

第五章 采购需求

序号	产品名称	指标项	技术要求
1	数据集成平台	适配多种数据资源	<p>1、▲支持适配各种关系型数据实时与异步接入，包括但不限于 Oracle、MySQL、SQL Server、PostgreSQL、MongoDB、Sybase、Teradata 和 DM、GBase、GaussDB；</p> <p>2、▲支持文件搬移与文件解析功能；支持文件夹或目录增量同步功能；支持针对文件夹的子文件夹及文件增、删、改操作的实时增量同步功能</p>
		基础功能	<p>3、支持数据探查功能，即支持对多数据源的表进行数据探查，提供不少于 3 种探查要素，至少包括数据格式、数据值范围、数据值分布。</p> <p>4、ETL 组件库支持数据抽取、关联、排序、脱敏、清洗转换、聚合、装载等功能。</p>
			<p>5、支持邮件告警、短信告警功能，支持自定义条件来发送邮件或短信告警。</p>
			<p>6、支持批量创建作业功能；支持在 ETL 任务同步时，批量作业通过一个任务管理多张表的数据同步；支持同步至目的库时增加入库时间。</p>
		易用性	<p>7、▲支持异构数据库在无目的表的情况下自动建表功能。</p> <p>8、▲支持自动断点续传功能。当遇到网络故障或服务器故障修复后，自动重启中断的交换任务，从故障点续传数据。</p>
		传输性能	<p>9、▲结构化小数据（单行 200 字节），至少满足一种关系型数据库之间的传输速度不低于 102MB/s，投标人自行匹配测试硬件，对测试硬件性能无强制性要求。</p>
2	数据治理平台	元数据管理	1、支持元模型管理功能，包含内置通用元模型及自定义元模型，可创建、编辑、删除自定义元模型。
		数据血缘	2、▲支持全链路数据血缘的查看及分析，支持全流程数据血缘（包含采集、加工、共享分发）；能够展示细粒度的数据血缘关系，包括不限于数据表级、字段级数据血缘关系展示。
		数据地图	3、支持按照多种条件组合查询数据表的功能，查询条件包括但不限于数据表描述、业务标签、创建者、生命周期、大小、项目空间名称、分区表属性等。
		数据标注	4、支持管理数据元、字典表、限定词的能力。

		准	5、▲数据标准支持多租户（工作空间）实例，每租户实例可以使用独立的数据标准模板
		数据模型	<p>6、支持在线建模功能；能够以在线的方式管理逻辑模型和物理模型，并支持对不同模型基于主题域实现分组、编目管理。</p> <p>7、▲支持可视化建模功能，支持数据模型管理、模型加工任务管理功能，能够自动生成数据加工任务，配置任务加工方式包括不限于多表联合、多表连接、自定义映射、多表去重。</p>
3	大数据存储计算平台	服务能力	1、采用分布式存储技术，能实现海量数据存储能力，可支持 PB 级的存储量。
		计算能力	<p>2、支持 Spark 计算框架，提供 SparkSQL、SparkAPP 处理接口，能实现数据分布式并行计算。支持运行 Java/Python/Scala 编程语言脚本。</p> <p>3、支持基于 SQL 的批数据处理模式，支持内置函数库，提供包括数学函数、字符串函数、日期函数和窗口函数、聚合函数的数据分析功能。</p>
		安全性	4、▲提供将物理资源划分为逻辑资源池功能，防止资源抢占，提升资源使用效率，即不同的租户可以独占不同的资源池，也可以共享一个资源池。
4	数据开发平台	数据清洗	<p>1、▲支持内置的数据清洗和过滤组件。过滤组件应包括但不限于：长度过滤、正则过滤、空值过滤、比较过滤；清洗组件应包括但不限于：全角半角转换、类型转换、字典标准化、日期标准化、身份证校验转换、去除特定字符、去除空白字符、字符串截取、手机号校验、条件判断、表达式处理、数据去重；自定义组件应包括但不限于：SQL、JS 自定义清洗，规则库。</p> <p>提供具有 CMA（中国计量认证）标识或 CNAS（中国合格评定国家认可委员会认证）标识的检验报告关键页（含首页、盖章页及对应功能描述页）或国家认证的第三方权威机构出具的检测、检验类证明材料。</p> <p>2、▲支持规则库相关功能。支持配置清洗规则，应提供去除空白、空值过滤、全角半角转换、去除特定字符、正则校验、长度校验等内置规则等；支持规则自定义；支持源表配置，对源表数据进行 where 条件过滤；支持目的表配置，对输出字段进行字段映射，并配置写入规则和去重规则；支持问题数据配置，配置问题数据输出到问题表或丢弃。</p>

		易用性	3、▲支持通过图形化的拖拉拽方式将不同类型节点连线组织开发成工作流的功能，可实现零编码交互，工作流画布支持自动布局。 4、提供一站式集成开发环境，满足大数据环境下的开发、流程调度和生产运维需要。
5	数据共享交换平台	资源目录管理	1、支持目录编制、目录提交、目录修改、目录导入与导出、目录删除、目录发布与版本管理等资源目录管理功能。 2、支持流程管理能力，能够支持自定义审批流程，支持管理流程设计功能，支持创建、复制、编辑、导入、导出、删除等功能；支持流程的版本管理功能。
		资源服务使用	3、支持资源服务订阅功能，用户订阅后可以获取针对资源的各种不同服务；支持为订阅者设置订阅期限，订阅期满后，须重新购置订阅才可继续使用服务；支持审批机制，审批后才可以使用服务。
		资源服务管理	4、▲信息资源目录支持通过统一编制模板、统一发布流程、统一申请流程实现数据编目的能力。要求使用统一流程编制库表、文件、API、流式数据源（需支持至少一种流式数据源）。
			5、支持数据表转服务能力，支持 SQL 语句录入和图形模式，支持根据主流关系型数据库的数据元信息自动生成 RESTfulAPI 功能，包括 MySQL、Oracle 和 PostgreSQL。
		服务性能横向（集群）扩展	6、▲表转 API 能力不低于 3000QPS，投标人自行匹配测试硬件，对测试硬件性能无强制性要求。 提供具有 CMA（中国计量认证）标识或 CNAS（中国合格评定国家认可委员会认证）标识的检验报告关键页（含首页、盖章页及对应功能描述页）或国家认证的第三方权威机构出具的检测、检验类证明材料。
		服务运维监控	7、支持信息统计功能，支持可按照目录、交换、服务等维度统计，支持流程超期告警、数据归集超期告警、数据源超期告警功能。 8、支持告警查看功能，支持集群状态、CPU、内存告警信息，支持告警配置功能。 9、支持按照目录、交换、服务等多维度进行信息统计。 10、支持用户操作审计功能。
		实施与培训	提供详细的实施方案与进度计划，实施方案中需要列出本次项目团队成员构成与团队成员能力资质证书（信息系统项目管理师证书、系统架构设计师证书、统集成项目管理工程师证书、软件设计师证书），以及相关实施人员社保证明；合同签订后按照进度计划 20 个工作日内完成系统部署实施工作； 提供培训方案、培训计划、培训资料列表，含不低于 2 次系统使用及维护培训，免费提

		供一年的远程培训咨询。
	售后服务期限及内容	免费质保期 2 年，质保期内提供不少于 3 人全职驻场开发及运维服务（驻场人员须分别具有信息系统项目管理师证书、系统架构设计师、系统分析师、软件设计师、数据库系统工程师，合同签订后如驻场人员无法提供证书或人员与投标文件不一致的，视为中标人违约，采购人有权终止合同），在此期间软件产品的故障排除、功能升级均、技术咨询提供免费支持。