

2019年5月17日

科创板人工智能系列之二：海天瑞声
计算机
主要财务指标（单位：百万元）

	2016A	2017A	2018A
营业收入	84.23	119.07	192.66
(+/-)	-	41.37%	61.80%
营业利润	12.98	39.67	72.01
(+/-)	-	205.52%	81.53%
归属母公司净利润	10.29	34.15	67.14
(+/-)	-	231.90%	96.61%
EPS（元）	-	1.17	2.24

● 海天瑞声是科创板已受理的人工智能企业之一。海天瑞声是人工智能数据资源服务商，主营业务是人工智能数据资源产品和服务的研发与销售，主要产品及服务包括数据资源定制服务、数据库产品和数据资源相关的应用服务。公司为下游客户提供完成整个人工智能系统开发所需的端到端、一站式数据研发及相关服务，服务链条贯穿上下游人工智能技术开发全周期。公司产品线已包含全球 130 余个主要语种及方言，在欧洲建立的语言研究中心支持了全球化业务的拓展和实施。

● 人工智能数据资源产品和服务的市场规模正不断扩大。人工智能目前正处于商业化的初级阶段，随着应用场景的不断拓展，人工智能产业链的专业化分工将愈加清晰。人工智能算法模型的不断迭代将衍生出大量的训练数据资源和服务需求，带来数据资源产品和服务市场规模的不断扩大，专业化的数据资源服务商将在产业链上扮演重要的角色。

● 2017 年度中国智能语音企业市场海天瑞声排名第七。根据中国语音产业联盟出具的《中国智能语音产业发展白皮书 2017-2018》，入围该报告中国智能语音企业市场规模排行的智能语音企业共 8 家，前三强分别为科大讯飞、百度和苹果。公司位居第 7 名，是唯一入围的数据资源服务企业。

● 优质客户资源凸显公司实力。公司是国内人工智能基础数据资源领域较少的具有全球化支持和交付能力，并且可以在智能语音、计算机视觉和自然语言领域实现全面业务覆盖的数据资源提供商。公司客户包括阿里巴巴、腾讯、百度、微软、三星等大型科技公司，科大讯飞、商汤科技、云知声、海康威视等人工智能企业，以及中国科学院、清华大学等科研机构。

● 对标公司估值：对标公司 Appen 的 PE-TTM 为 62.78 倍。整体行业产业规模的增长和细分需求的丰富，将共同促使市场需求不断扩大。海天瑞声将数据处理技术与人工智能技术进行有效整合，深入开展计算语言学并将其融入到语音数据资源库的研发中，预计海天瑞声上市后市场将会给予较高估值。

● 风险提示：研发能力未能匹配市场需求的风险、数据安全相关风险、新业务拓展不及预期的风险等等。

分析师：徐鹏
 执业证书编号：S1050516020001
 电话：021-54967573
 邮箱：xupeng@cfsc.com.cn

华鑫证券有限责任公司
 地址：上海市徐汇区肇嘉浜路 750 号
 邮编：200030
 电话：(86 21) 64339000
 网址：http://www.cfsc.com.cn

目录

一、公司概况	3
1.公司简介	3
2.我国正成为世界人工智能领域的新增长极	4
3.主要客户	5
二、财务数据	6
三、公司股权结构	6
四、对标公司分析及同类公司估值	7
五、风险提示	7

图表 1 人工智能技术架构示意	3
图表 2 数据资源定制服务内容	3
图表 3 数据库产品	3
图表 4 2017 年度中国智能语音企业市场份额	4
图表 5 2015-2020 年全球人工智能市场规模（亿元）	5
图表 6 中国人工智能市场规模（亿元）	5
图表 7 2018 公司前五大客户	5
图表 8 2018 年公司营收划分（按产品）	6
图表 9 公司营收及净利润(2016-2018 年)	6
图表 10 公司股权结构	6
图表 11 募集资金用途	7
图表 12 主要竞争对手及部分重要同行业公司估值（截止 2019 年 5 月 15 日）	7

