

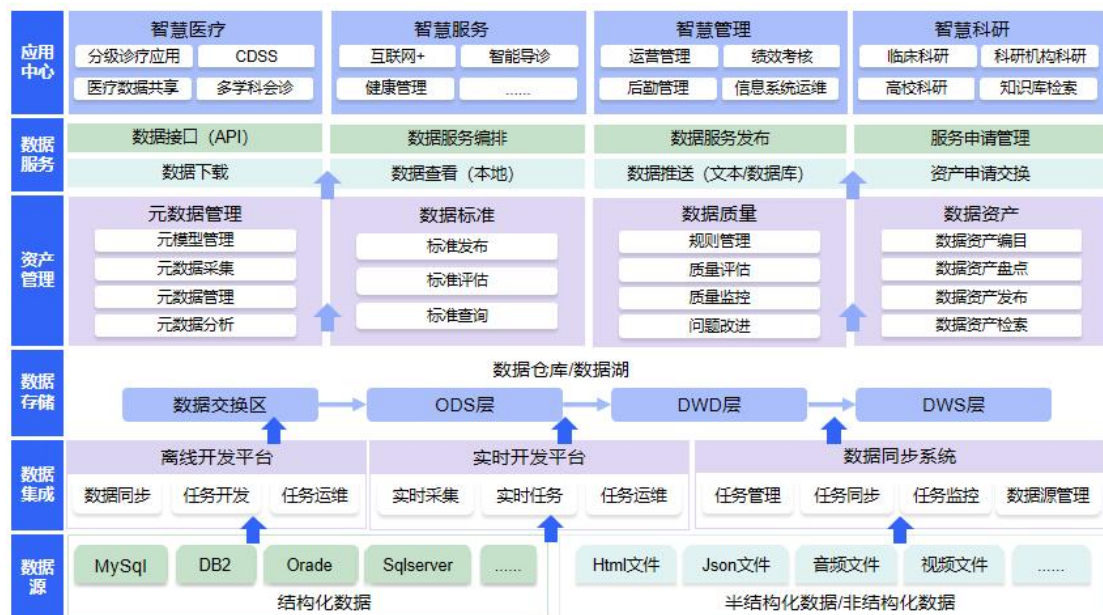
湖仓一体化平台

产品功能介绍

一、系统建设目标

湖仓一体化一站式大数据开发平台是由多个产品组成的一款实用敏捷的数据全生命周期产品。聚焦于大数据采集、交换、处理、存储、分析、管控、数据开发等数据处理及数据应用全环节，涵盖企业数据业务的全链路解决方案，包含离线和实时数据集成、数据治理、数据安全、数据服务、数据资产管理等子系统，可以助力企业数字化转型，为不同行业用户构建数据资产，实现数据价值变现且持续增值，以打破“信息孤岛”消除“数据烟囱”为理念，帮组企业把数据用起来，以发挥数据最大价值。

平台为各类业务系统和物联网设备，提供数据、消息、服务和设备集成，对业务应用屏蔽底层系统差异，提供统一服务支撑，帮助开发者快速开发整合现有系统的能力，满足数字化进程中丰富的联接需求的弹性分布式 iPaaS 平台。

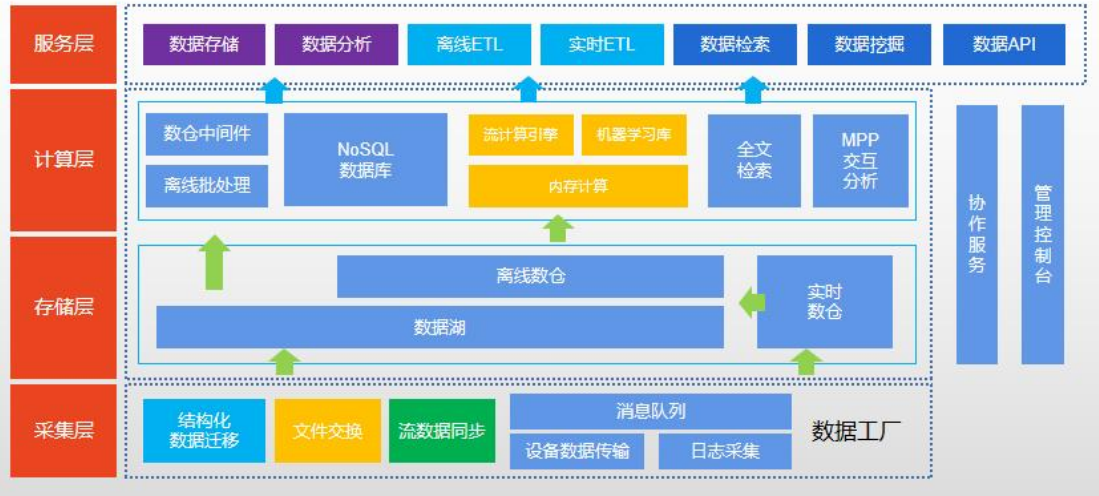


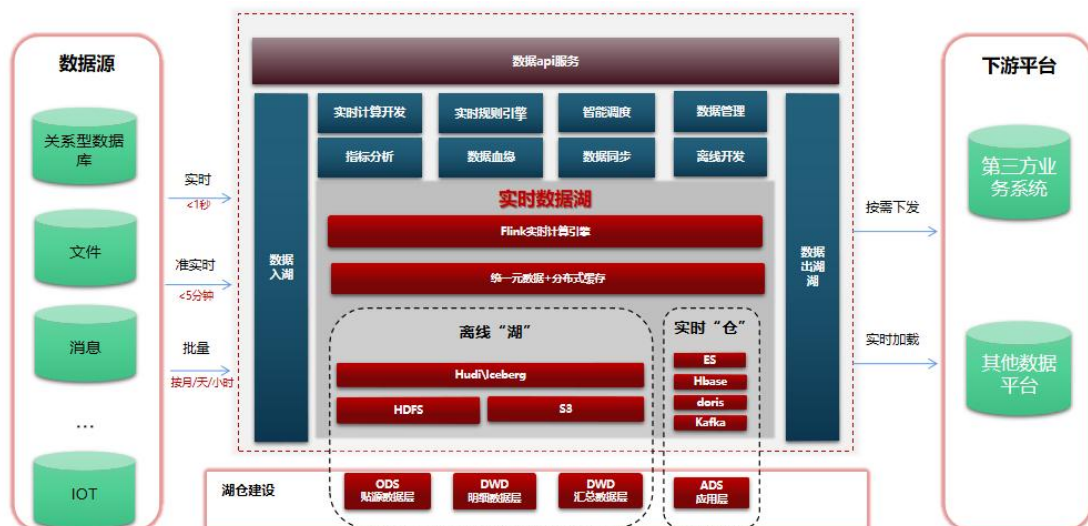
整体架构图



平台分层架构图

湖仓一体平台





湖仓一体架构图

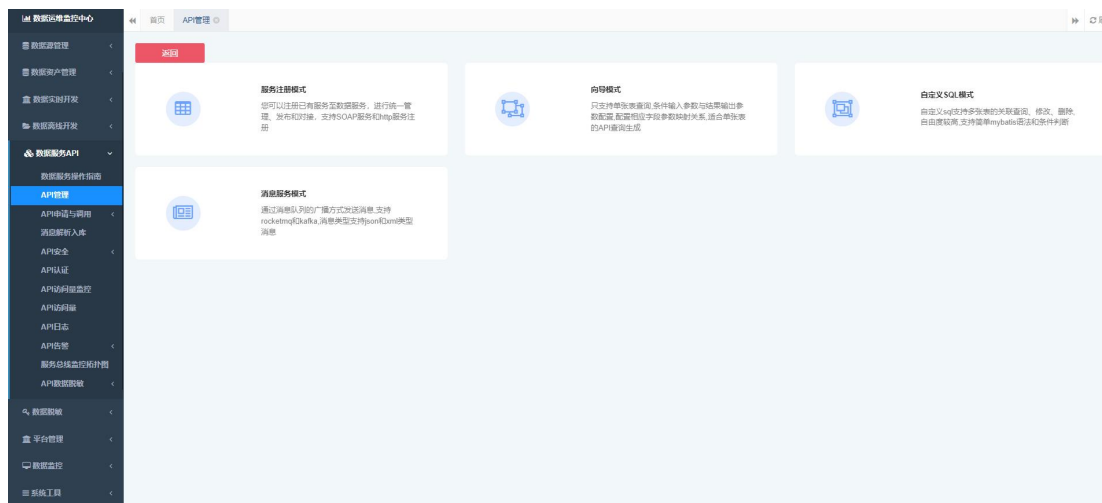
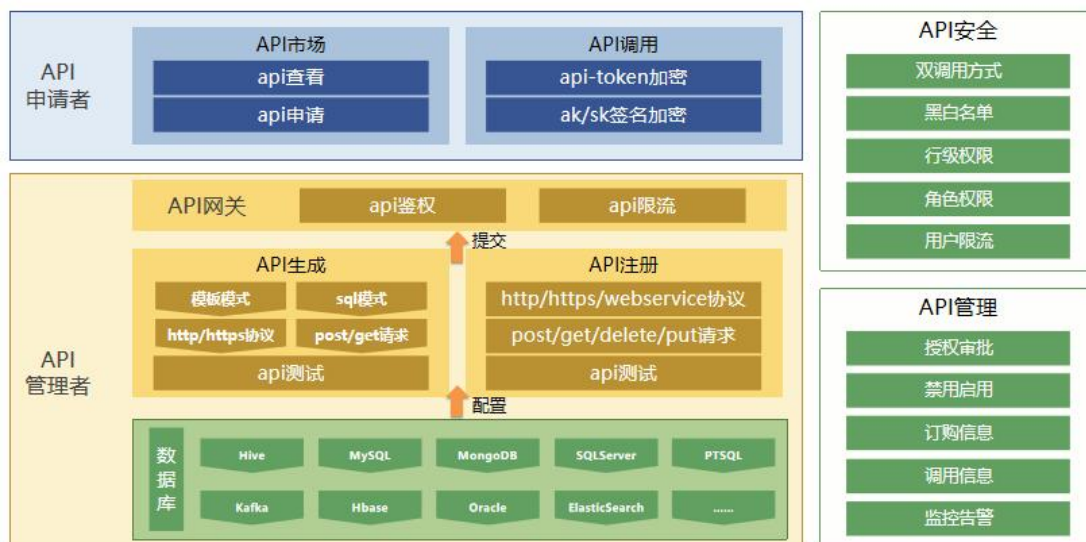
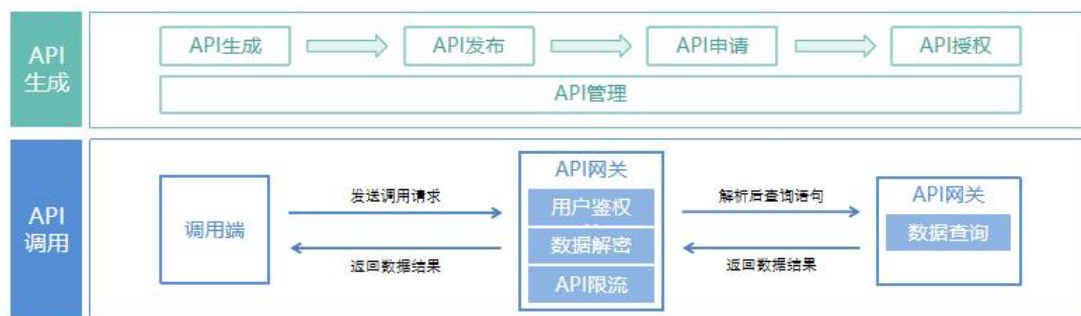
二、产品矩阵

