ICS
CCS 点击此处添加 CCS 号

DB11

北　京　市　地　方　标　准

DB 11/T XXXX—XXXX

数据流通跨域管控技术要求

Technical requirements for cross-domain management and control of data transaction

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX－XX－XX 发布　　XXXX－XX－XX 实施

北京市市场监督管理局　　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。

本文件由北京市政务服务和数据管理局、北京市经济和信息化局提出。

本文件由北京市政务服务和数据管理局、北京市经济和信息化局归口管理并组织实施。

本文件起草单位：北京国际大数据交易所有限责任公司。

本文件主要起草人：

数据流通跨域管控技术要求

1 范围

本文件规定了数据流通跨域管控的技术要求，描述了数据流通跨域流程和数据流通跨域的管控要求。

本文件适用于数据流通跨域管控方案的设计、部署和使用，也可为测试评估提供参考。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

域 domain

在一组统一的策略下运行的、有明确边界的一组实体。相关方可以根据数据流通场景、合规要求、安全假设等实际情况，按照安全策略，所属主体等维度去划分不同的域，如安全域、主体域等。

[来源：GB/T 25069-2020，3.741，有修改]

示例1：安全域指遵从共同安全策略的一组实体。

示例2：主体域是单一主体设置策略下的一组实体。

3.2

实体 entity

现实世界中独立存在的对象。

[来源：GB/T 40651-2021，3.10]

3.3

数据 data

任何以电子或其他方式对信息的记录。数据在不同视角下被称为原始数据、衍生数据、数据资源、数据产品和服务、数据资产、数据要素等。

3.4

原始数据 raw data

在一次计算任务中的输入信息，由任务的数据提供方提供。

3.5

跨域管控 cross-domain management and control

数据离开提供方的域后，数据提供方仍然能够有效地控制数据的流转和使用，避免其被滥用、窃取或者非预期的使用。

3.6

属主 owner

对某个资源拥有控制管理权限的实体。

[来源：GB/T 25069-2022，3.137]

3.7

标识 identifier

用于无歧义地标识对象的全局唯一值。

[来源：GB/T 25069-2022，3.138，有修改]

4 缩略语

下列缩略语适用于本文件。

API:应用程序编程接口（Application Programming Interface）

IPSec:互联网安全协议(Internet Protocol Security)

TEE：可信执行环境(Trusted Execution Environment)

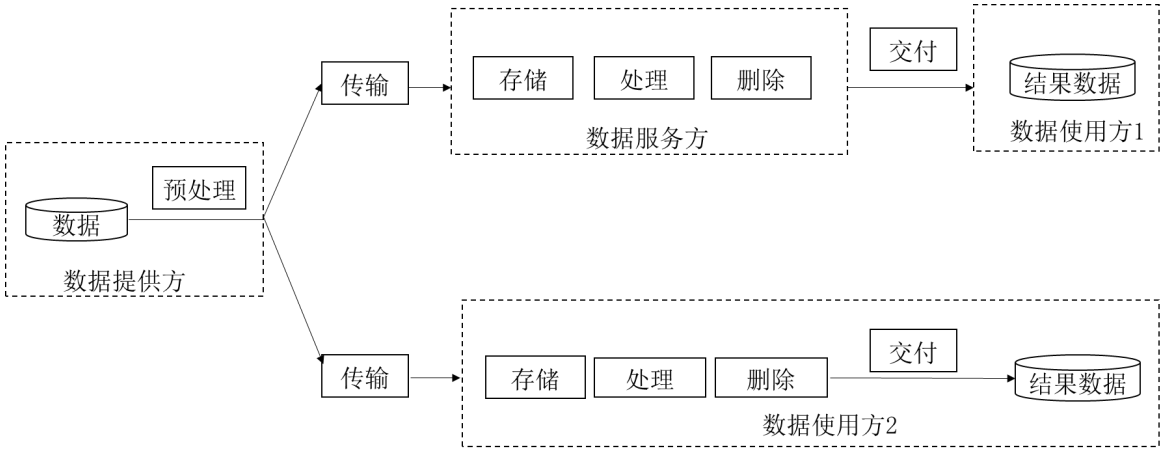
TPM:可信平台模块（Trusted Platform Module）

TPCM:可信平台控制模块（Trusted Platform Control Module）

5 数据流通跨域

5.1 概述

数据流通跨域是指以数据作为流通对象，按照一定规则从数据提供方的域内流出，以便数据使用方能进行加工使用，数据流通跨域框架如图1所示。数据可以直接流通到数据使用方域内，也可以通过数据服务方进行分析、聚合、加工等处理后将结果数据提供给数据使用方。数据流通跨域过程中的关键环节包括：数据预处理、数据传输、数据存储、数据处理、结果交付和数据删除等过程。



注:此图虚线框表示不同参与方的域，实线框表示数据处理功能模块。

图 1 数据流通跨域示意图

5.2 数据预处理

数据提供方在数据流通到域外之前，在本域内基于流通技术要求对数据进行的处理：

a) 在基于秘密分享的安全多方计算中，数据预处理将数据分割成多个分片，以便后续将部分数据分片提供给其他参与方用于计算；

b) 在基于联邦学习的数据流通中，在本域内对数据进行训练，生成用于数据流通的模型，在后续的数据流通中将模型的部分参数提供给其他参与方；

c) 在去标识化数据流通中，在本域内对数据中的个人标识信息进行替换、删除或泛化等操作后，在后续的数据流通过程中将去标识化数据提供给其他参与方；
d) 在同态计算的数据流通中，在本域内对数据进行同态加密后，在后续流通过程中将加密后的数据提供给其他参与方进行计算；
e) 在基于可信执行环境的数据流通中，在本域内对数据进行加密后，在后续流通过程中将加密后的数据提供给其他参与方。加密的数据只能在硬件隔离的环境内解密并进行计算。

