

合同编号: H201203B

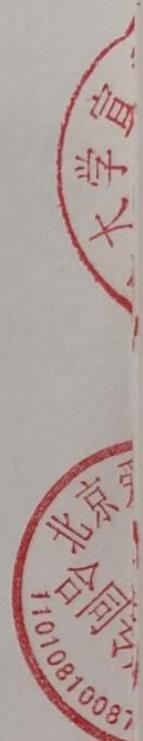
招标编号: TC200V0B2

首都医科大学宣武医院

整合业务数据搭建科研平台项目

合 同 书

甲方名称: 首都医科大学宣武医院
乙方名称: 北京爱医声科技有限公司
签订地点: 中国 北京



[首都医科大学宣武医院(以下简称“甲方”)]

其注册地址在：北京市西城区长椿街 45 号,

联系人：梁志刚，电话：010-83198999；

[北京爱医声科技有限公司(以下简称“乙方”)]

其注册地址在：北京市海淀区学清路 10 号 1 号楼 A 座 1903,

联系人：赵宁，电话：010-62622292。

首都医科大学宣武医院（以下简称“甲方”）就首都医科大学宣武医院整合业务数据搭建科研平台项目中所需信息化软件购置委托中招国际招标有限公司在国内公开招标（招标编号：TC200V0B2）。经评标委员会评定，北京爱医声科技有限公司（以下简称“乙方”）为该项目中标单位（附件二：中标通知书）。甲乙双方经过平等协商，根据《中华人民共和国合同法》等相关法律法规的规定，达成如下条款，并由双方共同遵守。

第一条 项目工作内容

首都医科大学宣武医院整合业务数据搭建科研平台项目。项目工作内容及技术要求详见附件一《技术要求及实施方案》。

第二条 项目完成时间

合同签订后 10 个月内完成软件供货、本地化改造、平台搭建及系统上线。

第三条 合同价格及付款方式

1、合同标的及合同总价（单位：人民币万元）

序号	产品名称	品牌/规格型号	数量	单价 (万元)	小计 (万元)
1	数据整合平台	爱医声/Aihealthx_zh v 1.0	1 套	80	80
2	临床数据分析系统	爱医声/Aihealthx_bi v 1.0	1 套	20	20
3	院内/院间数据联盟系统	爱医声/Aihealthx_lm v 1.0	1 套	20	20
4	基于区块链的数据全流程追踪系统	爱医声/Aihealthx_qu v1.0	1 套	20	20
5	智能随访系统	爱医声/Aihealthx_sf v 1.0	1 套	59	59
合同总价：大写人民币壹佰玖拾玖万元整（小写：¥1,990,000 元）					

合同总价中已包括软件产品供货、系统开发、安装、调试及培训、售后服务等费用。与本合同有关应征收的一切税费均由乙方承担。

2、付款条件

合同签订生效之日起 15 个工作日内，甲方向乙方支付合同总价的 50%，计人民币玖拾玖万伍仟元整（小写：¥995,000 元）；

项目整体验收合格后 15 个工作日内，甲方向乙方支付合同总价的 40%，计人民币柒拾玖万陆仟元整（小写：¥796,000 元）；

项目验收合格满一年后 15 个工作日内，甲方向乙方支付合同总价的 10%，计人民币壹拾玖万玖仟元整（小写：¥199,000 元）；

乙方应在甲方每次付款前向甲方提交相应数额的正式发票。

3、账户信息

【乙方账户信息】

单位名称：北京爱医声科技有限公司

纳税人识别号：91110108MA00AWHA9Q

地址及电话：北京市海淀区学清嘉创大厦 A 座 1903, 010-62622292。

开户行：上海浦东发展银行北京清华园支行

账号：91310154800002921

【甲方开票信息】

单位名称：首都医科大学宣武医院

纳税人识别号：121100004006883851

地址及电话：北京市西城区长椿街 45 号，010-83198899

开户行：北京银行总行营业部

账号：01090520500120109034255

开户名：首都医科大学宣武医院

第四条 质量保证

乙方应严格遵照本合同技术要求，按照国家有关标准和规定完成软件供货、系统开发、调试、部署、安装、培训及售后服务，按照合同规定向甲方提交符合技术要求的成果。同时，乙方为本项目提供 1 年期的免费维护服务，自系统验收合格之日起开始计算。

维护期内建立维护档案及维护热线，乙方提供 7×24 小时电话技术支持。除现场常驻人员可随时取得联系外，亦可通过热线电话向公司服务中心进行咨询。

若非现场维护无法解决，要求乙方现场服务时，乙方保证 2 小时内到达甲方现场。

第五条 提交成果

乙方应根据本合同附件一《技术要求及实施方案》的要求向甲方提供列示的项目成果。

所提供的文档类型和文档质量需符合软件工程规范，其中电子文档是成果不可分割的部分。除了产品本身外，在本项目建设过程中，形成的所有程序乙方必须向甲方提供源代码及源代码介质，程序代码归甲方所有，乙方此前已有产品的代码除外。

