ALIBABA CLOUD

阿里云

专有云敏捷版数据智能场景

数据资源平台

产品简介

产品版本：V1.2.1

文档版本：20221212

阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。

2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。

3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。

4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。

5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。

6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	危险 重置操作将丢失用户配置数据。
警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	警告 重启操作将导致业务中断，恢复业务时间约十分钟。
注意	用于警示信息、补充说明等，是用户必须了解的内容。	注意 权重设置为0，该服务器不会再接受新请求。
说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击**设置**> **网络**> **设置网络类型**。
粗体	表示按键、菜单、页面名称等UI元素。	在**结果确认**页面，单击**确定**。
`Courier字体`	命令或代码。	执行 `cd /d C:/window` 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	`bae log list --instanceid` *Instance_ID*
[] 或者 [a|b]	表示可选项，至多选择一个。	`ipconfig [-all|-t]`
{} 或者 {a|b}	表示必选项，至多选择一个。	`switch {active|stand}`

目录

1.DIStack整体介绍 08
2.产品优势 09
3.产品架构 10
4.功能特性 11
5.应用场景 25
6.数据资源平台DataQ产品简介 27
6.1. 什么是数据资源平台 27
6.2. 产品优势 29
6.2.1. 核心能力 29
6.2.2. 用户价值 35
6.3. 功能特性 36
6.3.1. 研发工作台 36
6.3.1.1. 数据同步 36
6.3.1.2. 空间数据管理 38
6.3.1.3. 数据查询 40
6.3.1.4. 数据标准 44
6.3.1.5. 数据建模 46
6.3.1.6. 数据开发 47
6.3.1.7. 智能标签 52
6.3.1.8. 空间服务 57
6.3.1.9. 统一服务 58
6.3.1.10. 通用资源 62
6.3.1.11. 运维监控 63
6.3.1.11.1. 任务中心 63
6.3.1.11.2. 全景监控 67
6.3.1.12. 质量评估 74

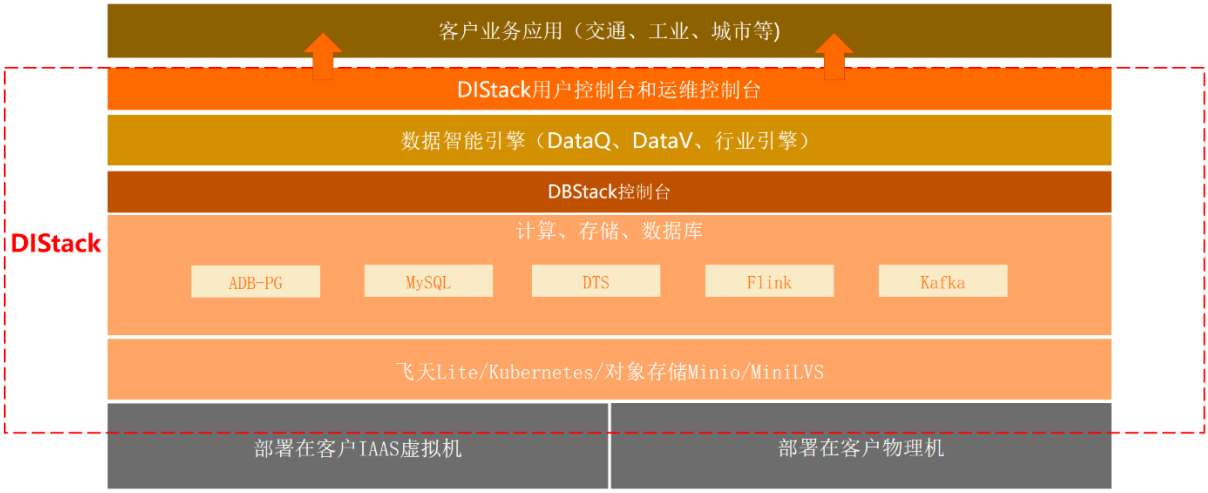
6.3.1.13. 解决方案 76
6.3.2. 资产运营工作台 77
6.3.3. 数据资产中心 81
6.3.4. 画像分析 84
6.3.5. 数据探索 91
6.3.5.1. 首页 91
6.3.5.2. 数据中心 92
6.3.5.3. 算子中心 96
6.3.5.4. 模型中心 97
6.3.5.5. 任务中心 99
6.3.5.6. 用户中心 102
6.3.6. 我的资产 105
6.3.7. 系统设置 107
6.4. 应用场景 110
6.5. 基本概念 112
6.6. 使用限制 115
7.DataV数据可视化 产品简介 118
7.1. 什么是DataV数据可视化 118
7.2. 产品优势 119
7.3. 产品架构 120
7.3.1. 功能架构 120
7.3.2. 系统架构 121
7.4. 功能特性 121
7.5. 应用场景 124
8.交通云控平台产品 127
8.1. 功能特性 128
8.2. 应用场景 129
9.Hdata产品简介 132

9.1. 什么是医疗健康数智平台 132
9.2. 产品优势 133
9.2.1. 核心能力 133
9.2.2. 用户价值 143
9.3. 功能特性 143
9.3.1. 业务工作台 143
9.3.1.1. 数据治理 143
9.3.1.2. 医疗数据资产 146
9.3.1.3. 智能业务能力 146
9.3.2. 研发工作台 153
9.3.2.1. 资产加工 153
9.3.2.1.1. 数据同步 153
9.3.2.1.2. 数据查询 155
9.3.2.1.3. 自助分析 159
9.3.2.1.4. 数据标准 159
9.3.2.1.5. 数据建模 162
9.3.2.1.6. 数据开发 163
9.3.2.1.7. 智能标签 167
9.3.2.1.8. 统一服务 172
9.3.2.1.9. 通用资源 176
9.3.2.2. 运维监控 177
9.3.2.2.1. 任务中心 177
9.3.2.2.2. 全景监控 181
9.3.2.3. 质量评估 188
9.3.2.4. 解决方案 191
9.3.3. 运营工作台 192
9.3.4. 数据资产中心 196
9.3.5. 画像分析 199

9.3.6. 系统设置 206
9.4. 应用场景 211
9.5. 基本概念 211
9.6. 使用限制 214
10.感知数据工厂产品简介 216
10.1. 产品介绍 216
10.2. 基本概念 216
10.3. 应用场景 217
10.4. 功能介绍 219
11.空间治理平台 251
11.1. 什么是空间治理平台 251
11.2. 产品优势 252
11.3. 产品架构 252
11.4. 功能特性 253
11.5. 应用场景 256

1.DIStack整体介绍

DIStack是Apsara Data Intelligence Stack的英文简称，中文名称叫数据智能混合云平台。DIStack是面向垂直行业和边缘数据智能市场的跨产品组合，覆盖离线、实时、在线等常见计算架构（ adb-pg、flink、kafka、spark等），内置丰富智能场景和行业知识（DataQ和云控平台等行业引擎），具有功能完整和体积小巧的产品特性，支持虚拟化和物理机等多种部署方式，具备软件和软硬一体两种产品输出形态，帮助客户快速、安全、经济的实现企业数智化转型。



2.产品优势

面向产业边缘计算数据智能市场，DIStack具备场景化、轻量化和一体化和三大产品优势：



1、场景化

•依托于阿里云城市大脑的经验、算法和算力实践，针通用的数据中台场景，形成数据资产定义、加工、管理、服务能力；

•依托于阿里云城市大脑的经验、算法和算力实践，针对不同细分行业场景（比如交通管理场景），形成模型表达、行业参数、路网编码、评价指标等标准化行业模型；

•具备低成本优势，快速切入垂直行业应用，提供数据+智能的核心竞争力，满足下沉市场更多边缘的数据智能行业应用场景。

2、轻量化

•具有功能完整、体积小巧的计算和存储引擎，覆盖离线、在线、实时多种常见计算架构；
•云平台管控更轻量、更精简，做到0公摊；
•具备软硬件解耦能力，支持在主流虚拟机部署，满足在客户虚拟化云平台轻量部署需求。

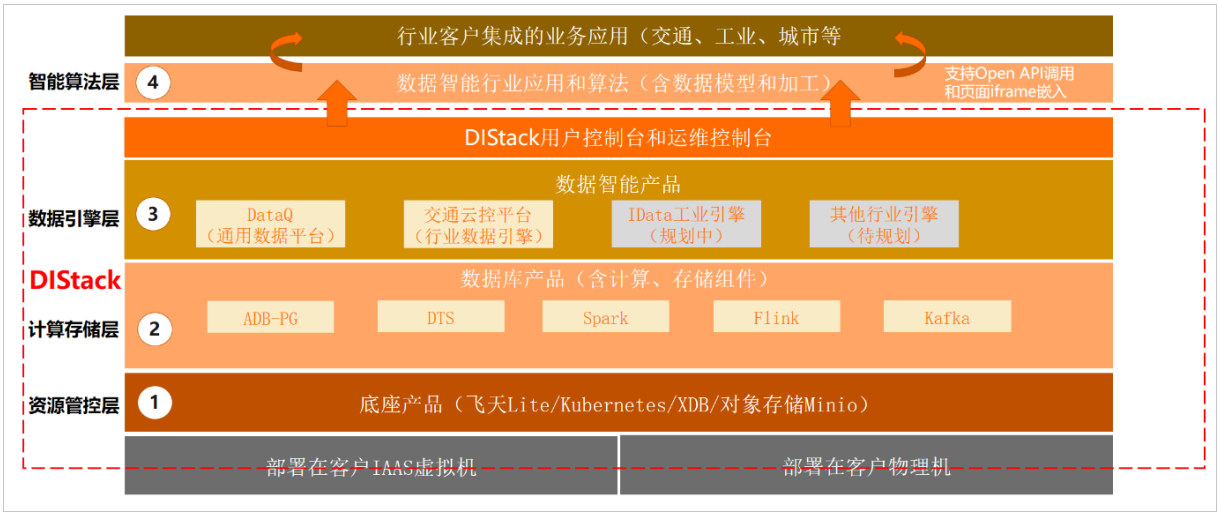
3、一体化

•支持软硬一体化部署，开箱即用；
•稳定的云平台计算和存储引擎，调优的通用和行业数据引擎，加速行业数据智能应用一体化开发；
•内置行业解决方案，贴行业应用一体化输出，支持算法工程师快速定制客户数据需求，提升对客户业务服务能力。

3.产品架构

本文介绍DIStack产品架构。

兼顾平台稳定性和业务敏捷性考虑，DIStack采用产品分层架构设计：
1、资源管控层
底座基于天基和Kubenets云原生技术，支持主流虚拟机，物理机多种资源部署和管理，产品稳定性高；
2、计算存储层
提供核心计算和存储能力（ADB-PG、Spark、Flink、Kafka），能够高效、低成本、稳定支持行业算法场景；
3、数据引擎层
提供稳定和高质量的行业数据智能引擎能力，包括通用的数据资源平台DataQ，具有行业数据特性的交通云控平台等，向数据算法工程师提供完善的数据操作用户控制台，向运维工程师提供整合的云平台运维控制台。
4、智能算法层
贴近行业客户需求，通过算法引擎和数据模型敏捷迭代开发适配，应用数据加工作业等操作，能快速支持行业客户解决方案落地，为客户业务应用提供被集成能力和API，快速实现行业应用智能化。



4.功能特性

本文介绍DIStack功能特性。

1、资源管控层

包括飞天服务、PaaS基础服务、XDB数据库引擎实例。主要功能是最基础的管理职能，如包括Kubernetes的服务、鉴权、基础元数据存储等，支持主流虚拟机，物理机多种资源部署和管理。

一级模块	二级模块	功能描述
账号管理	账号管理	支持创建、删除、锁定账号、修改密码。
	角色管理	支持查看不同的角色类型。
主机管理	主机管理	支持展示主机列表信息，主机详情信息，并支持修改主机信息。
	启停分配实例	支持开启和停止分配实例。
	下线主机	支持手动下线主机。
	主机监控	支持监控在主机上运行的进程，了解每个进程对资源的占用情况。
性能监控	报警规则管理	支持自定义报警规则。
	报警渠道管理	支持通过不同的渠道获取报警信息。
	查询历史报警	支持查询单个或所有报警历史信息。

2、计算存储层

包括核心计算和存储组件：ADB-PG、DTS、Spark、Flink、Kafka，能够高效、低成本、稳定支持行业算法场景。

1）云原生数据仓库AnalyticDB PostgreSQL版

一级模块	二级模块	功能描述
	冗余架构	采用全冗余架构，无单点故障，每个关系型数据库实例均实现主从热备。
	平滑升级扩容能力	支持在线平滑升降级，计算能力、存储容量和总IO带宽同步线性扩容。

