一些非上市公司的创业公司。落到上市公司层面，我们建议可以关注一下杭叉集团，公司昨天开了三季报的业绩会，重点提及了他们在叉车单车智能方向的一些布局，以及未来在物流场景里面去做更多智能化布局的事情。公司质地也非常好，是叉车的一个龙头公司，估值也比较低，所以这是一个可关注的标的。
上市公司层面可以关注的另一个标的，我们认为是中邮科技
，这家公司其实是一个应用
方，作为中国邮政旗下的硬科技平台，可以在应用端利用现在已经开始突破的技术，首先给中国邮政体系内去做到降本增效、获得收益，另外也可以拓展到中国邮政体系外的市场，也是可以去关注的标的。
除了物流这个场景之外，我们找到的另外一个场景其实就是杰克科技了。
杰克作为一
家做缝纫机的公司，质地也非常好，发布的机器人目前在跟模板机结合、跟订扣专机以及锁眼专机相结合方面，准确率也已经完全没有问题，效率也逐步在提升，其实是可以去实现在纺织服装这个行业里面机器替人的目标。我们知道纺织服装这个行业目前的工人人数还是非常多的，全球将近1亿的纺织工人。随着特别是中国目前想找一个有几千人的纺织工厂已经很难了，缺人缺劳动力还是一个现实存在的问题，所以在中国市场上是有实在的机器替人的需求。这个行业在往东南亚转移，往东南亚转移的话，对于全球龙头的服装品牌去选择合作的代工厂，也是希望看到他们装备能力的提升。所以在这个情况下，如果有杰克的新一代AI产品以及机器人产品，也会助力这些服装代工厂去绑定他们的下游品牌客户。所以我们非常看好它后续AI缝纫机跟机器人的
放量。具体的量，大家如果有更关心的，也可以会后跟我们取得联系，去做更深入的讨论。
今天我们所汇报的机器人的一些更新，特别是机器人的scaling law，包括中国的大模型巨头在具身智能机器人方向的布局，就跟大家汇报到这里，感谢大家的收听。

--- 物理AI催化汽零大涨，后市如何展望？20261028.pdf ---
全文总结：
汽零板块调整复盘：板块自去年9月下跌至今近一年，主要经历两个阶段——对2026年传统主业业绩承压的担忧，以及对2026年机器人出货量预期的下修，导致估值持续承压。
当前反弹核心原因：机构对汽零板块持仓比例极低，叠加物理AI催化，板块指数当日上涨4个多点，机器人板块涨幅居前。
传统主业判断：国内Q3同比下滑近20%，批发销量尚未转正，略弱于此前预期；但龙头已现拐点，比亚迪7、8月已实现同比转正，9月同比增速爬升至20%左右，判断2026年上半年为主业底部。
出口维度更乐观：2027年4月起出口退税率由13%降至9%，有望推动一季度抢出口；若出口占比从今年40%提升至五六成以上并保持双位数增长，明年批发销量同比转正预期较大。
机器人产业趋势：物理AI有望在1至2年内迎来重大突破；国内外龙头交付量快速增长，中国龙头占据全球人形机器人出货量大部分份额，海外龙头出货量仍很少，中间存在预期

差。
配置观点：综合业绩支撑、机构持仓偏低、物理AI催化及2027年出货量放量预期，当下配置汽零板块位置不错，物理AI上游零部件机会值得重点关注，国内出货量较大的龙头及国内零部件也有反弹机会。
会议实录：
各位投资者大家晚上好，我是开源汽车邓建全，今天晚上给大家更新一下今天整个汽零大涨，我们整体基本面的变化以及观点。
首先，整个汽车零部件板块应该说已经跌了很久了，大家都能看得到，从去年9月份跌到现在，已经跌了快接近一年的时间。复盘来看，下跌的时间维度主要分为两个阶段：第一个阶段是大家对于整个2026年传统主业业绩承压的担忧，因为从2025年下半年开始，整个内需开始有点环比走弱，加上2026年有一些扰动因素，比如新能源购置税开始减半增收，大家就对于主业的业绩比较担忧。第二个阶段是对于整个2026年机器人出货量的整体下修，也导致了汽零板块估值承压。因为大部分市场还是按照像特斯拉年化、龙头100万台的出货量去预估的，而从2026年年初到现在的周度排产来看，离我们刚刚讲的市场预期还是比较远的。所以在这两个因素共同作用下，整个汽零板块调整得比较多，幅度也比较大。
当下整个市场对于汽零的持仓比例非常低，这也是今天整个汽零指数涨得比较多的核心原因，即机构整体持仓比较低。今天整个汽零指数涨了大概4个多点，其中机器人板块涨幅比较大。

所以我们当下很多领导也比较关心怎么看接下来的行情，我今天也给大家讲讲我们整体的观点。第一个，还是围绕我们刚刚讲到的两个核心变量：一个是传统主业，一个是机器人的出货量对汽零业绩增量的弹性，我们还是围绕这两个核心变量给大家做分析。
先说传统主业。首先，我们大部分人还是觉得国内需求今年比较差，从实际数据来看，国内确实如此，包括Q3整个国内还是同比有接近20%左右的下滑，也导致批发销量还没开始转正，当然这也比我们今年上半年对于下半年批发销量转正的预期弱了一点，主要是国内整体在Q3的表现还是有一点点弱。那么当下再去看整个2027年，首先国内这边还不好判断，因为国内涉及到的变量，包括2027年以旧换新的补贴力度，这些因素现在还没有完全清晰，所以对国内的判断现在来看还有一点点不确定。
但是我们觉得2027年出口这个维度也有一些变化，包括2027年4月份开始出口退税从13%降到9%，这个变化也会推动整个2027年一季度出口量预期出现比较明显的抢出口情况；又加上市场可能也会预期2028年以后可能会有进一步退税的情况。所以对整个2027年，我们认为当下去看出口的表现，我们相对来讲是比较乐观的。但是从我们了解的情况来看，市场目前大部分预期还是比较弱的，大部分人可能会认为今年的出口增速比较高、基数也比较高，所以整个2027年出口增长的确定性，现在很多人还是相对比较悲观。但是从我们最近的草根调研来看，来自出口的增量应该还是可以用更加积极的预期去看的。所以我们觉得明年出口如果能够从今年的40%占比提升到百分之五六十以上，并且还能够保持双位数的增长，那么明年批发销量能够同比转正的预期应该还是比较大的。
所以我们觉得首先主业今年见底的确定性相对来讲还是比较高的，从龙头也可以看得出来，像比亚迪从今年七八月份已经开始实现同比转正，到今年9月份同比增速已经爬到了20%左右的增长。龙头的增速也预示着整个行业从开发商角度来看转正的概率也比较大。所以我们认为，2026年上半年主业这边的承压应该相对来讲是一个底，接下来2026年Q3、Q4以及2027年，我们判断向上的概率也比较大。从最近有些零部件披露的Q3业绩也可以看得出来，有些公司已经开始走出下滑的态势，例如中联安全这类公司已经开始同比出现比较好的拐点。当然大部分公司也有这种情况，有一些来自份额提升，有一些来自AIDC的增量，有一些来自机器人的增量等等。所以我们觉得从业绩角度看，2026年上半年应该是一个相对比较低的位置。
机器人这边，是这一轮物理AI催化的核心逻辑点。催化的核心原因是，在最近一两个月，特别是十月份，有几个世界物理AI大会，一个是在南京举办的，一个是在10月22日榆树举办

的物理大会。在这些大会上，也提出了一些行业趋势的判断，例如语数科技也提出，物理AI大概率会在一到两年之内迎来比较重大的突破。所以这些产业领导给出的判断，也是当下估值的一个催化。

从量的维度来看，整个行业不管全球龙头不断地爬产，还是国内在2026年迎来继续往上提的态势。像国内的语数，今年上半年整体交付量同比增长速度也非常快，全年大概率能有翻倍以上的增长；去年全年人形机器人出货量大概在5500台，今年上半年就已经达到去年的水平，5900台，所以全年翻倍的难度应该不是很大。除了语数之外，另外几个龙头，譬如智元，去年全年大概也就三四千台，今年上半年出货量已经达到8400台，所以全年大概率也是大几百的增长。譬如优必选这些，整体来看交付量也非常不错，今年上半年全尺寸的人形机器人交付同比去年也大幅度提升。所以我们觉得，不管是全球龙头的排产，还是国内几个龙头的交付量提升，从数据
来看整个行业的趋势还是在不断演进的。
当然，当下机器人主要应用的场景还是一些偏商业类的为主，例如工业场景、应急场景、城市治理场景，相对来讲是一些特定、重复性比较多的应用场景为主，未来大概率也会逐步从工业场景往更多场景去提升覆盖面。所以我们觉得从量的维度来看，也是支撑这一次反弹的比较大的核心。
其中有一个数据也供大家参考，因为过去炒机器人的时候，其实很多领导还是盯着全球龙头，但在这一轮是否有预期差，我们觉得也是值得大家关注的。因为从2026年整体全球人形机器人交付量来看，中国这些龙头占据了大部分份额
，国内刚刚讲的那几家，语数、智元、优必选
这些，应该占据了全球大部分的人形机器人出货量；海外的龙头，不管是特斯拉、Figure还是波士顿动力，他们整体在今年上半年的出货量还是非常少的。所以从绝对份额的角度来看，中国的企业应该是遥遥领先的。所以我们觉得，未来炒物理AI终端应用这个维度，是否一定还是盯着海外龙头的技术演进来看，我们觉得也可能会存在一定的预期差。因为毕竟机器人本身也是支架领域技术的复用，虽然支架这个维度我们过去可能跟得比较紧，但在基建这个维度也不排除能够反超的可能性。所以我们觉得从产业维度来看，也是能够看得到的，这一轮反弹存在着中长期产业趋势反弹的支撑。
所以我们觉得综合来看，当下机器人板块包括物理AI衍生出来的整车以及零部件这些方
向，整个基本面还是向上的
。包括刚刚提到的政策，其实大家可以看一下业绩，像比亚迪二季
度的业绩已经开始拐头往上走了。所以我们觉得有业绩支撑，加上机构整体持仓也比较低，加上在物理AI应用领域的催化，大家展望整个2027年的出货量依然能够比较快速地放量。所以我们觉得当下去看整个物理AI上游零部件带来的机会，也是值得大家重点关注的。这是整个板块的逻辑点。至于个股，因为时间有限，我就不一一给大家展开了。当然从这两天的盘面来看，零部件大部分都迎来比较大幅度的反弹，除了机器人相关的之外，很多汽车零部件反弹的幅度也不错。我们觉得之前可能大家核心盯的都是谁拿了哪些数量、或者拿了哪些定

点，当下去看整个零部件板块，我们觉得也不仅仅可能是这些提到的标的有机会，国内出货量比较大一点的国内龙头，也有可能带来国内零部件一定的反弹
。综合来看，我们认为从业绩
也好、从行业趋势也好，当下去配整个汽零板块应该还是一个不错的位置。
以上就是我今天给大家汇报的核心观点。至于个股，如果感兴趣的领导，也欢迎跟我们私下进一步交流。以上是我今天汇报的内容，感谢大家的收听。

--- 物理AI专家对话-数采会是下一个风口吗？20261027.pdf ---
全文总结
本次电话会由中邮证券机械组组织，围绕物理AI与数据采集赛道邀请上海德麦肯总经理张总、德玛科技AI总监尚总进行交流。核心讨论了物理AI与数据采集的最新产业变化、数据采集的

几种模式及优劣势、无本体数据采集的成本与效率、模型与数据的关系、竞争格局与壁垒，以及德玛在具身智能方向的布局。

一、物理AI与数据采集的产业变化

对具身智能产业来说，数据和模型一直是相伴相生的两个胜负手。从去年开始，具身智能从实验室走向生活中，主要集中在商业表演领域，但大家更关注的还是真实物理世界的操作。对操作而言，现在最大的短板就在于数据；

行业里有几种采集模式：

第一种是基于本体（有本体）的操作，即操作人员佩戴传感设备，远程遥控操作机器人本体，在真实场景中采集数据。这种技术最成熟，但缺点也比较明显：最大的问题是需要先购买和部署大量真机，导致成本比较高；同时需要聘请专业的数据采集人员佩戴设备采集，并付出比较高的采集费用；

第二种是通过构建仿真环境，在虚拟环境中生成仿真数据。仿真数据最大的问题是，它生成的数据在进行模型训练时，会带来从仿真到真实环境的差距，行业里称为sim-to-realgap；

第三种是这两年兴起的无本体数据采集，主要由美国MIT等顶尖高校提出的论文，以及在此基础上带来的新技术。它最大的好处是不需要部署真实的机器人本体，通过简单的夹爪、或者佩戴头环，就可以相对简便地采集数据；还有一点是这种采集方式不需要搭建采集环境，主要是采集人员在真实的工作场景中采集真实的人类数据。目前行业里主要就是这么几种数据采集方式。

二、今年以来数据需求的边际变化

特别是从今年三季度以来，能明显感觉到对物理AI数据的需求，从原来单纯的具身模型公司需要数据做训练，转向通用大语言模型公司也开始寻求物理AI的数据来提升整个模型的能力；

数据、模型加上算力这三个要素，是整个AI行业的三驾马车或三个基座。算力主要由算力卡主导，包括英伟达及国内硬件公司；对模型公司来说，现在最稀缺的就是数据；

之前的数据采集成本和规模效率都非常低，随着今年无本体数据采集模式兴起之后，大家看到了数据规模化非常好的前景，行业里称为scalinglaw。无本体数据采集一是采集成本比较低，二是采集规模和效率非常高，使数千万小时甚至上亿小时规模的批量数据供给成为可能；

像OpenAI、国内的通用大语言模型公司，在模型到了一定天花板之后，也开始从纯语言模型向多模态图像、再进一步向物理AI模型演进。所以从今年三季度开始，市场上对数据的需求迎来了一个爆发。

三、不同场景的解决方案与落地优先级

像德玛现在做的物流领域机器人，大脑可能是具身模型，再配合一些应用场景偏技能包的数据采集和训练，打造成比较实用的机器人。面对更复杂的场景，可能还需要物理AI大模型层面的一些突破；

模型的能力主要受限于算力，另外最大的瓶颈就是数据。数据的短缺导致模型对通用场景的泛化操作能力现在还没有到能够真实落地的阶段；

在与主要模型公司交流中，模型公司有两个诉求：一是具身智能需求非常旺盛、有付费能力、大家能算得过来账；二是在这些场景里还存在大量的简单重复性劳动。符合

这两个条件的行业梳理下来，可能家庭相对复杂一点，因为家庭场景最多变；工业和
物流这两个场景，是现在模型厂商普遍看好、基于当前模型能力和数据条件优先迎来
具身智能落地的。德玛的数据采集目前也集中在工业和物流这两个场景，家庭场景也
有在采。家庭场景最大的好处是数据供给相对充足，可以通过德玛构建的社会众包采集网
络，由大量家庭主妇或居家人员佩戴采集设备在家里采集；工业和物流则与自身资源禀
赋有关——德玛本身有非常好的物流生产场景，已在德玛自己的工厂里部署了第一人
称无本体数据采集设备，采集工人真实的劳动场景，并通过自建的数据管线打包成规
模化的数据集。物流场景当前还有大量相对简单的空包、上包、分拣、码垛等场景，
需要工人做简单重复性劳动，德玛在这些场景部署了无本体设备采集数据。工业和物
流场景的数据，目前在行业里也是很多模型公司所需要的。

四、模型技术路线：从VLA转向世界模型

模型和数据一直是相辅相成的。过往最早大家押注VLA技术路线，是因为整个行业数据短缺，VLA是当时条件下数据最充足、可以基于语言发展出来的技术路线；
随着新的数据路线进一步明确，规模化数据供给所带来的变化，就是大家可能更多地
从传统VLA转向以世界模型、端到端技术为代表的模型路线；

