TC609

全国数据标准化技术委员会技术文件

TC609-6-2026-XX

数据基础设施 数据沙箱公共服务平台技术要求

Data infrastructure technical requirements for data sandbox public service platform

（征求意见稿）

2026-XX-XX 发布　　2026-XX-XX 实施

全国数据标准化技术委员会　发布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由全国数据标准化技术委员会（SAC/TC609）提出并归口。

本文件起草单位：

本文件主要起草人

数据基础设施 数据沙箱公共服务平台技术要求

1 范围

本标准规定了数据基础设施数据沙箱公共服务平台的技术要求，包括公共数据、企业数据、个人数据在存储、流通、利用场景的安全使用和共享，并提出了数据导入导出、空间隔离、透明访问、数据加密、泄露防护、使用控制、远程映射、用后即焚、数据核验、数据共享、联合开发、身份认证和API对接等安全能力要求、使用要求、服务要求及审计要求。数据沙箱公共服务平台可部署于数场、可信数据空间、数据元件、数联网等多种基础设施路线中，用于数据托管、数据联合开发、数据使用、数据共享、数据流转、数据治理和数据交付，也可作为数据基础设施独立部署，用于数据安全内控及数据安全流通。

本标准适用于数据基础设施数据沙箱公共服务平台的规划、建设、运营、维护和评估等环节，旨在为数据基础设施相关方提供全面的技术指导，确保公共数据、企业数据、个人数据的流通性、安全性、保密性、可控性、完整性和可用性。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GM/T 0028-2024 密码模块安全技术要求

GM/T 0039-2024 密码模块安全检测要求

GB/T 22239-2019 信息安全技术 网络安全等级保护基本要求

GB/T 43697-2024 数据安全技术 数据分类分级规则

GB/T 31168-2014 信息安全技术 云计算服务安全能力要求

GB/T 35274-2017 信息安全技术 大数据服务安全能力要求

GB/T 35273-2017 信息安全技术 个人信息安全规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

数据基础设施 data infrastructure

从数据要素价值释放的角度出发，面向社会提供数据采集、汇聚、传输、加工、流通、利用、运营、安全服务的一类新型基础设施，是集成硬件、软件、模型算法、标准规范、机制设计等在内的有机整体。

3.2

数据沙箱 data sandbox

数据沙箱是一种存储、隔离和管理数据的安全环境，在该环境中，数据使用方或数据应用可按照使用控制约定对数据进行安全处理、分析、建模等操作，数据的使用、导出受到严格的控制，从而保证数据安全的前提下实现数据的共享和利用。

3.3

数据沙箱公共服务平台 data sandbox public service platform

数据沙箱公共服务平台是面向公众提供安全数据沙箱服务的数据基础设施，支持多个用户组透明安全地利用数据沙箱进行分布式数据安全存储、多方共管、数据共享、联合开发、使用控制、租户隔离、审计监管和存证溯源等功能实现。

3.4

多方共管 multiparty co-management

是指多个主体对同一份数据实现共同管理，任何一方都无法独立地使用数据。

4 概述

数据沙箱是一种存储、隔离和管理数据的安全环境，在该环境中，数据使用方或数据应用可按照使用控制约定对数据进行安全处理、分析、建模等操作，数据的使用、导出受到严格的控制，从而保证数据安全的前提下实现数据的共享和利用。

数据沙箱

支持数据导入导出、数据上传、数据加密、远程映射、数据隔离、透明访问、泄露防护、用后即焚、使用控制、数据核验、数据共享、联合开发、身份认证和API对接等能力，并具备数据可用不可见和防止超级管理员非法复制等特点。多个数据沙箱可组成分布式或集群式数据安全使用环境，支持数据安全存储、多方共管、数据共享、联合开发、使用控制、租户隔离、审计监管和存证溯源等能力。数据沙箱是数据领域的关键技术之一，可为用户提供一种安全的数据隔离环境，是数据基础设施中的重要组成部分。

数据沙箱支持网络传输、文件拷贝、数据流等方式的数据导入导出；并支持多种业务的计算引擎在授权情况下，透明访问数据沙箱隔离环境中的数据，如：大数据业务、人工智能业务、数据库等业务；数据沙箱可支持加密存储、碎片化存储和分布式存储等方式，并实现多租户和多用户的数据隔离；数据沙箱可本地化部署和SaaS化服务方式部署，两种方式均支持沙箱管理系统对沙箱导出数据进行审计和输出管控。

数据沙箱公共服务平台由沙箱管理系统和数据沙箱共同构成，两者协同提供一体化服务。

沙箱管理系统中包含密钥管理、存储管理、共享管理、权限管理、负载管理、审计管理等功能，部署于数据基础设施公共服务平台，用于支撑数据沙箱的数据使用、数据共享和数据流转。

数据沙箱可在数场、可信数据空间、数据元件、数联网等多种基础设施路线中部署应用，作为联合加工环境；还可作为可信数据空间的数据使用控制环境；也可作为数据基础设施数据安全流通和安全内控的组件独立部署，用于支撑业务应用的数据存储、开发、加工、处理、使用、共享、流转和交付。

5 定位

数据沙箱公共服务平台作为国家数据基础设施中的一类业务节点，应符合NDI-TR-2025-01及NDI-TR-2025-02确定的国家数据基础设施及互联互通整体架构及基本约束，符合目录标识、身份登记、接口要求的三统一规范，并继承国家数据基础设施业务节点的基本要求，如图1所示。

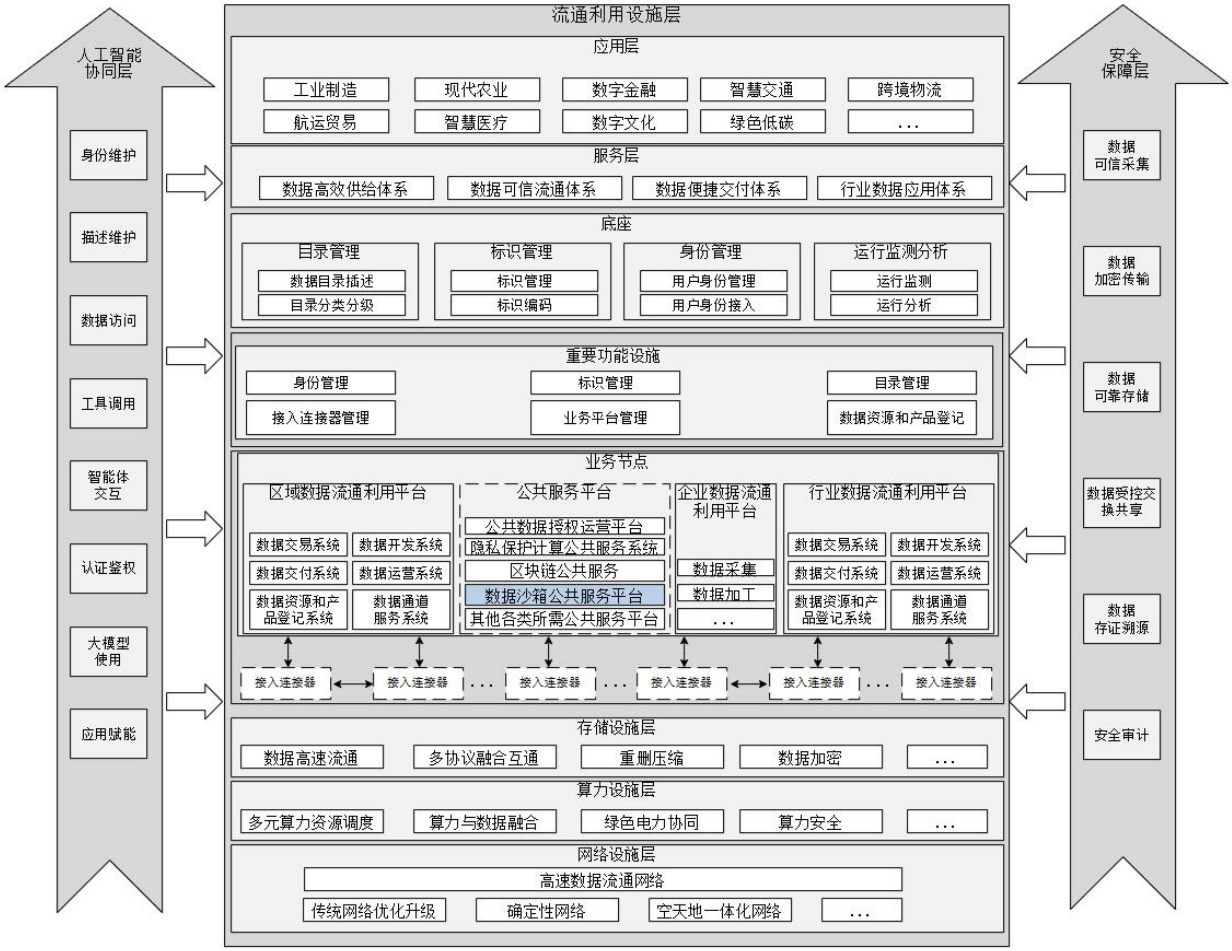

图 1　数据沙箱公共服务平台在数据基础设施中的位置

6　总体框架

数据沙箱公共服务平台由沙箱管理系统和数据沙箱两部分组成。沙箱管理系统具备密钥管理、存储管理、共享管理、权限管理、负载管理和审计管理等能力，部署于数据基础设施公共服务平台，用于支撑数据沙箱的数据使用、数据共享和数据流转。数据沙箱公共服务平台逻辑上属于数据基础设施业务节点，用于支撑业务应用的数据存储、开发、加工、处理、使用、共享、流转和交付。数据沙箱数据导入导出方式包括网络传输、文件拷贝和数据流等方式，如图2所示：