5.3 数据传输

数据通过在线或离线等多种方式，从数据提供方传输到数据服务方或数据使用方。

5.4 数据存储

数据保存在不同参与方的存储介质中，如数据服务方的服务器或数据使用方的本地设备。

5.5 数据处理

将数据进行加工、聚合、分析等操作，得到用于优化生产经营的结果数据。
a) 数据聚合：将多主体的数据进行汇聚、重新格式化，并以汇总形式呈现的处理活动；
b) 数据加工：数据加工一般会涉及数据自身的改变，需要先读取数据，并经过变换、转换、纠错、编码、分析、挖掘、脱敏等数据操作生成新数据的处理活动；
c) 数据分析：通过特定的技术和方法，对数据进行整理、研究、推理和概括总结，从数据中提取有用信息、发现规律、形成结论的处理活动。

5.6 数据交付

数据处理之后会依据约定的方式将数据交付给数据使用方：
a) 直接交付，数据使用方通过API接口、数据报告等形式下载数据集，如全量数据集、筛选数据集、单个数据项等；
b) 基于计算的交付，数据使用方通过使用控制、隐私保护计算等方式获得结果数据，如模型结果、求交的数据集等。

5.7 数据删除

数据使用方和数据服务方将数据及相关信息进行移除或销毁，确保其不能通过推断、重建等方式恢复数据。

6 管控技术要求

6.1 管控框架

数据流通跨域场景下，数据离开数据提供方域后，除了通用数据安全风险外，在域间流通的各个环节都存在安全风险（见附录A），会导致数据泄露、滥用等。数据流通跨域管控的框架图如图2所示，数据提供方灵活地配置加工使用策略，明确数据的使用时间、用法用量等规则，结合标识和属主、权限管理和审计溯源等控制面功能模块管控数据流通跨域，确保数据处理与策略保持一致。数据流通参与方利用一种或多种技术手段（见附录B），在数据预处理、传输、存储、加工使用和结果分发等关键步骤中执行策略，保证参与方不能在非授权的情况下直接或间接获得数据提供方的数据，而且只能按照约定范围、用法来使用数据，防止数据泄露和不当使用，场景详见附录C。

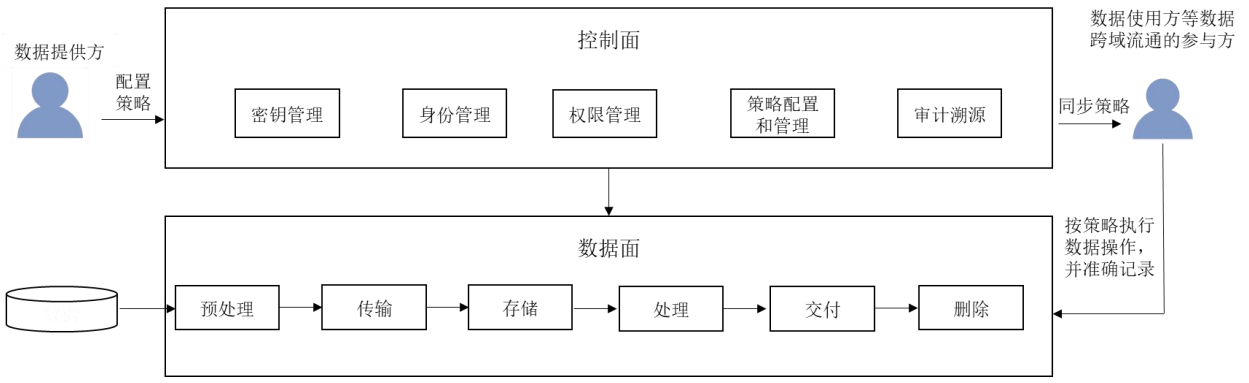


图 2　跨域管控框架图

6.2　数据面要求

6.2.1　数据预处理

数据流通跨域管控过程中，数据预处理符合以下要求：

a）　对涉及个人信息，应取得有效授权或进行匿名化处理，对于无法满足前述条件的，进行去标识化处理，使其在不借助额外信息的情况下无法识别特定自然人；

b）　进行数据传输之前，应验证数据接收方身份的真实性，仅当真实性验证通过时才提供数据；

c）　进行数据传输之前，应验证数据接收方环境的安全性，仅当安全性验证通过时才提供数据；

d）　应确定数据的使用策略，具体加工使用策略应符合6.3.4的要求；

e）　以上活动应均在数据提供方域内执行。

6.2.2　数据传输

数据流通跨域管控过程中，数据传输符合以下要求：

a）　应对接收节点进行身份验证；

b）　应采用校验技术或密码技术保证通信过程中数据的完整性；

c）　应采用密码技术保证通信过程中数据的保密性。

6.2.3　数据存储

数据流通跨域管控过程中，数据存储符合以下要求：

a）　数据和对应的加工使用策略在存储时应保持关联关系一致性；

b）　仅获得数据提供方授权的用户按照加工使用策略的规则访问数据；

c）　应采用校验技术或密码技术保证数据和加工使用策略存储过程中的完整性；

d）　应采用密码技术保证数据和加工使用策略在存储过程中的保密性；

e）　应实施容灾备份和存储介质安全管理，定期开展数据恢复测试和灾难恢复演练，对备份数据的有效性和可用性进行检查和恢复验证。

6.2.4　数据处理

6.2.4.1　处理的过程

数据流通跨域管控过程中，处理过程符合以下要求：

a) 应确保加工使用过程不能泄露数据提供方非授权以外的信息泄露；
b) 加工使用过程中产生的衍生数据应按照数据提供方约定或者加工使用策略约定的方式和范围进行分发；
c) 当使用基于密码学技术手段保护计算安全时，密码技术的安全强度应不低于一定的阈值，如128比特；
d) 加工使用结果的输出方式应和加工使用策略保持一致性；
e) 当存在大量或敏感数据的导出时，需要数据提供方明确授权审批；
f) 应在结束后及时清理加工使用痕迹；
g) 应记录数据加工使用的过程，确保数据提供方能够通过审计回溯数据加工使用是否如预期。

