360数据质量操作指南

## 首页

系统登录后，会默认进入首页，若从其他菜单页面进入首页，点击菜单左侧菜单首页即可



左边红框数据为随周期和数据时间变化的区域，展示当前周期数据时间内的数据质量整体情况，包括稽核

情况，问题情况，稽核历史趋势，告警处理历史趋势，每个数字上面，均提供链接入口查看明细信息。

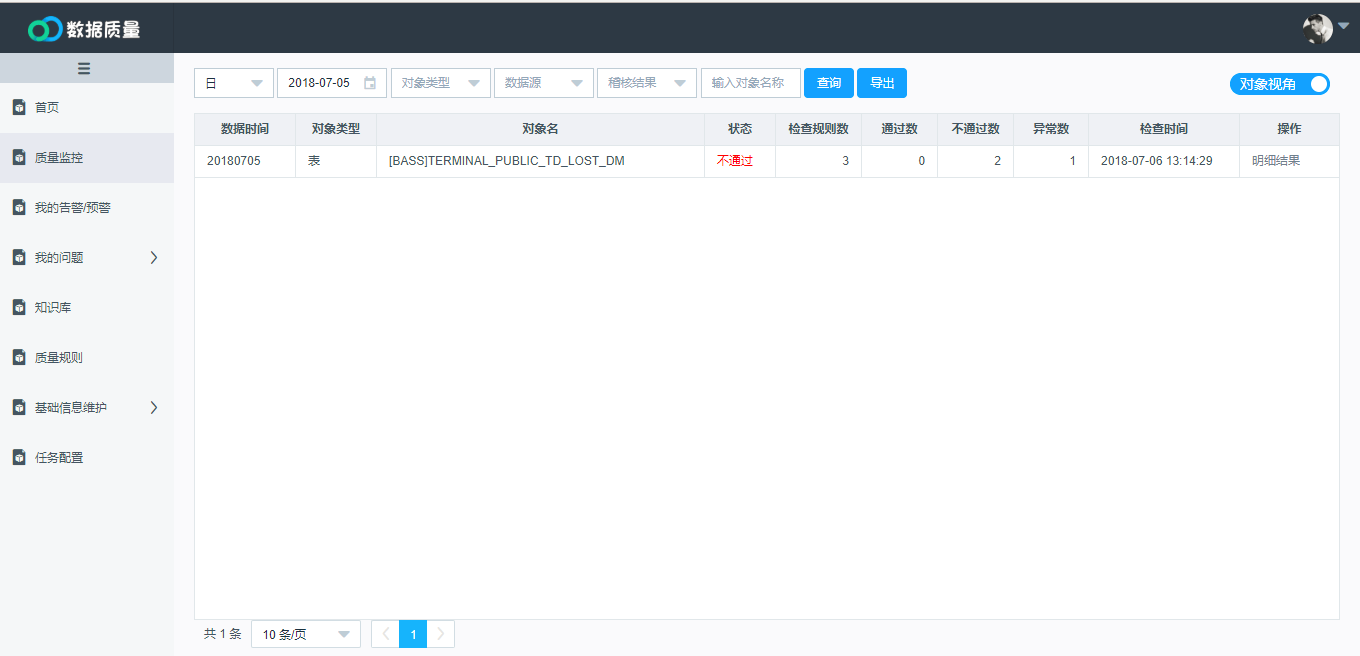
趋势图和表格支持联动

右边红框数据，展示全部的告警，问题，最严重告警，最严重预警，最新告警，最新预警等信息，支持

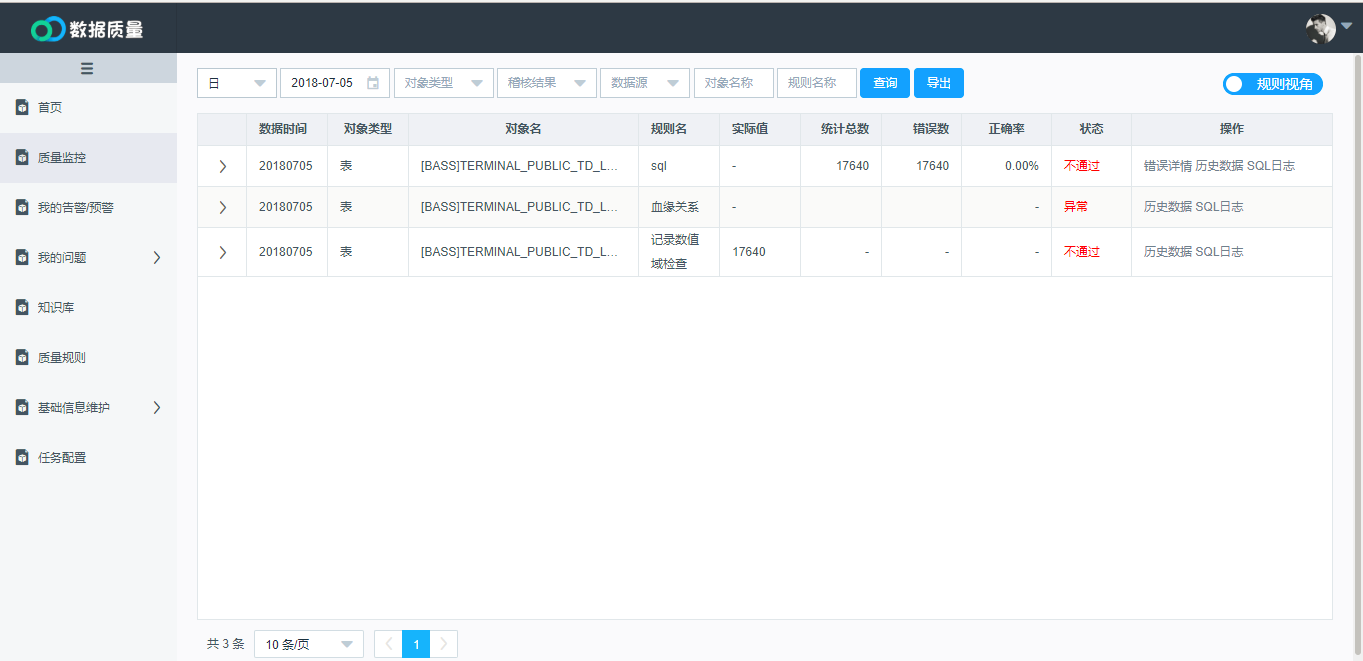
链接明细展示和更多展示

## 质量监控

点击质量监控，能够查看数据质量监控报告。搜索框输入对应的查询条件，点击查询按钮，能够查询出对应条件的质量报告。根据查询条件展示出报告列表，点击导出按钮，能够成功导出对应查询条件下的报告列表。

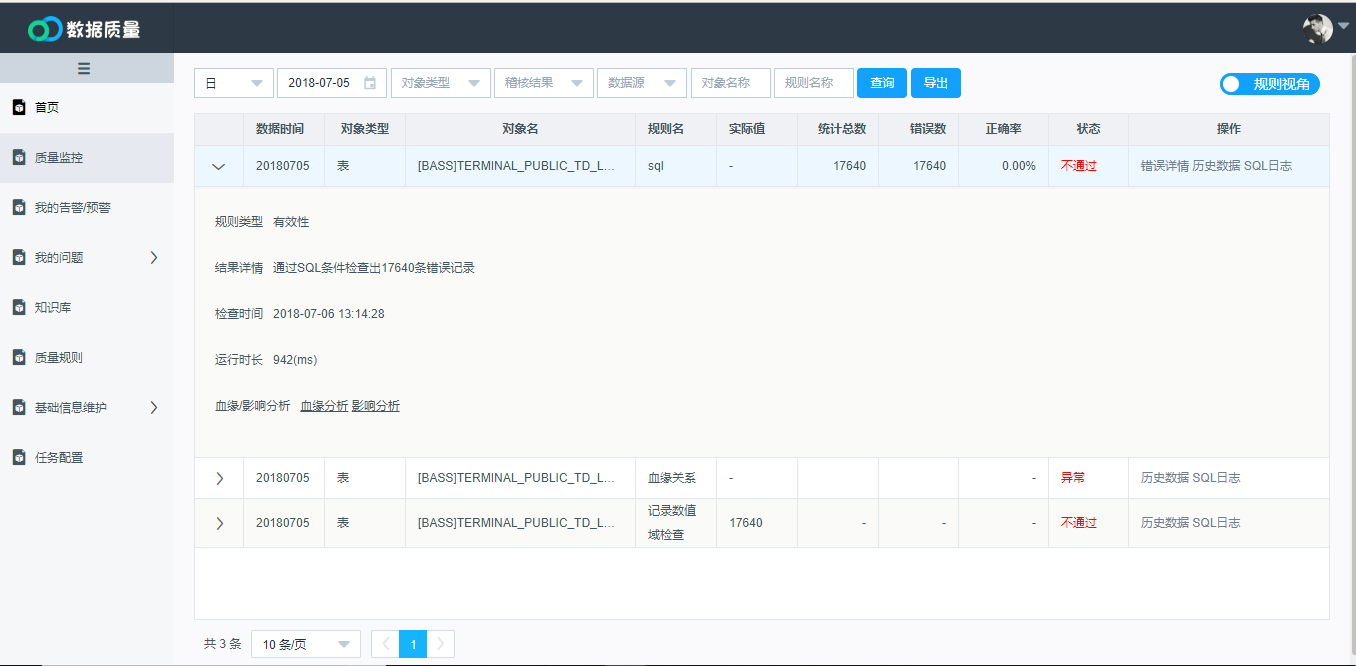


切换右上角的对象视角和规则视角，可以由不同视角查看报告。

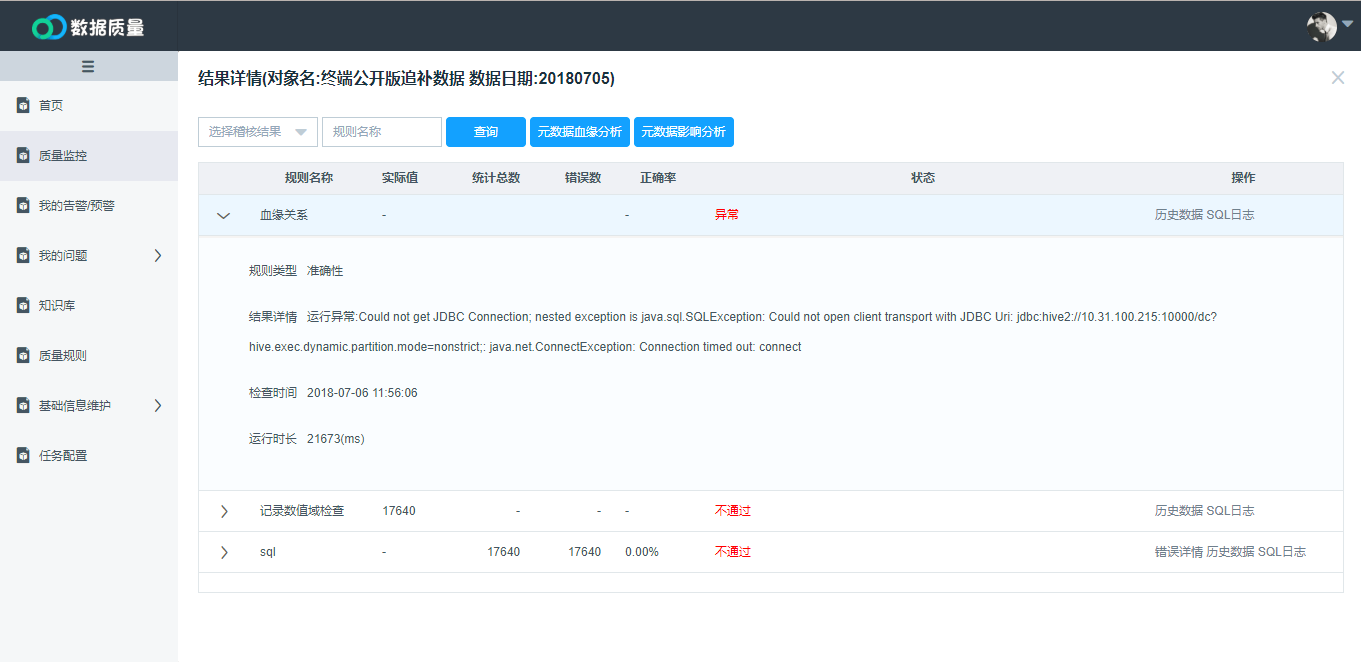


在按照规则视角展示的情况，点击左侧折叠按钮展开，可以查看该条规则的质量详情。并支持

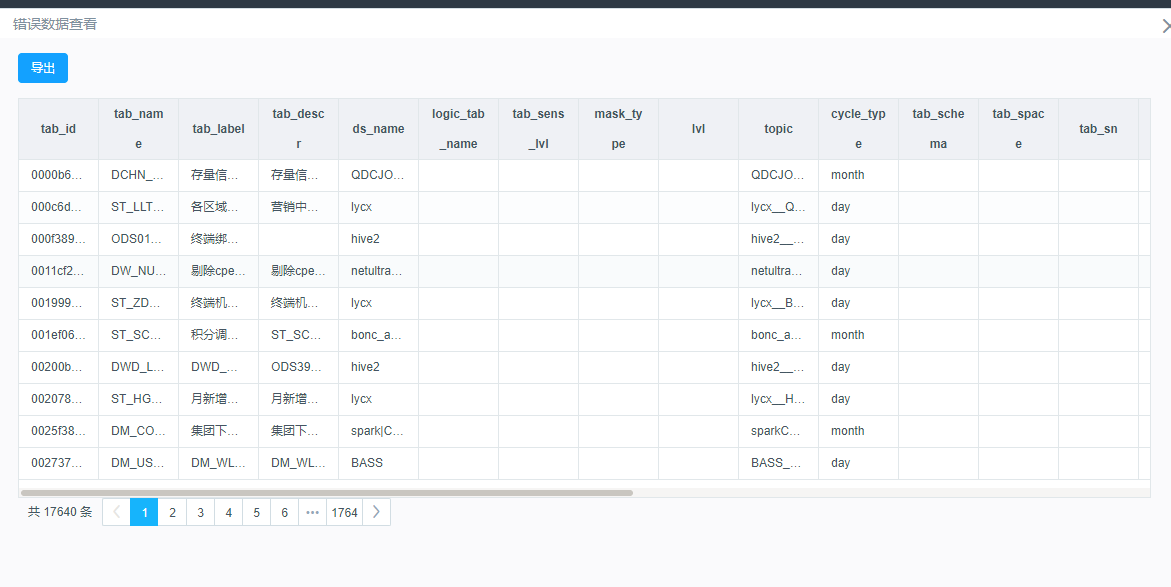
血缘分析和影响分析的查看



在按照对象视角展示的情况，点击“操作”栏的明细结果，进入详细报告。在搜索框输入查询条件，点击查询按钮，能够成功查询出对应条件的结果。



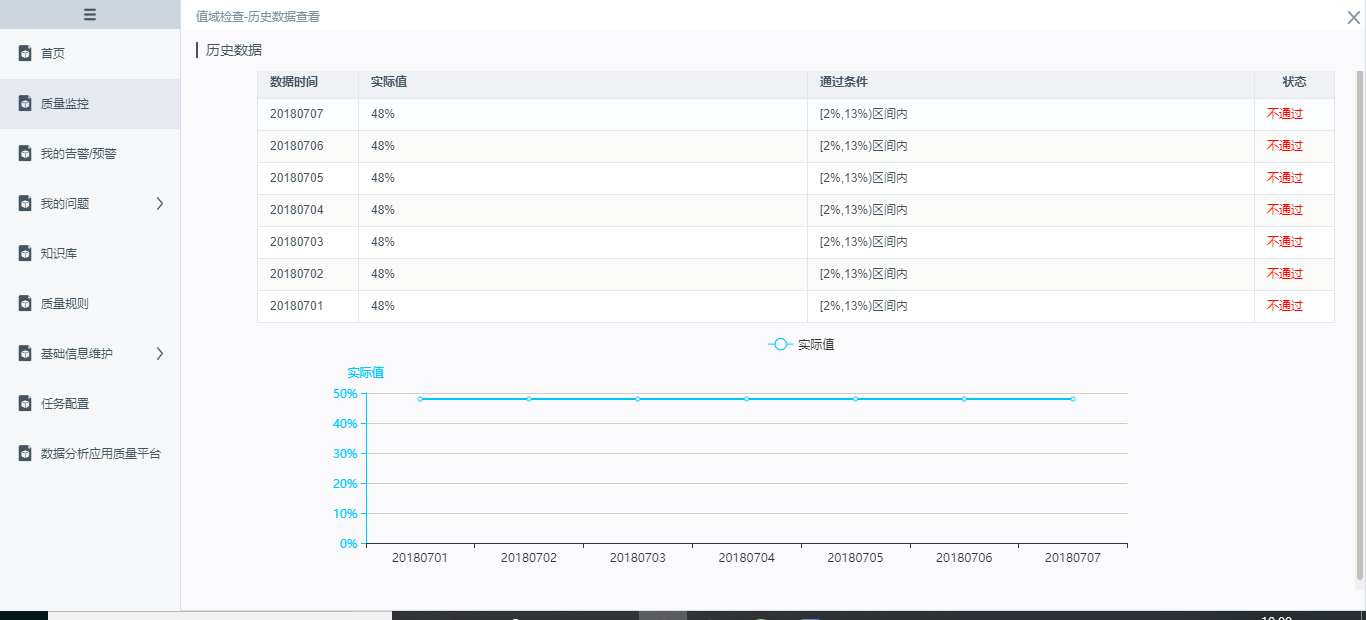
点击“错误详情”进入，页面展示具体的错误数据。点击导出按钮，能够成功导出所有错误数据。



