

海天瑞声 (688787)

AI 数据领军企业，全球化布局打造第二成长曲线

买入 (首次)

2025 年 06 月 22 日

证券分析师 王紫敬

执业证书: S0600521080005

021-60199781

wangzj@dwzq.com.cn

盈利预测与估值	2023A	2024A	2025E	2026E	2027E
营业总收入 (百万元)	170.01	237.08	345.39	482.65	643.88
同比 (%)	(35.33)	39.45	45.68	39.74	33.41
归母净利润 (百万元)	(30.39)	11.34	31.69	55.70	90.63
同比 (%)	(203.16)	137.31	179.54	75.77	62.71
EPS-最新摊薄 (元/股)	(0.50)	0.19	0.53	0.92	1.50
P/E (现价&最新摊薄)	(187.62)	502.88	179.90	102.35	62.90

股价走势



投资要点

- 国内 AI 训练数据解决方案领军企业:** 海天瑞声是 AI 数据行业首家科创板上市公司, 致力于为 AI 企业、研发机构提供 AI 数据集及服务。董事长和实控人贺琳曾任职于中国科学院声学研究所, 管理层技术背景强。中移投资是海天瑞声首发前股东之一, 公司持续与中国移动保持紧密战略协同。公司业务主要涵盖智能语音、计算机视觉、自然语言等方向。2024 年, 大模型加速促进收入修复, 高毛利产品收入占比提升带动盈利质量改善。2024 年公司营收 2.37 亿元, 同比增长 39%, 归母净利润 0.11 亿元, 扭亏为盈。
- 数据标注产业迎来政策催化, 大模型加速成长。** 2025 年 1 月, 国家发展改革委等 4 部门发布《关于促进数据标注产业高质量发展的实施意见》, 提出到 2027 年, 数据标注产业规模大幅跃升, 年均复合增长率超过 20%。大模型驱动数据需求剧增。scaling law 证明了通过预训练使用更多的数据和算力, 能够让 AI 模型的性能大幅提升。大模型提升对高质量专业数据集需求, 对数据集需求正从“通用知识”向“专业知识”延伸拓展, 带动数据标注市场需求, 根据中商情报网, 2025 年中国数据标注市场规模有望突破 100 亿元。目前市场以品牌数据服务商和需求方自建团队为主, 市场集中度有所提升。
- 从 Scale AI 看海天瑞声的成长曲线。** 美国 Scale AI 专注于为人工智能和机器学习应用提供高质量的数据标注和管理服务。2024 年 Scale AI 以 138 亿美元的估值完成了一轮 10 亿美元融资, 英伟达、亚马逊、Meta、英特尔等科技巨头纷纷入局。公司主要服务大厂和政府客户, 外包标注团队, 并加强自动化标注水平, 陆续斩获上亿美元美国国防部大单, 由数据标注服务深入到 AI 模型应用落地。
- 海天瑞声: 全球化布局, 加强政府和产业合作。** 2024 年, 公司通过在新加坡投资设立控股公司的方式拓展海外首个自主可控的数据交付基地, 并积极拓展智能终端等领域海外客户, 2024 年公司境外收入同比增长 89.53% 至 1.14 亿元。公司持续深化政企合作, 已经和河北保定、湖南长沙等地政府合作成立合资公司, 布局数据标注基地全周期运营, 并发布了 DeepSeek 行业智能体解决方案, 从数据标注拓展 AI 应用落地, 打开成长新空间。
- 盈利预测与投资评级:** 高质量数据集是人工智能发展的基础, 公司是 A 股稀缺的专注提供 AI 训练数据解决方案的领军企业, 未来有望持续获得互联网、央国企等客户订单, 深度受益于 AI 大模型开发迭代和应用落地, 我们预计公司 2025-2027 年营收分别为 3.45/4.83/6.44 亿元, 归母净利润分别为 0.32/0.56/0.91 亿元, 首次覆盖, 给予“买入”评级。
- 风险提示:** 政策推进不及预期, 行业竞争加剧, 项目回款不及预期, 研发进展不及预期。

市场数据

收盘价(元)	94.50
一年最低/最高价	40.08/151.48
市净率(倍)	7.65
流通 A 股市值(百万元)	5,700.73
总市值(百万元)	5,700.73

基础数据

每股净资产(元,LF)	12.35
资产负债率(% ,LF)	8.22
总股本(百万股)	60.33
流通 A 股(百万股)	60.33

相关研究

内容目录

1. 国内 AI 训练数据解决方案领军企业	4
2. 数据标注产业迎来政策催化，大模型加速成长	6
3. 从 Scale AI 看海天瑞声的成长曲线	9
3.1. Scale AI: 从数据标注到应用落地，服务大厂和政府客户	9
3.2. 海天瑞声：全球化布局，加强政府和产业合作	11
4. 盈利预测与估值	13
5. 风险提示	14

