

2019年05月15日

海天瑞声拟科创板上市，有望成为人工智能数据服务国内龙头？

新三板策略报告

证券研究报告

## 科创板受理公司巡礼系列之（九）

### ■内部挖掘：提供基础产品，海天瑞声何以市场份额前十？

- ✓ 公司成立于2005年05月11日，从事为人工智能全产业链中的企业及科研机构提供数据资源产品和服务，业务领域涵盖了智能语音、计算机视觉、自然语言等人工智能核心领域。2018年公司营业收入为1.92亿元(+89.78%)；净利润为6474.16万元(+96.61%)。
- ✓ 海天瑞声是全球领先的人工智能数据资源服务商，连续多年荣获“国家高新技术企业”及“中关村高新技术企业”称号，2018年被评为国家规划布局内重点软件企业、以及中关村高成长Top100企业。
- ✓ 人才优势较为突出，公司高管及核心人员大多毕业于清华、北大、中科大、复旦等院校，大部分曾在微软、英特尔、IBM、中科院等业内领先成熟企业与研究机构担任人工智能领域技术研发与管理的领导职务。员工全部接受过高等教育，本科及以上学历合计占比超过95%。
- ✓ 营收主要来自于数字资源定制服务，2018年对营收贡献超60%；三大人工智能领域中重点渗透智能语音领域，2018年对营收贡献超80%；主要面向境内销售，境外客户遍布日本、美国等地区，2018年境外收入占比达到31%。客户覆盖BAT、微软、三星等知名机构。
- ✓ 本次拟公开发行人民币普通股1000万股，拟募资金7.25亿元，用于“天籁”自主研发产品扩建项目、一体化技术支撑平台建设项目、研发中心建设项目、业务管理平台建设项目及补充流动资金。

### ■中观思维：行业多领域融合应用，市场或将高速发展？

- ✓ 公司所在细分行业为“人工智能数据资源开发及服务行业”，属于人工智能行业，是国家支持的战略行业。当今的人工智能技术以机器学习、深度学习为核心，在智能语音、计算机视觉、自然语言等应用领域迅速发展，催生新的人工智能产业链与架构形态。
- ✓ 随着人工智能技术的不断发展和在各个行业的应用落地，客户对个性化定制以及更加全面高效的技术平台的需求攀升，推动人工智能市场规模快速增长。近年来国家对高科技人才的大力投入以及相关政策的不断推出实施，2018年国内市场规模为415.5亿元(+75%)，中国国际人工智能人才投入总量8.9%位列第二。

### ■竞争格局：国内暂无龙头企业，公司有望赶超国外强劲对手 Appen

- ✓ 人工智能数据资源开发及服务行业中的国内企业主要有除海天瑞声、慧听科技、标贝科技。行业内企业提供的人工智能数据资源产品及服务在技术领域、应用场景方面各有差异。慧听科技、标贝科技均为非上市公司。目前尚无业务一致、具有业务较强可比性的国内可比上市公司。
- ✓ 近些年海天瑞声快速发展，有望凭借自身较好盈利能力，继续缩短与国外竞争对手 Appen 的差距，公司有望在未来把握住机遇，成为业内龙头。
- ✓ 营业收入和净利润方面海天目前的体量较小，营收仅为 Appen 的 1/9。但是海天 34.85%利润率是 Appen 11.45% 利润率的近 3 倍。这充分体现出海天瑞声的较好盈利能力，且随着人工智能市场规模的不断扩张，同 Appen 在规模上的差距会进一步减少。

### ■风险提示：上游竞争加剧的风险，下游需求无法及时满足的风险

诸海滨

分析师

SAC 执业证书编号：S1450511020005

zhuhb@essence.com.cn

021-35082086

### 相关报告

从问询函看行业技术路径—

—科创板申报企业首轮问询2019-05-05  
点评

第一物业：绿色科技物业龙头，业绩稳定可持续增长 2019-05-05

赛普健身：健身教练培训龙头企业，经营成果保持高增2019-05-05  
长

广州塔：打造地标品牌 IP，首创高塔预约分时段观光模式 2019-05-05

新三板日报（新三板年报季即将收官，齐鲁银行利润超2019-04-30 20亿）

## 内容目录

1. 写在前面：人工智能数据资源服务商，海天瑞声有望再上一层楼？ .....	4
2. 内部挖掘：提供数据资源基础产品，海天瑞声优势何在？ .....	5
2.1. 纵观公司整体：全球领先的 AI 数据资源服务商，产品线已含 130+语言 .....	5
2.1.1. 主营业务：数字资源产品+服务，位居中国智能语音市场前十 .....	5
2.1.2. 股权结构：股权相对集中，创始人合计控股 45.41% .....	6
2.1.3. 人员构成：董监高几乎均为名校出身，近七成员工为研发技术人员 .....	6
2.2. 探究产品领域：主营业务三大板块，涵盖 AI 三大核心领域 .....	8
2.2.1. 核心产品：重点渗透智能语音领域，数据资源定制服务营收占比超 60% .....	8
2.2.2. 应用领域：智能语音领域收入占比超 80%，计算机视觉项目毛利率较高 .....	11
2.2.3. 经营模式：业务核心在研发产品与提供服务上，非核心采用对外承包模式 .....	13
2.3. 剖析竞争优势：人才优势奠定技术优势，BAT 为其大客户 .....	14
2.4. 解析经营业绩：净利润年复合增长率 155%，期间费用率逐年递减 .....	17
2.4.1. 营收情况：营业收入稳定增长，近 30%源于境外 .....	17
2.4.2. 期间费用：近三年费用率逐年下降，汇兑损益+利息收入低令财务费用大涨 .....	18
2.5. 分析募集方案：拟募集 7.25 亿元，用于平台建设与产品研发 .....	19
3. 中观思维：行业多领域融合应用，市场或将高速发展？ .....	20
3.1. 行业情况：技术多领域融合衍生，逐渐形成完整产业链 .....	20
3.1.1. 行业特点：核心业态的技术突破，是推动产业升级的核心驱动力 .....	20
3.1.2. 当下现状：深度学习、丰富数据、强大算力，共同助推行业高速发展 .....	21
3.2. 发展前景：国内 AI 市场增速将达到 75%，国际人才投放量已位居第二 .....	21
3.2.1. 市场规模：全球人工智能产业规模将达 6800 亿元，中国或成新增长极 .....	21
3.2.2. 人才投入：国家大力投入高科技人才，中国占世界总量 8.9%居于第二 .....	22
3.2.3. 政策发布：政策利好，行业有望成为国家支柱产业 .....	23
3.3. 竞争格局：未来有望成为国内龙头企业，赶超国外强劲对手 Appen .....	24
3.3.1. 国内主要企业：没有龙头企业，竞争壁垒较强 .....	24
3.3.2. 境外强劲对手：体量差距较大，但盈利能力与数据库种类并不逊色 .....	25
4. 潜在风险：上游竞争加剧，下游需求恐无法及时满足 .....	26

## 图表目录

图 1：2015-2020 年全球人工智能市场规模及预测 .....	4
图 2：公司业务产品线 .....	5
图 3：人工智能技术架构示意 .....	5
图 4：海天瑞声股权结构 .....	6
图 5：2018 年公司员工职能构成情况 .....	7
图 6：2018 年公司员工学历构成情况 .....	7
图 7：2016-2018 年公司各业务营业收入占比情况 .....	9
图 8：2016-2018 年公司各业务毛利占比情况 .....	9
图 9：2018 年数据资源定制服务细分领域收入情况（万元） .....	9
图 10：智能语音领域产量、销量、毛利率情况 .....	9
图 11：2018 年数据库产品细分领域收入情况（万元） .....	10
图 12：智能语音数据库产品产量、销量、毛利率 .....	10
图 13：主要产品及服务在智能语音领域的应用情况 .....	11
图 14：主要产品及服务在计算机视觉领域的应用情况 .....	12

图 15: 海天瑞声主要产品及服务在自然语言领域的应用情况.....	12
图 16: 公司主要盈利模式.....	13
图 17: 数据资源开发模式.....	13
图 18: 海天瑞声人工智能数据资源开发一站式解决方案示意.....	15
图 19: 2016-2018 年平均年研发投入占收入比重.....	15
图 20: 2016-2018 年公司境外收入占比.....	16
图 21: 2018 年主要客户及营收情况.....	16
图 22: 2016-2018 年主营业务收入及增长率.....	17
图 23: 2016-2018 年归母公司净利润及增长率.....	17
图 24: 2016-2018 年净利润及增长率.....	17
图 25: 2016-2018 年公司毛利率.....	17
图 26: 2018 年公司按区域分类的主营业务收入情况 (万元) .....	18
图 27: 2016-2018 年销售费用及占营业收入比重.....	18
图 28: 2016-2018 年管理费用及占营业收入比重.....	18
图 29: 2016-2018 年研发费用及占营业收入比重.....	18
图 30: 2016-2018 年财务费用及占营业收入比重.....	18
图 31: 2016-2018 年期间费用及占营业收入比重.....	19
图 32: 人工智能产业链.....	20
图 33: 人工智能技术架构.....	20
图 34: 人工智能产业生态.....	20
图 35: 人工智能技术发展的三大核心要素.....	21
图 36: 2015-2020 年全球人工智能市场规模及预测.....	22
图 37: 2016-2018 年中国人工市场规模及增速.....	22
图 38: 2018 年各国国际人工投入总量.....	23
图 39: 世界各国人工智能战略与政策发布情况图.....	23
图 40: 行业未来发展趋势.....	24
图 41: 对手 Appen 公司的经营特点.....	25
表 1: 公司现任董事基本情况.....	7
表 2: 公司核心技术人员及管理人员介绍.....	8
表 3: 主要产品及服务按类型分布.....	11
表 4: 公司目标客户.....	14
表 5: 近年公司人工智能领域奖项.....	14
表 6: 公司拟投资项目.....	19
表 7: 行业内主要企业.....	24
表 8: 与 Appen 非财务指标对比.....	25
表 9: 2018 年与境外可比公司的营收、净利润对比.....	25





本号仅用于**招募付费会员**

添加请备注：**任职企业+职位**

**无入会需求勿扰**

## 1. 写在前面：人工智能数据资源服务商，海天瑞声有望再上一层楼？

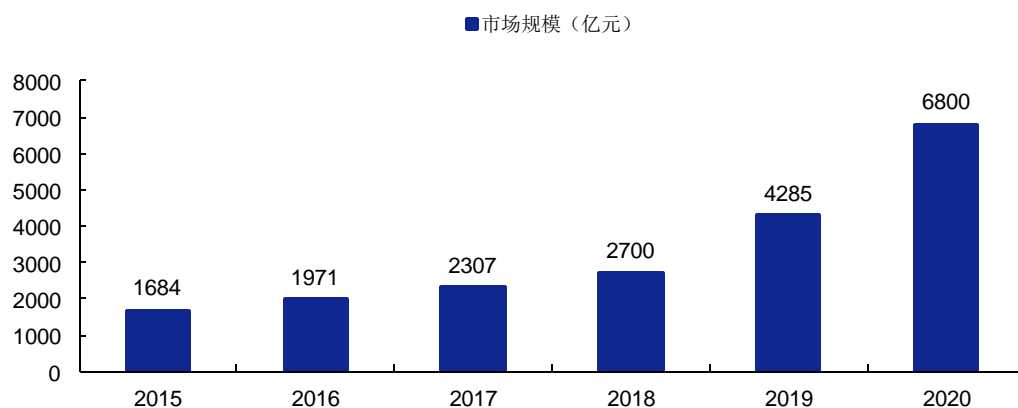
北京海天瑞声科技股份有限公司本次拟发行 A 股新股 1000 万股，拟募资金 7.25 亿元，募集资金主要针对平台的建设以及新产品的研发拓展，围绕主营业务进行。

北京海天瑞声科技股份有限公司成立于 2005 年 05 月 11 日，是全球领先的人工智能数据资源服务商。下有子瑞安、香港海天瑞声两家全资子公司，海天瑞声三河、荣成两家分公司。公司从事为人工智能全产业链中的企业及科研机构提供数据资源产品和服务，业务领域涵盖了智能语音、计算机视觉、自然语言等人工智能核心领域。海天瑞声连续多年荣获“国家高新技术企业”及“中关村高新技术企业”称号，2018 年获评为国家规划布局内重点软件企业、以及中关村高成长 Top100 企业。

当今的人工智能技术以机器学习、特别是深度学习为核心，在智能语音、计算机视觉、自然语言等应用领域迅速发展，催生新的人工智能产业链与架构形态，从而引发链式突破，为传统行业带来全面重构，推动经济社会各领域从数字化、网络化向智能化加速跃升。

随着人工智能技术的不断发展和在各个行业的应用落地，客户对个性化定制以及加全面、高效的一体化人工智能数据处理技术支撑平台的需求攀升，推动人工智能市场规模快速增长。并且，近年来国家对人才的大力投入以及相关政策的不断推出实施，也凸显了人工智能行业在未来的巨大发展潜力。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。

