

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：2023-013

投资者关系活动类别	<input checked="" 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 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	招商证券资产管理有限公司 吴彤 上海重阳投资管理股份有限公司 陈心 合众资产管理股份有限公司 田玉铎 东方证券股份有限公司自营 张开元 申万菱信基金管理有限公司 龚云华 富荣基金管理有限公司 郭梁良 百年人寿保险股份有限公司 许娟娟 基金管理股份有限公司 刘伟林 海南悦溪私募基金管理合伙企业（有限合伙） 李晓光 长安基金管理有限公司 徐小勇 申万宏源证券资产管理有限公司 陈旻 鹏华基金管理有限公司 于凯 长江养老保险股份有限公司 邓锐 龙远投资 李声农 青骊投资管理（上海）有限公司 赵栋 上海混沌投资(集团)有限公司 曹映昊
会议时间	2023年8月31日

会议地点	海天瑞声会议室
上市公司接待人员姓名	证券事务代表：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、《生成式人工智能服务管理暂行办法》对公司的利好的点在哪？</p> <p>一方面，《生成式人工智能服务管理暂行办法》更加强调数据来源和数据处理过程的安全合规性，对训练数据处理活动提出明确的合规目标和要求。例如涉及个人信息的，需要确保取得个人同意，不得侵害他人依法享有的知识产权等等；同时，增强训练数据的真实性、准确性、客观性、多样性等等，这就要求相关数据服务企业具有完备的数据安全保障体系，不仅要确保数据来源合法合规，同时在数据处理过程中还要具备相应的数据处理经验和技能（例如，如何通过清洗和预处理技术排除训练数据中噪声和偏差，使得用于训练的数据不带有民族、国别、地域等歧视内容），进一步增加了数据服务的难度和门槛。以上数据安全标准将有助于行业门槛提升以及对专业第三方数据服务公司的依赖程度。</p> <p>2、智能驾驶目前的进展和未来的目标是什么？</p> <p>23年上半年，公司进一步升级了智能驾驶的数据服务能力，推出了新一代智能驾驶数据处理平台 DOTS-AD。新平台可支持更多场景、更多传感器、以及更前沿的自动驾驶感知算法需求，可以覆盖市面上主流的 3D、2D-3D 融合标注需求以及更为前沿的数据处理能力(包括:3D 点云语义分割、BEV 鸟瞰图标注、300 万大点云以及 4D 标注等);同时，算法中台功能更加完善，2D 算法准确率持续提升，3D 算法功能逐步完善；此外，公司成功引入最新 SAM 开源模型，并基于数据预标注实际需求，对模型进行了优化升级，有效提升了 2D 语义</p>

分割项目中的降本增效能力。

未来，公司将紧抓智能驾驶行业发展机遇，全方位布局和提升智能驾驶业务核心能力，前瞻性布局、迭代智能驾驶数据处理平台，优化算法中台中枢能力，同时加大数据安全体系建设，全方位保障数据生产流程安全合规。同时，基于已经获取的测绘资质，进一步拓展数据服务环节，开展采标一体化服务，提升业务毛利水平。

3、科大讯飞也在做智能驾驶的业务，我们和科大讯飞的关系？

科大讯飞是公司多年来的优质客户，公司为科大讯飞提供的产品及服务主要集中在智能语音领域，包括语音识别数据集产品或语音识别数据定制服务。未来，随着大模型等新一代人工智能技术驱动产业发展，公司也将积极拓展自身能力边界，力争为各类客户提供更多新产品、新服务。

4、是否考虑接入大模型来提升自身的数据生产效率？

智能化标注能力是数据服务企业的核心竞争力之一，公司一直致力于不断提升数据标注的智能化水平，目前公司已通过自行研发以及 API 接入两种方式，探索将大模型接入公司一体化数据处理平台，以提高数据处理过程中的人机协作效率，辅助公司的数据生产。自研大模型方面，公司已在针对大模型预训练数据集设计与处理技术进行初步研究和规划，并基于研究成果开展了 CommonCrawl、中文书籍等适用于预训练阶段的数据的获取与清洗工作，形成了各项技术的框架方向；开展大模型评测技术调研，完成基础框架设计，形成可行性结论。另一方面，公司已将部分开源或提供 API 接口的

大模型接入公司一体化数据处理平台。例如，公司已在智能驾驶平台 DOTS-AD 中接入开源的语义分割模型 SAM（Segment Anything Model），并基于数据预标注实际需求，对模型进行了优化升级，有效提升了 2D 语义分割项目中的降本增效能力。

5、公司定增项目二数据生产垂直大模型的进展和未来规划如何？

为进一步提升公司智能化标注水平，公司正在积极推进数据生产垂直大模型的建设工作。截至目前，公司针对大模型预训练数据集设计与处理技术进行了初步研究和规划，并基于研究成果开展了 CommonCrawl、中文书籍等适用于预训练阶段的数据的获取与清洗工作，形成了各项技术的框架方向；同时，开展大模型评测技术调研，完成基础框架设计，形成可行性结论。

未来，公司会基于实际市场需求，在大模型基础技术研究、大模型算法研发、以及平台工具建设等方面，持续推进数据生产垂直大模型的建设。

6、数据标注行业未来会有什么样的发展趋势？

首先是更加智能化，即通过拓展算法覆盖的场景以及算法预识别的准确率等，持续提升机器参与程度以及人机协作效率，降低数据处理成本。

其次，随着 AI 技术不断革新，应用行业以及场景不断增加，各行业、各领域数据安全规范逐渐落地将成为趋势，对于以数据生产为主营业务的数据服务企业，数据安全及合规能力将成为数据服务能力的核心评价维度，成熟的安全合规管理体系将成为重要评价标准，能持续跟踪法律环境变化，积极响应监管政策的企业将具有更强的市场竞争力。

此外，随着境内、外企业的全球化扩张成为确定

	<p>性趋势以及各类客户群体扩张步伐加速，多语种能力作为支撑企业顺利出海的核心要素之一，重要意义更加凸显，具有强大语言研究能力的数据服务企业将获得更多商业机会。</p> <p>另外，随着数据服务向多元化、多类型、多场景持续发展，充足、稳定且高质量的数据处理团队储备、以及更加智能化的资源配置能力，将成为数据高效、稳定交付的重要保障。</p>
附件清单（如有）	
日期	2023 年 9 月 4 日