

解决方案实践

软通动力数据中台及数据治理服务解决方案实践

文档版本 1.0
发布日期 2023-12-15



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

1 方案概述	1
2 资源和成本规划	4
3 实施步骤	5
3.1 元数据用户操作手册	5
3.1.1 用户管理	5
3.1.2 模型架构	10
3.1.3 适配器	15
3.1.4 采集任务	17
3.1.5 功能列表	21
3.2 数据资产用户操作手册	22
3.2.1 信息架构	22
3.2.2 物理模型	28
3.2.3 数据标准	33
3.2.4 资产注册	36
3.2.5 系统管理	38
3.2.6 功能列表	39
3.3 数据集成平台操作手册	40
3.3.1 平台界面	40
3.3.2 数据集成	43
3.3.3 数据调度	51
3.3.4 功能列表	53
3.4 数据地图平台操作手册	54
3.4.1 数据全景	54
3.4.2 资产注册	59
3.4.3 数据运营与质量	61
3.4.4 标签管理	62
3.4.5 系统管理	64
3.4.6 功能列表	68
3.5 数据生命周期平台操作手册	69
3.5.1 总览	69
3.5.2 数据源管理	70
3.5.3 数据归档管理	72

3.5.4 系统管理.....	76
3.5.5 功能列表.....	81
3.6 DataLab 用户手册.....	82
3.6.1 数据源.....	83
3.6.2 数据集.....	84
3.6.3 数据探索.....	85
3.6.4 数据可视化.....	87
3.6.5 系统管理.....	95
3.6.6 功能列表.....	98
4 附录 背景信息.....	99
5 A 修订记录.....	102

1 方案概述

应用场景

软通动力数据中台及数据治理服务解决方案采用“轻量化、高解耦、全场景、零代码”的架构设计，具备多部署模式、灵活产品组合、无缝集成对接等核心特性，且能够提供数据中台咨询、实施及运维专业端到端服务。

1. 应用场景

数据使能咨询、数据集成与开发、工业物联数据采集与分析、数据治理与数据安全实施、BI应用开发、行业AI应用开发、行业数据应用开发、数据与数据平台运营运维

2. 业务效果

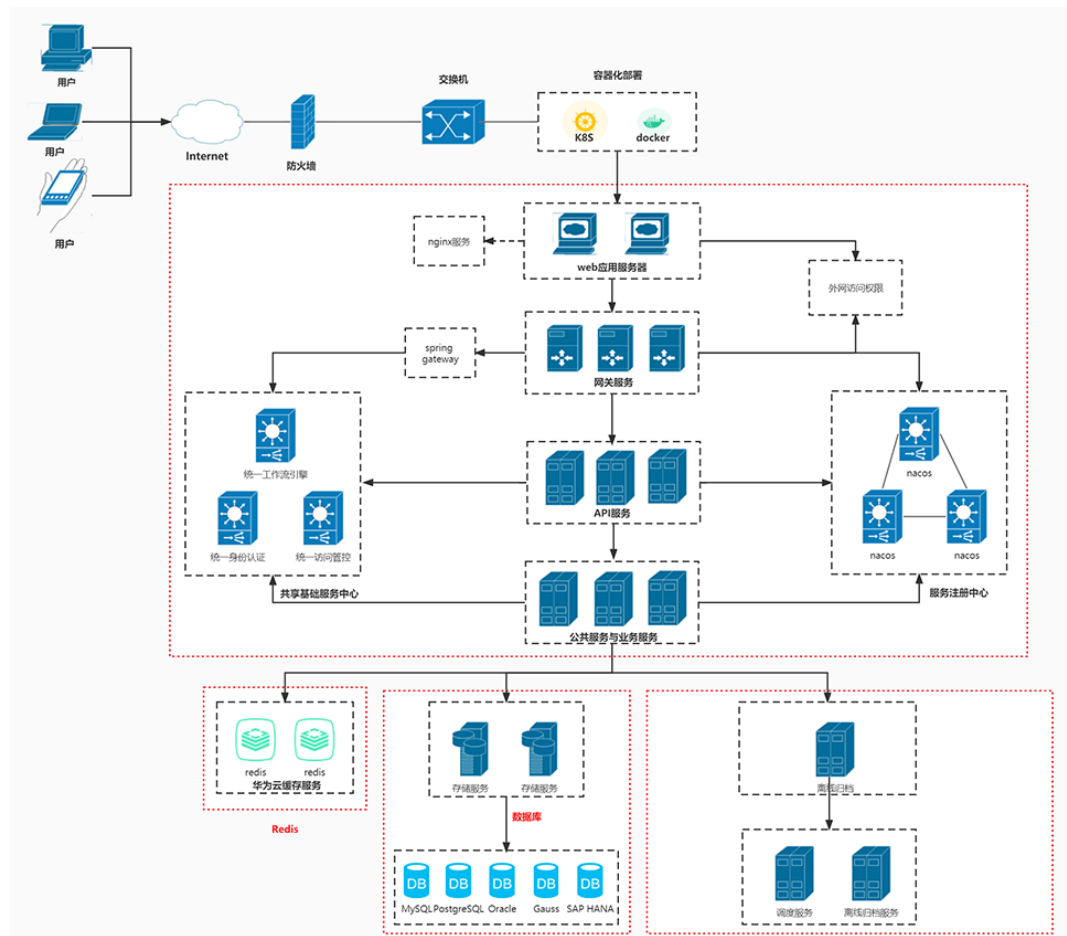
- 价值实现：围绕业务需求，长治久安地构建起涵盖资产、安全、标准、质量、管控等5大体系的企业数据管理机制；
- 价值应用：基于业务流程场景，开启企业数智之路，打通数据资产识/取/用/管/运/维“最后一公里”；
- 成本收益：集成管理数据资源，实现低成本业务创新；
- 效率收益：业务与IT协同创新，实现多应用并行开发

方案架构

图 1-1 业务架构



图 1-2 部署架构



架构描述:

- 本架构基于软通动力“云上中台 重明”数据中台企业级数据平台工具链解决方案，围绕业务需求，构建资产、安全、标准、质量、管控等5大企业数据管理体系，基于业务流程场景，开启企业数智之路；
- 平台采用主流的容器化和微服务架构，具备实现分布式、高可用、故障自动转移、负载均衡、资源可以动态按需扩容等优势；
- 支持公有云、私有云、混合云和本地化多种方式部署；
- 支持自研数据库部署。

方案优势

- 产品轻量级，覆盖数字平台全链路：和华为云DGC深度集成形成联合解决方案，满足不同类型企业在业务数据的汇聚、治理、运营、消费等场景下的差异化需求；
- 紧密贴合客户需求，灵活组合产品簇套件，具有高性价比与价格优势，价格平均比市面上同类产品低10%；
- 零代码平台：界面配置、自助拖拽、自由编排，2-3天可以熟练上手；
- 服务优势：由产品团队提供轻咨询、驻场实施、定制化开发、运维等端到端服务。

2 资源和成本规划

表 2-1 资源和成本规划

云资源	规格	数量	每月费用 (元)
VPC	网段选择172.16.0.0/16, 其他采用默认配置	1	00.00
Sub net	网段选择172.16.0.0/24, 其他采用默认配置	1	00.00
安全组	根据需要开通入方向3306等端口	1	00.00
DG C	企业版: 5,000次/天的数据开发调度, 附带8作业并发, 最大1.5Gbps带宽的数据集成能力, 并且支持管理5,000个技术资产、100个数据模型。	1	2,0000 0.00
ECS	X86计算 通用计算型 s3.xlarge.2 4核 8GB; 通用型SSD 40GB	2	3323.0 0
CD M	cdm.xlarge; 16核/32GB 10/4 Gbit/s 32 并发任务	1	26780. 00
DW S	dwsx2.2xlarge.m7 云数仓 X86 8 vCPUs 64 GB * 3 节点; 1000 GB * 2 (副本) * 3 节点	1	155069 .28
总计: 385172.28			

3 实施步骤

- 3.1 元数据用户操作手册
- 3.2 数据资产用户操作手册
- 3.3 数据集成平台操作手册
- 3.4 数据地图平台操作手册
- 3.5 数据生命周期平台操作手册
- 3.6 DataLab用户手册

3.1 元数据用户操作手册

3.1.1 用户管理

图 3-1 用户管理

用户登录名	员工工号	用户名	用户类型	用户状态	创建时间	用户创建人	修改时间	修改人	操作
test	002	myh	数据管家	正常					编辑 删除
zhj	001	zhj	普通用户	正常			2021-12-27 14:45	admin	编辑 删除
ceshi	ceshi	ceshi	普通用户	禁用	2021-12-24 15:20	admin	2021-12-24 15:21	admin	编辑 删除

用户分为三个角色：系统管理员/数据管家/普通用户

- 系统管理员：项目初始化角色不可编辑
- 数据管家：该角色只能系统管理员来创建，该角色拥有元数据管理平台上数据的查询和编辑权限

- 普通用户：系统管理员和数据管家都可以创建该角色，该角色只具备元数据管理平台数据查询权限

图 3-2 新建用户

新建用户

* 用户登录名: 请输入...

* 用户密码: 请输入...

* 员工工号: 请输入...

* 用户姓名: 请输入...

* 用户类型: 系统管理员 数据管家 普通用户

提交 取消

图 3-3 编辑用户

+ 新增用户

用户登录名	员工工号	用户姓名	用户类型	用户状态	创建时间	用户创建人	修改时间	修改人	操作
test	002	myh	数据管家	正常					编辑 删除
zhj	001	zhj	普通用户	正常			2021-12-27 14:45	admin	编辑 删除
ceshi	ceshi	ceshi	普通用户	禁用	2021-12-24 15:20	admin	2021-12-24 15:21	admin	编辑 删除

图 3-4 编辑用户信息

编辑 X

* 用户登录名:

* 员工工号:

* 用户姓名:

* 用户类型: 系统管理员 数据管家 普通用户

图 3-5 删除用户

+ 新增用户

用户登录名	员工工号	用户姓名	用户类型	用户状态	创建时间	用户创建人	修改时间	修改人	操作
test	002	myh	数据管家	正常					编辑 删除
zhj	001	zhj	普通用户	正常			2021-12-27 14:45	admin	编辑 删除
ceshi	ceshi	ceshi	普通用户	禁用	2021-12-24 15:20	admin	2021-12-24 15:21	admin	编辑 删除

图 3-6 数据源

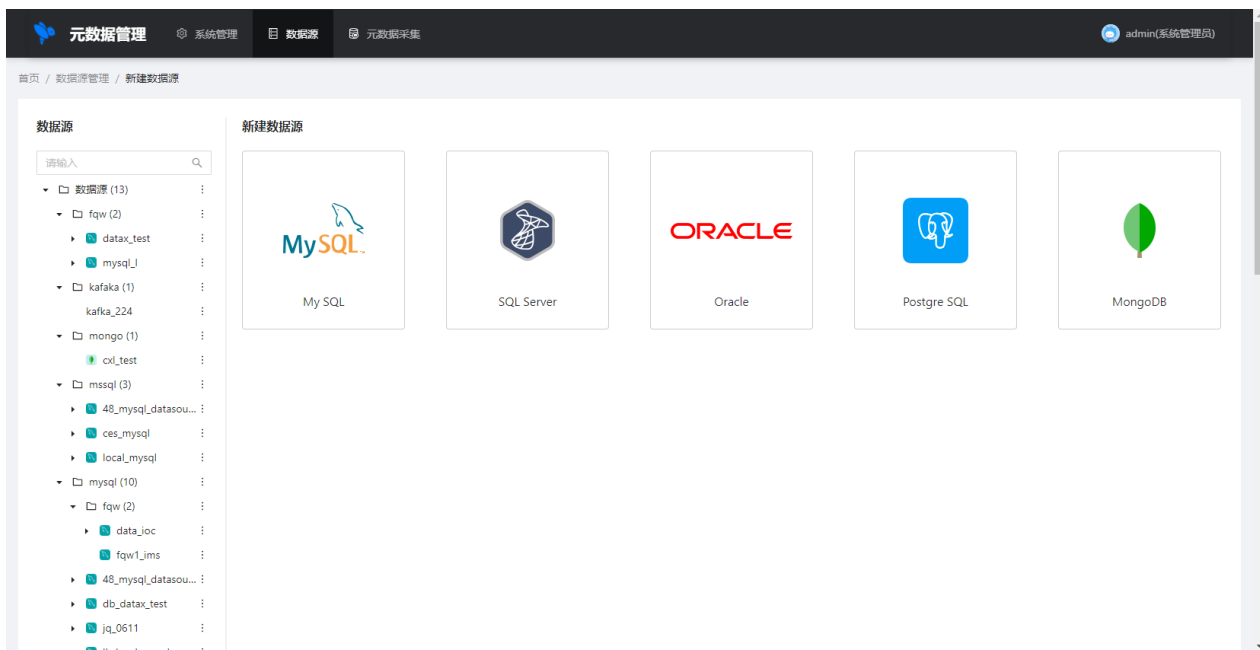


图 3-7 新增数据源

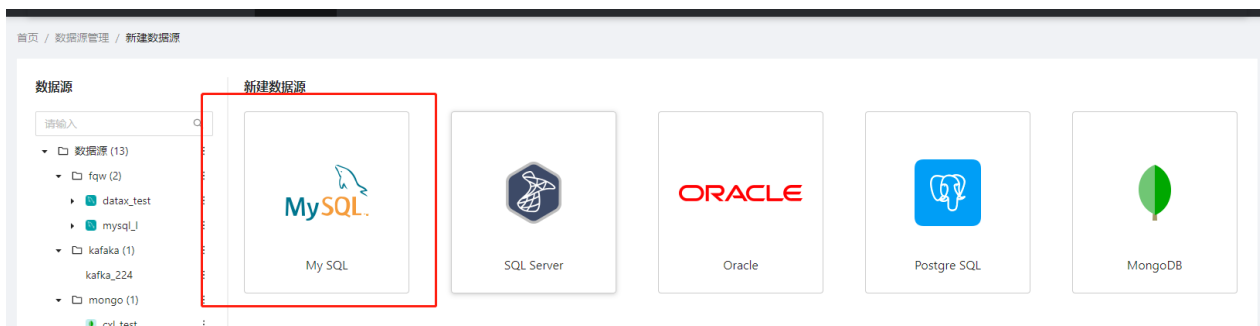


图 3-8 录入数据源信息

添加数据源 X

1 测试连接 2 选择目录 3 完成

* 数据源名称:

* 连接URL:
格式: jdbc:mysql://IP:端口/服务名

* 用户名:

* 密码:

数据源描述:

图 3-9 编辑数据源

首页 / 数据源管理 / 查看数据源

数据源

请输入

- 数据源 (13)
- fqw (2)
- datax_test**
- mysql_1
- kafaka (1)
- kafka_224
- mongo (1)
- cd_test
- mssql (3)
- 48_mysql_datasou...
- ces_mysql
- local_mysql
- mysql (10)
- fqw (2)
- data ioc

* 数据源名称:

* 连接URL:

* 用户名:

* 密码:

数据源描述:

3.1.2 模型架构

图 3-10 模型架构

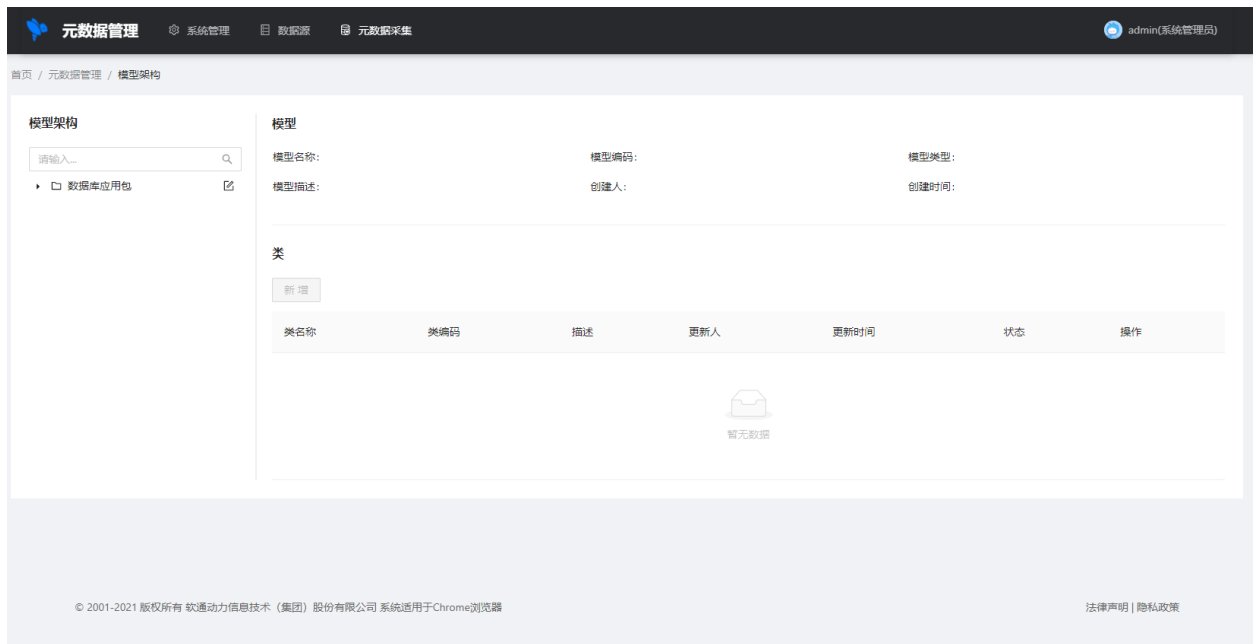
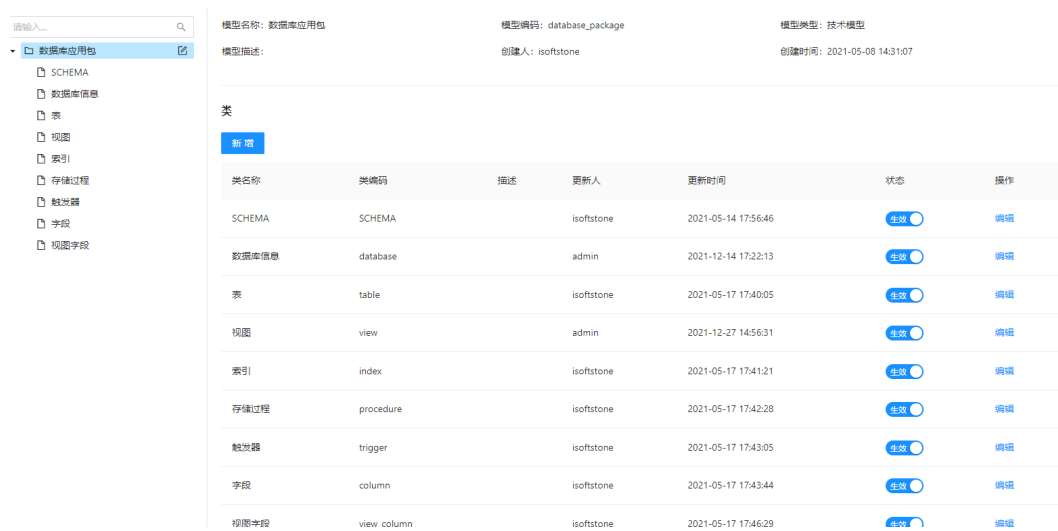


图 3-11 模型架构--类



模型架构中类信息的维护

图 3-12 新增类

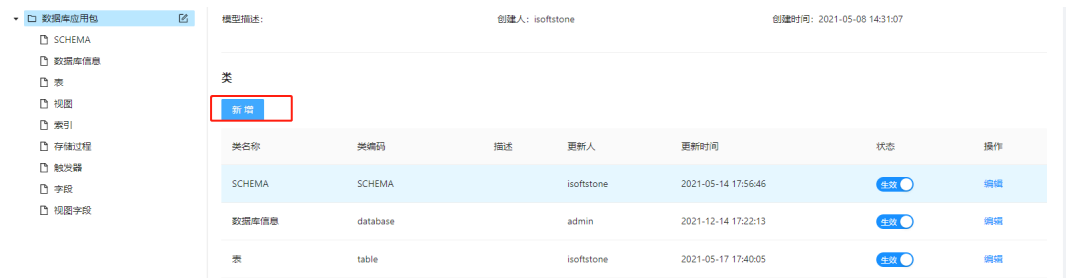


图 3-13 录入类信息

新增

* 类名称:

* 类编码:

* 是否生效: 是 否

描述:

图 3-14 编辑类

类

新增

类名称	类编码	描述	更新人	更新时间	状态	操作
SCHEMA	schema		admin	2021-12-27 16:40:52	<input checked="" type="radio"/> 生效	编辑
数据库信息	database		isoftstone	2021-05-14 16:52:48	<input checked="" type="radio"/> 生效	编辑
表	table		isoftstone	2021-05-17 17:40:05	<input checked="" type="radio"/> 生效	编辑

图 3-15 修改类信息

修改 X

* 类名称:

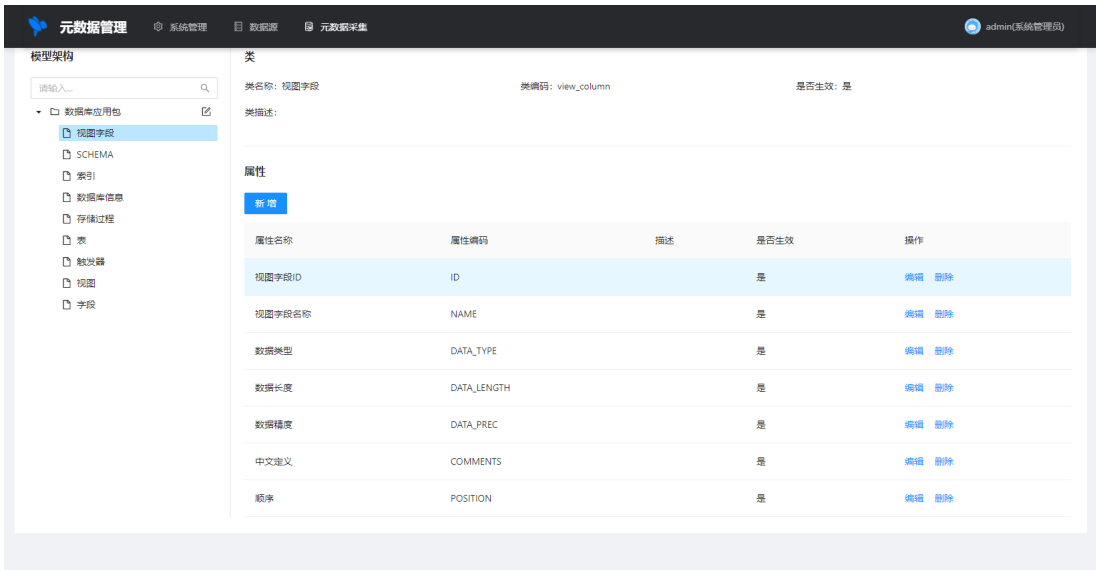
* 类编码:

* 是否生效: 是 否

描述:

在适配器中该类已经被添加了sql语句的，不能编辑修改，如要修改，必须将该类下面对应的sql语句都删除掉。

图 3-16 类属性



单击按钮弹出页面进行修改。

图 3-17 新增类属性



图 3-18 录入类属性信息

新增 X

* 属性名称:

* 属性编码:

* 是否生效: 是 否

描述:

图 3-19 编辑类属性

属性

[新增](#)

属性名称	属性编码	描述	是否生效	操作
视图字段ID	ID		是	编辑 删除
视图字段名称	NAME		是	编辑 删除

图 3-20 修改类属性信息

修改 X

* 属性名称:

* 属性编码:

* 是否生效: 是 否

描述:

图 3-21 删除类属性

属性

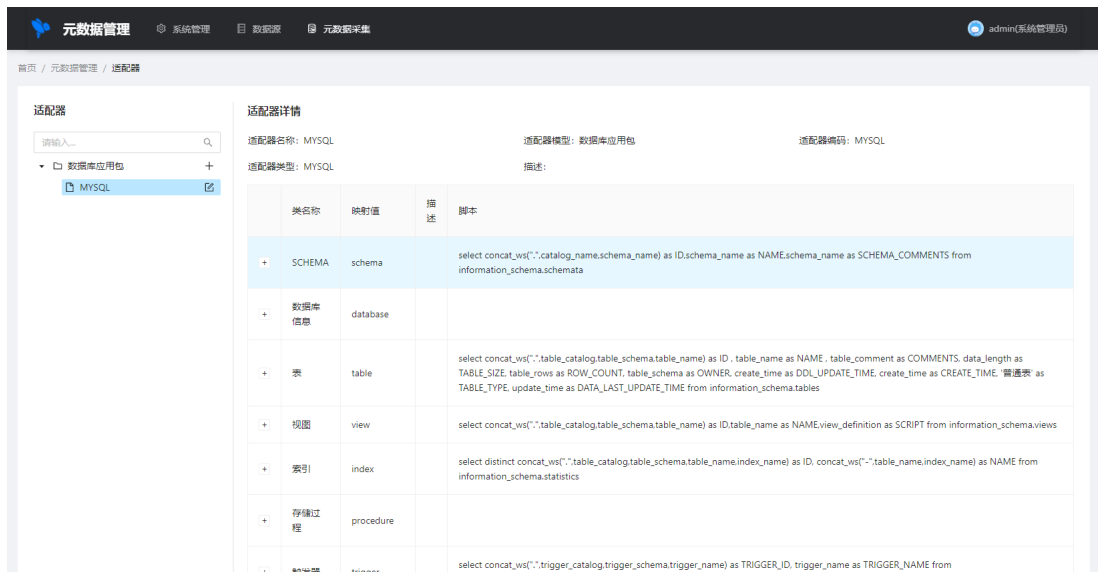
属性名称	属性编码	描述	是否生效	操作
视图字段ID	ID		是	<input type="button" value="编辑"/> <input type="button" value="删除"/>
视图字段名称	NAME		是	<input type="button" value="编辑"/> <input type="button" value="删除"/>

