多模态数据科研服务平台

项目需求书

注意事项:《项目需求书》中标注有“▲”号的条款为基本条款，需要保留。

# 一、 项目概况

|  |  |  |
| --- | --- | --- |
| **招标内容** | **最高限价（人民币）** | **交付时间** |
| 多模态数据科研服务平台 | 100万元 | 自合同签订之日起3个月内完成开发、验收并交付使用。 |

# 二、 采购内容

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **项目** | **子项** | **技术参数** | **数量** | **单位** |
| 1 | 技术开发服务 | 数据采集与预处理子系统 | 详见“四、技术要求” | 1 | 项 |
| 2 | 数据治理子系统 | 1 | 项 |
| 3 | 数据管理与安全子系统 | 1 | 项 |
| 4 | 数据分析与应用子系统 | 1 | 项 |
| 5 | 科研数据云服务子系统 | 1 | 项 |
| 6 | 计算服务器 | 计算服务器 | 用于科研数据分析、计算 | 1 | 台 |
| 7 | 实施部署 | 实施部署 | 软硬件系统部署 | 1 | 项 |
| 8 | 运维支持 | 运维支持 | 运维服务1年 | 1 | 项 |
| 9 | 认证培训课程 | 认证培训课程 | 培训与指导系统用户快速掌握并了解平台使用方法 | 1 | 期 |

# 三、 建设要求

1. 充分利用医院现有数据资源和软硬件资源，通过大数据技术、人工智能技术，构建一个集各类多模态科研相关数据的采集、关联、治理、管理、分析与应用于一体的高质量数据资源平台。涵盖的数据资源范围包括但不限于：临床门诊数据、住院数据、诊断数据、检查检验数据、放射影像数据、常规组织病理与细胞病理检查数据、分子基因检查数据、病理诊断数据、随访数据等多模态、全流程数据；
2. 构建一个支持各类科研应用获取数据，分析数据、使用数据的科研应用生态支撑平台。
3. 建设的成果包括但不限于以下应用或功能：