一级模块	二级模块	功能描述
	支持虚拟专有网络VPC	支持虚拟专有网络，可支持指定虚拟网络创建数据库实例。
	数据库运维	提供数据库自主诊断、慢查询分析，提供全面的健康状态以便用户自动化运维。
	管理门户和API	提供用户自服务门户和API接口，用户可自行创建不同规格的关系型数据库实例，并提供关系型数据库实例的在线扩容、自定义备份、数据恢复、性能监测分析、异常告警、日志管理等功能。
	统一账号和认证体系	支持对云数据库服务采取统一账户、统一认证，确保用户信息的安全，避免用户面对复杂的账户、认证场景。
	参数设置	支持对数据库的参数进行设置。
	版本升级	支持升级数据库内核小版本。
	实例管理-创建实例	支持创建实例。
	实例管理-释放实例	支持删除实例。
	实例管理-重启实例	支持重启实例。
	实例管理-连接实例	支持通过客户端、驱动、工具连接数据库实例。
	实例管理-变配实例	支持通过增加计算组，进行实例的变配。
	实例管理-实例详情	支持查看实例详情，查看配置、连接地址/端口、节点等信息。
	日志管理-慢查询日志明细	支持查看慢查询日志明细。
	日志管理-慢查询日志搜索	支持查找慢查询日志。
	日志管理-慢查询日志统计与分析	支持慢查询日志统计与分析。

一级模块	二级模块	功能描述
数据库管理	连接实例（DMS）	支持通过DMS链接数据库。
	实例监控	支持实例级、节点级多指标监控指标采集与可视化展示。
	实例报警	支持实例级、节点级报警，预置默认报警规则，可自定义报警规则和报警等级，以及报警渠道。
		支持数据管道服务，把Kafka日志数据同步到AnalyticDB PostgreSQL中，让用户可以轻松构建基于流式数据的分析和应用。
	外部数据管理	

一级模块	二级模块	功能描述
数据导入	支持离线平台的批量导入，数据直接装载到二级分区，提高数据导入效率；支持关系型数据库的导入与实时同步；支持流计算实时入库；同时数据也支持导出至对象存储和离线平台。	数据管理
维度表	支持维度表（又称为复制表）和事实表，事实表支持两级分区策略，且支持按不同关键字段的多级分区。	
存储过程	支持存储过程。	

一级模块	二级模块	功能描述
UDF	支持UDF。	
分布式事务	支持分布式事务。	
数据分区	支持数据分区。	
外部表	支持外部表。	
分布式架构	支持分布式架构。	
线性扩展	支持线性扩展。	
常用分析函数	支持常用分析函数。	
数据导出	支持数据导出。	
物化视图	支持物化视图。	
行存储	支持行存储。	
列存储	支持列存储。	
数据冗余	支持数据冗余。	
数据自动分布	支持数据自动分布。	
分布式join	支持分布式join。	
资源队列	支持资源队列。	
并发控制	支持并发控制。	
表空间管理	支持表空间管理。	

一级模块	二级模块	功能描述
数据库对象DDL/DML操作	支持数据库对象DDL/DML操作。	
账号管理	创建账号（高权限账号、普通账号）	支持创建账号（高权限账号、普通账号）。
	删除账号	支持删除账号。
	重置密码	支持重置密码。
	修改账号权限	支持修改账号权限。
	重置高权限账号	支持重置高权限账号。
	服务授权账号	支持服务授权账号。

2）数据传输服务DTS（含Spark、Flink、Kafka组件）

一级特性	二级特性	特性描述
Kafka	提供Kafka分布式流处理平台	• 支持的版本为1.0 • 支持发布和订阅消息流 • 支持集群部署 • 支持多下游数据消费
Flink	提供分布式流数据流引擎	• 支持的版本为flink-1.12.0 • 支持kafka和adb-pg作为源和目标的流式数据加工功能
Spark	提供数据批处理引擎	• 支持的版本为spark-3.0.3 • 支持MySQL、PolarDB-O、ADB PG作为源和目标的离线数据加工功能
安全	不同账号下的资源隔离	通过独立的进程、文件进行租户的实例和数据隔离。例如，禁止用户读写实例的操作系统文件，确保用户无法接触其他用户的数据。
告警监控	支持配置任务告警	支持配置告警并查看全量任务的监控信息。

3、数据引擎层
提供稳定和高质量的行业数据智能引擎能力，包括通用的数据资源平台DataQ，具有行业数据特性的交通云控平台等。
1）数据资源平台DataQ

数据资源平台-基础组件	系统设置	工作组管理	支持工作组的创建管理、支持工作组下云计算资源、成员的添加、管理。
		用户管理	支持对账号相应的租户管理员、租户访客、租户资产管理员、租户运维等角色进行添加、管理。
		权限列表	支持查看系统中各功能模块中各角色对应的权限列表。
		我的申请	支持申请人查看自己提交的申请单，并提供按审批状态筛选申请信息和撤回申请等功能。
		我的审批	支持审批人查看需要自己审批的申请单，并提供对审批单的查看、通过、拒绝等功能。
		应用管理	支持应用创建、管理功能，支持应用的成员管理功能，提供应用密钥管理功能、Open API 授权等功能。
		组织管理	支持树形组织的创建、管理，支持组织内成员的添加、管理功能。
	解决方案	整体导入导出	支持将工作组内已加工完成的数据标准、数据模型、数据加工场景、数据服务API配置等信息，进行版本化解决方案封装。支持「全量」、「按范围关联」两种方式灵活创建解决方案，并可获取逻辑表的逻辑一致性、物理一致性结果，以进行打包前检查。 支持解决方案包导入并实例化至工作组。
	任务中心	全局概览	支持查看任务流、任务、任务实例的统计信息，并提供以图形化的方式展示历史实例运行信息。
		任务流	支持查看任务流的属性信息，包括任务流中包含的任务节点、任务节点信息、任务节点的最新实例等。 支持查看任务流的部署、运行日志信息。
		实时任务	支持查看实时同步任务的属性与运行日志。
		离线任务	支持查看离线计算任务的属性信息，支持任务节点的补数据操作。 支持查看离线同步实例、离线计算实例的属性与运行日志。

		补数据计划	支持查看补数据计划、补数据计划实例中各节点的实例信息和运行日志。
	通用资源	自定义函数	支持UDF函数管理，并在计算节点中使用，便于快速对数据进行处理。
		函数资源	支持对自定义函数中所使用的代码包的新建和管理。
数据资源平台-数据上云	数据同步	数据源管理	支持对数据同步的数据源端和目标端进行统一的注册和管理，支持关系型数据库、文件、消息队列等类型数据源的配置，并对所填写数据源的连通性性进行有效性验证，支持同步工作组内数据源。
		数据模板管理	支持数据模板创建，添加数据字段及字段类型，并提供编辑、删除等功能。
		离线数据同步	支持单表离线任务、批量离线任务的新建。 支持自定义选择已经注册的数据源端数据和目标端，支持多种常见离线同步链路，并提供全量任务自动建表、同名字段自动映射、任务调度配置功能。
		实时数据同步	支持目标端为Datahub的单表实时任务的新建。支持Orcala CDC、MySQL binlog、SQLserver CDC三种实时数据同步模式。提供完善的实时数据接入配置等功能，有效保障实时任务接入的稳定性和可控性。
	数据探查	物理表预览	支持对多种云计算资源中的物理表进行查询，提供数据字段及数据内容预览，条件筛选功能。
		物理表探查	支持面向多种云计算资源中的表及字段进数据自动探查，提供表级别、字段级别探查结果图表化展示，可为开发人员和数据应用人员提供全面直观的数据情况呈现。（数据资源平台-智能数仓-高级版）
		逻辑表命名规则	支持对逻辑表命名标准的规范设计，可基于逻辑表命名属性，对逻辑表命名规则进行配置，规范统一后续模型设计的命名标准。
		逻辑表命名属性	提供对逻辑表命名组成中的一些常用属性进行自定义配置功能，如刷新频率、增量、统计周期。

数据资源平台-智能数仓	数据标准	数据元管理	提供对数据元的标准定义和统一管理，通过分析业务流程，抽象关键业务对象、业务对象属性，配置数据元的类型、长度、质量校验函数和引用数据字典，为逻辑表标准化设计及数据治理过程中的质量监控提供规则依据。 支持数据元导入、导出，方便行业经验积累和跨项目复制。
		指标管理	支持指标类型数据元的定义和管理，可通过选择指标属性的方式，快速新建指标，供逻辑表设计时进行使用。
		指标属性管理	支持指标定义的相关属性管理，包含时间修饰、业务修饰、度量，帮助快速新建指标数据元。
		维度管理	提供对维度表的统一管理，通过分析业务流程、抽象关键业务对象、业务对象属性，配置维度属性的字段类型、长度、质量校验函数和引用数据字典，为模型标准化设计及数据治理过程中的质量检测提供依据，并提供导入、导出、删除、编辑等功能。
		数据字典管理	支持创建并维护逻辑表中使用的数据字典，方便用户在定义数据元或维度属性时进行数据字典引用。 支持数据字典物理化、数据字典导入、导出，方便数据字典使用与跨项目复制。
		质量校验函数	支持质量检测函数自定义，方便用户在定义数据元或维度属性时进行质量规则引用，以支持后续对字段的质量规则的快速配置和校验。
		业务应用	对逻辑表数据模型支撑的应用系统信息进行创建、管理。
		业务域	对逻辑表数据模型来源于哪些系统信息进行创建、管理。
		数据域	支持对数据域的定义与管理，供逻辑表设计时进行使用。
		表属性	支持对逻辑表的表属性做自定义设计，补充业务需求的表信息描述，以便在模型加工过程中拓展并丰富模型的自定义属性，为数据资产管理提供更高灵活性
		字段标准属性	支持对数据元、指标、维度的属性进行自定义配置和管理，补充业务需求的表信息描述，拓展字段标准的元信息

	数据建模	逻辑表管理	对逻辑表进行层次结构设计、通过新建逻辑表、Excel导入和云计算资源扫描的方式新建逻辑表。 支持对创建的逻辑表进行管理，包括物理化逻辑表、查看逻辑表详情、导入导出逻辑表、复制逻辑表、查看逻辑表一致性等。
		逻辑表定义	支持通过可视化的形式创建模型的基本信息、字段信息、关联信息。 支持DDL语句新建字段、引用现有表字段、引用标准库字段、手动添加字段等方式快速创建逻辑表字段及字段信息。
		质量规则	支持表级、字段级、自定义SQL数据质量规则。支持MaxCompute、Hive、RDS for MySQL等多种存储类型的质量规则。支持根据关联字段标准，自动创建质量规则。（数据资源平台-智能数仓-高级版）
		逻辑表物理化	支持根据逻辑表自动生成并保存不同云计算资源类型的DDL语句。 支持将DDL已存在的逻辑表物理化到指定数据库，并对物理化执行结果信息进行展示
		数据模型设计	支持通过图形化拖拽的方式在画布中建立表之间的关联关系，并支持编辑表关联关系的属性。
		一致性对比	支持物理一致性比对，对比当前逻辑表和已物化物理表的字段的一致性。 支持逻辑一致性比对，对比当前逻辑表与数据开发中引用节点中数据结构字段的一致性。
	数据基础库	数据基础库	支持标准化基础数据的管理维护，以便在模型加工过程中使用行业通用化标准化数据，进行后续数据加工和内容参照。
	质量评估	质量检测计划管理	可批量选择需要周期执行且已配置好质量规则的数据模型，配置周期检测任务和调度，进行持续质量检测。同时支持对计划任务的管理、启动、停止和试跑。（数据资源平台-智能数仓-高级版）
		质量检测任务管理	支持对质量检测任务运行实例进行管理，可查看任务详情和日志，并可将检测任务导出明细。（数据资源平台-智能数仓-高级版）
		数据质量报告	支持以周期性数据质量检测结果为基础，提供从多个维度对数据模型质量进行监控并生成报告。（数据资源平台-智能数仓-高级版）

		物理表监控明细	支持以物理表为维度，查看物理表在数据开发和质量计划检测中最新的数据质量结果和历史质量记录。（数据资源平台-智能数仓-高级版）
数据资源平台-AIMaster	数据开发	公共节点	支持公共节点类目管理。支持批量、流式等多种类型的公共节点。 支持对公共计算节点脚本、公共计算节点输入、公共计算节点输出、公共计算节点参数等进行配置和管理。 支持查看公共节点关联的场景，展示公共节点的使用情况。支持公共节点的版本管理。 （数据资源平台-AIMaster-轻量版不支持自定义算法节点）
		测试场景	支持测试场景类目管理。 支持测试场景引用公共节点，支持在测试场景中创建场景内本地节点。 支持数据处理流程编排（多个计算节点组合成一个任务流），用户可灵活的通过可视化方式自定义任务流。 支持批量、流式等多种类型的节点在一个场景内的混合编排。支持配置批量计算节点的调度属性。支持测试环境中场景的部署、运行和发布，并将测试环境中调试通过的场景发布到生产环境。（数据资源平台-AIMaster-轻量版不支持自定义算法节点）
		线上场景	支持生产环境中场景(线上场景)的运行和下线，对各节点的脚本信息、计算资源、存储资源、调度周期、输入输出等详细信息的查看。（数据资源平台-AIMaster-轻量版不支持自定义算法节点）
数据资源平台-智能标签	智能标签	标签模型	面向业务的数据内容编排，构建数据视图，实现标签库设计。支持业务数据OLT（对象、关系、标签）建模能力，并提供建模画布，支持对象类目管理、标签类目管理。 支持标签的自定义属性设计，实现标签元数据信息拓展。 标签的全生命周期管理，支持标签开发、上线/下线和资产公开/私有状态的管理。
		标签工厂	提供衍生标签加工的功能，支持类SQL语言开发，并提供编辑、删除、查看等功能。支持衍生标签的任务管理，设置标签任务调度周期，查看任务日志，以及任务产生物理表的数据探查。（数据资源平台-智能标签-高级版）
		标签加速	支持将构建的标签模型统一加速至指定的分析型数据库形成大宽表，并进行宽表物理化，优化后续分析场景查询效率。 支持多个加速计划，可选择全量/部分标签，自定义加速至不同云计算资源
		分析引擎	提供对业务对象的标签查询，支持类SQL语言，获得查询结果，记录查询过程。 支持将查询语句封装为API供外部应用调用。支持标签API基础信息设置，API版本管理、列表管理和目录设计。