3.1.3 适配器

新增适配器信息

- 适配器名称: 必填
- 适配器编码: 必填 不能重复
- 适配器类型: 单选 (MYSQL/ORACLE/POSTGRES/MSSQL)
- 类脚本: 选填

图 3-22 适配器

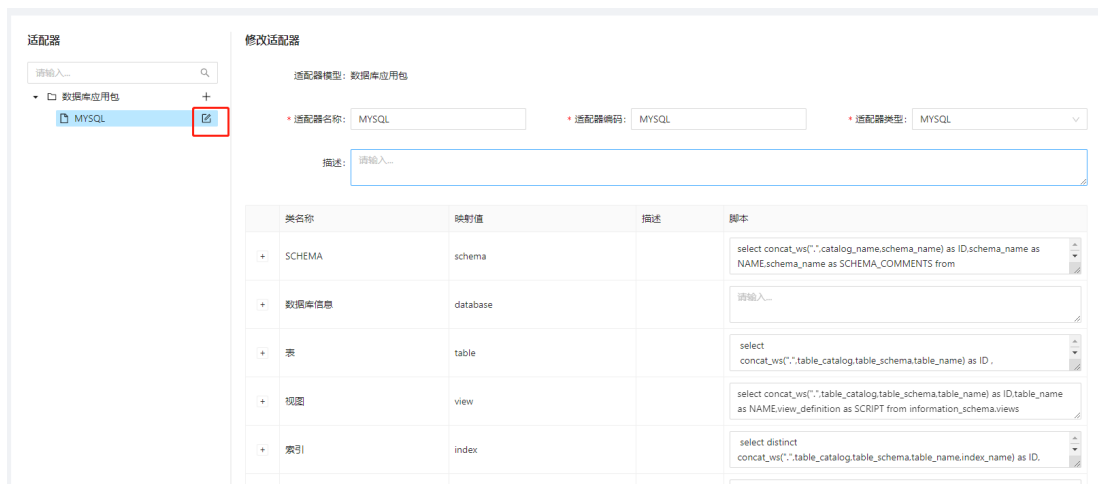


可以对已经创建好的适配器信息进行编辑，字段校验规则参考新增。

修改特殊处理规则：

1. 新增适配器脚本时，如果该适配器已经创建了采集任务，所有已经创建的采集任务及对应的子任务都需要重新部署
2. 修改已经存在的适配器脚本，该适配器对应的所有的采集任务中对应的可执行脚本命令需要重新生成，但不影响本生采集任务的部署状态

图 3-23 适配器编辑



3.1.4 采集任务

图 3-24 采集任务

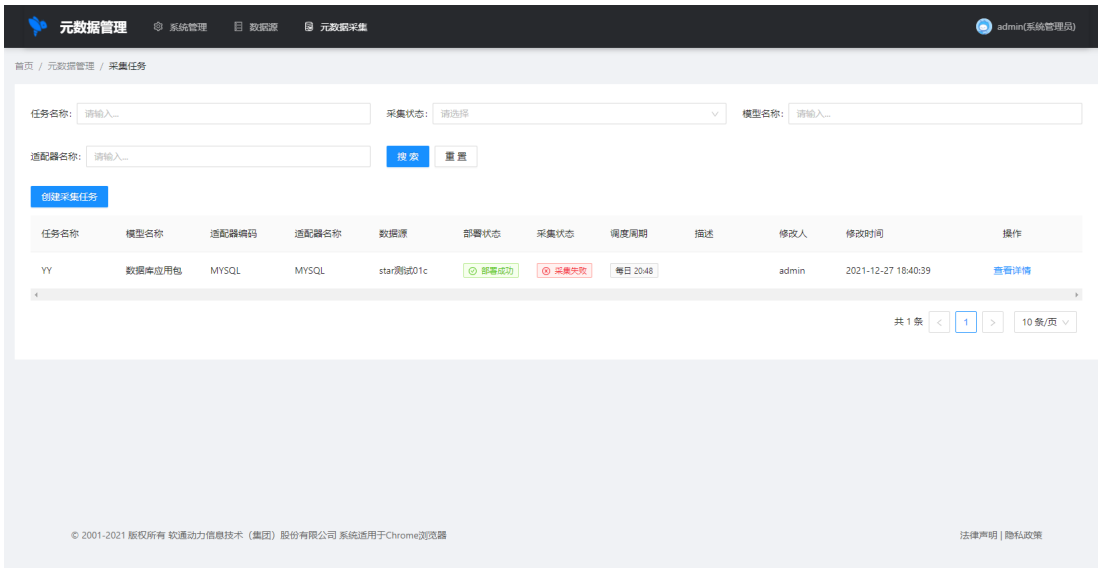


图 3-25 新增采集任务



图 3-26 录入采集任务信息

新建采集任务 X

* 任务名称:

描述:

* 数据源:

* 模型名称:

* 适配器:

* 开始执行时间:

* 调度周期:

- 任务名称: 必填 不可重复
- 描述: 选填
- 数据源: 必填 选择已经存在的数据源
- 模型名称: 必填 选择已经存在的模型
- 适配器: 必填 选择模型下存在的适配器
- 开始执行时间: 必填 采集任务首次开始执行时间
- 调度周期: 必填 采集任务的调度频率选择

图 3-27 部署采集任务

任务名称	模型名称	适配器编码	适配器名称	数据源	部署状态	采集状态	调度周期	描述	修改人	修改时间	操作
YY	数据库应用包	MYSQL	MYSQL	star测试01c	✔ 部署成功	❌ 采集失败	每日 20:48		admin	2021-12-27 18:40:39	查看详情
cs001	数据库应用包	MYSQL	MYSQL	mysql_yun_94	⚠ 未部署	⚠ 未采集	每日 17:02		admin	2021-12-28 17:02:44	编辑 启动部署 查看详情

图 3-28 编辑采集任务



任务名称	模型名称	适配器编码	适配器名称	数据源	部署状态	采集状态	调度周期	描述	修改人	修改时间	操作
YY	数据库应用包	MYSQL	MYSQL	star测试01c	部署成功	采集失败	每日 20:48		admin	2021-12-27 18:40:39	查看详情
cs002	数据库应用包	MYSQL	MYSQL	mysql_yun_94	未部署	未采集	每日 17:10		admin	2021-12-28 17:07:04	编辑 启动部署 查看详情
cs001	数据库应用包	MYSQL	MYSQL	mysql_yun_94	部署成功	采集失败	每日 17:02		admin	2021-12-28 17:03:42	查看详情

图 3-29 修改采集任务



修改采集任务

* 任务名称: cs002

描述: 请输入...

* 数据源: mysql_yun_94

* 模型名称: 数据库应用包

* 适配器: 10

* 开始执行时间: 2021-12-28 17:11

* 调度周期: 每日 17:10

取消 确定

未部署的采集任务才可以编辑

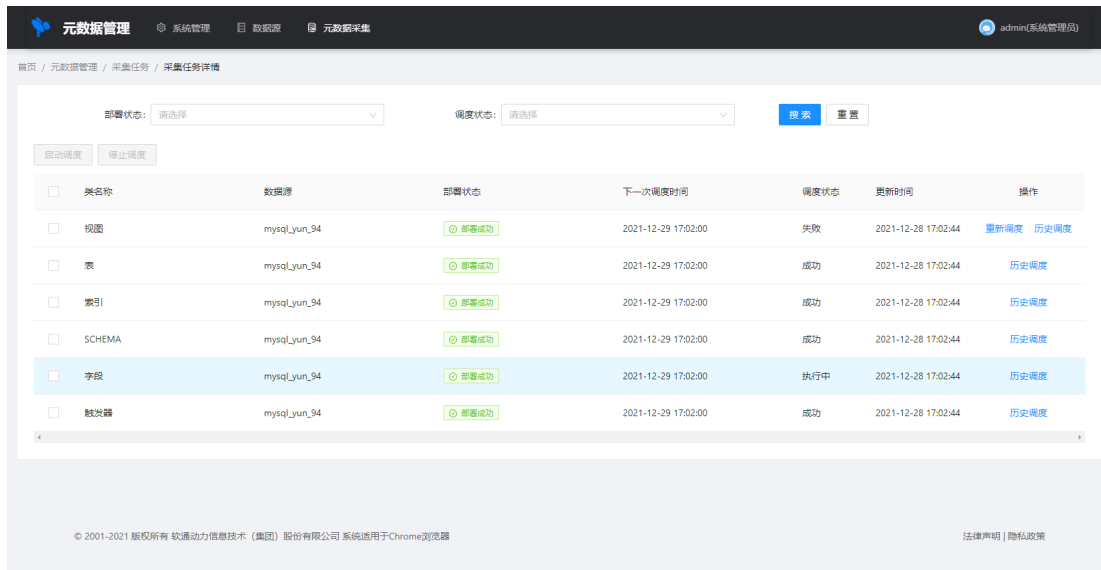
采集任务重新部署：部署失败的采集任务可以进行重新部署。

图 3-30 采集子任务-1



任务名称	模型名称	适配器编码	适配器名称	数据源	部署状态	采集状态	调度周期	描述	修改人	修改时间	操作
YY	数据库应用包	MYSQL	MYSQL	star测试01c	部署成功	采集失败	每日 20:48		admin	2021-12-27 18:40:39	查看详情
cs002	数据库应用包	MYSQL	MYSQL	mysql_yun_94	未部署	未采集	每日 17:10		admin	2021-12-28 17:07:04	编辑 启动部署 查看详情
cs001	数据库应用包	MYSQL	MYSQL	mysql_yun_94	部署成功	采集失败	每日 17:02		admin	2021-12-28 17:03:42	查看详情

图 3-31 采集子任务-2



子任务重新部署

部署失败的子任务可以重新进行部署。

图 3-32 子任务重新调度



调度失败的子任务可以重新调度

图 3-33 子任务历史调度查询-1



图 3-34 子任务历史调度查询-2



调度历史列表中，可以查看每次调度运行的详细日志信息，日志信息如下：

图 3-35 日志信息



3.1.5 功能列表

表 3-1 元数据功能列表

软件名称	元数据
用户管理	管理员用于创建和维护平台的用户信息，并分配响应的角色
数据源	用于管理企业的数据库连接信息
模型架构	为元数据采集制定标准的模型框架

软件名称	元数据
适配器	适配不同的关系型数据库，针对不同的模型架构制定不同的采集脚本
采集任务	创建元数据的采集任务，从不同的数据源中根据不同的模型和适配器，采集所需的元数据信息，包括过程中的数据实体（系统、库、表、字段的描述）

3.2 数据资产用户操作手册

3.2.1 信息架构

图 3-36 信息架构



搜索

对主题域/主题/业务对象进行搜索。

主题域操作

- 新建主题域：在页面左侧信息架构树根节点下新建主题域。

图 3-37 新建主题域



- 编辑主题域：选中要修改的主题域，可重命名/删除主题域

图 3-38 编辑主题域



主题操作

- 新建主题：在选中的数据库节点下新建主题。

图 3-39 新增主题



- 编辑主题：选中要修改的主题，可重命名/删除主题

图 3-40 编辑主题



业务对象操作

- 新增业务对象：在选中的主题节点下新建业务对象。

图 3-41 新增业务对象



- 编辑业务对象：选中要修改的业务对象，可重命名/删除业务对象

图 3-42 编辑业务对象



实体属性操作

- 新增实体属性

步骤1 选中页面左侧信息架构树最末级节点（“业务对象”节点），右侧表格上出现“新增实体”按钮。

步骤2 单击“新增实体”按钮，跳转进入新增实体页面。

步骤3 填写实体信息，新增属性，单击“提交”按钮完成操作。

----结束

图 3-43 新增实体属性-1



图 3-44 新增实体属性-2



- 编辑实体属性：单击右侧表格中的“编辑”，跳转进入编辑实体页面。

图 3-45 编辑实体属性-1



图 3-46 编辑实体属性-2

The interface is titled '修改实体' (Edit Entity). It includes a sidebar for '信息架构' (Information Structure) and a main form with the following fields:

- 主题域: 测试主题域 (Domain)
- 主题: 测试主题 (Subject)
- 业务对象: 测试业务对象 (Business Object)
- 逻辑实体中文名: 测试1 (Logical Entity Chinese Name)
- 用途分类: IT用途 (Usage Category)
- 数据分类: 维度数据 (Data Category)

Below the form is the '属性设置' (Attribute Settings) section, which includes a table:

属性中文名	属性英文	业务属性描述	引用数据标准
<input type="checkbox"/> 测试1	test1	请输入...	DS202111151518202
<input type="checkbox"/> 测试2	test2	请输入...	DS2021111810205911
<input type="checkbox"/> 测试3	test3	请输入...	DS2021111810275712
<input type="checkbox"/> 测试4	test4	请输入...	DS2021110100003

- 批量导入：批量实体信息填入下载的模板后上传，完成批量导入。

图 3-47 批量导入-1

The interface is titled '数据资产管理' (Data Asset Management). It includes a sidebar for '信息架构' (Information Structure) and a main form with the following fields:

- 主题域: 测试主题域 (Domain)
- 主题: 测试主题 (Subject)
- 业务对象: 测试业务对象 (Business Object)

Below the form is the '批量导入' (Batch Import) section, which includes a table:

操作	信息架构名称	逻辑实体中文名	逻辑实体英文名	用途分类	数据分类
<input checked="" type="checkbox"/>	测试主题域/测试主题/测试业务对象	测试1	test1	IT用途	维度数据

图 3-48 批量导入-2

The dialog box is titled '导入数据' (Import Data). It includes a close button (X) and two buttons:

- ↓ 下载模板 (Download Template)
- ↑ 选择文件 (Select File)

At the bottom, there are two buttons: '取消' (Cancel) and '确定' (Confirm).

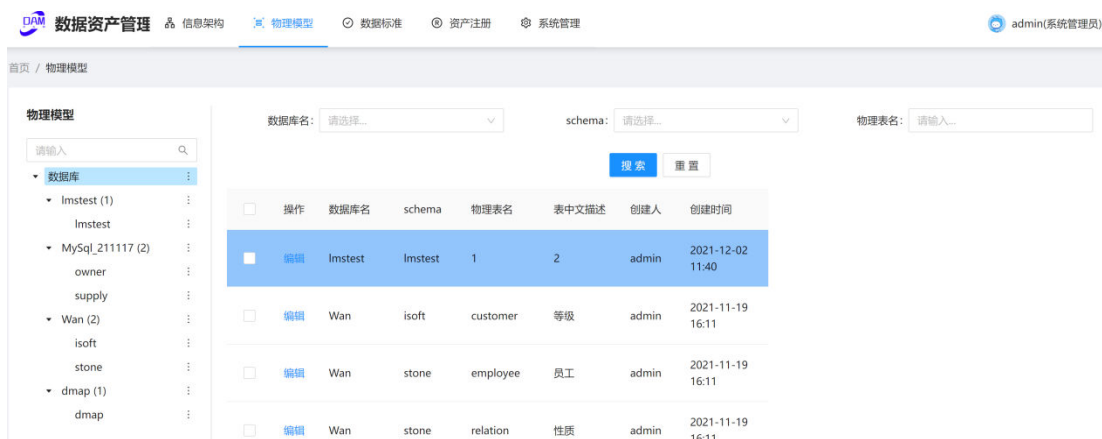
- 批量删除：批量删除所勾选的实体，默认置灰。

图 3-49 批量删除



3.2.2 物理模型

图 3-50 物理模型



搜索

对数据库名/schema/物理表名进行搜索。

数据库操作

- 新增数据库：在物理模型树根节点下新建数据库。

图 3-51 新增数据库



- 编辑数据库：选中要修改的数据库，可重命名/删除数据库

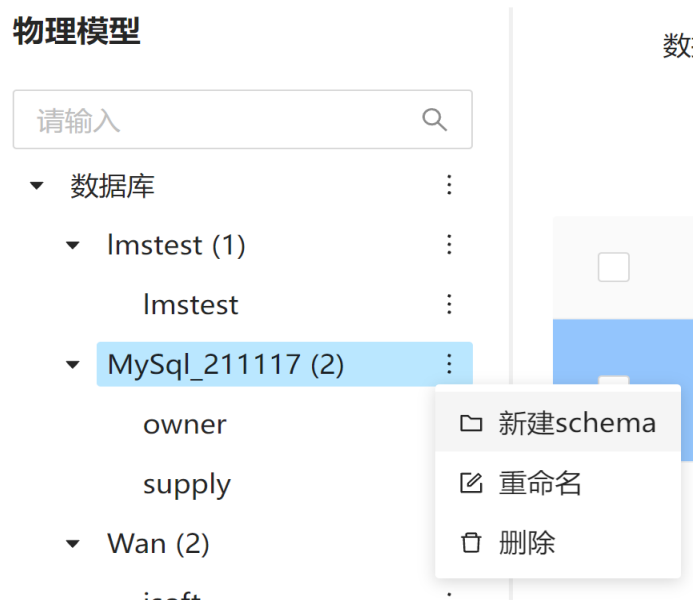
图 3-52 编辑数据库



schema 操作

- 新增schema：在选中的数据库节点下新建schema。

图 3-53 新增 schema



- 编辑schema：选中要修改的schema，可重命名/删除schema

图 3-54 编辑 schema



物理表操作

- 新增物理表

步骤1 选中页面左侧物理模型树最末级节点（“schema”节点），右侧表格上出现“新增物理表”按钮。

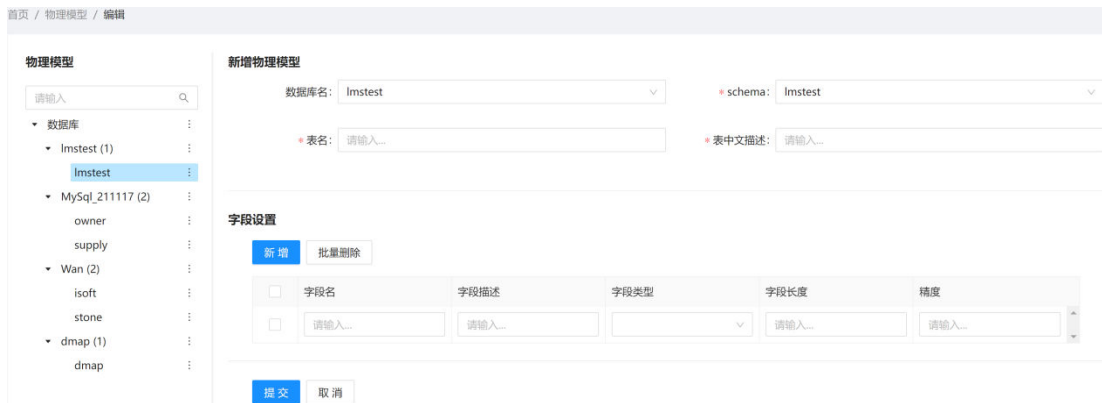
步骤2 单击“新增物理表”按钮，跳转进入新增物理表页面。

步骤3 填写表信息，新增字段，单击“提交”按钮完成操作。
----结束

图 3-55 新增物理表



图 3-56 填写物理表信息



- 编辑物理表：单击右侧表格中的“编辑”，跳转进入编辑物理表页面。

图 3-57 编辑物理表-1

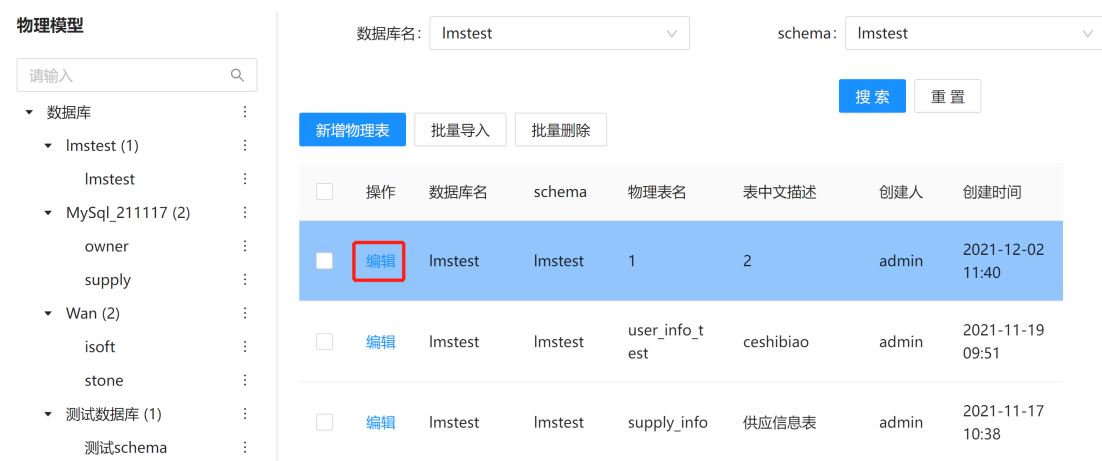
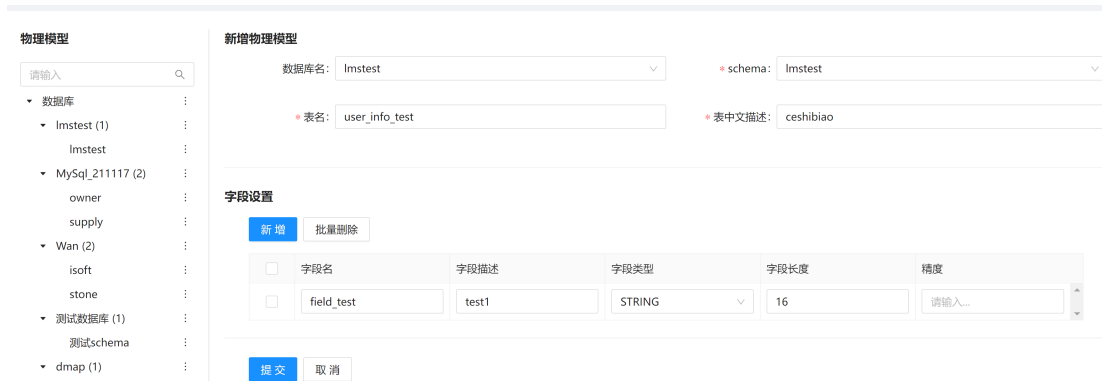


图 3-58 编辑物理表-2



- 批量导入：批量物理表信息填入下载的模板后上传，完成批量导入。

图 3-59 批量导入-1

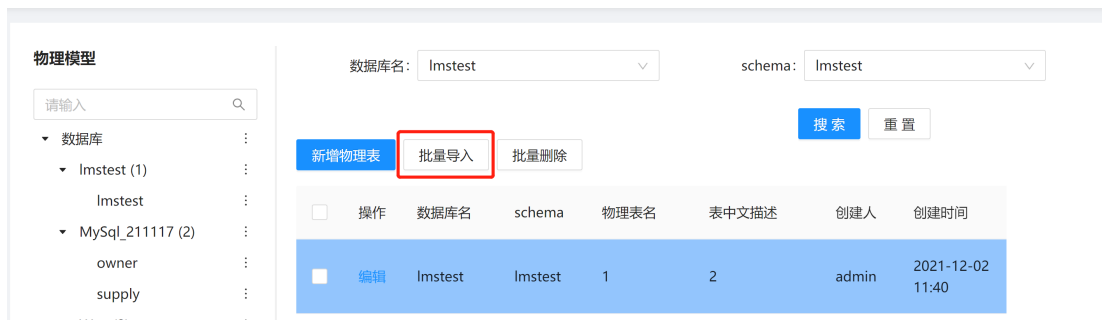
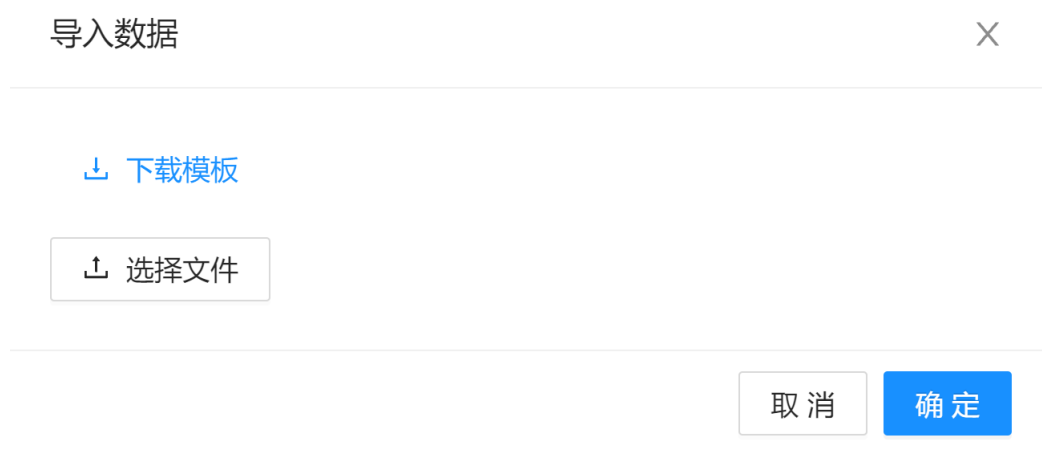


