一、项目概述

1、项目名称：基于医疗大语言模型的智算平台建设项目

2、服务项目背景介绍：上海交通大学医学院附属新华医院创建于1958年，是新中国成立以来上海自行设计建设的首家综合性医院。60余年发展路，新华人励精图治，打造了一家儿科与成人学科门类齐全且特色鲜明，门急诊量常年位居上海市级医院前茅，有能力为患者提供0-100岁+，全生命周期呵护的三级甲等医院。新华医院先后获得全国先进集体、全国卫生系统先进单位、全国卫生系统卫生文化建设先进单位、上海市文明单位和上海市五一劳动奖状等荣誉称号。

为响应国家宏观政策和基于医院高质量发展战略要求，面向未来的“智慧医院”应顶层规划，提升医院各类评级的等级，有必要搭建支撑医院高质量发展平台。上海交通大学医学院附属新华医院作为一家现代化综合医院，贵院具备卓越的医疗水平和先进的设备。随着医院规模和就诊人数的不断增加，各方对医疗服务质量的要求也在不断提高，传统的信息化建设面临着一定的瓶颈。因此，拟通过建设基于医疗大语言模型的智算平台，在此基础上搭建各类智能化应用，以提升医院的影响力和服务质量，实现群众获得感不断增强、医院持续健康发展的良好局面。

通过基于医疗大语言模型的智算平台建设项目，建设医疗大语言模型和医疗智能体开发平台。医疗大语言模型为底层基础，利用高质量医学语料数据训练而成，能够处理各类复杂医学任务，对外提供医疗大模型能力，能与医生、患者等人群进行更加自然的对话。医疗智能体开发平台提供全面的智能体管理与创建功能，支持用户轻松构建多轮对话和以工作流为核心的AI智能体。平台具备可视化的内容编排和流程搭建工具，使得用户能够直观地设计智能体的交互逻辑。在完成智能体创建后，用户将智能体无缝集成到现有业务流程中，从而提升业务效率和优化用户体验。

智算平台的建设旨在帮助医院快速、便捷地部署和集成医疗大语言模型，构建定制化的智能体解决方案。该平台支持医院根据不同的需求和业务场景，灵活调整和配置智能体，确保高效的运营。

目标1.作为医院级医疗大模型基础设施

帮助医院快速部署和集成医疗大语言模型，构建定制化的智能体解决方案，满足不同医疗场景下的高效运营需求。系统提供API接口，方便与医院的其他信息系统进行对接，实现信息流的无缝衔接，提高医疗服务的整体效率。

目标2.提供高质量的语料库构建与搜索增强（RAG）技术

支持医院构建专属语料库，结合医疗大语言模型，提升信息检索的精准度和灵活性。通过先进的检索增强生成（RAG）技术，平台能够确保AI系统准确理解并调用医院核心知识和数据。系统提供自然语言处理能力，能够增强问答功能，支持医生和患者在查询医疗信息时，获得更准确、更智能的反馈。这种深度融合不仅可以提升医生决策的效率，还能提升患者的自助服务体验，促进精准医疗的实现。

目标3.将医疗大语言模型赋能已有业务，打造大模型智能应用

智能体可无缝集成至医院现有业务流程，为医院提供全面的运营服务和数据分析支持。通过智能体与业务流程的结合，平台能够帮助医院增强现有应用能力，提升运行效率。此外，系统可以采集和分析运营数据，为医院的管理层提供可视化报告和决策支持，帮助识别潜在的业务改进点和提高整体用户体验。通过这种集成，医院不仅能够优化资源配置，提升工作效率，还能增强患者满意度，最终推动医院的全面数字化转型。

3、项目建设地点：上海交通大学医学院附属新华医院

4、项目建设周期：合同生效后3个月内

5、付款方式：

（1）本项目自签约后开始实施，乙方按照项目实施计划进行项目实施工作，并经过甲方阶段性验收后（产品上线），根据医院付款流程，向乙方支付本合同项目总金额40%的合同款。

