附件2

贵州省农村信用社联合社数据资产管理平台及数据血缘分析平台原型测试内容

一、系统技术参数验证

（一）系统架构：产品需支持信创部署方案，操作系统包括统信、麒麟，CPU架构包括海光X86、鲲鹏ARM，中间件包括宝蓝德或东方通，数据库包括达梦、金乌或高斯数据库。应用系统须支持两地三中心架构下同系统多应用的安装部署、在线运行和同时使用，不受授权码（License）、用户数量及部署主机CPU数量等因素的限制。在逻辑层面和物理层解耦，采用REST/GraphQL/gRPC等标准化接口，存在异步通信采用消息队列或事件驱动。平台满足我社多级法人（省联社、审计中心、法人社/行）体系下业务管理和使用，能支持我社管辖范围内发起设立机构（如：村镇银行）的使用并满足应用和数据安全隔离性要求。

（二）技术选型：选用的技术栈成熟、稳定、与项目需求高度匹配、满足我社《开源软件推荐产品目录》要求，且具备良好的扩展性和兼容性，包括不限于社区活跃度、企业支持、行业采用率、兼容性、符合行业趋势方面。

（三）性能表现：根据性能要求，在预期的并发用户情况下，系统平均响应时间和最大响应时间合理，系统响应时间满足在1.5秒以内。其中：请求转发时间不超过0.2秒，其他业务处理响应时间不超过1.3秒。支持实时联机业务并发用户数不小于500户。稳定性与容错能力。系统运行过程中CPU、内存等资源使用率保持在合理范围内，在业务最高峰或批处理运行期间CPU占用率和内存使用率不应超过70%，内存使用率不超过80%。支持7\*24小时不间断运行。

（四）前台兼容性：主要测试平台对各个浏览器的兼容性，我社浏览器主要以IE8及以上版本、火狐、谷歌浏览器、360浏览器为主且满足分辨率要求。支持国际化（i18n）。

（五）参数化程度：参数配置功能可以通过页面配置实现而非硬编码，配置项设置合理。前端的展示风格可定制，支持低代码/拖拽式配置，业务规则支持规则引擎，多数据源支持动态切换。

二、系统业务功能演示

（一）数据资产管理平台

1.元数据管理。支持各主流信创数据库，如GaussDB、TDSQL、达梦、OceanBase等的数据源接入，支持传统数据库，如oracle、mysql、PostgreSQL的数据源接入，支持JSON报文、CVS文件、DAT文件等数据源接入。能够实现源系统的码表采集，自动实现源系统元数据与数据标准的对标比对及贯标检查。

2.数据标准管理。支持统一的词根词库管理能力，可进行用语、单词、域、代码等数标不同领域的分类管理。支持将基础数据标准与源头业务系统进行自动对标分析，形成贯标情况报告，支持将指标数据标准与BI报表、监管报表进行自动对比分析，形成贯标情况报告。

3.数据质量管理。支持进行数据质量评分规则设定，支持动态进行源头数据质量评估、问题发现、问题认责、问题清单分发、问题跟踪等功能，支持进行数据质量在线考核。

4.库表模型管理。支持进行数据库逻辑模型、物理模型设计及开发，并生成DDL与E-R图，支持导入源体系数据库表结构并进行逆向工程，逆向为模型数据，支持与数据标准进行对标分析联动。

5.数据安全管理。支持以智能推荐+人工确认审核的方式对源头业务系统、数据中台系统、数据应用系统相关数据进行安全分类分级，支持对敏感数据进行识别、管理数据脱敏加密策略。

6.数据资产管理。支持以驾驶舱、目录、地图等形式展示企业全量数据资产，支持进行数据资产的智能化识别与盘点，支持维护、上下线数据资产，支持数据资产的编录与打标。

7.其他特色功能。

（二）数据血缘分析平台

1.全域数据采集。多源异构数据支持覆盖传统数据源（如 MySQL、Oracle、Hive、Spark、Kafka、BI工具等）、新兴技术栈（如 ClickHouse、Doris、TiDB等）、主流信创数据源（如GaussDB、TDSQL、达梦、OceanBase等）。

2.智能解析引擎。支持国内外市场主流、信创等大部分数据库方言的语法解析、SQL解析，自动生产表、字段级血缘，兼容复杂嵌套查询。支持Java、Spark SQL、PySpark等代码中的操作逻辑识别，实现代码解析。支持解析帆软等主流BI报表的关联关系解析。支持对接我社ETL工具，实现解析抽取、转换。支持手动补全血缘断点和调整数据血缘关联关系。

3.自动化血缘解析与可视化血缘图谱。支持复杂 SQL 语句（如存储过程、嵌套查询、CASE-WHEN 逻辑）的字段级血缘提取，准确识别字段间的直接依赖和间接依赖关系。自动解析Java、Spark SQL、PySpark等代码中的数据处理逻辑，生成算子级血缘图谱，定位数据转换的具体步骤。对于可视化 ETL 工具和 API 接口，通过元数据采集或流量分析实现血缘自动补全，减少人工干预。支持分钟级实时血缘更新，通过监听数据源变更事件或调度系统日志，动态刷新血缘关系，确保数据链路与业务变化同步。提供活跃血缘与静默血缘区分能力，自动识别长期未更新的“僵尸”链路，辅助资源优化与风险预警。提供模式级、实体级、表级、字段级多层级粒度血缘视图、多层级展开上下游数据关系，标记区分数据流向，支持三维动态图谱或热力图展示核心数据链路的繁忙程度，帮助快速定位高价值或高风险节点。

4.影响分析。支持正向影响分析，如展示字段或表的变更，自动计算下游受影响的指标、报表、API 接口数量，并生成影响范围报告。支持反向溯源分析，针对异常数据，快速定位其上游来源，如某个 ETL 任务或数据源，并展示历史变更记录，帮助排查问题根因。

5.其他特色功能。