**项目需求书**

**一、项目背景**

随着人工智能业务的快速发展，我单位为推进深圳（东部）人工智能产业公共服务平台（第二阶段）建设，该阶段规模为3万P(FP16）先进算力集群，现需委托第三方公司开展全面且深入的可行性研究，为决策提供有力依据，以满足自然语言大模型、无人驾驶、具身智能、多模态大模型等对高性能算力的需求。

**二、工作内容及要求**

（一）市场调研要求

1.设计调研问卷和调研计划

设计调研问卷需要明确目的，结构合理，包括开头、主体和结尾三部分，要求语言简洁明了，问题适当且有效，具有逻辑性和连贯性。调研计划设计则要求明确调研目标、调研对象，选择合适的调研方法，制定调研时间和进度安排，确定调研样本量，规划数据收集和整理方式，考虑调研预算，并设置质量控制措施。

2.企业规模调研算力需求

根据企业规模调研，调研不少于15家在自然语言大模型、无人驾驶、具身智能、多模态大模型等领域具有代表性的企业，了解其当前的算力使用现状、未来3年的算力需求预测，以及对算力性能、成本、可靠性等方面的特殊要求。分析不同企业在训练和推理阶段的算力需求差异，如无人驾驶领域对实时性、低延迟的要求，具身智能领域对异构算力、大小脑融合架构的需求，多模态大模型对大规模集群算力、存储和网络协同的要求等。调研企业的规模应包括年营业收入10亿元以下、10-50亿元、50亿元以上规模的企业。

3.根据应用类型调研

调研收集不少于20个人工智能应用和模型，包括自然语言大模型（deepseek等）、无人驾驶、具身智能、多模态大模型等领域的主要算法、模型框架及其对算力的适配性，评估市场主流算力设备（如NVIDIA H200、GB100等GPU，英特尔酷睿Ultra系列处理器等）在支持这些算法和模型方面的性能表现和优化潜力。

4.根据算力建设地调研

调研3个不同算力中心建设情况，调研对象中应包含“东数西算”算力枢纽节点。调研其现有的算力基础设施规模和布局，包括数据中心的数量、服务器的总算力、存储容量等，评估其是否具备进一步扩展和升级的潜力。了解当地算力基础设施的技术水平和设备更新换代情况。考察算力中心所在区域的电网架构和输电能力，评估其是否能够满足大规模算力基础设施的用电需求，以及在用电高峰期的供电稳定性和可靠性。同时，分析电力成本，为算力中心的长期经济运行提供电力成本参考，以降低运营成本。

5.算力芯片供应调研：

调研先进算力芯片，推荐合理算力芯片。包括芯片制造的产能规模、技术工艺水平、设备先进程度等。了解供应商的生产能力可以预测其能否满足市场需求的增长以及应对订单波动的能力。了解各领域对芯片的性能、功耗、可靠性等方面的具体要求，以及未来需求的增长速度。研究芯片供应商和分销商的库存管理策略，分析其库存水平、库存周转率、安全库存设置等指标。分析芯片市场需求是否存在季节性变化规律，以及不同季节的市场需求特点。

（二）竞品分析

1.竞品选择与分类

选取国内外不少于10家具有代表性的算力服务提供商或拥有大型算力集群的企业作为竞品分析对象，其中至少包括5家市场份额排名靠前的企业，以及5家在技术创新、运营模式等方面具有特色的新兴企业。确保所选竞品中涵盖在无人驾驶、具身智能、多模态大模型领域有较多应用和实践经验的企业。

2.竞品量化对比

性能指标：对竞品的算力集群在算力密度（每单位面积提供的算力值）、网络带宽（Mbps或Gbps）、存储容量（PB级以上）等关键性能指标进行详细对比，列出各竞品的具体数值及排名情况。同时，针对无人驾驶、具身智能、多模态大模型等应用场景，对比各竞品在推理延迟、任务吞吐量、模型精度保持等方面的性能表现，评估其对不同业务需求的适配程度。

成本指标：分析竞品的建设成本（包括硬件采购、网络部署等方面的初始投资）、运营成本（如电力消耗、冷却系统运行、人员维护等费用），以柱状图或折线图等形式直观展示各竞品的成本结构差异。重点研究其在无人驾驶、具身智能、多模态大模型等业务上的成本优化策略和经济效益。