		标签探查	支持低代码方式查询对象的标签及标签内容数据，支持复合多标签的内容筛选，支持模糊搜索或精确匹配，获得查询结果，记录查询过程。 支持根据结果快速生成可视化图表，支持图表的基础设置。
		标签统计	支持标签生产和使用情况统计，包含标签总数、对象总数、API总数统计 支持按类目的、按对象的标签数统计；支持对象标签的生产和使用统计；提供按不同类别查看标签使用情况。
数据资源平台-统一服务	统一服务	API管理	支持对服务API的生命周期管理，包括新建、编辑、调试、上线、下线、删除、日志查看。 支持数据服务的封装，支持MySQL、PostgreSQL、Analytical DB、OTS等云计算资源上的数据服务封装。 支持三方API的注册、数据上报服务的封装。
		服务报表	支持服务被调用后的报表统计功能。服务调用汇总支持用户按照天、周、月的时间范围下，按照小时、天的统计粒度统计复用调用汇总信息，包括总调用次数、平均响应时间、成功调用次数。 服务调用明细支持单次调用的明细信息查看，包括：调用时间、TraceID、响应状态、响应时间、请求参数、失败原因。
		实时监控	监控最近24小时内服务的实时调用趋势及异常趋势。 支持查看监控时间点的服务调用概览，及服务调用明细。
数据资源平台-全景监控	全景监控	监控概览	监控任务、监控链路、监控业务分组等统计信息概览。
		业务链路	支持以业务数据链路为视角进行监控，可查看当前业务链路中节点的运行情况、告警信息。
		监控明细	支持查看任务节点监控结果，可细节掌握每个监控节点的最新状态、任务日志和历史状态分布情况。 支持查看监控任务节点的最近一次告警信息及历史告警信息，并支持手动设置告警处理暂停时间状态。
		监控管理	支持监控任务管理，针对不同来源的任务进行统一监控任务的配置管理。 支持通过批量创建、手动创建的方式完成监控任务创建。 支持业务链路管理，按照业务视角自定义配置业务监控链路，可快速开启或关闭业务链路。 支持业务链路的导入导出。
		监控设置	支持业务分组管理、告警订阅配置、负责机构、人员管理、监控端管理。 支持通过钉钉机器人、系统内置标准协议的方式订阅告警消息。

数据资源平台-数据DNA	运营工作台	资产注册	支持平台侧需管理的数据资产注册功能，用户可选择需要管理的数据资产来源作为管理输入，支持对MaxCompute/Hive/Datahub数据表资源进行主动订阅；可自动订阅研发工作台的上线的数据表、标签、API等资产的信息。
		资产管理	支持资产对象管理，对待管理的数据表、标签、API资产进行编目分类，并提供查询、公开、隐藏等功能。 支持资产目录管理，提供自定义目录配置功能，通过设置多级目录，以便数据资产进行后续编目。
	数据资产中心	概览	提供数据资产概览信息，以可视化方式了解数据表、标签、API等已接入的数据资产的量级及趋势等指标信息。
		资产目录	支持资产目录列表，列表展示公开的资产目录，支持数据表/标签/API目录展现及关键词搜索。 支持数据表详情，展示数据表资产基本信息、字段信息、数据预览、血缘关系、数据质量等信息。 支持标签详情，对标签基本信息进行展现，并提供申请标签使用权限的功能。 支持API详情，展示API基本信息、接口信息、血缘信息，并提供申请API使用权限的功能。
		我的资产	支持展示自有权限或申请授权成功的数据表、标签、数据服务API资产，并提供资产详情查看等功能。 支持展示有管理权限的数据表、标签、数据服务API数据资产，并提供资产详情查看等功能。 支持对已对其他用户授权的数据表、标签、数据服务API等数据资产进行列表展示。
数据资源平台-画像分析	智能搜索	/	提供对象关系上个体标签、个体临时标签、群体的搜索功能，可快速找到目标个体或群体。 查询结果支持查看群体和个体详情，群体查询结果支持群体分析。
	标签市场	/	支持公共标签浏览，以及权限申请；支持标签管理，如“我可使用的标签”，“我的管理的标签”。支持业务标签创建，通过拖拉拽方式编辑计算逻辑创建枚举型、数值统计型标签，方便业务用户沉淀业务经验。（数据资源平台-智能标签-高级版）
	标签圈群	/	支持自定义标签筛选条件，圈定目标群体功能。支持关联对象的组合筛选。支持筛选条件的交并差计算，进一步精准圈群。群体圈定结果支持重新圈定、发布群体、发布API、可下载到本地Excel、设置圈选结果展示标签、分析等功能。

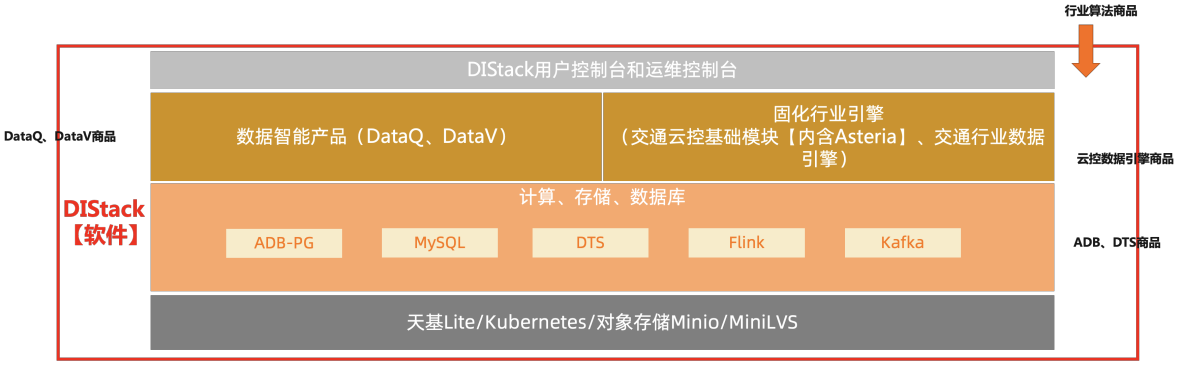
	群体管理	/	支持群体列表管理：支持对群体信息进行新建、删除、更新、状态变更、查看、编辑、查看不同类型群体个数和占比等操作。新建群体支持标签圈群、导入本地文件和群体计算三种方式。 支持群体运维：可发布静态和动态群体，并提供在数据库一次性和周期落表的任务运维管理。 支持群体计算：通过图形化的计算方式可实现两个群体交并差计算得到新的群体，并提供对群体结果支持查看详细群体、发布API、静态、动态和实时群体。
	群体画像	/	提供对群体或个体进行多维度分析，及群体对比分析，个体对比分析。 支持保存分析结果和下载分析报表。

2）交通云控平台数据引擎Asteria（包含全网模型基础算法功能）

一级特性	二级特性	描述
交通云控平台基础组件	开放平台	提供云控用户体系、角色权限、登录控制等功能。
	数据融合引擎	云控平台核心引擎，提供数字化评估、数据接入、数据服务、数据质量监控等功能。
	空间构建	提供交通基础数据的地图展示、编辑、空间语义设置、空间挂载等功能。
	设备管控	提供设备总览、协议市场、边缘节点运维等功能。
交通全网模型-交管	交管数据模型、业务算法	提供交警业务所需数据模型、服务接口和后台处理算法，包含所需地图、路况、轨迹数据。（非标部署）

5.应用场景

DIStack通过使用云原生精简的混合云底座，专属计算节点等手段，集成adb-pg、flink、kafka、spark等各计算引擎，组合DataPaaS和交通、工业等行业产品解决方案，形成功能完整、体积小巧，面向垂直行业客户，边缘数据智能应用市场的跨产品集合，DIStack一期支持数据中台和轻量交管两个主要应用场景。后续，DIStack逐步迁移行业算法，支持城市交通管理、高速运营管理、公共交通运营、海陆空交通枢纽、工业大脑边缘云等行业应用场景。



一、通用的数据中台场景

快速构建数据仓库，有效治理数据质量，实现政企各部门数据的业务协同和共享。DIStack一期全部支持以下核心场景：

- 数据同步：通过数据同步功能，将不同业务系统数据汇聚到统一的存储计算引擎，实现数据的初步融合；
- 数据标准：管理数据标准和构建数据模型，将数据标准贯彻到数据质量分析、保障及检查的全过程中，将散乱的多源异构数据加工成标准、干净的数据资产，确保数据的完整性、一致性、准确性和可用性；
- 数据建模：通过数据建模模块提供的各种数据开发工具，实现数据的清洗、加工和转换；
- 资产管理：通过资产运营功能实现资产的注册、编目、上架，在资产中心方便各个部门的使用人员搜索数据资源并申请数据权限。

二、不同细分行业场景

DIStack一期默认支持轻量级交管暨路网数字还原场景，后续不同细分场景通过在DIStack基础上，逐步迁移行业算法和应用来支持。

1、城市交通管理

- 路网数字还原：标精/高精矢量地图路网还原，提供路档查询和编辑工具，交通运行事件融合精细认知。【一期已支持】
- 车路协同控制：提供全局协同信号优化、可变车道控制等路侧控制方案，以及车内红绿灯、行人/事件预警和车速诱导等车载控制方案。【后续待支持】
- 出行诱导服务：通过诱导屏/情报板等路侧设备发布交通事件、图/文诱导路径和建议车速。同时，车内互联网导航系统可上报交通事件、拥堵路段信息。【后续待支持】
- 重点车辆管理：可以结合信用评估，驾驶人安全码管理两客一危重点车辆，路径规划及违法实时预警。【后续待支持】

2、高速运营管理【后续待支持】

- 路网数字还原：百米精度路网还原，动态信息还原，交通气象预警，三急一速隐患点分析。
- 车路协同控制：10+车路协同场景库，路侧控制和车载控制联动，提升通行效率，降低事故率。
- 应急事件推演：评估事件影响，推荐交通管控策略，通过出行平台发布，并对策略效果仿真评估。

- 精准在线计费：实时还原车辆路径，收费稽核。

3、公共交通运营【后续待支持】

- 全量出行还原：职住通勤客流还原、出行方式分析、景区客流分析预测。
- 公交运行评价：线网评价（线网密度/站点覆盖率/可达性）、线路评价（满载率/运行速度/可靠性/首发站准点率）、站点评价（到站时间预测/到站准点率）。
- 接驳优化推荐：动态公交调度、公交线网优化、公交排班优化、地铁运营与检修排班。
- MaaS出行服务：个性化联程方式出行规划、一码通乘服务、出行核验、用户运营管理。

4、海陆空交通枢纽【后续待支持】

- 场面数字还原：飞机动线轨迹还原、保障节点自动识别、车辆轨迹还原、路段巡检监测、站台作业识别、船号船型识别、船舶轨迹还原、航道密度识别。
- 安全事件感知：红线区域清空与锥桶摆放识别、飞机泊位识别预警、机场围界入侵检测、人员着装识别、车辆越界与停止线识别、施工辅助盯防、站台安全监控、海上动态感知。
- 运行时刻推测：飞机滑行路径规划、可变滑行时间预测、保障节点时间预测、施工持续时间预测、站台安全风险预测、船舶航行/集卡路径规划、桥吊作业时间预测、龙门吊作业事件预测。
- 运力资源调度：机位智能分配、人员智能排班、车辆智能调度、船舶智能避碰、泊位智能分配、集卡车辆智能调度、龙门吊桥吊智能调度。

6.数据资源平台DataQ产品简介

6.1. 什么是数据资源平台

阿里云数据资源平台是数据资产定义、加工、管理、服务的全流程平台，提供数据同步、数据查询、数据标准、数据建模、数据加工、质量评估、标签构建、业务模型构建、资产管理、数据服务、画像分析等功能，为智能数据应用持续稳定供给全量、标准、干净、智能的数据资源。

数据资源平台需覆盖数据同步、数据标准建模及数据质量检查、数据开发、数据标签体系构建，基于标签数据的群体分析、专家业务模型构建、全流程任务监控告警、数据服务化、数据资产管理等核心能力，提供标准化程度高、易用性强的一站式大数据管理平台。

依托数据资源平台，可设计高质量的标准化数据模型，减少重复开发工作，用户可全面了解数据质量、数据使用情况和系统运行情况，并从业务视角更直观地使用并探索数据，更高效地从数据中获取业务价值。

