

证券简称：海天瑞声

证券代码：688787

北京海天瑞声科技股份有限公司

(Beijing Haitian Ruisheng Science Technology Ltd.)

(北京市海淀区成府路 28 号 4-801)

海天瑞声

2023 年度向特定对象发行 A 股股票 方案论证分析报告（修订稿）

二〇二三年十月

北京海天瑞声科技股份有限公司（以下简称“海天瑞声”或“公司”）是上海证券交易所科创板上市的公司。为满足公司业务发展的资金需求，增强公司资本实力，提升盈利能力，根据《中华人民共和国公司法》《中华人民共和国证券法》和《上市公司证券发行注册管理办法》（以下简称“《注册管理办法》”）等有关法律、行政法规、部门规章或规范性文件和《公司章程》的规定，公司编制了 2023 年度向特定对象发行 A 股股票方案论证分析报告（修订稿）。

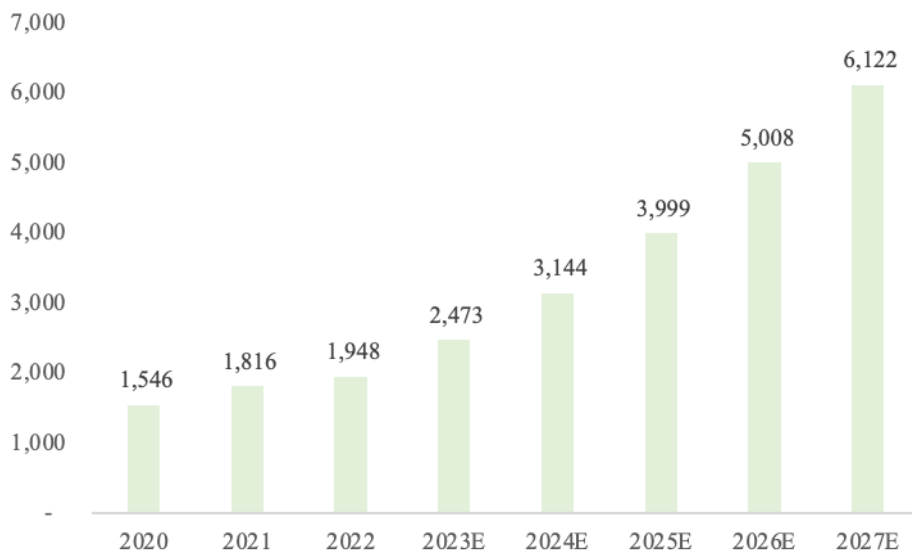
一、本次向特定对象发行的背景

（一）全球人工智能行业快速发展，我国正加速布局 and 应对

随着互联网、云计算、物联网以及穿戴设备的发展，数据要素价值不断释放，数字化转型成为大势所趋。目前，全球主要国家均高度重视发展数字经济，纷纷出台战略规划，重塑数字时代的国际竞争新格局。人工智能作为数字经济发展的底层核心技术之一，成为数字经济发展的关键战略抓手。2022 年底 ChatGPT 的出现，掀起又一波人工智能发展热潮，以无监督学习模式为代表的预训练+人类反馈强化学习所构成的大模型技术路线的落地意味着人工智能开启发展新范式，基础模型能力通过预训练及基于人类反馈的强化学习得到不断解锁，以解决海量开放式任务。

根据艾瑞咨询数据，2022 年中国人工智能产业规模达 1,948 亿元，预计 2027 年市场规模将达到 6,122 亿元，年复合增长率为 25.6%，主要与智算中心建设以及大模型训练等需求拉动的 AI 芯片市场、无接触服务需求拉动的智能机器人及对话式 AI 市场等快速增长相关。

图：2020-2027 年中国人工智能产业规划（单位：亿元）



数据来源：艾瑞咨询

人工智能大模型因其良好的通用性、泛化性和迁移性，有助于推动人工智能进入大规模落地应用，已成为人工智能发展新赛道。同时其强大的理解和生成能力，将驱动人工智能技术加速与实体产业融合，并深刻改变未来人类的生活和工作方式，发展大模型技术成为全球各国比拼科技实力，提升经济效率，拉升经济增长的重要动能之一。目前，国际巨头纷纷布局以大模型为核心的通用人工智能产业，产业进入加速发展期。在这一信息技术重点领域，我国与国际巨头存在一定差距，正加速布局和应对。国内众多研究机构、企业积极研究生成式 AI 大模型技术的最优路径，并进行产品发布。近期，在国内科技及投资各领域的高度关注下，百度、商汤、阿里巴巴、华为、科大讯飞、360、京东、字节跳动等企业均有所行动。

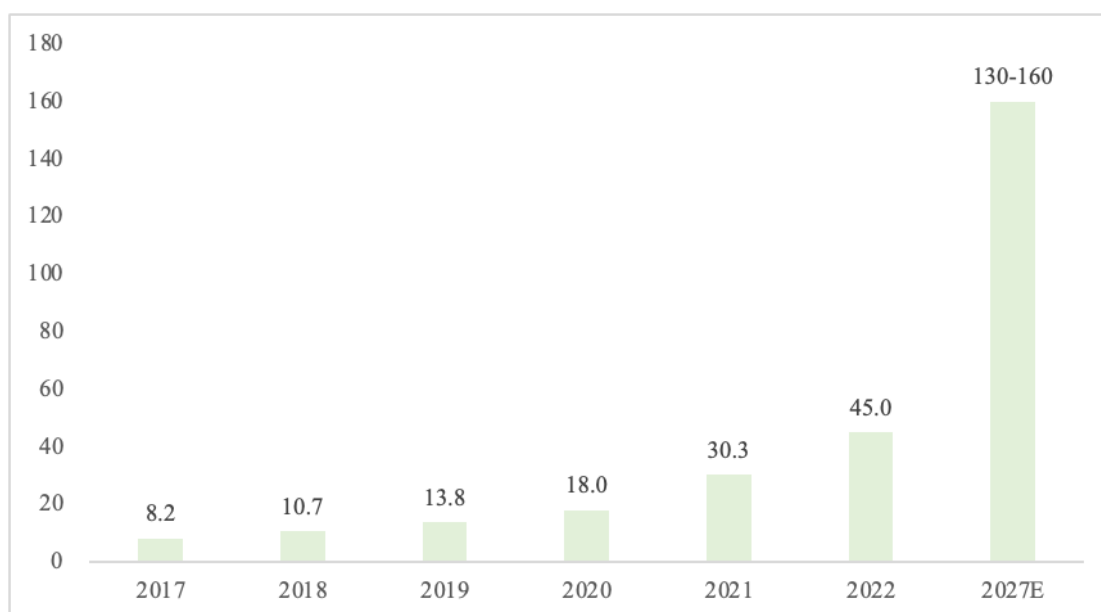
我国在“十四五”期间，针对人工智能的未来发展陆续出台了相关指导方案和激励政策，对人工智能的整体发展方向和技术发展重点做出重要规划，同时提出加强算法创新与应用、推动算力基础设施建设、完善数据基础支撑体系等关键建议，倡导未来不断夯实产业发展新基础。全国各地亦陆续出台多项数据政策，其中，《北京市促进通用人工智能创新发展的若干措施》明确提出要“系统构建大模型等通用人工智能技术体系：开展大模型创新算法及关键技术研究；加强大模型训练数据采集及治理工具研发；建设大模型评测开放服务平台；构建大模型基础软硬件体系。推动通用人工智能技术创新场景应用。”《北京市加快建设具有全球影响力的人工智能创新策源地实施方案（2023-2025 年）》提出“到 2025 年，

