

AI 基础数据服务行业研究 买入(维持评级)

行业深度研究(深度)

证券研究报告

计算机组

分析师: 孟灿 (执业 S1130522050001)

mengcan@gjzq.com.cn

分析师: 陈奕骄 (执业 S1130523020001)

chenyyijiao@gjzq.com.cn

AI 掘金潮下的“燃料”，新场景催化数据采标需求加速释放

行业观点

- AI 基础数据服务商脱胎于专业数据采标分工需求。国内 AI 基础数据服务产业主要包括上游-数据产生及产能资源、中游-训练数据生产、下游-AI 算法研发三大产业环节。数据工程是 AI 工程基础环节，目前正处于市场格局渐趋清晰，新老技术迭代、下游需求加速释放的关键节点。
- AI 快速落地叠加数据量指数级跃升，2025 年国内百亿规模可期。场景侧，AI 已在金融、医疗、交通、安防等多个垂直场景深度落地，且应用场景拓展势能强劲，伴生数据海量增长；产品侧，据信通院，自 OpenAI 于 2020 年推出 GPT-3 以来，超大预训练模型参数数量、训练数据规模按照 300 倍/年的趋势增长，客观推动高质量数据需求指数级扩张。据艾瑞咨询，2025 年我国 AI 基础数据服务市场或达 101 亿元，相较 2022 年翻倍。
- AI 大模型催生高要求和新需求，专业化数据集及 AI 训练师需求利好海天瑞声等优质专业数据解决方案提供商。AI 大模型时代无监督/半监督训练重回主流，“基础模型+微调”成为 AI 开发新范式，RLHF 微调技术催生两大新兴需求。一是预训练阶段催生 AI 厂商对标注完善、清洗完备的各类专业化场景数据集的需求；二是交互奖励阶段催生 AI 厂商对于具备专业事实判断与规范价值判断的人工智能训练师的需求。目前两类新兴需求尚未得到充分有效的市场供给。伴随超大规模预训练模型推动训练数据指数级跃升，市场蓝海亟待填补。
- 自动驾驶训练数据需求方兴未艾，海天瑞声发力布局。自动驾驶约占我国 AI 基础数据服务市场规模的 35%，系第一大下游场景，2025 年市场规模或达 25 亿元。自动驾驶领域的业务类型可分为车载摄像头采集的 2D 图像数据标注与激光雷达采集的 3D 点云数据标注。目前业内客户对于全栈式闭环数据解决方案需求较高。
- 国内市场集中度趋势性收敛，海天瑞声具备领先优势。伴随需求侧 AI 垂直场景及专业化需求凸显，数据需求量提升但数据采标业务门槛提高；供给端国内数据安全标准及相关资质要求提升，相关解决能力向头部厂商聚集。以海天瑞声、Appen 为代表的品牌数据服务商未来将替代中小型供应商成为市场主要供应力量。据 IDC，2021 年海天瑞声在国内 AI 基础数据服务行业市占率高达 12.9%，位居第一。Appen 在覆盖区域、语种/方言覆盖能力等方面更具优势。
- 海外对标：Appen，全球 AI 基础数据服务龙头。公司近五年营收 CAGR 达 27.2%，主营业务包括数据采集、数据预处理与模型评价三大类，业务类型齐全。主要客户包括 Google、Amazon、Microsoft 等知名 IT 巨头，客户质量较高。全球范围看，Appen 位居全球 AI 基础数据服务行业收入体量首位，员工量能及标注技术均领先同业。他山之石可以攻玉，对于国内厂商而言，积极卡位全球 AI 头部客户将有助于获得大模型预训练的业务红利；持续拓宽业务矩阵，开辟模型评价模块主动挖掘客户数据需求将有助于纵向做深客户价值；夯实数据采标技术，紧跟行业趋势发展包括机器标注在内的先进数据采标技术将有助于筑牢自身业务护城河。

投资建议

- AI 基础数据服务行业，需求侧 AI 垂直场景及专业化需求凸显，数据需求量提升但数据采标业务门槛提高；供给端国内数据安全标准及相关资质要求提升，相关解决能力向头部厂商聚集。推荐关注国内业内领军海天瑞声。

风险提示

- 产业需求快速迭代带来研发投入回收不确定性风险；训练数据泄露引发的安全性风险；市场竞争加剧导致利润受到挤压的风险。

内容目录

1. AI 基础数据服务脱胎于专业数据标注需求，2025 年国内百亿规模可期.....	5
1.1 产业图谱：AI 基础数据服务脱胎于专业数据采标分工需求	5
1.2 发展历程：站在人工标注市场出清与机器标注迭代的十字路口	5
1.3 市场规模：AI 快速落地叠加数据量指数级跃升，2025 年国内百亿规模可期.....	6
1.4 政策支持：近五年国家政策加持国内 AI 产业发展	8
2. AI 大模型催生高要求新需求，专业化数据集及 AI 训练师需求利好优质专业数据提供厂商	10
2.1 场景特质：AI 大模型时代无监督/半监督训练成为主流，RLHF 催生新兴需求.....	10
2.2 场景需求：预训练阶段高质量专业数据集+交互奖励阶段人工智能训练师.....	11
2.3 场景价值：超大规模预训练模型推动训练数据指数级跃升，市场需求持续延展	12
3. 自动驾驶领域训练数据需求方兴未艾，海天瑞声发力布局.....	12
3.1 场景特质：自动驾驶客户需求全栈式闭环数据解决方案	12
3.2 场景需求：自动驾驶领域数据标注分为 2D 图像标注与 3D 点云标注.....	14
3.3 场景价值：自动驾驶训练数据市场方兴未艾，2025 年国内市场规模或达 25 亿元.....	14
4. 国内市场集中度趋势性收敛，海天瑞声具备领先优势.....	15
4.1 发展趋势：品牌商价值效应凸显，国内市场集中度持续收敛	15
4.2 竞争格局：业内主要玩家发展各有侧重，Appen 及海天瑞声具备领先优势.....	16
4.3 他山之石：Appen 发展复盘，全球 AI 基础数据服务巨头的崛起之路.....	19
5. 投资建议	20
6. 风险提示	21

图表目录

图表 1: AI 基础数据服务商脱胎于专业数据标注分工需求.....	5
图表 2: 数据工程约占 AI 工程 80%时长, Garbage in garbage out 效应显著.....	5
图表 3: 72%的全球受访者认为至少需要超过 10 万条训练数据进行模型训练.....	6
图表 4: 96%的全球受访者在训练模型的过程中遇到训练数据质量不佳、效率不足的问题.....	6
图表 5: AI 数据工程发展历程及展望.....	6
图表 6: 2025 年全球 AI 市场规模或达 2,218.7 亿美元.....	7
图表 7: 2022 年国内 AI 市场规模或达 116 亿美元.....	7
图表 8: 2018~2025 年全球数据量 CAGR 或达 26.9%.....	7
图表 9: 2018~2025 年中国数据量 CAGR 或达 30.4%.....	7
图表 10: 2025 年我国 AI 基础数据服务市场或达 101 亿元, 相较 22 年翻倍.....	7
图表 11: 2019 年我国 AI 基础数据服务各类型数据需求占比.....	7
图表 12: 近五年国家政策加持国内 AI 产业发展.....	8
图表 13: AI 大模型传统数据标注需求下降.....	10
图表 14: OpenAI 开展 RLHF 三步骤.....	11
图表 15: RLHF 需求: 预训练阶段高质量专业数据集+交互奖励阶段人工智能训练师.....	12
图表 16: 超大规模预训练模型参数及数据规模近年来指数级跃升.....	12
图表 17: 自动驾驶场景中的 AI 基础数据服务需求.....	13
图表 18: 自动驾驶客户需求全栈式闭环数据解决方案.....	13
图表 19: 自动驾驶场景中的 AI 基础数据服务业务类型.....	14
图表 20: 2020 年自动驾驶约占 AI 基础数据服务市场规模的 35%.....	15
图表 21: 2025 年我国自动驾驶 AI 基础数据服务市场规模或达 24.9 亿元.....	15
图表 22: 2019 年我国 AI 基础数据服务市场竞争格局.....	15
图表 23: 2019 年我国 AI 基础数据服务市场 CR5.....	15
图表 24: 业内主要品牌数据商.....	16
图表 25: 业内主要玩家指标对比.....	16
图表 26: 智能语音类训练数据产品对比.....	17
图表 27: 通用语音识别训练数据产品对比.....	17
图表 28: 车载语音识别训练数据产品对比.....	18
图表 29: 特色语音识别训练数据产品对比.....	18
图表 30: 自然语言训练数据产品对比.....	18
图表 31: 计算机视觉训练数据产品结构对比.....	19
图表 32: AI 基础数据服务巨头, 近五年营收 CAGR 达 27.2%.....	19
图表 33: Appen 业务类型齐全, 客户质量较高.....	20

图表 34: Appen 处于员工量能及标注技术矩阵头部位置.....	20
图表 35: 推荐关注国内 AI 基础数据服务领军者海天瑞声.....	21

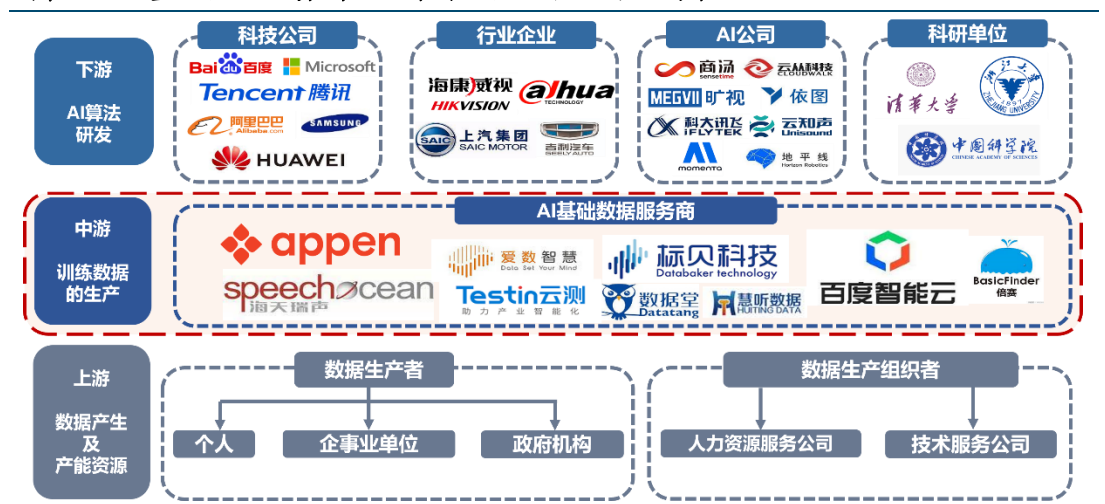
1. AI 基础数据服务脱胎于专业数据标注需求，2025 年国内百亿规模可期

1.1 产业图谱：AI 基础数据服务脱胎于专业数据采标分工需求

国内 AI 基础数据服务产业主要包括上游数据产生及产能资源、中游训练数据生产、下游 AI 算法研发三大产业环节。其中，部分产业环节重合度较高，AI 基础数据服务商主要脱胎于专业数据采标分工需求。