图表目录

图 1: 海天瑞声发展历程.....	4
图 2: 海天瑞声股权结构 (截至 2025 年 4 月)	4
图 3: 海天瑞声产品服务矩阵.....	5
图 4: 训练数据生产过程示意图.....	5
图 5: 公司 2021-2024 年各业务收入 (单位: 百万元)	6
图 6: 公司 2021-2024 年各业务毛利率	6
图 7: 公司 2021-2024 年营业收入及同比增速	6
图 8: 公司 2021-2024 年归母净利润	6
图 9: 数据标注七大试点城市.....	7
图 10: 使用 Scaling Law 在 HumanEval 上准确预测 GPT-4 的通过率.....	7
图 11: 2021—2025 年中国数据标注产业市场规模趋势.....	8
图 12: 2023 年中国 AI 基础数据服务供给方的市场份额	9
图 13: Scale AI 主要产品.....	10
图 14: 海天瑞声 DeepSeek 行业智能体解决方案	12
表 1: Scale AI 历史融资情况.....	10
表 2: 海天瑞声标准化数据集产品储备情况.....	12
表 3: 海天瑞声盈利预测 (单位: 百万元)	13
表 4: 可比公司估值 (截至 2025 年 6 月 19 日)	14

1. 国内 AI 训练数据解决方案领军企业

AI 训练数据解决方案提供商。海天瑞声成立于 2005 年，是我国最早从事 AI 训练数据解决方案提供商之一。海天瑞声作为 AI 数据行业首家科创板上市公司，致力于为 AI 企业、研发机构提供 AI 数据集及服务。公司向全行业提供多语言、跨领域、跨模态的人工智能数据及相关数据服务，涵盖智能语音（语音识别、语音合成等）、计算机视觉、自然语言等多个核心领域，覆盖全球超 200 个主要语种及方言。

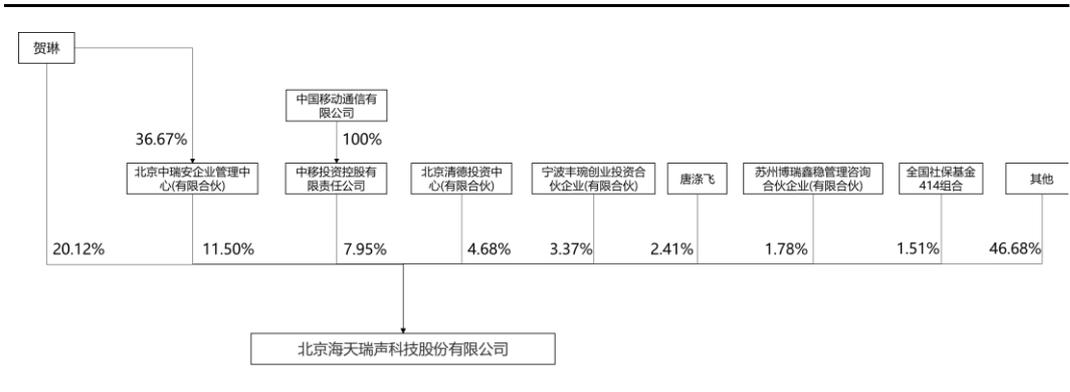
图1：海天瑞声发展历程



数据来源：公司官网，东吴证券研究所

管理层技术实力强，中移投资持续赋能。公司股权结构相对稳定，董事长贺琳直接和间接共持有公司约 24% 的股权，为公司创始人和实控人。贺琳女士毕业于北京大学，曾任职于中国科学院声学研究所，从事语音识别、语音合成、汉语语言理解、语音心理测试等方面的研究工作，先后参与了“汉语人-机语音对话系统工程”、“汉语语音参数合成”等国家自然科学基金重点项目、国家八六三计划智能计算机主题项目和中国科学院“八五”重大项目。现任总经理李科毕业于清华大学，获得信息与信号处理专业硕士学位，曾任 IBM 中国技术开发中心软件工程师，管理层技术实力强劲。中移投资是海天瑞声首发前股东之一，公司持续与中国移动保持紧密战略协同。

图2：海天瑞声股权结构（截至 2025 年 4 月）



数据来源：公司官网，东吴证券研究所

公司业务按照训练数据类型，分为智能语音、计算机视觉、自然语言等。盈利模式包括以下三类：

1) **定制服务**: 公司根据客户需求提供定制训练数据集并收取服务费。在此种模式下, 公司享有服务费收入, 不享有最终生成的训练数据的知识产权, 不可将此类业务生产的训练数据向其他客户重复销售。

2) **标准化产品**: 公司开发自有知识产权的训练数据集产品, 通过销售训练数据集产品的使用授权许可, 获取让渡资产使用权收入。此类训练数据集一经开发完成, 可多次销售并获取授权许可收入。

3) **训练数据相关的应用服务**: 公司基于生产的训练数据提供算法模型相关的模型拓展及训练服务, 通常以软件授权或软硬件一体化形式交付算法模型拓展、开发成果, 获取让渡资产使用权收入和技术服务收入, 以及少量硬件销售收入。以某大型科技公司客户项目为例, 客户研发了特定语音识别算法模型, 需要根据算法模型的实际场景(如法院庭审场景)开发落地应用。公司承担了部分落地应用拓展相关的开发工作, 围绕客户的算法模型和接口开发, 最终协助客户算法模型实现多个麦克风收集庭审语音内容并实时转成文字记录入系统的功能。

