

二、 开标一览表

开标一览表(单位均为人民币元)

项目名称：南宁市社会管理图像信息报警联网系统一期、二期及三期第一阶段运维及升级改造项（适用服务类）

项目编号：NZC2026-G3-990054-NNSZ

分标：分标、分标、分标

投标人名称：中国移动通信集团广西有限公司

序号	服务名称	具体服务内容	数量 ①	单价(元)②	单项合价(元) ③=①×②	提交服务 成果时间
1.	算力应用服务	<p>主要提供大模型开发支撑服务，包含数据管理、智能体开发、资源管理等功能，支持实现对大模型和智能体的开发和应用，以及提供 2 套算力支撑服务，单套配置满足显存 250GB，显存带宽 1.8TB/s,FP 16 性能 300TFLOPS，适配主流大模型框架（如 PyTorch、TensorFlow），支持多节点分布式训练与混合精度优化。具体如下：</p> <p>一、大模型开发平台</p> <p>提供 2 节点及管理中心授权以及 1 年原厂电话邮件远程支持服务，包含数据管理、模型部署等功能。</p> <p>（一）数据管理</p> <p>1、数据集管理：数据集的创建、</p>	1	1940000.00	1940000.00	自合同签订之日起 25 日内到货，45 日内完成服务提供



	<p>编辑、发布、删除等功能。支持数据在线批量标注，对数据质量进行评估。</p> <p>2、文本管理：从各种文档转化为大模型训练数据。支持原始文档格式包含不限于pdf、doc、pptx、xml、epub、mp3、md、html、png。可配置隐私、异常信息、格式、去重的过滤。</p> <p>3、数据抽取：基于文本自动生成高质量大模型微调数据，可配置长度、正则、打分、去重过滤。</p> <p>4、数据扩充：对已有微调数据进行数据增强，获得同类型、同主题的高质量微调数据。</p> <p>5、数据过滤：对已有文本数据和微调数据进行清洗，清除乱码、异常字符、敏感信息、隐私信息等。提升数据质量。</p> <p>6、数据评分：针对微调数据进行自动化打分，评估数据质量。</p> <p>7、数据库支持：支持智能体对接主流数据库及国产化数据库（包含不限于达梦、南大通用、人大金仓、瀚高等）。支持配置结果可视化、大模型总结、统计结果等。</p> <p>（二）模型部署</p> <p>1、模型训练：支持用户对预置的</p>				
--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--

大模型进行开发、训练和微调。
支持主流的开源大模型，包含不限于 DEEPSSEK、通义千问、豆包等主流开源模型。

2、可视化开发环境：提供包括但不限于 Jupyter、Webshell 等在线开发环境，支持 0.x 张卡 (GPU 显存隔离/MIG)、单卡/多卡，单机/多机等不同算力环境创建。

3、图形化大模型微调界面：支持微调任务的创建、编辑、删除等操作。支持在线配置微调参数，包括训练轮数、micro batch size、学习率、加速方式、梯度累积、学习率调节器、验证集比例、优化器等。支持查看微调日志，模型收敛过程可视化。

4、断点续训：在微调任务异常中断的情况下，具备自动从上一次任务基础上继续训练的能力，避免从头开始训练。

5、自动模型评估：支持人工和自动化评估，支持衡量不同数据集或参数下的大模型表现。

通过配置模型、数据集、超参数 (包含不限于 do_sample、top_p、top_k、temperature 等)

	<p>启动评估任务；支持查看模型评估结果，包括任务信息、评估分数、评估结果和模型评估记录。</p> <p>评估结果可以进行统计查看，包括：平均推理时间、总 tokens、错误率、TTF、total time、生成持续时间等指标。</p> <p>6、模型推理部署：提供服务管理能力。提供模型部署和镜像部署两种方式，灵活适配业务实际需求。</p> <p>7、模型服务管理：包括服务上线、下线、更新、删除等。支持查看服务详情，包括服务基本信息、资源信息、服务调用地址、推理服务条件。支持服务图像信息，图像信息信息包括 cpu 利用率、内存使用率、内存已用、内存总量、部署网络传输速度、部署网络接收速率、请求数、请求成功率、请求延迟。</p> <p>8、智能体能力配置：支持大语言模型和多模态模型，支持智能体发布、分享和删除。</p> <p>9、对话智能体：支持表单形式配置对话智能体，设置模型、Prompt、开场白、知识库、工具和推荐问题等。</p>				
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--





	<p>10、智能体编排：支持可视化的智能体编排，让用户通过简单的拖拽操作进行智能体业务编排。用户可以根据业务需求，定义符合流程编排的参数配置。</p> <p>配置检索策略：支持全文解析和图文解析两种策略。允许用户自定义文档分块大小、多级分块，可配置检索策略、引用数量、相似度阈值等；支持 3+种向量编码模型进行知识编码；支持在线测试。</p> <p>12、GPU 调度设置：支持基于 WEB 页面的 GPU 细粒度调度设置，包括：（1）MIG 模式：每张卡独立配置 MIG 方案；（2）显存隔离：按 GPU 显存进行任意大小的切分，最小支持 1GB 显存；（3）复用率：按 GPU 上并发任务个数进行配置。</p> <p>（三）其他功能</p> <p>1、管理视图：支持以集群视图、节点视图、加速卡视图查看集群整体运行情况，提供不同层级的图像信息查看，提供 GPU、CPU、内存、IB/RoCE 网络、存储等关键设施的性能曲线、图表展示。</p>				
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--