图 1：2015-2020 年全球人工智能市场规模及预测



资料来源：招股说明书，安信证券研究中心

综上所述，我们认为，随着人工智能技术的飞快发展，对于人工智能数字资源的需求攀升，人工智能行业迎来新机遇。基于此，我们推出本专题，对人工智能行业的发展情况进行深入分析，做出如下思考：

- 内部挖掘：结合公司的主营业务、股权结构与人员构成，作为提供基础层产品与服务的海天瑞声，是到达如何市场份额前十的地位的？
- 中观思维：随着人工智能技术的应用领域拓宽，市场或将飞速发展？
- 从产品、财务、研发角度出发，对比可比软件、信息服务公司，海天瑞声究竟有何独特优势？

## 2. 内部挖掘：提供数据资源基础产品，海天瑞声优势何在？

### 2.1. 纵观公司整体：全球领先的 AI 数据资源服务商，产品线已含 130+语言

#### 2.1.1. 主营业务：数字资源产品+服务，位居中国智能语音市场前十

北京海天瑞声科技股份有限公司成立于 2005 年 05 月 11 日，是全球领先的人工智能数据资源服务商，致力于为人工智能全产业链中的企业及科研机构提供数据资源产品和服务，业务领域涵盖了智能语音、计算机视觉、自然语言等人工智能核心领域。海天瑞声连续多年荣获“国家高新技术企业”及“中关村高新技术企业”称号，2018 年获评为国家规划布局内重点软件企业、以及中关村高成长 Top100 企业

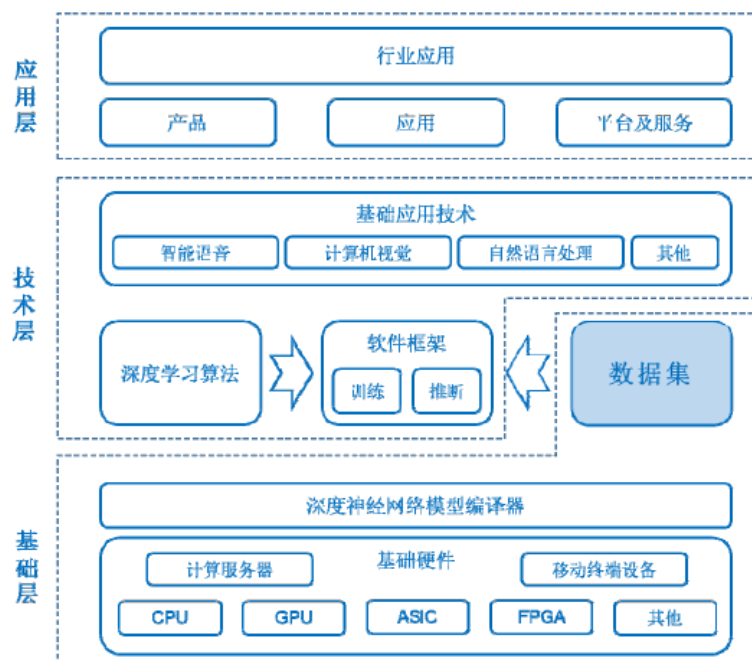
图 2：公司业务产品线



资料来源：公司官网，安信证券研究中心

公司已具备全球化的业务支持与交付能力，产品线已包含全球 130 余个主要语种及方言，是我国智能语音产业中市场份额前十名的企业，也是唯一入围的数据资源服务企业。目前公司客户累计数量近 400 家，基本覆盖了主要的大型科技公司、人工智能企业及科研机构。《中国语音产业联盟白皮书 2017-2018》

图 3：人工智能技术架构示意



资料来源：招股说明书，安信证券研究中心

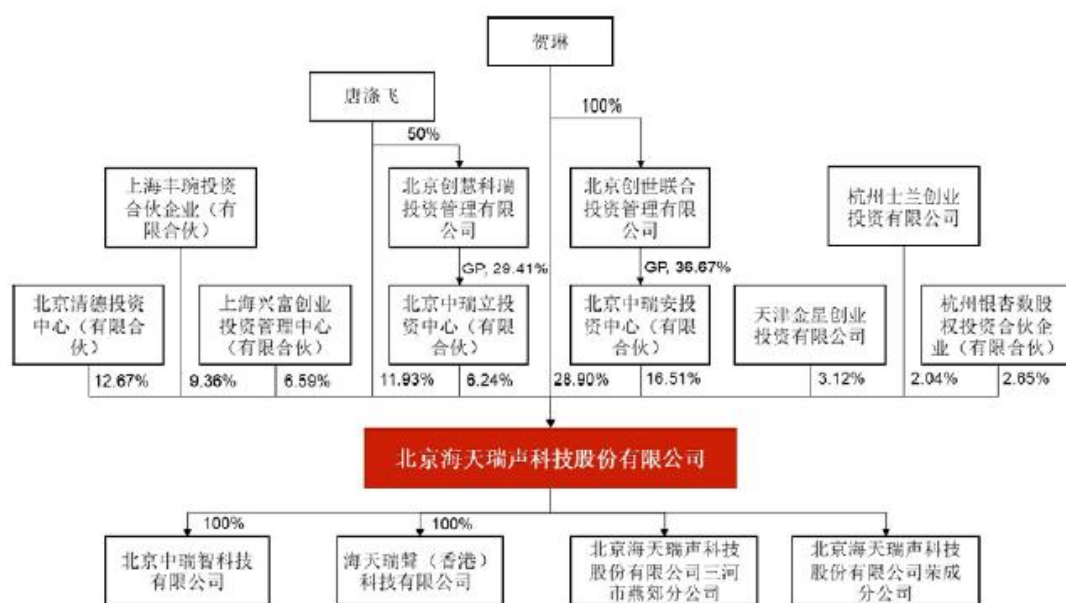
海天瑞声所提供的人工智能数据资源产品及服务属于人工智能产业链的基础层，是自主研发人工智能技术的企业与机构必需的基础生产要素，其数量多寡和质量高低将会直接影响到人工智能产业链内企业的研发周期、产品性能和可扩展性。

### 2.1.2. 股权结构：股权相对集中，创始人合计控股 45.41%

截至本招股说明书签署日，公司的股权结构如上图所示。公司共有 10 名股东，其中 2 名为自然人股东，2 名为员工持股平台，6 名为机构股东。经穿透合并计算，公司股东数量未超过 200 人。

公司创始人贺琳女士直接持有海天瑞声 866.97 万股股份（占发行前总股本的 28.90%），并通过控制中瑞安间接控制海天瑞声 495.41 万股股份（占发行前总股本的 16.51%），合计控制公司 45.41% 的股份，为公司实际控制人。

图 4：海天瑞声股权结构



资料来源：招股说明书，安信证券研究中心

### 2.1.3. 人员构成：董监高几乎均为名校出身，近七成员工为研发技术人员

公司高管及核心人员大多毕业于清华、北大、中科大、复旦等院校，大部分曾在在微软、英特尔、IBM、中科院等业内领先成熟企业与研究机构担任人工智能领域技术研发与管理的领导职务。

现公司共有 7 名董事，其中 3 名为独立董事。公司董事由股东大会选举产生，任期 3 年，任期届满可连选连任。监事会现有 3 人，其中 2 名为外部监事，人资行政经理任员工监事，3 人分别毕业于复旦、浙大、中南财大。此外，海天瑞声另有 3 名高级管理人员和 2 名核心技术人员，是清华、东大、复旦、北航、北邮等名校的高材生。

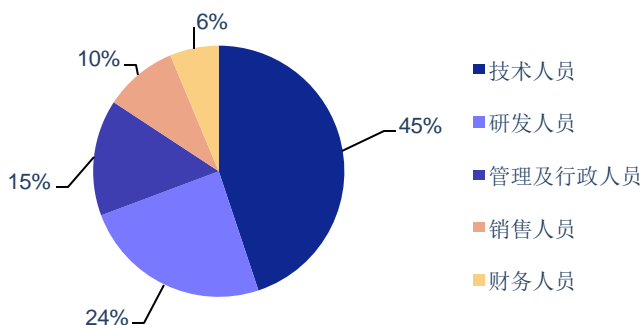
表 1：公司现任董事基本情况

姓名	在本公司任职	简历
贺琳	董事长	1968 年 7 月出生，中国国籍，毕业于北京大学。1987 年至 2017 年 7 月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，完成《计算机汉语口语输出系统的设计与实现》等论文十余篇，参与编写《汉语语音合成-原理和技术》著作
唐涤飞	董事、总经理	男，1970 年 12 月出生，中国国籍，有美国永久居留权，硕士，毕业于中国科学技术大学，获得通信与电子系统专业工学硕士学位。1997 年至 2010 年，历任微软中国研发中心中文技术部项目经理、微软亚洲工程院资深项目经理；2010 年至 2013 年，任阿里巴巴云计算公司数据应用与服务、无线平台的产品及项目总监。
志鹏	董事	女，1972 年 6 月出生，中国国籍，无境外永久居留权，硕士，于清华大学获得计算机科学与技术专业工学学士学位，于剑桥大学获得哲学硕士学位。1996 年至 2004 年任职于微软公司，2006 年至 2011 年任职于 TLcomCapital，2012 年至 2013 年，任职于昆山青石投资管理有限公司，2013 年至今任清控资产管理有限公司专务副总裁
吕思瑶	董事、副总经理、财务负责人、董事会秘书	女，1983 年 8 月出生，中国国籍，无境外永久居留权，硕士，毕业于北京大学光华管理学院，获得会计学专业学士学位，中国人民大学财政与金融学院金融学硕士学位，中国注册会计师、国际注册内审师。2005 年至 2011 年，任普华永道中天会计师事务所审计师、高级审计师、经理；2011 年至 2012 年，任英特尔（中国）内部控制经理；2012 年至 2014 年，任北京新合作连锁超市有限公司财务总监；2015 年至 2016 年，任英特尔（中国）信用管理高级经理。
仪晓辉	独立董事	男，1968 年 10 月出生，中国国籍，无境外永久居留权，硕士，毕业于北京大学无线电子学系信息中心研究院，获得工学硕士学位。2004 年至 2008 年，历任联想集团中国区 Thinkpad 业务部总经理、中国笔记本业务总经理、联想集团副总裁、大中华及俄罗斯区商用事业部总经理以及联想新兴市场集团下属香港、台湾和韩国区副总裁；2012 年至 2014 年，任惠普公司全球高级副总裁、打印与信息产品集团中国区总裁。2017 年 11 月至今任公司独立董事。
张东晖	独立董事	男，1967 年 7 月出生，中国国籍，无境外永久居留权，学士，毕业于清华大学，获得工学学士学位。2014 年至 2015 年代表阿里巴巴任恒生电子董事；2015 年 11 月至今，任金华几何科技有限公司监事，为该公司创始人之一。2018 年 11 月至今任公司独立董事。
杨涛	独立董事	男，1975 年 2 月出生，中国国籍，无境外永久居留权，学士，毕业于首都经济贸易大学，获得审计学专业学士学位，中国注册会计师。现任优途天下（北京）房车有限公司副总经理。2017 年 11 月至今任公司独立董事。

资料来源：招股说明书，安信证券研究中心

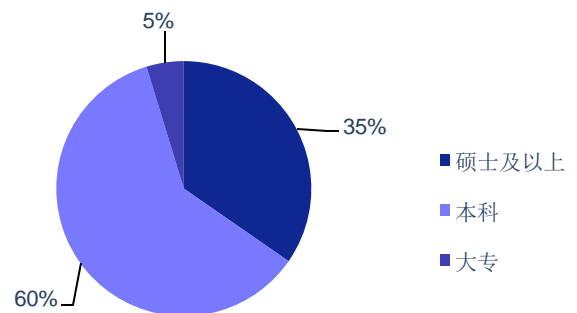
截止至 2018 年 12 月 31 日，公司（含分公司及子公司）共有员工 127 人，其中研发人员 31 人，占员工总数的 24.41%。技术人员 57 人，占员工总数的 45%。另外，公司硕士及以上学历的员工为 44 人，占员工总数的 34.65%，本科学历 77 人，占比为 60.63%，合计占比超过 95%。