第六条 知识产权

1. 乙方应保证甲方在使用本合同项下的货物或货物的任何一部分时，免受第三方提出的侵犯其知识产权的索赔或诉讼。

2. 如果乙方所提供的货物，被第三方诉称侵犯了该方知识产权或任何其它权利，甲方以书面方式通知乙方，乙方应负责处理这一指控并应以乙方的名义向起诉方提出抗辩。由此可能产生的一切法律责任和经济责任由乙方承担。甲方将尽可能地对乙方抗辩给予协助，由此发生的费用由乙方承担。

3. 本合同项目成果以及由此现成的知识产权归甲方所有，包括但不限于著作权、专利申请权、专利权、非专利技术和其它智力成果的权利和利益。涉及乙方产品的知识产权归乙方所有。

第七条 保密条款

1. 甲乙双方在未征得对方同意前，不得将对方提供的有关合同或任何合同条文、计划或资料提供给与履行本合同无关的任何人。即使向与履行本合同有关的人员提供，也应承担保密义务。在验收前，乙方应向甲方提交与本项目相关的全部资料。

2. 对于在合同履行过程中所获得或了解的任何秘密，任何一方均应承担保密

义务，未经对方书面许可，另一方不得利用或披露。

违约金外，乙方有权要求甲方支付合同总价 10% 的违约赔偿金。如果合同继续履行，乙方履行本合同的期限应相应顺延。

第十条 不可抗力

1. 甲乙双方任何一方由于受不可抗力事件影响而不能履行合同时，履行合同的期限可以延长，延长的期限应相当于事件所影响的时间。

2. 任何一方因发生不可抗力事件履行合同受阻，应在发生不可抗力事件后 7 天内以书面方式通知对方，并于不可抗力事件发生后 14 天内将有关当局出具的证明文件用特快专递或挂号信寄给对方审阅确认。一旦不可抗力事件的影响持续 90 天以上，双方应通过友好协商达成进一步履行合同的协议，双方也可解除合同。

第十一条 仲裁

1. 合同实施过程中与合同有关的一切争端应通过友好协商解决。如果友好协商开始后 60 天不能解决，任何一方都可向北京仲裁委员会申请仲裁；

2. 在仲裁期间，除正在进行仲裁的部分外，本合同其它部分应继续执行。

第十二条 合同生效及其他

1. 本合同一式肆份，甲乙双方各执贰份，自双方签字盖章之日起生效。

2. 本合同附件为本合同不可分割的组成部分并与本合同具有同等法律效力。

3. 任何对合同条款的变更或修改均需双方同意并签订书面文件，该文件作为整个合同不可分割的组成部分，并具有同等法律效力。

4. 本合同的未尽事宜将友好协商解决，甲乙双方可以对本合同进行修改与补充，并在双方签署补充合同后生效，补充合同与本合同具有同等法律效力。

(以下无正文)

甲方：首都医科大学宣武医院

(签章)

授权代表(签字):

签订日期: 2020 年 12 月 3 日

乙方：北京爱医声科技有限公司

(签章)

授权代表(签字):

签订日期: 2020 年 12 月 3 日

1. 系统测试：系统开发工作完成后 5 个工作日内，由双方按照合同要求对系统总体功能和性能进行测试并出具测试报告。系统测试结果必须满足合同规定的技术和功能要求，如系统测试不合格，乙方在双方约定的时间内进行完善后进行下一次测试，与之相关的费用由乙方承担。

系统测试前，乙方应向甲方提供系统整体的测试方案。

2. 系统验收：系统测试合格后，乙方提出书面申请，并将与本合同有关的相关技术资料提交给甲方，经甲方确认材料齐全后，由甲方按照本合同及其附件要求，在 15 日之内组织完成验收，逾期不验收则视为验收合格。

验收标准以本合同约定的技术指标、软件工程规范和惯例为依据。

第九条 违约责任

1. 乙方因自身原因未能按照合同规定的时间提交成果，应向甲方支付误期赔偿费，每周按合同总价的 5‰ 累计，误期赔偿费的最高限额为合同总价的 5%，同时乙方应继续履行本合同所规定的义务。

2. 乙方因自身原因延迟交付成果超过 90 天，甲方有权终止合同，乙方应返还甲方已支付的合同款项。除前款所约定的误期赔偿费外，甲方有权要求乙方支付合同总价 10% 的违约赔偿金。

3. 本合同中约定的软件系统，如在设计和功能等方面不符合合同技术要求，或证实存在技术缺陷，甲方有权要求乙方在双方约定的期限内修补完善，如果仍不能达到合同要求，甲方有权终止合同，并要求乙方返还甲方已支付的合同经费，同时乙方应承担相关赔偿责任。