人工智能基础理论研究取得突破；关键核心技术基本实现自主可控，其中部分技术与应用研究达到世界先进水平；人工智能高水平应用深度赋能实体经济，促进经济高质量发展”的目标，并进一步提出了“自然语言、通用视觉、多模态交互大模型等形成完整技术栈；生成式产品成为国内市场主流应用和生态平台”等具体目标。

（二）人工智能基础数据服务业规模不断提升，大模型技术发展带来 AI 范式变革，正催生更多数据及其服务模式新需求

在人工智能产业链中，算法、算力和数据共同构成技术发展的三大核心要素。过去十年，人工智能产业以算法为中心，随着算法趋于开源，数据的重要性愈发凸显。在人工智能模型从技术理论到应用落地的过程中，需要依赖大量的训练数据，相较于以模型为中心的训练方法，以海量数据为中心的训练方法能够提升模型推断结论的可靠性。万亿 GB 量级的数据随着互联网、云计算、物联网、大数据等发展源源不断地产生，但数据质量参差不齐，对海量的复杂数据进行深入挖掘、输出，进而激活和释放数据的深层价值也成为数据市场的发展重点。根据德勤数据，2022 年中国人工智能基础数据服务市场规模为 45 亿元，2027 年规模将达到 130-160 亿元，年复合增长率为 23.6%-28.9%。

图：2017-2027 年中国人工智能基础数据服务市场规模（单位：亿元）



数据来源：德勤

随着人工智能进入大模型时代，数据需求和数据服务模式不断提升，数据的质量以及数据清洗的工程化能力会显著拉开大模型预训练阶段的效果差距；同时，更多模型或将采用类强化学习模式来进行特定领域或特定方向上的优化迭代，以使得机器能够以更加接近于人类期望的方式提供答案输出。对于大模型训练而言，不仅需要持续获取大规模、高质量、多模态、多场景、多垂向的数据，更须具备持续迭代的高质量数据清洗和标注策略，以不断提升包括预训练（Pre-training）、模型微调（Fine-tune）及奖励模型（Reward Model）等过程中所需数据（例如提示（Prompt）类数据）的质量，确保语言类和常识性知识之外的其他垂直领域的应用场景的能力提升，为大模型精确性、通用性及泛化能力的实现奠定坚实基础。

在以上背景下，一方面，大规模、高质量数据集重要性凸显，成为模型训练效果的核心支撑之一。除了在通用基础能力方面需要大量高质量数据训练外，大模型算法训练的需求正逐渐从通用基础能力建设向垂直领域拓展，数据需求向专业化方向发展。为加速实现人工智能产业化落地，行业将衍生出更多垂直场景的数据需求，大模型将通过不断学习各个专业领域的行业高质量数据，实现更广阔的垂直拓展。另一方面，AI 发展所面对的数据前沿性及工程化技术的挑战也较为凸显。长期看，只有 AI 数据处理技术的不断拓新与发展，才能及时适应甚至超前引领大模型技术和应用的发展。

（三）各地推动数据基础制度建设，数据要素市场迎来新发展机遇

近年来，我国数字经济蓬勃发展，数据要素因具有基础性战略资源和关键性生产要素的双重属性，相关市场规模持续增长。尤其在《中共中央、国务院关于构建数据基础制度更好发挥数据要素作用的意见》出台后，我国系统性布局了数据基础制度体系的“四梁八柱”，加速了数据流通交易和数据要素市场发展，进一步推动了公共数据、企业数据、个人数据合规高效流通使用。为更好响应中央号召，北京、上海、广州、深圳、杭州等地数据政策陆续出台，逐步构建了多层次、多元化数据要素市场生态体系。

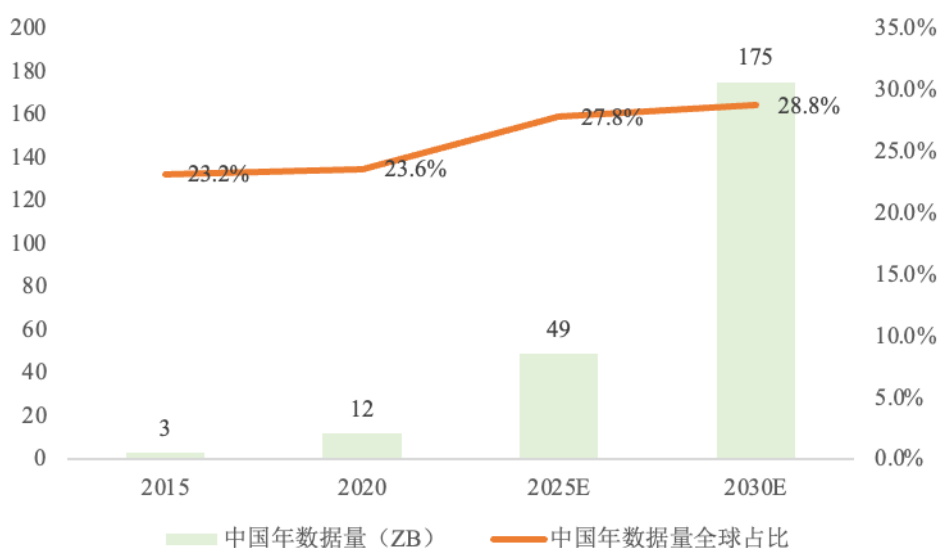
以北京为例，《北京市促进通用人工智能创新发展的若干措施》和《关于推进北京市数据专区建设的指导意见》指出，北京市要加快建设“数据基础制度先行先试示范区”（以下简称“先行先试示范区”），探索打造数据训练基地，归集高质量基础训练数据集，推动数据要素高水平开放，提升本市人工智能数据标注

