

华泰联合证券有限责任公司
关于北京海天瑞声科技股份有限公司
2024 年半年度持续督导跟踪报告

保荐机构名称：华泰联合证券有限责任公司	被保荐公司简称：海天瑞声
保荐代表人姓名：张鹏	联系电话：010-56839300
保荐代表人姓名：葛青	联系电话：010-56839490

根据《证券法》《证券发行上市保荐业务管理办法》《上海证券交易所科创板股票上市规则》等有关法律、法规的规定，华泰联合证券有限责任公司（以下简称“华泰联合证券”或“保荐机构”）作为北京海天瑞声科技股份有限公司（以下简称“海天瑞声”、“公司”或“发行人”）首次公开发行股票保荐机构，对海天瑞声进行持续督导，并出具本持续督导跟踪报告。

一、保荐机构和保荐代表人发现的问题及整改情况

无。

二、重大风险事项

公司目前面临的风险因素主要如下：

（一）核心竞争力风险

1、核心技术快速迭代的风险

由于人工智能行业整体发展较为迅速，应用领域和应用场景不断拓展，下游需求随着行业发展不断发生变化，公司进行数据开发所需的各项技术也面临着快速迭代更新的风险。如果公司无法保持对行业发展趋势的洞察能力以及对研发人才的有效组织和研发经费的经济投入，则可能导致公司在行业技术快速迭代过程中无法持续保持技术先进性和技术优势，进而对公司的客户拓展、产品创新和经营业绩造成不利影响。

2、核心技术人员流失风险

随着行业的持续发展，行业内企业之间对于高端人才的竞争日益激烈，如果

公司无法持续加强核心技术的培养及引进并为核心技术人员提供有竞争力的激励机制和薪资待遇，则将存在核心技术人员流失的风险，公司的技术水平、研发能力也将受到不利影响。

3、核心技术失密风险

公司的数据产品和服务均以核心技术为基础。公司制定了严格的保密制度并严格执行，但上述措施仍无法完全避免公司核心技术的失密风险。如果公司相关核心技术的内控和保密机制不能得到有效执行，则可能导致公司核心技术失密的风险。

4、训练数据标准化产品无法实现授权销售的风险

公司根据对算法模型应用领域、行业发展趋势、市场需求等的评估和研判，自行设计并开发多种训练数据集标准化产品，开发完成后根据客户实际需求进行销售，授权给客户使用。在该种业务类型下，公司开发训练数据集产品，承担开发费用，并拥有训练数据集产品的知识产权。如果公司开发的训练数据集产品不能满足下游客户对训练数据的要求，则可能存在无法实现授权销售的风险，进而对公司未来的经营业绩产生不利影响。

5、数据安全相关风险

公司主要从事训练数据的研发设计、生产及销售业务，所提供的产品和服务主要以数据的形式体现。一方面，随着公司业务的快速发展和规模的持续扩张，原料数据采集与数据加工的数量持续增长；另一方面，包括《数据安全法》、《个人信息保护法》等在内的与数据安全、个人信息保护相关的法律规章体系逐步完善。如果将来公司未能根据法律规章的更新要求及时调整现行业务开展方式、公司的数据安全管理体系研发升级未能跟上业务发展的需要、或客户未能遵守训练数据产品保护相关商业约定，则公司可能面临生产经营不符合法律规章的要求、训练数据产品被泄露、盗版等数据安全相关风险或可能产生诉讼纠纷，进而给公司的经营带来不利影响。

(二) 经营风险

1、收入主要来自老客户，新客户收入占比较低的风险

报告期内，公司收入主要来自老客户、占比达到 91.54%，新客户收入占比相对较低。公司正致力于维护老客户、拓展新客户，若老客户降低对公司产品、服务的采购量或公司未能维持与老客户的合作关系，则公司的经营情况将受到不利影响；若新客户拓展情况不达预期，则公司的经营情况也将受到不利影响。

