

证券代码： 300418

证券简称： 昆仑万维

昆仑万维科技股份有限公司投资者关系活动记录表

编号： 2023-005

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 checked="" 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 type="checkbox"/> 其他（请文字说明其他活动内容）
参与单位名称及人员姓名	来自东吴证券、海通证券、招商证券、华泰证券、广发证券、兴业证券、民生证券、中信建投、国泰君安、安信证券、首创证券、信达证券、中泰证券、开源证券、国信证券、高盛、银华基金、诺安基金、中信保诚、中邮创业、太平洋人寿、工银国际、荷宝投资等 84 位机构投资者
时间	2023 年 10 月 30 日 16:00-17:00
地点	电话会议
上市公司接待人员姓名	昆仑万维 董事长 方汉先生 昆仑万维 战略副总裁 牟竹云先生
投资者关系活动主要内容介绍	<p>方总回顾公司 2023 年前三季度整体情况：</p> <p>2023 年前三季度，我们坚定不移地执行“AI in” AGI 与 AIGC 战略。公司实现营业收入的稳健增长，营收达到 36.8 亿元，同比增长 8%；实现经营性现金流 7.6 亿元，同比增长 33%。同时，公司的海外业务收入占比进一步提升至 84%；整体毛利率达 80%，继续保持在较高水平。</p> <p>我们深知研发是生成式 AI 行业的立足之本。今年前三季度，公司不断加大研发投入，研发费用增长至 6.2 亿元，同比增长 28%。硬件侧，截至本季度末，采购及租赁芯片已到货约 6,000 张，另外还有约 3,000 张芯片待交付。同时，为完成“算</p>

力基础设施—大模型算法—AI 应用”全产业链布局，公司宣布通过增资方式控股 AI 算力芯片企业北京艾捷科芯科技有限公司。艾捷科芯拟开发一款可编程、具有高性能的 NPU 产品，同时应用于模型训练及推理。人才侧，计算机视觉和机器学习领域的国际顶级专家颜水成教授于今年 9 月 1 日宣布加盟昆仑万维，与昆仑万维创始人周亚辉一起出任天工智能联席 CEO，并兼任昆仑万维 2050 全球研究院院长。同时，昆仑万维将在新加坡、伦敦、硅谷三地建立 2050 全球研究院研究中心，针对人工智能、生物智能等前沿技术领域展开研究。

我们的“天工”大模型也取得了亮眼的成绩。2023 年 9 月，“天工”大模型凭借卓越的产品技术创新荣登 2023 年《财富》中国最佳设计榜；同月，我们通过了中国信通院的可信 AI 评估，并被评选为大模型工作组副组长单位。10 月，我们凭借在 AI 领域卓越的创新力入选 2023 福布斯中国 AI 算法创新应用企业 TOP 30。从大模型的具体能力表现来看，在权威推理榜单 Benchmark GSM8K 测试中，“天工”以 80%的正确率脱颖而出，大幅领先 GPT-3.5(57.1%)和 LLaMA2-70B(56.8%)，推理能力达到全球领先水平；文本理解方面，在基于 SuperCLUE 综合性测评基准开展的测试中，天工大模型文理兼顾，跻身中文闭源模型第一梯队；多模态方面，在腾讯优图实验室联合厦门大学开展的多模态大语言模型测评中，“天工”综合得分排名第一。

我们抓住了技术突破带来的新机遇，在应用端进行了全面布局。我们对“天工”APP 进行了全面的升级迭代，整合 AI 搜索、AI 阅读、AI 创作等核心功能，覆盖工作、学习、生活等众多应用场景。AI 搜索方面，我们将传统搜索与 ChatGPT 式的问答结合在一起，帮助用户获取更高质量的信息，提升学

习工作效率和生产力；AI 阅读方面，“天工”APP 可以通过 AI 一键生成文章的摘要，帮助用户快速了解文章主旨、重点和关键细节。AI 创作方面，“天工”APP 满足学术教育、职场文档、创意写作、广告营销等不同场景需求。

AI 游戏方面，公司首款 AI 游戏《Club Koala》于 8 月 25 日亮相德国科隆国际游戏展。考拉俱乐部搭载的 UGC 游戏编辑器 Koala Editor 帮助用户实现零代码游戏开发，大幅降低游戏创作门槛。同时 Club Koala 还引入了拥有自我意识及记忆能力的 AI NPC，为玩家提供更具沉浸感的游戏体验。未来，我们还将推出全新的 AI 动漫、AI 社交以及 AI 音乐的产品。