库规模和质量，并建设针对重大领域、重点区域或特定场景建设专题数据区域，吸纳市场主体和数据、技术、资本等多元要素参与。北京市陆续出台的多项文件旨在打破数据壁垒，推动数据融合利用，加快推动公共数据开放，促进数据要素流通，激发数字市场创新活力，释放和发展数字化生产力，打造多层次数据要素市场，成为具有竞争力和影响力的数字产业集群。按照“政府引导、市场运作、创新引领、安全可控”的原则，“先行先试示范区”有望成为国际领先的数据要素高效流通核心枢纽。

（四）我国具有拥有海量丰富的数据资源，但数据质量亟待提升

我国各行业的数据资源较为丰富，根据艾瑞咨询数据，2015 年-2030 年中国数据量规模由 3ZB 将增长至 175ZB，预计 2030 年中国数据量约占全球的 28.8%，年复合增长率约为 31%。

图：2015-2030 年中国数据量规模及全球占比



数据来源：艾瑞咨询

虽然中国数据资源丰富，但由于数据挖掘不足，以及大量数据无法在市场上自由流通等原因，优质中文数据集仍然稀缺。以 ChatGPT 为例，其模型训练数据中，中文数据来源不足千分之一。目前，国内头部科技企业主要基于公开数据集以及自身特有的数据进行大模型训练，但由于中文优质数据质量以及数据资源的制约，国内大模型的能力与以 ChatGPT 为代表的国际大模型相比仍存在一定差距。

国内缺乏高质量数据集的主要原因包括当前国内数据挖掘和数据治理的力度不足、资金投入较大；数据流通与数据安全保障措施不够健全；国内市场缺乏开源意识，大量数据无法在市场上自由流通；国内相关公司成立较晚，数据积累较少；学术领域中文数据集受重视程度低以及国产数据集市场影响力及普及度较低等。从原始数据到可被应用的数据集产品，需要经历数据集结构设计、数据获取、数据处理（包括数据清洗、数据标注/优化等）等过程，以形成可供使用的优质数据集，国内数据服务市场的发展有助于缓解中文数据集数量不足和质量欠佳等问题。

二、本次发行证券及其品种选择的必要性

（一）本次发行证券的品种

本次发行的股票种类为境内上市人民币普通股（A 股），每股面值为 1.00 元。

（二）本次发行证券品种选择的必要性

1、丰富公司训练数据集产品类别，拓展 AI 大模型训练数据集产品及服务领域，探索新型业务

公司作为全球领先的人工智能训练数据提供商，一直深耕人工智能行业，在基础数据服务领域居于龙头地位，自成立以来，公司致力于为各类机构提供深度学习算法模型开发训练所需的专业数据集，所提供训练数据涵盖智能语音、计算机数据、自然语言处理等多个核心领域。

为更好实现公司业务扩张战略，公司在保障传统业务稳健发展的同时，不断探索寻求新的业绩增长点。公司将基于过往海量数据服务经验，结合以大模型发展为代表的行业前沿需求，研发生产大模型领域专业数据集，丰富公司数据集产品类别，拓展 AI 大模型训练数据领域服务能力，力争将大模型训练数据等创新业务打造成为具有潜在高增长价值的新型业务板块。并藉此提升行业内面向大模型训练数据集的类别和质量，实现基于大模型通用能力和垂直领域数据的支撑和训练学习，协助实现个人数据、企业数据、公共数据等各类高价值数据资源汇聚。

2、巩固公司核心技术壁垒，构建大模型数据处理技术通用化能力，构建长期技术实力支撑并提升数据服务综合竞争力

数据集生产各项技术与能力和一体化技术支撑平台是公司核心技术的重要体现。本次募集资金投资项目的建设将基于公司在深度学习阶段数据集生产所积累的 know-how，自主研发海天瑞声数据生产垂直大模型，构建大模型数据处理技术通用化解决方案能力，实现完整、可持续迭代的大模型数据技术框架和数据策略；并以数据生产垂直大模型为核心，升级海天瑞声一体化技术支撑平台中包括设计、获取、清洗、标注、安全管理、质控评测等不同的环节的功能，对大模型时代的数据处理需求形成全面支撑。以上举措均将进一步提高公司在人工智能基础数据服务领域的智能化水平，巩固公司的核心技术壁垒，形成长期技术实力支撑。

此外，本项目有助于进一步优化公司的数据处理技术，促进数据资源处理经验的进一步沉淀，长期来看，可以大幅提高公司的数据处理能力、效率，提升服务范围 and 水平，适应人工智能发展的新阶段，获得有效长期的发展动力，进一步巩固和提升公司在数据服务领域的竞争力。

三、本次发行对象的选择范围、数量和标准的适当性

（一）本次发行对象选择范围的适当性

本次发行的发行对象为不超过三十五名（含三十五名）符合中国证监会规定条件的特定投资者，包括符合规定条件的证券投资基金管理公司、证券公司、信托公司、财务公司、保险机构投资者、合格境外机构投资者以及符合规定的其他法人、自然人或其他合格的投资者等。其中，证券投资基金管理公司、证券公司、合格境外机构投资者、人民币合格境外机构投资者以其管理的二只以上产品认购的，视为一个发行对象；信托公司作为发行对象的，只能以自有资金认购。