图 3-60 批量导入-2



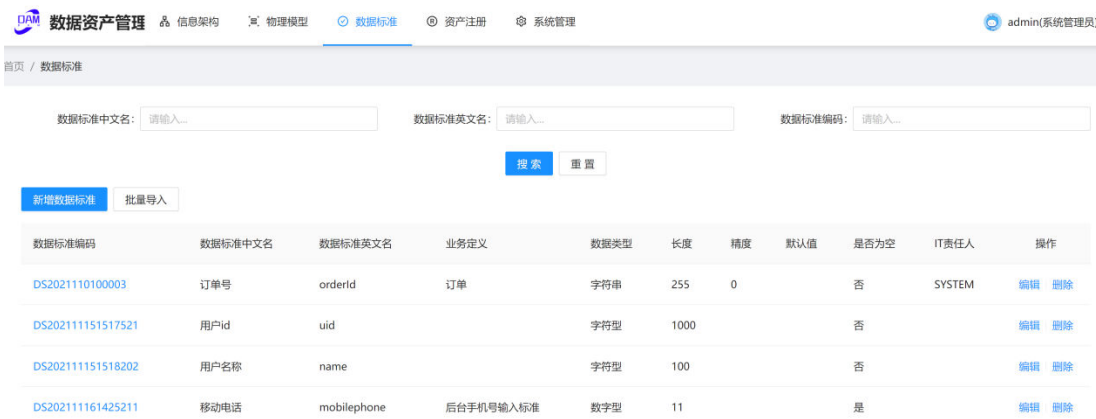
- 批量删除：批量删除所勾选的物理表，默认置灰。

图 3-61 批量删除



3.2.3 数据标准

图 3-62 数据标准



搜索

对数据标准中文名/数据标准英文名/数据标准编码进行模糊搜索。

新增数据标准

图 3-63 新增数据标准

新建×

业务属性

* 数据标准中文名:

* 数据标准英文名:

业务定义:

技术属性

* 数据类型:

长度: 精度:

默认值: 是否允许为空: 是 否

建议字段名:

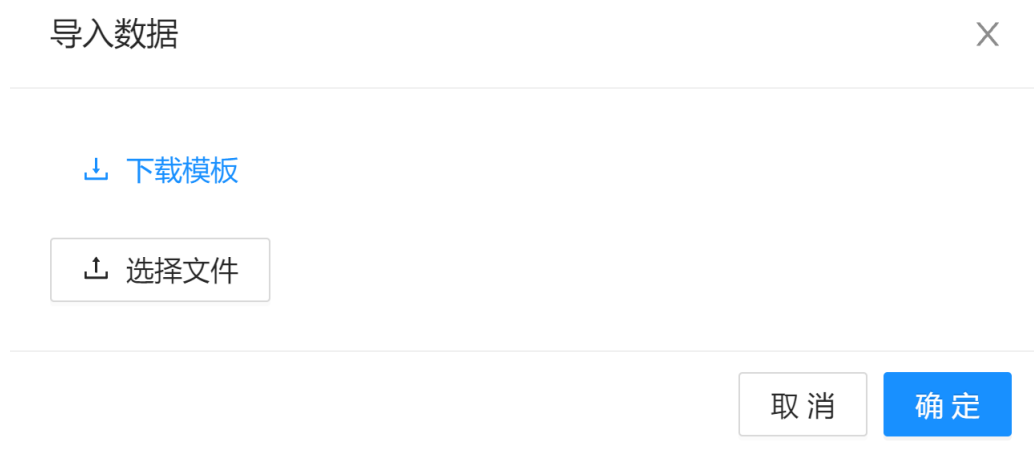
管理属性

IT责任人:

批量导入

批量数据标准信息填入下载的模板后上传，完成批量导入。

图 3-64 批量导入



批量删除

批量删除所勾选的关系清单，默认置灰。

编辑

单击按钮弹出页面进行修改。

图 3-65 编辑

编辑✕

业务属性

* 数据标准中文名:

* 数据标准英文名:

业务定义:

技术属性

* 数据类型:

长度: 精度:

默认值: 是否允许为空: 是 否

建议字段名:

管理属性

IT责任人:

删除

删除单个数据标准。

3.2.4 资产注册

关系注册

注册主题域、主题、业务对象、逻辑实体与技术资产之间的关系

图 3-66 关系注册

注册类型: 逻辑实体与技术资产的关系

* 主题域: 请选择...

* 业务对象: 请选择...

* 数据库名: 请选择...

* 物理表: 请选择...

* 主题: 请选择...

* 逻辑实体: 请选择...

* Schema: 请选择...

是否注册属性: 是 否
是否选择注册属性与字段的关系

描述: 请输入...

关系清单

查看/修改/批量导出已注册的关系清单和属性关系，并能新增注册/批量注册/删除/批量删除属性关系。

图 3-67 关系清单

新增注册 批量注册 批量导出 批量删除

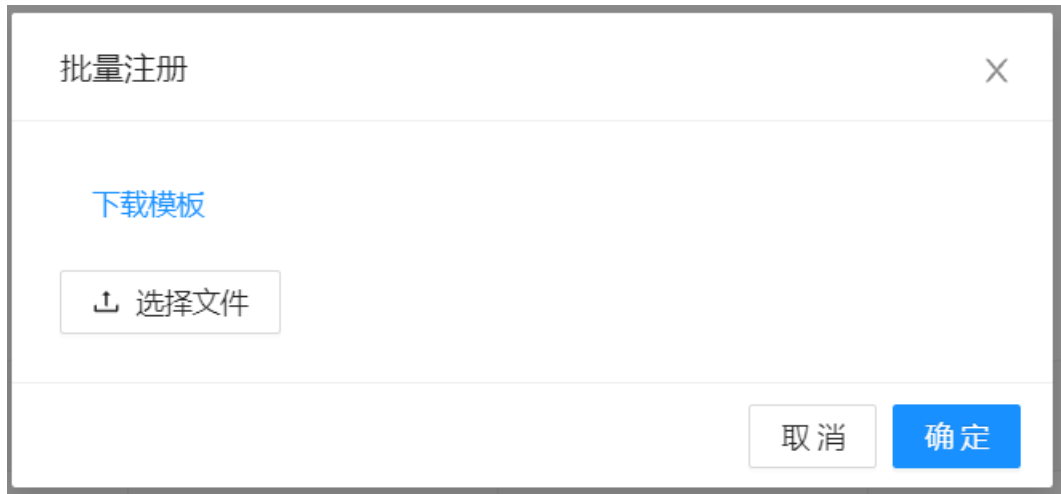
输入逻辑实体, 物理表搜索...

<input type="checkbox"/>	主题域	逻辑实体ID	逻辑实体	物理表	创建人	创建时间	更新人	更新时间	操作
<input type="checkbox"/>				product	admin	2021-11-22 10:11	admin	2021-11-19 16:40	修改 删除 属性关系
<input type="checkbox"/>				supply_info	admin	2021-11-22 09:53	admin	2021-11-19 16:40	修改 删除 属性关系
<input type="checkbox"/>	生产	202137311761718107	手机	customer	admin	2021-11-19 16:50	admin	2021-11-19 16:50	修改 删除 属性关系
<input type="checkbox"/>				user_info	admin	2021-11-19 14:13	admin	2021-11-19 16:40	修改 删除 属性关系
<input type="checkbox"/>				sys_param_info	admin	2021-11-19 14:13	admin	2021-11-19 16:40	修改 删除 属性关系

共 5 条 < 1 > 10 条/页

- 新增注册：跳转至关系注册页面
- 批量注册：批量关系注册信息填入模板后上传，完成批量注册。

图 3-68 批量注册



- 导出：导出勾选的关系清单到本地
- 批量删除：批量删除所勾选的关系清单，默认置灰。
- 搜索：对逻辑实体/物理表名称进行模糊搜索。
- 修改：跳转至关系注册页面进行修改。
- 删除：删除单个关系清单。
- 属性关系：展示已注册关系的逻辑实体中的字段属性描述。

图 3-69 属性关系

业务属性	字段名称
	name
	level

3.2.5 系统管理

可新建系统用户，查看、维护系统用户的相关信息、账户状态和角色。

用户管理

- 新建用户：创建系统新用户

图 3-70 新建用户

新建用户

* 用户登录名: 请输入...

* 用户密码: 请输入...

* 员工工号: 请输入...

* 用户姓名: 请输入...

* 用户类型: 系统管理员 数据管家 普通用户

提交 取消

- 操作：操作功能位于列表最右端，可对系统用户进行查看、编辑或者删除操作。

我的导入

展示已导入模板文件的详细信息，提供搜索功能。

图 3-71 我的导入

数据资产管理 信息架构 物理模型 数据标准 资产注册 系统管理 admin(系统管理员)

文件名: 请输入... 文件类型: 请选择... 导入结果状态: 请选择...

搜索 重置

文件名	模板类型	导入时间	导入结果	详情
-----	------	------	------	----

3.2.6 功能列表

表 3-2 数据资产功能列表

软件名称	数据资产
信息架构	用于创建转换企业的实际业务架构为逻辑实体
物理模型	创建从数据库到表字段信息的物理表模型
数据标准	提供用于创建业务数据的技术标准的功能

软件名称	数据资产
资产注册	注册主题域、主题、业务对象、逻辑实体与技术资产之间的关系
系统管理	提供用户管理及查看导入信息功能

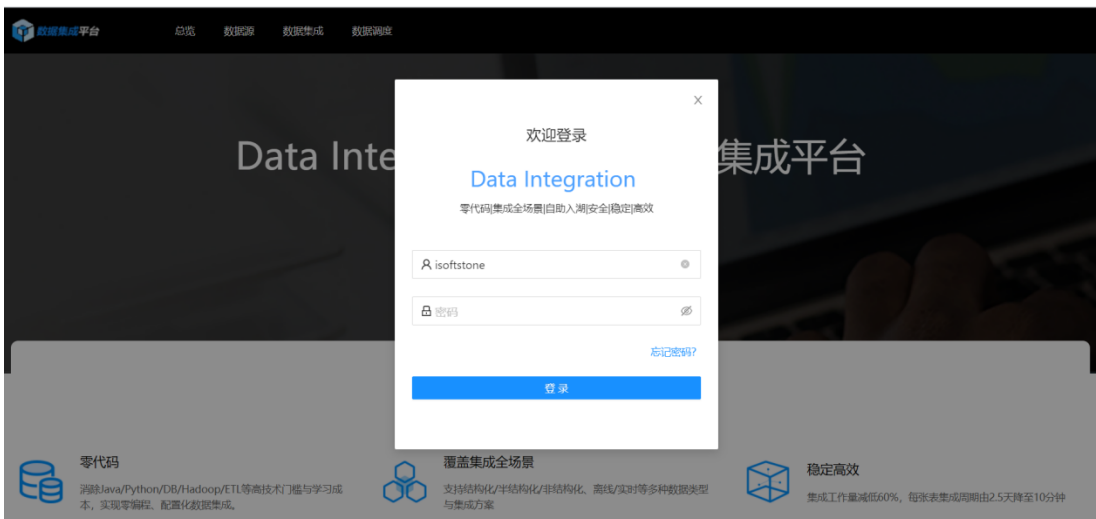
3.3 数据集成平台操作手册

3.3.1 平台界面

首页

输入用户的账号和密码，登录进主页面。

图 3-72 登录界面



总览

总览主要分为三大块内容：分别是集成任务数量图表展示，数据库类型占比图表，调度异常监控图表以及调度异常数据展示。

图 3-73 总览

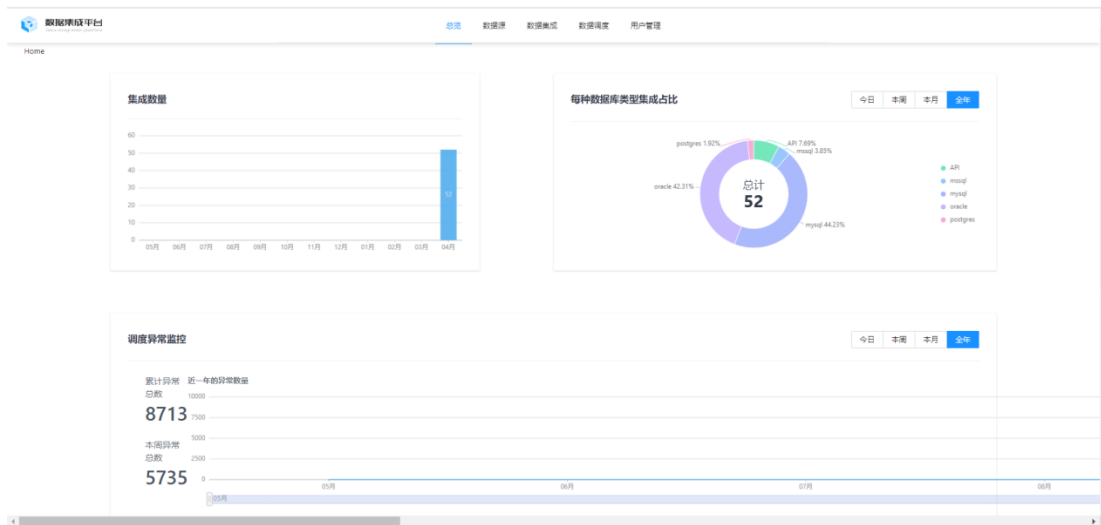


图 3-74 调度异常数据

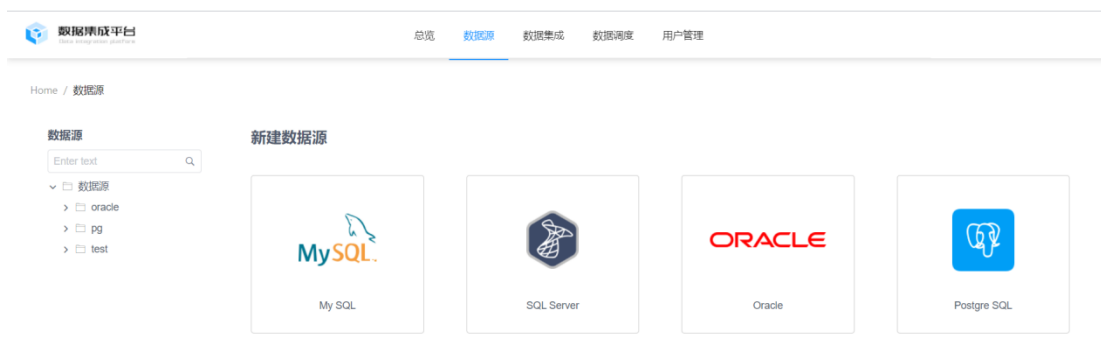
The table displays a list of scheduling anomalies with the following columns: 序号 (Serial Number), 调度异常任务ID (Scheduling Anomaly Task ID), 异常任务责任人 (Anomaly Task Responsible Person), 异常运行类型 (Anomaly Running Type), 异常运行开始时间 (Anomaly Running Start Time), and 异常运行结束时间 (Anomaly Running End Time). The table contains 10 rows of data, all with the same task ID (R2021040600010) and responsible person (dongbin). The running type is consistently '离线-增量更新' (Offline-Incremental Update). The start and end times vary slightly across the rows, indicating multiple instances of the task.

序号	调度异常任务ID	异常任务责任人	异常运行类型	异常运行开始时间	异常运行结束时间
1	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:27:40
2	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:33:39
3	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:37:39
4	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:42:39
5	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:47:39
6	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:52:39
7	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 18:57:40
8	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 19:02:39
9	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 19:07:39
10	R2021040600010	dongbin	离线-增量更新	2021-04-06 18:22:39	2021-04-06 19:12:39

数据源

数据源主要包含两块：左侧树状导航展示数据源路径、右侧可进行新建数据源操作。

图 3-75 新建数据源



右侧可以查看或编辑数据源配置信息，包含数据源连接串、用户名、密码、数据源名称、数据源描述等；

图 3-76 编辑数据源



单击“新建数据源”，添加一个新的数据源，选择数据源类型，这条记录会同步显示在左侧树状导航中。

图 3-77 新建数据源-1

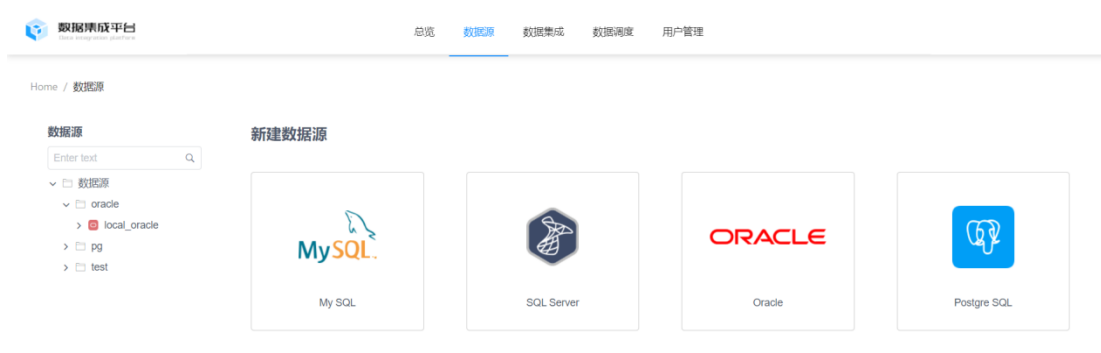


图 3-78 新建数据源-2



3.3.2 数据集成

数据集成主要包含三块：离线集成，实时集成，集成任务。

图 3-79 数据集成



离线集成

主要为结构化集成和非结构化集成两种，结构化集成主要是离线集成部署方式，非结构化集成为API部署方式。

1. 结构化集成

图 3-80 结构化集成



步骤1 集成作业

图 3-81 集成作业



- 任务名称：自定义数据
- 需求类型：目前只支持新增
- 源数据源：为2.3页签数据源，可以单击新增保存新数据源，并且选择数据源后可以单击查看，数据源的相关信息
- 目标数据源：要部署到的目标库数据源，为2.3页签数据源
- 目标表名：为用户手工输入，用户填写有效目标表名后，部署环节会在目标库，schema下面自动建表
- 是否自定义SQL：用户可以自定义初始化，周期任务条件，填写完后必须单击校验且通过方可流转到下一步骤

选择否：为默认的全量，增量初始化和周期任务条件
选择是：用户可以自定义初始化，周期任务SQL条件。

步骤2 参数设置

单击新增，支持常量模式和SQL模式，相当于定义全局变量，可以写入条件sql，用全局变量的方式填入到初始化和周期任务sql里面。

图 3-82 参数设置



图 3-83 参数类型



- 部署方式为全量，初始化抽取SQL为用户必填

图 3-84 部署方式-全量



- 部署方式为增量，初始化sql，增量抽取SQL为必填

图 3-85 部署方式-增量



步骤3 MAPPING映射

页面输入源库，schema，表数据后会带出来源表字段，用户可勾选要部署字段并校验通过。支持复制字段，全部字段和部分字段部署集成功能。

图 3-86 MAPPING 映射



步骤4 集成模式

支持离线全量和离线增量集成

图 3-87 集成模式



步骤5 集成调度

----结束

支持自定义集成时间范围和调度时间。

图 3-88 集成调度



2. 非结构化集成

请求接口方式，集成非结构化任务。

图 3-89 非结构化集成



- 类型：目前支持API拉取，必填
- 任务名称：自定义任务名称，必填
- 增量模式：模式分全量和增量，必填
- API地址：请求API地址，必填
- 请求类型：目前支持GET，POST请求，必填。
- 内容类型：下拉框选择Application/json，必填。
- 超时时间：单位ms，非必填，设置超时时间。
- 一天多调：选择范围为1，2，3，分别对应每天定时调度频率，必填。

- 认证：下拉框选择范围为No Auth，Sgov，Basic Auth三种认证方式，必填。
- 请求参数：自定义请求参数，分静态参数和动态参数，支持删除。

图 3-90 请求参数

参数名称	参数类型	参数值	参数格式	参数位置	操作
<input type="text" value="请输入参数名称"/>	<input type="text" value="请选择参数..."/>	<input type="text" value="请输入参数值"/>	<input type="text" value="请输入参数格式"/>	<input type="text" value="请选择参数..."/>	<input type="text" value=""/>

测试

- 测试：在线校验，并提示校验结果。

实时集成

实时集成方式任务，源端数据变化实时同步到目标端。

步骤1 集成作业

图 3-91 集成作业-1

数据集成平台

总览 数据源 数据集成 数据调度 用户管理

1 集成作业

- 1 集成作业
 - * 任务名称
 - * 源数据源
 - * 源Schema
 - * 源表名
 - * 目标数据源
 - * 目标Schema
 - * 目标表名
- 2 MAPPING映射

图 3-92 集成作业-2



- 任务名称：自定义任务名称
- 源数据源：选择2.3步骤数据源，支持在线新增和查看。
- 源schema：跟据源数据源，选择源端schema，选择了数据源之后自动带出
- 源表名：选择源数据源和源schema后，填写源表，存在提示源端表数据
- 目标数据源：选择目标端数据源
- 目标schema：选择目标端数据源后自动带出
- 目标表名：自定义目标表名，存在同名校验。

步骤2 MAPPING映射

----结束

填写完实时任务参数数据，根据源表字段，可勾选要集成字段，并在线校验。

图 3-93 MAPPING 映射



集成任务

1. 结构化集成

图 3-94 结构化集成

The screenshot shows the '数据集成平台' (Data Integration Platform) interface. The main content area displays a table of integration tasks under the '结构化集成' (Structured Integration) tab. The table has columns for '任务ID', '任务名称', '源Schema', '源表', '部署信息', '查看调度', '需求类型', '申请日期', '目标Schema', and '目标表'. The tasks listed include various data synchronization and integration jobs with their respective statuses (e.g., '部署成功', '部署失败').

任务ID	任务名称	源Schema	源表	部署信息	查看调度	需求类型	申请日期	目标Schema	目标表
S2021050800003	同步dis_product...	data_integration	dis_product_line_info1	部署成功	查看	新增	2021-05-08 16:27	data_integration	dis_product_line_info2
S2021050800002	同步mysql_table...	data_integration	data_integration_api_L...	部署失败	查看	新增	2021-05-08 16:22	data_integration	fqw_test1
S2021050700005	zhj_1	WANIUNIUN	CUSTOMER	部署成功	查看	新增	2021-05-07 17:54	data_discovery	CUSTOMER_copy
S2021050700004	同步mysql_table...	data_integration	data_integration_api_L...	部署成功	查看	新增	2021-05-07 16:15	data_integration	data_integration_api_L...
S2021050700003	同步mysql_table...	data_integration	data_integration_api_L...	部署成功	查看	新增	2021-05-07 16:04	data_integration	data_integration_api_L...
S2021050700002	mysql_add_test_...	data_master	data_integration_task...	部署成功	查看	新增	2021-05-07 11:14	data_test_48	data_integration_task
S2021050700001	mysql_48_test_0...	data_master	data_integration_task...	部署成功	查看	新增	2021-05-07 10:45	data_master	data_integration_asss...