图 5：2018 年公司员工职能构成情况



资料来源：招股说明书，安信证券研究中心

图 6：2018 年公司员工学历构成情况



资料来源：招股说明书，安信证券研究中心

公司核心技术人员均为计算机、信息工程、自动化等领域的专业人士，并在人工智能行业从业多年，具备良好的产业背景和丰富的研发设计经验，保障了公司的研发与创新能力。

**表 2：公司核心技术人员及管理人员介绍**

姓名	职务	专业背景和具体贡献
唐涤飞	董事、总经理	拥有 20 余年人工智能和软件行业的研发和管理经验，领导 跨国团队在语音语义、数据应用和移动平台等技术、产品、服务取得重要突破，是 6 项美国、中国专利的发明人。加入海天瑞声后，带领团队打造了以基础算法研究、技术支撑平台，算法技术与工具平台相结合验证数据库效果为核心的公司技术战略布局，确立了公司数据库产品在业内的竞争优势
李科	副总经理	主导研发了公司一体化人工智能数据处理技术支撑平台。此平台拥有超过 300 多个技术工具或模块，实现了数据开发的智能化和高效率。在此平台上，开发了 1,600 余个人工智能数据库，涵盖智能语音、计算机视觉、自然语言等 领域，支持 130 多个语种或方言，处理了超过 80 万小时语音时长，3,000 多万张图片、2,000 多万行文本数据。负责公司知识产权保护工作，获得 50 余项软件著作权。
郝玉峰	技术总监、首席科学家	负责通过语音识别、语音合成、自然语言处理、图像处理等 AI 技术手段，提高公司数据库制作过程中的工作效率，提升数据库质量，主持或参与了 30 余项软件著作权的研发工作，任职期间设计研发多语种语料清洗和语料设计流程与相关算法；基于 VAD 技术和 SNR 计算技术的录音质量监控技术、基于 ASIO 的多通道语音采集软件、以及基于 ASR 的语音自动切分系技术等信号处理技术。以上算法和技术被广泛地应用于公司数据设计、开发和质量控制流程，构成了公司一体化人工智能数据处理技术支撑平台的主要技术来源。是中国中文信息学会语音信息专业委员会委员、中国人民大学信息学院工程硕士导师。
黄宇凯	技术总监	重构一体化人工智能数据处理技术支撑平台的软件架构；组织开发并改进数据处理各环节所需工具以高效支撑智能语音、计算机视觉、自然语言处理等公司业务；遴选公司人工智能基础算法的研究成果并转化为软件系统使之应用于数据开发业务。
曹琼	计算机视觉部门经理	设计并实现使用计算机视觉数据自动检测算法与人工标注结合的人机交互数据生产方案，主导开发了计算机视觉数据处理通用平台，并在实际项目中得到应用，提高了生产效率，保障了数据的隐私安全性。其成果被应用于公司计算机视觉业务中，成为公司竞争优势之一。是“目标检测方法及设备”、“图像标注方法与装置”等 14 项中国、美国及日本专利发明人。
邵志明	高级研发工程师	从事语音信号处理、语音合成、说话人识别、语种识别、自然语言处理等技术研究，其主导研发的语音多通道对齐技术和 ASR 唤醒词数据库自动转写系统，在相应的业务上节省了三分之一的人工成本。

资料来源：招股说明书，安信证券研究中心

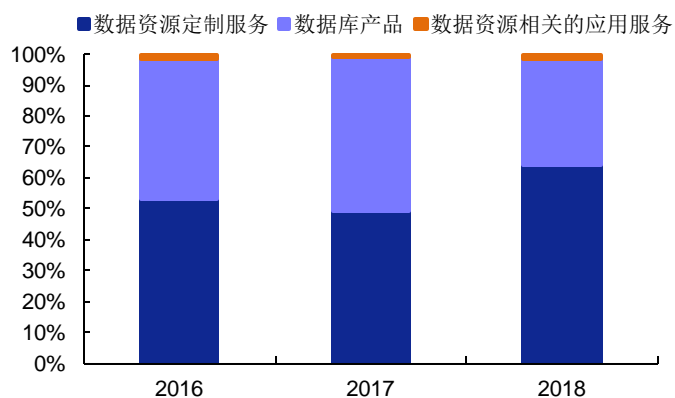
## 2.2. 探究产品领域：主营业务三大板块，涵盖 AI 三大核心领域

### 2.2.1. 核心产品：重点渗透智能语音领域，数据资源定制服务营收占比超 60%

公司属于“新一代信息技术领域”中的“人工智能”领域，主营业务包括数据资源定制服务、数据库产品及数据资源相关的应用服务三大板块。主要产品及服务覆盖了智能语音、计算机视觉及自然语言处理等各领域，广泛应用于人工智能算法模型的开发、训练、拓展、优化等环节。其中，数字资源相关的应用服务及数字资源定制服务占据了公司绝大部分的营业收入，2018 年数字资源定制服务的营收占比超 60%。

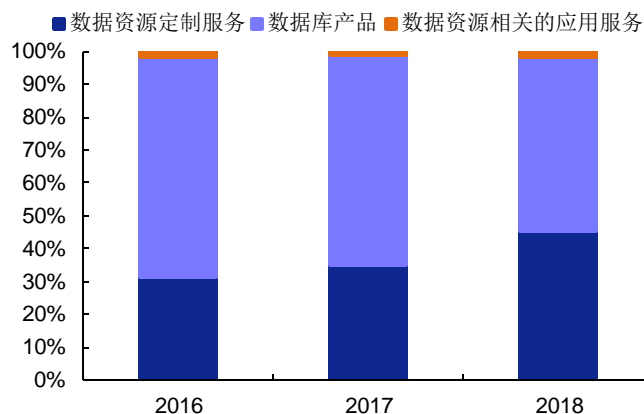
公司主营业务的毛利主要来源于数据库产品和数据资源定制服务，这两类业务均表现出较强的盈利能力，构成了公司持续盈利能力的基础。公司数据库产品无相应营业成本发生，毛利率 100%，因此在该类业务收入规模小于数据资源定制服务的情况下，贡献毛利高于数据资源定制服务。

图 7：2016-2018 年公司各业务营业收入占比情况



资料来源：招股说明书，安信证券研究中心

图 8：2016-2018 年公司各业务毛利占比情况



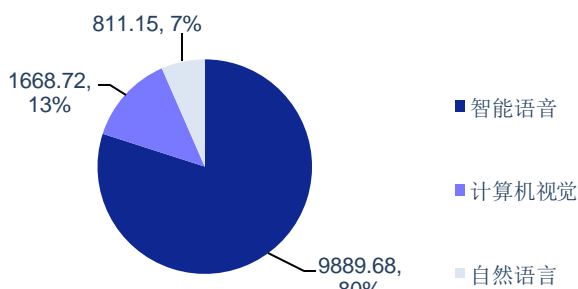
资料来源：招股说明书，安信证券研究中心

➤ **数据资源定制服务：以销定产，营收同比增长 96.43%，智能语音领域为最主要来源**  
报告期内，公司数据资源定制服务业务规模持续增长，2016-2018 年收入规模分别为 4471.89 万元、6297.23 万元和 12369.55 万元，三年复合增长 66.31%；近年来收入占比也基本呈现上升趋势，2016 年、2017 年及 2018 年占营业收入比重分别为 53.09%、52.89% 及 64.20%。

在智能语音、计算机视觉及自然语言三大领域中，智能语音领域为公司数据资源定制服务最主要的来源，2016 年、2017 年及 2018 年占数据资源定制服务。整体收入比重分别达到 83.97%、77.72% 及 79.95%。

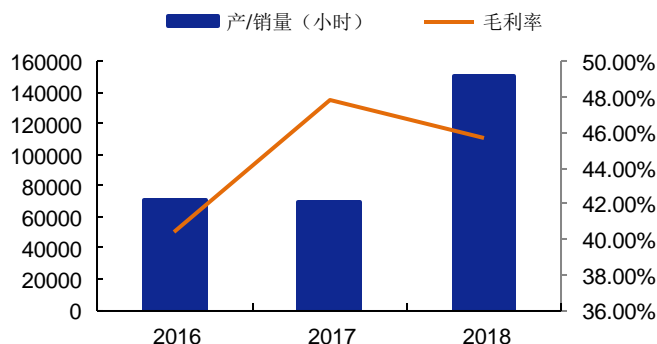
数据资源定制服务是公司根据客户需求提供的定制数据库开发服务及数据处理服务，属于“以销定产”模式。2018 年，智能语音类数据资源定制服务销量大幅增长，主要原因是阿里巴巴、腾讯等主要客户对智能语音数据资源需求的增长。同时平均价格有所下降，是由于 2018 年以来中文人工智能领域的快速发展导致中文业务大量涌现，占比显著提升，但相对单价较低所致。

图 9：2018 年数据资源定制服务细分领域收入情况（万元）



资料来源：招股说明书，安信证券研究中心

图 10：智能语音领域产量、销量、毛利率情况



资料来源：招股说明书，安信证券研究中心

➤ **数据库产品：营收增速放缓，毛利率三年维持 100%**

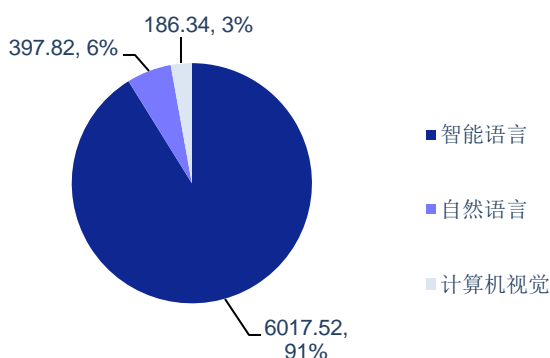
报告期内，公司数据库产品业务收入稳步上升，产品收入主要由智能语音数据库产品贡献，

2016 年、2017 年及 2018 年收入规模分别达到 3,826.77 万元、5,489.31 万元及 6,601.67 万元，占营业收入比重分别为 45.43%、46.10%及 34.27%，2017 年和 2018 年分别同比增长 43.45%和 20.26%。

数据库产品的收入规模在 2016 年和 2017 年与数据资源定制服务接近，但在 2018 年增速放缓，主要是由于部分公司客户随着其人工智能业务的深入发展，数据资源定制化服务需求增长迅速，但在未来仍将作为公司主营业务的重要组成部分之一，与数据资源定制服务共同构成公司完整的数据资源服务体系。

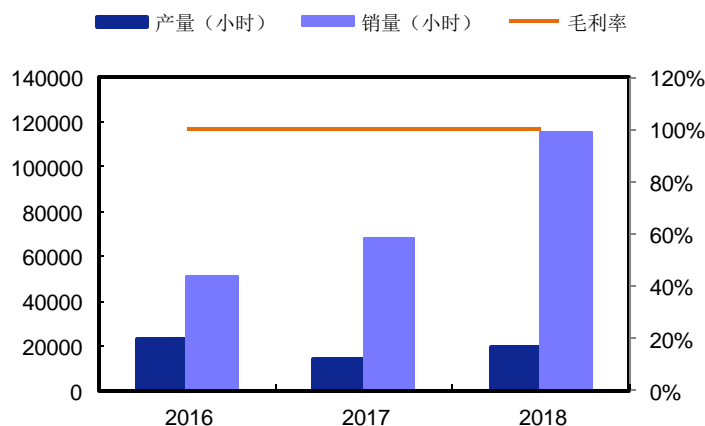
在公司的数据库产品中，智能语音领域为收入贡献最多的数据库产品领 2016 年、2017 年及 2018 年占数据库产品销售收入比重分别达到 88.75%、89.65%及 91.15%。销量与价格的变动原因同智能语音类数据资源定制服务的变化原因一致。

图 11：2018 年数据库产品细分领域收入情况（万元）



资料来源：招股说明书，安信证券研究中心

图 12：智能语音数据库产品产量、销量、毛利率



资料来源：招股说明书，安信证券研究中心

一个数据库产品研发完成后，公司可多次销售其使用授权许可，因此上表数据库产品的销量大于产量；数据库产品所发生的支出在当期全部计入研发费用，无相应的营业成本，毛利率为 100%

➤ 数据资源相关的应用服务：营收占比不足 2%，无显著影响

数据资源相关的应用服务是公司基于自身数据资源优势所提供的增值性服务和延伸服务，收入规模较小，2018 年取得营业收入 294.55 万元，占比 1.53%，对公司整体业务和营业收入不构成显著影响。