最终发行对象将在本次发行申请获得上交所审核通过并由中国证监会同意注册后，按照相关法律法规的规定及监管部门要求，由公司董事会或董事会授权人士在股东大会的授权范围内，与保荐机构（主承销商）根据询价结果协商确定。

若国家法律、法规对向特定对象发行 A 股股票的发行对象有新的规定，公司将按新的规定进行调整。

所有发行对象均以人民币现金方式按照同一价格认购本次发行的股票。

本次发行对象的选择范围符合《注册管理办法》等法律法规的相关规定，发行对象的选择范围适当。

（二）本次发行对象数量的适当性

本次发行的发行对象数量不超过三十五名（含三十五名），发行对象的数量符合《注册管理办法》等相关法律法规的规定，发行对象数量适当。

（三）本次发行对象标准的适当性

本次发行对象应具有一定的风险识别能力和风险承担能力，并具备相应的资金实力。本次发行对象的标准符合《注册管理办法》等法律法规的相关规定，本次发行对象的标准适当。

四、本次发行定价的原则、依据、方法和程序的合理性

（一）本次发行定价的原则和依据

本次发行采取询价发行方式，定价基准日为发行期首日。

本次发行价格不低于定价基准日前 20 个交易日（不含定价基准日，下同）公司股票交易均价的 80%（即“发行底价”）。定价基准日前 20 个交易日公司股票交易均价=定价基准日前 20 个交易日公司股票交易总额/定价基准日前 20 个交易日公司股票交易总量。

若公司股票在定价基准日至发行日期间发生派息、送股、资本公积转增股本等除权除息事项，本次发行底价将按以下办法作相应调整。调整公式为：

派发现金股利： $P1=P0-D$

送红股或转增股本： $P1=P0/(1+N)$

两项同时进行： $P1=(P0-D)/(1+N)$

其中， P_0 为调整前发行底价， P_1 为调整后发行底价，每股派息/现金分红为 D ，每股送红股或转增股本数为 N 。

本次发行的最终发行价格将在公司本次发行申请获得上交所审核通过并经中国证监会作出同意注册决定后，由董事会或董事会授权人士根据股东大会的授权和保荐机构（主承销商）根据询价结果协商确定，但不低于前述发行底价。

本次发行定价的原则及依据符合《注册管理办法》等法律法规的相关规定，本次发行定价的原则合理。

（二）本次发行定价的方法和程序

本次向特定对象发行股票的定价方法和程序均根据《注册管理办法》等法律法规的相关规定，已经公司召开董事会、股东大会审议通过，并将相关公告在交易所网站及指定的信息披露媒体上进行披露。

本次发行定价的方法和程序符合《注册管理办法》等法律法规的相关规定，本次发行定价的方法和程序合理。

综上所述，本次发行定价的原则、依据、方法和程序均符合相关法律法规的要求，合规合理。

五、本次发行方式的可行性

（一）本次发行方式合法合规

1、本次发行符合《证券法》规定的发行条件

公司本次发行未采用广告、公开劝诱和变相公开的方式，符合《证券法》第九条第三款之规定。

2、公司不存在《注册管理办法》第十一条规定的不得向特定对象发行股票的情形：

（1）擅自改变前次募集资金用途未作纠正，或者未经股东大会认可；

（2）最近一年财务报表的编制和披露在重大方面不符合企业会计准则或者相关信息披露规则的规定；最近一年财务会计报告被出具否定意见或者无法表示

意见的审计报告；最近一年财务会计报告被出具保留意见的审计报告，且保留意见所涉及事项对上市公司的重大不利影响尚未消除。本次发行涉及重大资产重组的除外；

（3）现任董事、监事和高级管理人员最近三年受到中国证监会行政处罚，或者最近一年受到证券交易所公开谴责；

（4）上市公司及其现任董事、监事和高级管理人员因涉嫌犯罪正在被司法机关立案侦查或者涉嫌违法违规正在被中国证监会立案调查；

（5）控股股东、实际控制人最近三年存在严重损害上市公司利益或者投资者合法权益的重大违法行为；

（6）最近三年存在严重损害投资者合法权益或者社会公共利益的重大违法行为。

3、公司募集资金使用符合《管理办法》第十二条的规定：

（1）符合国家产业政策和有关环境保护、土地管理等法律、行政法规规定；

（2）除金融类企业外，本次募集资金使用不得为持有财务性投资，不得直接或者间接投资于以买卖有价证券为主要业务的公司；

（3）募集资金项目实施后，不会与控股股东、实际控制人及其控制的其他企业新增构成重大不利影响的同业竞争、显失公平的关联交易，或者严重影响公司生产经营的独立性；

（4）科创板上市公司发行股票募集的资金应当投资于科技创新领域的业务。

（二）本次发行程序合法合规

本次向特定对象发行股票已经公司董事会、股东大会审议通过，董事会、股东大会决议以及相关文件均在交易所网站及中国证监会指定的信息披露媒体上进行披露，履行了必要的审议程序和信息披露程序。

根据有关规定，本次发行方案尚需经上海证券交易所审核通过并经中国证监会同意注册。

综上所述，本次向特定对象发行股票的审议程序合法合规，发行方式可行。

六、本次发行方案的公平性、合理性

公司本次向特定对象发行股票的方案经董事会、股东大会审慎研究后通过，发行方案的实施将进一步增强公司资本实力，优化资产负债结构，通过新业务的开展提升公司的抗风险能力，符合全体股东利益。

本次向特定对象发行方案及相关文件已在中国证监会指定信息披露媒体上进行披露，保证全体股东的知情权。

综上所述，本次发行方案已经董事会及股东大会审议通过，认为该发行方案符合全体股东利益；本次发行方案及相关文件已履行了相关披露程序，保障了股东的知情权；同时，本次发行方案将在股东大会上接受参会股东的公平表决，具备公平性和合理性。

