

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：2026-002

投资者关系活动类别	<input type="checkbox"/> 特定对象调研 <input checked="" 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	景顺长城 陆哲皓 东吴证券 张文佳 东方证券 宋鑫宇 鹏华基金 甘洋科 鹏华基金 闫思倩 中信建投 张敏 华商基金 王云翥 煜德投资 程跃 中信建投基金 赵坚 中信证券自营 宋文开 国新国证基金 范贵龙 诺安基金 刘慧影 开源证券 刘逍遥 Brilliance 胡镜誉 国君资管 范明 君阳私募 张春龙 凯恩投资 熊晓峰

星泰投资	曹秉超
中信期货	魏巍
中银基金	赵建忠
易川资产	冯强
名禹资产	朱细汇
PLEAID	Simon
仁灏	王贵州
东吴证券	田华
东吴证券	张文佳
博时基金	郭子琪
五矿自营	吴秋如
玄卜投资	夏一
东北自营	吴江涛
东吴自营	杨默曦
东吴证券	王紫敬
名禹资产	王友红
平安资管	万淑珊
国泰投信	Albert
正德泰投资	钟华
复星保德信	李心宇
东方睿石	唐谷军
国寿安保	张帆
富兰克林华美	陈宏毅
金鹰基金	陈颖
元兹投资	黄颖峰
Infini	刘书豪
博时基金	孔祥瑞
野村资管	江俊晨
铭基博远	祝泉

	恒复投资 刘强 明河投资 王蒙 汇添富 劳杰男 长安基金 徐小勇 中信期货 魏巍 南方基金 陆宇轩 肇万资产 李大瑞 东吴证券 常海 国新自营 高伟绚 宁远资本 张辛洋
会议时间	2026 年 1 月 15 日 2026 年 1 月 19 日 2026 年 1 月 28 日 2026 年 2 月 2 日 2026 年 2 月 3 日 2026 年 2 月 4 日 2026 年 2 月 6 日
会议地点	线上说明会、公司会议室
上市公司接待人员姓名	董事会秘书：张哲 投资者关系负责人：袁璐
投资者关系活动主要内容介绍	<p>1、公司在境外基地方面的布局计划如何？</p> <p>公司于 2024 年整合了位于东南亚的一个超过 1000 人的标注基地。该基地主要提供内容审核和数据标注服务。预计 2025 年该基地将贡献千万级美元收入。此次布局标志着公司境外交付能力建设进入了新阶段。我们计划于 2026 年在东南亚地区建设第二个本地化交付基地，预计到 2026 年底，境外基地总人数将再新增 500 人左右。</p> <p>该类基地建设为公司未来承接中国一线科技公司</p>

的出海业务、以及北美头部客户的定制化大型订单提供了关键的基础保障。

2、境外传统训练数据业务的驱动因素是什么？

过往多年，公司持续为境外多家头部科技大厂的全球人工智能产品的本地化及出海提供关键的多语种、多模态数据支撑。随着全球化 AI 应用场景的快速落地，市场对高质量、多语种、场景化训练数据的需求持续提升。具体而言，驱动需求的产品线主要包括但不限于：

（1）多语种语音识别数据：服务于智能助手、客服机器人等产品的全球化部署与口音适配。

（2）多语种手写体数据：支持金融票据识别、表单处理、手写笔记数字化等应用在不同语言文字区域的准确理解。

（3）多语种文本数据：涵盖自然语言理解、内容审核、机器翻译等任务所需的多语言文本语料库。

公司凭借在多语言、多模态数据处理领域长期积累的全球供应链管理能力和技术 know-how，正在持续获取并交付此类项目，从而推动境外数据业务的快速发展。

3、政府类业务布局和进展如何？

目前，我们与地方政府的合作已形成清晰的模式：一是基于地方特色数据（如农业、文旅）建设行业高质量数据集；二是搭建可信数据空间，保障数据安全；三是参与建设数据交易平台，促进数据流通；四是向下游延伸，开展场景化模型开发与应用落地。

在项目进展方面，公司已与成都、长沙等国家级数据标注基地试点城市建立合作，同时与广西、呼和浩特、保定、彭州等地方政府已启动具体项目推进。今年第三季度，公司已完成呼和浩特首批高质量行业数据集，以

及首批广西东盟语料库数据的交付。

4、请问公司如何看待具身智能数据业务？海天在该方向上有具体的计划吗？

公司非常看好具身智能数据领域作为高增长的新兴赛道，已组建具身智能数据专项团队开展相关业务，并开始在全国多个城市启动专项公司的选址。

一方面，国家已将具身智能列为未来产业予以重点培育。另一方面，控制机器人认知决策等“大脑”仍处于发展早期，大模型在泛化能力上仍有不足，因此想实现规模化落地将依赖大量高质量、细颗粒度的训练数据，因此这一领域市场潜力巨大。