服务质量指标：评估各竞品在服务方面的优势和不足。特别是针对无人驾驶、具身智能、多模态大模型等领域的客户反馈和服务支持情况，分析其对客户满意度和业务稳定性的影响。

（三）国家政策分析

1.政策收集与梳理

全面收集国家层面近3年来出台的与算力基础设施建设相关的政策文件，包括但不限于“东数西算”工程、算力产业发展行动计划、数据中心建设规范等政策，整理成详细的政策清单，注明各政策的发文部门、发布时间、核心内容等关键信息。重点关注政策中对无人驾驶、具身智能、多模态大模型等新兴领域算力支持的相关条款和引导方向。

2.政策量化解读与应用

资金支持政策：对国家和地方政府提供的算力产业专项补贴、税收优惠等政策进行量化分析，计算公司新建算力集群可获得的潜在补贴金额、税收减免额度等，并评估其对公司投资回报率的影响程度。结合自然语言大模型、无人驾驶、具身智能、多模态大模型等领域的产业发展重点，分析相关政策在支持这些领域算力项目建设方面的具体适用性和扶持力度。

能耗指标政策：研究相关政策对算力集群能耗指标（如PUE限制值、能耗双控目标等）的具体要求，分析公司在不同建设地区和运营模式下满足能耗政策的可行性和成本投入。探讨在保障无人驾驶、具身智能、多模态大模型等业务高效运行的前提下，如何优化算力集群的能耗表现，以符合政策规定和可持续发展要求。

产业带动政策：结合地方政策对算力产业带动当地数字经济发展的期望和要求，估算公司算力集群建成后对周边地区的产业辐射效应，包括但不限于带动就业人数、促进相关产业产值增长等可量化指标。分析其对无人驾驶、具身智能、多模态大模型等产业生态建设的推动作用，以及对当地相关产业的协同发展和创新带动效应。

（四）技术方案可行性分析

1.技术指标量化评估

对不同的算力集群架构设计进行量化评分，评估指标包括但不限于计算单元与存储单元的传输延迟（微秒级）、网络交换设备的转发效率（每秒转发的数据包数量）、集群的整体扩展能力（可支持的最大卡数及扩展后的性能下降比例）。结合自然语言大模型、无人驾驶、具身智能、多模态大模型等业务的特点和需求，分析各架构在支持这些业务的高并发、低延迟、大规模数据处理等方面的优劣势和适应性。

对关键设备进行选型评估时，明确各设备的技术指标范围，如服务器的CPU主频（GHz）、内存带宽（GB/s）、GPU加速卡的算力性能（TFLOPS或GFLOPS），并建立设备选型的量化评估模型，综合考虑性能、兼容性、性价比等因素进行打分排序，提出最优设备组合方案。针对自然语言大模型、无人驾驶、具身智能、多模态大模型等不同应用场景，推荐适合的设备配置和技术方案，以满足其对算力的多样化需求。

对算力资源管理云平台进行评估，评估包括对不同类型算力资源的管理能力，对异构芯片算力的适配兼容能力，以及资源的分配、调度和监控，评估与“中国算力网”等国家项目的对接能力，评估平台对算力资源的编排能力，评估兼容适配模型的能力，包含多模态大模型应用能力。

2.技术风险量化评估

分析不同技术方案可能面临的技术风险因素，如技术成熟度风险、供应商依赖风险、技术更新换代风险等，对每种风险的发生概率进行量化评估（以百分比表示），并结合风险发生后可能造成的损失程度（以货币金额或性能下降比例衡量），计算各技术方案的综合技术风险值，为技术方案的选择提供风险依据。特别关注自然语言大模型、无人驾驶、具身智能、多模态大模型等领域技术的快速发展和演进趋势，评估现有技术方案在未来应对技术变革和业务升级方面的风险和挑战。

（五）项目建设与运营规划

1.建设规划量化

制定详细的算力集群建设进度计划，以甘特图形式呈现项目的各个阶段和关键里程碑节点，明确各阶段的起止时间和责任人。在建设过程中，充分考虑无人驾驶、具身智能、多模态大模型等业务的特殊需求和时间节点要求，确保项目能够按计划满足业务发展的算力需求。

详细核算项目建设所需的各项费用，包括硬件设备采购费用、软件授权费用、机房建设费用、网络接入费用、人员培训费用、项目管理费用等，以表格形式列出各项费用的估算依据和详细金额，形成完整的项目投资估算报告，为项目的成本控制提供准确依据。结合无人驾驶、具身智能、多模态大模型等业务的算力需求特点和规模，合理分配项目预算，确保关键设备和技术的投入能够满足业务发展的要求。