1、数据服务平台

数据服务管理是一款支持 API 快速生成并对外输出的数据服务产品，为客户提供安全稳定、低成本、易上手的数据开放共享服务。可以为企业各种应用提供统一的数据访问服务，从而消除各种应用系统与数据中心的直接耦合性，可以将数据中台的数据共享给其他系统，无缝对接业务生产系统，主要解决数据对外快速共享的场景。

通过 API 接口的快速开发平台，不需要专业开发人员，以低代码的方式，通过配置实现 API 接口的快速开发部署。支持 API 服务的全生命周期管理，覆盖 API 创建、第三方 API 注册、发布、审批、下线、历史版本管理的整个生命周期。帮助用户规范 API 开发流程、管理 API 文档信息的同时追溯服务从生产到下线再到版本管理的各个环节。

平台提供四种 API 注册类型，满足 API 注册端的多样性，对已开发的 API 进行测试访问，查看 API 是否正常。测试正常的接口通过部署的 API 网关，完成 API 的对外开放准备。



数据源管理

数据资产管理

数据实时开发

数据离线开发

数据服务API

数据服务操作指南

API管理

API申请与调用

消息解析入库

API安全

API认证

API访问监控

API访问量

API日志

API告警

服务总线系统升级

API数据源

数据源管理

平台管理

数据监控

系统工具

首页

API管理

编辑API自定义SQL

1 基础属性

2 参数配置

3 完成

基本信息

所属项目: 数据库

API名称: apiSQL

API中文名称: api自定义sql增加

API path: /apiSQL

API描述: api自定义sql增加

API参数

请求协议: http

请求方式: POST

参数类型: body-json

安全与限制策略

每分钟调用次数上限 QPS/min: 10

安全组: API安全

返回

保存

数据中台

数据源管理

数据资产管理

数据实时开发

数据离线开发

数据服务API

数据服务操作指南

API管理

API申请与调用

消息解析入库

API安全

API认证

API访问监控

API访问量

API日志

API告警

服务总线系统升级

API数据源

首页

API管理

编辑API自定义SQL

1 基础属性

2 参数配置

3 完成

是否是查询类sql: ☒

API配置语句

SELECT
a.IDENTIFIER,
a.NAME,
b.CODE,
b.MEANING
FROM
ds_value_table a
left join
ds_value_table_domain b on a.id = b.VALUE_TABLE_ID and a.STATE = 1 and b.STATE = 1
WHERE
1=1
<if test="name != null and name != "">
and a.name like concat('%', #{name}, '%')
</if>
<if test="Identifier != null and Identifier != "">
and a.Identifier like concat('%', #{Identifier}, '%')
</if>
<if test="categoryId != null and categoryId != "">
and a.category_id = #{categoryId}
</if>

返回

上一步

保存

数据源管理

数据资产管理

数据实时开发

数据离线开发

数据服务API

数据服务操作指南

API管理

API申请与调用

消息解析入库

API安全

API认证

API访问监控

API访问量

API日志

API告警

服务总线系统升级

API数据源

首页

API管理

编辑API自定义SQL

1 基础属性

2 参数配置

3 完成

API测试

返回结果JSON

请求方式: POST

返回类型: JSON

输入参数

categoryId: string 否

name: string 否

Identifier: string 否

返回

上一步

提交

01. {
02. "msg": "操作成功",
03. "code": 0,
04. "data": {
05. "data": [
06. {
07. "CODE": "21",
08. "IDENTIFIER": "CV05.01.007",
09. "MEANING": "左肩后(LSCP)",
10. "NAME": "胎方位代码表",
11. },
12. {
13. "CODE": "04",



API调用流程图

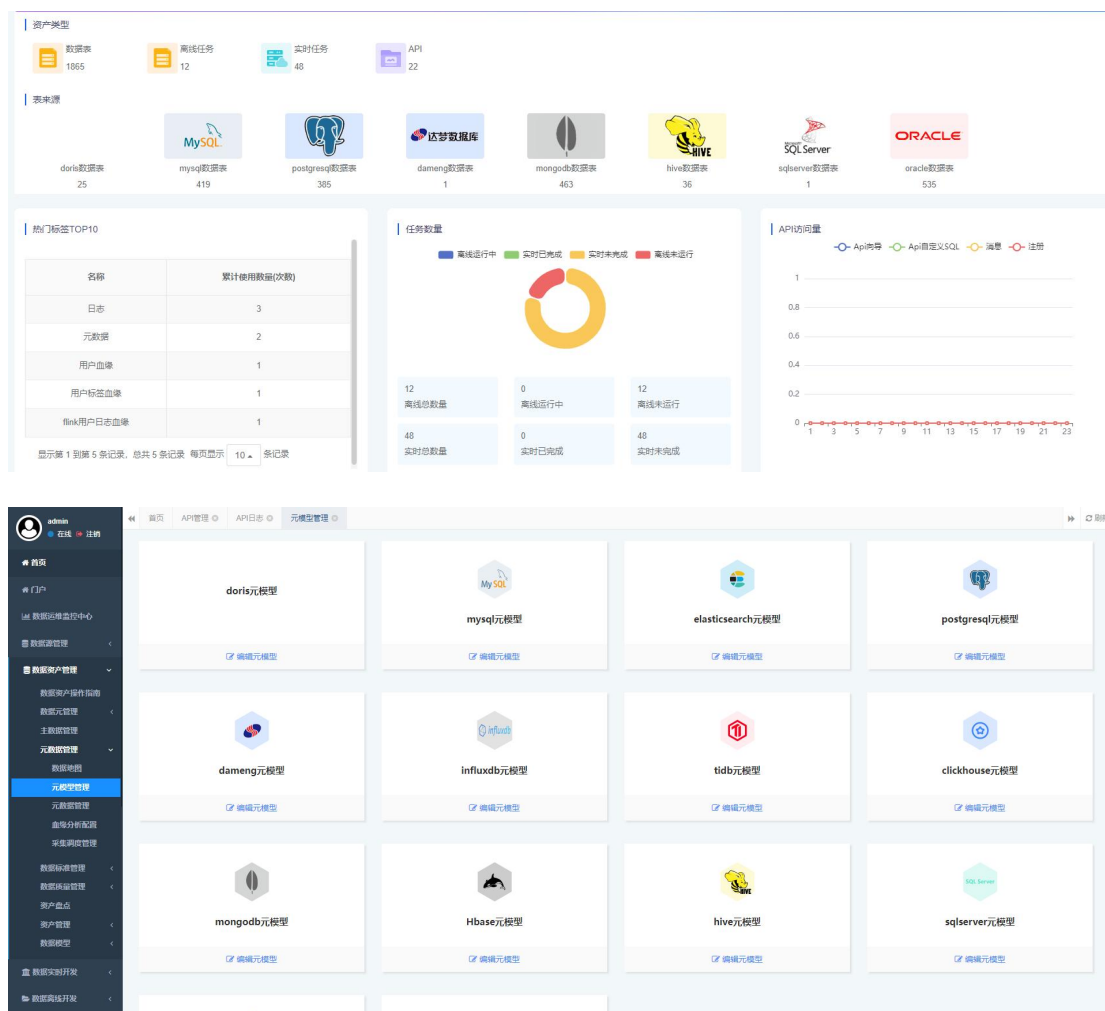
```
graph LR; A[医疗云平台] -- "请求数据" --> B[服务总线kafka]; B -- "订阅数据" --> C[HIS医院信息系统];
```