2、新业务拓展风险

随着人工智能行业持续快速地增长，下游客户对高质量、结构化、大规模训练数据的需求将持续增加。公司能否紧跟下游需求变化，加深对下游行业基础技术的理解并开发出更多的业务机会，一方面受到行业发展状况、市场需求变化以及市场竞争状况等因素影响，另一方面也取决于公司对于下游众多新领域新场景的理解认知以及合作拓展能力等诸多因素。因此，公司新业务的开拓可能不及预期或者遇到其他不利因素，进而对公司未来的经营业绩产生不利影响。

3、采购成本上升风险

报告期内，公司进行采购原料数据采集、标注所支付的数据服务费占全年与主营业务相关的采购总额的比重较高，达到 76.10%。随着城乡生活成本的持续提升，前述采购价格可能存在持续上升的趋势，若公司的技术发展带来的效率提升无法抵消前述采购成本上升的影响，可能会对公司的经营业绩带来一定不利影响。

（三）行业风险

1、市场竞争加剧的风险

公司所在的细分领域为 AI 基础数据服务领域。近年来，参与该行业的公司不断增多，由于整体行业较新，行业标准和法律监管仍处于规范阶段，因此整体行业目前处于高速发展且格局未定的状态，导致市场竞争加剧。公司如果不能持续有效地制定并实施业务发展规划，则将受到竞争者的挑战，从而面临市场竞争加剧而导致的经营状况下滑、市场地位下降和可能失去主要客户的风险，进而影响公司的盈利能力和发展潜力。

2、业务发展受下游人工智能领域发展状况影响较大的风险

公司主要为 AI 产业链上的各类机构提供算法模型开发所需的训练数据，服

务于下游人工智能行业发展。公司的经营情况、业务发展与下游人工智能领域的发展状况息息相关。若未来人工智能领域景气度下降，或因技术更迭使得对训练数据的市场需求发生变动，将对公司业绩产生较大影响。

（四）宏观环境风险

1、经营业绩受税收优惠政策影响较大的风险

报告期内，公司享受了增值税即征即退、高新技术企业和研发费用加计扣除等一系列税收优惠，经营业绩受税收优惠政策影响较大。如果未来公司所享受的税收优惠政策出现不利变化，或者公司不再符合税收优惠的条件，则公司将无法持续享受税收优惠或存在所享受的税收优惠减少的可能性，并进而对公司未来经营业绩和现金流水平造成不利影响。

2、汇率波动风险

海天瑞声作为一家面向全球的训练数据服务商，拥有数量众多的境外客户，遍布美国、韩国、日本等各地区，报告期内公司境外收入占比为 47.28%，该等收入使用外币进行结算，并受人民币汇率水平变化的影响。人民币汇率可能受到国内外政治、经济环境等因素的影响，存在波动风险，进而对公司的经营业绩产生不利影响。

3、中美贸易摩擦风险

公司产品和服务的主要出口地包括美国、日韩等国家或地区。报告期内，公司出口美国的营业收入占当期公司营业收入的比重为 24.38%。公司相关境外业务始终严格遵守中国及他国法律。但国际局势瞬息万变，近年来中美贸易摩擦不断，如果中美双方出台新的贸易保护主义措施，可能会影响美国市场相关客户将数据提供给公司用于标注，减少与公司进行训练数据方面的业务合作，公司相关业务可能会受到约束，将会对公司的生产经营产生不利影响。

三、重大违规事项

无。

四、主要财务指标的变动原因及合理性

本报告期，公司主要财务数据及指标情况如下：

单位：元 币种：人民币

主要财务数据	本报告期	上年同期	本期比上年同期增减(%)
营业收入	92,426,380.35	74,460,903.49	24.13
归属于上市公司股东的净利润	416,421.26	-17,241,392.69	不适用
归属于上市公司股东的扣除非经常性损益的净利润	-3,675,849.60	-23,104,588.43	不适用
经营活动产生的现金流量净额	15,676,856.20	-23,873,756.60	不适用
主要财务数据	本报告期末	上年度末	本期末比上年同期末增减(%)
归属于上市公司股东的净资产	753,379,145.86	782,293,983.51	-3.70
总资产	787,789,765.71	824,507,109.18	-4.45
主要财务指标	本报告期	上年同期	本期比上年同期增减(%)
基本每股收益(元/股)	0.01	-0.29	不适用
稀释每股收益(元/股)	0.01	-0.29	不适用
扣除非经常性损益后的基本每股收益(元/股)	-0.06	-0.38	不适用
加权平均净资产收益率(%)	0.05	-2.10	增加 2.15 个百分点
扣除非经常性损益后的加权平均净资产收益率(%)	-0.48	-2.82	增加 2.34 个百分点
研发投入占营业收入的比例(%)	28.86	43.26	减少 14.40 个百分点