|  |  |  |
| --- | --- | --- |
| 类别 | 应用 | 主要功能开发 |
| 数据采集与预处理子系统 | 多源数据采集 | 支持将院内多源异构数据进行聚合，包括各业务系统数据HIS、LIS、PACS、病案、病理、内镜、心电、护理、随访等 |
| 多模态非结构化数据采集 | 支持病理图像、病理数字切片、影像、生信等非结构化数据的采集 |
| 多模态数据存储 | 支持多模态非结构化数据的集中式和分布式存储，支持平台授权方式统一调用和访问 |
| 数据预处理 | 支持数据清洗（去重、智能补全或纠错等），数据标准化、归一化、脱敏加密等处理，对采集后的数据进行提取和转换，便于后续处理和分析。 |
| 数据治理子系统 | 医学术语标准化 | 支持医学术语标准映射，术语维护、术语对照、术语概念归一、术语管理。 |
| 后结构化处理 | 支持通过医学自然语义处理、机器学习等技术，基于医学术语标准化应用模块，对采集数据的结构化、标准化和归一化等自动处理、监控数据质量等功能 |
| 元数据管理 | 对采集的元数据进行管理，建立元数据仓库。提供元数据搜索及查询，元数据版本查询，方便用户查找和理解数据，追溯元数据演化过程。 |
| 数据质控 | 支持对数据抽取质检，保证源和目标数据的一致性；结合医学术语标准化应用模块、指标理解等，对数据进行有效性检查、逻辑校验等系统性核查；支持人工数据质疑核查功能，并全流程处理留痕，实现数据版本管理。 |
| 非结构化数据治理 | 利用机器学习和人工智能技术，对非结构化数据进行自动分类（如按数字病理切片的取材部位、疾病分期分型等）。 |
| 非结构化数据标注 | 支持对病理数字切片、图像等非结构化数据进行标注，并支持对标注数据的关联搜索；支持外部AI标注数据的对接和导入。 |
| 数据管理与安全子系统 | 数据存储与管理 | 支持结构化数据和非结构化数据采用不同存储管理策略；支持院内各架构存储资源池，异构存储统一接入；支持非结构化数据安全分布式存储；支持各种存储协议与实时扩容。 |
| 数据安全交换 | 平台多模态数据的安全交换引擎和安全访问引擎，实现医院内网中不同系统之间、医院内网与互联网之间的安全交换与共享；安全交换引擎实现结构化数据安全交换和安全同步；安全访问引擎实现数字切片等非结构化数据授权访问功能，包括授权、交换、到期销毁等交换全过程管理。 |
| 数据脱敏加密 | 支持对数据中的敏感信息通过脱敏或加密规则进行访问，实现敏感隐私数据的可靠保护；支持对系统脱敏字段列表展示并进行条件查询；支持对数据进行加密处理。 |
| 系统配置与权限管理 | 包括用户账户管理系统，支持多级角色和权限设置，角色管理，部门管理等基础管理功能；字典管理、日志管理、数据授权审批管理、下载管理等系统管理功能 |
| 数据分析与应用子系统 | 字段集管理 | 根据我国卫健委发布的行业标准、WHO肿瘤相关指南和本部门实际工作中的具体要求，设置平台标准化字段集，并根据不同业务域进行管理。 |
| 统一数据解析系统 | 支持数字病理、影像等非结构化数据的统一解析；支持数字病理国内外各主流生信图像格式。 |
| 智能结构化处理 | 采用大模型、NLP等技术，结合平台标准字段集要求，对数据进行结构化处理。保证数据标准化、规范化。 |
| 高级检索系统 | 支持精确检索与模糊检索、支持全文检索；支持非结构化数据的精确检索；支持检索结果分组、排序、统计等功能；支持自定义检索方案保存、调用等 |
| 数据洞察与统计 | 支持数据洞察分析，了解指标值域分布情况；支持自定义指标进行统计分析；支持多种可视化形式展示数据统计结果；支持数据的合并展示等。 |
| 数据驾驶舱 | 支持平台数据的概览展示；支持平台数据主要指标的特征概况展示；支持平台用户、访问等数据的可视化展示。 |
| 数据保存与导出 | 支持用户检索结果数据保存；支持用户自定义检索条件自动数据订阅；支持用户自定义字段导出；支持EXCEL、PDF等多种导出格式 |
| AI辅助应用接入支持 | 支持病理AI辅助诊断应用的无缝接入；支持病理AI诊断结果数据导入并查看。 |
| 个人中心 | 支持用户收藏病例数据、切片数据等功能；支持用户自定义保存数据表；支持用户科研项目数据申请、审批及管理等功能。 |
| 数据接口中心 | 支持对院内其他系统提供数据接口服务；支持对院外合规授权后提供数据调用接口服务；支持对检索、解析阅片等提供功能调用接口服务；接口服务需满足脱敏、安全等要求。 |
| 科研数据云服务子系统 | 科研数据云服务平台 | 支持在院内多模态数据科研服务平台获得合规授权后，在互联网端实现数据的访问和调用；支持云服务平台相关数据权限、用户系统、脱敏策略、数据服务内容及功能等由院内服务平台统一管理配置;实现数据服务的多终端同步访问。 |
| 多模态数据云存储及云解析系统 | 支持科研数据云服务平台专属云存储服务；支持病理数字切片、影像等多模态数据的统一解析服务。 |
| 计算服务器 | 计算服务器 | 用于科研数据的人工智能计算、分析 |

# 四、 技术要求

以下技术要求中，包含▲的条款需提供相关文件或证书或产品系统界面截图或投标人自拟的技术说明等相关证明材料，并加盖原厂公章。

（一）数据采集与预处理

1. 支持从医院HIS、LIS、电子病历等系统中采集患者基本信息、诊断信息、治疗信息等；
2. 数据采集支持定时抽取、全量抽取、增量抽取。支持采集优先级、采集时间、采集周期管理；
3. 支持不同数据类型的数据集成，不同数据库格式、文本格式、JSON格式等。支持结构化数据、半结构化或非结构化数据的解析与融合；
4. 支持数据采集日志的查看和管理；
5. ▲支持组织病理、细胞病理数据采集，支持病理诊断报告、样本描述、病理图像、病理切片等数据的采集；
6. ▲支持分子基因数据采集，分子基因测序结果、基因表达数据、样本类型、分子检测报告等；
7. 支持影像数据、其他检验检查数据以及随访数据采集；
8. ▲支持病理数字切片数据、病理图像数据、影像数据等非结构化数据的集中式存储或分布式存储；
9. ▲支持非结构化数据的统一授权、加密调用和访问。
10. 支持对数据进行去重、补全、纠错等处理，支持对数据进行标准化、归一化等预处理。

