



大数据管理

产品文档





文档目录

产品概述

产品概述

功能特性

应用场景

产品优势

快速入门

步骤一开通数据管理

步骤二分配用户角色

步骤三进入数据管理

操作指南

新建数据源

标签管理

数据库的创建及发布

新建数据库

发布到测试

发布到生产

查看发布申请

发布申请撤回

发布审批

数据表的创建及发布

新建数据表

表发布到测试

表发布到生产

表查看发布申请

表发布申请撤回

表发布审批

数据权限管理

权限申请

申请记录

撤销申请

权限审批

数据授权

数据搜索

数据目录

业务元数据管理



分类管理

元模型管理

元数据维护

元数据关联配置

数据质量

规则配置

质量查询

质量告警

常见问题



产品概述

产品概述

最近更新时间: 2019-10-28 01:47:59

数据管理是一种全方位的企业数据治理服务，提供云端传统关系型数据库、MPP数据库、Nosql数据库、大数据平台数据仓库的统一管理，提供数据地图、数据质量、数据血缘分析能力。通过统一元数据管理解决金融、政府等行业数据孤岛、数据质量、数据合规使用问题。助力客户掌握数据全景，打通采数、识数、取数、用数全流程，实现数据管理无死角。

功能特性

最近更新时间: 2019-11-26 15:30:29

统一元数据管理

- 大数据云内部数据和外部数据源统一元数据管理，打通整体元数据，可实现平台内的数据和外部数据源通过统一SQL查询引擎统一进行查询的功能，统一进行数据访问权限管理等。支持Hive、MPP、Oracle、MySQL、ES、HBase、kafka、Redis、图数据库、时序数据库、FTP、对象存储等多种异构数据源的数据统一管理。
- ### 库表管理

- 支持创建、删除、修改数据库，支持创建、修改、删除表等操作。支持库表模型设计，库表可由设计态发布到测试环境和生产环境，支持一个库/表同时发布到多个数据源。
- 支持Hive、MySQL、Oracle、MPP、HBase、ES库表创建；支持kafka topic创建；支持对象存储Bucket创建。

权限管理

- 支持管理Hive、MPP、Oracle、MySQL、ES、HBase、kafka、图数据库、对象存储等数据源的数据权限，支持数据权限的在线申请和审批。

全方位数据地图

- 支持库表基本信息、分区、存储量分布等各类统计信息查看
- 支持以技术目录、业务目录、标签目录等多种方式查看元数据信息
- 支持按关键字全文检索所有类型元数据信息，支持按技术目录、业务目录、标签目录等多种高级方式筛选元数据

数据质量监控

- 支持批量设置数据质量检核规则，对数据处理流程全链路各个环节进行数据质量的监控管理。支持质量检核告警记录查询与管理。

数据血缘管理

- 从数据采集/同步，到数据ETL、统计分析、结果查询等数据产生的全链路对数据血缘进行分析，便于了解数据流链路情况。



数据热度分析

- 借助元数据管理、统一SQL查询等功能，可集中收集分析平台用户对各项数据的读写热度等情况，可以有效保障资源的合理使用。



应用场景

最近更新时间: 2019-11-26 15:30:29

数据发现与维护

- 用户可通过数据目录发现自己感兴趣的数据，查看具体数据库表详情，业务数据详情；
- 开发人员维护数据表结构、创建技术检核质量规则；
- 业务人员维护业务数据，创建业务检核质量规则；
- 有权限的用户可对具体的表、字段、业务数据打标签。

数据权限审批

- 管理人员进行元数据变更审批，数据权限审批，质量规则发布审批。



产品优势

最近更新时间: 2019-11-26 15:30:29

统一元数据管理

- 多种异构数据源元数据同步与托管
- 在线库表模型设计与多源发布
- 灵活的业务元模型定义, 统一数据标准
- 全域数据权限统一管理
- 口径一致的元数据访问接口服务

快速数据识别与发现

- 数据资产一键检索定位
- 技术、业务、标签多角度构建数据资产目录
- 库、表、字段、元模型详情查看
- 标签、热度、质量指标快速定位有用数据
- 表级、字段级数据权限开放申请

企业级数据安全合规

- 严格的数据权限申请与审批流程
- 严格的元数据变更管理流程
- 逻辑层、物理层双层数据鉴权机制
- 数据密级管理与脱敏处理

增强企业级特性



- 全方位数据地图查看企业数据分布
- 字段级血缘助力数据回溯与影响分析
- 基于生命周期的数据智能调度



快速入门

步骤一开通数据管理

最近更新时间: 2019-10-28 01:53:00

将租户账号提供给运营人员进行数据管理服务开通。开通流程详见购买指南部分。



步骤二分配用户角色

最近更新时间: 2019-11-26 15:30:29

租户或系统管理员用户（由租户创建）新建数据管理相关用户，并分别授予数据管理相关角色。目前数据管理提供全局数据管理员、数据资产管理员、项目数据管理员三种管理角色。角色说明如下：

全局数据管理员：

全局数据管理员主要对元数据权限进行管理，拥有平台所有数据权限和数据管理组件的所有功能权限。数据源管理、审核数据库发布申请、业务元数据分类管理、业务元模型审批是全局数据管理员的专属功能。

数据资产管理员：

数据资产管理员主要对数据管理组件的业务数据进行管理，包括业务元模型管理、定义所管元模型业务元数据管理人员、审核业务元数据发布申请、管理数据标签、审核业务元数据关联关系（业务关联业务、业务关联技术）的发布申请。

项目数据管理员：

项目数据管理员主要对项目内的数据进行管理，包括审批本项目内技术元数据访问权限申请、审批本项目内技术元数据发布申请、数据库维护及发布申请、审核质量规则发布申请。

步骤三进入数据管理

最近更新时间: 2019-11-12 08:18:39

登录大数据云租户控制台后，点击左侧的【数据管理】菜单直接进入数据管理页面进行作业。

平台概览

资源管理

- 已购产品
- 资源组管理

项目空间

- 我的项目
- 所有项目

数据管理

运维中心

组织管理

- 个人中心
- 用户管理
- 群组管理

组织信息

用户数	群组数
29	5

[个人中心](#) [用户管理](#) [群组管理](#)

产品信息

已开通	未开通
8	0

[已购产品](#) [选购产品](#)

项目信息

我的项目	所有项目
51	51

最近使用的项目 [更多项目 >](#)

di_pro

- 流计算
- 数据采集
- 数据集成
- 离线计算
- 数据服务

proj713

- 数据采集
- 数据集成
- 离线计算
- 数据挖掘



通过点击右上方的【控制台首页】按钮可返回大数据云控制台的平台概览页。

The screenshot shows the 'Data Catalog' (数据目录) section of the 'Jinshan Big Data Cloud' (金山大数据云) interface. The left sidebar contains navigation options like '数据源管理', '元数据管理', and '数据目录'. The main area displays details for a table named 'di_oracle_table1' under the 'oracle' database. The '基本信息' (Basic Information) tab is active, showing fields like '表名' (Table Name), '表负责人' (Table Owner), '数据库类型' (Database Type), and '所属项目' (Project). A '数据标签' (Data Tag) section allows for tagging the table with various attributes like '商品属性' (Product Attribute) and '外观属性' (Appearance Attribute). A '权限申请' (Apply for Permissions) button is visible in the top right corner of the table details panel.

表名:	di_oracle_table1	创建时间:	2019-07-05 00:28:20
表负责人:	ccb_test	生命周期:	all
数据库类型:	oracle	真实名称:	DI_ORACLE_TABLE1
所属项目:	di_pro		

操作指南

新建数据源

最近更新时间: 2019-11-15 06:52:55

登录大数据云平台，点击【数据管理】，进入【数据源管理】页面

平台概览

资源管理

项目空间

数据管理

运维中心

组织管理

个人中心

组织信息

产品信息

项目信息

用户数 34

群组数 15

已开通 8

未开通 0

我的项目 61

所有项目 61

最近使用的项目

lir_0702

流计算 数据采集 数据集成 离线计算 数据挖掘 数据服务

点击新增【数据源按钮】

数据管理 > 数据源管理

新增数据源 批量删除

数据源类型: 全部

输入数据源名称查询

数据源名称	数据源类型	测试环境链接信息	生产环境链接信息	管理模式	创建时间	数据源描述
ss	elasticsearch	http访问地址: http://10.69.66.215:9200	http访问地址: http://10.77.0.24:9200	托管元数据	2019-07-16 21:46:00	测试和生产都成功

点击选择需要添加的数据源



说明：数据管理产品支持以下多种数据源

序号	数据名称	托管模式	数据源级别	备注
1	Oracle	托管	数据库	自动注册数据库，支持用户在该库下创建数据表
2	MySQL	托管	实例/数据库	实例级支持用户创建数据库和在库下创建数据表；数据库级自动注册数据库，支持用户在该库下创建数据表。
3	MPP	托管	数据库	支持用户创建模式，支持用户在模式下创建数据表
4	Hive	托管	实例/数据库	实例级支持用户创建数据库和在库下创建数据表；数据库级自动注册数据库，支持用户在该库下创建数据表。
5	HBase	托管	实例	支持用户创建命名空间（类似数据库）和在命名空间下创建表
6	kafka	托管	实例	支持用户创建topic，平台自动创建相应的库
7	ES	实例	实例	支持用户创建索引（类似数据库）和在索引下创建类型（类似数据表）
8	Redis	非托管	实例	后台自动创建一个库
9	ArangoDB	托管	实例	支持用户创建数据库和数据表（document和edge）



10	FTP	非托管	实例	不支持创建任何文件
11	COS	托管	实例	支持创建bucket（类似数据库）和在bucket下创建数据交换接口（类似数据表）

填写配置（以Hive为例，点击选择Hive，填写新增Hive数据源对话框中的配置）。![]

(<http://imgxxfb.yun.ccb.com//raw/2af19c3d7348280c3309bb2325c8d2d8.png>) ![]

(<http://imgxxfb.yun.ccb.com//raw/6511034f992540a44a95d6f9cfd11750.png>)

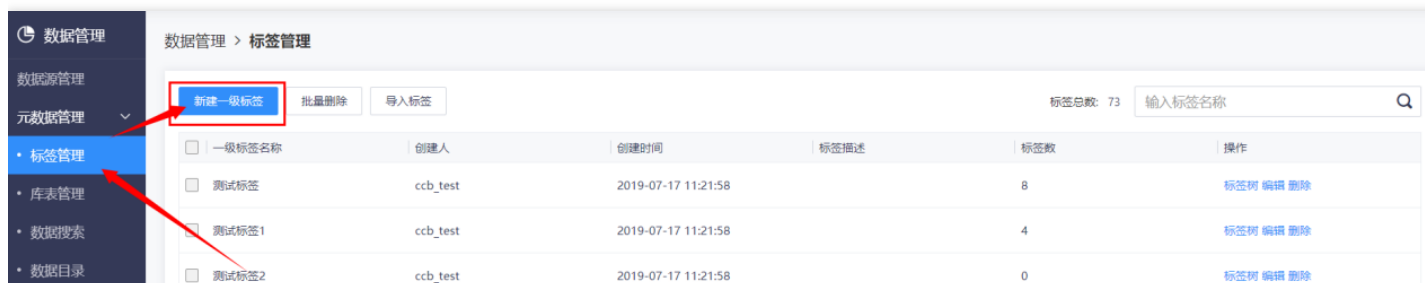
配置	说明
数据源名称	数据源名称仅支持英文、数字、下划线，且不能以数字和下划线开头，最大字符100字符
数据源描述	对数据源进行简单描述，不得超过500个字符。
地址	Hive的地址信息。
端口	Hive的端口号。
thrift地址	填写?thrift地址。
用户名	Hive登录用户名。
密码	Hive登录密码。