API调用记录

应用名称	用户名	订阅地址	发送/推送时间	响应时间	方向	执行结果	操作
HIS医院信息系统	俞文彦	http://192.168.31.115:18081/api/kafka/getDrugInfo	2022-09-21 11:09:17	2022-09-21 11:09:17	订阅方	成功	
医疗云平台	凌鹏飞	-	2022-09-21 11:09:15	2022-09-21 11:09:16	发送方	成功	
HIS医院信息系统	潘慧千	http://192.168.31.125:9199/api/Yest	2022-09-21 11:08:54	-	订阅方	成功	重新订阅

2、数据资产管理

企业数据资产门户，全域数据资产地图，在识别出自身数据资产的基础上，进一步构建数据资产目录，帮助用户更好的理解、使用以及分析数据。盘清数据资产的能力，理清数据的来龙去脉，以体系化的方式实现数据的可用、好用，用较小的数据成本获得较大的数据收益，为业务应用和数据获取夯实基础。数据资产主要包含元数据采集、元数据搜索、数据血缘、数据模型、数据资产编目、资源上下架、资产市场、资源申请审批等能力。



数据资产管理

数据资产操作指南

数据元管理

主数据管理

元数据管理

数据标准管理

数据质量治理

资产盘点

资产管理

主题管理

资产编目

资源挂接

审批管理

资产市场

我的数据资产

数据资产统计

资产使用普及统计

系统主题

数据中台

院长决策

电子病历

OA

ERP

OA

资产名称: 请输入资产名称

资产代码: 请输入资产代码

资产状态: 所有

搜索

重置

+ 添加资产

搜索

刷新

打印

序号	资产名称	资产编目代码	资产状态	上架状态	更新时间	操作
1	数据中台资产1	DM22092941691	草稿	未上架	2022-09-29 11:34:51	编辑 申请发布 删除 打标
2	数据中台资产3	DM22101051224	已发布	已上架	2022-10-10 14:13:45	编辑 申请变更 删除 打标
3	数据中台资产2	DM22092941718	草稿	未上架	2022-09-29 11:35:18	编辑 申请发布 删除 打标