- 新建：用户操作后跳转页面，新建集成结构化集成任务
- 部署：结构化集成任务订单部署任务，正在部署或者部署完成订单不可操作部署。
- 重新部署：针对订单部署失败订单，重新部署操作。并且只能部署失败状态的订单。
- 导出：勾选列表订单批量导出，不勾选默认为当前页面所有订单
- 删除：针对未部署订单，可以执行删除操作
- 刷新：刷新页面订单的最新数据。
- 搜索：支持任务ID，任务名称或源表名称条件搜索，模糊搜索和精确搜索。
- 分页：支持选择当前页面展示数量，可以手工翻页上一页下一页，也可以跳转到任意页面列表订单操作
- 查看调度：操作后跳转到调度页面，查看订单的调度详情数据
- 详情：操作后可以查看到结构化集成订单的详情数据

2. 非结构化集成

图 3-95 非结构化集成

任务ID	任务名称	API地址	调度周期	认证	目标数据库	目标表
API2021050600002	renewewe	http://10.28.117.221:8080/tit...	1次	none	data_integration	data_integration_api_table2
API2021050600001	rrrr	http://10.28.117.221:8080/tit...	2次	none	data_integration	data_integration_api_table1
API2021043000001	20210430	http://10.28.117.221:8080/tit...	2次	none	data_integration	tb_user_111
API2021041200001	API_titles_0412_01	http://10.28.117.221:8080/tit...	3次	none	data_integration	API_titles_0412_01
API2021040600003	jq_api_0406	http://10.28.117.221:8080/tit...	2次	none	data_integration	jq_api_0406
API2021040600002	测试部署任务001	http://10.28.117.181:8011/in...	2次	none	superset_test	测试部署表名
API2021040600001	JQ111	http://10.28.117.221:8080/tit...	2次	none	data_integration	JQ_1111

- 新建：操作后跳转到非结构化集成新建任务页面
- 导出：勾选列表订单批量导出，不勾选默认为当前页面所有订单
- 刷新：刷新页面订单的最新数据
- 搜索：支持任务ID，任务名称或源表名称条件搜索，模糊搜索和精确搜索。

3.3.3 数据调度

调度管理

调度管理为离线任务部署管理页面，所有离线订单的部署任务汇总调度维护。部署任务订单状态成功，失败，并可以查看历史调度和操作日志。

图 3-96 调度管理-1

任务ID	任务名称	任务状态	集成频率	调度状态	计划时间	责任人	操作
S2021050800003	同步dis_product_line	运行中	CRON : 04 27 16 * ...	成功	2021-05-12 16:27:04	SYSTEM	执行 历史调度 操作日志
S2021050700005	zhj_1	运行中	CRON : 36 53 17 * ...	成功	2021-05-12 17:53:36	SYSTEM	执行 历史调度 操作日志
S2021050700004	同步mysql_table004	运行中	CRON : 25 14 16 * ...	成功	2021-05-12 16:14:25	SYSTEM	执行 历史调度 操作日志
S2021050700003	同步mysql_table002	运行中	CRON : 53 02 16 * ...	成功	2021-05-12 16:02:53	SYSTEM	执行 历史调度 操作日志
S2021050700002	mysql_add_test_0507	运行中	CRON : 17 14 12 * ...	失败	2021-05-12 12:14:17	SYSTEM	执行 历史调度 操作日志
S2021050700001	mysql_48_test_0507	运行中	CRON : 04 49 10 * ...	失败	2021-05-12 10:49:04	SYSTEM	执行 历史调度 操作日志
S2021050600001	同步mysql_table001	运行中	FIX_RATE : 7200	失败	2021-05-12 10:52:53	SYSTEM	执行 历史调度 操作日志
API2021050600002	renewewe	运行中	CRON : 0 6 1 * * ?	成功	2021-05-13 01:06:00	SYSTEM	历史调度 操作日志

图 3-97 调度管理-2

任务ID	任务名称	任务状态	集成频率	调度状态	计划时间	责任人	操作
S2021050800003	同步dis_product_line	运行中	CRON : 04 27 16 * ...	成功	2021-05-12 16:27:04	SYSTEM	执行 历史调度 操作日志
S2021050700005	zhj_1	运行中	CRON : 36 53 17 * ...	成功	2021-05-12 17:53:36	SYSTEM	执行 历史调度 操作日志
S2021050700004	同步mysql_table004	运行中	CRON : 25 14 16 * ...	成功	2021-05-12 16:14:25	SYSTEM	执行 历史调度 操作日志
S2021050700003	同步mysql_table002	运行中	CRON : 53 02 16 * ...	成功	2021-05-12 16:02:53	SYSTEM	执行 历史调度 操作日志
S2021050700002	mysql_add_test_0507	运行中	CRON : 17 14 12 * ...	失败	2021-05-12 12:14:17	SYSTEM	执行 历史调度 操作日志
S2021050700001	mysql_48_test_0507	运行中	CRON : 04 49 10 * ...	失败	2021-05-12 10:49:04	SYSTEM	执行 历史调度 操作日志
S2021050600001	同步mysql_table001	运行中	FIX_RATE : 7200	失败	2021-05-12 10:52:53	SYSTEM	执行 历史调度 操作日志
API2021050600002	renewewe	运行中	CRON : 0 6 1 * * ?	成功	2021-05-13 01:06:00	SYSTEM	历史调度 操作日志

- 启动：任务状态为停止方可启动，默认置灰不可选
- 停止：运行状态才可以停止，默认置灰不可选
- 重新初始化：运行中状态订单可以执行，默认置灰不可选
- 导出：勾选列表订单批量导出，不勾选默认为当前页面所有订单
- 刷新：刷新页面订单的最新数据。
- 搜索：支持任务ID条件搜索，可以模糊搜索或者精确条件
- 列表订单操作：维护集成任务订单，操作字段可以查看历史调度，操作日志数据。并且集成任务部署完成了之后，可以手工操作执行调度，输入调度参数

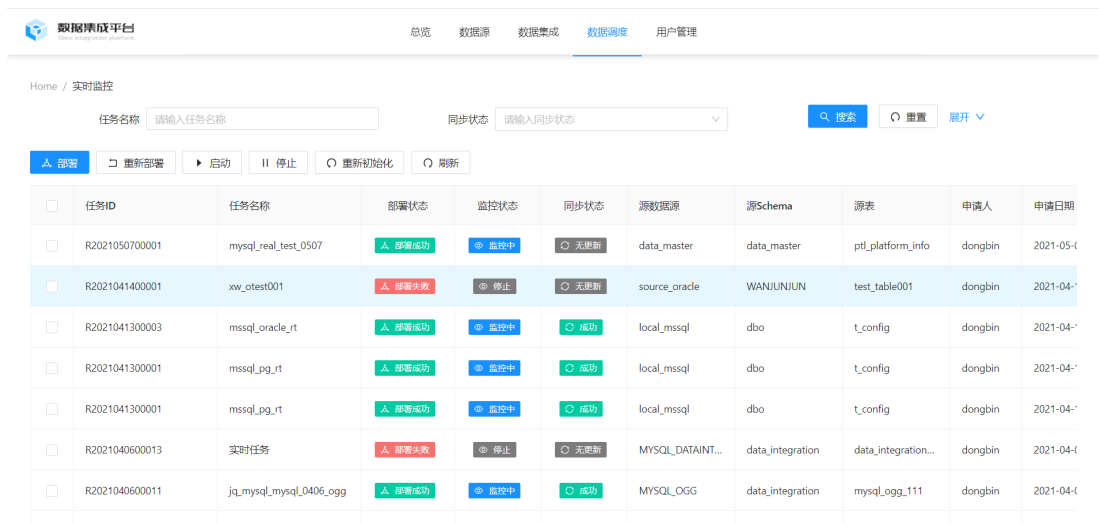
图 3-98 列表订单操作



实时监控

实时集成任务部署订单，汇总维护实时任务部署订单。可以查看历史记录，并手工触发部署任务

图 3-99 实时监控



- 条件搜索：支持任务名称，同步状态，源数据源，源表，申请人，目标数据源，目标表单一条件或者组合条件搜索，并可模糊搜索/精确条件搜索
- 部署：任务状态为未启动可以手工触发部署
- 重新部署：任务状态为失败方可手工触发，重新部署功能
- 启动：只能启动监控停止或失败且部署成功的任务
- 停止：只能停止监控中切部署成功的任务
- 重新初始化：任务状态不限状态可以执行重新初始化功能
- 刷新：刷新当前页面，同步最新页面数据
- 列表订单操作：字段操作可以查看历史记录，以及部署任务状态

3.3.4 功能列表

表 3-3 数据集成平台功能列表

软件名称	数据集成平台
软件功能项	功能描述
portal登录	用户登录/退出、产品宣传视频、产品介绍
首页	产品优势介绍，应用场景，功能说明以及帮助文档
总览	集成任务数据分析以及时间轴线统计
数据源	新建/编辑/删除数据源、数据源搜索、数据源树状展示
离线集成	集成全量，增量任务，并支持自定义SQL，生成集成任务
实时集成	实时集成任务编写提交，生成实时任务
集成任务	集成任务管理页面，针对集成任务订单部署，停止订单，导出列表数据，支持任务ID，任务名称，数据源条件搜索

软件名称	数据集成平台
调度管理	集成任务订单对应调度管理页面，可手工启动停止，手工触发调度，并传入时间范围。也可以查看历史调度和操作记录
实时监控	实时任务订单对应调度管理页面，可手工启动停止，手工触发部署，支持历史记录查看

3.4 数据地图平台操作手册

3.4.1 数据全景

数据全景支持根据关键字或者主题域分组来查询数据资产信息，帮助用户能快速找到需要的数据。

图 3-100 数据全景



- 搜索：通过关键字搜索所需的资产信息
- 搜索列表：展示搜索到的技术资产和业务资产

图 3-101 搜索列表



- 技术资产详情：展示所选技术资产名的物理表、逻辑实体、字段、数据评论等详细信息

图 3-102 技术资产详情

数据地图 | 数据全景 | 资产注册 | 数据运营 | 数据质量 | 标签管理 | 系统管理 | 主题域分組责任人001

首页 / 数据全景 / 搜索列表 / 技术资产详情

技术资产详情

do_ao_entity_table_rel

系统名称: 数据地图	数据库名称: DMAP	Schema名称: dmap
物理实体类型: BASE TABLE	更新频率: 1天3调	是否已经入湖: 否
源系统表: do_ao_entity_table_rel	数据更新时间:	总存储量: 2.08G
IT责任人: liuqibin	记录数: 6	

逻辑实体

逻辑实体编码	逻辑实体中文名	逻辑实体英文名	数据管家	是否主数据	安全密级
	资产移交-清理	资产移交-清理			内部公开

字段

字段名称	关联属性名称	字段备注	数据类型	数据长度	主键标志	可控标志
rel_id		关系表_ID	decimal		PR	
entity_id		逻辑实体_ID	varchar	255	MU	
table_id		表ID	varchar	255	MU	
table_type		物理对象类型.0表, 1视...	decimal			
system_id		IT 系统ID	decimal			
is_source_sys		是否源系统	char	1		
is_source		是否可信源	char	1		
maintain_by		维护人	varchar	100		
rel_by		关联人	varchar	100		
rel_date		关联时间	datetime			
secret_level		密级	varchar	50		
modify_by		最后修改人	varchar	100		
modify_date		最后修改时间	datetime			
record_date		记录创建时间	datetime			
delete_flag		逻辑删除标志	char	1		
create_by		创建人	varchar	100		
last_update_date		记录最后更新日期	timestamp			

数据点评

我的评论 ★★★★★ 5

0 / 200

提交评论

大家评论: ★★★★★ 4.7

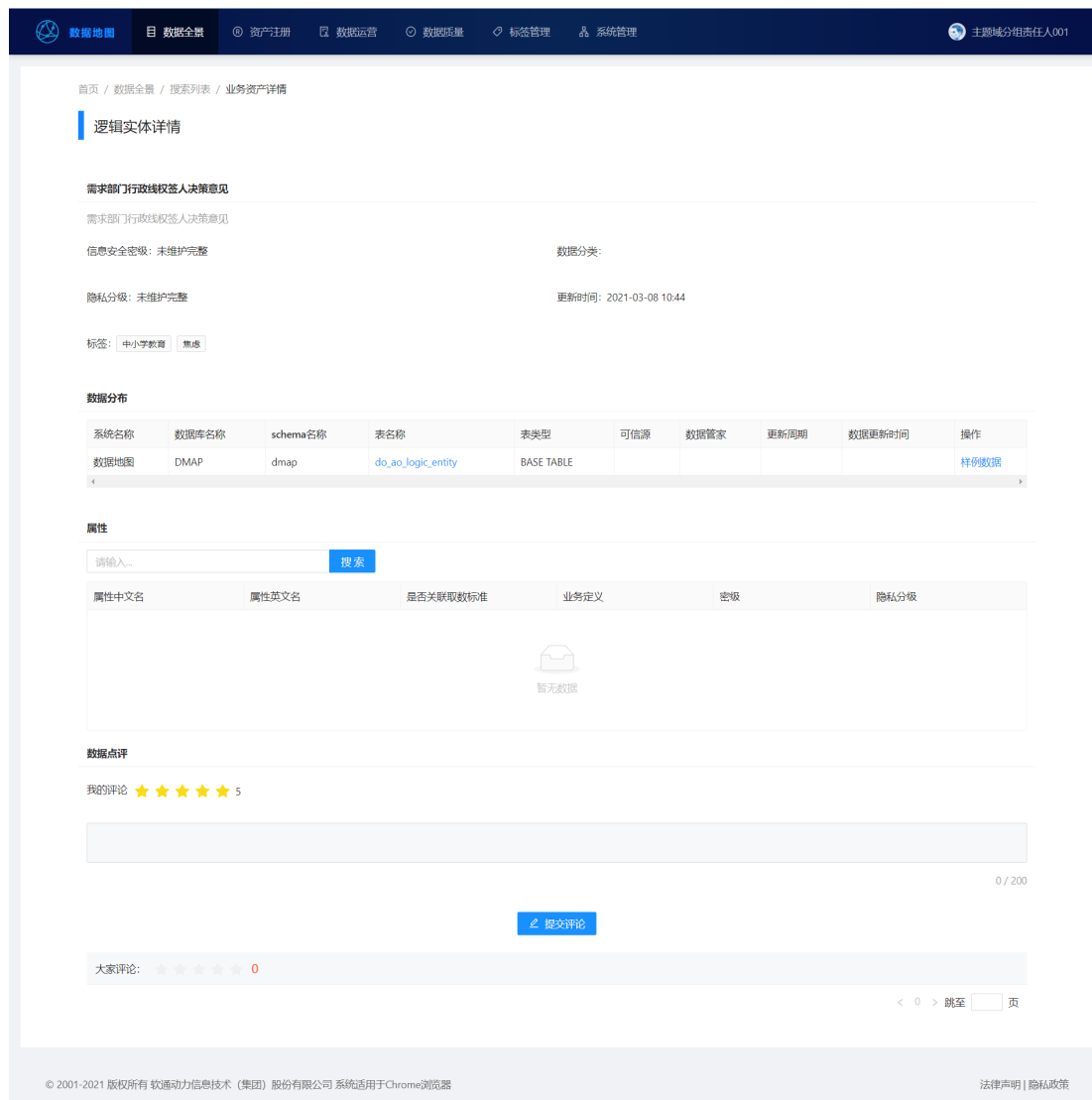
- ★★★★★ 好 isoftstone 2021-04-30 10:30
- ★★★★★ 11 isoftstone 2021-04-21 11:11
- ★★★★★ 123 isoftstone 2021-04-21 11:10

< 1 > 跳至 页

© 2001-2021 版权所有 软通动力信息技术(集团)股份有限公司 系统适用于Chrome浏览器 法律声明 | 隐私政策

- 逻辑实体详情：展示所选业务资产名的业务对象、数据分布、属性、数据评论等详细信息；通过单击下钻查看对应的技术资产详情和样例数据。

图 3-103 逻辑实体详情



- 样例数据：展示所选逻辑实体中对应数据表的样例数据

图 3-104 样例数据

.d...	meta_from	entity_define	delete_flag	last_updat...	object_ow...	entity_path	record_date	auditor	assets_cata...	entity_id	e:
		与会人共享...				Common/...				1	
		数据集-IT产...		161512857...		Manage BT...	161512857...			10	
		数据集-用...		161512857...		Manage BT...	161512857...			11	
		数据集-BO...		161512903...		IPD/IPD分...	161512903...			12	
		数据集-BO...		161512903...		IPD/IPD分...	161512903...			13	
		数据集-CC...		161512903...	fanqiwang	IPD/IPD分...	161512903...			14	
		数据集-CFG...		161512903...		IPD/IPD分...	161512903...			15	
		数据集-CFG...		161512903...		IPD/IPD分...	161512903...			16	
		数据集-Rel...		161512903...		IPD/IPD分...	161512903...			17	
		拟发放录用...		161517147...		Manage HR...	161517147...			18	

- 元数据列表：由数据全景首页，搜索栏下方的主题域分组进入，展示各主题域分组的主题域、业务对象和逻辑实体信息。

图 3-105 元数据列表

数据地图 | 数据全景 | 资产注册 | 数据运营 | 数据质量 | 标签管理 | 系统管理 | 主题域分组责任人001

首页 / 数据全景 / 元数据列表

元数据列表

请输入... 搜索

主题域组: 全部 **Manage HR** Common Lead to Cash 运营分析数据 经营分析数据 IPD Manage BT&IT

主题域: 全部 招聘调配 绩效管理 人员基础应用 组织基础 组织发展和有效性 资源部署与流动 HR基础服务 组织基础应用 人员基础 人才识别与发展
学习管理 员工关系与工作体验 干部管理 人与组织的关系 HR分析域 绩效管理

业务对象: 逻辑实体

候选人基本信息

业务定义: 候选人基本信息 归属领域: Manage HR
数据OWNER: 数据来源: 康系统

人员基础配置

- 业务对象详情：展示所选业务对象名的基础信息、逻辑实体等详细信息。

图 3-106 业务对象详情



- 逻辑实体详情：（同业务对象详情）

3.4.2 资产注册

通过逻辑实体与技术资产的关系注册，明确技术资产对应的逻辑实体、业务对象、主题域、主题域组，并定义属性与字段的关系、每个属性的密级和隐私，来确保数据的安全性。

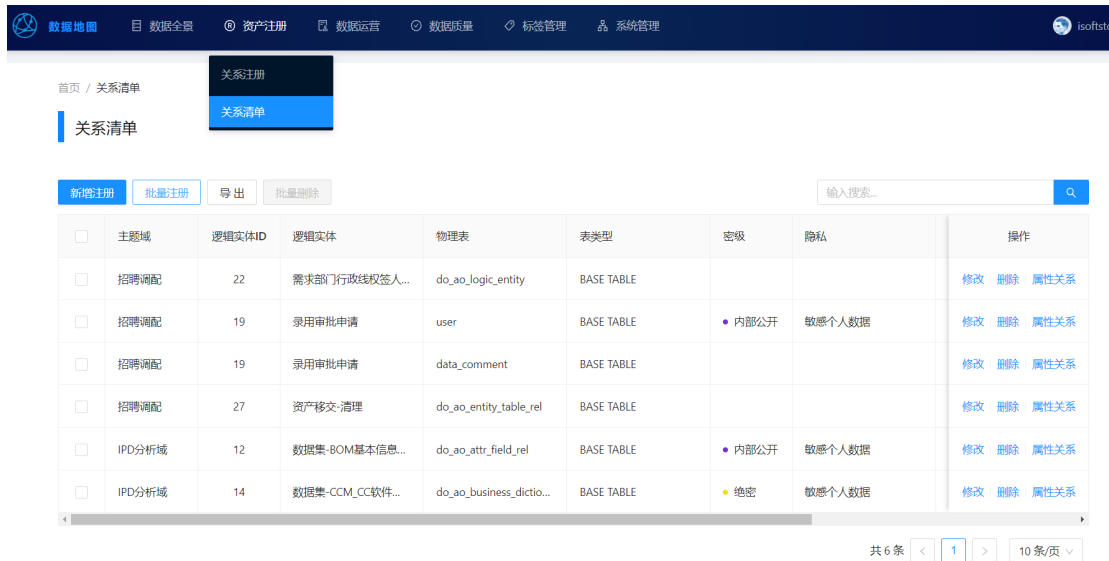
- 关系注册：注册主题域组、业务对象、逻辑实体与技术资产之间的关系。

图 3-107 关系注册



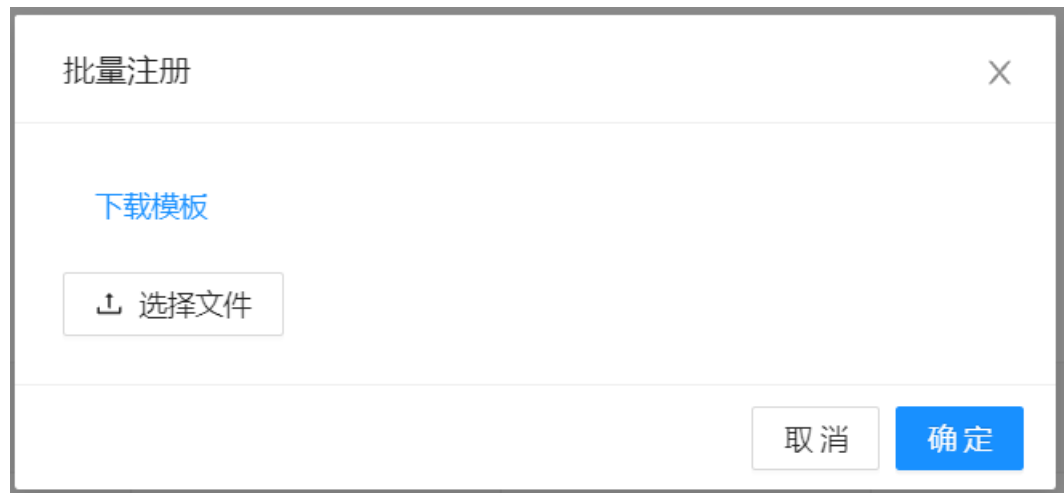
- 关系清单：查看/修改/批量导出已注册的关系清单和属性关系，并能新增/批量注册/删除/批量删除关系

图 3-108 关系清单



- 新增注册：跳转至关系注册页面
- 批量注册：批量关系注册信息填入模板后上传，完成批量注册。

图 3-109 批量注册



- 导出：导出勾选的关系清单到本地。
- 批量删除：批量删除所勾选的关系清单，默认置灰。
- 搜索：对逻辑实体名称进行模糊搜索。
- 修改：跳转至关系注册页面进行修改。
- 删除：删除单个关系清单。
- 属性关系：展示已注册关系的逻辑实体中的字段属性描述。

图 3-110 属性关系

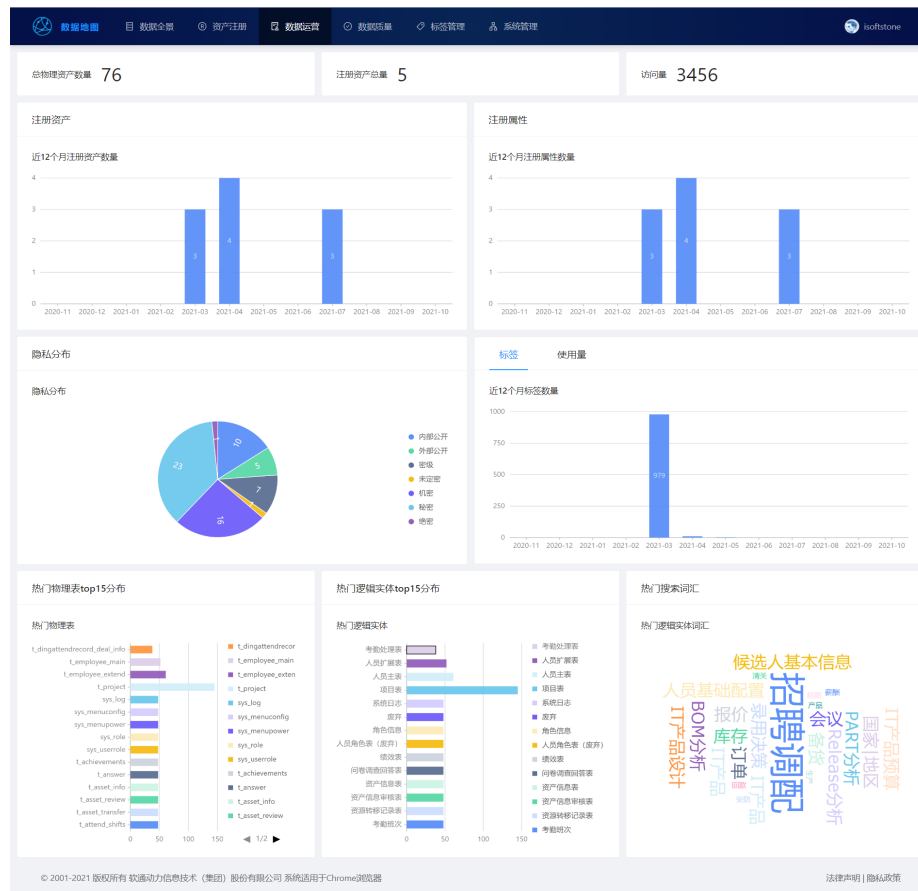
业务属性	字段名称	密级	隐私	描述	创建人	创建时间	更新人	更新时间
HC号	uid	外部公开	商业联系个...	商业联系个...	isoftstone	2021-04-21 11:27	isoftstone	2021-04-21 11:27
一级业务属性	password	内部公开	商业联系个...	商业联系个...	isoftstone	2021-04-21 11:27	isoftstone	2021-04-21 11:27
二级业务属性	salt	外部公开	敏感个人数据	敏感个人数据	isoftstone	2021-04-21 11:27	isoftstone	2021-04-21 11:27
三级业务属性	username	未定密	一般个人数据	一般个人数据	isoftstone	2021-04-21 11:27	isoftstone	2021-04-21 11:27

3.4.3 数据运营与质量

数据运营

展示数据的注册、使用情况；同时将热门物理表、逻辑实体和热搜词直观展示给用户，以供参考决策。

图 3-111 数据运营



- 总物理资产数量：系统中已存在的总的物理表数量
- 注册资产总量：系统中已注册关系的资产数量
- 访问量：访问系统的用户数量
- 注册资产：展示近12个月中注册的关系资产的数量分布
- 注册属性：展示近12个月中注册的属性关系的数量分布
- 隐私分布：查看属性关系数据中字段的隐私属性分布
- 标签/使用量：展示近12个月标签的创建数量和使用量
- 热门物理表top15分布：展示物理表使用最频繁的TOP15
- 热门逻辑实体top15分布：展示逻辑实体使用最频繁的TOP15
- 热门搜索词汇：展示用户搜索最频繁的内容

数据质量

给用户提供一个反馈数据问题的渠道，加强用户体验感、参与感；同时给数据质量核验增加用户渠道，共同打造一个准确，高质量的一个数据查询平台。

图 3-112 数据质量

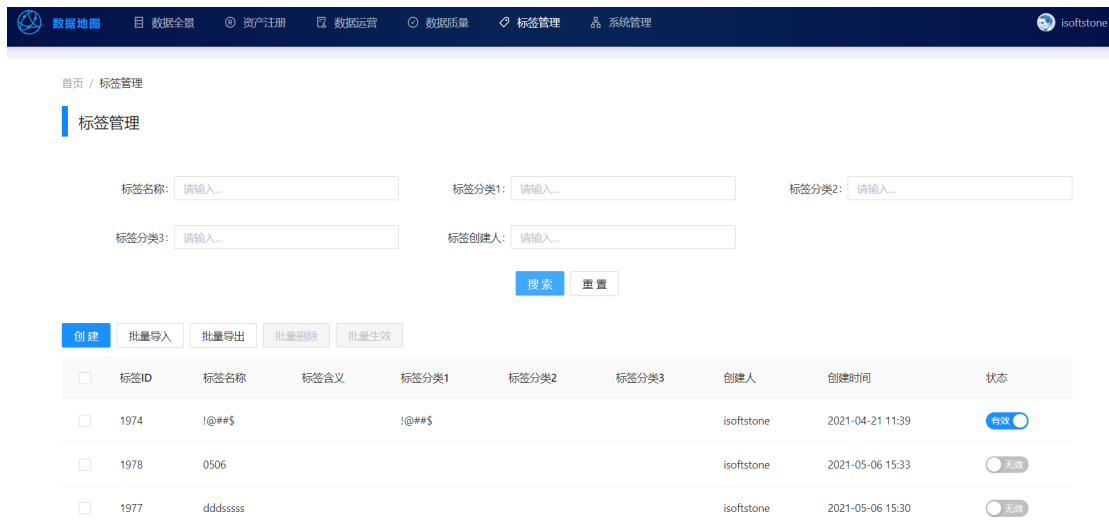
The screenshot shows a web interface for '数据质量' (Data Quality). At the top, there is a navigation bar with tabs for '数据地图', '数据全景', '资产注册', '数据运营', '数据质量', '标签管理', and '系统管理'. The '数据质量' tab is active. Below the navigation bar, there is a breadcrumb '首页 / 数据质量' and a title '数据质量'. A progress indicator shows four steps: 1. 问题反馈 (Problem Feedback), 2. 问题处理 (Problem Processing), 3. 问题验收 (Problem Acceptance), and 4. 问题闭环 (Problem Closure). The current step is 1. The form contains several input fields: '主题域组' (Theme Domain Group) and '主题域' (Theme Domain) are dropdown menus; '业务对象' (Business Object) and '逻辑实体' (Logical Entity) are also dropdown menus; '描述' (Description) is a text input field. Below these fields, there is a label '下一级问题处理人:' (Next level problem handler:). At the bottom of the form, there are two buttons: '提交' (Submit) and '重置' (Reset).

