临港新片区公共数据管理与赋能平台项目

需 求 文 件

二〇二四年

**一、项目概况**

**（一）项目背景**

2023年3月7日国务院机构改革中，成立了国家数据局，负责协调推进数据基础制度建设、统筹数据资源整合共享和开发利用、统筹推进数字中国、数字经济、数字社会规划和建设等。国家数据局的设立凸显出我国对数据要素市场发展的高度重视，表明国家大力发展数字经济的决心，也标志着我国数字经济正式进入专业性、实质性的落地阶段。在此背景下，临港新片区管委会数据处应运而生。临港新片区管委会数据处的成立有利于加强临港数据基础制度，推进数据标准化、规范化、法治化建设，完善数据安全和质量管理体系，保障数据安全和合规性；促进整合各部门、行业、领域的数据资源，建立统一的数据目录、数据仓库、数据平台等，推动实现数据资源的共享和开放，加强数据资源的高效利用和价值创造。为保障相关数据工作的开展，建设临港新片区公共数据管理与赋能平台，提升临港数据管理与赋能效能，沉淀临港数据资产也成为支撑临港新片区数据发展建设工作中至关重要的一环。临港新片区行政服务中心（大数据中心）承担临港新片区政务服务，数据归集共享，底座算力建设，网络安全建设，数字化项目管理和总体运维等工作。为了满足临港城市发展中逐步凸显的数字化基础能力需求，现拟通过建设临港新片区公共数据管理与赋能平台，为临港新片区的数据管理工作提供基准工具赋能与标准化流程支撑，驱动城市数据工作体系化、流程化、高效化、智慧化。

目前，临港新片区大数据中心已建设的公共数据管理系统，作为支撑公共数据归集共享的核心业务系统，是临港新片区内各部门数据交换的主要通道。其应用场景包括：应用系统备案登记、数据归集编目展示、数据申请审核共享以及物联感知和视频流设备数据的申请审核共享等。其主要功能模块包括：（1）数据资源目录：提供多维主题分类、资源编目发布、资源智能检索、自动化服务生成、个性化流程配置、工单化运营管理等；（2）数据服务总线：汇聚管理已有数据接口，可基于已有数据自动形成标准化的开放接口，对已有接口提供管理、维护；（3）数据服务门户：提供给数据需求方在线的数据查阅、申请、下载功能。

**（二）建设目标**

为夯实临港新片区数据资源基础管理能力，加强公共数据的有效性、合规性和安全性，提升数据资源开发利用的效率与统筹性，建设临港新片区公共数据管理与赋能平台，形成分布式监测采集网络，实现提升多源全类型的数据汇聚和任务调度，实现多类型数据的存储计算，提升数据源访问的准入控制，访问权限分配及分级管控。

本次项目建设临港新片区公共数据管理与赋能平台，包括数据管理与服务平台以及数据治理能力赋能平台两个子平台，以不同的视角与工作维度，使能临港新片区数据治理与管理工作更加高效。

**二、技术要求**

**（一）平台总体设计要求**

（1）先进性要求

在方案设计和技术架构方面应具备先进性、合理性。系统建设应基于统一的整体架构，采用先进的、成熟的、可靠的技术，满足业务管理要求，保证数据的一致性、完整性及业务应用的连贯性、整体性。

（2）兼容性要求

系统的开发与部署环境由临港新片区提供信创云环境。系统应适配麒麟、达梦、金蝶、东方通等国产操作系统、中间件（以上品牌仅供参考，非指定限制，具体以实际部署要求为准），实现系统的稳定运行。供应商应当提供与本期项目中涉及临港新片区公共数据管理系统已建设部分的兼容性方案说明（包括系统对接、UI优化、部分功能调整升级等），提供兼容性相关保证或证明材料。

（3）扩展性要求

随着新的业务需求的不断产生，平台需支持基础数据模型、应用分析模型、前端应用的扩展性；需支持在统一系统架构中服务器、存储、I/O设备等的可扩展性。

（4）易用性要求

系统软件的功能设计必须充分考虑用户的使用习惯，体现设计方便实用的特性和优点，提高用户使用的方便与效率。

（5）安全性要求

应达到安全保护等级三级以上要求，符合《信息安全技术信息系统安全等级保护基本要求》（GB/T 22239-2008）与政务大数据的相关安全要求，按照国家标准、行业标准、安全规范等实现平台的安全管理。

**（二）项目建设内容**

（1）软件定制开发，包含以下子系统：

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **子平台** | **定制开发需求** | **子模块内容** |
| 1 | 数据治理能力赋能平台 | 治理开发集群 | SQL开发 |
| 2 | 可视化治理任务开发 |
| 3 | 监控运维 | 数据库运维监控 |
| 4 | 数据接入运维监控 |
| 5 | 数据服务运维监控 |
| 6 | 云产品监控 |
| 7 | 物理设备监控 |
| 8 | 报表查询 |
| 9 | 运行资源管理 |
| 10 | 告警列表 |
| 11 | 告警配置 |
| 12 | 故障处置流程管理 |
| 13 | 流程引擎平台 | 流程对接 |
| 14 | 流程工单 |
| 15 | 任务管理 |
| 16 | 数据管理与服务平台 | 数据汇聚管理平台 | 数源管理 |
| 17 | 数据详情 |
| 18 | 总体情况 |
| 19 | 任务创建 |
| 20 | 数据资源管理平台 | 数据资产管理 |
| 21 | 数据建模管理 |
| 22 | 数据标准管理 |
| 24 | 数据质量管理 |
| 25 | 应用元数据管理 |
| 26 | 数据服务平台 | 数据工作台 |
| 27 | 数据服务门户 |
| 28 | 数据运营平台 |
| 29 | 数据大屏 |
| 30 | 数据授权运营平台 | 被授权运营主体管理 |
| 31 | 授权运营授权管理 |
| 32 | 授权运营数据流通管理 |

（2）软件产品采购，包含以下软件产品：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **子平台** | **产品采购需求** | **采购需求内容** | **数量** | **单位** |
| 1 | 数据治理能力赋能平台 | 存储计算支撑平台 | 分布式数据仓库 | 1 | 套 |
| 2 | 分布式Nosql数据库 |
| 3 | 实时计算 |
| 4 | 数据检索 |
| 5 | 基础存储管理运维 |
| 6 | 安全管理对接 |
| 7 | 存储计算管控平台 |
| 8 | 时空数据库 |
| 9 | 数据分级分类加密存储 |
| 10 | 汇聚能力集群 | 数据接入 | 1 | 套 |
| 11 | 数据转换 |
| 12 | 目标端配置 |
| 13 | 数据及任务导出 |
| 14 | 操作配置 |
| 15 | 任务监控 |
| 16 | 可视化能力集群 | 可视化分析引擎 | 1 | 套 |
| 17 | 可视化图表库 |
| 18 | 可视化仪表盘 |
| 19 | 可视化分析导出 |
| 20 | 可视化大屏引擎 |
| 21 | 可视化大屏组件管理 |
| 22 | 可视化大屏资源管理 |
| 23 | 数据管理与服务平台 | 数据安全平台 | 基础信息管理 | 1 | 套 |
| 24 | 统建纳管管理 |
| 25 | 分级分类与敏感数据管控 |
| 26 | 安全审计与阻断 |

**（三）项目建设功能性要求**

**3.1平台总体架构需求**

临港新片区公共数据管理与赋能平台总体架构包含数据治理能力赋能平台和数据管理与服务平台两个子平台。其中，数据治理能力赋能平台主要面向数据开发者，向开发者提供标准化的工具，以实现数据治理和开发效能的提升（包括存储计算支撑平台、汇聚能力集群、治理开发集群、监控运维、可视化能力集群、流程引擎平台）；数据管理与服务平台主要面向数据管理者，以实现数据管理和数据服务能力的提升（包括数据汇聚管理平台、数据资源管理平台、数据服务平台、数据授权运营平台、数据安全平台）。最终通过临港新片区公共数据管理与赋能平台的建设，综合赋能数据存储、数据汇聚与更新、数据任务开发、数据运维管理、数据调度使用、数据可视化分析与填报等多业务场景。



**图：平台总体架构图**

**3.2 平台接口需求**

（1）临港新片区公共数据管理与赋能平台需与临港新片区统一身份认证支撑系统进行对接，包括统一用户认证，单点登录等信息同步。

（2）临港新片区公共数据管理与赋能平台需与市级平台和临港新片区内相关业务系统进行对接，实现数据联动和业务协同。

**3.3平台功能需求**

临港新片区公共数据管理与赋能平台包含数据治理能力赋能平台和数据管理与服务平台。其中，（1）数据治理能力赋能平台需建设存储计算支撑平台、汇聚能力集群、治理开发集群、监控运维、可视化能力集群、流程引擎平台等模块。（2）数据管理与服务平台需建设数据汇聚管理平台、数据资源管理平台、数据服务平台、数据授权运营平台、数据安全平台等模块。

投标方应根据以下功能需求描述，实现相关功能。其中涉及项目产品采购软、硬件产品选型的，应符合项目功能与技术指标要求，并优先考虑行业有成功应用的产品，且须充分考虑与可利旧系统或设备的互联互通、信息共享及统一管理要求。下述需求描述中带▲号的为软、硬件产品选型需实现的重点要求。