4. 系统保证期（系统验收合格起 1 年）内系统发生故障时，乙方接到甲方通知后，未能在规定时间内到达现场处理，乙方应向甲方支付延误赔偿费，每延误一次，赔偿合同总价的 1‰，延误赔偿费的最高限额为合同总价的 5%。

5. 甲方延期付款时（正当拒付除外），应向乙方支付延期违约金，每周按延期付款金额的 5‰ 累计，延期违约金的最高限额为合同总价的 5%。

6. 甲方延期付款时间超过 90 天，乙方有权终止合同，除前款所约定的延期违约金外，乙方有权要求甲方支付合同总价 10% 的违约赔偿金。如果合同继续履行，乙方履行本合同的期限应相应顺延。

附件一 技术要求及实施方案

一、 软件功能及技术要求

(一) 基于区块链的数据整合平台

平台以固定的时间间隔从本院已有的 **CDR** 数据库中抽取绝大部分有用数据，并将重要业务资产与数据流通行为进行区块链上链操作，保证抽取来的数据与源数据完全一致。在此基础上以可视化方式完成数据的清洗与转换作业（**ETL**），包括数据的规范化、标准化、类型转化与数据去重等。同时平台支持非结构化电子病历基于自然语言识别的结构化转换，对疾病名称、临床症状、药品名称、检验检查等医学术语进行较高识别率的识别抽取，并且在识别的过程中自动通过机器训练达到模型优化以进一步提高识别率。平台也能通过词表匹配+实体识别算法解析的方式识别病历中的各类医学术语与临床指标，用于构建基于知识图谱的医学术语集维护管理工具。另外平台支持 **OCR** 文字识别录入功能，可以用于将的纸质病历进行电子化和结构化存储。平台可以抽取基于门诊医生工作站的结构化门诊电子病历系统的数据。系统还支持各类型病历报告表（**CRF**）的定制与数据抽取，并以结构化方式整合进平台数据库。

平台可对数据资产进行标准化定义，规范管理数据资产、算法资产、课题资产 3 类数据资产。

平台将重要的数据资产和数据流通行为进行上链，支持行程个人资产库。

平台可将两方及以上参与方进行课题共建时，提供联邦学习技术能力，进行联邦建模，训练，输出结果。

1. 数据采集

支持 **CDR** 中 70% 数据采集。

支持 **T+1** 采集间隔。

支持重要数据资产及数据审批、使用流程通过区块链技术进行上链，支持采集过程，发布、申请、使用、管理和展示全流程跟踪管理。

2. 数据清洗转换（ETL），支持以下清洗和转换

空值处理：

将信息表中的属性分为数值属性和非数值属性来分别进行处理。如果空值是数值型的，就根据该属性在其他所有对象的取值的平均值来填充该缺失的属性值；如果空值是非数值型的，就根据统计学中的众数原理，用该属性在其他所有对象的取值次数最多的值(即出现频率最高的值)来补齐该缺失的属性值。

规范数据格式：

基础数据标准化是的信息化重要工作之一，建立集中、规范统一的基础数据标准，是保证医院信息化系统正常运行的前提条件。此外，统一编码也是医院的一项重要的基础管理工作对医院管理标准化具有促进作用。通过建立标准化制度，使各业务科室能够协同工作，消除重复性劳动，大幅度提高工作效率。

拆分数据：

把逻辑上是统一整体的数据分割成较小的、可以独立管理的物理单元进行存储，以便于重构、重组和恢复，以提高创建索引和顺序扫描的效率。拆分数据可以使数据库的开发人员和使用者具有更大的灵活性。

数据类型转换：

通常分为自动类型转换与强制类型转换。自动类型转换是编译器默默地、隐式地进行的一种类型转换，不需要在代码中体现出来；强制类型转换是程序员明确提出的、需要通过特定格式的代码来指明的一种类型转换。

数据过滤：

数据的过滤在数学建模中占有很重要的地位。它是数学建模的第一步，我们只有得到好的数据才能保证得出的结果的真实性与准确性。而在实际的问题中我们遇到的数据量往往是巨大的。为了保证我们所用的方法能够在原始数据的支持下得以实现，我们必须对数据进行过滤，使得解决方法简单化。同时，我们又要保证过滤出来的数据具有代表性，使得到的结果更加准确与真实。

字段合并：

可以认为是拆分数据的逆过程，把可进行连接的两个或多个字段按照一定的规则进行合并，以减少数据维度，提高统计分析效率。

重复数据删除：

一种数据缩减技术，通常用于基于磁盘的备份系统，旨在减少存储系统中使用的存储容量。它的工作方式是在某个时间周期内查找不同文件中不同位置的重复可变大小数据块。重复的数据块用指示符取代。高度冗余的数据集从数据重复删除技术的获益极大，用户可以实现 10 比 1 至 50 比 1 的缩减比。而且，重复数据删除技术可以允许用户的不同站点之间进行高效，经济的备份数据复制。