点击【连通性测试】按钮，测试通过后点击【完成】。

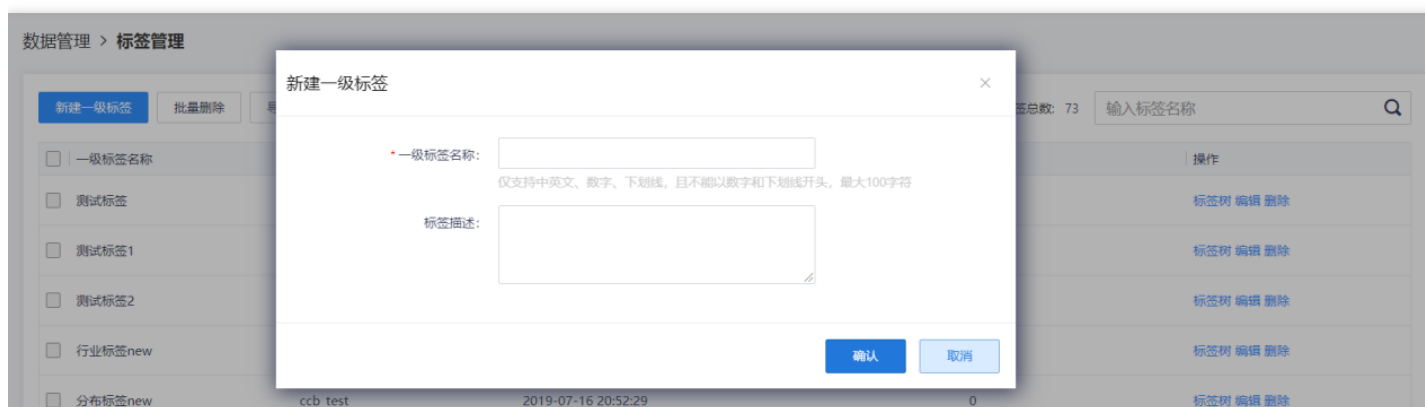
标签管理

最近更新时间: 2019-11-12 07:43:26

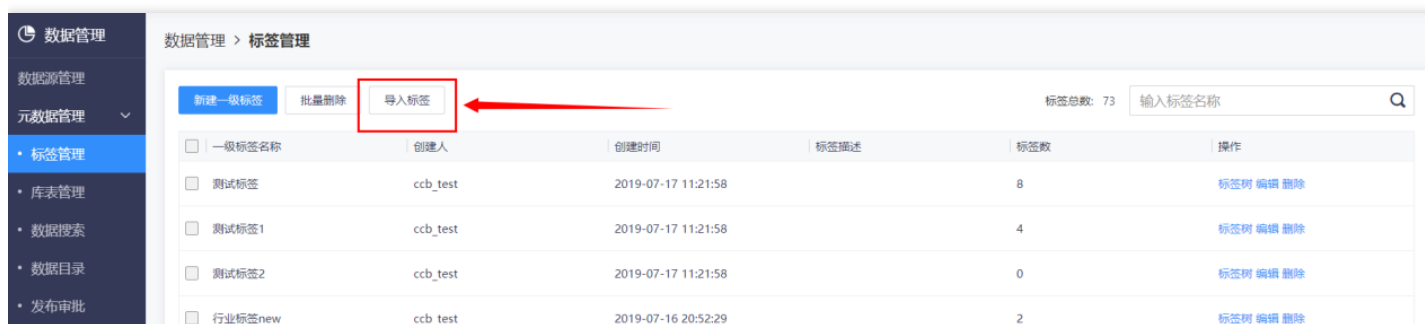
新建一级标签 点击选择【标签管理】，点击【新建一级标签】



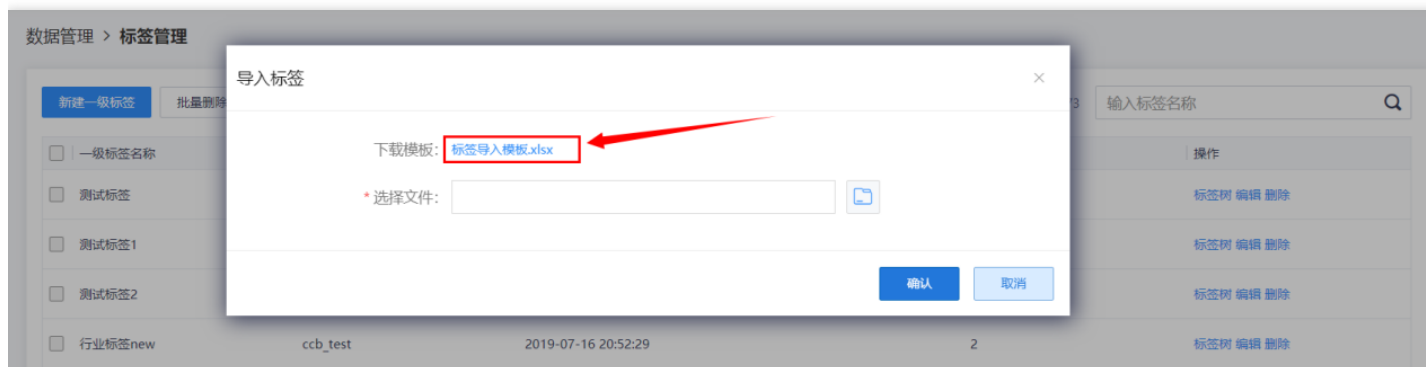
填写一级标签名称，点击【确认】。



说明：一级标签名称仅支持英文、数字、下划线，切不能以数字或下划线开头，最大100字符。导入标签 点击【导入标签】

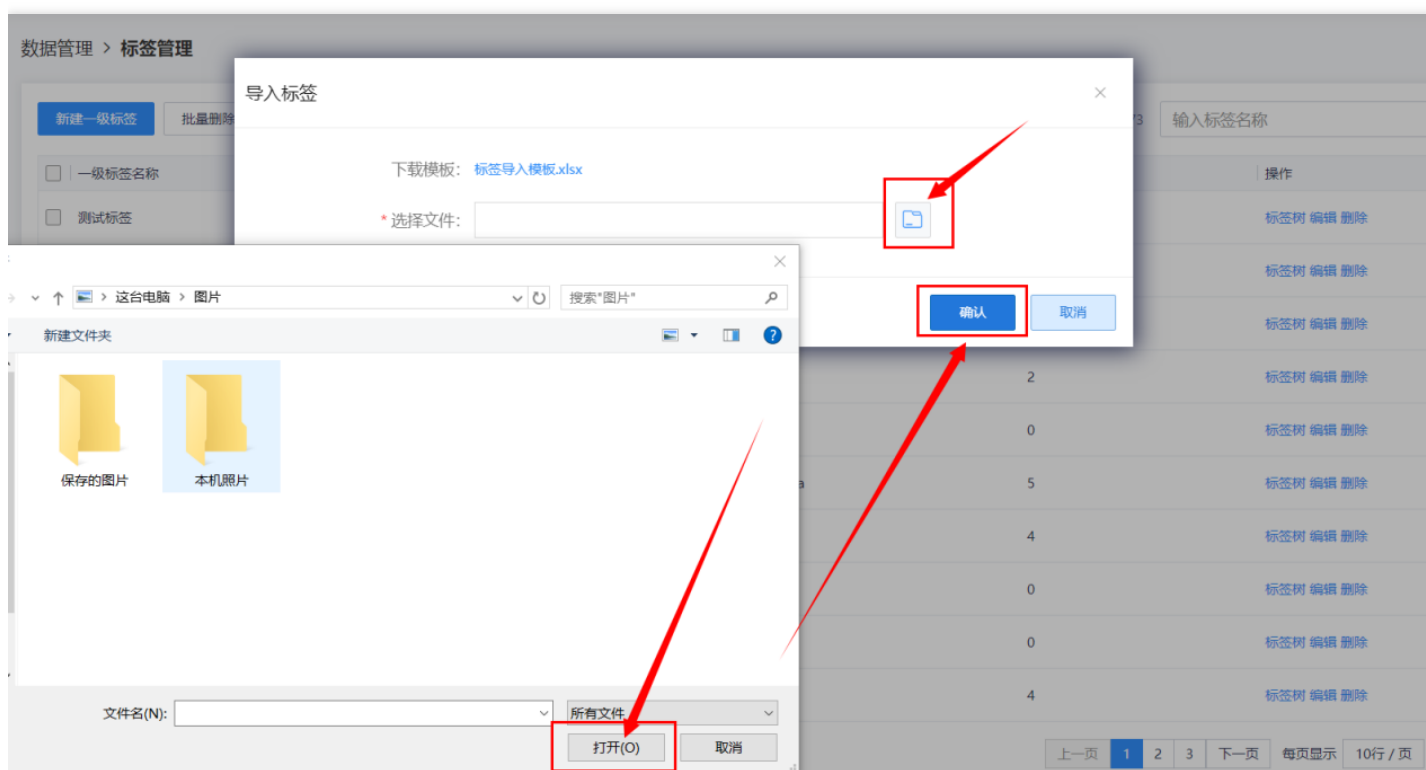


点击下载【标签导入模板.xlsx】，按照要求填写各级目录名称，点击文件夹图标在您的计算机上浏览找到填写好的模板进行导入，点击【确认】



下图为下载的模板文件。

	A	B	C	D	E	F	G	H	I	J	K	L
1	一级分类	二级分类	三级分类	四级分类	五级分类	六级分类						
2	行业标签										
3	领域标签											
4	分布标签											
5												
6												





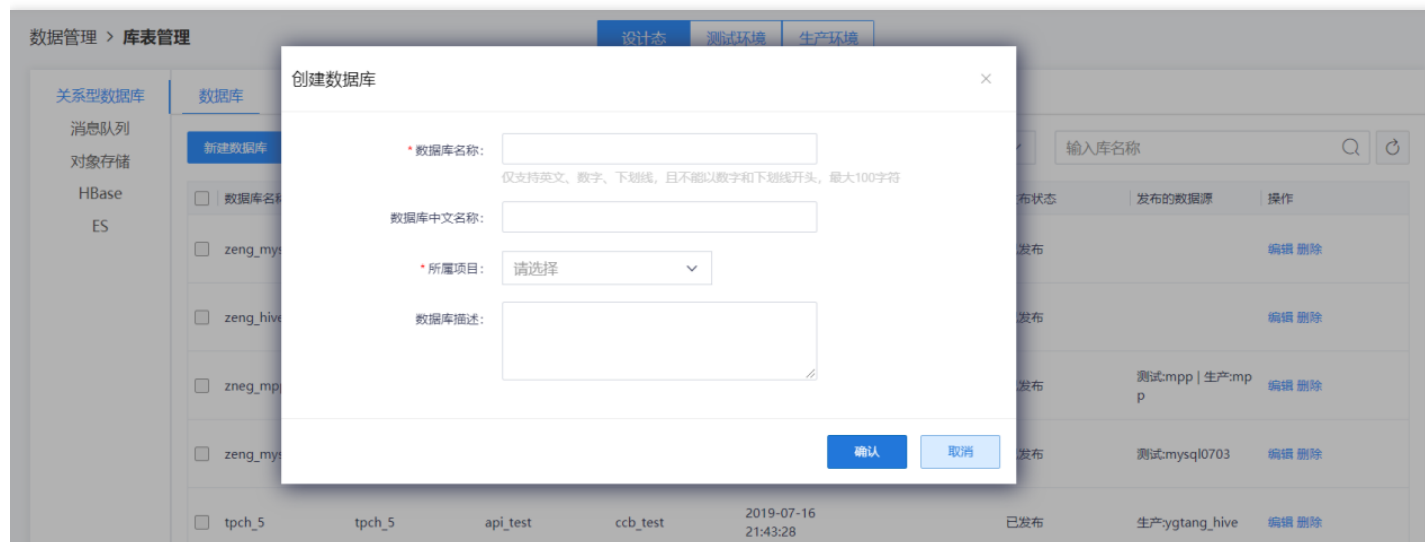
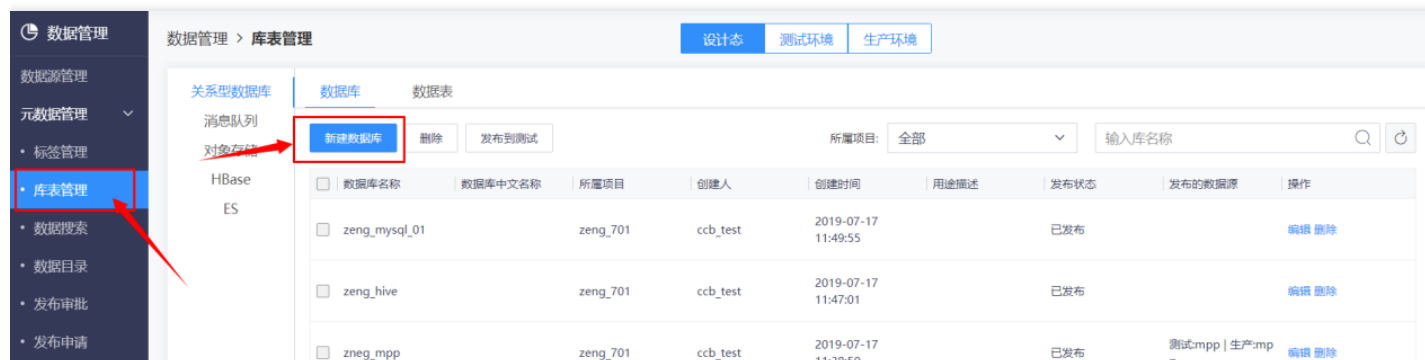
数据库的创建及发布

最近更新时间: 2019-10-28 02:03:01

新建数据库

最近更新时间: 2019-11-12 07:38:17

数据源创建完成后, 进入库表管理页面。 点击【库表管理】按钮, 点击【新建数据库】按钮



填写新建数据库对话框中的配置。

配置	说明
数据源名称	支持英文、数字、下划线,且不能以数字和下划线开头,最大100字符。
数据库中文名称	数据库中文名称, 最多输入100个字符。
所属项目	选择所属项目。
数据库描述	对数据库进行简单描述, 不得超过200个字符。

点击【确认】按钮, 提示创建成功。

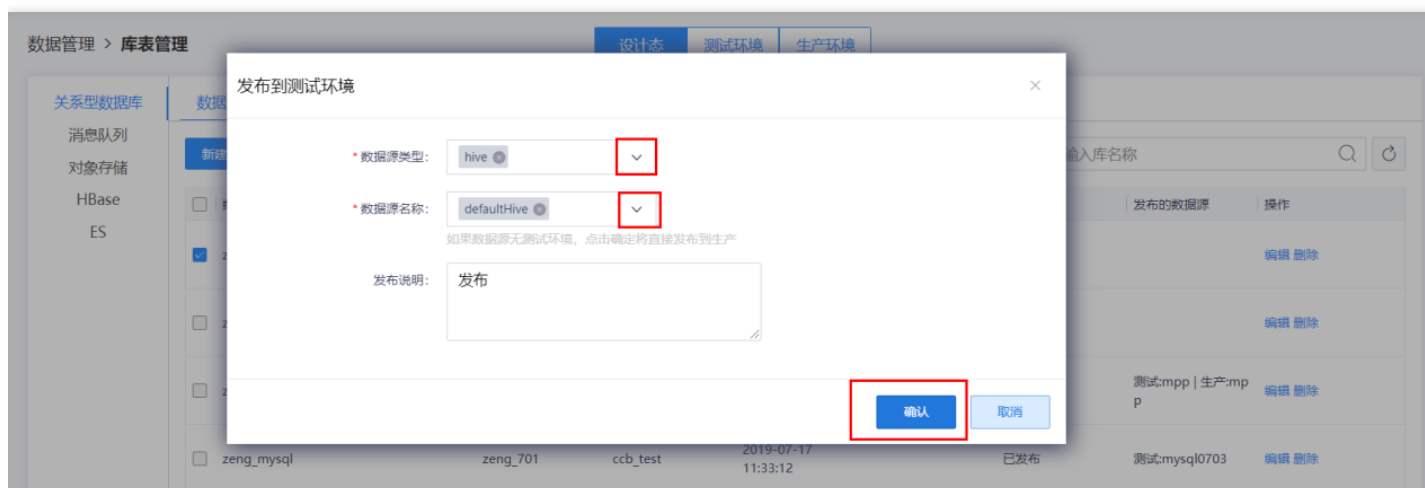
发布到测试

最近更新时间: 2019-11-12 07:38:17

新建数据库完成后, 点击【发布到测试】



选择数据源, 以Hive数据源为例, 点击【确认】按钮提示发布成功。





管理员审批通过后数据库将在【库表管理】中【生产环境】【我设计的库表】下显示。

数据管理 > 库表管理

设计态 测试环境 生产环境

我设计的库表 我有权限的表 项目账号的库表 项目有权限的表

数据源种类: [v] 数据源名称: [v] 所属项目: [v] 输入库名称 [Q] [刷新]

库名称	库中文名称	所属项目	数据源种类	数据源名称	创建人	创建时间	用途描述	操作
test_index_lir0717		test0716_lir	elasticsearch	ss	ccb_test	2019-07-17 15:56:45		显示表
estestdb717		lir_0702	elasticsearch	ss	ccb_test	2019-07-17 15:50:31		显示表
test_0717		test0717	elasticsearch	essssssssssssssssssssssss ssssssssssssssssssssssss ssssssssssssssssssssssss ssssssssssssssssssssssss ssssssss	ccb_test	2019-07-17 15:07:10		显示表



查看发布申请

最近更新时间: 2019-11-12 07:32:33

点击【发布申请】按钮可以查看发布的申请号。

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回
3990	2019-07-17 15:07:07	lir_0702	查看	撤回

点击【申请号】可以查看发布明细

申请号	申请时间	申请项目	流程状态	操作
3992	2019-07-17 15:08:58	lir_0702	查看	撤回

返回 申请编号 3992 发布明细

序号	发布对象	发布类型	发布内容
2777	test_tb_lir0717	update	oracle0702数据源下修改: test_tb_lir0717



点击【已结束的申请】查看已结束的申请记录

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 **已结束的申请**

所属项目: 全部 申请状态: 全部 请选择日期

申请号	申请时间	申请项目	处理人	处理时间	流程状态
3995	2019-07-17 15:30:50	api_test	ccb_test	2019-07-17 15:30:50	查看
3993	2019-07-17 15:13:09	test0717	ccb_test	2019-07-17 15:13:09	查看
3990	2019-07-17 15:07:53	test0717	ccb_test	2019-07-17 15:07:53	查看
3988	2019-07-17 15:06:23	lir_0702	ccb_test	2019-07-17 15:06:24	查看
3986	2019-07-17 15:05:30	lir_0702	ccb_test	2019-07-17 15:05:33	查看
3984	2019-07-17 14:49:10	test_0716	ccb_test	2019-07-17 14:49:11	查看
3983	2019-07-17 14:48:46	test_0716	ccb_test	2019-07-17 14:48:46	查看

发布申请撤回

最近更新时间: 2019-11-12 07:21:13

对于已经提交的发布到测试或生产的申请可以进行撤回 点击【发布申请】

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回

点击选择【测试环境】或【生产环境】查看相应申请。

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

数据管理 > 发布申请

数据空间 测试环境 生产环境

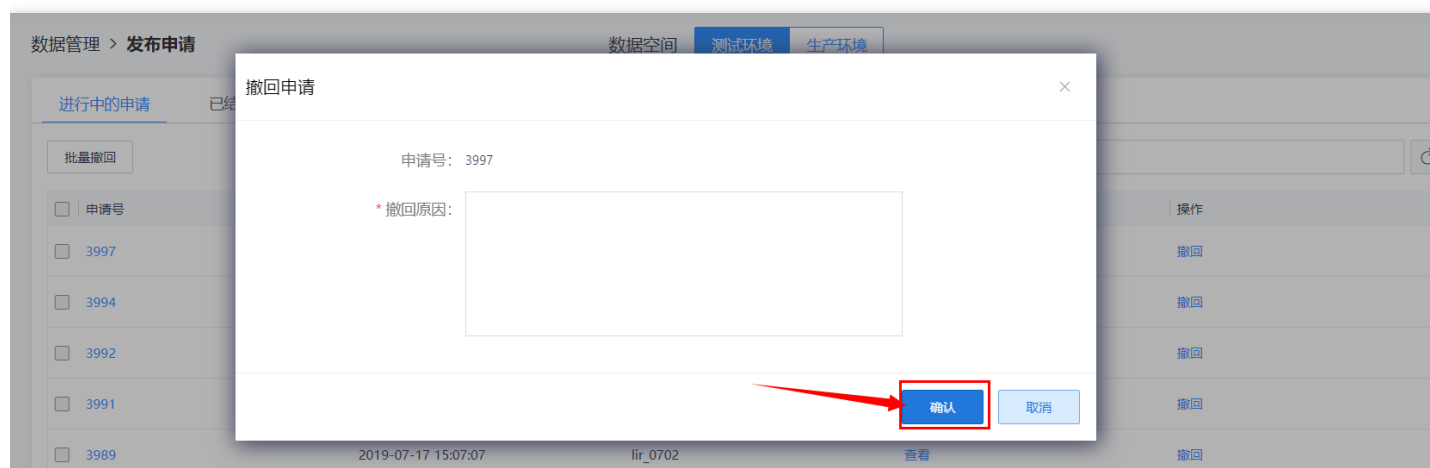
进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回