**3.3.1数据治理能力赋能平台**

**（1）存储计算支撑平台**

本项目需采购1套存储计算相关产品，用于平台的存储计算支撑工作。采购内容需包括：分布式数据仓库、分布式Nosql数据库、实时计算、数据检索、基础存储管理运维、安全管理对接、存储计算管控平台、时空数据库、数据分级分类加密存储。投标方需提供至少6个以上生产环境服务器节点和配套的测试环境服务器节点的使用许可，并承诺后期扩展中提供服务。平台软件架构可支撑后续扩展到300个服务器节点以上。存储计算支撑平台为整个数据治理能力赋能平台提供分析计算和存储的引擎和工具，是整个数据治理能力赋能平台的底层功能支撑，应能与国内自主成熟可控的数据治理平台和数据交换平台集成兼容。

（a）分布式数据仓库

* SQL引擎

需支持标准ANSI SQL2003语法。

需支持标准JDBC、ODBC连接。

需支持完整的数据类型，包括：BOOLEAN、INT、FLOAT、DOUBLE、DECIMAL、CHAR、VARVHAR、STRING、DATE、TIME、BINARY等。

需支持标准DDL操作数据库对象，包括：数据库、表、视图等。

需支持多种数据格式，包括：ORC、TXT、CSV、Parquet等。

需支持多种数据源类型，包括：Hbase、ES等。

需支持多种类型分区，包括：单值、Range、Interval等。

需完善的中文支持，数据库名、表名、字段名等都可使用中文。

需兼容开源HIVE，可与HIVE元数据互操作，可执行HIVE开发的批处理程序。

需支持业务平滑接入，内置数百个与关系数据库语义一致的UDF。

需支持FIFO调度、抢占式调度、公平调度等多种调度模式。

需提供图形化SQL运维与调优工具。

需支持通过SQL方式跨组件数据快速流转。

* 数据管理与审计支撑

需支持数据字典（Data Dictionary）。

需支持元数据管理和数据血缘跟踪。

需支持安全审计管理，记录所有安全和权限相关的操作，可追溯检索详细审计历史。

* 运维与高可靠

需支持计算服务高可用。

需支持集群高性能扩展，能支撑数千节点的分布式计算与存储。

需支持基于容器技术的集群安装、部署、在线升级。

需提供易用的运维工具，能便捷的启停集群，监控集群状态，实时告警。

* 分布式事务与高性能处理

需支持分布式事务处理，保障原子性、一致性、隔离性、持久性 (ACID)。

需支持事务隔离，采用可序列化快照隔离技术，并保证事务处理的并发性能。

需支持列式存储数据的高性能增删改，单节点每秒上万条记录。

需支持基于成本的优化器(cost based optimizer)，性能领先于开源SparkSQL和MPP数据库。

需支持高并发SQL路由规则，自动负载均衡。

* 权限分配支持

需支持角色访问权限控制(RBAC)，能便捷的配置对服务、数据库、图的操作权限。

需支持数据库安全认证，CAS LDAP，单点登录，增删改查用户、用户组。

需支持多租户管理，提供弹性的资源共享，保障数据、资源、应用间的安全隔离。

需支持权限管理，包括：增删改查库级、用户级、表级、行列级权限。

* 多语言支持与数据迁移

需支持标准Oracle PL/SQL存储过程语法。

需支持标准IBM DB2 SQL PL存储过程语法。

需支持标准Teradata SQL语法和方言。

需支持从关系数据库(Oracle、DB2、MySQL等)准实时(秒级)增量同步数据。

需支持数据稽核(data audit)，提高导入数据的质量。

需支持Furion调度，用树形结构管理队列，管理细化到Task，调度策略更优。

* 数据联邦与优化

需支持高级调度管理混合负载，在多租户场景下，高优先级用户可更优先调度获取资源。

需支持小文件智能合并，提升数据读写性能。

需支持基于DB-Link的数据联邦功能。

* 高性能湖仓存储

需支持基于分布式数据管理系统的高性能湖仓统一存储格式。

需支持数据自适应分布，无需手动分桶。

需支持冷热数据分级存储。

需支持创建数据快照，可以高性能访问快照数据，并进行复杂分析。

需支持冷热数据分级Compcation策略，保障系统吞吐性能。

需支持批量、实时数据接入，数据入库即可加工分析。

▲需支持动态扩展能力，具备扩容不中断业务、缩容不中断业务功能。

（b）分布式Nosql数据库

* 海量数据的存储与检索

需支持基于rowkey的高速数据检索。

需支持基于堆外的内存管理模型，提升稳定性。

需支持SQL-bulkload工具，能超高速写入数据。

需支持DSTool工具，能快速修复在线集群状态。

需支持Kerberos加密，细粒度权限管控。

需支持Region的动态拆分和合并。

* 实时数据入库

需支持混合存储，可使用SSD加速数据读写性能；

需支持标准ANSI SQL2003语法。

* 高并发的查询业务

需支持全局索引；

需支持JSON/BSON文档存储功能。

* 非结构化数据的存储与分析

需支持视频图像数据存储。

需支持BLOB对象存储。

（c）实时计算

* 准实时ETL

需支持Kafka等多种方式输入输出，允许自定义数据源和数据目标。

需支持标准ANSI SQL2003语法和流式SQL (Streaming SQL)。

需支持低延时事件驱动模式，管理事件的时间序列。

需支持滑动时间窗口、区间时间窗口、会话窗口。

需支持多个实时流的关联分析、流和表的关联分析。

需支持基于实时流的数据安全管控。

* 管理与运维

需支持流任务和应用开发的资源隔离管控、用户权限控制、生命周期管理。

需支持自动的背压流控机制 (Back Pressure)。

需支持数据不丢不重，保证数据的一致性和可靠性。

需支持任务自动负载均衡，集群高可用，故障自动恢复。

* 实时预测性分析

需支持SQL存储过程语法 (Streaming Procedural Language)。

需支持线性代数、矩阵、向量空间、线性变换、有限维线形方程组等功能。

需支持实时流数据挖掘功能 (Streaming Machine Learning)。

* 实时统计分析和复杂事件处理

需支持分布式规则引擎 (Rule Engine)，使用SQL自定义规则，兼容Drools导出规则。

需支持通过SQL方式实现复杂事件处理，实现复杂实时事件监控 (Streaming CEP)。

需支持高速的分布式缓存，能缓存多层级细粒度的中间指标。

需支持通过SQL方式实现流式实时累计指标统计、数据分析加速功能。

（d）数据检索

总体要求：▲需支持非结构化数据检索、全文检索、组合索引查询、中英文多语言支持、聚合分析、索引定义与管理、数据导入、数据导出功能。需支持相关度排序、基于字段值的排序、组合排序功能。

