

证券代码：002230

证券简称：科大讯飞

科大讯飞股份有限公司投资者关系活动记录表

编号：20250213

投资者关系活动类别	<input type="checkbox"/> 特定对象调研 <input type="checkbox"/> 分析师会议 <input type="checkbox"/> 媒体采访 <input type="checkbox"/> 业绩说明会 <input type="checkbox"/> 新闻发布会 <input type="checkbox"/> 路演活动 <input type="checkbox"/> 现场参观 <input checked="" type="checkbox"/> 其他 民生证券：AI 尖峰时刻看龙头 科大讯飞电话会
参与单位名称及人员姓名	博时基金 施永辉、博时基金 谢泽林、交银施罗德基金 柳轩、交银施罗德基金 朱维缜、中银基金 蔡国栋、中银基金 杨成、光大保德信 陈卓钰、光大保德信 安鹏、光大保德信基金 林晓枫、摩根士丹利 雷志勇、华夏基金 施知序、富国基金 蒲梦洁、中国人民养老保险 孙楠、东方红资产 郑华航、银华基金 方建、鹏华基金 杨飞、永赢基金 任桀、新华资产 陈朝阳、中信建投 梁斌、华安基金 刘潇、东兴基金 高怡宁、兴业基金 张超、长盛基金 郭堃、东兴基金 周昊、民生证券 胡又文、长江证券 王旺、长江证券 李昊明、长江证券 卞曙光、浙商证券 许运凯、上海海宸投资 金朝晖、上海肇万资产 崔磊、安联保险 徐彪、乾璐投资 陈少楠、金科投资 刘宇、上海健顺投资 邵伟、北京暖逸欣私募基金 王涛、国寿安保 谢夫、国泰基金 邵雪丹、北大方正 陈发、九泰基金 赵万隆、华英证券 池诚、长江证券 王毅、中银国际证券 陆莎莎、中信资管 宋文开、上海肇万资产 崔磊、金科投资 高洁、上海复胜资产 赵超、海金(大连)投资 黄樊林、峰岚 蔡荣转、明世伙伴私募基金 孙勇、进门财经 钟嘉敏
时间	2025.02.13
地点及方式	线上电话会
上市公司接待人员姓名	科大讯飞董事、副总裁、董事会秘书江涛

问答交流纪要

问：近来，DeepSeek 引发了国内外各界极高的关注和热议，怎样看待 DeepSeek 的成功？

答：深度求索公司近期发布的 DeepSeek V3 和 DeepSeek R1 大模型，围绕高效推理和经济训练的创新算法，以较低成本实现了大语言模型的性能效果突破。DeepSeek 团队是梁文锋先生在前期商业化成功的基础上（DeepSeek 系源于运用数学与人工智能进行量化投资的对冲基金公司——幻方量化，管理规模 600 亿），依托国内顶级高校年轻人打造的一个有技术信仰、有现金支撑、人才密度极高的团队。DeepSeek 用实实在在的创新成果和技术突破，打破了西方世界对中国科技创新路径的刻板印象，也证明了今天的中国是完全有可能成为顶尖人才成长的沃土、原创性创新的策源地，提振了中国科技创新的信心。

问：DeepSeek 在 2025 年 1 月 20 日发布的 R1 快速对标 OpenAI 的 o1 大模型，公司在对应的深度推理技术创新上目前处于什么阶段？

答：1、DeepSeek 模型从基本理论框架来看，还是 Transformer 和强化学习等深度神经网络理论，训练算力设施还是基于英伟达的 A800 和 H800，但其在工程优化上实现了出色的率先突破。

2、DeepSeek R1 快速对标 o1 的主要创新是其提出并实现的 R1-Zero 强化学习训练方案，即直接在 DeepSeek-V3-base 预训练模型上进行大规模强化学习训练，可在数学、代码等推理任务上得到接近 o1 模型的效果，而不依赖任何有监督微调（SFT），这是当前国内外公开已知的、在大规模任务上获得成功效果的

首个发布成果，之前学术界已有相关研究，但因为预训练基座模型不够好而使得强化学习的收益不好，这一成果具有较强的创新性，大大减少了数学、代码等任务训练对人工标注推理过程数据的需求。