6.2.4.2 处理的应用

数据流通跨域管控过程中，对数据进行处理的应用符合以下要求：
a) 代码实现应与加工使用策略保持一致；
b) 应支持对应用程序及其依赖进行完整性证明，为通过验证的应用分配唯一的标识；
c) 应确保应用标识与应用之间的关联是一致且完整；
d) 应支持向数据提供方提供完整性验证，如远程证明；
e) 涉及多个加工使用应用时，一个加工使用应用要对相连接的其他加工使用应用进行鉴别，鉴别的依据要能够防止被篡改和替换。

6.2.4.3 处理的环境

数据流通跨域管控过程中，处理环境符合以下要求：
a) 应支持向数据提供方提供环境安全性证明，如系统组件、关键配置符合数据提供方预期；
b) 应抵御内部人员的攻击，如平台运维人员；
c) 应在不泄露原始数据的前提下，支持使用合成数据、混淆数据等方法进行数据产品和服务的研发；
d) 应防止外部恶意软件和攻击者的攻击，如侧信道攻击；
e) 应设置身份鉴别、访问控制、隔离存储、加密、脱敏等安全措施，并确保环境中的漏洞已得到处置；
f) 应提供隔离的开发测试环境和生产环境。

6.2.5 数据交付

数据流通跨域管控过程中，数据交付符合以下要求：
a) 结果应按照与数据提供方提前约定的方式和范围进行分发；
b) 应对结果接收方身份或节点进行验证，验证通过后才分发结果；
c) 应对输出结果进行保护，只有指定的结果使用方才可得到结果；
d) 应支持一个或多个结果使用方，每个结果使用方按照加工使用策略得到的相同或不同的结果数据；
e) 结果数据宜能够抵御差分攻击，防止根据多次的计算结果反推个体敏感信息。

6.3 控制面要求

6.3.1 身份管理

数据流通跨域过程中，身份管理符合以下要求：

a) 应采用口令、密码技术、生物技术等一种或多种组合的鉴别技术对用户进行身份鉴别；
b) 数据流通跨域过程中，数据流通跨域参与方，数据处理的应用，以及数据提供方提供的数据都应该具有唯一标识确定，标识与参与方、数据应用以及数据资源的对应关系不能被篡改、伪造。

6.3.2 权限管理

数据流通跨域管控过程中，权限管理支持数据提供方实现对数据的有效授权和鉴权，符合以下要求：
a) 应支持定义多类角色身份，并支持按类分配不同的权限；
b) 应支持针对不同数据和应用服务定义具体的可见权限、访问权限和操作权限，对于已授权的开放资源，支持数据拥有方关闭数据授权；
c) 宜支持基于角色的访问控制（RBAC）、基于属性的访问控制（ABAC）等模型，实现细粒度的权限管理；
d) 宜支持权限审核功能，定期检查和审计用户权限配置，防止权限滥用和权限泄露；
e) 数据加工使用的授权应由所属的数据提供方生成或者以数据提供方认可的方式生成；
f) 应确保授权的真实性和完整性；
g) 授权的内容应包含但不限于数据、使用范围、使用对象、加工使用策略和授权有效期等；
h) 在数据加工使用、访问前，应进行鉴权操作；
i) 鉴权操作应具有完整的校验链条，包含但不限于请求身份、数据信息、数据提供方签名、使用策略等，确保授权信息是由真实的数据提供方签发的且数据的请求内容和授权内容是一致的等当授权、鉴权操作是计算方平台或第三方提供时，要保证授权、鉴权操作不能被平台或第三方的内容人员伪造、篡改或绕过。

6.3.3 密钥管理

数据流通跨域过程中，密钥管理符合以下要求：
a) 应支持密钥的生成，包括对称密钥及非对称密钥；
b) 应支持密钥的安全存储，包括根密钥、派生密钥等。在存储过程中应支持密钥及其属性的完整性进行保护；
c) 应支持按照每个密钥生成时设定的用途进行使用；
d) 应支持对密钥进行访问控制；
e) 应支持对密钥的使用情况进行记录；
f) 应支持密钥的更新，密钥更新时应变更密钥版本，被更新的旧密钥应进行停用或销毁；
g) 应支持密钥停用，已停用的密钥不应重新使用，仅可应用解密或验证历史数据；
h) 应支持密钥的销毁，对密钥销毁时，应销毁该密钥的所有副本，密钥销毁应具备不可逆性。

6.3.4 策略配置和管理

6.3.4.1 策略配置

6.3.4.1.1 概述

数据在跨域流通过程中，数据提供方可以设置加工使用策略，包含时间、使用者、用法和用量、使用通知等一种或多种维度信息的管控。

6.3.4.1.2 时间

加工使用策略支持时间维度的管控，符合以下要求：

a） 限定数据访问时间，如只在特定时间段内允许访问数据；
b） 将数据使用限制在特定时间间隔内，如在指定的时间间隔内允许或禁止数据的使用；
c） 将数据使用限制在特定时间持续内，如只允许在数据接收后的一定时间内使用。