	<p>2、租户管理：支持设置系统管理员、组管理员、普通用户角色权限。</p> <p>3、知识库：支持文档知识库和问答知识库，支持知识库的创建、编辑、合并、拆分和删除。支持应用挂载多个知识库。支持问答结果溯源知识库中的引用文档和引用段落。</p> <p>二、平台配套硬件</p> <p>提供 AI 训推服务器 2 台，25GE 接入交换机 1 台。</p> <p>（一）AI 训推服务器</p> <p>1、国产机架式，配套标准机柜导轨一副。</p> <p>2、配置 2 个海光处理器 C86-4G，单颗处理器核心数 48 核，单颗处理器 96 线程，主频（基准频率）2.6GHz，支持同步多线程技术，内置安全处理器。通过中国信息安全测评中心的安全可靠测评。</p> <p>3、配置 512GB DDR5-5600MHZ 内存，配置 24 个内存插槽。</p> <p>4、配置 2 块 960G SSD 硬盘和 4 块 3.84TB NVMe SSD 硬盘；硬盘托架具备 Raid 重建时不可移出硬盘提示指示灯，避免误操作。</p>				
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--



<p>5、配置 1 个 RAID 阵列卡，最高 256MB L3 Cache ， 支持 RAID0/1/5/6/10。</p> <p>6、配置 4 个千兆电口，4 个 25G 光口（选配多模光模块）。</p> <p>7、最大支持 12 个标准 PCIe 4.0 插槽，非存储、网络等设备专用插槽。</p> <p>8、配置 4 个铂金热插拔电源模块，配置冗余热插拔冗余风扇。</p> <p>▲9、硬件管理：主机前面板提供故障诊断功能，具有对系统/内存/电源/风扇/温度/网络等关键部件的故障诊断报警功能；能够分别提示硬盘故障、系统运行故障、风扇及温度故障、网络故障、内存故障和电源故障；</p> <p>▲10、配置虚拟 KVM 功能，可实现与操作系统无关的远程对服务器的完全控制，包括远程的开机、关机、重启、更新 Firmware、虚拟光驱、虚拟文件夹等操作，提供服务器设备管理、健康日记管理、屏幕自动/手动截图，提供电源图像信息 BMC 系统独立于业务操作系统，可通过 BMC 管理页面或者前后面板 RST 按键独立重启 BMC 系统。</p>				
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--

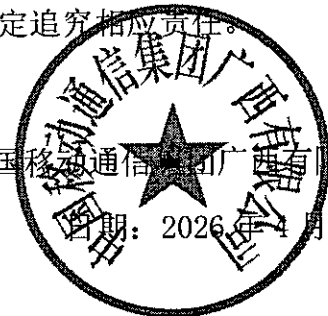
	<p>▲11、具备 BMC 管理功能，支持 DHCP、静态 IP 设置网络功能，支持设备日志记录，BMC/BIOS flash (ROM) 硬件冗余、升级失败后自动回退或切换至另一片 Flash 启动。支持配置硬盘数据安全擦除功能软件。</p> <p>12、支持 BMC 界面集成 redfish 开发调试工具，并支持客户自定义功能，根据用户自定义 API 进行 Redfish/Restful 自动注册，返回值可按用户设置键值返回，即 Redis 数据库中的 Key Value 进行返回。</p> <p>13、配置 4 张兼容 CUDA 的国产化 GPU 算力卡，显存 64G，显存带宽 1.8 TB/s，FP16 性能 80 TFLOPS。单卡最高支持 128 路编码和 96 路解码，兼容 HEVC、H.264、AV1 等多种视频格式，支持 8K 分辨率。</p> <p>14、含国产化操作系统。</p> <p>（二）25GE 接入交换机配置如下：</p> <p>1、设备关键芯片采用国产化芯片；</p> <p>2、交换容量 8Tbps，包转发率 2400Mpps（若官网有 A/B 值，以 A</p>				
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--	--

	值为准); 3、配置 48 个 25GE SFP28 接口, 6 个 100GE QSFP28 接口;提供组网所需全部光模块和线缆; 4、配置双电源;配置安装导轨; 5、三年维保				
报价合计(包含税费等所有费用): (大写) 人民币壹佰玖拾肆万元整 (¥1940000.00 元)					

注:

- 1、投标人需按本表格式填写, 不得自行更改, 也不得留空, 如有多分标, 按分标分别提供开标一览表, 必须加盖投标人有效电子公章, 否则其投标作无效标处理。
- 2、本表内容均不能涂改, 否则其投标作无效标处理。
- 3、如为联合体投标, “投标人名称”处必须列明联合体各方名称, 并标注联合体牵头人名称, 且盖章处须加盖联合体各方公章, 否则其投标作无效标处理。
- 4、以上表格要求细分项目及报价, 填写具体服务内容, 否则其投标作无效标处理。
- 5、特别提示: 采购机构将对项目名称和项目编号, 中标供应商名称、地址和中标金额, 主要中标标的的名称、服务范围、服务要求、服务时间、服务标准等予以公示。
- 6、符合招标文件中列明的可享受中小企业扶持政策的投标人, 请填写中小企业声明函。注: 投标人提供的中小企业声明函内容不实的, 属于提供虚假材料谋取中标、成交, 依照《中华人民共和国政府采购法》等国家有关规定追究相应责任。

投标人名称(电子签章): 中国移动通信集团广西有限公司



日期: 2026年4月13日