数据标准化：

主要包括数据同趋化处理和无量纲化处理两个方面。数据同趋化处理主要解决不同性质数据问题，对不同性质指标直接加总不能正确反映不同作用力的综合结果，须先考虑改变逆指标数据性质，使所有指标对测评方案的作用力同趋化，再加总才能得出正确结果。数据无量纲化处理主要解决数据的可比性。数据标准化的方法有很多种，常用的有“Min—Max 标准化”、“Z-score 标准化”和“按小数定标标准化”等。经过标准化处理，原始数据均转换为无量纲化指标测评值，即各指标值都处于同一个数量级别上，可以进行综合测评分析。

在线可视化实现数据的清洗转换：

将上述的 ETL 过程全部以可视化方式呈现，用户可以便捷、直观地进行所需要的复杂操作。

3. 病历结构化

➤ 基于规则的电子病历结构化

支持科室、疾病名、药名等主要信息抽取，准确率需达到 95%。

➤ 基于 NLP 的电子病历结构化

支持疾病名称、临床症状、药品名称、检验检查等医学术语进行识别抽取。

非特定科室的抽取准确率达到 80% 以上。

指定科室通过模型优化训练准确率达到 90% 以上。

支持通过词表匹配+实体识别算法解析病历中的医学术语。

支持通过词表匹配+实体识别算法解析病历中的数值指标、时间频率等内容信息。

将文本形式的病历转化为标准化的表格形式，用于医学研究等后续相关应用。

支持 OCR 文字识别录入功能：即对文本资料进行扫描，然后对图像文件进行分析处理，获取文字及版面信息的过程。

支持自然语言识别处理，并构建基于知识图谱的医学术语集维护管理工具。

4. 医学知识图谱术语维护

支持面向专科医学术语标准化的知识图谱维护。

可用于协助医生对常见医学术语的不同用语表达进行在线维护。

5. 病例报告表(Case Report Form, CRF)

病例报告表(Case Report Form, CRF)是一套集 ETL 数据抽取、表单设计器、动态表单为一体的病例报告表单管理系统，通过该系统可以有效的整合用户各类数据，快速应对用户新需求及变更，从多维度、多切入点梳理报告数据，为整个 CRF 数据表单管理提供高效、易用、安全、便捷的环境。

CRF 模型：用于在线管理、设计病历报告的工具，可以通过数据字典、模块组、模块、布局元素、自定义控件等组装表单，结合动态数据源绑定表单内容与数据源关系，已达到自动录入的目地。

CRF 字典：CRF 模型的重要组成元素，可以将系统数据规范整理，方便复用。CRF 字典包含文本、数值、日期、时间、单选、复选、单选下拉、多选下拉、表格等多种类型，而且可配置每种类型的特性数据。

CRF 字典参数：CRF 模型的最小颗粒度，为 CRF 字典做数据支撑。同样也是起到规范、整合、复用的作用。

CRF 数据查询：通用的数据查询模块，所有已录入的 CRF 表单都可在此查询，该模块具备表单切换、动态条件查询、自定义列、导出、打印等功能。这只是系统查询 CRF 数据的一个入口，还可通过 CRF 数据中心和患者二个维度查询。

组别管理：在 CRF 管理中，分组是必不可少的。通过组别管理可以将医生、护士以及负责的 CRF 表单进行绑定管理。以便后续在使用时能更有效的查找、管理、书写。

患者入组：对组级别内患者进行管理，可以进行入组、退组患者等操作。使 CRF 表单管理中根据不同患者维护不同的 CRF 模板的关键。

CRF 数据中心是 CRF 数据录入的入口之一，可以根据分组、ID 等模式筛选患

者，对患者所具有的 CRF 模型进行数据录入，该功能是根据 CRF 模型通过 CRF 动态表单技术组装而成，具有丰富的录入控件，符合用户需求的数据模型。数据源可支持访问各系统数据通道，已整合了该患者的所有相关医疗数据，只需要点选就可以完成数据的自动录入。对于只有单次数据的可自动完成录入，大大提升了病历书写效率。

（二）临床数据分析系统

在完成了数据整合以后，对临床数据进行系统性分析成为了可能。系统大致可分为用户管理层、项目管理层与数据模型层三个部分。用户管理层中，使用者可以维护更新自己在系统里的个人信息，并在系统内根据自己所分配到的角色权限进行相应的操作，系统管理员可以对这些角色权限进行相应的调整与分配，角色目前分为科研人员、科主任、科研部、信息科与分管院长，不同角色对应不同的系统权限。在项目管理层中，使用者可以在科研系统门户查看与下载最新的公告与文档，收发各类课题相关信息，并编辑自己的科研成果发布状态。同时根据不同的角色权限，对所有的项目进行流程化规范化管理，包括项目的申报、项目的变更、项目的结束、项目成员的管理、项目经费的管理、项目评审的管理和项目成果的管理。在数据模型层，支持使用者根据各类临床数据建立对应的规范化标准化的科研数据子中心，在该子中心下使用者可以进行各种形式的数据浏览与数据的筛选查询，以及根据不同需要进行基础或复杂的统计分析应用和数据可视化输出。