6.3.4.1.3 使用者

加工使用策略支持使用者维度的管控，符合以下要求：
a） 指定数据使用者所在地域，如只允许特定国家或地区的人使用数据；
b） 将数据使用限制在指定的设备上，如数据只能在特定设备上被访问；
c） 将数据使用限制在特定系统或应用程序组上，如只允许在完整性证明后的环境和应用中使用数据；
d） 将数据使用限制在特定一方或用户组上；
e） 限制数据使用的环境状态，如合同终止或者防火墙关闭时禁止使用数据等。

6.3.4.1.4 用法和用量

加工使用策略支持用法和用量维度的管控，符合以下要求：
a） 使用后超过即禁止继续访问；
b） 数据使用次数不超过一定次数；
c） 允许或禁止数据的特定操作，如读取、分发、打印、删除、显示等；
d） 使用算法；
e） 在特定事件发生时限制数据使用；
f） 使用数据后删除；
g） 结果导出方式，如API、报告、模型、禁止导出等。

6.3.4.1.5 使用通知

加工使用策略支持使用通知维度的管控，符合以下要求：
a） 记录数据使用信息，如数据提供方要求记录数据向数据消费端传输的记录；
b） 当数据被使用时通知数据提供方；
c） 当数据被使用时通知一方或特定用户组。

6.3.4.2 策略管理

数据流通跨域管控过程中，策略管理支撑数据提供方灵活地配置加工使用策略，符合以下要求：
a） 应提供加工使用策略模板；
b） 应支持用户自定义加工使用策略；
c） 应支持数据提供方对加工使用策略的增加、修改、撤回等操作；
d） 应支持数据提供方查询特定数据所关联的加工使用策略列表。

6.3.5 审计溯源

数据流通跨域管控过程中，审计溯源支撑数据提供方对数据流通的各个环节进行监控和审计，符合以下要求：
a） 应制定安全审计策略和审计日志管理操作规范，记录数据处理活动日志，为数据提供相关安全事件的处置、应急响应和事后调查提供证据支撑；
b） 应对数据采集汇聚、传输、存储、加工使用和结果分发等跨域管控全流程的数据处理、权限管理和人员操作等进行记录；

c） 记录内容应包括但不限于：时间、事件主体、事件客体、事件描述、事件结果等；
d） 日志留存时间应符合根据数据安全等级相关标准规定；
e） 应采用访问控制、备份，防篡改等技术手段对记录文件进行保护。

附 录 A
（资料性）
数据跨域流通安全风险

A.1 数据预处理风险

数据预处理阶段建立数据与数据提供方的对应关系，存在以下风险：

a)数据和数据提供方对应关系被篡改；

b)数据提供方身份被冒充；

c)数据标识符和数据对应关系被篡改；

d)数据提供方标识和数据提供方身份的对应关系被篡改；

e)数据来源不清晰，准确度低。

注：该风险会贯穿整个数据生命周期。

A.2 数据传输风险

数据传输阶段，存在以下风险：

a)不安全的网络环境导致数据被拦截、窃取或篡改；

b)传输协议被破解导致数据泄露；

c)参与方内部人员恶意拦截、窃取或篡改数据。

A.3 数据存储风险

数据存储阶段，存在以下风险：

a)存储在未加密的或不安全的系统中导致数据被未经授权的人访问；

b) 内部人员如运维人员，可能接触数据，导致数据权限被滥用或被恶意解密。

A.4 数据处理风险

数据处理阶段，存在以下风险：

a)数据提供方对数据进行授权，授权被伪造、篡改；

b)授权可能存在有效期限制，过期后继续使用数据将构成侵权行为；

c)数据处理之前绕过鉴权；

d)鉴权系统可能存在安全漏洞，被攻击者利用以获取未授权的数据访问权限。内部员工可能滥用权限，绕过鉴权机制，导致数据泄露；

e)数据计算过程中，存在计算过程被其他参与方窥探，即数据泄露风险，或未按照约定的算法或用途使用，导致数据被滥用。

A.5 数据交付风险

数据交付阶段，存在以下风险：

a)计算结果不按照提前约定的规则分发；
b)依据多轮计算结果进行差分攻击，反推出敏感信息。

附 录 B
（资料性）
跨域管控技术和方案

B.1 技术和方案列表

在数据流通的不同关键环节，可以采用表B.1所示的各种技术、技术组合和方案实现提供方数据跨域管控。

表B.1 数据流通跨域管控技术和方案

层面	关键环节	安全技术	技术说明
数据面	数据采集	数据源认证与授权	数据源认证：通过数字证书（如 X.509）或 OAuth 等机制验证数据源身份，确保数据来源可信。 授权机制：采用 RBAC 或 ABAC 控制数据采集范围，结合动态授权技术按需调整权限。
		数据脱敏与匿名化	实时脱敏：在采集阶段对敏感数据进行遮蔽或部分替换处理。 去标识化：移除或替换可直接识别个人身份的字段。 添加噪声：对数据添加噪声或模糊处理，保护隐私。如：差分隐私。
	数据传输	链路加密	链路加密是一种网络层或传输层的加密技术，它对数据在每一段传输链路上进行加密保护。数据在链路的起点加密，到达链路的终点解密，中间的每一段链路都需要重新加密和解密。如：SSL/TLS、IPSec。
		端到端加密	端到端加密是一种应用层加密技术，它在数据发送端进行加密，并在接收端解密，确保在整个传输过程中只有通信双方能够访问明文数据。如：PGP、Signal Protocol、SSH。
		网络隔离与安全路由	利用专用网络、虚拟专用网（VPN）或 SD-WAN 实现跨域数据流通安全隔离。
		传输完整性校验	利用摘要算法、数字签名等方式验证传输数据的完整性，防止篡改。 如：SM3、MD5、SHA256 算法
	数据存储	数据加密	对存储在磁盘、数据库或其他存储介质上的数据进行加密，防止未授权访问。 如： 使用对称加密算法（如 AES、SM4）对存储数据进行加密。 数据库加密（如 MySQL、PostgreSQL 的内置加密功能）。 硬盘加密（如 BitLocker、dm-crypt）。