3.4.4 标签管理

标签管理

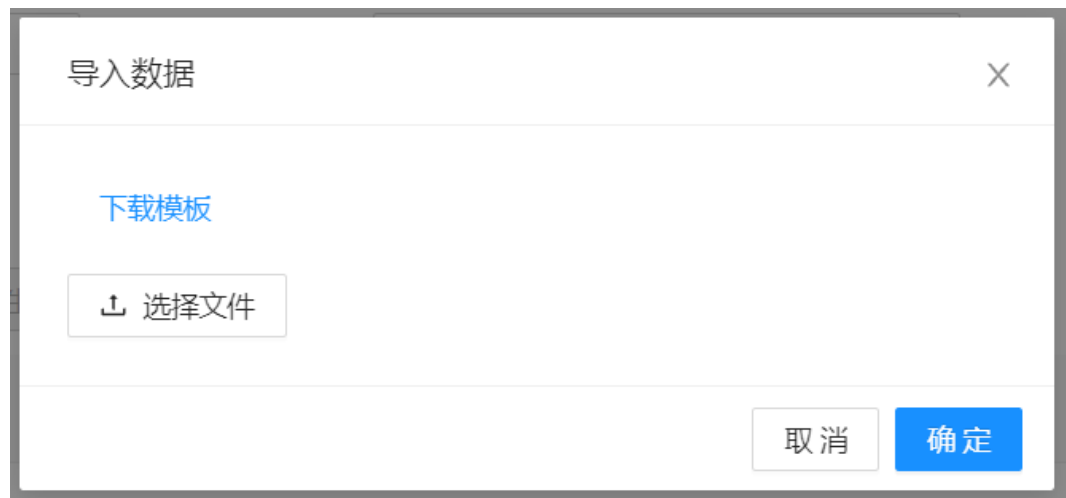
创建、删除标签，修改标签状态，供标记数据资产时使用。

图 3-113 标签管理



- 创建：创建新的标签
- 批量导入：批量创建新的标签

图 3-114 导入数据



- 批量导出：批量导出勾选的标签
- 批量删除：批量删除不需要的标签
- 批量生效：批量修改勾选的标签的状态为有效状态
- 状态：修改标签的状态

资产标签

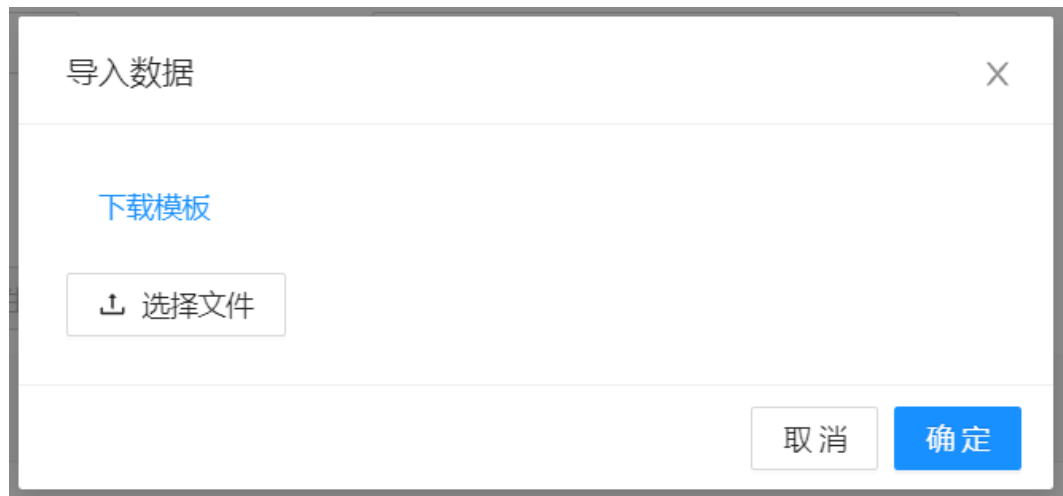
对资产数据进行标记，便于用户依照标签分类快速获取数据并进行分析。

图 3-115 资产标签



- 创建：创建新的资产标签
- 批量导入：批量创建新的资产标签

图 3-116 导入数据



- 批量导出：批量导出勾选的资产标签
- 批量删除：批量删除勾选的资产标签
- 批量生效：批量修改勾选的资产标签状态为生效状态
- 状态：修改资产标签的状态

3.4.5 系统管理

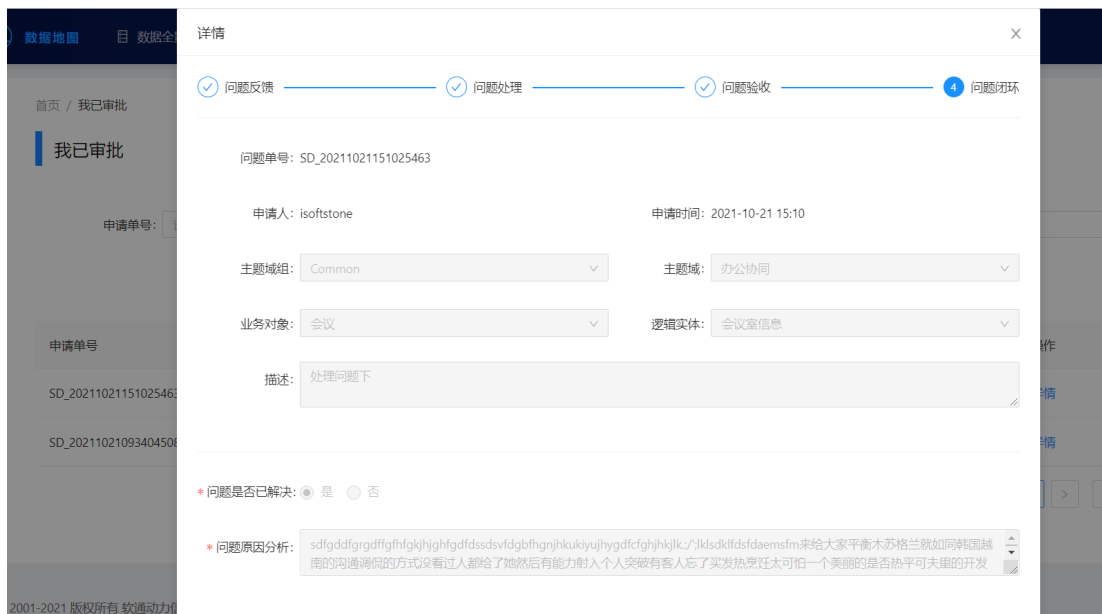
用户可处理待办，查看电子流，管理系统用户。

- 我已审批：查看用户已经参与过的电子流的信息，可在操作列进入详情查看电子流中问题详细信息。

图 3-117 我已审批

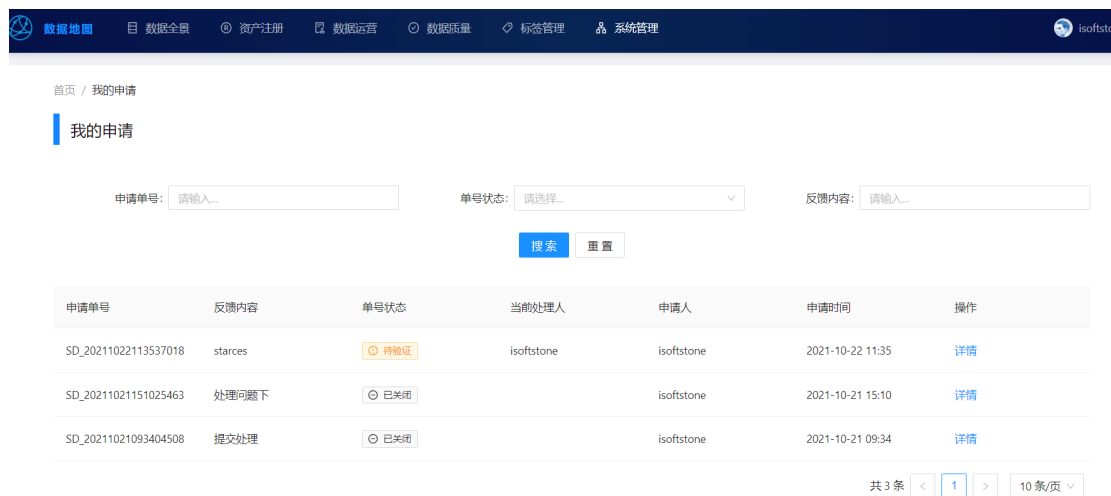


图 3-118 审批详情



- 我的申请：查看用户发起申请的电子流的信息，可在操作列进入详情查看电子流中问题详细信息。

图 3-119 我的申请



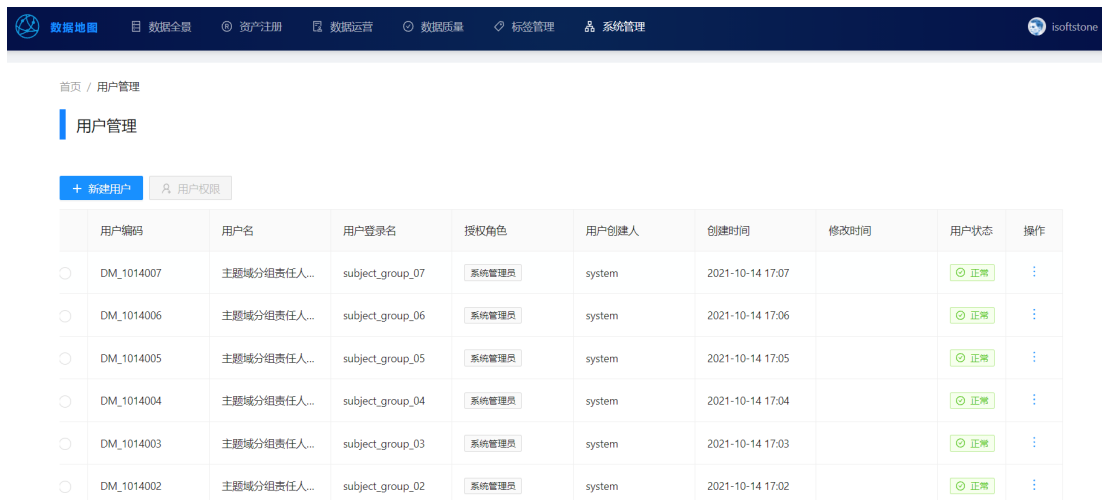
- 我的待办：查看待用户处理的电子流的信息，可在操作列进入详情来处理电子流，问题处理人不能驳回电子流。

图 3-120 我的待办



- 用户管理：可新建系统用户，查看、维护系统用户的相关信息、账户状态和系统权限。

图 3-121 用户管理



The screenshot shows a web interface for user management. At the top, there is a navigation bar with icons for '数据地图', '数据全景', '资产注册', '数据运营', '数据质量', '标签管理', and '系统管理'. Below the navigation bar, the page title is '用户管理'. There are two buttons: '+ 新建用户' and '用户权限'. The main content is a table with the following columns: '用户编码', '用户名', '用户登录名', '授权角色', '用户创建人', '创建时间', '修改时间', '用户状态', and '操作'. The table contains 7 rows of user data.

用户编码	用户名	用户登录名	授权角色	用户创建人	创建时间	修改时间	用户状态	操作
DM_1014007	主题域分组责任人...	subject_group_07	系统管理员	system	2021-10-14 17:07		正常	
DM_1014006	主题域分组责任人...	subject_group_06	系统管理员	system	2021-10-14 17:06		正常	
DM_1014005	主题域分组责任人...	subject_group_05	系统管理员	system	2021-10-14 17:05		正常	
DM_1014004	主题域分组责任人...	subject_group_04	系统管理员	system	2021-10-14 17:04		正常	
DM_1014003	主题域分组责任人...	subject_group_03	系统管理员	system	2021-10-14 17:03		正常	
DM_1014002	主题域分组责任人...	subject_group_02	系统管理员	system	2021-10-14 17:02		正常	

- 新建用户：创建系统新用户

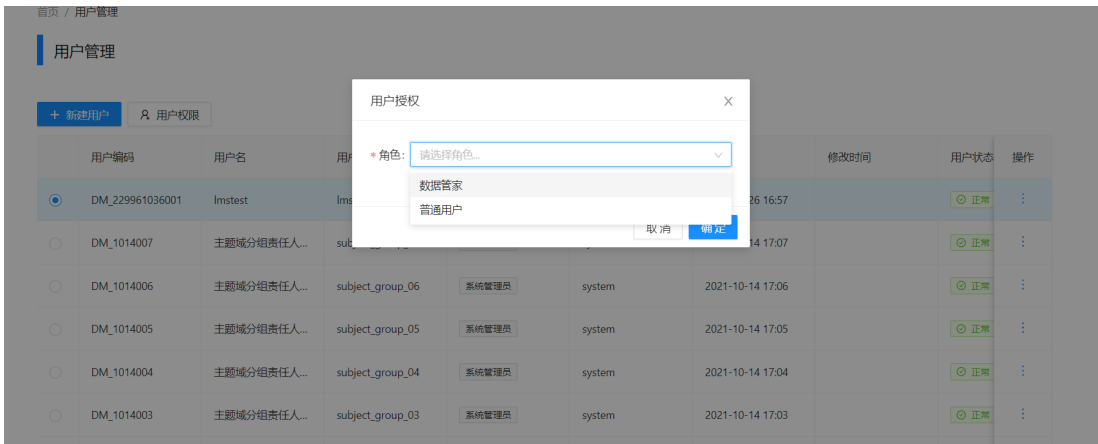
图 3-122 新建用户



The screenshot shows a '新建用户' (New User) form. The form is divided into several sections. The first section contains '用户姓名' (User Name) and '用户登录名' (User Login Name), both with '请输入...' (Please enter...) input fields. The second section contains '密码' (Password) and '用户状态' (User Status). The password field has '请输入...' and a clear button. The user status has radio buttons for '正常' (Normal) and '禁用' (Disabled). The third section contains '用户联系电话' (User Contact Phone) and '用户联系邮箱' (User Contact Email), both with '请输入...' input fields. The fourth section contains '用户年龄' (User Age) with a numeric input field containing '1', and '用户所属公司' (User Company) with '请输入...' input field. The fifth section contains '用户所属部门' (User Department) and '用户职位' (User Position), both with '请输入...' input fields. The final section is '备注' (Remarks) with a '请输入...' input field.

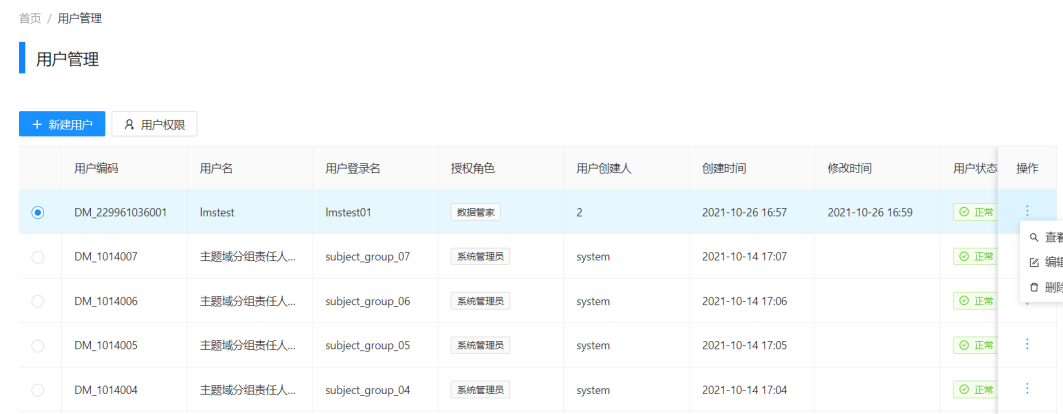
- 用户权限：为创建的系统新用户赋予使用权限

图 3-123 用户权限



- 操作：操作功能位于列表最右端，可对系统用户进行查看、编辑或者删除操作。

图 3-124 操作



3.4.6 功能列表

表 3-4 数据地图平台功能列表

软件名称	数据地图(Dmap)
数据全景	输入关键字，搜索技术资产/业务资产信息；展示主题域分组
数据全景-元数据列表	展示全部主题域分组，以及主题域分组下的主题域、业务对象和逻辑实体
数据全景-业务对象详情	查看业务对象的详细信息
数据全景-逻辑实体详情	查看、评论逻辑实体的详细信息，查看样例数据
数据全景-技术资产详情	查看、评论技术资产的详细信息

软件名称	数据地图(Dmap)
资产注册-关系注册	注册逻辑实体与技术资产的关系
资产注册-关系清单	查询已经注册的关系清单，提供修改删除导出功能
数据质量	对数据质量问题，可以提交审核
标签管理-标签管理	提供标签创建删除置失效功能
标签管理-资产标签	查询已经打上标签的资产
系统管理-我已审批	用户已参与过的电子流信息
系统管理-我的申请	用户发起申请的电子流
系统管理-我的待办	待用户处理的电子流
系统管理-用户管理	新增、维护、删除系统用户，查看系统用户信息

3.5 数据生命周期平台操作手册

登录界面：由右上角的按钮登录系统

图 3-125 登录界面



3.5.1 总览

查看当前用户所属应用下的资产数据、归档数据、销毁数据及归档任务统计信息。

图 3-126 数据看板

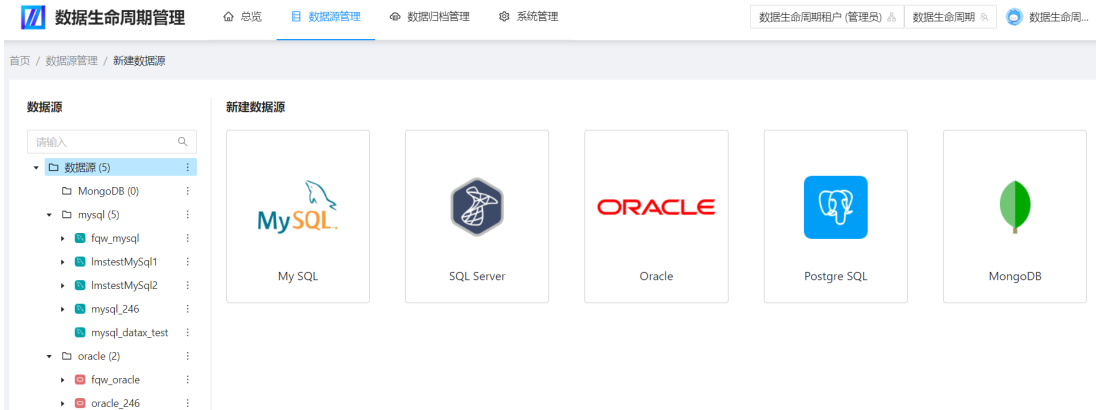


- 总资产：当前用户所属应用下的资产总数，明细可在2.4.2.2节所属领域资产页面中查看。
- 已归档：已归档的物理表数量及容量。
- 销毁资产：已销毁的归档成功的物理表数量及容量
- 归档任务统计：归档任务运行状态的数据统计
- 异常调度：归档失败的任务数据统计，分日/周/月/年四个时间维度统计
- 异常详情：归档失败的任务明细列表，可由任务名下钻至任务详情页面

3.5.2 数据源管理

- 创建、维护数据归档的目标数据源。

图 3-127 数据源管理



- 新建/重命名/删除目录

图 3-128 编辑数据源



- 新建数据源：新增数据归档的目标数据库名称、连接串、访问数据库的账号及密码

图 3-129 新建数据源



3.5.3 数据归档管理

任何事物都具有一定的生命周期，数据也不例外。从数据的产生、加工、使用乃至消亡都应该有一个科学的管理办法，将极少或者不再使用的数据从系统中剥离出来，并通过核实的存储设备进行保留，不仅能够提高系统的运行效率，更好的服务客户，还能大幅度减少因为数据长期保存带来的储存成本。数据生命周期一般包含在线阶段、归档阶段（有时还会进一步划分为在线归档阶段和离线归档阶段）、销毁阶段三大阶段，管理内容包括建立合理的数据类别，针对不同类别的数据制定各个阶段的保留时间、存储介质、清理规则和方式、注意事项等。

归档推荐：展示超出了一定容量的或者访问量低的数据库表，为用户提供归档推荐；单击归档操作可对该表进行数据归档。

图 3-130 归档推荐

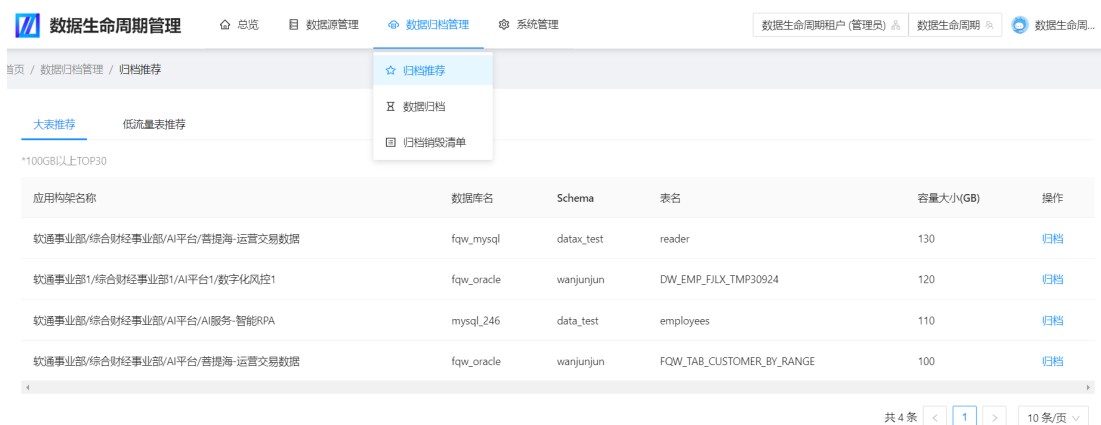


图 3-131 大表推荐

The screenshot shows the 'Data Lifecycle Management' interface with the 'Data Archiving Management' tab selected. Under the 'Large Table Recommendation' sub-tab, a table lists tables with sizes greater than 100GB. The table has columns for application framework name, database name, schema, table name, size in GB, and an archive action.