1. 用户管理

➤ 用户

支持用户新增。

支持用户删除。

➤ 登录

输入用户名、密码和验证信息进行登录。

支持验证码校验。

➤ 退出

退出到登录页面。

➤ **个人信息维护**

修改个人资料（学历、学位、职位等）。

修改个人密码。

2. 权限管理

➤ **科研人员权限**

接受信息。

查看课题进展。

申报课题/项目。

申请课题/项目相关数据使用权限。

提交信息。

科研数据使用和管理。

查看个人课题/项目数据使用情况，并生成分析报表（相关数据还支持周报、月报、季报、年报等阶段性查询方式）。

➤ **科主任权限**

课题审查。

课题/项目相关数据使用权限审查。

查看本科室课题/项目进展，并生成分析报表（相关数据还支持周报、月报、季报、年报等阶段性查询方式）。

支持自定义科室核心数据，并按照不同的保密程度进行审批。

➤ **科研处权限**

课题审查。

课题预警。

组织课题/项目申报。

跟踪课题/项目进展情况。

查看各科室课题/项目进展，并生成分析报表（相关数据还支持周报、月报、季报、年报等阶段性查询方式）。

支持定义数据中间表。

➤ **信息中心权限**

数据恶意使用预警。

跟踪课题/项目数据使用情况。

通过课题/项目相关数据使用权限。

查看各科室课题/项目数据使用情况，并生成分析报表(相关数据还支持周报、月报、季报、年报等阶段性查询方式)。

➤ **分管院长权限**

申报课题审查。

查看各科室课题/项目进展，并生成分析报表。

查看各科室课题/项目数据使用情况，并生成分析报表(相关数据还支持周报、月报、季报、年报等阶段性查询方式)。

3. 科研系统门户

➤ **通知公告**

查看医院、科研部（科教部）发布的最新公告信息。

查看设置可见并邀请加入的课题/项目。

➤ **文件下载**

下载科研部（科教部）发布的文档及模板文件。

文件大小不超过 20M。

➤ **收发消息**

接受各审批人员发送给科研人员的消息。

给其他课题/项目负责老师发送消息。

➤ **预警信息**

接收课题/项目时间进度等预警消息。

接受恶意使用数据预警信息。

➤ **我的科研动态**

个人科研成果展示（如课题/项目、论文、著作、专利、获奖情况等）。

4. 科研项目管理

➤ **科研人员/科主任对项目业务管理**

科研项目数据权限申请。

科研项目成员管理。

结果数据管理。

课题/项目申报。

课题/项目立项。

课题/项目变更。

课题/项目中检。

课题/项目结项。

➤ 科研部对科研项目的规范管理

项目发布。

项目申请。

专家评审。

研究过程管理。

经费管理。

课题的申报。

➤ 科研部对科研成果的规范管理

学术论文管理。

学术著作管理。

科研专利管理。

科研奖项管理。

科研失信行为记录。

5. 数据中心模型

➤ 梳理各个主题数据，建立标准化、规范化的科研数据中心

病人基本信息。

门诊病历。

处方记录。

检查数据。

治疗数据。

护理数据。

6. 数据浏览

支持数据目录浏览。

支持数据结构浏览。

支持数据内容浏览。

7. 数据筛选提取

支持单表数据提取。

支持字段过滤。

支持多表联合。

支持多条件过滤。

支持数据转换。

支持列拆分和合并。

支持 SQL 语句提取。

8. 基础统计分析

支持求和计算。

支持平均值计算。

支持中位数计算。

支持众数计算。

支持标准差计算。

支持方差计算。

支持极值计算。

支持相关分析计算：

相关分析是研究两个或两个以上处于同等地位的随机变量间的相关关系的统计分析方法。例如，人的身高和体重之间；空气中的相对湿度与降雨量之间的相关关系都是相关分析研究的问题。相关分析与回归分析之间的区别：回归分析侧重于研究随机变量间的依赖关系，以便用一个变量去预测另一个变量；相关分析侧重于发现随机变量间的种种相关特性。相关分析在医疗、工农业、水文、气象、社会经济和生物学等方面都有应用。

支持卡方检验计算：

卡方检验是用途非常广的一种假设检验方法，它在分类资料统计推断中的应用，包括：两个比率或两个构成比比较的卡方检验；多个比率或多个构成比比较的卡方检验以及分类资料的相关分析等。