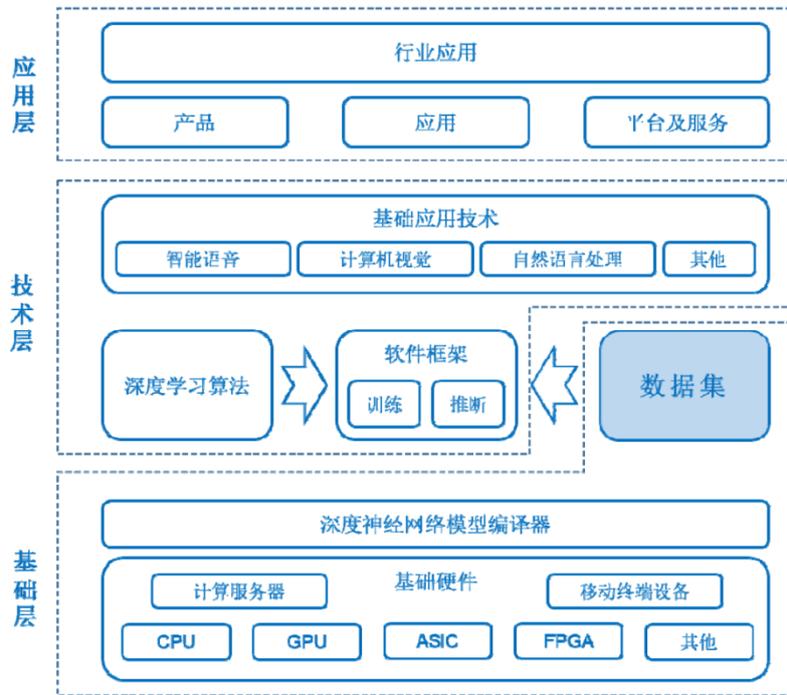
一、公司概况

1. 公司简介

海天瑞声是人工智能数据资源服务商，主营业务是人工智能数据资源产品和服务的研发与销售，主要产品及服务包括数据资源定制服务、数据库产品和数据资源相关的应用服务。

海天瑞声所提供的人工智能数据资源产品及服务隶属于人工智能产业链的基础层，是自主研发人工智能技术的企业与机构必需的基础生产要素，其数量多寡和质量高低将会直接影响到人工智能产业链内企业的研发周期、产品性能和可扩展性。

图表 1 人工智能技术架构示意



资料来源：公司招股说明书申报稿

图表 2 数据资源定制服务内容

序号	应用领域	内容简介	用途
1	智能语音领域	发行人根据客户相关要求，针对特定的应用领域，进行定制化的语音识别数据库、语音合成数据库、词典资源数据库、文本资源数据库、图像资源数据库和视频资源数据库等的设计及开发；或对客户提供的数据资源进行处理，最终形成符合客户需求的数据资源产品。	用以训练、开发、拓展算法模型，并应用到相应的人工智能领域。
2	计算机视觉领域		
3	自然语言领域		

资料来源：公司招股说明书申报稿

图表 3 数据库产品

序号	领域	产品示例	内容简介	用途
1	智能语音领域	语音识别数据库	发行人根据对人工智能算法模型应用领域、行业发展趋势、市场需求等的评估和研判，设计开发的由不同内容和结构的智能语音、计算机视觉、自然语言等领域的数据库资源组成的通用型数据库产品，发行人拥有自主知识产权。	用以训练、开发、拓展算法模型，并最终应用到相应人工智能领域。
		语音合成数据库		
2	计算机视觉领域	图像资源数据库	一般情况下，公司的一个数据库产品所提供的数据库资源即可满足一般规模的智能语音应用系统上线所需要的算法模型训练数据。	
		视频资源数据库		
3	自然语言领域	文本资源数据库		
		词典资源数据库		

