11月6日上午鹏城实验室和云安宝公司在鹏城实验室召开了《面向隐私保护的大数据挖掘平台》项目申报讨论会，会上根据申报指南把研究任务和考核指标进行了分解，如下。请各合作单位参考以下内容，根据自己的优势尽快编制可行性研究报告。

研究基于程序浮动的大数据挖掘与隐私保护技术。

1. 研究隐私保护前提下的数据分析和模型训练工作台体系结构。（数据挖掘模型）

考核指标：

1. 支持 3 种以上格式数据文件导入，亿数量级数据分析结果秒级响应
2. 支持对数据分析模型的自动化选择与自动化模型参数调优
3. 支持适配 3 种以上主流大数据平台重要组件
4. 研究对自有数据和外部数据源融合分析场景的隐私保护以及数据分析浮动程序的可视化辅助和自动生成技术。

考核指标：

1. 支持对自有和外部数据源融合分析场景的隐私保护
2. 项目完成时须研制完成支持多租户数据隔离与访问控制，支持多租户协同工作流绘制、模型训练、可视化报表生成。
3. 研究数据传递中的反隐私隐藏技术。

3.1研究基于自然语言处理技术、图像文字检测与识别等智能反隐私隐藏技术，可对包括结构化、半结构化和非结构化数据进行敏感数据识别和脱敏；

3.2研究对图片隐写，模型隐藏敏感数据等行为的智能识别与审查；

3.3形成可供应用调用和集成的引擎SDK。

4.研究面向数据交易场景的隐私保护技术。

3、4考核指标：

1. 支持识别常见个人身份信息等敏感数据；
2. 支持识别隐匿在大数据平台中的敏感数据，
3. 支持识别和还原隐匿在训练模型中的敏感数据（GAN？）
4. 研制完成 5 种以上的敏感数据动态脱敏方式
5. 研究基于白名单机制的行为和内容审核技术，对数据访问、交易和分析等行为的留痕审计，研发安全的数据交易流转系统；
6. 面向政务数据开放场景，聚焦教育、交通、环境、医疗、商业等重点领域，提出适用的数据隐私保护体系。

其他考核指标：

项目执行期内新申请发明专利或软件著作权 10 项以上；发表高水平学术论文不少于 5 篇；向国内外标准化组织提交标准草案不少于 2 项。