图3: 海天瑞声产品服务矩阵



数据来源: Wind, 东吴证券研究所

图4: 训练数据生产过程示意图

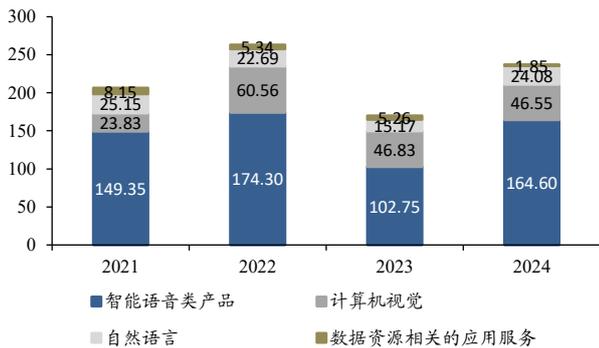


数据来源: Wind, 东吴证券研究所

大模型加速促进收入修复, 高毛利产品收入占比提升带动盈利质量改善。2022年, 受益于车厂大力发展智能驾驶, 计算机视觉业务收入同比有154%增长, 2023年, 受到

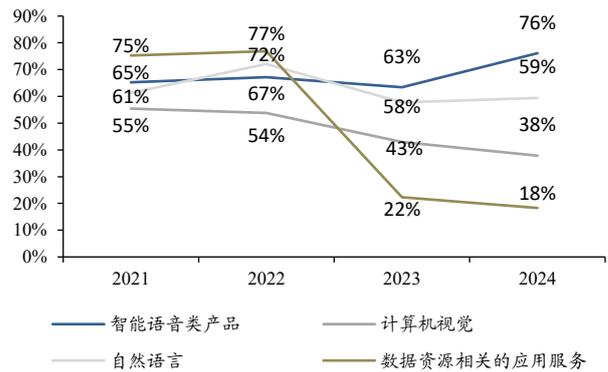
境外部分客户阶段性裁员、业务方向及研发节奏周期性调整影响，以及数据出境相关法规落地实施的阶段性影响，境内部分客户预算阶段性减缓，行业内竞争加剧，导致收入同比下滑。2024年，大模型技术加快发展，国内外大厂加大对多模态数据投入，以支撑在 AI 终端、内容生成等领域的应用，自然语音、自然语言需求同比显著增加。大模型促进了对独特数据集产品的需求，公司高毛利的数据集产品收入占比提升，驱动整体毛利率增加，盈利质量改善，2024年公司营收 2.37 亿元，同比增长 39%，归母净利润 0.11 亿元，扭亏为盈。

图5：公司 2021-2024 年各业务收入（单位：百万元）



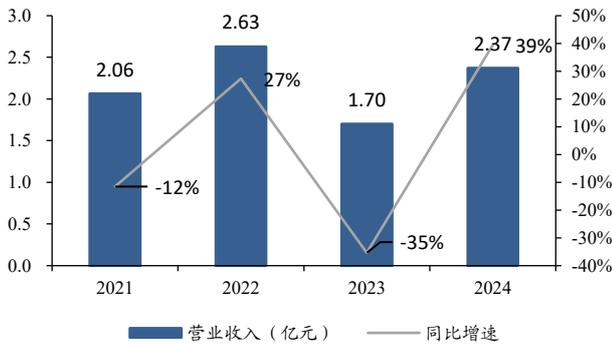
数据来源：Wind，东吴证券研究所

图6：公司 2021-2024 年各业务毛利率



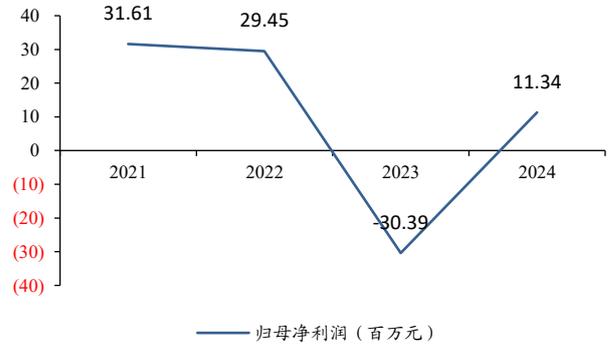
数据来源：Wind，东吴证券研究所

图7：公司 2021-2024 年营业收入及同比增速



数据来源：Wind，东吴证券研究所

图8：公司 2021-2024 年归母净利润



数据来源：Wind，东吴证券研究所

2. 数据标注产业迎来政策催化，大模型加速成长

政策促进数据标注产业高质量发展，目标到 2027 年产业规模 CAGR 超过 20%。2025 年 1 月，国家发展改革委等 4 部门发布《关于促进数据标注产业高质量发展的实施意见》，提出到 2027 年，数据标注产业专业化、智能化及科技创新能力显著提升，产业

规模大幅跃升，年均复合增长率超过 20%，培育一批具有影响力的科技型数据标注企业，打造一批产学研用联动的创新载体，建设一批成效明显、特色鲜明的数据标注基地，形成相对完善的数据标注产业生态，构建创新要素聚集、产业链上下游联动、区域协同发展的新格局。

数据标注七大试点城市快速推进，带动数据标注行业相关产值超过 83 亿元。2024 年 5 月，国家数据局在第七届数字中国建设峰会上公布成都、沈阳、合肥、长沙、海口、保定、大同为全国数据标注试点城市，截至 2025 年 3 月，已形成医疗、工业、教育等行业的高质量数据集 335 个；赋能 121 个国产人工智能大模型研发；引进和培育标注企业 223 家；标注从业人员达 5.8 万人；带动数据标注行业相关产值超过 83 亿元。