上述主要财务数据及财务指标的变动原因如下：

1、报告期内，公司实现营业收入 9,242.64 万元，较上年同期增长 24.13%，主要系受益于大模型技术的快速革新以及相关政策的持续推动，全球人工智能产

业驶入加速发展轨道，国内外科技巨头持续加大人工智能领域投入，相关训练数据需求快速增长。同时，公司持续加大全球市场投放力度，通过品牌升级、营销推广等多元化战略积极拓宽海外市场空间，带动境外业务收入强力复苏。此外，公司不断拓展智能驾驶服务能力边界，抢抓车企出海机遇，积极探索布局境外数据采集业务；同时继续保持有竞争力的研发投入，持续优化迭代智能驾驶数据处理平台以及智能化标注能力，以更好适应以 BEV 技术为主、稳步向 OCC 技术方向探索的高阶智能驾驶技术需求。与此同时，受益于多模态技术的发展，大模型业务板块收入快速提升，成为继智能驾驶业务后公司又一重要的增长引擎。受上述各项积极因素带动，公司整体营业收入实现同比增长。

报告期内，归属于上市公司股东的净利润为 41.64 万元，实现归属于上市公司股东的扣除非经常性损益的净利润-367.58 万元，分别较上年同期增长 1,765.78 万元、1,942.87 万元。主要是因为报告期内整体营业收入增长的同时，标准化数据集产品收入占比也持续提升，同时公司整体数据交付能力在管理能力、平台技术提升、供应链进一步整合和扩展等因素的共同推动下，使训练数据定制服务的毛利率也同比显著增长，共同驱动公司整体毛利率增加至 70.34%。同时，在研发投入方面，为积极应对产业需求的新兴变化，公司继续在大模型数据领域进行研发投入，但随着首发募投项目的结项，公司在传统深度学习数据集建设等方向已达到阶段性成熟状态，整体研发投入强度呈现自然回落；此外，在销售及管理投入方面为进一步提升整体运营效率，公司进行了资源的合理配置和流程优化，使得销售费用以及管理费用有效降低，以上因素共同驱动归属于上市公司股东的净利润同比大幅改善；同时，以政府补助、理财收益为代表的非经常性收益下降，导致归属于上市公司股东的扣除非经常性损益的净利润的同比增幅稍窄。

2、报告期内，经营活动产生的现金流量净额为 1,567.69 万元，同比增加 3,955.06 万元。主要是因为报告期内，公司加大对应收账款的催款力度，使得销售标准化数据集产品、提供定制化数据服务等收到的现金同比上期增加约 2,024.56 万元；同时由于管理资源的合理配置和流程优化人工成本同比上期减少，驱动本期支付给职工及为职工支付的现金减少约 1,401.29 万元。

3、报告期内，基本每股收益、稀释每股收益较上年同期均增加 0.30 元/股，

扣除非经常性损益后的基本每股收益较上年同期增加 0.32 元/股。主要系报告期内归属于上市公司股东的净利润及归属于上市公司股东的扣除非经常性损益的净利润增加所致。