2.2.2. 应用领域：智能语音领域收入占比超 80%，计算机视觉项目毛利率较高

表 3：主要产品及服务按类型分布

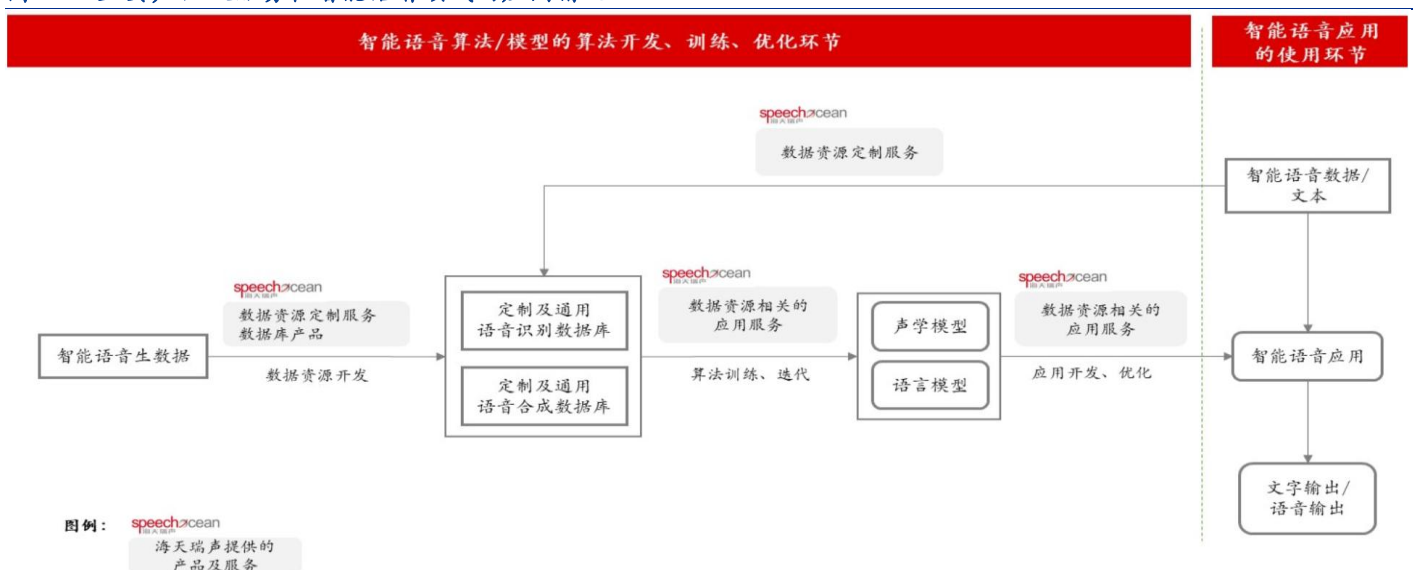
产品	应用领域	产品示例	内容简介	用途
数据资源定制产品	智能语音领域	/	公司根据客户相关要求,针对特定的应用领域,进行定制化的语音识别数据库、语音合成数据库、词典资源数据库、文本资源数据库、图像资源数据库和视频资源数据库等的设计及开发;或对客户提供的数据库资源进行处理,最终形成符合客户需求的数据资源产品。	用以训练、开发、拓展算法模型,并应用到相应的人工智能领域。
	计算机视觉领域			
	自然语言领域			
数据库产品	智能语音领域	语音识别数据库	公司根据对人工智能算法模型应用领域、行业发展趋势、市场需求等的评估和研判,设计开发的由不同内容和结构的智能语音、计算机视觉、自然语言等领域的数据资源组成的通用型数据库产品,公司拥有自主知识产权。 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。	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		
	计算机视觉领域	图像资源数据库		
		视频资源数据库		
	自然语言领域	文本资源数据库		
		词典资源数据库		
数据资源相关的应用服务	/	/	协助客户实现人工智能算法模型的识别率提升、语言种类拓展和垂直应用领域拓展等,助力人工智能技术及应用的设计、开发和领域拓展。	

资料来源：招股说明书，安信证券研究中心

➤ 智能语音：数据收集+处理，多项技术拥有计算机软件著作权

智能语音产品领域主要包括语音识别、语音合成等业务方向。语音识别指公司根据客户需求，将音频数据及文本数据将结构化组合，形成可用于进行语音识别相关人工智能算法或模型的开发、训练的数据资源产品。语音合成指公司根据客户需求完成语音合成数据资源设计、数据采集、数据处理等工作，终形成可用于语音合成相关人工智能算法或模型开发、训练的语音合成数据资源。在智能语音领域，公司拥有的多项核心技术：多通道录音技术、移动设备上的语音采集技术、基于 C/S 架构的大规模语音处理平台、基于 C/S 架构的音素标注技术等已累计取得的 18 项计算机软件著作权。

图 13：主要产品及服务在智能语音领域的应用情况

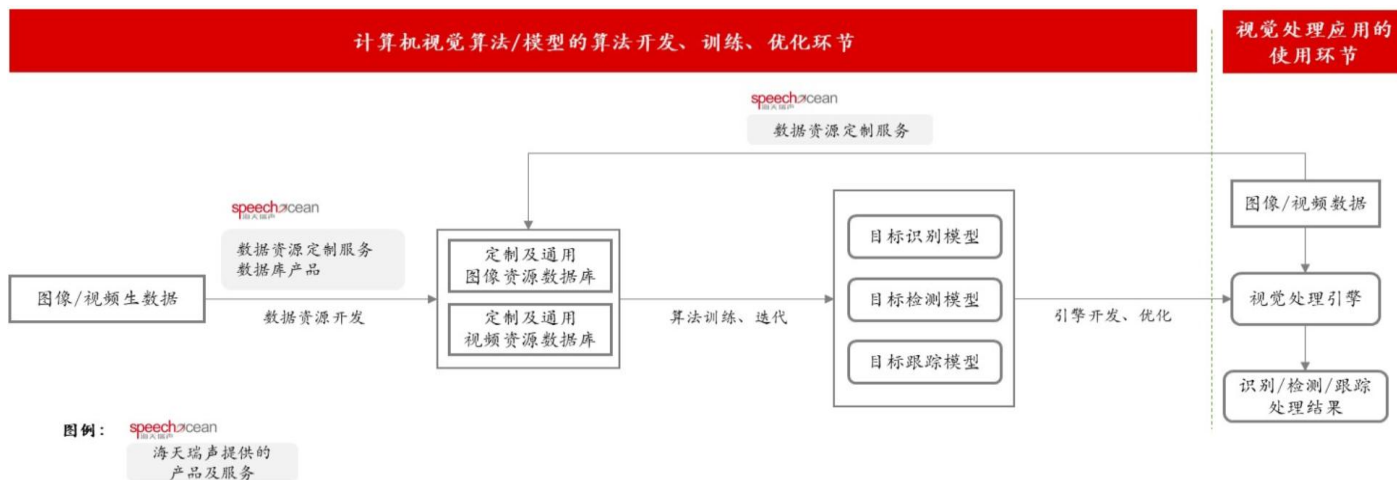


资料来源：招股说明书，安信证券研究中心

➤ **计算机视觉：视觉资源的标注与分割，产品毛利率较高**

计算机视觉是指公司根据客户需求进行计算机视觉数据资源采集，并对图片、视频等计算机视觉数据进行客观要素标注、像素分割、主观要素标注、特定需求标注等处理，形成可用于计算机视觉领域人工智能算法或模型的开发、训练的数据资源产品。公司在计算机视觉领域拥有人脸检测和识别、物体识别、光学字符识别、场景分割等多项核心技术，利用专门的计算机视觉数据处理通用平台，提高了生产效率，保障了数据的隐私安全性。由于计算机视觉中涉及到像素分割、内容理解相关的主观标注等工作，相对复杂度较高，所需技术处理环节较多，因此产品毛利率较高。

图 14：主要产品及服务在计算机视觉领域的应用情况



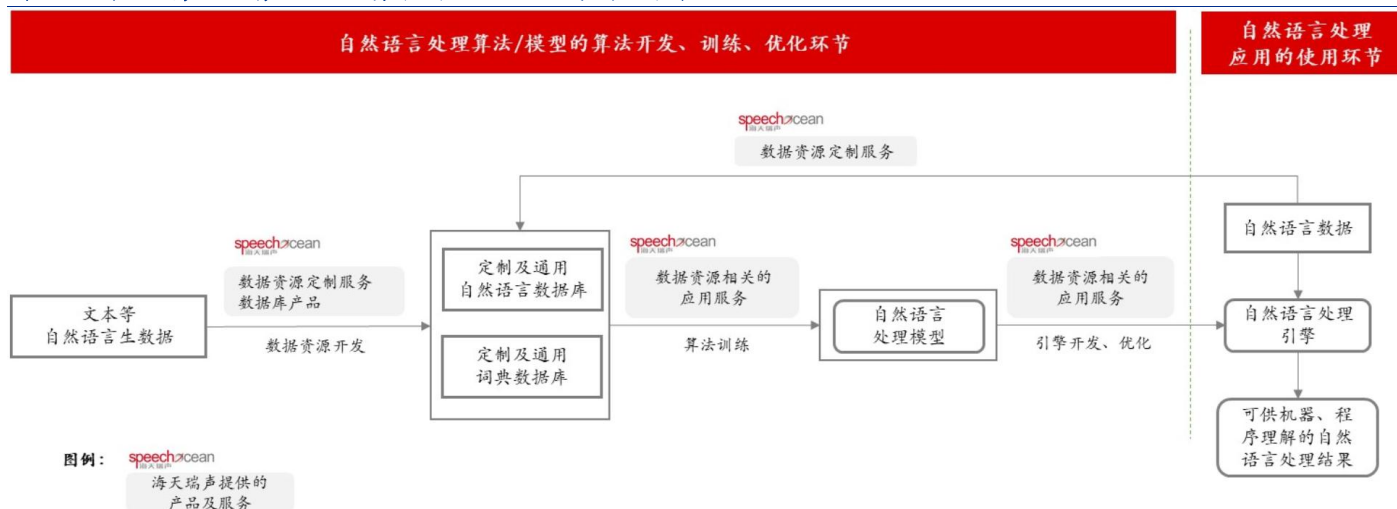
资料来源：招股说明书，安信证券研究中心

➤ **自然语言：130 余个语言数据库，拥有丰富资源与经验**

自然语言是指公司为客户提供一体化自然语言数据资源定制服务，根据客户需求进行自然语言数据资源采集，并对相关自然语言数据进行相关处理服务，形成可用于自然语言处理领域人工智能算法或模型的开发、训练的数据资源产品。

公司可提供 130 余个语种和方言的发音词典，并且积累了丰富的语言学家团队资源，成员遍布世界各地的学校和研究机构，在该领域具备丰富经验。公司拥有多项自然语言相关技术，其中多语言分布式文本处理技术已取得一项计算机软件著作权。

图 15：海天瑞声主要产品及服务在自然语言领域的应用情况



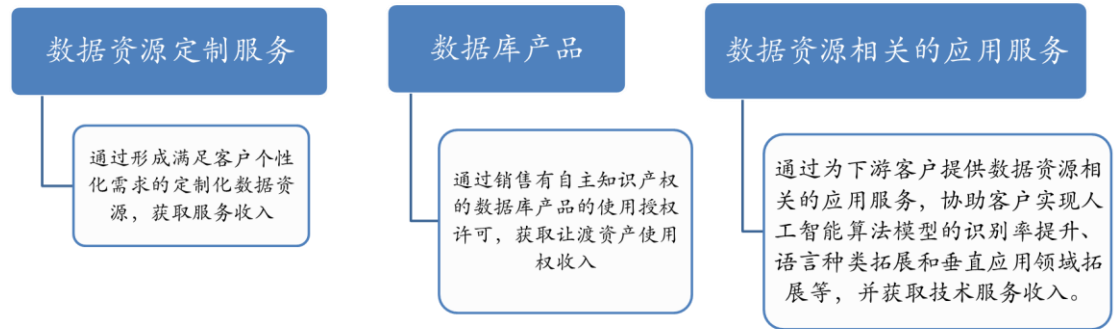
资料来源：招股说明书，安信证券研究中心

2.2.3. 经营模式：业务核心在研发产品与提供服务上，非核心采用对外承包模式

➤ 盈利模式：收取服务与授权使用费

公司主要通过向人工智能产业链上的各类机构提供数据资源定制服务、数据库产品及数据资源相关的应用服务，并收取相应的服务费或数据库授权使用费。

图 16：公司主要盈利模式

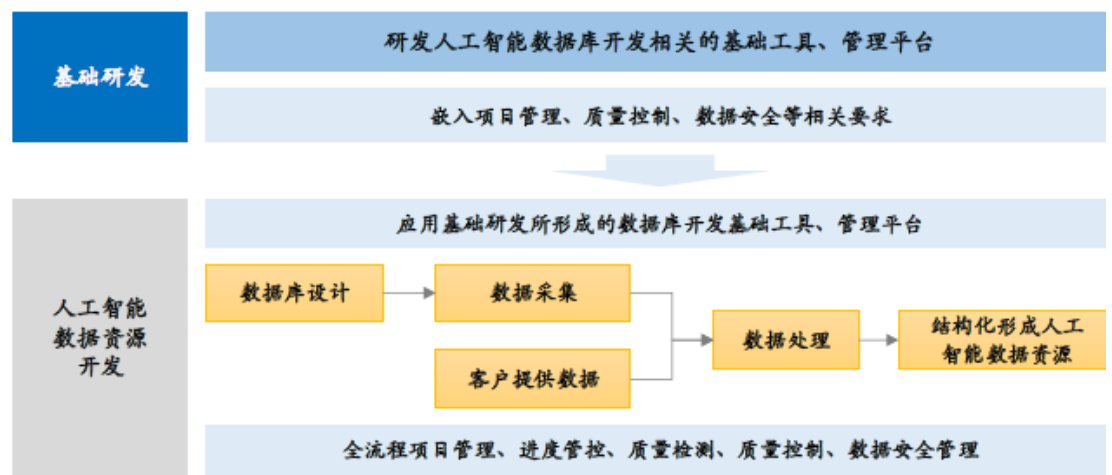


资料来源：招股说明书，安信证券研究中心

➤ 生产或服务模式：基础开发+数据资源开发

公司为客户提供的产品或服务价值均体现为人工智能算法模型开发、训练、拓展、优化所需的数据资源。公司人工智能数据资源的开发模式，主要涉及基础研发、数据库设计、数据采集、数据处理、质量控制、数据安全等过程。