应用框架名称	数据库名	Schema	表名	容量大小(GB)	操作
软通事业部/综合财经事业部/AI平台/普提海-运营交易数据	fqw_mysql	dtax_test	reader	130	归档
软通事业部1/综合财经事业部1/AI平台1/数字化风控1	fqw_oracle	wanjunjun	DW_EMP_FILX_TMP30924	120	归档
软通事业部/综合财经事业部/AI平台/AI服务-智能RPA	mysql_246	data_test	employees	110	归档
软通事业部/综合财经事业部/AI平台/普提海-运营交易数据	fqw_oracle	wanjunjun	FQW_TAB_CUSTOMER_BY_RANGE	100	归档

图 3-132 低流量表推荐

The screenshot shows the 'Data Lifecycle Management' interface with the 'Data Archiving Management' tab selected. Under the 'Low Flow Table Recommendation' sub-tab, a table lists tables based on their most accessed time in the top 30 over the last half year. The table has columns for application framework name, database name, schema, table name, size in GB, last access time, and an archive action.

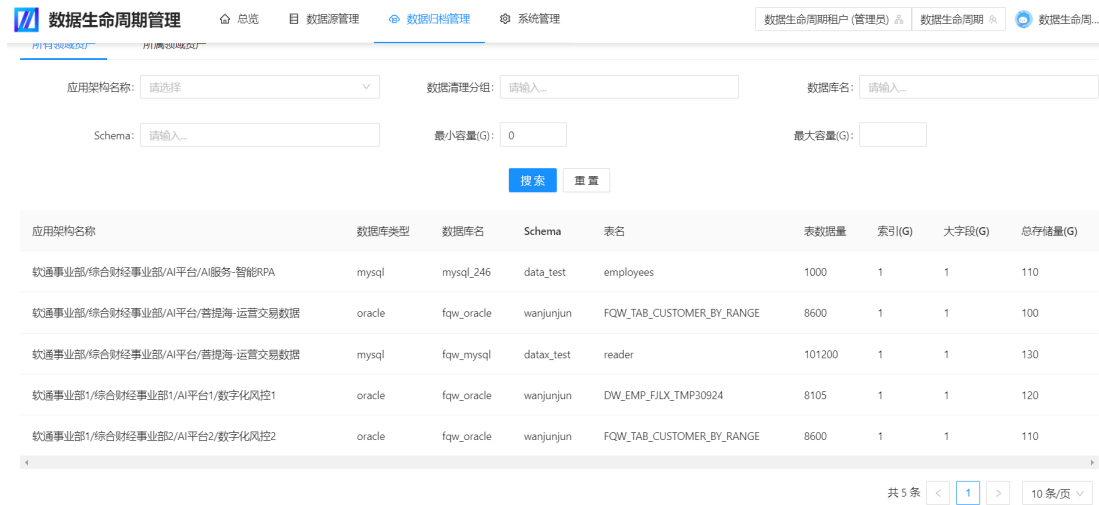
应用框架名称	数据库名	Schema	表名	容量大小(GB)	最近访问时间	操作
软通事业部/综合财经事业部/AI平台/普提海-运营交易数据	fqw_mysql	dtax_test	reader	130	2021-01-25 01:15:08	归档
软通事业部/综合财经事业部/AI平台/普提海-运营交易数据	fqw_oracle	wanjunjun	FQW_TAB_CUSTOMER_BY_RANGE	100	2021-02-25 01:15:08	归档
软通事业部1/综合财经事业部1/AI平台1/数字化风控1	fqw_oracle	wanjunjun	DW_EMP_FILX_TMP30924	120	2021-03-15 01:15:08	归档
软通事业部/综合财经事业部/AI平台/AI服务-智能RPA	mysql_246	data_test	employees	110	2021-03-25 01:15:08	归档

数据归档

查看所有领域资产，可对所属领域的资产进行创建归档任务，配置归档策略等操作。

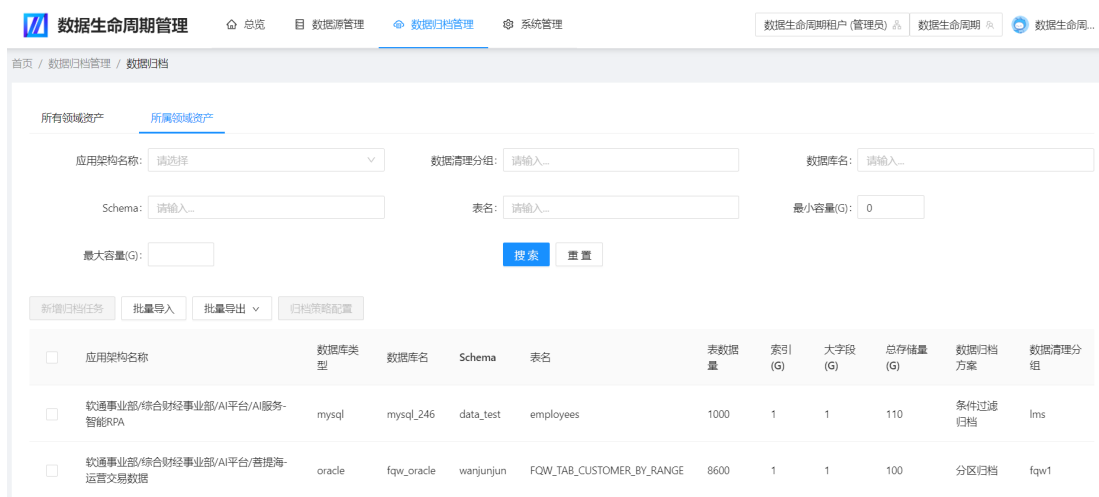
- 所有领域资产：搜索、查看系统所有领域的资产列表。

图 3-133 所有领域资产



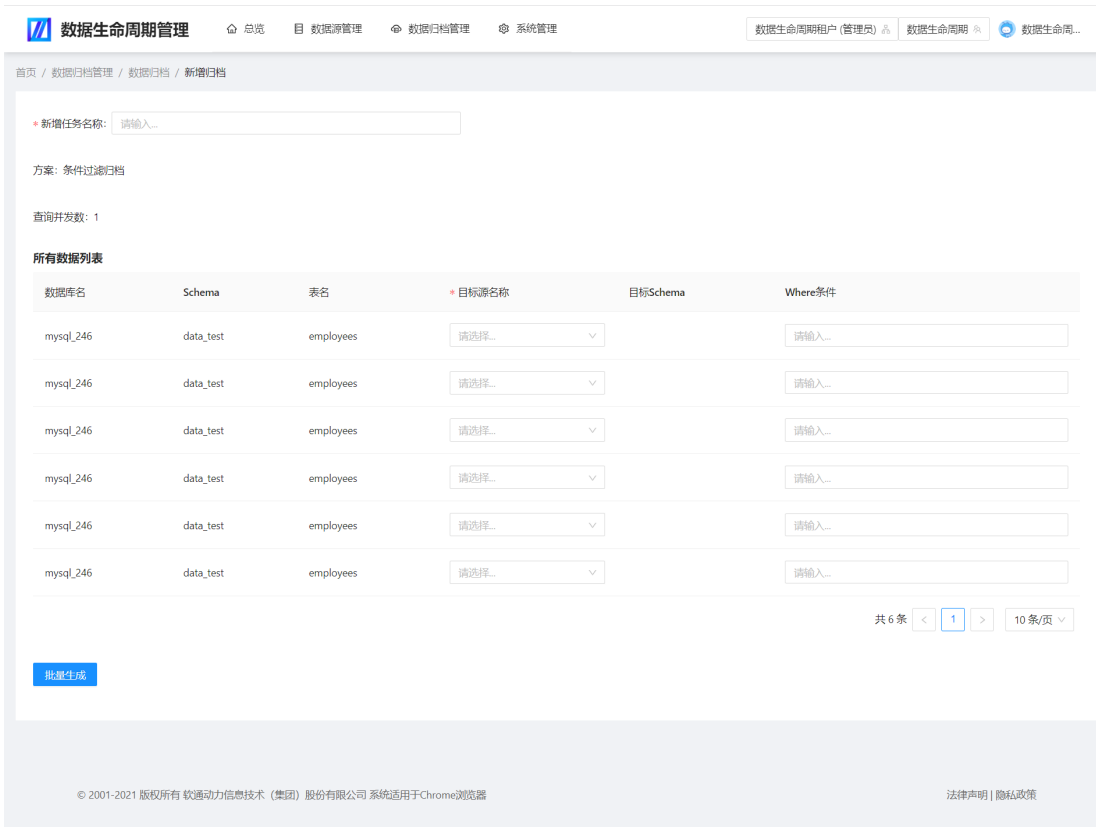
- 所属领域资产：查询、查看用户所属领域的资产，可进行归档操作。

图 3-134 所属领域资产



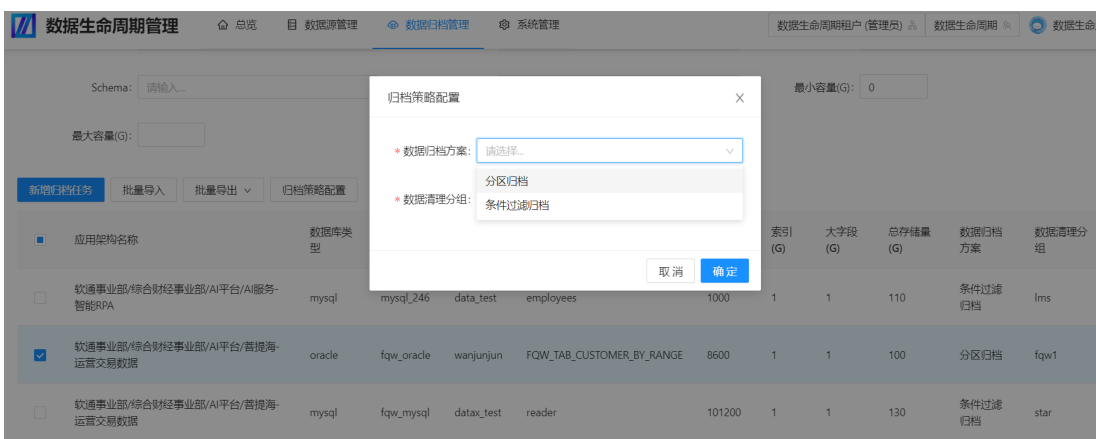
- 新增归档任务：创建所选资产的数据归档任务。

图 3-135 新增归档任务



- 批量导入：批量导入用户所属领域的资产清单。
- 批量导出：批量导出用户所属领域的资产清单。
- 归档策略配置：配置资产的数据归档策略。

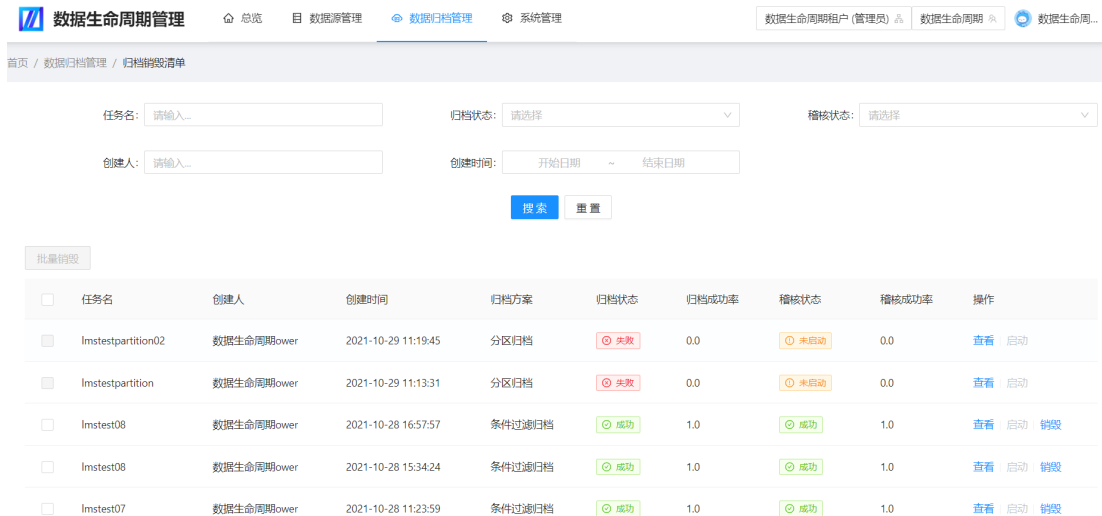
图 3-136 归档策略配置



归档销毁清单

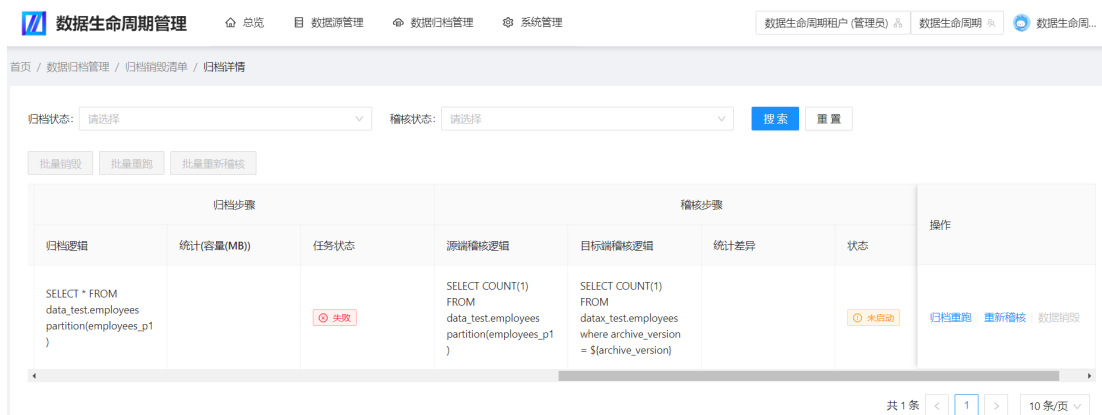
搜索、查看归档任务及稽核状态，可对数据源中已归档的数据进行销毁。

图 3-137 归档销毁清单



- 批量销毁：可批量销毁已归档成功的源表数据。
- 查看：可进入归档详情页面，查看任务归档明细信息。

图 3-138 归档详情

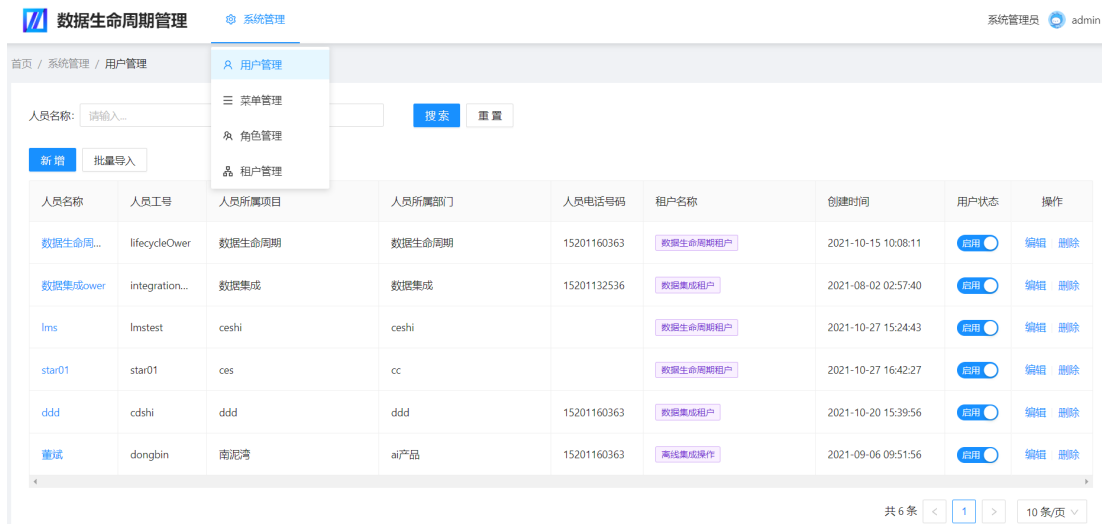


- 批量销毁/重跑/重新稽核：对同一任务下多个归档明细数据进行批量处理。
- 任务状态：可单击查看归档任务状态信息。
- 状态：可单击查看数据稽核状态信息。
- 归档重跑：可再次对数据进行归档操作。
- 重新稽核：可对归档成功的数据再次进行稽核。
- 数据销毁：可销毁已归档成功的源表中的数据。
- 启动：可启动运行新建的归档任务。
- 销毁：可销毁任务中全部已归档的源表中的数据。

3.5.4 系统管理

系统管理员可对生命周期系统的用户、菜单、角色、租户进行管理。

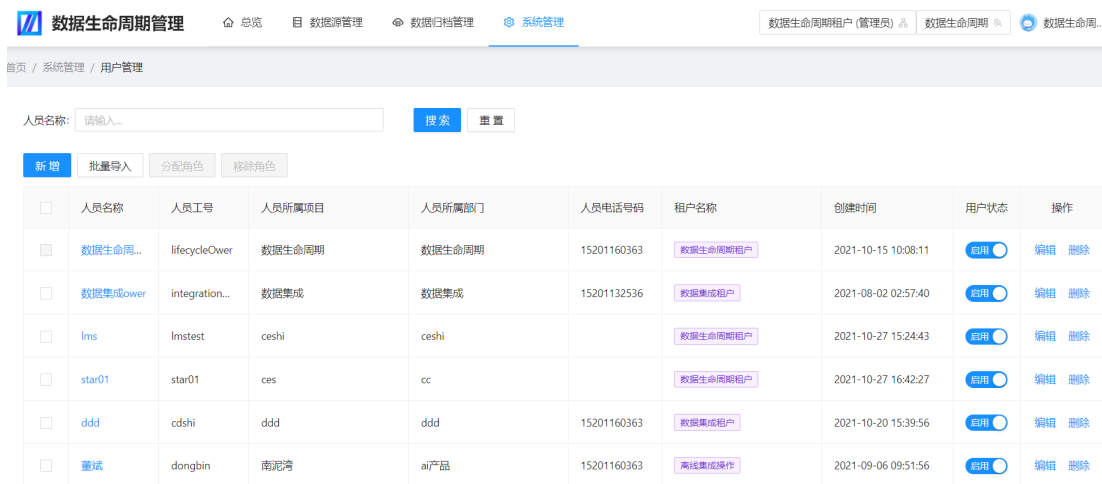
图 3-139 系统管理



用户管理

可新增、批量导入新用户，分配用户角色、修改用户状态、编辑维护用户信息以及删除用户，模糊搜索用户信息。

图 3-140 用户管理

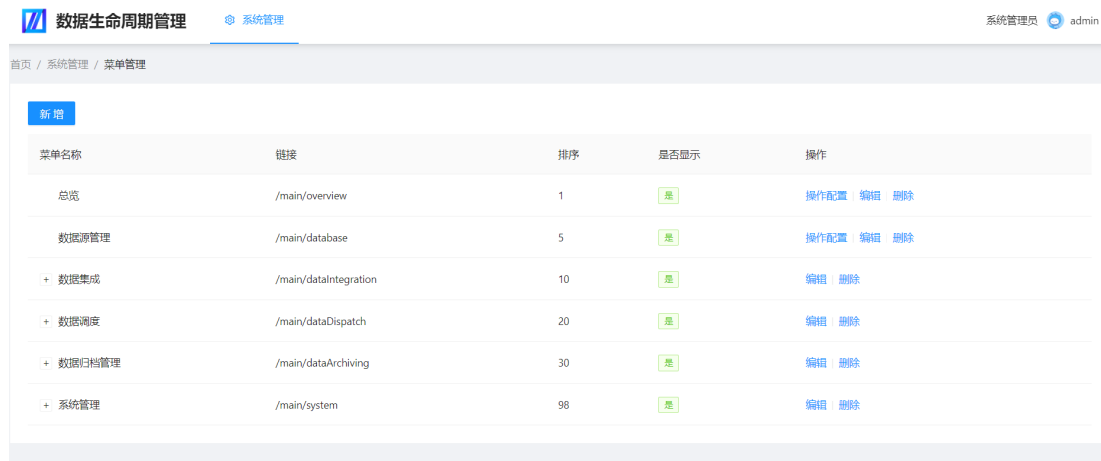


- 新增：新增单个系统用户。
- 批量导入：批量新增系统用户。
- 分配角色：分配用户角色权限。
- 移除角色：移除用户角色权限。
- 用户状态：启用/禁用当前用户。
- 编辑：维护用户信息。
- 删除：删除系统用户。

菜单管理

对数据生命周期系统的菜单进行管理。

图 3-141 菜单管理



菜单名称	链接	排序	是否显示	操作
总览	/main/overview	1	<input checked="" type="checkbox"/>	操作配置 编辑 删除
数据源管理	/main/database	5	<input checked="" type="checkbox"/>	操作配置 编辑 删除
+ 数据集成	/main/dataIntegration	10	<input checked="" type="checkbox"/>	编辑 删除
+ 数据调度	/main/dataDispatch	20	<input checked="" type="checkbox"/>	编辑 删除
+ 数据归档管理	/main/dataArchiving	30	<input checked="" type="checkbox"/>	编辑 删除
+ 系统管理	/main/system	98	<input checked="" type="checkbox"/>	编辑 删除

- 新增：新增菜单栏，配置菜单所在的节点，菜单名称等信息。

图 3-142 新增



新增

* 上级菜单: 请选择...

* 菜单名称: 请输入... * 链接: 请输入...

* 菜单类型: 目录 功能 * 是否显示: 是 否

* 图标: 请输入... 排序: 0

描述: 请输入...

