**BISDK 访问用户**

1. **数据字典**

**1.1表级**

英文名：dwd\_evt\_bisdk\_visit\_dm

中文名：BISD访问用户

表类型：快照表

ETL调度频率：日表

**1.2字段级**

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段中文名 | 备注 |
| terminal\_os | STRING | 终端操作系统 |  |
| device\_name | STRING | 设备型号 |  |
| rom\_ver | STRING | ROM版本 |  |
| app\_ver | STRING | 应用版本 |  |
| imei\_md5 | STRING | IMEI\_MD5 |  |
| package\_name | STRING | 包名 |  |
| dev\_up\_id | STRING | 开发者编号 |  |
| app\_spread\_channel | STRING | 应用推广渠道 |  |
| bisdk\_net\_type | STRING | BISDK网络类型 |  |
| carrier | STRING | 运营商 |  |
| user\_ip\_addr | STRING | 用户IP地址 |  |
| client\_resolution | STRING | 客户端分辨率 |  |
| first\_access\_page\_time | STRING | 首次访问页面时间 |  |
| final\_access\_page\_time | STRING | 最后访问页面时间 |  |
| quit\_access\_page\_name | STRING | 退出访问页面名称 |  |
| entering\_access\_page\_name | STRING | 进入访问页面名称 |  |
| access\_page\_total\_cnt | INT | 访问页面总数 |  |
| access\_total\_duration | INT | 访问总时长(秒) |  |
| access\_id | STRING | 访问编号 |  |
| session\_id | STRING | 会话编号 |  |
| access\_app\_src | STRING | 访问应用来源 |  |
| imei | STRING | 设备编号 |  |
| server\_time | STRING | 服务器时间 |  |
| etl\_time | VARCHAR(30) | ETL时间 |  |

1. **数据来源**

**2.1．来源于数据打点源表:**

数据打点源表 : ODS\_BISDK\_DATA\_HM (打点源表)

**2.2．其他依赖的表:**

dwd\_ref\_cloudservice\_package\_par\_ds (包名参考表):存储包名和业务分区的对应关系

1. **数据逻辑**

**3.1．主体逻辑**

**App打点上报的数据 通过采集服务器收集，收集之后的数据推送到ODS,并通过jar包解析成3张小时表，此表是汇总小时表一天的数据**

**3.2．字段逻辑**

| 字段名 | 处理逻辑 |
| --- | --- |
| terminal\_os | 直抽，手机操作系统 |
| device\_name | 直抽，手机机型。列如：HUAWEI TAG-AL00、G621-TL00 |
| rom\_ver | 直抽，手机固件版本。 |
| app\_ver | 直抽，app版本号 |
| imei\_md5 | 直抽，设备id，采用md5加密 |
| package\_name | 直抽，app包名 |
| dev\_up\_id | 直抽，开发者id |
| app\_spread\_channel | 直抽，推广渠道，例如华为应用市场、小米应用市场的渠道号等 |
| bisdk\_net\_type | 直抽，网络系统名称 |
| carrier | 直抽，网络提供商。例如：中国电信、中国移动、中国移动3G |
| user\_ip\_addr | 直抽，IP地址 |
| client\_resolution | 直抽，屏幕分辨率 |
| first\_access\_page\_time | 直抽，首次访问时间，格式：yyyy-MM-dd HH:mm:ss,精确到秒，页面最后访问时间 |
| final\_access\_page\_time | 直抽，最后一次访问时间，格式：yyyy-MM-dd HH:mm:ss,精确到秒 |
| quit\_access\_page\_name | 直抽，最后退出的页面名称。目前大部分数据为空 |
| entering\_access\_page\_name | 直抽，首次进入的页面名称。目前大部分数据为空 |
| access\_page\_total\_cnt | 直抽，共访问了几个页面，整数，访问页面总数 |
| access\_total\_duration | 直抽，这个消息包，总共访问的页面时长。这个消息包，总共访问的页面时长 |
| access\_id | 直抽，上报数据包的消息ID |
| session\_id | 直抽，会话ID |
| access\_app\_src | 直抽，来源ID |
| imei | 设备id,做了格式化。没有做非法imei。 |
| server\_time | 采集服务器接收到数据上报的时间，用dateutil函数做了时间标准化 |
| etl\_time | ETL时间字段。记录数据入库的服务器时间。 |

**3.3．ETL逻辑**

基本是直抽的逻辑。只有个别字段经行数据标准化的处理。

1. **其他**

**4.1． imei处理过程**

4.1.1标准化：revertDeviceId(deviceIdFormat(imei))