目前数据资源平台提供以下功能：



- **研发工作台**
  - **数据同步**：数据同步可实现离线、实时多源异构数据的便捷同步或接入，系统可提供完善的数据接入配置、 数据模板配置、数据同步任务运行监控等功能，有效保障数据接入的稳定性和可控性，满足各类平台、数据源及应用系统间的数据汇聚需求。
  - **空间数管**：空间数据管理支持单文件、多文件、数据库等多种形态，矢量、栅格、瓦片、倾斜摄影等多种类型，不同格式的多源异构空间数据的接入和管理；支持本地上传、从OSS导入等多种方式添加数据；添加数据的同时支持各类型空间数据的元信息自动解析，支持按空间对平台所有空间数据进行统一检索、统计和管理。
  - **数据探查**：提供对云计算资源中物理表的探查，快速了解物理表详情及分布情况。支持面向多种云计算资源中的表及字段进行预览和表数据自动探查，提供表级别、字段级别探查结果图表化展示。
  - **数据标准**：提供逻辑表标准、字段标准（数据元、指标、维度、数据字典等）管理能力。帮助用户通过设计标准数据元素，定义关键业务对象、业务对象属性及值域定义，并规范标准数据字典，制定并管理平台遵循的统一数据标准，帮助平台管理者和数据管理者管控治理后数据的一致性和数据质量。
  - **数据建模**：帮助用户在数据标准的约束下构建数据模型，将数据标准贯彻到数据质量分析、保障及检查的全过程中，将散乱的多源异构数据加工成标准、干净的数据资产，确保数据的完整性、一致性、准确性、可用性，通过客观量化评估指标帮助客户了解数据治理工作进程，指引数据治理工作的螺旋式上升过程。为了完成在云上积累可运营数据资源的目标，提供数据充分融合、数据高质量可用的必要保障。

- **数据开发**：支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理，同时提供数据加工、算法服务任务流程开发、编排与调试、上线、部署、维护等功能。支持流式计算、批量离线计算等类型的计算节点在一个工作流中统一编排，通过可视化操作界面，通过拖拉拽的方式连接计算节点迅速实现数据加工流程编辑。支持通过空间算子编辑器零代码实现对空间数据的处理，包括地址空间化（依赖达摩院地址标准化产品）、坐标转换、矢量分析和计算、栅格分析和计算等。空间计算节点可作为节点任务统一编排到工作流，实现空间数据和非空间数据的融合治理及治理链路固化。
- **智能标签**：可将治理后的数据以业务化视角进行建模、查看、管理及使用，并提供业务衍生标签的自定义功能，为上层应用提供统一的标签数据目录和标签调用接口，沉淀上层应用制作的模型标签，实现高价值标签共享复用，形成标签运营生态。
- **空间服务**：支持将空间数管中的各类型空间数据发布成行业标准的空间服务，支持发布OGC WMTS、WMS、WFS服务；TMS（MVT）矢量瓦片服务；S3M、3DTiles等标准的三维服务。同时支持通过SLD文件配置地图服务的样式风格。
- **统一服务**：是一款数据中台建设过程中的数据服务化组件，面向数据开发者提供覆盖各个加工阶段统一体验的、便捷的数据查询转服务、服务管理、服务运维能力；面向数据资产管理者提供服务的统计分析、服务用量统计分析能力，实现数据中台建设后半场“数据应用”的有效落地，支撑数据智能应用的高效开发。
- **运维监控**：对数据资源平台中的数据同步、数据开发、标签加工等任务进行运维管理。按业务链路配置监控场景，支持自定义监控链路和告警消息配置，可快速实现从数据生产到业务应用的全链路编排及溯源。
- **质量评估**：对数据模型中建立的逻辑表进行质量规则配置和管理，支持根据配置的质量规则在数据开发环节自动生成质量检测节点和设置自定义质量检查计划。提供面向不同计算资源多种类型质量规则，可通过数据质量监控报告展现系统整体数据质量概览，和多维度细分数据的质量情况。
- **解决方案**：对数据资源平台中已沉淀的数据标准、数据模型、数据加工场景、数据服务API配置、云计算资源配置等数据资产进行导入导出，便于在不同环境中快速搭建数据治理工作的初始化。

- **资产运营**
  - **资产注册**：支持平台侧需管理的数据表资产自动化注册功能，可对多个数据源端注册信息进行展示，并通过手动刷新方式同步元数据；同时，可自动上报数据表、标签和API资产的信息。
  - **资产目录**：支持按资产类型分别维护资产，帮助资产目录内容提供者对各类资产进行管理维护，并针对资产进行编目分类及发布上下线管控。
- **数据资产中心**
  - **资产概览**：提供统一的数据资源视图，以可视化的方式展示数据表、标签、API和指标等数据资产的总量与增量的信息。
  - **资产目录**：对已接入上线的数据资产，可在资源目录对资产进行搜索、目录列表查看及资产详情信息展示。
- **画像分析**：可基于标签的多维度分析，按照业务需求圈定特定群体范围，实现对复杂群体数据的智能化统计、筛选、加工、沉淀，建立单实体360度全方位个性化档案，通过关系图谱展示、可视化线索分析、相似人员智能化推荐等功能实现用户对于目标群体的精准定位。
- **数据探索**：数据探索面向行业客户/业务人员，提供工具内容一体化的业务模型构建平台，实现低代码、可视化构建全场景专家模型，重塑大数据服务创新模式。平台通过将业务数据沉淀为智能数据或智能算子，以可视化拖拉拽和简单图形化条件设定进行模型编排，支持离线、在线、实时全场景，采用简化建模过程、提高模型运行效能、融合智能化算法等方式帮助用户将数据与业务结合起来，不断积累和沉淀专家业务模型，服务于日常事件挖掘、实时预警事件、在线风险识别。
- **我的资产**：支持我可使用的、我可管理的、我已授权的数据资产及资产信息进行展示、管理。
- **系统设置**：数据资产平台的基础功能，包含账号管理、系统设置。该功能帮助用户快速、便捷完成用户角色及相应权限等基础配置，并实现必要的系统管理与控制，保障模块正常运转。

6.2. 产品优势

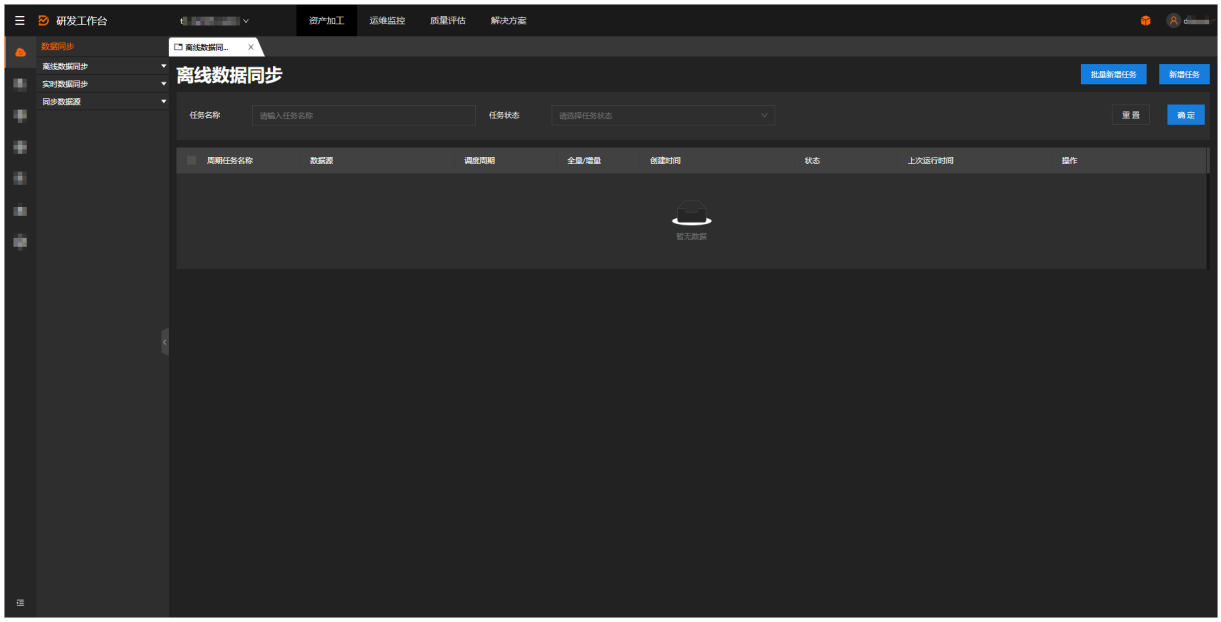
6.2.1. 核心能力

数据资源平台为用户提供了一站式数据资产定义、生产、管理与服务平台，提供企业级数据资产构建能力和一致性使用体验，助力客户快速构建数据智能平台，实现数据资源统一管理，挖掘潜在规律，优化业务决策，让大数据真正的驱动客户业务。 本文主要介绍数据资源平台的核心能力。

一站式大数据开发平台，提供一致性开发体验

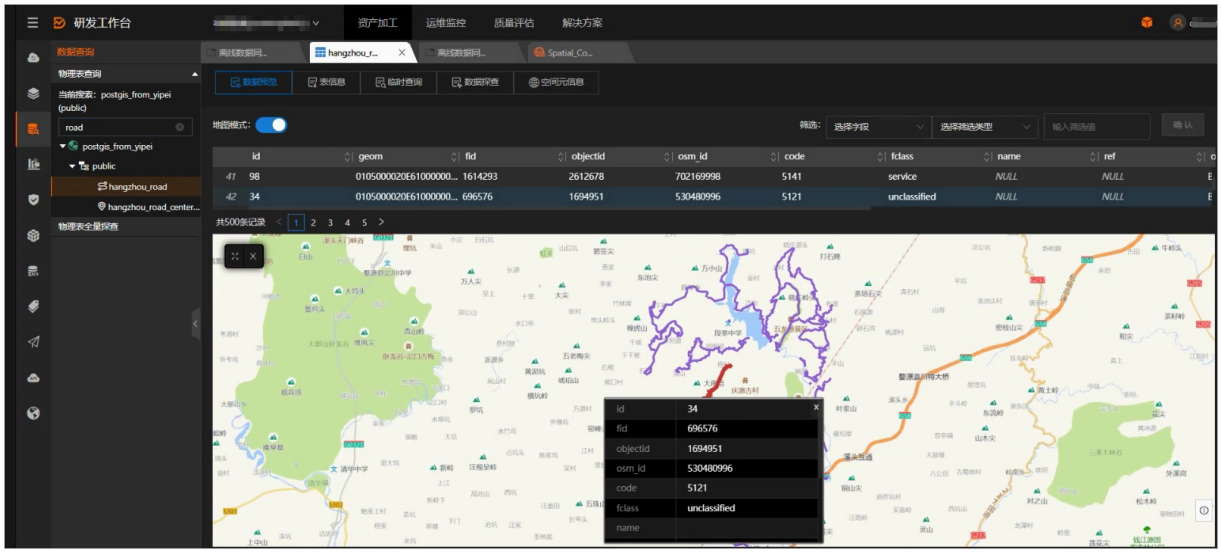
一站式数据同步

- 支持离线/实时数据传输。
- 周期同步任务调度统一管理。
- 主流空间数据库全面支持同步汇聚（超图SDX+，ArcGIS SDE，PostGIS，Ganos）。



自动灵活的数据探查

- 物理表预览和表数据自动探查。
- 多种探查规则和算法，可快速实现表级别、字段级别探查结果图表化展示。
- 支持空间数据的地图可视化查看。



强大的数据加工能力

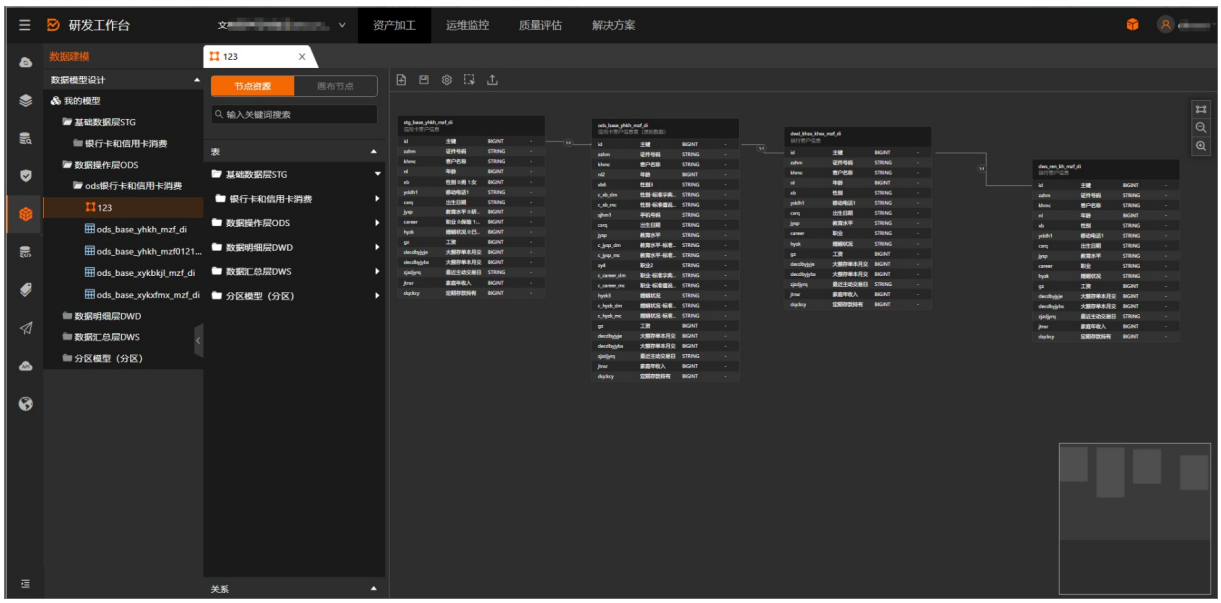
- 离线、实时、在线多类型计算节点。
- 不同类型计算节点在同一个任务流中混合编排调度。
- 支持空间算子编辑器，零代码针对空间数据进行计算、处理和分析。