点击【撤回】，输入撤回原因，点击【确认】。



发布审批

最近更新时间: 2019-11-12 06:52:49

管理员进入【数据管理】页面，点击【发布审批】。

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回
3992	2019-07-17 15:08:58	ccb_test	lir_0702	查看	通过 驳回
3991	2019-07-17 15:08:44	ccb_test	test0717	查看	通过 驳回
3989	2019-07-17 15:07:07	ccb_test	lir_0702	查看	通过 驳回
3987	2019-07-17 15:05:48	ccb_test	test0717	查看	通过 驳回
3985	2019-07-17 14:58:09	ccb_test	lir_0702	查看	通过 驳回

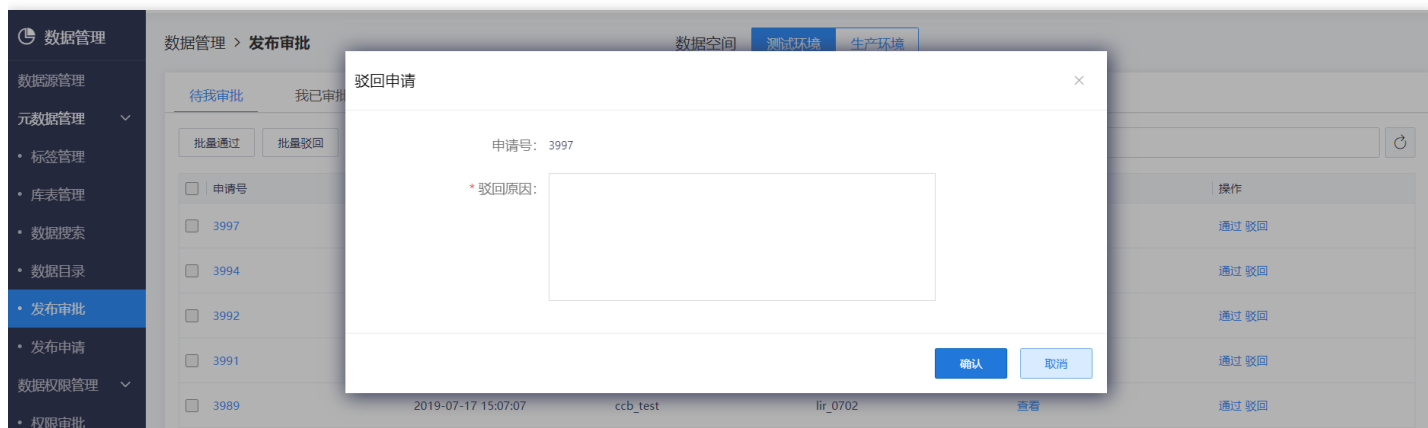
选择【测试环境】或【生产环境】查看环境下的待审批内容。

【待我审批】模块，查看并审批其项目下待审批记录，包括表申请

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回
3992	2019-07-17 15:08:58	ccb_test	lir_0702	查看	通过 驳回
3991	2019-07-17 15:08:44	ccb_test	test0717	查看	通过 驳回
3989	2019-07-17 15:07:07	ccb_test	lir_0702	查看	通过 驳回
3987	2019-07-17 15:05:48	ccb_test	test0717	查看	通过 驳回

点击【申请号】查看详细信息，点击【查看】可以查看流程状态。

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回



点击【我已审批】菜单，查看已经处理过的申请记录，包括表。





数据表的创建及发布

最近更新时间: 2019-10-28 02:11:41

新建数据表

最近更新时间: 2019-11-12 06:52:49

数据库创建完成并审批通过后, 点击【库表管理】中【数据表】选项卡



点击【新建数据表】按钮



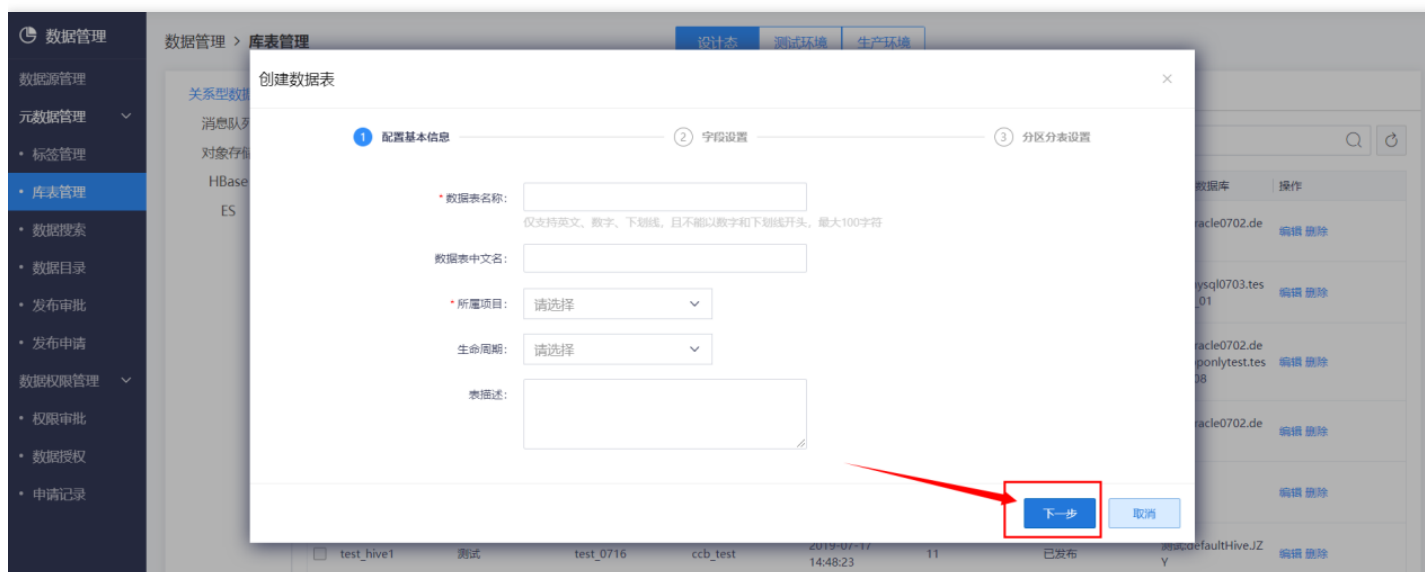
填写创建数据表配置信息



配置	说明
----	----

数据表名称	仅支持英文、数字、下划线，且不能以数字和下划线开头，最大100字符。
数据表中文名称	表的中文名称。
所属项目	当前表所处的项目。
描述	当前表的简要说明。
生命周期	支持一个月、一季度、一年、三年、永久、自定义四种选项。

单击【下一步】按钮






填写新建表页面中填写字段设置。



配置	说明
字段名称	字段英文名，由字母、数字、下划线组成。
字段类型	数据类型（字符串、整数、浮点数、日期、时间、日期时间）。



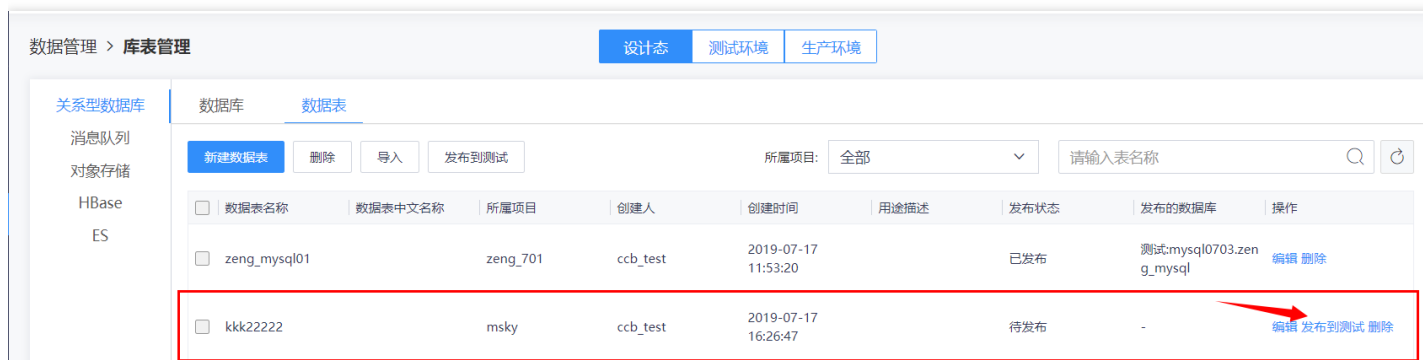
字段长度	字段长度设置。
是否为主键	是或者否
字段脱敏	无、默认脱敏法
字段描述	字段详细描述。
操作	上移、下移、删除。

单击【下一步】按钮  进行分区设置  点击【完成】按钮  提示创建成功，并且列表自动刷新展示刚刚新建的表。

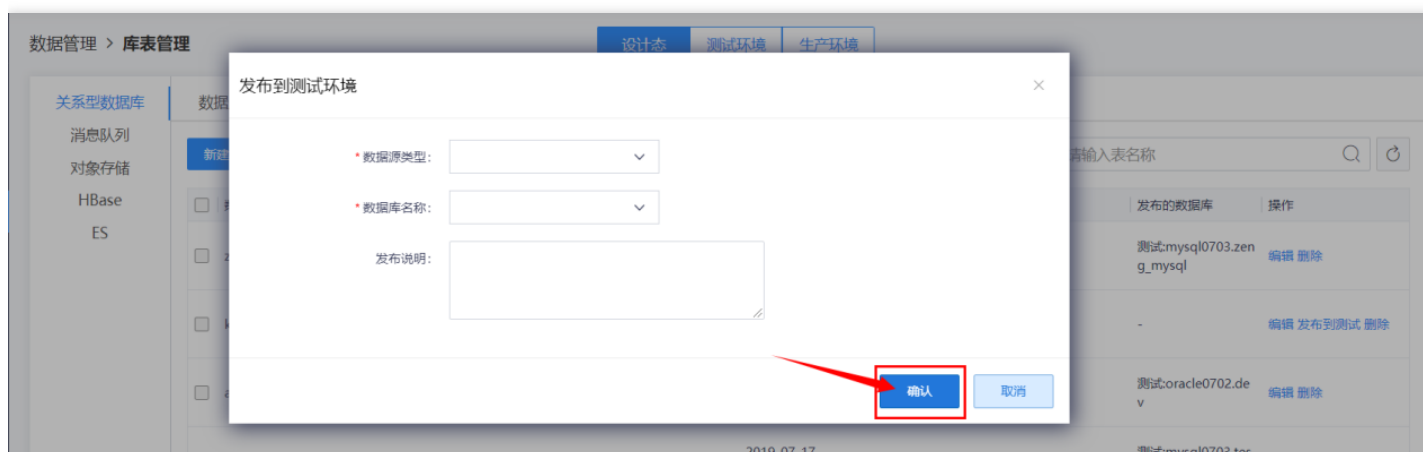
表发布到测试

最近更新的时间: 2019-11-12 06:52:49

发布到测试 新建数据表完成后, 在列表点击【发布到测试】按钮



选择数据源, 数据库, 点击【确认】按钮

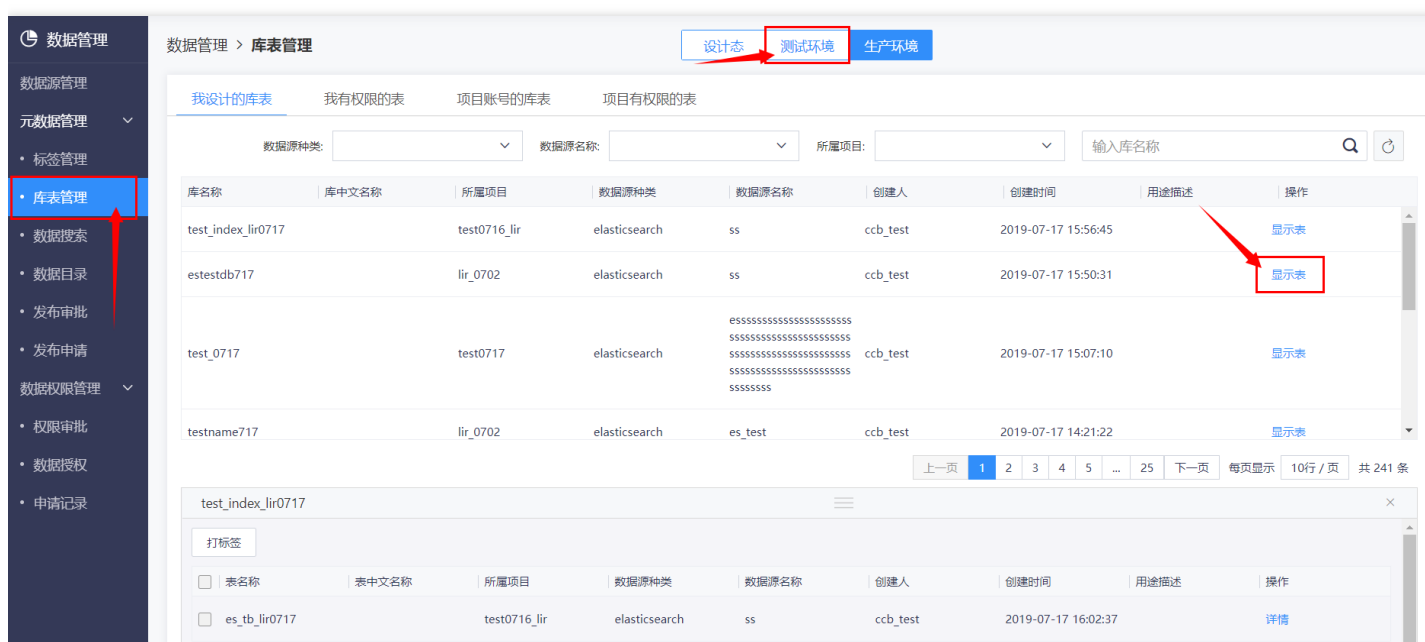


提示发布成功。

点击【确认】按钮，提示发布成功。



管理员审批通过后，在【库表管理】中【生产环境】下【我设计的库表】点击【显示表】显示





表查看发布申请

最近更新时间: 2019-11-12 06:52:49

点击【发布申请】按钮可以查看发布的申请号。

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回
3990	2019-07-17 15:07:07	lir_0702	查看	撤回

点击【申请号】可以查看发布明细

申请号	申请时间	申请项目	流程状态	操作
3992	2019-07-17 15:08:58	lir_0702	查看	撤回

返回 申请编号 3992 发布明细

序号	发布对象	发布类型	发布内容
2777	test_tb_lir0717	update	oracle0702数据源下修改: test_tb_lir0717



点击【已结束的申请】查看已结束的申请记录

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 **已结束的申请**

所属项目: 全部 申请状态: 全部 请选择日期

申请号	申请时间	申请项目	处理人	处理时间	流程状态
3995	2019-07-17 15:30:50	api_test	ccb_test	2019-07-17 15:30:50	查看
3993	2019-07-17 15:13:09	test0717	ccb_test	2019-07-17 15:13:09	查看
3990	2019-07-17 15:07:53	test0717	ccb_test	2019-07-17 15:07:53	查看
3988	2019-07-17 15:06:23	lir_0702	ccb_test	2019-07-17 15:06:24	查看
3986	2019-07-17 15:05:30	lir_0702	ccb_test	2019-07-17 15:05:33	查看
3984	2019-07-17 14:49:10	test_0716	ccb_test	2019-07-17 14:49:11	查看
3983	2019-07-17 14:48:46	test_0716	ccb_test	2019-07-17 14:48:46	查看

