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
2023年度向特定对象发行A股股票预案（修订稿）

海天瑞声

（Beijing Haitian Ruisheng Science Technology Ltd.）
（北京市海淀区成府路 28 号 4-801）

二〇二三年十月

公司声明

1、公司及董事会全体成员保证本预案内容真实、准确、完整，并确认不存在虚假记载、误导性陈述或重大遗漏。

2、本预案按照《上市公司证券发行注册管理办法》等法规及规范性文件的要求编制。

3、本次向特定对象发行股票完成后，公司经营与收益的变化由公司自行负责；因本次向特定对象发行股票引致的投资风险，由投资者自行负责。

4、本预案是公司董事会对本次向特定对象发行股票的说明，任何与之相反的声明均属不实陈述。

5、投资者如有任何疑问，应咨询自己的股票经纪人、律师、专业会计师或其他专业顾问。

6、本预案所述事项并不代表审批机关对于本次向特定对象发行股票相关事项的实质性判断、确认、批准或核准，本预案所述本次向特定对象发行股票相关事项的生效和完成尚待上海证券交易所审核通过并经中国证监会作出予以注册决定。

特别提示

本部分所述的词语或简称与本预案“释义”中所定义的词语或简称具有相同的含义。

1、本次发行相关事项已于2023年6月21日经公司第二届董事会第十七次会议、2023年7月7日经公司2023年第一次临时股东大会审议通过，并于2023年10月23日经公司第二届董事会第二十次会议审议修订。根据有关法律法规的规定，本次发行尚需上海证券交易所审核通过并经中国证监会作出予以注册决定后方可实施。

2、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其他符合相关法律、法规规定条件的法人、自然人或其他合格的机构投资者等。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经上海证券交易所审核通过并由中国证监会作出予以注册决定后，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。

3、本次发行的股票数量按照募集资金总额除以发行价格确定，同时本次发行股票数量不超过本次发行前公司总股本60,325,180股的20%即12,065,036股（含本数），最终发行数量上限以中国证监会予以注册决定要求的发行数量上限为准。在前述范围内，最终发行数量将在本次发行由中国证监会作出予以注册决定后，由公司董事会或董事会授权人士根据公司股东大会的授权和发行时的实际情况与本次发行的保荐机构（主承销商）协商确定。

在本次发行的董事会决议公告日至发行日期间，若公司发生送红股、资本公积金转增股本、股权激励、股票回购注销等事项引起公司股份变动，本次发行股票数量的上限将进行相应调整。