		完整性校验	利用摘要算法(如：SM3、MD5、SHA256 算法)、数字签名(如：RSA、DSA、ECDSA、SM2)等方式验证传输数据的完整性，防止篡改。
	数据加工和使用	隐私保护计算	一种数据处理技术框架，旨在保护数据隐私，在不暴露原始数据的前提下进行计算或分析。包含技术范围：联邦学习、多方安全计算、可信执行环境、差分隐私等
		可信计算	一种更广泛的技术框架，旨在通过硬件和软件的协同，确保计算系统的完整性、安全性和可信度。
		数据空间	一个用于数据共享与交换的生态系统，提供了标准化的数据存储、交换、管理和隐私保护机制。
		沙箱环境	一种隔离的虚拟环境，用于数据的安全分析和处理，确保原始数据不泄露，分析结果可控。
		密态计算	通过综合利用密码学、可信硬件和系统安全的可信隐私保护计算技术，其计算过程实现数据可用不可见，计算结果能够保持密态化，以支持构建复杂组合计算，实现计算全链路保障，防止数据泄漏和滥用。
	结果分发	数据水印	动态水印：对每个分发的数据嵌入唯一标识符（如用户 ID、时间戳等），用于追踪泄露源。 隐形水印：嵌入数据文件的底层结构中，不影响数据使用，但能追踪和标识分发来源。 防篡改水印：结合水印和数字签名，防止分发的数据被篡改。
		数据防泄露	结合规则和 AI 分析,实时检测和阻止敏感数据通过不安全的方式分发。若检测到异常分发行为（如篡改、水印被移除），立即终止分发并使数据失效。
控制面	权限管理	细粒度权限控制	基于角色的访问控制（RBAC）：根据接收者角色分配权限。 基于属性的访问控制（ABAC）：根据用户属性（如部门、地区）和环境条件（如时间、IP 地址）动态分配权限。
		动态授权	实时评估分发请求，动态调整授权策略，确保分发操作符合实时安全策略。
		密态胶囊	通过将加密数据与使用规则绑定，确保数据在流通与使用全程受控。可信环境会对数据使用规则进行验证，并严格按照策略执行，实现对数据泄露与滥用的动态防控。
		零信任架构	在每次分发请求中重新验证用户身份和权限，避免过期授权导致的安

			全隐患。
	审计溯源	数据水印	在数据中嵌入隐形标识，用于标记数据来源或授权对象。
		区块链	使用区块链记录数据流通的全流程，确保数据来源可信、路径透明。
		操作日志记录	记录数据的每次访问、修改和传输行为，用于审计与溯源。

B.2 应用技术

B.2.1 隐私保护计算

隐私保护计算能够在保证数据提供方不泄露原始数据的前提下，对数据进行分析、处理和使用，其中：

a) 联邦学习是一种分布式机器学习技术或机器学习框架。在保证数据提供方在本地的前提下，与数据使用方协作联合建模；
b) 多方安全计算是基于密码学的算法协议，如同态加密，混淆电路，不经意传输和秘密分享等，实现在保护原始数据的情况下，多个参与方共同计算一个目标函数，且每一方仅获取自己的计算结果，无法从交互数据中推测出其他方的输入数据；
c) 可信执行环境通过软硬件方法在中央处理器中构建隔离的安全区域，从技术上实现了对其内部加载的程序和数据的机密性与完整性保护。将多个参与方的数据经安全信道汇聚到可信执行环境内进行融合计算；
d) 密态计算是指通过综合利用密码学、可信硬件和系统安全的可信隐私保护计算技术，其计算过程实现数据可用不可见，计算结果能够保持密态化，以支持构建复杂组合计算，实现计算全链路保障，防止数据泄漏和滥用。

B.2.2 可信计算

可信计算（Trusted Computing，TC）通常以TPM（TPCM）等硬件芯片为可信根，验证从BIOS到操作系统、应用软件的启动链条，确保启动的软件没有被恶意篡改，保障计算环境的安全性。TPM一般还提供远程验证功能，通过该功能远程客户端可以确认与其交互的计算机平台是否使用了安全软件，保证数据使用逻辑的完整性。

B.2.3 可信数据空间

可信数据空间以数字合约、使用控制技术为核心，以数据跨主体流通使用的可信（符合预期）为目标。通过数字合约技术描述特定参与方对数据内容、使用方式、使用次数等流通利用行为预期并达成共识；通过集成在特定软硬件环境中的使用控制技术对算法、应用进行控制和审计，实现数据访问、分析、计算等行为的管控，保证数据的流通利用过程符合预期。

B.2.4 数据沙箱

沙箱技术是一种用于隔离正在运行程序的安全机制，其目的是限制不可信进程或不可信代码运行时的访问权限。

a) 虚拟化技术：虚拟化技术是一种使计算资源如服务器、存储设备、网络资源等实现逻辑上的分割与抽象的方法。该技术允许单个物理设备上运行多个独立的虚拟实例，每个实例均拥有