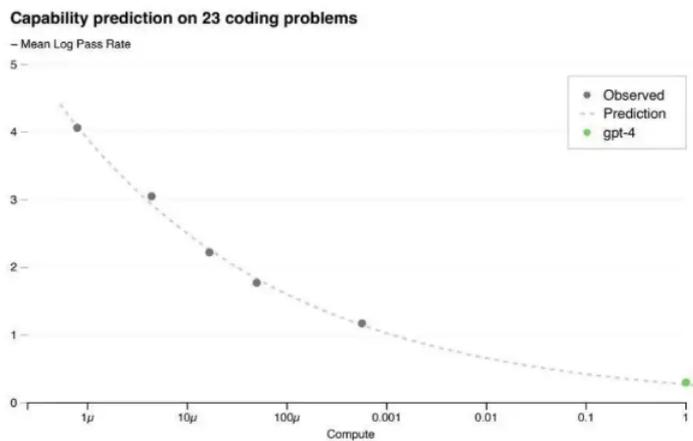
图9：数据标注七大试点城市



数据来源：中财经国际数字经济研究院，东吴证券研究所

大模型驱动数据需求剧增。数据是训练 AI 能力的“原料”之一，大模型的“涌现”能力出现和海量的训练数据密不可分，scaling law 作为过去两年多时间里 AI 大模型赛道风起云涌的基础，证明了通过预训练使用更多的数据和算力，能够让 AI 模型的性能大幅提升。根据 AI 基础数据服务厂商 LXT 对 322 家有 AI 项目经验的美国企业的调研，训练数据的资金投入占这些企业的 AI 整体建设投入的 15%。

图10：使用 Scaling Law 在 HumanEval 上准确预测 GPT-4 的通过率



数据来源：机器之心，东吴证券研究所

在 2025 年第八届数字中国建设峰会上，由国家数据局指导的全国数据标准化技术委员会提出，将强化标准引领，分三类建设高质量数据集。一类为“通用数据集”，包含面向社会公众、无需专业背景即可理解的通用知识，主要用于支撑通用模型落地应用。

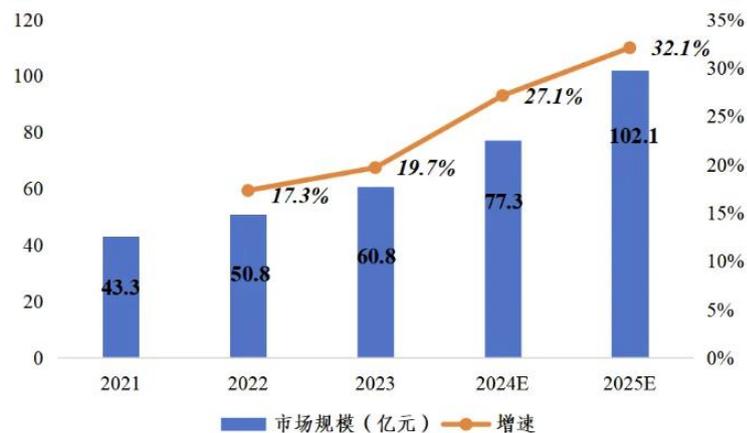
一类为“行业通用数据集”，包含面向行业从业人员、需要一定专业背景才能理解的行业领域通用知识，主要用于支撑行业模型落地应用。

一类为“行业专识数据集”，包含面向特定业务场景相关人员、需要较深的专业背景才能理解的行业领域专业知识，主要用于支撑业务场景模型落地应用。

AI 大模型对数据集需求正从“通用知识”向“专业知识”延伸拓展，提升数据标注门槛，带动数据标注市场需求。随着大模型在各行业落地，对于如医学、编程、竞赛等专业数据集的需求大幅增加，而覆盖全面或行业细分的数据集建设往往需要大量的时间、精力，也需要对于数据供应市场的深入了解，通过简单的爬虫或者购买公开数据集难以满足需求，因此大模型厂商，或者应用落地方往往会寻求数据标注专业厂商的合作。

根据中商情报网数据显示，2023 年中国数据标注市场规模达 60.8 亿元，较上年增长 19.7%。据预测，2025 年中国数据标注市场规模有望突破 100 亿元。