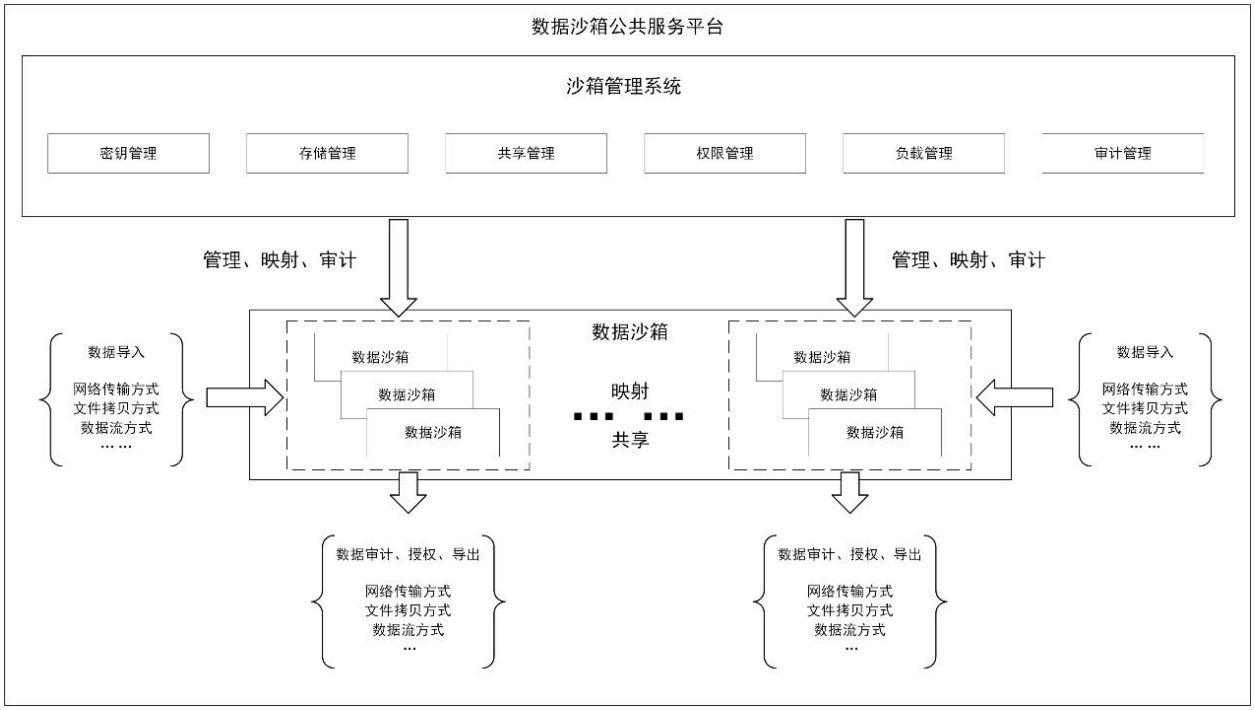

图 2　数据沙箱公共服务平台架构图

6.1　沙箱管理系统

沙箱管理系统包括密钥管理服务、存储管理服务、共享管理服务、权限管理服务、负载管理服务和审计管理服务，密钥管理服务和存储管理服务可部署于不同的主体或服务节点。根据参与方的数量和实际需求，主体或服务节点可以是一个，也可以是多个。每个主体或服务节点分别由不同的组织或人员采用平权方式共同管理。

沙箱管理系统中的密钥管理服务和存储管理服务用于接收数据沙箱上传的密文、密钥及元数据，并响应数据沙箱的密文、密钥及元数据处理请求，按需提供数据。该系统负责数据沙箱租户密钥和存储空间的管理（包括创建、删除）及状态监控。

共享管理服务用于同主体、跨主体、跨空间、跨地域的数据共享管理。

权限管理服务用于数据沙箱租户的接入、授权和验证，以及安全策略的管理。

负载管理服务用于管理和解析密钥管理模块的服务IP地址、端口等信息。

审计管理服务用于收集密钥管理服务、存储管理服务、共享管理服务、权限管理服务和负载管理服务的日志，并对数据沙箱数据使用、共享、流转情况进行审计、追溯、监管以及全局态势感知；另外，审计管理服务在得到存储管理服务和密钥管理服务授权后，还可受控审计和查看数据沙箱内的数据内容。

6.2　数据沙箱

数据沙箱是一种存储、隔离和管理数据的安全环境，一方面对接沙箱管理系统，实现数据、策略、状态等交互；另一方面对接业务应用，实现环境隔离和使用控制。数据沙箱支持单节点、分布式、集群式、分层式和多节点互通等工作模式，支持的数据对象包括文本、图片、视频、模型、数据集、数据库等。

7　部署方式

数据沙箱公共服务平台中沙箱管理系统与数据沙箱之间支持以下两种部署方式，如图3所示：

集中部署模式：多个/多组数据沙箱集中部署,通过运营管理界面、应用或工具执行数据导入、处理、分析、建模、导出等操作；

混合部署模式：数据沙箱、沙箱管理系统中的组件也可以部分集中部分分布的混合式部署。例如数据基础设施侧部署沙箱管理系统、部分数据沙箱，数据提供方或使用方侧在连接器部署部分沙箱及数据存储系统，通过连接器执行数据导入、处理、分析、建模、导出以及数据存储等操作。

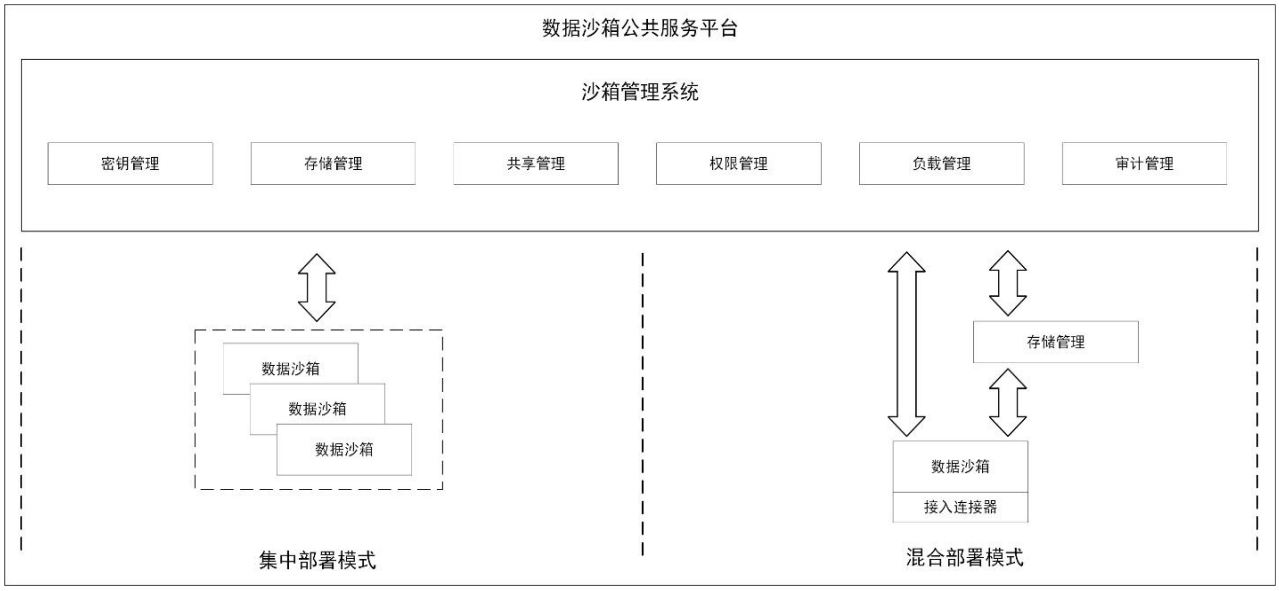

图 3　数据沙箱公共服务平台中沙箱管理系统与数据沙箱之间交互模式

8　服务能力

数据沙箱公共服务平台应提供以下服务：

8.1　安全数据操作服务

数据沙箱公共服务平台为数据提供方、使用方、监管方等多主体提供易用的数据沙箱开发环境。各方可在数据沙箱执行数据加工、计算、模型训练与推理等操作任务，实现数据防泄露导入与受控导出。

8.2　数据共管服务

数据沙箱公共服务平台支持数据导入沙箱后，按照约定，多方主体对同一数据资产享有平等、易用的共同管理权限，各方可通过数字合约等方式约定数据的使用方法与授权工具，开展数据处理活动。

8.3　可信数据流通服务

数据沙箱公共服务平台依托沙箱数据映射、用后即焚等能力，结合数据基础设施统一身份、统一目录、统一标识，为跨主体、跨地域的数据处理、分析与建模等操作提供可信流通支持。

8.4　运营管理与合规服务

数据沙箱公共服务平台为提供方、使用方、监管方等主体提供全景仪表盘，以实时监测数据沙箱的数据流通量、运行状态及异常情况等关键指标；同时，遵循《数据安全法》、《个人信息保护法》等法规要求，提供对数据沙箱内部数据的存证、抽检、溯源、取证及审计能力。

9 平台功能要求

9.1 认证接入环节

认证接入是平台数据生命周期的起始环节，涵盖身份认证与数据采集导入两大基础能力。其中，身份认证是贯穿全生命周期的基础安全保障——所有接入平台的主体：数据提供方、使用方、运营方、监管方等均须在本环节完成身份验证，验证结果作为后续存储、使用、共享、导出、监管各环节的授权依据。数据导入则负责将外部源数据安全接入沙箱环境，在认证通过的主体之间建立可信数据传输通道。两者协同，为平台全生命周期的数据安全流转奠定可信基础。