独立的操作系统和应用程序环境。通过实现计算资源的逻辑隔离，虚拟化技术确保了各虚拟实例间的操作互不干扰，从而提供了物理级别的安全隔离；

b) 访问控制技术：访问控制技术是一套用于规范和管理用户及系统对资源访问权限的框架。它通过实施访问策略和权限设置，确保只有经过验证和授权的个体能够访问或操作指定的资源。访问控制技术通过限制对敏感数据和关键操作的访问，有效防止了未经授权的数据访问和篡改，确保了数据的完整性和机密性。同时，该技术还有助于防止恶意软件的横向移动和攻击行为，提升了系统的整体安全性；

c) 防躲避技术：防躲避技术是一系列旨在识别、监控和阻断恶意行为逃避安全检测的措施。这些技术能够检测到恶意软件试图隐藏其行为、修改系统状态或绕过安全控制的企图，并采取相应的拦截和响应措施。防躲避技术通过增强安全监控系统的能力，确保了对恶意行为的持续可见性和可控性。该技术有助于提高对新型和未知威胁的检测率，从而有效降低了安全威胁对系统的影响，保障了系统的安全性和稳定性。

B.2.5 密态传输技术

数据传输加密的目的是对传输中的数据流加密，通常有线路加密和端到端加密两种。线路加密侧重对保密信息通过线路采用加密密钥提供安全保护。端到端加密指加密信息由发送端自动加密，然后以不可读、不可识别的方式穿过网络，信息到达目的地后，被解密成为可读数据。

数据跨域传输需要在不同网络域之间安全、可靠地传输数据的过程。在这个过程中，确保数据的机密性、完整性和可用性是至关重要的。在技术方面，通常会采用一系列的安全通讯协议来保护数据在传输过程中不被未授权访问或篡改，如SSL/TLS,TLCP,VPN，如附录A所示。 以下是一些常用的安全通讯协议。

B.2.6 数据密态胶囊

数据密态胶囊将加密后的数据与其使用规则绑定，保证管控策略如约执行。数据提供方在对数据加密时，应确认该可信环境会对数据使用规则等进行验证，并严格按照策略执行。

B.2.7 区块链

区块链是一种不可篡改、安全可信的去中心化分布式账本，区块链上的数据按照区块结构存储，并通过链的方式顺序连接。

在区块链上发行数据权属凭证，可以固化每个数据资产的唯一标识、映射资源存储的地址、绑定权益信息。通过智能合约调用记录及变更资产与归属者的权属关系，其映射的实物资产或虚拟资产不需要在链上存储，仅作为映射资产的所有权数字凭证，后续用于各项数据资产的确权验证。

B.2.8 数据水印

数据水印是指从原始环境向目标环境进行敏感数据交换时，通过一定的方法向数据中植入水印标记，从而使数据具有可识别分发者、分发对象、分发时间、分发目的等因素，同时保留目标环境业务所需的数据特性或内容的数据处理过程。数据水印具有隐蔽性、可追溯性、确定性等特点。

附 录 C
（资料性）
数据跨域流程场景

C.1 媒体和品牌数据安全屋归因分析场景

媒体提供用户在媒体的访问数据，品牌提供会员下单数据，分别以密文的形式同步到中立平台方的沙箱中，由数据处理方提供数据分析代码，在沙箱中计算得到统计级的广告归因报告，指导品牌主广告策略优化，如图 C.1。该场景中主要使用数据跨域管控技术详见表C.1。

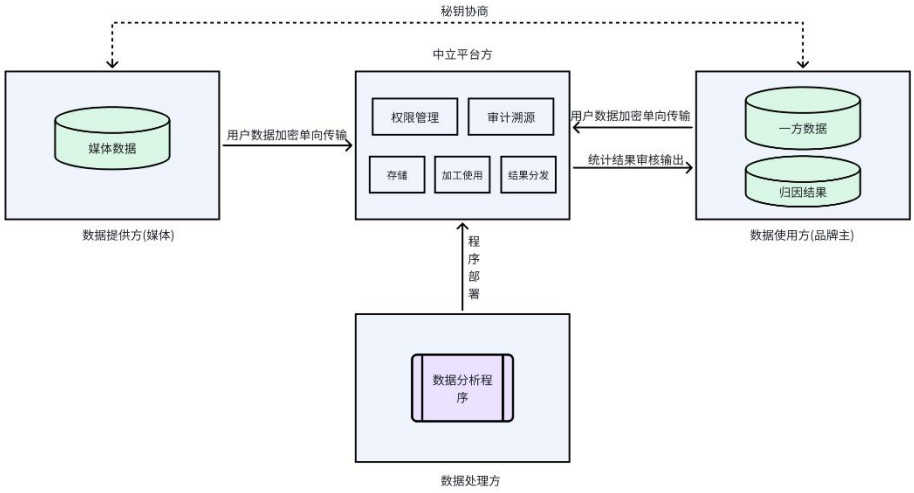


图 C.1 媒体和品牌数据归因场景下数据流通跨域管控示意图

表 C.1 应用的跨域管控技术

关键环节	安全技术	技术说明
数据采集	数据源认证与授权	通过数字证书和 IP 白名单等确保数据来源可信
	数据脱敏与匿名化	仅传输必要字段，且针对用户 ID 进行去标识化和加密处理
数据传输	链路加密	采用 SSL/TLS 等技术进行链路加密
	端到端加密	采用单独的加密算法实现端到端加密
	网络隔离与安全路由	利用 IP 白名单限制实现跨域数据流通安全隔离。
	传输完整性校验	利用摘要算法、数字签名等方式验证传输数据的完整性，防止篡改。
数据存储	数据加密	对存储在磁盘、数据库等存储介质上的数据进行加密，防止未授权访问。