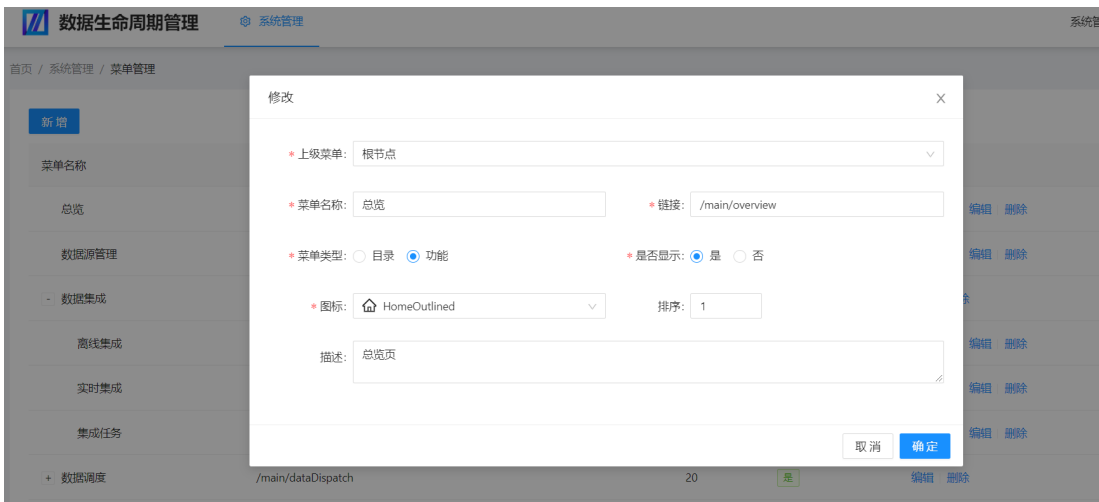
- 操作配置：配置功能菜单的操作权限。

图 3-143 操作配置



- 编辑：修改菜单信息。

图 3-144 编辑

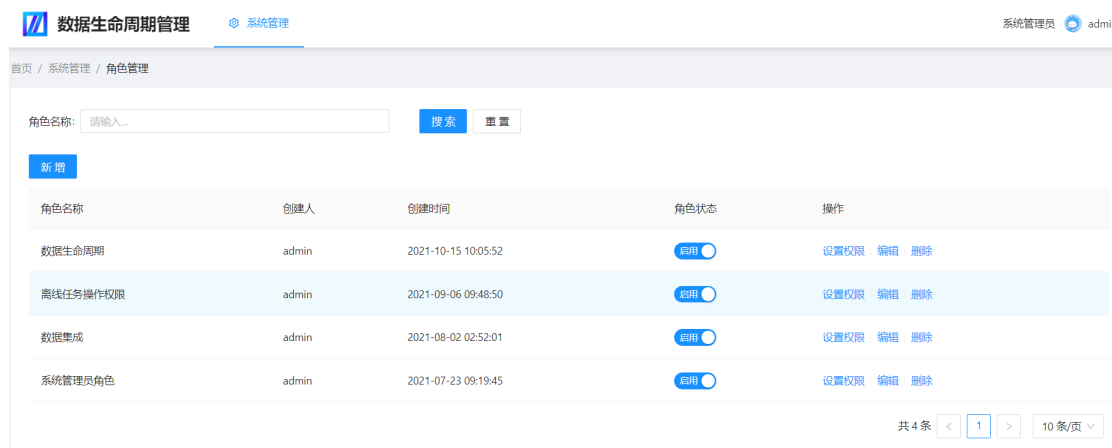


- 删除：删除多余菜单。

角色管理

可新增、编辑、删除系统角色，维护角色状态及权限。

图 3-145 角色管理



- 新增：新增系统角色。

图 3-146 新增

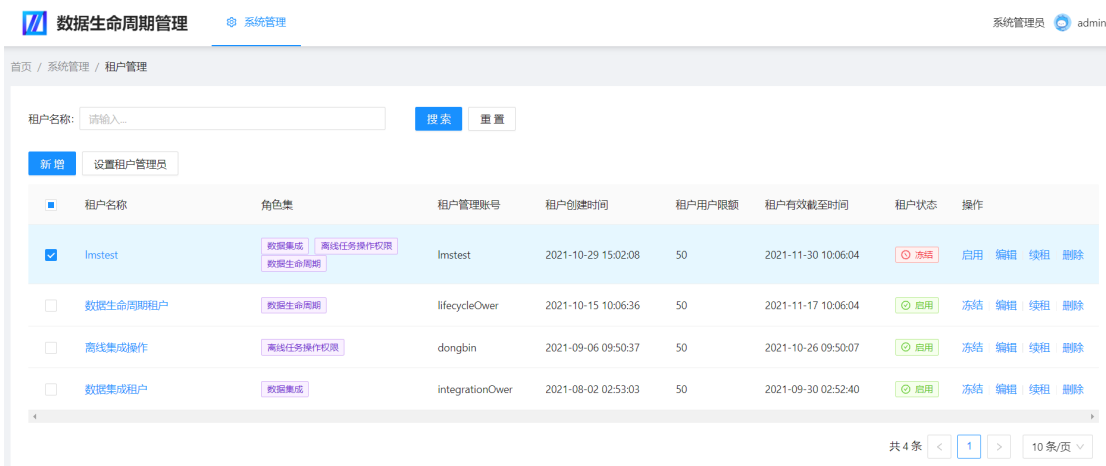


- 角色状态：启用/禁用所选角色。
- 设置权限：设置角色的权限范围。
- 编辑：编辑角色信息。
- 删除：删除角色。

租户管理

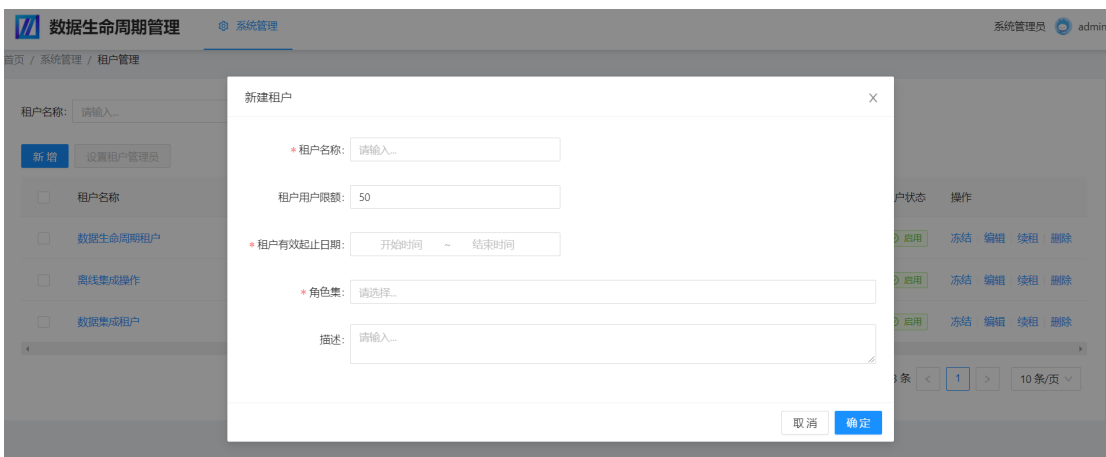
搜索、查看系统租户信息，可进行新增/删除租户，维护租户信息，延续租期，冻结异常租户等操作。

图 3-147 租户管理



- 新增：新增系统租户。

图 3-148 新建用户



- 设置租户管理员：为选择的租户指定管理员。
- 冻结/启用：冻结、启用所选租户。
- 编辑：编辑维护租户信息。
- 续租：延长租户租期时间范围。

3.5.5 功能列表

表 3-5 数据生命周期平台功能列表

软件名称	数据生命周期平台
总览	资产数据、归档数据、销毁数据及归档任务统计信息
数据源管理	创建、维护数据归档的目标数据源

软件名称	数据生命周期平台
数据归档管理-归档推荐	推荐用户进行归档的列表清单
数据归档管理-数据归档	查看数据资产，对所属领域的资产进行归档操作
数据归档管理-归档销毁清单	查看任务归档信息，销毁已归档成功的源数据
系统管理	管理员对系统的用户、菜单、角色和租户进行管理和维护

3.6 DataLab 用户手册

输入用户的账号和密码，登录进主页面。

图 3-149 登录界面



总览

总览主要分为五大块内容：分别是登录用户所属租户的基本信息展示，所属租户运营指标展示，所属租户最近操作资产清单列表展示，帮助文档列表和当前租户管理员基本信息展示。

图 3-150 数据总览



3.6.1 数据源

数据源主要包含两块：左侧树状导航展示数据源路径、右侧可进行新建数据源操作。

图 3-151 数据源



单击“数据源”，展开列表，可看到该数据源下的所有表；右侧可以查看或编辑数据源配置信息，包含数据源连接串、用户名、密码、数据源名称、资源权限、责任人、数据源描述等；

图 3-152 数据源信息



单击“新建数据源”，添加一个新的数据源，这条记录会同步显示在左侧树状导航中。

图 3-153 添加数据源

添加数据源

1 测试连接 2 选择目录 3 完成

* 连接URL 请输入...

* 用户名 请输入... * 密码 请输入...

* 数据源名称 请输入... * 资源权限 请选择

* 责任人 请选择

数据源描述 请输入...

测试连接 取消

3.6.2 数据集

数据集主要包含两块：左侧树状导航展示数据集路径、右侧可进行新建数据集操作。

图 3-154 数据集

DataLab

数据集

新建数据集

数据源导入 文件导入

* 数据集/类名 请选择

* 数据集名称 请输入...

数据集描述 请输入...

* 保存路径

A_user_1_dataset

保存 重置

数据集可以直接从数据源在线导入，也可以离线导入本地文件。

图 3-155 数据集导入

DataLab

数据集

新建数据集

数据源导入 文件导入

上传 选择上传的csv文件 下载模板

* 数据集名称 请输入...

数据集描述 请输入...

* 保存路径

A_user_1_dataset

保存 重置

单击“数据集”，可看到该数据集详情和元数据信息。

图 3-156 数据集详情

The screenshot shows the DataLab interface with the '数据集' (Dataset) tab selected. The main content area displays a table of dataset details. The table has columns for provinceName, provinceCnCode, city_curedCount, countryName, city_zipCode, cityEnglishName, city_suspectedCount, province_deadCount, continentName, province_confirmedCount, and province_suspectedCount. The data rows include entries for Belgium, Russia, Ukraine, Poland, Mexico, Hungary, Bolivia, India, New Zealand, and Czech Republic.

provinceName	provinceCnCode	city_curedCount	countryName	city_zipCode	cityEnglishName	city_suspectedCount	province_deadCount	continentName	province_confirmedCount	province_suspectedCount
比利时			比利时				17792	欧洲	603159	0.0
俄罗斯			俄罗斯				46453	欧洲	2625848	0.0
乌克兰			乌克兰				14998	欧洲	885039	0.0
波兰			波兰				22676	欧洲	1126700	0.0
墨西哥			墨西哥				113019	北美洲	1229379	0.0
匈牙利			匈牙利				6784	欧洲	276247	0.0
玻利维亚			玻利维亚				5562	欧洲	177665	0.0
印度尼西亚			印度尼西亚				16653	亚洲	611631	0.0
新西兰			新西兰				1998	欧洲	93734	0.0
捷克			捷克				9450	欧洲	575422	0.0

图 3-157 元数据

The screenshot shows the DataLab interface with the '元数据' (Metadata) tab selected. The main content area displays a table of metadata. The table has columns for field name, Chinese name, English name, field type, field description, and index number. The data rows include entries for continentName, continentEnglishName, countryName, countryEnglishName, provinceName, provinceEnglishName, province_zipCode, provinceCnCode, province_confirmedCount, province_suspectedCount, and province_curedCount.

字段名称	中文名称	英文名称	字段类型	字段描述	排序序号
continentName		continentName	VARCHAR		1
continentEnglishName		continentEnglishName	VARCHAR		2
countryName		countryName	VARCHAR		3
countryEnglishName		countryEnglishName	VARCHAR		4
provinceName		provinceName	VARCHAR		5
provinceEnglishName		provinceEnglishName	VARCHAR		6
province_zipCode		province_zipCode	VARCHAR		7
provinceCnCode		provinceCnCode	VARCHAR		8
province_confirmedCount		province_confirmedCount	VARCHAR		9
province_suspectedCount		province_suspectedCount	VARCHAR		10
province_curedCount		province_curedCount	VARCHAR		11

3.6.3 数据探索

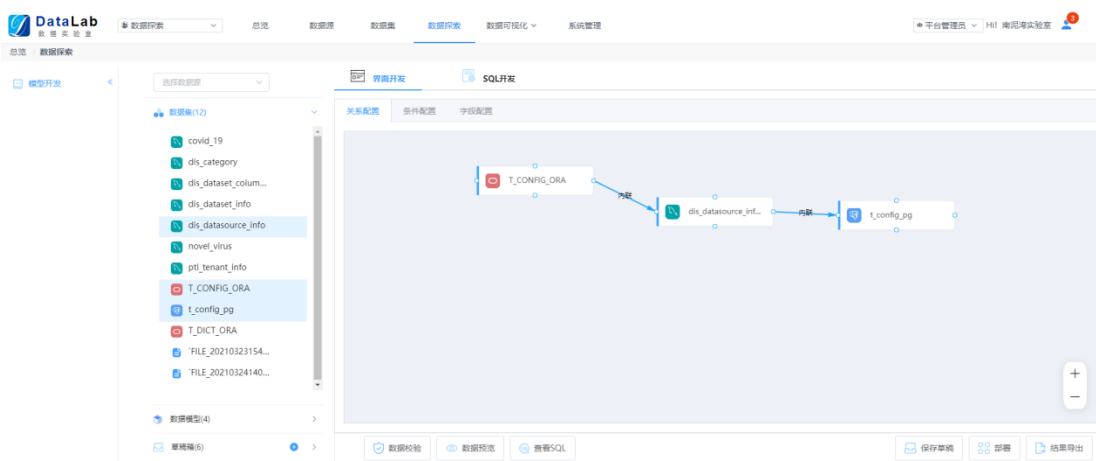
数据探索通过关联加工数据集，进行数据建模。数据建模有界面开发和sql开发两种方式。