若国家法律、法规及规范性文件对本次发行的股票数量有新的规定或中国证监会予以注册的决定要求调整的，则本次发行的股票数量届时相应调整。

释　义

在本预案中，除非文义另有所指，下列简称具有如下含义：

一、基本定义		
公司/本公司/海天瑞声/发行人/上市公司	指	北京海天瑞声科技股份有限公司
本预案	指	北京海天瑞声科技股份有限公司2023年度向特定对象发行A股股票预案
本次向特定对象发行/本次发行	指	北京海天瑞声科技股份有限公司2023年度向特定对象发行A股股票的行为
A股	指	人民币普通股
定价基准日	指	本次发行的发行期首日
股东大会	指	北京海天瑞声科技股份有限公司股东大会
董事会	指	北京海天瑞声科技股份有限公司董事会
监事会	指	北京海天瑞声科技股份有限公司监事会
中国证监会/证监会	指	中国证券监督管理委员会
上交所	指	上海证券交易所
《公司法》	指	《中华人民共和国公司法》
《证券法》	指	《中华人民共和国证券法》
《科创板上市规则》	指	《上海证券交易所科创板股票上市规则》
《注册管理办法》	指	《上市公司证券发行注册管理办法》
《公司章程》	指	《北京海天瑞声科技股份有限公司章程》
元、万元、元/股	指	人民币元、人民币万元、人民币元/股
控股股东、实际控制人	指	贺琳
中端安	指	北京中端安投资中心（有限合伙），为公司股东
二、专业术语		
训练数据	指	通过采集与处理等步骤形成的、结构化的、可供人工智能算法模型训练使用的数据
算法	指	解题方案的准确而完整的描述，是一系列解决问题的清晰指令，算法代表着用系统的方法描述解决问题的策略机制
算力	指	计算能力
人工智能/AI	指	英文为Artificial Intelligence，缩写为AI，是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门技术科学
生成式人工智能/生成式AI	指	生成式人工智能（Generative AI）是使用机器学习和深度学习算法，从现有数据中学习并生成新内容的一种人工智能技术。它可以生成各种形式的数据，如图像、音频、视频和文本，并在虚拟现实、自然语言处理、艺术创作和模拟实验等领域有广泛应用
大模型	指	大模型或是Foundation Model（基础模型），指通过在大规模宽泛的数据上进行训练后能适应一系列下游任务的模型
数据生产垂直大模型	指	海天瑞声数据生产垂直大模型，是公司利用大模型理论和算法，自主打造的面向人工智能数据基础服务的技术底座，实现面向多个下游数据任务的设计与处理等规则及内容的多维生成能力模型
多模态	指	多模态机器学习（Multi-modal Machine Learning）是一种机器学习方法，用于处理多种类型的数据，如图像、文本、音频等。它通过将来自不同模态的数据进行整合和联合建模，以提取和学习不同模态之间的相关信息和特征。多模态数据是指用于多模态机器学习训练使用的数据形态数据类型
深度学习（DL）	指	英文为Deep Learning，缩写为DL，是从机器学习中的人工神经网络发展而来的多层人工神经网络模型的训练方法是一种基于人工神经网络的机器学习，其中多层处理被用来从数据中逐步提取更高层次的特征
人类反馈强化学习	指	英文为Reinforcement Learning from Human Feedback，缩写为RLHF，是一种机器学习方法，结合了强化学习和人类反馈，以优化人工智能系统的行为。这种方法通过引入人类评估者的反馈来指导模型的学习过程，以确保模型的输出更符合人类的期望和价值观
人工智能基础数据、训练数据、人工智能基础数据服务	指	人工智能基础数据、训练数据均指通过采集与处理等步骤形成的、结构化的、可供人工智能算法模型训练使用的数据；人工智能基础数据服务指为AI算法训练及优化提供的数据集设计、数据采集/获取、清洗、标注/优化、评测等服务
ChatGPT	指	美国OpenAI研发的聊天机器人程序，ChatGPT是人工智能技术驱动的自然语言处理工具，它能够通过理解和学习人类的语言来进行对话，还能根据聊天的上下文进行互动，真正像人类一样来回交流，甚至能完成撰写邮件、视频脚本、文案、翻译、代码，写论文等任务
智能语音	指	实现人机语言的通信，包括语音识别技术（ASR）和语音合成技术（TTS）
语音识别（ASR）	指	Automatic Speech Recognition，是一种将语音信号自动转换为文本的技术，是实现智能人机交互的关键技术之一
语音合成（TTS）	指	Text to Speech，即“从文本到语音”，是一种将文字信息转换为自然语音的技术
自然语言处理	指	英文为Natural Language Processing，缩写为NLP，是计算机科学领域与人工智能领域中的一个重要方向，研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法
计算机视觉	指	英文为Computer Vision，缩写为CV，是一门研究如何使机器“看”的科学，是指用摄影机和电脑代替人眼对目标进行识别，跟踪和测量等
原料数据	指	Raw Data，即未经加工处理的原始数据
采集	指	使用软件工具、录音设备、图像拍摄设备等收集、读制和获取原料数据的过程
标注	指	训练数据的生产步骤之一，指对原料数据进行加工，进而生成符合算法模型训练要求的训练数据的过程
GB	指	千兆字节，是数字信息存储容量的一个单位。它被用来衡量文件、文档和数据存储设备的大小。一个千兆字节是1024兆字节
TB	指	太字节，是数字信息存储容量的一个单位。一个太字节是1024个千兆字节
ZB	指	Zettabyte，是数字信息存储容量的一个单位。一个Zetta字节是100万太字节。它代表了一个巨大的数据存储容量，通常在讨论大规模数据中心、全球数据流量或主要组织或行业的存储需求时使用
OpenAI	指	OpenAI是一家美国人工智能研究实验室和公司，在人工智能的各个领域进行前沿研究
OCR	指	是指电子设备（例如扫描仪或数码相机）检查纸上打印的字符，通过检测暗、亮的模式确定其形状，然后用字符识别方法将形状翻译成计算机文字的过程
Prompt	指	提示文本，提供模型完成下游任务所需的上下文等信息，以帮助模型更好地理解任务需求，从而更好地完成下游任务
预训练/Pre-training	指	预训练是通过在大规模数据集上进行初始训练，学习通用特征和表示的机器学习方法。预训练模型可用于后续任务，并在多个领域应用广泛
模型微调/Fine-tune	指	微调，是一种常用于预训练模型的模型优化技术，基于已经训练好的预训练模型进行微调，使模型能够更好地适应目标任务
奖励模型/Reward Model	指	奖励模型（Reward Model）是在强化学习中使用的一种关键概念。它定义了在特定任务中，智能体所追求的目标和奖励的方式。奖励模型通过指定在智能体采取行动后，所获得的奖励信号来引导智能体的学习和决策过程。奖励模型可以基于任务的目标设定，对智能体的行为进行正向奖励或负向惩罚，从而影响智能体的策略优化和决策选择
智算中心	指	人工智能计算中心，指为人工智能（AI）开发和应用提供所需算力服务、数据服务和算法服务的机构
Stable Diffusion	指	图文生成模型，Stability AI公司（英国一家人工智能企业）于2022年提出开发的一个深度学习模型，用于从提示信息的自然语言描述中生成图像