五、核心竞争力的变化情况

（一）研发优势

公司拥有同行业具有竞争力的综合研发实力，截至报告期末，公司研发团队规模为65人，研发能力全面覆盖平台工具开发、算法研究、产品设计等多条主线。公司核心研发人员均曾服务于全球顶尖科技企业或毕业于名校，组成了公司坚实的研发领导核心力量。同时，多年来，公司积累了较为完备的综合性、一体化数据处理平台及工具体系，覆盖智能语音、计算机视觉、自然语言等全业态领域，可服务于市面上绝大多数数据处理场景需求，包括但不限于智能驾驶、智能家居、个人助手、内容生成等。公司还设置了专门的产品研发部门及AI+研发部门，可前瞻性挖掘和布局新兴市场需求，抢占市场先机。此外，基于公司在训练数据领域多年积累下的大量、多元化数据以及数据处理经验，通过算法平台团队与数据业务团队无缝衔接，公司可更好完成预识别算法性能提升，实现少量数据注入->算法训练->标注员反馈->模型生成->处理大量数据的自闭环体系，助力公司做大规模、提升效率、降低成本。同时，近20年数据行业的从业经验积累，使得公司具备较强的数据生产工程化能力，可助力公司大幅缩短数据研发及生产周期，为公司抢占市场提供了有力支撑。上述因素共同驱动公司构建较高的行业技术壁垒。

（二）产品优势

通常来说，标准化数据集产品具有可即时获取、价格相对优惠等特点，因此市场需求旺盛。但由于标准化数据集产品是先于客户需求形成，需要数据服务商具备较强的市场趋势前瞻能力、以及财务实力进行提前投入，因此行业中具备标准化数据集产品规模化生产能力的服务商较为有限，该方向具备较高竞争壁垒。

对于公司而言，标准化训练数据集产品的知识产权由公司享有，具有一次性研发生产、可重复多次销售的特点，可显著助力公司毛利水平提升，实现训练数据产品的规模化效应，因此也是公司区别于众多竞争对手的一个优势亮点。

公司基于近20年数据服务经验、以及对客户需求的深刻理解，持续开拓标准化数据集，自有知识产权的训练数据产品储备超过1,650个，全面覆盖智能语音、计算机视觉、自然语言等多条业务线。

（三）语音语言学能力优势

随着境外客户全球化扩张以及境内客户出海成为新的行业发展趋势，多语种能力成为数据服务领域的核心竞争力之一，而语音语言学储备则是实现多语种拓展的重要前提。经过多年积累，公司已沉淀下深厚的语音语言学基础研究成果，并已将其运用至构建高质量智能语音、以及多模态训练数据。在语音语言学基础研究领域，公司拥有丰富的多语种语言学家团队资源积累和多语种发音词典积累，以及不断创新的发音词典构建技术和流程技术。截止报告期末，公司已经拥有超过205个语种/方言的覆盖能力，不仅包括含英、法、德、意、西、日、韩等常见语种，还包括东南亚、一带一路等国家地区的罕见小语种，尤其在亚洲小语种、中东欧小语种的服务上具备竞争优势。同时，公司已积累下超过130个多语种的发音词典，覆盖波斯尼亚语、塞尔维亚语、巽他语、尼泊尔语、奥利亚语、基隆迪语、茨瓦纳语、达利语等小语种，累计词条数超过1,200万条，可支撑构建高质量智能语音、以及多模态训练数据，是公司的主要竞争壁垒及核心技术之一。

（四）数据安全及合规能力优势

海天瑞声自成立之初即把数据安全、数据合规视为生命线，公司在多年的数据风险识别和管理实践中，已经形成了较为成熟的安全、合规管理体系。资质方面，公司拥有 ISO27001 信息安全管理体系认证、ISO27701 隐私信息管理体系认证、国家信息系统安全等级保护三级备案、北京市规划和自然资源委员会行政许可乙级测绘资质等。行业参与方面，公司入选中共中央网络安全和信息化委员会办公室“人工智能企业典型应用案例”，相关业务和数据安全实践受到《人民日报》等权威媒体报道和关注。公司参加北京市经济和信息化局组织的企业首席数据官素养能力培训，董事长贺琳成为北京市首批首席数据官。公司为中国信通院数据安全推进计划成员单位，董事兼副总经理李科连续入选该计划数安智库专家；其所参与、领导的公司数据安全工作，受到《北京日报》等权威媒体报道。此外，