9.1.1 身份认证功能

身份认证功能是指数据沙箱对拟接入的数据提供方、数据使用方、运营方、服务方、应用方及监管方等主体的身份进行验证，确保其身份真实、合法、可信的功能，认证机制应符合数据基础设施互联互通的相关要求。该功能应满足以下要求：

a) 应支持符合数据基础设施互联互通基本要求中所规定的身份认证协议与流程；
b) 应支持设备实例间的双向身份认证，并在认证通过后建立加密传输通道；
c) 宜支持基于硬件国密 UKEY、动态令牌或同等安全强度的密码技术进行多因素身份认证；
d) 宜支持将硬件 UKEY 与会话状态进行绑定，当 UKEY 装置被拔出或离线时，数据沙箱应立即终止当前会话或停止运行。
e) 当同一身份连续多次认证失败时，系统应自动对该账户实施临时或永久锁定，并支持配置失败次数阈值、锁定时间及解锁方式。

9.1.2 数据导入功能

数据导入功能是指将原始数据通过指定的对接方式，从外部源传输并持久化至数据沙箱安全环境的功能。数据沙箱安全环境应支持配置导入数据的本地存储与远端存储两种模式。该功能应满足以下要求：

数据接入方式

a) 应支持通过网络方式导入数据，包括但不限于 JDBC/ODBC（SQL）、REST API、NoSQL 原生协议、FTP 等；
b) 应支持实时数据流方式的数据导入；
c) 应支持基于接入连接器或可信执行环境的数据导入。

数据存储处理

a) 当配置为本地存储模式时，应支持导入数据的实时加密及本地存储；
b) 当配置为远端存储模式时，应支持导入数据的实时上传及远端存储；
c) 应支持将数据上传至采用分布式和多副本架构的存储服务；
d) 宜支持通过数据库 SQL 协议向数据库型存储服务实时上传数据。

9.2 加密存储环节

加密存储是保障数据静态安全的核心手段。平台对导入数据、再生数据及临时数据进行实时加密或碎化加密处理，并通过密钥管理服务统一托管密钥及元数据，通过存储管理服务实现密文的持久化、备份与同步，确保存储态数据的机密性与可用性。

9.2.1 数据加密功能

数据加密功能是指数据沙箱对内部导入数据、再生数据及临时数据进行加密处理，并以密文形式存储于本地或远端存储服务的功能。该功能应满足以下要求：

a) 应支持使用国密算法或同等安全强度的其他密码算法，对导入数据、再生数据及临时数据进行实时加解密；可结合硬件密码卡实现。
b) 宜支持数据碎化技术，能够对文件数据、密钥数据、数据库字段等实施随机、实时的碎片化处理。若采用该技术，文件级碎化可达到非等长、字节级粒度。
c) 宜支持每个文件使用独立的随机密钥进行数据加密和碎化；该随机密钥应经主密钥加密后上传至密钥管理服务；主密钥的更新应符合 GM/T 0028-2024《密码模块安全技术》中“敏感安全参数管理”的规定。
d) 宜支持依据文件的分级分类标签，自动选择相应的加密强度（如密钥长度、算法类型）及安全级别的存储管理服务，实现差异化安全存储。
e) 数据沙箱与密钥管理服务之间的通信链路，应采用国密算法或同等安全强度的密码算法进行加密保护。
f) 数据加密与碎化处理应达到以下性能指标：非结构化数据实时碎化加解密吞吐量 ≥ 80 MB/s，合并延迟 ≤ 60 ms；非结构化数据实时碎化吞吐量 ≥ 100 MB/s，合并延迟 ≤ 50 ms；结构化数据 TPM ≥ 50000；半结构化数据 OPS ≥ 12000。

9.2.2 密钥管理功能

密钥管理功能是指沙箱管理系统对公共服务平台加密数据所涉及的密钥及元数据实施全生命周期管理的功能，涵盖生成、存储、分发、使用、隔离、轮换、备份、恢复、撤销及审计等环节。该功能应满足以下要求：

基础功能

a) 应支持密钥及元数据的实时读写，满足数据沙箱内加密操作对密钥的高效访问需求；
b) 应支持生成用于数据加密的安全随机数；
c) 应支持基于数据沙箱租户的密钥及元数据逻辑隔离，确保公共服务平台不同租户间密钥互不可见。

安全管控

a) 应支持对主密钥的全生命周期管理，至少包括使用、轮换和撤销操作；
b) 应支持对密钥的使用、轮换、撤销等操作实施授权与审批流程，确保密钥操作经合法授权；
c) 应支持对主密钥的使用、轮换、撤销等操作进行日志记录，并将日志实时同步至审计管理服务。

运维保障

a) 应支持密钥及元数据的备份与恢复机制，确保密钥数据的可靠性与可恢复性；
b) 宜支持与关联密钥管理服务之间的密钥及元数据同步、导入和导出；
c) 应支持记录密钥及元数据在指定周期内的变更历史，并可配合数据恢复、审计、监管及追溯等指令进行数据操作；
d) 应提供密钥管理相关的 API 接口，至少包括管理操作、身份验证、日志输出及状态查询等功能；
e) 应支持获取和查看密钥管理功能的健康状态，并具备异常自动告警功能。

9.2.3 存储管理功能

存储管理功能是指沙箱管理系统对公共服务平台的加密数据所涉及的密文或碎片化密文进行全生命周期管理的功能，涵盖存储、访问、映射、隔离、同步、备份、恢复及审计等环节。该功能应满足以下要求：

核心存储与访问功能

a) 应支持对密文或碎片化密文进行安全存储，并能在密文状态下执行必要的处理与运算；

b) 应支持实时响应数据沙箱对密文或碎片化密文的访问请求与映射请求。

隔离与数据保护

a) 应支持基于数据沙箱租户的密文或碎片化密文逻辑隔离，确保不同公共服务平台租户间存储数据互不可见；

b) 应具备密文或碎片化密文数据的备份与恢复机制，确保数据的可靠性与可恢复性。

数据同步与流转

a) 应支持与关联存储管理服务之间的部分或全部数据同步、导入和导出；

b) 应支持配置同步策略，并具备在数据冲突时依据预设规则进行自动处理的功能，以确保数据一致性。

运维与审计保障

a) 应支持获取和查看存储管理功能的健康状态，并具备异常自动告警功能；

b) 应支持记录密文或碎片化密文在指定周期内的变更历史，并可配合数据恢复、审计、监管及追溯等要求进行数据操作；

c) 应提供与存储管理相关的对外服务接口，支持管理操作、状态查询及日志输出等基本功能。

9.3 使用防护环节

使用防护是保障数据在用阶段安全的关键环节。平台通过数据隔离构建独立的运行环境，通过透明访问使授权应用无感接入，通过泄露防护阻止数据非授权外发，通过用后即焚确保内存明文及时清除，通过使用控制对输出行为实施实时策略拦截，并支撑多方联合开发场景下的安全协同。上述功能共同构成数据使用过程的纵深防御体系。

9.3.1 数据隔离功能

数据隔离功能是指数据沙箱依据用户、数据、应用、网络等维度的隔离策略，构建相互独立的运行与访问环境，确保不同租户、不同数据集及不同应用间的操作互不影响的功能。该功能应满足以下要求：

a) 应支持不同沙箱用户之间的数据隔离，确保各用户只能访问其被授权的数据；

b) 应支持同一沙箱用户下不同数据集之间的隔离；

c) 应支持同一沙箱用户下不同应用之间的数据隔离，防止应用间的非授权数据访问；

d) 应支持同一沙箱用户下不同网络区域或网段之间的数据隔离。

9.3.2 透明访问功能

透明访问功能是指数据沙箱提供的一种功能，允许已授权的应用在不改变自身代码、配置及运行环境的前提下，对沙箱内的数据执行增、删、改、查操作。该功能应满足以下要求：

a) 应支持对不同数据形态的透明访问，具体包括：

1) 对文件的透明访问：授权应用可直接对沙箱内文件进行读写，无需迁移或修改应用；

2) 对数据库的透明访问：授权应用可直接连接沙箱管控的数据库实例进行操作，无需修改应用的数据库驱动或连接配置；

3) 对 API 的透明访问：授权应用可按原有的调用方式（包括 URL 路径、请求方法、参数格式）调用沙箱管控的 API，无需进行任何修改。

b) 应支持大数据生态组件的透明访问，包括但不限于 HDFS、HBase、Spark、Kafka、ClickHouse 等；
c) 应支持对人工智能训练及推理应用的透明访问。

9.3.3 泄露防护功能

泄露防护功能是指数据沙箱确保其内部数据（包括导入数据、再生数据及处理过程中产生的临时数据）仅能被授权应用在沙箱内部使用，严禁以任何形式非授权复制或导出至沙箱外部存储或外部环境的功能。该功能应满足以下要求：

a) 应覆盖导入数据、再生数据、共享数据以及处理过程中产生的临时数据的泄露防护；
b) 应确保沙箱内数据可用但不可非授权复制、不可非授权导出；
c) 当数据需要在多个数据沙箱间通过网络进行传输时，应对该传输过程实施泄露防护，确保数据在传输链路上不被明文截获或导出；
d) 应支持将授权的数据处理应用所产生的日志信息回流至数据沙箱内部存储，防止因日志外泄导致的敏感数据泄露。