美国的戴安娜（Diana）在今年8月第一个发布了基于百万第一人称无本体数据的模
型，可以让机器人在厨房、家庭里展现出不错的泛化操作效果；国内像上个月做世界
模型的一家企业（自然意志），也发布了国内第一个基于百万小时数据的端到端世界
模型；
从行业来看，国内与国外的模型代差在进一步缩短，大家都已经从原来的VLA技术路
线转向了世界模型。世界模型路线更强调端到端，更强调整体操作的通用和泛化；
除了具身智能公司、通用大语言模型公司，大家都开始陆续把人类真实数据特别是第
一人称无本体数据加入到模型里，在这种物理AI前景上看到了明显的曙光。

五、通用大模型与物理AI的融合、真机数据趋势

通用大语言模型公司和具身智能公司，从不同方向都在朝着最终物理AI的通用目标前进。像OpenAI为代表，以及前两天在宇树大会上王兴兴也提到，他们已经开始把通用大
语言模型部署到宇树的机器上，从原来的信息AI大模型开始转向物理AI方向；
通用大语言模型公司对千卡万卡这种算力的基建操作和使用更加熟悉，拥有庞大的算
力，也能非常好地消化数据。从目前情况看，大家对数据的使用非常乐观，在物理AI方向上都展现出非常好的进展；
关于物理AI大脑能否进化到完全不需要真机训练：终极目标肯定是希望用更高的效率
使用数据。当前技术路线上，大家都在逐渐弱化真机数据
，因为真机数据成本高、采集

效率慢；但当前技术条件下大家对真机数据的使用还比较熟悉，整体上是在逐渐减少真机数据；

像今年2月英伟达的一篇scale论文里提到的，也是从原来100%真机，弱化到10%或20%这样一个相对少的真机数据。近期与国内外模型厂商交流，大家的目标也都是希望把真机数据尽可能降低。现在主要目标是：通过真实世界的无本体数据训练一个基模，再搭配少量真机数据做后训练和对齐，这样可以适配到不同的机器人本体上；

相当于Windows系统可以适配到不同的电脑上，但仍然需要少量的驱动程序，真机数据现在的价值主要在这里，它可以把模型能力对齐到不同的机器人硬件本体上。所以趋势上，真机数据在逐渐减少。

六、德玛的应对与技术布局

德玛一直在积极部署多模态数据，从去年的（真机）训练场到现在的无本体数据采集。今年4月，德玛在国内行业里第一个发布了阿林（Ego）无本体解决方案，包括完整的采集硬件和采集数据管线，是全自动化的数据采集管线；

德玛也在积极针对行业变化布局数据业务，当前已经构建了较为广泛的中保（众包）采集网络，努力构建真实场景的数据集，与行业技术趋势紧密同步，而且是国内少有的能够部署全模态数据——从真机数据到无本体数据、再到真实场景后训练部署的公司，在这方面布局非常领先；

无本体（无场景）数据即第一人称数据采集：佩戴轻量的头环和手环，在真实场景里由工作人员（如工人、家庭主妇）佩戴设备，采集人类的真实操作数据；

主持人理解：模型发展可预期地会使真机训练和数据采集这一块蛋糕变小，但德玛通过不断强化自身竞争力、实现技术层面进步，去拿到更多份额；

德玛除了真机数据之外，已较早部署第一人称无本体数据；

未来机器人模型的主教材应该是无本体（第一人称无本体）数据，德玛在国内已比较早部署；真机数据相当于维生素，还是需要的，但从规模上比不上无本体数据；

整体数据市场规模，未来持比较乐观的态度。对模型公司来说，研发费用支出主要在两大块：一是算力，二是数据；数据对整个具身智能产业链而言，应该是模型厂商投资的主要方向。

七、真机数据与无本体数据的优缺点对比

真机数据：好处是比较直接，采集的就是机器人的关节运动角度等数据，在模型训练时使用难度比较小。缺点非常明显：采集成本很高、采集效率很低。首先要买一个机器人本体，当前要花费十数万到数十万；第二要雇专门的人在搭建场景里采数据；因为要

遥控操作一个机器人来采，每天能采集的数据规模比较小，一个专业采集人员加一个
机器人本体，每天能采集的有用数据可能就一两个小时。所以经过最近两年发展，真
机（遥操）数据规模并没有达到行业需要的规模化数据；
无本体数据采集：德玛也部署了阿林（Ego）解决方案，对应无本体数据采集。这种
采集方式不依赖机器人本体，不需要购买机器人——人类本身就在生产数据，从第一
性原理出发，人类每天都在生产大量数据，只是没有被记录下来。所以开发了一系列
采集硬件，比如一个头环戴在头上，各行各业的工作人员在操作的时候，头环可以以
人的第一视角记录手部操作和胳膊运动方式，再经过数据管线加工成机器人模型训练
能使用的数据集；
这种方式采集效率非常高：正常一个工作岗位按每天工作8小时算，算上间断或休息
时间，每天能积累5~6小时的有效数据规模，而且这个工作岗位在生活中无处不在，
千行百业都可以佩戴设备采集数据。这种新方式给行业和模型性能提供了源源不断的
数据来源，就是行业里补充的scalinglaw（规模效应）；

最新的技术路线就是无本体数据采集，可以快速积累起数千万甚至上亿小时的数据
集。
八、无本体数据采集的成本结构与设备
无本体数据采集的主要成本构成：一是一次性的数据采集设备，成本在万元级别，未
来有希望降低到千元级别；二是在真实场景里给工作人员少量的采集补贴；
每小时采集成本从原来的数千元每小时降低到数十元每小时，单位时间能降低100
倍。原先部署一个真机遥操，成本非常高，主要是人员工资和设备一次性投入，投入
一台本体和采集设备就要达到40万元；
设备是电子设备，正常使用两到三年不成问题；
随着设备进一步量产、规模效应和研发费用分摊，单纯硬件成本会下降较快。希望构
建一个万人甚至10万人的采集规模；
目前德玛阿林（Ego）解决方案的主要设备，是一个包含多目相机的头环，戴在头上
以人的第一视角记录人的操作过程，通过算法把人的手部关节进行还原，记录手部操
作方式，变成模型可用的数据——这对应人的学习过程，因为人对物理世界的认知超过50%甚至60%都是通过视觉记录的；
未来还需要进一步补齐力觉、触觉等信息。行业里有各种解决方案，比如做视觉手套或力觉手套，也有通过肌肉电极（EMG）方式记录来反推手部力量。目前设备主要是视频、

音频这些数据信息，未来随着产业链进一步成熟，还会补齐触觉、温度，甚至机器人
模型训练能对应的人类感官信息，进一步去对齐。
九、力觉与触觉数据的发展节奏
当前的数据采集是在为机器人模型准备教科书，当前教科书还处在通识教育阶段，所
以主要以视觉为主；
力觉最主要的问题是传感器价格还没下来，像力觉采集手套单副就要数万元，而且可
用性比较差，大概用1到2周就会坏掉，所以传感器件还没有准备成熟；
对模型厂商来说，现在最聚焦的还是让模型具备初步的泛化和操作能力，所以还是以
视觉为主。随着产业和模型能力进一步成熟，就会从通识教育的教材转向专门的教
材，这是产业和行业相辅相成的一个过程。
十、数据规模与模型规模的关系
前期行业里VLA模型或早期的真机数据规模，都在数千小时到1万小时这个规模；

最新的无本体方向，因为采集的数据规模能快速拉起来，通过佩戴简单设备采集人的
手部操作，能快速把数据规模推到数十万小时甚至数百万小时，所以当下比较领先的
模型都会以百万小时的人类数据作为一个行业门槛；
主持人问：明后年随着世界模型发展，能否把这个百万小时降低到十万小时？专家指出，
未来的小时数应该是加大的，可以类比大语言模型的参数规模来理解：现在模型规模
还比较小，只能吃比较少的数据，具备初步的通用和泛化操作能力，它需要的数据规
模应该在上亿小时；
关于针对某个具体场景（如物料搬运）的训练规模：按无本体训练来讲，可能也需要10
万个小时的数据采集和训练；
训练时长与数据规模的概念需要区分：不是"训练时长"的问题，而是需要预先采集更
多的数据，用来做预训练或后训练，对模型进行训练与调整。类比大语言模型，参数
规模会越来越大；模型规模越大，能"吃下"更多的数据，泛化的通用操作能力越强。
未来大模型规模会越来越大、能力越来越强，在工业、物流、生活各种场景的部署进
程会加快。
十一、德玛的场景解决方案与合作伙伴
一个具身智能机器人或通用机器人在一个场景落地，需要：第一，有一个操作性非常
好的本体；第二，有一个能力不错的模型，并有足够的数据去训练它；最后，有一个
非常好的场景去部署这个具身智能；
德玛与行业里整体的本体厂商和模型厂商都建立了非常好的合作，最近已与蚂蚁、宁

波签署战略合作关系。蚂蚁、宁波主要聚焦机器人世界模型、或者说机器人大脑；
德玛对场景有非常好的认识，在场景里积累了大量的数据，在这个数据上与行业里领
先的模型厂商和本体厂商合作。德玛的目标是让机器人真正在场景里能够用起来，定
位是把场景、数据、模型、本体做一个整合，形成非常好的解决方案，借助整个市场
渠道和服务网络，真正在客户场景里落地和应用；
德玛当下聚焦两点：一是采集足够多的数据积累，对场景有更好的理解和数据；二是
以非常开放的态度，与行业里本体做得非常好的公司（如优必选、宇树等）和大脑做
得非常好的公司合作，整合成一个真正能够落地的解决方案；
关于蚂蚁、宁波模型的参数级别，尚总补充：今年年底会发布垂类模型，训练采用的
是多模态、多样性的数据，数据质量非常高，因此需要的数量级并不大。训练出来的
效果是，在分拣、搬运这类比较简单的场景里能达到七八十分的效果，即能开始有一
些简单应用。这就像小孩学走路、学跑步一样，刚开始可以进行一些简单应用，然后再通

过POC客户端部署时，根据实际场景做进一步调整和优化。在部署和做POC的过程中，也
不断积累实际数据和应用。如果模型刚出来时能用大概70分的效果，POC测试完成之
后达到80分甚至更高，才好向客户端部署。所以后续需要的数量，不一定需要自己去
专门采集，客户端也可以不断积累；
从70分到80分需要的数量比以前更高，80到90或更高水平，需要的数量级可能是呈指
数级提升；
关于蚂蚁、宁波合作的模型具体参数信息，需要等到12月份正式发布时再对外说明。
十二、行业竞争格局与壁垒
数据对模型作用非常直接也非常重要。目前数据行业里，参与的除了德玛这类专业数据公司以外，各家的机器人、模型公司也都推出了自己的数据方案，但目前来看主要偏
自用，同时也是积累对数据采集的认知，并把认知反馈给类似蜜蜂这样的专业数据公
司；
整个数据行业格局：最顶层是客户，即模型公司；中间层是行业主流的数据采集公
司；最下层是拥有场景的服务公司，或者原来的人力资源公司，它们能在各种场景里
快速组织起比较多的人力资源，把设备部署下去，实现数据的回流和采集；
数据生意的壁垒，主要在三方面：
一是要对模型有充分的认知，与模型厂商保持比较好的沟通，知道什么样的数据对模型有用。前期出现过看起来数据门槛不高、在场景里佩戴设备快速把数据采上来，但实际上这些数据大部分没有用的情况。所以第一个关键，是要对

模型和数据两者的关系、什么样的数据能对模型的泛化和智能能力有比较好的提升有理解；
二是能够快速把数据采集能力部署下去、扩展开来，实现规模化的数据，达到百万小时、千万小时这样的数据规模。模型公司对数据需求非常巨大，但模型厂商不可能无限制对接很多数据公司，可能还是希望绑定少数几家，实现规模化的数据供给；
三是能够有很多场景，并在这些场景里，根据模型公司的要求，把数据从规模上采集上来。这需要在行业里有比较多的积累或者资源。以德玛为例，目前在工业场景、物流场景本身相对有优势，无论在自有工厂还是物流行业客户方面，都是过往在行业里的积累。
十三、关于世界模型进入后的竞争担忧

主持人担心：如果海外世界模型已经很发达、能用于物料搬运等物流场景，稍加少量技能包和实景演示就能很快学会、实操效果很好，是否有这种可能性；
专家类比大语言模型和自动驾驶：最终模型追求的关键目标肯定是通用和泛化，需要学习各种各样的数据、各种各样的能力，去充分认识和理解物理世界。所以即使到今天，各家大语言模型公司还在追求数据规模，自动驾驶公司也在通过不停路测、甚至用户真实操作来回流数据；
具身智能所需要的数据规模，可能相比自动驾驶和大语言模型要成百上千倍，所以数据需求是一个长期的过程；
在物流和分拣场景里，不仅仅是要具备基础的搬运、上包这些操作能力，也需要去理解面对不同形状、不同材质、不同颜色的各种包裹或对象，对它有一个更好的理解。
所以未来模型对数据的需求，是一个长期的过程；
尚总补充：现在德玛做测试是基于比较简单几个工况，如分拣、上包、码垛。按照模型公司做，出来的效果可能就是五六十分、六七十分。但如果真要实现对人的期待、达到90分、95分甚至更高，刚开始部署达到80多分就OK，但后面达到90多分的时候，是指数级的数据需求。也包括自动驾驶到现在都还没有完全成熟，其实是一个长期的需求过程。模型公司做出来的效果比较OK，但真正达到完全替代、或者比人做得更好，这个时间会是很长期的。而且现在简单几个工况解决了，未来还有更复杂的工况、比较长流程的工作流要面对，数据需求会更大；
关于"同样一个场景达到90分能力"的疑问，专家解释：真正能落地的技术路线和自动驾驶、大语言模型类似，底层是一个神经网络架构，本质是一个调校良好的神经网络，用大量真实数据去训练，让模型学会理解物理世界、理解操作的因果关系，掌握更多操作技能。现在模型的数据量都是一些不那么精细的操作，只是让模型学会一些基本

操作；
类比让小孩从牙牙学语到认识世界，知道推一个瓶子水会洒、推一个桌子桌子能移动，有一个事件反馈的过程。未来希望机器人在真实场景替代人工做非常精细化的工作，这个场景就需要更多多模态数据，所以数据规模会越来越大。并不是说只在物流场景操作就只给物流数据，而是需要各种泛化数据，让模型展现非常好的对世界的理解——可能在其他场景学到的物理规律，能在物流场景理解和运用；未来的基模，模型规模会越来越大。目前阶段受限于算力，提升模型规模大概在数十B，即几百亿参数规模；未来可能像大语言模型一样，规模会到千亿、万亿级别。所谓数据规模，是指需要更大规模的数据量，而不是需要更多的训练时长。

十四、公司经营数据

主持人询问德玛肯的业务数据（年收入、年利润、未来订单和收入指引），尚总表示暂不太方便公开，等年报、三季报之后内部讨论后再看如何分享。

会议实录：

主持人（中邮证券机械组刘卓）：

各位投资者，大家早上好，我是中邮证券机械组的刘卓。欢迎大家参加中邮机械组织的关于物理AI和数据采集的专家电话会。本场电话会，我们有幸请到上海德玛肯（德麦琴）的总经理张总和德玛科技的AI总监尚总来参加。非常感谢两位领导的时间。

我先简单介绍一下，上海德玛肯是目前德马科技的（控股）子公司，也是国内比较知名的、做具身智能机器人数据采集服务的公司。现在对于德马科技的具身智能战略来讲，它的战略地位相当于智元的蜜蜂，是具身智能战略的重要一环，是具身智能未来战略的一个重要版图。