* 海量文本存储与检索

需支持海量数据全文搜索。

需支持海量数据分析能力，并保证结果正确性。

需支持分区表，可构建超大规模数据集。

需支持标准ANSI SQL2003语法和搜索扩展SQL语法。

需支持堆外内存管理技术，提升单节点数据服务能力至100TB。

* 非结构化数据的存储与分析

需支持非结构json/object等数据的存储和检索能力。

* 搜索引擎

能够提供更优异的高负载数据写入能力。

需支持分布式一致性协议Raft，提升集群规模上限至200实例以上。

需支持聚合函数下压技术，提升聚合分析计算性能。

需支持读写分离技术，同时支撑实时写入和实时高并发检索业务。

需支持生命周期管理技术。

能够提供跨中心部署能力。

需支持冷热数据分离技术，为用户提供高性价比存储方案。

▲需支持数据备份、数据恢复、数据高可用、集群扩/缩容、服务高可用功能。

（e）基础存储管理运维

* 集群部署与运维

能够基于容器和kubernetes，全平台服务容器化图形化部署，实现完整的资源隔离。

需支持通过界面操作实现系统服务配置。

需支持在一个Manager添加和管理多套集群，每个集群具有不同的物理节点和服务。

需支持基于NTP服务同步集群时间。

需提供License服务健康监控、起停等操作，可通过二维码进行License的操作。

需支持集群或者单个服务的滚动升级，提供不停机滚动升级能力。

Manager需支持高可用。

需支持操作系统版本异构的集群。

* 服务管理、监控与告警

需提供服务全生命周期管理，支持对内部角色提供添加、删除、服役、退役以及迁移等操作，支持不同级别的监控状况检查。

需支持节点和机柜的添加与删除，提供迁移节点上的服务能力。

需提供基于RBAC模型管理用户。

需提供统一的身份认证和单点登录服务，并提供一键为所有服务开启安全的便捷操作。

需提供节点物理资源和服务的各项指标信息，可以设定不同告警级别的阈值，需支持以邮件或者短信形式发送告警信息。

* 组件管理

需提供应用市场化的组件市场，用户可以发现新版本和新组件，按需选择升级或者安装。

内置基础组件，需支持上架开发组件。

（f）安全管理对接

* 用户认证对接

需支持多种认证接口，如LDAP，KERBEROS，CAS，OAUTH，OIDC，Access Token等。

需支持单点登录SSO，单点登出SLO。

需支持接入第三方认证，如Active Directory，JDBC，OAuth等。

需支持跨域信任，方便多集群数据流动。

* 授权与资源管理

需支持基于角色的权限管理（RBAC）。

需支持基于属性的权限管理（ABAC）。

需支持权限穿透，打通上层表结构与底层数据的权限。

需支持管理员权限细化，三权分立。

需支持细粒度权限管理：行级与列级权限管理。

需支持YARN队列等资源管理。

需支持HDFS、配额管理。

* 审计

需支持文件、关系数据库等审计。

（g）存储计算管控平台

* 存储计算调度系统

需实现数据存储计算资源的分配能力，对存储计算资源进行统一管理和调度。

需提供多样化的应用控制器以支持包括无状态应用、数据库、大数据等分布式应用的控制和管理。

容器云存储，需支持大容量的本地存储管理能力。

容器云网络，需支持面向容器云平台的网络管理能力，支持 Underlay 和 Overlay 网络两种模式以及多集群网络点对点通信。

四层及七层软件负载均衡，需支持基于浮动 IP 的高可用以及真实源 IP 等功能。

镜像仓库，需提供容器镜像和 Chart 文件管理能力。

自动弹性扩展，能够根据服务的性能指标或者定时来做应用做横向或纵向扩展。

图形化的安装工具，需支持向导方式安装容器云平台。

* 存储计算资源管控平台

需提供项目管理、产品部署和销毁、自定义产品、用户管理、权限管控、工单审批等服务模块。

需提供运营分析服务模块，可查看数据湖内各租户的使用情况、资源使用负载情况、主机运行情况等。

租户创建与管理，需支持自主的租户创建、配额设置以及生命周期管理。

需提供资源隔离的多租户能力，完全隔离的租户资源管理机制，各个租户的服务之间互相不影响。

工单与通知系统，需提供一个工单和通知系统，用于业务申请和审批等，可以自定义审批流程和权限管控。

用户权限及角色管理系统，需提供可自定义用户的角色和组，对用户的权限进行管理和控制。

细粒度计费系统，需提供基于事件的计费系统，按照小时计费，支持自定义定价和折扣策略。

日志运维系统，需集成对各个系统组件和应用的日志收集和分析能力，优化日志链路保证了链路的高吞吐。

图形化的集群指标监控和管理工具，需提供包含大量实际运维中提炼的系统组件的监控指标、告警规则等。

基于RBAC的用户安全体系，需提供通过网络隔离的多租户模型，保证租户间的安全；租户内支持RBAC做细粒度的权限划分。

网络安全管理，需支持独立的网络地址保证服务的安全，可自定义网络防火墙规则对网络通信做安全管控。

存储安全管理，需支持存储资源的RBAC管控，保证数据的安全可靠和合规性。

（h）时空数据库

* 时空数据存储架构

分布式数据库需采用Shared Nothing（完全无共享）的分布式架构；不存在单点性能瓶颈和单点故障（SPOF），所有节点无共享，拥有对等计算能力。

单一分布式数据库需同时支持OLTP、OLAP及混合型业务应用，具备强大的在线事务处理能力，同时支持大规模联机数据分析处理。

单一分布式数据库支持组建超大规模的数据库集群，最大集群规模数达到100节点以上。

集群需支持协调节点和数据节点，协调节点响应用户请求，存储节点存储数据；支持每种角色自动扩容。

需支持自动的分布式数据存储管理，基于可配置的分布字段，支持数据分片以及均匀分布。

需支持视图、索引、存储过程、函数、触发器、主外键、自定义函数等数据库高级特性。

需具备数据多副本容错技术，提供数据的冗余保护，需保证副本数据强一致性。

需可实现自动故障探测和容灾处理，当集群中的某个节点失效时，可被集群中的其它节点接管。

需支持任意节点数的集群在扩展节点时，新增节点通过网络接入集群，数据重新分布，工作负载在节点之间动态分发。

需支持自定义数据类型，函数(普通函数、聚合函数)等。

需支持基于自定义数据类型的分布式存储管理，支持自定义数据类型作为分片字段。

需支持自定义函数下推执行，充分利用集群各节点性能。

* 高性能

应具备强大的处理和存储能力，支持PB级以上的数据存储和查询能力。

事务隔离级别为读已提交，需支持高并发场景下的事务处理。多个用户操作同一条记录时，任何情况下读、写互不影响，不允许脏读。

DML操作需采用无限制行级锁方式，任何情况下不允许锁升级，大幅降低死锁的可能性，若出现死锁，能自动解锁。

需支持大对象存储管理，每个对象字段的数据最大可达4TB，大对象统一在系统表中管理，可以通过API对大对象进行读/写操作。

需支持分区表，支持透明的分区忽略功能，查询语句可自动忽略与查询条件无关的分区以减少IO和提高性能，无须修改应用。

需支持水平表分区技术，包括范围分区、列表分区等，以实现对大数据量的管理和访问性能优化，部分分区离线不能影响其他分区的使用。

需支持动态B树索引、函数索引、全文索引等索引技术，提供快速的查询处理能力。

基于代价的路径优化技术，需提供完整的SQL执行计划。

* 兼容性

需支持LINUX等多种操作系统平台。

需支持X86通用架构的标准PC服务器。

需支持国产CPU及服务器，如龙芯、飞腾等。

需支持国产操作系统，如中标麒麟、银河麒麟、普华等。

需兼容Oracle的各类对象，支持PL/SQL语法进行存储过程/函数开发。

* 其他

需支持当前最流行的数据库标准ANSI/ISO SQL92、ANSI/ISO SQL99等。

需支持C API、ODBC、JDBC、ADO.NET接口规范。

数据库字符集应支持多语种，如英文、中文等，必须完全支持中文国家标准的中文字符的存储处理。

需提供角色、系统权限、对象权限等权限控制手段。

▲需支持影像与格网对象管理，包括对象原生数据类型、对象构造函数、对象输入/输出、对象空间索引、图像处理、代数与分析计算、对象编辑、对象空间参考转换功能。

▲需支持移动对象管理，包括移动对象空间分析、对象空间统计、对象时空索引、对象时空关系判断、对象时空统计、对象相似度分析功能。

（i）数据分级分类加密存储

需支持通过基于申请已有密码机服务，实现建立加密数据空间，为平台云上应用提供数据加密存储服务，提供机密保护。

需支持与数据安全平台获取数据安全分级分类信息，对敏感数据进行加密存储。

需支持加密存储、完整性检验、访问控制、密码资源、双卡校验、双因子认证、数据隔离、密钥管理、加密方式、国密算法、安全加固等功能。

需支持连接模式、部署模式、虚拟化连接、在线升级、备份与恢复、快照、告警与监控、终端日志审计、日志审计等功能。

**（2）汇聚能力集群**

本项目需采购1套数据汇聚相关产品，用于平台的数据归集工作。采购内容需包括：数据接入、数据转换、目标端配置、数据及任务导出、操作配置、任务监控。投标方需提供汇聚能力集群的使用许可，保障项目所涉及实施人员开展数据汇聚工作，并承诺后期扩展中需提供服务，且在数据归集过程容量扩展中不设置使用许可限制。平台需支持离线同步和近实时同步归集。相关归集过程的操作规范应遵循《临港新片区公共数据归集指南》。

（a）数据接入

需提供数据加载、数据同步、文件加载、实时数据接入、非结构化数据加载，需支持通过全量数据加载、增量数据接入；需支持在接入过程中提供向导化配置；需支持通过清单展示接入任务，并可对接入任务进行验证、调试。数据接入核心功能应包括：

需支持从达梦、DB2 等传统关系数据库的数据导入。

▲需支持hive、Hbase、Kafka、API接口等数据源的数据接入。并支持多表/整库迁移，支持基于时间戳增量同步，支持CDC实时同步。

需支持CSV、定长文件、JSON 、XML文件的导入。

需支持增量文件导入，实现准实时的数据同步。

需支持跨集群的数据导入。

（b）数据转换

需实现数据的清洗，加工，包括但不限于字段映射功能、数据关联、集合操作、聚合操作、过滤、去重等。需支持全局新增字段、全局字段名映射规则、全局字段类型映射规则、单表字段映射规则、数据转换规则、配置执行顺序调整。

（c）目标端配置

需提供接入任务中目标端配置能力，支持自动建表、非自动建表、自定义配置。

需支持设置最大并发度。

需支持清空目标表的配置，需支持选择清空目标表的数据、删除目标表。当为增量加载时，还需支持清理指定位点的数据。

（d）数据及任务导出

* 数据导出

需支持直接导出数据至传统关系数据库。

需支持导出多种数据格式，需支持平台中所有类型的表，普通ORC，HBase，ES表等。

需支持跨集群导出。

* 任务导出

需提供数据任务导出、同集群导入、跨集群导入。

▲需支持导出资源包（包括：作业流、批作业、实时作业、程序、数据源等）并可以导入到目标项目中，并可以查看资源包历史记录，进行在线投产、离线投产，查看投产记录。

（e）操作配置

* 版本管理

需提供版本管理功能，支持对汇聚接入任务进行版本维护。提供查看版本、回滚等操作功能。

* 发布下线

需支持对数据接入汇聚任务进行发布，自动检测该任务是否具备发布条件，检测内容包括非法格式、缺失项。

需支持提供发布审批功能，根据业务需要，发布或下线任务流需可进行提交审批申请。

* 调试

需提供数据汇聚接入任务调试。

* 运行

需支持对已发布的数据加载任务进行执行操作，手动执行时提供参数弹框，支持对任务参数进行查看与调整。

* 全局配置

需提供全局配置，支持配置内容包括：名称、描述、调度周期、超时取消等。

（f）任务监控

* 任务操作

需提供以下任务操作功能，以进行汇聚接入任务的维护：查看、执行/重跑、错误恢复、取消执行等。

* 任务运维

需提供任务运维功能，提供任务信息展示与维护能力。支持从平台全局视角、调度引擎视角统计运维数据。支持调度引擎的资源使用情况监控，包括：磁盘、CPU、内存、心跳，支持统计作业流失败的情况，可将监控信息以API的方式提供给外部系统使用。