9.3.4 使用控制功能

使用控制功能是指数据沙箱对授权应用的数据输出行为实施实时管控,确保所有对外输出的结果均符合预设安全策略或数字合约要求的功能。该功能应满足以下要求：

a) 应支持基于安全策略或数字合约，对数据的输出行为进行使用控制；
b) 应支持基于进程名称或进程哈希值，对授权应用进行身份识别与访问控制；
c) 应支持对操作系统超级管理员账户的数据使用行为进行控制；
d) 应支持在通过 SQL 或 API 接口输出数据时，根据安全策略或数字合约进行实时过滤与脱敏处理；
e) 应支持在通过 SQL 或 API 等接口进行数据查询时，将精确数值替换为模糊化结果；
f) 应支持按时间周期（如日、周、月）设置数据使用方的操作次数与操作量阈值，并在超出阈值时自动停止其相应权限；
g) 宜支持对显示或输出的数据内容添加水印。

9.3.5 用后即焚功能

用后即焚功能是指数据沙箱在内存中对密文数据进行解密后使用，并在使用完毕后，按预设策略使用无意义数据或其他数据覆盖该明文所在内存区域的功能。该功能应满足以下要求：

a) 对于按需解密的数据，应在访问结束后自动执行内存区域的用后即焚，即清除或覆盖其明文内容；
b) 应支持通过手动或预设策略两种方式，设置并控制内存中明文数据的清除时间；
c) 应确保清除操作执行后，明文数据在内存中无残留，无法通过技术手段恢复；
d) 对于用户记分、临时评分等临时性数据处理操作，其明文数据仅在内存中解密使用，不写入磁盘或持久化存储。

9.3.6 联合开发功能

联合加工功能是指多个数据提供方及数据使用方在数据沙箱构建的安全环境中,将各自数据进行融合处理与联合计算，并由各方共同管控加工过程、中间数据及最终输出结果的功能。该功能应满足以下要求：

环境创建与管理

a） 应支持根据业务架构或组织架构，创建多方共管的联合加工沙箱环境；
b） 应支持在单机、分布式或集群等不同架构模式下部署联合加工环境；
c） 应支持联合加工环境中多方数据、模型及应用的网络映射供数。

加工与管控功能

a） 应支持多方数据的相互融合计算，同时支持按需实现数据间的隔离；
b） 应支持联合加工各方对加工过程中产生的再生数据进行多方分存；
c） 应支持联合加工各方对可输出的结果和内容进行共同管控，在未获得各方一致同意或授权时，任何一方不得单独持有、查看或导出加工结果。

技术与应用支撑

a） 应支持对接联邦学习、多方安全计算、可信执行环境等隐私计算技术，并在此基础上进行数据联合加工；
b） 应支持基于大数据应用、人工智能应用及数据库应用开展数据联合加工。

9.4 共享流通环节

共享流通支撑数据在跨主体、跨空间、跨地域间的安全流转。远程映射作为底层技术手段实现按需拉取数据，数据共享作为管理行为实现授权交付，共享管理则对共享行为实施策略审批与权限治理。三者协同，确保数据在流通过程中可控、可追溯。

9.4.1 远程映射功能

远程映射功能是指数据沙箱通过请求或推送方式，建立与远端存储管理服务及密钥管理服务的虚拟访问关系，使沙箱能够在本地内存中直接检索和读取远端数据，而无需预先将全量数据导入沙箱的功能。该功能应满足以下要求：

a） 应支持在沙箱内存中构建面向全部映射数据的检索功能，并正确响应应用对数据的读取请求；
b） 应支持按需通过网络从远端存储服务请求或接收映射数据；
c） 应支持通过配置方式或自动化策略，实现映射数据的预请求或预接收；
d） 应支持原始数据导入沙箱后的实时数据映射；
e） 应支持原始数据无需预先导入沙箱的扫描式实时数据映射，即直接对远端存储设备中的数据进行扫描并按需映射；
f） 应支持结构化数据映射和非结构化数据映射；
g） 应确保映射数据在传输和访问过程中的一致性、完整性和可用性。

9.4.2 数据共享功能

数据共享功能是指数据提供方通过数据沙箱的存储管理服务，向数据使用方提供数据访问与获取的功能。该功能应满足以下要求：

a） 应支持跨主体、跨空间、跨地域的数据共享；
b） 应支持通过文件、目录、SQL 接口及 API 接口等多种方式实现数据共享；
c） 应支持基于密钥管理服务和存储管理服务的安全流转式数据共享，确保共享过程中数据的保密性与完整性；
d） 应支持以单个数据沙箱为粒度进行整体数据共享；
e） 应支持数据沙箱间一对多和多对一的共享关系；
f） 应支持在分布式或集群式架构下实现数据共享。

9.4.3 共享管理功能

共享管理功能是指沙箱管理系统对数据提供方的数据共享行为实施策略控制、流程审批、权限治理与审计追溯的功能。该功能应满足以下要求：

管理范围与粒度

a) 应支持对同主体、跨主体、跨空间及跨地域的数据共享进行统一管理；

b) 应支持以文件、目录及数据沙箱为粒度单位，实施差异化的数据共享管理；

c) 应支持一对多分发和多对一汇聚两种共享拓扑模式下的权限与策略管理。

安全管控机制

a) 应支持对多方共享策略进行协同管理，包括策略的制定、协商与统一配置；

b) 应支持共享策略的审批机制，确保共享行为经合法授权后方可执行；

c) 应支持对联合加工数据协同管理，包括策略的联合制定及联合授权；

d) 应支持随时撤销已授予的数据共享访问权限，权限撤销后应立即生效。

审计追溯

a) 应支持对共享管理操作进行完整的日志记录，日志至少包括操作主体、操作内容、操作时间及操作结果；

b) 日志记录的保存期限应不少于六个月，并支持按需导出与审计分析。

9.5 导出交付环节

导出交付是数据离开沙箱的最终环节。平台对导出行为实施严格的审核、授权与核验机制，确保数据仅在合法合规前提下安全输出至沙箱外部，并留存完整的审计记录，实现数据安全闭环。

9.5.1 数据导出功能

数据导出功能是指数据使用方在受控条件下,将数据沙箱内的指定数据安全地转移至沙箱外部环境的功能，整个过程须经过数据提供方审核、授权，并留存完整审计记录。该功能应满足以下要求：

导出管控流程

a) 应支持数据提供方管理员对导出数据的内容、时间、范围及数量进行审核、授权与控制；

b) 应支持对网络导出目的地的身份和 IP 地址进行核验，核验通过后方可执行导出操作。

数据保护与溯源

a) 宜支持基于导出文件扩展属性设置数据标识，用于标识数据的来源、密级和使用限制；

b) 宜支持对文档、图片、音频、视频及数据库等类型的导出内容添加水印或溯源标识；

c) 应支持导出文件携带完整性核验哈希值，以便接收方在获取数据后验证数据的完整性与防篡改性。

审计保障

a) 应确保每次导出操作均生成完整的审计记录，内容至少包括导出主体、导出时间、导出数据范围、审核人及授权结果。

9.6 运维监管环节

运维监管覆盖数据全生命周期的保障与治理。权限管理为各环节提供精细化访问控制，负载管理保障服务高可用，审计管理记录全量操作行为并支持溯源，数据核验在导入、存储、共享、导出等环节验证数据完整性，API对接提供标准化外部集成接口。上述功能共同支撑平台的稳定运行与合规监管。

9.6.1 权限管理功能

权限管理功能是指沙箱管理系统对各类主体（提供方、使用方、运营方、服务方、应用方及监管方）访问其内部数据及资源的操作权限，实施精细化分配、控制与全生命周期治理的功能。该功能应满足以下要求：

管理对象

a) 应支持数据沙箱账号的全生命周期管理，包括账号的创建、授权、验证及删除；
b) 应支持对系统管理员、审计管理员、策略管理员等预设角色的权限进行配置与管理。
c) 策略模型与控制
d) 应支持基于文件、SQL 接口及 API 接口的读写权限管理，实现对数据资源的精细化访问控制；
e) 应支持对数据使用行为进行权限管理，控制维度至少包括应用类型、时间窗口、访问范围及用量上限；
f) 应支持基于网络进出策略的权限管理，管控网络层级的访问行为；
g) 应支持对数据导出策略实施权限管理与审批管理，确保导出行为经合法授权。

协同与共管

a) 应支持多方权限策略的协同管理，实现跨主体场景下的统一权限治理；
b) 应支持数据沙箱安全策略的单方管理或多方共管，以满足不同协作模式下的管控需求。

运维与审计保障

a) 宜支持权限管理策略的可视化编辑、批量导入导出及策略版本管理；
b) 应支持对权限管理操作进行完整的日志记录，日志至少包括操作主体、操作内容、操作时间及操作结果；
c) 日志记录的保存期限应不少于六个月，并支持按需导出与审计分析。

9.6.2 负载管理功能

负载管理功能是指沙箱管理系统对密钥管理服务、存储管理服务等后端依赖服务实施访问入口解析、请求分发调度及准入控制，以保障服务访问的连续性、高效性与安全性的功能。该功能应满足以下要求：

服务解析与准入

a) 应支持对密钥管理服务、存储管理服务等后端服务的访问入口信息进行解析，至少包括 IP 地址及端口；
b) 应支持对数据沙箱及存储管理服务的访问请求实施准入控制，确保仅合法请求被分发至后端服务。

负载调度与故障转移

a) 应支持基于多维度参数的综合负载解析，参数维度应至少包括负载率、响应时间、地理位置及自定义优先级；
b) 应支持根据负载解析结果进行请求分发，并在服务节点发生故障时自动切换至可用节点，确保服务连续性。