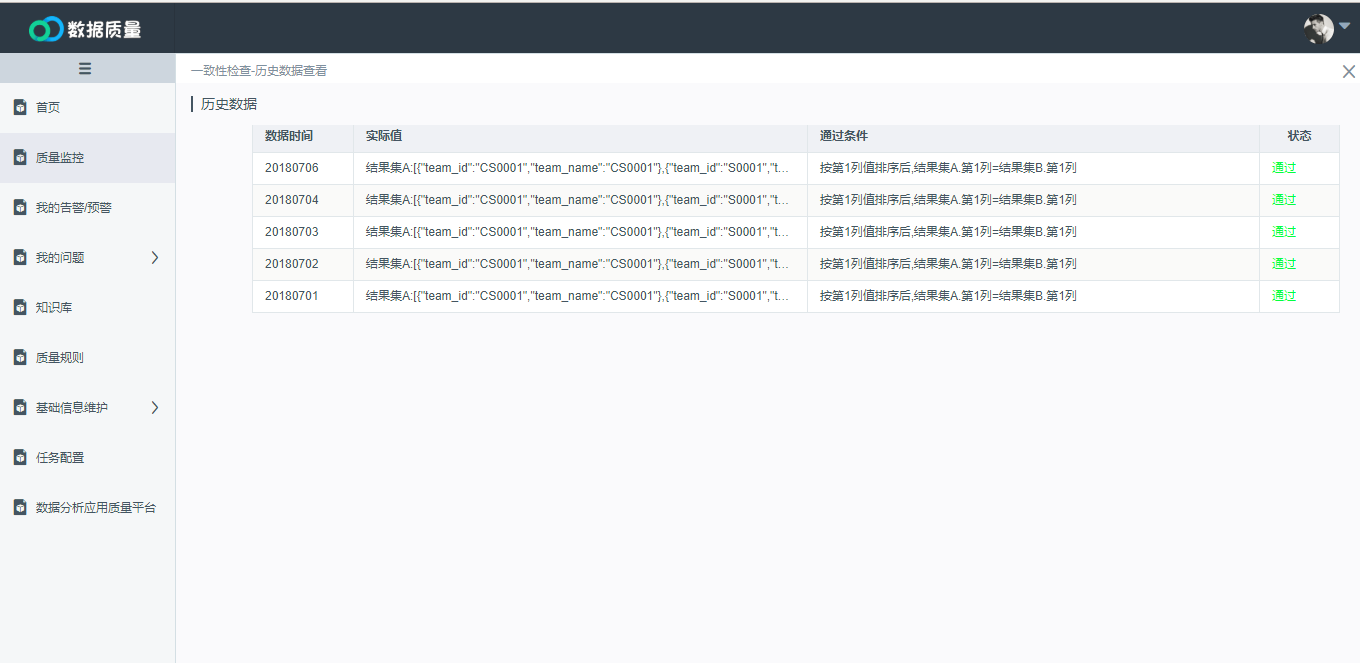
点击“历史数据”进入，页面展示出历史数据及趋势图。根据不同的结果类型展示不同的界面。

如果是数值类型，则展示表格和该数值的历史趋势，如果为非数值类型，则仅展示表格。如果为多条记录类型，则展示正确率和检查总数的表格的趋势。

数值类型：



非数值类型：

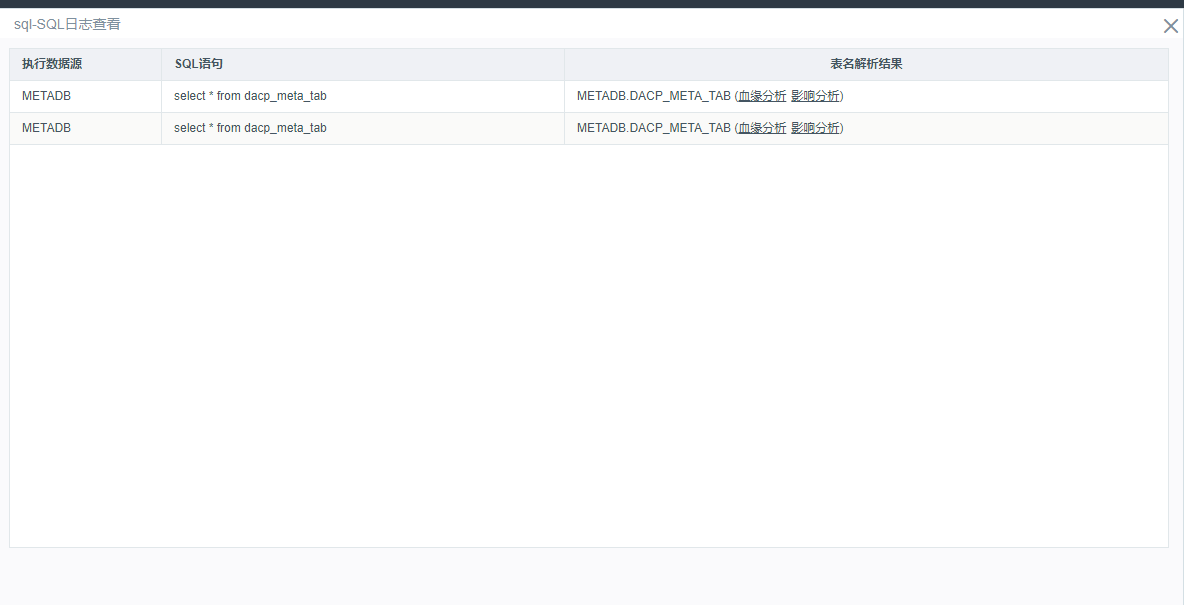


多条记录类型：



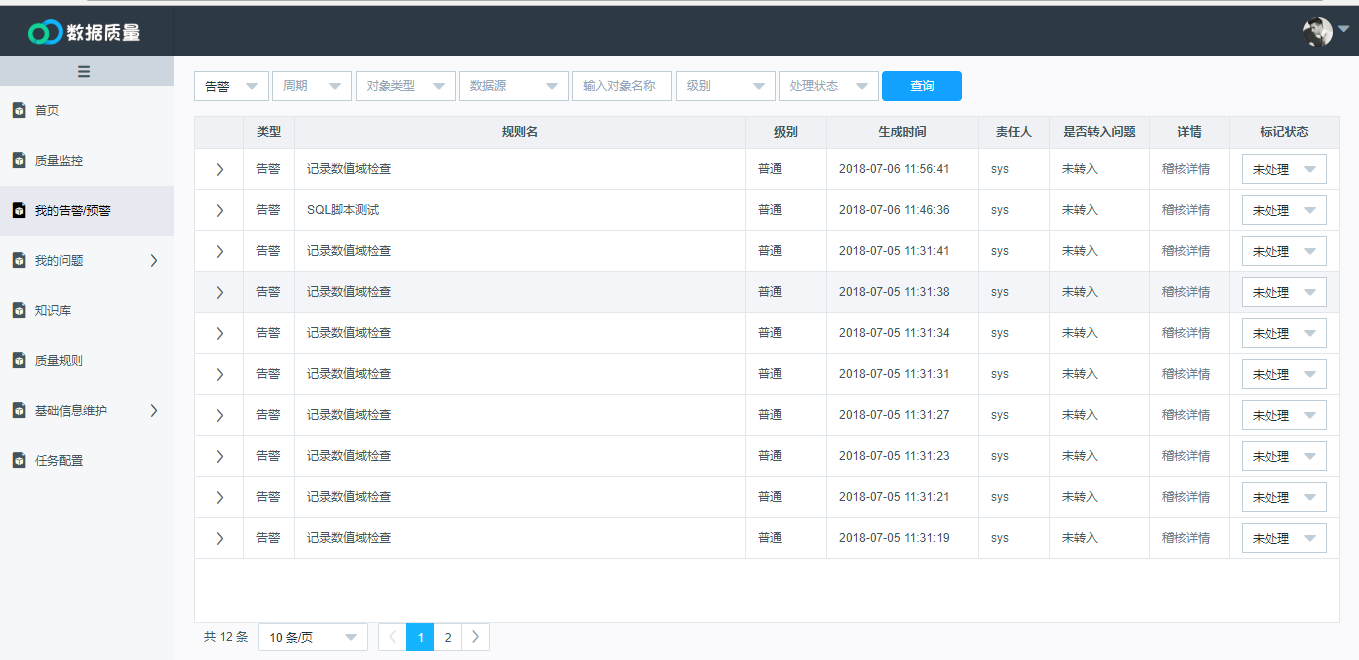
点击SQL日志，可以查看结果执行过程的SQL明细，并支持解析SQL得到表名，用于查看表的血缘分析

和影响分析



## 我的告警/预警

点击左侧“我的告警/预警”菜单，进行我的告警/预警功能页面，能够查看当前自己的告警和预警情况，



点击稽核详情可查看，稽核中更明细的信息



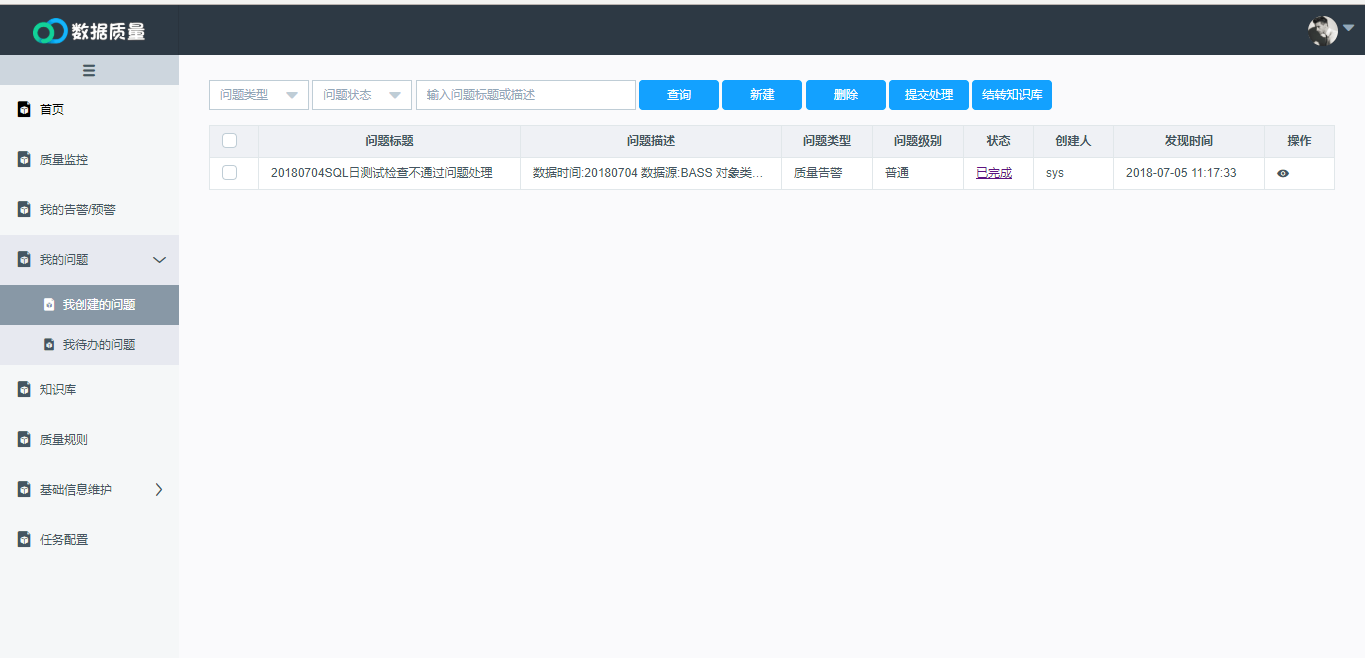
若该告警/预警未生成问题，则可以直接标记状态，已生成问题的告警/预警，需要等问题处理结束后，自动标记状态为已处理