图 11：2021—2025 年中国数据标注产业市场规模趋势



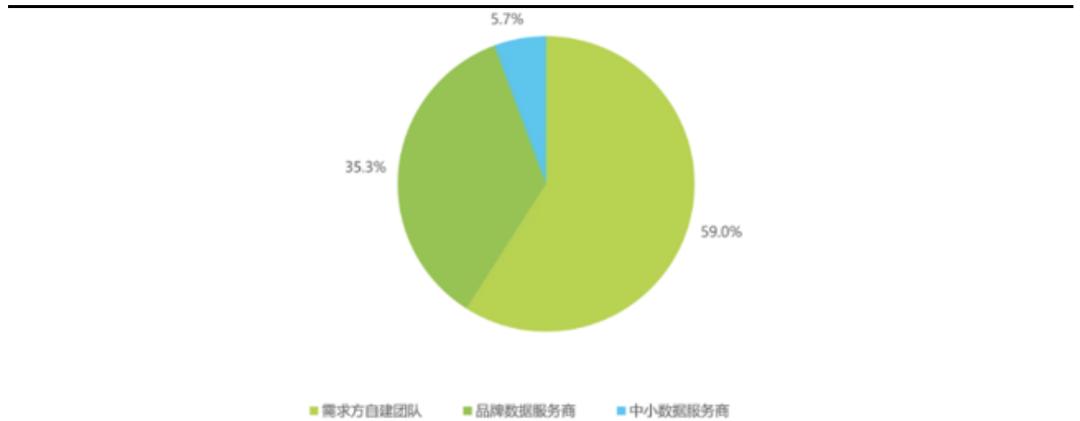
数据来源：中商情报网，东吴证券研究所

品牌数据服务商和需求方自建团队为主，市场集中度提升。传统 AI 数据标注市场竞争激烈，而大模型、智能驾驶等新兴项目体量较大需要较强的综合服务能力，智驾部分客户回款周期较长，较多中小数据服务商已退出市场；在大模型、智能驾驶等新兴 AI 算法及对应标注方式快速迭代时期，为保障信息安全，较多需求方通过自建团队满足数据服务需求，然而随着大模型对专业知识需求量增加，大厂自建团队覆盖多元化的垂类深度数据的难度和成本陡增，品牌数据服务商通过深度挖掘客户需求，可以打造高质量的垂类标准化数据集，竞争优势凸显。

未来随着品牌数据服务商的数据版权的丰富、专业能力的提升、标注方法的成熟，品牌数据服务商将承接更多的数据服务需求。据艾瑞调研统计，2023 年中国 AI 基础数

据服务行业的 CR4（前四大企业的市场份额）为 22.0%，市场仍较为分散。相比 2019 年 14.3% 的 CR4，中国 AI 基础数据服务市场在 2023 年的集中度显著提升。

图12：2023 年中国 AI 基础数据服务供给方的市场份额



数据来源：艾瑞咨询，东吴证券研究所

3. 从 Scale AI 看海天瑞声的成长曲线

3.1. Scale AI: 从数据标注到应用落地，服务大厂和政府客户

从单一数据标注到端到端解决方案。美国 Scale AI 由华人创始人 Alexandr Wang 和 Lucy Guo 创立，专注于为人工智能和机器学习应用提供高质量的数据标注和管理服务。Scale AI 前身是 Scale API，一个提供按需人工服务的平台，帮助企业完成内容审核、数据提取、预约安排等任务。企业只需一行代码，就能调用人工劳动力完成算法无法处理的任务。不久之后，Scale AI 发现自动驾驶行业对于数据审核和标注的需求量非常大，公司逐步形成了以数据标注为核心，同时帮助客户收集、清理、标注和管理大规模高质量数据，优化机器学习模型的业务模式。随着技术的发展，公司从单一的数据标注服务转型为提供端到端的 AI 解决方案。产品主要分成 4 大类：数据标注（Annotate）、管理和评估（Manage & Evaluate）、自动化（Automate）和合成（Generate）。

图13: Scale AI 主要产品

场景	产品	简介
数据标注 (Annotate) 数据标注业务是 Scale AI 的核心业务，标注可以来自人工标注或接近正确的 ML 模型，重点是在保持标注质量的前提下，增加由算法而非人工的标注量。 其中 Scale 3D Sensor Fusion 和 Scale Mapping 是特定于自动驾驶、机器人以及 AR/VR 行业。	Scale Rapid	可以在用户上传数据后加快数据标签的形成，要求数据由 Scale 标注。Scale Rapid 正在试图赶上 SuperAnnotate 在创建工作流和质量方面的领先地位。
	Scale Studio	提供标注工具，用户可以批注、监控和提高数据质量，要求用户自带标注人员。
	Scale 3D Sensor Fusion	3D 传感器数据标注。
	Scale Image	图像标注。
	Scale Video	视频标注。
	Scale Text	文档标注。
	Scale Audio	人工音频标注和分类。
	Scale Mapping	地图标注，可自定义。
	Scale Catalog	电商和零售团队可用于商品标注，AI 生成产品图。
管理和评估 (Manage & Evaluate)	Scale Nucleus	2020 年 8 月推出的数据集管理平台，用户可以在 Nucleus 中可视化查看标签性能、搜索特定标签条件下的图片、直接微调/修正标签、比较不同版本 ML 模型的准确度指标、查找相似图片、协作等，该产品为 Scale 拓展了更多小企业客户。
	Scale Launch	帮助用户在生产中部署和交付 ML 模型，可以配置和扩展 CPU 和 GPU 计算，进行影子部署、A/B 测试和 rollback，实时监控 ML 的性能、延迟时间等。Scale Launch 是 Scale AI 试图赶上 DataLoop 或 V7 等公司的地方。
自动化 (Automate)	Scale Content Understanding	帮助用户分析内容（比如人物表情是开心/难受）并将内容分类，以及标记敏感内容。
	Scale InstantML	用户只需上传数据，即可生成用户要求的 ML 模型，比如用于文本分类/图像分割/物体检测的模型。当用户上传更多数据或随着时间推移，模型会不断改进。
生成 (Generate)	Scale Spellbook	可以构建、比较和部署大语言模型应用程序。用户上传数据和 prompt，即可比较 LLM 模型并进行微调，最后可以一键部署到 API 接口。
	Scale Synthetic	2022 年推出，可以利用计算机算法生成多样化合成数据，但使用合成数据训练模型可能不会使得 ML 模型的性能得到提高。

数据来源: 36kr, 海外独角兽, 东吴证券研究所

2024 年 5 月，Scale AI 以 138 亿美元的估值完成了一轮 10 亿美元融资，投资方涵盖 20 多家机构和个人，英伟达、亚马逊、Meta、英特尔等科技巨头纷纷入局。2025 年 6 月，Scale AI 宣布已获得一笔来自 Meta Platforms 的重大投资，相应投资令该公司估值达到 290 亿美元，在获得投资的同时，该公司联合创始人兼首席执行官 Alexandr Wang 将加入 Meta。Meta 此次支付了 143 亿美元，取得 Scale AI 约 49% 的股权，这是 Meta 史上第二大交易，仅次于 2014 年以 190 亿美元收购 WhatsApp。