9.6.3 审计管理功能

审计管理功能是指沙箱管理系统对内部数据操作行为、内容及流转过程实施完整记录、追溯与合规审查，确保所有操作可问责的功能。该功能应满足以下要求：

日志采集与存储

a) 应支持获取存储管理服务、密钥管理服务及负载管理服务的数据行为日志，并对沙箱数据操作行为进行审计；
b) 应支持受控查看数据沙箱内的数据和日志，并对数据内容及操作日志进行审计和溯源；
c) 日志记录的保存期限应不少于六个月，并支持按需导出与审计分析。

多方存证与防篡改

a) 宜支持多方日志记录的存证与追溯，确保任何单方篡改日志不会影响其他方日志的完整性与真实性，并支持多方存证比对和联合追溯；
b) 宜支持对历史数据提供方、数据加工方及联合加工方的操作存证与身份追溯。

9.6.4 数据核验功能

数据核验功能是指数据沙箱对导入、存储、共享、同步、导出等环节中的数据，通过哈希值对比、密文对比、接口对比等手段，触发式地验证其完整性与一致性的功能。该功能应满足以下要求：

a) 应支持对沙箱内的文件进行基于哈希值的完整性核验；
b) 应支持对不同存储服务间的同步密文数据进行一致性比对核验；
c) 应支持通过接口对数据源与数据目标进行一致性比对核验，核验内容应包括文件内容、文件属性及 SQL 记录等；
d) 应确保上述核验功能覆盖导入、存储、共享、同步、导出等数据全生命周期环节。

9.6.5 API 对接功能

API对接功能是指数据沙箱通过提供标准化的应用程序编程接口，与已授权的第三方应用或系统进行数据交互与能力集成的功能。该功能应满足以下要求：

核心对接功能

a) 应支持通过 API 实现数据的批量导入与导出，并遵循数据分级分类的安全管控要求；
b) 应支持为被授权的第三方应用或系统提供以下管理类接口：对接第三方身份认证系统，实现统一身份认证；输出数据沙箱的运行状态日志；批量获取目标文件的历史数据提供方、加工方及联合加工方的追溯信息。

接口安全要求

a) 所有 API 接口均应实施访问控制机制，对调用方身份进行验证，并对接口访问权限进行授权管理；
b) 所有 API 接口均应采用加密传输协议，确保数据在传输过程中的机密性与完整性。

接口运行保障

a) 应支持 API 的版本管理，确保接口升级或变更时向后兼容；
b) 应支持对 API 调用实施速率限制，防止因过度调用导致服务降级；
c) 应明确 API 的请求与响应格式标准，并优先采用主流规范的接口协议。

10 平台数据流通关系视图

沙箱数据流通关系视图中主要包括操作关系视图和映射关系视图。

10.1 操作关系视图

平台操作关系视图中包括数据导入、数据使用、数据共享、数据流转和数据导出的操作关系，如图4所示：

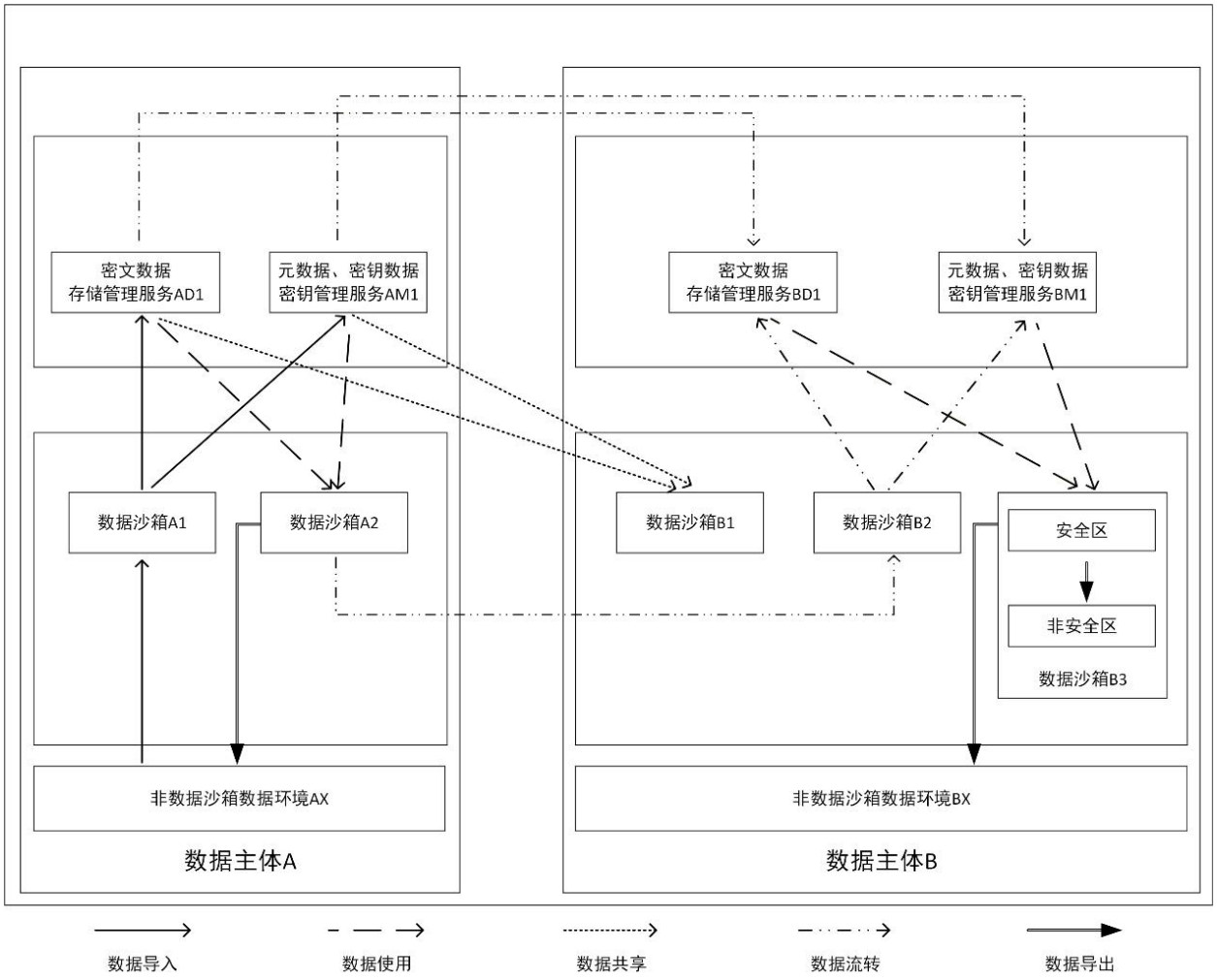

图 4　平台操作关系视图

平台操作关系如下：

数据主体A的非数据沙箱数据环境AX将数据写入数据沙箱A1，数据沙箱A1再将数据写入数据主体A数据沙箱密文密钥服务AD1和AM1的操作，称为“数据沙箱数据导入”；

数据主体A的数据沙箱A2读写数据主体A数据沙箱密文密钥服务AD1和AM1数据的操作，以及数据主体B的数据沙箱B3读写数据主体B数据沙箱密文密钥服务BD1和BM1数据的操作，称为“数据沙箱数据使用”；

数据主体B的数据沙箱B1读写数据主体A数据沙箱密文密钥服务AD1和AM1数据的操作，称为“数据沙箱数据共享”；

数据主体A的存储管理服务AD1和密钥管理服务AM1对数据主体B的BD1和BM1服务的密文、密钥传输操作，称为“数据沙箱密文密钥数据流转”；

数据主体A的数据沙箱A2对数据主体B数据沙箱数据沙箱B2的数据传输操作，称为“数据沙箱数据流转”。

数据主体A的数据沙箱A2和数据主体B的数据沙箱B3将数据复制到非数据沙箱数据环境AX和BX的操作，以及数据主体B的数据沙箱B3将数据沙箱内数据复制到数据沙箱外的操作，称为“数据沙箱数据导出”。

10.2　映射关系视图

平台数据映射可实现数据提供方对数据使用方的快速数据传递，数据沙箱数据映射关系视图中分别表示了导入式数据映射和扫描式数据映射的映射关系，扫描式数据映射又分为扫描式明文数据映射和扫描式密文数据映射，如图5所示：