（二）数据治理

1. 支持医学术语标准映射，术语维护，术语概念归一及管理；
2. 支持医学自然语义处理、机器学习对数据进行结构化、标准化、归一化处理；
3. ▲支持运用AI大模型技术对病理报告、影像报告等进行全量解析，并结合知识图谱进行解读表达；
4. 对采集的元数据进行管理，建立元数据仓库。提供元数据搜索及查询，元数据版本查询，方便用户查找和理解数据，追溯元数据演化过程；
5. 支持对数据抽取质检，保证源和目标数据的一致性；
6. 支持结合医学术语标准、指标理解等，对数据进行有效性检查、逻辑校验等系统核查；
7. 支持人工数据质疑及核查；
8. 支持数据核查全流程处理留痕，实现数据的版本管理；
9. ▲利用人工智能技术，对病理数字切片等非结构化数据进行处理，实现病理数字切片按检查指标、取材部位、病种等要求进行分类管理；
10. ▲支持对病理数字切片、图像等非结构化数据进行标注，并支持标注数据的关联搜索。支持外部AI标注数据的导入和查看。

（三）数据存储与安全交换

1. 支持结构化数据和非结构化数据采用不同存储管理策略；
2. ▲支持院内各架构存储资源池，异构存储统一接入；
3. 支持非结构化数据安全分布式存储；
4. 支持各种存储协议与实时扩容；
5. ▲提供通过公安部计算机信息系统安全产品检验检测报告；
6. ▲对数据的交换过程提供区块链上链存证服务，保证每一笔数据交换情况可溯源、可确认，防止非法的数据获取和篡改。（需提供电子数据区块链存证证明）。

（四）系统及权限管理

1. ▲支持以病例、切片、分子、影像等不同维度，配置构建数据表方案；
2. 支持用户管理、角色管理、部门管理等基础管理功能；
3. 支持字典管理、日志管理等系统管理功能；
4. 支持对数据中的敏感信息通过脱敏或加密规则进行访问，实现敏感隐私数据的可靠保护；
5. 支持对系统脱敏字段列表展示并进行条件查询、策略设置等管理；
6. 支持角色权限管理，支持根据角色进行数据的权限与策略配置、平台功能的策略配置；
7. ▲支持院外访问权限设置管理，针对数据类型进行配置院外访问策略；
8. 支持数据申请与审批管理，具有审批权限的管理员可管理数据的导出、下载、院外访问等数据申请。

（五）高级检索

1. ▲基于医学自然语义处理技术，对结构化、半结构化、非结构化数据进行检索，支持复杂的检索条件并支持全文检索；
2. ▲支持不同数据集类型采用不同检索条件配置方案，根据条件字段的属性自动匹配逻辑运算规则、值域范围、输入方式等；
3. ▲支持数字病理切片等非结构化数据精确搜索，支持按切片指标名、指标值、判读结果性质、取材部位等指标，实现切片文件的精确定位访问；支持切片关联病例、患者信息、分子检测、影像等其他临床数据的多维展示和访问；
4. ▲支持按基因名、变异类型、检测项目、变异解读等分子数据结合病种、患者信息、临床信息等精确检索病例数据；
5. 支持以条件树形式进行多级组合条件筛选，组合条件之间支持“and \or ”多层逻辑关系的嵌套组合；
6. 支持单个检索条件的逻辑运算规则设置，包括：大于、等于、小于、包含、不包含等规则；
7. 支持对搜索结果进行实时统计并可视化展示，支持用户设置统计条件；
8. 支持用户对搜索结果按行展开操作，对选中行的数据进行详细显示；
9. 支持搜索结果按列操作，对选中列进行搜索操作，实现搜索结果的二次筛选，同时支持对选中列的排序、隐藏等操作；
10. ▲支持搜索结果单元格内容展开操作，对选中单元格的详细数据内容进行查看，例如：同一病例的相关报告数量、切片数、基因数、诊断信息等字段，选中单元格后，可对相关指标涉及的数据进行详细查阅；
11. ▲支持自定义条件树，并保存为常用检索条件方案，支持用户选择调用；
12. ▲支持对搜索结果进行自定义分组显示，支持分组条件的层级嵌套组合；
13. ▲支持对搜索结果进行自定义排序显示，支持排序条件的组合；
14. 支持用户设置并保存搜索结果展示方案，包括：字段的隐藏/显示设置、字段的展示顺序设置、每页显示数量等；
15. ▲支持搜索结果数据中转功能，用户可选择每次搜索结果中的部分数据放入数据中转站，搜索结束后对中转数据进行一次性保存、导出操作；
16. 支持搜索结果数据的导出功能，支持用户自定义导出字段，支持根据权限匹配数据导出的脱敏策略，支持EXCEL、PDF等多种导出格式；
17. 支持搜索结果保存为个人数据表，可对数据表进行命名和添加说明操作，保存数据表的同时，保存该数据表的检索条件方案、数据统计、分组方案等；
18. ▲支持数据订阅功能，用户自定义检索条件后，当平台有新增符合条件的数据时，自动加入用户保存的数据表。