接下来我们正式开始今天的电话会。首先请教一下张总，因为最近我们看到物理AI以及数据采集整个讨论的热度非常高，请张总跟我们分享一下，从产业角度来讲，最近物理AI和数据这一块有什么比较大的变化？

张总：

好的，谢谢主持人。对具身智能产业来说，数据和模型一直是相伴相生的两个胜负手。从去年开始，具身智能行业里从实验室走到我们的生活中，主要是集中在机器人商业表演领域；其实大家更关注的还是之前在真实物理世界的操作。对操作而言，现在最大的短板就在于数据。

行业里有几种采集模式：

第一种是基于本体（有本体）的操作，简单来说就是操作人员佩戴一些传感设备，远程遥控操作机器人本体，在真实场景中去采集数据。这种技术是最成熟的，但它的缺点也相对比较明显，最大的问题是要先购买和部署大量的真机，由此导致成本比较高；同时也需要聘请专业的数据采集人员，佩戴设备来采集，也要付出比较高的采集费用。

第二种采集方式是通过构建仿真环境，在环境中构建仿真的实体，然后在虚拟环境中生成仿真数据。仿真数据最大的问题是，它在进行模型训练的时候，会带来从仿真到真实环境

的差距，行业里称为sim-to-realgap。
第三种采集方式是这两年兴起的无本体数据采集
，主要是由美国MIT等顶尖高校提出的论
文，以及在此基础上带来的新技术。它最大的好处是不需要部署真实的机器人本体，通过简
单的一些夹爪、或者佩戴头环，就可以相对简便地采集数据。还有一点是这种数据采集方

式不需要搭建采集环境，主要是由采集人员在真实的工作场景中去采集真实的人类数据。
目前行业里主要就是这么几种数据采集方式。
主持人：
明白，感谢张总。现在数据整个技术发展也非常快，技术成熟速度也较之前有很大提升。比如今年以来，德玛肯有没有感受到训练数据的订单变化？另外，最近OpenAI也参与到具身智能大脑层面的领域，包括它展示的机器人动作等，可能在大脑层面给具身智能机器人带来非常大的变化。那它的模型使用、以及整个行业的促进，对大家数据采集这一块的热度有没有明显增加？
张总：
是的，特别是从今年三季度以来，我们能明显感觉到，对物理AI数据的需求，从原来单纯
的具身模型公司需要用数据做训练，转向通用大语言模型公司也开始寻求物理AI的数据来
提升整个模型的能力。应该说，数据和模型、再加上算力这三个要素，是整个AI行业的三
驾马车或者三个基座。算力主要是由算力卡主导，像英伟达、像国内的一些硬件公司；那么对模型公司来说，现在最稀缺的就是数据。
之前的数据，主要是前面提到的，采集的成本、规模和效率都非常低。随着今年这种无本
体数据采集模式兴起之后，大家看到了数据规模化的非常好的前景，行业里称为scaling
law。在AI行业里，大家对数据的需求和规模化的需求，一直梦寐以求。无本体的数据采集，一是采集成本比较低，二是采集规模和效率非常高。整个行业对数据的需求，可能是数千
万小时甚至上亿小时的规模，这种无本体的数据采集方式，使得规模化的数据供给成为一
种可能性。所以，不但具身智能公司对新的数据、规模化所带来的模型推动有需求；像
OpenAI、国内通用大语言模型公司，它的模型到了一定天花板之后，也开始从纯语言模
型，向多模态的图像、再进一步向物理AI模型演进。所以从今年三季度开始，市场上对数
据的需求迎来了一个爆发。
主持人：
明白。德玛肯的业务场景覆盖非常广泛，工业场景甚至家用场景都非常多。面向不同场景，你们的解决方案、或者说行业层面的解决方案有什么大的不同吗？比如德玛现在做的物流领域机器人，大脑可能是具身模型，再配合一些应用场景偏技能包的数据采集和训练，打造成比较实用的机器人。面对其他可能更复杂的场景，是不是大脑层面还需要一些物理AI大模型的突破？现在整个行业层面大概是什么情况？
张总：
刚才提到模型能力，主要受限于算力，另外最大的瓶颈就是数据。数据的短缺，导致模型对

通用场景的泛化操作能力现在还没有到能够真实落地的阶段
。在我们跟行业里主要做模型的

公司交流时，模型公司有两个诉求：第一，具身智能的需求非常旺盛，能够有付费能力、
大家能算得过来账；第二，在这些场景里面还存在着大量的简单重复性劳动，这是跟模型
当前发展阶段有关。符合这两个条件的行业，我们梳理下来，可能家庭会相对复杂一点，因
为家庭场景是最多变的；而工业和物流这两个场景，是现在模型厂商普遍看好、能够基于
当前模型能力和当前数据条件，优先迎来具身智能落地的。所以我们的数据采集目前也是
集中在工业和物流这两个场景，当然家庭的场景我们也有在采。
家庭场景最大的好处是数据供给相对会充足一些，我们可以通过德玛构建的社会众包
采集网络，有大量的家庭主妇或者居家人员，通过佩戴我们的采集设备和在家里采集。而
工业和物流，这跟我们自身的资源禀赋也有关系。德玛本身有非常好的物流生产场景，我
们已经在自己德玛的工厂里面部署了一些第一人称的无本体数据采集设备，在这些场景里
面可以采集工人真实的劳动场景，把这些场景通过我们自主构建的数据管线系统，打包成
规模化的数据集。物流场景现在最大的好处是，在当前这个阶段里，还有大量相对简单的、像
空包、上包、分拣、码垛这些相对简单的场景，这些场景目前还是需要大量工人去做简单重复性
劳动，我们也是在这些场景里面部署了我们的无本体设备来采集数据。工业和物流场景的
数据，目前在行业里也是被很多模型公司所需要的。
主持人：
明白。现在像海外那些，比如OpenAI做的Astra，还有英伟达、谷歌搞的那些模型的路线，咱们
怎么看？它们的优劣势如何？现在感觉哪一种可能会比较快跑出来？
张总：
模型和数据一直都是相辅相成的
。其实过往最早大家押注VLA技术路线，也是因为整个行业里
数据短缺，我们选择了在当时条件下数据最充足的、可以基于语言发展出来的VLA技术路线。那
么随着新的数据路线进一步明确，规模化数据供给所带来的变化，就是对模型的变化——
大家可能更多地从传统VLA，转向以世界模型为代表的、世界模型和端到端技术为代表的模
型路线。所以大家能看到，美国的戴安娜（Diana）在今年8月份第一个发布了基于百万第
一人称无本体数据的模型，可以让机器人在厨房、家庭里面展现一个不错的泛化操作效
果。国内的企业，像上个月国内一家做世界模型的公司，也发布了国内第一个基于百万小时
数据的端到端世界模型。从行业来看，国内与国外的模型代差也在进一步缩短，大家都已

经从原来的VLA技术路线转向了世界模型。世界模型路线更强调端到端，更强调整体操作的
通用和泛化。所以整体上从国内国外来看，除了具身智能公司和通用大语言模型公司，大家
都开始陆续把人类的真实数据、特别是第一人称的无本体数据加入到模型里面，让我们在
对物理AI的前景上看到了明显的曙光。

主持人：
明白。像OpenAI搞的通用大模型用在机器人身上感觉弄得还挺好的，那后面会不会又带来一些技术上的变革？跟您再请教一下。
张总：
应该说，这种通用大模型公司，大家从不同的方向都在朝着最终的通用物理AI这个目标在
前进。从我们现在看到的一些情况来看，像OpenAI为代表，前两天在宇树的大会上，宇树的
王兴兴王总也提到，他们已经开始把通用大语言模型部署到宇树的机器上。应该说，从不同
的路线上，大家都开始从原来的信息AI大模型，开始转向物理AI这个方向。那么通用大语言模型公司，对千卡万卡这种算力的基建操作和使用会更加熟悉一些，他们拥有庞大的算力，
也能够非常好地消化数据。对于现在看到的情况来看，大家对数据的使用还是非常乐观的，所以在物理AI这个方向上都展现出非常好的进展。
主持人：
明白。物理AI的大脑，它是能进化到根本就不需要真机训练这种地步吗？还是也绕不开真机训练？
张总：
最终的终极目标，大家肯定都是希望能够用更高的效率来使用数据。当前的技术路线来
看，大家都在逐渐弱化真机数据，因为真机数据的缺点刚才也提到了，成本比较高、采集
效率比较慢。但在当前的技术条件下，大家对真机数据的使用还是比较熟悉的，所以整体
上是在逐渐减少真机数据。像最早今年2月份英伟达的一篇scale论文里提到的，也是从原来的100%真机，现在弱化到10%或者20%这样一个相对少的真机数据。那么近期来看，我们跟国内外的一些模型厂商交流，大家的目标也都是希望把真机数据尽可能降低。它现在的主要目标是
通过真实世界的无本体数据来训练一个基模，然后再搭配少量的真机数据来做后训练和对
齐，这样可以适配到不同的机器人本体上
，它主要作用是在这个地方。相当于打一个简单的
例子，我们原来的Windows系统可以适配到不同的电脑上，但仍然需要少量的驱动程序。那
么真机数据现在的价值主要是在这个地方，它可以把模型的能力对齐到不同的机器人硬件
本体上去。所以趋势上，真机数据在逐渐减少。
主持人：
明白。那您的意思是，模型的发展趋势是在努力减少真机训练。这样的话，我们的业务对应到真机训练这个层面，您觉得如何应对这样一个可能对我们不利的趋势？

张总：
德玛一直在积极部署多模态的数据。从去年的（真机）训练场到现在的无本体数据采集，
应该说我们在今年4月份，在国内行业里第一个发布了阿林（Ego）无本体解决方案，包括
完整的采集硬件和采集数据管线，它是一个全自动化的数据采集管线
。我们也在积极针对行
业的变化来布局我们的数据业务。当前我们已经构建了一个较为广泛的中保（众包）采集网
络，努力构建真实场景的数据集。这是跟行业技术趋势紧密在走的，而且是国内少有的能
够部署全模态数据——从真机数据到无本体数据、再到真实场景的后训练部署，德玛在这
方面布局非常领先。
主持人：
明白。您说的那个无场景数据是指哪方面？
张总：
无场景数据就是我刚刚提到的第一人称数据采集。我们佩戴轻量的头环和手环，在真实场
景里面有工作人员，比如工人、家庭主妇，去佩戴这个设备，去采集人类的真实操作数
据。
主持人：
明白。我可不可以这么理解：因为模型的发展，可能可预期的真机训练和数据采集这一块蛋糕会变少，但我们的竞争力也非常（强），就是通过不断强化自身竞争力，实现技术层面的进步，然后拿到更多的份额。可以这么理解吗？
张总：
对。应该说，除了主持人刚刚提到的真机数据之外，德玛已经比较早地部署了现在的这种
第一人称无本体数据。
主持人：
明白。那无本体数据和真机数据，还是在训练层面。您觉得后面的模型训练，虽然说无本体的训练也绕不开，但后面整个真正实现真正意义上的物理AGI时，可能只需配合少量真机数据，就能达到一个非常强的泛化的大脑，是这样吗？
张总：
应该说，未来机器人的主粮或者主教材应该是无本体的数据，即第一人称的无本体数据，
德玛已经在国内比较早地部署了。那么它的真机数据相当于维生素，还是需要的，但是从
规模上来说，真机数据比不上无本体的数据。整体的数据市场规模，我们未来是一个比较
乐观的态度。对模型公司来说，它的研发费用支出主要在两大块：第一块就是算力，第二
块就是数据；数据对于模型厂商，在整个具身智能产业链里，应该是投资的主要方向。

主持人：
明白。能不能请您再分享一下真机和无本体的优缺点，开头也讲过一下，能不能再展开分享一

下？
张总：
好的。真机数据的好处就是它比较直接，它采集的就是机器人的关节运动角度等数据，在模型训练时使用的难度会比较小。但它的缺点也非常明显，就是采集成本会很高、采集效率会很低。首先要买一个机器人本体，当前也要花费十数万到数十万这个水平；第二个还要雇一个专门的人在一个搭建的场景来采数据；同时它每天采的数据，因为它是遥控操作一个机器人来采，所以每天能采集的数据规模会比较小。一个专业的采集人员加上一个机器人本体，每天能够采集的有用数据可能就是一到两个小时。所以经过最近两年的发展，整个真机数据、就是遥操数据，它的规模并没有达到行业里需要的规模化的数据。
今年以来兴起的新的技术是无本体数据采集。德玛也部署了阿林（Ego）解决方案，它对应的就是无本体数据采集。这种采集方式不依赖于机器人的本体，不需要去购买机器人。我们人类本身就在生产数据，我们从第一性原理角度出发，人类每天都在生产大量的数据，只是没有被记录下来，所以我们开发了一系列的采集硬件。我们现在开发的是一个头环戴在头上，那么我们各行各业的工作人员在操作的时候，这个头环可以以人的第一视角方式，把我们手部的操作和胳膊的运动方式记录下来，再经过我们的数据管线加工成机器人模型训练能够使用的数据集。这种方式，大家可以想象，我们各行各业的工作人员或者操作人员，他在正常工作的时候就可以把数据记录下来，这种方式的数据采集效率会非常高。正常一个工作岗位，每天工作按8小时算，算上一些间断或者休息时间，每天能够积累的数据规模也能达到5~6个小时的有效数据规模。而且，这个工作岗位在我们生活中是无处不在的，千行百业都可以佩戴这个设备采集数据。所以这种新的方式，给行业、给我们的模型性能提供了源源不断的数据来源，这就是我们行业里补充的scalinglaw，就是规模效应。
主持人：
明白。最新的技术路线就是这种无本体数据采集，可以快速积累起数千万甚至上亿小时的数据集。目前无本体数据采集的成本结构、成本构成大概是怎么样的？
张总：
无本体数据采集的主要成本构成，一个是一次性的数据采集设备，它的成本也是在万元级别，万元甚至未来有可能会降低到千元级别这样一个设备。然后同时在采集人员方面，在

真实场景里，我们给工作人员少量的采集补贴。所以它每小时的采集成本，从原来的数千
元每小时降低到数十元每小时，单位时间能降低100倍。
主持人：
数千元降低到数十元，单位时间能降低100倍，是吗？这么夸张吗？
张总：
对的。原先我们部署一个真机遥操，它的成本会非常高，主要一个是人员工资，一个是设
备的一次性投入，投入一台本体和采集设备就要达到40万元。
主持人：
明白，您是比较之前的有机器人本体的情况下，对吧？
张总：
对的对的。
主持人：
那现在您说的那些采集硬件，是一次性的吗？大概使用寿命是多长？
张总：
它是一个电子设备，正常一个设备使用2~3年不成问题。
主持人：
明白。您说的是万元成本，就是一套原来大概万元，然后现在已经降到千元了。原来一套数据采集成本要购买一个机器人本体，可能是几十万元，现在我们购买一个数据采集器，也就万元左右的水平。
张总：
哦，明白。原来的一套数据采集成本要购买一个机器本体，可能是几十万元；现在我们购
买一个数据采集器，成本也就是万元左右。
主持人：
然后您说未来可能会降到千元，就是在现在的基础上还能继续有很大的下降空间吗？
张总：
对的，因为现在整体的规模还没有完全铺开，我们希望是构建一个万人甚至10万人这样一
个采集规模。那么随着设备的进一步量产，它的规模效应，研发费用分摊之后，这个单纯
的硬件（BOM）成本会下降比较快。
主持人：
明白，就是现在整个硬件其实还是在一个降价的通道里面，对吧？