表1: Scale AI 历史融资情况

融资日期	融资轮次	融资金额	企业估值	投资方
2025 年 6 月	/	143 亿美元	290 亿美元	Meta
2024 年 5 月	F 轮	10 亿美元	138 亿美元	Accel、AMD、Meta、ServiceNow、Tiger Global Management、亚马逊、英特尔、思科
2021 年 4 月	E 轮	3.25 亿美元	70 亿美元	Coatue、Dragoneer Investment Group、Founders Fund、Greenoaks Capital、Index Ventures 等
2020 年 12 月	D 轮	1.55 亿美元	35 亿美元	Tiger Global Management
2020 年 4 月	C+轮	/	/	Next Play Capital
2019 年 8 月	C 轮	1 亿美元	/	Accel、Coatue、Founders Fund、Index Ventures、Spark Capital、Thrive Capital
2018 年 8 月	B 轮	1800 万美元	/	Accel、Index Ventures、Y Combinator
2017 年 5 月	A 轮	450 万美元	/	Accel、Arena Ventures、Y Combinator
2016 年 8 月	种子轮	12 万美元	/	Y Combinator

数据来源: 企查查, 百度百科, IT 之家, 东吴证券研究所

服务大厂和政府客户。Scale AI 的客户群体非常广泛，涵盖了自动驾驶、金融科技、政府机构（如美国国防部）、零售、电商等多个行业。在 GPT-2 时代，Scale AI 便与 OpenAI 展开了合作，迎来大模型浪潮之后，公司收获了 Meta、微软等一众客户，还积极拓展 G 端市场，业务迅速向政府部门渗透。根据华尔街见闻，Scale AI 2024 年营收大约 8.9 亿美元，预计 2025 年将超过翻倍——至 20 亿美元的水平。

重视政府合作，由数据标注服务深入到 AI 模型应用落地。2020 年，Scale AI 获得首份国防部大单，金额达到 9100 万美元，主要协助美国军方开发、测试和更新用于机器学习与 AI 的数据标注集，支持军事 AI 模型的训练。2022 年，公司获得 2.49 亿美元的国防部合同，专注机器学习、深度学习和神经网络测试，涵盖图像分析和自然语言处理。2023 年，公司为国防部首席数字与 AI 办公室（CDAO）提供数据管理和标注服务，支持联合部队多模态数据需求，提升军方数据质量，优化机器学习算法。公司 Donovan 平台是第一个在政府机密网络上部署的大语言模型，开发商将其定位为“人工智能决策平台”，可以处理来自不同来源的情报数据，并向人类官员提供建议。

外包人工标注团队，提升自动化标注水平。Scale AI 的标注外包人员主要从委内瑞拉、肯尼亚、菲律宾等工资水平相对较低的国家招聘，毛利较高。此外，公司一直致力于提升自动化标注水平，利用机器学习模型（如目标检测、语义分割算法）对原始数据进行初步标注，减少人工工作量。

3.2. 海天瑞声：全球化布局，加强政府和产业合作

全球拓展 AI 客户，进展显著。公司 2024 年积极拓展智能终端等领域海外客户，并且随着《促进和规范数据跨境流动规定》等法规出台，数据出境有章可循，公司业务拓展更加便捷，2024 年公司境外收入同比增长 89.53% 至 1.14 亿元。

海外设立交付基地，深化全球布局。2024 年，公司通过在新加坡投资设立控股公司的方式拓展海外首个自主可控的数据交付基地，在订单、产能两端同步实现增量覆盖，不仅有助于培育新的营收增长点，更为下一步提升海外定制化业务能力奠定基础。

深化政企合作。2025 年 3 月，国家数据局党组书记、局长刘烈宏主持召开数据标注领域民营企业座谈会，与海天瑞声、数据堂、百度等 11 家民营企业代表座谈交流。海天瑞声已经和河北保定、湖南长沙等地政府合作成立合资公司，布局数据标注基地全周期运营。2025 年，雨花国资集团与海天瑞声签约，合资成立长沙雨花数据产业集团，海天瑞声持股 20%；保定市莲池区与海天瑞声合作，成立河北数海科技有限公司，以市场化机制承接中国移动梧桐大数据标注基地全周期运营。

2025 年 4 月，中国—东盟人工智能创新合作中心项目签约仪式暨企业座谈会在南宁举行，海天瑞声联合创始人、CEO 李科表示，广西在开展面向东盟的人工智能合作方面具有独特优势，公司将在人工智能语料库建设、数字广西等领域加大投入力度，加快

共建东盟语料库, 构建可信数据空间, 积极参与中国—东盟人工智能创新合作中心建设。

加强高质量数据集建设。截至 2024 年 12 月 31 日, 公司已向下游客户提供了累计超过 9,500 次/个定制或标准化训练数据集, 并已应用于个人助手、语音输入、内容生成、智能家居、机器人、语音导航、智能客服、智能播报、语音翻译、移动社交、虚拟人、智能驾驶、智慧医疗、智慧教育、智慧交通、智慧城市、智慧金融、机器翻译、智能问答、信息提取、情感分析、OCR 识别等 22 类创新应用领域。随着公司进一步加强地方合作, 拓展数据标注覆盖深度和广度, 在 AI 大模型和应用落地加速趋势下, 公司标准化数据集产品收入有望快速增长。