对已转入问题的告警/预警，支持问题链接查看

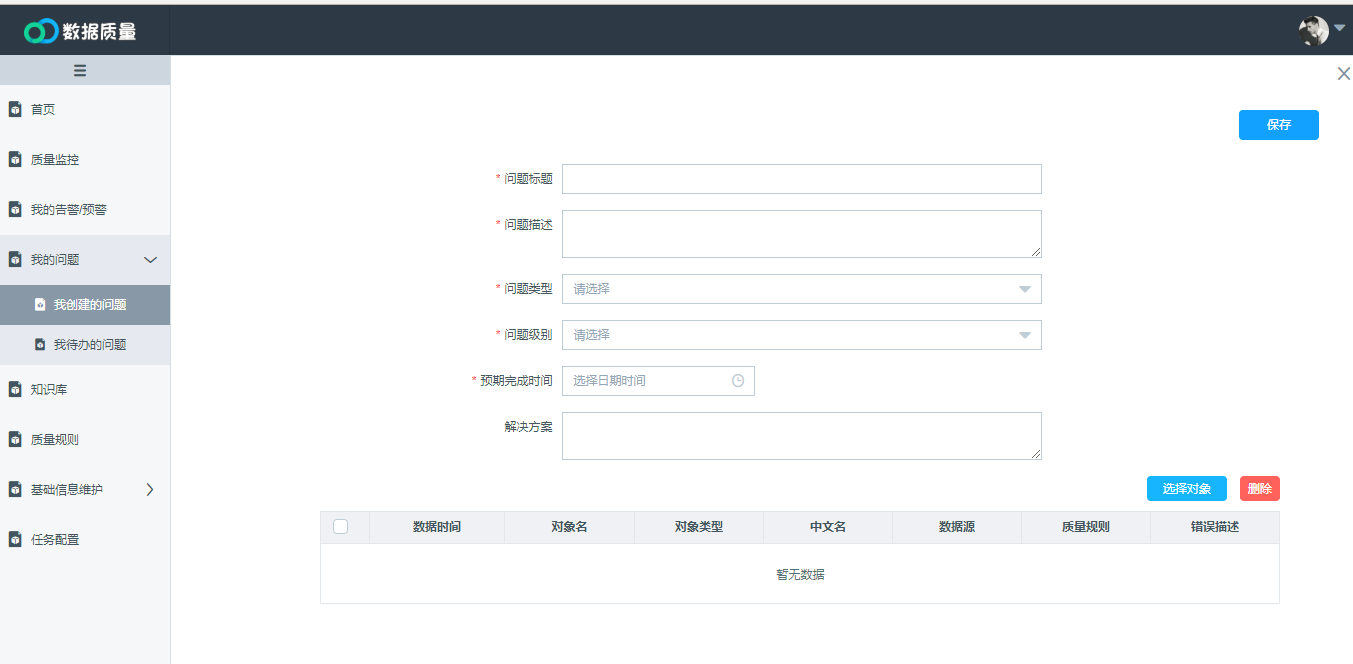
## 我的问题

* + 1. 我创建的问题

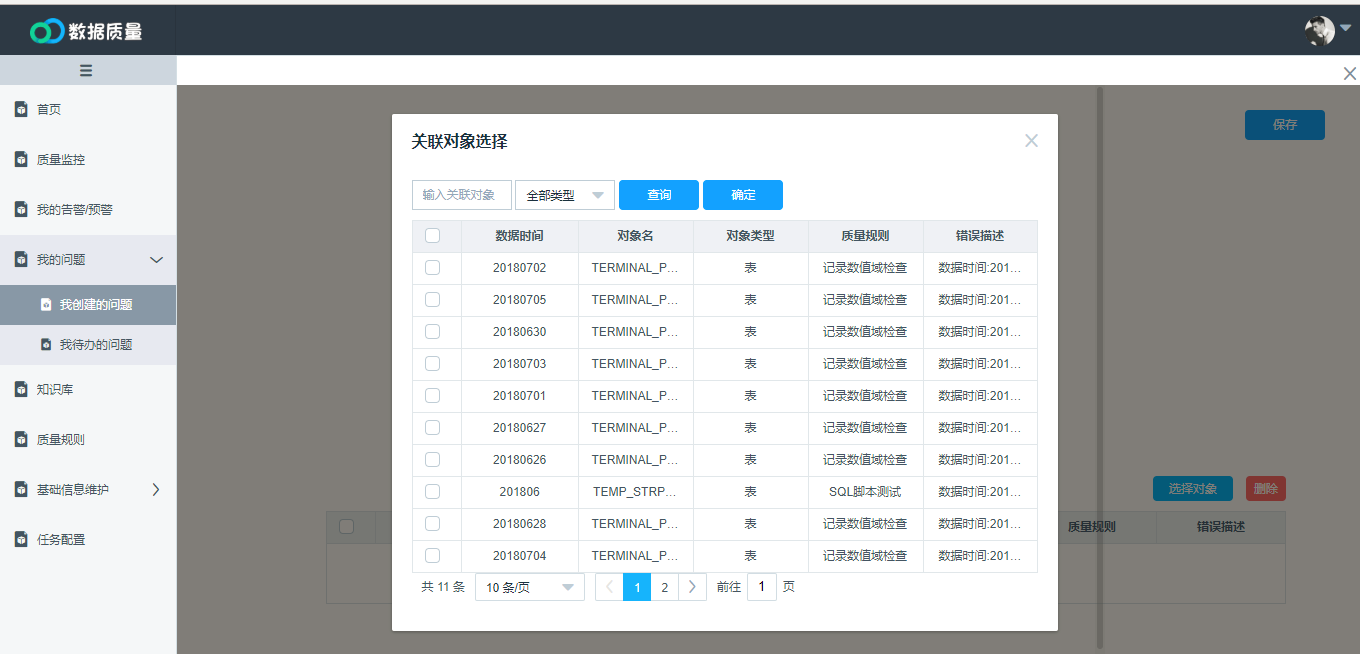
点击菜单-我的问题-我创建的问题，进入功能页面

、

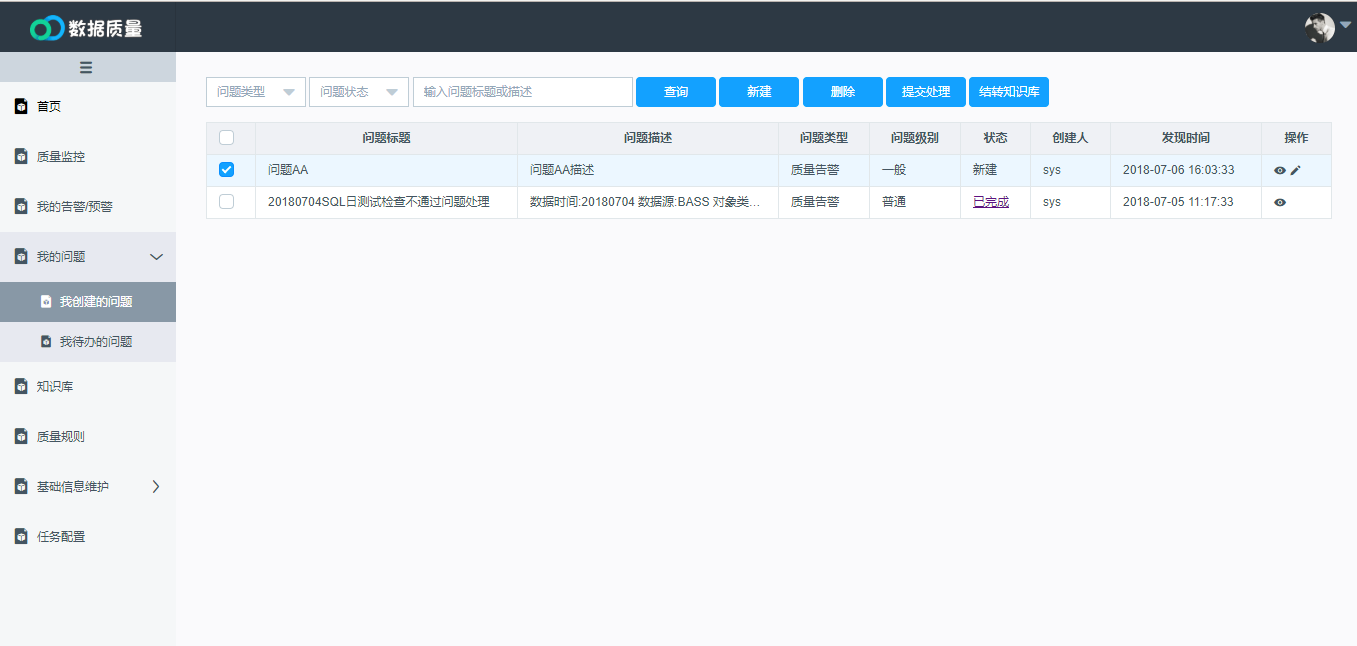
点击新建，进入新建问题



选择对象，关联告警/预警



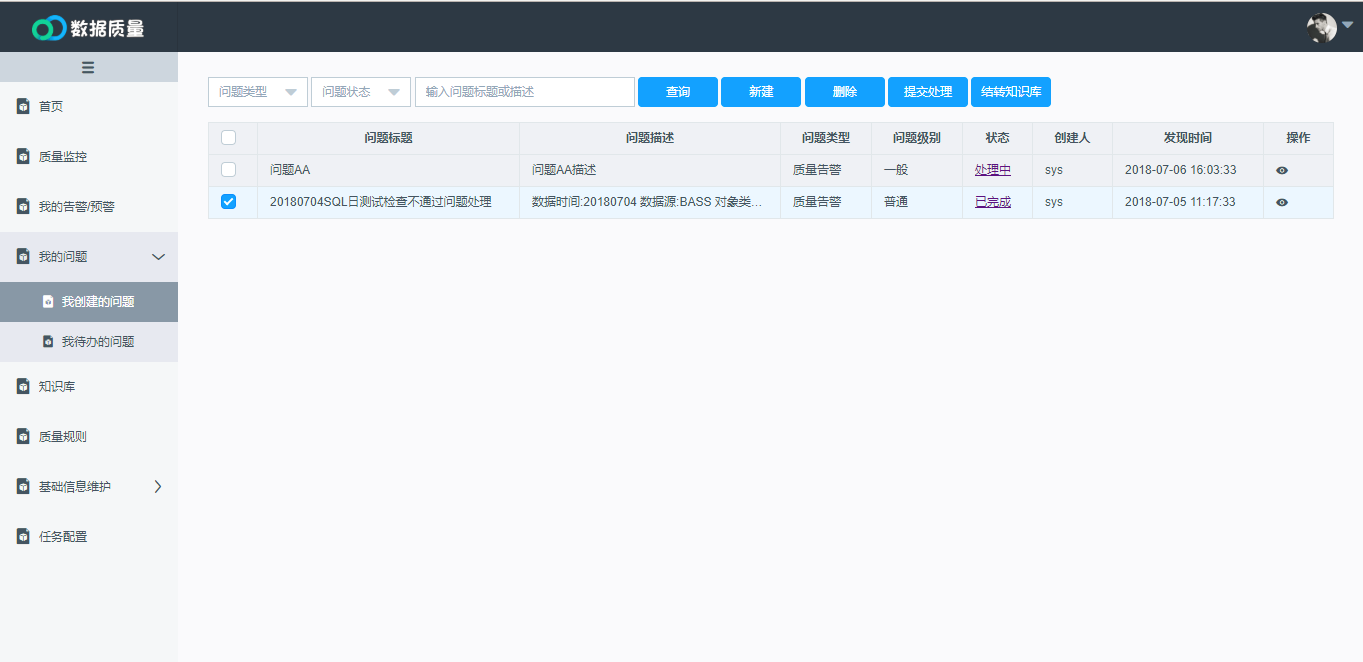
勾选新建的问题，进入流程提交处理页面



输入任务标题，任务描述，选择任务责任人，点击提交

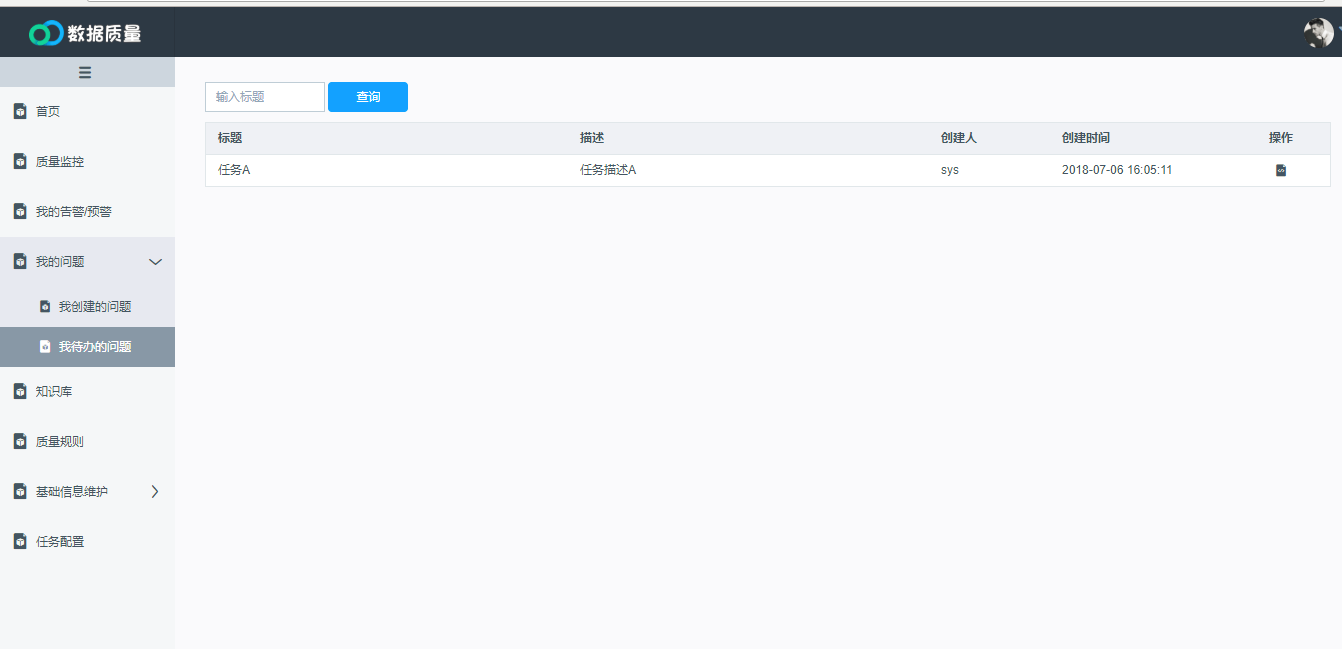


对已完成的任务，可以直接结转知识库

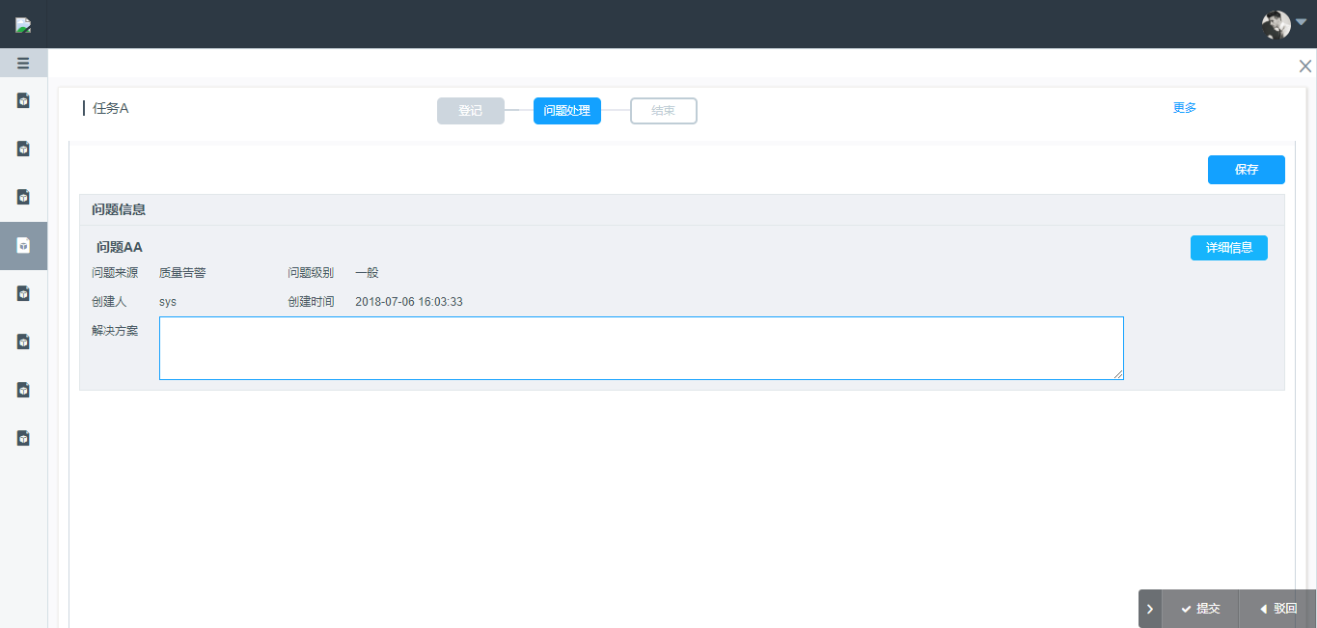


* + 1. 我待办的问题

点击菜单-我的问题-我待办的问题，进入功能页面



点击处理操作按钮，可以进入问题处理页面

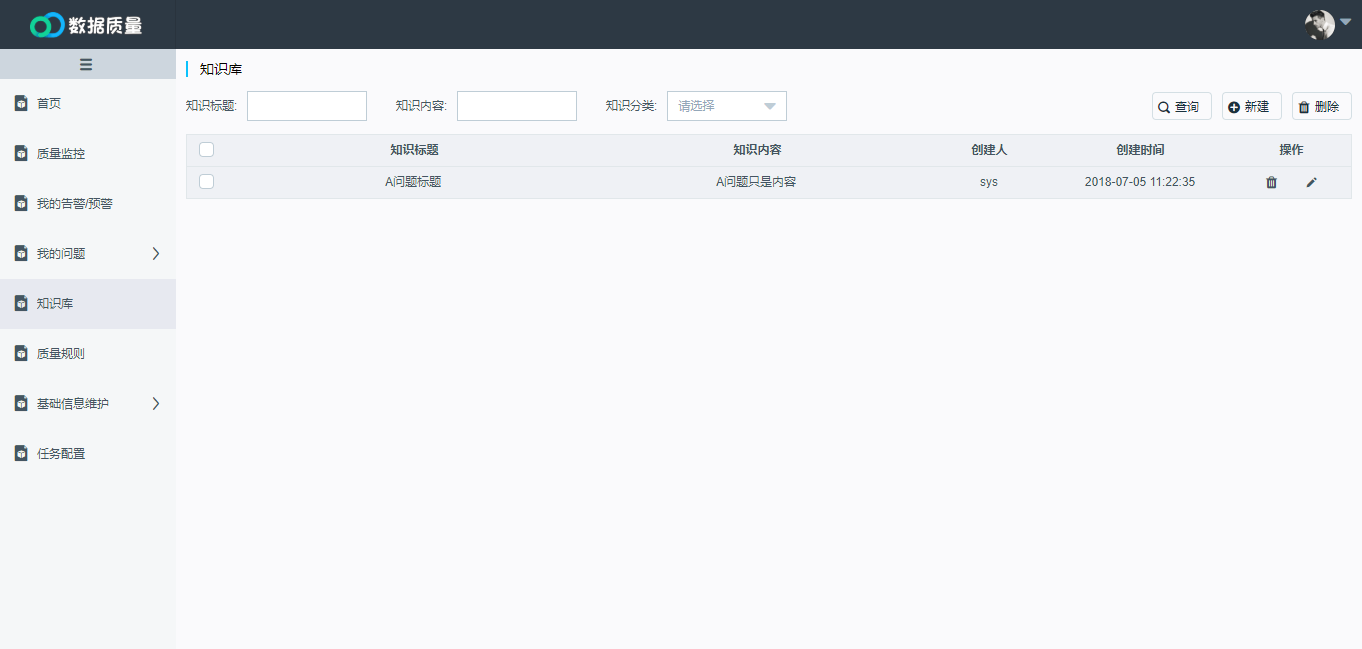


输入问题的解决方案，保存后点击提交按钮，输入流程备注，将该问题结束，问题