显示第 1 到第 3 条记录, 总共 3 条记录 每页显示 10 条记录

主题: 系统主题

资产编目总数5

已发布编目数1

已上架编目0

上架数据资源总量0

主题目录资产编目统计top10

近一年已上架资产编目变化图

上架资产编目挂接数据源统计

资产编目上下架情况统计

资产编目使用Top5

资产总申请次数22

资产审批次数21

API总使用次数82

API当月使用次数0

项目总数2

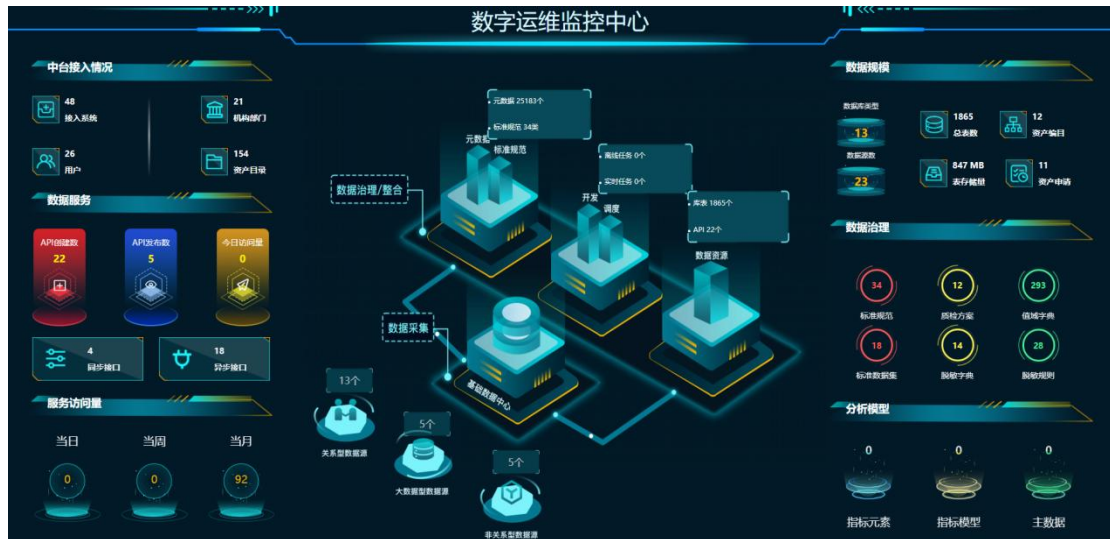
元数据TOP10

领域TOP10

API TOP10

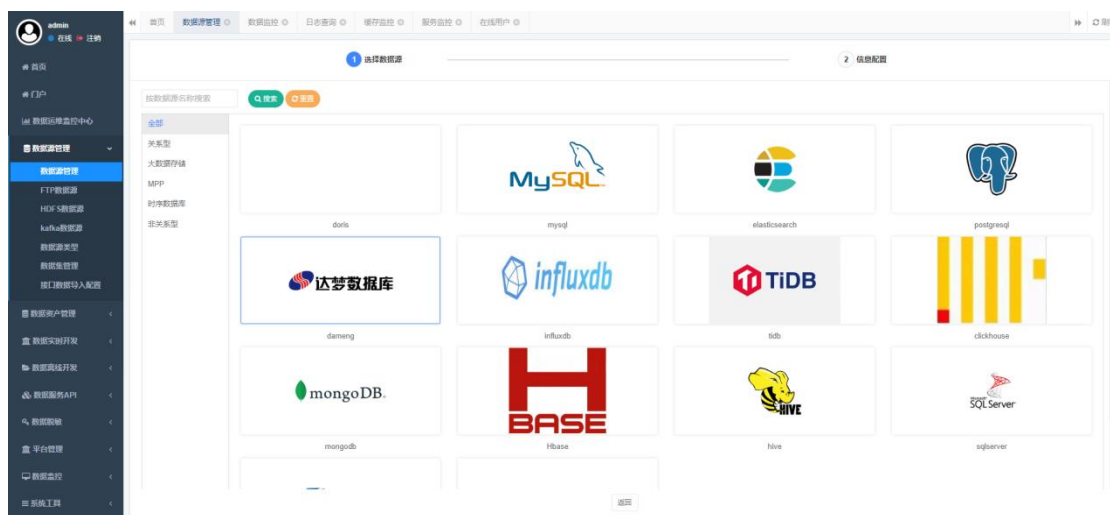
主题TOP10用户

项目TOP10

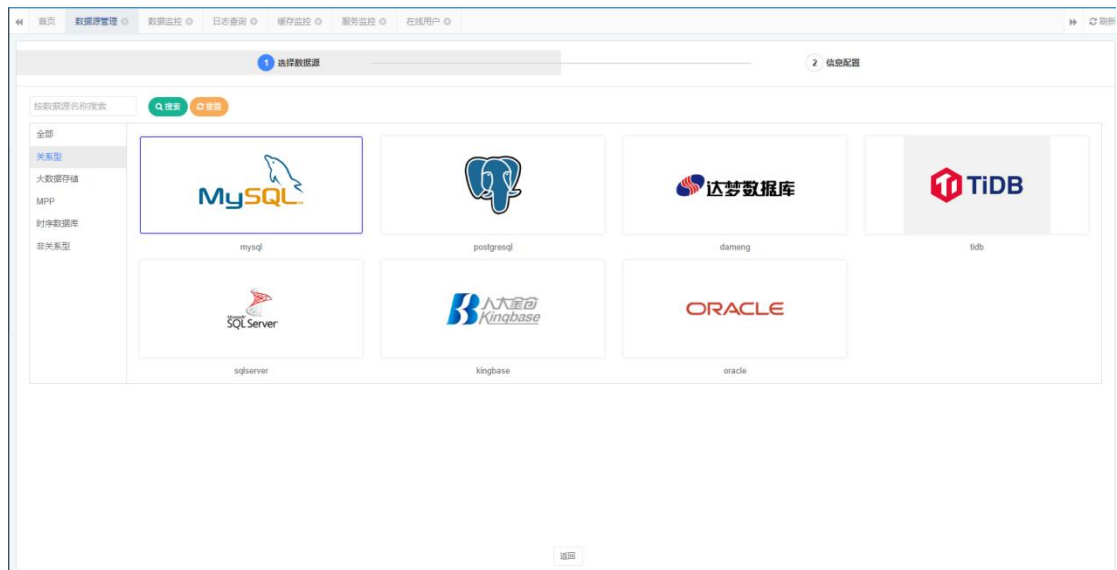


2.1 数据源管理

现在平台中支持的数据源类型除了包括 mysql oracle 等常用关系数据源外，还支持达梦，金仓等国产化数据库，以及 hbase hive 等大数据存储类型。

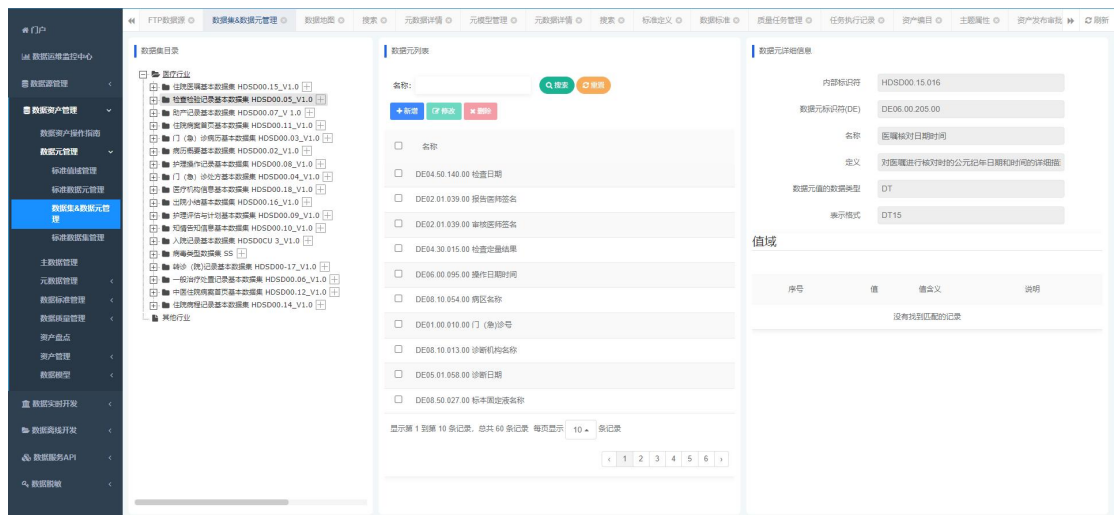


在数据源管理中可以对每种数据源进行性能指标的监控，查看数据表存储、索引大小等信息



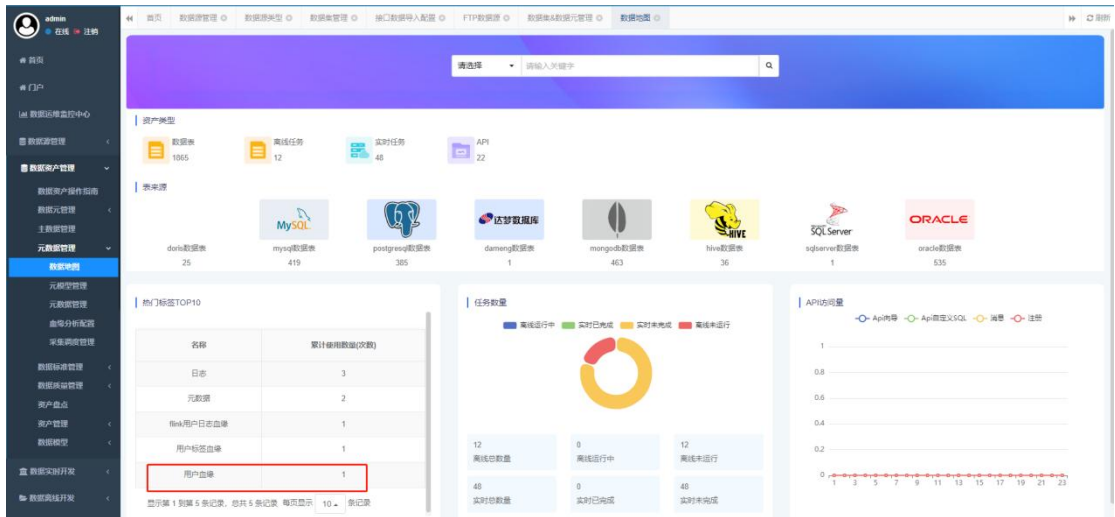
2.2 数据集数据元

数据元一般用来对各行业的数据进行自身规范化的定义，规划好行业数据元之后，可以为行业构建出统一的数据模型奠定基础。



2.3 元数据管理

通过元数据管理可以快速理清数据资源，了解数据来龙去脉，构建数据地图。



This view shows the detailed structure of the 'std_users' table. It includes a search bar and tabs for '表结构' (Table Structure), '数据预览' (Data Preview), '血缘关系' (Lineage), '版本变更' (Version Change), and '商品应用' (Product Application). The table structure is as follows:

字段名	字段中文名	字段描述	数据类型	字段标签
id			varchar	添加标签
name			varchar	添加标签
sex			varchar	添加标签
idcard			varchar	添加标签
money			varchar	添加标签
address			varchar	添加标签
age			int	添加标签
create_time			varchar	添加标签
phone			varchar	添加标签
role_id			varchar	添加标签

显示第 1 到第 10 条记录, 总共 10 条记录, 每页显示 10 条记录



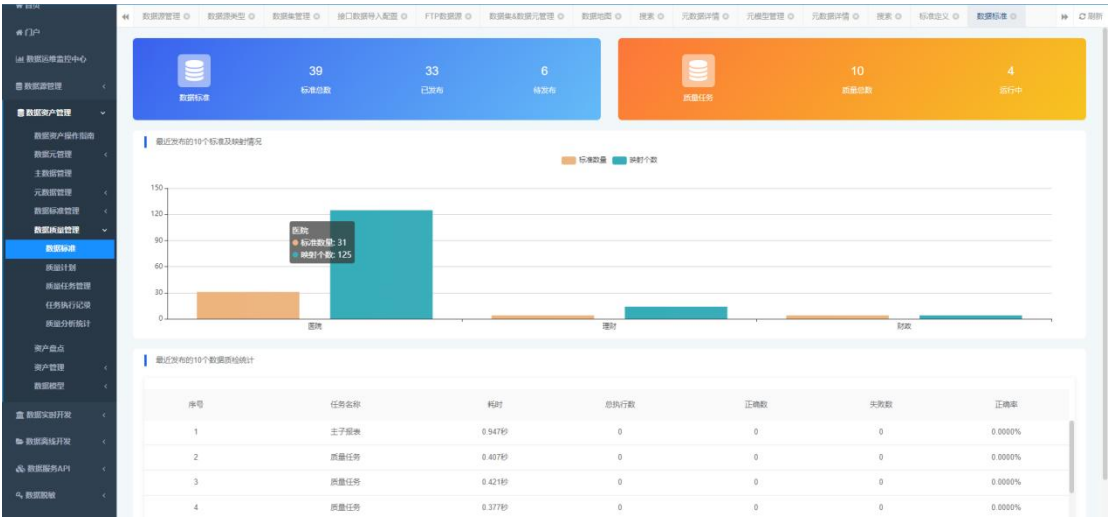
2.4 数据标准管理

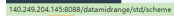
通过数据标准功能，可以制订数据的完整性、有效性、一致性、规范性的检测标准，为数据质量检查提供标准依据



2.5 数据质量管理

以数据标准为数据检核依据，数据质量可以监控到数据开发全流程中涉及各个数据实体，实现数据质量问题的全面筛选和精准定位，以可视化方式实现对数据质量的全局掌控和管理。数据质量主要包含数据标准管理、质量监控以及质量报告等能力。





1、资产主题管理

主题用于对资产进行分类管理。从组织机构、物理数据库、行业主题、业务主题、安全主题等多角度对数据进行分类，常见的分类标准可以按照企业内部的业务进行划分，也可以根据项目进行划分。

2、数据资产目录

从资产业务分类的角度，为数据模型、元数据、主数据、API 数据接口等数据资产，定义数据资产编码，通过类目，标签等方式进行分类和挂载，为数据资产进行重新的组织和定义，能够让数据变得可阅读和易理解，对外提供查看、申请授权使用的服务。

admin

在线 | 注销

我的

门户

数据运营中心

数据源管理

数据资产治理

数据资产持有制度

数据元管理

主数据管理

元数据管理

数据标准管理

数据质量治理

资产盘点

资产管理

主数据管理

资产编目

资源注册

审核管理

资产冻结

我的数据资产

数据资产统计

资产使用情况及统计

协同管理

接口数据导入记录

FTP数据源

数据集市数据元管理

数据地图

搜索

元数据管理

元数据详情

元数据详情

数据标准

数据标准管理

任务执行记录

资产编目

安全主题

切换

资产名称: 资产代码: 资产状态: 所有

Q 搜索

重置

公开

内部

私有商业

序号

资产名称

资产编目代码

资产状态

上架状态

更新时间

操作

1

测试-待编

DM22100851083

已发布

未上架

2022-10-31 10:03:42

编辑 申请变更 删除 打印

2

安全资产3

DM22092930336

已发布

未上架

2022-09-29 14:37:47

编辑 申请变更 删除 打印

3

公开演示1

DM22092952118

草稿

未上架

2022-10-09 10:45:40

编辑 申请发布 删除 打印

4

test2

DM22120157631

已发布

已上架

2022-12-01 16:00:31

编辑 申请变更 删除 打印

5

安全资产1

DM22092930264

草稿

未上架

2022-09-29 10:37:45

编辑 申请发布 删除 打印

6

测试-yit

DM22100960146

已发布

已上架

2022-10-09 16:45:08

编辑 申请变更 删除 打印

7

安全资产2

DM22092930302

已发布

已上架

2022-09-29 10:38:23

编辑 申请变更 删除 打印

8

test1

DM22120156845

已发布

已上架

2022-12-01 15:34:05

编辑 申请变更 删除 打印

9

安全资产-yit

DM22101050720

已发布

已上架

2022-10-10 14:05:21

编辑 申请变更 删除 打印

显示第 1 到第 9 条记录, 总共 9 条记录 每页显示 10 条 记录

3、发布资产和变更资产

将维护好的资产编目进行发布操作后才可以进行对该资产编目进行数据挂载。资产维护人员点击申请发布按钮后，待资产审核人员同意审批后，该编目将会由草稿状态变为已发布状态，此过程中以及发布成功之后，该资产编目的信息将不可编辑。