表发布申请撤回

最近更新时间: 2019-11-12 06:52:49

对于已经提交的发布到测试或生产的申请可以进行撤回 点击【发布申请】

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回

点击选择【测试环境】或【生产环境】查看相应申请。

数据管理 > 发布申请

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

数据管理 > 发布申请

数据空间 测试环境 生产环境

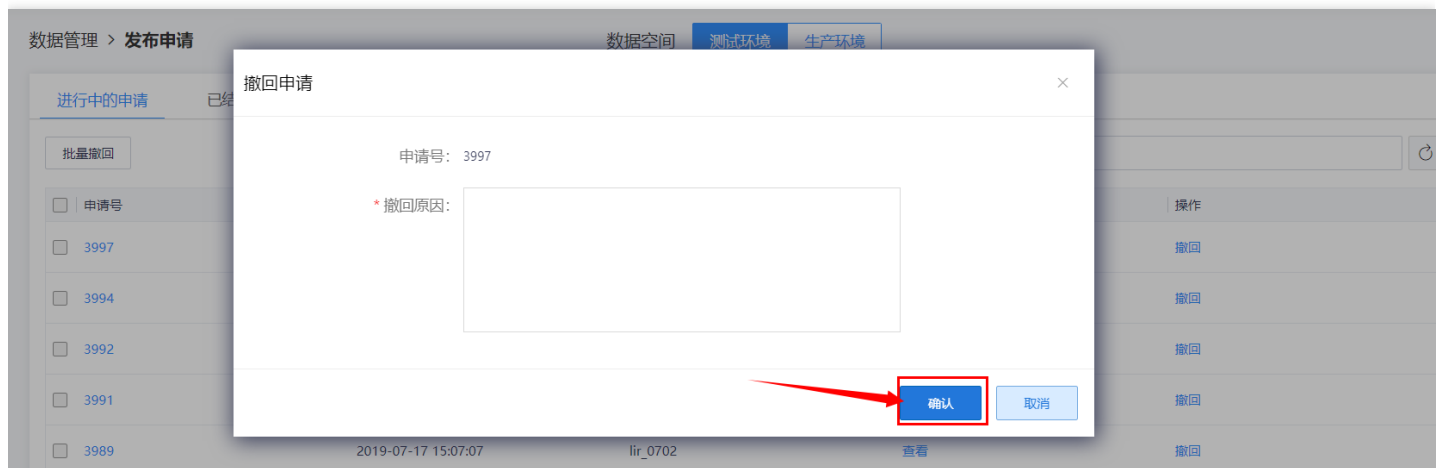
进行中的申请 已结束的申请

批量撤回 所属项目: 全部 请选择日期

申请号	申请时间	申请项目	流程状态	操作
3997	2019-07-17 15:40:42	lir_0702	查看	撤回
3994	2019-07-17 15:18:47	lir_0702	查看	撤回
3992	2019-07-17 15:08:58	lir_0702	查看	撤回
3991	2019-07-17 15:08:44	test0717	查看	撤回



点击【撤回】，输入撤回原因，点击【确认】。



表发布审批

最近更新时间: 2019-11-12 06:52:49

管理员进入【数据管理】页面，点击【发布审批】。

数据管理 > 发布审批

数据空间 [测试环境](#) [生产环境](#)

待我审批 我已审批

批量通过 批量驳回

申请人: 全部 所属项目: 全部 请选择日期

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回
3992	2019-07-17 15:08:58	ccb_test	lir_0702	查看	通过 驳回
3991	2019-07-17 15:08:44	ccb_test	test0717	查看	通过 驳回
3989	2019-07-17 15:07:07	ccb_test	lir_0702	查看	通过 驳回
3987	2019-07-17 15:05:48	ccb_test	test0717	查看	通过 驳回
3985	2019-07-17 14:58:09	ccb_test	lir_0702	查看	通过 驳回

选择【测试环境】或【生产环境】查看环境下的待审批内容。

数据管理 > 发布审批

数据空间 [测试环境](#) [生产环境](#)

待我审批 我已审批

【待我审批】模块，查看并审批其项目下待审批记录，包括表申请

数据管理 > 发布审批

数据空间 [测试环境](#) [生产环境](#)

待我审批 我已审批

批量通过 批量驳回

申请人: 全部 所属项目: 全部 请选择日期

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回
3992	2019-07-17 15:08:58	ccb_test	lir_0702	查看	通过 驳回
3991	2019-07-17 15:08:44	ccb_test	test0717	查看	通过 驳回
3989	2019-07-17 15:07:07	ccb_test	lir_0702	查看	通过 驳回
3987	2019-07-17 15:05:48	ccb_test	test0717	查看	通过 驳回

点击【申请号】查看详细信息，点击【查看】可以查看流程状态。

申请号	申请时间	申请人	申请项目	流程状态	操作
3994	2019-07-17 15:18:47	ccb_test	lir_0702	查看	通过 驳回

点击【驳回】按钮，弹出驳回申请框，输入审批意见点击【确认】按钮提示审批成功。



点击【我已审批】菜单，查看已经处理过的申请记录，包括表。





数据权限管理

最近更新时间: 2019-10-28 02:23:14

数据权限管理主要包括两方面：一租户或者数据管理员直接给项目或者用户添加表或者列级别的白名单，二其他角色的人员申请项目或者个人的权限，数据管理员审批通过、驳回或者通过后回收权限。说明：数据管理产品支持以下角色与权限

角色名称	角色权限
全局数据管理员	数据源管理
	审核数据库发布申请
	业务元数据分类管理
	业务元模型审批
数据资产管理员	业务元模型管理
	定义所管元模型业务元数据管理人员
	审核业务元数据发布申请
	管理数据标签
	审核业务元数据关联关系（业务关联业务、业务关联技术）的发布申请
业务人员	业务元数据新增、编辑、删除、发布
	权限范围内的业务元数据和技术元数据关系建立
	配置权限范围内的质量规则及发布申请
项目数据管理员	审批本项目内技术元数据访问权限申请
	审批本项目内技术元数据发布申请
	数据库维护及发布申请
	审核质量规则发布申请
开发人员	技术元数据创建、编辑、删除、发布
	权限范围内的业务元数据和技术元数据关系建立
	配置权限范围内的质量规则及发布申请

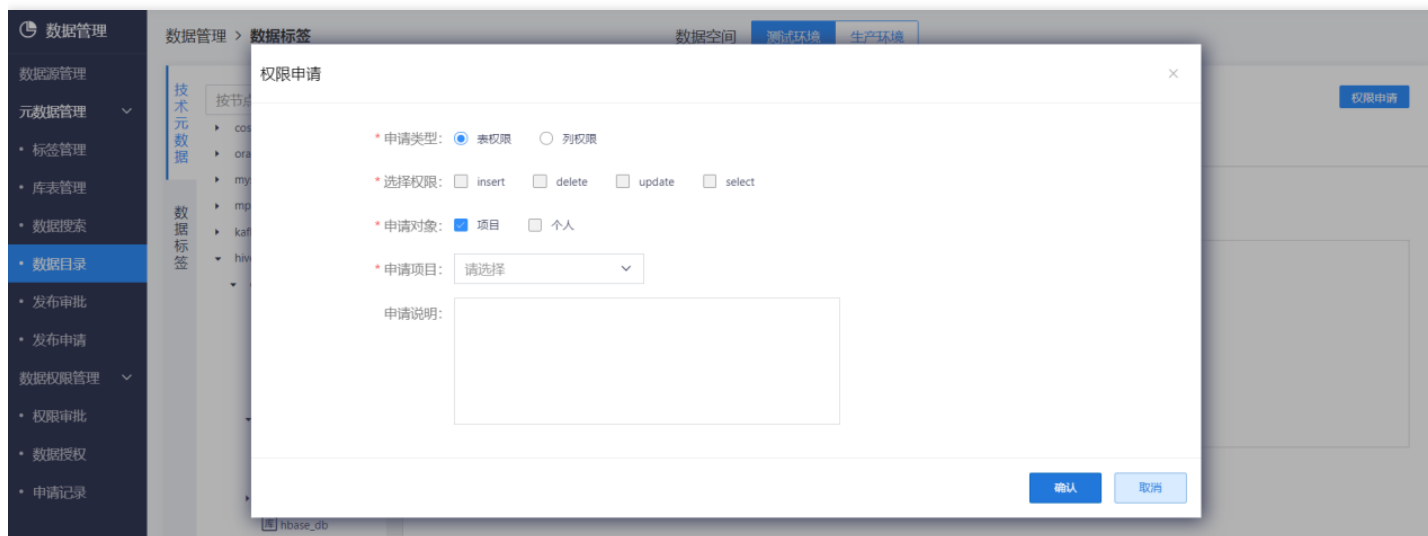


所有角色	申请技术元数据访问权限
	访问数据目录、数据搜索、数据详情页

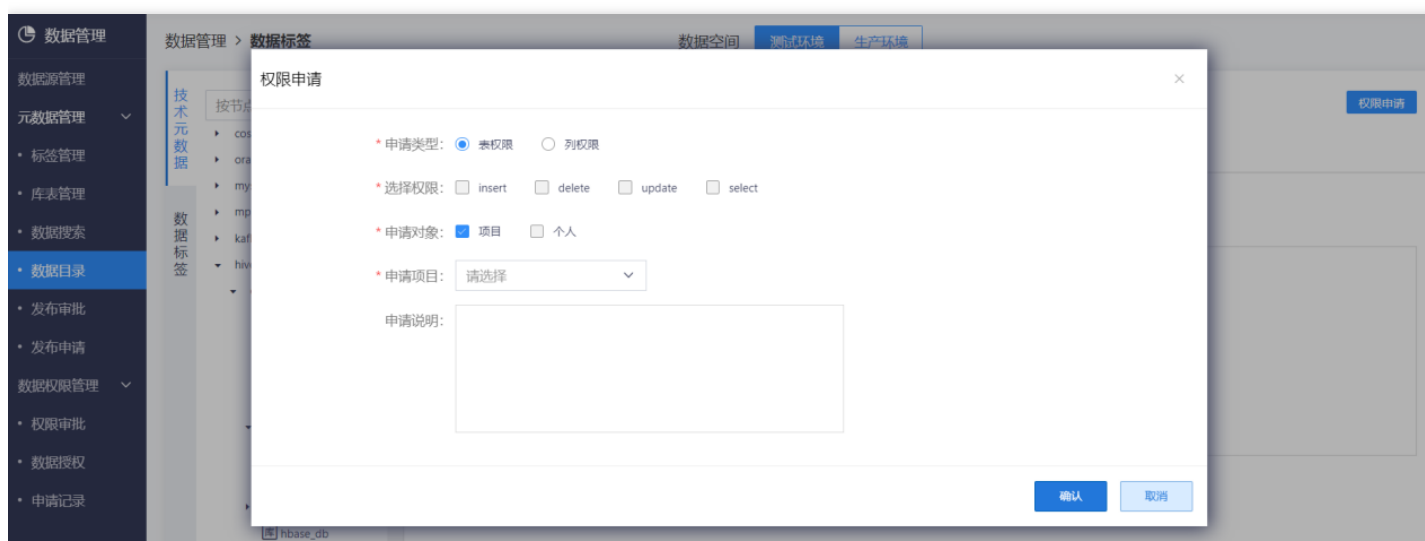
权限申请

最近更新时间: 2019-11-12 06:52:49

除数据管理员角色以外的其他角色的人员（包括：项目管理员，开发人员，运维人员，业务人员等）进入数据管理模块，点击【元数据管理】下面的【数据目录】选中数据源（以Hive为例），点击表名称，数据空间展示表详情。



点击【权限申请】按钮，弹出权限申请的页签，选择申请对象，选择权限，选择申请对象，选择申请项目，点击【确认】按钮，权限申请成功



说明：1. 申请类型为列权限时，申请对象只能为个人；申请类型为表权限时，申请对象可以为项目，个人或者项目



和个人。2. kafka类型数据源，没有列权限，只有表权限，不需要选择申请类型；kafka类型数据源的权限选择只有两个选项：insert和select。3. 申请个人和项目的表权限，会同时生成两条申请记录，审批记录也是两条。

申请记录

最近更新时间: 2019-11-12 03:43:04

进入数据管理模块【数据权限管理】->【申请记录】查看提交的权限申请。

数据管理 > 申请记录

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 请选择申请项目 请选择申请时间

申请号	申请时间	审批人	申请项目	流程状态	操作
1195	2019-07-17 10:44:09	ygtang6	api_test	查看	撤回
1169	2019-07-16 14:34:54	ygtang6	Test_0715	查看	撤回
1167	2019-07-16 10:09:08	ygtang6	Test_0715	查看	撤回
1156	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1157	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1158	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1159	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1160	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回

点击【申请号】查看记录详情。

数据管理 > 申请记录

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

批量撤回 请选择申请项目 请选择申请时间

申请号	申请时间	审批人	申请项目	流程状态	操作
1195	2019-07-17 10:44:09	ygtang6	api_test	查看	撤回

数据管理 > 申请记录 > 记录详情

序号	权限级别	数据源类型	数据库名	数据表名	数据列名	申请的权限
1	表	oracle	dev	oralce_line_item_1	all	UPDATE,SELECT,INSERT,DELETE

点击【查看】查看流程状态。



数据管理 > 申请记录 数据空间 测试环境 生产环境

[进行中的申请](#) [已结束的申请](#)

请选择申请项目 请选择申请时间

<input type="checkbox"/>	申请号	申请时间	审批人	申请项目	流程状态	操作
<input type="checkbox"/>	1195	2019-07-17 10:44:09	ygtang6	api_test	查看	撤回

点击【已结束的申请】查看已结束的申请记录。

数据管理 > 申请记录 数据空间 测试环境 生产环境

[进行中的申请](#) [已结束的申请](#)

申请时间:

<input type="checkbox"/>	申请号	申请时间	申请项目	审批人	流程状态	操作
<input type="checkbox"/>	1058	2019-07-13 19:01:45	0715_testreview	test_dba	查看	删除
<input type="checkbox"/>	1048	2019-07-13 16:18:40	0715_testreview	test_dba	查看	删除
<input type="checkbox"/>	1025	2019-07-13 15:25:07	lir_0702	test_dba	查看	删除
<input type="checkbox"/>	1026	2019-07-13 15:25:07	lir_0702	test_dba	查看	删除
<input type="checkbox"/>	1027	2019-07-13 15:25:07	lir_0702	test_dba	查看	删除