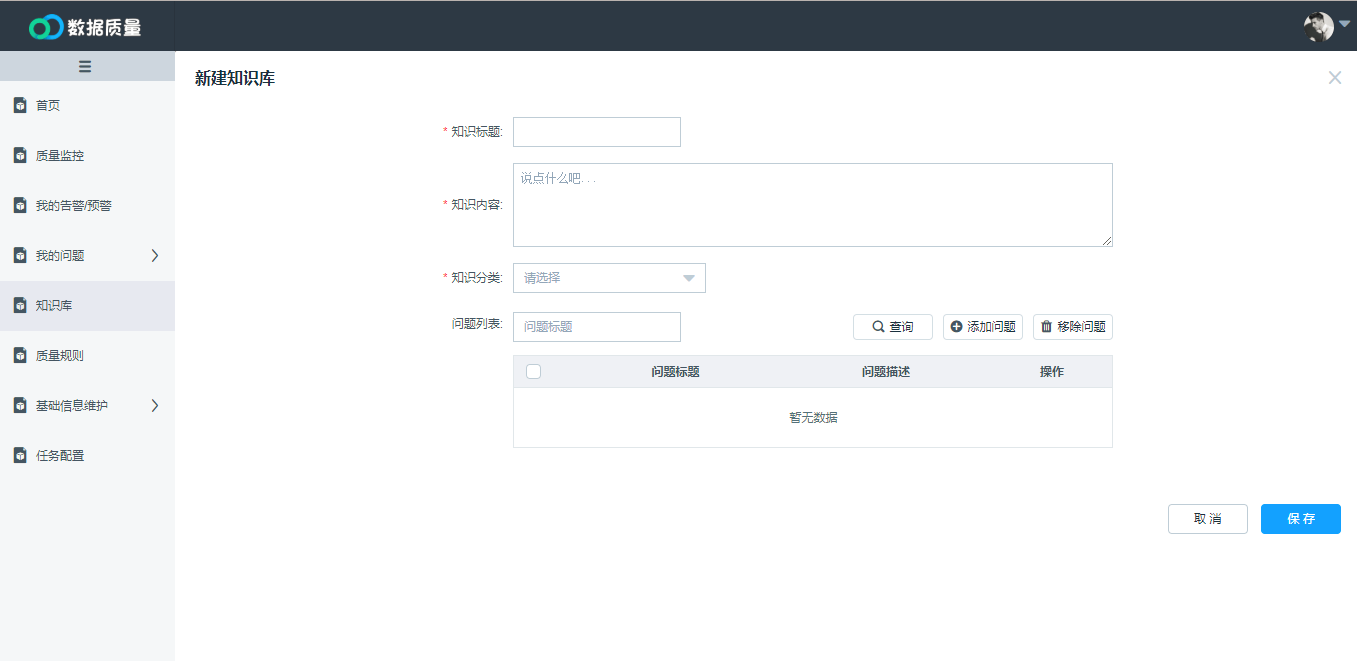
结束后，问题状态变为已完成

## 知识库

点击知识库菜单，进入知识库功能页面



点击新增进入知识库的新增，输入知识标题，知识内容，选择知识分类，添加问题，点击保存即可保存知识库

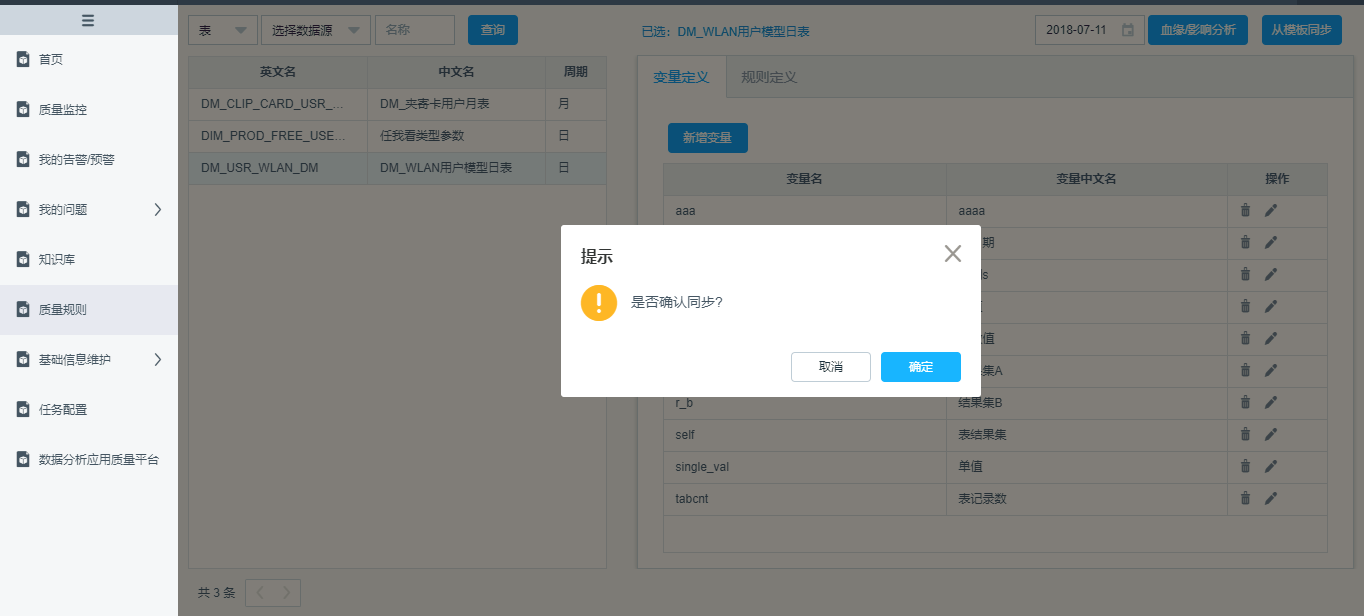


## 质量规则

点击质量规则菜单，进入质量规则管理界面。输入查询条件点击查询按钮，能够成功查询出对应条件的规则。右上角从模板同步按钮，选中一个对象，同步之后会同步变量和规则。



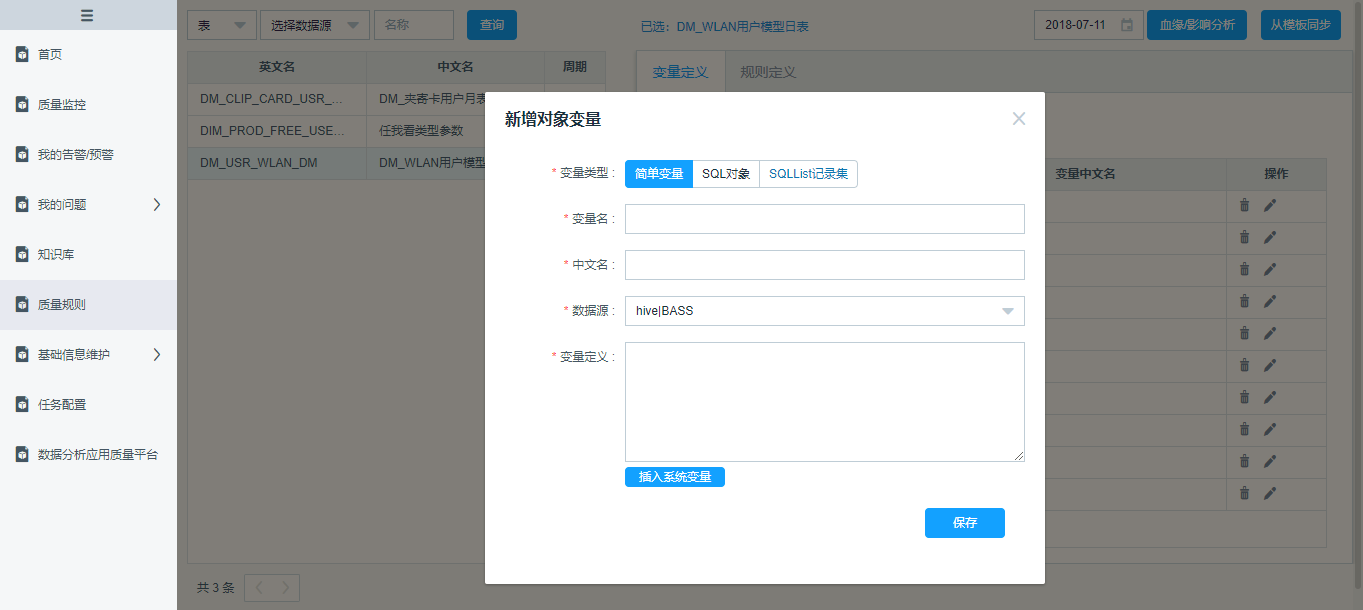
选中一个对象，点击从模板同步按钮，弹出确认提示框，点击确定能够成功同步，基于安全性考虑，如果已存在变量或规则，则同步操作不会进行，即不会覆盖已有变量或规则。点击取消按钮，则不进行同步。



选中一条规则，进行变量定义，可以新增、编辑、删除变量。



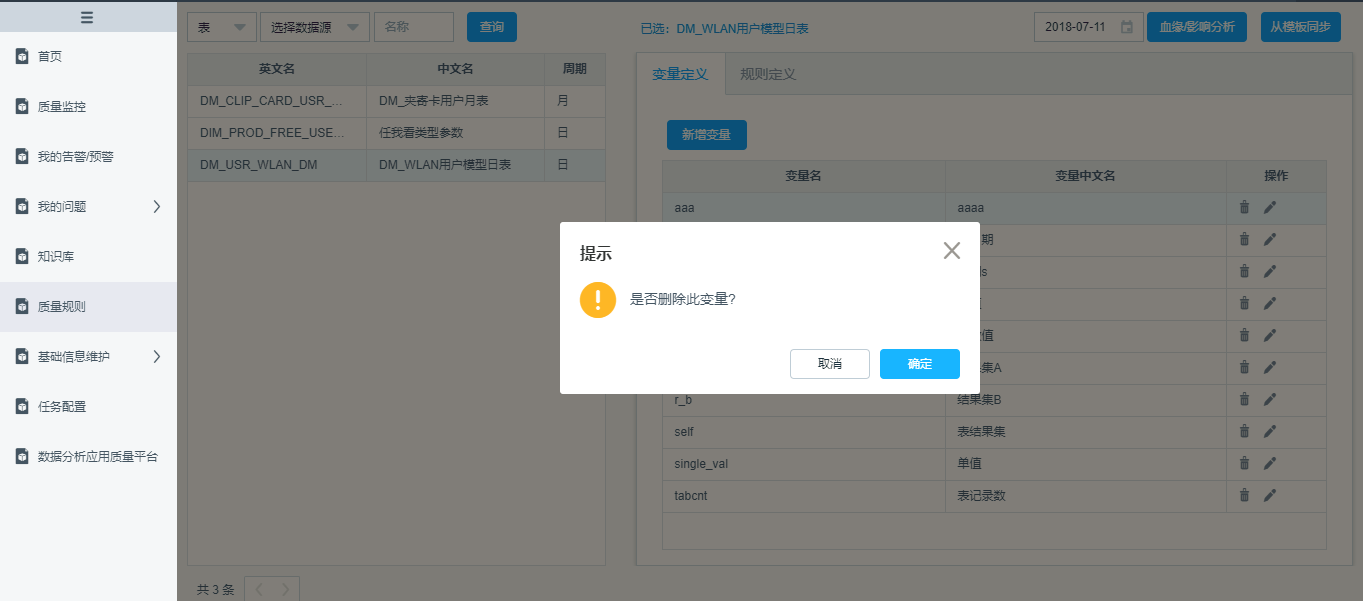
点击新增变量，可以新增简单变量、SQL对象、SQLlist记录集三种变量类型。填写变量名、中文名、数据源、变量定义，点击保存按钮，能够成功新增变量。