遵循数据标准的数据建模，实现数据质量闭环控制

可视化、符合标准的数据建模

- 可视化ER模型、维度模型设计，自动生成物理表DDL。
- 数据标准贯彻到模型、质量、保障及检查的全过程。
- 自定义表属性、自定义数据元属性等元模型设计。



全面、准确的数据质量评估报告

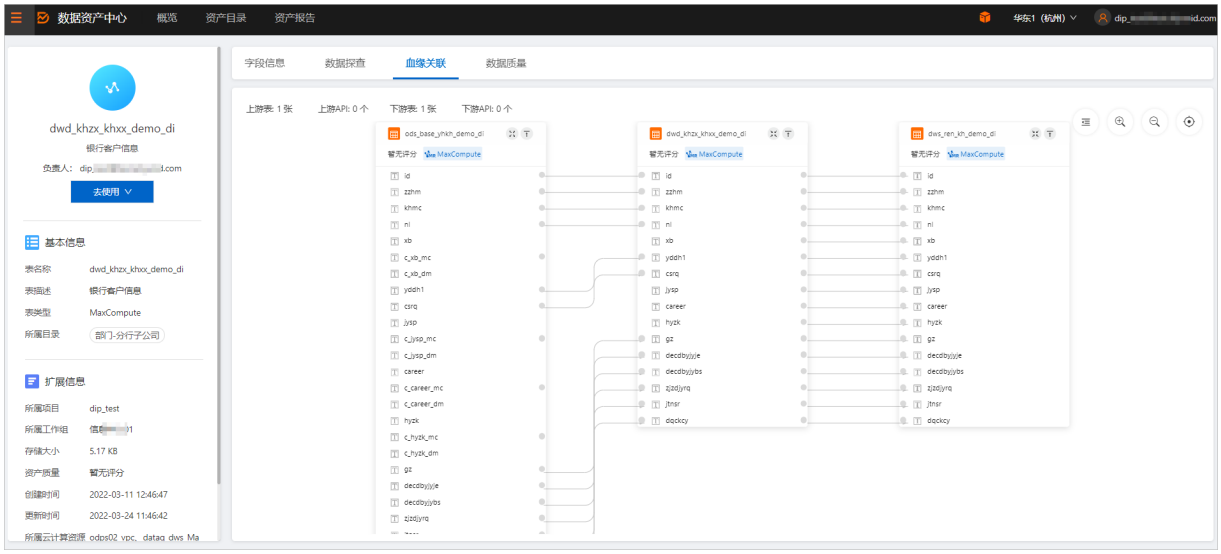
- 多种类型数据源自定义表及字段质量规则配置。
- 支持自定义SQL质量规则。
- 多维度数据质量评估，可视化质量评估报告。



数据资产的管理、服务与敏捷交付，使用数据资产更便捷

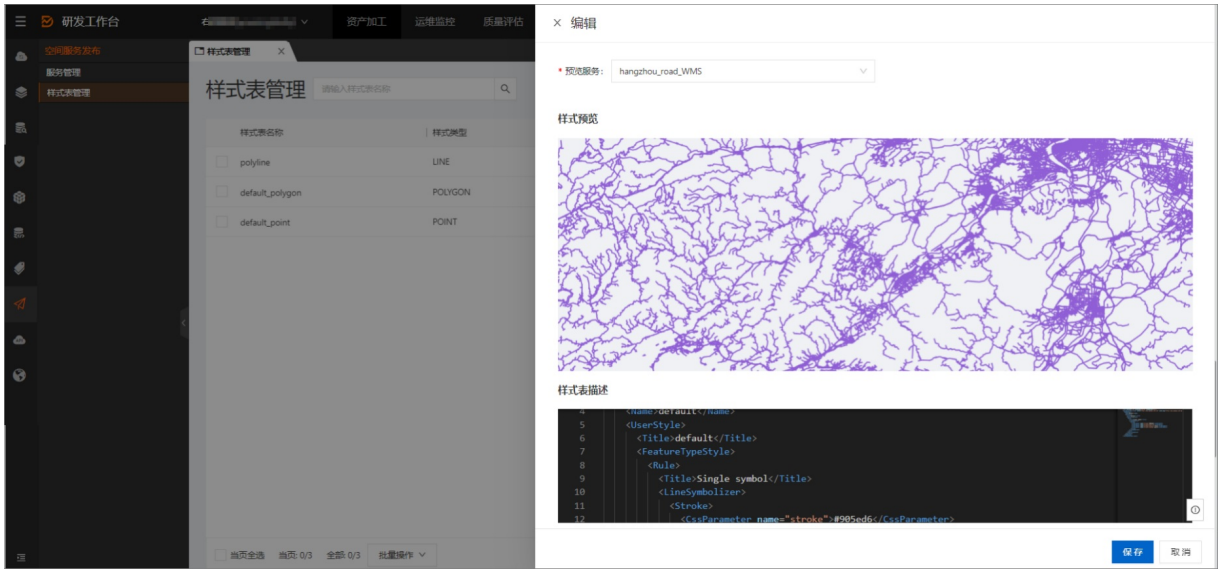
数据资产智能管理

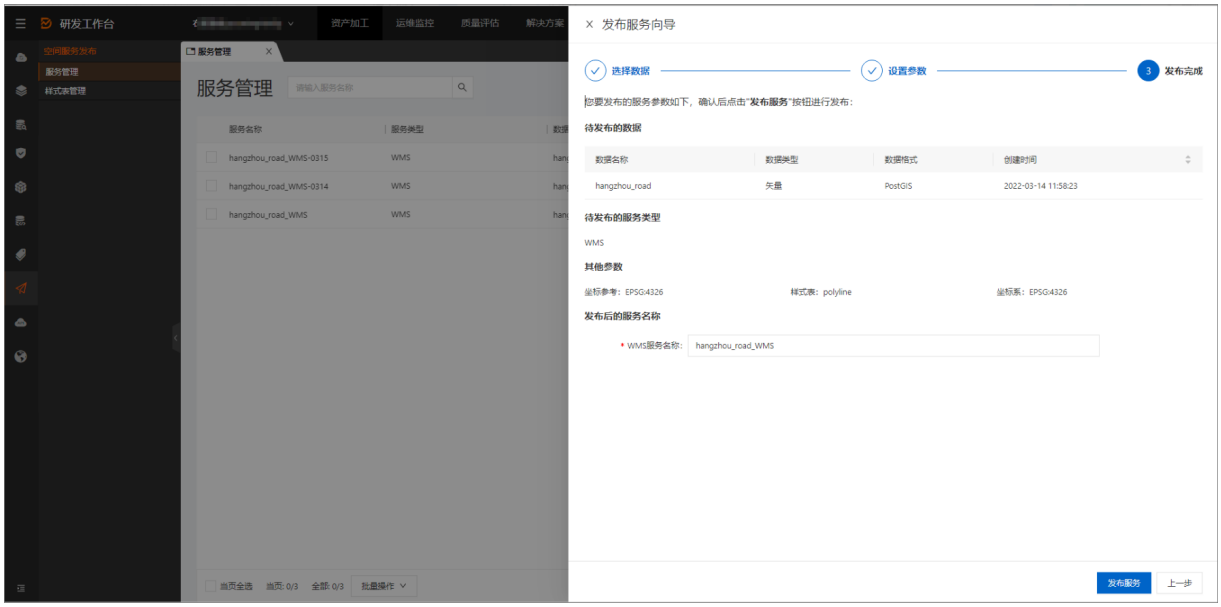
- 表资产、标签资产、API服务资产和指标资产的目录搜索和详情展现。
- 表资产、标签资产、API资产的使用申请和审批管理。
- 字段粒度的端到端数据血缘。



标准空间服务发布

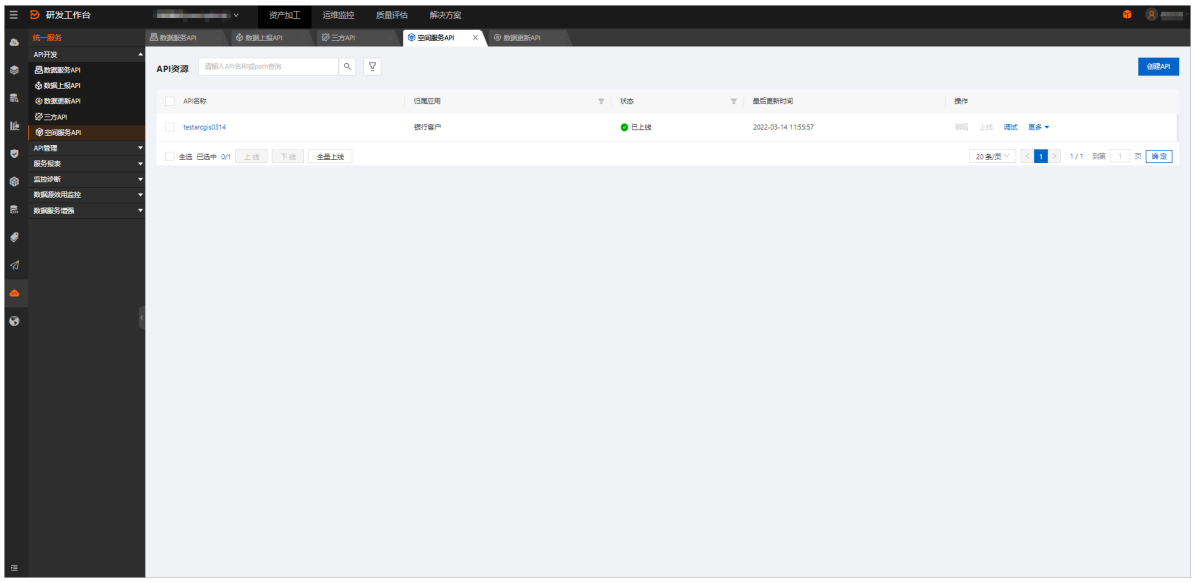
- 支持OGC WMTS、WMS、WFS标准的瓦片服务、地图服务、要素服务的发布。
- 支持TMS（MVT）矢量瓦片服务发布。
- 支持S3M、3DTiles标准的三维服务发布。
- 支持SLD标准的地图服务样式风格配置。





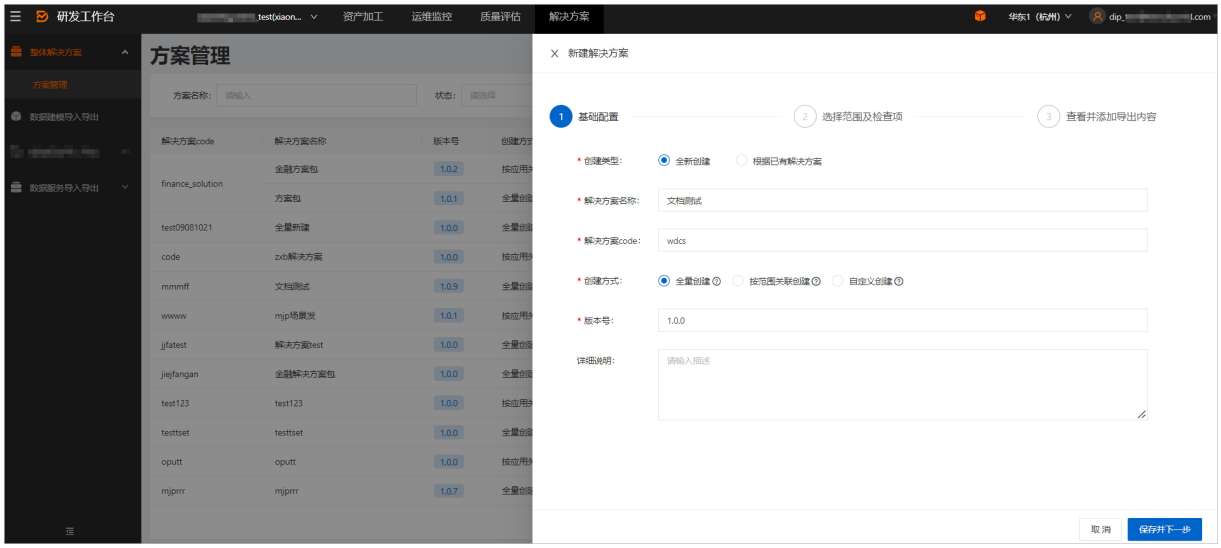
统一数据服务封装

- 实时、离线、半结构化、非结构化、消息队列的数据服务封装。
- 支持向导模式和SQL模式等的服务封装。
- 简单身份认证和加密签名身份认证模式。
- 数据服务调用流量管控。