图 17：数据资源开发模式



资料来源：招股说明书，安信证券研究中心

在基础研发过程中，公司始终秉承基础研究与实际应用紧密结合的原则，持续研究、开发、完善人工智能数据资源开发所需的基础工具和管理平台；在人工智能数据资源开发环节，公司通过数据的采集与处理后形成结构化的数据资源。

最后，公司充分发挥自身在人工智能数据资源方面的优势，为下游客户提供数据资源相关的应用服务，协助客户实现人工智能算法模型的提升与拓展，助力人工智能技术及应用的设计、开发和领域拓展。

➤ **采购模式：主要对外采购**

在数据采集与处理过程中，由于生数据与标记为公司业务的非核心环节，公司主要向人力资源服务公司、咨询公司等进行对外采购模式，具备合理性。此外，在个别业务中，由于数据标记量较小，公司也存在向个人供应商采购生数据标记服务的情形。

➤ **销售模式：直销+顾问式服务方式**

公司主要的销售模式与行业普遍的销售模式相同，即采用直接对接并服务客户的直销模式进行营销。基于多年的行业经验和知识储备，公司主要以高品质的数据资源产品及服务吸引客户，并在不断服务的过程中提升服务价值和客户黏度。

**表 4：公司目标客户**

客户类型	具体公司
大型科技公司	阿里巴巴、腾讯、百度、微软、三星等
人工智能企业	科大讯飞、商汤科技、云知声、海康威视等
科研机构	中国科学院、清华大学等

资料来源：招股说明书，安信证券研究中心

## 2.3. 剖析竞争优势：人才优势奠定技术优势，BAT 为其大客户

在行业实践过程中，公司自主开发了一体化人工智能数据处理技术支撑平台，在基础研究、平台工具、数据资源开发三个维度均积累下核心技术。公司先后获得国家重点软件企业、国家高新技术企业、中关村高新技术企业等资质，并成为中国人工智能产业发展联盟理事单位、中国语言产业联盟理事会员单位、中关村高新技术企业协会理事单位。

**表 5：近年公司人工智能领域奖项**

时间	颁奖单位	奖项
2017 年	中国信息协会、信息化观察网	年度人工智能数据服务领域领军企业
2017 年	美国 IT 企业联盟杂志 APACCIO Outlook	亚太唯一一家人工智能基础数据供应商入选“年度 25 家最具推荐价值人工智能企业”
2018 年	中国语言产业联盟	智能语音创新大赛行业先锋奖
2018 年	中国信息协会、信息化和软件服务网	2018 年人工智能行业最佳产品

资料来源：招股说明书，安信证券研究中心

➤ **人才优势：核心技术人员占比 4.7%，技术背景行业经验过硬**

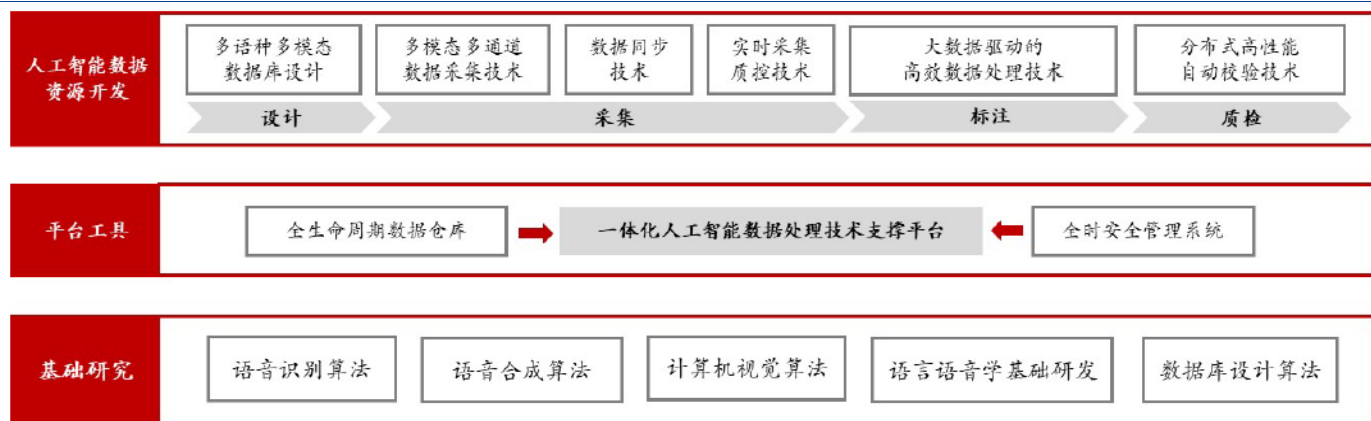
公司高管及核心人员大多毕业于清华、北大、中科大、复旦等一流院校，大部分曾在在微软、英特尔、IBM、中科院等业内领先成熟企业与研究机构担任人工智能领域技术研发与管理的领导职务。2018 年公司核心技术人员 6 名，占公司总员工总数的比例达 4.7% 这些核心技术人员均为计算机、信息工程、自动化等领域的专业人士，将积累的技术经验和管理经验应用于公司实践，并通过传帮带培养了一批专业而精干的中层技术团队，有效提升了企业的技术水平和规范化运作水平，实现了对客户的快速响应、高品质交付。

➤ **技术优势：积累了大量核心技术，有机搭建一站式解决方案**

公司重视技术研发的投入和创新，始终保持着对国际先进技术的追踪，并与清华大学、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。公司通过多年的业务实践和创新自主开发了一体化人工智能数据处理技术支撑平台，在基础研究、平台工具、数据资源开发三个维度下均积累核心技术，有机搭建起公司的人工智能数据资源开发一站式解决方案。其中较为突出的有多通道录音技术、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。针对其核心技术及产品，公司及其子公司已提交 8 项专利申请，其中 5 项已进入实质审查阶段；并已取得计算机软件著作权 92 项。

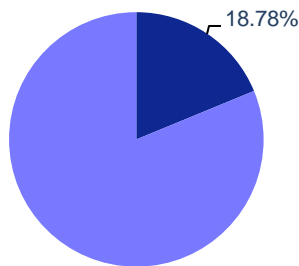
图 18：海天瑞声人工智能数据资源开发一站式解决方案示意



资料来源：招股说明书，安信证券研究中心

公司不断加大研发投入，2016-2018 年公司研发费用分别为 2174.92 万元、2527.99 万元和 2734.53 万元，呈逐年上升趋势，近三年平均年研发投入占收入比重为 18.78%。

图 19：2016-2018 年平均年研发投入占收入比重



资料来源：招股说明书，安信证券研究中心

截至目前公司已经拥有包含语料清洗、语音合成录音、文本正则化、移动设备手写体采集、命名实体检测与标注等近百个模块的一体化人工智能数据处理技术支撑平台，在数据资源的设计与开发、质量评测、结构设计与安全存储、训练效果指标、全球语言学研究等方面在国内外同行业中处于领先地位。

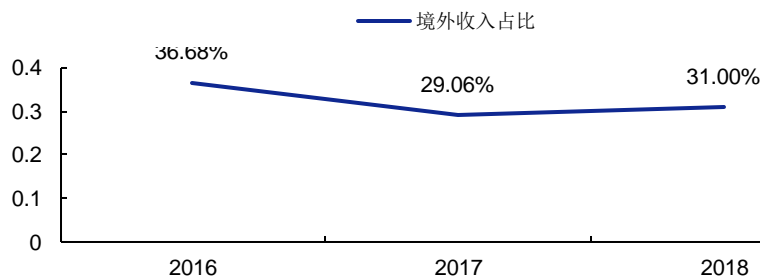
➤ **多场景及国际化优势：提供多场景、多语种的人工智能数据服务**

公司目前产品线已涵盖全球 130 余个主要语种及方言，并且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。目前，公司来自日本、韩国、美国等人工智能研发活跃区域的境外收入保持了持续稳定增长。

2016 年、2017 年及 2018 年，公司境外收入占比分别为 36.68%、29.06%和 31.00%。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。

图 20：2016-2018 年公司境外收入占比



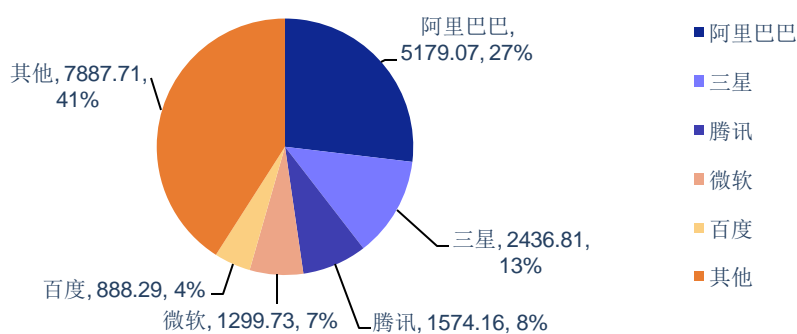
资料来源：招股说明书，安信证券研究中心

➤ **客户资源优势：优质资源客户遍布全球，覆盖 BAT、微软、三星等知名机构**

经过多年的发展，公司积累了深刻的行业洞见和突出的技术优势，与国内外诸多知名机构形成了长期稳定的合作关系，赢得了众多大型优质客户的认可，例如阿里巴巴、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。

公司一直秉持以客户利益为先的原则，利用自身的技术优势，将数据处理技术与人工智能技术进行有效整合，为客户购买数据资源服务和数据库产品提供了较为精准的效果预测和展示来帮助客户提高效率，节省成本。并且全部采用直接销售模式和顾问式的服务方式，与客户建立了密切的沟通、联系和互动。2016 年、2017 年及 2018 年，公司向前五名客户合计销售额占当期营业收入比重分别为 47.40%、51.22%及 59.06%，不存在对单一客户的依赖。

图 21：2018 年主要客户及营收情况



资料来源：招股说明书，安信证券研究中心

➤ **管理优势：全面化精细化的数据开发流程管理**

公司通过历史上在基础研究、平台工具、数据资源开发三个维度积累下的核心技术，自主研发了人工智能数据资源开发一体化解决方案和管理流程，充分提升了开发效率、服务质量及数据安全性，可保障企业提供高效率的数据资源定制服务、高质量的数据库产品、和高水准的数据资源应用相关服务。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。

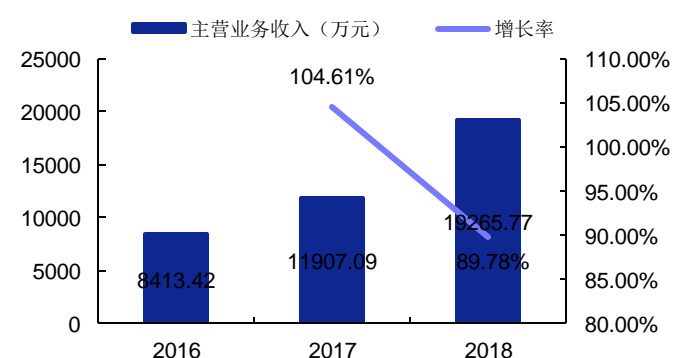
## 2.4. 解析经营业绩：净利润年复合增长率 155%，期间费用率逐年递减

海天瑞声的主营业务是为人工智能产业链上的各类机构提供数据资源定制服务、数据库产品和数据资源相关的应用服务。经过多年的经营，海天瑞声已形成了覆盖多领域、多场景、多语种的产品和服务体系。公司客户主要为大型科技公司、人工智能企业及科研机构，为国内人工智能领域的高速发展提供了重要支持与助力。

### 2.4.1. 营收情况：营业收入稳定增长，近 30%源于境外

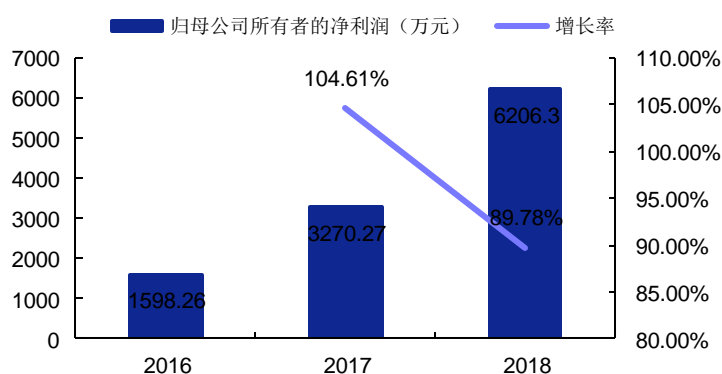
2016 年度、2017 年度及 2018 年度，公司的营业收入分别为 0.84 亿元、1.19 亿元及 1.92 亿元，公司营业收入逐年上升，2017 年、2018 年营业收入增幅分别为 104.62%、89.78%。2018 年归母净利润达 0.62 亿元，同比增长 89.78%；净利润为 6474.16 万元，增长率达到 96.61%。2018 年度，公司综合毛利率为 64.77%，相比上一年 72.26%，毛利率有所下降。