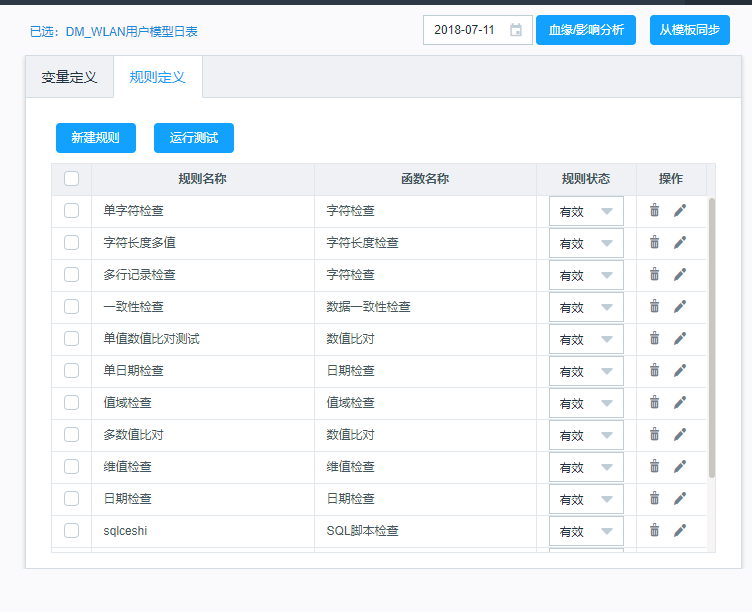
点击编辑按钮，可以编辑变量名、中文名、数据源、变量定义，点击保存按钮能够成功修改变量。



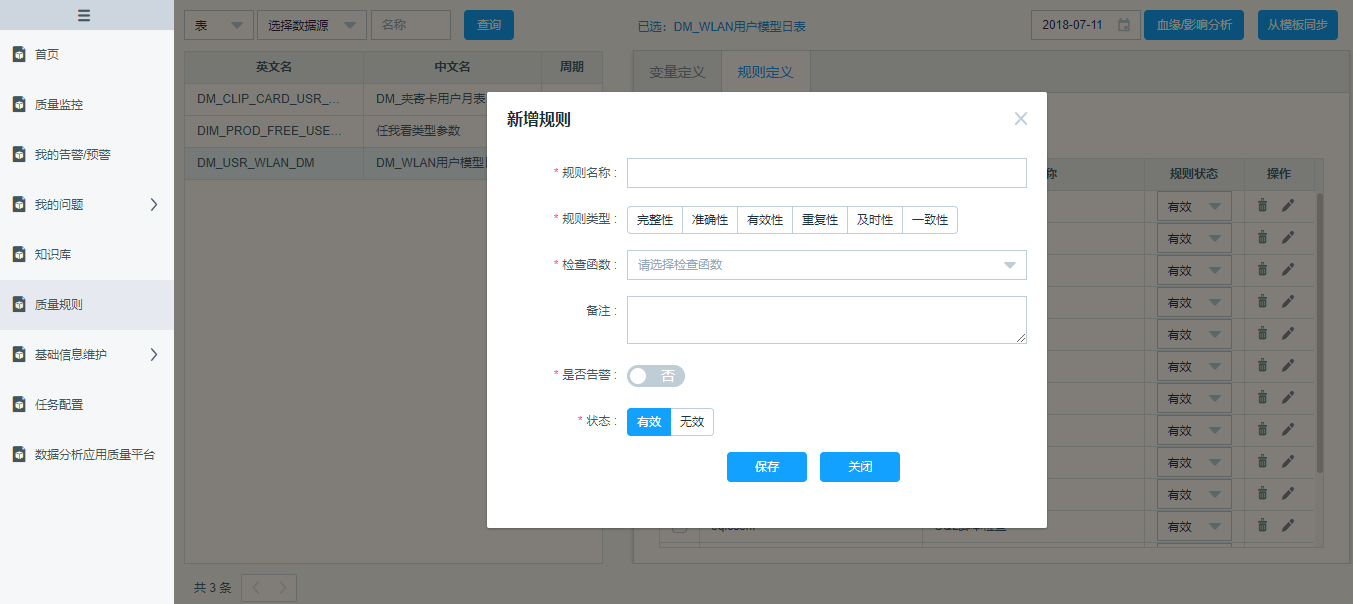
点击删除按钮，弹出确认删除框，点击确定按钮，成功删除变量，点击取消按钮，删除失败。



选中一条规则，进行规则定义，可以新增、编辑、删除规则以及对规则运行测试。



点击新建规则，填写规则名称、规则类型，选择检查函数，填写对应的函数条件，点击保存按钮，能够成功保存新增规则。



说明：

新增规则时，如果勾选了是否告警，则会展开告警相关的输入项，包括告警级别，是否转问题中心，

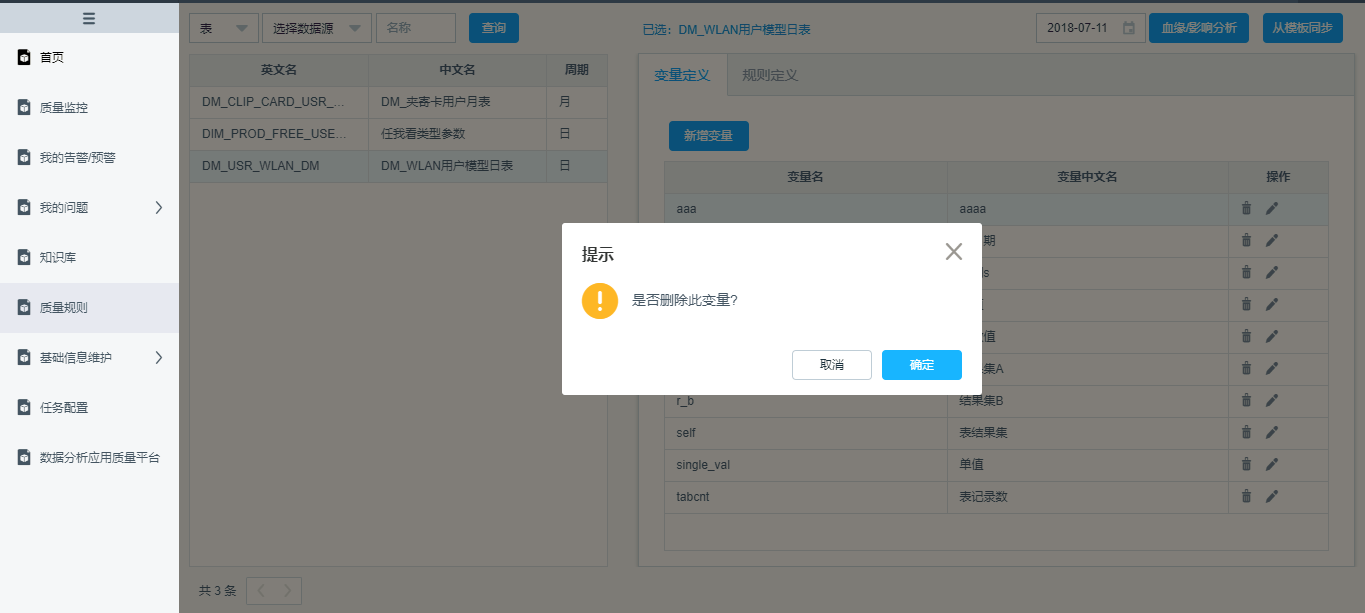
接收人。其中告警级别和接收人必填。是否转问题中心如果选择为是的，则该规则发生告警时，会自动生成问题并流转至所选的接收人。此时问题的创建者为系统管理员sys。



点击编辑按钮，编辑规则名称、规则类型，选择检查函数，填写对应的函数条件，点击保存按钮，能够成功修改规则。



点击删除按钮，弹出确认删除框，点击确定按钮，成功删除规则，点击取消按钮，删除失败。



选中一条规则，点击运行测试，进入执行测试页面。该页面可以修改数据时间，执行测试，查看运行日志和运行结果。



点击执行测试按钮，成功输出运行日志。



点击查询运行结果按钮，展示运行结果页面，下面展示运行结果数据。



点击运行结果数据左侧折叠按钮展开，可以成功查看运行结果详情。



**变量及函数说明：**

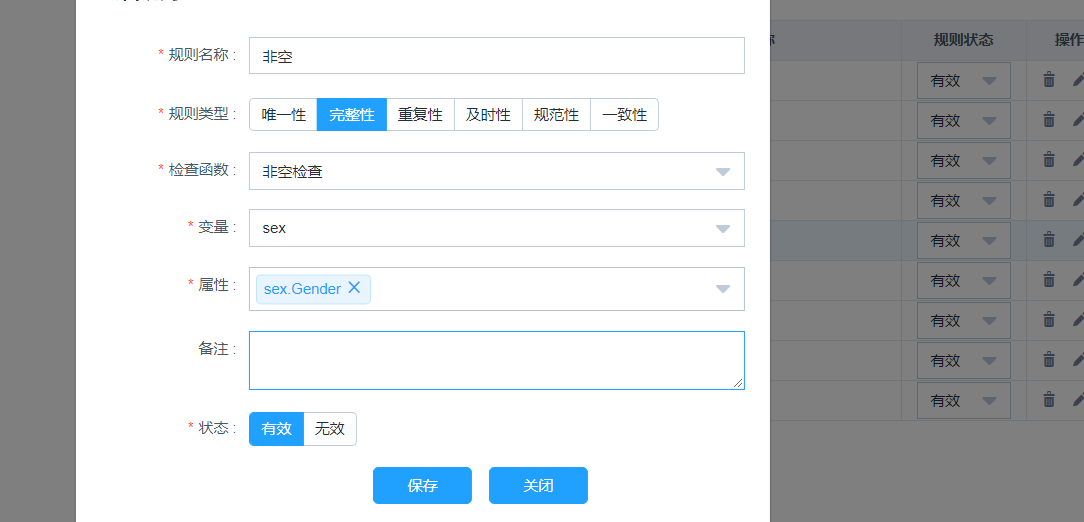
1. 变量类型说明：
2. 简单变量：通过SQL语句在数据库中查询出的一个值，例如select value from table where op\_time=’${op\_time}’(op\_time是运行质量检查时的批次号参数)，结果是一个值；
3. SQL对象：通过SQL语句在数据中查询出的一条(一个或者多个字段)数据，例如：select value1,value2,value3 from table where op\_time=’${op\_time}’，结果是一条数据；

SQL对象在定义的时候，需要定义属性，也就是需要定义字段，后期在使用的时候会直接指定到这个属性（字段）；

1. SQLList记录集：通过SQL语句在数据库中查询出的多条数据，例如：select value1,value2,value3 from table，结果是多条数据；

SQL对象在定义的时候，需要定义属性，也就是需要定义字段，后期在使用的时候会直接指定到这个属性（字段）；

1. 质量检查函数说明：
2. 非空检查：检查的对象，简单变量（单值），sql对象.属性(单值),同一sqllist记录集的1个或多个属性，支持表的多列字段，联合非空，即可a is null and b is null



规则类型：作为后期统计分类使用，按照检查类型进行分类；

变量：被检查的数据集变量

属性：被检查的数据集变量中的某个字段；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 数据一致性检查：检查对象为实体表数据，数据集A必须为sqllist记录集(多行多列),数据集B也必须为sqllist记录集(多行多列)



规则类型：作为后期统计分类使用，按照检查类型进行分类；

数据集A：被用来比较一致性的数据集变量；

数据集B：用来比较一致性的另一个数据集变量；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 维值检查：检查的对象:sqllist记录集.属性，维值为逗号分隔常量，或者单列多行值（即sqllist记录集.属性)

规则类型：作为后期统计分类使用，按照检查类型进行分类；

属性：被检查的数据集变量中的某个字段；

维值类型：（1）常量：维值指定为某几个固定的值；

（2）变量：维值指定为某个字段；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 数值比对：检查的对象:简单变量（单值），sql对象.属性(单值），sqllist记录集.属性，基准值，必须为常量或者单值(即，简单变量（单值），sql对象.属性(单值))

属性：被检查的数据集变量中的某个字段；

基准值类型：（1）常量:基准值指定为某个固定的值；

（2）变量:基准值指定为某个字段；

误差值：指定某个误差范围；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 值域检查：同数值比对，只是阈值必须为常量，暂不支持变量

规则类型：作为后期统计分类使用，按照检查类型进行分类；

属性：被检查的数据集变量中的某个字段；

条件：确定值域检查的范围，根据检查需求选择；

阈值：根据检查范围指定的一个临界值；

时间段配置：可以配置单日期、日期范围、周期，取值可以是固定的值或范围；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 字段唯一性检查：检查的对象就只能是同一sqllist记录集的1个或多个属性，支持一个表的一个字段或者多个字段联合起来检查

变量：被检查的数据集变量；

属性：被检查的数据集变量中的某个字段；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. 波动性检查：当期值和同/上期值，必须为单值的(即 简单变量（单值），sql对象.属性(单值))

当期值：被用来比较波动性的数据集变量；

同/上期值：用来比较波动性的另一个数据集变量；

波动范围：被检查变量的最小和最大波动范围；

状态：有效：在执行时会运行，无效：在执行时会忽略；

1. SQL脚本检查：直接通过sql检查出不符合的数据

规则类型：作为后期统计分类使用，按照检查类型进行分类；

数据源：用来选择检查对象的数据库；

错误记录SQL语句：用来输入错误记录的SQL语句；

全部记录SQL语句：用来输入全部记录的SQL语句；

状态：有效：在执行时会运行，无效：在执行时会忽略。

1. 字符检查：检查的对象:简单变量（单值），sql对象.属性(单值），sqllist记录集.

属性：被检查的数据集变量中的某个字段；

条件：确定字符检查的范围，根据检查需求选择

条件值：要比较的字符值

1. 日期检查：检查的对象:简单变量（单值），sql对象.属性(单值），sqllist记录集.

属性：被检查的数据集变量中的某个字段；

条件：确定日期检查的范围，根据检查需求选择

阈值：要比较的日期值

1. 字符长度检查：检查的对象:简单变量（单值），sql对象.属性(单值），sqllist记录集.