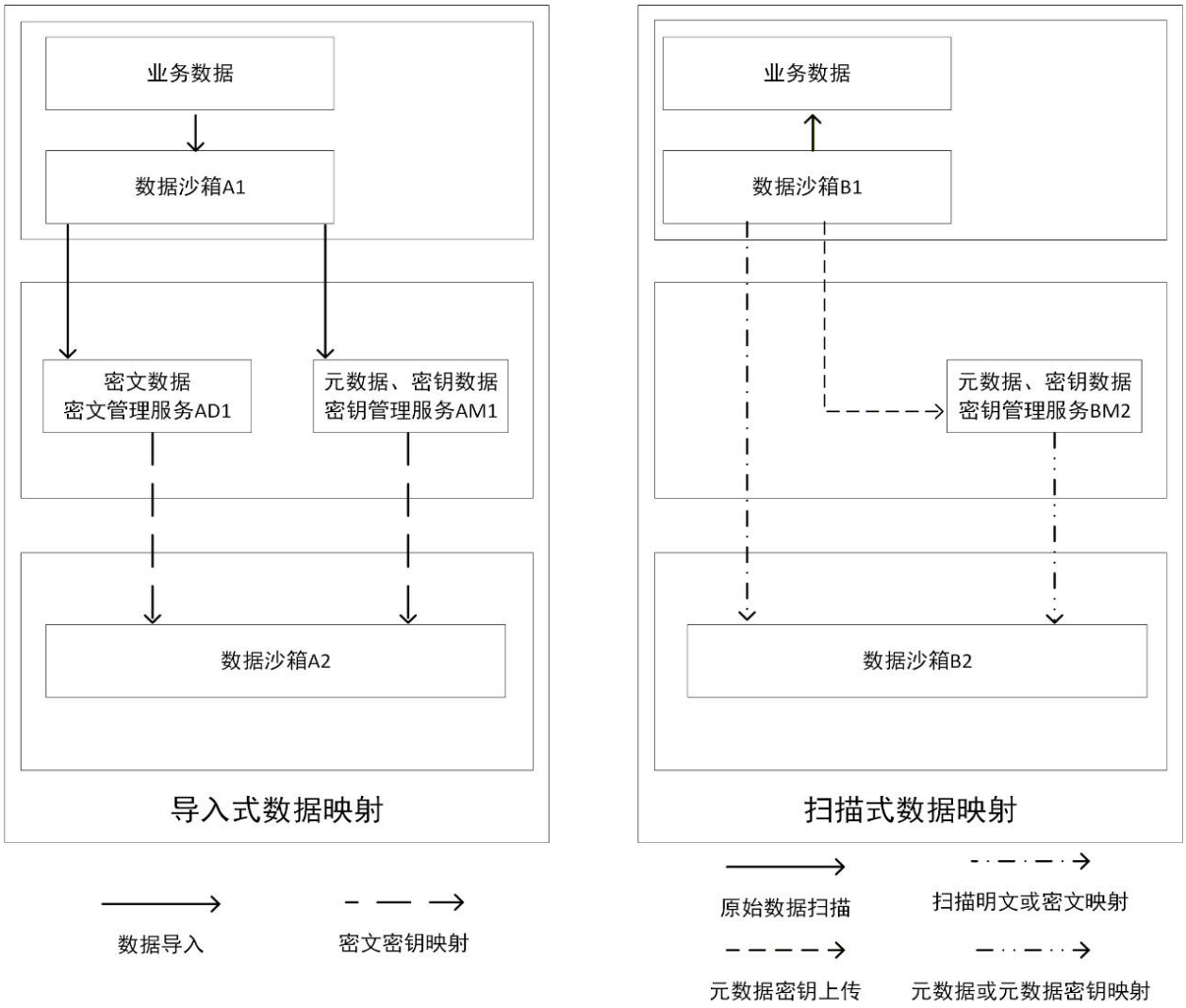

图 5　平台数据映射关系视图

数据沙箱映射关系如下：

数据沙箱A1将业务数据导入数据沙箱密文密钥服务后，再由数据沙箱密文密钥服务将数据元素对应关系和对应内容按需同步到数据沙箱A2的操作，称为“导入式数据映射”；

数据沙箱B1扫描业务数据并将扫描数据的元数据提取写入数据沙箱密文密钥服务后，再由数据沙箱B1和密钥管理服务BM2将数据元素对应关系和对应明文数据内容按需同步到数据沙箱B2的操作，称为“扫描式明文数据映射”；

数据沙箱B1扫描业务数据并将扫描数据的元数据和加密密钥写入数据沙箱密文密钥服务后，再由数据沙箱B1和密钥管理服务BM2将数据元素对应关系、对应加密数据内容、元数据和密钥按需同步到数据沙箱B2的操作，称为“扫描式密文数据映射”。

附 录 A
（资料性）
数据沙箱公共服务平台典型使用方式

A.1 受控访问与安全使用

该方式是基本的使用方式，核心特征是“数据不落地，使用可追溯”。

核心特点：

a) 数据驻留服务端：所有数据通过导入、映射或共享进入沙箱后，始终以密文形态驻留在服务端存储，使用者无法将明文原始数据和中间结果下载至本地。

b) 精细化权限管控：访问权限不限于“使用方等各方能看什么”，而是细化到可使用的时间窗口、操作频次、数据范围及用量上限，超出阈值即自动阻断。

c) 透明无感使用：授权使用方可沿用原有工具和操作习惯直接访问数据，无需修改应用代码或切换环境，对数据已被管控这一事实几乎无感知。

d) 操作全程防泄露：明文数据仅在内存中实时解密使用，复制、外发等操作被系统级阻断，确保使用过程中数据不被非授权导出。

e) 结果受控输出：分析结果并非随意导出，须经多方审核流程确认不包含原始敏感信息，输出时可强制添加溯源信息并携带完整性校验哈希值。

f) 操作全链可审计：所有访问与操作行为被完整记录，形成不可篡改的审计日志，支持事后追溯与合规审查。

A.2 多方协作与联合加工

该方式是跨主体数据协作的使用方式，核心特征是“数据不出域，价值共创造”。

核心特点：

a) 环境多方共建：协作环境由多个数据提供方和使用方共同创建与管理，任何单一方无权独自决定环境配置。

b) 数据逻辑融合、物理隔离：各方数据可在沙箱内进行融合加工，但底层存储基于租户实施逻辑隔离，确保提供方间原始数据互不可见。

c) 结果多方共管：联合加工产生的中间数据与最终结果，由数据沙箱碎化分存至各方共同管控。在未获得一致授权的情况下，任何一方不得单独查看、持有或导出。

d) 策略协同制定：数据的访问权限、使用范围、输出条件等安全策略，由各方协同协商后统一配置，而非由单方强行设定。

e) 数据可验证：支持数据核验技术，确保提供数据符合约定，输出数据可被各方验证其合规性与完整性。

f) 操作多方存证：联合加工过程中的每一步操作，日志被多方分别存证，任何单方篡改均可被比对发现，确保协作过程不可抵赖。

A.3 数据服务与安全开放

该方式是将数据作为服务向外提供的方式，核心特征是“原始数据不出门，数据产品可交付”。

核心特点：

a) 接口化交付：数据不以文件形式直接交付，而是通过 SQL 接口、API 接口等形式提供查询与分析服务，使用者只能调用接口获取结果。

b) 实时管控与过滤：每次接口调用均受安全策略实时管控，敏感字段可被动态脱敏，查询结果可被过滤处理后再返回。

c) 用量可计量、可限制：可按日、周、月等周期设置调用次数与数据访问量上限，到达阈值自动停止服务，实现数据服务的量化管控。
d) 模糊化输出：对于非精确查询场景，支持将精确值自动替换为模糊化结果，在保护隐私的同时满足趋势分析需求。
e) 服务状态可监控：数据提供方可实时查看数据服务的健康状态、调用情况，并接收异常告警，做到服务状态全程可视。
f) 权限可随时收回：已授予的访问权限支持随时撤销，撤销后立即生效，确保数据主权始终在提供方手中。

A.4 安全运维与态势感知

该方式是面向管理者的使用方式，核心特征是“特权操作受管控，全局风险可感知”。
核心特点：

a) 多权分立：密钥管理员、密文管理员、审计管理员、策略管理员等角色权限相互制约，防止单一角色拥有过大权力。
b) 特权行为受控：即使是操作系统超级管理员，其对沙箱内数据的访问和使用行为同样受沙箱策略管控，无法绕过安全机制。
c) 全局统一视图：可在一个界面中查看所有沙箱实例、存储服务、密钥服务的运行状态与健康度，实现统一运维。
d) 实时风险感知：汇聚各服务实时日志，对数据导入、使用、流转、导出等环节进行全局态势监控，支持自定义阈值预警。
e) 内容级审计：不仅记录操作日志，还可受控查看沙箱内的实际数据内容，对数据操作进行内容级审计与溯源。
f) 日志防篡改存证：审计日志支持多方存证，任何单方无法篡改，为合规审查与事故溯源提供可靠依据。

附 录 B
（资料性）
数据沙箱公共服务平台典型应用场景

B.1 原始数据防复制交付应用场景

本场景描述数据提供方将数据安全导入公共服务平台，授权数据使用方在受控沙箱环境中完成数据处理与分析，并审核输出结果的全过程。

数据提供方通过数据沙箱X1的数据导入功能，将原始数据加密上传至存储管理服务，密钥由密钥管理服务统一托管。数据使用方经身份认证与权限管理后，进入受提供方管控的数据沙箱环境X2，在内存中对已授权数据解密后进行分析、建模等处理。操作过程受泄露防护功能保护，处理结果暂存于沙箱内部存储。

加工完成后，数据使用方可在沙箱内受控使用经脱敏或过滤后的数据，也可提交结果数据导出申请。经数据提供方审核与授权后，通过数据沙箱X3的数据导出功能将结果数据安全交付至使用方。如图B.1所示：