- 上游主要包括数据生产者和数据生产组织者，主要提供原料数据的采集服务。
- 中游主要由基础数据服务商构成，通过数据处理能力和项目管理能力完成训练数据集结构设计、数据加工和质量检测等工作，为下游客户提供训练数据产品和服务。AI 基础数据服务整体可分为两大类，一种是具备自有的标注基地或全职标注团队，这类企业也参与产业上游部分直接提供产能资源，另一种是依靠众包或外包模式，专注于数据产品的开发与项目执行。
- 下游包括科技公司、行业企业、AI 公司和科研单位等，主要负责算法研发。部分下游 AI 公司拥有自主的标注工具，也可通过 AI 中台获取一些通用标注工具，少数数据需求大的企业还孵化了自主的数据服务团队。

图表1：AI 基础数据服务商脱胎于专业数据标注分工需求

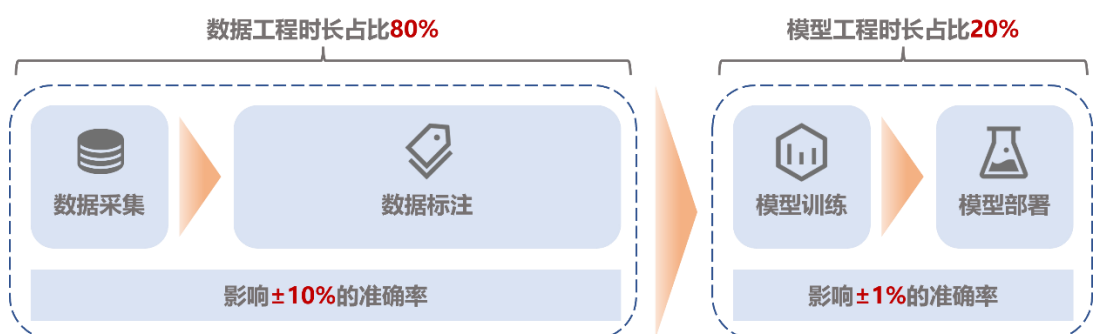


来源：海天瑞声招股说明书，艾瑞咨询《2019 年中国人工智能基础数据服务行业白皮书》，国金证券研究所

1.2 发展历程：站在人工标注市场出清与机器标注迭代的十字路口

理论层面：数据工程系 AI 工程基础环节，核心在于高效的数据标注。AI 工程=数据工程+模型工程。其中，数据工程主要包括数据采集与数据标注，约占 AI 工程时长的 80%；模型工程主要包括模型训练与模型部署，约占 AI 工程时长的 20%。数据工程是 AI 工程的前置且基础环节，直接影响到模型的质量与精度。数据工程的核心在于高效的数据标注，Garbage in garbage out 效应显著。

图表2：数据工程约占 AI 工程 80% 时长，Garbage in garbage out 效应显著

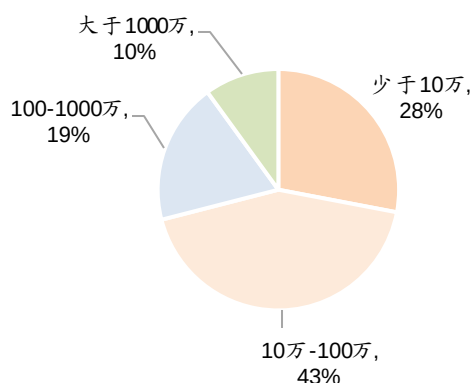


来源：整数智能，国金证券研究所

实践层面：AI 模型训练数据需求规模大，训练数据质量不佳、效率低下情况普遍。据

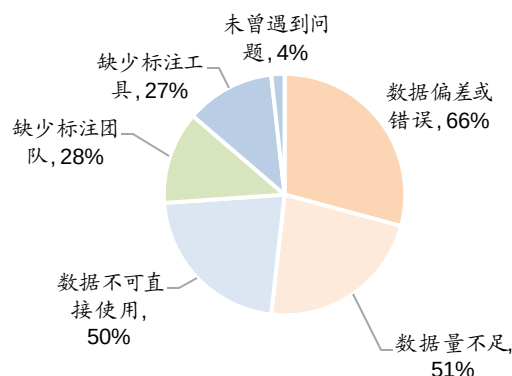
Dimensional Research 全球调研报告, 72%的受访者认为至少使用超过 10 万条训练数据进行模型训练, 才能保证模型有效性和可靠性, 96%的受访者在训练模型的过程中遇到训练数据质量不佳、数量不足、数据标注人员不足等难题。为应对训练数据所带来的多方面挑战, AI 企业开始从第三方购买原料数据收集、训练数据生产和数据专家咨询等服务。

图表3: 72%的全球受访者认为至少需要超过10 万条训练数据进行模型训练



来源: 海天瑞声招股说明书转引自 Dimensional Research, 国金证券研究所

图表4: 96%的全球受访者在训练模型的过程中遇到训练数据质量不佳、效率不足的问题

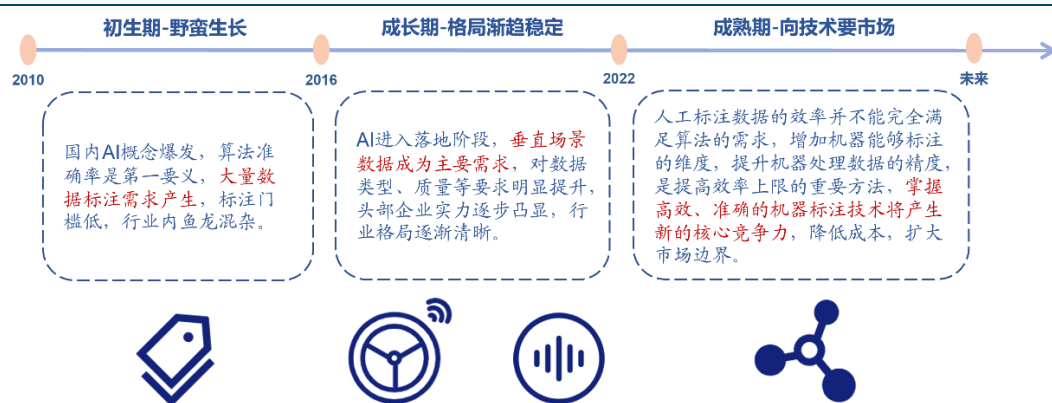


来源: 海天瑞声招股说明书转引自 Dimensional Research, 国金证券研究所

AI 数据工程发轫于 AI 产业落地元年, 系 AI 下游应用的基础且必备环节。目前行业处于市场格局渐趋清晰, 新老技术面临迭代, 下游需求加速释放的关键节点。

- 产业初生期 (2010~2016): 2010 年语音识别和计算机视觉领域产生重大突破, 国内开始萌生 AI 概念。后续数年, 早期的 AI 基础数据服务门槛较低, 质量参差不齐。
- 产生成长期 (2016~2022): 近五年来, 供给侧高烈度的业内竞争加速市场出清, 需求侧对产业落地以及垂直场景的定制化数据采集需求逐渐凸显。最终引致行业头部企业浮出水面, 行业格局逐渐清晰。
- 产业成熟期 (2022~至今): 2022 年以来, AIGC 产品集中爆发, 高 level 自动驾驶需求加速释放。传统人工标注的效率已不能完全满足算法需求, 行业护城河转向自动化机器标注技术, 预计产业将进入向技术要市场的新阶段。

图表5: AI 数据工程发展历程及展望

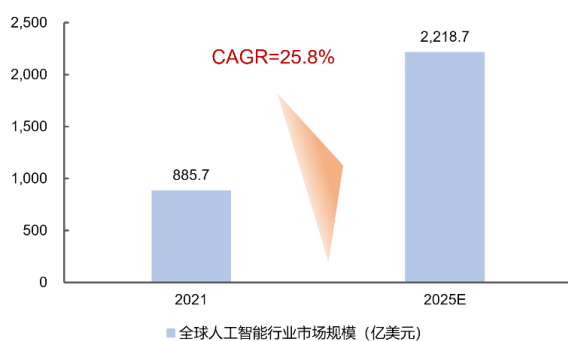


来源: 艾瑞咨询《2019 年中国人工智能基础数据服务行业白皮书》, 国金证券研究所

1.3 市场规模: AI 快速落地叠加数据量指数级跃升, 2025 年国内百亿规模可期

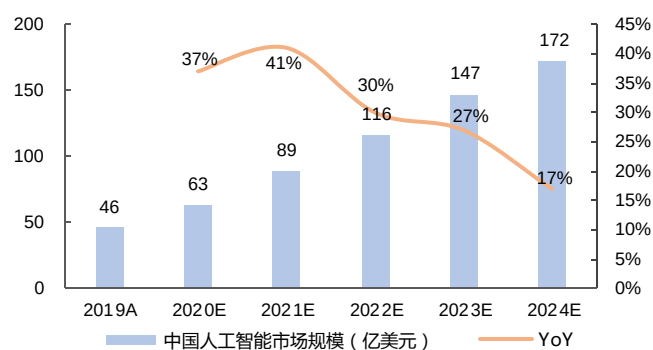
国内 AI 市场规模超百亿美金, 约占全球市场 10% 份额。目前, AI 已在金融、医疗、交通、安防等多个垂直场景深度落地, 且应用场景拓展势能强劲, 商业化进程加速。从全球市场看, 据 IDC, 2021 年全球 AI 产业规模达 885.7 亿美元, 预计 2025 年将达到 2,218.7 亿美元, CAGR 高达 25.8%。从国内市场看, 据 IDC, 2022 年我国 AI 产业规模或达 116 亿美元, 预计未来数年内仍保持两位数增长。以 2021 年计, 国内 AI 市场规模约占全球 10%, 国内市场成长潜力巨大, 国内企业出海空间广阔。