**（3）治理开发集群**

（a）SQL开发

* 数据源选择

需支持通过数据库对象资源查看数据源的数据预览信息。

需支持在SQL脚本数据链接中选择数据源执行SQL。

* 数据库导航

需提供数据库导航栏，可查看存在的连接、元数据信息、创建元数据、指定默认数据库、对比元数据对象等功能。

* SQL编辑器

需提供便捷的SQL编写和调试环境，支持用户进行SQL代码的编辑、调试和执行，并在需要时提供异常告警。

需支持参数执行，可配置SQL执行参数配置及参数默认值设置。

* SQL执行器

需提供SQL编辑器模块用于执行SQL语句的功能，确保系统能够执行SQL语句并提供相应的结果展示和错误提示功能，以及历史SQL语句的查看功能，从而增强系统的数据处理和监控能力。

需提供执行记录功能，可查看此SQL文件的所有历史执行记录。

* 数据编辑器

需提供多种数据管理接口，方便用户对数据进行查看和编辑。

需提供目录树支持查看数据库对象，展示内容包括数据源，数据库，表，视图，存储过程。

需支持选择数据表进行数据预览，支持选择字段后进行数据下载。

需支持数据库对象列表中手动刷新功能。

需支持展示执行结果查看和下载，展示每一条SQL执行的时间点、所耗时间，以及影响的结果行数，可以将查询结果导出为EXCEL或者CSV。

* 版本管理

需支持对SQL进行编辑后进行版本提交，输入版本描述后可生成新版本。

需支持查看SQL文件的历史版本，查看版本详情、进行版本对比以及版本回退。

* 函数管理

需支持系统预置函数、用户自定义函数功能，同时需与存储支撑平台调通，并对函数进行规范化管理。

需支持在进行SQL开发时，智能提示信息中包含预置函数和自定义函数，可在SQL语句中使用。

* 自助建表

需支持页面新建表结构，在自助建表文件内通过界面操作配置建表内容。

需支持以离线文件导入，或其他交互友好，高效协同的形式，批量自助建表。

（b）可视化治理任务开发

平台需提供可视化任务调度工具以进行数据开发过程中的工作流设计、任务调度与任务管理。

* 图形化操作

需提供图形化的操作平台，为用户提供设计任务流、调试任务、触发和调度策略等功能。并且具有丰富的分析能力，通过提供依赖关系、执行历史、甘特图等图表，帮助诊断作业流的执行状况。

* 团队协作

需具备完整的版本控制，能够需支持作业流在团队间共享协助；能通过平台整合权限管控，以实现更高的安全性。

* 多任务支持

需支持代码文件、数据处理任务、嵌套任务流等任务类型。

* 任务流开发

需提供任务流调度设计与管理的图形化工具，支持Shell、SQL等脚本，数据加载或同步、数据质量检查等任务类型，以及嵌套任务流任务。

* 任务配置

需提供任务配置功能，实现开发中与已完成任务的任务配置与管理，提供：任务状态、任务编辑、任务依赖、任务执行、任务流配置、任务发布、任务监控。

**（4）监控运维**

监控运维为运维人员提供以数据链路为中心的进度、质量、安全等层面的可视化监控、闭环化问题处理工具，从而保障向外进行数据共享开放的运营过程中，各项数据关联任务的准确性。

（a）数据库运维监控

需对接存储计算支撑平台，展示项包括：用户管理、权限管控、工单审批等服务；运营分析服务，可查看数据湖内各租户的使用情况、资源使用负载情况、主机运行情况等。

（b）数据接入运维监控

需对接汇聚能力集群。数据接入任务展示项包括：调度时间、开始执行时间、执行时长、执行版本、执行模式、源表名称、目标表名称、描述、运行状态。

（c）数据服务运维监控

需对接数据服务平台。数据服务任务展示项包括：提供服务任务列表，点击服务名称，可查看当前任务的执行情况。需提供服务任务的执行情况的统计以及可视化分析。

（d）云产品监控

需提供全量云产品的概览目录，目录中根据计算、网络、数据库、大数据、存储进行大类划分。

 需支持用户选择云产品类型后，对应产品实例列表界面，列表展示内容包括实例名称、实例ID、归属项目、实例关键指标、实例状态等。

需支持对应实例的面板界面，面板中展示实例的关键信息和性能指标的历史趋势图，用户可以自定义设置时间范围查看并将结果一键导出。

（e）物理设备监控

需提供全量物理设备的概览目录，目录中根据服务器、网络设备、存储设备进行大类划分，展示用户权限项目的已分配的设备类型和对应数量，若存在告警实例则以红色数字备注。

需提供对应物理设备列表界面，列表展示内容包括设备名称、设备SN、归属项目、设备关键指标、设备状态、健康状态。

需提供对应设备的面板界面，面板中展示设备的关键信息和性能指标的历史趋势图，用户可以自定义设置时间范围查看并将结果一键导出。

（f）报表查询

需提供自定义的报表查询功能，用户可以根据产品、设备类型来配置查询模板，勾选相应的实例、项目，配置查询的汇聚方式和时间范围、勾选查询的指标，实时进行查询并将结果以报表的形式导出。

（g）运行资源管理

需提供云资源的管理界面，展示字段包括云产品的基础资源配置信息、项目管理相关信息、应用管理相关信息。

需提供物理设备的管理界面，展示字段包括物理设备的基础资源配置信息、项目管理相关信息、应用管理相关信息，资源模型之间基于拓扑关系进行连接，用户可以管理资源实例间的上下联关系。用户可以选择对应字段进行模糊检索。

（h）告警列表

需支持至少四级告警级别，能对每类产品的告警级别数量统计，当出现新增告警且用户未确认操作时系统进行提醒，点击数量可显示该级别告警分类统计及处理情况。

（i）告警配置

需支持通过webhook、SNMP、API等多种方式对接外部系统，配置告警派发的规则逻辑，实时触发告警并产生运维工单。

需支持根据采集到的监控指标，设定相应的阈值，配置对应的触发周期和清除规则生成对应的告警。

需支持配置对应告警的派发对象和通知对象。

（j）故障处置流程管理

需提供工单自动派发功能，接收其他模块或系统派发的工单。另外，用户也可以手动发起报障。

需提供故障处置流程管理，对接节点的先后、存续关系以及处置动作均可编排。

需提供故障工单列表。用户可以通过故障单号、资产编号、处理进展、报修人、处理人等字段对列表进行过滤，列表中的展示字段可以自定义配置。

**（5）可视化能力集群**

本项目需采购1套数据可视化相关产品，用于平台的数据可视化展示和分析能力建设。采购内容需包括：可视化分析引擎、可视化图表库、可视化仪表盘、可视化分析导出、可视化大屏引擎、可视化大屏组件管理、可视化大屏资源管理。投标方需提供可视化能力集群的使用许可，保障项目所涉及实施人员开展数据可视化工作，并承诺后期扩展中需提供服务，且在数据归集过程容量扩展中不设置使用许可限制。可视化能力集群需集成数据分析、编辑、展示的成熟软件系统，应能与国内自主成熟可控的数据治理平台和数据交换平台集成兼容。

（a）可视化分析引擎

需提供成熟的可视化分析引擎，需支持基于提供的组件、数据通过拖曳的方式快速构建分析报表，并实现基于后台数据动态化的数据分析。

需提供拖曳分析、图表切换、筛选过滤器、分析字段设置、高级分析计算、图表交互、图表数据排序、图表个性化配置。

（b）可视化图表库

需实现针对多字段和数值维度的组合分析可视化图表库，为用户提供多维分析支撑。实现基于地理位置的时空大数据的可视化分析，支撑各业务场景数据在地图上不同类型的可视化呈现。通过自定义图表实现第三方开源图表接入。

需提供普通图表库，包括但不限于表格、指标卡、计量图、瀑布图、箱线图、3D散点图、3D折线图、旭日图、漏斗图、雷达图、面积图、桑基图、双轴图等。

需提供GIS经纬度地图，主要包括气泡图、热力图、海量点图、统计图、轨迹图、动态轨迹图和相关地图配置等。

需支持自定义图表的构建能力，支持JavaScript等语言的自定义开发，通过接口方式与第三方图表库（D3图库、Echarts图库等）进行对接，自主开发定义图表展现形式，开展维度和数据的配置展示。