3、DeepSeek R1 并不是唯一高水平的国产深度推理大模型。2024 年 9 月 OpenAI o1-preview 发布以后，国内技术领先的大模型厂商也在快速跟进。例如，科大讯飞基于此前在评语模型、强化学习方面的积累，在 2024 年 10 月就实现了蒙特卡洛树搜索（MCTS）等算法创新并跑通了深度推理模型训练算法，因为只能使用国产算力，我们额外花费 2 个月完成华为昇腾 910B 算力的适配和优化并大幅提升训练效率（训练效率相比 A100 从 55% 提升到 85%），于今年 1 月 15 日完成了星火深度推理模型 X1 的训练和发布。虽然当前 X1 模型很小（仅 130 亿参数），但依靠我们算法和数据优势，星火 X1 已达到可以和 OpenAI o1-preview 对标的水平；随着国产算力的逐步到位，模型参数增加后，我们完全有信心实现数学答题和过程思维链能力全面对标甚至超过 OpenAI o1。此外，结合与中国教科院等诸多教育专家的合作成果，我们将使得 X1 的思维链结果更加符合教育教学的专业要求，在教育、医疗等领域实现更大的应用赋能价值。

问：DeepSeek 实现了大模型训练和推理成本的大幅降低，讯飞星火在大模型训练和推理的软硬件工程化方面有什么样的积累，如何实现对成本的有效控制？

答：1、DeepSeek 团队围绕 DeepSeek V3 和 R1 的训练和推理降本，实现了多个有效的软硬件结合深度工程优化创新，包括 MLA（通过低秩矩阵对 KV 矩阵进行压缩显著降低 GPU 显存的消耗）、MTP（通过多帧预测提升数据利用率及训练效果和推理效率）、DualPipe（通过更底层的 PTX 指令优化英伟达 GPU

计算和通信的并行)、FP8 混合精度训练(业界首次使用 8Bit 量化浮点数方式实现超大规模模型的训练,以往都是使用 16Bit 量化),最终 DeepSeek 基于 2048 张英伟达 H800 卡,约 560 万美元就能完成 V3 模型的一次完整预训练。另外, DeepSeek R1 因为涉及基于大量数据的强化学习迭代训练,其成本未公布。

2、和 DeepSeek 类似,科大讯飞受限于有限的算力资源,过去几年关于星火大模型训练和推理成本效率的持续优化也做了大量投入,和 DeepSeek 直接使用英伟达 H800 卡上开展的各种工程优化不同,科大讯飞选择了更难的全国产算力路线。从 2023 年 5 月起科大讯飞就联合华为先后攻克了万卡高速互联组网、计算通信隐藏、训练推理强交互、高吞吐推理优化以及国产算子优化等一系列难题,将通用大模型、类 o1 的深度推理模型等的训练效率对标 A100 均从最初的 30%-50% 优化达到了 85%-95% 以上;以 DeepSeek 非常看重的万卡网络通信带宽的利用率为例,DeepSeek 报道的是 93%,而讯飞星火做到了 95%。训练成本也是讯飞星火大模型的竞争优势,我们单次训练成本和 DeepSeek V3 在一个层次,而且后面优化路径也很明确,目前以昇腾 910B 为代表的国产硬件还不支持 FP8 精度训练以及 DualPipe 的实现,随着国产算力在底层能力上进一步提升,讯飞星火的训练成本还有较大的下降空间。

3、需要特别说明的是,大模型对算力的需求为训练和推理两个方面,而训练实现的难度远大于推理。这就是虽然陆续有公司宣布可以在国产算力平台上提供 DeepSeek 在内等大模型的推理服务,但仍只有讯飞星火一家是训练和推理均在国产算力上进行的,并且仅用 1 万张 910B 国产算力卡,科大讯飞不仅用比友商少一个数量级的算力取得了大模型研发上的第一梯队成果,而且做了大量国产