七、本次发行对原股东权益或者即期回报摊薄的影响以及填补的具体措施

根据《国务院关于进一步促进资本市场健康发展的若干意见》（国发[2014]17号）、《国务院办公厅关于进一步加强资本市场中小投资者合法权益保护工作的意见》（国办发[2013]110号）和《关于首发及再融资、重大资产重组摊薄即期回报有关事项的指导意见》（中国证券监督管理委员会公告[2015]31号）的相关规定，为保障中小投资者知情权，维护中小投资者利益，公司就本次向特定对象发行股票相关事项对即期回报摊薄的影响进行了认真分析并提出了具体的填补回报措施，相关主体对公司填补回报措施能够得到切实履行作出了承诺。具体情况如下：

（一）本次向特定对象发行摊薄即期回报对公司主要财务指标的影响

经公司第二届董事会第十七次会议、2023年第一次临时股东大会审议通过，并经第二届董事会第二十次会议审议修订，本次向特定对象发行股票募集资金总额为 66,600.00 万元；本次发行前公司总股本为 60,325,180 股，本次发行股份数量不超过本次发行前上市公司总股本 60,325,180 股的 20%，即不超过 12,065,036

股（含本数）。

本次发行完成后，公司总股本和归属于母公司股东权益将有一定幅度增加，公司即期及未来每股收益和净资产收益率面临下降的风险。为充分保障投资者的利益，公司对本次融资的必要性、合理性、本次募集资金投资项目与公司现有业务的关系、以及公司填补回报的具体措施进行分析以及作出相关承诺。

1、测算假设和前提

1、假设宏观经济环境、产业政策、行业发展状况、产品市场情况及公司经营情况等方面未发生重大不利变化。

2、假设本次向特定对象发行于 2023 年 11 月底实施完成，该完成时间仅为公司用于本测算的估计，最终以经中国证监会同意注册并实际发行完成时间为准。

3、本次向特定对象发行股票数量不超过本次发行前公司总股本的 20%，即不超过 12,065,036 股（含本数），募集资金总额不超过 66,600.00 万元（含本数）。因此，假设本次向特定对象发行股票数量为 12,065,036 股，该发行股票数量仅为公司用于本测算的估计，最终以经中国证监会同意注册后实际发行股票数量为准。

4、不考虑本次向特定对象发行募集资金运用对公司生产经营、财务状况（如营业收入、财务费用、投资收益）等的影响。

5、公司 2022 年度归属于上市公司股东的净利润为 2,945.41 万元，扣非后归属于上市公司股东的净利润为 1,014.91 万元。

6、假设公司 2023 年度归属于上市公司股东的净利润、扣除非经常性损益后归属于上市公司股东的净利润分别在上年同期的基础上按照下降 20%、持平和增长 20%的业绩变动幅度测算。

7、在预测公司 2023 年末总股本时，以本次向特定对象发行前公司总股本为基础，同时仅考虑本次向特定对象发行对总股本的影响，不考虑其他可能产生的股权变动事宜（如资本公积转增股本、股票股利分配、可转债转股）。

上述假设仅为测算本次向特定对象发行股票对公司即期回报主要财务指标的摊薄影响，不代表公司对 2023 年经营情况及财务状况的判断，亦不构成盈利

预测，2023 年公司收益的实现取决于国家宏观经济政策、行业发展状况、市场竞争情况、公司业务发展状况等诸多因素，存在较大不确定性。投资者不应据此进行投资决策，投资者据此进行投资决策造成损失的，公司不承担赔偿责任。

2、对公司即期回报的摊薄影响

基于上述假设和说明，公司测算了本次向特定对象发行股票对公司的每股收益等主要财务指标的影响如下：

项目	2022年度/2022年12月31日	2023年度/2023年12月31日	
		本次发行前	本次发行后
总股本（股）	42,800,000	60,325,180	72,390,216
假设一：公司2023年度归属于母公司股东的净利润和扣除非经常性损益后归属于母公司股东的净利润较2022年度下降20%			
归属于母公司股东净利润（万元）	2,945.41	2,356.33	2,356.33
扣除非经常性损益后归属于母公司股东的净利润（万元）	1,014.91	811.93	811.93
基本每股收益（元/股）	0.69	0.39	0.38
稀释每股收益（元/股）	0.69	0.39	0.38
扣除非经常性损益后基本每股收益（元/股）	0.24	0.13	0.13
扣除非经常性损益后稀释每股收益（元/股）	0.24	0.13	0.13
假设二：公司2023年度归属于母公司股东的净利润和扣除非经常性损益后归属于母公司股东的净利润与2022年度持平			
归属于母公司股东净利润（万元）	2,945.41	2,945.41	2,945.41
扣除非经常性损益后归属于母公司股东的净利润（万元）	1,014.91	1,014.91	1,014.91
基本每股收益（元/股）	0.69	0.49	0.48
稀释每股收益（元/股）	0.69	0.49	0.48
扣除非经常性损益后基本每股收益（元/股）	0.24	0.17	0.17
扣除非经常性损益后稀释每股收益（元/股）	0.24	0.17	0.17
假设三：公司2023年度归属于母公司股东的净利润和扣除非经常性损益后归属于母公司股东的净利润较2022年度增长20%			
归属于母公司股东净利润（万元）	2,945.41	3,534.50	3,534.50
扣除非经常性损益后归属于母公司股东的净利润（万元）	1,014.91	1,217.89	1,217.89

项目	2022年度/2022年12月 31日	2023年度/2023年12月31日	
		本次发行前	本次发行后
基本每股收益（元/股）	0.69	0.59	0.58
稀释每股收益（元/股）	0.69	0.59	0.58
扣除非经常性损益后基本每股收益 （元/股）	0.24	0.20	0.20
扣除非经常性损益后稀释每股收益 （元/股）	0.24	0.20	0.20

注：每股收益按照《公开发行证券的公司信息披露编报规则第9号——净资产收益率和每股收益的计算及披露》规定计算，测算2023年度相关数据及指标时，不考虑已授予限制性股票的发行、回购、解锁及稀释性影响，不考虑可能发生的权益分派及其他因素的影响。

（二）本次发行摊薄即期回报的特别风险提示

由于本次募集资金到位后从投入使用至募投项目投产和产生效益需要一定周期，在募投项目产生效益之前，股东回报仍然依赖于公司现有的业务基础，由于公司总股本增加，本次发行后将可能导致公司每股收益指标下降。本次向特定对象发行A股股票当年存在摊薄公司即期回报的风险。敬请广大投资者理性投资，并注意投资风险。