（c）可视化仪表盘

需具备通过选择已有的数据资源，利用系统内置的丰富的可视化图表，支持组合、排版、布局等需求。

需提供可视化仪表管理工具，支持对仪表盘创建、编辑、删除、共享、移动、复制、全屏、隐藏、分享等操作。

（d）可视化分析导出

需提供多种分析服务方式，实现分析图表导出、仪表盘导出和API接口管理等功能，主要包括EXCEL数据、图片格式、PDF格式、WORD格式、带有布局框架和数据配置格式的仪表盘文件、数据接口等。

（e）可视化大屏引擎

需支持数据配置管理，提供数据源配置、查看与管理。

需支持业务分组管理，提供分组的创建、编辑、删除、统计和预览展示等。

需支持可视化大屏管理，提供以项目的维度对可视化大屏进行管理，支持大屏模板的创建、配置、导入、导出、共享等。

需支持可视化大屏编辑器，提供大屏的编辑操作，支持基于系统数据资源的配置、设计，并生成可视化大屏。

（f）可视化大屏组件管理

需提供多种类型的可视化大屏组件管理，以满足不同场景下的可视化大屏定制需求。组件包括但不限于：常规图表组件、布局组件、数据组件、地图组件、媒体组件、交互组件、关系组件、装饰组件以及用户自定义组件。

（g）可视化大屏资源管理

需提供可视化大屏资源管理，资源包括但不限于：字体资源库、图片资源库、视频资源库以及图标管理、文件管理、培训教程管理等。

**（6）流程引擎平台**

通过统一的流程引擎平台的流程集成、服务集成、数据集成、Portal框架等SOA架构与相关技术，实现以流程为中心的流程/服务/UI的整合，需支持在现有系统既有能力之上的业务创新；建设统一的组织机构和用户权限管理，适应政务业务需求不断演进导致的组织机构灵活变更，并使流程具备穿越各IT系统的能力，强化各委办数据协同。

（a）流程对接

* 流程引擎

需提供流程引擎功能，支持基于预定的流程进行流程配置、任务配置，并支持多流程版本。支持流程事件、定时调度、工作日历等功能。主要完成由管理平台向开发平台管控流程的定制。

* 流程配置门户

需支持业务流程管理、业务流程定制、业务资源管理等功能，支持对已形成的流程进行管理与导入导出。

* 业务实施框架

提供业务化配置框架，抽象常用的业务化配置，供业务人员线上使用，支持流程属性配置，包括流程图标、流程实例名称模板、流程启动权限等，启动权限等。支持对PC和移动端的业务框架实施。

* 个人任务办理

需支持查看我的待办、我的已办、我的申请、我的草稿、我的待阅、我的已阅等工作项查询。

需支持工作项批量处理。

需提供丰富的API，供外部集成使用。

（b）流程工单

* 我的工单

需提供我的待办功能：对我的待办消息进行集成、提醒，供业务人员集中处理。

需提供我的已办功能：对我的待办消息进行集成、记录和展示，供查询、回溯。

需提供我的申请功能：对我的申请记录进行集成、展示，以追踪进度。

需提供我的草稿功能：对我的流程草稿进行记录，以再次编辑、使用。

（c）任务管理

需提供流程实例管理功能：查看、修改、检索流程信息、流程实例、流程节点、流程图。

需提供应用流程管理功能：查看应用下的流程，并进行迁移操作。

需提供参与者日历功能：查看、配置管理流程参与者。

需提供流程业务配置功能：查看、修改流程在业务中的顺序、图标、时限等信息。

**3.3.2数据管理与服务平台**

**（1）数据汇聚管理平台**

（a）数源管理

需支持提供已接入的各数据资源，通过列表形式展示各个数据资源的情况，包括数源名称、数据来源、接入形式等。

（b）数据详情

需支持查看数源详情，包括接入任务情况、接入周期、接入类型、数据示例信息。

（c）总体情况

需支持对数据接入总量、接入数据类型等信息进行查看。

（d）任务创建

需支持对任务进行创建后形成工单流转至汇聚能力集群，并由相关人员实现数据资源接入。

**（2）数据资源管理平台**

（a）数据资产管理

* 数据资源目录管理

需支持数据资源目录展示、检索查询、查看详细信息。

* 数据资源基础

需支持按主题域划分资产目录、数据资源维度主题分类、数据资源分类管理。

* 资产编目

需支持数据资产的注册、管理、维护以及批量导入导出。

需支持基于逻辑模型实体与技术元数据创建数据资产目录，资产属性可以与数据标准、标签进行关联映射，数据资产分类能够自定义灵活制定。

* 资产流程

需支持对新增数据资产进行发布，发布后的资产可对接至服务门户进行浏览与申请。

* 数据地图

需支持以表格和图形的形式，展示及查询各部门或系统资源目录编制情况和数据资源共享情况、数据需求、数据责任数量情况、数据共享时效指标。

* 标签管理

需支持进行数据标签分类、数据标签新增、数据标签关联，服务于数据的精细化管理。

（b）数据建模管理

* 模型设计

需支持提供数据模型建设能力与数据模型上线审批的管理能力。

* 模型应用

需支持模型查询、模型检核、模型导入导出功能。

* 模型维护

需支持数据模型的版本管理、变更审批功能。

（c）数据标准管理

* 标准分类管理

需支持标准信息管理及展示、检索、批量导入。

* 标准计划书管理

需支持标准计划书信息管理及展示、检索、下载。

* 数据元标准管理

需支持数据元标准信息管理及展示、检索、审批、导入导出。

* 代码标准管理

需支持代码标准信息管理及展示、检索、导入导出。

* 业务术语管理

需支持业务术语信息管理及展示、检索、导入导出。

* 指标标准管理

需支持指标标准基本管理、关联管理、指标维度管理。

* 通用标准代码

需支持内置国家与行业标准，供平台用户直接使用。

* 标准映射管理

需支持数据项映射信息管理。

* 标准落地

需支持数据标准推荐、核标、展示。

* 数据标准总览

需提供标准统计信息，包括标准的发布情况及被引用情况。

（d）数据质量管理

需完成对原有公共数据管理系统中数据质量功能向本次建设系统所在信创环境的迁移，需迁移模块包括质量标准配置、数据源检核配置、质量规则配置、问题数据、检核类别、调度任务。

此外，需新增模块具体说明如下：

* 数据质量总览

需具备数据质量监测、质量规则、数据质量问题展示。

* 质量规则管理

需支持质量规则分组、规则模板管理、指标信息维护。

* 质量规则映射管理

需支持选择指标建立质量规则与数据标准进行关联映射。

* 质量监测

需支持检核任务管理、检核规则实施、检核规则日志。

* 质量告警

需支持告警配置、告警结果、邮件服务器配置，检核规则关联告警后自动生成告警信息。

* 质量问题管理

需支持质量问题清单、质量问题整改。

* 质量报告

需支持基于检核规则、质量报告模板，自动生成相关数据质量报告。

* 知识库管理

需支持知识库内容管理、检索；需支持线下文档上传、维护、打标签。

（e）应用元数据管理

需完成对原有公共数据管理系统中技术元数据管理功能向本次建设系统所在信创环境的迁移。

此外，针对应用元数据管理，需包含以下模块：

* 应用元数据概览

需支持提供数据统计、表单统计功能，整体展示应用元数据情况。

* 配置管理

需支持业务目录添加与应用添加；需支持将元数据中数据库的表/字段数据与应用系统中的数据源进行关联和同步。

需支持通过应用元数据采集工具模拟系统访问，自动化采集系统菜单、表单、请求、SQL、表等应用元数据信息。

* 应用元数据管理

需支持对应用元数据链路的分析展示；需支持导出PDF功能。

* 数据管理

需支持将表/字段信息从元数据中同步后，在数据补录中展现。

* 数据分析

需支持对表、字段、功能、表单维度建立分词索引及分析。

需支持应用元数据的全链分析、影响分析、血缘分析三种维度的链路展示。

**（3）数据服务平台**

（a）数据工作台

需为各系统提供统一入口，支持待办工作提示、工作导览及快速入口。

需提供数据库管理、数据汇聚管理、数据资源管理、数据共享管理等完整的数据工作业务流程导览。

中标方需在建设过程中先行给出数据工作台深化设计，包括但不限于工作流程、页面布局、高保真原型图、DEMO等，需与临港新片区大数据中心确认设计后完成数据工作台建设。

（b）数据服务门户

需完成对原有数据服务门户功能向本次建设系统所在信创环境的迁移，需迁移模块包括：服务产品管理、服务申请、服务管理、服务监控、服务审计。

此外，需新增模块具体说明如下：

* UI设计优化

中标方应在项目实施过程中提供数据服务门户的设计优化方案、对外服务业务逻辑优化方案。包括但不限于：设计理念、功能布局图、高保真原型图、门户DEMO等，其中UI框架需同时兼容现有VUE2和VUE3框架。UI设计优化方案需经临港新片区大数据中心确认后再完成建设。