4、资源挂接

对审核通过的资产编目进行数据资源挂接，挂接内容包括 API 数据、元数据、标准值域。并发起上下架申请流程。

5、资产市场

已发布并审核通过的各类数据资产，会在资产市场中进行展示，在资产市场中当前登录人可以查看具有权限访问主题和类目下已上架的各类资产编目，并对相应的资产编目进行使用申请，申请通过后，在个人资产门户中可以进行查看并使用。

资产名称: 请输入资产名称

资产代码: 请输入资产代码

资产状态: 所有

Q 搜索

重置

序号	资产名称	资产编码	上架状态	更新时间	操作	挂接资源
1	测试-徐鹏	DM22100851083	已上架	2022-10-31 10:03:42	下架	-
2	安全资产3	DM22092938336	未上架	2022-09-29 14:37:47	上架	操作
3	测试-ytt	DM22100960146	已上架	2022-10-09 16:45:08	下架	-
4	安全资产2	DM22092938302	已上架	2022-09-29 10:38:23	下架	-
5	安全资产-ytt	DM22101050720	已上架	2022-10-10 14:05:21	下架	-

显示第 1 到第 5 条记录, 总共 5 条记录 每页显示 10 条记录

数据资产管理中心

数据资产元数据管理 数据地图 搜索 元数据详情 元模型管理 元数据详情 搜索 标准定义 数据标准 质量任务管理 任务执行记录 资产编目 主题属性 资产发布审批 刷新

数据源管理 数据资产目录 数据资产操作指南 数据元管理 主数据管理 元数据管理 数据标准管理 数据标准管理 资产盘点 资产管理 主题管理 资产编目 资源管理 数据管理 资产发布审批 资产上下架审批 资产使用审批 资产市场 数据资产目录 数据资产统计 资产使用统计 数据管理 数据模型

待审批的 已审批的

资产名称: 请输入资产名称 资产代码: 请输入资产代码 申请状态: 所有 Q 搜索 重置

序号	资产名称	资产代码	申请人	申请状态	申请时间	操作
1	测试测试	DM22100850870	管理员	申请发布	2022-10-08 14:09:01	发布

显示第 1 到第 1 条记录, 总共 1 条记录 每页显示 10 条记录

数据资产管理中心

数据资产元数据管理 数据地图 搜索 元数据详情 元模型管理 元数据详情 搜索 标准定义 数据标准 质量任务管理 任务执行记录 资产编目 主题属性 资产发布审批 刷新

数据源管理 数据资产目录 数据资产操作指南 数据元管理 主数据管理 元数据管理 数据标准管理 数据标准管理 资产盘点 资产管理 主题管理 资产编目 资源管理 数据管理 资产发布审批 资产上下架审批 资产使用审批 资产市场 数据资产目录 数据资产统计 资产使用统计 数据管理 数据模型

待审批的 已审批的

资产名称: 请输入资产名称 资产代码: 请输入资产代码 申请状态: 所有 Q 搜索 重置

序号	资产名称	资产代码	申请人
1	测试测试	DM22100850870	管理员

显示第 1 到第 1 条记录, 总共 1 条记录 每页显示 10 条记录

标签

系统信息

数据资产代码: DM22100850870 数据资产名称: 测试测试

数据资产类型: OA Q

基础性

数据资产类型: 数据资产

来源系统: 数据中台 发布日期: 发布

更新周期: 请选择

信息资源提供方属性

数据资产提供方: 提供方联系方式: 提供方内部部门: 提供方开放属性

共享类型: 无条件共享 是否向社会开放: 是

信息资源分类

完成



6、数据资产统计

该功能可以对各个主题下数据进行统计。可以切换主题进行查看。

数据资产看板：资产编目总数、已发布数、已上架数、已上架编目挂架的数
据资源总数、资产编目上下架情况统计、上架资产编目挂载资源统计。



7、资产使用情况统计

查看当前资产的总申请数

查看当前所有资产的审批次数

查看已申请使用的 API 的总使用次数

查看已申请使用的当前 API 使用次数

查看当前项目的总数

以用户为维度统计当前用户所拥有资产数量情况

以主题为维度统计当前维度下元数据、API、标准值域、编目数量

以项目为维度统计当前维度下元数据、API、标准值域、项目人员数量

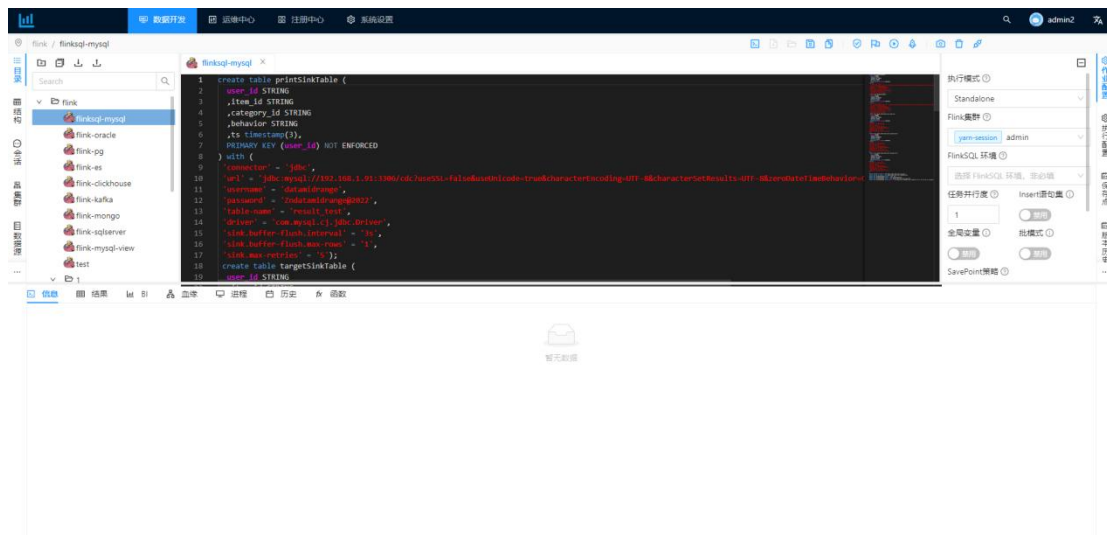


3、数据开发管理

数据开发管理平台是一个数据实时同步+实时数据处理+实时数据开发的数据平台，提供低代码形式的开发模式，降低企业大数据开发的技术门槛，提升数据开发效率

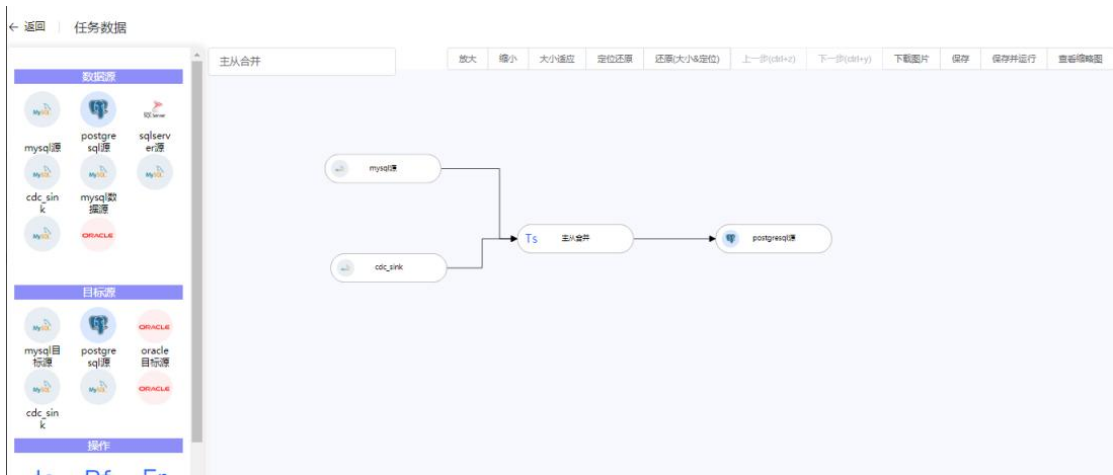
系统是基于 Apache Flink 构建的高性能、一站式实时数据开发平台，其包含实时开发、调试、实时运维监控等功能，支持多元异构数据源接入，提供面向PB 级别海量数据的秒级分析能力。

兼顾多种数据源实时采集能力,提供 Flink 大规模计算能力,支持 FlinkJar、FlinkSql、Mysql、Oracle、Clickhouse 等任务类型，提供丰富的开发功能和完善的权限管控，保障数据开发的易用性和灵活性。



4、数据同步系统

对异构数据源实时高性能的一站式数据同步系统。用于企业多源数据融合，解决关系型数据库、非关系型数据库及大数据平台等异构数据源之间，大数据量高并发下的交互和数据同步问题，实现企业数据的共享。具备高并发、高容错，自动负载均衡能力，提供灵活的任务管理功能，全面支持主流的开源数据库及商业数据库以及消息类中间件，非侵入式数据同步，基于数据库日志文件，无需安装插件，秒级同步 DML 和 DDL 操作，可实现秒级实时数据传输能力，支持同构/异构数据源之间批量、增量、实时的数据迁移服务；应支持整库数据迁移，快速实现大量库表全量、增量数据同步。