属性：被检查的数据集变量中的某个字段；

条件：确定字符检查的范围，根据检查需求选择

条件值：要比较的长度阈值

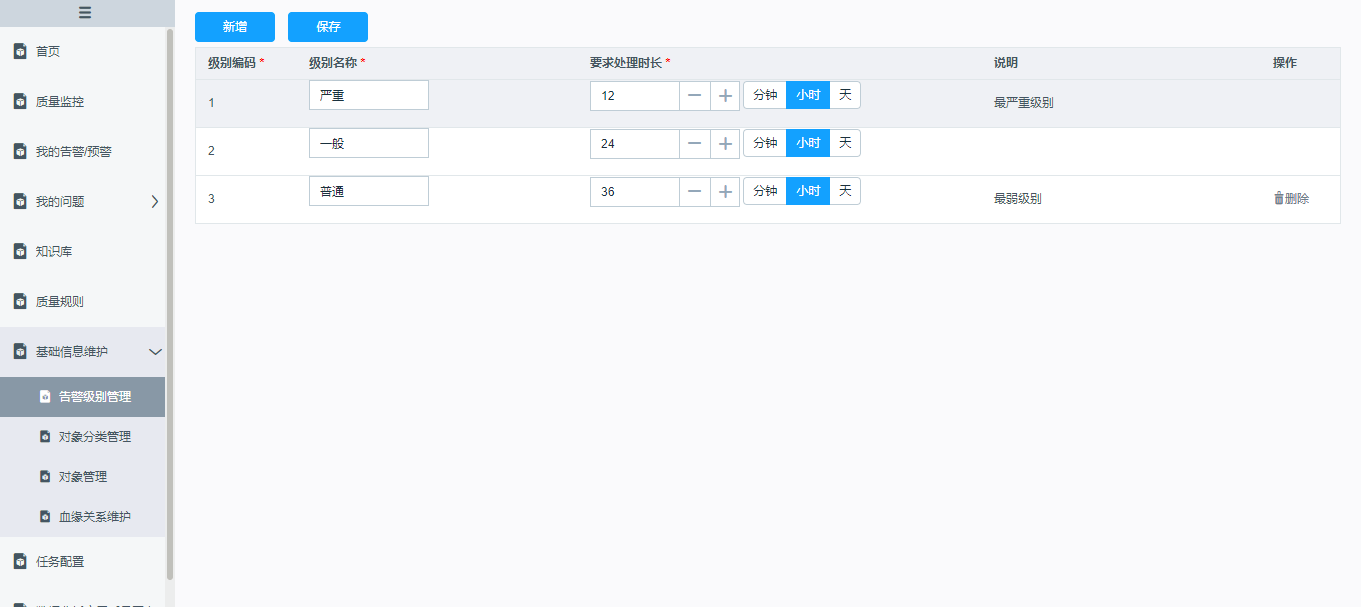
常用变量说明：

|  |  |  |
| --- | --- | --- |
| 数据日期:20180713 | 数据日期：201807 | 数据日期 2018 |
| 当天（YYYYMMDD）：  ${dayid} 替换后值20180713  当天（YYYY-MM-DD）：  ${day\_id} 替换后值为 2018-07-13 | 当月(YYYYMM)：  ${monthid} 替换后值:201807  当月(YYYY-MM)  ${month\_id} 替换后值 2018-07 | 当年：  ${year}替换后值 2018 |
| 昨天(YYYYMMDD)：  ${pdayid}替换后值为20180712  昨天(YYYY-MM-DD):  ${p\_day\_id} 替换后值为 2018-07-12 | 上月(YYYYMM)：  ${pmonthid} 替换后值 201806  上月(YYYY-MM)：  ${p\_month\_id} 替换后值2018-06 | 上年：  ${pyearid} 替换后值 2017 |
| 上月同天(YYYYMMDD)  ${p1month} 替换后值为 20180613  上月同天(YYYY-MM-DD)  ${p1monthid} 替换后值为 2018-06-13 | 上年同月(YYYYMM)：  ${pyearid}${month} 替换后值 201707  上年同月(YYYY-MM)  ${pyearid}-${month} 替换后值 2017-07 |  |
|  |  |  |

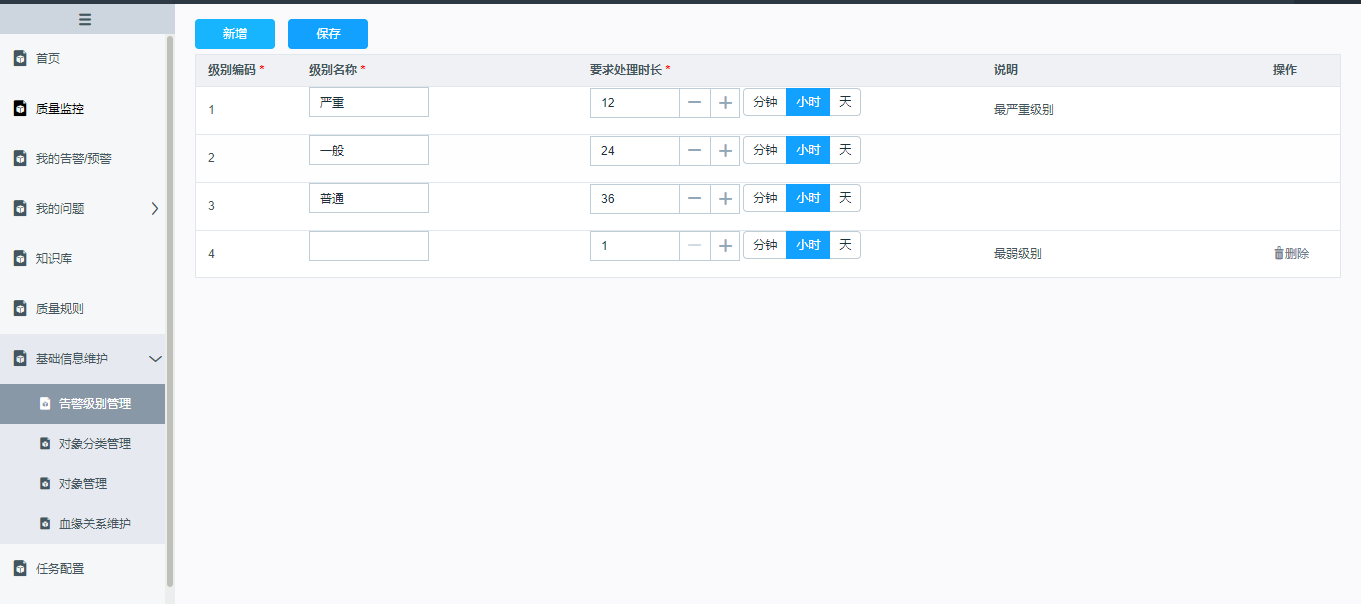
## 基础信息维护

* + 1. 告警级别管理

点击基础信息维护-告警级别，进入功能页面



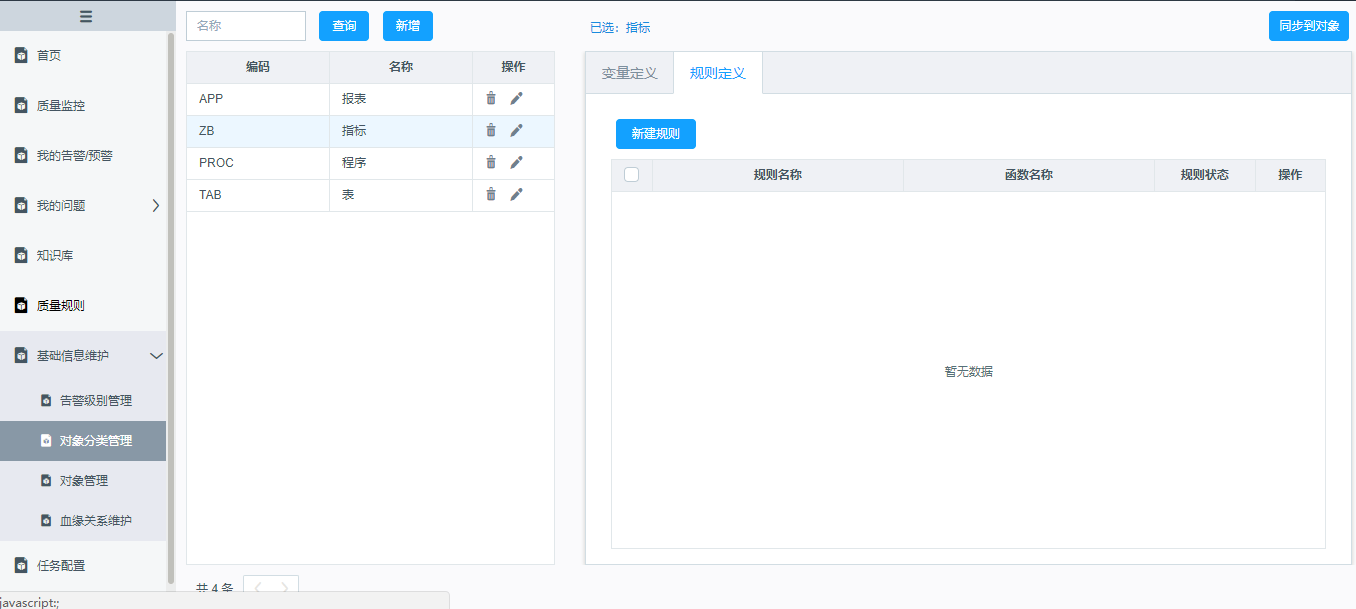
点击新增，则增加一行输入级别名称，要求处理时长并选择单位，点击保存可保存告警级别数据



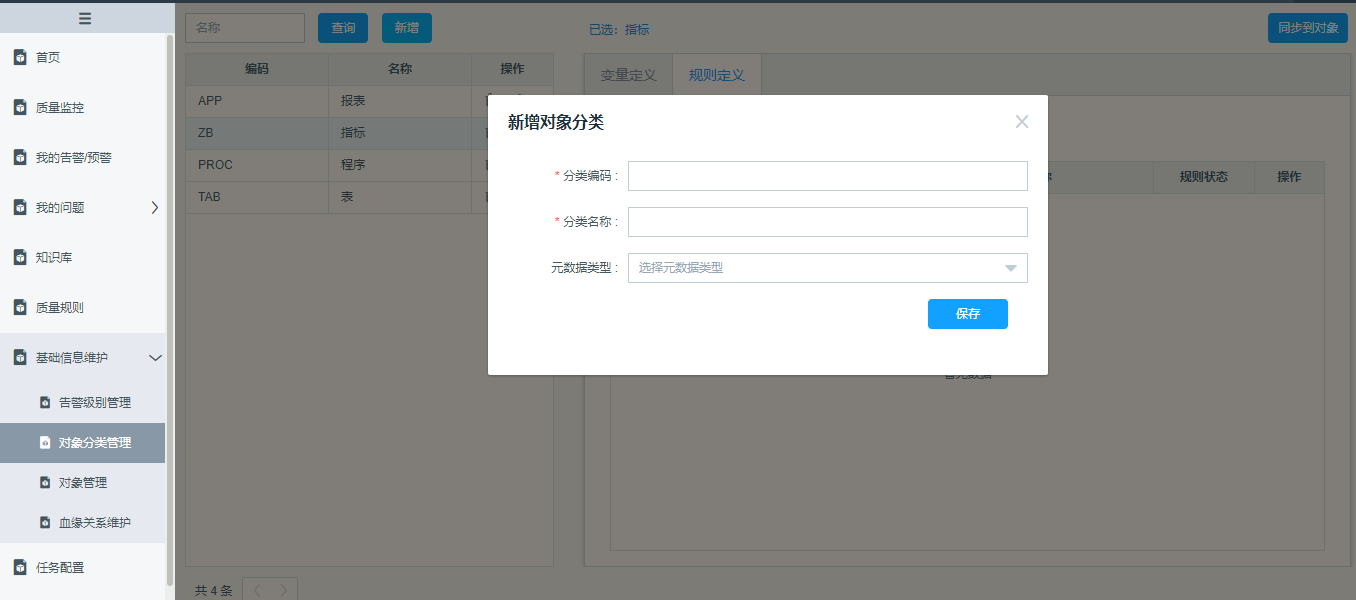
点击删除，则前台会删除该条记录，再点保存，刚才删除的记录入库生效。为保证级别编码的连续性，仅能编码值最大的记录。

* + 1. 对象分类管理

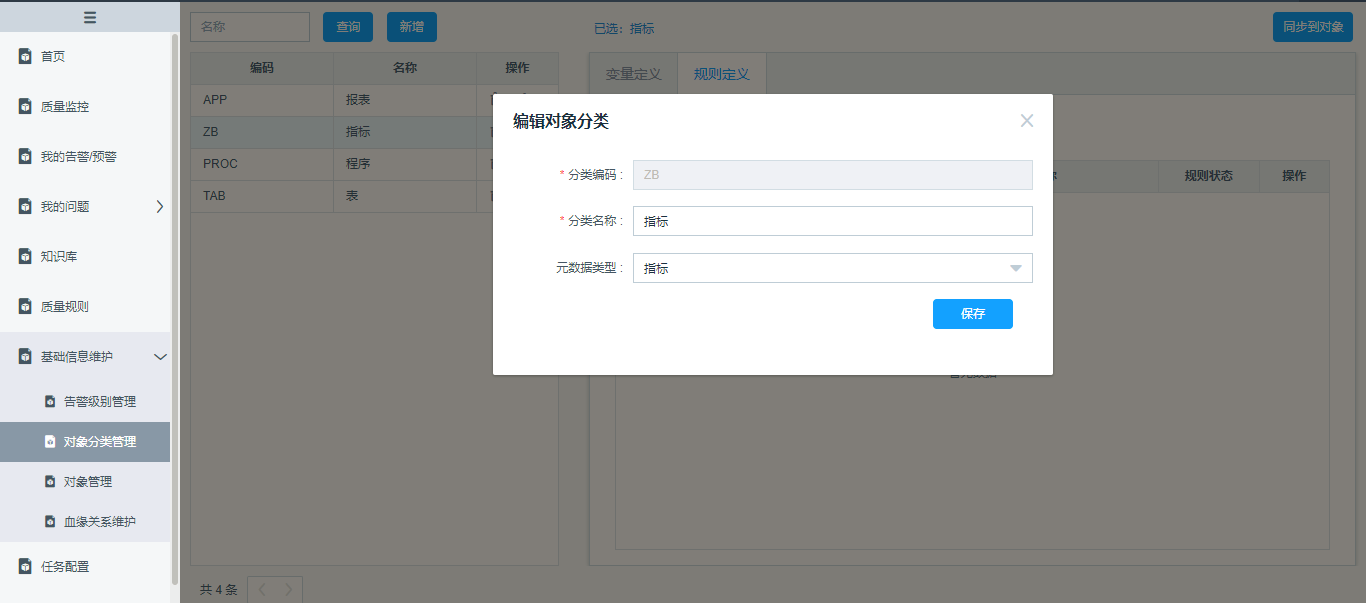
点击基础信息维护-对象分类管理，能够查看已有的对象分类。可以进行对象分类的新增、编辑、删除。输入名称点击查询按钮，能够成功查询出对应的对象分类。



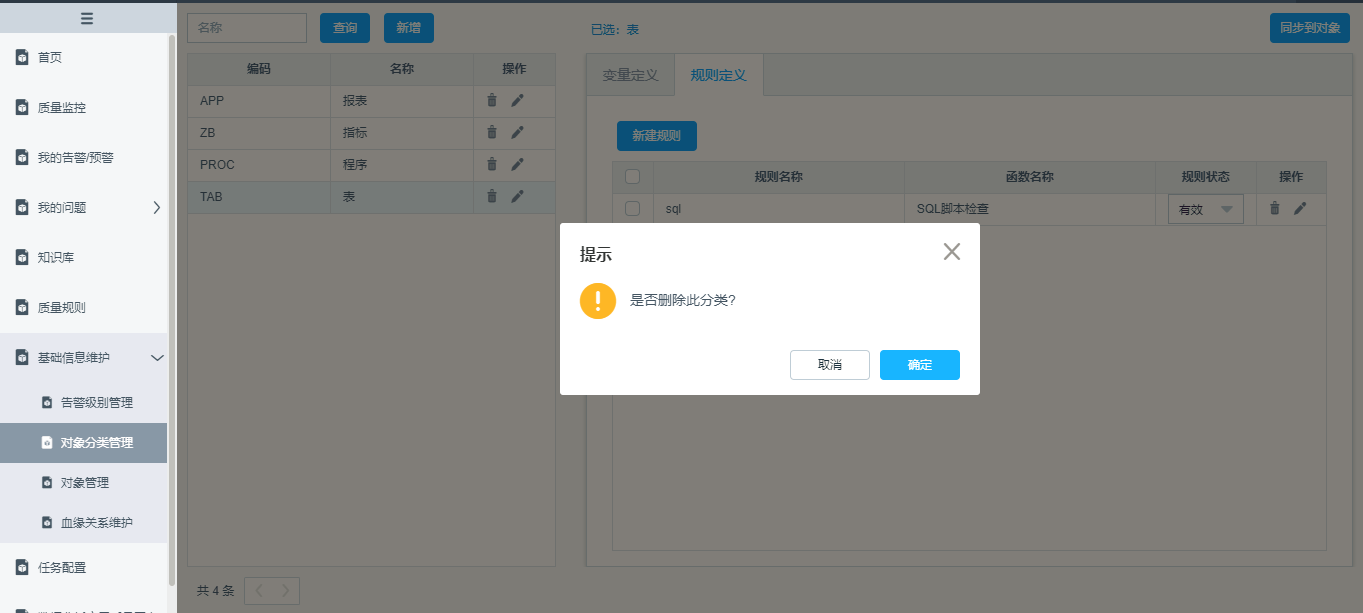
点击新增按钮，输入分类编码、分类名称、元数据类型，点击保存按钮，能够成功新增对象。



