# 取数系统设计

## 角色设计

待定

## 模型设计

### 表模型设计

### 后台执行模型设计

## 功能模块设计

## 界面设计

### 数据源管理

#### 数据源添加

数据源列表

标签聚合分类

### 用户群管理

用户群新增

包含要素

用户群列表

用户群集合

### 自助取数

创建取数

任务列表

下载管理

### 任务管理-调度管理

取数任务

报表任务

ETL任务

### 文档管理

帮助文档

技术分享

优秀站点

### 系统管理

权限管理

页面管理

后台管理

## 技术实现

前端:vue+element-ui

后台:springboot+mybatis

调度:xxl-job

## 详细设计

### 需求描述

### 业务流程设计

#### 2.1 取数任务状态和操作设计

状态设计:

1. 任务创建-未提交是属于草稿状态

2. 任务已提交-未执行是属于已提交状态

3. 任务执行中-为完成是属于执行中状态

4. 任务执行完成是属于已完成状态

5. 任务在执行中或已提交状态时,点击取消,任务状态: 执行中->草稿,已提交->草稿状态

6. 任务执行过程中出现异常,任务状态: 执行中->异常状态

7. 后期添加下载审批权限时增加几个状态 审批中 未通过 可下载 需要审批的任务状态变化:已完成->审批中->未通过/可下载 ,不需要审批的任务状态变化:已完成->可下载

  { key: '0', display\_name: '草稿' },

  { key: '1', display\_name: '已提交' },

  { key: '2', display\_name: '执行中' },

  { key: '3', display\_name: '已完成' },

  { key: '4', display\_name: '异常' },

  { key: '10', display\_name: '审批中' },

  { key: '11', display\_name: '未通过' },

  { key: '20', display\_name: '可下载' }

操作设计:

创建任务

保存 – 状态变为草稿

编辑 – 草稿/异常状态允许的操作

提交 – 草稿状态允许的操作

删除 – 草稿/已完成/异常状态允许的操作

取消 – 已提交/执行中允许的操作

复制 – 任何状态都可以的操作,复制之后的新任务是草稿状态

审批 – 单独页面控制,-待定

#### 2.2 创建取数任务业务流程设计

直接进入创建取数任务界面的入口:

1. 任务列表中创建按钮
2. 创建取数菜单
3. 查看取数任务的时候的复制按钮

隐式进入

1. 查看取数任务
2. 编辑取数任务
3. 审核取数任务

页面元素:

1. 取数任务名称
2. 说明

创建取数任务进入时:

保存时保存到五个流程id上

1.显示的时候根据流程ID再次查询 --- 流程太复杂

2.可以直接返回一个json --- 解析复杂

先使用1. 后期改为2,可以导入导出取数任务配置

数据源

*//flowValue1 选择的数据源编码  
//flowValue2 选择的数据源名称  
//flowValue3 选择的数据源中文名称  
//flowValue4 选择的数据源别名*

用户群

*//flowValue1 选择的用户群  
//选择的用户群交集并集方式  
//flowValue3 排除的用户群  
//flowValue4 排除的用户群交集并集方式*

*flowValue5 前台构造和解析使用的json串*

条件

*//flowValue1 选择的条件  
//flowValue2 排除的条件  
//flowValue3 前台构造和解析使用的json串*

*flowValue5 前台构造和解析使用的json串*

*一个条件json串包含的元素:*

*首先是单个条件包含的元素:*

1. *关系*
2. *前括号*
3. *指标信息和作用指标上的处理(函数,自定义)*
4. 运算符
5. 条件值信息*和作用*条件值*上的处理(函数,自定义)*
6. 后括号
7. 排序

标签

*//flowValue1 选择的标签编码  
//flowValue2 选择的标签名称  
//flowValue3 选择的标签中文名称  
//flowValue4 选择的标签配置*

执行计划

*//flowValue1 执行引擎  
//flowValue2 执行周期类型  
//flowValue3 选定日期  
//flowValue4 执行sql*

结果

*//flowValue1 文件格式  
//flowValue2 文件分隔符  
//flowValue3 是否生成客户群  
//flowValue4 结果保存方式*

6个流程json都保存在后台了,简单处理下就可以导出为一个大的json文件.方便配置导入导出,复制等.

#### 2.3 取数sql生成

分为两类

1. 开发模式:直接使用前台编写的sql.谨慎使用
2. 正常模式:
3. 数据源问题, 使用指定关联字段和默认关联字段
4. 用户群问题, 用户群只能使用默认关联字段

生成sql情况:

1. 只有单个数据源

使用默认主键

1. 有多个数据源

其中一个默认主键,其他使用自定义主键

数据源:

单个

多个

用户群的主键不能修改,只能和默认主键的数据源关联

用户群:

单个

多个

筛选 user\_id in

排除 user\_id not in

交集 a left join b on a.id = b id where a.id != null and b.id != null

并集 a full join b on a.id = b.id

筛选条件:

and

or

括号

展示指标:

转义 --原有方案,生成带转义字段的取数表

不转义

用户群指标

展示高级功能暂时不做

Sql格式:

--交集 a left join b on a.id = b id where a.id != null and b.id != null

--并集 a full join b on a.id = b.id

Null 不参加运算

Select 数据源展示指标,用户群展示指标,自定义指标

From

数据源主表 a --暂时默认第一个是主表 .无数据源的情况呢

Left Join 数据源表 b on a.key=b.key

Left Join 数据源表 c on a.key=c.key

Left join 筛选用户群集合d on a.key = d.key –直接拼sql不太好处理,组成集合,条件也不好处理====改为关联多次,并集= (a is not null or b. is nut null)