同时，海外信息分发及元宇宙平台等业务全面升级 AI 功能。Opera 原生浏览器 AI 助手 Aria 推出三大细节功能，分别为“Refiner”“Compose”和“My Style”。用户无需自行输入提示词指令，即可得到优化后的内容。Aria 在 iOS 终端正式上线，这标志着 Aria 现已覆盖包括 Mac、Windows、Linux、Android 及 iOS 所有主要平台。公司元宇宙入口 Opera GX 也已全面接入 Aria，为广大玩家带来最前沿的 AI 使用体验。

此外，为更好地聚焦于 AGI 与 AIGC 业务的发展，我们将剥离所持绿钒新能源权益。未来除昆仑基金少数股权财务投资外，不再进行任何人工智能上下游行业以外的其他行业的控股型或长期投资。展望未来，我们将专注公司核心业务，致力于成为一家全球领先的人工智能科技企业。

问答环节

问：公司 C 端和 B 端应用当前的进展及未来的展望，包括年内的重要节点及明年的主要期望如何？

答：公司布局 AI 大模型、AI 搜索、AI 游戏、AI 社交、AI 动漫及 AI 音乐等六大方向，目前 AI 大模型、AI 搜索及 AI 游戏已经逐步落地。

大模型方面，天工大模型保持高频迭代并取得重要突破。“天工”的逻辑推理能力、文本理解能力、多模态能力在多个全球知名测评集中表现突出。逻辑推理方面，在权威推理榜单 Benchmark GSM8K 测试中，“天工”以 80%的正确率大幅领先 GPT-3.5（57.1%）和 LLaMA2-70B（56.8%），推理能力达到全球领先水平；多模态方面，在腾讯优图实验室联合厦门大学开展的多模态大语言模型测评中，“天工”综合得分排名第一，而 LLaMA 的多模态大语言模型排名第 12。今天，我们也正式宣布开源天工 Skywork-13B 系列，并配套开源了 600GB、150B Tokens 的超大高质量中文数据集。Skywork-13B 系列在多个权威评测与基准测试上都展现出同等规模模型的最佳效果，其中中文能力尤为出色。具体来看，在主流英文能力评测集 MMLU 和中文能力测评集 C-Eval 测试中 Skywork-13B 系列大模型正确率均超过 60%，远超 LLaMa-2 13B 模型的 36.6%和 36.5%；Skywork-13B 在中文、科技、金融、政务等领域表现均高于其他开源模型。

AI 搜索方面，公司 8 月 23 日正式发布中国第一款 AI 搜索产品“天工 AI 搜索”。该产品既有 ChatGPT 式的问答交互，又可以像普通搜索引擎列出相关链接。弥补了之前 ChatGPT 不能联网、结果难以考证、利用数据库训练信息更新慢等问题，有效帮助用户提升工作效率，全面重塑中文搜索体验。自发布以来，“天工 AI 搜索”持续优化升级，现已可处理超长上下文。同时，为了向用户呈现更加全面、优质的搜索内容，我们不断扩展信息来源，目前搜索结果已覆盖图片以及视频等模态内容。此外，我们引入了全新的 AI 档案功能，用户可以提前设定好自己的职业、关注领域、兴趣爱好以及搜索用途等信息，

AI 搜索将结合用户的知识面和认知水平反馈更加个性化的搜索结果。

AI 游戏方面，我们的首款 AI 游戏《Club Koala》于 8 月 25 日正式亮相德国科隆游戏展。这将是一款以 UGC 为主打的高 DAU 游戏，它的 Koala Editor 游戏编辑器面向没有编程经验的普通玩家，使得玩家可以自行制作关卡及 Party Game。同时，这款游戏通过引入大量 AIGC 元素为用户带来全新的 AI 游戏体验。我们还实现了更具智能性和多样性的 AI NPC，为玩家提供更逼真、更具沉浸感的模拟游戏玩法。商业化方面，则主要通过玩家充值购买时装等道具以及 VIP 服务，从而完成变现。