资料来源：公司招股说明书申报稿

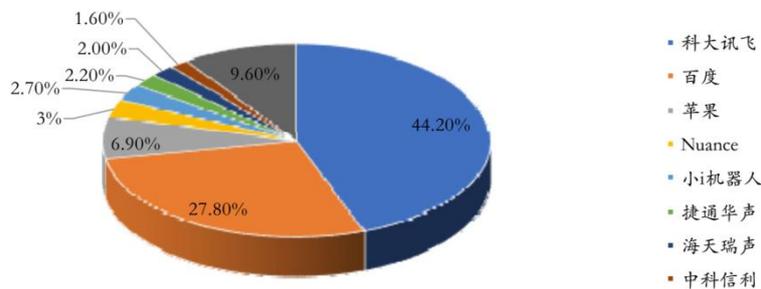
海天瑞声是我国最早专业从事人工智能数据资源产品和服务研发与及销售的主要企业之一。可以为下游客户提供完成整个人工智能系统开发所需的端到端、一站式数据研发及相关服务，服务链条贯穿上下游人工智能技术开发全周期。

公司与清华大学、北京语言文化大学等高校联合举办人工智能竞赛及研究合作项目。公司积累了多通道录音技术、基于 C/S 和 B/S 架构的数据处理技术、移动设备数据采集技术、音素边界预测技术、物体识别技术等人工智能核心技术。

海天瑞声是国内人工智能基础数据资源领域较少的具有全球化支持和交付能力的公司，是国内为数不多的可以在智能语音、计算机视觉和自然语言领域实现全面业务覆盖的数据资源提供商。目前产品线已涵盖全球 130 余个主要语种及方言，可以为客户提供在智能语音、无人驾驶、智慧城市等应用领域的的数据资源解决方案。公司在欧洲建立的语言研究中心有力地支持了全球化业务的拓展和实施。随着全球化的技术整合浪潮以及“一带一路”战略的顺利推进，具备全球化支持和交付能力的国内厂商在人工智能领域将获得更多的机会和更高的成长速度。

公司是我国智能语音产业中市场份额前十名的企业。根据中国语音产业联盟出具的《中国智能语音产业发展白皮书 2017-2018》，入围该报告中国智能语音企业市场规模排行的智能语音企业共 8 家，其中海天瑞声位居第 7 名，是唯一入围的数据资源服务企业。除海天瑞声以外的 7 家企业均为大型科技公司或人工智能企业，位于人工智能产业链的应用层及技术层，主营业务中与人工智能相关的部分主要系人工智能技术及应用开发；而海天瑞声处于基础层、主要为人工智能产业链中处于应用层及技术层的机构提供数据资源及相关服务。

图表 42017 年度中国智能语音企业市场份额



资料来源：《中国智能语音产业发展白皮书 2017-2018》，公司招股说明书申报稿

公司客户遍布全球，赢得了众多大型优质客户的认可。客户有阿里巴巴、腾讯、百度、微软、三星等大型科技公司，科大讯飞、商汤科技、云知声、海康威视等人工智能企业，以及中国科学院、清华大学等科研机构。公司自主开发了一体化人工智能数据处理技术支撑平台来实现全面化精细化数据开发流程管理。基于该平台，公司在 2018 年开发的数据总量超过 60,000 人次，处理的语音时长超过 26 万小时，包含 50 个语种或方言。

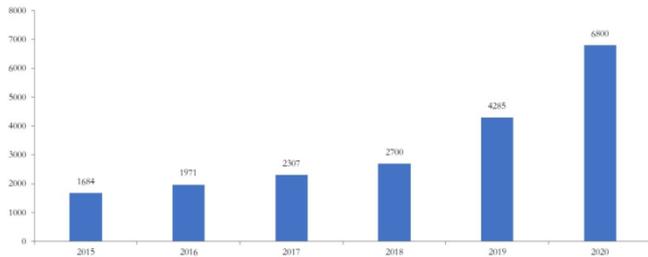
2.我国正成为世界人工智能领域的新增长极

近年来基于深度学习的智能语音、计算机视觉、自然语言处理等技术开始向各个应用领域渗透，全球人工智能产业规模快速增长。2017 年，全球人工智能产业规模达到 2,307 亿元，预计 2020 年全球人工智能市场规模将达 6,800 亿元。