注：本预案中所列数据可能因四舍五入原因，而与根据相关单项数据直接相加之和在尾数上略有差异。

第一节 本次向特定对象发行股票方案概要

一、发行人基本情况

公司名称：	北京海天瑞声科技股份有限公司
英文名称：	Beijing Haitian Ruisheng Science Technology Ltd.
注册地址：	北京市海淀区成府路28号4-801
法定代表人：	贺琳
总股本：	60,325,180股
成立日期：	2005-05-11
电话号码：	86-10-62660772
传真号码：	86-10-62660892
公司网址：	www.datacoesnai.com
电子信箱：	ir@datacoesnai.com
股票简称：	海天瑞声
股票代码：	688787.SH
股票上市交易所：	上海证券交易所
经营范围：	一般项目：技术服务、技术开发、技术咨询、技术交流、技术转让、技术推广；软件开发；软件销售；计算机软硬件及辅助设备零售；计算机软硬件及辅助设备批发；信息处理服务；数据处理；互联网信息服务；（除依法须经批准的项目外，凭营业执照依法自主开展经营活动）许可项目：第二类增值电信业务；互联网信息服务；（依法须经批准的项目，经相关部门批准后方可开展经营活动，具体经营项目以相关部门批准文件或许可证件为准。）（不得从事国家和本市产业政策禁止和限制类项目的经营活动。）

二、本次向特定对象发行的背景和目的

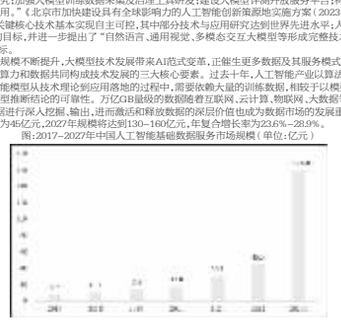
（一）本次向特定对象发行的背景

1、全球人工智能行业快速发展，我国正加速布局和应对

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，提费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重要战略抓手。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，拉开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战略出台了相关指导方案和激励政策，对人工智能的体系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，倒导未来不断开发产业发展新基础。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。

2、人工智能基础数据服务应用有其特殊性，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，相应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理能力和泛化能力，并在多个领域应用广泛。目前，以互联网为代表的数字原生企业，在长期的业务发展过程中，通过积累海量的用户数据，以及高质量数据为中心的训练方法能够提升模型推理能力和泛化能力，并在多个领域应用广泛。

国内缺乏高质量数据的主要原因包括当前国内数据控制权的归属不清、产权保护不足、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）本次向特定对象发行的目的

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。

三、本次向特定对象发行股票的发行对象及其与公司的关系

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其他符合相关规定的法人、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作出予以注册决定后，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。

（四）定价基准日、定价原则和发行价格

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，P0为调整前发行底价，P1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予以注册决定后，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。