（2）本项目自验收合格之日起，且甲方在收到乙方开具的有效发票后，根据医院付款流程，向乙方支付本合同项目总金额40%的合同款。

（3）本项目剩余的合同款按实际维保期月份数/合同约定维保月份数折算，实际未维保月份不予支付。甲方在收到乙方开具的有效发票后，根据医院付款流程，向乙方支付。

二、项目内容

通过基于医疗大语言模型的智算平台建设项目，建设医疗大语言模型和医疗智能体开发平台。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 项目 | 平台/系统 | 功能/模块 | 子功能/模块 | 数量 |
| 基于医疗大语言模型的智算平台建设项目 | 医疗大语言模型 | 模型知识体系 | 多来源高质量医学训练数据 | 1套 |
| 多维度医学领域覆盖 |
| 模型理解能力 | 文本理解与多轮对话 | 1套 |
| 意图识别与意图分类 |
| 语义推理与医学逻辑判断  |
| 模型知识库挂载能力 | 知识库插件集成  | 1套 |
| 文档库集成能力 |
| 长程记忆存储与长文本能力 | 长程记忆存储能力 | 1套 |
| 长文本能力 |
| 大模型工程化能力 | 高可塑性提示工程优化 | 1套 |
| 智能化语言风格与格式调整 |
| 医疗智能体开发平台 | 医疗智能体创建 | 可视化智能体编排与管理 | 1套 |
| 可视化Prompt编辑器 |
| 大模型参数配置 |
| 富文本编辑器 |
| 自定义Suggest Prompt |
| TTS语音合成 |
| ASR语音识别 |
| 文档解析与多模态支持 |
| 智能体背景设置 |
| 多类型技能挂载 |
| 实时预览与日志 |
| 医疗智能体发布及运营 | 多渠道发布管理 | 1套 |
| 数据可视化运营管理 |
| 对话日志与统计 |
| 语料管理库与配置 | 语料库管理 | 1套 |
| 语料库向量化配置 |
| 语料库召回相关配置 |
| 语料库召回测试 |
| 工作流绘制与管理 | 工作流可视化绘制 | 1套 |
| 工作流智能体编排 |
| 工作流列表管理 |
| 多功能插件管理 | 插件定义和使用 | 1套 |
| 多维度插件分析 |
| 用户管理中心 | 多角色权限管理 | 1套 |
| 统一广播与个性化订阅 |
| 安全管理中心 | 安全监控管理 | 1套 |
| 安全策略管理 |
| 模板集市管理 | 官方医疗场景模板 | 1套 |
| 自定义模板与一键复制 |
| 消息中心 | 平台通报与协作提醒 | 1套 |
| 多维度过滤与归档 |