* 信息资源管理

需支持资源浏览、信息资源分区、应用系统配置、服务器节点配置、数据源配置功能。

* 服务节点管理

需支持通过页面配置增减服务节点数量，需具备批量扩容能力；对于新增的服务节点，支持在线连通性测试。

* 服务产品管理（升级）

需新增支持按照主题授权功能，通过将主题分类授权给用户，实现用户访问权限控制。

* 服务申请（升级）

需新增支持空间数据服务模块、视频数据服务模块。

需新增支持申请数据汇聚能力模块，通过流程化配置方式，将数据从数据源向目标数据库迁移。

* 服务管理（升级）

需新增支持模型建设内容同步；需新增支持生成服务文件、查看下载API服务文档。

需新增支持对接数据安全平台的分级分类与数据敏感信息，对数据服务的返回值进行加密或脱敏。

* 服务监控（升级）

需新增支持对提供服务的IT资源、服务状态、设备状态进行监控告警，以及告警规则配置。

（c）数据运营平台

* 运营执行流程

需支持运营流程配置以及运营工单的审批、派发、跟踪，需支持对运营结果物的查看和审批。

* 运营过程管理

需支持运营服务项的管理；需支持运营工作情况和人员绩效的查询统计。

* 运营报告分析

需支持运营指标设计、页面布局设计、数据接入。

需支持人员 、工单、日常需求等维度的运营分析报表。

（d）数据大屏

需建设数据汇聚情况大屏、数据治理情况大屏、数据共享情况大屏。

中标方在实施过程中应结合临港新片区数据建设实际情况，给出数据大屏设计，包括展示数据类别、数据大屏高保真原型图、数据大屏DEMO等，需经临港新片区大数据中心确认后再完成建设。

**（4）数据授权运营平台**

（a）被授权运营主体管理

需支持被授权运营主体准入申请、被授权运营主体授权续期。

需支持对被授权运营主体进行监管，经评估审核后暂停或终止运营，在平台上进行公示。

（b）授权运营数据流通管理

需支持被授权运营主体申请数据产品上架审核、下架管理。

系统应预留接口，获取被授权运营主体所建设数据授权运营可信域中的系统日志，并需支持对数据产品信息进行抽查、审计，保障授权运营安全性。

（c）授权运营授权管理

需具备对数据服务平台中已发布的公共数据资源目录进行数据分级浏览与查询，支持对示例数据的下载、数据产品的预览以及数据服务文档的下载。

需支持被授权运营主体对公共数据提出申请，支持授权运营管理单位对相关申请进行审核。

需支持通过数据服务的形式提供相关数据至被授权运营主体建设的可信域，对该可信域进行安全审计，保证原始服务数据不出可信域。

**（5）数据安全平台**

本项目需采购1套数据安全相关产品，用于平台的数据安全管理工作。采购内容需包括：基础信息管理、统建纳管管理、分级分类与敏感数据管控、安全审计与阻断。投标方需提供数据安全平台的使用许可，并承诺后期扩展中提供服务，且在数据归集过程容量扩展中不设置使用许可限制。数据安全平台应能与国内自主成熟可控的数据治理平台和数据交换平台集成兼容。

（a）基础信息管理

需具备组织管理、用户管理、权限配置、身份体系对接等功能模块。

需支持录入数据管理单位、数据使用单位、数据提供单位以及数据治理单位等组织信息，实现对临港新片区管委会数据关联单位的组织、用户、权限等基础信息管理。

需支持数据安全组织机构管理、权属职责管理，并根据不同的岗位信息实现系统与数据权限的分配。

需支持对接临港新片区管委会现有身份体系。

（b）统建纳管管理

需具备数据库隔离、虚拟地址账号、访问客户端、服务器管理、数据库管理、资源申请审批、数据库管理、项目列表、项目详情、应用详情等模块。

需支持数据库可用不可见，支持通过数据库的虚拟访问地址，经过中转后访问目标数据库；需支持通过数据库的虚拟访问账号，对数据库内的内容进行访问。

需支持资源管理，支持提供服务器资源、数据库运维人员、数据库建表的管理；支持以项目为单位提出资源申请，需支持统一审批和管理。

需支持业务库管理，支持项目的创建与维护、项目信息与外部系统进行同步；支持项目详情、应用详情信息展示。

（c）分级分类与敏感数据管控

需具备数据源管理与资产发现、重要数据识别、数据分级分类、分级分类报告、数据动态脱敏、数据静态脱敏等模块。

需具备数据动态脱敏、数据静态脱敏功能。对已完成分级分类并关联脱敏规则的数据，需支持与汇聚能力集群、服务门户等模块对接，实现数据生产过程合规，对外服务时数据的自动脱敏。

需具备将数据分级分类结果通过接口或手动导出方式，同步给存储计算支撑平台的数据分级分类加密存储模块，以协助存储计算支撑平台协同加密机设备，遵循临港大数据中心现有密码管理体系完成对敏感数据的加密工作。

（d）安全审计与阻断

需具备日志数据采集、日志数据审计、风险展示、安全审计、风险行为阻断、威胁告警、风险识别引擎、订阅规则告警等模块。

需具备对系统整体运行中产生的各项操作日志进行审计，实现对风险行为的研判展示；同时通过全方位的数据库审计手段，实现对风险及未授权行为的阻断，在此基础上生成告警信息并进行推送。

日志审计需支持多类型数据采集；需支持被动采集、主动采集等多样化日志接入；需支持终端agent单独管理；需支持多维度日志审计、关联分析、告警；需提供标准化接口，支持第三方厂商的安全日志接入及数据同步。

安全审计需支持对数据库操作行为、操作人进行监控审计和告警；需支持告警引擎配置，对于应用程序、访问对象、访问操作等识别并告警。

**（四）项目建设非功能性要求**

**4.1软件系统性能要求**

（1）平台综合性能

（a）数据处理速度，每秒处理数据量应≥100MB；

（b）系统响应时间，平均响应时间应＜2秒；

（c）系统吞吐量，每秒处理请求数应≥1000个。

（2）平台健壮性

（a）运行稳定性，全年宕机时间总和应＜24小时；

（b）系统升级，系统升级成功率应≥95%。

（3）分布式数据仓库

在存储计算支撑平台中采购的分布式数据仓库产品应满足以下性能指标要求：

（a）▲具备百TB量级数据高性能分析能力，100TB级别TPC-DS测试成绩不超过8小时。

（b）每个CPU Core的平均扫描速度不低于100万条/秒，平均聚合速度不低于100万条/秒；支持列式存储数据的高性能增删改，单节点每秒上万条记录；支持对接实时流引擎，支持100w/s的实时流数据对接，并进行即席分析。

（c）具备批处理能力，提供7\*24稳定高效的计算引擎，满足大数据量处理和复杂逻辑运算需求。

（4）分布式Nosql数据库

在存储计算支撑平台中采购的分布式Nosql数据库产品应满足以下性能指标要求：

（a）按关键字检索单表记录延时小于200ms，单个节点并发度超过1000。

（b）按关键字检索多表关联记录延时小于1s，单个节点并发度超过1000。

**4.2信息技术自主创新要求**

本项目所使用的技术需具有跨平台、高可靠、可扩展、可移植、可维护等特性，采用的体系架构符合当前业界主流技术路线；各子系统及配套应用应能够运行在国产操作系统、中间件及国产数据库等基础软件提供的运行环境上，技术路线及选型符合国产化技术发展需要。

**4.3商用密码应用要求**

平台系统需与临港大数据中心现有密码体系进行对接，实现现有密码的统一鉴权管理。同时，需满足商用密码应用评估要求。

**4.4软件测试要求**

平台系统需通过国家认可的第三方机构的软件测试，获取软件测试报告。

**4.5等保测试要求**

平台系统需满足等保三级测评要求。

**4.6信息安全要求**

在项目实施过程中，应结合临港新片区大数据中心统一安全保障要求，开展信息安全管理工作。应对系统运行进行充分的安全评估，对潜在的系统运行风险建立各种预防措施。应围绕系统内数据安全隐患进行充分评估，建立充分的数据安全保障工作体系。

在平台安全性方面应满足以下要求：

（1）数据隐私，数据泄露事件发生率应＜1%；

（2）安全漏洞，每月安全漏洞数量应≤3个。

**4.7项目资源申请要求**

本项目建设所需的基础运行环境，包括服务器硬件资源，存储，操作系统,数据库，中间件，安全资源等均由上海市电子政务云提供，临港新片区大数据中心统一分配。

投标方应结合现有数据资源情况，并充分考虑临港新片区未来发展规划中所面临的数据存储、计算、分析需求，给出合理的政务云资源申请资源清单。

**4.8** ▲内容汇总

下表为上述需求描述中带▲号的为软、硬件产品选型需实现的重点要求汇总，供应商需根据下述▲指标进行响应，标明偏离情况，提供▲指标技术参数偏离表，并且提供满足该指标的相关证明材料（需提供相应的第三方测试报告、适配证明、加盖原厂公章的产品官方技术资料白皮书或官方网页截图等证明文件），相关证明文件中必须以突出方式圈出能够证明满足该指标要求的内容。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **序号** | **子平台** | **模块** | **子模块** | **技术要求** |
| 1 | 数据治理能力赋能平台 | 存储计算支撑平台 | 分布式数据仓库 | ▲需支持动态扩展能力，具备扩容不中断业务、缩容不中断业务功能。 |
| 2 | ▲具备百TB量级数据高性能分析能力，100TB级别TPC-DS测试成绩不超过8小时。 |
| 3 | 数据检索 | ▲需支持非结构化数据检索、全文检索、组合索引查询、中英文多语言支持、聚合分析、索引定义与管理、数据导入、数据导出功能。支持相关度排序、基于字段值的排序、组合排序功能。 |
| 4 | ▲需支持数据备份、数据恢复、数据高可用、集群扩/缩容、服务高可用功能。 |
| 5 | 时空数据库 | ▲需支持影像与格网对象管理，包括对象原生数据类型、对象构造函数、对象输入/输出、对象空间索引、图像处理、代数与分析计算、对象编辑、对象空间参考转换功能。 |
| 6 | ▲需支持移动对象管理，包括移动对象空间分析、对象空间统计、对象时空索引、对象时空关系判断、对象时空统计、对象相似度分析功能。 |
| 7 | 汇聚能力集群 | 数据接入 | ▲需支持Hive、Hbase、Kafka、API接口等数据源的数据接入。并支持多表/整库迁移，支持基于时间戳增量同步，支持CDC实时同步。 |
| 8 | 数据及任务导出 | ▲需支持导出资源包（包括：作业流、批作业、实时作业、程序、数据源等）并可以导入到目标项目中，并可以查看资源包历史记录，进行在线投产、离线投产，查看投产记录。 |