（五）发行数量

本次发行的股票数量按照募集资金总额除以发行价格确定，同时本次发行股票数量不超过本次发行前公司总股本60,325,180股的20%即12,065,036股（含本数），最终发行数量上限以中国证监会予以注册决定要求的发行数量上限为准。在前述范围内，最终发行数量将在本次发行由中国证监会作出予以注册决定后，由公司董事会或董事会授权人士根据公司股东大会的授权和发行时的实际情况与本次发行的保荐机构（主承销商）协商确定。

在本次发行的董事会决议公告日至发行日期间，若公司发生送红股、资本公积金转增股本、股权激励、股票回购注销等事项引起公司股份变动，本次发行股票数量的上限将进行相应调整。

若国家法律、法规及规范性文件对本次发行的股票数量有新的规定或中国证监会予以注册的决定要求调整的，则本次发行的股票数量届时相应调整。

（六）限售期

本次发行的发行对象认购的股份自发行结束之日起6个月内不得转让。法律法规、规范性文件对限售期另有规定的，依其规定。

本次发行结束后，发行对象取得的本次发行的股份由于公司送红股、资本公积金转增股本等原因所衍生取得的公司股份，亦应遵守上述限售期安排。限售期结束后发行对象减持认购的本次发行的股票将按届时有效的中国证监会及上交所的有关规定执行。

（七）募集资金金额和用途

本次发行股票募集资金总额不超过66,600.00万元（含本数），扣除相关发行费用后将用于投资以下项目：

单位：万元

序号	项目名称	项目投资总额	拟投入募集资金金额
1	AI大模型训练数据集建设项目	38,337.36	35,507.25
2	数据生产垂直大模型研发项目	40,651.64	31,092.75
合计		78,989.00	66,600.00

在本次发行募集资金到位前，公司将根据募集资金投资项目的实际情况，以自筹资金先行投入，并在募集资金到位后按照相关法律、法规规定的程序予以置换。募集资金到位后，若扣除发行费用后的实际募集资金净额少于拟投入募集资金总额，在本次发行募集资金投资项目范围内，公司将根据实际募集资金数额，按照项目的轻重缓急等情况，调整并最终决定募集资金的具体投资项目、优先顺序及各项目的具体投资金额，募集资金不足部分由公司自筹解决。

若本次发行募集资金总额因监管政策变化或发行注册文件的要求予以调整的，则届时将相应调整。

（八）本次发行前滚存未分配利润的安排

本次发行前公司滚存的未分配利润，由本次向特定对象发行股票完成后的新老股东按照本次发行后的股份比例共享。

（九）上市地点

本次发行的股票将在上交所科创板上市。

（十）决议的有效期

本次发行决议的有效期为自公司股东大会审议通过之日起十二个月。

五、本次发行是否构成关联交易

截至本预案出具日，本次发行尚未确定具体发行对象，最终是否存在因关联方认购公司本次向特定对象发行股票构成关联交易的情形，将在发行结果公告的发行情况报告书中披露。

六、本次发行是否导致公司控制权发生变化

本次发行前，公司控股股东贺琳控制大学城，贺琳直接持有公司12,137,615股股份（占总股本的20.12%），并通过控制中端安间接控制公司6,935,779股股份（占总股本的11.50%），合计控制公司31.62%的股份，为公司实际控制人。若假设本次发行股票数量为发行上限12,065,036股，则本次发行完成后（仅考虑本次发行导致的公司股份数量变化），公司的总股本为72,390,216股，贺琳、中端安分别占公司股本的16.77%、9.58%，贺琳合计控制公司26.35%的股份，仍为公司的控股股东、实际控制人。因此，本次发行不会导致公司的控制权发生变化。

七、本次向特定对象发行股票的实施是否可能导致股权分布不具备上市条件

本次向特定对象发行股票方案的实施不会导致股权分布不具备上市条件。

八、本次发行方案已经取得有关主管部门批准的情况以及尚需报批的程序

本次向特定对象发行股票相关事项已于2023年6月21日经公司第二届董事会第十七次会议、2023年7月7日经公司2023年第一次临时股东大会审议通过，并于2023年10月23日经公司第二届董事会第二十次会议审议修订。尚需履行以下审批：