三、技术需求

（一）医疗大语言模型

（1）模型知识体系

1.多来源高质量医学训练数据

1.1模型以广泛且高质量的医学知识数据进行训练，数据范围涵盖医学教材、权威医学指南、临床路径、药品库、疾病库、体检报告等。

1.2利用多种指令数据集进行模型训练，包括医疗指令数据集、通用领域指令数据集、医疗领域知识类数据集、通用领域知识类数据集等。

2.多维度医学领域覆盖

2.1模型可覆盖内科、外科、妇产科、儿科、影像科、检验科、护理学、公共卫生学、药理学等多专业子领域，为临床与科研提供知识支持。

2.2对医学术语、药品通用名/商品名等关键信息具有较高识别与理解能力。

（2）模型理解能力

1.文本理解与多轮对话

1.1具备强大的自然语言理解能力，能够识别医学领域的专业术语、缩写以及上下文关联含义。

1.2在多轮对话中能保持对话上下文信息，记忆用户先前提供的病症描述、检查结果等，提供连贯、准确的答复。

2.意图识别与意图分类

2.1能够将自然语言文本解析为机器可理解的结构化信息，并自动识别患者、医生或其他医疗人员的真正意图。

2.2针对不同的咨询类型（如问诊、科普、用药咨询、流程咨询等）进行分类，并选择对应的专业回答策略。

3.语义推理与医学逻辑判断

3.1具备语义推理能力，可对临床路径、用药方案、检验结果等进行综合判断，帮助医生或患者进行初步的信息筛选或逻辑分析。

3.2可根据疾病与症状之间的关联关系，给出差异化诊断思路或建议供就诊科室参考。

（3）模型知识库挂载能力

1.知识库插件集成

1.1模型可支持与外部知识库或插件服务进行挂载。

1.2支持通过API或向量检索等方式实现对海量专业医学资料的实时查询与总结。

2.文档库集成能力

2.1支持与文档库进行对接。

2.2支持动态调用最新病例库、用药方案库、临床指南库等。

（4）长程记忆存储与长文本能力

1.长程记忆存储能力

1.1内置长程记忆存储能力，可在与用户的长期交互过程中保留并更新关键上下文信息（如历史病历、对话记录、检查结果等），为下一步对话提供连续性支持。

1.2支持根据交互的实际需求动态调整记忆的深度和更新频率，在多次对话中精准把握用户需求。

2.长文本能力

2.1具备长文本处理和理解能力，能够高效解析并提取文本中的关键信息，支持对多段、长篇文本的高质量理解与分析。

2.2在面对长文本时，能够灵活应对并保持文本间的逻辑连贯性，快速识别文本中的上下文关系和重要数据点。

（5）大模型工程化能力

1.高可塑性提示工程优化

▲1.1支持提示工程能力：具备提示工程效果优化能力，支持进行提示词修改和提示工程优化。需提供“医疗大模型提示词配置软件”或类似名称的计算机软件著作权证书。

1.2支持对话语气、专业程度、回答形式等多维度参数的配置与调优。

1.3支持调用外部编排引擎或工作流，在提示词中嵌入上下文信息，让模型回答更具逻辑性与准确度。

2.智能化语言风格与格式调整

2.1根据使用场景需求，自动调整回复内容的语言风格（如通俗化解释、专业学术表达等）、详略程度和格式化输出。

2.2可配置文本输出格式，如分点陈述、表格列举、医学报告模板等。

（二）医疗智能体开发平台

（1）医疗智能体创建

1.可视化智能体编排与管理

1.1支持拖拽式配置对话流程、业务流程，灵活定义节点关系。

1.2支持创建简单对话型智能体，也支持创建复杂工作流型智能体。

1.3支持直接应用官方或团队共享的智能体模板。

2.可视化Prompt编辑器

2.1支持实时编辑并查看Prompt效果，支持调试、测试模型的回复结果。

2.2支持内置典型医疗场景的Prompt库，便于快速编写问诊、诊断类提示词。

2.3支持个性化Prompt模板创建、复用。

3.大模型参数配置

3.1支持不同大模型的接入与切换，可对接本地部署LLM、私有化大模型或公有云API。

3.2支持对模型输入及输出相关参数进行配置。

3.3支持对模型相关参数进行配置，包括生成随机性、重复惩罚参数等。

4.富文本编辑器

4.1支持以所见即所得方式配置智能体开场白。

4.2支持开场时预置引导性问题，提升用户交互效率。