张总：
是的，这种新兴的采集方式，应该说从今年的年终开始到现在，是一种新的采集方式，所
以它设备的单价，未来还有降价的空间。
主持人：
明白。那目前需要用到的主要采集设备都有哪些？有没有未来可能会增加的一些设备？
张总：
目前对于德玛的阿林（Ego）解决方案来说，它主要设备是一个包含多目相机的头环，戴在
我们头上，以人的第一视角方式去记录人的操作过程，相当于是一个视觉的记录方式，然

后再通过算法把人的手部关节进行还原，把手部的操作方式记录下来，变成模型可用的数
据。这对应人的学习过程，因为我们人对物理世界的认知，可能超过50%、甚至60%都是通
过视觉来记录的。那么未来还需要进一步补齐的是力觉、触觉这些信息。当然这个行业里
也有各种解决方案，有的是做视觉手套或者力觉手套，也有用肌肉电极（EMG）的方式去记
录，来反向推算手部的力量。主要就是目前的设备，主要是视频、音频这些数据信息。未
来随着产业链的进一步成熟，还需要进一步补齐一些触觉、温度，甚至机器人模型训练能
够对应的人类这些触觉、感官信息，进一步去对齐。

主持人：

明白。现在其实很多工业和物流场景还是非常需要力觉和触觉，因为它拿不同重物，还是需要这一块。那之前没有去这么做的一个原因是技术没有攻克，还是什么原因？现在这一块的需求会变得非常迫切吗？

张总：

当前应该说，如果把对比来说，我们采集数据是在为机器人模型准备教科书，当前的教科
书还是一个通识教育阶段，所以主要是以视觉为主。那么力觉最主要的问题是，现在传感
器的价格还没有下来，像力觉采集手套单副就要数万元，然后它的可用性也比较差，大概
也就用个1周到两周就坏掉了。所以这些传感器件还没有准备成熟。所以现在行业里，我们
对模型厂商来说，最聚焦的还是让模型具有初步的泛化和操作能力，所以还是以视觉为
主。那么随着产业的进一步成熟和模型能力的进一步成熟，我们就会从通识教育的教材转
向专门的教材。所以这是产业和行业相辅相成的一个过程。

主持人：

明白。说到模型的情况，再回到之前请教您的问题：您觉得目前可能训练的时长达到1万个小时这种级别吗？

张总：

模型的训练时长，原来前期行业里这些VLA模型或者早期的真机数据规模都在数千小时到1
万小时这个规模。那么最新的方向，我们的无本体，因为它采集的数据规模能够快速拉起
来，我们刚刚说通过佩戴一些简单的设备来采集人的手部操作，能够快速把数据的规模推
到数十万小时、甚至数百万小时，所以当下比较领先的模型，它们都会有一个行业门槛，
就是百万小时的人类数据。

主持人：

百万小时人类数据。那这个数字在明后年，随着世界模型的发展，会把这个百万小时降到比如十万小时吗？

张总：

未来的小时数应该是加大的。这个小时数我们可以类比大语言模型的参数规模来理解，它
不是完全相同，但可以类比。现在是模型的规模还比较小，只能吃比较少的数据，具备初
步的通用和泛化操作能力，它需要的数据规模应该在上亿小时。
主持人：
明白。假如我们针对某一个具体的场景，比如说物料搬运这个场景，以目前的模型能力来讲，如果机器人要在物料搬运这个场景实现非常实用的操作，它需要的训练角时（数据规模）可能达到10万小时，对吧？
张总：
就按照无本体训练来讲的话，可能也需要10万个小时的数据采集和训练。
主持人：
那么随着模型的发展，可能到明后年再针对物料搬运这个场景，会不会把这个训练时长大幅压缩呀？有这个可能性吗？
张总：
你指的是我们模型训练的能力，还是数据规模？
主持人：
我不懂，我就是说模型的泛化能力是在强化的，现在原来是VLA模型，现在慢慢往世界模型去转，世界模型也在强化。那现在模型可能没有那么发达的情况下，你训练它可能需要（较长时间）才能做物料搬运；那明后年模型都非常发达了，你训练的这个时长会不会就大幅压缩，可能只需要1万个小时，差不多就能用了？

张总：
这个逻辑可能有点问题。应该说，随着模型能力的增加，它需要的数据规模、也就是小时
数会越来越多。相当于对人来说，他需要看更多的书，才能够学会更多的知识和技能。所
以未来在物流和工业这个场景，机器人能够展现一个非常好的操作能力，它可能需要的数
据规模是千万甚至上亿小时这样一个规模，它能够吃下更多的数据。
主持人：
就是同样的场景，还需要（采集）更多的数据是吗？
张总：
这个不是训练时长的概念，是我们预先采集更多的数据，拿这个数据去做预训练或者后训
练，对模型进行训练跟调整。类比来说，就相当于大语言模型，它的参数规模会越来越
大。所以未来大语言模型的规模会越来越大，能力会越来越强，那么在工业、物流、生活
各种场景，它的部署进程会加快。
主持人：
明白您的意思了。就是找到一个非常强的、有泛化能力的大脑，然后需要更多的数据和训练，才能实现泛化能力，然后更多场景都能用。那现在我们德玛做物流场景，现在的方案是怎么样的？模型是我们自己弄的，然后现在做一些数据采集和训练，这个路线会就这样持续执行下去吗？
张总：
一个具身智能机器人或者一个通用机器人，能够在一个场景落地，它需要的可能第一是有
一个非常好的操作性的本体，第二是有一个能力不错的模型，然后有足够的数据去训练

它，最后一步就是我们有一个非常好的场景去部署这个具身智能。那么对德玛来说，我们
跟行业里的整体厂商和模型厂商都建立了非常好的合作，大家最近关注到我们跟蚂蚁、宁
波也签署了战略合作关系。蚂蚁、宁波主要聚焦的就是机器人的世界模型、或者说机器人
的大脑。那么对德玛来说，我们对场景有非常好的认识，我们在场景里积累了大量的数
据，在这个数据上，跟行业里领先的这些模型厂商和本体厂商去合作。德玛的目标是让机
器人真正在场景里能够用起来，所以我们的定位是把前面提到的场景、数据、模型、本体
做一个整合，形成一个非常好的解决方案，借助我们的整个市场渠道和服务网络，真正能
够在客户的场景里去落地和应用。所以当下我们德玛聚焦的点，一个是采集足够多的数据
积累，我们对场景更好的理解和数据；第二个是以一个非常开放的态度，与行业里面做得
非常好的、像优必选、像宇树等等这种行业里本体做得非常好的公司，和大脑做得非常好
的公司，把它整合成一个真正能够落地的解决方案。

主持人：
明白。蚂蚁、宁波他们的模型方案大概是什么样的参数级别？大概是怎么样的一个想法？
尚总：
我来补充一下，刚才刘总问的那个问题。我们今年年底应该会发布垂类的模型出来。这个模
型训练采用的这个数据是多模态的、多样性的数据，数据质量非常高。如果训练这个模型
的话，其实需要的数量级并不是很大。我们会训练出来一个效果，就是说在分拣、在搬运
这种比较简单的场景里面，我们会达到七八十分的效果，就是能够开始有一些简单的应
用。这个也像小孩子去学习走路、学习跑步一样，刚开始的时候，他可以简单地去进行一
些应用，然后我们再通过POC客户端去部署的时候，再根据实际的一个场景去做进一步的调
整、还有优化。其实在部署和做POC的过程之中，我们也不断积累实际的数据和应用。然后
未来，如果我们可能刚出来的时候就能用大概70分的一个效果，然后我们要把它提升到80
分，POC测试完成之后达到80分甚至更高的效果，我们才好向客户端去进行部署。所以我
们后续需要的这个数量，不一定就是需要自己去专门采集，可能客户端那边也可以不断去
积累。另外，如果从70分到80分，它需要的数量就比以前可能更高；然后80到90或者更高
的一个水平的话，它需要的数量级可能是呈一个指数的提升。刚才您提到蚂蚁联播（宁

波）的这个数据、它模型跟我们合作模型的具体参数或者什么一个情况，这个信息的话，
我觉得还是要等到我们12月份正式发布，因为这个事情我们现在还不太合适对外说。
主持人：
明白，感谢张总，很清楚。那我再请教一下张总，现在数据采集这个行业的竞争格局是怎么样的？参与的这个玩家主要有哪些？
张总：
数据对模型的作用是非常直接也是非常重要的。所以目前数据这个行业里面，参与的除了
像我们这样专业的数据公司以外，各家的机器人、模型公司也都推出了自己的数据方案；
但目前来看，一个是主要偏自用，另外一个是积累对数据采集的一些认知，同时把认知反
馈给类似蜜蜂这样专业的数据公司。所以整个数据行业的格局，最顶层就是我们的客户，
是模型公司；然后中间层就是行业的主流数据采集公司；最下层是拥有场景的这种类似服
务公司、或者原来这些人力资源公司，他们可以组织起、在各种场景里能够快速组织起比
较多的人力资源，然后把我们的设备给部署下去，实现一个数据的回流和采集。行业大概
是这样一个构成。
主持人：
明白。那整个数据生意的壁垒，您认为主要在哪些方面？

张总：
数据这个生意，第一，我们认为首先是要对模型有充分的认知，知道模型厂商、和模型厂
商保持一个比较好的沟通，知道什么样的数据对模型是有用的。前期也出现过看起来数据
门槛不高，我们在场景里面去佩戴一个设备，然后快速把数据采上来，但实际上这些数据
大部分是没有用的情况，前期行业里一些中小公司这样做。所以第一个，我们认为是要对
模型和数据两者的关系、在技术上什么样的数据对模型的泛化和智能能力有比较好的提
升，这个理解我们认为是非常关键的。第二，是能够快速把数据的采集能力部署下去、扩
展开来，实现一个规模化的数据，实现百万小时、千万小时这样一个数据规模。应该说这
跟其他行业一样，现在模型公司对数据的需求是非常巨大的，但是模型厂商不可能去无限
制地对接很多的这种数据公司，可能还是希望绑定少数几家，然后实现规模化的数据供
给。第三，还有一点我认为非常关键，就是能够有很多的场景，然后在这个场景里面，根
据模型公司的要求，把数据从规模上采集上来。
主持人：

明白。第三点，通过什么方法可以进行比较好的控制和管理？
张总：
对场景，就是要在行业里有比较多的积累或者资源。以德玛为例，我们目前的数据在像工
业场景、物流场景，这些场景我们本身相对有优势，无论在自有工厂，还是我们有非常多
的物流行业客户，这就是过往在行业里的一些积累情况。
主持人：
明白。您说到可能还需要配合不同的模型进行数据的适配。那这又回到我之前纠结的问题：假如海外很多世界模型都已经非常发达，比如英伟达的GR00T（力控）进来了、能用了，那会不会经过他们很多的信息（预训练）创建之后的模型就已经非常能打了？如果用在物料搬运这些物流场景里面，可能稍微用一些为他（定制）的小一点的技能包，然后稍微做一些实景的操作演示，他基本上就学会了，然后很快实操效果就已经非常好了，有这种可能性吗？因为如果现在找国内一些模型厂商合作训练，可能做分拣、做物料搬运还要训练1万个小时或者几千个小时，最终只能用来做物料搬运；那如果后面世界模型进来了，它已经具备泛化能力了，只是不会某一些技能，学习一些特定技能以后，就可能非常实用了。我不知道有没有表达清楚这个意思。
张总：
我们来类比来看，一个是大语言模型，一个是自动驾驶。最终模型追求的肯定还是通用和泛
化，它需要学习各种各样的数据、各种各样的能力，去充分地认识和理解物理世界。所以
即使到今天，各家大语言模型公司还在追求数据规模，自动驾驶公司也在通过不停路测、

甚至用户真实的操作来回流数据。那么具身智能所需要的数据规模，可能是要成百上千
倍，相比较于自动驾驶和大语言模型。所以这个数据的需求，应该是一个长期的过程。那
么在物流和分拣这个场景里面，它不仅仅是要具备基础的搬运、上包这些操作能力，也需
要去理解我们面对不同形状、不同材质、不同颜色的各种包裹或者对象，对它有一个更好
的理解。所以说，未来模型对数据的需求，我们认为是一个长期的过程。
尚总：
刘总，这个问题我也稍微补充一下。因为现在我们基于德玛在做测试，也是基于比较简单的
几个工况，比如分拣、上包，还有码垛这些。这些场景的话，如果按照模型公司做，我们
出来这个效果可能就是五六十分、六七十分的那个效果。但是如果我们真的让他能够实现
对人的期待，达到90分、95分甚至更高的话——因为刚开始我们部署，你达到80多分就
OK；但是后面达到90多分的时候，这个指数级的一个数据需求，也包括刚才张总提到的，
就像无人自动驾驶到现在这个阶段都还没有说非常完全的成熟，其实这是一个非常长期的
需求过程。可能模型公司做出来一个效果比较OK，但是它能够达到真正的完全替代、或者

比人做得更好，这个时间我们觉得会是很长期的。另外，现在如果是简单的几个工况解决
了，那么还有未来更加复杂的工况，还有比较长流程的工作流，这些事情又要我们去面
对，数据集的需求会越大。
主持人：
我接着那个。因为之前咱们预期达到替代人的90分效果，可能比如在今年或者明年就能实现一个突破，但是从产业角度去了解，可能现在的模型能力还是有一点（距离）。其实我想知道的是，同样一个场景达到90分的能力，从第一性原理出发去理解这个事情，我理解张总刚才说的意思，就是说可能同样的场景达到同样90分的能力，随着模型的发展，反倒需要更多的训练时长，这就感觉非常不太好理解。因为从第一性原理出发，随着模型的进步和发展，同样的场景达到90分的能力，应该需要的训练时长越来越短，对吧？但如果是随着模型的发展，同样的场景达到同样90分的能力，训练时长反而要越来越长，这个怎么去理解呢？
张总：
是这样子的。真正能落地的技术路线，它和自动驾驶、大语言模型是类似的，它底层是一
个神经网络的架构，它本质就是一个调校良好的神经网络，然后用大量的真实数据去训练
这个模型，让这个模型学会理解这个物理世界，理解操作的因果关系，然后掌握更多的操
作技能。所以现在模型的数据量，还都是一些非常、怎么说呢，就是不是那么精细的一些
操作，只是让这个模型学会一些基本的操作。类比来看，就相当于让一个小孩子从牙牙学
语到认识这个世界，知道我推一个瓶子，瓶子的水会洒出来；我推一个桌子，桌子能够移
动，有一个事件的反馈。那么未来我们希望机器人是在真实场景去替代人工，去做一些非
常精细化的工作。那么这个场景你就需要更多的、除了视觉之外更多多模态的数据，所以
这个数据的规模是会越来越大的。并不是说我只是在物流场景操作、我只为物流的数据给
它就可以的，其实是需要各种各样的泛化数据，让这个模型展现一个非常好的对世界的理
解。所以说它的数据量，不仅仅是物流场景这些数据，它可能是人类世界的各种真实场景
的数据都会给这个模型，那么它可能在其他场景学习到的物理规律，在物流这个场景能够
理解和使用。所以我们说未来的基模，它的模型规模会越来越大。目前的阶段，因为算力
受限，提升模型规模大概都是在数十B，就相当于几百亿这个参数规模；未来它可能像大语
言模型一样，整个模型的规模会到千亿、万亿这样一个规模。所以我指的数据规模是这样
子，它并不是说需要更多的训练时长，而是指需要更大规模的数据量。
主持人：
明白，理解您的意思。感谢张总。我想问一下尚总，现在方便公开德玛肯的一些业务层面的数据吗？比如大概的年收入情况、年利润情况，然后未来订单和收入的指引，这一块能方便吗？

尚 总 ：
这个暂时还不太方便，等到我们这边内部讨论，等年报之后、三季报之后，我们这边讨论之后，再看怎么去做分享，现在还不太合适。
主 持 人 ：
明白，感谢感谢。我这边暂时也没有别的（问题）了，我们要不开放一下线上提问。
（会议助理播报提问方式）
主 持 人 ：
张总，我看暂时没有提问，那今天（会议就到这里）。