支持 t 检验计算：

t 检验，亦称 student t 检验（Student's t test），主要用于样本含量较小（例如 $n < 30$ ），总体标准差 σ 未知的正态分布。t 检验是用 t 分布理论来推论差异发生的概率，从而比较两个平均数的差异是否显著。

9. 复杂统计分析

支持主成分分析：

主成分分析（Principal Component Analysis, PCA）通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的变量，转换后的这组变量叫主成分。在实际课题中，为了全面分析问题，往往提出很多与此有关的变量（或因素），因为每个变量都在不同程度上反映这个课题的某些信息。

支持聚类分析。

聚类分析指将物理或抽象对象的集合分组为由类似的对象组成的多个类的分析过程。它是一种重要的人类行为。聚类分析的目标就是在相似的基础上收集数据来分类。聚类源于很多领域，包括数学，计算机科学，统计学，生物学和经济学。在不同的应用领域，很多聚类技术都得到了发展，这些技术方法被用作描述数据，衡量不同数据源间的相似性，以及把数据源分类到不同的簇中。

支持关联分析：

关联分析又称关联挖掘，就是在交易数据、关系数据或其他信息载体中，查找存在于项目集合或对象集合之间的频繁模式、关联、相关性或因果结构。或者说，关联分析是发现交易数据库中不同商品（项）之间的联系。

支持分类分析：

分类分析，简单地说就是把数据分成不同类别。分类方法的选择还和数据量的大小很有关系，如果数据量比较小，也就是说用于判断数据分布的样本量可能不足的情况下，选择一些相对简单的方法会更好一些，因为如果选择复杂的方法，很容易在样本不足的情况下产生过拟合的情况（样本不足而方法复杂，可能很多

噪音和误差会被误认为是特征); 反之, 如果有大量的数据, 那么可以更优先选择复杂的、集成的方法, 这样能提升模型的整体性能。

支持回归分析:

在统计学中, 回归分析 (**regression analysis**)指的是确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。回归分析按照涉及的变量的多少, 分为一元回归和多元回归分析; 按照因变量的多少, 可分为简单回归分析和多重回归分析; 按照自变量和因变量之间的关系类型, 可分为线性回归分析和非线性回归分析。

10. 数据/图像导出

支持饼形图导出。

支持折线图导出。

支持柱形图导出。

支持条形图导出。

支持散点图导出。

支持列表数据导出。

(三) 院内/院间数据联盟系统

平台应用区块链的分布式存储技术与联邦学习的跨平台建模的思想, 不但支持院内多科室间的数据共享, 更能支持院际间的多中心数据共享, 而且在多中心的协同工作中, 本院的数据并没有离开本地的数据中心, 院际多中心间以加密方式交换模型参数和计算结果的形式进行联合建模、训练和测试, 形成融合的机器学习模型, 满足医疗科研大数据需求, 将从根本上解决数据共享和数据安全问题。在此基础上, 无论是院内的数据共享系统还是院际的数据联盟系统, 都可以实现对项目的申报、查看、协同、发布等全流程管理。

1. 院内数据共享

支持院内多课题/项目协同及数据共享。

支持申请/加入开放性院内多课题/项目。

支持申请查看和使用开放性院内多课题/项目数据,

支持基于区块链进行院内数据使用的全流程追踪。

2. 院际数据联盟

支持院外多临床中心课题/项目协同。

支持院外多临床中心通过联邦学习技术在数据不出本院的基础上实现数据共享。

支持申请/加入开放性院外多临床中心课题/项目。

支持申请查看和使用开放性院外多临床中心课题/项目数据。

支持数据可基于区块链技术进行医院之间数据使用的全流程追踪

(四) 基于区块链的数据全流程追踪系统

平台应用了区块链的数据可溯性特点，对异常数据和非法数据都可以进行实时地自动检测，并将检测的结果自动上报系统管理员，同时保持跟踪处理状态。同时基于区块链的分布式存储，对数据的完整性保持了高度的检测。在将数据由CDR整合进科研平台的过程中，系统自动对数据进行脱敏处理，所有涉及到患者个人隐私的数据都将被自动过滤掉，保障了患者的隐私权不收侵犯。在院间与院际内进行数据传输的工作中，所有传输的参数都采用了同态加密，确保通信全过程的数据安全。在此功能的保障下，可在实现院际科研合作中，在数据不出门的情况下，也可以共同训练数据模型，完成科研任务。系统可以对以上全流程以任务调度的方式进行可视化的管理操作，最大化提高系统管理的便捷性与高效化。

1. 数据质量控制

异常数据的自动检测：

检测离群点或异常值是数据挖掘的核心问题之一。数据的爆发和持续增长以及物联网设备的传播，使我们重新思考处理异常的方式以及通过观察这些异常来构建的应用场景。进行异常值检测的方法有很多，主要有以下几种大的分类：传统统计方法、机器学习方法、深度学习方法、其他方法等。