图 22：2016-2018 年主营业务收入及增长率



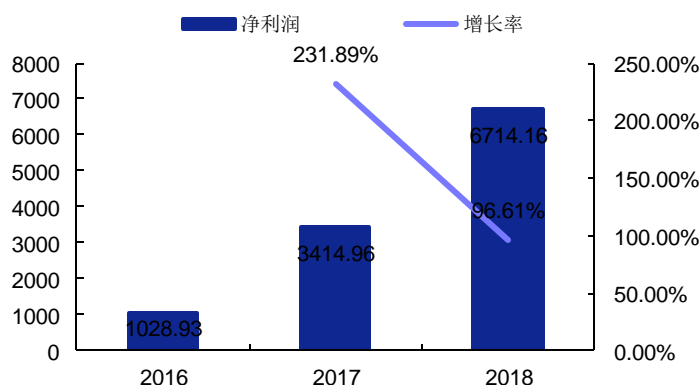
资料来源：招股说明书，安信证券研究中心

图 23：2016-2018 年归母公司净利润及增长率



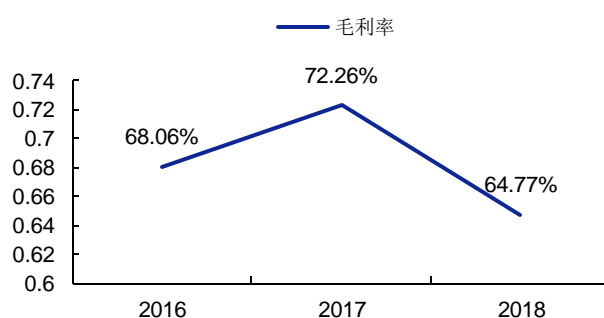
资料来源：招股说明书，安信证券研究中心

图 24：2016-2018 年净利润及增长率



资料来源：招股说明书，安信证券研究中心

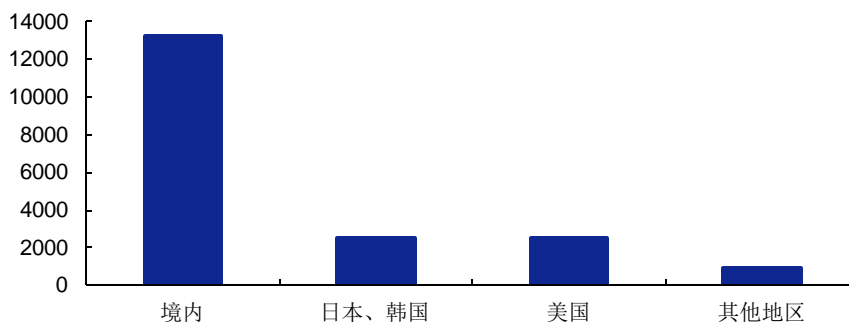
图 25：2016-2018 年公司毛利率



资料来源：招股说明书，安信证券研究中心

报告期内，公司主营业务收入大部分来源于境内，2018 年境内收入 1.33 亿元，占主营业务收入 69%。同时，公司也存在一定规模的境外收入，其中以日本、韩国、美国等地区为主。

图 26：2018 年公司按区域分类的主营业务收入情况（万元）

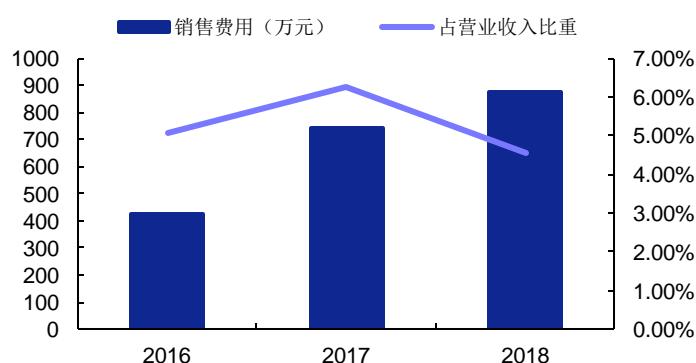


资料来源：招股说明书，安信证券研究中心

#### 2.4.2. 期间费用：近三年费用率逐年下降，汇兑损益+利息收入低令财务费用大涨

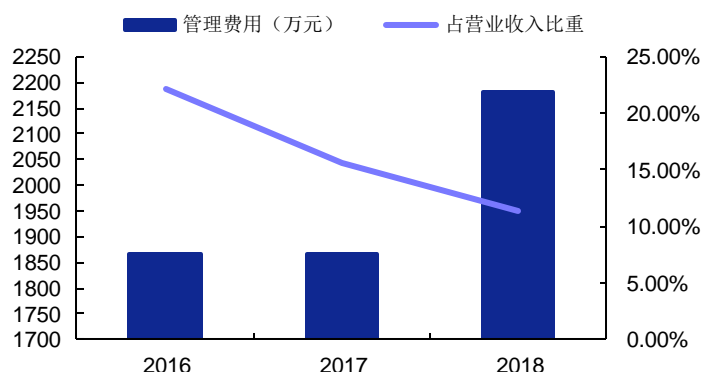
2016 年度、2017 年度及 2018 年度，公司期间费用合计金额分别为 4,441.01 万元、5,132.92 万元和 5,902.38 万元，占营业收入的比例分别为 52.73%、43.11%和 30.64%，呈现逐年下降趋势，主要是由于管理费用率和研发费用率逐年降低所致。

图 27：2016-2018 年销售费用及占营业收入比重



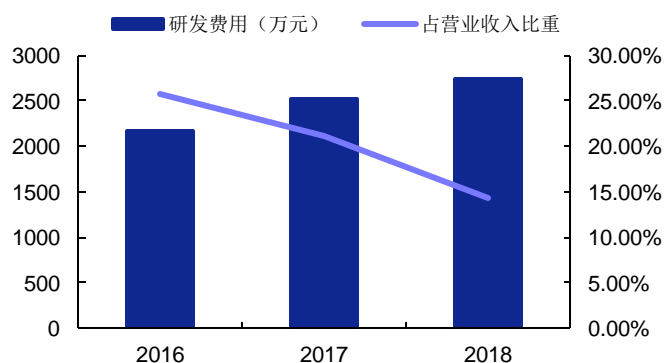
资料来源：招股说明书，安信证券研究中心

图 28：2016-2018 年管理费用及占营业收入比重



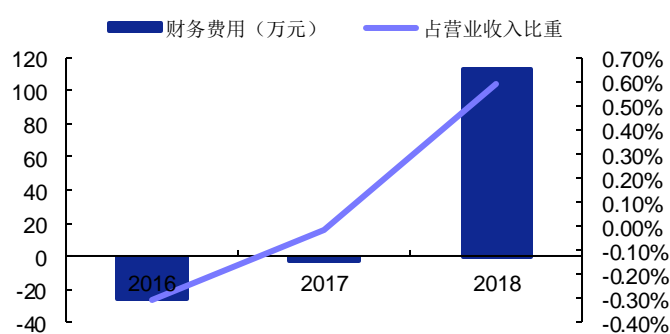
资料来源：招股说明书，安信证券研究中心

图 29：2016-2018 年研发费用及占营业收入比重



资料来源：招股说明书，安信证券研究中心

图 30：2016-2018 年财务费用及占营业收入比重



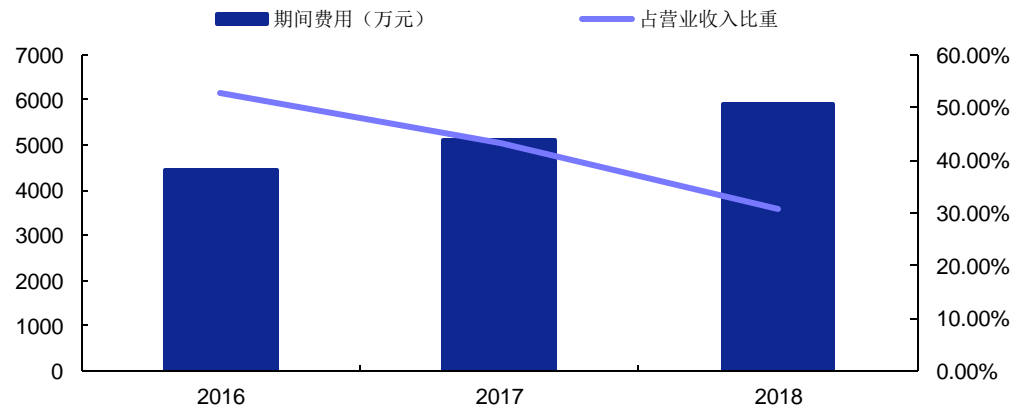
资料来源：招股说明书，安信证券研究中心

公司 2018 年期间费用相对较低，占营业收入的比例较小，主要由于管理费用具有显著的规模效应，并不会随着公司营业收入的快速增长而同比例增长，因此管理费用率显著下降。并

且，由于研发费用内研究数据库产品支出的降低以及日益逐渐凸显的技术优势，研发费用率逐渐下降。2018年财务费用大幅增长是由于汇兑损益较大，并且利息收入较低导致。

对于研发费用，尽管费用率下降，投入金额一直上涨。2016年至2018年，公司不断加大研发投入，累计研发投入达到7,437.45万元，平均年研发投入占收入比重为18.78%。支付研发人员的职工薪酬以及数据库产品研发生产过程采购非核心技术环节的生数据采集、标记服务所支付的数据服务费为研发费用的主要构成部分，2018年占比85.25%。

图 31：2016-2018 年期间费用及占营业收入比重



资料来源：招股说明书，安信证券研究中心

## 2.5. 分析募集方案：拟募集 7.25 亿元，用于平台建设与产品研发

本次拟公开公司人民币普通股 1,000 万股，拟募资金 7.25 亿元，用于“天籁”自主研发产品扩建项目、一体化技术支撑平台建设项目、研发中心建设项目、业务管理平台建设项目及补充流动资金。募集资金投资项目均围绕公司主营业务进行，符合公司的发展战略。

表 6：公司拟投资项目

项目名称	总投资金额 (万元)	拟投入募集资金 (万元)
“天籁”自主研发产品扩建项目	21928.14	21928.14
一体化技术支撑平台建设项目	21047.81	21047.81
研发中心建设项目	16238.6	16238.6
业务管理平台建设项目	3327.91	3327.91
补充流动资金	10000	10000

资料来源：招股说明书，安信证券研究中心

**“天籁”自主研发产品扩建项目：**计划新开发 137 个数据库，其中 68 个语音数据库、19 个计算机视觉数据库和 50 个词典数据库。新开发数据库的应用领域将涉及语音助手、智能家居、智能驾驶、智慧城市、智能硬件和数字娱乐等人工智能应用广泛、发展广阔的行业。

**一体化技术支撑平台建设项目：**在公司现有数据资源研发平台和工具的基础上，通过扩大研发团队，投入先进的开发设备和软件，建设测试实验室，升级数据处理工具及平台，重构一体化技术支撑平台，为公司的数据资源服务业务拓展提供支持。

**研发中心建设项目：**拟聚焦于语音识别、声纹识别、语种识别、语音合成、自然语言处理、计算机视觉识别、计算机视觉分割与处理 7 个方向的前沿技术，针对公司语音识别数据与服务、语音合成数据与服务、自然语言处理技术服务、计算机视觉数据与服务 4 个业务方向，完善公司产品从研发到服务于客户整个流程中的技术应用，保证未来公司的持续竞争力。

**业务管理平台建设项目：**拟投入技术开发团队和相应的软硬件设备，升级现有的业务管理平台，对项目管理的流程进行优化和固化，从而提升公司的项目管理能力，并完善业务流程中的各项环节，提升总体业务管理效率。同时，本项目将在平台中着重开发知识库管理模块，以期加强公司的知识复用能力。平台建成后，将大大提升公司业务团队的工作效率。

### 3. 中观思维：行业多领域融合应用，市场或将高速发展？

#### 3.1. 行业情况：技术多领域融合衍生，逐渐形成完整产业链

##### 3.1.1. 行业特点：核心业态的技术突破，是推动产业升级的核心驱动力

公司所在细分行业为“人工智能数据资源开发及服务行业”，属于人工智能行业，是国家支持的战略行业。人工智能是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