图表6: 2025 年全球 AI 市场规模或达 2,218.7 亿美元



来源: 海天瑞声招股说明书转引自 IDC, 国金证券研究所

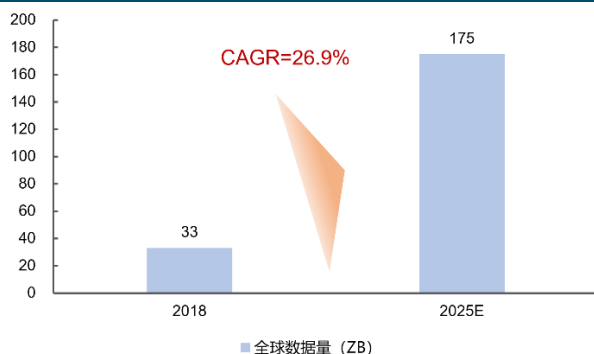
图表7: 2022 年国内 AI 市场规模或达 116 亿美元



来源: 海天瑞声招股说明书转引自 IDC, 国金证券研究所

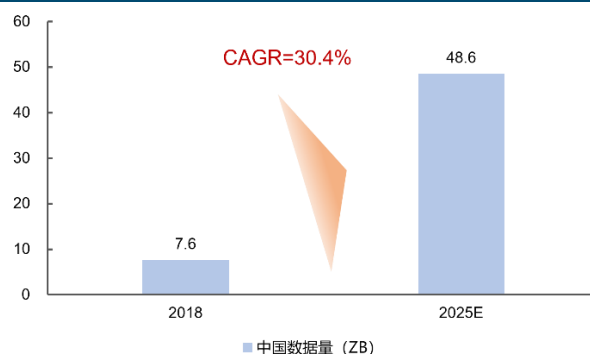
全球数据量呈指数式增长, 中国数据量增速跑赢全球。据 IDC, 全球每年生产的数据量将从 2018 年的 33ZB 猛增至 2025 年的 175ZB, 其中结构化数据仅占到全部数据量的 20%, 其余 80% 都是以文件形式存在的非结构化和半结构化数据, 日志文件、机器数据等占非结构化数据的 90%, 产生了源源不断的数据清洗与标注需求。相比之下, 中国的数据量增速领跑全球, 平均每年增速比全球快 3%。2018 年, 中国的数据量为 7.6ZB, 占全球总量的 23.4%, 预计到 2025 年将增至 48.6ZB, 占全球总量的 27.8%, CAGR 高达 30.4%。

图表8: 2018~2025 年全球数据量 CAGR 或达 26.9%



来源: 海天瑞声招股说明书转引自 IDC, 国金证券研究所

图表9: 2018~2025 年中国数据量 CAGR 或达 30.4%

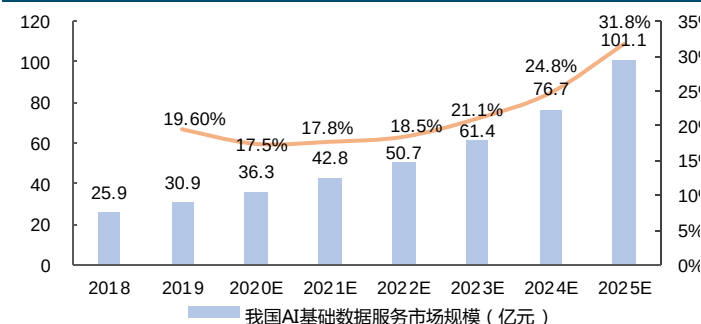


来源: 海天瑞声招股说明书转引自 IDC, 国金证券研究所

中国 AI 基础数据服务行业市场规模 2025 年有望突破百亿。一方面, 随着算法模型、技术理论和应用场景的优化和创新, AI 产业对训练数据的拓展性需求和前瞻性需求均快速增长; 另一方面, 随着业内对训练数据需求类型的增加以及对服务标准要求的提高, 产业链的专业化分工将愈加清晰, 专业化的训练数据服务提供商将扮演更加重要的角色。

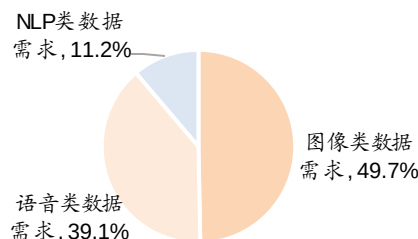
据艾瑞咨询, 2019 年中国 AI 基础数据服务行业市场规模达 30.9 亿元, 其中图像类、语音类、NLP 类数据需求规模占比分别为 49.7%、39.1%和 11.2%。预计 2025 年国内 AI 基础数据服务行业市场规模将突破 100 亿元, 年复合增长率高达 21.8%。

图表10: 2025 年我国 AI 基础数据服务市场或达 101 亿元, 相较 22 年翻倍



来源: 艾瑞咨询《中国 AI 基础数据服务行业发展报告 (2020)》, 国金证券研究所

图表11: 2019 年我国 AI 基础数据服务各类型数据需求占比



来源: 艾瑞咨询《中国 AI 基础数据服务行业发展报告 (2020)》, 国金证券研究所

1.4 政策支持：近五年国家政策加持国内 AI 产业发展

近五年国家系列政策推动国内 AI 产业蓬勃发展。《“十四五”规划》指出要加快数字化发展，建设数字中国，同时打造数字经济新优势，充分发挥海量数据和丰富应用场景优势，促进数字技术与实体经济深度融合，赋能传统产业转型升级，催生新产业新业态新模式，壮大经济发展新引擎。在国家顶层设计的支持下，我国 AI 基础数据服务行业稳步发展，行业训练资源库等细分应用领域的产业价值逐步凸显。

图表12：近五年国家政策加持国内 AI 产业发展

实施时间	颁布主体	主要法律法规及行业政策	相关内容
2017 年 2 月	国家发改委	《战略性新兴产业重点产品和服务指导目录》	人工智能首次进入指导目录名单，包括面向社会开放的文本、语音、图像、视频、地图及行业应用数据等多类型海量训练资源库和标准测试数据集在内的公共数据平台成为人工智能方向的重要子方向之一。
2017 年 7 月	国务院	《新一代人工智能发展规划》	到 2020 年，总体技术和应用与世界先进水平同步；到 2025 年，基础理论实现重大突破，部分技术与应用达到世界领先水平；到 2030 年，理论、技术与应用总体达到世界领先水平，成为世界主要人工智能创新中心。
2017 年 10 月	十八届中央委员会	《十九大报告》	人工智能写入十九大报告，将推动互联网、大数据、人工智能和实体经济深度融合。
2017 年 12 月	工业和信息化部	《促进新一代人工智能产业发展(2018-2020 年)》	《行动计划》从推动产业发展角度出发，结合“中国制造 2025”，对《新一代人工智能发展规划》相关任务进行了细化和落实，信息技术与制造技术深度融合为主线，以新一代人工智能技术，以技术的产业化和集成应用为重点，推动人工智能和实体经济深度融合。
2018 年 1 月	国家电子技术标准化研究院	《人工智能标准化技术白皮书》	全面推进人工智能标准化工作，促进产业发展
2018 年 3 月	国务院	《2018 年国务院政府工作报告》	人工智能再次被列入政府工作报告，加强新一代人工智能研发应用；在医疗、养老、教育、文化、体育等多领域推进“互联网+”发展智能产业，拓展智能生活。
2018 年 4 月	教育部	《高等学校人工智能创新行动计划》	到 2020 年，基本完成适应新一代人工智能发展的高校科技创新体系和学科体系的优化布局；到 2025 年，高校在新一代人工智能领域科技创新能力和人才培养质量显著提升；到 2030 年，高校成为建设世界主要人工智能创新中心的核力量力和引领新一代人工智能发展的人才高地。
2018 年 10 月	中共中央政治局	人工智能发展现状和趋势第九次集体学习	人工智能是引领新一轮科技革命和产业变革的战略性技术，具有溢出带动性很强的“头雁”效应。人工智能正在对经济发展、社会进步、国际政治经济格局等方面产生重大而深远的影响。加快发展新一代人工智能是我们赢得全球科技竞争主动权的重要战略抓手，是推动我国科技跨越发展、产业优化升级、生产力整体跃升的重要战略资源。
2018 年 11 月	工业和信息化部	《新一代人工智能产业创新重点任务揭榜工作方案》	基于《促进新一代人工智能产业发展三年行动计划(2018-2020 年)》中部署的重点任务和目标，人工智能揭榜工作将在 17 个方向及细分领域，征集并遴选一批掌握关键核心技术、具备较强创新能力的创新主体，以突破产业发展的短板和瓶颈，加快我国人工智能产业

			与实体经济深度融合。
2019 年 3 月	中央全面深化改革委员会	《关于促进人工智能和实体经济深度融合的指导意见》	促进人工智能和实体经济深度融合，要把握新一代人工智能发展的特点，坚持以市场需求为导向，以产业应用为目标，深化改革创新，优化制度环境，激发企业创新活力和内生动力，结合不同行业、不同区域特点，探索创新成果应用转化的路径和方法，构建数据驱动、人机协同、跨界融合、共创分享的智能经济形态。
2019 年 8 月	科学技术部	《国家新一代人工智能开放创新平台建设工作的指引》	明确指出“开放、共享”是推动我国人工智能技术创新和产业发展的重要理念，鼓励开放创新平台面向细分领域建设标准测试数据集，促进数据开放和共享，形成标准化、模块化的模型、中间件及应用软件，以开放接口、模型库、算法包等方式向社会提供软硬件开放共享服务。
2019 年 8 月	科学技术部	《国家新一代人工智能创新发展试验区建设工作指引》	明确指出试验区建设以促进人工智能与经济社会发展深度融合为主线，重点任务包括加强网络基础设施、大数据基础设施、计算基础设施建设，提升传统基础设施的智能化水平，形成支撑新一代人工智能广泛应用的基础设施体系。
2020 年 3-4 月	中共中央政治局常委会、国家发改委	中央政治局常委会会议重要讲话	加快新型基础设施建设（“新基建”）进度，新基建范围包括以人工智能、云计算、区块链等为代表的新技术基础设施，以数据中心、智能计算中心为代表的算力基础设施等。
2020 年 4 月	中共中央国务院	《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》	首次将“数据”作为市场化要素写入国家顶层设计级别文件，提出要加快培育数据要素市场，发挥数据在市场化配置中的作用。
2020 年 8 月	国家发改委、科技部、工业和信息化部	《国家新一代人工智能标准体系建设指南》	加强人工智能领域标准化顶层设计，推动人工智能产业技术研发和标准制定，促进产业健康可持续发展。
2021 年 3 月	十三届全国人大四次会议	《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景纲要》	《十四五规划》指出要加快数字化发展，建设数字中国，同时打造数字经济新优势，充分发挥海量数据和丰富应用场景优势，促进数字技术与实体经济深度融合，赋能传统产业转型升级，催生新产业新业态新模式，壮大经济发展新引擎。同时指出要加强关键数字技术创新应用：聚焦高端芯片、操作系统、人工智能关键算法、传感器等关键领域；建设重点行业人工智能数据集，发展算法推理训练场景。
2022 年 7 月	科技部等六部门	《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》	意见指出我国人工智能仍存在对场景创新认识不到位，重大场景系统设计不足，场景机会开放程度不够，场景创新生态不完善等问题，需要加强对人工智能场景创新工作的统筹指导。
2023 年 2 月	北京市经济和信息化局	《2022 年北京人工智能产业发展白皮书》	截至 2022 年 10 月，北京拥有人工智能核心企业 1,048 家，占我国人工智能核心企业总量的 29%，数量位列全国第一。核心区域产业集聚能力全国第一，已经形成了全栈式的人工智能产业链。