--- 六大行业共谈：物理AI20261027.pdf ---
全文总结
一 、 策 略 视 角 ： 为 什 么 在 当 前 时 点 关 注 物 理 A I
从策略角度看，当前市场正处于剧烈震荡、资金面承压的过程中，资金面压力是近期市场波动的核心原因之一。但上周国有资金通过大额申购ETF承接抛压，向市场注入了一定流动性：沪深３００ＥＴＦ（华泰柏瑞）上周申购净流入约２７８亿元，科创５０ＥＴＦ（华夏）上周申购净流入
１７０亿元，创业板ＥＴＦ（易方达）上周申购净流入２１０亿元。
在这样的背景下，市场底部确实
可以初步确认，但目前仍然处于反弹后高波震荡的框架之内，仍在再平衡的框架之内，后续
并非线性上涨。一方面筹码集中度相较7月大幅改善，另一方面美债利率下行，全球宏观流动性风险得到一小部分释放；但也要注意，近期长短美债利率下行的背后，其实有一部分是对AI趋势和回报率下降的担忧所引发的避险，即分子端出现了逻辑上的一些问题。
因此，从投资逻辑上讲，微观结构虽然已有一定程度的缓和，国有资金承接抛压、资金面压力有所缓解，但目前的筹码结构和催化角度仍处在再平衡、反弹后高波震荡的框架里，后续并非线性上涨。在这样的框架下，需要重点关注具备二次点火与叙事张力的方向，而物理ＡＩ正
好同时具备筹码比较干净、以及二次点火和叙事张力这两点。
物理ＡＩ是ＡＩ的下一个应用方向，具备比较大的增长潜力；它也是机器人的智能核心，
能让机器人理解物理世界并做出合理行动。
本次会议邀请开源证券研究所六大行业首席，分
别从计算机、电子、机械、军工、汽车、中小盘角度分享物理AI的投资机会。
二 、 计 算 机 ： 物 理 A I 是 全 球 共 识 的 大 产 业 趋 势
物理ＡＩ是全球科技巨头的共识，不是某一家公司单独去做的全球工程，谷歌、Meta、
OpenAI、特斯拉，以及国内的字节等都会下场做，这代表着科技下一个发展的方向。物理AI最核心有三块：世界模型、ＶＬＡ、数据原生或数字仿真，三者共同构成物理AI。物理AI的定义，实
际上是对所有物理世界发展规律的理解，要求载体（机器人、车、机械臂等）根据当前物理状态判断下一步可能发生什么物理状态变化，并据此做出动作反馈。
物理ＡＩ这次根本性的变化，是因为大语言模型在一定程度上突破了物理世界与文字世
界（数据世界）的边界。大语言模型本质上根据上文输出预测下一段文字，是推理结果，与物理AI是完全不同的训练体系：数据不一样、训练方式不一样、输出结果也不一样。随着GPT-6 Astra的发布，通过大量数据投喂、加上一些物理专业知识，再把结果转化成电机驱动信号，基本就能够解决、替代VLA的部分功能。7月发布Astra之后，一些具身智能人员使用Astra大模型放

在宇树的机器人上，已经成功完成了一些命令的调用和执行。这意味着全球所有大语言模型公司

都可以下场去做世界模型或具身大脑，在千亿级美元资本开支之下，每家必然投入重注，因为模型能力再提升已到瓶颈，而把资本开支放到物理AI或世界模型上，会打开更大的市场空间。
大模型事件模型要解决的问题：用的还是Transformer架构，与原有体系类似，不同的是训练数据，数据是未来每家公司的核心生产要素，谁能拿到更高质量、更多的数据，谁就能
做出更好的模型。数据体系包括：真机数据（最真实）、Ego数据、UMI数据，再加上数字仿真，共同构成整个数据体系。
围绕数据的投资思路：第一，Ego数据的采集，一级供应商如蜜蜂做得比较好，其口袋式、手上采集设备需要加装大量传感器，国内做六维力、IMU、视触觉的供应商都可能参与；第二，真机数据，利好有本体的公司，如宇树、上纬新材、中天科技等，因为有本体就可以拿本体去采集真机数据；第三，数据孪生、测试合成，如索辰、风语筑等公司，1:1还原真实设计的数据，也是核心方向。
数据市场非常大，理论上世界模型能力提升没有边界，数据的规模也没有边界，用每年几亿、十几亿、二十亿小时衡量都可以，但很难做详细测算，我们认为每家基本数据会在大几千
万小时的级别，对应的数据市场有非常大的增量空间。
执行端：整个物理AI最终利好特斯拉，因为特斯拉有全球最好的数据，弱点在大模型，而OpenAI把模型问题解决了，所以未来特斯拉在世界模型这块肯定是全球做的最好的之一，更
利好特斯拉的产业链。特斯拉在执行端基本上领先其他公司3到5年左右，未来商业化前景会更明确、业绩的确定性和弹性会更确定。另外，OpenAI、Meta等未来都会做机器人，他们会找
新的供应链，与特斯拉供应链重合度大概率不会那么高
，因为现在给特斯拉做机器人执行端
的产能基本都打满了，没有空余产能给其他厂商；所以会诞生很多给OpenAI、Meta供货的公司。
OpenAI目前的方案基本确定下来是两种路线：一种是下半身为三角支架、上半身为半身人形、有两只手臂、每边手臂大概3个手指，该方案已逐步推进，手臂是外购的，但手由OpenAI自己把控，主要用机器人采集真实数据，明年可能会推出自己的整体方案。包括国内的小米等公司也可能下场做机器人，大家都会看特斯拉的V3，V3.1发布之后很多公司可能会模仿特斯拉的路径，类似于新能源时代。明年全球人形机器人出货量可能是一个非常大的爆发
：今年上半年国内、
全球人形出货量可能也就三四万台，今年全年、明年全球可能是六七十万台的量级，基本是快十倍的增长，很多确定性参与这个环节的公司的业绩都会充分受益。
看好的标的：第一，OpenAI链，从传动、微型传动角度，中大力德做得最好，在微传动这块绝对领先；利群精密，因为OpenAI背后可能有很多公司能够参与；创新精密可能通过天机的环节参与。第二，数采环节，要么有Ego出货量能力，要么传感器市场份额很高、价值量很

大，奥比中光是比较好的标的；此外福莱新材、安培龙等做传感器的公司也可能受益。第三，气动（"气念"应为气动/丝杠环节），从市场份额和与市场率的价值角度，科森、恒立、福赛
目前非常明确、进入到逐步业绩兑现且非常核心的环节；目前整个预期差最大的可能是光洋股份等新进入、具有绝对竞争力的公司。后续市场可能交易的环节：一是手部的灵巧手，因为需要灵巧手去采集，是重点要解决的方向；二是具有高质量数据的公司
，因为有些公司本身的高
质量数据可以作为数据价值并列，类似于中控技术这样的公司。
三、电子（电信板块）：端侧电池与机器人皮肤
物理AI对电子板块主要有两个此前偏冷门的细分领域的变化。第一是数采装备的便携式电池，它与大家理解的端侧电池有一定区别，但有一定相似性，生产厂家也是交叉的。数采此前很多是插线式、拖着一根长线，但今年很多机器人公司在前期开发模型、做测试时，数采是硬性刚

需。行业内如豪鹏科技、中兴通讯（"中外关羽"按原文应为其他公司，此处按豪鹏科技为
主）等公司都做了很多端侧布局，尤其豪鹏今年整个机器人板块的电池收入可能大几千万、
接近一个亿，主要收入来源就是数据采集，市占率相当高
。数据采集领域对下游客户处于成
本不敏感阶段，所以这大几千万收入带来的利润率水平非常高，后续增速也非常快，对消费电池板块来讲是一个非常好的新业绩增长点。大家此前更多预期是机器人本体上电池的放量，但现在领先一步的是数据采集领域。另外这两家公司同时也有端侧电池概念，像穿戴类眼镜、戒指及其他可穿戴设备，随着AI发展会有很多新应用，需要小型化、高功耗的电池，端侧电池有明显通胀逻辑。
第二是数采领域的机器人皮肤。近期随着数据采集、前期数据需求，皮肤也充当了比较重要的角色，比如温度、压力、仿生等方面的新变化，逻辑与数采应用电池一样，都是数采领域优先放量。相关公司包括福莱新材、汉威科技，最近表现相当亮眼。整体上与之前看机器人皮肤时的预期差并不是特别大，但行业有很多积极变化，包括海外大展会、海外客户积极对接，随着数采及物理AI领域爆发，后面还是会迎来更好发展。目前这些板块相关标的前期经历了很大回调，无论从筹码还是从基本面角度，很多票都跌到一个非常深的位置，建议积极关注。
四、机械：物理AI的基点突破与本体、气电、数采三方向
这一波物理AI行情主要来源于两方面。第一，OpenAI的GPT-6Astra新模型
，从功能角度
突破了此前大语言模型对物理世界理解较差的局限，不仅对物理世界理解加成很多，在集成智能驱控层面作用也比较大，各个本体公司实测下来，解决了很多此前模型不具备的驱控问题。我们与硅谷投具身智能、投大模型的VC交流，大家觉得对物理世界理解的突破是一个断点式突破，需要大量训练才可能出现基点变化，Astra的出现是时隔两年之后在具身智能相关模型上发现
的断点型突破，有人把这种变化称为物理AI的一个基点
。往后看，它的出现不仅是这么简

单，更多可能会激发其他模型公司、包括OpenAI后期对物理AI训练的投入，形成一个生态式投入，有可能会带来后面1~2年整个物理AI模型的突破进展
。王兴兴最近也发表言论，认为物理
AI未来1~2年可能会被击穿，"击穿"采用的就是大规模训练下的断点突破。
在这样的变化之下，市场对机器人、具身智能板块的理解，从原来只是做展示性、娱乐
性简单任务的机器人，变成了有可能具有类似手机安卓系统这样的操作平台的物理硬件，
换句话说变得真正智能化，并可能有相对统一的智能化平台系统。长期看，如果这样的模型存在并持续迭代，本体的使用价值会大幅提升，智能化程度也会大幅提升。首先受益的是本体厂商，这也是宇树、上纬近期涨得比较多的核心原因——他们把身体训练好了，大脑交给Astra或其他大模型去驱动，是很好的分工。
其次是垂直类本体公司，现在市场也非常关注，比如服装缝纫领域的杰克科技，以及融创智能（包装、简单装配领域）。为什么这些小的本体公司受关注？核心原因是在目前数据量相对有限的情况下，做通用大类模型或大范围通用落地挺难，所以具有闭环数据、私域数据，同
时在专用行业里专用设备能力强的公司具有稀缺优势
。类似杰克、融创，甚至其他有专用设
备的公司，比如复星、华商达、国储等，都值得重点去看。
第三是气电（"替炼"应为特斯拉/T链），特斯拉毫无疑问是把本体训练得最好、复杂程

度最领先的一家。特斯拉近期在中国产业链加单，周产100台往周产几百台去靠，按此推算
到明年特斯拉可能在全球拥有1万台以上存量的OptimusV3机器人，核心是为自己明年的数
采和量产做较大准备。T链公司比较多，我们比较推荐在T里站位比较好、同时SKU占比往上走的，比如福莱、思林（"思林"按原文），这两个定点比较多；还有三花智控、恒立，这两个
是大的T链中心，也非常值得看。
第三类，长期看模型发展，数据采集非常重要，数采类公司最近表现也很好，一方面是
传感器、电子皮肤的公司，另一方面是专门做数采的公司，后者有标的是华丽科技值得重点关注；前者我们认为除了市场现在炒得比较热的福莱新材之外，像天安新材、大美股份这类竞争公司也被严重低估，值得重点关注。
五、军工：装备正向研发、数字孪生与战场具身智能三大方向
物理AI在军工板块主要有三个方向的较大应用前景：一是装备的正向研发，二是训练以及合成数据、数字孪生等，三是战场具身智能，包括无人作战平台未来向具身智能转变。
第一是装备的正向研发。目前很多军工装备研发周期偏长，需要巨大物理量输入，前期研发占用大量时间成本。有了物理AI之后，在飞行器高速飞行、气动布局、新材料开发、跨尺度研发等方面，如果借助物理AI的输入，能够巨大提速研发进程。目前行业内已开始用物理AI方式设计隐身、做电子仿真，在某些领域已能看到应用雏形。

第二是数字孪生。已有部分领域应用，包括能科科技之前在生产端和实战端做的一些数字孪生技术。军工很多模拟仿真此前积累了相当宝贵的数据，此前最大问题在于没有办法给出非常准确的物理量，受限较大；如果物理AI能突破这层限制，在很多数据荒的领域就可能会有
较大突破，包括提供更准确的红蓝方数据、无人作战等方面的指挥决策数据，能够根据实时物理环境做到更准确输入，这是精准检验武器装备是否有效、是否达到作战效能的重要应用方向。以前更多企业在做比较简单的数字仿真或训练用模型，现在如果有物理AI加入，可能会把模拟仿真数据提高到另一个维度，增强的模拟仿真在整个军工系统的开发、应用和验证过程中比较重要，能够极大缩短周期、提高整体效能，目前有比较大的经济价值。
第三是具身智能。包括装备本身的智能交互，以及无人作战平台等装备本身的协同，也是我们认为能够率先落地的方向。目前无人作战已成为全球比较重要的主题，经过几次较大冲突，可以发现无人装备平台已显示出巨大的经济价值，包括可以不对称打击，以及在非常低的战损情况下去执行复杂任务。此前市场比较关注的无人战机、僚机、巡飞弹、蜂群，以及地面的无人狗、无人车等，都是九三阅兵之后市场比较关注的方向。如果有物理AI加持，可能将这些无人
平台本身的作战效能成倍放大，我们一般把这类平台叫做倍增器，物理AI当下可能就是最
重要的一个倍增器，可以将雷达、周围数据传输平台集成到某个具身智能体上，实现整个装备的互联互操作，可能彻底改变当下战场形态，也是我们认为当下具身智能在军工领域、无人平台上可能落地最快的方向。此外，"十五五"期间整个无人装备有望快速发展，装备比例有望显
著提升。九三阅兵已显示出机器狗、无人机、UUV等大型新型无人装备，但与海外相比仍有较大差距，海外如弗兰（按原文）等公司已陆续获得一些较大订单；如果物理AI跟海外相比又拉开显著差距，可能"十五五"期间无人机不会是我们想象中的简单平台，会承担更多具身智能任务，从过去比较辅助的装备向核心作战力推进。
相对受益的标的，包括原来做微波的、做无人的，如国睿科技（"国国电子"按原文）、高

德红外、复旦微电等，我们认为当下可能率先落地；数据应用方面，能科科技是目前看到的在物理AI上与军工密切合作最多、有望成为数据上比较重要支点的应用方向。

六、汽车：机器人+智驾两条主线，数采与整机代工两个新方向

物理AI在汽车板块的核心落脚点主要是两个：机器人和智驾，这两个也是AI应用层面率先落地的方向。特别是智驾已经从过去三五年发展得比较快，到2027年也有很多L3车型逐步落地的预期。所以物理AI层面，机器人加智驾应该是两条主线。在这个背景下，除了上一轮机器人行情炒的硬件类零部件方向之外，还有新的方向值得关注。