非法数据的自动检测：

基于对程序源代码的静态分析方法，针对程序中的非法计算故障，提出了一种自动测试方法。以非法计算故障的表现形式为基础，将静态分析程序源代码和

识别非法计算的过程模型化，建立了静态检测非法计算所需的计算模型，将非法计算故障静态检测过程自动化，给出了非法计算自动测试算法。该方法通过对程序源代码的静态分析，实现了非法计算故障检测的自动化。实验结果表明，该自动测试方法能有效地提高检测非法计算的准确率和自动化水平。

异常与非法数据的自动上报及跟踪处理。

将检测到的异常数据与非法数据自动上传到系统管理员处，并实时记录处理结果。

数据完整性的检测：

随着存储数据的增加，访问人数的增多，增长了对数据完整性的潜在需求。为了确信信息系统内的数据和网络是否处于已知的正确状态，以一种基于 B/S 结构的数据完整性检测(简称 DNI)方法，研究如何利用 DNI 保护信息系统安全运行，提升信息系统安全的效率，并采用 Socket(套接字)技术来编程，实现一种基于 Client/Server 的数据完整性检测系统。

2. 数据脱敏与加密

对敏感信息进行脱敏处理，确保患者隐私信息（比如姓名、地址、手机号、证件号等信息）不出现在科研数据中心，保护患者隐私。

数据传输过程进行同态加密，确保通信过程的数据隐私安全。

支持运用联邦学习技术在保证数据不出门的情况下，共同训练算法模型。

3. 任务调度

支持数据采集、数据清洗、数据转换、数据质量控制等功能按照一定调度顺序定时执行。

任务调度间隔最短支持 1 小时。

支持失败重试。

支持 DAG 图调度任务：

DAG 数据结构跟踪基本块中值和变量的计算和赋值；块中使用的来自别处的值表示为叶子结点；值上的操作表示为内部结点；新值的赋值表示为将目标变量或临时变量的名字附加到表示赋值的结点上；DAG 可用于对数学和计算机科

学中得一些不同种类的结构进行建模。

（五）智能随访系统

智能随访只限应用于科研平台的随访业务，不涵盖院内普通随访业务。

1.技术要求

医疗语音识别服务需求： 提供人工智能语音自动随访功能，可根据患者的回答问题语音进行语音识别，需要在院内搭建私有云医疗语音识别服务器引擎，医疗语音识别服务具有很好的抗噪性，可以识别复杂声学特定环境噪音判断所处环境（如鼠标击声、键盘输入声、咳嗽声、呼吸声、饮水声）等，帮助分析患者当前状态。

平台语音识别技术识别率达 97.5%及以上，具有百万级别的医学专有名词，覆盖医院各科室的病历术语、医学符号、医学单位、医学文献，针对医疗场景进行算法优化。

医学自然语言理解服务需求： 需要在院内搭建医学自然语言理解服务器引擎，对识别的自然语言进行处理，自动对应随访问卷的对应标准选项，服务可对同义文本进行基于自然语言文本相似度结果打分，可以对非同义词进行基于自然语言文本相似度结果打分，可以基于中文关键字规则进行同义短文本识别，可以基于拼音关键字规则进行同义短文本识别。

病程管理： 平台支持对患者全病程的管理，实现对患者的全病程个案追踪，通过人工智能技术，智能发起追踪随访、复诊提醒、病情关怀等业务。

随访方式： 平台智能随访系统需支持 AI 语音随访、AI 微信随访、AI 短信随访、人工电话等多种随访方式。

2.随访任务

提供随访任务快捷处理入口；

提供随访异常情况的展现和处理入口。

提供患者诊疗信息的查阅及完成随访问卷的全过程记录及宣教。

提供对已完成随访的展现及查询。

3.患者管理

出院患者管理: 针对出院患者可基于出院信息（患者姓名，性别，年龄，籍贯，身份证号，住院 ID，住院号，入院时间，出院日期，出院科室，出院病区，主治医生，住院费用，出院诊断，手术名称，联系人姓名，联系人电话）等条件进行患者筛选和检索。

门诊患者管理: 支持查看历次门诊诊疗信息，包括就诊记录、门诊诊断、检验、检查等；

按照时间顺序对历次门诊诊疗记录进行排列，同时可查看门诊就诊次数。

基本信息维护: 基本信息维护内容：患者姓名，性别，年龄，籍贯，身份证号，住院 ID，住院号，入院时间，出院日期，出院科室，出院病区，主治医生，住院费用，出院诊断，手术名称，联系人姓名，联系人电话。