图 32：人工智能产业链

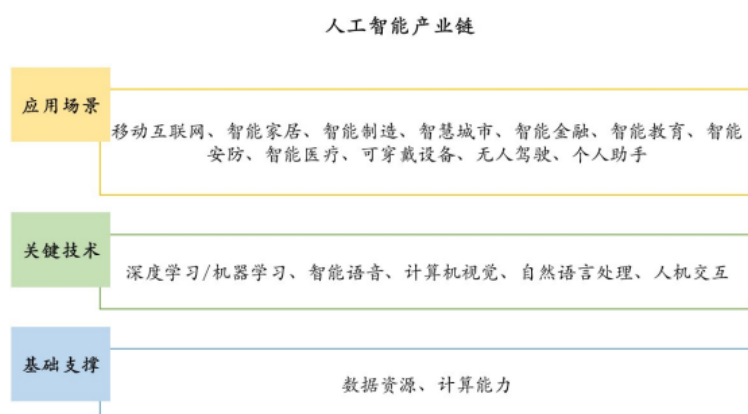
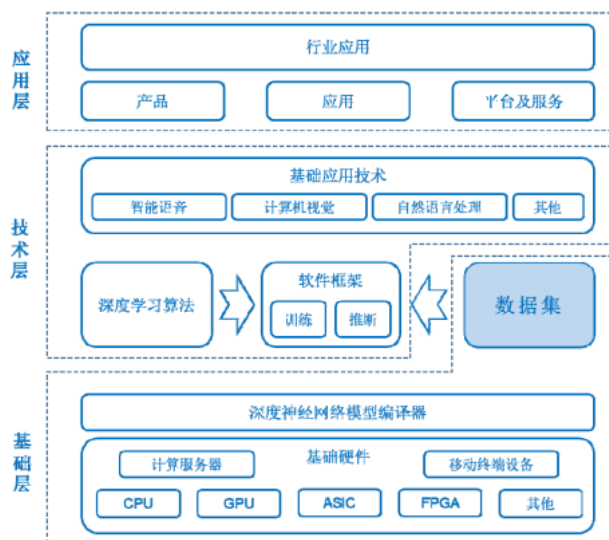


图 33：人工智能技术架构

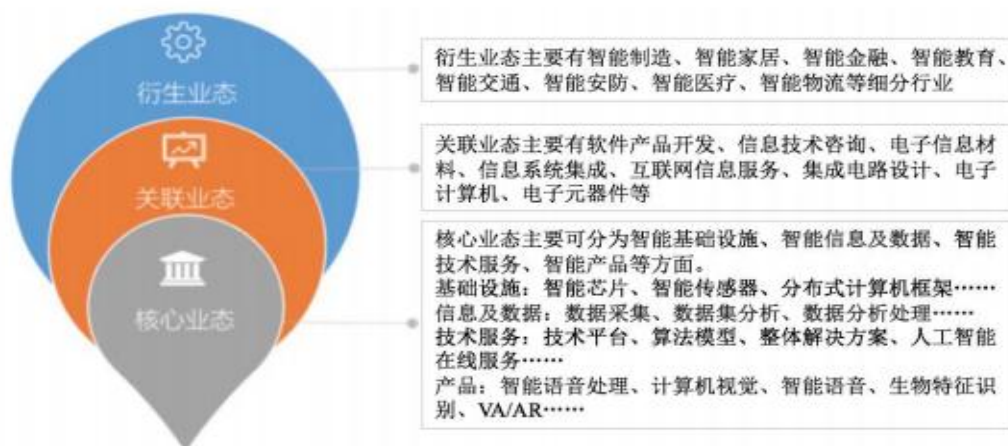


资料来源：招股说明书，安信证券研究中心

资料来源：招股说明书，安信证券研究中心

基础层主要为人工智能技术提供计算能力以及数据输入，技术层在基础层上开发算法模型，并通过数据训练和机器学习建模开发面向不同应用领域的技术，如智能语音、计算机视觉和自然语言处理等，在应用层将人工智能技术与应用场景结合起来，实现商业化落地。

图 34：人工智能产业生态



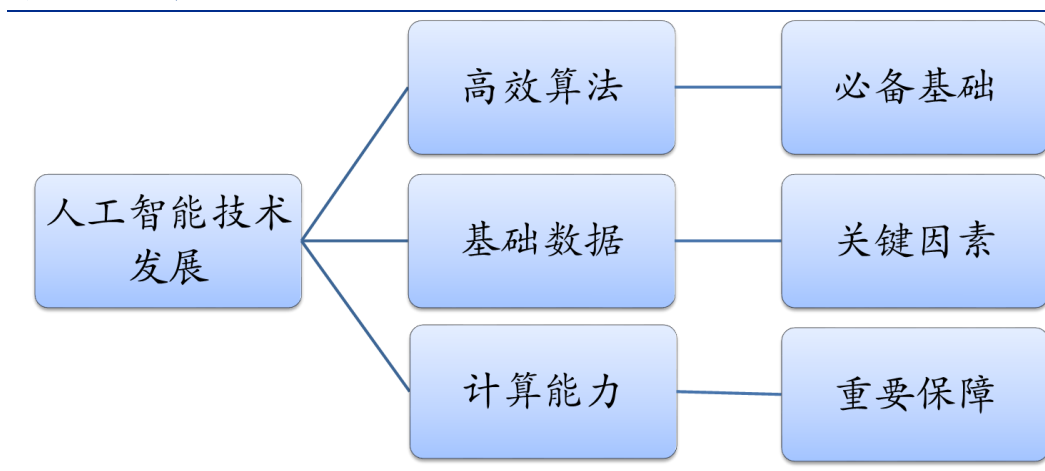
资料来源：招股说明书，安信证券研究中心

人工智能产业生态主要分为核心业态、关联业态、衍生业态三个层次。在核心业态中的基础设施与技术服务提供了重要保障，有助于人工智能与行业领域逐渐开展融合合作，将改变甚至重新塑造传统行业，逐渐形成关联业态与衍生业态。从人工智能产业进程来看，核心业态中的技术突破是推动产业升级的核心驱动力。核心业态中的数据资源、运算能力、核心算法将共同发展，掀起人工智能第三次新浪潮。

### 3.1.2. 当下现状：深度学习、丰富数据、强大算力，共同助推行业高速发展

经过 60 多年的演进，特别是在移动互联网、大数据、超级计算、脑科学等新理论新技术以及经济社会发展强烈需求的共同驱动下，人工智能加速发展，呈现出深度学习、跨界融合、人机协同、群智开放、自主操控等新特征。

图 35：人工智能技术发展的三大核心要素



资料来源：招股说明书，安信证券研究中心

- 当下，行业采用深度学习算法，超越了传统机器的各种浅层学习算法，利用数据处理网络形成特征与分类之间的关联规则。近年来，以深度学习算法为代表的人工智能技术快速发展，在计算机视觉、语音识别、自然语言等领域都实现了突破。
- 数据集的丰富性和大规模性对算法训练非常重要，获取海量而优质的应用场景数据是实现机器精准识别的关键基础。深度学习算法中的训练数据量与人工智能算法模型的准确性密切相关，数据量越大，模型推断效果越好、有用性越强。因此，经处理的结构化数据是人工智能持续发展的基础性动力。
- 高速和大规模的算力能对数据不断地迭代、试错以挖掘最优的关联规则，传统的 CPU 难以满足大规模和高速率的计算需求。随着图形处理器 (GPU)、专用芯片、分布式计算等各种更加强大算力的发展，此类高效计算能力将大幅推动人工智能发展。

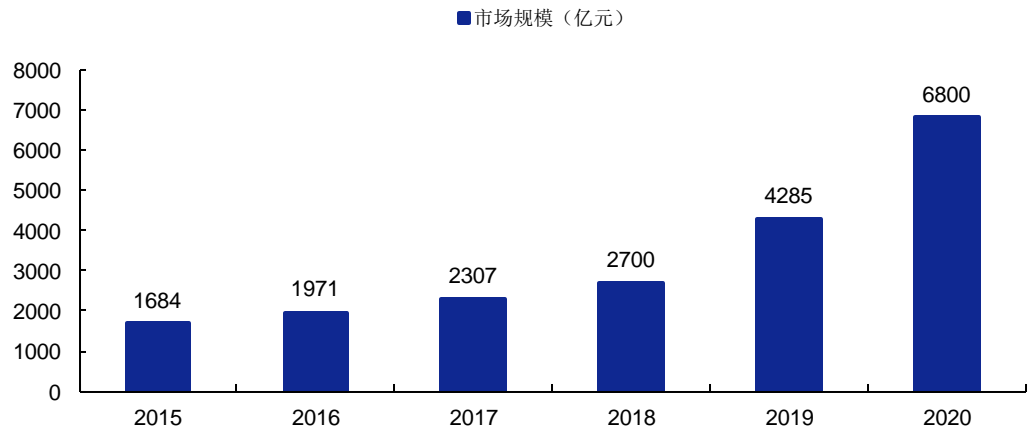
## 3.2. 发展前景：国内 AI 市场增速将达到 75%，国际人才投放量已位居第二

### 3.2.1. 市场规模：全球人工智能产业规模将达 6800 亿元，中国或成新增长极

当下，人工智能在深度学习、海量数据和高性能计算的支撑下，现已进入产业化应用初期。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，各国纷纷出台政策与规划，加大资本投入，实现了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。

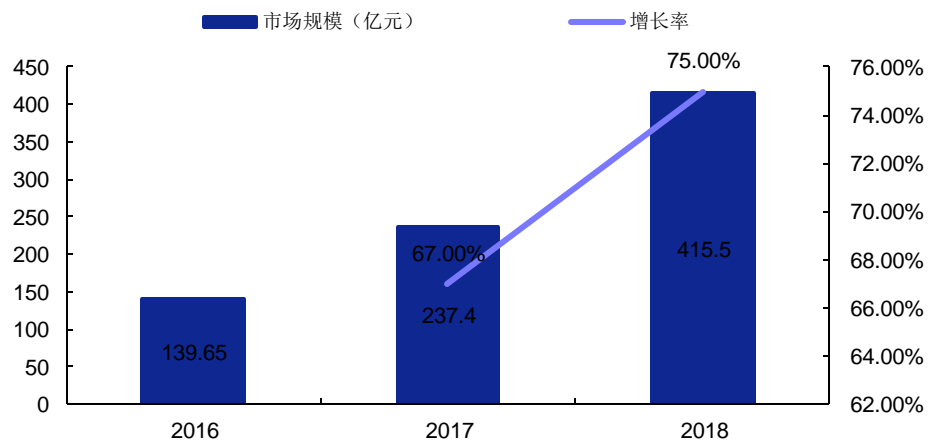
图 36: 2015-2020 年全球人工智能市场规模及预测



资料来源:招股说明书,安信证券研究中心

当前,我国正在成为世界人工智能领域的新增长极。2017 年我国人工智能市场规模达到 237.4 亿元,相较于 2016 年增长 67%。随着人工智能各项技术的不断成熟以及各类应用场景的落地,预计在 2018 年,国内人工智能市场增速将达到 75%,整体规模将达到 415.5 亿元。

图 37: 2016-2018 年中国人工智能市场规模及增速



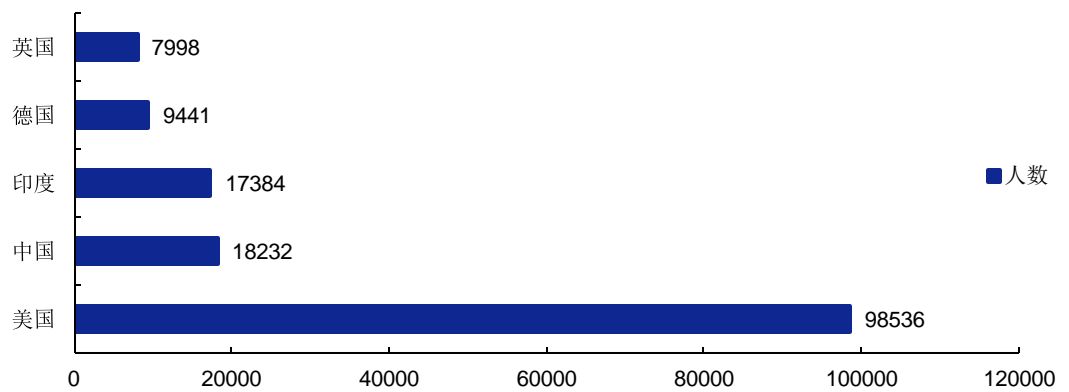
资料来源:《中国人工智能发展报告 2018》,安信证券研究中心

### 3.2.2. 人才投入: 国家大力投入高科技人才, 中国占世界总量 8.9% 居于第二

2018 年美国在人工智能人才投入量独占鳌头,累计高达 28536 人,占据世界总量的 13.9%。中国国际人工智能人才投入总量位列第二,累计达 18232 人,占美国人数的 65%,占据世界总量的 8.9%。