此外，AI 社交与 AI 动漫产品预计于年内上线，全新 AI 音乐产品也在研发中，各 AIGC 垂类应用落地均在稳步推进。

问：如何看待芯片端的限制性政策？天工大模型的训练和推理是否可以兼容不同的算力芯片方案？公司投资的自研芯片进展如何？

答：首先，昆仑万维算力芯片储备充足。我们从 2020 年就开始大模型的研发，所以在芯片采购方面领先于同行和其他创业企业。截至本季度末，公司此前采购及租赁的芯片已经到货约 6,000 张，另外还有约 3,000 张待交付。

其次，训练和推理算法基本上以每半年提升 10 倍的速度进行优化。从长期来看，我们认为储备的算力可以应对下一代多模态甚至 MoE 大模型的研发。

在自研领域，为应对潜在的芯片限制，公司投资控股了艾捷科芯。艾捷科芯计划开发一款可编程、高性能的 NPU 产品，同时应用于模型训练和推理。核心团队由芯片研发、集成电路、人工智能、大语言模型等领域的知名专家学者组成，其中蒋博士是芯片领域的领军人物，担任清华大学企业家协会半导体行

业分会会长，同时担任国内安路科技等多家半导体企业的董事，并主导投资了盛科通信、派瑞股份、云英谷等高科技企业，是芯片和半导体领域知名企业家与投资人。

问：公司的大模型训练和推理成本大概是什么水平？在商业化之前，计划将成本优化到什么水平？会采取哪些措施？

答：首先在硬件端，自 2020 年起，我们便前瞻性地进行算力储备，现已构建了自主可控的算力集群。目前公司芯片储备达 9,000 张，预计能够满足未来 1~2 年除视频 AIGC 之外的大模型算力需求。

其次，与大模型不同，应用端调用的模型是经过针对性优化的小模型，对算力的消耗相对较小，公司现有的硬件储备足以支撑其应用。

另外，关于推理和训练的优化工作，全球范围内都在进行。我们可以观察到，基本上我们所有的推理成本都在快速下降。在海外推出的基于 UGC 社区的 C 端产品形态中，大部分用户是使用阅读、收听或观看 AIGC 内容的 C 端用户，而真正进行 UGC 创作的用户占比相对较小。我们认为这种新的商业形态有望进一步降低我们对训练和推理成本的需求。从我们的业务模式来看，我们在商业化方面对算力成本的依赖和需求非常低。

问：天工大模型和天工 AI 搜索相关的用户数据，以及后续的迭代和商业化计划如何？与艾捷科芯的合作进展以及后续合作方向如何？

答：首先，由于国内市场向 C 端用户提供服务需要在监管部门完成备案，因此我们的天工 AI 搜索目前仍处于内测阶段。我们正在积极配合监管的备案审查工作，并同时准备了一系列针对天工 AI 搜索的营销活动。我们希望在正式完成备案后，再与大家分享相关数据。

	<p>第二，艾捷科芯专注于 AI 大模型配套芯片的研发，尤其聚焦于生成式 AI 的智能计算。其核心团队汇集了来自芯片研发、集成电路、人工智能及大型语言模型等领域的顶尖专家和学者。公司创始人蒋毅敏博士是芯片领域的领军人物。他于 1996 年毕业于清华大学电子工程系，并在美国马里兰大学取得了电机工程硕士和博士学位。蒋毅敏博士在卫星、无线通信、数字电视、可配置矢量数字信号处理器等领域有重要创新突破。艾捷科芯算是蒋毅敏博士的第二次创业。</p> <p>艾捷科芯旨在开发一款可编程的、具有高性能的 NPU 产品。与传统 AI 芯片相比，生成式 AI 芯片的研发难度更高。首先，AI 应用的计算量很大，需要更高的并行处理能力，而且处理的数据量很大，因此对内存带宽有更高的需求，数据读取速度要更快。最后，基于 Transformer 结构的大模型大量采用了注意力机制，与基于卷积神经网络和递归神经网络的 AI 芯片相比，生成式 AI 芯片中的神经元间连接众多，因此需要强大的扩展能力。我们可以看出生成式 AI 今天面临的挑战主要有三个方面，一是如何提高计算速度；二是如何优化存储；三是如何有效地进行互联。艾捷科芯的创新点即在存储密度与读取速度之间做权衡，为了在技术上取得新的突破，将考虑采纳一系列创新的架构设计。</p>
附件清单(如有)	无
日期	2023 年 10 月 31 日