撤销申请

最近更新时间: 2019-11-12 03:43:04

登录账户，点击选择【数据权限管理】 -> 【申请记录】 -> 【进行中的申请】，找到想要撤回的申请号，点击右侧【撤回】。

数据管理 > 申请记录

数据空间 测试环境 生产环境

进行中的申请 已结束的申请

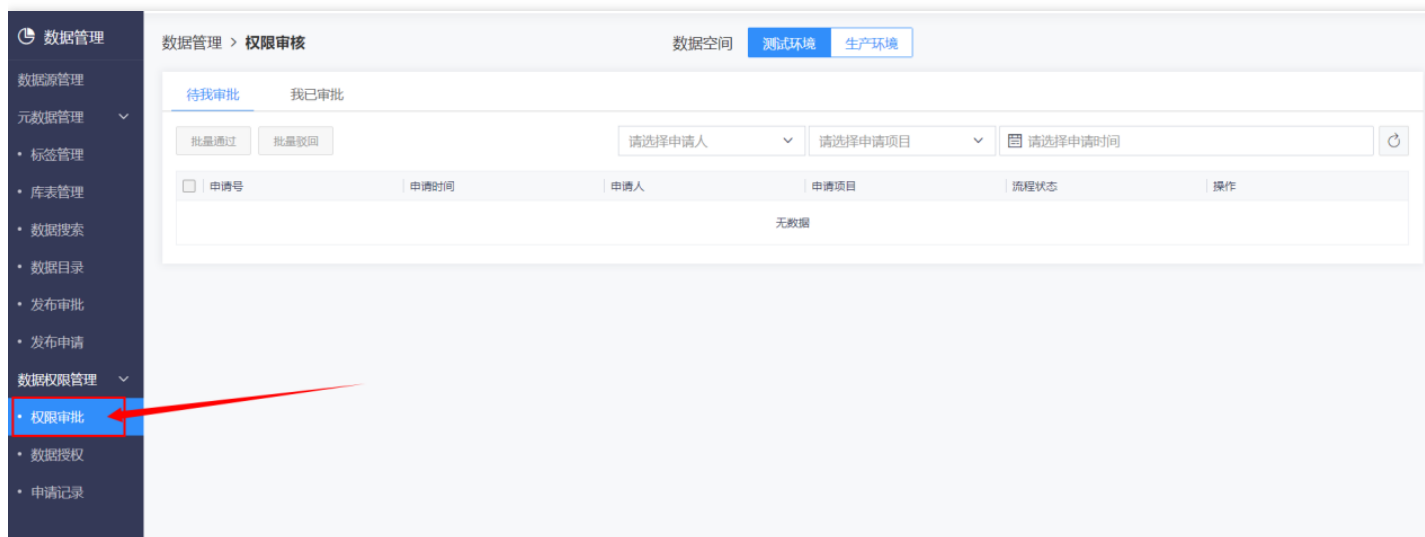
批量撤回 请选择申请项目 请选择申请时间

申请号	申请时间	审批人	申请项目	流程状态	操作
1195	2019-07-17 10:44:09	ygtang6	api_test	查看	撤回
1169	2019-07-16 14:34:54	ygtang6	Test_0715	查看	撤回
1167	2019-07-16 10:09:08	ygtang6	Test_0715	查看	撤回
1156	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1157	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1158	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回
1159	2019-07-15 14:38:48	ygtang4	Test_0715	查看	撤回

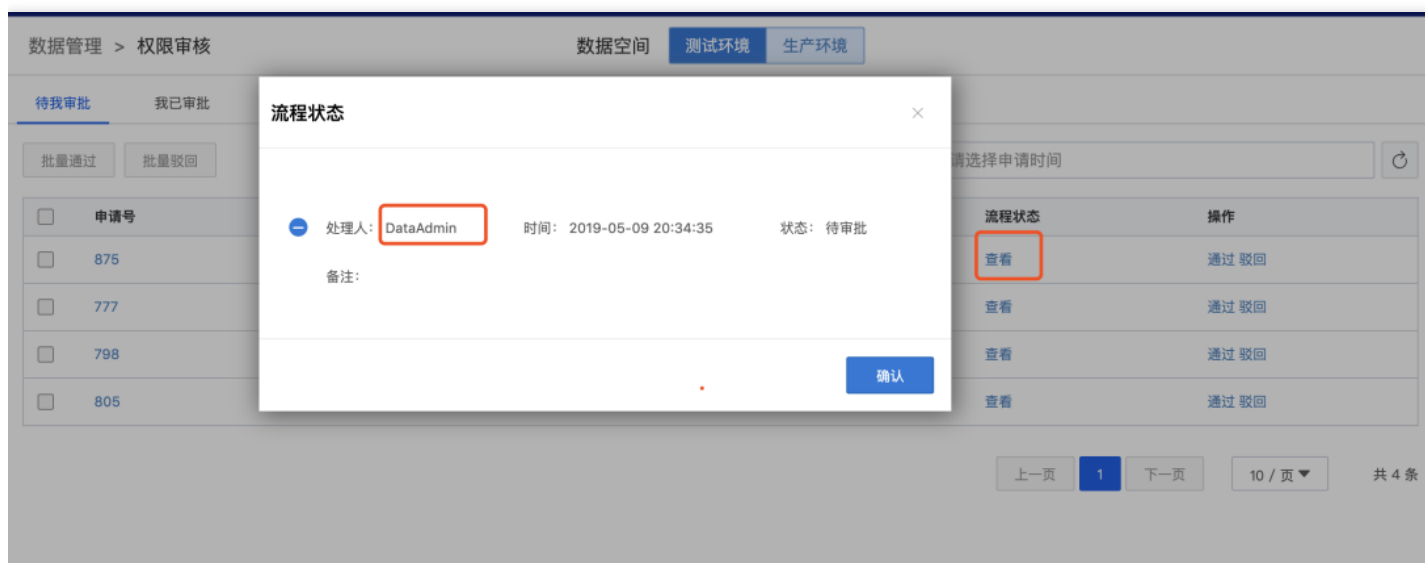
权限审批

最近更新时间: 2019-11-12 03:43:04

使用数据管理员账号登录进入数据管理模块，点击【数据权限管理】，展开下拉列表中点击【权限审批】



根据申请号选择提交的申请，点击【流程】可查看提交申请的用户，备注中显示提交原因。



点击【通过】（或者【驳回】）审批该申请。说明：通过申请，如果申请的时候选择的是项目，该表的项目白名单会新增一条数据，项目管理员对该数据表有对应的权限；如果选择的是个人，用户白名单新增一条数据，申请人对



该表有对应的权限。

数据管理 > 权限审核

数据空间 测试环境 生产环境

待我审批

批量通过

申请号

<input type="checkbox"/>	875	
<input type="checkbox"/>	777	
<input type="checkbox"/>	798	
<input type="checkbox"/>	805	

操作

通过 驳回
通过 驳回
通过 驳回
通过 驳回

1 下一页 10 / 页 共 4 条

通过申请

申请号: 875

审批意见: 通过

确认 取消

数据授权

最近更新时间: 2019-11-12 03:43:04

使用数据管理员或者租户的账号登录，进入【数据管理】模块，点击【数据权限管理】，展开的下拉框中点击【数据授权】，数据空间显示数据授权列表。

表名	表中文名称	所属项目	数据源种类	数据源名称	创建人	公开状态	白名单	操作
es_test_mzq_t1		api_test	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置
zeng_mysql01		zeng_701	mysql	mysql0703	ccb_test	未公开	查看	公开设置 白名单设置
es_tb_lir0717		test0716_lir	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置
test_es_sink_1		lir_0702	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置
testestbl1550		lir_0702	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置
cc		lir_0702	cos	cos0705	ccb_test	未公开	查看	公开设置 白名单设置
a2		lir_0702	oracle	oracle0702	ccb_test	未公开	查看	公开设置 白名单设置

白名单设置 选中需要授权的表，点击【白名单设置】按钮，弹出白名单设置标签，选择授权类型，选择权限，选择授权对象，授权项目和人都支持多选，点击【确认】

表名	表中文名称	所属项目	数据源种类	数据源名称	创建人	公开状态	白名单	操作
es_test_mzq_t1		api_test	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置

白名单设置

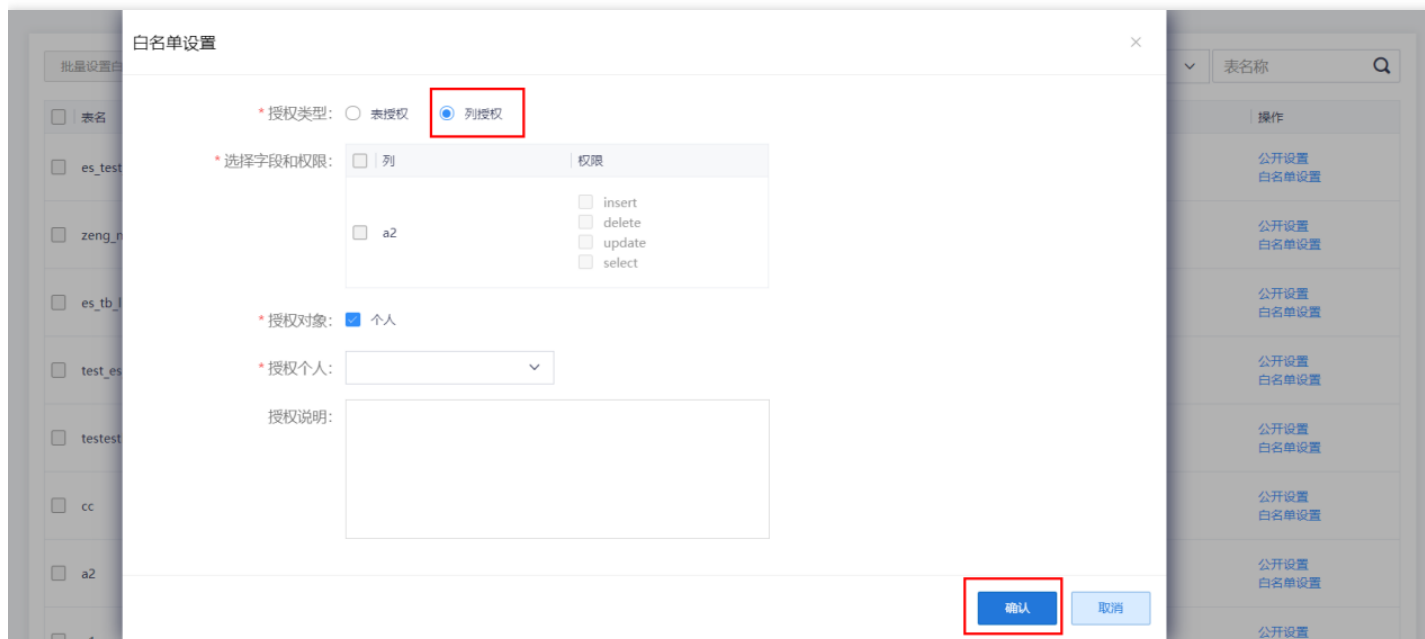
* 授权类型: 表授权 列授权

* 选择权限: insert delete update select

* 授权对象: 项目 个人

* 授权项目:

授权说明:



说明：授权类型为表权限时，授权对象可以为项目，个人或者项目和个人；授权对象为列权限时，授权对象只可以为个人；个人和项目均支持多选和搜索。授权成功后，点击【查看】按钮，可以看到授权的项目白名单和个人白名单。

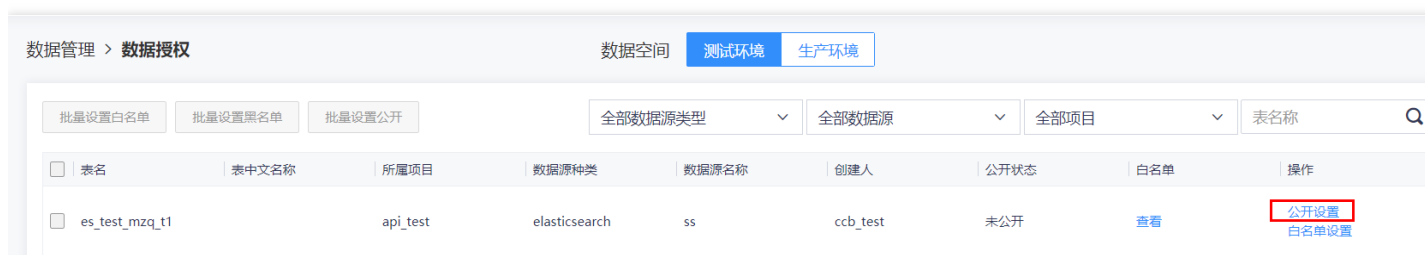
<input type="checkbox"/>	表名	表中文名称	所属项目	数据源种类	数据源名称	创建人	公开状态	白名单	操作
<input type="checkbox"/>	es_test_mzq_t1		api_test	elasticsearch	ss	ccb_test	未公开	查看	公开设置 白名单设置

说明：查看项目白名单和用户白名单，点击【新增白名单】按钮，可以给表增加白名单，操作类似白名单设置。

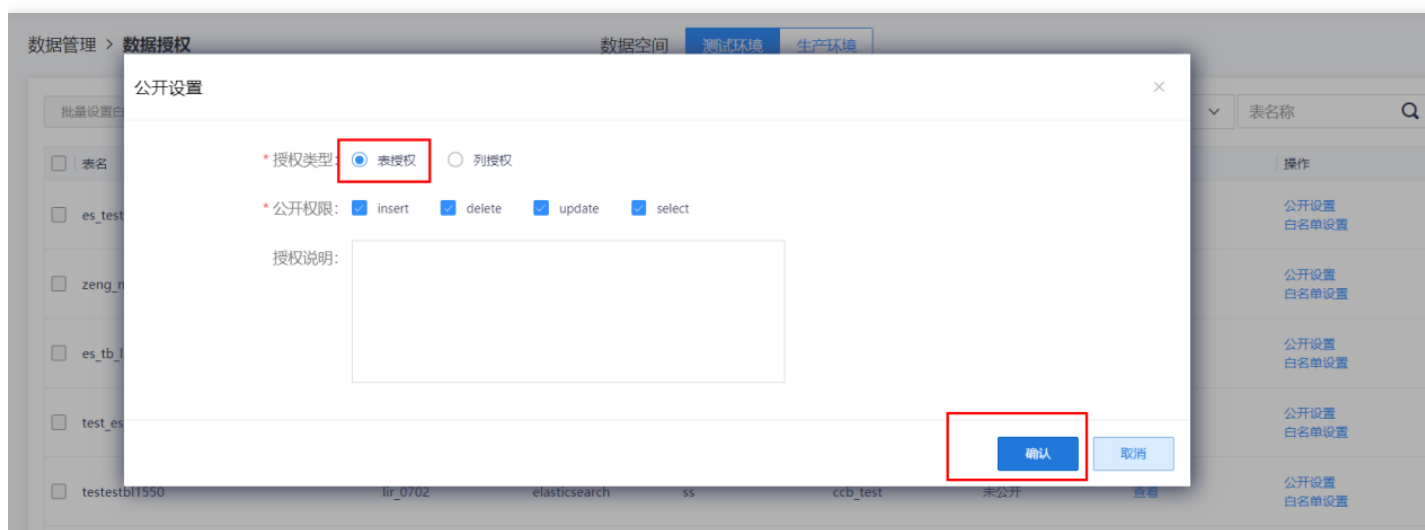




公开设置 进入【数据权限管理】->【数据授权】，找到表名点击【公开设置】。



选择授权类型和公开权限，点击【确认】。





数据管理 > 数据授权

数据空间 测试环境 生产环境

公开设置

* 授权类型: 表授权 列授权

* 选择字段和权限:

列	权限
<input type="checkbox"/> 列	<input type="checkbox"/> insert
<input type="checkbox"/> a2	<input type="checkbox"/> delete
	<input type="checkbox"/> update
	<input type="checkbox"/> select

授权说明:

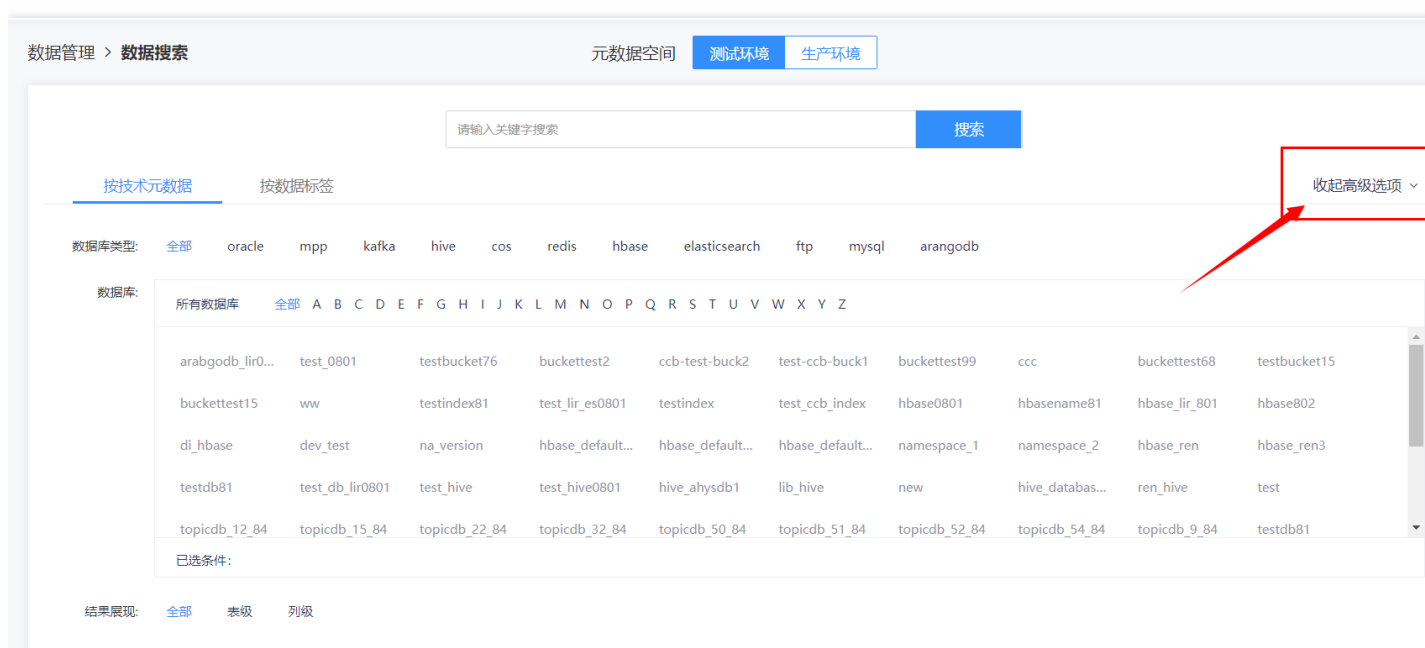
数据搜索

最近更新时间: 2019-11-12 03:43:04

点击【数据搜索】，在屏幕上放选择搜索内容所属的元数据空间（测试环境或生产环境），输入关键字即可进行简单搜索



如需提高搜索效率，点击【展开高级选项】显示如图菜单



如果选择【按技术元数据】，您可以通过限定数据库类型找到相应的数据库缩小搜索范围，如果选择【按数据标



签】，您可以按照【标签管理】处设置的标签进行查找，在屏幕下方选择结果的展现形式（全部，表级或列级）。

按技术元数据 按数据标签 收起高级选项

数据库类型: [全部](#) [oracle](#) [mpp](#) [kafka](#) [hive](#) [cos](#) [redis](#) [hbase](#) [elasticsearch](#) [ftp](#) [mysql](#) [arangodb](#)

数据库: 所有数据库 [全部](#) [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

arangodb_lir0...	test_0801	testbucket76	buckettest2	ccb-test-buck2	test-ccb-buck1	buckettest99	ccc	buckettest68	testbucket15
buckettest15	ww	testindex81	test_lir_es0801	testindex	test_ccb_index	hbase0801	hbasename81	hbase_lir_801	hbase802
di_hbase	dev_test	na_version	hbase_default...	hbase_default...	hbase_default...	namespace_1	namespace_2	hbase_ren	hbase_ren3
testdb81	test_db_lir0801	test_hive	test_hive0801	hive_ahysdb1	lib_hive	new	hive_databas...	ren_hive	test
topicdb_12_84	topicdb_15_84	topicdb_22_84	topicdb_32_84	topicdb_50_84	topicdb_51_84	topicdb_52_84	topicdb_54_84	topicdb_9_84	testdb81

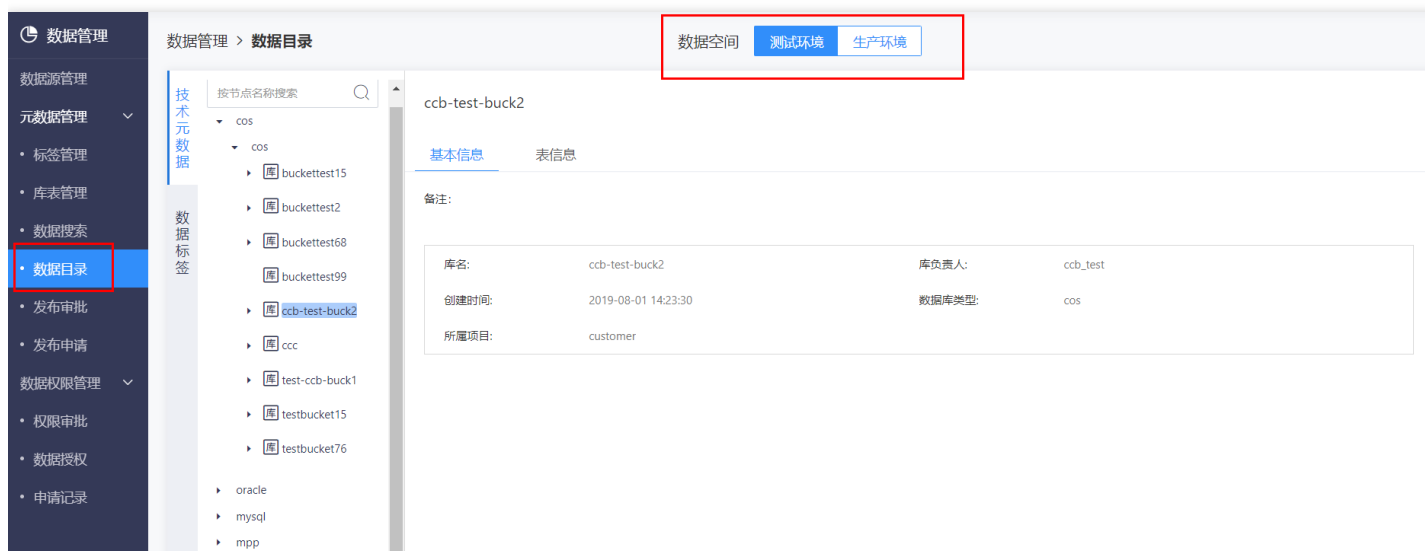
已选条件:

结果展现: [全部](#) [表级](#) [列级](#)

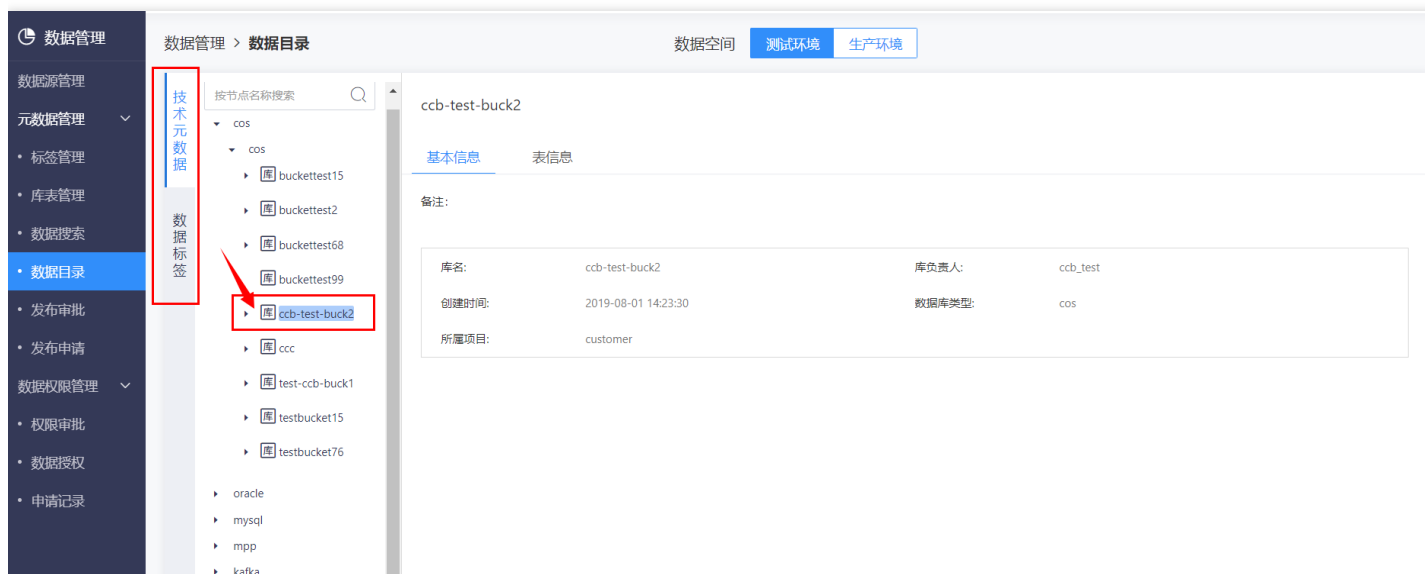
数据目录

最近更新时间: 2019-11-12 03:43:04

您可以通过【数据目录】功能按照技术元数据、数据标签、业务元数据查看数据库表信息和业务数据信息。点击【数据目录】，在页面上方选择数据空间。



在屏幕右侧选择【技术元数据】或【数据标签】查看相应目录，点击下箭头展开下一级，点击数据库名称查看基本信息。



点击【表信息】可以查看该数据库中各数据表的基本信息。



ccb-test-buck2

基本信息 **表信息**

批量权限申请

<input type="checkbox"/>	表名称	表中文名	项目所属	创建人	创建时间	操作
<input type="checkbox"/>	bbcsv		customer	ccb_test	2019-08-02 10:27:26	查看详情 权限申请
<input type="checkbox"/>	ccb_input2	测试2	customer	ccb_test	2019-08-01 14:38:32	查看详情 权限申请

上一页 1 下一页 每页显示 10行 / 页 共 2 条

点击左侧目录树的表，可以查看该数据表的详细信息，如基本信息、字段信息、数据预览、权限拓扑等。

数据管理 > 数据目录 数据空间 **测试环境** 生产环境

按节点名称搜索

- oracle
- mpp
- kafka
 - defaultKafka
 - topicdb_102_206
 - test_0904
 - topicdb_10_206
 - ccb_test_sou
 - kafka_source**
 - sink_php_jsor
 - source_php
 - source_php_j
 - source_php_r
 - test_json_moi

kafka_source_1

基本信息 字段信息 数据预览 权限拓扑

表中文名:		表负责人:	uat_test
数据库类型:	kafka	所属项目:	ss_test_php
创建时间:	2019-08-30 16:21:22	分区数量:	1
数据格式:	csv	csv的分割符号:	,
生命周期:	7天	备注:	
数据标签:	<input type="text" value="+"/>		

点击字段信息中具体某个字段后面的【详情】按钮，可以查看该字段的详细信息



数据管理 > 数据目录 数据空间 **测试环境** 生产环境

按节点名称搜索

- oracle
- mpp
- kafka
 - defaultKafka
 - topicdb_102_206
 - test_0904
 - topicdb_10_206
 - ccb_test_sou
 - kafka_source**
 - sin_k_php_isor
 - source_php
 - source_php_j
 - source_php_r

kafka_source_1 权限申请

基本信息 字段信息 数据预览 权限拓扑

字段名称	字段中文名	字段类型	字段密级	字段脱敏	字段描述	字段长度	操作
name	-	STRING	5	否	-	50	详情
age	-	NUMBER	5	否	-	9	详情
ctime	-	NUMBER	5	否	-	9	详情
cost	-	DECIMAL	5	否	-	10,2	详情

上一页 **1** 下一页 每页显示 10行 / 页 共 4 条

[返回](#) name

基本信息 权限拓扑

字段中文名:		所属表:	kafka_source_1
所属项目:	ss_test_php	创建时间:	2019-08-30 16:21:22
字段类型:	STRING	字段长度:	50
加密等级:	5	脱敏类型:	0
备注:		数据标签:	<input style="border: 1px solid #ccc; padding: 2px 5px;" type="button" value="+"/>



业务元数据管理

最近更新时间: 2019-10-28 02:33:01

分类管理

最近更新时间: 2019-11-12 03:43:04

业务元数据的分类，如数据标准、数据字典、业务术语等，由全局数据管理员维护。启用的业务元数据分类全员可见，未启用的只有全局数据管理员可见。可视化业务包是默认生成的，主要用于可视化BI。左侧菜单选中业务元数据，默认进入分类管理tab页。点击【新增分类】按钮弹出设置框，输入分类名称和描述，点击【确认】即可。



新增的业务分类需要启用才能生效，点击操作列的【启用】按钮启用确认提示，点击【确定】即可。当前启用状态的分类不支持修改和删除。

数据管理 > 业务元数据			
分类管理 元模型管理 元数据维护			
新增分类			
分类名称	分类描述	状态	操作
分类演示		停用	启用 修改 删除
test904	test904	启用	停用
OOO		启用	停用

元模型管理

最近更新时间: 2019-11-12 03:43:04

元模型隶属于业务分类，如数据标准分类下可以包含多个标准元模型。元模型管理提供元模型的新建、发布、下线、修改、授权、更改责任人、删除功能。由数据资产管理员进行元模型的管理和维护。新建元模型支持自定义模板和导入模板两种方式。自定义模板需要用户手动逐个填入模板字段及相关定义。导入模板支持下载模板文件，将模板字段定义批量写入模板文件，然后上传到平台。

元模型名称	元模型类别	元模型名称	元模型说明	新建方式	操作
测试元模型9041					下线 修改 授权 更改责任人
测试元模型名称904					下线 修改 授权 更改责任人
帆软测试20190831002					下线 修改 授权 更改责任人
帆软测试20190831001					下线 修改 授权 更改责任人
import_test11					下线 修改 授权 更改责任人
ss56					下线 修改 授权 更改责任人
帆软测试2019083001	可视化业务包	uat_test		2019-08-30 19:35:22	待发布更新 发布 下线 修改 授权 更改责任人
测试12	test1829	uat_test		2019-08-30 12:16:58	待发布 发布 修改 授权 删除 更改责任人
帆软测试20190830	可视化业务包	uat_test		2019-08-30 13:03:18	已发布 下线 修改 授权 更改责任人
阿范德萨范德萨	test1826	uat_test		2019-08-29 20:13:56	待发布 发布 修改 授权 删除 更改责任人

自定义模板类似定义一张表，需要定义该元模型包含的字段内容及相关属性。此处的字段属性跟建表时类似，此处不再赘述。需要说明的是“是否展示字段”，其含义是：元模型的默认展示字段，不展示的字段将会被隐藏，搜索结果不展现此类字段的内容，主要为了防止定义的字段太多，无法展示，且展示意义不大的问题。

新建元模型



① 配置基本信息

② 元模型字段

字段名称	字段中文名称	字段类型	字段长度	是否为主键	是否允许为空	是否展示字段	字段描述	操作
name	标准名称	字符串	50	是	否	是		下移 删除
		请选择		是	是	是		上移 删除

新增字段

上一步

完成

导入模板方式提供模板下载和上传，可以批量设置元模型字段：

新建元模型



① 配置基本信息

② 元模型字段

下载模板：[元模型导入模板.xlsx](#)