公司对相关财务数据的假设仅用于计算相关财务指标，不代表公司对经营情况及趋势的判断，也不构成对公司的盈利预测或盈利承诺。投资者不应根据上述假设进行投资决策，投资者据此进行投资决策造成损失的公司不承担赔偿责任。

（三）本次向特定对象发行股票的必要性和合理性

1、本项目建设是响应国家建立数据基础制度，落实北京建设“先行先试示范区”的必然选择

党的十八大以来，习近平总书记屡次强调建设数字中国以及构建数据要素的重要性，并明确指出数据是新的生产要素，是基础性资源和战略性资源，也是重要生产力。为进一步推动国家数字经济发展，发挥数据要素在经济发展中的重要价值，我国推出《中共中央、国务院关于构建数据基础制度更好发挥数据要素作用的意见》，从顶层设计角度，在数据产权、流通交易、收益分配、安全治理等方面构建了数据发展的基础制度和规划纲要，以促进数据合规高效流通使用，充

分发挥中国海量数据规模和丰富应用场景优势，赋能实体经济，激活数据要素的潜能。

北京市则率先开展国家数据基础制度“先行先试示范区”建设，2023年5月发布的《北京市促进通用人工智能创新发展的若干措施》指出，充分发挥政府引导作用和创新平台催化作用，整合创新资源，加强要素配置，营造创新生态，提升高质量数据要素供给能力，归集高质量基础训练数据集。2023年7月发布的《关于更好发挥数据要素作用进一步加快发展数字经济的实施意见》提出，“发展数据生产服务业，支持企业开展数据采集、清洗加工、存储计算、数据分析、数据标注、数据训练等数据生产服务，支持企业研发建设数据生产线，推进数据生产自动化”。公司作为人工智能基础数据服务领域具有较强国际竞争力的国内头部企业，有义务和责任积极响应北京建设“先行先试示范区”的号召，通过本项目的实施有效助力数据要素市场培育，推动数字经济创新发展，为北京市加快建设全球数字经济标杆城市提供助力。

2、本项目建设是践行国家规范生成式人工智能产品要求的重要举措

生成式人工智能产品因其复杂性可能带来社会风险、技术伦理风险、企业商业秘密和个人信息泄露风险、虚假信息风险、知识产权侵权风险及其他潜在风险。为了更好地促进生成式人工智能技术健康发展和规范应用，国家网信办等七部门于2023年7月出台了《生成式人工智能服务管理暂行办法》，该办法从内容合规、数据来源合法性、知识产权及商业秘密保护、虚假信息防范等方面，对生成式人工智能产品提出了全方位的合规要求。该办法明确提出，“提供者应当依法开展预训练、优化训练等训练数据处理活动”、“采取有效措施提高训练数据质量，增强训练数据的真实性、准确性、客观性、多样性”。

根据前述规定，数据获取、数据处理的高标准意味着数据获取难度及处理成本将大幅增加，以预训练阶段为例，由于大量数据来源应合法合规，需投入大量成本完成数据获取。因此，出于成本与数据集质量的平衡性考量，在大模型训练中，大模型厂商通常会选择与专业的第三方数据集厂商合作，由专业第三方提供的合规、高质量数据集或相关解决方案将成为践行国家规范生成式人工智能产品要求的重要举措。

3、本项目建设是支撑大模型训练，提升大模型输出能力的有效方式

随着人工智能应用场景日益丰富、产品智能化要求的不断提升，数据需求逐渐向海量、高质量、多元化方向演进。从自然数据源简单收集、获取的数据资源，通常无法直接满足大模型的训练需求，需经专业化的数据分类设计、清洗、加工处理，形成相应的工程化数据，以供大模型训练使用。一般而言，符合大模型训练标准的数据需具备质量高、规模大、样本丰富等三个特点。首先，海量具有无毒害性、公平性等高质量特征的数据集能够提高模型效果（例如，精度与可解释性），并且减少收敛到最优解的时间；其次，在强化学习阶段，原始数据由于存在信息量低、含有噪声或需补齐等问题，使用前需要进行数据对齐等诸多微调操作，优秀的指令数据集能够帮助大模型更好的泛化适配更多下游任务。再次，数据丰富程度能够显著提高大模型的泛化能力，减少过拟合情况的发生，达到更优的模型效果。

当前国内数据资源虽然丰富，但优质的中文大模型训练数据仍然稀缺，中文大模型训练数据数量与质量，受国内产业环境、数据积累程度、数据运营生态等因素影响，与全球领先国家仍存在一定差距，使得国内大模型难以拥有足够专业的数据资源进行训练。本项目通过提供覆盖预训练、强化学习及应用拓展阶段的海量、高质量专业数据集，更好的支撑大模型训练，提升大模型输出能力。

4、本项目建设符合公司“夯实传统业务，探索新型业务”的战略目标

为更好实现公司业务发展战略，公司在保障人工智能基础数据业务稳健发展的同时，不断探索寻求新的业绩增长点。如前文所述，数字经济时代下，数据要素市场发展前景广阔，大模型等人工智能技术已成为国家科技发展的重要抓手，但国内数据仍存在数据质量差、各领域数据无法流通等问题制约了人工智能行业的发展。公司将基于过往的数据服务经验，结合行业前沿需求，积极拓展大模型训练数据服务领域，力争将大模型训练数据等创新业务打造成为具有潜在高增长价值的新型业务板块。

5、本项目建设是落实国家科技创新发展战略的重要举措

人工智能是战略性新兴产业的重要组成部分，对我国经济发展和提升国家战略安全具有重要意义。在世界政治经济格局加速重构的影响下，未来逆全球化趋势仍将延续。全球产业合作格局重构、国际分工体系全面调整，关键环节的国际竞争将加剧，我国在关键核心技术上的问题愈发突出，战略性新兴产业的产业链安全稳定存在潜在隐患。因此，我国需要进一步集中优势资源，在重点领域加快突破一批关键核心技术，助力提升我国新兴产业的产业链关键环节、关键领域、关键产品的安全保障能力，保障国家战略安全。

