

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：2026-004

<b>投资者关系活动类别</b>	<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）
<b>参与单位名称及人员姓名</b>	兴业证券 蒋佳霖、罗池婧 中银基金 郭艳超 中信保诚基金 吴振华 华夏基金 郭琨研 信达澳银基金 伏涛、郭敏 西部利得基金 吴海健 博时基金 郭子琪 上银基金 王婷、杨东朔 国寿安保基金 张帆 交银人寿 王玉英、周捷 长江资管 王毅 国华兴益资管 李灿 明世伙伴基金 王宇昊 聆泽投资 翟云龙 绿地金控 许迟 嘉泽私募 李湧

正圆投资	王伟平
国信证券深圳分公司	张浩冉
复星保德信人寿	李心宇
合众资产	胡纪元
星泰投资	曹秉超
丰琰基金	钟盾
弘洛私募	胡超
浙江国信投资	姜超
华泰证券	郭雅丽、徐诚伟
弘泽基金	蒋茜
西部利得基金	张昭君
古木投资	卢长利
红杉中国	闫慧辰
中国联通	寇士琪
昱奕资管	马李仪
磐石资产	朱怡琳
物产中大	刘嫁辉
国联民生	郭新宇
东方红	蒋娜、张伟锋、秦绪文
博时基金	田俊维
平安资产	邵进明
中信建投	方子箫
国金证券	刘琳琳、李可夫
东方红	张明宇
新华资产	张喜贺
交银施罗德	杨浩、武家和
华创证券	吴鸣远、周志浩
广发基金	冯骋
和谐汇一	王延飞、凌晨、章溢漫

	易方达 陈开熠
会议时间	2026年3月3日 2026年3月4日 2026年3月5日 2026年3月6日 2026年3月10日 2026年3月12日 2026年3月13日 2026年3月14日 2026年3月18日
会议地点	线上说明会、线下交流
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2025年收入增长的驱动因素是什么？</p> <p>全球人工智能产业从“技术突破”迈向“规模化落地”，公司训练数据业务迎来强劲增长，其中视觉业务收入增长超过200%，文本业务收入增长超过100%，语音业务收入亦保持合理水平。驱动这一高增态势的核心动能来自四重引擎：一是国内外科技巨头加速全球纵深布局，带动各类模态的多语种数据爆发增长；二是视觉大模型持续突破，推动高精度、场景化视觉数据投入持续加码；三是交互体验迈向拟人化，多情感、多角色、多音色的语音数据需求持续释放；四是垂直行业渗透加速，文旅、医疗、教育等领域垂类大模型落地提速，专业化、场景化数据需求日益凸显。</p> <p>与此同时，AI生态的参与主体日趋多元，除传统科技企业外，政府与国央企积极入局共建。政府机构依托数据资源优势，通过可信数据空间建设打通数据孤岛，围绕当地特色产业构建高质量语料体系。25</p>

年，公司已为广西、四川等地打造多类特色数据集，并同步推进湖南、内蒙古等地的专项数据集建设，相关收入有望成为公司新的价值增长高地。此外，国务院国资委深入推进央企“AI+”专项行动，以运营商为代表的央企加快布局自主可控 AI 大模型，数据资源投入力度持续加大，为公司带来新的增量空间。

此外，公司新拓展的东南亚数据交付体系已进入成熟运营阶段，面向海外客户的定制化数据服务能力持续提升，为境外收入贡献了可观的持续性增量。

除上述训练数据业务外，公司积极尝试向下游应用环节延伸，成功拓展数字人研发及音色模型平台重大项目，以数千数字人形象及丰富音色赋能区域乡村振兴与文化传播，该应用类收入也正成为公司业务版图中的有益补充。

以上多重增长动能叠加共振，共同驱动公司 2025 年度营业总收入实现显著增长。

## 2、公司在境外基地方面的布局计划如何？

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。该基地主要提供数据标注服务，2025 年该基地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 年在东南亚地区继续推进自主可控的数据基地建设，预计到 2026 年底，境外基地总人数将再新增 300-500 人左右。

该类基地建设为公司未来承接中国一线科技公司的出海业务、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。

## 3、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。

#### 4、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。25 年，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。

5、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？

公司高度看好具身智能数据这一新兴赛道，已将其作为战略重点进行前瞻布局，并组建了专项团队加快业务落地。

从行业前景看，具身智能已被国家列为未来产业重点培育方向。当前，机器人“大脑”认知决策等核心能力仍处于发展早期，大模型泛化能力尚未成熟，规模化落地亟需海量高质量、细颗粒度的训练数据支撑，市场空间广阔。

从市场需求看，国内外头部科技企业与本体厂商已启动规模化数据采集，需求覆盖真实场景交互数据、第一人称视角数据、umi 数据、仿真数据等。与此同时，部分地方政府正积极规划“具身智能训练场”，以工程化方式生产高质量数据，赋能产业应用。

目前，公司已与国内某头部具身厂商签署真机数据采集合约，并与多家头部具身厂商、科技大厂及地方政府启动订单对接，正稳步推进样例数据方案的设计与落地，为具身智能数据业务的规模化发展蓄力筑基。

6、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、ISO27701 个人信息安全管理体系认证，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。

#### 7、客户是否会自建数据团队？

有一些客户是会有自建团队的，客户自建团队主要解决其自身的部分数据需求，如敏感数据等，但受专业化分工的影响，客户仍然会大量购买数据服务提供商的数据，尤其是那些需要投入较高研发力量的复杂数据，以充实其算法模型训练的规模性需求。

8、训练数据产品和服务的定价模式、收费模式是什么样的？价格变动趋势如何？

定制服务定价模式：一般采用成本加成定价法。公司根据客户的具体服务需求预估项目成本，在预估成本的基础上，参考公司制定的指导毛利率水平，结合项目技术难度、复杂程度、时限要求等进行报价，并根据市场环境与客户协商，最终确定价格。

产品定价模式：一般采用需求导向定价法。公司综合考虑训练数据集的开发支出、市场需求程度、预计未来重复销售的频率等因素，制定产品标准价格及价格区间，在销售过程中，根据客户的实际需求情况，以价格区间为基础向客户报价，经双方协商确定最终销售价格。训练数据产品通常以单个数据集为单位进行定价，定价比较灵活。

价格走势主要由市场的供需关系决定。如果某类数据为市场稀缺数据，例如具有较高进入壁垒的多模态、虚拟人等前沿类数据需求、或传统业务里的多语种数据，都可在一定时间内维持较高的溢价水平。但在较为成熟的细分方向，比如中文智能语音数据领域，确实存在进入者增多、价格竞争的情况。因此，未来公司将主攻有较高技术壁垒，存在较大毛利空间的细分场景，尽力避免价格竞争带来的过度消耗。

9、训练数据产品、定制服务对公司有怎样的意义？

拥有海量自有知识产权训练数据产品集群是公司区别于众多竞争对手的显著优势，公司已沉淀下超1,800个自有知识产权的数据集产品，通过开发大量通用型、复卖率高的标准化产品数据集为公司的规模化和高利润率提供了保障。

	<p>训练数据定制服务是公司收入的重要来源，而且在提供训练数据定制服务过程中，公司会接触到各种类型的数据，帮助公司了解行业最新的技术路线和需求，同时，该等数据会在公司的采集加工平台上进行处理，对公司的算法提升和平台的工程化能力都提供了很好的学习资源，是研发迭代的较强助力。</p>
附件清单（如有）	
日期	2026年3月18日