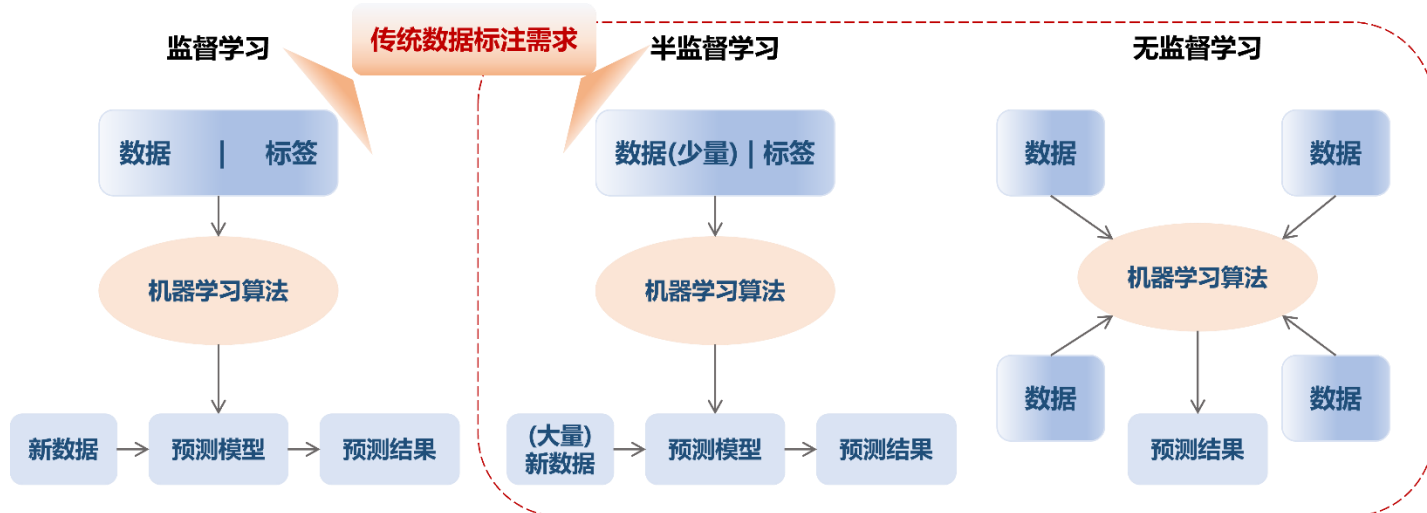
来源：海天瑞声招股说明书转引自中国政府网，中国政府网，国金证券研究所

2. AI 大模型催生高要求新需求，专业化数据集及 AI 训练师需求利好优质专业数据提供厂商

2.1 场景特质：AI 大模型时代无监督/半监督训练成为主流，RLHF 催生新兴需求

大模型时代无监督/半监督训练成为主流。AI 模型的训练方法主要包括监督学习和无监督学习两种典型方式，后随模型训练数据量的增加衍生出半监督学习方法。AI 训练方法的发展历经“监督-无监督-监督-无监督/半监督”4 个阶段，在目前的大模型阶段，无监督/半监督训练再次成为主流。监督学习与无监督学习的主要区别在于是否使用带有人工标注的数据集训练数据，半监督学习则是使用大量未标注数据+少量标注数据进行训练。

图表13：AI 大模型传统数据标注需求下降

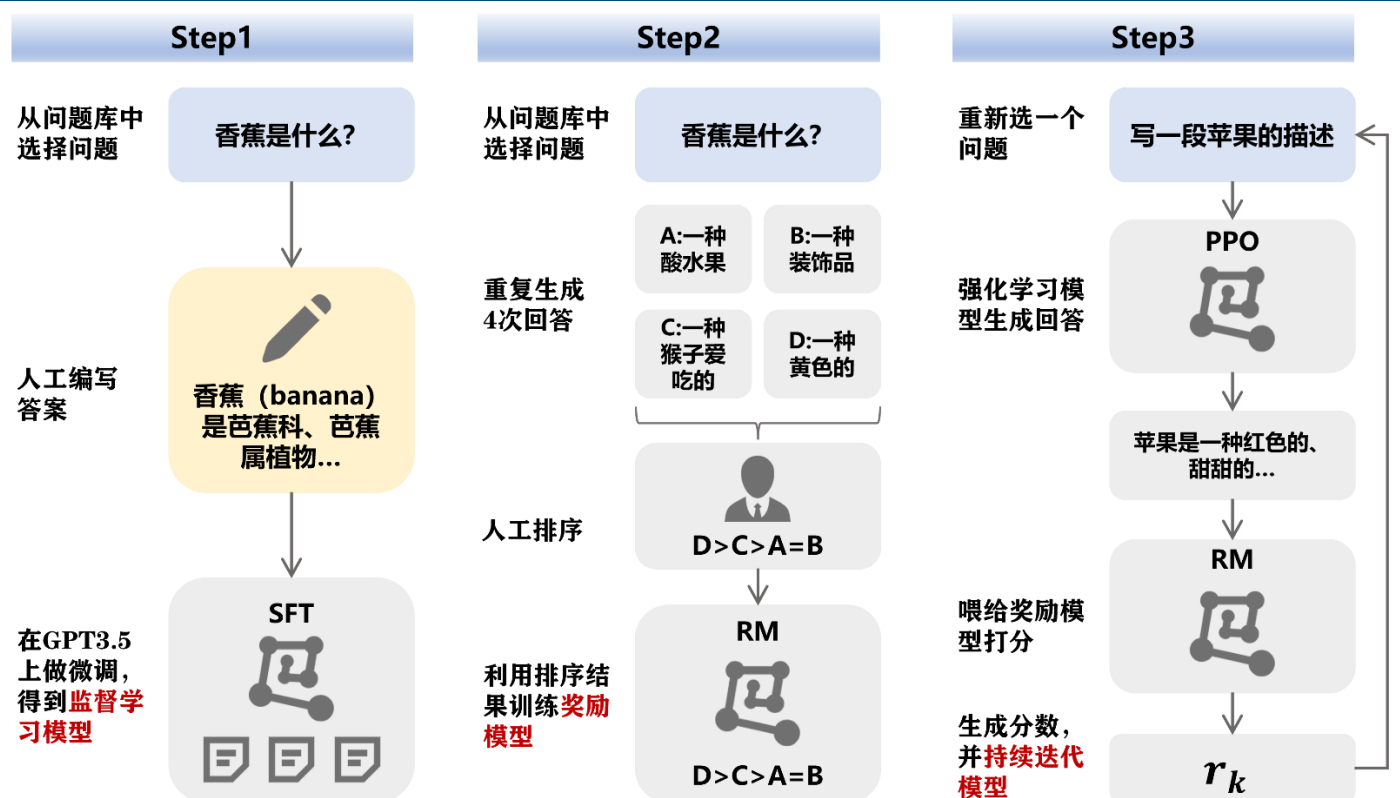


来源：华章科技，阿里云开发者社区，国金证券研究所

大模型时代“基础模型+微调”成为 AI 开发新范式，RLHF 微调技术催生更高要求的数据标注需求。AI 大模型由海量数据通过无监督学习训练得到，本身并不能直接应用于具体任务，必须经过微调才可投入应用。微调是指基于大规模基础模型，在现有训练得到的模型参数之上，针对特定任务类型、应用特定场景的数据对模型进行二次训练。通俗来说，大规模基础模型为 AI 提供了基础知识，而微调则是让 AI 获特定领域知识，并赋予其组织、应用知识的能力。以近日备受关注的 ChatGPT 为例，在其微调技术 RLHF（强人工反馈）系统当中：

- 第一步，预训练阶段。模型首先需要在标注完备的大数据集上进行预训练，得到监督学习模型。
- 第二步，交互奖励阶段。模型与专业的人工智能训练师进行交互，专业标注人员会对 ChatGPT 生成的回答进行标注、评估和反馈，给出一个针对回答的分数或者标签。这些标注数据可以作为强化学习过程中的“奖励函数”来指导 ChatGPT 的参数调整，得到奖励模型。
- 第三步，迭代优化阶段。基于奖励模型的奖励函数以 PPO（一种使用两个神经网络的强化学习算法）的方式微调监督学习训练出来的生成模型，基于强化学习 loss 持续迭代生成模型，最终帮助模型进行强化学习和不断优化。

图表14: OpenAI 开展 RLHF 三步骤



来源: 数字时氦转引自 OpenAI, 国金证券研究所

2.2 场景需求: 预训练阶段高质量专业数据集+交互奖励阶段人工智能训练师

伴随业界大模型市场竞争的白热化, RLHF 系统也将得到进一步的推广及迭代使用, 从而带来两大类新兴数据标注需求。一是预训练阶段催生 AI 厂商对于标注完善、清洗完备的各类专业化场景数据集的需求; 二是交互奖励阶段催生 AI 厂商对于具备专业的事实判断与规范的价值判断的人工智能训练师的需求。

目前, 上述两类新兴需求尚未得到充分有效的市场供给, 利好海天瑞声这类优质专业数据解决方案提供商。针对 RLHF 预训练阶段需求, 由于传统数据采标厂商的主流商业模式以销售工具系统和标注服务为主, 所以普遍缺少自有数据, 较少经营出售精准数据集的服务。针对 RLHF 交互奖励阶段需求, 聚焦专业垂类的模型训练师则更为稀缺, 市场蓝海亟待业务开拓。