公司是我国人工智能数据服务领域的龙头提供商，本项目以研发数据生产垂直大模型为核心，并基于该生产垂直大模型对数据集生产的强大支撑能力，升级海天瑞声一体化技术支撑平台，持续以自主可控的技术与平台为我国人工智能技术与产业发展提供支撑。本项目的建设是公司落实国家科技创新发展战略的重要举措。

6、本项目建设是巩固公司的核心技术壁垒，构建长期技术实力支撑的必然手段

随着人工智能从深度学习阶段走向大模型阶段，对数据集的需求发生较大变化，大模型时代所需的数据集具有数量多、质量高、多模态的特点：在无监督学习阶段，模型所需的数据量巨大，数据合规要求较高；在迁移学习阶段，模型所需的数据质量较高，并需要以相关领域know-how作为模型输入。此外，随着多模态大模型的不断发布，语音、文本和视频图像数据等多种类别的数据集需求增多。

数据集生产能力和一体化技术支撑平台是公司核心技术的重要体现。目前ChatGPT等模型执行通用生成任务的效果证明了采用大模型来进行数据集的生成是可行的。本次募集资金投资项目的建设，将积累沉淀公司深度学习阶段数据集生成的know-how，自主研发数据生产垂直大模型，进一步提高公司在基础数据服务领域的智能化水平，巩固公司的核心技术壁垒，构建长期技术实力支撑。

7、本项目建设是优化公司数据集生产流程，提升数据服务综合竞争力的有力途径

大模型训练数据集的生产流程包括设计、获取（模型生成）、清洗、标注、安全管理、质控评测等不同的环节。系统化的开发平台和专业化的软件处理工具对应对大模型时代的数据处理需求和全流程支撑至关重要。本项目有助于进一步优化公司的数据处理技术，促进数据资源处理经验的进一步沉淀，长期来看，可以大幅提高公司的数据处理能力、效率，提升服务范围和水平，适应人工智能发展的新阶段，获得有效长期的发展动力，进一步巩固和提升公司在数据服务领域的竞争力。

（四）本次募集资金投资项目与公司现有业务的关系及公司从事募投项目在人员、技术、市场等方面的储备情况

1、本次募集资金投资项目与公司现有业务的关系

本次募集资金拟用于“AI大模型训练数据集建设项目”和“数据生产垂直大模型研发项目”，募集资金投资项目均围绕公司主营业务进行。

AI大模型训练数据集建设项目将拓展训练数据产品体系，公司顺应人工智能大模型的发展趋势，大幅提升训练数据产品的数据规模、广度、质量和精度，扩大产品覆盖的场景和领域，是前瞻性布局大模型领域、拓展潜在高增长价值的新型业务板块的重要举措；数据生产垂直大模型研发项目以研发海天瑞声数据生产垂直大模型为核心，升级海天瑞声一体化技术支撑平台，进一步延伸大模型领域综合数据服务能力，旨在提高公司业务能力的维度、广度、效率和水平，巩固公司的核心技术壁垒，构建长期技术实力支撑，不断提高公司核心竞争力。

此外，本次募投项目建设符合公司“夯实传统业务，探索新型业务”的战略目标。为更好实现公司业务发展战略，公司在保障人工智能基础数据业务稳健发展的同时，不断探索寻求新的业绩增长点。如前文所述，数字经济时代下，数据要素市场发展前景广阔，大模型等人工智能技术已成为国家科技发展的重要抓手，但国内数据仍存在数据质量差、各领域数据无法流通等问题制约了人工智能行业的发展。公司将基于过往的数据服务经验，结合行业前沿需求，积极拓展大模型训练数据服务领域，力争将大模型训练数据等创新业务打造成为具有潜在高增长

价值的新型业务板块。

2、公司从事募投项目在人员、技术、市场等方面的储备情况

（1）人员储备

公司拥有高素质的管理、运营及研发团队，公司高管及核心研发人员大多毕业于清华、北大、复旦等一流院校，大部分曾在微软、阿里巴巴、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。截至2023年6月30日，公司研发人员达到77人，经验丰富的技术团队为本项目的执行提供了人才保证。

（2）技术储备

截至2022年底，公司拥有算法模型框架16个、算法模型数量超过200个，公司自然语言理解算法支持包括语义理解、情感分析和意图识别等能力，语音识别算法支持语种58个，计算机视觉算法支持几十大类、上百小类的物体识别。公司在智能语音、自然语言、计算机视觉领域均有多年算法积累，该等算法模型能够全面支撑公司多个领域数据生产活动的开展。

（3）市场储备

公司自2005年成立以来，始终致力于挖掘行业客户需求，解决客户痛点，通过在智能语音、计算机视觉、自然语言等领域的技术积累，获得全球众多客户认可，截至2022年底，公司累计客户数量已达到810家。公司现有客户包括阿里巴巴、腾讯、百度、科大讯飞、海康威视、字节跳动、微软、亚马逊、三星、中国科学院、清华大学等全球主流企业、教育科研机构以及政企机构。产品广泛应用于客户所研发的个人助手、智能音箱、语音导航、搜索服务、短视频、虚拟人、智能驾驶、机器翻译等多种产品相关的算法模型训练过程中。公司的存量客户与本次募集资金投向新业务领域的客户重合程度较高，本次募集资金投资项目具有稳定的客户资源和市场储备。

（五）公司应对本次向特定对象发行摊薄即期回报采取的措施

为保护广大投资者的合法权益，降低本次向特定对象发行可能摊薄即期回报的影响，公司拟采取多种措施保证本次向特定对象发行募集资金有效使用、有效防范即期回报被摊薄的风险。公司填补即期回报的具体措施如下：

1、强化募集资金管理，提高募集资金使用效率

根据《上市公司监管指引第2号——上市公司募集资金管理和使用的监管要求》《科创板上市规则》等法律法规的要求，结合公司实际情况，公司已制定《募集资金使用管理办法》，明确了公司对募集资金专户存储、使用、用途变更、管理和监督的规定。募集资金将存放于公司董事会决定的专项账户集中管理，做到专款专用，以保证募集资金合理规范使用。

2、积极落实募集资金投资项目，助力公司业务发展

本次募集资金投资项目的实施，将推动公司业务发展，加强公司主营业务的研发能力及生产能力，进一步提高公司市场竞争力，为公司的战略发展带来积极影响，发行募集资金到位后，公司将积极推进募集资金投资项目，从而降低本次发行对股东即期回报摊薄的风险。