2.运营规划量化

设计算力集群的运营管理模式，明确运维团队的人员编制和岗位职责，制定运维流程和考核指标，如服务器年平均故障率（%）、网络故障修复平均时长（分钟）、客户投诉处理及时率（%）等，确保集群的稳定运行和高效利用。针对无人驾驶、具身智能、多模态大模型等业务的高可用性和低延迟要求，制定专门的运维保障措施和服务质量标准，提高业务运行的可靠性和客户满意度。

制定算力集群的商业运营策略，包括服务定价模型、市场推广计划、客户拓展目标等，对项目的收益进行量化预测，如预计每年的算力租赁收入、增值服务收入等，编制财务报表，进行盈利能力分析和偿债能力分析，评估项目的盈利能力和投资回报率。结合无人驾驶、具身智能、多模态大模型等领域的市场潜力和业务发展趋势，制定针对性的市场拓展策略和客户合作方案，提高项目的市场竞争力和收益水平。

（六）财务分析

1.投资估算细化与量化

对算力集群建设所需的硬件设备、软件采购、机房建设、网络接入、人员培训、项目管理等各项费用进行详细估算，列出投资清单，并分析投资构成和成本控制措施，以柱状图或饼图等形式直观展示各项费用在总投资中的占比情况。结合无人驾驶、具身智能、多模态大模型等业务的算力需求规模和技术选型，合理优化投资结构，降低投资成本。

2.收益预测量化

根据算力集群的运营规划和市场分析，预测项目未来5年内运营期内的收入来源和盈利水平，包括算力租赁收入、增值服务收入等，以表格形式按年度列出各项收入的具体预测数值，并对预测数据的依据和假设条件进行详细说明，使收益预测更具合理性和可信度。充分考虑自然语言大模型、无人驾驶、具身智能、多模态大模型等领域市场的增长潜力和业务拓展机会，合理预测相关业务对项目收益的贡献，评估项目的长期盈利能力和投资价值。

3.财务评价指标量化计算

运用财务评价指标如内部收益率（IRR）、投资回收期（PP）、净现值（NPV）等对项目财务可行性进行评估，详细列出各项指标的计算公式和过程，计算出具体的数值结果，并与行业基准值或公司要求的目标值进行对比分析，判断项目的投资价值和经济效益。结合无人驾驶、具身智能、多模态大模型等业务的收益特点和风险因素，综合评估项目的财务健康状况和可持续发展能力。

（七）风险评估

1.风险识别与量化

全面识别算力集群项目建设与运营过程中可能面临的技术风险、市场风险、政策风险、管理风险、安全风险等各类风险因素，对每种风险的发生概率进行量化评估（以百分比表示），并结合风险发生后可能造成的损失程度（以货币金额或性能下降比例衡量），建立风险矩阵，确定各风险的优先级和关注重点。重点关注无人驾驶、具身智能、多模态大模型等领域技术的快速发展可能带来的技术替代风险、市场需求变化风险以及相关政策法规调整风险等。

2.风险应对措施量化

针对识别出的风险因素，制定相应的风险应对策略，如风险规避、风险减轻、风险转移、风险接受等，并对风险应对措施的成本和效益进行量化分析，评估风险应对措施的投入产出比，确保风险应对措施的有效性和经济性。针对无人驾驶、具身智能、多模态大模型等领域业务的特点和风险情况，制定专项的风险应对预案，如技术研发备份计划、市场拓展调整策略、政策法规跟踪机制等，降低风险对项目的不利影响。

**三、成果要求**

（一）提交一份完整详细的可行性研究报告，内容包括但不限于以上工作内容所涉及的各个方面，报告应逻辑清晰、数据准确、分析深入、结论明确，能够为公司决策提供有力依据，特别是对自然语言大模型、无人驾驶、具身智能、多模态大模型等领域的算力需求和相关技术方案的可行性研究部分应详细阐述，提供具体的量化数据和分析结论。

（二）报告应以纸质版和电子版形式同时提交，纸质版需装订成册，电子版应包括可编辑的Word文档和PDF格式文档，电子文档应按照目录结构进行分类整理，便于查阅和使用。

（三）合同签订后，两个月内需输出可行性研究报告

**四、资质要求**

（一）具有独立法人资格，具备有效的营业执照

（二）本项目不接受联合体投标，投标方不得将项目转包或分包给其他单位或个人。

**五、其他要求**

具备专业的技术团队和咨询团队，团队成员应具备人工智能、计算机、通信、经济管理等相关专业背景，拥有丰富的算力集群建设和咨询经验，特别是在自然语言大模型、无人驾驶、具身智能、多模态大模型等领域有相关项目经验或研究成果的团队成员优先。