算力平台上的无人区的适配和效率优化，追求国产算力的极致效率，这比在英伟达上的实现更需要技术实力和战略勇气。

问：DeepSeek 的火爆出圈，对公司相关的大模型业务开展会有哪些影响？

答:1、DeepSeek 火爆出圈在全球范围内是一次深刻的人工智能用户普及教育：DeepSeek 引发了全球媒体的广泛报道，从科技媒体到主流媒体，都对 DeepSeek 进行了介绍和分析，使人工智能的话题得到更广泛的传播，提升了公众对人工智能的整体认知度，让大量原本对人工智能了解有限的人开始接触和使用 AI 产品。同时，DeepSeek 的火爆吸引了大量云服务厂商、算力企业、芯片厂商、应用端企业等产业链上下游积极参与，中国人工智能产业生态建设的速度进一步加快，人工智能的赋能价值和产业规模加速发展，对科大讯飞这样长期投入培育人工智能市场的企业有很大的综合好处。

另外，DeepSeek 在英伟达卡上的工程化探索，对我们进一步提升国产算力基础上大模型性价比有很好的借鉴作用；DeepSeek 验证的长思维链可读化输出，有助于整个产业链降低在数据推理路径标注的投入，未来大模型在需要复杂推理的应用落地过程会进一步加快。当前，行业众多“深水区”问题及需求已经来到了突破解决的前夜，大模型技术阶跃发展是最关键和明确的曙光。讯飞星火会坚定地走全国产化路线和面向行业刚需落地，底座大模型会持续对标国内外最高水平。

2、Deepseek 的火爆，让更多的行业客户对大模型技术和应用有了更加广泛的关注和期待。就具体业务来看，通过数据飞轮驱动的知识积累和场景创新的业

务领域，如教育、医疗等领域，人工智能应用的规模化落地会进一步加快，讯飞在这些领域拥有综合的先发优势会进一步得到发挥。

讯飞星火当前在教育、医疗等行业都是市场占有率第一的大模型。鉴于讯飞底座大模型的技术实力、“用户—场景数据—商业模式”应用数据飞轮和行业理解优势，未来我们会形成更稳固的行业领先优势。

以教育为例，DeepSeek R1 作为慢思考类深度推理模型，适合解决数学、代码等问题，但在面向教育等行业场景的深度应用方面，光靠基础模型还不够，还需要行业知识 RAG、场景定制优化工具链、多模态能力等配套组件及解决方案才能在具体场景中端到端应用。这些方面讯飞的场景和专业积累会更加明显。我们在业界公开算法基础上又进一步实现了思维链数据合成、过程奖励模型等方面的新技术创新，正在训练的星火 X1 新版本预计在 3 月内完成，我们很有信心实现数学答题和过程思维链能力全面对标甚至超过世界一流同行。此外，我们正在将 X1 深度推理模型进一步结合与中国教科院等诸多教育专家的合作，通过专业知识反思及长思维链强化学习结合，生成符合教育教学需求的“教学思维链”，提高复杂场景推理的逻辑正确性、专业性、可解释性，并将率先用于教师助手、批改、辅学等全系产品创新，预计在 2025 世界数字教育大会期间发布教育专属大模型和创新应用，引领中国教育数字化和智能化方向，为世界打造生成式人工智能教育应用示范样板。

3、DeepSeek 的火爆对讯飞在 G 端和 B 端业务上带来的影响，相比教育医疗会更加复杂一些。部分 G 端客户当前处于对 DeepSeek 的技术创新和产业影响处于关注和学习阶段，短期内对正在实施的智算中心的建设节奏有一定的滞后影

响，但从长期来说，各地政府对国产智算中心等新型基础设施建设、产业生态体系完善、拓展 AI+新型工业化应用场景、加快政府的数智化治理转型等方面的投入会进一步加大，这些对讯飞的 G 端业务都会带来好处。在以央国企为代表的 B 端企业市场，DeepSeek 的开源免费，一些企业在 AI 落地项目中大模型底座建设资金投入比例上会更加谨慎，但越来越多的企业在 AI 应用和构建行业专属大模型需求的激活后，讯飞在人工智能企业应用落地中的自主可控优势、专属模型训练工具链、场景落地能力的优势，有可能会使讯飞在央国企业务合作收入的规模增长上形成更大的促进。

总体来说，DeepSeek 的出现促进了市场对大模型应用的热度，为讯飞等已经有提前布局的企业带来了市场加快成熟的新机会。我们更加有信心依托纯国产化的星火大模型底座竞争力、数据飞轮和场景创新，能够在大型模型的应用上持续深化护城河和差异化竞争优势。