公司积极参与行业发展，入选国家标准《信息技术 人工智能 术语》核心参编单位、《人工智能数据标注平台技术要求和测试方法》核心参编单位，数据安全能力获得认可并产生广泛、积极的行业影响。综上所述，本报告期公司核心竞争力未发生重大不利变化。

六、研发支出变化及研发进展

单位：元 币种：人民币

	本报告期	上年同期	变化幅度（%）
费用化研发投入	26,673,210.96	32,213,644.59	-17.20
资本化研发投入	-	-	-
研发投入合计	26,673,210.96	32,213,644.59	-17.20
研发投入总额占营业收入比例（%）	28.86	43.26	减少 14.40 个百分点
研发投入资本化的比重（%）	-	-	-

报告期内，公司研发投入合计 2,667.32 万元，较去年同期减少 17.20%，研发投入占营业收入的比例为 28.86%，较上年同期减少 14.40 个百分点。其主要系为积极应对产业需求的新兴变化，公司继续在大模型数据领域进行研发投入，但随着募投项目的结项，公司在传统深度学习数据集建设等方向已达到阶段性成熟状态，整体研发投入强度呈现自然回落。

报告期内，公司新申请软件著作权 5 项。截至报告期末，公司共获得专利授权 37 项，获得登记的软件著作权 173 项。

七、新增业务进展是否与前期信息披露一致

不适用。

八、募集资金的使用情况及是否合规

经中国证券监督管理委员会于 2021 年 7 月 13 日出具《关于同意北京海天瑞声科技股份有限公司首次公开发行股票注册的批复》（证监许可〔2021〕2366 号），同意公司首次公开发行股票注册的申请，并经上海证券交易所同意，公司首次向社会公众公开发行人民币普通股 1,070 万股，发行价格为 36.94 元/股，募集资金

总额为人民币 395,258,000.00 元，扣除各项发行费用人民币 58,901,108.97 元后，实际募集资金净额为人民币 336,356,891.03 元。上述募集资金已全部到账并经上会会计师事务所（特殊普通合伙）审验出具“上会师报字（2021）第 8543 号”《验资报告》。

截至 2023 年 12 月 31 日，首次公开发行募集资金投资项目已经结项，募集资金投资项目节余资金已经永久补充流动资金，公司募集资金专用账户余额为 0。

九、控股股东、实际控制人、董事、监事和高级管理人员的持股、质押、冻结及减持情况

报告期内，公司董事、监事、高级管理人员直接持有公司股份的持股、质押、冻结及减持情况如下：

单位：股

姓名	职务	期初直接持股数	期末直接持股数	报告期内股份增减变动量	增减变动原因
贺琳	董事长	3,467,890	3,467,890	-	-
王晓东	总经理	29,261	29,261	-	-
吕思遥	董事、副总经理、财务负责人、董事会秘书	43,350	43,350	-	-
李科	董事、副总经理、核心技术人员	55,262	55,262	-	-
黄宇凯	董事、副总经理、技术总监、核心技术人员	48,450	48,450	-	-
贾琦	董事	0	0	-	-
耿学锋	董事	0	0	-	-
秦勇	独立董事	0	0	-	-
杨华	独立董事	0	0	-	-
张蔷薇	独立董事	0	0	-	-
张小龙	监事会主席	0	0	-	-
许洁萍	监事	0	0	-	-
李娟	职工监事	0	0	-	-
郝玉峰	副总经理、技术总监、核心技术人员	2,570	2,570	-	-
曹琼	核心技术人员	7,750	7,750	-	-
邵志明	核心技术人员	0	0	-	-
合计	/	3,654,533	3,654,533	-	-

截至 2024 年 6 月底，根据公司提供的说明，公司控股股东、实际控制人、董事、监事及高级管理人员持有的公司股份不存在质押、冻结的情况。

十、上海证券交易所或保荐机构认为应当发表意见的其他事项

无。

（以下无正文）

（本页无正文，为《华泰联合证券有限责任公司关于北京海天瑞声科技股份有限公司 2024 年半年度持续督导跟踪报告》之签字盖章页）

保荐代表人：

张鹏

张鹏

葛青

葛青

华泰联合证券有限责任公司



2024年9月13日