第一个是新方向数采。这偏计算机行业看的，但汽车行业也有一些公司在这个维度有布局。一是阿尔特，主业是做第三方整车设计研发，在物理AI方向布局了两条线：第一条线是跟光轮智能合作成立合资公司，一块做仿真以及物理AI数据的收集和测评（该合资公司于今年6月25号签约成立，光轮智能主要做仿真数据模拟，阿尔特提供汽车制造场景数据的收集，以及用收集回来的真实数据去做仿真校准，两者优势互补）；第二条线是跟富士康成立合资公司富特康，做整机的机器人代工（今年8月签约成立），主要为整个产业形成闭环，是最后商业闭环的最后一环，延续阿尔特此前在汽车领域设计整车交钥匙的全服务优势，现在机器人也能做到交钥匙的整个方案，从硬件、软件到数据反馈形成一整套商业服务。阿尔特作为国内规模最大的第三方独立汽车设计公司，在整车设计维度有比较多数据，也有整车研发、零部件以及很多实力收集回来的数据优势，这是其布局该方向的核心优势。此外，阿尔特主业今年上半年已开始回暖，此前因客户信用贬值问题导致过去两年业绩承压，但今年上半年规模利润同比实现大幅提升。

二是中国汽研，主业是做汽车检测，即新车上市公告前的检测，也包括各种仿真测试，是国内体量比较大的汽车场景数据采集、仿真测评检测机构，手上包括智驾的数据能够比较好地向物理AI领域延展。它有一个仿真云平台，在仿真数据积累维度有比较多优势，不管是在智驾领域落地，还是机器人领域，都能延续过往优势。除汽车领域检测外，中国汽研也有机器人检测，旗下在重庆有国家级的机器人检测与评定中心，是中西部唯一一个国家级的机器人第三方检测机构，检测覆盖原型机器人、工业机器人、AGV、特种机器人，包括核心零部件如灵巧手等都够服务，所以本身也有机器人领域布局。

第二个新方向是整机代工。机器人迎来逐步商业化量产过程中，除了数采方式之外，整机代工也值得关注。整个汽车板块也有一些公司在不同领域布局，有些是工业机器人、有些是人形机器人、有些是轮式，此前这个维度也有一些布局：一是阿尔特，跟富士康成立合资公司后也会做整机代工；二是浙江先通（按原文），此前跟浩瀚星空、启腾、机建等成立控股合资公司布局机器人代工；三是均胜电子，旗下均普智能也有一些机器人包括人形机器人的业务布局；四是中鼎股份，旗下也有子公司做整机代工。所以我们认为在AI应用领域落地的变化过程中，整机代工维度也可以引起大家关注。

七、中小盘：看好视觉与触觉两个方向

视觉方面，视觉是物理AI最关键的感知入口之一。物理AI要解决的不仅只是生成内容，而是需要让系统在真实的物理空间当中完成对物体、场景和动态关系的感知、理解与交互。在全部各种传感模态里面，视觉几乎是唯一一个能够同时提供高空间分辨率、丰富语义信息、连续实时信息和非接触测量能力的感知方式。视觉之所以能占据核心位置，不是因为它其他模态不重要，而是因为视觉的信息密度最高、覆盖面最广：力觉、触觉可以告诉机器人接触到了什么、接触力是否合适，惯性器件可以告诉它自身姿态和加速度如何变化，激光雷达可以提供距离和轮廓，但这些模态往往难以单独回答眼前到底是什么物体、边界在哪里、是否存在缺陷、姿态如何、是否已发生细微形变等问题。尤其在工业场景和具身场景里，往往要求同时理解类别、位置、姿态、深度、表观缺陷、相对关系和动态变化，这些是视觉擅长的领域。更重要的是，物理

AI所需的数据不是孤立的图片，而是连续、多模态耦合、携带物理量的真实交互数据流，视觉既能生成2D图像，也能通过结构光、双目、ToF点云、动捕等方式提供3D空间信息，还能和IMU反馈处理数据融合。

从产业链逻辑来看，视觉的价值已经从过去工业自动化里的反馈工具，逐步升级为模型训练的数据入口，这意味着后续企业竞争不再只是比拼相机、镜头性能，而是比拼能否稳定采集到高质量数据、能否高效标注和训练、能否沉淀行业经验、能否在真实场景中持续回流样本并形成部署闭环。所以谁能掌握高质量真机视觉数据，谁就更有机会把握这一轮物理AI的机会。视觉方向主要推荐奥普特、凌云光、天准科技（"眼睑科技"按原文应为天准科技）。

奥普特的布局可从感知硬件、算法与数据、执行与协同、场景与生态四方面展开。第一层感知硬件，即物理AI的眼睛：公司自主产品线已全面覆盖视觉算法库、智能视觉平台、工业AI光源、光源控制器、工业镜头、工业相机、读码器、3D传感器、测量系统及工业传感等，是国内少数能够把机器视觉核心软硬件做得相对比较全的厂商。第二层算法与数据，决定采下的数据能不能变成可训练、可回流资产：目前已形成以Deep3、DeepCloud和DSC为核心的工业视觉AI软件平台体系——Deep3面向本地设备与现场工作站，覆盖数据管理、标注、训练、评估到部署；DeepCloud面向企业级多工厂应用，承担模型版本管理、算力调度、远程运维以及边缘设备批量下发；DSC面向工业数据不足、异常样本稀缺的问题，提供高保真工业图像与缺陷生成，依托公司千万级工业高质量数据积累，可对缺陷类型、位置、形态、数量、纹理和背景条件做可控生成。第三层执行与协同，决定把"看得见"推进到"做得准"：公司此前控股收购东莞泰来切入精密运动部件，构建从视觉感知到运动执行到智能决策的一体化产品体系，2026年上半年东莞泰来贡献营收约9590万元。第四层产品与生态，核心是检验前面积累的能力有没有真的跑出来：今年5、6月份，奥普特跟越疆科技开展了一系列合作，5月签署战略合作协议，双方基于奥普特机器视觉全栈技术能力以及越疆的软硬多体具身智能平台，共同打造"手眼脑"一体化的具身工业解决方案，并在生态共建方面融合越疆的数据采集平台和奥普特的视觉数据能力，建设工业具身智能数据集。

凌云光主要卡位物理AI的"数据+训练+测评"的基础设施，并主要通过动捕系统打开生成空间。从产品维度看，凌云光已形成视觉器件、视觉系统、智能装备以及工业软件矩阵，并通过收购JAI强化高端成像与海外渠道；子公司元客视界（按原文）可以通过FZMotion光学动捕系统切入机器人感知数据底座。解决方案维度，已构建多模态数据采集、机器人运动性能评估、机器人"大脑"训练三大方案，并在光学动捕基础之上融合多视点视觉、第一视角设备以及力反馈手套，实现从数据采集、动作训练到真机部署的链路覆盖。商业化维度，动捕系统也有相关收入，2025年FZMotion收入同比增长70%，2026年上半年FZMotion收入继续同比增长80%以上，表明这块需求、商业化做得比较不错。生态维度，凌云光跟德塔智能研发全身大模型相关技术，并跟华为云合作开展CloudRobo全量数据采集合作，有望强化其平台兼容性以及客户覆盖。

天准科技通过收购华兴智能，以全栈自研的视觉硬件为底座，在物理AI的数据入口环节完成了硬件到算法到数据到场景的闭环，这一点跟奥普特很像。硬件端，其视觉光源、工业2D相机、工业3D相机、读码器、工业镜头、视觉控制器等核心硬件全部全自主全套研发；软件算法端，自主软件Smart X（按原文）集成了标定、定位、测量、识别、缺陷检测与深度学习分类等功能，再配套视觉控制器，形成硬件加软件加方案的交互闭环；数据端，公司已在光伏、医药的缺陷检测，锂电焊缝检测等AI质检项目中跑通从样本到训练到部署的项目积累，其全量样本与算法迭代也构成真机数据侧的积累；应用端，华兴智能目前覆盖3C、新能源汽车、半导体、智能物流、医疗设备等多个行业，同时还在往半导体光控方向延伸。

触觉方面，触觉是物理AI从"看见世界"到真正能够"操作世界"的关键闭环。视觉告诉

机器人物体在哪里、长什么样，但只有触觉才能告诉它能否接触、接触后发生了什么、力是否过大、是否打滑、材料软硬如何、是否已经抓到位，这是机器人从演示级动作走向工业装配、家庭服务抓取的核心分水岭。从产业演进角度看，当前物理AI已经在视觉、语言、规划层面快速进步，但真正制约后续机器人物理AI落地的并不是能不能理解任务，而是接触瞬间能不能做好动作——机器人知道去抓杯子，和真正稳定抓住杯子，中间差的是接触反馈；知道要插入零件，和真正完成精密装配，中间差的是力控以及位置的微调。从数据属性来看，触觉的重要性还在于
它的稀缺性：视觉数据和语言数据可以从互联网大规模抓取，但触觉数据只能来自真实的接触和真实的受力过程，没有办法被纯仿真或网络爬虫轻易替代。所以后面机器人训练的语料不仅仅是视频，而是带接触、带受力、带材质信息的多模态交互样本。
触觉方面，我们重点看好一个新方向——数据采集手套。数据采集手套是连接人类操作经验与灵巧手操作模型的关键接口，短期受益于整个具身智能数据的建设，中期来自于灵
巧手以及后续复杂操作任务的放量。主要受益标的是达诺电子（"达入电子"按原文），主
要凭借在小微电子功能件、结构性器件、3C装配自动化领域的相关能力，把相关产品应用到数据采集手套环节。

会议实录：
主持人（策略）：各位投资者，今天由电信首席刘小龙总、中小盘首席周佳总以及汽车首席邓建全总，共同给投资者分享和汇报物理AI的投资机会。
从策略角度来讲，为什么在这个时候我们认为物理AI可以进入大家的投资目光呢？因为市场确实在一个剧烈震荡的过程中，资金面的承压是近期市场波动的核心原因之一。但上周国有资金通过大额申购ETF承接抛压，向市场注入了一定的流动性，包括沪深３００ＥＴＦ（华泰柏瑞）上周
申购净流入大概２７８亿元，科创５０ＥＴＦ（华夏）上周申购净流入１７０亿元，创业板ＥＴＦ（易
方达）上周申购净流入２１０亿元
。我们认为在这样的背景之下，市场底部确实可以初步确认，
但我们一直都在强调，目前仍然在反弹后高波震荡的框架之内、仍在再平衡的框架之内，后续并非线性上涨。
一方面，筹码集中度相较7月份大幅改善；另一方面，美债利率下行，全球宏观流动性风险得到一小部分释放。但也要注意，近期长短美债利率下行的背后，其实有一部分是对AI趋势和回报率下降担忧所引发的避险，就是分子端出现了逻辑上的一些问题。所以从投资逻辑上讲，微观结构虽然已经有了一定程度的缓和，国有资金承接抛压、资金面压力有所缓解，但目前的筹码结构和催化角度仍然处在再平衡、反弹后高波震荡的框架里，后续并非线性上涨。在这样的有机会反弹的框架下，我们需要重点关注二次点火、具备叙事张力的方向，而物理ＡＩ方向正好
一方面筹码比较干净，另一方面确实具备二次点火和叙事张力的结合。
物理ＡＩ是ＡＩ的下一个应用方向，具备比较大的增长潜力；它也是机器人的智能核心，
能让机器人理解物理世界并做出合理的行动。
为了更深入地探讨当下到底什么是物理AI、它
的投资机会在哪里，我们邀请到了开源研究所六大行业首席，给各位投资者详细分享物理AI的投资机会。
计算机首席张悦：各位老师大家好，感谢大家聆听本次电话会，我是开源计算机的首席张悦。今天我主要先做一个开场，讲一下整个物理AI投资的脉络或者逻辑，包括我们看好的方向和接下来可能存在机会的标的。

第一个是物理AI大逻辑，我认为这是一个大的产业趋势。
为什么是大的产业趋势？第
一，全球所有科技巨头都在做这件事，不光是谷歌、Meta、OpenAI，还有特斯拉，以及国内的字节等等，大家都会下场去干物理AI。所以世界模型是一个全球共识，不是某一家公司单独去
想的全球工程，每个科技巨头都在做，这一定代表着科技下一个发展的方向。

第二，物理AI本质上是一个产业的变化。物理AI最核心的有三点：第一是世界模型，第二是VLA，第三是数据原生或数字仿真，这三者共同构成物理AI。
物理AI的定义，实际上是对
所有物理世界发展规律的理解，它要求载体（机器人也好、车也好、机械臂也好）根据当前物理状态，判断下一步可能会发生什么样的物理状态变化，并据此做出动作上的反馈，这就是物理AI。
物理AI这次根本性的变化，是因为大语言模型在一定程度上突破了物理世界与文字世
界、或者说数据世界的边界。我们以前讲的通用大语言模型是一个数据世界，不管是OpenAI、Claude还是GPT，本质上都是一个大语言模型，它要做的是根据上文的文字输出预测下一段文字，是一个推理结果，跟物理AI完全不一样，完全不是一个训练体系：第一，数据不一样；第二，训练方式不一样；第三，输出结果也不一样。所以在这之前，通用大模型完全不可能切入到物理AI的事情上，或者说特斯拉这一条体系、与智谱、MiniMax这些是一套体系，两者中间是完全隔开的。
但随着GPT-6 Astra的发布，通过大量数据的投喂，也包括一些物理专门知识，再加上把它的结果转化成电机驱动的信号，再加上这个环节，基本上就能够解决、完全替代VLA的一些部分功能。今年7月份，发布完Astra之后，有些具身智能的人员使用了Astra大模型，放在宇树的机器人上，已经成功完成了一些命令的调用和执行。所以从这个角度来看，它意味着什么呢？意味着全球所有的大语言模型公司都可以下场去做世界模型或者具身大脑，在这种情况下，每
家公司必然就会投入重注，为什么呢？因为以OpenAI或者OpenAI为例，每家公司每年的资本开支基本都是千亿级别、美金级别，这个资本开支非常庞大，未来的空运这块，每家已经进入到一个群雄逐鹿的状态。比如在模型这块再做出几倍的增长，其实难度相当大；所以下一步如果模型能力再提升是瓶颈，那么把钱放到物理AI或世界模型这块，会打开这些公司更大的市场空间，这也是每个大模型公司都要发力、或者说下注去做的一件事，所以全球都在做这件事。
第二，刚刚说的OpenAI的事件模型，大模型可能会做事件模型，要解决哪些问题？第一，它用的还是Transformer架构，这跟原有体系是类似的；不一样的就是训练的数据。这里的数据是一个非常大的空间，因为每家包括OpenAI、包括我们跟梁（文锋）也说过，包括字节等，他们是没有数据的；字节是有视频数据，视频数据可以进行投喂，但这还不够，最好的数据应该是：第一是真机数据，真机数据是最真实的；第二是Ego数据；第三是UMI数据。这三块数据再加
上数字仿真，构成了整个数据体系。这个数据是未来每家的一个核心生产要素，谁能拿到更
高质量的数据、谁能拿到更多的数据，就能够做出比别人更好的世界模型，这是每家要加注
的点。

对于一个大模型公司来说，数据怎么去获得呢？第一种就是Ego数据，你要么得让第三方公司给你做Ego数据的采集，要么自己做Ego数据，就是买很多Ego设备给你做。所以第一条围绕数据的投资思路就是Ego数据的采集，目前Ego的一级供应商，比如蜜蜂
做得比较好；在蜜蜂这
个环节，它的口袋式、手上的采集设备，需要加上大量的传感器，这些传感器国内有很多供应商