（六）数据应用

1. 根据我国卫健委发布的行业标准、WHO肿瘤相关指南和科室实际工作要求，设置并管理本平台各类字段集，包括：基础信息字段集、临床信息字段集、病理科研字段集、影像科研字段集等标准字段集，也可以根据专病科研的特点和具体要求，设定专病科研字段集；
2. 对每个字段进行字段图标、中英文名、标准名映射、定义说明、值域、业务域、检索排序规则、数据提取规则、来源等多维标签管理；
3. 采用大模型、NLP等技术，对病历报告、影像报告、病理诊断报告、分子诊断报告等非结构化大文本类数据进行数据挖掘，并结合平台标准字段集要求，进行结构化处理；
4. ▲支持病理数字切片、影像文件等非结构化数据的统一解析，其中病理数字切片解析要求支持以下数据格式，包括但不限于svs、tiff、ibl、kfb、sdpc、mdsx、mrxs、scn、ndpi。
5. ▲支持多家病理AI辅助诊断应用的接入，从平台直接发起AI诊断任务并支持AI辅助诊断结果数据的回传并显示，要求支持的AI算法厂商不少于3家，支持的病种不少于3家；
6. 个人数据中心，支持用户科研项目的数据申请需求，支持个人数据申请全部记录、审批状态、查阅详情等管理功能；
7. 支持用户合规授权的数据及状态保持院内、院外访问同步；
8. 支持用户收藏数据，包括病例数据、切片数据、影像数据等；
9. 支持用户保存和管理自定义数据表；
10. 支持对院内其他系统提供标准数据接口、功能接口服务，包括可授权数据、检索接口、解析阅片接口等；
11. 支持在合规授权后，为院外使用多模态科研数据提供安全、可靠的数据接口、功能接口服务，实现数据资源安全互联互通；
12. ▲数据接口、功能接口服务需满足脱敏、安全等要求。

（七）数据可视化与统计

1. 实现数据驾驶舱相关功能要求，支持平台数据的可视化展示；
2. 支持平台全量数据概览展示；支持平台数据主要指标的特征概况展示；
3. 支持数据统计的时间范围动态调整，以及年、月、日时间颗粒度调整；
4. ▲支持数据按亚专科、病种、部位、分型、分期、分级等指标进行分类统计及可视化展示；
5. 支持用户自定义展示各维度指标数据，包括患者信息、病种部位、检查指标、检查类型、诊疗时间等；
6. ▲支持病理数字切片数据按检查指标、取材部位等具体指标值进行分类统计及可视化展示；
7. ▲支持分子基因数据按检查项目、检测平台、样本类型、基因名、变异类型及解读等具体指标进行分类统计和可视化展示；
8. 支持日志分析概览，包括用户的PV/UV统计，访问趋势统计和TOP访问用户排名统计等；
9. 支持平台对外数据接口、功能接口服务的调用统计与可视化展示；
10. 支持院外访问、科研数据云服务平台的访问数据可视化展示。

（八）科研数据云服务

1. 建设部署科研数据云服务平台，要求和院内平台功能及样式等使用体验保持一致；
2. 支持云服务平台相关数据权限、用户系统、脱敏策略、数据服务内容及功能等由院内科研数据平台统一配置管理；
3. ▲实现科研数据服务的PC端、手机端的多终端同步访问；
4. 支持科研数据云服务平台的专属云存储服务；
5. ▲支持病理数字切片、影像文件等多模态数据的统一云解析服务；
6. ▲支持合规授权数据的院外下载功能，并提供病理数字切片多种数据格式的标准化转换服务；
7. ▲支持本地存储对接云存储模式，支持对接云存储镜像模式和缓存模式，支持的云存储应包括移动云、天翼云、亚马逊云、阿里云等三个以上云厂商对象存储；
8. ▲互联网云端进行科研数据访问时，可支持多种安全机制下的访问方式，包括但不限于（1）已交换至云端的安全脱敏数据直接访问，（2）数据不出院情况下，被授权用户访问院内指定的安全脱敏数据，如医疗影像数据等，并实现实时解析下的高速访问；
9. 支持合规授权的数据进行脱密、加密分享，支持用户对分享数据的自定义密码、有效期等访问权限设置。