**（五）项目管理及实施要求**

**5.1项目管理要求**

投标方需根据项目建设需要，设立相关组织机构，进行项目管理，包括但不限于项目的质量、进度、费用、安全、环保等过程的管理。

资质要求：投标方应提供自己公司企业综合能力情况，包括基本情况和资质证书情况以及介绍自己公司近2年以来承接的有效的数据平台建设或数据服务类项目业绩。

**5.2项目实施要求**

（1）本项目工期：自合同签订之日起12个月完成验收（包含3个月试运行）。

（2）投标方应在投标文件中提供详细的实施方案，应包括项目组织机构、项目实施流程、项目实施进度、项目实施安排。

（3）项目开发过程中，中标方必须提供项目管理文档，以便采购方及时了解项目进展情况。

（4）投标方应向采购方提供项目管理人员和技术人员配置情况。投标方技术人员至少12人及以上，其中50%以上人员需具有相关软件开发经验。在开发阶段，中标方需派项目经理携至少3名开发人员，常驻现场。

本项目的人员配置需满足以下基本配置要求：

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **岗位名称** | **最低人数要求** | **人员岗位要求** |
| 1 | 项目经理 | 1  | 具有硕士研究生及以上学历，计算机相关专业；具有中级工程师职称及以上，请提供证明文件；具有相关大数据项目实施和管理经验；要求仅为1人，全职，且未经采购方书面许可不得变更。投标人在项目经理岗位设置多人的，取其中客观分合计得分最低的一人进行评审，且多余人员不计入项目总人数。 |
| 2 | 技术负责人 | 1 | 具有硕士研究生及以上学历，计算机相关专业；需具有相关大数据项目经验，需熟悉大数据平台整体架构和技术实现，熟悉数据治理相关工作；未经采购方书面许可不得变更。 |
| 3 | 驻场开发工程师 | 3 | 具有本科及以上学历，具有软件开发或大数据项目相关经验；提供驻场开发工程师详细名单。 |
| 4 | 非驻场开发工程师 | 5 | 具有本科及以上学历，具有软件开发或大数据项目相关项目经验；提供非驻场开发工程师详细名单。 |
| 5 | 实施工程师 | 2 | 具有本科及以上学历，具有信息化项目实施或运维经验；提供实施工程师详细名单。 |
| 合计 | 12 |  |

注：投标方必须提交完整的项目团队人员名单、身份证号码、职称、联系方式、项目实施经历、在本项目中的职务及任务。在投标文件中需提供项目团队人员相关资质证明材料以及最近一个季度任意一个月的在投标方本单位依法缴纳社保费的证明。

（5）中标方应与采购方相关业务与技术人员一同完成本项目的深化调研分析阶段的工作；需求分析结束后，采购方将再进一步提出设计、开发、测试等详细要求，中标方应按要求编制有关说明书。

（6）本招标文件提出项目需求，是投标方编制投标文件和报价的主要依据，但采购方有权根据实际业务需求，对招标文件规定的需求进行细化和优化，最终的需求以各方确认的细化和优化需求规格说明书为准，并以此需求规格说明书作为验收依据。

（7）中标方在本项目建设与运行维护期内应配合临港新片区大数据中心及临港新片区大数据中心指定的数据运营服务团队，进行包括但不限于：数据汇聚、数据治理、资产管理、数据服务等工作。主要配合内容为：任务示例编制、平台使用解惑、平台功能微调等。

**5.3安装和调试要求**

（1）负责本项目范围内平台系统的现场安装部署、集成、测试和调试，保证系统功能、性能要求的实现。在安装、配置和测试、调试过程中，中标方应对最终用户技术人员所提出的技术问题，给予满意的答复。

（2）要求有完整的安装和配置程序，具有详细的系统安装配置说明手册、用户使用说明书和系统维护说明书。系统实际安装与操作必须与说明书描述一致。

（3）要求具有完整的系统测试计划，包括根据用户需求编写的测试用例，合理的测试方案和测试方法。要求保留完整的测试报告。

（4）中标方负责解决系统集成中全部技术问题，对用户单位项目建设中碰到的其他技术问题，有责任和义务提供咨询和帮助。

**5.4售后服务要求**

（1）中标方应具有驻场实施服务能力，应提供至少**1**年的项目免费维保服务，要求提供不少于1名驻场维护人员并具备独立承担本系统日常维护的技术能力。维保期从项目最终验收合格之日算起。在维保期内，出现软件系统的故障，中标方应研究其故障原因，并迅速修复，直至满足采购方原定的要求为止。

在开发和维护期间中标单位人员发生变动时必须经过采购方的同意，否则按每人以合同总价2%的标准在项目费用中扣除。

（2）维保服务的响应时间。若采购方提出故障申请，中标方须提供7×24小时电话服务，维保服务响应时间为30分钟，并在2小时内派专业技术人员到达现场进行免费维保服务。故障原因在8小时内无法排除的，中标方在1个工作日内提交解决软件系统故障的方案。通过与用户交涉承诺将以最快的时间将故障排除。中标方还须提供系统应用软件的故障处理、维护和现场巡检等服务，以及其他的技术需支持工作。在维保期内，与软件系统维保等相关的费用由中标方负责。

（3）维保期后的服务要求。维保期结束后，中标方仍应提供与维保期内相同质量的售后服务，其维保的相关费用由采购方自行承担。

（4）投标方可视自身能力在投标文件中提供更优、更合理的维保服务承诺。

（5）投标方在其投标文件中须对售后服务作出明确的承诺，包括服务时限、故障响应处理时限、应用恢复处理时限等。

**5.5项目培训要求**

（1）本项目的培训，是指投标方对临港新片区公共数据管理与赋能平台使用和维护该软件的人员进行培训。

（2）培训由投标方负责师资及教材，并负责组织实施。投标方必须派出具有相关专业资格和实际工作经验的教师及辅导人员授课。

（3）投标方应详细制定人员培训方案，包括培训目的、培训时间安排、人数、次数、培训课程、培训师资情况、培训组织方式等。

（4）培训费用计入总报价。

**（六）项目验收**

**6.1验收条件**

模块验收合格条件为：

（1）试运行时，平台功能和性能满足合同要求；

（2）系统测试和试运行时出现的问题已被解决；

（3）平台各模块通过指定的第三方适配测试和密保测试；

（4）已提供合同要求的全部文档资料。

**6.2验收方式**

项目验收分初步验收和最终验收。

（1）初步验收

系统开发调试完毕后，中标方应对系统的整体性能和功能进行自检，自检结果必须符合招标文件要求及合同中的相关条款，中标方应负责组织专业技术人员进行系统的安装调试，解决安装中出现的问题。系统功能开发完成、自检无误后，中标方向采购方提交初验收申请，经采购方初验合格同意后，系统进入试运行。

（2）试运行

在试运行期间，由于软件系统功能质量等造成某些指标达不到要求，允许中标方更换、修复、修改等，在全部达到要求时可进行最终验收。

（3）最终验收

项目终验前，由中标方提交验收申请，经采购方同意后，组织进行终验，验收结果必须符合招标文件要求及合同中的相关条款。终验期间，如有发现系统运行有问题，中标方应无条件重新检测并调试系统直至最终验收合格交付使用。

**6.3验收内容**

验收内容应包括软件功能、性能的验收。包括软件功能的可用性，技术框架的稳定性及可扩展性，所用自主产品的是否安全可靠等内容。

项目验收产品交付物：约定的可稳定运行的系统产品、应用软件定制开发部分系统代码。

项目验收相关文档如下，媒介为纸质或电子：

|  |  |
| --- | --- |
| **序号** | **文档名称** |
| 1 | 《临港新片区公共数据管理与赋能平台项目需求规格说明书》 |
| 2 | 《临港新片区公共数据管理与赋能平台项目概要设计说明书》 |
| 3 | 《临港新片区公共数据管理与赋能平台项目详细设计说明书》 |
| 4 | 《临港新片区公共数据管理与赋能平台项目接口设计规范》 |
| 5 | 《临港新片区公共数据管理与赋能平台项目部署手册》 |
| 6 | 《临港新片区公共数据管理与赋能平台项目操作与运维手册》 |
| 7 | 《临港新片区公共数据管理与赋能平台项目测试方案与测试报告》 |
| 8 | 《临港新片区公共数据管理与赋能平台项目试运行报告》 |
| 9 | 《临港新片区公共数据管理与赋能平台项目培训方案》 |
| 10 | 《临港新片区公共数据标准管理规范》 |
| 11 | 《临港新片区公共数据质量规范》 |
| 12 | 《临港新片区公共数据质量考评细则》 |