5、数据脱敏管理

基于数据中台的数据脱敏管理系统，能有效将客户敏感信息进行保护，避免敏感信息在使用时发生泄露，更能保障客户的使用隐私。系统采用高性能、高扩展性的敏感数据自动发现动态数据脱敏和静态数据脱敏功能，可实现自动化发现源数据中的敏感数据，并对敏感数据按需进行脱敏规则变形，避免敏感数据泄露。同时，脱敏后的数据保持了数据一致性和业务关联性，应用于开发测试环境、数据交换、数据分析、数据共享等场景。



1)、脱敏策略管理

(1) 表级敏感策略

表级敏感策略制定模块针对数据中心各个数据实体定义安全级别，用于表级权限控制。数据中心可以根据各个系统数据实体定义数据灵活定义敏感度分级策略。

(2) 字段级敏感策略

字段级敏感定模块针对数据实体之上再次细分字段级数据敏感级别控制，用于服务于数据开放场景下的同一数据源的多种开放策略。

(3) 数据脱敏规则

在数据权限模型的基础上，还可以对数据定义脱敏规则。该模块可以灵活的设置新增脱敏规则，脱敏规则与多种脱敏算法相结合，可以实现各行业场景各种敏感数据的业务需求，脱敏规则和脱敏算法支持用户自定义扩展。

（4）脱敏字典

系统内置了大量的各行业的业务数据字典，字典可以根据用户的行业和实际的场景进行增加扩展，脱敏字典用于仿真替换时使用，以保证脱敏后的数据保持了数据的一致性和业务的关联性。

（5）灵活脱敏数据分发

支持广泛的数据脱敏分发方式，支持数据库到数据库、数据库到文件、文件到文件、文件到数据车四种完全不落地的脱敏方式；

2）、脱敏算法

丰富的脱敏能力和脱敏数据以假乱真，支持多种脱敏算法，支持随机映射、固定映射、替换、加减值、范围随机、截断、截取、加密、格式化脱敏等多种算法；脱敏功能能最大限度的保证数据原始特征保持业务规则关联性，保持数据逻辑一致性，确保交付可用、可靠的高质量数据；

（1）遮蔽脱敏：对敏感数据的全部或部分内容用“*”或者“#”等字符进行遮蔽，导致敏感数据全部或部分不可见

（2）仿真随机：采用和原数据结构相同，内容相近的随机内容进行随机替换，确保数据格式不变

（3）仿真替换：采用和原数据结构相同的数据进行替换，对相同的原数据，脱敏后数据也相同

（4）乱序脱敏：打乱原数据的排列顺序，但是不改变原数据的值，是破坏数据关联性的脱敏方式



3)、敏感数据自动感知

数据中台中存放着海量的数据，对于数据的使用方，很难完全清楚，哪些数据属于敏感数据，自动识别功能，可以让用户方便的了解自己的数据中有哪些涉及互敏感数据。对于敏感数据的识别字典可以根据用户所在的行业进行设置扩展。系统利用各类敏感信息规则和算法，通过自动扫描发现的方式高效、方便、全面的获取的感知信息，支持灵活的配置方式自动探测集成到数据中台的各类数据源中的敏感信息字段，感知后的数据可以设置是否自动进行脱敏；

4)、动态脱敏

一般用在生产环境，动态脱敏可以实现不同用户拥有不同的脱敏策略，在访

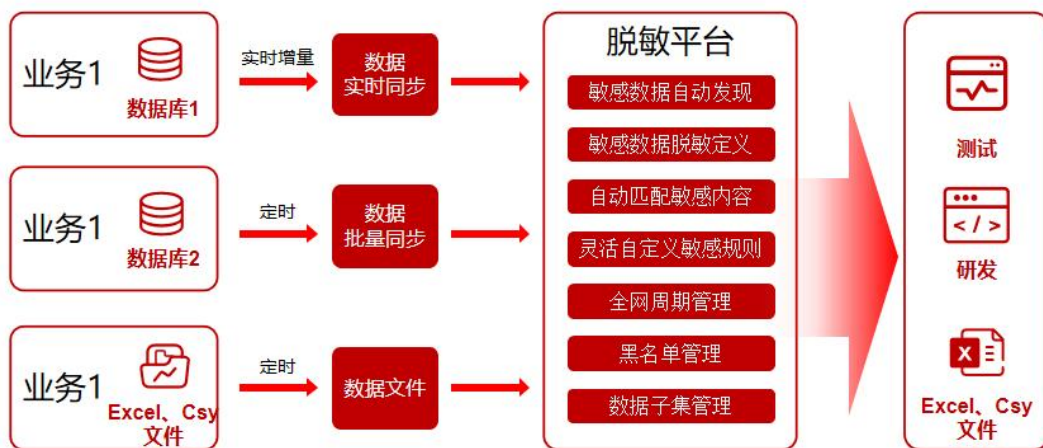
问敏感数据当时进行脱敏，用来解决在生产环境需要根据不同情况对同一敏感数据读取时需要进行不同级别脱敏的问题。

各业务系统如果需要根据不同的使用人员、角色权限进行敏感数据的展现时，可通过动态脱敏功能针对不同级别的人员进行权限管理，只需要配置应用、配置第三方应用中对应计划的的分权策略，策略包含用户、ip、时间段、角色等。各业务系统通过接口调用，便可实现针对不同系统，根据设定的脱敏策略将敏感数据进行脱敏，实现不同权限的人员查看同一数据时，反馈不同的数据信息；



5、静态脱敏

具备高性能、高扩展性的数据屏蔽和脱敏功能，采用专门的脱敏算法对敏感数据进行变形、屏蔽、替换、随机化、加密，将敏感数据转化为虚构数据，隐藏了真正的隐私信息，但是具备可逆脱敏能力，同时保证脱敏数据之间的一致性、关联性。为数据的安全使用、共享提供了基础保障。同时无需改变相应的业务系统逻辑，实现了企业低成本、高效率、安全的使用生产的隐私数据。在敏感数据从生产环境脱敏完毕之后再在非生产环境使用，一般用在非生产环境，如开发、测试、外包和数据分析等环境。



6、指标管理系统

一站式企业数据指标管理平台，构建企业数据指标体系，是一款涵盖指标定义、指标建模、指标可视化分析，实现指标全流程、一体化管理的系统。旨在帮助各类企业建设指标体系，疏通指标数据，落地指标管理，发挥指标最大价值。通过指标的规范化定义、标准化开发，搭建企业数据指标体系，落地指标数据结果，消除数据二义性，实现指标数据的可视、可用、可管。

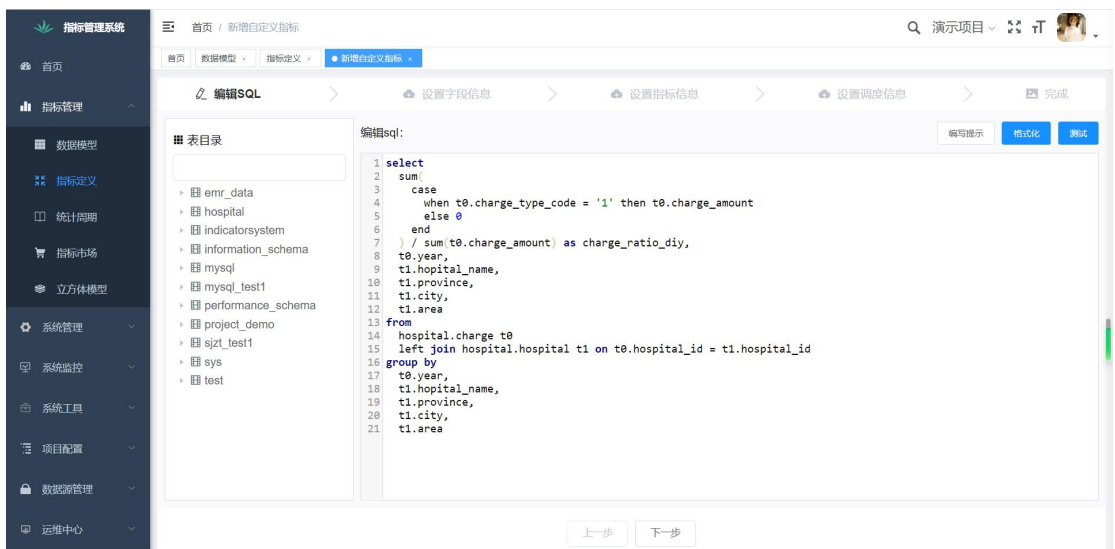
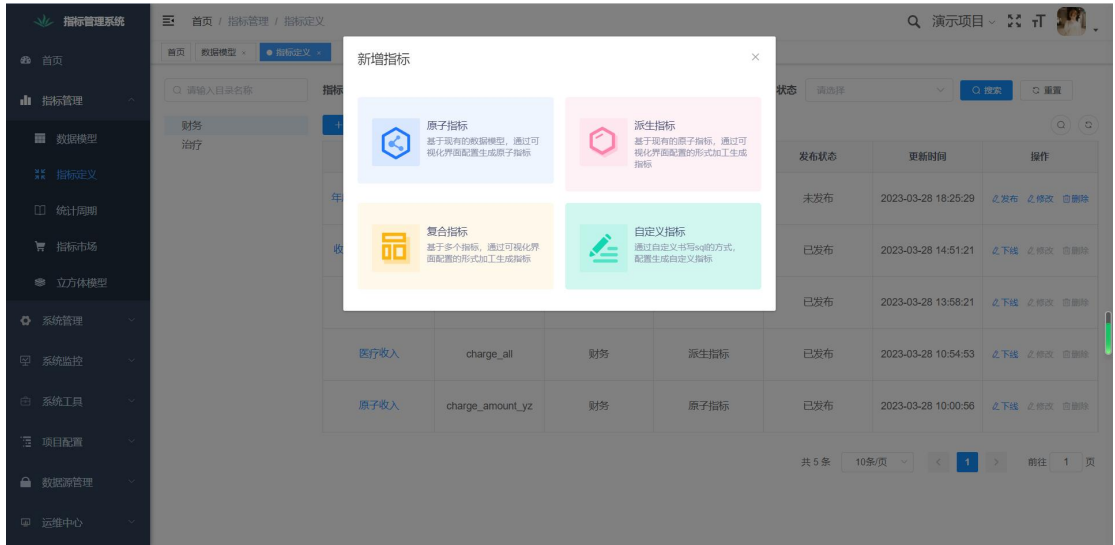
将分散在不同系统的各类指标集中管理，并用统一的标准进行约束，清晰展现用户指标一览表及各指标的统计方法、数据来源、统计口径等信息，整个指标体系一览在目。