5.自定义Suggest Prompt

5.1支持在模型回复后，自动生成推荐追问和/补充问题。

5.2支持配置相关Prompt逻辑，帮助用户进一步深入了解咨询内容。

6.TTS语音合成

6.1支持集成TTS引擎，将文本内容以多种语音风格进行播报。

6.2支持可扩展多语言音色，如普通话、英语等。

7.ASR语音识别

7.1支持将用户语音输入转换为文本并实时传递给智能体模型。

7.2支持通过医疗专用词表优化语音识别准确率。

8.文档解析与多模态支持

8.1支持用户上传病例、检验报告、影像资料（PDF、图片等）进行解析。

8.2支持与文档、图片OCR插件集成，结构化输出核心信息。

9.智能体背景设置

9.1支持上传智能体对话界面的背景图并进行缩放、位置调整。

9.2支持医疗场景主题皮肤（医院Logo、统一设计风格等）。

10.多类型技能挂载

10.1支持一键挂载知识库、工作流、插件等不同能力。

10.2支持智能体在对话中调用对应技能。

11.实时预览与日志

▲11.1调试预览：支持对编排配置实时生效调试预览；支持查看调试相关详情信息，包含对话涉及的模型参数、运行情况、运行时间等内容；需提供系统截图作为证明材料。

11.2支持调试日志的阅览、筛选和导出，支持对调试情况的整体统计指标展示。

（2）医疗智能体发布及运营

1.多渠道发布管理

1.1支持将智能体集成至已有的业务流程以提升业务效率，优化用户体验。

1.2支持以网页发布智能体，支持二维码形式分享入口。

▲1.3支持以API和Web SDK形式发布智能体。需提供系统截图作为证明材料。

1.4支持阅览发布历史版本的相关配置内容。

1.5支持智能体发布历史的版本回退。

▲1.6支持智能体发布后对相关发布渠道的参数进行管理；需提供系统截图作为证明材料。

2.数据可视化运营管理

2.1支持对发布后的产品相关运营数据进行可视化展示。

▲2.2支持智能体发布后的用户使用相关信息的统计、展示、筛选和导出；需提供投标截止日在有效期内的软件著作权证书证明。

3.对话日志与统计

3.1支持记录用户与智能体的所有交互进行记录。

3.2支持交互记录的检索、筛选与导出。

（3）语料库管理与配置

1.语料库管理

1.1支持创建个人语料库，支持创建团队语料库并要求其他用户共同编辑。

1.2支持创建语料库，支持本地上传语料库文档，支持对语料库内文档的列表展示和筛选。

2.语料库向量化配置

2.1支持按照语义段落、页面、句子级粒度切分文档。

2.2支持查看并编辑切分结果，以确保医学内容的准确性。

3.语料库召回相关配置

3.1语料库检索策略配置，支持针对不同语料选择语义、全文、混合搜索方式。

3.2支持配置top K和score阈值参数以过滤低相关结果。

3.3支持对自定义无召回情况的回复配置。

4.语料库召回测试

4.1支持对知识库配置进行召回测试，以实时调整相关配置。

4.2支持针对知识库检索配置进行模拟查询，快速调优参数。

（4）工作流绘制与管理

1.工作流可视化绘制

1.1支持通过可视化编排的方式，以直观的拖拽操作对相关功能节点进行串联，从而实现复杂、稳定的业务流程。

▲1.2支持创建以工作流作为主体流程的智能体，可支持多轮对话；需提供系统截图作为证明材料。

1.3支持创建用于具体场景的业务过程、任务或活动的工具型自动化流，可在普通智能体的编排中添加和引用。

1.4支持集成后的工作流智能体自动启动并开始流程。

▲1.5支持大语言模型生成（LLM）、医学文本结构化、条件判断等节点；需提供系统截图作为证明材料。

▲1.6支持绘制交互式表单，并支持对表单填写内容进行数据提取或者分析；需提供系统截图作为证明材料。

▲1.7支持定义会话变量，支持以交互的形式填写会话变量；需提供系统截图作为证明材料。

1.8支持工作流节点的动态参数传递和实时调试。

1.9支持配置过程中相关问题的检查和提醒。

1.10支持画布toolbar、绘制变更自动保存和后退、相关快捷键操作。

2.工作流智能体编排

2.1支持创建、复制或引用现有工作流作为智能体主体流程。