（九）计算服务器

* 1. 提供高性能计算服务器一台，用于科研数据的计算、分析与人工智能应用支持等。
  2. 要求服务器性能不低于以下参数：CPU不低于96核（建议性能不低于AMD Ryzen Threadripper PRO 7995WX CPU），内存不低于512G，显存不低于32G（建议支持CUDA的 NVIDIA GPU），固态硬盘空间不低于4T，机械硬盘空间不低于 10T。

# 五、安全要求

（一）应用安全管理建设

所建应用包括用户管理、用户访问安全和应用操作安全，满足国家及我院对应用安全管理相关规定。

（二）数据安全管理要求

数据安全管理包括数据传输安全、数据加工存储安全、数据备份等，信息系统在建设的过程中应该有具体的数据安全管理方案和措施，满足国家及我院对数据安全管理的相关规定。

（三）实施安全管理建设

根据我院对医院信息系统实施安全管理相关文件的基本要求，做好实施安全管理建设，配合好信息系统等级保护测评及整改相关工作。

# 六、 商务要求

（一）报价要求

1. 投标供应商的投标总报价应当包括本项目完整实施所需的安装调试、标配工具、运输保险、培训、售后运维服务、各项税费（增值税专用发票）及合同实施过程中不可预见费用等一切费用。如有缺项、漏项，视为已包含在总报价中。
2. 投标总报价不得高于采购最高限价。

（二）供货要求

1. 投标供应商所供产品均应为原厂全新的原装产品，无侵权行为、无任何缺陷隐患，在中国境内可依常规安全合法使用；
2. 产品必须得到制造商认证，符合国家或国际的质量标准；
3. 招标人有权对中标产品进行第三方的检测，如中标产品未能达到招标文件要求的性能参数，可要求中标供应商无偿更换或单方面终止合同。

（三）交付时间

合同签订后相关硬件产品30个日历日内完成安装、调试，验收合格后正式交付招标人使用。相关软件系统在3个月内完成相关的需求分析、设计、研发、测试、上线、试运行和验收工作。

（四）交付地点

由招标人指定地点，免费送货上门，免费安装调试合格。

（五）实施服务

为保证项目质量，保证措施、服务响应及时、项目管理（包括项目组织管理、人员管理、技术管理、进度管理、资料管理等），在实施期内，提供不少于2个现场驻点服务人员。项目实施需做好用户测试系统原有应用与数据迁移，确保无缝过度可正常使用。

（六）验收要求

1. 软件安装、调试、测试完成后正式移交招标人，试运行30个日历日后，如系统运行正常，由中标供应商提出验收申请并出具项目实施情况验收报告。
2. 交付验收标准依次序对照适用标准为：①符合中华人民共和国国家安全质量标准、环保标准或行业标准；②符合招标文件、投标文件及招标人认可的合理最佳配置、参数及各项要求。

# 七、 服务要求

（一）售后服务

1. 保修期自项目验收合格之日起计算，质保、运维服务期1年，投标供应商应在投标文件中对保修期的具体期限做出承诺，保修期内对所供产品提供免费上门保修保用服务。
2. 保修期内负责对所供产品实行包退、包维护保养，所有产品的维修、版本升级均为免费。
3. 驻场服务结束后，投标供应商应指定一名以上专业技术人员为招标人提供售后咨询和技术服务，提供7×10技术支持服务、5×8在线技术支持服务。故障报修的响应时间为2小时，如远程无法解决的，应在12小时内到达现场（特殊情况由双方协商）。
4. 保修期内，所有维修服务均为上门服务。中标人至少每半年一次定期派工程师到招标人现场维护和巡查，并免费提供正常使用下的维修及保养服务。
5. 应提供系统扩充、升级方面的技术支持服务。

（二）培训要求

1. 基本培训要求：针对系统管理人员提供产品的配置、操作及维护培训，主要内容为平台系统的基本结构、功能与特性，日常使用操作、维护与管理，常见故障的排除，紧急情况的处理等，直到业主工程师熟悉使用。培训时间不少于1天。
2. 用户使用培训：针对业务人员使用操作培训，主要内容为平台系统基本情况与功能介绍、实践操作，技术指导等。培训时间不少于3天。
3. 培训地点：平台系统安装所在地，具体由招标人确定，人数不限。
4. 维保期内提供不少于3次的集中培训服务。

# 八、 服务要求

由中标人与招标人双方协商确定，以签订合同所订立的付款方式进行付款。付款前，中标人需向招标人提供与合同金额等额的正式增值税专用发票。