	完整性校验	利用摘要算法、数字签名等方式验证传输数据的完整性，防止篡改。
数据处理	沙箱环境	由中立的第三方提供隔离的虚拟环境，用于数据的安全分析和处理，确保原始数据不泄露，分析结果可控。
结果分发	数据防泄露	实时检测和阻止敏感数据通过不安全的方式分发。若检测到异常分发行为，立即终止分发并使数据失效。

C.2 数据运营服务平台场景

数据运营服务平台可信数据流通方案是基于一系列技术手段，让数据资源在数据提供方、运营方和数据使用方之间以及安全计算环境下，实现无风险可信流通的解决方案。该场景数据流程详见图B.2。按照不同数据提供方划分，原始数据分为数据提供方1（内部数据源）和数据提供方2（外部数据源）两种情况。对于数据提供方1，依托自有数字底座/数据中台产品进行初始数据的处理加工；对于数据提供方2，则依托第三方数据中台进行初始数据的加工和准备。经过合规处理的初加工原始数据，根据数据密级的不同，分为低密数据和高密数据两种处理流程。低密数据可直接发布到数据运营服务平台进行数据产品上架和数据目录展示；而高密数据则通过数据空间连接器送入可信计算中台的鲲信一体机进行计算处理。用户在数据运营平台购买数据产品并获得使用权限后，可通过平台下发元数据、SQL模型等信息和指令，可信计算中台在接收到各方提供的初加工原始数据后，加密推送至目标计算节点，然后加载到通用计算引擎，执行SQL模型，完成安全计算。最终，授权的数据使用方以API或文件形式获取密态计算结果，从而实现“原始数据不出域，数据可用不可见”的数据流通范式。主要使用数据跨域管控技术详见表B.2。