人工智能行业居于产业链上游,国家对于此类高科技人才有着高投入、高关注。伴随着科技的不断提升以及新兴企业的不断萌芽,人工智能行业在国内外都有着巨大的发展潜力。

图 38：2018 年各国国际人工投入总量

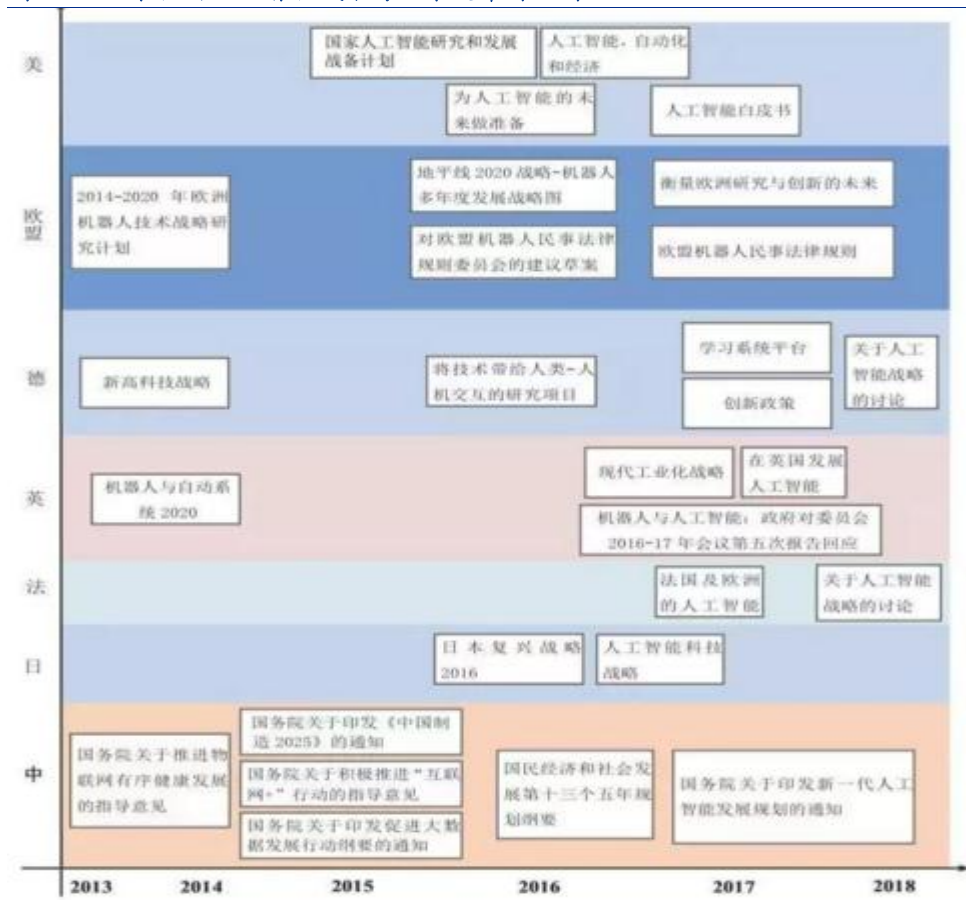


资料来源：《中国人工智能发展报告 2018》，安信证券研究中心

### 3.2.3. 政策发布：政策利好，行业有望成为国家支柱产业

美国发布的政策着力点在于应对人工智能蓬勃发展的大趋势，欧盟以及德、英、法为代表的欧洲国家，着重关注的是人工智能带来的伦理和道德。日本的人工智能政策发布较晚，政策预期在国家层面建立起相对完整的人工智能研发促进机制。中国人工智能政策在初期偏向于互联网领域，因此相关产业偏重于应用层，如计算机视觉、自然语言处理、智能机器人和语音识别。因此，目前中国的人工智能战略重点在于强化人工智能对科技、经济、社会发展和国家安全的全面支撑。

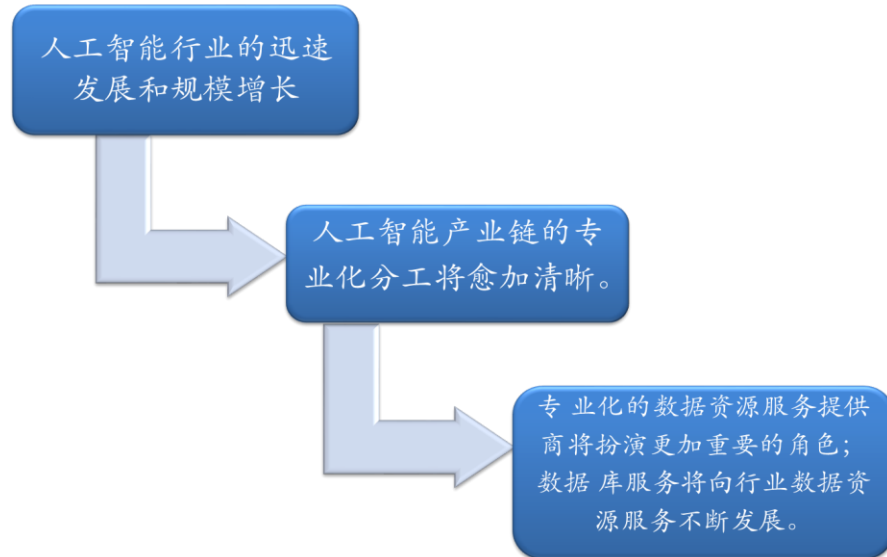
图 39：世界各国人工智能战略与政策发布情况图



资料来源：招股说明书，安信证券研究中心

未来数年，预计国内外人工智能行业产业规模仍将快速增长。从公司所处的人工智能细分领域来看，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。这些发展趋势都对数据库产品和服务带来了更多元化的发展方向和更丰富的需求。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。

图 40：行业未来发展趋势



资料来源：招股说明书，安信证券研究中心

### 3.3. 竞争格局：未来有望成为国内龙头企业，赶超国外强劲对手 Appen

#### 3.3.1. 国内主要企业：没有龙头企业，竞争壁垒较强

2015 年-2016 年前后，人工智能技术进入广泛应用期和高速发展期，产生了众多的企业。行业内尚未有形成垄断的龙头企业，公司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销售的主要企业之一，拥有多年的业务实践和创新，技术实力、业务规模均居行业领先地位，形成了较强的竞争壁垒。

表 7：行业内主要企业

行业内主要企业	简介
慧听科技	北京慧听科技有限公司成立于 2011 年，公司业务包括语音识别、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数据库制作、音乐数据制作标注、语音质量评测等。
标贝科技	标贝（北京）科技有限公司成立于 2016 年，总部位于北京。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、图像识别等人工智能数据服务。

资料来源：招股说明书，安信证券研究中心

人工智能数据资源开发及服务行业中的国内企业主要有除海天瑞声、慧听科技、标贝科技。行业内企业提供的人工智能数据资源产品及服务在技术领域、应用场景方面各有差异。慧听科技、标贝科技均为非上市公司。

### 3.3.2. 境外强劲对手：体量差距较大，但盈利能力与数据库种类并不逊色

Appen 于 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，已于 2015 年 1 月 7 日于澳大利亚证券交易所上市。作为全球前十的技术公司，Appen 是全球领先的机器学习和人工智能高质量人类注释数据集的领导者，是行业中的佼佼者，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，帮助他们开发、增强和利用依赖自然语言和机器学习的产品。

图 41：对手 Appen 公司的经营特点



资料来源：招股说明书，安信证券研究中心

表 8：与 Appen 非财务指标对比

	海天瑞声	Appen
数据库种类	语音、计算机视觉、自然语言	语音、计算机视觉、自然语言
语种/方言	130 余个	180 个
员工数量	127 人	513 人
主要客户	大型科技公司，如阿里巴巴、腾讯、百度、微软、三星等；人工智能企业，如科大讯飞、商汤科技、云知声、海康威视等；科研机构，如中国科学院、清华大学等	微软等大型科技公司、汽车厂商及政府

资料来源：招股说明书，安信证券研究中心

作为国内为数不多的可以在语音、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，基于公司自主研发的创新技术，海天瑞声在语种以及数据库种类这两方面与境外上市公司 Appen 差异不大，可以体现出公司在语言学研究、数据库的数据资源设计与处理方面的技术优势。

从主要客户上可以看出海天瑞声经过多年的发展，公司积累了深刻的行业洞见和突出的技术优势，与国内外诸多例如阿里巴巴、腾讯、百度等机构形成了长期稳定的合作关系，赢得了众多大型优质客户的认可。

虽然海天瑞声的员工数量不多，但是这从另一方面体现出海天瑞声技术人员的质量高，有显著的人才优势，能够有效升级技术水平，抓住全球人工智能快速发展的机遇。

表 9：2018 年与境外可比公司的营收、净利润对比

财务指标	海天瑞声	Appen
营业收入	19,265.8 万元	36,428.9 万澳元
净利润	6,714.2 万元	4,172.8 万澳元

资料来源：招股说明书，安信证券研究中心

营业收入和净利润方面海天目前的体量较小，营收仅为 Appen 的 1/9。但是海天 34.85% 利润率是 Appen 11.45% 利润率的近 3 倍。这充分体现出海天瑞声的较好盈利能力，且随着人工智能市场规模的不断扩张，同 Appen 在规模上的差距会进一步减少。

#### 4. 潜在风险：上游竞争加剧，下游需求恐无法及时满足

业绩波动风险：报告期内，受益于人工智能行业整体的高速发展，公司呈现出较高的成长性，营业收入从 2016 年的 8,422.86 万元增长至 2018 年的 19,265.77 万元，复合增长率达 51.24%。虽然当前公司的经营业绩呈现高速增长态势，但各期增长速度仍有一定波动。如果公司未来不能及时提供满足下游客户需求的产品和服务，公司未来业绩将存在大幅波动

市场竞争加剧风险：人工智能行业属于国家支持的战略行业，而人工智能数据行业位于产业链上游，处于快速增长期，具有广阔的发展前景。近年来，参与该行业的公司不断增多，人工智能数据公司大量涌现，一方面推动行业快速发展，另一方面也导致行业竞争加剧。公司如果不能持续有效地制定并实施业务发展规划，保持技术、产品和服务的领先性，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和市场地位下降的风险，进而影响公司的盈利能力和发展潜力。

## ■ 分析师声明

诸海滨声明，本人具有中国证券业协会授予的证券投资咨询执业资格，勤勉尽责、诚实守信。本人对本报告的内容和观点负责，保证信息来源合法合规、研究方法专业审慎、研究观点独立公正、分析结论具有合理依据，特此声明。

## ■ 本公司具备证券投资咨询业务资格的说明

安信证券股份有限公司（以下简称“本公司”）经中国证券监督管理委员会核准，取得证券投资咨询业务许可。本公司及其投资咨询人员可以为证券投资人或客户提供证券投资分析、预测或者建议等直接或间接的有偿咨询服务。发布证券研究报告，是证券投资咨询业务的一种基本形式，本公司可以对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向本公司的客户发布。

## ■ 免责声明

本报告仅供安信证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因为任何机构或个人接收到本报告而视其为本公司的当然客户。

本报告基于已公开的资料或信息撰写，但本公司不保证该等信息及资料的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映本公司于本报告发布当日的判断，本报告中的证券或投资标的价格、价值及投资带来的收入可能会波动。在不同时期，本公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，本公司将随时补充、更新和修订有关信息及资料，但不保证及时公开发布。同时，本公司有权对本报告所含信息在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以本公司向客户发布的本报告完整版本为准，如有需要，客户可以向本公司投资顾问进一步咨询。

在法律许可的情况下，本公司及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或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务，提请客户充分注意。客户不应将本报告为作出其投资决策的惟一参考因素，亦不应认为本报告可以取代客户自身的投资判断与决策。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，无论是否已经明示或暗示，本报告不能作为道义的、责任的和法律的依据或者凭证。在任何情况下，本公司亦不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告版权仅为本公司所有，未经事先书面许可，任何机构和个人不得以任何形式翻版、复制、发表、转发或引用本报告的任何部分。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“安信证券股份有限公司研究中心”，且不得对本报告进行任何有悖原意的引用、删节和修改。

本报告的估值结果和分析结论是基于所预定的假设，并采用适当的估值方法和模型得出的，由于假设、估值方法和模型均存在一定的局限性，估值结果和分析结论也存在局限性，请谨慎使用。

安信证券股份有限公司对本声明条款具有惟一修改权和最终解释权。

■ 销售联系人

上海联系人	朱贤	021-35082852	zhuxian@essence.com.cn
	孟硕丰	021-35082788	mengsf@essence.com.cn
	李栋	021-35082821	lidong1@essence.com.cn
	侯海霞	021-35082870	houhx@essence.com.cn
	潘艳	021-35082957	panyan@essence.com.cn
	刘恭懿	021-35082961	liugy@essence.com.cn
	孟昊琳	021-35082963	menghl@essence.com.cn
	苏梦		sumeng@essence.com.cn
	孙红	18221132911	sunhong1@essence.com.cn
	北京联系人	温鹏	010-83321350
姜东亚		010-83321351	jiangdy@essence.com.cn
张莹		010-83321366	zhangying1@essence.com.cn
李倩		010-83321355	liqian1@essence.com.cn
姜雪		010-59113596	jiangxue1@essence.com.cn
王帅		010-83321351	wangshuai1@essence.com.cn
夏坤		15210845461	xiakun@essence.com.cn
深圳联系人		胡珍	0755-82528441
	范洪群	0755-23991945	fanhq@essence.com.cn
	杨晔	0755-23919631	yangye@essence.com.cn
	巢莫雯	0755-23947871	chaomw@essence.com.cn
	王红彦	0755-82714067	wanghy8@essence.com.cn
	黎欢	0755-23984253	lihuan@essence.com.cn

安信证券研究中心

深圳市

地址： 深圳市福田区深南大道 2008 号中国凤凰大厦 1 栋 7 层

邮编： 518026

上海市

地址： 上海市虹口区东大名路 638 号国投大厦 3 层

邮编： 200080

北京市

地址： 北京市西城区阜成门北大街 2 号楼国投金融大厦 15 层

邮编： 100034