图 3-158 数据探索



- 界面开发：界面开发采用拖拽方式，把来自不同数据源的数据集进行关联。不用编写代码，就能轻松完成数据集的关联查询与分析。

图 3-159 界面开发



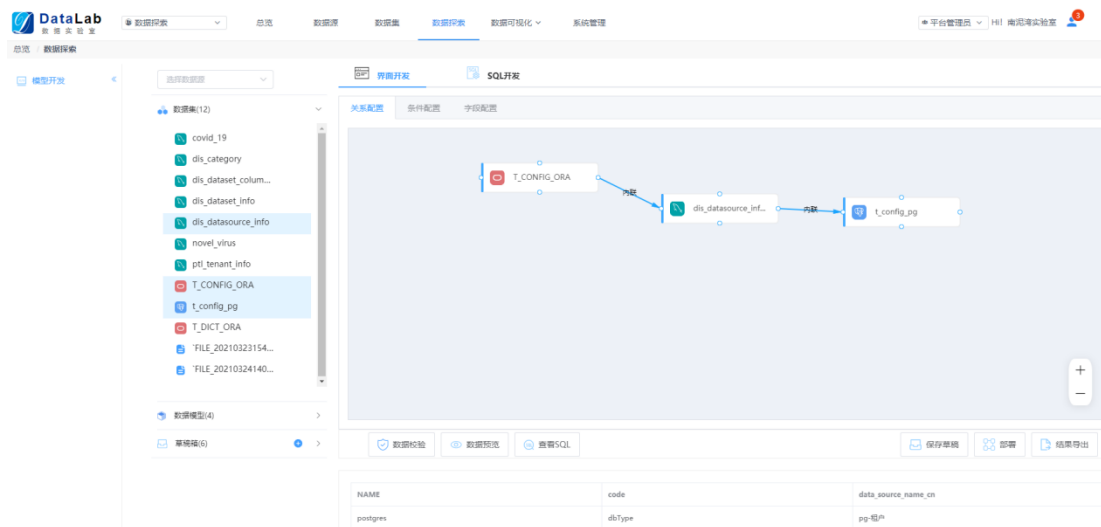
通过关联关系配置，实现不同的应用场景。

图 3-160 关系配置-1



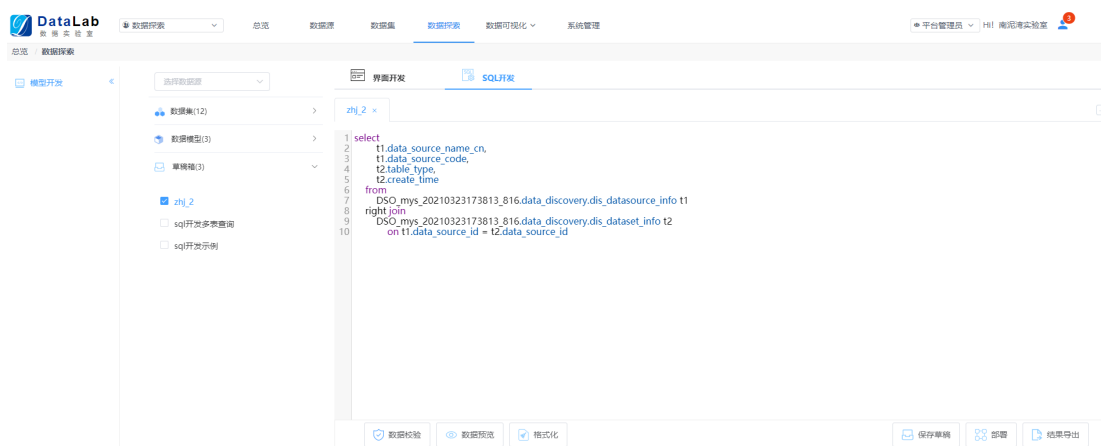
可以对关联加工后的数据模型进行数据预览，结果导出。

图 3-161 关系配置-2



- SQL开发：SQL开发针对IT人员，提供编写代码的方式对数据集进行逻辑加工，同样支持数据预览，结果导出。

图 3-162 SQL 开发



3.6.4 数据可视化

数据可视化分为可视化图表和可视化看板两部分。

图表

可视化图表列表中展示所有的图表信息，可根据不同的搜索条件去查找图表。

图 3-163 图表



图 3-164 编辑图表



图 3-165 编辑图表属性



图 3-166 批量删除图表



图 3-167 新建图表



通过选择数据集，并对图表类型、时间范围、查询指标、过滤条件等项做好配置，得到可视化图表。同时，还能查看查询结果和数据样本。

图 3-168 选择数据集

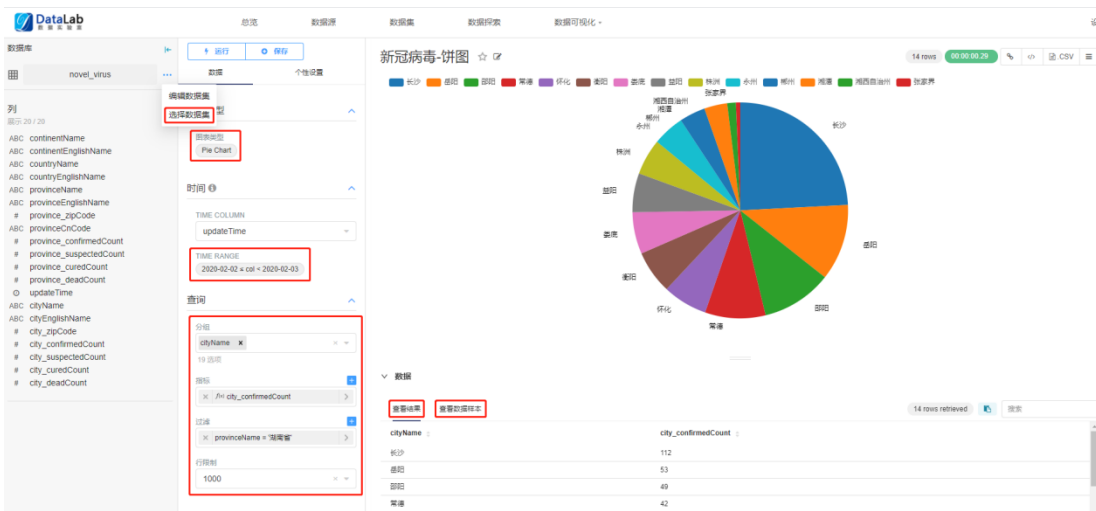


图 3-169 个性设置

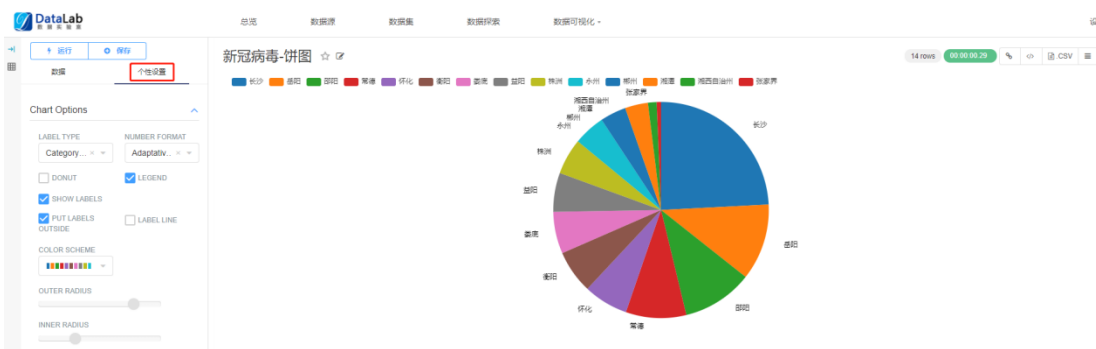


图 3-170 保存图表



看板

可视化看板列表中展示所有的看板信息，可根据不同的搜索条件去查找看板。

图 3-171 看板



图 3-172 编辑看板



图 3-173 基本信息-权限



图 3-174 基本信息-颜色



图 3-175 自定义看板-1

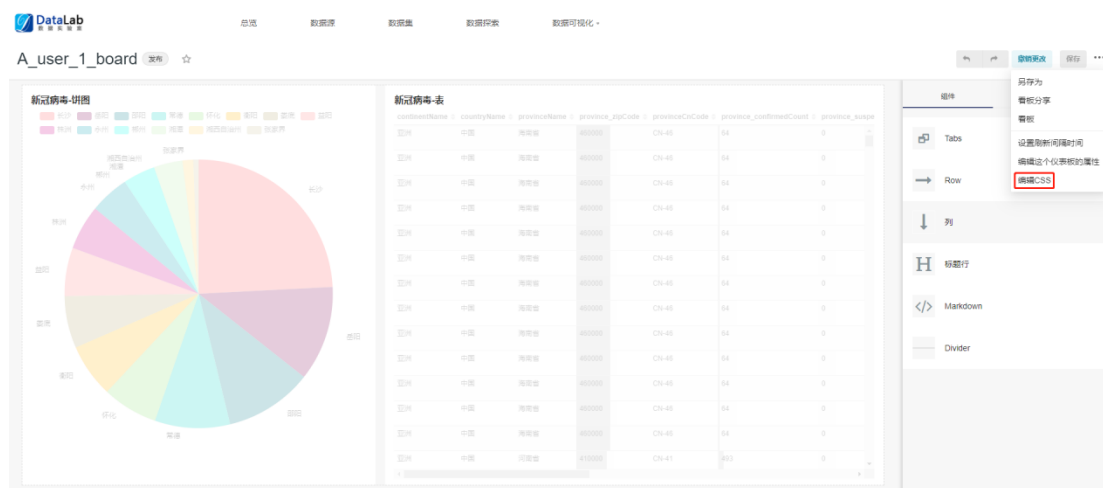


图 3-176 自定义看板-2

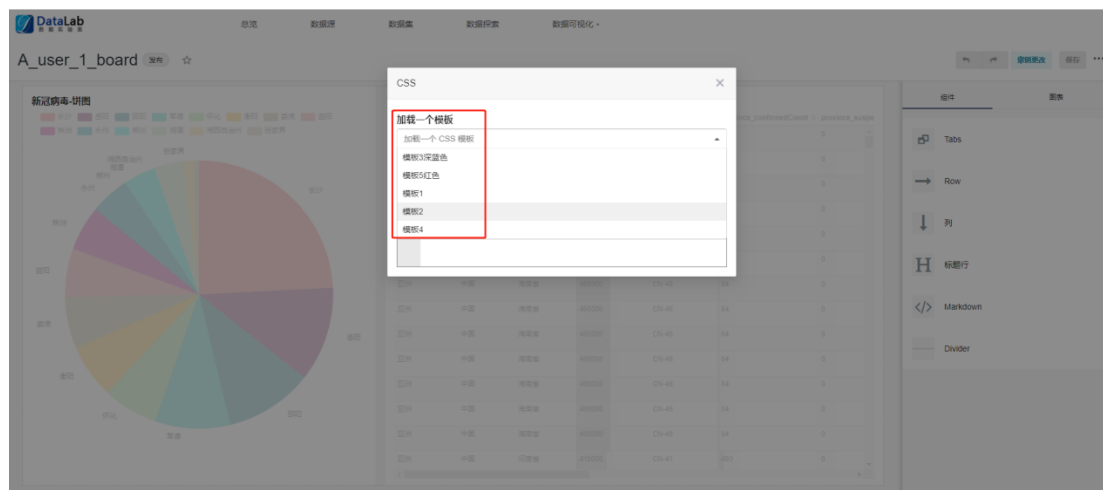


图 3-177 看板-批量删除和导出



图 3-178 新建看板



图 3-179 自定义看板-拖拉组件布局

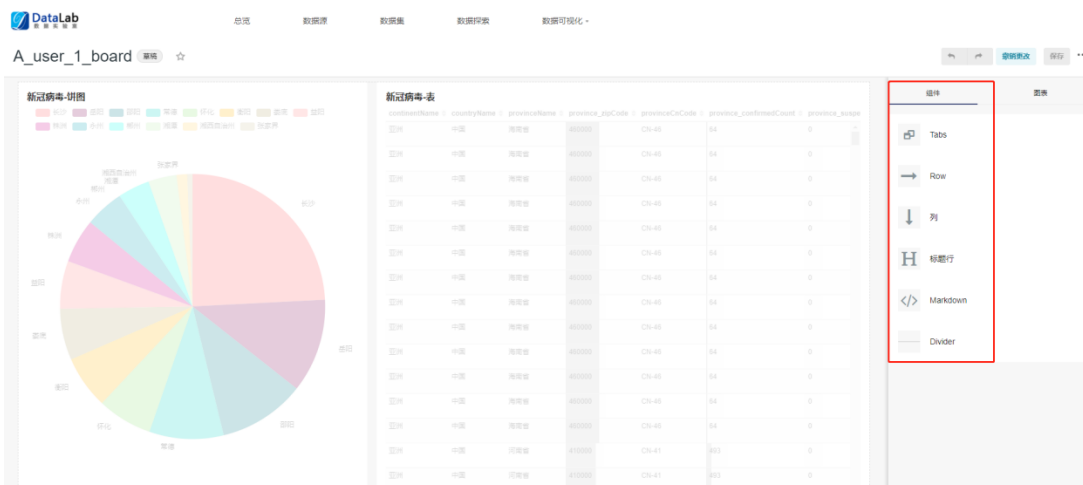


图 3-180 自定义看板-拖拉图表

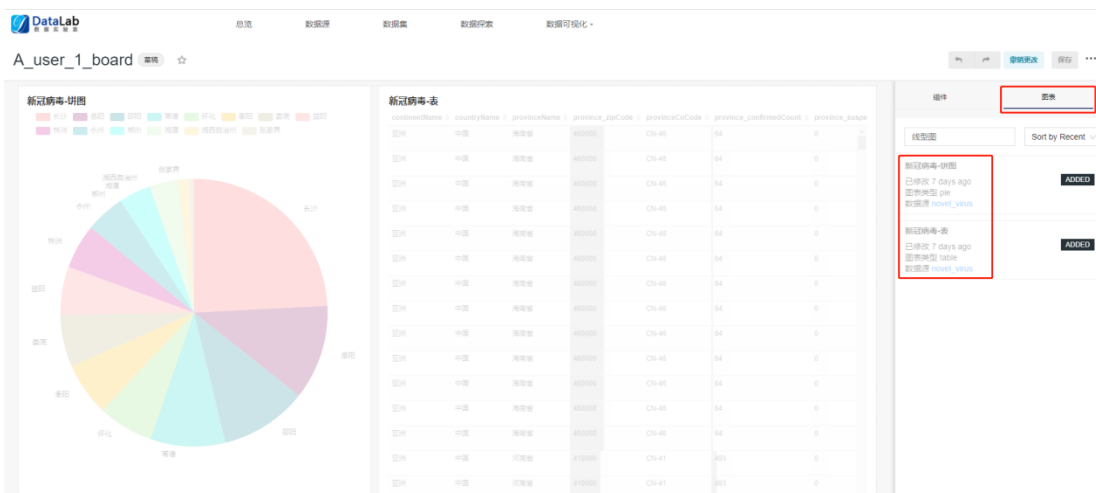


图 3-181 保存看板

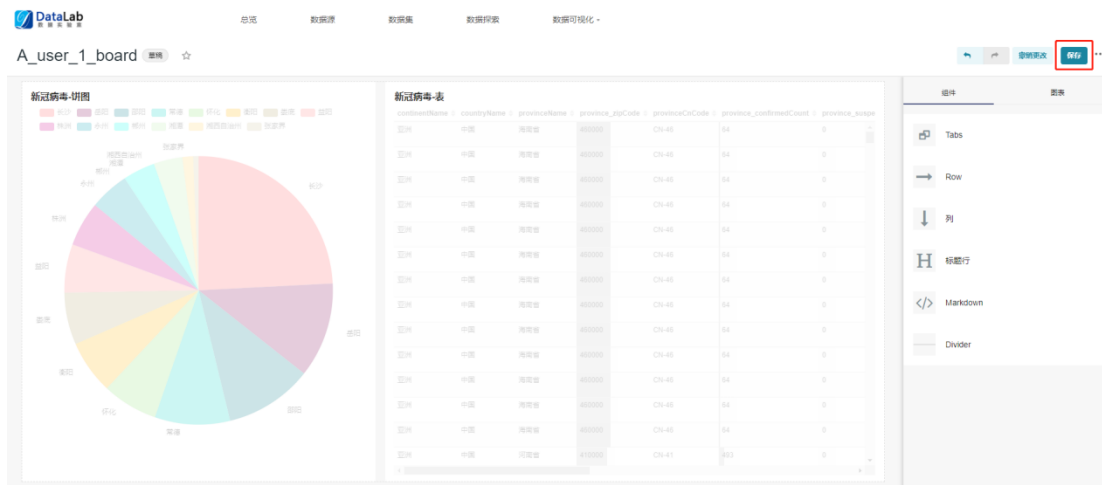


图 3-182 发布看板

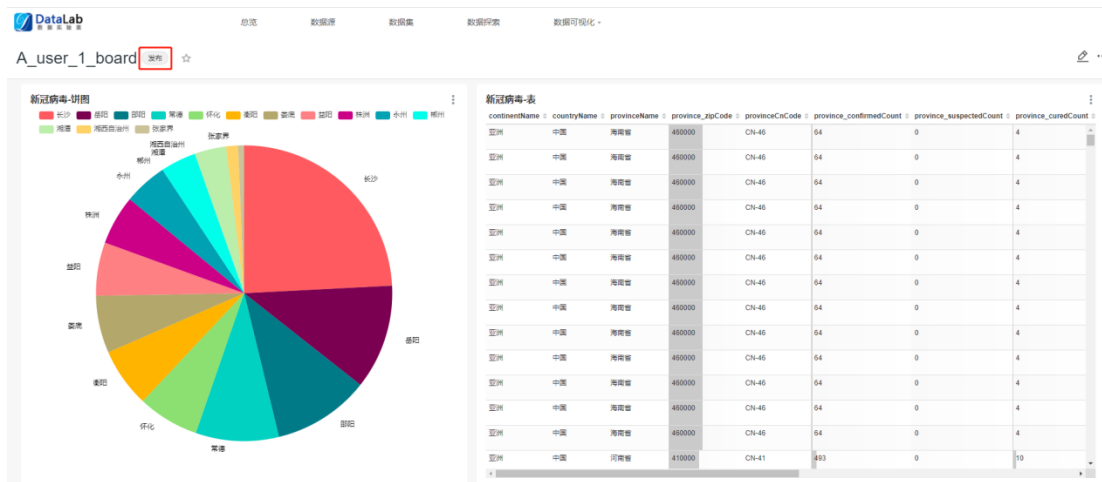


图 3-183 另存为看板

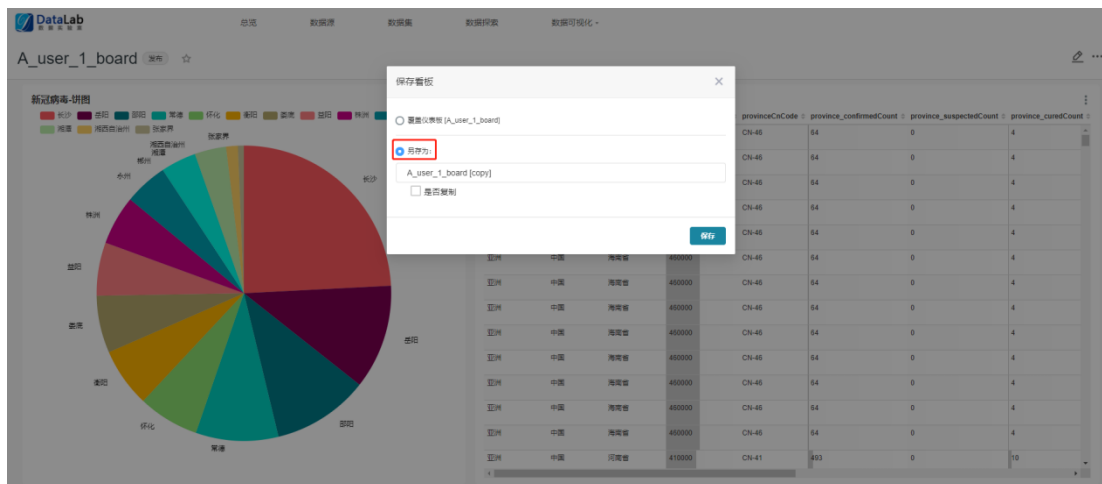
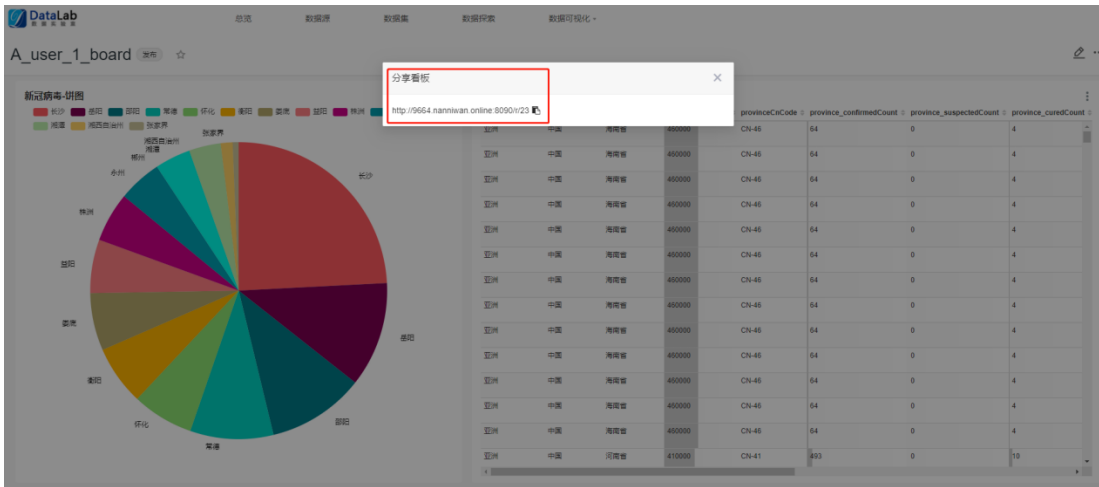


图 3-184 分享看板



刷新看板：可以手动刷新看板，使看板数据动态更新。

图 3-185 设置自动刷新频率

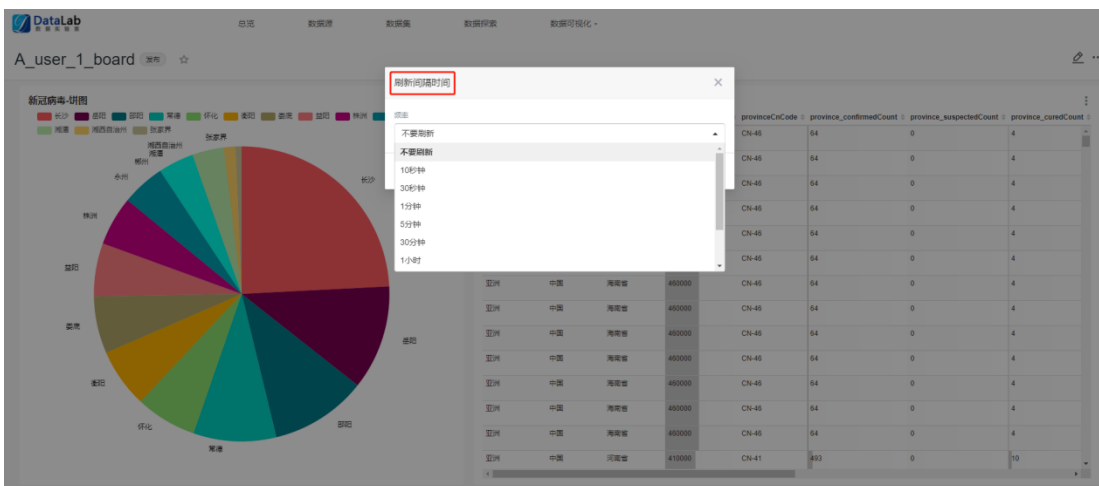


图 3-186 看板图片下载

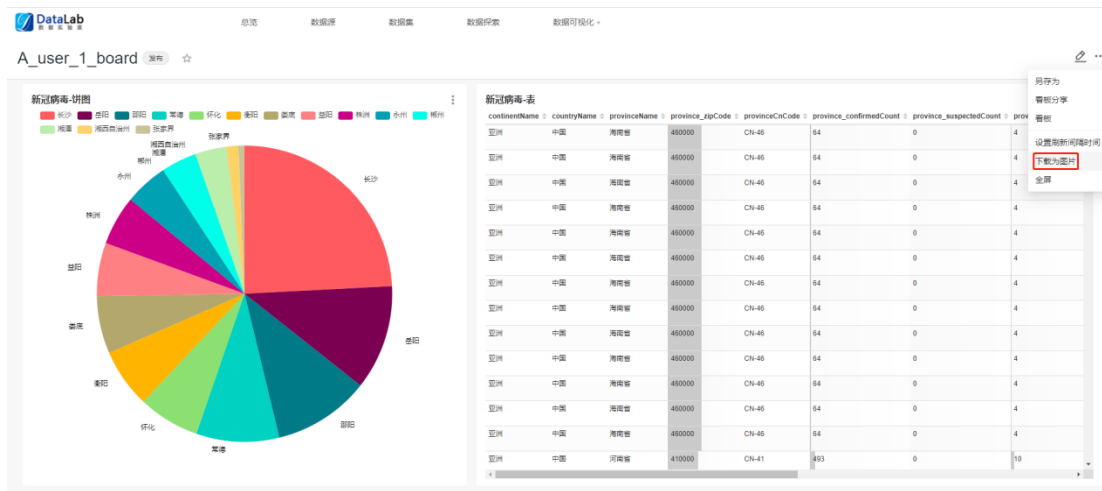
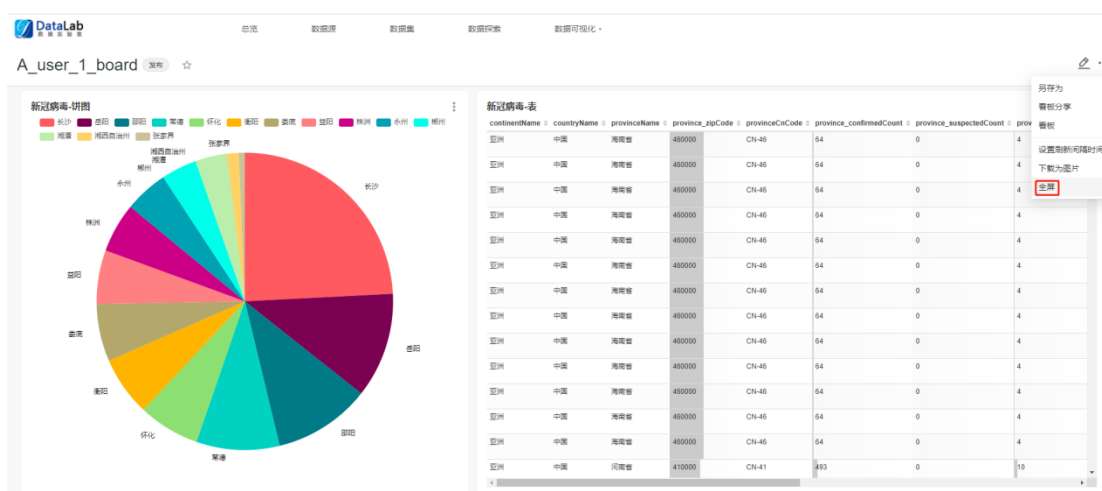


图 3-187 看板全屏展示



3.6.5 系统管理

系统管理包含租户管理和用户管理。以租户隔离数据，以用户隔离可视化看板。

图 3-188 系统管理

The screenshot shows the '租户管理' (Tenant Management) section of the system. It includes a sidebar with '租户管理' and '用户管理' (User Management) options. The main area displays a table of tenants with columns for tenant ID, name, creator, owner, creation time, user count, expiration time, status, and actions.

租户编号	租户名称	租户创建人	租户Owner	租户创建时间	租户用户数	租户有效截止时间	租户状态	操作
TN_2021032514950_766	B_tenant	软通动力	B_owner	2021-03-25 15:48:51	10	2021-04-08 10:21:56	启用	⋮
TN_20210325154831_262	A_tenant	软通动力	A_owner	2021-03-25 15:48:32	10	2021-10-30 00:00:00	启用	⋮
TN_20210325143938_106	主数据	软通动力	主数据用户1	2021-03-25 14:39:40	10	2021-05-31 00:00:00	启用	⋮
TN_20210309175334_845	superset	软通动力	superset_user_1	2021-03-09 17:53:34	10	2021-04-23 00:00:00	启用	⋮
TN_20210309162157_544	数据研发	软通动力	wanjuan	2021-03-09 16:21:57	20	2021-10-31 00:00:00	启用	⋮

租户管理

租户列表展示所有租户信息，可对租户信息进行查看、编辑、冻结、续租、删除等操作。

图 3-189 租户管理



新建租户：平台管理员有新建租户权限，在新建租户前要新建用户。可以指定一个用户为租户OWNER。

图 3-190 新建租户-1



图 3-191 新建租户-2

新建租户

* 租户名称

请输入...

* 租户OWNER

请选择

* 租户用户限额

10

租户状态

启用

* 租户有效截止时间

请选择截止时间...

* 选择角色集

请选择

租户描述

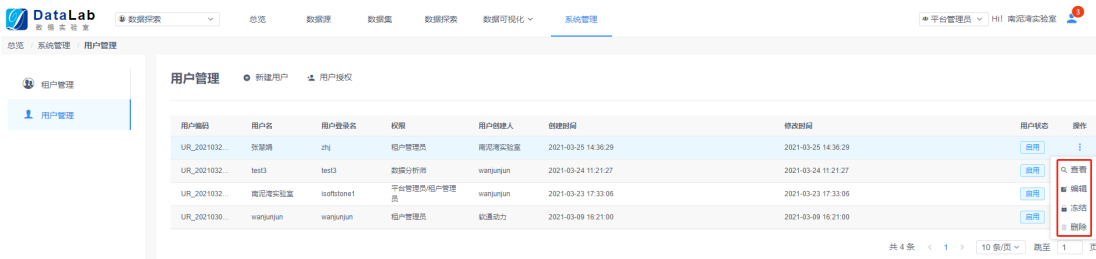
请输入...

提交 取消

用户管理

用户列表展示所有用户信息，可对用户信息进行查看、编辑、冻结、删除等操作。

图 3-192 用户管理



新建用户，并给用户授权，赋予不同的角色。

图 3-193 新建用户-1

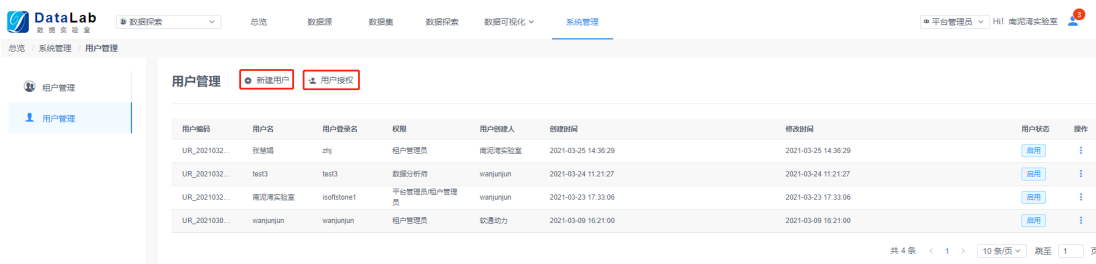


图 3-194 新建用户-2

新建用户

* 用户姓名

* 用户登录名

* 用户密码

* 租户状态 启用

用户联系电话

用户联系邮箱

用户年龄

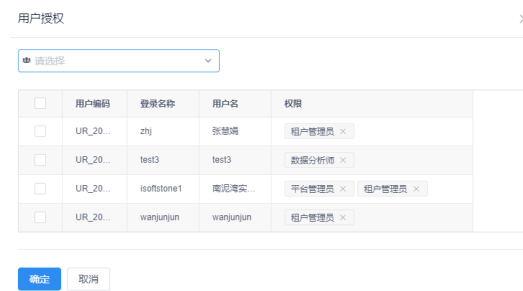
用户所属公司

用户所属部门

用户职位

备注

图 3-195 新建用户-3



3.6.6 功能列表

表 3-6 DataLab 功能列表

软件名称	DataLab
软件功能项	功能描述
portal登录	用户登录/退出、产品宣传视频、产品介绍
总览	展示登录用户所属租户基本信息、所属租户运营指标、所属租户最近操作资产清单列表、帮助文档列表、当前租户管理员基本信息
数据源	新建/编辑/删除数据源、数据源搜索、数据源树状展示
数据集	在线数据源导入数据集、离线本地文件导入数据集、编辑/删除数据集、数据集搜索、数据集树状展示
数据探索	拖拉组件进行数据集关联、关联关系配置、查看SQL、编写SQL代码进行数据集关联、数据校验、数据预览、结果导出、数据模型草稿保存、数据模型部署
可视化图表	图表列表展示、图表搜索、图表基本信息编辑/修改、批量删除图表、图表授权、新建/编辑/删除图表、图表随配置条件的变化联动、查询结果查看、数据样本查看、图表个性化设置、图表保存
可视化看板	新建看板、拖拉组件布局看板、拖拉图表制作看板、看板保存、看板发布、另存为看板、分享看板、刷新看板、设置看板自动刷新频率、编辑看板属性、编辑CSS样式、图片下载、全屏展示
租户管理	租户列表展示、租户信息查看/编辑、新建/冻结/续租/删除租户、给租户授权
用户管理	用户列表展示、用户信息查看/编辑、新建/冻结/删除用户、给用户分配角色

4 附录 背景信息

表 4-1 数据集成实施

类别	步骤	实施说明
数据集成实施	开发测试	<ul style="list-style-type: none">● 数据开发任务梳理：根据数据湖数据集成方案、数据湖数据集成实施指南及待入湖数据情况，梳理需要开发的数据集成任务和脚本，并明确入湖开发分工及节奏。● 数据入湖开发实施：按照数据集成方案和实施指南进行实施，开发数据集成任务或脚本，并完成测试。● 系统集成测试：各数据链路迭代完成后，对全链路进行数据一致性校验、链路可靠性测试、数据入湖性能测试，并由最终用户依据业务需求进行测试。● 数据集成支撑：根据数据入湖需求，协调获取如下支撑和保障：<ol style="list-style-type: none">1. 完成数据采集所需要的技术支撑，包括账号、权限，网络，文件格式，文件内容，技术文档等；2. 客户方协调业务、技术人员，以及第三方人员（如数据应用分析团队）按需提供支撑。● 实时（含流式）数据入湖：按照数据集成打样设计中实时（含流式）数据入湖内容，开展相应数据开发、参数配置、任务调度和链路调测等，确保打样链路畅通，能够顺利完成数据入湖。

类别	步骤	实施说明
	部署上线	<ul style="list-style-type: none"> ● 上线计划制定：经与相关干系人沟通、确认，制定上线计划，明确上线范围和里程碑节点，经客户审批后实施 ● 上线相关方案制定：制定上线失败回退方案、数据入湖运维方案、上线操作手册等，并对客户相关人员进行培训。 ● 上线检查实施： <ol style="list-style-type: none"> 1. 清理测试阶段的数据和与商用无关的数据； 2. 检查系统中的各项配置是否正确，并导出当前配置文件作为副本； 3. 收回所有人的接入权限，删除测试用户的注册信息并关闭白名单； 4. 检查上线的操作手册、脚本是否完备等。 ● 上线实施： <ol style="list-style-type: none"> 1. 建设客户业务系统、中间库和数据集成平台的网络策略，网闸策略； 2. 按照接入方案，执行集成作业，对各个业务系统全量和增量数据进行接入。 ● 上线试运行：开展上线试运行支持，监控数据接入链路运行情况，并对接入数据一致性等进行监控，及时解决上线发现的问题，并输出上线试运行报告。

表 4-2 数据治理实施

类别	实施条目	实施内容说明
数据治理实施	数据标准实施	根据咨询和设计输出的数据标准，在DataArts上落地数据标准，并完成数据标准发布。
	物理建模实施	根据设计环节输出的物理模型，完成数据各层的物理建模实施，并检查实施结果。
	数据指标开发	根据指标设计方案，配置原子指标、衍生指标和复合指标以及对应的ETL脚本，其他还包括：新建时间限定和维度配置。
	脚本/作业开发	<p>按需完成各数据层的ETL脚本开发、数据同步，并基于各层数据的时间顺序、频率的要求，完成作业开发和调度执行。</p> <ol style="list-style-type: none"> 1. 完成ODS层->DWI层的数据同步，在此过程中完成数据的清洗转换，以及多源异构数据的整合； 2. 完成DWI层->DWR层的数据同步，在DWR层将数据按分析/主题维度汇聚，并根据业务需求进行不同粒度的轻度汇总； 3. 完成DWR层->DM层的数据同步，对应数据应用系统的需求，在DM层将数据按应用维度进行汇聚。
	数据服务开发	根据数据服务设计方案，基于数据使能平台的数据服务模块完成API开发，包括：API生成、API注册、API审核、API发布。

类别	实施条目	实施内容说明
	数据质量实施	根据需要监控的业务指标或数据质量，选择系统预置的数据质量模板或开发相应的数据质量分析脚本，并配置作业定期执行。通过已定义的数据质量规则对目标数据进行检查。
	数据资产实施	基于数据目录、数据表两个维度管理对应的数据权限；基于数据密级以及脱敏策略实现数据安全管控；针对各类数据源实现元数据采集管理。

表 4-3 数据地图平台实施

类别	步骤	实施说明
数据地图平台	开发实施	<ul style="list-style-type: none"> 信息架构集成：从DGC获取五层信息架构，通过云服务配置同步到数据地图。 数据湖资产同步：根据数据湖数据集成方案，同步数据湖中已有数据。 数据血缘关系导入：以搜集到的数据链路为输入，导入到数据地图。 API资产管理：采集和维护API资产，并进行配置展现。 报表资产管理：按照调研内容，遵循数据资产管理要求，采集已有报表资产。并完成与现有报表系统的对接，实现报表资产的自动化采集。 <p>从数据湖到数据地图平台，各数据链路迭代完成后，对全链路进行数据一致性校验、链路可靠性测试，并由最终用户依据业务需求进行测试。</p>
	部署上线	<ul style="list-style-type: none"> 上线计划制定：经与相关干系人沟通、确认，制定上线计划，明确上线范围和里程碑节点，经客户审批后实施 上线相关方案制定：制定上线失败回退方案、数据地图平台运维方案、上线操作手册等，并对客户相关人员进行培训。 上线检查实施： <ol style="list-style-type: none"> 清理测试阶段的数据和与商用无关的数据； 检查系统中的各项配置是否正确，并导出当前配置文件作为副本； 收回所有人的接入权限，删除测试用户的注册信息并关闭白名单； 检查上线的操作手册、脚本是否完备等。 上线实施：按照接入方案，执行数据资产接入，对数据湖，以及各个业务系统报表资产和API资产进行接入。 上线试运行：开展上线试运行支持，监控数据资产运行情况，并对接入数据资产一致性等进行监控，及时解决上线发现的问题，并输出上线试运行报告。

5 A 修订记录

发布日期	修订记录
2023-12-15	第一次正式发布。