* 选择文件：



上一步

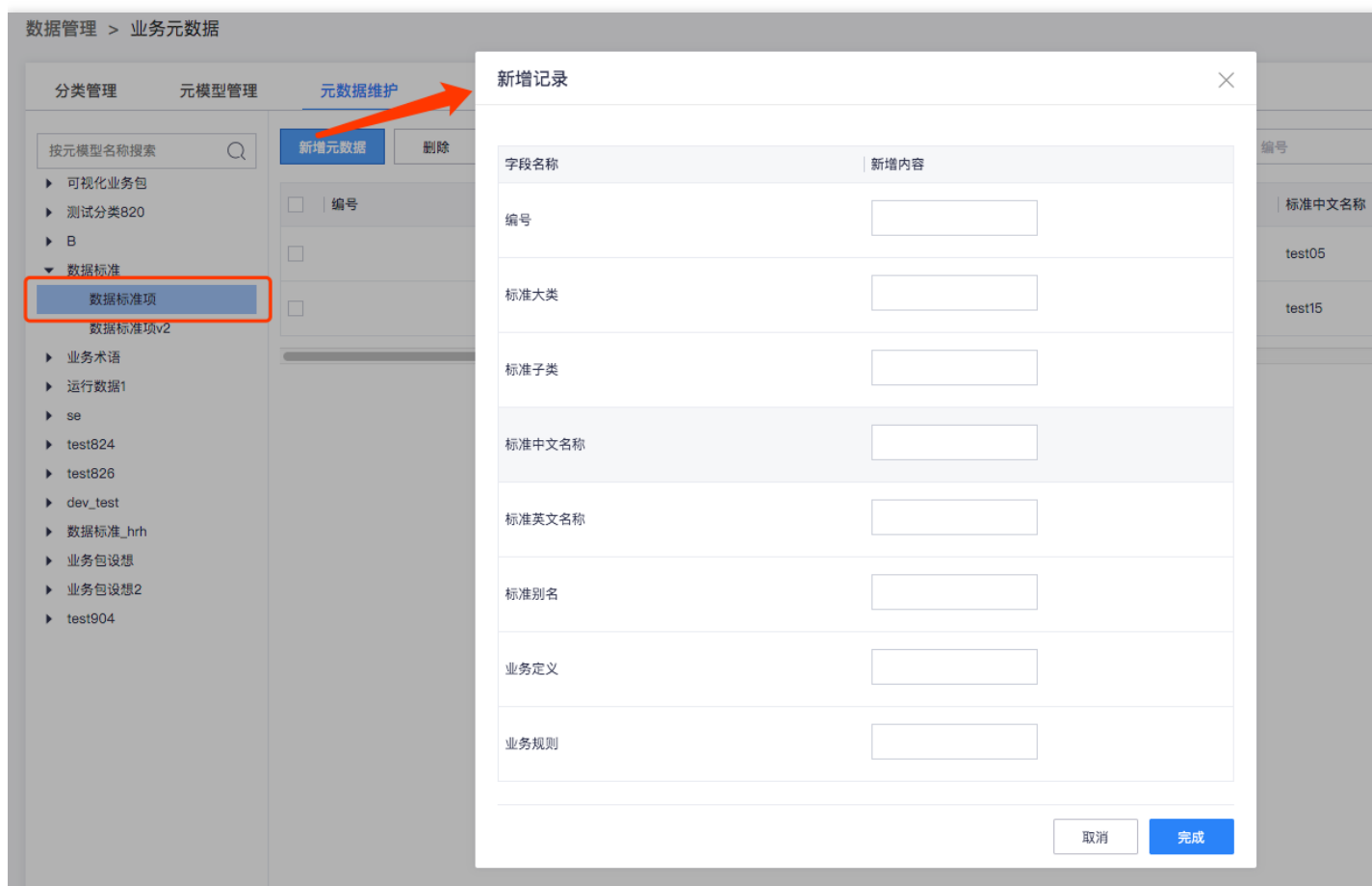
完成

新建的元模型为待发布状态，可点击操作列下的【发布】按钮进行发布。发布元模型需要全局数据资产管理审批。元模型的发布申请记录可以在发布申请列表看到。审批通过的元模型全员可见。发布的元模型不用后可以下线，下线的元模型其他人不可见，只有元模型创建者和上级管理员可见。发布后的元模型支持修改，修改后元模型状态由已发布变为待发布更新。点击【发布】按钮将走变更发布流程。元模型可以授权给其他人，授权的用户才能向该元模型里添加业务数据。当人员离职或变更时支持修改元模型的责任人。将元模型转交给其他数据资产管理维护。下线的元模型支持删除。

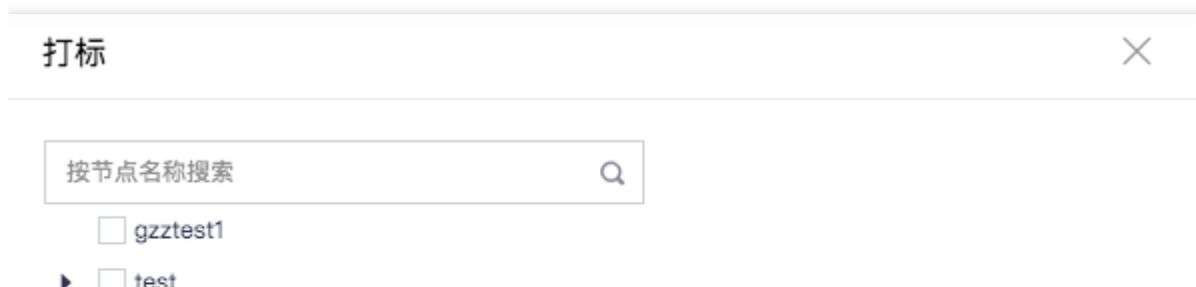
元数据维护

最近更新时间: 2019-11-12 03:43:04

定义了业务元模型下一步就可以添加业务数据。切换到元数据维护tab页，左侧目录树会显示发布的业务分类和元模型。选中具体的元模型右侧区域显示元模型定义的展示字段。点击【新增元数据】可单条新增业务数据记录。点击【导入】可向该元模型批量导入业务数据。只有授权的用户方可添加业务元数据。



新增的业务数据需要发布才能生效，可点击具体记录操作列下的【发布】按钮单条发布，或选中每条记录前面的复选框，点击上方【发布】按钮进行批量发布。业务元数据发布需要元模型创建人审批。发布后的业务数据支持下线。若发布的业务数据需要变更，可直接进行修改，修改后的业务数据需要重新发布。已下线的业务数据支持删除。点击操作列下的【历史】按钮可查看该条业务数据的历史变更记录。已发布的业务数据支持打标签，点击操作列下的【打标】按钮弹出打标界面。





—

- test111
- test333
- ▶ 一级
 - 一级故障标签
 - 一级缺陷标签
- ▶ 乙
- ▶ 分布2
 - 分布标签
 - 分布标签1
- ▶ 北京
- ▶ 测试
- ▶ 甲
- ▶ 行业2
- ▶ 行业标签
- ▶ 行业标签1
- ▶ 设备
- ▶ 部门树
- ▶ 领域标2
 - 领域标签
 - 领域标签1

元数据维护页面默认展示元模型的展示字段（新建元模型时定义），要看其他字段内容可点击最右侧的【设置图标】，调整要展示的字段，如下图：



自定义表格

编号 标准大类 标准子类 标准中文名称 标准英文名称

标准别名 业务定义 业务规则 发布状态 创建时间 更新时间

创建人 修改人

取消 确认

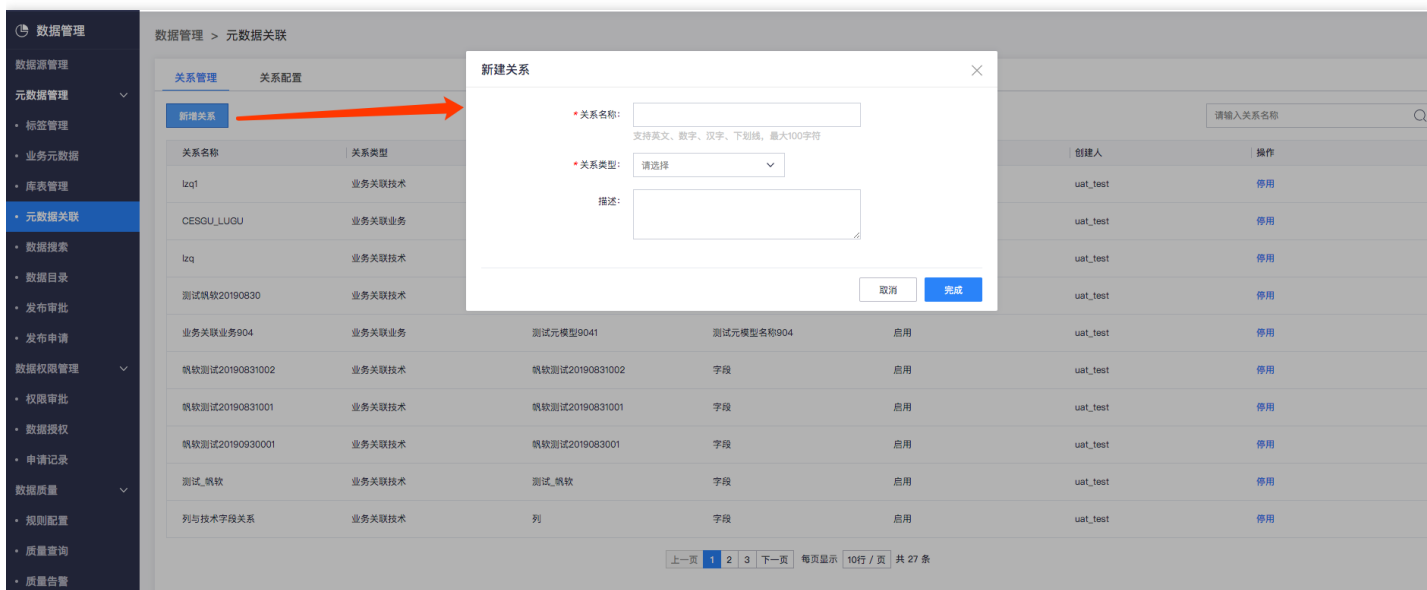
发布状态	操作
待发布	发布 修改 删除 历史
已发布	打标 下线 修改 历史
已发布	打标 下线 修改 历史

上一页 1 下一页 每页显示 10行 / 页 共 3 条

元数据关联配置

最近更新时间: 2019-11-12 03:43:04

数据管理支持技术元数据管理和业务元数据管理，以及技术元数据与业务元数据、业务元数据与业务元数据之间的关联关系配置。关联关系配置需要先定义关系，然后基于定义的关系配置具体的元数据关联。关系定义了关系类别（目前提供业务关联业务、业务关联技术两种关系类型）以及源实体和目标实体。业务的实体为具体的元模型，技术的实体和表或字段。左侧菜单切换到元数据关联页面，默认进入关系管理tab页。点击【新增关系】按钮可新建关系。



新建的关系需要启用才能生效，生效的关系通过【停用】可以失效，即关系不可用。停用状态的关系方可删除。



关系定义了之后可以进行关系配置。切换tab页到关系配置，系统默认会选中一个关系，下面展示该关系定义的源实体。可通过下拉列表切换关系类别和关系名称，下面的内容会相应联动变化。

关系管理 **关系配置**

关系类别: 业务关联技术 关系名称: 数据标准关联技术 **选择关系**

按定义的关键字段搜索 **关系定义的源实体** **发布**

数据标准项

标准大类	标准子类	标准中文名称	标准英文名称	标准别名	业务定义	业务规则	编号
test03	test04	test05	test06	test07	test08	test09	
test13	test14	test15	test16	test17	test18	test19	

上一页 1 下一页 每页显示 10行 / 页 共 2 条

点击源实体的任意一行可设置该行内容与目标实体的关联关系。弹出的目标实体页面支持按已发布的关联和未发布的关联查看。选中源实体记录要关联的目标实体记录，系统会自动保存。

数据管理 > 元数据关联

关系管理 **关系配置**

关系类别: 业务关联技术

按定义的关键字段搜索

数据标准项

标准大类	标准子类
test03	test04
test13	test14

表

未发布 数据库

所属表名称	所属库名称	数据源名称
<input checked="" type="checkbox"/> zzz	uat-bucket68	cos
<input checked="" type="checkbox"/> br09	topicdb_35_206	defaultKafka
<input checked="" type="checkbox"/> test0822	uat-bucket26	cos
<input checked="" type="checkbox"/> sparksqlTest_daxia	daxia_db1	mpp
<input type="checkbox"/> kc_house_nornal	dsp830_db	mpp
<input type="checkbox"/> cos_ren017	uat-bucket89	cos
<input type="checkbox"/> testmm	test_conn	hbase_online
<input type="checkbox"/> ren_826_tb2	ren_826	defaultHive
<input type="checkbox"/> tb_829	db_829	mpp

上一页 1 2 3 4 5 ... 18 下一页 每页显示 10行 / 页 共 177 条

关闭目标实体弹框，点击右侧【发布】按钮，即可发布刚才建立的关联关系。发布也需要走审批流程，关联关系的发布申请由项目数据管理员进行审批。



数据管理 > 元数据关联

发布

请确认是否发布

取消 确认

发布

数据标准项

标准大类	标准子类	标准中文名称	标准英文名称	标准别名	业务定义	业务规则	编号
test03	test04	test05	test06	test07	test08	test09	
test13	test14	test15	test16	test17	test18	test19	

上一页 1 下一页 每页显示 10行 / 页 共 2 条



数据质量

最近更新时间: 2019-10-28 02:38:52

数据质量模块当前支持对MPP、Oracle、Hive、MySQL四种关系型数据库进行质量检核。提供质量规则配置、质量任务查询、质量告警管理等功能。

规则配置

最近更新时间: 2019-11-15 07:10:03

当前质量规则配置提供五种技术检核算子和五种业务检核算子。具体说明如下：

检核类型	检核算子	算子级别	算子说明
技术检核	乱码检查	字段级	检核字符串类型的字段中是否含有乱码，有乱码的话将乱码替换为自定义的字符串
	转义字符检查	字段级	检核字符串类型的字段中是否含有转义字符，有的话将转义字符替换为自定义的字符串，默认替换为空格。
	内容长度检查	字段级	检核字符串类型、整型、浮点型字段内容长度是否超过类型定义的最大长度，超过则拒绝数据写入。
	内容格式检查	字段级	检核日期、时间类型字段的格式是否正确，不匹配的数据则拒绝写入。
	内容范围检查	字段级	检核字符串类型的字段是否包含规定以外的内容，如仅包含数字、仅包含字母、仅包含数据和字母，不满足条件产生告警。
业务检核	取值范围检查	字段级	检查指定对象的取值是否在指定的范围内，检查对象可以是字段或是字段加函数处理后的值，如字段长度、字段的一部分等；指定范围可以是输入的单个值、集合等。?例如：员工信息表中的证件类型代码是否在企业级数据字典域编号中。
	取值约束配置	字段级	检查当指定对象（A）的记录值为某一个/一组值时，另一指定对象（B）必须为特定的一个/一组值。例如：当对公客户类型为“事业机构”时，具备法人资格标志应为“是”；当对公客户类型为“个体工商户”、“外国政府驻华机构”时，具备法人资格标志应为“否”。
	数据格式检查	字段级	检查指定对象的格式是否符合特定规律的格式要求。例如：组织机构代码由9位数字组成。?
	非空检查	字段级	检查指定对象是否满足非空性要求，即所有记录值中不存在空值。例如：信息表中机构名称不能为空。
	唯一性检查	字段级	检查指定对象是否满足唯一性要求，即所有记录值中不存在重复记录。例如：机构信息表中机构编号不能重复。

操作说明如下：新建规则 业务人员、开发人员可以新建规则。菜单切换到数据质量下面的规则配置，点击【新增规则】按钮弹出创建规则配置框，如下图： ![]

(<http://imgxxfb.yun.ccb.com//raw/8eaf1dff770205c887ca4d92022e6d47.png>) 各字段说明如下：



输入信息	说明
规则名称	自定义规则名称，不能以下划线开头,支持数字,字母,中文和下划线组合，最多输入 100 个字符
检测对象	可以点击【新增】按钮选择检测对象，可以批量选择多个字段，最多选择26个检测对象，且当前不能夸项目选择。选错的对象可以删除。
检核类型	分技术检核和业务检核。不同的检核类型包含的算子会有差别。
检核算子	封装了具体检核的逻辑，具体说明见上面的算子说明。同时只能选一个算子。
告警阈值	分阈值下限和阈值上限，不同的阈值对应的告警级别不同。系统当前默认当检核的指标小于阈值时不产生告警；大于阈值下限，小于阈值上限时产生‘警告’级别告警；大于阈值上限产生‘严重’级别告警。
告警对象	告警对象为具体的用户，可以指定多个用户。
告警方式	短信、邮件告警，系统会自动往指定用户的电话或邮箱里发告警信息，但当指定的用户没有在用户管理模块配置电话或邮箱时，告警将无法发送。
描述	对规则的简单描述信息，最多输入 100 个字符

规则发布 新建的规则发布后才能生效，发布需要走申请审批流程，由库表所在项目的数据管理员进行审批。不用的规则可以下线。下线同样需要走审批。规则试运行 当前版本发布的业务检核规则支持试运行，用生产的几条数据进行试跑，可以批量试运行。试运行成功点击【试运行成功！点击查看结果】链接可以跳转到质量查询页面进行查看。 规则修改、删除 未发布的规则支持修改和删除；已发布的规则支持修改阈值和告警信息，修改后无需再走审批。规则查看 展开左侧的目录树，可以根据表或字段查看该表下所有字段或当前字段配置的所有规则。