图表15: RLHF 需求: 预训练阶段高质量专业数据集+交互奖励阶段人工智能训练师

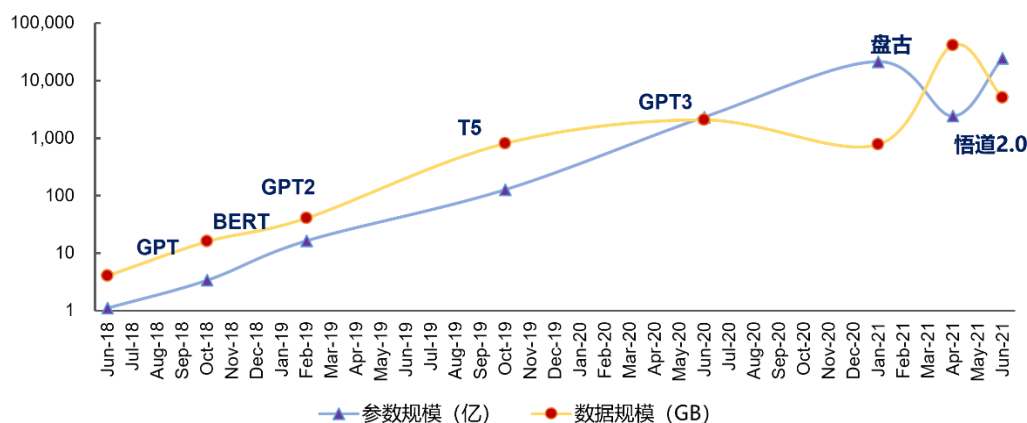


来源: 数字时氦转引自 OpenAI, 海天瑞声公司官网, 国金证券研究所

2.3 场景价值: 超大规模预训练模型推动训练数据指数级跃升, 市场需求持续延展

超大规模预训练模型推动训练数据指数级跃升。自 OpenAI 于 2020 年推出 GPT-3 以来, 谷歌、华为、智源研究院、中科院、阿里巴巴等企业和研究机构相继推出超大规模预训练模型, 包括 Switch Transformer、DALL·E、GPT-4、盘古、悟道 2.0、紫东太初和 M6 等。目前, 预训练模型参数数量、训练数据规模按照 300 倍/年的趋势增长, 继续通过增大模型和增加训练数据仍是短期内主流演进方向, RLHF 技术的推广使用或将推动训练数据市场需求持续延展。

图表16: 超大规模预训练模型参数及数据规模近年来指数级跃升



来源: 中国信通院《人工智能白皮书 (2022 年)》, 国金证券研究所

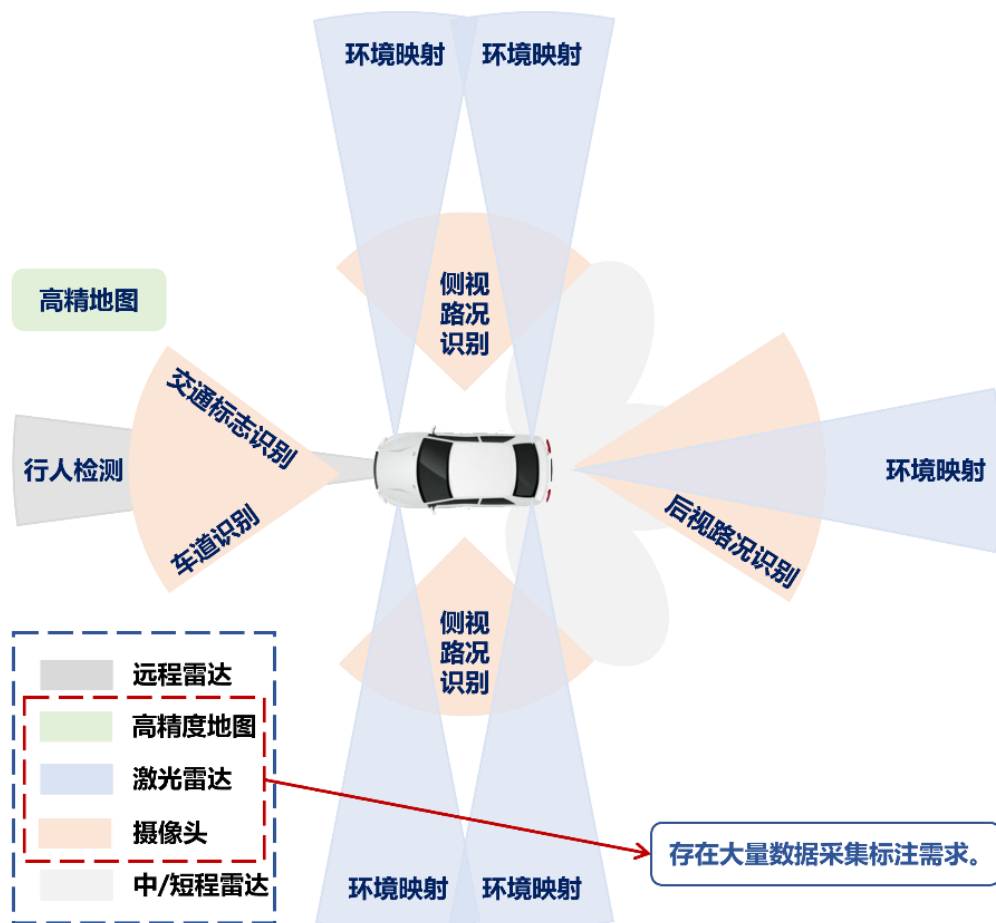
3. 自动驾驶领域训练数据需求方兴未艾, 海天瑞声发力布局

3.1 场景特质: 自动驾驶客户需求全栈式闭环数据解决方案

自动驾驶场景对于 AI 数据服务需求较为刚性。自动驾驶基础数据主要是道路交通图像、障碍物图像、车辆行驶环境图像等, 需求方以科技公司、汽车厂商和高精地图厂商为主。L3 级别以上的自动驾驶系统需对雷达、摄像头等传感器采集的点云和图像数据进行抽取、处理和融合, 构建车辆行驶环境, 为预测和决策做依据。目前自动驾驶的视觉技术主要应

用有监督的深度学习，是基于已知变量和因变量推导函数关系的算法模型，需要大量的标注数据对模型进行训练和调优。近几年，汽车厂商在 ADAS 和自动驾驶方向的投入明显，对于数据的采集和标注需求也逐年增加，汽车厂商有望成为需求主力。

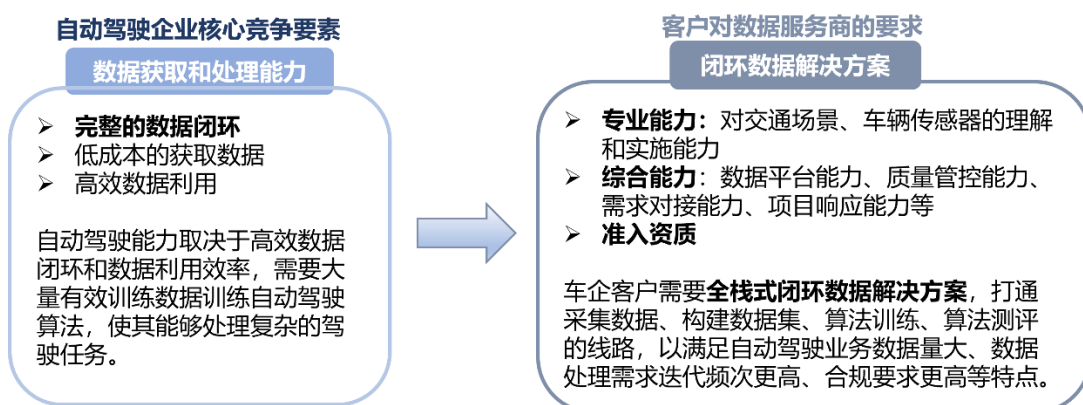
图17：自动驾驶场景中的 AI 基础数据服务需求



来源：艾瑞咨询《2019 年中国人工智能基础数据服务行业白皮书》，国金证券研究所

自动驾驶领域对于基础数据服务商提出更高要求，业内客户需要全栈式闭环数据解决方案。数据获取和处理能力是自动驾驶企业的核心竞争要素之一，自动驾驶能力取决于高效的数据闭环和数据的利用效率，并能利用大量有效数据训练智能驾驶算法。因此，自动驾驶客户要求数据服务商能够提供闭环数据解决方案，以满足智能驾驶业务数据处理量大、数据处理需求迭代频次高等特点，专业知识、服务经验及准入资质将成为衡量的重要标准。

图18：自动驾驶客户需求全栈式闭环数据解决方案



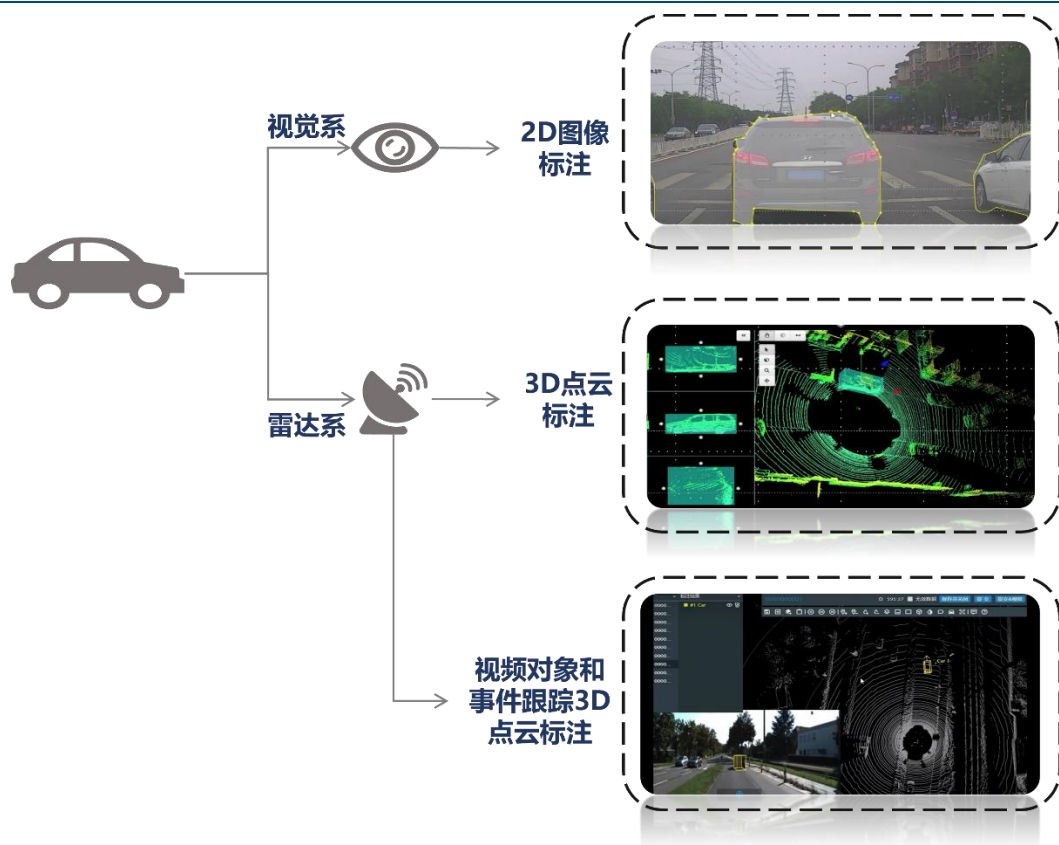
来源：海天瑞声 2021 年报，艾瑞咨询《2019 年中国人工智能基础数据服务行业白皮书》，国金证券研究所

