|  |  |
| --- | --- |
| 产品名称Product name | 密级Confidentiality level |
|  | 秘密 |
| 产品版本Product version |  |
|  |

终端云数据中心管理规范

|  |  |  |  |
| --- | --- | --- | --- |
| Prepared by  拟制 |  | Date  日期 |  |
| Reviewed by  评审人 |  | Date  日期 |  |
| Approved by  批准 |  | Date  日期 |  |
| Authorized by  签发 |  | Date  日期 |  |



Huawei Technologies Co., Ltd.

华为技术有限公司

All rights reserved

版权所有 侵权必究

（REP01T02 V1.21/ IPD-CMM V3.0 / for internal use only）

（REP01T02 V1.21/ IPD-CMM V3.0 / 仅供内部使用）

**目** **录**

第1章 概述 3

第2章 数据中心管理规范 4

2.1 数据采集规范 4

2.1.1 规范 4

2.2 数据处理规范 4

2.2.1 规范 5

2.3 数据应用规范 5

2.3.1 规范 5

2.4 管理和协作规范 5

2.4.1 规范 5

# 概述

终端云数据中心上下文如下图所示：（备注：这里要显示4个数据源，一个数据接口，一个批量通道）

1. 终端云数据中心上下文

从图可见，终端云数据中心数据来源包括四部分：

* 客户端用户行为信息。用户行为信息通过应用内嵌人的日志收集SDK或网页内嵌入的JS收集，客户端收集的日志信息通过HTTP/HTTPS等协议即时上报或周期上报。
* 服务端日志：数据中心通过FTP/SFTP等定期从生产系统（服务端）获取日志信息。
* 业务DB数据
* 第3方数据

收集到原始数据后，数据分析开发人员利用各种工具在数据中心完成数据处理，处理结果通过提供三种应用方式：

* Web界面查询。数据中心用户通过Web界面查看数据处理结果。例如，查看报表，OLAP分析、即席查询，数据产品等。
* 二次数据开发。业务运营人员基于数据中心提供的各种工具开发各种产品。
* 数据通道开放。数据中心以安全的数据通道向其它数据应用业务开放数据。提供两种开放方式：数据批量传输和数据接口。

另外，数据中心本身还提供系统管理功能，系统管理员利用此功能完成系统管理和维护。

数据中心管理规范

数据的处理过程大致可分为三个阶段：数据采集、数据处理、数据应用，下面分别描述在各个阶段的数据中心管理规范。

数据采集规范

数据采集阶段对应将服务端日志信息、业务DB数据、客户端用户行为信息、第三方数据采集到数据中心的过程。

### 规范

1. 业务系统中的用户及内容相关信息数据在备份到到数据中心前不得删除，这些数据都必须接入到数据中心进行存储，分析和应用。数据中心和业务方共同协商数据接入计划并付诸实施。
2. 鼓励业务系统预留数据收集接口（数据采集接口包括需要采集的内容、内容格式、存储方式，传输方式但不限于这些），业务系统的采集进口要符合由数据中心和业务双方预先约定接口规范。
3. 数据采集接口变更应由双发协商并确认后才能实施，业务方负责业务数据采集，数据中心负责传输通道、工具和接收处理。特殊情况下，数据采集过程影响生产系统时，应允许生产系统暂时关闭数据采集功能。
4. 数据中心负责数据整个采集体系的建设和运维，包括采集客户端通道的建设和维护，确保采集体系的顺畅运行。
5. 数据采集过程应尽可能降低对生产系统的性能影响，集中传输在业务量低谷时间段进行。采集系统要异步进行，采集系统的变更，事故要尽量不影响业务系统的正常运行。
6. 数据中心负责采集到的业务相关数据的安全性和可用性。
7. 用户敏感信息（如姓名、银行卡号、手机号码等）以加密方式传输和存储在数据中心中。

数据处理规范

数据处理是数据中心的核心功能，数据处理主要由数据分析开发人员完成。

### 规范

1. 禁止各业务自建数据仓库平台，数据中心有责任提供平台能力、分析能力给各业务方，并负责数据平台的运维和数据安全生产。
2. 鼓励各业务方经营/运营决策及其他相关数据处理诉求统一提交给数据中心，以便数据中心提供更高效的平台和数据产品（如报表、推荐引擎、营销引擎等），支持业务的数据化运作。
3. 数据中心负责建立高效的，稳定的，结构化核心数据模型，并提供和维护核心模型的说明，以便BI团队和业务分析师高效的对数据进行分析。
4. 原则上各业务只能接触到个业务自己的数据（包括原原始数据和生成数据），要跨业务处理数据需要目标数据所属业务方的授权，数据中心可以Access到所有数据。
5. 数据中心对业务方数据处理提供技术支持（如二次开发工具等）。
6. 业务方的数据处理用到数据中心的人力、软件、计算和存储资源会计入业务方的成本之中，供业务和数据中心进行成本核算。
7. 数据中心的主要数据处理职责还包括：负责对跨业务的数据需求进行分析、对共有模式化分析提供分析模板或是产品、整理全业务的用户profile、研发公共的数据产品平台、提供数据开放接口支持开放生态建设等。

数据应用规范

数据中心的数据提供三种应用方式：结果浏览、二次开发和数据通道开放。

### 规范

1. 数据中心数据除了3种数据应用方式，原则上是与外界隔离的，确保数据安全。
2. 数据应用的权限根据应用所基于的数据来划分。数据中心提供的数据结果浏览，数据接口都支持用户鉴权，批量数据走专有数据通道，需要单独申请权限，数据中心记录3个月的审计信息，保证数据应用的安全。
3. 数据通道直接加密，保证传输的安全性。

管理和协作规范

系统管理是数据中心正常运行的基础，主要由数据中心维护人员操作执行。协作是指业务方和数据中心如何合作确保数据中心和业务双方的高效合作。

### 规范

1. 数据中心的原始数据不得删除，以确保信息的不丢失和原始性。
2. 数据中心要确保数据安全，提供数据的定期备份，及容灾能力
3. 数据中心负责管理维护数据平台，保证数据平台的不间断服务及服务水平，并保证平台的安全性。
4. 数据中心与个业务方之间应建立接口人制度，双方接口人对各自团队及彼此负责，职责是明确需求，负责承诺，协调双方人员进行高效合作。