点击编辑按钮，可以修改分类名称、元数据类型，点击保存按钮，能够成功修改该对象。



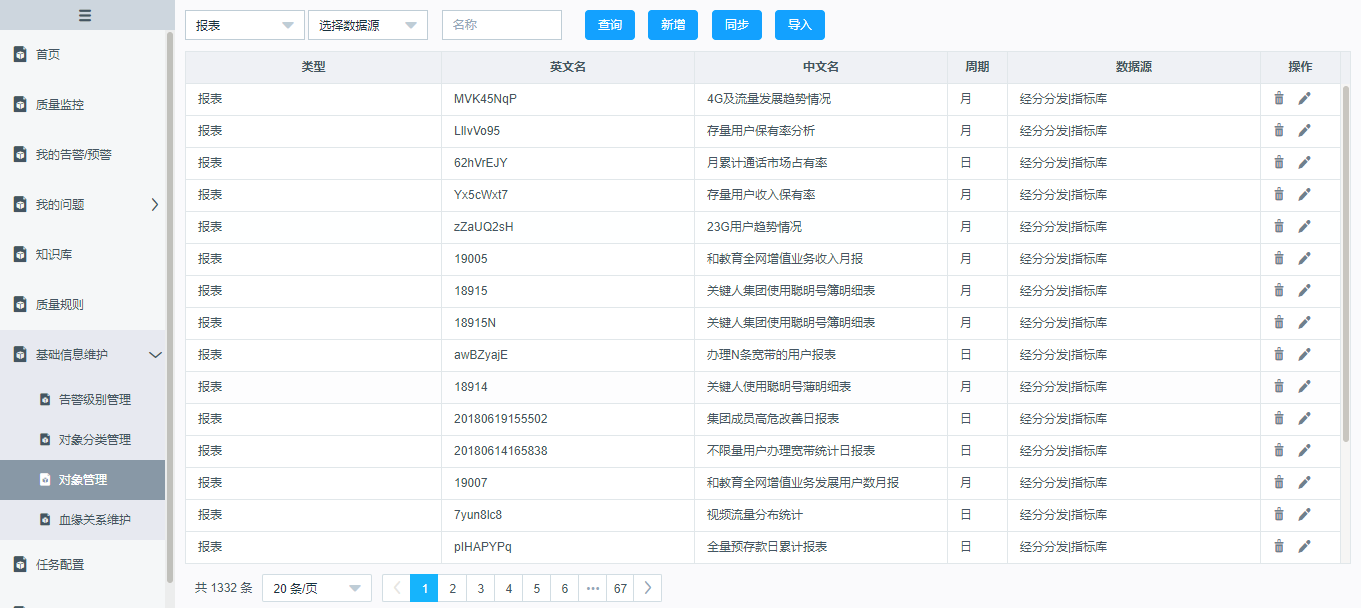
点击删除按钮，弹出确认框，点击确认删除，成功删除该对象，点击取消删除，不能删除该对象。



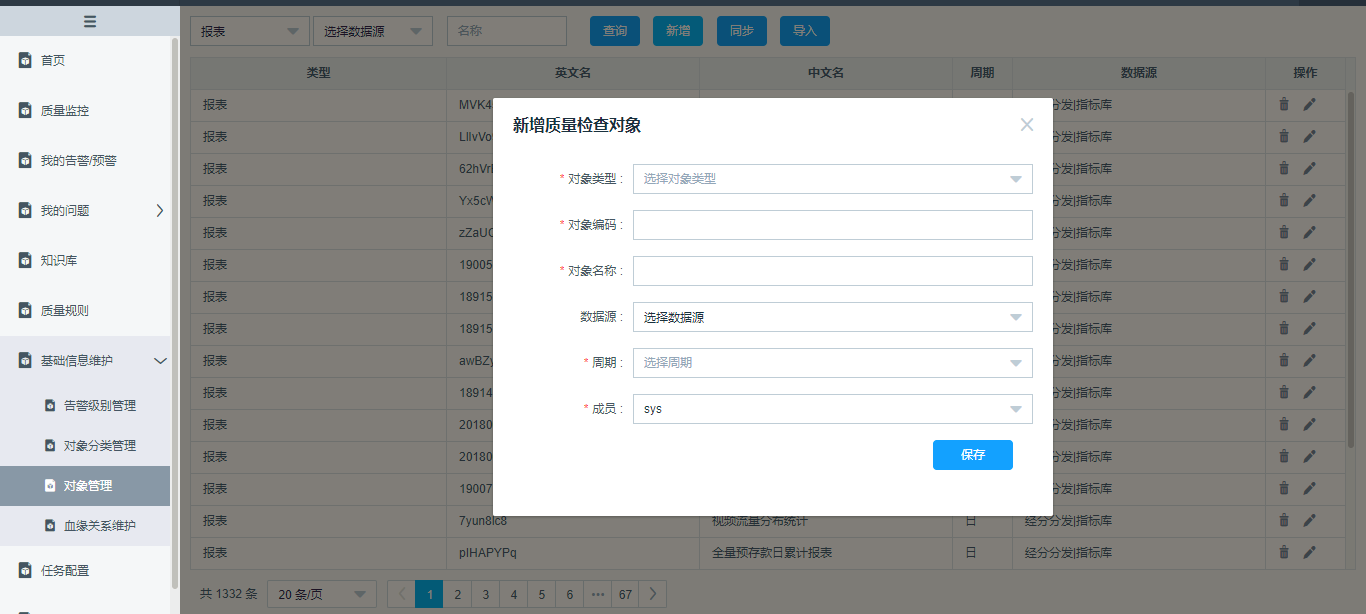
选中一个分类，进行变量定义，可以新增、编辑、删除变量。也可以进行规则定义。此处设置的变量定义规则定义仅仅作为该分类对象的模板变量定义和模板规则定义。

* + 1. 对象管理

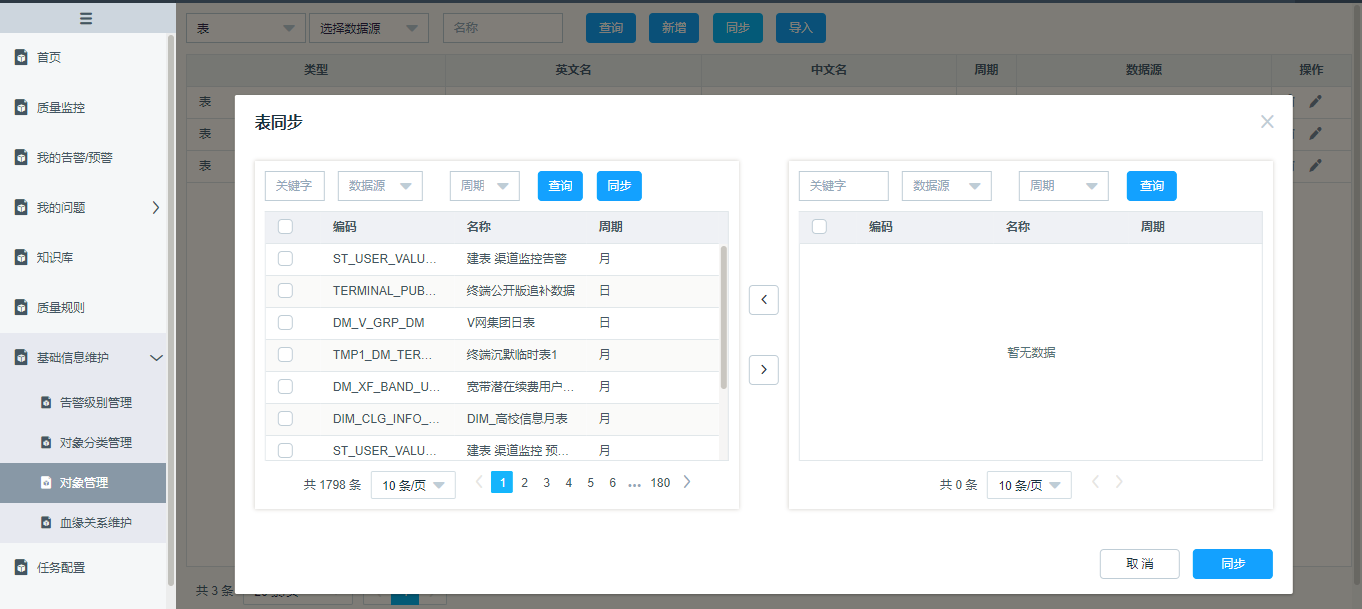
点击基础信息维护-对象管理，能够查看已有的对象。可以进行对象的新增、同步、导入、编辑、删除。输入搜索条件点击查询按钮，能够成功查询出对应的对象。



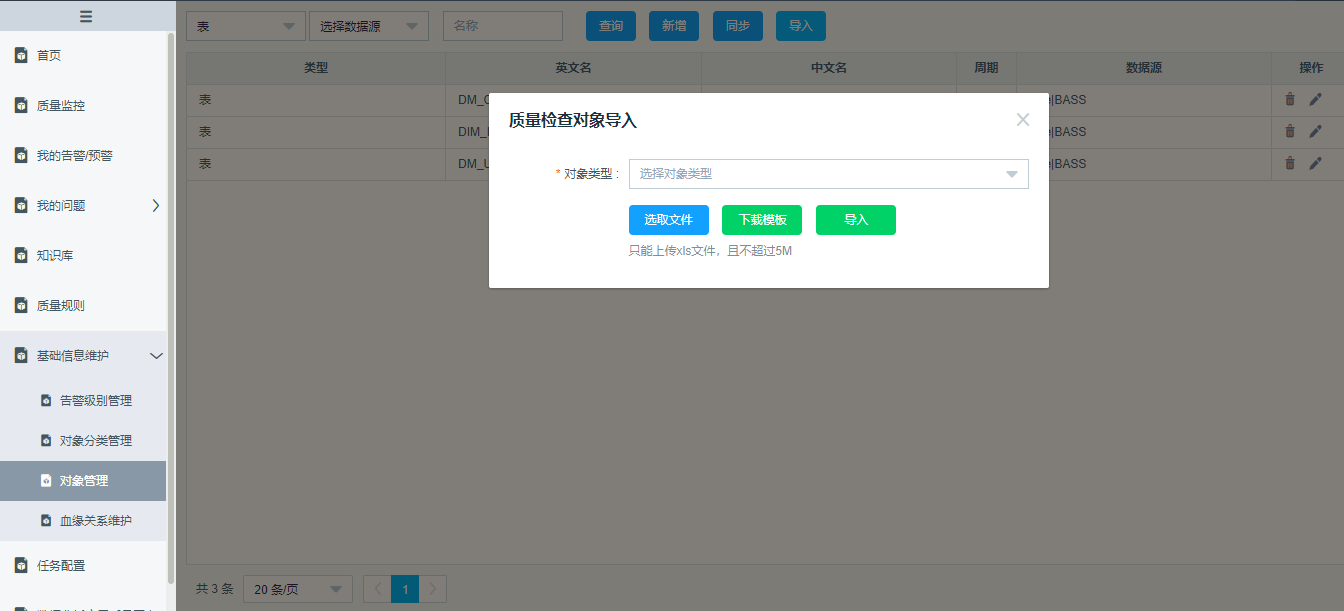
点击新增按钮，选择对象类型、填写对象编码、对象名称、选择数据源、周期，点击保存按钮，成功新增质量检查对象。



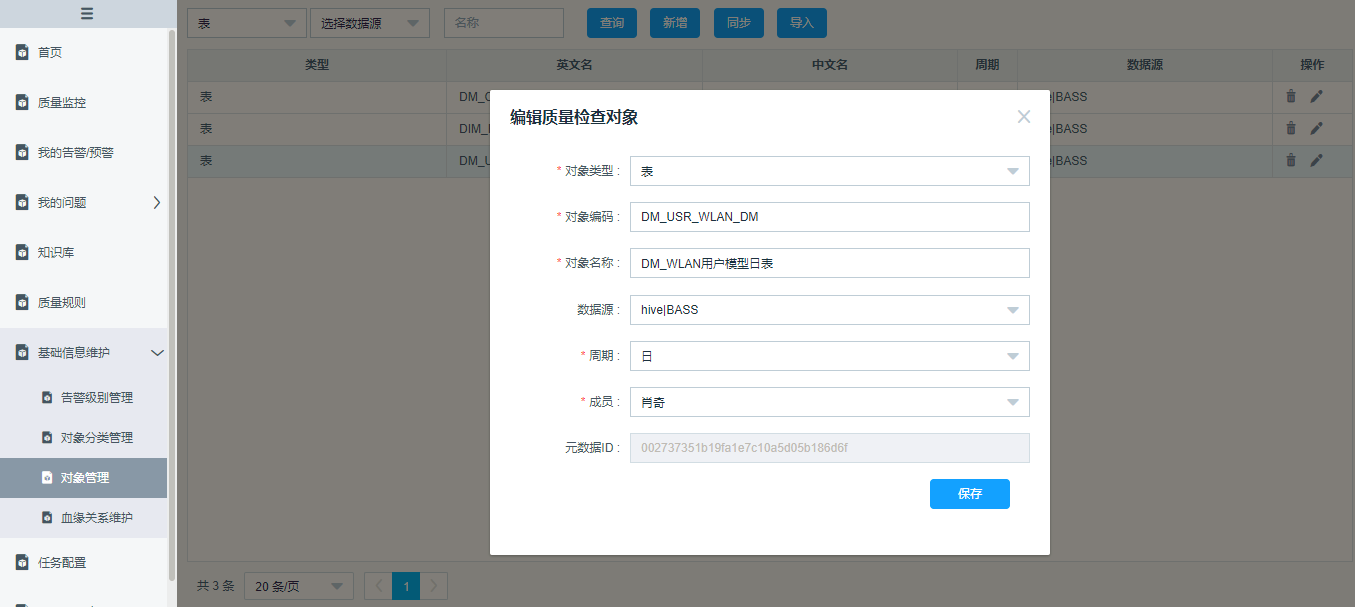
点击同步按钮，根据对象类型弹出相应元数据类型的同步选项框，点击左侧同步，同步左侧查询的所有的对象，点击右侧同步，同步右侧所有的对象。