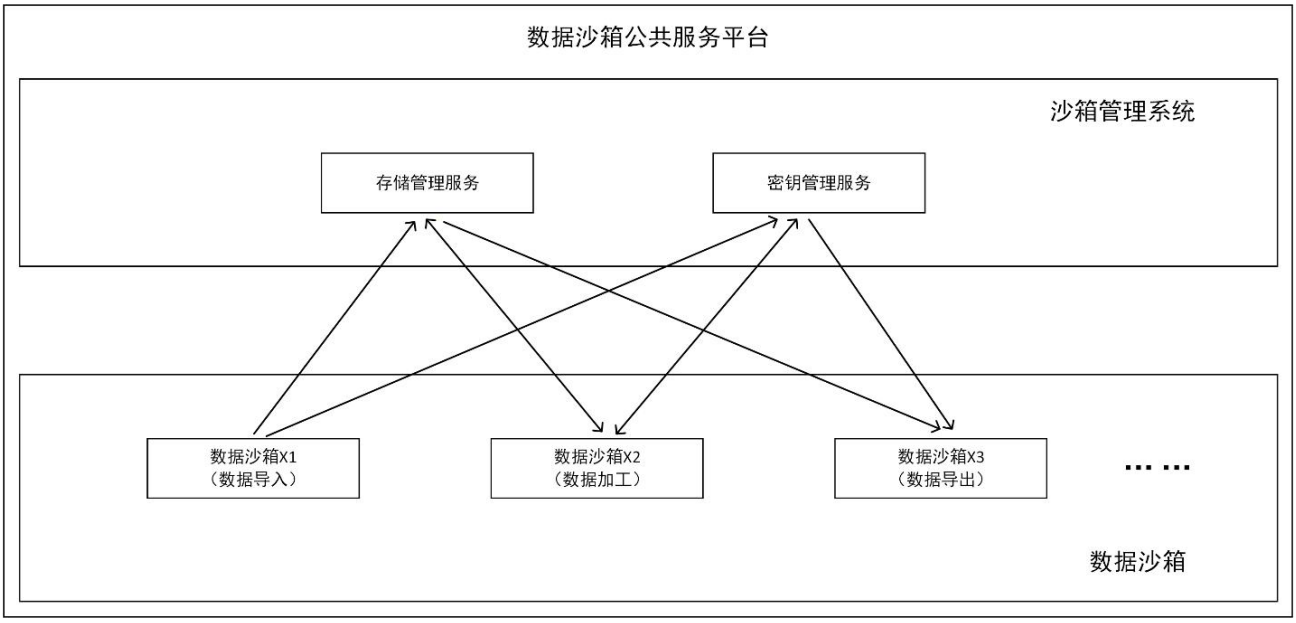

图 B.1 原始数据防复制交付场景示意图

B.2 多方数据联合加工应用场景

本场景描述多个数据提供方与数据使用方共同创建融合沙箱环境，实现多方数据安全联合加工，并对产出结果实施共同管控的全过程。

数据提供方X与Y分别通过数据导入功能，将各自原始数据加密存储至自管的存储管理服务XA和YA中。双方通过远程映射功能，将各自的密文数据映射至由数据提供方及使用方三方共同创建的联合加工沙箱环境Z2。在该环境中，数据在内存中被实时解密后进行联合加工、处理与分析。运算产生的新结果数据，将被实时碎化加密分存至X、Y、Z各自管控的存储管理服务XA、ZA、YA中，使数据提供方与使用方均对新结果数据拥有共同管控的权利。

数据使用方Z可直接在联合加工沙箱中查询或使用已授权数据；当需要导出结果数据时，使用方Z提申请，由数据提供方X和Y共同审核与授权后，使用数据沙箱Z3将结果数据导出。如图B.2所示：

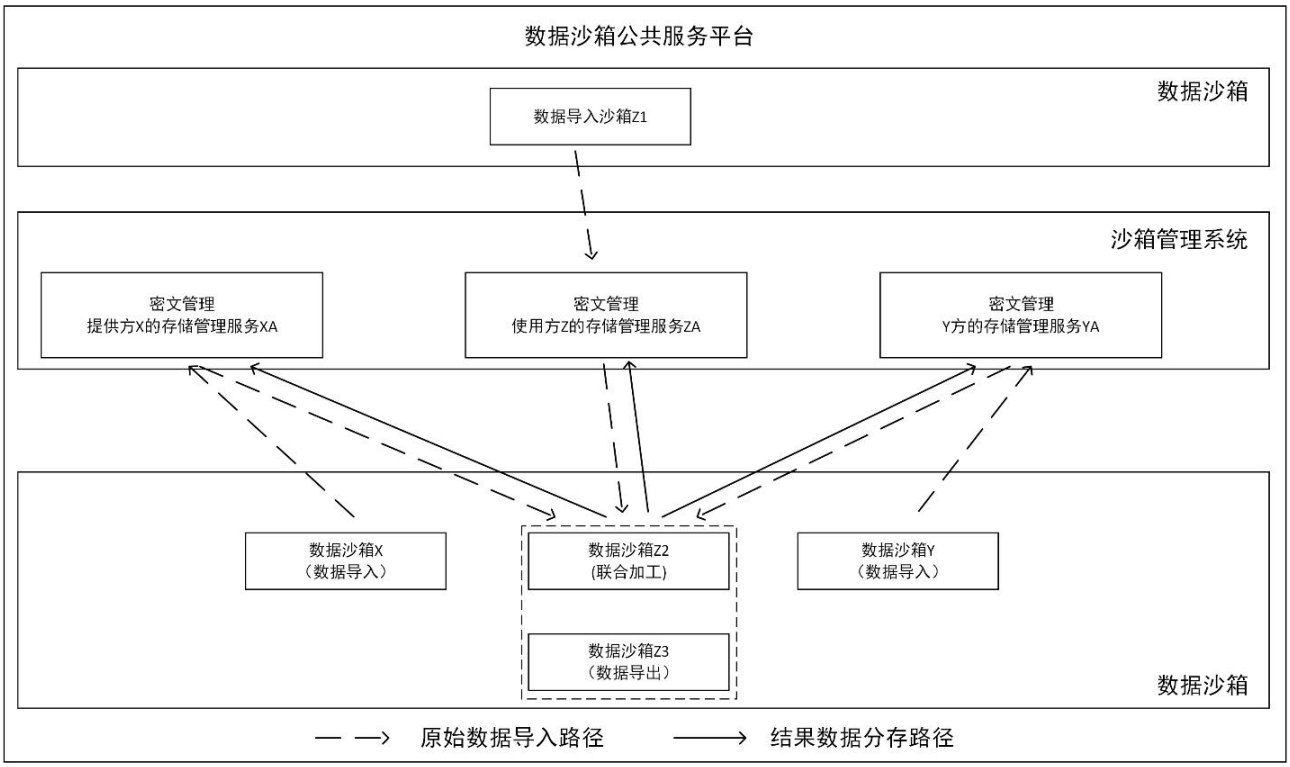

图 B.2 多方数据联合加工应用场景示意图

B.3 人工智能数据安全防护应用场景

本场景描述在人工智能业务中，数据沙箱对高质量数据集的上传、加工、训练、推理及模型导出等环节提供全流程数据安全保障的应用方式。

在人工智能全程供数用数模式下，各环节的数据安全由数据沙箱公共服务平台统一管控，具体流程如下：

数据准备与加工阶段：数据提供方通过数据导入功能，将高质量数据集加密存储至存储管理服务A1，密钥由密钥管理服务M1统一托管。需加工时，通过远程映射功能将数据集映射至指定的加工沙箱环境X2n，加工完成后的结果数据加密回存至存储管理服务A1。

模型训练阶段：加工后的数据集密文按需同步至模型训练环境的密文存储服务A2。训练任务启动时，数据沙箱X3n将数据实时解密后供给模型训练设备使用，训练完成的模型文件由沙箱加密后回存至存储管理服务A2。

模型推理与导出阶段：训练完成的模型密文按需同步至推理服务环境的密文存储服务A3，推理设备通过数据沙箱X4n获取解密后的模型文件执行推理任务。当模型需要导出时，须经相关参与方共同授权，通过数据沙箱X5n导出功能完成完整性核验后安全交付。

此外，数据沙箱公共服务平台也可针对上述环节中的特定节点（如仅加工环节、仅训练环节或仅推理环节）提供针对性保护，灵活适配不同业务需求。如图B.3所示：

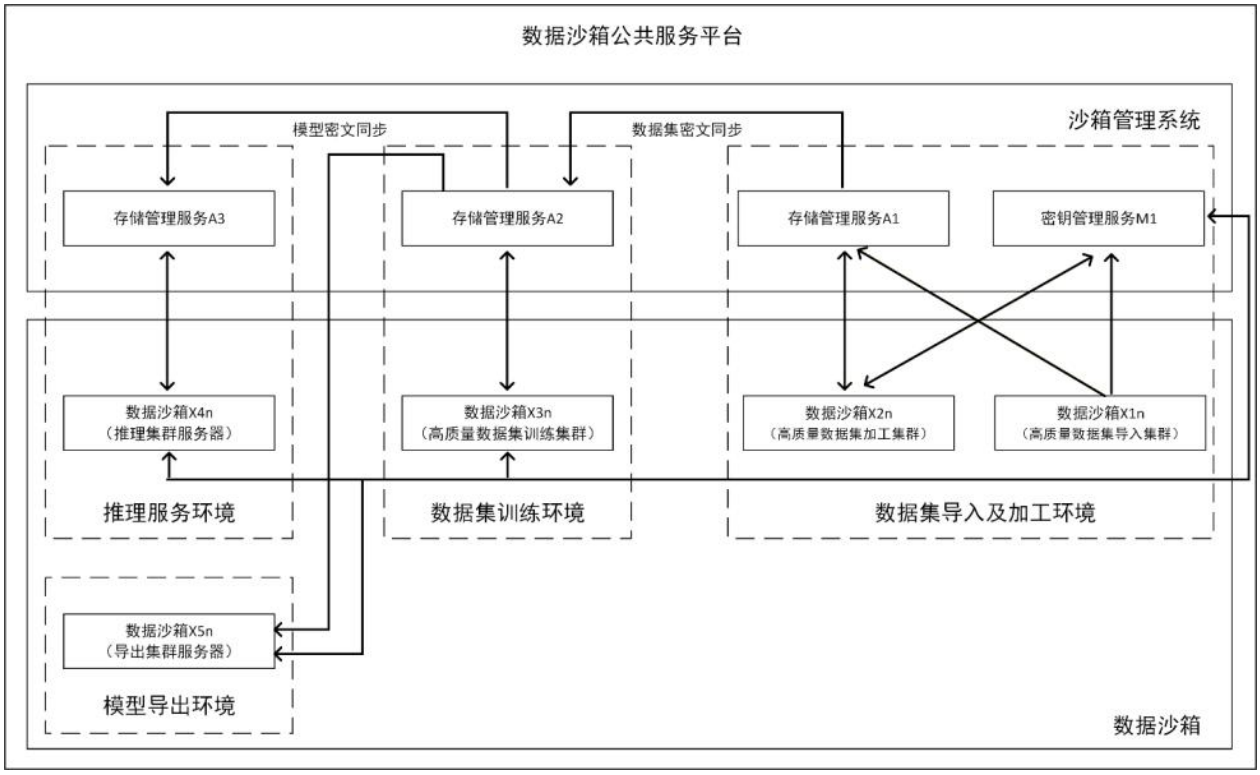

图 B.3 人工智能数据安全防护应用场景示意图

B.4 跨境数据安全管控应用场景

本场景描述在跨境业务中，数据提供方通过境内密钥管理服务与存储管理服务，实现对境外数据沙箱的远程管控，保障出境数据在境外使用过程中的安全性与可控性。本场景提供两种典型模式。