表2: 海天瑞声标准化数据集产品储备情况

业务类型	应用领域	储备量(截至 2024 年末)	储备量(截至 2023 年末)
智能语音	个人助手、语音输入、智能家居、智能客服、机器人、语音导航、智能播报、语音翻译、移动社交、虚拟人、智慧金融等	1,179 个	1,088 个
计算机视觉	智能驾驶、移动社交、虚拟人、智慧金融、智慧交通、智慧城市、OCR 识别等	177 个	163 个
自然语言	通用 AI 大模型、机器翻译、智能问答、信息提取、情感分析等	360 个	307 个
合计		1,716 个	1,558 个

数据来源: Wind, 东吴证券研究所

从数据标注拓展 AI 应用落地, 打开成长新空间。2025 年 4 月, 海天瑞声携 DeepSeek 行业智能体解决方案亮相中国移动云智算大会, 海天瑞声 DeepSeek 智能体解决方案通过多 Agent 协同架构、行业模型能力提升、8 倍标注效率提升, 大幅降低企业 AI 应用门槛, 使企业能够在保障核心数据安全的前提下, 快速构建专属智能体。公司从数据标注业务介入, 拓展到为客户提供 AI 应用落地解决方案, 有望打开成长新空间。

图14: 海天瑞声 DeepSeek 行业智能体解决方案



数据来源: 海天瑞声公众号, 东吴证券研究所

4. 盈利预测与估值

关键假设:

- 智能语音:** 终端客户加大 AI 布局促进智能语音数据标注需求, 公司竞争力领先, 我们预计 2025-2027 年公司智能语音业务收入增速为 40%/35%/30%, 智能语音业务标准化产品占比相对较高, 我们预计 2025-2027 年业务毛利率稳定在 75%。
- 计算机视觉:** 央国企客户有望加强视觉大模型投入, 公司和中国移动等头部客户深入合作, 我们预计 2025-2027 年公司计算机视觉业务收入增速为 60%/50%/40%, 毛利率稳定在 38%。
- 自然语言:** 大模型垂直行业应用落地催化对自然语言个性化数据需求, 公司和头部厂商多年合作, 我们预计 2025-2027 年公司自然语言业务收入增速为 60%/50%/40%, 维持 60%的毛利率水平。

表3: 海天瑞声盈利预测 (单位: 百万元)

		2023	2024	2025E	2026E	2027E
智能语音	收入	102.75	164.60	230.44	311.09	404.42
	同比增速	-41%	60%	40%	35%	30%
	毛利率	63%	76%	75%	75%	75%
	收入占比	60%	69%	67%	64%	63%
计算机视觉	收入	46.83	46.55	74.48	111.72	156.41
	同比增速	-23%	-1%	60%	50%	40%
	毛利率	43%	38%	38%	38%	38%
	收入占比	28%	20%	22%	23%	24%
自然语言	收入	15.17	24.08	38.53	57.79	80.91
	同比增速	-33%	59%	60%	50%	40%
	毛利率	58%	59%	60%	60%	60%
	收入占比	9%	10%	11%	12%	13%
训练数据相关的应用服务	收入	5.26	1.85	1.94	2.04	2.14
	同比增速	-1%	-65%	5%	5%	5%
	毛利率	22%	18%	20%	20%	20%
	收入占比	3%	1%	1%	0%	0%
总收入		170.01	237.08	345.39	482.65	643.88
	同比增速	-35%	39%	46%	40%	33%
	毛利率	56%	66%	65%	64%	64%

数据来源: Wind, 东吴证券研究所测算

基于上述预测, 我们预计公司 2025-2027 年营收分别为 3.45/4.83/6.44 亿元, 同比增

长 46%/40%/33%，归母净利润分别为 0.32/0.56/0.91 亿元，同比增长 180%/76%/63%，对应 EPS 分别为 0.53/0.92/1.50 元。

可比公司估值与投资建议：我们选取同样在数据治理、数据集开发等方面有相关布局的易华录、星环科技、拓尔思作为可比公司，可比公司 2025 年 PS 均值为 15 倍。海天瑞声作为 A 股稀缺的专注提供 AI 训练数据解决方案的领军企业，未来有望持续获得互联网、央国企等客户订单，我们认为可以给予更高估值。高质量数据集是人工智能发展的基础，公司将深度受益于 AI 大模型开发迭代和应用落地，首次覆盖，给予“买入”评级。

表4：可比公司估值（截至 2025 年 6 月 19 日）

证券代码	可比公司	市值 (亿元)	营业收入 (亿元)			PS		
			2025E	2026E	2027E	2025E	2026E	2027E
300212.SZ	易华录	147	8.85	12.39	16.60	17	12	9
688031.SH	星环科技-U	53	4.71	5.87	7.04	11	9	8
300229.SZ	拓尔思	150	8.97	10.27	11.67	17	15	13
	平均					15	12	10
688787.SH	海天瑞声	59	3.45	4.83	6.44	17	12	9