质量管理 > 规则管理

检核算子	监测对象	创建人	创建时间	状态	操作
乱码检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/09/05 17:53:09	已发布	详情 下线 修改
取值范围检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/09/05 16:07:53	未发布	详情 发布 修改 删除
非空检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/09/04 20:26:20	已下线	详情 发布 修改 删除
乱码检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/08/29 23:42:23	已发布	详情 下线 修改
数据格式检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/08/28 21:07:45	未发布	详情 发布 修改 删除
非空检查	mppTest_0824 / HBFDL_YEAR / DIYU	uat_test	2019/08/26 11:35:20	已发布	详情 下线 修改

共 6 条

质量查询

最近更新时间: 2019-11-12 03:43:07

当前版本质量规则在数据管理的质量模块发布之后需要到数据集成去使用规则。数据集成的同步、加工任务会根据需求运行相关的检核规则，并生成质量检核任务。成功运行的质量检核任务会同步到数据管理的质量查询页面，该页面按照任务维度展示检核结果，可查看具体某次检核任务的质量检核结果。如下图：

数据管理 > 数据质量 > 质量查询

状态: 全部 数据源: 全部 数据库: 全部 所属项目: 全部 任务类型: 全部

业务时间: 请选择日期 执行时间: 请选择日期 输入表名称: 输入任务ID: 清空

任务ID	任务名称	任务类型	所属项目	数据表名	数据库名	数据源名	责任人	业务时间	执行时间	状态	规则数	异常规则数	操作
3107	testyewu109	数据集成	test_lir_0816	testtb@827	testdb827	mysql		2019-09-05 14:31:15	2019-09-05 14:31:16	正常	1	1	详情
3106	testyewu109	数据集成	test_lir_0816	testtb@827	testdb827	mysql		2019-09-05 14:15:58	2019-09-05 14:15:58	正常	1	1	详情
3105	gaozz_rule_201905_1	数据集成	dl_pro_uat	hrh_table3	hrh_test_quality	mysql	uat_test	2019-09-05 12:09:35	2019-09-05 12:09:35	正常	1	0	详情
3099	gzz_test_rule_2	数据集成	dl_pro_uat	dl_tech_source	dl_mysql_check	mysql	uat_test	2019-09-05 11:47:32	2019-09-05 11:47:32	正常	1	1	详情
3098	testyewu109	数据集成	test_lir_0816	testtb@827	testdb827	mysql		2019-09-05 11:21:24	2019-09-05 11:21:24	正常	1	1	详情
3097	0905	数据集成	dl_pro_uat	dl_tech_source	dl_mysql_check	mysql		2019-09-05 11:12:56	2019-09-05 11:12:56	正常	1	0	详情
3096	0905	数据集成	dl_pro_uat	dl_tech_source	dl_mysql_check	mysql		2019-09-05 11:11:36	2019-09-05 11:11:36	正常	1	0	详情
3095	的个人提供	数据集成	dl_pro_uat	hrh_table2	hrh_test_quality	mysql		2019-09-05 10:56:27	2019-09-05 10:56:27	正常	1	1	详情
3094	test905	数据集成	test_lir_0816	testtb@241927	testdb824	mpp		2019-09-05 10:50:59	2019-09-05 10:50:59	正常	1	1	详情
2421	kkkkkkk	数据集成	test_lir_0816	ren828_tb01	ren828_db	defaulthive	libangzhu	2019-09-04 19:48:12	2019-09-04 19:48:12	告警	1	1	详情

质量任务可以按照任务状态、数据源类型、数据源名称、所属项目、任务类型、业务时间、执行时间进行过滤，按具体的表或任务ID进行搜索。

状态: 全部 数据源: mpp数据库 数据库: 全部 所属项目: 全部 任务类型: 全部

业务时间: 请选择日期 执行时间: 请选择日期 输入表名称: 输入任务ID: 清空

定位到的质量任务支持点击操作列的【详情】按钮查看详情：

表名: testtbl8241927 > 任务详情 [返回](#)

规则名称	检核类型	规则模板	监测对象	阈值下限	阈值上限	异常数据 (条数)	操作
testmpp技术检核	技术检核	内容范围检查	testdb824/testtbl8241927/name	1	2	62	规则详情 历史结果
testmpp仅包含字母	技术检核	内容范围检查	testdb824/testtbl8241927/name	1	2	89	规则详情 历史结果
test内容长度检核_2	技术检核	内容长度检查	testdb824/testtbl8241927/name	1	2	90	规则详情 历史结果
test转义字符检核_1	技术检核	转义字符检查	testdb824/testtbl8241927/name	0	1	0	规则详情 历史结果
testmpp乱码检查	技术检核	乱码检查	testdb824/testtbl8241927/name	0	1	0	规则详情 历史结果
testmpp仅包含数字和字母	技术检核	内容范围检查	testdb824/testtbl8241927/name	0	1	61	规则详情 历史结果

上一页 1 下一页 每页显示 10行 / 页 共 6 条

详情里展示了该任务跑的具体有哪些规则，以及规则相关的一些配置信息和每条规则在本次任务检核中检测到的异常数据，点击异常数据列的具体数字可查看具体的异常数据，如下图：

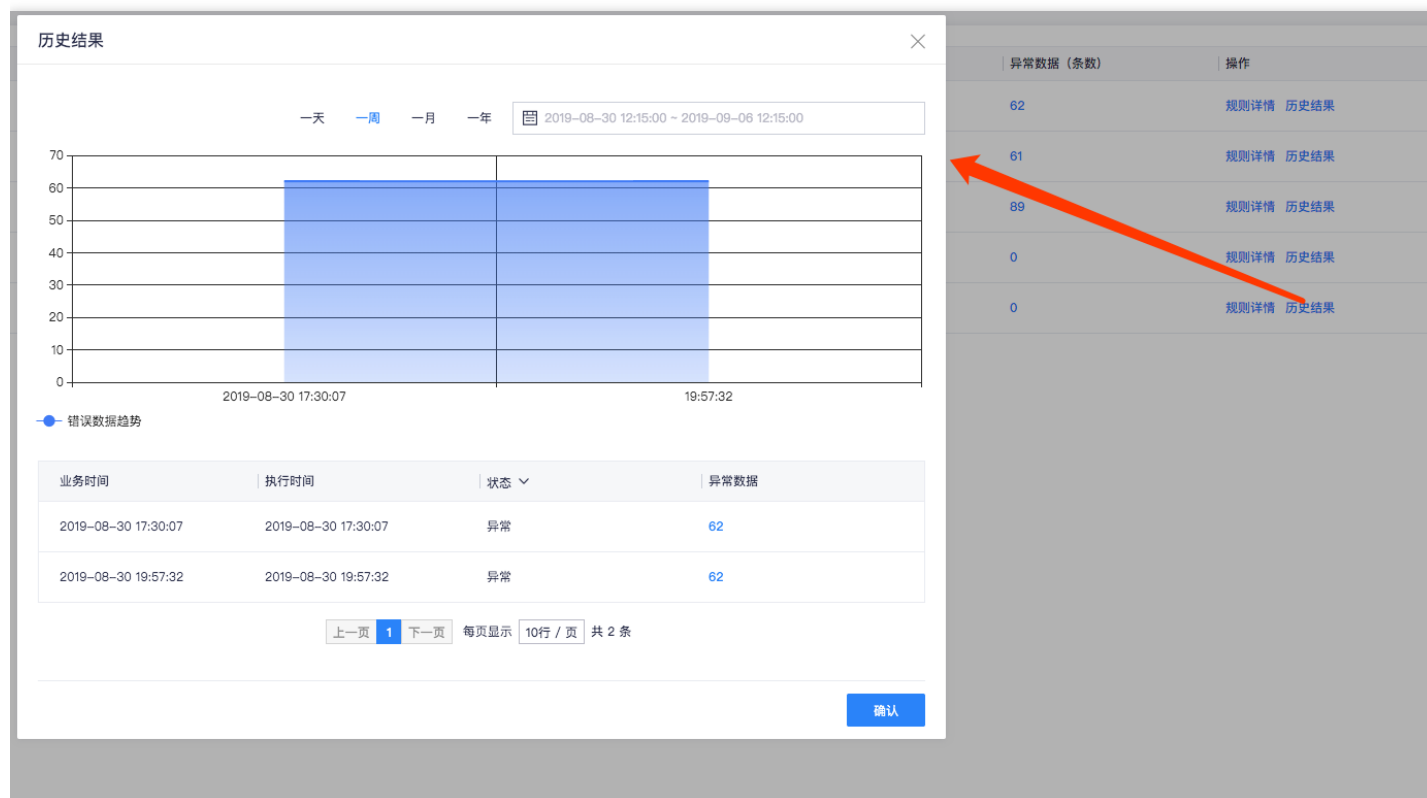
异常数据

异常数据	异常数据描述
@@@ 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name @@@ 不符合内容范围检查规则
张三 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name 张三 不符合内容范围检查规则
张三 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name 张三 不符合内容范围检查规则
张三 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name 张三 不符合内容范围检查规则
李四 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name 李四 不符合内容范围检查规则
张三 23.2 2019-08-24 20:00:00 2019-08-24 14:20:00	testdb824@ testtbl8241927@ 2019-08-28 19:59:02@ name 张三 不符合内容范围检查规则

上一页 1 2 3 4 5 6 7 下一页 每页显示 10行 / 页 共 62 条

确认

每个任务的每条规则支持查看历史检核结果。点击操作列【历史结果】按钮弹出历史结果框，如下图所示：





质量告警

最近更新时间: 2019-11-12 03:43:07

质量检核任务产生的告警信息统一在质量告警页面展现。可通过勾选【只看失败告警】查看发送失败的告警记录。也可通过告警级别、告警ID、规则名称、告警接收人搜索具体的告警信息。

- 数据管理
- 数据源管理
- 元数据管理
- 标签管理
- 业务元数据
- 库表管理
- 元数据关联
- 数据搜索
- 数据目录
- 发布审批
- 发布申请
- 数据权限管理
- 权限审批
- 数据授权
- 申请记录
- 数据质量
- 规则配置
- 质量查询
- 质量告警

数据质量 > 质量告警

只看失败告警 告警级别: 全部 告警ID: 输入告警ID

告警ID	规则名称	告警时间	告警方式	告警接收人	告警级别	告警信息	告警状态
743	onlyCheck0111111111	2019-09-04 19:49:24	短信	libangzhu	警告	onlyCheck0111111111告警通知,算子类型为:唯一性检查!告警信息如下: 告警时间为:2019-09-04 19:49:23 告警数据量为:8 告警阈值下限为:1 告警阈值上限为:10 告警级别为:警告 数据源为:defaultHive,数据库为:ren828_db,告警表为:ren828_tb01 告警字段为:a5 请尽快处理!	失败
742	onlyCheck0111111111	2019-09-04 18:10:14	短信	libangzhu	警告	onlyCheck0111111111告警通知,算子类型为:唯一性检查!告警信息如下: 告警时间为:2019-09-04 18:10:13 告警数据量为:8 告警阈值下限为:1 告警阈值上限为:10 告警级别为:警告 数据源为:defaultHive,数据库为:ren828_db,告警表为:ren828_tb01 告警字段为:a5 请尽快处理!	失败
713	mysql仅包含数字	2019-08-30 21:25:12	邮件	uat_test	警告	mysql仅包含数字告警通知,算子类型为:内容范围检查!告警信息如下: 告警时间为:2019-08-30 21:25:11 告警数据量为:6 告警阈值下限为:1 告警阈值上限为:10 告警级别为:警告 数据源为:mysql,数据库为:ren828_db,告警表为:ren828_tb02 告警字段为:a6 请尽快处理!	失败
712	oracle_01_ren_内容长度	2019-08-30 20:00:16	邮件	uat_test	警告	oracle_01_ren_内容长度告警通知,算子类型为:内容长度检查!告警信息如下: 告警时间为:2019-08-30 20:00:15 告警数据量为:1 告警阈值下限为:1 告警阈值上限为:9 告警级别为:警告 数据源为:oracle,数据库为:ccb,告警表为:ren828_tb01 告警字段为:a1 请尽快处理!	失败
676	oracle_01_ren_内容长度	2019-08-30 18:35:30	邮件	uat_test	警告	oracle_01_ren_内容长度告警通知,算子类型为:内容长度检查!告警信息如下: 告警时间为:2019-08-30 18:35:29 告警数据量为:1 告警阈值下限为:1 告警阈值上限为:9 告警级别为:警告 数据源为:oracle,数据库为:ccb,告警表为:ren828_tb01 告警字段为:a1 请尽快处理!	失败
593	唯一性检查_oracle_jlr829_6	2019-08-30 13:11:00	短信	test_dba,test_pro_dba,test_pro_admin	严重	唯一性检查_oracle_jlr829_6告警通知,算子类型为:唯一性检查!告警信息如下: 告警时间为:2019-08-30 13:11:00 告警数据量为:125 告警阈值下限为:3 告警阈值上限为:10 告警级别为:严重 数据源为:oracle,数据库为:ccb,告警表为:testtb8241927 告警字段为:dd 请尽快处理!	失败
594	唯一性检查_oracle_jlr829_6	2019-08-30 13:11:00	邮件	test_dba,test_pro_dba,test_pro_admin	严重	唯一性检查_oracle_jlr829_6告警通知,算子类型为:唯一性检查!告警信息如下: 告警时间为:2019-08-30 13:11:00 告警数据量为:125 告警阈值下限为:3 告警阈值上限为:10 告警级别为:严重 数据源为:oracle,数据库为:ccb,告警表为:testtb8241927 告警字段为:dd 请尽快处理!	失败

常见问题

最近更新时间: 2021-09-16 10:08:31

Q: 什么是数据管理? A: 数据管理是一种全方位的企业数据治理服务, 提供云端传统关系型数据库、MPP数据库、Nosql数据库、大数据平台数据仓库的统一管理, 提供数据地图、数据质量、数据血缘分析能力, 助力企业全面管理数据资产。

Q: 为什么需要数据管理? A: 如果您的数据存在多个零散、异构的数据存储系统里, 如Oracle、HBase、Kafka、ES等, 想要对所有数据的元数据及权限进行统一管理, 您可能需要数据管理。如果您现在的数据系统封闭, 让用户发现数据、获取数据的成本很高, 您可能需要数据管理。如果您现在的数据质量不高, 缺乏相关标准和编码规范, 想要快速提升数据质量, 进行数据溯源与影响分析, 您可能需要数据管理。

Q: 数据管理提供哪些功能? A: 数据管理产品最大的亮点在于提供统一的元数据管理。无论是关系型数据库、分布式MPP数据库、NoSql数据库、图数据库、消息队列还是对象存储, 均能进行元数据映射管理, 并托管数据权限。同时支持数据标准、数据字典、业务术语等业务元模型管理与关联关系配置, 批量设置数据标签。在全域元数据的基础上数据管理提供数据地图、数据血缘、数据生命周期等高级应用, 并提供技术、业务多种数据质量检核规则, 全面提升数据管控、数据发现、数据共享效率, 保障数据质量提升和数据安全合规使用。

Q: 如何开通数据管理? A: 目前仅支持线下开通, 请联系商务人员, 会为您定制解决方案

Q: 数据管理不同版本有什么区别, 适用场景是什么? A: 数据管理标准版功能较少, 配置的资源相对较少, 支持的并发量较小, 适合小规模团队使用, 适用于侧重数据开发, 对数据管理无特别要求的场景。专业版提供数据标签、数据目录、数据搜索、数据血缘等更多功能, 能够快速定位数据, 以及发现数据之间的关系。配置的物理资源相对较多, 支持的并发数较多, 适合于中等规模数据量开发管理场景。企业版相对专业版来说提供业务元数据、元数据关联、数据地图、数据质量等更多高级功能, 能够更好的满足企业数据治理的场景, 全方位的了解企业数据资产。配置的物理资源是最高的, 支持的并发数也最多, 适合大型企业复杂的数据治理场景。对于一些特大企业, 如果以上版本配置的资源满足不了需求, 可进行定制化资源配置, 提供更多的服务资源。