个人随访主页: 患者基本信息，患者随访计划，患者住院诊疗记录，支持对当前选择的患者进行快捷的手动电话拨打、随访问卷/表单及满意度问卷的推送等操作。

4.随访管理

随访计划管理: 支持按出院患者来源、科室/病区、主治医生、出院诊断、用药记录、手术名称等附加筛选条件进行患者筛选，随访人员、计划随访时间、抽取比例的设置；

支持随访任务、宣教任务、提醒任务在同一随访路径下并设置。

时间条件: 系统可根据门诊、住院制定随访路径（支持制定随访日期，如门诊后 X 天、出院后 X 天等）

随访方式选择：电话随访，微信随访，短信随访，AI 电话随访

支持实时或指定时间点推送随访问卷/表单。

状态过滤: 转归情况为‘死亡’的患者自动不纳入随访任务。

支持过滤无效号码的患者。

随访患者筛选引擎: 所有需要随访的患者，均通过预先设置的随访计划条件，每天定时自动筛选符合条件的患者，支持按手术、疾病、用药情况筛选，支持按比例、数量抽取患者并生成相应的随访任务。

随访任务推送引擎：支持实时或指定时间点推送随访问卷/表单。

随访任务管理：随访任务列表：今日工作量（总数）、今日未完成、今日已完成、过期任务、异常随访（已完成问卷）； 用户可清晰看到随访方式、实际执行方式、回复状态、状态描述、随访指标异常状态、随访总体评价等

随访结果跟踪：可直观的展现患者的随访情况，例如是否回复，是否异常等；患者回复无异常情况，支持随访任务自动完成。支持对随访结果异常的患者进行回访沟通，明确异常情况，记录结果。

支持微信随访、短信随访下，患者回复有异常情况下，系统自动生成提醒内容并通知随访人员（短信），自动生成宣教任务并推送，或转交给其他异常处理人员。

5.AI 随访平台

应用 AI 技术，依据随访计划，自动对筛选的患者以电话的方式进行随访（通话过程录音），若患者随访异常，可自动增加下一次随访任务。随访过程中，患者有疑问，可以依据设定的问题触发路径自动推送健康宣教信息。

智能随访根据随访计划自动给患者拨打电话，通过与患者多轮人机对话完成对随访问卷的填写。整个随访过程可实时回放，每轮对话 NLP 需提供可信度分值，并提供人工对低可信答案的修正入口。

6.知识库管理

随访表单维护：提供科室/病区随访使用的随访表单分类维护。

健康宣教维护：提供对宣教中心使用的健康宣教分类维护；支持自主编辑内容、格式、排版，可引用图片、视频等。

提醒短语维护：提供对宣教中心使用的健康宣教分类维护；支持自主编辑内容、格式、排版，可引用图片、视频等。

满意度调查表维护：提供满意度调查使用的调查表分类维护；支持自主编辑题目，设置多种题型，设置选择项、题目分值、权重等。

随访路径维护：提供随访路径（随访日期，随访内容）预先分类设置。

健康宣教维护：提供宣教路径（宣教日期，宣教内容）预先分类设置。

表单编辑器：支持编辑题目，设置多种题型，设置选择项、题目分值、权重等。

表单题库：提供随访及满意度表单总单体维护，在表单编辑器中可直接应用。

7.统计中心

随访工作量统计：支持按科室、任务、个人多角度及时间范围统计随访工作量；支持数据导出，报告结果打印。

随访表单统计：支持按科室/病区、时间范围统计随访当个表单使用及恢复情况，并支持数据导出，报告结果打印；支持按科室/病区、计划、时间等对单个表单中使用的某个单体进行患者回复内容统计分析，并支持打印。

宣教统计：支持根据病区/科室、时间范围等按个人对宣教发送量，发送途径，患者查阅量进行统计分析；支持数据导出，报告结果打印。

短信统计：支持按病区/科室、时间对科室/人员进行短信发送量统计；支持数据导出，报告结果打印。

满意度统计：支持按患者、调查表内容及类别题、调查场景、异常情况及任务统计满意度情况；支持数据导出，报告结果打印。

二、项目验收文档

交付文档验收要求：

项目的工作内容及成果文档的提交应覆盖以下内容，电子文档是成果不可分割的部分。

1、系统的体系架构及描述。

2、提供的其它技术手册，包括：

需求分析报告（含软件功能需求与数据要求）；

系统实施方案；

软件培训资料；

程序安装维护手册；

软件使用操作手册；

项目验收报告。

附件二 中标通知书

中标通知书

北京爱医声科技有限公司：

在我公司组织的首都医科大学宣武医院整合业务数据搭建科研平台项目（招标编号：TC200V0B2）中，经评标委员会评标，确认贵公司为本项目的中标单位，中标金额为人民币 1,990,000.00 元(壹佰玖拾玖万元)。

请贵公司接此中标通知后按招标文件规定与采购人签订合同。

特此通知。