模式一：境内存储、境外映射

数据提供方通过数据沙箱X1，将数据上传至境内密钥管理服务M1及存储管理服务A1。境内数据通过远程映射功能直接映射至境外数据沙箱Y1，数据使用方在Y1中实时解密并使用数据，操作过程受泄露防护功能管控。

该模式的核心特征是数据始终以密文形式存储在境内，境外环境不落盘使用映射数据，适用于应用相对简单、数据吞吐量较小的场景。

模式二：境内管控、境外密文同步

数据提供方通过数据沙箱X2，将数据上传至境内密钥管理服务M2及存储管理服务A2。境内存储服务通过密文同步方式，将加密数据同步至境外存储管理服务A3。数据使用方通过境外数据沙箱Y2访问A3中的密文数据并实时解密使用，境外沙箱的密钥请求须经境内密钥管理服务M2授权。

该模式的核心特征是密文数据在境外就近存储，保障读写效率；同时密钥由境内管控，数据提供方可对境外沙箱的数据访问行为实施远程管控与审计，适用于数据吞吐量大、对访问延迟敏感的场景。如图B.4所示：

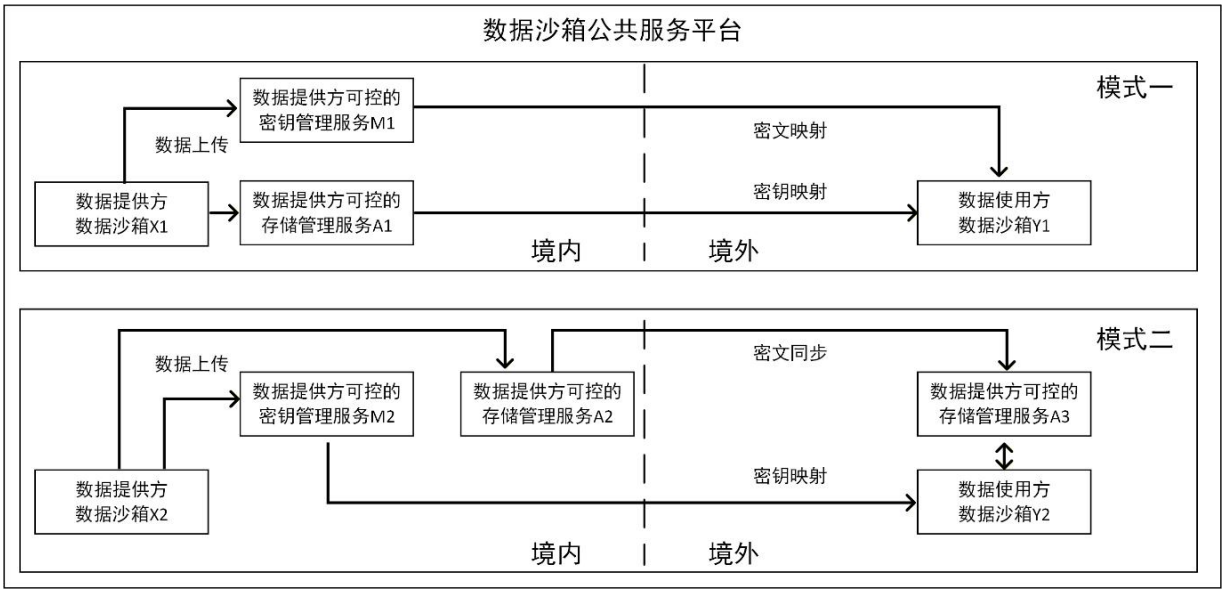

图 B.4 跨境数据安全管控应用场景示意图

B.5 跨地域分布式部署应用场景

本场景描述数据沙箱公共服务平台采用跨地域分布式部署模式（非中心化统一调度），各城市或地域独立部署存储服务与沙箱实例，通过集中管理的元数据及密钥实现跨地域数据的精准定位与按需映射。该模式为分散部署，不设置物理中心节点，以满足如国家数据集团归集各省数据等业务需求。

B、C、D、E各城市的数据分别存储于各自的存储管理服务中。A城市仅需使用各城市的一小部分数据时，通过公共服务平台集中管理的元数据及密钥，可快速定位目标数据所在位置，并通过远程映射功能直接获取相应的密文数据，无需等待全量数据传输。

A城市数据使用方通过本地数据沙箱即可对已映射数据进行查询、分析、处理等操作，操作过程不受数据物理存储位置的限制，实现跨地域的快速调数与安全用数。如图B.5所示：

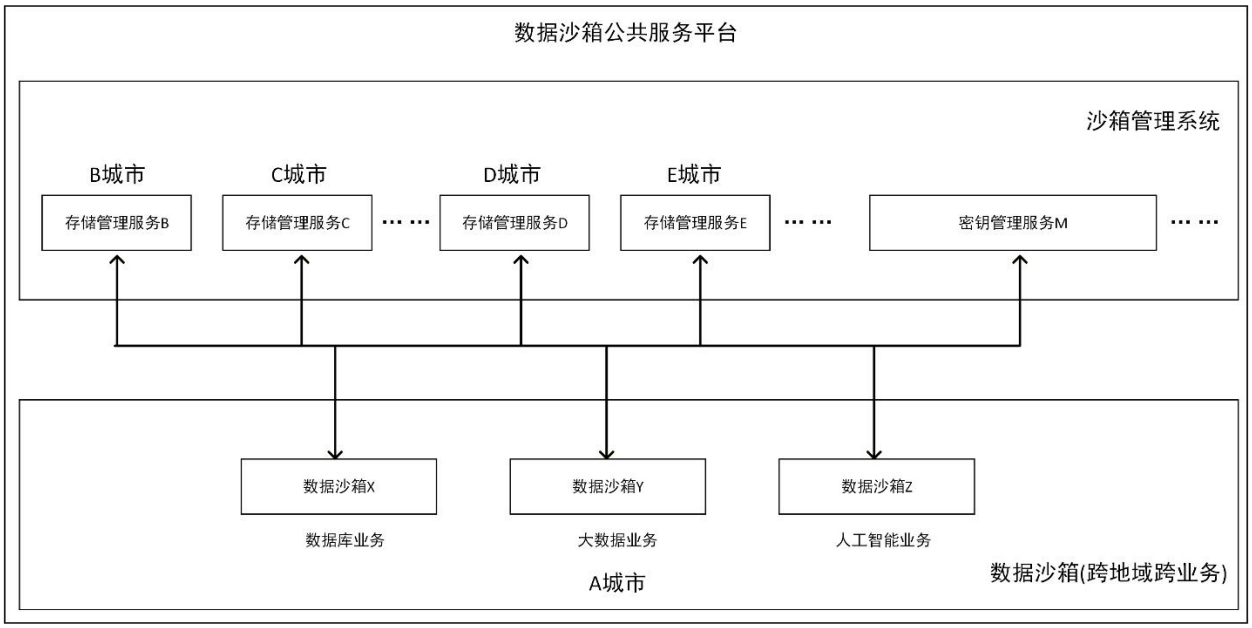

图 B.5 跨地域按需调数用数应用场景示意图

参 考 文 献

[1] 《国家数据基础设施建设指引》
[2] 《关于建立公共数据资源授权运营价格形成机制的通知》
[3] 《公共数据资源登记管理暂行办法》
[4] 《公共数据资源授权运营实施规范（试行）》
[5] 《关于完善数据流通安全治理 更好促进数据要素市场化价值化的实施方案》
[6] 《关于促进数据产业高质量发展的指导意见》
[7] 《关于促进企业数据资源开发利用的意见》
[8] 《可信数据空间发展行动计划（2024 —2028年）》
[9] 《国家数据标准体系建设指南》
[10]《“数据要素×”三年行动计划（2024—2026年）》
[11]《关于深入实施“东数西算”工程 加快构建全国一体化算力网的实施意见》
[12]《中华人民共和国网络安全法》
[13]《中华人民共和国个人信息保护法》
[14]《关键信息基础设施商用密码使用管理规定》
[15]《中华人民共和国数据安全法》
[16]《政务数据共享条例》
[17]《关于构建数据基础制度更好发挥数据要素作用的意见》
[18]《关于促进数据产业高质量发展的指导意见》
[19]《数据基础设施 参考架构》
[20]《数据基础设施 互联互通基本要求》
[21]《数据基础设施 接入连接器技术要求》
[22]《数据基础设施 用户身份管理和接入规范》
[23]《数据基础设施 安全能力通用要求》
[24]《数据基础设施 标识要求》
[25]《数据基础设施 数据目录描述规范》
[26]《可信数据空间 技术架构》
[27]《可信数据空间 数字合约技术要求》
[28]《可信数据空间 使用控制技术要求》