行业知识敏捷交付

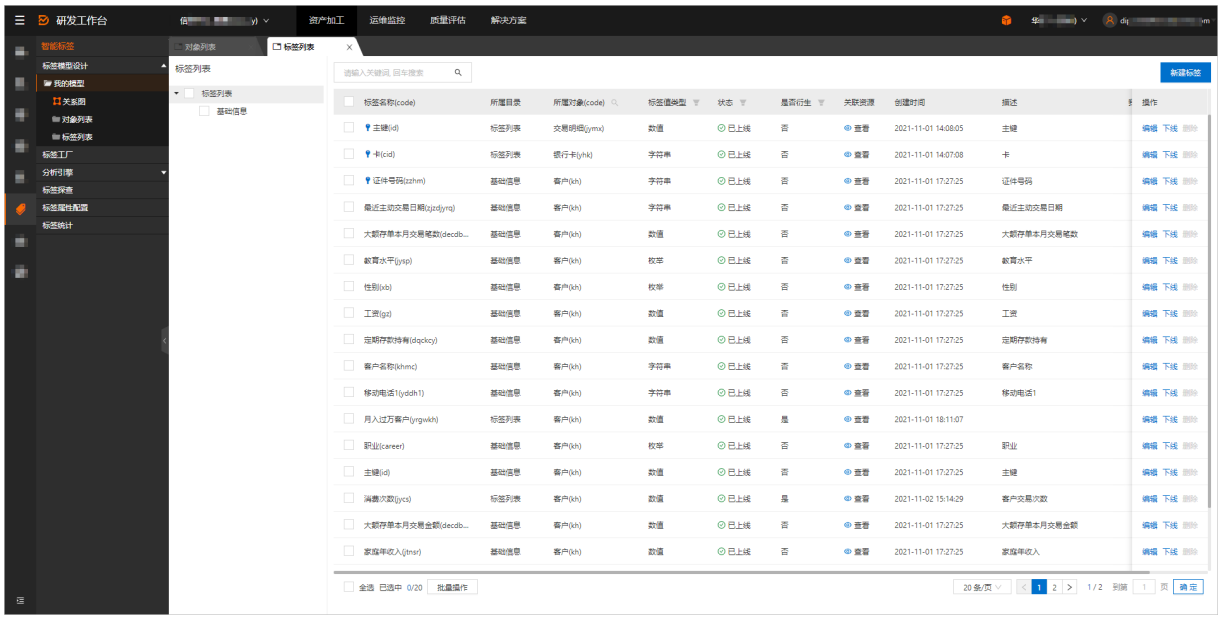
- 工作组下API服务、模型、节点、场景等全类型资产构建解决方案。
- 通过全局数据血缘关联资产灵活构建解决方案。



数据标签化，基于标签的可视化数据分析

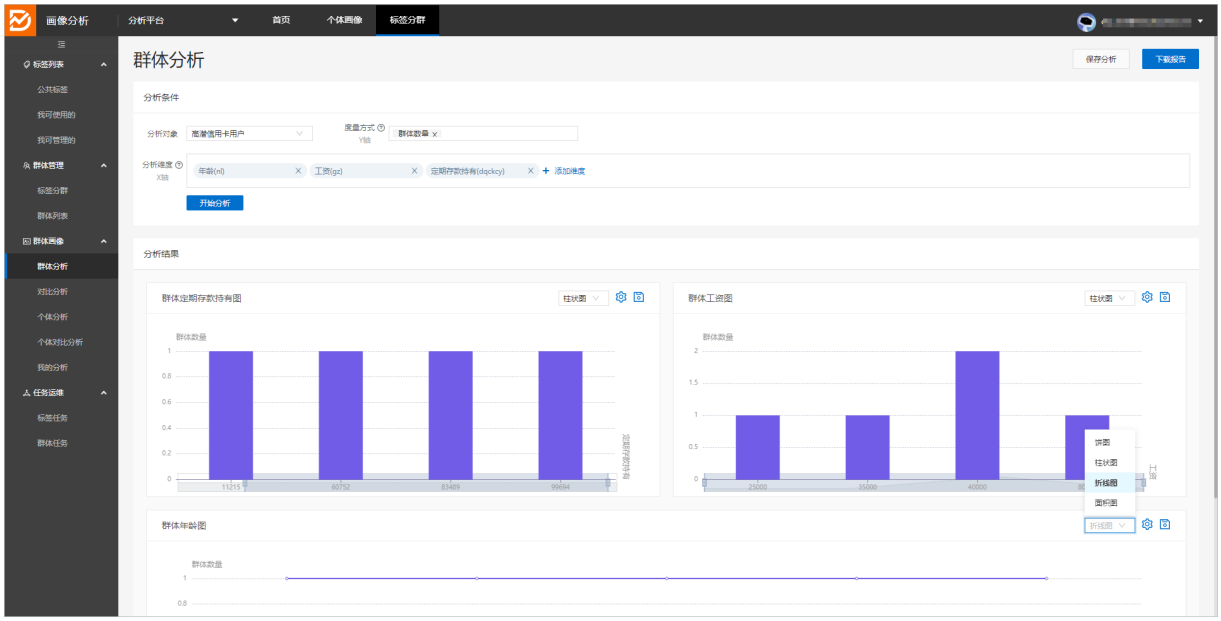
面向业务场景的标签体系构建

- 提供低门槛“宽表”构建能力，帮助上层应用构建业务模型，输出标签库和标签服务。
- 跨存储类型的数据源整合，降低数据同步运维成本。
- 零脚本积累标签画像资产，将数据开发与高频业务运营实现解耦，提高运营效率。



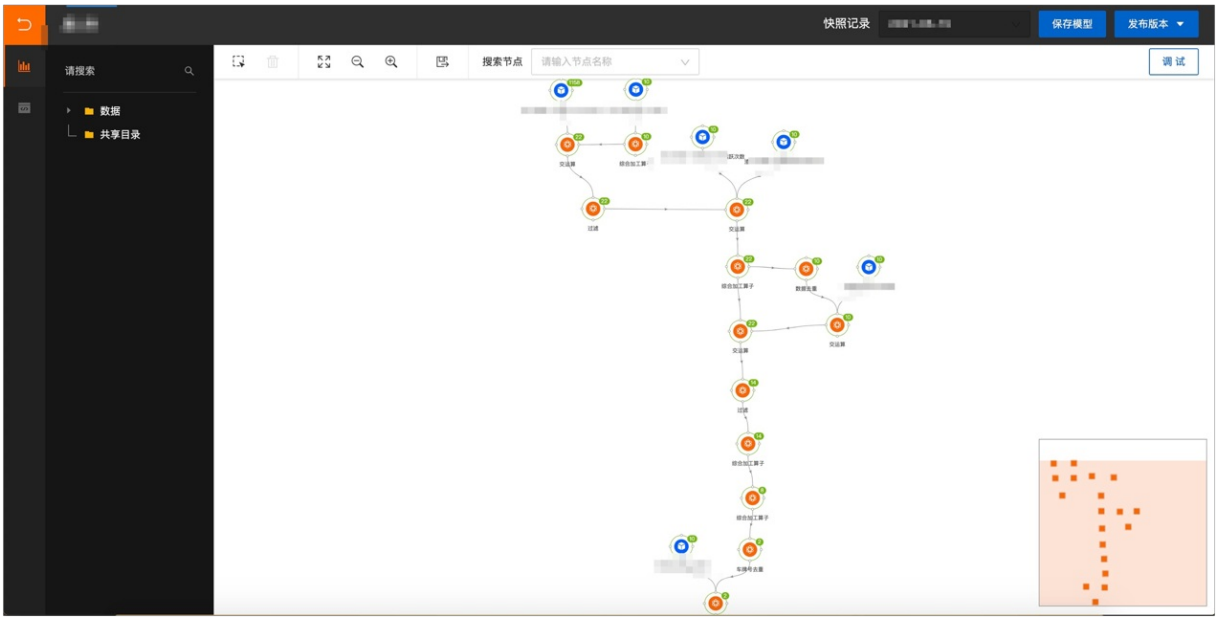
精准运营神器画像分析

- 客群画像分析，通过丰富标签，深入用户洞察，圈定目标群体，并实现精准画像的大数据分析。
- 灵活配置交互式界面：支持多维透视、群体计算等热门分析模型。
- 可沉淀的数据接口，快速集成标签数据、画像报告至各业务系统。



面向业务人员的低代码业务模型构建

- 实现在线、离线、实时模型全计算场景，提供模型编排、发布、服务实例的全生命周期管理能力。
- 实现跨行业通用层逻辑抽象，减少接入新行业/新场景的成本开销。
- 实现工具+内容的业务价值，沉淀行业算子/模型内容库，具备基础算子跨行业输出、业务模型跨行业交付的能力。



6.2.2. 用户价值

数据资源平台为用户提供了一站式数据资产定义、生产、管理与服务平台，提供企业级数据资产构建能力和一致性使用体验，助力客户快速构建数据智能平台，实现数据资源统一管理，挖掘潜在规律，优化业务决策，让大数据真正的驱动客户业务。 本文主要介绍数据资源平台的用户价值。



- **数据开发更简单**：平台提供的各种工具产品能够极大的简化数据开发过程，缩短数据治理周期，降低数据治理成本。通过标准化、精细化、规格化的智能数据生产流程，完成流水线式的数据生产作业，提升数据资源生产效率、消除数据供应品质差异。
- **数据服务更便捷**：能够赋予数据以业务价值，让各级用户能够直观的理解数据，并以此为基础向应用输出多样、便捷的数据服务，不断提升数据面向业务的价值。
- **数据应用更智能**：提供面向业务人员的无代码业务模型构建能力和数据分析能力，大大降低数据获取和分析门槛，让业务人员可以直接使用数据，积累沉淀业务模型，能够向上层应用提供更加智能的数据。
- **数据资产更清晰**：从宏观到微观助力数据管理方全面盘点数据资产，通过智能化全链资产血缘分析，理清战略数据资源，做到让管理者心中有数。
- **数据运营更高效**：遵循应用先行、以用带存、由存而通、因通促用的理念，实现城市数据运营，驱动客户业务创新。

6.3. 功能特性

6.3.1. 研发工作台

6.3.1.1. 数据同步

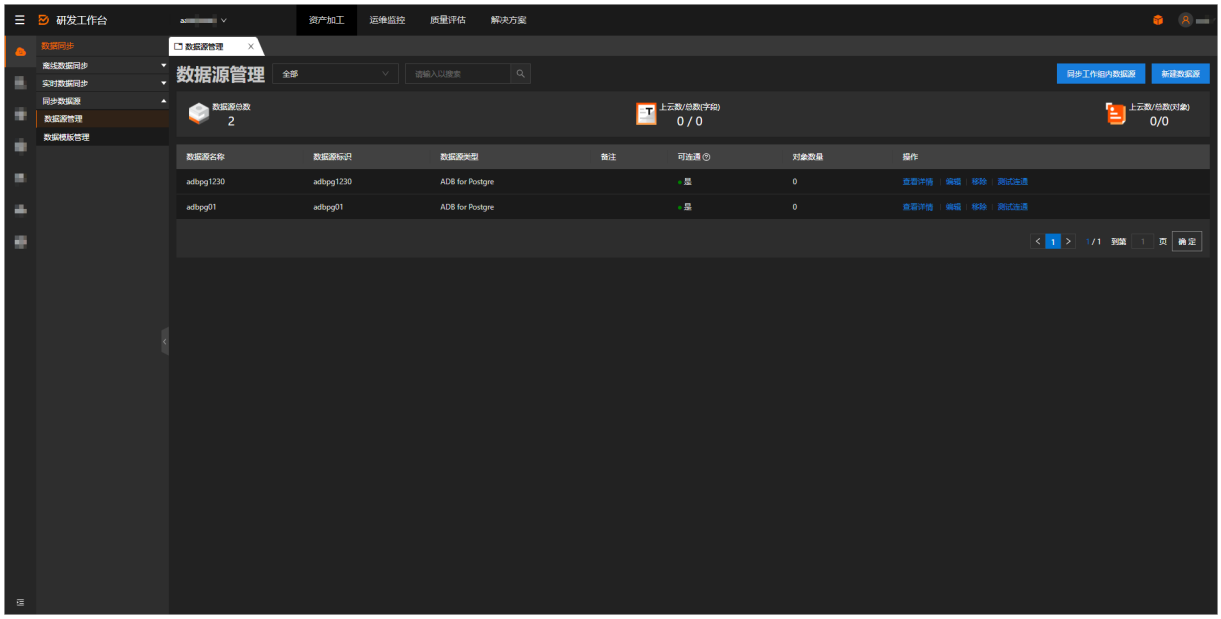
数据同步提供批量、实时多源异构数据的便捷同步或接入能力，系统提供完善的数据接入配置、任务运行监控等功能，有效保障数据接入的稳定性和可控性，满足各类平台、数据源及应用系统间的数据汇聚需求。同时支持对存储在基于ArcGIS、超图等主流GIS平台、开源PostGIS空间数据库中的各类空间数据进行同步汇聚。

通过数据同步的自定义同步任务配置能力，可实现快速创建周期数据同步任务和实时同步任务。用户可选择已经注册的数据源端数据，并选定目标端创建同步任务。系统提供目标端自动建表、源端目标端字段自动映射、批量创建同步任务和引用数据模块进行同步任务配置等多种便捷同步任务创建能力，帮助用户更便捷的完成数据同步配置。

在完成数据同步配置后，数据同步模块可对数据同步任务的执行和运行情况进行全面的监控运维。可对任务总数、线上任务、正运行任务等核心指标进行统计，并可以实现异常任务执行状态的监控，当前实例状态监控。对任务的开始时间、结束时间、耗时、速率、状态和执行日志等信息进行跟踪和展示，并支持周期任务重跑。