问：DeepSeek 开源后，有一些大模型创业公司会退出基座大模型的研发，讯飞未来在底座大模型的研发上会采取什么样的长期策略？

答：1、Deepseek 引发的全球关注，社会各界对通用人工智能有了更大的热情和期待，通用人工智能将更快迎来新的高潮，特别是在中国。大模型国产化龙头企业将优先受益。DeepSeek 的效应，将吸引以华为为代表（还包括寒武纪、海光、摩尔线程等）的国产算力厂商会投入更大资源支持国产大模型的适配，有利于国内人工智能的产业链更好发展；同时，DeepSeek 的效应，将吸引更多产业资源（如国产算力厂商的适配支持）、政府资源的支持的同时，一批原先瞄准通用大模型为目标的创业公司和互联网公司会因为资源无法支撑而退出角逐，转而做

特定方向的人工智能应用落地,这有助于我们国内人工智能产业资源的更合理分布。当前,科大讯飞已取得扎实的落地应用,已经与中国石油、中国移动、中国海油等央国企及 20 多个行业的头部企业一起发布行业大模型,覆盖 300 多个应用场景。此外,科大讯飞 2024 年取得“央国企大模型第一”的落地成绩,在中标项目数量和金额上都实现断层领先,在人工智能企业端应用方面也形成了领先的布局,教育医疗等行业的人工智能应用数据飞轮正在逐步加强。科大讯飞未来会继续保持“1+N”的战略投入定力即“1 个底座大模型+N 个行业大模型”的整体布局,底座大模型持续对标国内外最高水平,同时围绕赋能教育、医疗、金融、汽车、能源、司法、安全、工业互联网及 AI for Science 等关键领域构建 N 个重点行业大模型,积极追求用最先进和安全的人工智能赋能国计民生和国家安全领域。

2、DeepSeek 引发美国对中国人工智能发展高度关注之后,美国政府极大概率会对我国的英伟达算力获取和国产算力海外投片采取更严苛的限制措施,讯飞所构建的大模型全栈国产化技术更有战略意义和先发优势。大模型算力需求训练和推理两个方面对比,训练实现的难度远大于推理。当前所有全民下载的通用大模型 APP 中,只有讯飞星火一家是训练和推理均在国产算力上进行的,仅用 1 万张 910B 国产算力卡,科大讯飞不仅用比友商少一个量级的算力取得了大模型研发上的一系列成果,而且做了大量国产算力平台上的无人区的适配和效率优化。虽然陆续有公司宣布可以在国产算力平台上提供 DeepSeek 在内等大模型的推理服务,但目前仍只有讯飞星火一家是训练和推理均在国产算力上进行的。

3、DeepSeek 的火爆进一步表明当前通用人工智能前所未有的创新活力和发

展潜力，通用大模型的天花板仍然在被持续打开。通用人工智能是一场需要聚精会神的长跑，行稳，方可致远，我们会坚定不移的在大模型底座研发上持续投入并且不断输出最新成果。

科大讯飞正在训练的纯国产算力的星火 X1 新版本预计在 3 月内完成，预期就可以实现数学答题和过程思维链能力全面对标甚至超过 OpenAI o1，随着国家对人工智能自主可控算力的支持不断到位，我们对讯飞星火在通用人工智能的长跑中，有强大的技术后劲，能够不断对标世界一流水平甚至局部引领有充分信心。

问：讯飞开放平台 2 月 10 日宣布上线 DeepSeek 全线产品，公司对此有怎样的考虑？

答：讯飞开放平台始终秉持开放共享的理念，与国内外优秀 AI 厂商及研究机构形成同盟，核心组件的星辰 MaaS 平台积极拥抱开源生态，持续推进 AI 生态的合作与开放。自 2018 年起，讯飞开放平台已接入商汤、合合等 10 多家 AI 合作厂商，共计 98 项能力，平台总能力已达 806 项。讯飞开放平台 2 月 10 日宣布包括“满血版” DeepSeek-V3 和 DeepSeek-R1 在内的 DeepSeek 全系大模型正式上线，支持公有云 API 调用和专属模型一键部署。

而讯飞星火（即讯飞星火大模型）的战略生态位非常清晰：聚焦自主可控的底座能力和面向民生刚需以及赋能科研与生产的重点应用赛道，以长期主义精神保持战略定力，做全栈自主可控的大模型国家队。

日期

2025.02.13