3、不断完善公司治理，加强经营管理和内部控制

公司将严格遵循《公司法》《证券法》《科创板上市规则》等法律、法规和规范性文件的要求，不断完善公司治理结构，建立健全公司内部控制制度，促进公司规范运作并不断提高质量，维护公司整体利益，尤其是中小股东的合法权益。同时，公司将努力提高资金的使用效率，合理运用各种融资工具和渠道，控制资金成本，提升资金使用效率，节省公司的各项费用支出，全面有效地控制经营和管控风险，保障公司持续、稳定、健康发展。

4、严格执行利润分配政策，强化投资者回报机制

为完善公司利润分配政策，更好地维护全体股东及投资者利益，公司按照相关法律法规的要求并结合公司的实际情况，在《公司章程》中规定了有关利润分配的相关制度条款，明确了公司现金分红的具体条件和比例、股票股利发放条件

以及公司利润分配的决策程序和机制。为了进一步落实利润分配政策，公司还制定了《未来三年（2023年-2025年）股东分红回报规划》，持续优化投资者回报机制。公司未来将严格执行分红政策，在符合利润分配条件的情况下，积极推动对股东的利润分配，切实维护投资者合法权益，努力提升股东回报水平。

5、完善员工激励机制，加强队伍建设

公司将完善薪酬和激励机制，建立有市场竞争力的薪酬体系，引进市场优秀人才，建立与公司发展相匹配的人才结构，最大限度地激发员工积极性，挖掘员工的创造力和潜在动力，为公司的可持续发展提供可靠的人才保障。

6、持续完善公司治理，为公司发展提供制度保障

公司已建立、健全了法人治理结构，具有完善的股东大会、董事会、监事会和管理层的独立运行机制，设置了与公司生产经营相适应的组织职能机构，并制定了相应的岗位职责，各职能部门之间职责明确，相互制约。公司组织机构设置合理、运行有效，股东大会、董事会、监事会和管理层之间权责分明、相互制衡、运行良好，形成了一套合理、完整、有效的公司治理与经营管理框架。

公司将继续严格遵循《公司法》《证券法》以及《上市公司治理准则》等法律、法规和规范性文件的要求，不断完善公司治理结构，切实保护投资者尤其是中小投资者权益，为公司发展提供制度保障。

公司提示投资者，上述填补回报措施不等于对公司未来利润作出保证。

（六）相关主体关于本次向特定对象发行股票摊薄即期回报采取填补措施的承诺

1、公司董事、高级管理人员对公司填补回报措施能够得到切实履行做出的承诺

1、不无偿或以不公平条件向其他单位或者个人输送利益，也不采用其他方式损害公司利益。

2、对本人的职务消费行为进行约束。

3、不动用公司资产从事与本人履行职责无关的投资、消费活动。

4、由董事会或薪酬与考核委员会制定的薪酬制度与公司填补回报措施的执行情况相挂钩。

5、若公司后续推出股权激励方案，则未来股权激励方案的行权条件与公司填补回报措施的执行情况相挂钩。

6、自本承诺出具日至公司本次向特定对象发行 A 股股票实施完毕前，若中国证券监督管理委员会、上海证券交易所等证券监管机构作出关于填补回报措施及其承诺的其他新的监管规定，且上述承诺相关内容不能满足中国证券监督管理委员会、上海证券交易所等证券监管机构的该等规定时，本人承诺届时将按照中国证券监督管理委员会、上海证券交易所等证券监管机构的最新规定出具补充承诺。

7、本人承诺切实履行公司制定的有关填补回报措施以及对此作出的任何有关填补回报措施的承诺，若违反该等承诺并给公司或者投资者造成损失的，本人愿意依法承担对公司或者投资者的补偿责任。

8、作为填补回报措施相关责任主体之一，若违反上述承诺或拒不履行上述承诺，本人同意按照中国证券监督管理委员会、上海证券交易所等证券监管机构制定或发布的有关规定、规则，对本人作出相关处罚或采取相关管理措施。”

2、公司控股股东、实控人、董事长贺琳对公司填补回报措施能够得到切实履行作出的承诺

“1、不越权干预公司经营管理活动，不侵占公司利益，切实履行公司填补即期回报的相关措施。

2、不无偿或以不公平条件向其他单位或者个人输送利益，也不采用其他方式损害公司利益。

3、对本人的职务消费行为进行约束。

4、不动用公司资产从事与其履行职责无关的投资、消费活动。

5、由董事会或薪酬委员会制定的薪酬制度与公司填补被摊薄即期回报措施的执行情况相挂钩。

6、若公司未来推出股权激励计划，承诺拟公布的公司股权激励计划的行权条件与公司填补被摊薄即期回报措施的执行情况相挂钩。

7、自本承诺出具日至公司本次向特定对象发行 A 股股票实施完毕前，若中国证券监督管理委员会、上海证券交易所等证券监管机构作出关于填补回报措施及其承诺的其他新的监管规定，且上述承诺相关内容不能满足中国证券监督管理委员会、上海证券交易所等证券监管机构的该等规定时，本人承诺届时将按照中国证券监督管理委员会、上海证券交易所等证券监管机构的最新规定出具补充承诺。

8、本人承诺切实履行公司制定的有关填补回报措施以及对此作出的任何有关填补回报措施的承诺，若违反该等承诺并给公司或者投资者造成损失的，本人愿意依法承担对公司或者投资者的补偿责任。

9、作为填补回报措施相关责任主体之一，若违反上述承诺或拒不履行上述承诺，本人同意按照中国证券监督管理委员会、上海证券交易所等证券监管机构制定或发布的有关规定、规则，对本人作出相关处罚或采取相关管理措施。”

八、结论

综上所述，公司本次向特定对象发行股票具备必要性与可行性，发行方案公平、合理，符合相关法律法规的要求，本次向特定对象发行股票方案的实施符合公司发展战略，有利于进一步增强公司持续经营能力，符合公司及全体股东利益。

北京海天瑞声科技股份有限公司

董事会

2023 年 10 月 23 日