当前，国内、外多家头部科技企业已启动具身智能数据的规模采购，需求涵盖机器人在真实场景下的交互数据、构建仿真世界模型所需数据等。同时，部分地方政府也在积极规划建设“具身智能训练场”，旨在工程化生产高质量具身数据，赋能科研与产业应用。

目前，公司已与多家机器人本体厂商展开合作，并与若干家头部科技大厂及地方政府启动订单需求对接，正在推进相关样例数据交付与训练场方案的设计与落地，为下一步具身智能数据业务的加速发展奠定坚实基础。

5、请介绍一下公司在训练数据领域的竞争优势

经过多年发展与积累，公司逐步构建起了在行业内的竞争壁垒，核心竞争力主要体现在：

(1) 公司的业务模式是服务产品双模式，且产品化贡献显著，是收入和毛利的主要来源，标准化数据集的研、产、销体系是公司从业多年探索出来的业务模式，其复用性为公司的规模化和高利润率提供了保障。而保持这样的能力需要具备对行业需求的强判断力和较强

的资金实力。

(2) 技术平台能力：公司历来重视技术的研发，近年来更是加大研发投入的力度，全面提升公司的算法能力、平台能力、工程化能力，加深算法辅助能力与人工工作的结合，达到更佳的人机协同效率，这样能够做大规模、提升效率、降低成本。

(3) 供应链资源管理能力：公司通过长期建设的供应链体系，保障资源的获取，未来，公司会进一步加大供应链资源平台的建设，使人员管理、采标资源分配、质量检验、远程工作等各方面的能力得到显著提升，为客群拓展提供有力支撑。

(4) 数据安全及合规能力：数据安全及合规能力已经成为了衡量品牌数据服务商综合能力的重要指标。公司在多年数据风险识别和管理实践中，已形成了较为成熟的安全、合规管理体系。

公司全方位做好数据风险管控工作，通过了业内重要的 ISO/IEC 27001 体系认证、IS027701 个人隐私信息安全管理体系建设，形成了具有自身特色的数据安全与隐私保护整体解决方案。公司已获得了北京市规划和自然资源委员会行政许可，获得乙级测绘资质，为公司智能驾驶采集业务拓展奠定基础；同时开展信息系统等级保护测评工作，获得等保三级备案证明；目前，公司符合 GDPR、《数据安全法》、《个人信息保护法》等一系列国际通用与国内法律法规的管理规范要求，获得了业务领域合作客户的高度认可。

6、客户也在做自己的算法，也可以用算法做预标注，海天瑞声的价值在哪里？

目前行业内确实存在部分客户自身利用算法进行预标注的情况，但该类客户仍有大量数据需求依赖第三

方数据服务商解决。一方面，对于已经过客户算法预标注的数据，仍将依赖第三方服务商在其预标注结果基础上进行更高难度、更加精细化的处理，以达到更高的数据准确率。相比于客户的预标注算法，海天瑞声在部分场景下具有更优的算法预标注能力，因为公司算法迭代速度更快，我们是以人机耦合的方式进行数据处理，我们的预标注算法能够第一时间、甚至以近乎实时的方式得到人工处理的反馈，使得预标注算法能够在最短的时间内实现快速迭代，最大程度提升对数据的预标注效果。另一方面，由于其预标注范围仅局限于其过往积累的数据现象，客户算法通常无法覆盖所有类型的数据标注需求，因此，对于行业较为前沿、愈加丰富的数据需求或标注场景，客户仍将依赖专业的第三方数据服务机构进行数据处理。

7、公司的业务是否存在规模效应？

公司业务是存在规模效应的，一方面随着公司在研发方面加大投入，自研平台的能力逐步提升，可以赋能数据处理过程中的人机协作朝着更加智能化的方向前进，这就使得公司进行更大规模的数据生产成为可能。同时，数据产品的积累、平台以及工具的研发，在公司业务规模逐渐上升的情况下，相关的研发费用、管理费用将被摊薄；

从成本端看，数据生产的成本还有很大的下沉空间，对于成本控制我们会在两方面进行持续投入：一方面是继续加大技术投入，采用更为合理的人机协同比例完成数据处理任务，降低人员投入，提高处理效率；另一方面是加强供应链资源管理能力，扩大资源供给，降低单位成本。

此外，数据集产品一直是我们公司所坚持的重点方

	向，公司开发大量通用型、复卖率高的标准化产品数据集，反复给公司带来利润，也能实现训练数据产品的规模化效应。
附件清单（如有）	
日期	2026年2月9日