当前，我国正在成为世界人工智能领域的新增长极。截至到 2018 年 6 月，全球人工智能企业总数达到 4,925 家，中国（不含港澳台）人工智能企业总数为 1,011 家，

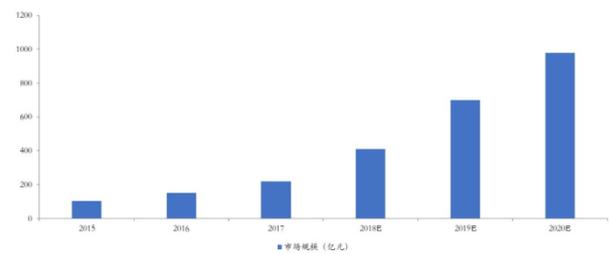
仅次于美国。根据清华大学《中国人工智能发展报告 2018》数据，2017 年中国人工智能行业投融资总额高达 277.1 亿美元，预计在 2018 年，我国人工智能市场整体规模将达到 415.50 亿元，增速将达到 75%。

图表 5 2015-2020 年全球人工智能市场规模（亿元）



资料来源：中国产业信息网，公司招股说明书申报稿

图表 6 中国人工智能市场规模（亿元）



资料来源：清华大学《中国人工智能发展报告 2018》，公司招股说明书申报稿

海天瑞声所在细分行业为“人工智能数据资源开发及服务行业”，属于人工智能行业。

人工智能算法模型的持续更新需要将衍生出大量的训练数据资源需求，各领域数据资源未来仍有较大的需求空间。根据 MckinseyGlobalInstitute 研究结果：在大多数应用场景中有效使用神经网络，需要大量的经标注训练数据和有效的计算基础设施，其中深度学习模型对训练数据的需求量更为显著。要发挥人工智能的完整潜能，则需要包括图像、视频及语音在内的多种数据资源，并要求算法、模型根据潜在的应用场景变化持续更新，相应地，人工智能算法模型所使用的训练数据也需要定期更新。

未来人工智能数据资源产品和服务的市场需求还将不断扩大。一方面，随着人工智能行业的迅速发展和规模增长，将带动数据资源产品与服务的需求量快速增长。另一方面随着行业内对数据资源产品与服务标准的提高和需求类型的丰富，人工智能产业链的专业化分工将愈加清晰，专业化的数据资源服务提供商将在人工智能产业链上的扮演更加重要的角色，数据库服务也将从通用数据资源服务向行业数据资源服务不断发展。

3. 主要客户

公司的目标客户主要为人工智能产业链上的各类机构。2018 年，公司前五大客户销售收入占比为 59.06%。其中，阿里巴巴为公司第一大客户，销售额占比为 26.88%。

图表 7 2018 公司前五大客户

序号	客户名称	销售额 (万元)	占比
1	阿里巴巴	5,179	26.88%
2	三星	2,437	12.65%
3	腾讯	1,574	8.17%
4	微软	1,300	6.75%
5	百度	888	4.61%
合计		11,378	59.06%

资料来源：公司招股说明书申报稿

本次公司申请在科创板上市，拟募集资金 7 亿元，主要用于“天籁”自主研发产品扩建、一体化技术支撑平台建设、研发中心建设等主业。

图表 11 募集资金用途

序号	项目名称	总投资额 (万元)	使用募集资金 投入金额(万元)
1	“天籁”自主研发产品扩建项目	21,928	21,928
2	一体化技术支撑平台建设项目	21,048	21,048
3	研发中心建设项目	16,239	16,239
4	业务管理平台建设项目	3,328	3,328
5	补充流动资金	10,000	10,000
合计		72,542	72,542