Left join 排除用户群集合 f on a.key=f.key – 可以直接拼sql

Where 1=1

And a.condition = value1 –筛选条件

And b.condition = value2

And c.condition = value3

And d.condition = valur4

And not (a.condition=value5) –排除条件

And d.key is not null

And f.key is null

Sql优化:

1. 用户群必须是静态用户,取数任务生成的用户群使用最新的结果数据
2. 排除用户群数量小/筛选用户群数量无标签时的情况,直接转化为条件语句,减少关联次数

#### 2.4 用户群生成

### 七 . 功能设计

#### 3.1 数据源选择模块

3.1.1 业务功能描述

用户可以选择数据源

支持选择多个数据源,后台配置用于关联的主键

数据源相当于按不同业务划分的用户信息表,也可以是包含用户所有标签的大宽表

选择的数据源会用于后面的标签选择

3.1.2 逻辑模型

|  |  |  |  |
| --- | --- | --- | --- |
| **功能名称：** | 1.数据源表 | 系统使用的数据源/标签源 | |
| **目标表** | | **数据源** | |
| tc\_gd\_datasource | |  | |
| **字段说明** | **字段名** | **数据类型** | |
| 数据源编码 | source\_id | varchar(10) |  |
| 数据源名称 | source\_name | varchar(50) |  |
| 数据源中文名称 | source\_name\_zh | varchar(200) |  |
| 分类 | source\_type | varchar(50) |  |
| 主键 | source\_key | varchar(50) |  |
| 别名 | source\_alias | varchar(10) |  |
| 有效状态 | status | varchar(2) |  |
| 备注 | remark | varchar(200) |  |
| 创建人 | create\_persion | varchar(200) |  |
| 创建时间 | create\_date | varchar(20) |  |

3.1.3 表汇总

1.数据源表

2.配置流程表

3.配置结果表

3.1.4 功能开发说明

1.首先从从数据源表中查询允许选择的数据源 /search

2.选择数据源后可以

3.1.5 角色及菜单路径

#### 3.2 用户群选择模块

3.2.1 业务功能描述

一定程度上类似数据源,但是又不完全相同,比数据源更加灵活

数据源一定是全量用户数据,只是标签的维度不同

用户群是全量用户的子集

业务人员可以选择使用单个/多个用户群基础上做筛选

也可以选择使用单个/多个用户群基础上再排除单个或多个用户群,再进行筛选

选择/排除的多个用户群可以做交集或并集处理

用户群来源包括

1.用户上传

2.取数生成

3.接口生成

用户上传

可上传只有用户标识的数据

也可以上传带标签的数据,上传处理为后台异步处理,上传时先预览部分数据,配置标签,支持csv/txt/xlsx文件,第一行为表头,可选择分隔符

可以提前生成好也可以临时上传

3.2.2 逻辑模型

3.2.3 表汇总

3.2.4 功能开发说明

1.用户群上传

3.2.5 角色及菜单路径

#### 3.3 配置筛选条件模块

3.2.1 业务功能描述

条件就是之前数据源和如果用户导入用户群中包含11的标签,再加上一定的判断条件形成的一个一个的单个条件,条件之间可以有一定的逻辑关系

基础筛选条件功能包括

1. 根据不同数据类型输入,如标签选择类,数值类,日期类
2. 可以添加and , or ,非 ,括号等逻辑判断
3. 可以有筛选条件,也可以有排除条件

高级筛选条件功能包括

1. 添加作用于标签上的函数
2. 添加标签之间有关联的函数
3. 标签分类
4. 抽样,随机取数

#### 3.4 选择展示标签模块

作用于每个标签级别的配置

先选择需要展示的标签,包括

数据源中的标签和用户群中导入的标签

可对每个标签进行函数计算,脱敏,单位转换等

可以添加自定义标签,对生成的用户打上特定标签

#### 3.5 配置执行计划模块

执行引擎可配置

可立即执行

也可以制定周期性执行,按天,按周,按月执行

#### 3.6选择结果保存方式模块

配置保存的文件类型

使用分隔符

是否生成用户群

结果保存方式:直接下载,指定ftp等

#### 3.7用户群创建/修改

用户群状态

用户群有不同的来源

处理状态有 未处理,处理中,已处理,处理失败

生效状态有 未生效,生效中,已失效

未处理,处理中, 处理失败, 未生效,生效中,已失效

1,2,3,4,5,6

两位控制法:

第一位表示处理状态

第二位表示生效状态

客户群的失效状态,定期更新

用户群存储方式

分表

1. 无用户标签的统一保存
2. 有用户标签的分表保存

是否校验

以打上标签的方式

离线数据和实时数据的数据偏差

1. 文件上传

固定数据源,最简单,优先实现

分为实时处理

2M或者一定条数以内的

和离线处理

大文件,支持预览部分数据

2. 自助取数生成

关联自助取数,涉及到周期任务,待详细设计

1. 创建取数任务的时候保存数据,状态为未生效

保存的数据有:

客户群基本信息

客户群标签信息

用户群标签问题:

自助取数生成的用户群标签----ck重复字段名称问题

1. 从数据源中取出的标签

名称保存为T+取数任务ID+源标签名称

1. 从用户群中取出的标签

名称保存为T+取数任务ID+标签序号, 如果用源标签名称的话,一个用户群多次生成,标签名称可能会超长

1. 后台执行时更新状态
2. 文件接口

涉及日月周期,下发时间等问题,待详细设计

1. 绑定数据源

数据更新等问题,待详细设计

…

#### 3.8用户群列表

查询

展示

导出

生失效控制

使用权限控制

修改删除操作入口

#### 其他

配置可以复制使用

保存历史数据,可观察历史数据变化

结果预览和抽样

文件加密

## 数据处理和接口

文件上传:springboot接口

文件下载:nginx代理