3.2 场景需求：自动驾驶领域数据标注分为 2D 图像标注与 3D 点云标注

自动驾驶领域的的数据可分为车载摄像头采集的 2D 图像数据和激光雷达采集的 3D 点云数据。一般而言,低 level 的自动驾驶技术以 2D 图像数据为主,3D 点云标注数据是中高 level 自动驾驶技术的基础训练数据,在自动驾驶领域中发挥着愈发重要的作用。3D 点云标注数据在自动驾驶领域的应用可以分为两个方面,一是基于场景理解和目标检测的实时环境感和处理,二是 SLAM (即时定位与地图构建) 加强定位。

- 2D 标注: 通过精确理解来自可见光摄像头的信息,寻找能够创建用于目标物体的可扩展边界框。
- 3D 点云标注: 通过识别和跟踪场景中的对象,了解汽车前方和周围的场景。将点云数据和视频流合并到要标注的场景中。
- 视频对象和事件跟踪 3D 点云标注: 锁定随时间移动的对象,并标注时间事件。在多帧视频和 LiDar 场景中跟踪进入和离开本体中的关注区域的对象(如其他汽车和行人)。在整个视频中,无论对象进入和离开视线的频率如何,都会保持对其特性的一致理解。

图表19: 自动驾驶场景中的 AI 基础数据服务业务类型

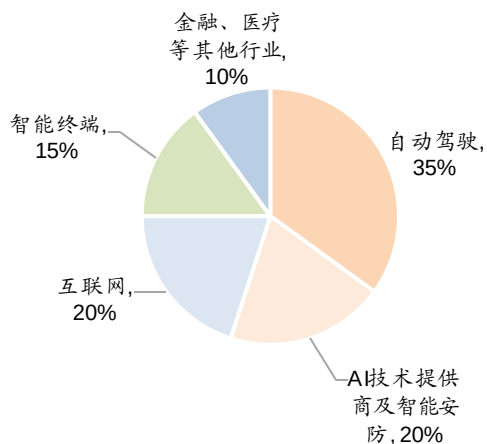


来源: Appen 公司官网, BodenAI, Fotune 激光雷达, 国金证券研究所

3.3 场景价值：自动驾驶训练数据市场方兴未艾，2025 年国内市场规模或达 25 亿元

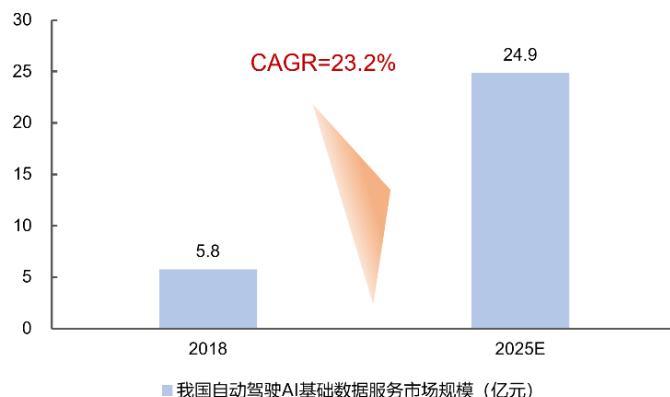
自动驾驶约占我国 AI 基础数据服务市场规模的 35%，2025 年市场规模或达 25 亿元。自动驾驶的视觉技术主要应用于有监督的深度学习,需要大量的标注数据对模型进行训练和调优。目前该领域的数据采集和标注需求已成为 AI 基础数据服务的主要下游之一。据 IDC, 2020 年我国自动驾驶领域占 AI 基础数据服务市场规模的 35%, 系第一大下游场景。另据艾瑞咨询, 2025 年我国自动驾驶 AI 基础数据服务市场规模或达 24.9 亿元, 预计 18~25 年 CAGR 高达 23.2%, 跑赢 AI 基础数据服务整体增速 21.8%。

图表20: 2020 年自动驾驶约占 AI 基础数据服务市场规模的 35%



来源: IDC, 国金证券研究所

图表21: 2025 年我国自动驾驶 AI 基础数据服务市场规模或达 24.9 亿元



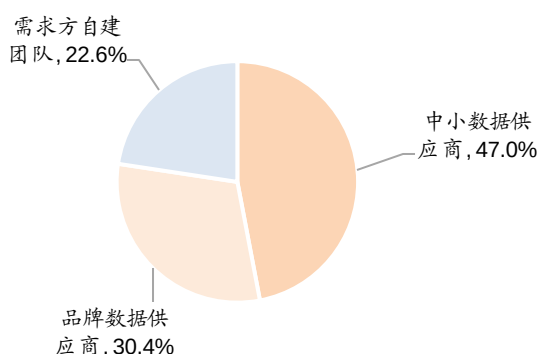
来源: 艾瑞咨询《2019 年中国人工智能基础数据服务行业白皮书》, 国金证券研究所

4. 国内市场集中度趋势性收敛, 海天瑞声具备领先优势

4.1 发展趋势: 品牌商价值效应凸显, 国内市场集中度持续收敛

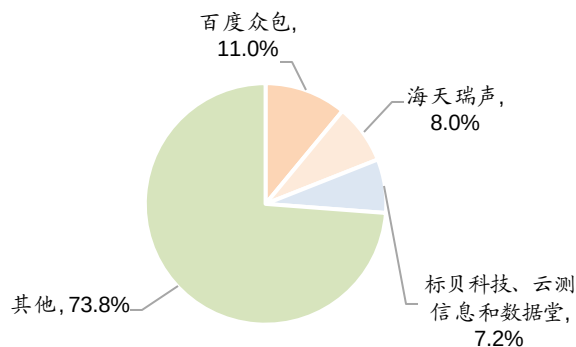
品牌数据服务商未来将替代中小型供应商成为市场主要供应力量。我国 AI 基础数据服务行业市场主体包括需求方自建基础数据团队与基础数据服务商(品牌数据服务商、中小数据服务商)。截至 2019 年, 中小数据供应商是主要供应力量, 占比高达 47%。其中百度众包、海天瑞声分居 top2, 占比分别为 11.0%、8.0%, 2019 年行业 CR5 仅为 26.2%。结合本文 1.2 部分分析可知, 近年间需求端垂直场景及专业化需求凸显, 业务门槛提高; 供给端竞争加剧挤压中低端业务利润空间, 加速市场出清。预计当前我国 AI 基础数据服务行业市场主要以品牌数据供应商与需求方自建团队为主, 行业龙头逐渐浮出水面, 中小供应商份额显著下降。

图表22: 2019 年我国 AI 基础数据服务市场竞争格局



来源: 艾瑞咨询《中国 AI 基础数据服务行业发展报告(2020)》, 国金证券研究所

图表23: 2019 年我国 AI 基础数据服务市场 CR5



来源: 海天瑞声招股说明书, 国金证券研究所

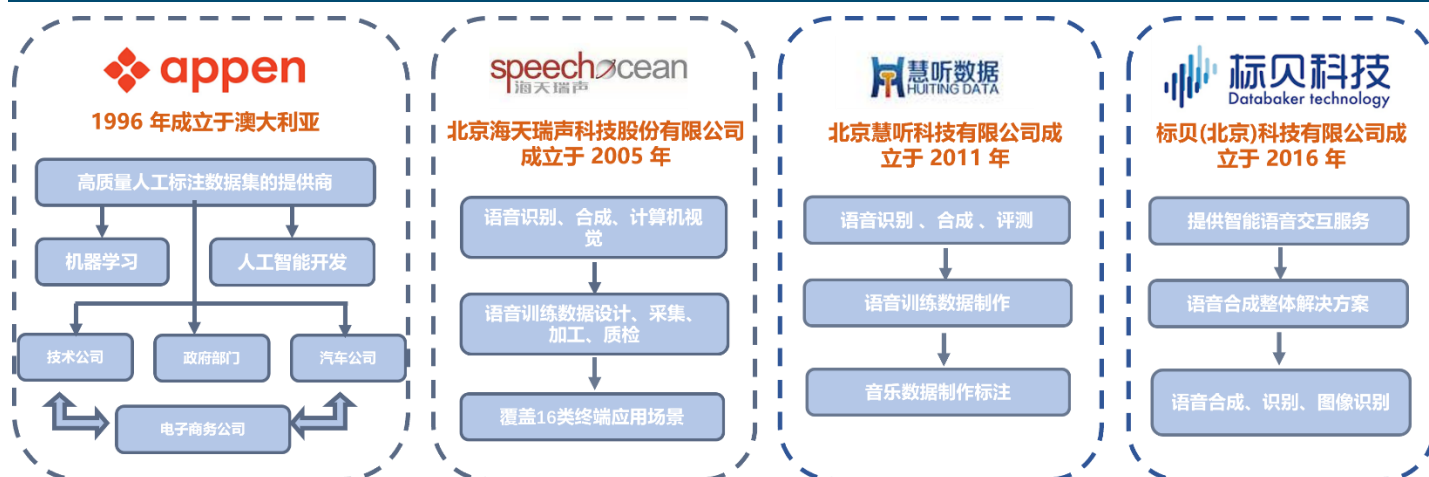
业内品牌数据商包括海外巨头 Appen、国内领军厂商海天瑞声、慧听数据、标贝科技等:

- **Appen:** 1996 年成立于澳大利亚, 面向机器学习和人工智能开发的高质量人工标注数据集, 于 2015 年 1 月于澳大利亚证券交易所上市。Appen 在采集并丰富语音、文字、图像和视频等各种数据类型上积累多年经验, 与全球技术、汽车和电子商务公司, 以及政府部门建立了合作。
- **海天瑞声:** 成立于 2005 年, 是国内领先的训练数据提供商。主要从事训练数据的研发设计、生产及销售业务。目前已成为同时具备核心技术、产品资源、优质客户为一

体的本土训练数据提供企业。核心收入来源为数据资源定制服务。

- 慧听科技：成立于 2011 年，业务包括语音识别、语音合成、语音评测、语言文本类、多媒体类等多领域数据制作，以及语音合成、语音识别、输入法系统的研发等。公司提供服务涵盖语音训练数据制作、音乐数据制作标注、语音质量评测等，经营模式包括定制开发和自有训练数据产品销售。
- 标贝科技：成立于 2016 年，主要提供智能语音交互相关服务，包括语音合成整体解决方案，以及语音合成、语音识别、图像识别等数据服务，经营模式包括定制开发和自有训练数据产品销售。