资料来源：公司招股说明书申报稿，华鑫证券研发部

四、对标公司分析及同类公司估值

人工智能数据资源开发及服务行业中的国内企业主要有慧听科技、标贝科技等，国外企业主要有 Appen（澳大利亚上市）等。Appen Limited，1996 年成立于澳大利亚，是面向机器学习和人工智能开发高质量人工标注数据集的提供商，Appen 在捕捉并丰富语音、文字，图像和视频等各种数据类型上积累了多年经验，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。慧听科技、标贝科技均为非上市公司，关键业务数据及财务指标均未披露。

对标公司 Appen 的 PE-TTM 为 62.78 倍。海天瑞声是我国最早专业从事人工智能数据资源产品和服务研发与及销售的主要企业之一，人工智能产业链的专业化分工将愈加清晰，整体行业产业规模的增长和细分需求的丰富，将共同促使市场需求不断扩大，预计海天瑞声上市后市场将会给予较高估值。

图表 12 主要竞争对手及部分重要同行业公司估值（截止 2019 年 5 月 16 日）

代码	公司简称	上市地	股价 (AUD/股)	总市值 (AUD亿元)	EPS TTM	PE TTM
APX.AX	Appen Limited	澳大利亚	24.27	29.35	0.38	63.87

资料来源：Yahoo Finance，华鑫证券研发部

五、风险提示

半导体行业的风险主要包括：研发能力未能匹配市场需求的风险、采购成本上升的风险、数据安全相关风险、新业务拓展不及预期的风险等等。

分析师简介

徐鹏：工学硕士，2013年6月加盟华鑫证券，主要研究和跟踪领域：TMT、可转债等。

华鑫证券有限责任公司投资评级说明

股票的投资评级说明：

	投资建议	预期个股相对沪深 300 指数涨幅
1	推荐	>15%
2	审慎推荐	5%—15%
3	中性	(-)5%— (+)5%
4	减持	(-)15%— (-)5%
5	回避	<(-)15%

以报告日后的 6 个月内，证券相对于沪深 300 指数的涨跌幅为标准。

行业的投资评级说明：

	投资建议	预期行业相对沪深 300 指数涨幅
1	增持	明显强于沪深 300 指数
2	中性	基本与沪深 300 指数持平
3	减持	明显弱于沪深 300 指数

以报告日后的 6 个月内，行业相对于沪深 300 指数的涨跌幅为标准。

免责声明

华鑫证券有限责任公司（以下简称“华鑫证券”）具有中国证监会核准的证券投资咨询业务资格。本报告由华鑫证券制作，仅供华鑫证券的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告中的信息均来源于公开资料，华鑫证券研究发展部及相关研究人员力求准确可靠，但对这些信息的准确性及完整性不作任何保证。我们已力求报告内容客观、公正，但报告中的信息与所表达的观点不构成所述证券买卖的出价或询价的依据，该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时结合各自的投资目的、财务状况和特定需求，必要时就财务、法律、商业、税收等方面咨询专业顾问的意见。对依据或者使用本报告所造成的一切后果，华鑫证券及/或其关联人员均不承担任何法律责任。

本报告中的资料、意见、预测均只反映报告初次发布时的判断，可能会随时调整。该等意见、评估及预测无需通知即可随时更改。在不同时期，华鑫证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。华鑫证券没有将此意见及建议向报告所有接收者进行更新的义务。

本报告版权仅为华鑫证券所有，未经华鑫证券书面授权，任何机构和个人不得以任何形式刊载、翻版、复制、发布、转发或引用本报告的任何部分。若华鑫证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，华鑫证券对此等行为不承担任何责任。本报告同时不构成华鑫证券向发送本报告的机构之客户提供的投资建议。如未经华鑫证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。华鑫证券将保留随时追究其法律责任的权利。请投资者慎重使用未经授权刊载或者转发的华鑫证券研究报告。

华鑫证券有限责任公司
研究发展部
地址：上海市徐汇区肇嘉浜路 750 号
邮编：200030
电话：(+86 21) 64339000
网址：<http://www.cfsc.com.cn>