为企业规范指标，统一指标口径，便于后续指标共享和应用，避免多方操作导致的指标混乱、数据不一致等问题。

随时查看指标数据，了解指标趋势，提供可视化的方式自由查看指标，可多种维度对指标进行分析。

摒弃复杂冗长的指标定义过程和枯燥的指标定义方法，指标定义配置步骤简单、界面清晰，业务人员轻松驾驭。

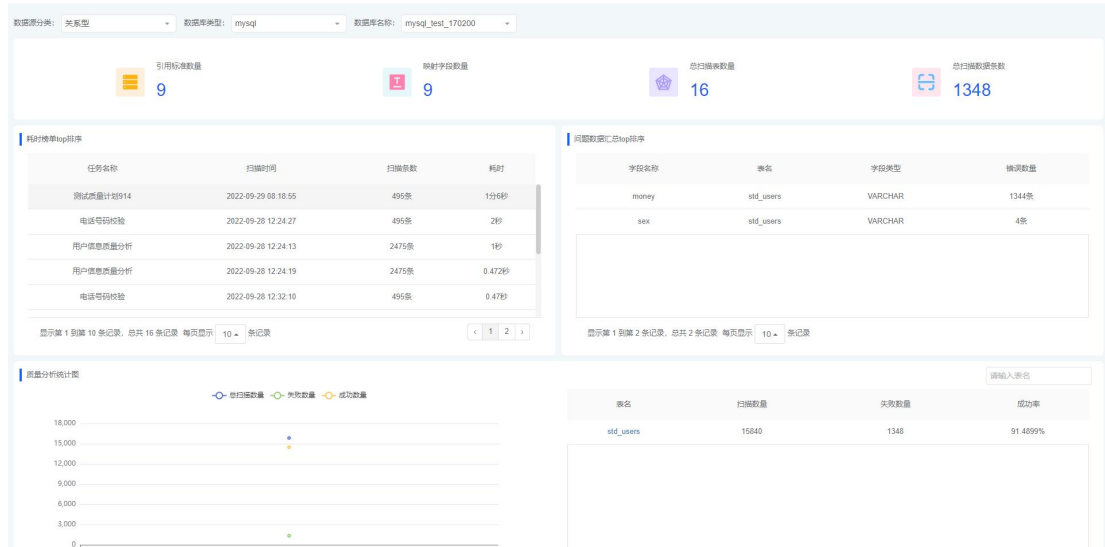


7、数据质量管理

检核数据质量的能力，通过数据标准功能，可以制订数据的完整性、有效性、一致性、规范性的检测标准，为数据质量检查提供标准依据。数据质量可以监控到数据开发全流程中涉及的各个数据实体，实现数据质量问题的全面筛选和精准定位，以可视化方式实现对数据质量的全局掌控和管理。依托数据质量的全库质量评估能力，可高效、方便的对整个数据库的数据质量进行检查与管理。通过向导、可视化的简易操作手段，将质量评估、质量检核与质量报告等工作环节进行流程整合，形成完整的数据质量管理闭环。

数据质量主要包含数据标准定义、标准映射、质量计划、质量任务管理、质量分析统计以及质量报告等能力。





数据运营监控中心

数据源管理 数据资产管理 数据元管理 主数据管理 元数据管理 数据标准管理 目录类型 标准定义 标准映射 数据质量治理 资产盘点 资产管理 数据实时开发 数据离线开发

标准定义

添加标准定义

业务属性

* 中文名称: 由中文、英文字母、数字、下划线 ()、and (&)、英文

* 英文名称: 由中文、英文字母、数字、下划线 ()、and (&)

标准编号: 由中文、英文字母、数字、下划线 ()、and (&)、英文

* 标准目录: 请选择归属目录

* 业务定义: 请输入

标准来源: 请输入

自定义属性: +添加

备注: 请输入

技术属性

设置自定义关联数据元: 自定义

数据类型: 字符型

数据长度: ~请选择 输入数值

是否允许为空: ☐ 可空 ☐ 不可空

是否允许重复: ☐ 可重复 ☐ 不可重复

默认值: 请输入

枚举范围: +引用 +自定义增加

属性值	名称	来源	操作
-----	----	----	----

数据运营监控中心

标准映射

标准目录

请输入标准名称进行搜索

中文名称: 测试y

名称 (引用): 名称 (引用)

身份证: 身份证

hive测试姓名: hive测试姓名

金额: 金额

性别: 性别

it: it

精度测试: 精度测试

标准3: 标准3

手机号 (引用): 手机号 (引用)

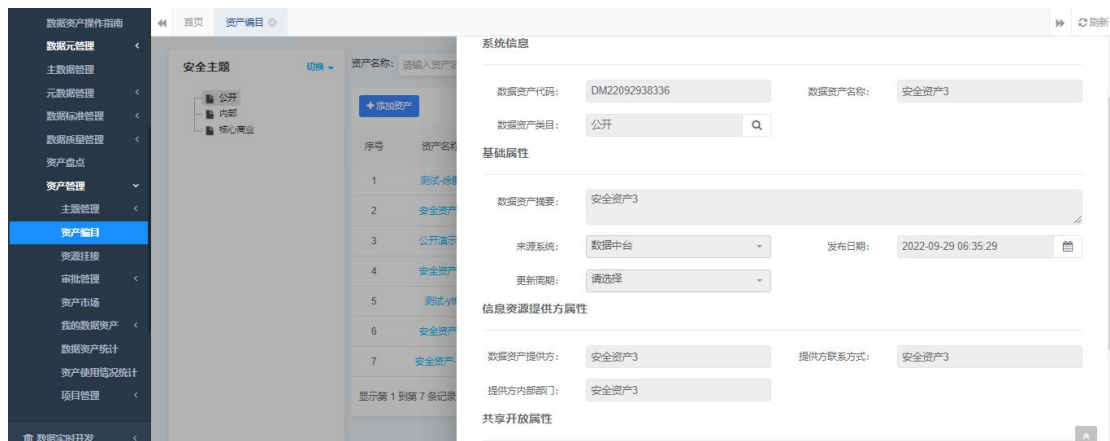
映射记录

金额 已发布

创建时间: 2022-08-15 13:38:02.63 最近发布时间: 2022-08-29 09:37:55.867 最近发布用户: admin

字段名	来源表	数据源类型	数据源	绑定时间	操作
money	std_user	postgresql	ptsql_test_170211	2022-08-30 11:54:06	解除绑定
MONEY	STD_USER	oracle	oracle-test	2022-08-18 14:56:01	解除绑定
money	std_users	mysql	mysql_test_170200	2022-08-18 14:55:45	解除绑定