图表24：业内主要品牌数据商



来源：海天瑞声招股说明书，国金证券研究所

4.2 竞争格局：业内主要玩家发展各有侧重，Appen 及海天瑞声具备领先优势

据 IDC，2021 年海天瑞声在国内 AI 基础数据服务行业市占率高达 12.9%，位居第一。其技术实力、语种/方言覆盖能力、专利及软著数量、成品训练集数量均位居行业前列。相比之下，Appen 在覆盖区域、语种/方言覆盖能力等方面更具优势。慧听科技与标贝科技则在音乐领域具备差异化业务覆盖能力。

图表25：业内主要玩家指标对比

	海天瑞声	Appen	慧听科技	标贝科技
国内市场占有率	AI 基础数据服务行业第二名，市场占有率为 12.9%(2021 年 IDC 数据)	/	/	/
技术实力概述	海天瑞声自主开发了一体化数据处理支撑平台，在基础研究、平台工具、训练数据生产三个维度下均积累核心技术，将多项具体核心技术整合为发行人特有的核心技术体系。	Appen 拥有人工智能辅助数据注释平台，在全球 170 多个国家与 100 多万名专业承包商合作，训练数据涵盖科技、汽车、金融服务、零售、医疗健康和政府等各个领域。	采用全程质量监控流程，执行完善的标注流程，配合保密管理手段，提供质量上乘的数据服务 20 余个。	拥有语音合成模型和算法，通过算法+专业的人工数据处理方式，为客户提供优质的语音合成服务。拥有 TOBI 标注体系，通过自主研发的 TTS 评测系统，为客户提供高质量的数据服务。
语种/方言覆盖能力	190 余个	超过 235 个	20 余个	10 余个
应用领域覆盖	智能语音、计算机视觉、自然语言	智能语音、计算机视觉、自然语言	智能语音、计算机视觉、自然语言、音乐	智能语音、计算机视觉、自然语言、音乐

已获得专利授权	26 项 (24 项发明专利、1 项实用新型专利及 1 项外观设计专利)	/	1 项	4 项
计算机软件著作权数量	159 项	/	14 项	30 项
拥有的成品训练数据集数量	1,054 个	291 个	45 个	179 个

来源：海天瑞声招股说明书、年报，国金证券研究所

注：①Appen、慧听科技、标贝科技数据：截至 2021 年 12 月 31 日；②海天瑞声数据：除特别标注外，均为截至 2022 年 6 月 30 日。

在语音类数据产品中，海天瑞声在结构方面的差异主要体现在覆盖的噪声类型、录音通道数量、录音文本内容类型等方面。相比之下，海天瑞声的训练数据产品结构整体而言可覆盖更丰富的录音文本内容类型、噪音环境类型、录音通道数量，Appen 则在稀有语言覆盖数量上更胜一筹。

图表26：智能语音类训练数据产品对比

训练数据集结构属性	海天瑞声	Appen	慧听科技	标贝科技
数量	579	155	19	86
覆盖语种/方言数量	108	92	19	10
录音场景类别	9	7	4	2
采集设备种类	10	4	5	2
采集通道覆盖数	6	4	4	1

来源：海天瑞声招股说明书，国金证券研究所

图表27：通用语音识别训练数据产品对比

训练数据集结构属性	海天瑞声	Appen	慧听科技	标贝科技
一训练数据产品覆盖的文本类型数量	1-18 类不等	1~10 类不等	1~3 类不等	1 类
文本类型	拼读词、命令词、数字串、自然数、时间、日期、人名、地名、歌曲名、机构名、长句等	数字串、自然数、字符串、人命地名企业名、命令词、申请词、拼音词、时间、日期长句等	命令短语、普通短语、日常短句	拼读词、命令词、数字串、自然数、时间、日期、人名、地名、歌曲名等
单一训练数据产品覆盖的噪声环境类别数量	1~6	1~6	1~4	1~3
噪声环境类型	车内、办公室、家庭、医院、地铁、游戏场餐厅/咖啡厅、商场、街道	办公室、家庭、街边、车内、公共场所、录音棚	办公室/宿舍/家、餐厅/咖啡厅、街道、车载	安静室内、车载、录音棚
单一训练数据产品覆盖的录音通道数量	1~6	1~5	1~6	1

来源：海天瑞声招股说明书，国金证券研究所

图表28：车载语音识别训练数据产品对比

训练数据集结构属性	海天瑞声	Appen	慧听科技	标贝科技
文本内容类型总数	55	8	/	36
文本内容类型示例	数字串、街道名称、控制词、地名、命令词、英语、数字串、时间、日期、度量衡、长句、自然口语等	数字、自然数、街道名称、通用指令和特定指令、字母串、地点、企业名称、自然语句等	/	导航找音乐，城市导航，机车系统控制，查火车，查酒店，查餐厅，查地铁，路况查询等
各训练数据产品覆盖的噪声环境类别数量	7	2	5	7
覆盖车载噪声环境具体类别	怠速、低速、低速噪音、高速、高速噪音、城市道路、城市道路噪音	怠速、高速	高速路况、低速路况、停车待机、车窗打开、车窗关闭	怠速、低速、低速噪音、高速、高速噪音、城市道路、城市道路噪音
各训练数据产品覆盖的录音通道数量	4	5	4	1

来源：海天瑞声招股说明书，国金证券研究所

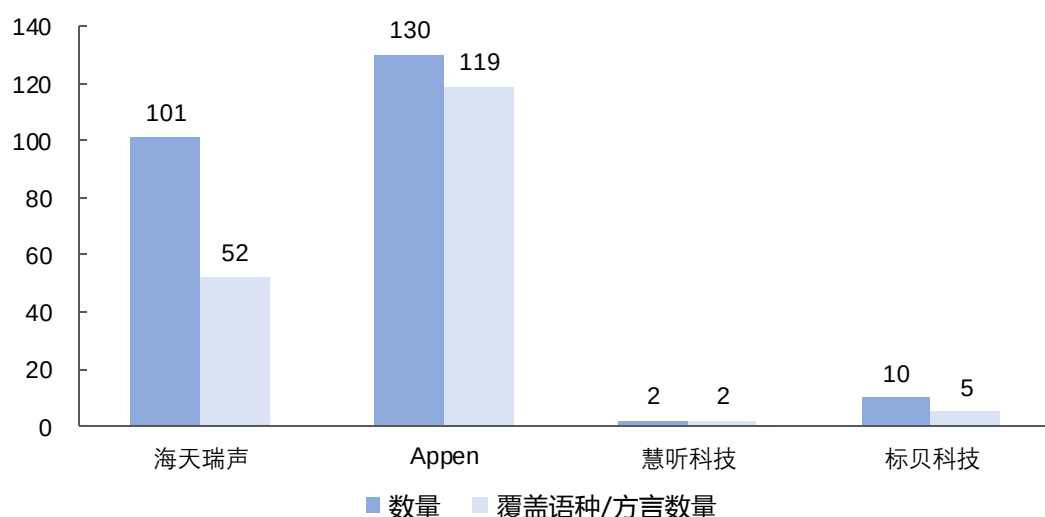
图表29：特色语音识别训练数据产品对比

训练数据集结构属性	海天瑞声	Appen	慧听科技	标贝科技
稀有语言覆盖数量	33	41	0	0

来源：海天瑞声招股说明书，国金证券研究所

在自然语言训练数据产品中，不同自然语言训练数据产品的结构差异主要体现在文本来源、内容、标注属性等方面。Appen 具备更为丰富的产品数量与更为全面的语言覆盖能力，海天瑞声在上述指标上紧随其后。

图表30：自然语言训练数据产品对比



来源：海天瑞声招股说明书，国金证券研究所

在计算机视觉训练数据产品中，海天瑞声在结构方面的差异主要体现在人像识别检测相关数据产品领域。主要玩家均可覆盖不同类型的人脸、姿势、场景等，因计算机视觉数据产品个性化特征较强，各类训练数据产品在具体的人脸类型、姿势范围、场景等方面具备一定差异。与 Appen 相比，海天瑞声的 OCR 训练数据产品可覆盖更多语种的光学字符识别需要，具备广泛性、强适用性特征。

图表31: 计算机视觉训练数据产品结构对比

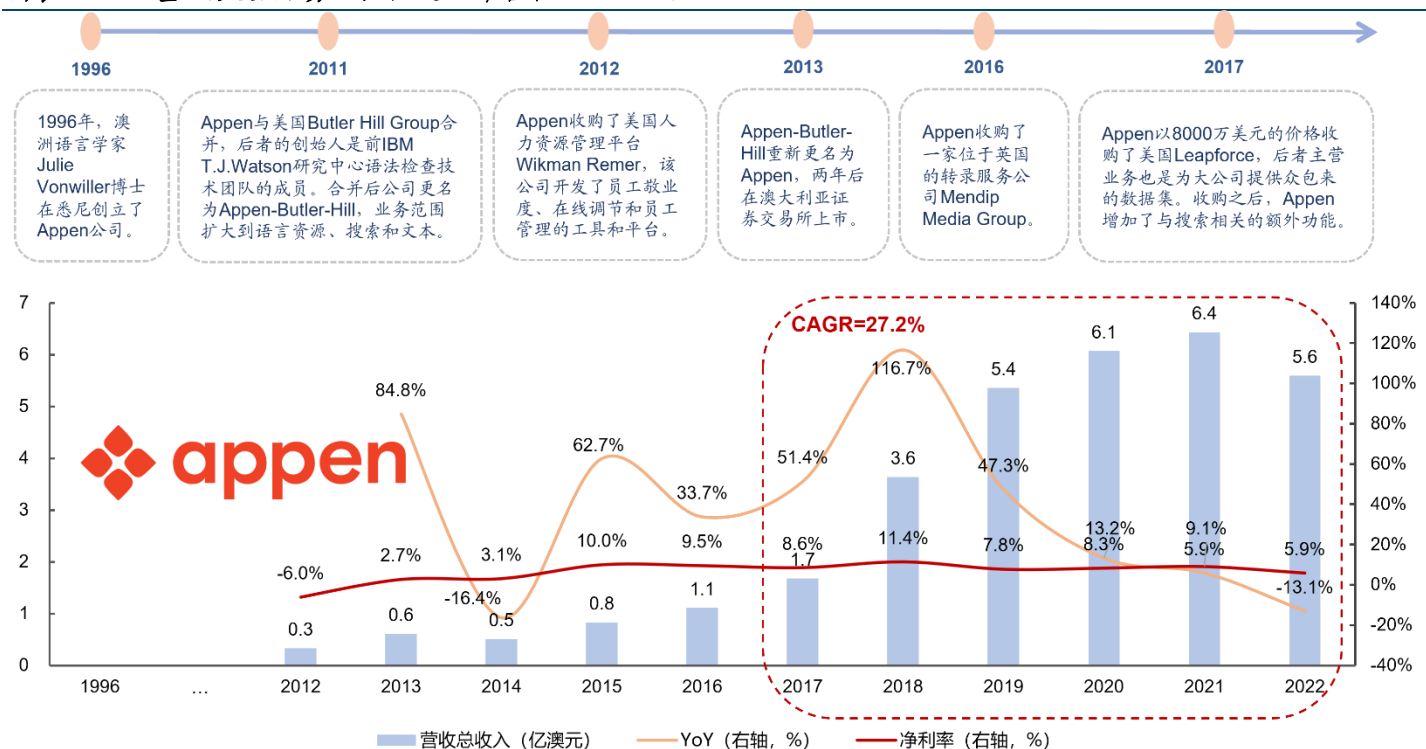
训练数据结构属性	海天瑞声	Appen	慧听科技
人像识别检测相关	包含了约 3 万不同人脸数据, 覆盖黑、白、黄、棕等不同肤色。覆盖闸机口安检库、手机拍照库、人物表情库等应用场景。	包含 1 万多张黑人在不同角度和光线下的人脸图片; 以及 1,000 名中国人身体动作视频等。	包含 2 万张人物轮廓图片, 覆盖不同年龄段男女老少、各类着装打扮, 各类行动姿势(站立、行走、蹲坐等)。
OCR (光学字符识别)	覆盖 10 余个语种, 覆盖菜单、路牌、收据、书籍等 17 类不同场景。	包含中文、泰语、芬兰语 3 个语种, 覆盖书籍、报刊、门票、路牌、菜单等 16 个不同场景。	/