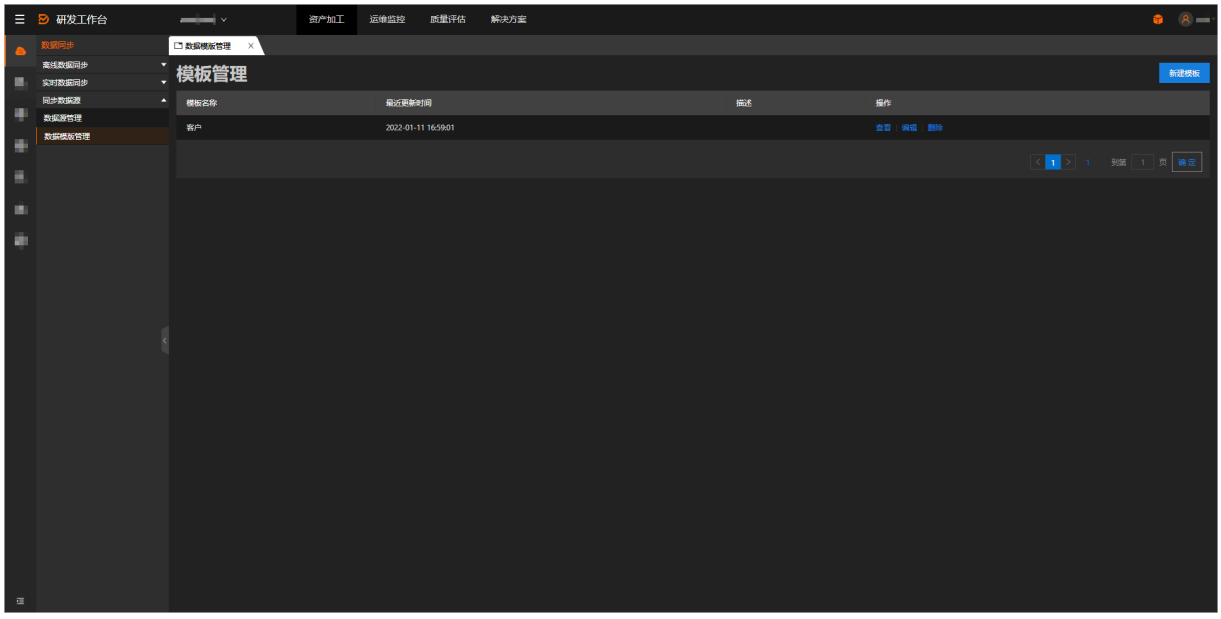
数据源管理

- 支持对数据同步的数据源端和目标端进行统一的注册和管理，支持对关系型数据库、文件、消息队列等多种类型数据源的配置，并可对所填写数据源的连通性进行有效性验证。支持对已配置联通的数据源进行元数据同步及数据对象查看，帮助用户了解元数据资产内容。
- 支持一键同步工作组内已配置的数据资源，快速对工作组已有资源在数据同步任务中进行应用。



数据模板管理

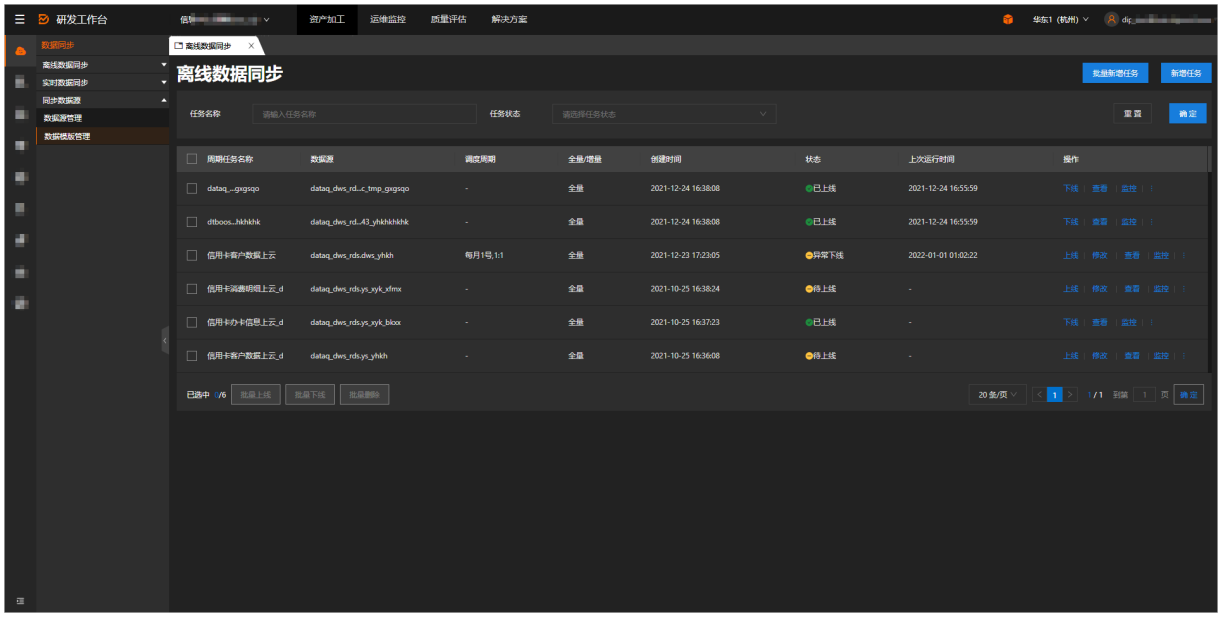
- 针对半结构化、非结构化数据，支持创建数据模板，支持自定义数据字段及字段类型，并提供编辑、删除等功能。
- 支持在离线数据同步、实时数据同步任务中引用已添加的数据模板，根据数据模板定义的数据结构执行数据同步任务。



离线数据同步

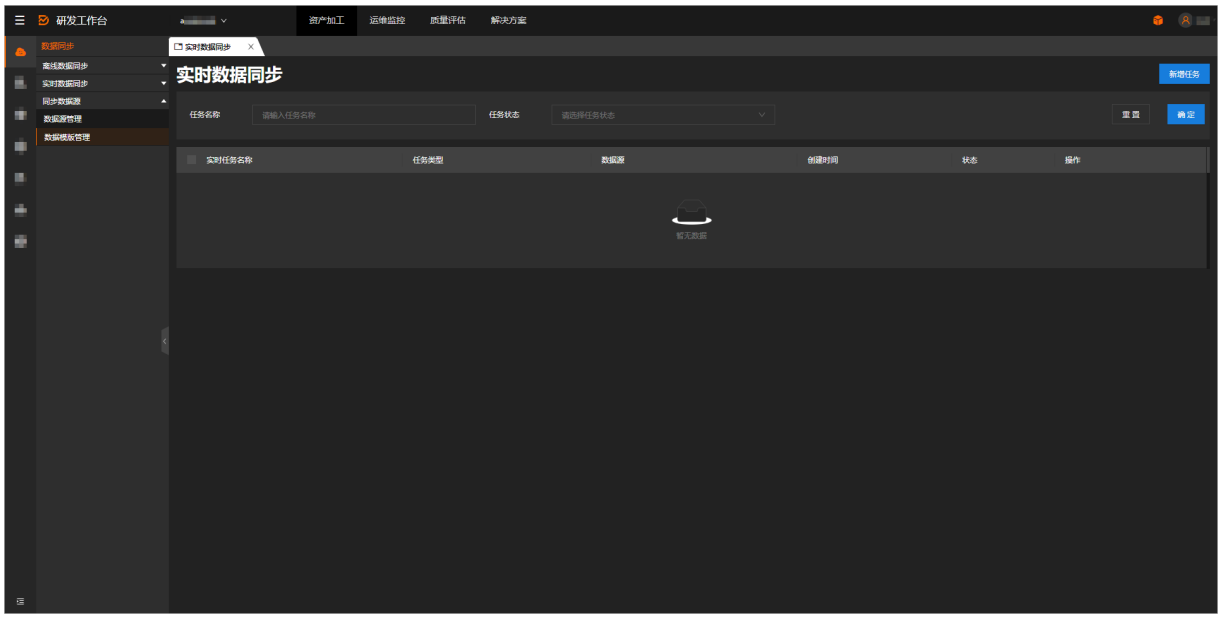
- 支持单表离线任务、批量离线任务的新建。

- 支持自定义选择已经注册的数据源端数据和目标端，支持多种常见离线同步链路，如PostGIS、Ganos、超图SDX和ArcGIS SDE的空间数据同步到PostGIS和Ganos中。
- 提供全量任务自动建表、同名字段自动映射、任务调度配置功能。
- 支持对离线任务实例的运行属性和运行日志进行监看。
- 支持对非增量离线同步实例进行实例重跑。



实时数据同步

- 支持目标端为DataHub的单表实时任务的新建。
- 支持Oracle CDC、MySQL binlog、SQL Server CDC、RabbitMQ、Kafka五种实时数据同步模式。
- 提供完善的实时数据接入配置等功能，有效保障实时任务接入的稳定性和可控性。

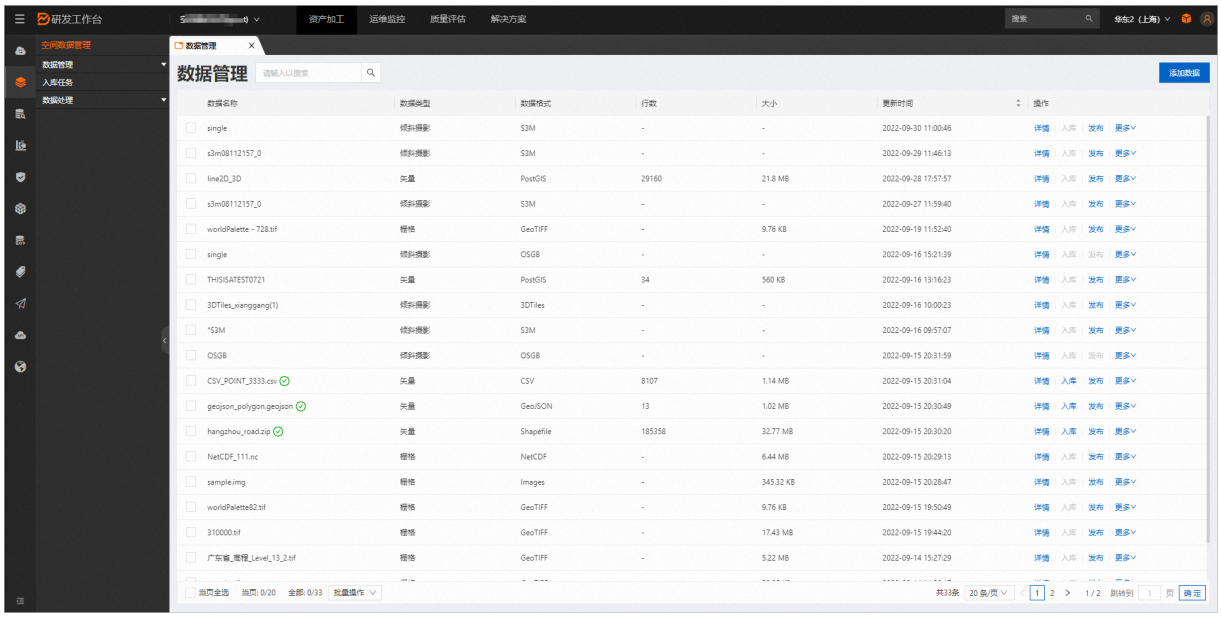


6.3.1.2. 空间数据管理

空间数据管理支持单文件、多文件、数据库等多种形态，矢量、栅格、倾斜摄影等多种类型，不同格式的多源异构空间数据的接入和管理；支持本地上传、从OSS导入等多种方式添加数据；添加数据的同时支持各类型空间数据的元信息自动解析，支持按空间对平台所有空间数据进行统一检索、统计和管理。