2.2支持配置以工作流为主体流程的智能体的开场白、文件上传等功能。

2.3支持以工作流为主体流程的智能体的功能调试和预览。

3.工作流列表管理

3.1工作流支持按名称、场景标签等进行搜索和分类。

3.2支持增删改查、复制工作流和查看工作流使用记录。

（5）多功能插件管理

1.插件定义和使用

1.1提供OCR解析、文档解析、医学图像关键字段提取等常见医疗场景功能。

1.2支持用户自定义插件工具。

2.多维度插件分析

2.1支持列表用户创建的插件，支持插件的搜索、复制、增删。

2.2支持插件被哪些智能体、工作流所调用相关使用信息的查看。

（6）用户管理中心

1.多角色权限管理

1.1支持创建院内账户、外部协作账户。

1.2支持角色分组（管理员、医生、护士、IT等）。

1.3支持灵活配置功能权限与资源配额。

1.4支持批量导入或导出用户列表。

2.统一广播与个性化订阅

2.1支持通过消息队列将平台公告推送相关用户。

2.2支持用户设置订阅偏好。

（7）安全管理中心

1.安全监控管理

1.1支持对触发安全监控的用户、被封禁用户进行查看和使用权限的操作。

1.2支持对监控记录进行搜索筛选、查看以及导出。

2.安全策略管理

2.1支持安全策略的版本管理和发布更新。

▲2.2支持配置当前智能体所用的安全策略内容。需提供投标截止日在有效期内的软件著作权证书证明。

（8）模板集市管理

1.官方医疗场景模板

1.1系统内置涉及医疗应用各类场景的第一方模板。

1.2支持官方医疗场景模板的更新和编辑。

2.自定义模板与一键复制

2.1支持将自定义创建好的智能体和工作流发布到模板集市。

2.2支持从模板集市复制内容并快速创建。

（9）消息中心

1.平台通报与协作提醒

1.1支持展示平台公告的变动内容。

1.2支持展示协作知识库的变动内容。

2.多维度过滤与归档

2.1支持对历史消息进行归档、筛选、查看，以便信息管理。

2.2支持对历史消息的导出，方便后续的查看和分析。

四、实施部署要求

（一）技术要求

1.算法性能：产品所采用的人工智能算法需在业务相关指标等上达到业务需求指标，以确保满足业务需求。

2.数据处理能力：具备高效的数据处理能力，能够处理业务场景下的数据规模、类型等，满足业务需求。

3.系统兼容性：可与现有业务系统无缝对接，兼容有业务系统的技术架构、操作系统等，如 Windows Server 操作系统、Oracle 数据库等。

4.接入国内外主流大模型API接口：项目所建设的平台支持接入国内外主流大模型API接口，如deepseek v3、deepseek r1、qwen、glm、日日新、大医等。

5.本地部署不低于70B规模的deepseek大模型。

6.数据要求：原始数据不出院

（二）交付和保障要求

1.本项目需要提供完善的项目实施方案，包含项目进度计划和保障措施等内容。

2.招标范围内的全部内容都由投标人负责，包括但不限于系统升级及安装、配置设备、安装、调试等。

3.项目建成系统安装完成后对招标方管理人员进行培训与技术指导，确保业务人员能够熟练使用产品。

★4.为保证综合服务能力，投标人须承诺派遣1名合格的技术人员或服务人员至用户指定的工作场所驻场，提供服务承诺函。

（三）售后服务要求

1.投标人具有固定的售后服务团队和软件开发服务团队：故障响应时间不超过1小时，非工作日及夜间故障响应时间为不超过2小时，需现场处理时技术人员2小时内到达现场，4小时内修复。免费维保期内每季度需要至少一次对系统优化和常规安全检查。

★2.本项目需提供自验收通过之日起≥1年7\*24小时免费维保服务。

三、商务要求

1. 最高限价

人民币199.00万元

2. 资格条件

（1）具有合法经营资质的独立法人、其他组织；

（2）未被“信用中国”网站（www.creditchina.gov.cn）列入失信被执行人名单、重大税收违法案件当事人名单、未被中国政府采购网（www.ccgp.gov.cn）列入政府采购严重违法失信行为记录名单。

（3）本项目不接受联合体投标。