在做，比如做六维力的、做IMU的、做视触觉的公司可能都会参与这一部分。第二个思路就是真机数据，真机数据第一你要有本体，这一块比较利好像宇树、上纬新材，包括中天科技等等做本体的公司，因为做本体的公司有本体，就可以拿它的本体去采集真机数据。第三个思路就是做数据孪生、做测试的合成，这一块也有比如像索辰、风语筑这些公司，他们也在做1:1还原真实设计的数据，这也是一个核心方向。
总体来说，整个数据这一块的思路是比较明确的。数据市场非常大，因为世界模型能力的提升理论上没有边界，那么数据的规模就没有边界；你用每年几亿、或者十几亿、二十亿小时去衡量其实都可以，但很难去做详细的测算。我认为每家的基本数据都会在大几千万小时的级
别，那么这种情况下，对应的数据市场就有一个非常大的增量环节，这是我们看好的第一个方向，必须要去解决数据这块的问题。
第二个方向是执行这一端。我们认为整个物理AI最终其实利好的是特斯拉，因为特斯拉有全球最好的数据，它的弱点在于大模型这块，但是OpenAI现在相当于把模型这个问题解决了，所以未来特斯拉的大脑应该会做得非常非常好，或者说它在世界模型这块，肯定在全球做到最好之一。所以这其实是更利好特斯拉的产业链。特斯拉在执行这端基本上领先其他公司3到5年，所以未来它商业化的前景反而会变得更加明确，执行端公司的业绩确定性和弹性会变得更加确定。
第三个就是，未来OpenAI、Meta等他们都会做机器人，所以他们会找新的供应链，新的供应链跟特斯拉大概率重合度不是那么高，因为现在给特斯拉做机器人执行端的产能基本都打满了，其实没有空余产能给其他厂商去使用。所以在这种情况下，我们认为国内也会诞生很多OpenAI、以及后续比如Meta的供应链公司。目前OpenAI这一块的方案基本上确定下来是两种路线：第一种是下半身是一个三角支架，上半身是一个半身的人，然后有两只手臂，每边手臂大概3个
手指，这个方案目前基本上已经逐步在进行了。这个机器人目前它的手臂是外购的，但是手是整个OpenAI自己去把控的，它主要是用这个机器人去采集一些真实的数据；那么明年的话，它可能会推出自己的整体方案。
除此之外，我们认为包括字节，包括很多公司可能都会下场去做机器人，所以会带来很多相关的机会，包括小米也是。所以其实大家都是在看特斯拉的V3，可能V3.1发布之后，很多公司就会模仿着特斯拉的路径去走，类似于新能源时代。所以我们也看到明年国内、全球的人形机

器人可能是有非常大的爆发的
：今年上半年国内整个量、全球人形出货量可能也就三四万台，
今年全年、明年可能就是一个六七十万台的全球量级，基本上是快十倍的增长，明年才是一个开始快十倍增长的行业，所以很多确定性能够参与这个环节的公司的业绩都是充分受益的。
整体来看，我们看好的标的也从刚刚几个环节去分享，我们也有一个大概的梳理。第一块是关于OpenAI链，目前我们认为从传动、微型传动的角度来说，中大力德
目前做得是最好的，
在微传动这块绝对是领先很多其他公司的，所以这应该是绕不开的一个环节。第二个就是利群精密，因为OpenAI后面可能有很多公司能够参与到这个环节中去；其他的就是包括创新精密，可能通过天机的环节去参与OpenAI。除此之外，我们比较看好数采这个环节
，数采环节我们认
为一个是你要么有Ego出货量的能力，要么是你的传感器市场份额很高，要么价值量很大，这一块我们觉得奥比中光是比较好的一个标的；除此之外，还有一些其他做传感器的，像福莱新材、安培龙这些环节，可能也会比较受益。
其次就是关于气动（丝杠）这块，我们认为目前从市场份额和与市场率的价值角度来说，科森、恒立、福赛目前是非常明确的、进入到逐步业绩兑现且非常核心的环节；目前整个预期差最大的可能还是像光洋股份这些新的、在其领域具有绝对竞争力的公司。所以整体来看，我们觉得物理AI的投资思路是比较明确的。后续我们认为市场可能会交易的环节，第一个是手部

的灵巧手，因为需要灵巧手去采集数据，灵巧手是一个重点要解决的方向，所以灵巧手可能是一个比较好的、后续市场去挖掘的思路；其次就是有高质量数据的公司，因为有些公司并不是采集数据，它本身就拥有高质量数据，可以作为数据价值并列，类似于中控技术这样的公司。整体来看，我们看好物理AI板块的投资价值，我们认为这是一个全球性共振的趋势。我这边大概是这样观点。

电子首席刘小龙：各位老师晚上好，感谢大家参加今天的会议。我是开源证券首席分析师刘小龙，我这边汇报一下物理AI，包括端侧数采等这个领域对于我们电子板块的一些积极变化。对我们电子板块，可能有两个细分领域此前偏冷门。第一是数采装备的便携式电池，它跟这个大家理解的端侧电池还是有一定区别，但确实有一定相似性，同时生产厂家也是交叉的。第二是数采领域涉及到的机器人皮肤等变化。

首先就是端侧电池。因为数采其实此前很多是插线式的，拖着一根长长的绳子。但今年很多机器人公司，虽然说大家看到的整个产业链感觉有一些兑现度的问题，但整体上数采领域是很多机器人本体厂商或者说一些下游应用企业，在前期开发模型、包括做测试的时候一个硬性的刚需。这一块我们像之前了解到行业里面的公司，包括像豪鹏科技，然后也包括一些相关公司，其实都做了很多端侧上的布局。尤其豪鹏今年整个机器人板块的电池收入可能大几千万、接近一个亿，这里面主要的收入来源就是数据采集，市占率也相当高。数据采集这块领域，因为对下游客户来讲现在处于一个成本不敏感的。所以这大几千万的收入带来的利润率水平是非常高的，后面的增速也非常快。所以这块对整个消费电池板块来讲，还是一个非常好的新的业绩增长点。因为大家之前可能更多的预期是机器人本体上电池的放量，但现在领先于他们一步的是数据采集领域。因为对于广大的厂商来说，使用这种带电池的设备肯定是更加方便、更加便利，采集到的数据，包括准确性等等，可能还是会有一定的提升。所以我们觉得这块可能是市场之前了解得比较少，重点公司主要就是豪鹏科技，最近表现也不错，另外还有珠海冠宇。然后这两个公司同时也有端侧电池的概念，像穿戴类的眼镜、戒指，包括一些其他可穿戴设备，最近市场关注度比较高，像OpenAI等等都有很多新的应用。随着AI的发展，这些新的设备、新的硬件、软件都被开发出来，大家肯定需要便利性，主要还是靠小型化的电池、高功耗的电池或者说掺硅含量比较高的电池，所以这里边端侧电池整个是有一个明显的通胀逻辑。这是第一大板块。

第二块就是数采里面可能像皮肤，皮肤其实也是大家之前关注度比较高的一个细分领域了，但最近确实因为随着数据采集、前期的一些数据需求，皮肤也在里面充当了一个比较重要的角色，比如温度、压力、仿生等等这方面的一些新变化，其实跟刚才我们提到的数采应用电池是一样的逻辑。所以在数采领域，优先放量，包括这些皮肤等等下游的应用。所涉及到我们这边比较相关的公司，包括像福莱新材，然后像汉威科技，最近的表现也是相当亮眼。这块我们觉得整体上跟之前我们在看机器人皮肤的时候，预期差并不是特别大，但是行业其实有很多积极的变化，包括海外的一些大的展会、海外一些客户的积极对接，所以这块我们觉得肯定还会继续，随着整个数采包括物理AI领域的爆发，后面还是会迎来更好的发展。

整体上，因为现在这些跟板块相关的一些标的，前期也经历了很大的回调，那无论是从筹码还是从基本面的角度，很多票都跌到一个非常深的位置，还是建议积极关注。以上就是我们电子板块整体的汇报，如果对个股有任何疑问，欢迎下来再联系我们，谢谢。

机械行业分析师：各位领导好，我给大家汇报一下机械设备这边对物理AI的一个看法。我们认为这一波的物理AI行情，可能主要来源于两个方面。

第一个方面就是来自于OpenAI的GPT-6Astra这个新的模型，从功能角度来说，它突破了之前大语言模型的一些局限。对于物理世界的理解，之前的模型是比较差的，而Astra这个模型不仅是对物理世界的理解加成了很多，它对于驱控层面上、具身智能驱控层面上的作用也相对

来说是比较大的。经过各个本体公司的实测下来，是解决了很多之前模型所不具备的、对于具身智能这样一个驱控作用的比较好的模型。应该说我们之前和硅谷的VC，投具身智能和投大模型的也做了一些交流，大家会觉得这个对物理世界理解的突破，实际上是一个断点式的突破，它需要

大量的训练才可能出现的基点变化。而Astra的出现，是在时隔两年之后，大家在这个具身
智能相关的模型上，发现了一个可以说断点型的突破，所以有人把这样的变化称为是一个
基点，就物理AI的一个基点。
往后看，我们觉得它的出现并不仅仅是这么简单，而更多的可能会激发起其他的模型公司，包括OpenAI后期对于物理AI训练这块的投入，可能会形成一个生态式投入。毕竟物理AI肯定是未来长期AI应用最大的一个方向，和数采最大的一个方向。我们认为有可能会带来后面1~2年整个物理AI模型的一个突破进展。我们看到王兴兴也在最近发表言论，认为物理AI未来1~2年可能会被击穿，采用"击穿"这个词，其实就是我们认为的大规模训练下的一个断点突破。
在这样的变化之下，市场现在对于整个机器人、具身智能板块的理解，就从之前的只是
做展示性或者娱乐性的、简单任务的机器人，变成了有可能会具有一个类似于手机里安卓
系统的这么一个操作平台的物理硬件
，换句话来说就变得真正智能化了，并且可能会有相对比
较统一的智能化的平台系统。从长期来看，如果这样的一个模型是存在的、能够持续迭代下去，那么本体的使用价值会大幅提升，本体的智能化程度在长期看也会大幅提升。所以我们认为第一首先是受益的是本体厂商
，这也是为什么宇树、上纬最近涨得比较多的核心原因，就是因为
他们把身体训练好了，那么大脑交给Astra或别的大模型去驱动，这是非常好的一个分工。
其次，我们看到本体的公司当中，现在出现了一些垂直类的本体公司，市场也非常关注，比如在服装缝纫领域的杰克科技，比如我在这里给大家推荐的融创智能，是在包装、简单装配领域的。为什么这些小的本体公司大家很关注呢？我们认为一个核心的原因，就是因为大家会认为在目前数据量相对来说有限的情况下，去做一个通用类的大模型、或者大范围通用的落地是挺难的一件事情。所以具有闭环的数据、私域的数据，同时在专用行业里面专用设备能力比较强
的这部分公司，是具有稀缺优势的
。顺着这个逻辑，我们会认为就像杰克、融创，甚至像其他
有一些专用设备的东西，比如复星、华商达，这些都是值得重点去看的，包括国储。
另外一个方向就是特斯拉链。因为特斯拉毫无疑问是把本体训练得最好、复杂程度最领先的一家。气电（T链）最近的产业变化也是有点大，它近期也是在中国产业链加单，周产100台往周产几百台去靠，按这样推算的话，到明年预计特斯拉可能会在全球拥有1万台以上存量的
OptimusV3机器人
。这样一个大的变化，核心其实是为它自己明年的数采和量产去做比较大的
准备。所以T链的公司，我们觉得也是近期市场非常关注的。因为T链公司比较多，我们还是比较推荐在T里站位比较好、同时SKU占比往上走的，比如像福莱新材、思林，这两个定点的比较多；同时，比如像三花智控、恒立，这两个我们认为是大的T链中心，也是非常值得看的，这是第二类T链的公司。

第三类，因为长期看模型的发展，数据的采集非常重要，所以数采类的公司最近我们
认为也是非常好的。一方面是它的传感器、电子皮肤的公司，一方面是专门做数采的公司，后者可能有个标的是华丽科技，值得重点关注。前者我们认为除了市场现在炒得比较热的福莱新

材之外，像天安新材、大美股份这一类的竞争公司也是被严重低估的，值得重点去关注。好，那我就主要汇报这里，谢谢。
军工行业分析师：各位领导好，汇报一下我们军工板块目前物理AI相关的一些应用。我们认为主要在三个方向上有比较大的一个应用前景：一个是装备的正向研发，另外是一些训练、合成的数据，包括像数字孪生等等，另外就是战场的具身智能，包括以前我们看到的一
些比较典型的无人作战平台，未来也是要向具身智能去转变的。这三个方向是我们看到的物理AI能够在军工里面有巨大应用前景、比较大潜力的应用方向。
首先是装备的正向研发。因为目前其实很多军工装备研发周期相对来说会比较偏长，整体来看因为需要巨大的物理量输入，因此在前期整个研发会占用比较多的时间成本。但是有了物理AI之后，其实我们在一些飞行器的高速飞行，包括气动布局等方面，还有包括一些新材料的开发，还有可能跨尺度的研发，这些如果能够借助物理AI的输入的话，其实是能够巨大地提速整个研发进程的。还有包括我们目前看到很多行业内已经开始用一些物理AI的方式去设计隐身，还有包括电子仿真等等，已经能够在某些领域看到一些应用的雏形。这是我们在研发这端就能够有一个比较巨大应用前景的方向。
第二个方向，其实我们看到是一些已经有部分领域应用的数字孪生，包括像能科科技之
前做的一些在生产端、还有包括在一些实战端，可能有了一些数字孪生的技术。军工的很多模拟仿真，其实之前也积累了相当宝贵的数据，而整个模拟仿真之前最大的问题就在于，我们没有办法给一个非常准确的物理量，相对来说是比较受局限的。而物理AI如果能够突破这层限制，我
们能看到其实在很多数据荒的领域，可能就会有比较大的突破，包括可能能够提供更准确的
红蓝方数据，还包括一些无人作战等方面的指挥决策数据，它能够根据实时的物理环境，可能做到更准确的输入。这个是能够精准地去检验一个武器装备是否有效、是否达到作战效能的一个非常重要的应用方面。我们觉得这个也是当下我们看到某些公司已经做了一些相对来说比较有前瞻性的方向。
以前我们看到可能更多的企业都是在做一些比较简单的数字仿真，还有包括一些可能训练用的模型等等。现在其实如果有物理AI加入的话，可能会把我们模拟仿真的数据提高到另外一个维度，那增强的模拟仿真在整个军工系统的开发、包括后续的应用，还有验证过程当中的一些比较重要的应用，能够极大地缩短或者是极大地去提高军工整体的效能，目前来看是有比较大的一个经济价值的。

最后一块应用，我们认为还是在具身智能上面，包括装备本身的一些智能的交互，还有包括可能无人作战平台等等这些装备他们本身的一个协同，这个也是我们认为能够率先落地的一个方向。目前无人作战其实已经变成全球比较重要的一个主题。我们经过这几次比较大的冲突，其实可以发现无人装备平台已经显示出了一个巨大的经济价值，包括可以不对称的打击，以及能够在非常低的战损情况之下，去执行一些比较复杂的一些任务。此前其实市场上比较关注的像一些无人战机，包括僚机、巡飞弹、蜂群，以及地面上的一些无人狗、无人车等等，这些都是在去年九三阅兵之后，市场比较关注的一些方向。如果有物理AI的加持，能够可能将这些无人平
台本身的作战效能成倍放大，我们一般都会把这些平台叫做倍增器，物理AI当下可能就是
最重要的一个倍增器。它可以可能将雷达、包括周围的一些数据传输的平台，集成到某个具身智能体上面，能够实现整个装备的互联互操作，这个可能能够彻底地改变当下战场的一些形态，也是我们认为当下整个具身智能在军工这个领域、在无人平台上面可能够落地最快的一个方向。
此外我们也认为整个"十五五"期间，整个无人装备可能都会有望快速地发展，整体的装