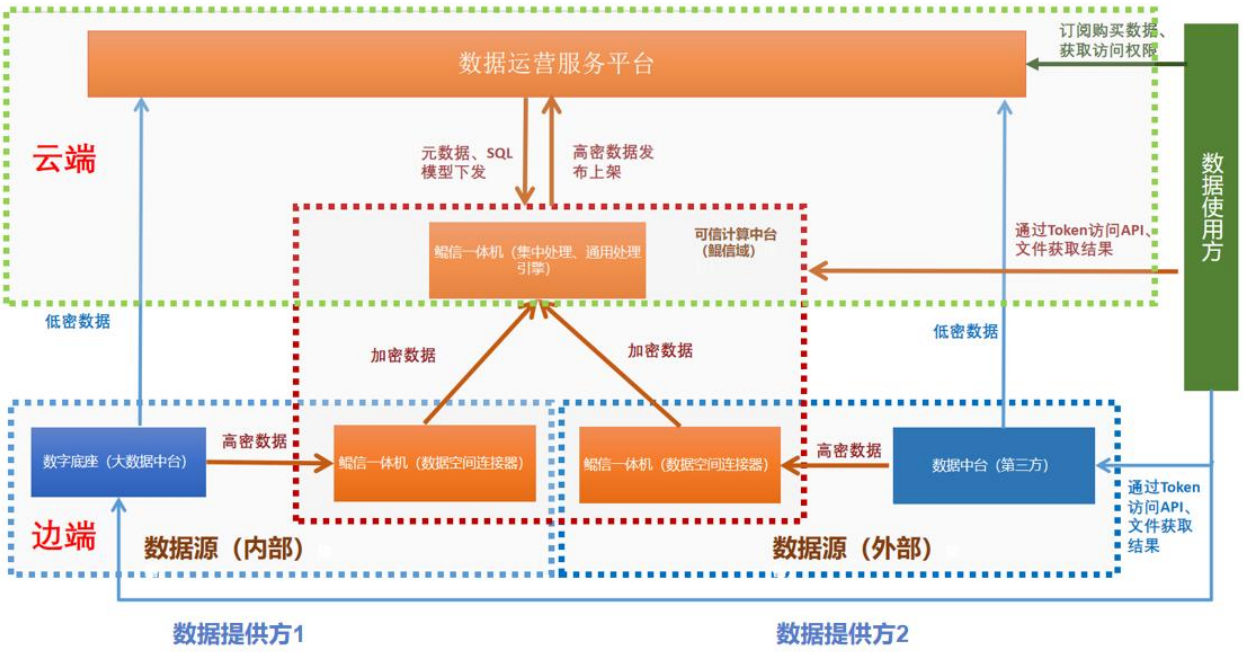


图 C.2 数据运营服务平台数据流通图

表 C.2 数据运营服务平台数据跨域管控技术

层面	关键环节	安全技术	技术说明
	数据传输	链路加密	链路加密是一种网络层或传输层的加密技术，它对数据在每一段传输链路上进行加密保护。数据在链路的起点加密，到达链路的终点解密，中间的每一段链路都需要重新加密和解密。如：SSL/TLS、IPSec。
		端到端加密	端到端加密是一种应用层加密技术，它在数据发送端进行加密，并在接收端解密，确保在整个传输过程中只有通信双方能够访问明文数据。如：PGP、Signal Protocol、SSH。
	数据加工和使用	隐私保护计算	机密计算是一种基于硬件的技术，它将数据、特定功能、应用程序，同操作系统、系统管理程序或虚拟机管理器以及其他特定进程隔离开来，让数据存储在可信执行环境中，即使是使用调试器，也无法从外部查看数据或者执行操作。TEE 确保只有经过授权的代码才能访问数据，如果代码被篡改，TEE 将阻止其继续进行操作。
		可信计算	一种更广泛的技术框架，旨在通过硬件和软件的协同，确保计算系统的完整性、安全性和可信度。
		数据空间	一个用于数据共享与交换的生态系统，提供了标准化的数据存储、交换、管理和隐私保护机制。
		沙箱环境	一种隔离的虚拟环境，用于数据的安全分析和处理，确保原始数据不泄露，分析结果可控。
		区块链	使用区块链记录数据流通的全流程，确保数据来源可信、路径透明。
		操作日志记录	记录数据的每次访问、修改和传输行为，用于审计与溯源。

C.3 公积金信用贷场景

C.3.1 被授权的数据运营单位汇聚公积金信用贷场景需要的相关原始数据，并对数据进行融合、治理，形成公积金信用主题数据库，在此基础上，根据具体数据需求，对主题数据进行加工处理，形成数据产品，供授权的银行进行数据查询或模型训练，在保证数据安全的前提下，满足银行对公积金信用数据的查询或模型训练需求，如图C.3所示。

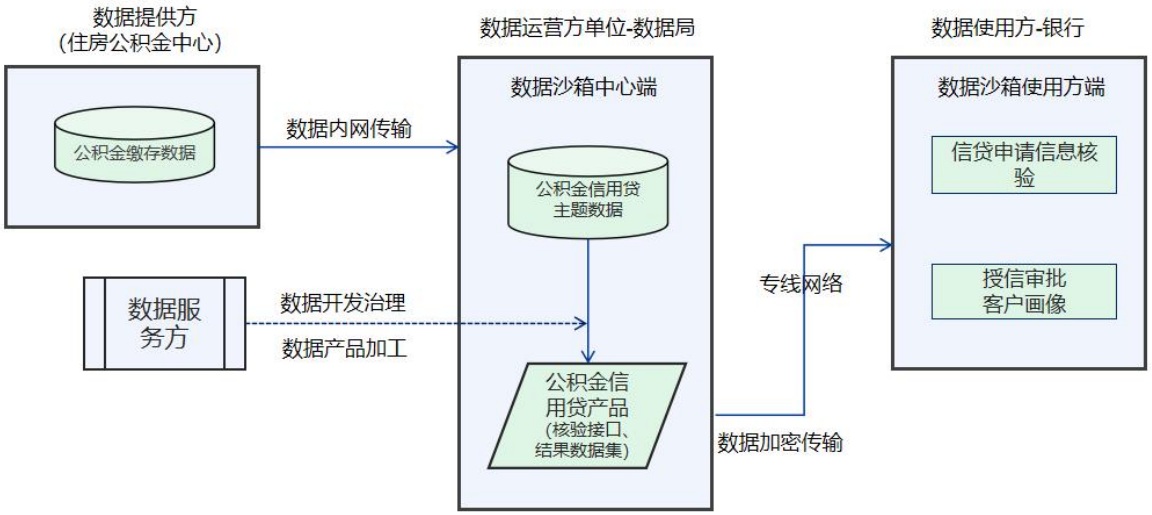


图 C.3 公积金信用贷数据流程图

C.3.2 参与方与活动详情为：

a) 数据提供方为公积金中心，主要基于合约对公积金缴存的原始数据的运营进行授权；

b) 数据运营方可以为授权的运营单位，如数据局等主体，也可以为公积金中心本身，主要负责制定并执行数据运营规则与管理规范，促进参与各方完成数据产品的加工和交付使用；

c) 数据服务方为委托的数据服务机构，主要对数据进行开发治理，并根据要求对数据进行加工，形成数据产品或数据服务；

d) 数据使用方为银行机构，主要通过核验查询，画像构建等方式使用公积金信用贷产品服务，从而完善自身的信贷申请信息核验、授信审批等业务。

C.3.3 实施效果：有效提升了银行的个人画像、企业画像、信用指数、信息核验等数据能力，帮助银行完善授信审批模型，精准评估市民和中小微企业的信用状况，降低市民信贷成本，提升中小微企业贷款可得性。并在数据流通使用过程中，有效防止了数据泄露和数据的滥用。

C.3.4 采用的数据跨域管控技术：

a) 数据沙箱技术，用于保障数据的开发安全，避免数据加工和使用环节的敏感数据泄露；

b) 数字签名技术，确保数据的来源可信和数据的完整性；

c) 数据泛化技术，对公积金敏感的金额数据进行泛化处理，避免敏感数据泄露的同时满足训练需求，实现数据的可算不可识；

d) 使用控制技术，与数据沙箱结合使用，有效控制数据使用方的使用范围和用量。

参 考 文 献

[1] GB/T 20945—2023信息安全技术 网络安全审计产品技术规范
[2] GB/T 25069—2020信息安全技术 术语
[3] GB/T 31167—2023信息安全技术 云计算服务安全指南
[4] GB/T 35273—2017信息安全技术 个人信息安全规范
[5] GB/T 36343—2018信息技术 数据交易平台 交易数据描述
[6] GB/T 37932—2019信息安全技术 数据交易服务安全要求
[7] GB/T 40651-2021信息安全技术 实体鉴别保障框架
[8] GB/T 42573—2023信息安全技术 网络身份服务安全技术要求
[9] JR/T 0223—2021金融数据安全数据生命周期安全规范