

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：2025-005

|             |  |
|-------------|--|
| 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<br><input type="checkbox"/> 媒体采访 <input checked="" type="checkbox"/> 业绩说明会<br><input type="checkbox"/> 新闻发布会 <input type="checkbox"/> 路演活动<br><input type="checkbox"/> 现场参观 <input type="checkbox"/> 电话会议<br><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 | 中信建投 张敏<br>国投瑞银基金 钟婷霞<br>上海启泰 谢哲<br>中信证券 Pinki<br>山东嘉信 李献刚<br>金石投资 崔文瀚<br>永赢基金 任桀<br>工银理财 刘荫涛<br>源和资产 杨耿<br>混沌投资 黎晓楠<br>九泰基金 袁多武<br>中信里昂 陶雪纯<br>华夏基金 胡斌<br>天风证券 徐杉<br>中欧基金 李珊<br>中信证券 韩林轩<br>建信养老金 高月  |

|               |  |
|---------------|--|
|               | 和基投资 赵伟<br>狐尾松资产 胡亚男<br>华泰资产 郑金镇<br>博时基金 谢泽林<br>盈峰资本 张贻军<br>长信基金 王子旭<br>上海砥俊资产 陈晖<br>明曜投资 曾昭雄<br>东财基金 罗申<br>中邮证券自营 陈昭旭<br>神农投资 孙一明<br>神农投资 齐飞翔   |
| 会议时间          | 2025 年 4 月 27 日<br>2025 年 4 月 28 日   |
| 会议地点          | 线上交流、现场交流  |
| 上市公司接待人员姓名    | 董事会秘书 张哲<br>投资者关系 袁璐   |
| 投资者关系活动主要内容介绍 | <p>1、2025 年第一季度，公司收入增长的驱动因素是什么？</p> <p>随着多模态大模型的快速迭代及行业应用渗透提速，公司计算机视觉业务和自然语言业务分别同比实现高速增长。其中，在国家对“AI+数据要素”政策同步发力的背景下，以运营商、互联网平台公司为代表的大型客户持续加码高质量图像/视频等多模态数据采购，为其通用多模态大模型训练提供有力支撑；同时，政务、法律合规等场景应用的落地，带动场景类文本数据需求快速增加。在全球化布局方面，公司去年在东南亚新增建设的数据交付体系已进入爬坡运营阶段，通过拓展海外定制服务市场，不仅带来了可观的增量收入，并有望</p> |

成为海外业务扩展新的战略支点。上述因素，共同驱动公司 2025 年第一季度营业收入显著增长。

2、目前公司是否有在尝试新的业务或者商业模式？

当前，在国家大力推进“人工智能+”行动和“数据要素 X”的战略指引下，公司正积极探索与实践数据产业新业务和新模式。一是按照国家推动公共数据资源的开发利用，发挥海天瑞声的技术优势，与多地政府、地方运营商等开展战略合作，共同探索数据要素市场化与产业化的创新路径，通过构建“数据可信空间”，协助地方政府打造安全、高效、合规的数据治理与流通体系，推动数据要素的价值释放。二是发挥海天瑞声的行业经验和积累，联合当地高校，培训和培养数据标注人才，提升就业率的同时夯实区域数字经济发展人才基础。三是，发挥海天瑞声的生态优势，助力地方及产业园区打造数据标注基地和构建数据标注产业新生态。

3、公司与运营商的合作进展如何？

在国家“AI+数据要素”战略的指引下，尤其是国务院国资委连续两年开年启动部署中央企业“AI+”专项行动以来，以运营商为代表的重点央企自 2024 年起加速布局通用+垂向大模型研发，带动了高质量图像、视频等训练数据的规模化采购需求。公司凭借在数据领域的核心优势，已快速成为运营商类客户重要的数据服务供应商。未来，随着以运营商为代表的重点央企在多模态大模型方向的持续加码，以及其基座大模型在更多传统行业的应用落地，预计相关数据需求将进一步增长，为公司收入带来持续的增长动能。

4、与客户自建数据团队相比，海天的优势是什么？

相较于客户自建团队，海天瑞声历来都是对接众多

大型科技公司、头部人工智能企业、科研院所等，获得的信息是广泛的，项目经验丰富，同时积累了大量的know-how，对数据的理解更广、更深刻，同时我们搭建了成熟的数据处理算法平台，通过更高效的人机交互实现降本增效，保证数据质量的同时能有效降低成本，为客户提供更高性价比的训练数据产品/服务。

5、行业的竞争格局现在及未来是什么样的？未来市场集中度将如何变化？

目前来看，市场上数据服务市场主要由品牌数据服务商、客户自建团队以及一些中小数据服务商构成。

未来，公司预判整个数据服务市场将进行重新洗牌，集中度将进一步提升。市场各类主体将会通过在技术研发投入、资源能力建设等主要方面的竞争，逐步淘汰掉那些研发能力弱、资源势力差的品牌服务商和中小玩家。此外，国家对于数据安全及合规要求的进一步趋严，会将那些不具备数据安全合规能力或尚未进行此方面布局的企业逐渐淘汰出局。

在客户自建团队部分，出于其自身对数据和业务的敏感性、保密性需求，可能会与品牌服务商长期共存。

6、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。

7、成本结构里最大的部分是什么？如何能够持续性的优化成本结构？

公司最大的成本就是原料数据采购费用，即：采集、标注成本。一方面，公司通过继续加大研发投入的力度，全面提升公司的算法能力、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，这样能够做大规模、提升效率、降低成本；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。

8、公司提供的训练数据整体解决方案中，各个环节的技术难度如何？

首先，训练数据集的设计和原料数据采集环节是存在相当的技术难度的，比如语音类采集，文本设计是否贴合实际场景、如何实现最小采集量且确保场景覆盖丰富度等因素均是设计和采集环节需要考虑和解决的；在视觉类采集方面，复杂的人像采集、物体影像采集，同样具有如何设计合理的数据浓度达到最小成本最高训练效果，如果是垂直行业数据集的采集，例如交通行业内的自动驾驶领域，则存在准入资质、技术难度（包括但不限于对于交通场景、车辆传感器等要素的综合理解和实施能力）等方面的门槛。

数据标注环节的难度在于面对大量的数据标注需

|          |   |
|----------|---|
|          | <p>求，如何快速的找到充足的资源，而且通过算法平台实现机器的辅助标注，并在人机协作过程中，寻找效率与质量的最佳平衡，在提升数据标注效率、保证数据质量的同时降低成本。</p> <p>9、定制数据逐渐积累，是否可以转化为自有的数据产品？</p> <p>客户定制服务涉及的训练数据在交付给客户并完成验收后，所有权完全转移给客户，海天瑞声是不能用于自身产品建设的，这一点是公司始终遵循的知识产权要求。</p> <p>在定制数据集的生产过程中，积累下来的经验、know-how 会帮助公司各方面能力的提升，例如工具平台因为处理了大量的定制数据集，使平台完善性有很大的增益，加强了公司的数据处理能力；再如，在一些情况下，公司在生产定制数据集时，也会根据对行业需求的判断，在保障数据权属划分清晰的前提下，利用团队管理、资源获取的便利性，同步安排额外的设计、采集和标注工作，完成产品数据集的开发。</p> |
| 附件清单（如有） |   |
| 日期       | 2025 年 4 月 29 日   |