**（七）功能点及架构方案需求**

投标方应按照本文件功能和非功能需求提供详细的功能点及架构技术方案书（简称“技术方案书”）。技术方案书中，供应商提供的各项设备及系统的功能、性能应完全符合招标方指明的标准，并满足或高于采购人指出的要求。对于本文件未规定的有关设备性能或属于投标方专有的系统性能指标，投标方应在其技术建议书中提出建议，并陈述其理由。

**（八）其他需求**

投标方需认真研读和评估以上需求，确保能提供以上需求的产品和服务，如需增补相关产品和服务，需在投标文件中列明清单和做出说明，并列入中标方的报价总价中。如有遗漏，中标方需对项目的完整性负责，并免费提供相关系统和服务。

**三、商务要求**

**（一）服务期限**

本项目工期为自合同签订之日起至2025年6月完成验收（包含3个月试运行）。

质保期为自项目验收通过之日起一年。

**（二）付款方式**

详见招标文件相关规定。

**（三）培训要求**

（1）本项目的培训，是指投标方对临港新片区公共数据管理与赋能平台使用和维护该软件的人员进行培训。

（2）培训由投标方负责师资及教材，并负责组织实施。投标方必须派出具有相关专业资格和实际工作经验的教师及辅导人员授课。

（3）投标方应详细制定人员培训方案，包括培训目的、培训时间安排、人数、次数、教材编写（列出教材基本内容）、培训课程（包括课程介绍）、培训师资情况（包括教师简历）、培训组织方式等。

（4）培训费用计入总报价。

中标方在质保期结束前向招标方人员提供免费指导和培训，质保期满后应根据招标方需要提供技术培训服务，如需收费，则由双方另行协商。

**（四）违约责任**

1.中标方未按照本合同约定的时间完成本项目各阶段工作的，每延迟一日，采购方有权自尚未支付的合同款项中扣除本合同服务费总金额1‰的逾期违约金。

2.除因采购方付款审批影响支付进度的情况外，如采购方未按照本合同约定的时间支付服务费，中标方书面催告采购方并给予不少于30个工作日的履行期限后采购方仍未支付的，中标方有权自该期限届满之日起要求采购方支付该阶段应付而未付合同价款每日1‰的逾期违约金。

3.免费质保期内，如中标方存在不及时提供保修服务等违约行为，每次采购方有权自支付合同余款中扣除合同总价款1‰的违约金。

4.本合同履行期间，发生下述情形时，采购方有权解除本合同，中标方除应退还采购方已支付的全部款项外, 还应支付合同总价款20%的违约金。中标方应同时将已完成的阶段成果移交给采购方，并将全部有关资料退还采购方。

（1）中标方未能按合同约定的日期(含协商延缓的日期)提交成果，延误时间超过15个工作日；

（2）因中标方工作的错误或遗漏造成成果质量损失，且中标方未按采购方要求采取及时有效的补救措施的；

（3）中标方虽如期提交最终成果，但连续二次未能通过采购方验收的；

（4）未经采购方书面在先同意，中标方擅自将本合同项目与任何第三方合作，或将本合同标的的全部或部分擅自转包给任何第三方，且未按采购方的要求终止与第三方的合同的；

（5）项目负责人未承担本合同项目实质性工作，或未经采购方书面在先同意，中标方擅自更换项目负责人的；

（6）发现中标方或中标方工作人员出现全部或部分丧失履行委托服务能力的情况，采购方有权通知中标方限期整改或更换工作人员，期限届满中标方并未进行改善的；

（7）中标方明确表示或者以自己的行为表明不履行本合同的。

5.上述违约责任各自独立且可累加。本合同所约定的违约金如低于因违约行为所造成的损失，违约方应补偿上述不足部分的损失。

**（五）知识产权：**

1.本项目全部研发成果的所有权、知识产权以及与之相关的所有权利归采购方所有；本合同签署前双方已经拥有的知识产权权利，仍归各自所有。

2.双方确定，采购方有权利用中标方按照本合同约定提供的研究开发成果，进行后续改进。由此产生的具有实质性或创造性技术进步特征的新的技术成果及其权利归属，全部由采购方享有。

3.中标方所提交的产品软件、应用软件及任何形式的数据、资料、软件等所有权归采购方拥有。

4.中标方所提交的产品软件、应用软件及任何形式的数据、资料、软件等不得侵犯任何第三方的知识产权或其他权利，由于该等文件、数据、资料、软件侵权所导致的任何索赔或责任均由中标方承担，同时，采购方有权要求中标方修改成果至不侵权或解除合同，并保留要求中标方赔偿相应损失的权利，包括但不限于上述索赔或责任所产生的诉讼费用、合理的律师费用、鉴定评估费用、调查费用、和解金额或生效法律文书中规定的赔偿金额。

5.未经采购方书面在先同意，中标方不得将本项目的中间成果和最终成果文件及任何数据、资料、软件向任何第三方转让、授权任何第三方使用或用于本合同项目外的其他项目。如发生以上情况，中标方应向采购方支付本合同总价款20%的违约金，如果违约金无法弥补采购方的损失，采购方有权要求中标方赔偿全部损失。

6.中标方完成本合同项目的研究开发人员享有在有关技术成果文件上写明技术成果完成者的权利和与采购方共同获得有关荣誉证书、奖励的权利。

**（六）保密条款**

1.双方应保守的秘密是指不为公众所知悉、能为权利人带来经济利益、具有实用性并经权利人采取保密措施的技术信息和经营信息，或由双方在履行本合同过程中明确指明为秘密的、法律所认可的任何信息，以及中标方在履行本合同过程中所获得或接触到的任何采购方内部数据资料。

2.双方应遵守国家的有关保密规定。妥善保管对方提供的资料，保守对方的各项秘密，并保护对方的知识产权。获取对方秘密的一方应当采取有效的保护方式防止秘密未经授权而被使用、传播或公开，仅可以将该秘密用于履行其在本合同项下的义务，且只能由相关的技术人员使用。

3.中标方应以保密方式处理在编制本项目成果文件过程中自采购方、采购方工作人员或采购方关联机构获得的相关信息、资料、图纸、数据等，或由采购方在履行本合同过程中明确指明为秘密的任何信息，以及中标方因本项目工作内容所直接或间接取得、处理或接触的任何其他资料。未经采购方书面在先同意，中标方不得向任何第三方透露与本项目有关的内容，或公开本项目的阶段性成果或最终成果内容。即使向与履行本合同有关的人员提供前述保密信息，也应限于履行合同的必需范围，且这些人员应接受至少与本条款同等严格的保密条款的约束。

4.中标方实施项目的程序应符合国家安全、保密的有关规定和标准。

5.上述保密义务不适用以下情况：

（1）获取的信息已被合法公开；

（2）获得信息拥有方书面许可并在该许可范围内披露；

（3）国家法律、法规规定的情形。

6.如投标方违反上述保密条款，招标方有权解除合同，投标方应在合同解除通知发出之日起10日内向招标方支付本合同服务费总金额20%的违约金，如果违约金无法弥补招标方损失，招标方保留要求投标方赔偿全部损失的权利。

7. 无论本合同或本合同其他条款是否有效，本保密条款始终约束双方。国家秘密的保密期限由国家有关部门确定，工作秘密的保密期限由采购方确定。

**四、投标报价**

1.本项目服务费采用包干制，应包括服务成本、法定税费和企业的利润。由企业根据招标文件所提供的资料自行测算投标报价；一经中标，投标报价总价作为中标单位与采购人签定的合同金额，合同期限内不做调整。

2.投标方应根据本企业的成本自行决定报价，但不得以低于其企业成本的报价投标；评标时，评标委员会认为投标方的报价明显低于其他通过符合性审查投标方的报价，有可能影响产品质量或者不能诚信履约的，应当要求其在评标现场合理的时间内提供书面说明，必要时提交相关证明材料；投标方不能证明其报价合理性的，评标委员会应当将其作为无效投标处理。

3.投标方的投标报价，应是本项目招标范围和招标文件及合同条款上所列的各项内容中所述的全部，不得以任何理由予以重复，并以投标方在投标文件中提出的综合单价或总价为依据。

4.除非政府集中采购机构通过修改招标文件予以更正，否则，投标方应毫无例外地按招标文件所列的清单中项目和数量填报综合单价或总价。投标方未填综合单价或总价的项目，在实施后，将不得以支付，并视作该项费用已包括在其它有价款的综合单价或总价内。

5.投标方应充分了解项目的位置、情况、道路及任何其它足以影响投标报价的情况，任何因忽视或误解项目情况而导致的索赔或服务期限延长申请将不获批准。

6.投标方不得期望通过索赔等方式获取补偿，否则，除可能遭到拒绝外，还可能将被作为不良行为记录在案，并可能影响其以后参加政府采购的项目投标。各投标方在投标报价时，应充分考虑投标报价的风险。