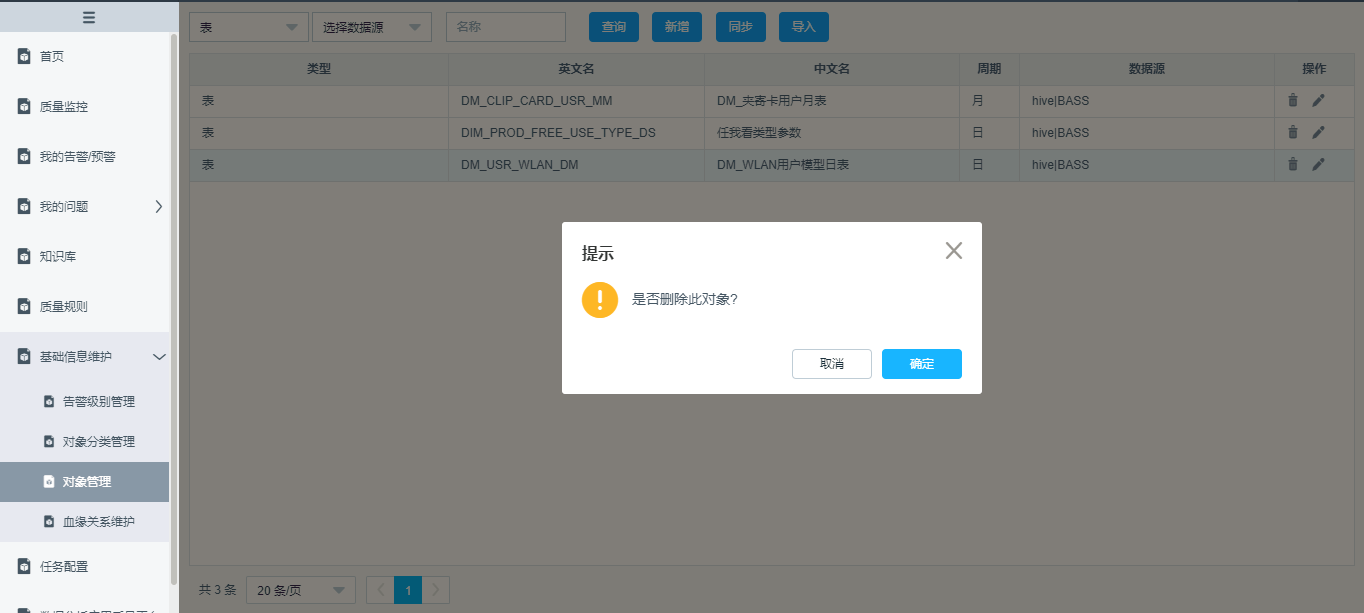
点击导入按钮，弹出导入框，选择导入对象类型，选取文件，可以成功导入对象管理。



点击编辑按钮，编辑对象类型、对象编码、对象名称、数据源、周期，点击保存按钮，成功修改质量检查对象。

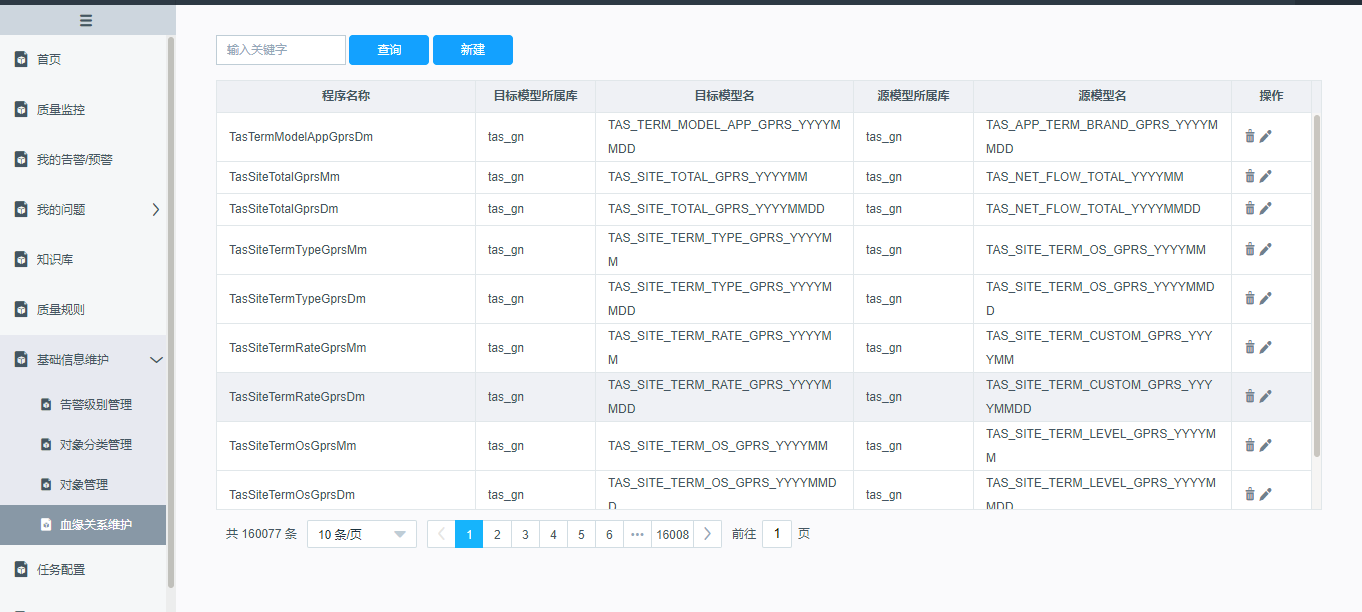


点击删除按钮，弹出确认删除框，点击确定按钮，成功删除对象，点击取消按钮，删除失败。

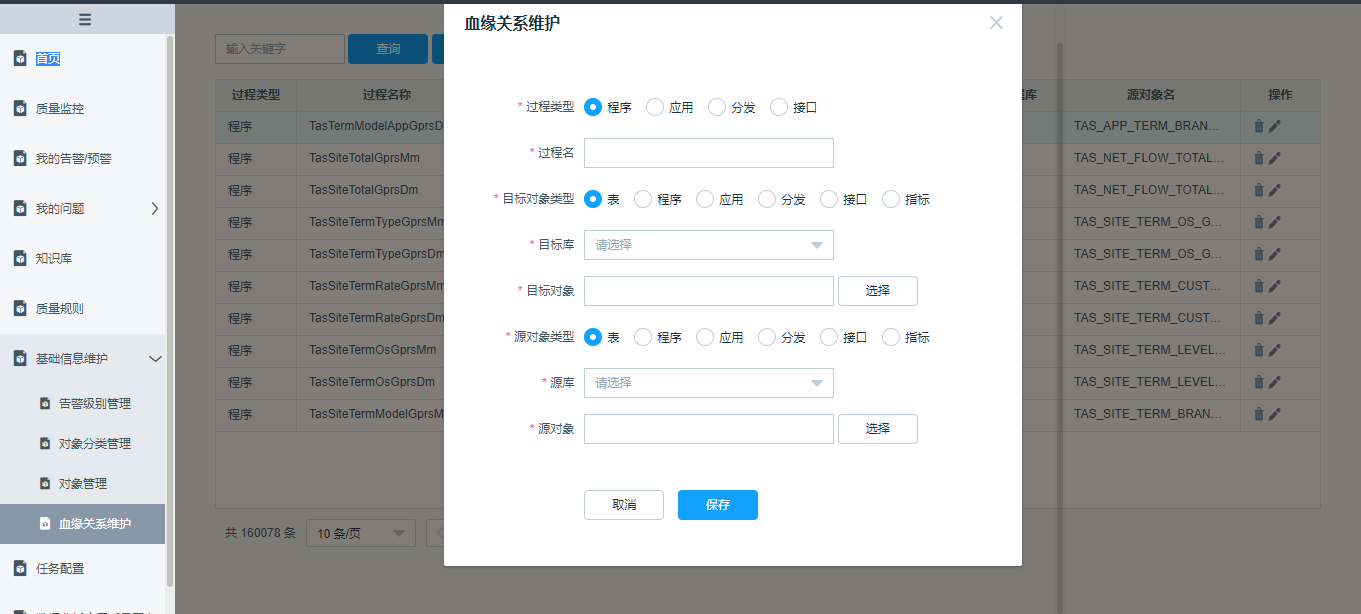


* + 1. 血缘关系维护

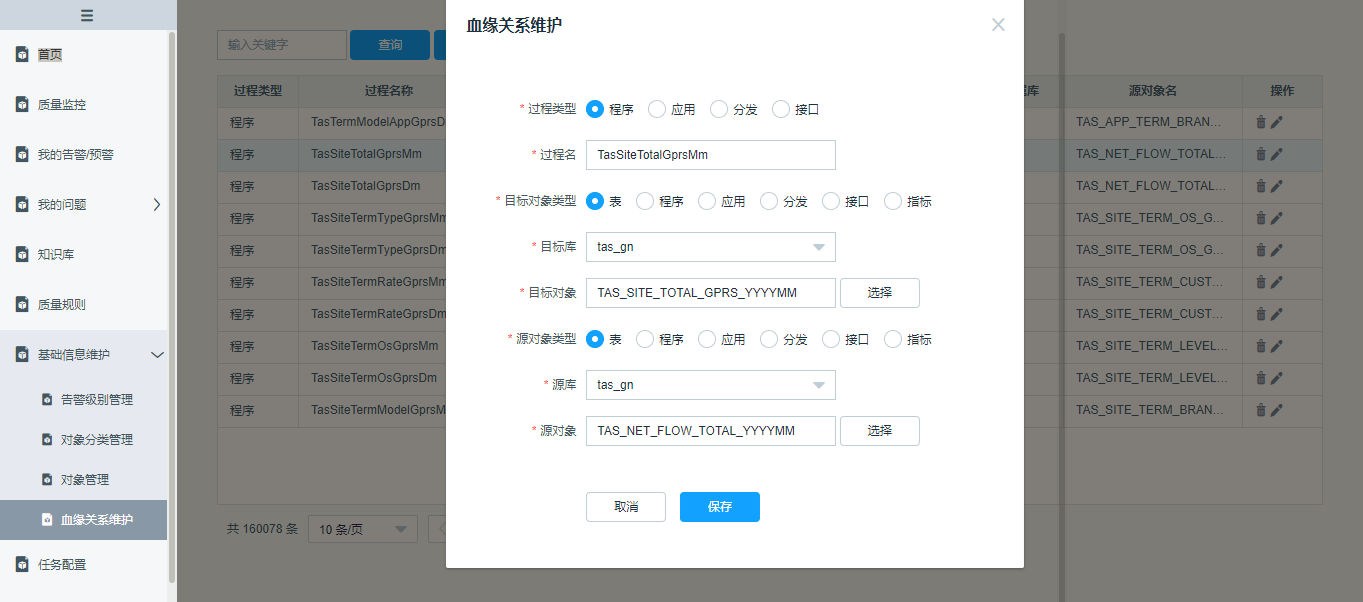
点击基础信息维护-血缘关系维护，可进入血缘关系维护界面，此界面暂时仅维护模型，程序之间的血缘关系



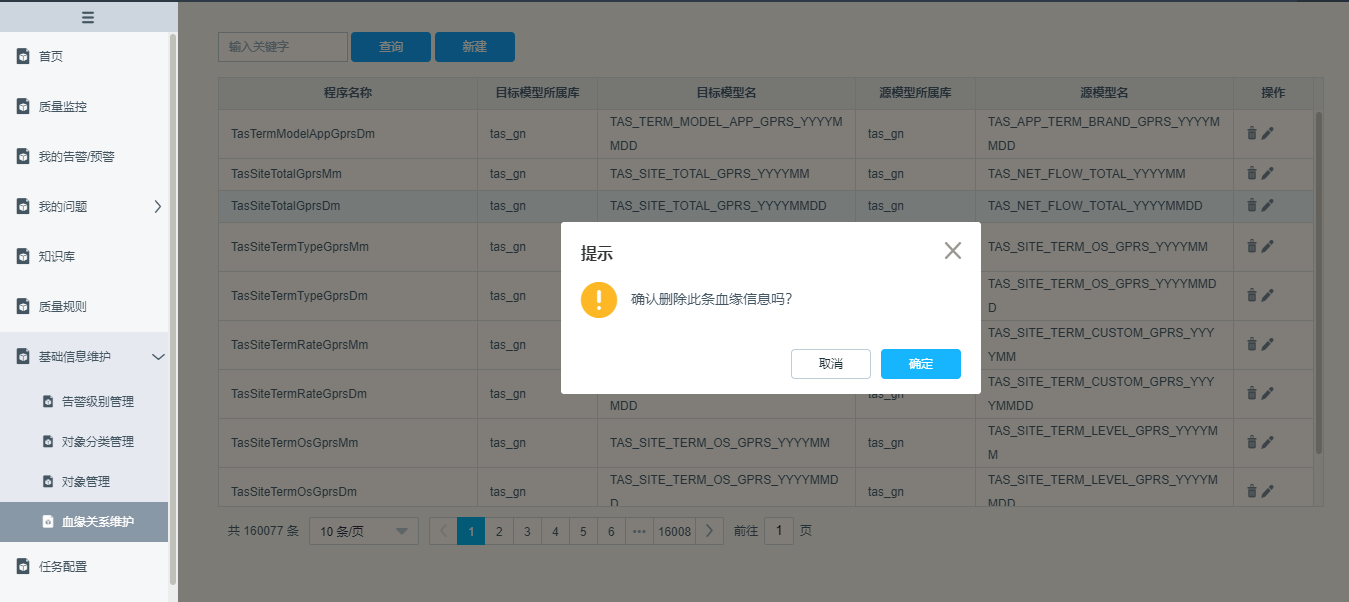
点击新建进行血缘新建



点击编辑，可编辑当前行的血缘关系



点击删除，可删除当前行的血缘关系



点击取消，则取消删除

## 任务配置

点击任务配置，进入任务配置管理页面，展示任务查询、任务新增、任务所有日志按钮，正常展示任务配置信息页面。



点击新增，弹出任务配置页，包括任务名称配置、任务组分配、程序类型、执行cron定时配置，填写完成正常保存成功新增的任务并支持自动调度。

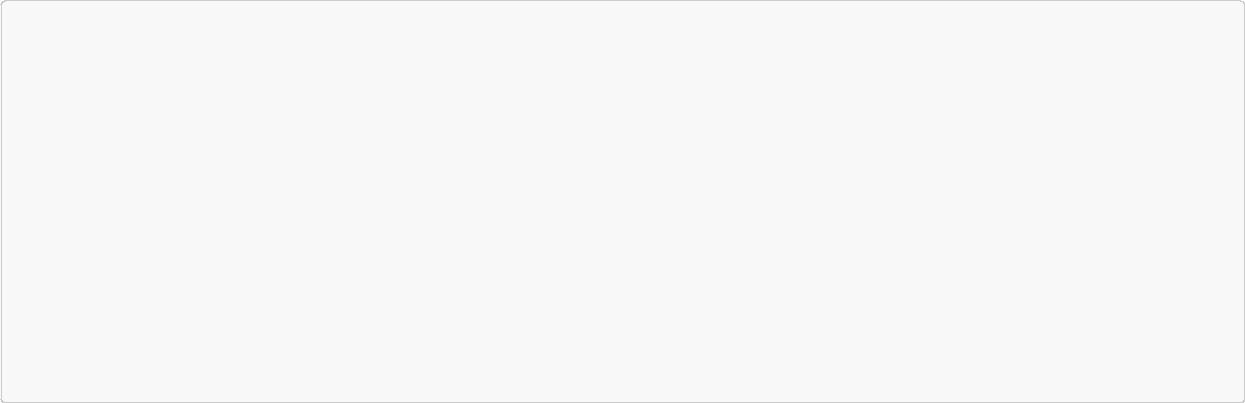


选中任务，点击暂停，正确暂停配置的任务，更改任务状态为已暂停。选中任务，点击继续，正确继续配置的任务，更改任务状态为正常并正常执行未完成的调度任务。选中任务，点击删除，成功删除配置的任务，并支持调度任务自动取消。

选中任务，点击日志，成功展示当前配置任务执行的任务日志信息



数据质量配置任务详细配置如下：



任务名称---新增时自定义填写，同步时同步为：数据运行质量

任务组----新增时自定义分组，同步时同步为：Default

程序类型----程序类型当前支持的为Java类。

Java类名---- com.asiainfo.dacp.datage.domain.dataquality.service.DqRunCheckJobService

调用方法----doCheck

方法参数----格式如下

小时周期：{"cycleType":"hour","differ":"-1"}，执行时间的上一个小时

日周期：{"cycleType":"day","differ":"-1"}，执行日期的头一天

月周期：{"cycleType":"month","differ":"-1"}，执行月份的头一月

年周期：{"cycleType":"year","differ":"-1"}，执行年份的头一年

也可以指定具体数据日期调度，比如这样：{"cycleType":"day","opTime":"20180423"}

备注：differ是相对于当前日期，往前推的周期

cron表达式---Quartz的cron表达式

数据质量日周期调度任务样例



数据质量月周期调度样例