备比例还是能够有望显著提升的
。去年的九三阅兵已经显示出了像机器狗、无人机，还有UUV
等这些大型的新型的无人装备。但其实我们跟海外相比，仍然有一个比较大的差距，市场其实也比较关注像海外的美股的公司，像弗兰，他们在去年已经开始陆陆续续地获得一些比较大的订单。那么当下如果物理AI我们跟海外相比的话，可能又会导致整体研发、还有包括具身智能上的应用，可能又会拉开一个比较显著的差距。我们可能也认为，"十五五"期间，无人机可能不会是大家想象中的那么简单的一个平台，它可能会承担更多的具身智能任务，也会从原来可能比较辅助的一个装备，像核心作战力一样去推进。所以这个板块我们觉得也是可以建议市场重点关注的。
那相对来说比较受益的板块、受益的标的，包括一些原来做微波的、做无人的，包括可能像
国睿科技、高德红外，还有包括像复旦微电
等等这些公司，我们认为是当下可能率先能够落地
的。而数据应用这方面，可能能科科技是我们目前看到的在物理AI上面和军工密切合作最多的，而且可能是有望能够成为数据上比较重要的一个支点的应用方向。我们军工板块的汇报大概就是这些。
汽车首席邓建全：各位投资者大家晚上好，我是开源汽车邓建全，我今晚给大家汇报一下整个汽车板块在这一轮机器人行情上面，它的核心的一些变量。这两天机器人板块、乃至零部件板块，其实开始从底部有一定程度的反弹了，我们觉得在当下去看整个机器人零部件、乃至整个汽车板块的零部件的话，还是有一些值得大家去关注的方向的。

首先，核心的催化就是因为物理AI大会，包括王兴兴也提到，在接下来的一两年，整个物
理AI大概率会迎来一个比较重大的突破时间窗口
，这是核心的催化。我们物理AI这个领域，
汽车板块的核心落脚点主要是两个：一个是机器人，一个是智驾
，这两个应该也是率先在AI应
用层面率先落地的。因为特别是智驾，已经从过去三五年以来发展得比较快了，到2027年也有很多L3的车型逐步落地的预期了。所以我们觉得物理AI这个层面，机器人加智驾应该是两条主线了。
在这个背景上面，除了我们觉得以往上一轮机器人行情炒的一些硬件类的零部件这些方向之外，我们觉得也有一些新的方向可以给大家做一个提示。第一个是数采，数采这个方向我们认为可能也是这一轮、未来三年可能要重视的。当然数采可能是偏计算机行业看的，但我们汽车行业也有一些公司在这个数采维度有一些布局。
第一个是阿尔特，大家可以重点关注一下。阿尔特本来主业是做第三方的整车设计研发，在物理AI这个方向，它布局了两条线：第一条线是跟光轮智能的合作，成立合资公司，一块
来做仿真以及物理AI数据的收集以及测评；第二条线是跟富士康成立合资公司，做整机的
机器人代工。这两条线其实也是比较符合当下市场对于物理AI应用层面落地的两条主线。
阿尔特为什么能够往这个方向去做，以及它在这个方向上面的一些优势，也是它布局的这个原因。因为它作为国内规模最大的第三方独立汽车设计公司，在整车设计这个维度，实际上有比较多的一些数据，它手上也有比较多整车研发，包括它也有一些零部件，也有很多实力收集回来的数据的优势，这也是它布局这个方向的核心优势之一。
它这两条线，我先讲第一条线。第一条线就是跟光轮智能以及宇信科技成立的合资公司，也是在今年6月25号签约成立的。首先光轮智能，大家都知道它主要是做仿真数据的模拟，它的优势在这里；而阿尔特提供的是包括汽车制造场景数据的收集，以及用收集回来的真实数据去做一个仿真的校准，所以这是阿尔特的一个优势。在这两个优势互补下面，他们未来去做整个数采的维度，应该也是有一定可以展望的空间的，这是它的第一条线。
第二条线，就是跟富士康成立的这个叫富特康的合资公司，也是在刚刚的8月份签约正式成立。成立这个富特康，它主要是要为整个产业形成一个闭环，因为刚刚我们讲的数采之后，它最

终还是要在这个机器人维度去做各种的校准，所以这也是最后商业闭环的最后一环。这个富特康也是延续了阿尔特之前的优势，阿尔特之前在汽车领域，它设计整个汽车之后，是达到交钥匙的全服务的；那么现在这个机器人，也是能够做到交钥匙的整个机器人方案，从硬件、软件，从数据反馈等等，都形成了最后的一整套商业服务。所以我们觉得阿尔特在这一轮的数采乃至整个物理AI应用领域，是可以关注的。

其实我们也看得到，阿尔特在今年的上半年已经开始有一定程度的回暖了。主业之前因为受到一些客户的信用贬值的问题，导致了过去两年的业绩有点承压，但在今年已经开始回暖了，像今年上半年它的规模利润也是同比实现了一个大幅度的提升。这是第一个给大家推荐的。
第二个就是中国汽研。中国汽研首先它的主业主要是做汽车检测，就是新车上市公告前的这种检测，也包括各种仿真测试。所以中国汽研作为国内体量比较大的汽车场景数据采集、仿真测评的检测机构，它手上拿的包括智驾的数据，应该也是能够比较好地向物理AI领域去延展的一个基础。对于中国汽研来讲，首先第一个，因为它有一个仿真云平台，所以它在仿真数据的积累这个维度，应该是有比较多的，所以它不管是在智驾领域的落地，还是说在机器人这个领域，应该也是能够延续它过往的比较大优势。所以我们觉得这个时间点去关注中国汽研，也是一个潜在的催化。
当然中国汽研本身，除了汽车领域的检测之外，它也有机器人的检测。因为本身我们也看得到，中国汽研其实旗下也有机器人的检测公司，是在重庆，是国家级别的机器人检测与评定中心，这个评定中心也是中西部唯一的一个国家级的机器人第三方检测机构，它检测覆盖的范围也包括原型机器人、工业机器人、AGV、特种机器人，包括一些核心零部件像灵巧手这种检测，它都能够去服务。所以它本身也是有机器人领域的布局的。所以我们觉得综合来看，中国汽研作为仿真数据比较有优势的整个检测机构，在同时有机器人检测机构的背景下面，这个公司也值得大家去关注一下，这是我们认为的第一条线。
第二条线，因为机器人在迎来逐步商业化量产这个过程中，像宇树最近股价涨得比较多，我们觉得除了刚刚讲到的数采这个方式之外，整机代工也是大家可以关注的。整机代工，
我们整个汽车板块也有一些公司在不同领域，有些是工业机器人，有些是人形机器人，或者一些是轮式的，在这个维度也有一些布局。像刚刚提到的第一个是阿尔特，因为阿尔特跟富士康成立合资公司之后，他也会去做整个人形机器人的整个代工。包括像零部件里边的浙江先通，那么
浙江先通之前也跟浩瀚星空、启腾、机建这些成立控股合资公司来布局整个机器人的代工。像均胜电子，均胜电子旗下的均普智能，也有一些机器人、包括人形机器人的业务布局。那像还有中鼎股份，中鼎旗下也有这样的一个子公司来做这种整个机器人的代工的。所以我们认为在AI应用领域落地的这么一个变化过程中，整机代工这个维度也是可以引起大家的关注的。
所以我们整体来看，就认为在当下物理AI应用领域，除了之前大家可能关注度比较大的这些零部件之外，数采以及整机代工这两个维度，也是可以建议大家重点去关注的。那么以上是我今天主要给大家汇报的主要内容，如果大家感兴趣的话，欢迎跟我们汽车团队进一步联系沟通，谢谢大家。

中小盘首席周佳：各位投资者大家晚上好，我是开源证券中小盘的分析师周佳。关于物理AI，我们目前主要看好视觉跟触觉这两个方向。
视觉方面，我们认为视觉是物理AI最关键的感知入口之一。因为物理AI要解决的不仅仅只是生成内容，而是需要让系统在真实的物理空间当中完成对物体、场景和动态关系的感知、理解与交互。而在全部的各种传感的模态里面，视觉几乎是唯一一个能够同时提供高空间分辨
率、丰富语义信息、连续实时信息和非接触测量能力的感知方式。
进一步展开来讲，在多种传感器里面，视觉之所以能够占到核心的位置，不是因为它其他模态不重要，而是因为视觉的信息密度最高、覆盖面最广。因为对机器人而言，力觉、触觉可以告诉它接触到了什么、接触力是否合适，惯性器件可以告诉它自身姿态和加速度如何变化，激光雷

达可以提供距离和轮廓，但这些模态往往难以单独回答，比如眼前到底是什么物体、边界在哪里、是否存在缺陷、姿态如何、是否已经发生了一些细微的形变等一些问题。尤其是在一些工业场景和具身场景里面，往往要求同时理解类别、位置、姿态、深度、表观缺陷、相对关系和动态变化，这些都是视觉擅长的领域。

那么更重要的就是，物理AI所需的数据也不是一个孤立的图片，而是连续、多模态耦合、携带物理量的一个真实的交互数据流。在这一点上，视觉既能够生成2D的图像，也能够通过结构光、双目、ToF点云、动捕等方式提供3D的空间信息，还可以和IMU反馈处理数据进行融合。

再从产业链逻辑来看，视觉的价值已经从此前工业自动化里面的一个反馈的工具，逐步升级为模型训练的一个数据入口。这也意味着后续如果企业竞争，不再单独只是PK你的相机、镜头的性能，而是PK你能否稳定采到高质量数据、能否高效标注和训练、能否沉淀行业的经验、能否在一些真实的场景当中持续回流样本，并且形成一个部署的闭环。所以后面是谁能够掌握高质量真机的视觉数据，谁就更有机会把握这一轮物理AI的机会。

那么在物理AI视觉方向，我们主要推荐我们团队覆盖的三个标的，包括奥普特、凌云光、天准科技。

第一个是奥普特。奥普特在物理AI方面的布局，我们可以分别从感知的硬件、算法与数据、执行与协同、场景与生态这四个方面去展开。第一层感知硬件，这是物理AI的眼睛，根据整个公司公告的数据，奥普特目前的自主产品线已经全面覆盖了视觉算法库、智能视觉平台、工业AI光源、光源控制器、工业镜头、工业相机、读码器、3D传感器、测量系统及工业传感等等，它是国内少数能够把机器视觉核心软硬件做得相对比较全的一个厂商。这是第一层。

第二层是算法与数据，因为它决定你采下的数据能不能变成可训练、可回流的资产。而奥普特目前已经形成了以Deep3、DeepCloud和DSC为核心的工业视觉AI软件平台体系，比如Deep3主要是面向本地设备与现场工作站，覆盖数据管理、标注、训练、评估到部署；DeepCloud主要面向企业级多工厂应用，承担模型版本的管理、算力调度、远程的运维以及边缘设备的批量下发；DSC主要是面向工业数据不足、异常样本稀缺的问题，提供高保真的工业图像与缺陷生成，然后依托整个公司千万级的工业高质量数据积累，可以对缺陷的类型、位置、形态、数量、纹理和背景条件做一个可控的生成。

那么第三层就是执行与协同，它主要决定的就是把看得见推进到做得准。奥普特在这一块，它核心就是之前控股收购的东莞泰来切入到精密运动部件里面，构建从视觉感知到运动执行到智能决策的一体化产品体系，2026年上半年东莞泰来贡献营收大概9590万元。

然后第四层就是产品与生态，这个核心是检验你前面积累的能力有没有真的跑出来。这核心就是整个今年5月份、6月份，奥普特跟越疆科技开展了一系列合作，奥普特跟越疆在5月份签署战略合作协议，双方主要是基于奥普特的机器视觉的全栈技术能力，以及越疆的软硬多体的具身智能平台，双方共同打造"手眼脑"一体化的具身工业解决方案。而且还要在生态共建方面，融合越疆的数据采集平台和奥普特的视觉数据能力，建设工业具身智能数据集。这是奥普特的逻辑。

第二个就是凌云光。凌云光主要是卡位了物理AI的"数据加训练加测评"的基础设施，并主要通过动捕系统打开生成空间。从产品维度来看，凌云光已经形成了视觉器件、视觉系统、智能装备以及工业软件矩阵，并且通过收购JAI强化它的高端成像与海外的渠道。另外，它的子公司元客视界还可以通过FZMotion光学动捕系统切入到机器人的感知数据的底座。然后在解决方案维度，凌云光目前已经构建了多模态数据采集、机器人运动性能评估、机器人"大脑"训练三大方案，并在光学动捕基础之上融合多视点视觉、第一视角设备以及力反馈手套，实现了从数据采集、动作训练到真机部署的一个链路覆盖。然后商业化维度，它的动捕系统也是有相关收入的，2025年FZMotion收入同比增长70%，2026年上半年FZMotion收入继续同比增长80%以上，

这表明它这一块的需求、商业化做得还是比较不错的。然后生态维度，凌云光也是跟德塔智能研发全身大模型的相关技术，并且跟华为云合作开展CloudRobo全量数据采集的合作，有望强化它这一个平台的兼容性以及它的客户覆盖。

那么第三家公司就是天准科技。天准科技核心就是通过收购华兴智能，然后以全栈自研的视觉硬件为底座，在物理AI的数据入口环节完成了硬件到算法到数据到场景的一个闭环，这一点是跟奥普特很像的。然后在硬件端，它的视觉光源、工业2D相机、工业3D相机、读码器、工业镜头、视觉控制器等核心硬件全都是它全自主的全套研发的。然后在软件算法端，它的自主软件Smart X也是集成了标定、定位、测量、识别、缺陷检测与深度学习分类等功能，然后再配套视觉控制器，形成硬件加软件加方案的交互闭环。然后在数据端，公司也已经形成了像光伏、医药的缺陷检测，锂电焊缝检测等一些AI质检项目当中，跑通了从样本到训练到部署的项目积累，它的全量样本与算法迭代，也构成了它的真机数据侧的积累。然后在应用端，华兴智能目前终端也覆盖了3C、新能源汽车、半导体、智能物流、医疗设备等多个行业，同时还在往半导体光控方向做一定的延伸。这是视觉方面的三个标的。

然后除了视觉之外，我们后续看好的方向就是触觉。我们认为触觉在物理AI里面，它算是一个从"看见世界"到真正能够"操作世界"的一个关键闭环。视觉，我们刚刚说，视觉要告诉机器人物体在哪里、长什么样，但只有触觉才能告诉它能否接触、接触后发生了什么、你的力是否过大、是否打滑、整体的材质的软硬如何、是否已经把东西抓到位。这也是机器人从演示级动作走向工业的装配、家庭服务进去抓取的一个核心分水岭。

那么进一步展开来讲，从产业演进的角度来看，当前物理AI其实已经在视觉、语言、规划层面快速进步，但真正制约后续机器人物理AI落地的并不是能不能理解任务，可能后面来说比较关键的就是你在接触的瞬间，你能不能做到正确的动作。比如说，机器人知道去抓杯子，和真正能够稳定抓住杯子，中间差的就是你接触的反馈；然后知道去相关插入零件，和真正完成精密装配，中间差的就是你的力控，以及你的位置的一个微调。

然后从数据的属性来看，触觉它的一个重要性还在于它的一个稀缺性。因为你的视觉数据和语言数据，你可以从互联网做一个大规模的抓取，但是触觉数据只能来自于你的真实的接触和真实的受力过程，你是没有办法被纯仿真或者网络爬虫轻易替代的。所以也就意味着，后面机器人训练的语料不仅仅是视频，而是带整个接触、带受力、带材质信息的一个多模态的交互样本。

那么在触觉方面，我们目前重点看好一个新的方向，就是数据采集手套这个新的方向。我们认为，数据采集手套是连接人类操作经验与灵巧手操作模型的关键接口，短期受益于整个具身智能数据的建设，中期来自于灵巧手以及后续复杂操作任务的放量。那么这里的受益标的主要是达诺电子，达诺电子主要是有机会凭借在小微电子功能件与结构性器件、3C装配的自动化领域的相关能力，把自己的相关产品应用到数据采集手套的环节。以上就是我们中小盘的主要汇报，谢谢大家。