来源: 海天瑞声招股说明书, 国金证券研究所

4.3 他山之石: Appen 发展复盘, 全球 AI 基础数据服务巨头的崛起之路

Appen 系 AI 基础数据服务行业全球巨头, 经历爆发式成长后收入利润短期双双回落。Appen 成立于 1996 年, 面向机器学习和人工智能开发的高质量人工标注数据集, 于 2015 年于澳大利亚证券交易所上市。2015~2019 年, 公司经历了爆发式成长期, 营收增速一度冲高至 100% 以上, 同期净利率接近 10%。2020~2022 年, 受到全球疫情及公司内生发展阶段调整影响, 2022 年公司营收规模相较 2021 年回落 13.1%, 净利率由 9.1% 回落至 5.9%。

图表32: AI 基础数据服务巨头, 近五年营收 CAGR 达 27.2%



来源: Wind, Appen 公司官网、公司年报, 晨哨并购, 国金证券研究所

Appen 业务类型齐全, 客户质量较高。Appen 主要业务类型包括数据采集、数据预处理与模型评价三大类。其中, 数据采集环节基于公司预先标记完备的海量数据集为下游客户提供大量其所需的高质量数据, 数据集语种覆盖范围领先优势较大。数据预处理环节, 公司具备全类型数据标注能力。模型评价环节, 公司通过用户测试及针对客户竞争对手的性能基准测试, 诊断模型潜在问题, 并为此提供能够优化模型的训练数据。公司主要客户质量较高, 包括 Google、Amazon、Microsoft 等知名 IT 巨头。

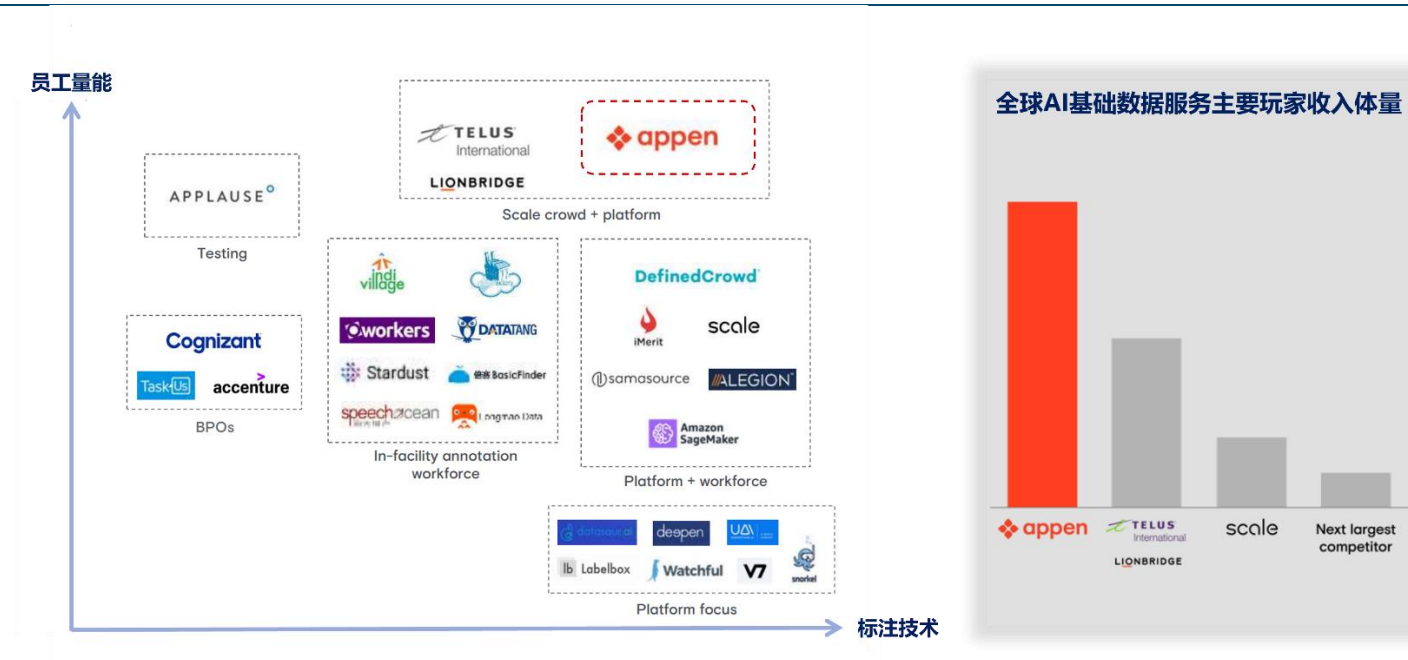
图表33: Appen 业务类型齐全, 客户质量较高



来源: Appen 投资者关系材料, 国金证券研究所

Appen 位居同业员工量能及数据标注技术前列。全球范围来看, Appen 位居全球 AI 基础数据服务行业收入体量首位。其员工量能(数量及质量)与数据标注技术领先程度均位居行业头部, 紧随之后的是 Telus 和 Lionbridge。以数据堂、海天瑞声为代表的国内头部厂商距离 Appen 等海外大厂仍存在一定差距。

图表34: Appen 处于员工量能及标注技术矩阵头部位置



来源: Appen 投资者关系材料, 国金证券研究所

5. 投资建议

AI 基础数据服务行业, 需求侧 AI 垂直场景及专业化需求凸显, 数据需求量提升但数据采集业务门槛提高; 供给端国内数据安全标准及相关资质要求提升, 相关解决能力向头部厂商聚集。推荐关注国内业内领军海天瑞声。

图表35：推荐关注国内 AI 基础数据服务领军者海天瑞声

股票代码	公司名称	推荐理由
海天瑞声	688787.SH	海天瑞声是国内领先的训练数据提供商，主要从事训练数据的研发设计、生产及销售业务。凭借丰富的技术积累和行业经验，公司已成为同时具备核心技术、产品资源、优质客户为一体的本土训练数据领军企业。公司主营专业训练数据集，现已全面布局智能驾驶数据服务领域，位居中国语音类基础数据服务领域首位。Wind 一致预计公司 22、23、24 年摊薄 EPS 分别为 2.21、1.25、2.72 元，对应 87X、153X、70X PE。

来源：Wind，国金证券研究所

注：盈利预测及估值来自 Wind 一致预期。

6. 风险提示

■ 产业需求快速迭代带来研发投入回收不确定性风险。

由于人工智能行业整体发展迅速，需求变化频繁，市场参与玩家进行数据开发所需的各项技术面临快速迭代更新的风险。如果市场主体无法保持对行业发展趋势的洞察能力及持续研发能力，可能失去技术领先优势，对行业格局造成不确定性扰动。

■ 训练数据泄露引发的安全性风险。

基础数据服务系 AI 产业的基础性环节，涉及海量原始数量的采标工作，受到行业数据安全监管要求变化的影响较大。若开发环节未能遵守最新的数据产品保护相关规定，行业可能面临训练数据产品被泄露、盗版等数据安全相关风险。

■ 市场竞争加剧导致利润受到挤压的风险。

在近期 ChatGPT 等 AIGC 概念持续催化的背景下，众多公司参与到市场竞争之中，市场出清进一步加速。中短期内业内主要玩家盈利能力或受到价格竞争持续挤压利润的影响。

行业投资评级的说明:

买入: 预期未来 3-6 个月内该行业上涨幅度超过大盘在 15%以上;
增持: 预期未来 3-6 个月内该行业上涨幅度超过大盘在 5%-15%;
中性: 预期未来 3-6 个月内该行业变动幅度相对大盘在 -5%-5%;
减持: 预期未来 3-6 个月内该行业下跌幅度超过大盘在 5%以上。

特别声明:

国金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。

本报告版权归“国金证券股份有限公司”(以下简称“国金证券”)所有,未经事先书面授权,任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发,或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发,需注明出处为“国金证券股份有限公司”,且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料,但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法,故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致,国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断,在不作事先通知的情况下,可能会随时调整,亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用,在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险,可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突,而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品,使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况,以及(若有必要)咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议,国金证券不就报告中的内容对最终操作建议做出任何担保,在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下,国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易,并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密,只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》,本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级(含 C3 级)的投资者使用;本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要,不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具,本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资,遭受任何损失,国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告,则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议,国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有,保留一切权利。

上海	北京	深圳
电话: 021-60753903	电话: 010-85950438	电话: 0755-83831378
传真: 021-61038200	邮箱: researchbj@gjzq.com.cn	传真: 0755-83830558
邮箱: researchsh@gjzq.com.cn	邮编: 100005	邮箱: researchsz@gjzq.com.cn
邮编: 201204	地址: 北京市东城区建内大街 26 号	邮编: 518000
地址: 上海浦东新区芳甸路 1088 号	新闻大厦 8 层南侧	地址: 中国深圳市福田区中心四路 1-1 号
紫竹国际大厦 7 楼		嘉里建设广场 T3-2402