数据来源：Wind，东吴证券研究所

注：海天瑞声盈利预测为东吴证券研究所预测，易华录、星环科技、拓尔思盈利预测取自 Wind 一致预期。

5. 风险提示

- 1. 政策推进不及预期：**数据标注相关政策推进受到多种因素影响，若推进不及预期，可能影响公司相关业务拓展。
- 2. 行业竞争加剧：**AI 训练数据市场增长较快，业务空间大、参与厂商多，未来行业竞争可能进一步加剧。
- 3. 项目回款不及预期：**AI 数据标注相关项目单个收入体量较大，项目确收节奏可能不及预期。
- 4. 研发进展不及预期：**AI 技术迭代快，若研发进展不及预期，可能影响公司相关产品服务竞争力。

海天瑞声三大财务预测表

资产负债表 (百万元)					利润表 (百万元)				
	2024A	2025E	2026E	2027E		2024A	2025E	2026E	2027E
流动资产	447	497	563	643	营业总收入	237	345	483	644
货币资金及交易性金融资产	328	324	321	320	营业成本(含金融类)	80	121	172	232
经营性应收款项	89	131	183	244	税金及附加	4	6	8	11
存货	23	35	50	67	销售费用	27	35	43	52
合同资产	1	2	2	3	管理费用	62	83	111	142
其他流动资产	5	6	7	9	研发费用	63	86	116	148
非流动资产	362	347	330	313	财务费用	0	0	0	0
长期股权投资	0	0	0	0	加:其他收益	5	7	10	13
固定资产及使用权资产	261	253	242	230	投资净收益	7	10	13	18
在建工程	46	37	30	24	公允价值变动	1	0	0	0
无形资产	8	9	9	11	减值损失	(2)	0	0	0
商誉	0	0	0	0	资产处置收益	0	0	0	0
长期待摊费用	1	1	1	1	营业利润	11	32	56	91
其他非流动资产	46	47	47	47	营业外净收支	0	0	0	0
资产总计	808	844	893	955	利润总额	11	32	56	91
流动负债	63	93	131	175	减:所得税	(1)	0	0	0
短期借款及一年内到期的非流动负债	3	3	3	3	净利润	11	32	56	91
经营性应付款项	23	35	50	67	减:少数股东损益	0	0	0	0
合同负债	14	20	28	38	归属母公司净利润	11	32	56	91
其他流动负债	23	35	50	67	每股收益-最新股本摊薄(元)	0.19	0.53	0.92	1.50
非流动负债	2	2	2	2	EBIT	4	32	56	91
长期借款	0	0	0	0	EBITDA	23	52	77	112
应付债券	0	0	0	0	毛利率(%)	66.46	65.04	64.41	63.94
租赁负债	2	2	2	2	归母净利率(%)	4.78	9.17	11.54	14.08
其他非流动负债	0	0	0	0	收入增长率(%)	39.45	45.68	39.74	33.41
负债合计	65	95	133	177	归母净利润增长率(%)	137.31	179.54	75.77	62.71
归属母公司股东权益	743	749	760	778					
少数股东权益	0	0	0	0					
所有者权益合计	743	749	760	778					
负债和股东权益	808	844	893	955					

现金流量表 (百万元)					重要财务与估值指标				
	2024A	2025E	2026E	2027E		2024A	2025E	2026E	2027E
经营活动现金流	29	17	33	58	每股净资产(元)	12.32	12.42	12.60	12.90
投资活动现金流	41	5	9	14	最新发行在外股份(百万股)	60	60	60	60
筹资活动现金流	(53)	(25)	(45)	(73)	ROIC(%)	0.52	4.22	7.33	11.70
现金净增加额	17	(3)	(3)	(1)	ROE-摊薄(%)	1.53	4.23	7.33	11.64
折旧和摊销	20	20	21	21	资产负债率(%)	8.06	11.30	14.88	18.54
资本开支	(31)	(4)	(4)	(4)	P/E(现价&最新股本摊薄)	502.88	179.90	102.35	62.90
营运资本变动	4	(24)	(31)	(36)	P/B(现价)	7.67	7.61	7.50	7.32

数据来源:Wind,东吴证券研究所,全文如无特殊注明,相关数据的货币单位均为人民币,预测均为东吴证券研究所预测。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准,已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司(以下简称“本公司”)的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下,本报告中的信息或所表述的意见并不构成对任何人的投资建议,本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下,东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易,还可能为这些公司提供投资银行服务或其他服务。

市场有风险,投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息,本公司力求但不保证这些信息的准确性和完整性,也不保证文中观点或陈述不会发生任何变更,在不同时期,本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有,未经书面许可,任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的,应当注明出处为东吴证券研究所,并注明本报告发布人和发布日期,提示使用本报告的风险,且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的,应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期(A 股市场基准为沪深 300 指数,香港市场基准为恒生指数,美国市场基准为标普 500 指数,新三板基准指数为三板成指(针对协议转让标的)或三板做市指数(针对做市转让标的),北交所基准指数为北证 50 指数),具体如下:

公司投资评级:

- 买入: 预期未来 6 个月个股涨跌幅相对基准在 15%以上;
- 增持: 预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间;
- 中性: 预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间;
- 减持: 预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间;
- 卖出: 预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级:

- 增持: 预期未来 6 个月内,行业指数相对强于基准 5%以上;
- 中性: 预期未来 6 个月内,行业指数相对基准-5%与 5%;
- 减持: 预期未来 6 个月内,行业指数相对弱于基准 5%以上。

我们在此提醒您,不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系,表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况,如具体投资目的、财务状况以及特定需求等,并完整理解和使用本报告内容,不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码: 215021
传真: (0512) 62938527
公司网址: <http://www.dwzq.com.cn>