1、本次向特定对象发行尚待上交所审核通过；

2、本次向特定对象发行尚待中国证监会作出予以注册决定。

第二节 董事会关于本次募集资金使用的可行性分析

一、本次向特定对象发行募集资金的使用计划

本次发行募集资金总额不超过66,600.00万元（含本数），扣除相关发行费用后将用于投资以下项目：

单位：万元

序号	项目名称	项目投资总额	拟投入募集资金金额
1	AI大模型训练数据集建设项目	38,337.36	35,507.25
2	数据生产垂直大模型研发项目	40,651.64	31,092.75
合计		78,989.00	66,600.00

在本次发行募集资金到位前，公司将根据募集资金投资项目的实际情况，以自筹资金先行投入，并在募集资金到位后按照相关法律、法规规定的程序予以置换。募集资金到位后，若扣除发行费用后的实际募集资金净额少于拟投入募集资金总额，在本次发行募集资金投资项目范围内，公司将根据实际募集资金数额，按照项目的轻重缓急等情况，调整并最终决定募集资金的具体投资项目、优先顺序及各项目的具体投资金额，募集资金不足部分由公司自筹解决。

二、本次募集资金投资项目的可行性分析

（一）项目概况

1、AI大模型训练数据集建设项目

项目的实施主体为北京海天瑞声科技股份有限公司。鉴于大模型训练数据通常具备数据规模大、数据质量高、数据类型丰富等特点，本项目拟建设AI大模型训练数据集，面向产用于通用型、及各种垂直领域大模型训练的海量、高品质数据集。本项目拟购置办公楼作为建设大模型训练数据研发生产基地，并购置数据采集、数据处理、数据存储和办公等软硬件设备，利用海量、高质量、多样化的公共数据资源、社会数据资源和第三方数据资源，通过数据采集/获取、清洗/分类/标准化、标注/优化、评测等全流程的任务执行进行高质量大模型训练数据集建设。

本项目将充分利用“先行先试示范区”在基础制度、数据供给等方面的先行先试政策，采用多元化的方式获取大规模原始数据；利用工程化的数据处理技术进行预训练阶段的数据清洗；采用人类反馈强化学习模式，基于微调和奖励模型训练的方法，以人类辨识少量的典型问题和标准答案与深度学习阶段基础性标注相结合的模式，生产出市场适用性较强的大模型训练数据集。

本项目建成后，将提供可供大模型训练和评测的不少于10个品类的专业数据集，显著提升行业内面向大模型训练数据集的类别和质量，协助实现公共数据、社会数据等各类高价值数据资源汇聚，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。

三、本次向特定对象发行股票的发行对象及其与公司的关系

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其他符合相关规定的法人、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作出予以注册决定后，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。

四、本次向特定对象发行股票方案概要

（一）发行股票的种类和面值

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）发行方式和发行时间

本次发行全部采用向特定对象发行A股股票的方式进行，公司将在通过上交所审核，并经中国证监会作出予以注册决定后的有效期内，选择适当时机向特定对象发行。

（三）发行对象和认购方式

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其他符合相关规定的法人、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作出予以注册决定后，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。

（四）定价基准日、定价原则和发行价格

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，P0为调整前发行底价，P1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予以注册决定后，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。

（五）发行数量

本次发行的股票数量按照募集资金总额除以发行价格确定，同时本次发行股票数量不超过本次发行前公司总股本60,325,180股的20%即12,065,036股（含本数），最终发行数量上限以中国证监会予以注册决定要求的发行数量上限为准。在前述范围内，最终发行数量将在本次发行由中国证监会作出予以注册决定后，由公司董事会或董事会授权人士根据公司股东大会的授权和发行时的实际情况与本次发行的保荐机构（主承销商）协商确定。

在本次发行的董事会决议公告日至发行日期间，若公司发生送红股、资本公积金转增股本、股权激励、股票回购注销等事项引起公司股份变动，本次发行股票数量的上限将进行相应调整。

若国家法律、法规及规范性文件对本次发行的股票数量有新的规定或中国证监会予以注册的决定要求调整的，则本次发行的股票数量届时相应调整。

（六）募集资金金额和用途

本次发行股票募集资金总额不超过66,600.00万元（含本数），扣除相关发行费用后将用于投资以下项目：

单位：万元

序号	项目名称	项目投资总额	拟投入募集资金金额
1	AI大模型训练数据集建设项目	38,337.36	35,507.25
2	数据生产垂直大模型研发项目	40,651.64	31,092.75
合计		78,989.00	66,600.00