显示第 1 到第 3 条记录, 总共 3 条记录 每页显示 10 条记录



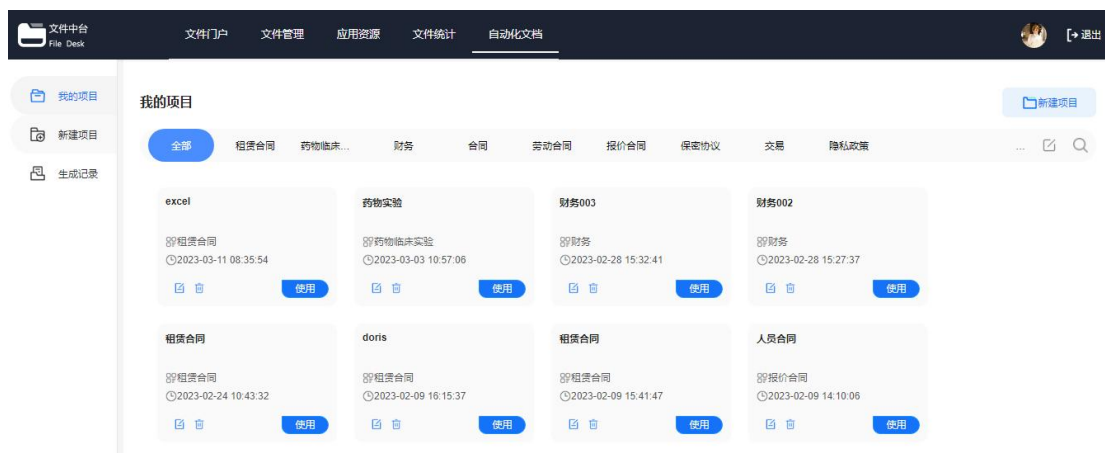
8、文件中台

实现企业和组织非结构化数据的统一存储、统一管理和统一服务。

通过丰富的 API 接口，打通各业务系统，提供强大的文件中台管理能力。横向打通了非结构化数据孤岛，构建起统一的非结构化数据中台，实现数据互通、业务融合。

通过平台提供的无需写代码的文档自动化系统，可以将复杂繁琐的文档制作工作转化为高效的自动化应用，只要您会用 WORD 写文档，您就会创建文档自动化工具。





三、核心功能

1、数据集成

实时数据传输平台是一个以设计、部署、调度、监控和管理 ETL 过程为核心功能的数据抽取汇聚系统，提供多种数据源（关系型数据、非关系型数据等）之

间灵活快速的数据实时离线同步，以及数据转换、数据预处理能力，并能方便的进行部署、调度及监控等管理活动，真正地提供一体化数据集成开发环境。

1、图形化操作界面

数据集成提供全流程编排可视化操作，基于 BS 架构兼容多种设备和浏览器访问，提供数据抽取、数据处理、数据加载工具，开发人员只需要在图形化界面中通过托拉拽的方式就可以轻松实现复杂的任务配置，不需做任何编码，业务人员即可轻松掌控数据流向，应对各种负责数据处理，降低数据集成难度及使用门槛。

2、强化数据处理

支持批量离线同步、实时同步、增量同步、全库迁移等能力。基于分布式数据流的数据处理模型，最大程度的发挥各步骤的处理能力；同时支持多实例并行执行。提供超过七十种数据处理转换组件集合，支持数据多样化性能处理，简化大批量数据集成中的人工操作。

3、作业统一调度

作业调度支持多时间维度(天、小时、分钟)的任务定时调度，只需简单几步便可完成数据增量抽取，配置简单、灵活、易用，支持灵活的、多角度的模型作业调度机制，包括实时调度、定时调度、时间间隔循环调度等，通过系统整合优化最终实现 All-in-One 的统一自动化调度管理。

4、全面的运行监控

系统提供了可视化的多维度作业运行监控，使得极佳的易用性获得最大限度的推广，提供运行环境多个执行器的资源监控，包括 CPU 占用、内存使用、磁盘空间等，通过统一监控，可合理分配任务的调度，提高数据处理效率。同时，当任务出现错误、执行器负载过高等场景时，数据集成支持通过预定义的方式告知用户实时告警信息。

5、非侵入式数据同步

基于数据库日志文件，无需安装插件，秒级同步 DML 和 DDL 操作，基于 CDC 数据变更捕获技术实现增量数据实时同步、处理，变更数据实时发送。

2、消息集成

消息集成基于分布式消息系统 Kafka 和 Rocket 双引擎，使用统一的消息接入机制，标准化的消息通道，提供可靠的、可持久化的、高吞吐量的准实时消息管道系统，广泛应用于数据流的实时生产、传输及消费。

1、安全可靠的消息传输机制

兼容原生的 Kafka 和 Rocket 特性，具备原生 Kafka 和 Rocket 所有消息处理特性，支持安全的消息传输，消息存储加密等措施加强网络访问控制，支持消息数据高可靠，支持消息持久化，多副本存储机制。

2、高性能准实时消息处理

消息集成作为数据存储的中间环节，提供高吞吐量的消息处理服务，延迟可缩小到毫秒 ms 级别。原生客户端没有提供查看 Topic 消息内容的机制，消息集成为用户提供了 Web 页面，可实时查看所有 Topic 中所有分区的任意偏移量的消息正文。消息查询主要提供毫秒级、可视化的集群中的消息查询能力，支持按照分区和生产时间进行过滤。

3、消费进度

对于较大数据量的 Topic 的消费进程，用户可以通过消息集成提供的 Web 页面实时查看消费者组消费的所有 Topic 的具体分区的消费进度。

4、Topic 重分配

在集群资源使用已超出系统配置的资源时，需要通过扩容 Kafka 节点来实现集群的资源扩容。新增的节点，只有在创建新的 Topic 才会参与工作，对于之前其他 Broker 节点上的分区是不会自动均衡的，不能达到负载的效果。当集群扩容以后，为了减轻原 Broker 节点的负载压力，通过消息集成的 Topic 重分配功能将 Topic 的所有分区重新分配到扩容后的所有节点完成负载均衡。

5、全方位消息指标监控

消息集成提供集群的各种监控指标，包括 Topic 堆积量，消息数量，消息容量，生产/消费请求失败数，被拒绝的字节数，消息转换数量，生产/消费请求数量等。

3、服务集成

服务集成作为连接上层应用系统与底层系统的桥梁，通过 API、数据包、业务模型等方式对上层应用提供服务，减低数据应用开发门槛，实现服务重用、服务沉淀、建立企业能力生态。数据服务可简单概括为将大数据平台中的数据共享给其他系统，或将某组织的数据共享给其他组织，主要解决数据对外快速共享的场景。

按照传统方式，生成 API 接口，往往需要后端开发人员通过 Java 或 Python 等语言进行编写。从开始生成到对外发布整个流程较长，且一个接口开发完成后，需要测试人员再进行测试验证，整个过程下来，投入成本较高。

基于数据中台的数据共享服务产品，其将数据服务能力进行封装，大部分功能在平台产品内部完成，面向用户的功能只是连接数据源、编写查询逻辑，大幅度缩短 API 流程，并降低开发成本。

4、数据资产

盘清数据资产的能力，理清数据的来龙去脉，以体系化的方式实现数据的可用、好用，用较小的数据成本获得较大的数据收益，为业务应用和数据获取夯实基础。数据资产主要包含元数据采集、主数据管理、数据血缘、数据资产目录、数据资产管理以及数据地图等能力。

5、数据质量

检核数据质量的能力，通过数据标准功能，可以制订数据的完整性、有效性、一致性、规范性的检测标准，为数据质量检查提供标准依据。数据质量可以监控到数据开发全流程中涉及的各个数据实体，实现数据质量问题的全面筛选和精准定位，以可视化方式实现对数据质量的全局掌控和管理。依托数据质量的全库质量评估能力，可高效、方便的对整个数据库的数据质量进行检查与管理。数据质量主要包含数据标准管理、质量监控以及质量报告等能力。

6、数据开发

全链路数据开发能力，将数据集成、离线分析、实时计算等模块深度融合，打造出全域数据开发和全链路的数据监控，让用户轻松能够看到整个开发状态和统计监控。数据开发主要包括数据实时集成、离线集成、实时计算、多维分析、业务流程、任务调度等能力。

四、平台优势

1、海量异构数据集成

企业业务系统的复杂性和多样性，一方面带来了数据的种类的多样性，如传统关系型数据库数据、消息、API、文件等对数据集成的功能完备性提出了挑战。另一方面也随着业务的发展，数据量也呈现爆发式的增长，对数据集成的效率也提出了很高的挑战。平台支持 30+种类的数据源，提供数据、消息、API、物联网协议等多种集成能力，包括针对批处理和实时场景的数据集成和应用程序高效安全的集成服务，实现一套集成平台完成海量异构数据的集成。

2、全生命周期的 API 治理

提供五种 API 注册类型，满足 API 注册端的多样性，对已开发的 API 进行测试访问，查看 API 是否正常，对异常的 API 进行相应的修复。测试正常的接口通过部署的 API 网关，完成 API 的对外开放准备。API 网关支持跨网的 API 开放，帮助企业快速实现跨网协作。授权功能对 API 进行访问权限授予和回收，生命周期结束的 API 进行删除。产品功能角度，提供 API 的注册、测试、部署、授权、编辑、删除等全生命周期的 API 管理。业务流程角度，融合集成平台提供应用系统层的数据库的 API 接入和生成能力，业务流程层提供可视化的服务编排的能力，消费体验层提供安全的授权、申请、订阅、审批控制能力，提供全生命周期的 API 管理、监控和治理能力。

3、先进的数据资产化理念

数据运营平台通过对数据资产的统一管理，以全局多维度资产视图，快速盘清资产数据的分布和统计情况，依据业务属性自定义对数据按照主题、分层、标签进行不同维度统计，清理全数据链路自动形成数据流转血缘关系图，实时了解数据的来龙去脉。

4、低代码开发模式

支持多人在线协作开发，支持 Java、MR、Spark、SQL、Shell、Flink 多种数据处理类型，作业开发支持 SQL 和画布拖拉拽多种方式，提供丰富的调度配置策略与海量的作业调度能力。

5、数据安全与隐私保护

系统采用高性能、高扩展性的敏感数据自动发现动态数据脱敏和静态数据脱敏功能，可实现自动化发现源数据中的敏感数据，并对敏感数据按需进行脱敏规则变形，避免敏感数据泄露。同时，脱敏后的数据保持了数据一致性和业务关联性。

6、非侵入式数据同步

非侵入式数据同步，基于数据库日志文件，无需安装插件，秒级同步 DML 和 DDL 操作，可实现秒级实时数据传输能力，支持同构/异构数据源之间批量、增量、实时的数据迁移服务；支持整库数据迁移，快速实现大量库表全量、增量数据同步；

