

证券代码：688787

证券简称：海天瑞声

北京海天瑞声科技股份有限公司

投资者关系活动记录表

编号：2026-009

投资者关系活动类别	<input type="checkbox"/> 特定对象调研 <input type="checkbox"/> 分析师会议 <input type="checkbox"/> 媒体采访 <input type="checkbox"/> 业绩说明会 <input type="checkbox"/> 新闻发布会 <input checked="" type="checkbox"/> 路演活动 <input type="checkbox"/> 现场参观 <input type="checkbox"/> 电话会议 <input type="checkbox"/> 其他（请文字说明其他活动内容）
参与单位名称及人员姓名	Haitong International Securities Company Limited ChenHaofei 陕西省国际信托股份有限公司 魏志鹏 平安基金管理有限公司 张荫先 中国国际金融股份有限公司 王凯 创金合信基金管理有限公司 张小郭 交银施罗德基金管理有限公司 沈楠 国泰海通证券股份有限公司 杨林 招商证券股份有限公司 彭小泉 荷宝香港有限公司 张子钊 中国光大银行股份有限公司 徐传明 摩根士丹利基金管理（中国）有限公司 李子扬 循远资产管理（上海）有限公司 田超平 平安基金管理有限公司 季清斌 格林基金管理有限公司 刘冬 山东省国际信托股份有限公司 杨卓 新华基金管理股份有限公司 俞佳莹

	<p>深圳博普科技有限公司 马迪</p> <p>建信理财有限责任公司 栾天悦</p> <p>东海证券股份有限公司 席红辉</p> <p>长城财富保险资产管理股份有限公司 胡纪元</p> <p>华安证券股份有限公司 陈涛</p> <p>银华基金管理股份有限公司 王卓立</p> <p>国寿安保基金管理有限公司 张帆</p> <p>武汉证国私募基金管理有限公司 陈阜东</p> <p>深圳市尚诚资产管理有限责任公司 黄向前</p> <p>中银基金管理有限公司 张令泓</p> <p>创金合信基金管理有限公司 梁雪</p> <p>宏利基金管理有限公司 石磊</p> <p>诺安基金管理有限公司 陈衍鹏</p> <p>中英人寿保险有限公司 袁野</p> <p>国泰海通证券股份有限公司 张盛海</p> <p>圆信永丰基金管理有限公司 田玉铎</p> <p>景顺长城基金管理有限公司 刘苏</p> <p>浙江浙商证券资产管理有限公司 鲁立</p> <p>华能贵诚信托有限公司 孙德龙</p> <p>信银理财有限责任公司 龙军</p> <p>嘉实基金管理有限公司 李涛</p>
会议时间	2026年6月7日
会议地点	线上交流
上市公司接待人员姓名	董事会秘书：张哲
投资者关系活动主要内容介绍	<p>1、一季度公司毛利率提升的原因是什么？</p> <p>2026年第一季度公司毛利率同比提升，主要得益于业务结构的积极变化，具体来说：高毛利的政府平台类业务在本季度显著起量。政府（G端）客户的可信数据空间建设，核心是提供标准化程度较高的数据处理平</p>

台及方案，相比传统定制化的数据采集与标注业务，其人力密集程度更低、交付周期更短、边际成本递减效应更明显，因此毛利率水平更高。

2026 年第一季度，随着公司前期在 G 端市场的全面布局进入收获期，这部分高毛利业务在收入结构中的占比明显提升，从而带动了整体毛利率的增长。

2、后训练和强化学习阶段，对数据标注需求有哪些变化？

在后训练和强化学习（RLHF/GRPO）阶段，模型的目标从“预训练期学会语言统计规律”转向“对齐人类偏好，这对标注数据的需求产生了深刻变化。

（1）从“量”到“质”与“复杂度”的转变

后训练阶段不再需要海量的原始文本，转而需要高质量、细粒度、领域专家级的标注数据。例如，SFT（监督微调）依赖精确的“标准答案”，而 RL 训练则需要能区分“好/较好/差”的偏好排序数据。

（2）聚焦“难例”与“长尾”场景

在 RL 阶段，“难例”（模型初始成功率低的数据）的价值远超易例，且能极大增强模型的泛化能力。标注重点从普遍规律转向稀缺、复杂的边界案例。

3、公司在具身数据领域的核心竞争优势是什么？

（1）硬件理解+驾驭壁垒：通过头部主流本体厂商合作经验，深度理解本体硬件性能特性、负载逻辑与异构适配要点，可快速解决硬件高故障率与稳定性问题；已在北京运营首个具身智能数据训练场，依托 100+ 台机器人实现高效数据产出。

（2）模型理解与落地壁垒：核心团队拥有大模型开发等资深背景，懂模型更懂数据；数据体系精准对齐机器人脑训练范式，提供具备高泛化能力的解决方案。

方案，支撑具身大脑模型训练。

(3) 综合数据工程化壁垒：不仅具备全球化数据采集与低成本交付资源；同时，自研的 AI 自动化标注平台与 7×24 小时算法质检，兼顾低成本与专家级质量。

4、不同机器人公司之间，是否存在可以共用的数据？

不同机器人公司之间确实存在可以共用的数据，但需要区分数据类型来看。

目前数据主要分为三类：真机数据、UMI 数据、第一人称视角数据。

从共用难度来看，真机数据很难直接共用，而 UMI 和第一人称视角数据相对更容易跨本体迁移。UMI 数据和第一人称视角数据不绑定具体机器人的关节参数和运动学模型，采集的是“操作过程”本身，理论上可以被不同机器人学习和泛化。

为什么真机数据很难共用？主要有 2 个原因：

一是硬件本体不统一。不同厂家机器人关节数量、自由度、末端执行器各不相同，A 公司的“抓取”数据没法直接拿去训练 B 公司的机械臂。

二是数据格式各自为政，行业缺乏统一的标准和“通用语言”。

所以行业正在推动的是：让 UMI 和第一人称这类“本体无关”的数据更多流通共用，同时通过技术对齐和标准统一，逐步解决真机数据的跨本体迁移问题。

5、具身智能数据服务的需求逻辑？

答：具身智能的核心在于赋予机器人具备理解与泛化能力的‘大脑’。这不仅是代码的堆砌，而是类似大模型‘涌现’的范式变革。由于当前仿真环境无法完美

	<p>模拟物理接触与长尾场景，海量、高泛化度的真实物理交互数据成为训练具身大脑的唯一‘燃料’，这是具身智能区别于传统自动化的根本逻辑。</p> <p>6、具身客户为何不自己做数据，而选择外部采购？</p> <p>答：数据工程是一项重人力、重流程、重管理的‘苦活累活’，并非算法厂商的核心基因。自建数据团队不仅边际成本高，且难以应对波峰波谷的弹性需求。我们定位为“具身智能行业的数据富士康”，通过规模化、标准化、可伴随客户迭代而同步迭代的整体数据供应链能力，帮助客户剥离非核心业务，使其聚焦于算法与本体研发，实现产业分工的最优解。</p> <p>7、具身智能数据赛道能做多久，是否像智驾一样只有短期机会？</p> <p>答：这是一个比自动驾驶更长、更宽的赛道。正如李飞飞教授所言，自动驾驶主要解决避障与导航，而具身智能需要解决非结构化环境下的复杂物理交互。当前我们甚至还未穷尽场景的定义，数据需求将伴随机器人进家庭、进工厂的全过程。我们定位为全行业的数据基础设施，旨在伴随行业全生命周期成长，而非短期套利</p>
附件清单（如有）	
日期	2026年6月8日