在本次发行募集资金到位前，公司将根据募集资金投资项目的实际情况，以自筹资金先行投入，并在募集资金到位后按照相关法律、法规规定的程序予以置换。募集资金到位后，若扣除发行费用后的实际募集资金净额少于拟投入募集资金总额，在本次发行募集资金投资项目范围内，公司将根据实际募集资金数额，按照项目的轻重缓急等情况，调整并最终决定募集资金的具体投资项目、优先顺序及各项目的具体投资金额，募集资金不足部分由公司自筹解决。

（七）募集资金用途

本次发行募集资金总额不超过66,600.00万元（含本数），扣除相关发行费用后将用于投资以下项目：

单位：万元

序号	项目名称	项目投资总额	拟投入募集资金金额
1	AI大模型训练数据集建设项目	38,337.36	35,507.25
2	数据生产垂直大模型研发项目	40,651.64	31,092.75
合计		78,989.00	66,600.00

在本次发行募集资金到位前，公司将根据募集资金投资项目的实际情况，以自筹资金先行投入，并在募集资金到位后按照相关法律、法规规定的程序予以置换。募集资金到位后，若扣除发行费用后的实际募集资金净额少于拟投入募集资金总额，在本次发行募集资金投资项目范围内，公司将根据实际募集资金数额，按照项目的轻重缓急等情况，调整并最终决定募集资金的具体投资项目、优先顺序及各项目的具体投资金额，募集资金不足部分由公司自筹解决。

（八）本次发行完成后，公司注册资本、股本总额将相应增加，公司将依法根据发行情况对《公司章程》中有关公司股本等相关条款进行相应调整，并办理工商变更登记。

（九）本次发行对股东结构的影响

本次发行前，公司控股股东贺琳控制大学城，贺琳直接持有公司12,137,615股股份（占总股本的20.12%），并通过控制中端安间接控制公司6,935,779股股份（占总股本的11.50%），合计控制公司31.62%的股份，为公司实际控制人。

若假设本次发行股票数量为发行上限12,065,036股，则本次发行完成后（仅考虑本次发行导致的公司股份数量变化），公司的总股本为72,390,216股，贺琳、中端安分别占公司股本的16.77%、9.58%，贺琳合计控制公司26.35%的股份，仍为公司的控股股东、实际控制人。因此，本次发行不会导致公司的控制权发生变化。

（十）本次发行对董事、监事、高级管理人员结构的影响

本次发行不对公司董事、监事、高级管理人员结构的重大变动情况。

截至本预案出具日，公司尚无对高级管理人员结构进行调整的计划。若公司拟调整高级管理人员结构，将根据有关规定，履行必要的法律程序和信息披露义务。

（十一）本次发行对业务结构的影响

本次向特定对象发行股票的募集资金投资的项目围绕公司主营业务开展，系对公司主营业务的拓展和完善，是公司完善产业布局的重要举措。本次发行完成后公司的业务结构不会发生重大变化。

二、本次发行完成后公司财务状况、盈利能力及现金流量的变动情况

（一）对财务状况的影响

本次发行完成后，公司的资本实力将进一步增强，总资产和净资产规模均有所增长，资产负债率将有所下降，营运资金得到进一步充实，有助于优化公司资产负债结构，提高公司抵御财务风险的能力，为公司后续发展提供良好保障。

（二）对盈利能力的影响

本次发行完成后，公司总股本将有所增加，由于募集资金投资项目从实施到经营效益完全释放需要一定时间，存在股权收益等指标在短期内被摊薄的风险。为保障中小投资者的利益，公司就本次向特定对象发行事项对即期回报摊薄的影响进行了认真分析，并制定填补被摊薄即期回报的具体措施。相关情况详见《海天瑞声关于2023年度向特定对象发行A股股票摊薄即期回报与填补措施及相关主体承诺（修订稿）的公告》。

本次募集资金是公司在现有主营业务基础上，结合市场需求和未来发展趋势，加大对公司核心主业重点产品及重要研究方向投