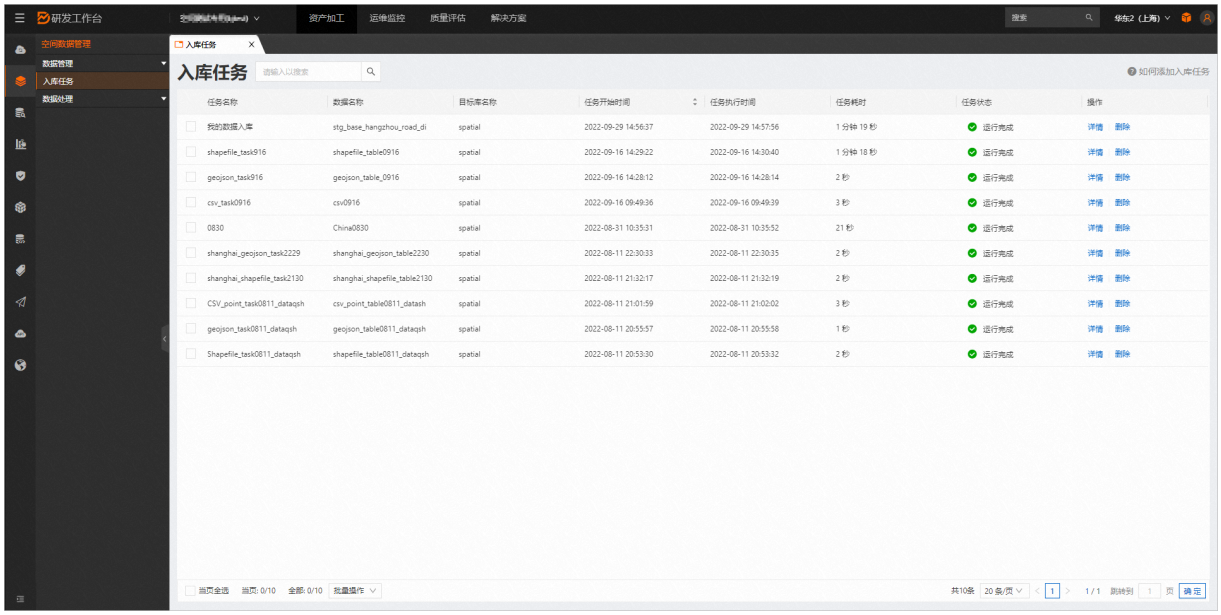
数据管理

- 支持添加单文件、多文件、数据库等多种形态，矢量、栅格、倾斜摄影等多种类型数据，同时可通过上传文件和基于OSS共享添加空间数据。
- 支持对添加后的数据列表批量删除、数据查询和服务发布。
- 支持查看基本信息和空间信息，包括几何形状、单位、要素个数以及坐标参考、空间范围等信息，方便对空间数据库进行空间检索、数据质检、血缘记录。



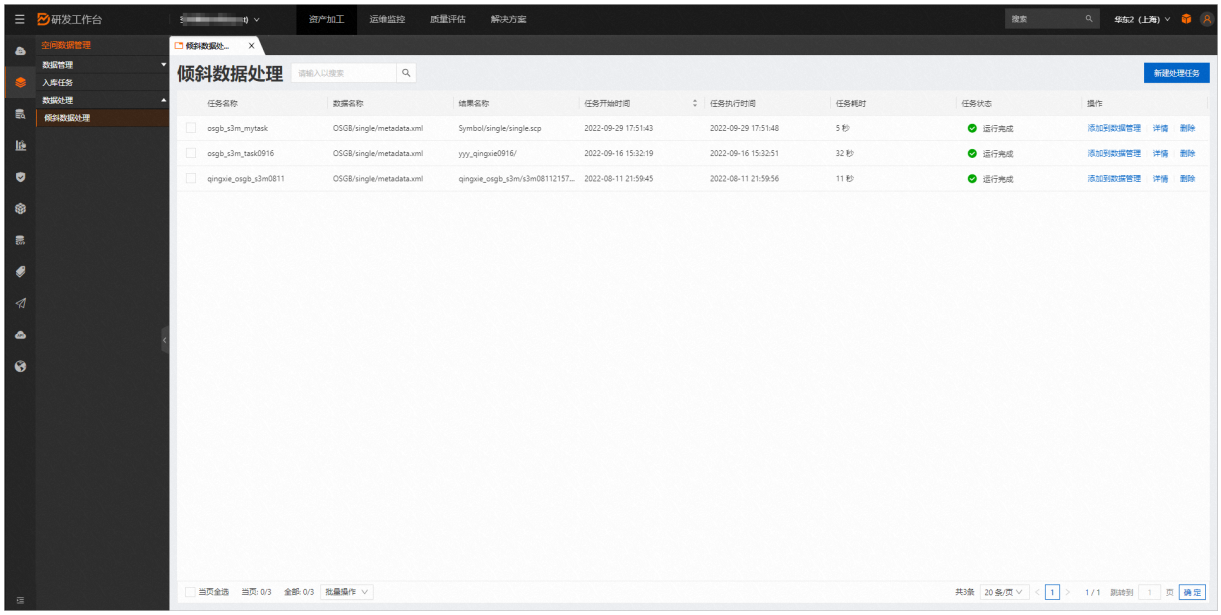
入库任务

- 支持将Shapefile、GeoJSON和CSV文件类矢量数据添加到系统默认的空间数据库或者指定的PostGIS、Ganos数据库，方便对空间数据库进行空间检索、数据质检、血缘记录，打通数据计算引擎链路，使得空间数据被用于数仓建设、标签管理等数据中台任务。
- 支持对入库任务进行管理。
- 支持查看入库任务详情，包括任务基本信息、入库目标表、任务执行状态等。



数据处理

- 支持倾斜数据处理能力，可以将OSGD格式数据转化为S3M数据，再将处理后的数据添加到数据管理中，以便被用于数仓建设、标签管理等数据中台任务。
- 支持对数据处理任务任务进行管理。
- 支持查看数据处理任务详情，包括任务的基本信息、处理的数据、数据坐标位置、结果数据、参数设置和任务执行状态。



6.3.1.3. 数据查询

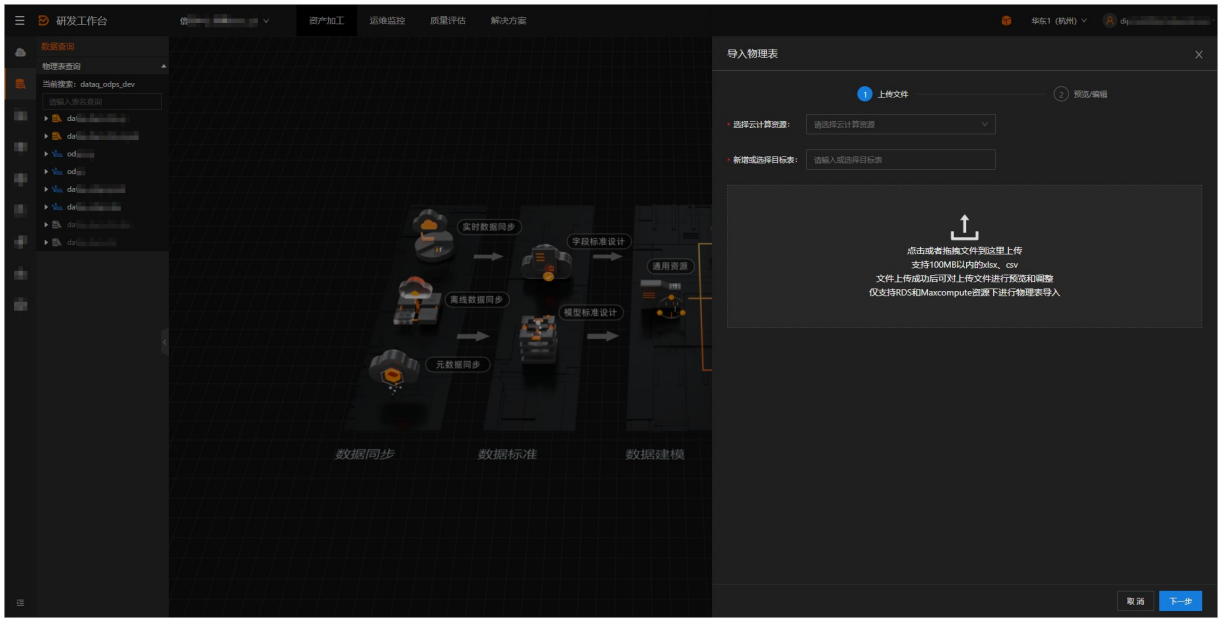
数据查询为用户提供对云计算资源中物理表基本情况的概览能力，帮助用户直观了解物理表基础信息、字段内容分布等情况，建立对数据表的宏观理解，为建立数据标准、定义数据模型、定义数据加工逻辑等工作提供基础依据。

如果数据表格存储的是空间数据，支持在表格的基础上开启地图模式，并实现表格和地图的双向联动。还能查看空间的元数据系统，以及支持SQL的方式进行查询浏览。同时支持对汇聚同步后的空间数据进行地图可视化

浏览，支持地图、属性联动，支持对空间元信息（坐标系、空间范围等）进行全方位查看。

物理表导入

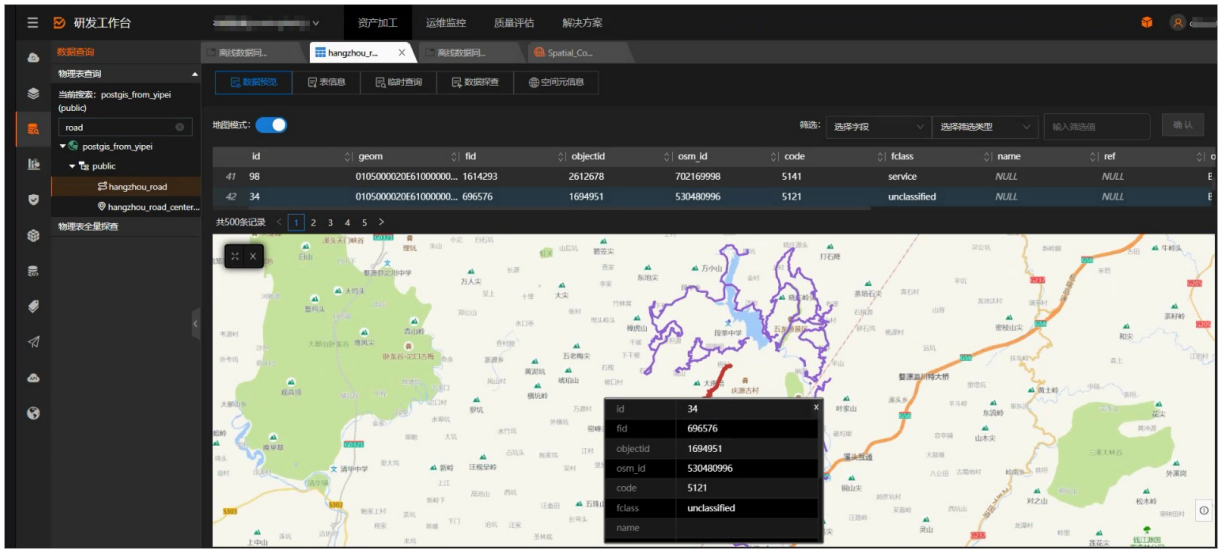
支持100 MB以内的XLSX、CSV文件快捷向RDS MySQL和MaxCompute资源进行导入，支持对要导入的文件进行预览和编辑。



数据预览

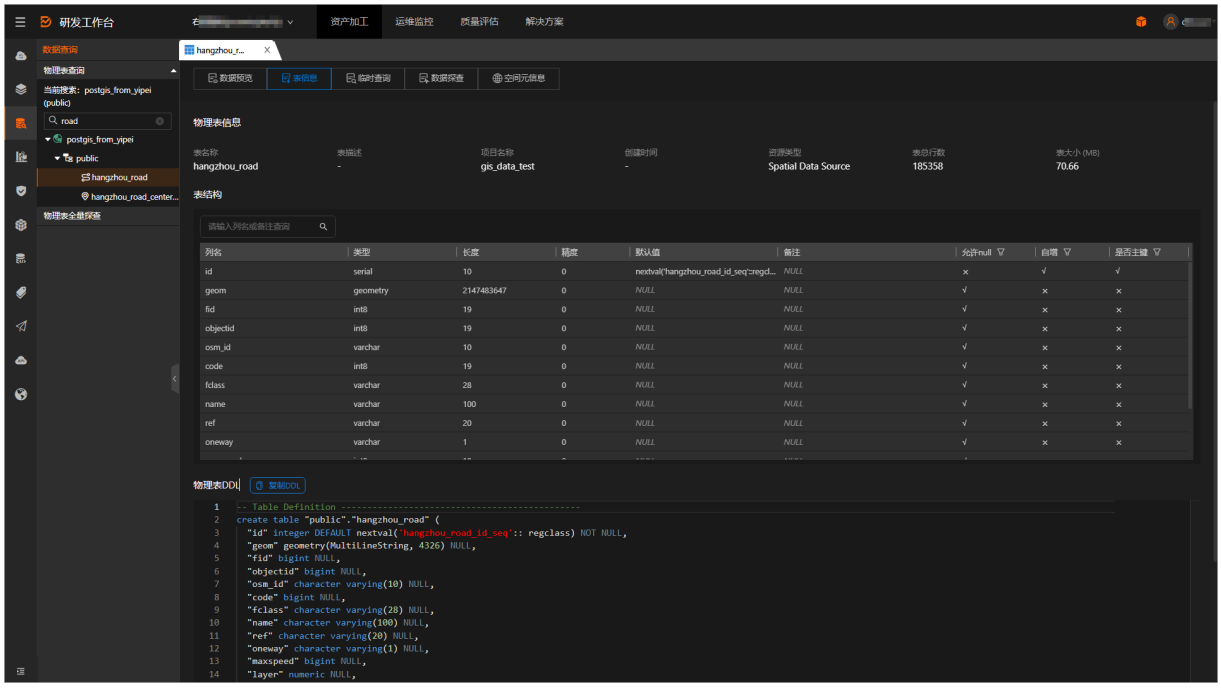
支持不同云计算资源数据的快速预览，可为开发人员和数据应用人员提供全面直观的数据呈现。

- 支持MaxCompute、RDS MySQL、PostgreSQL、AnalyticDB PostgreSQL、Hologres多种云计算资源进行物理表数据预览，可快速对各类数据资源下的物理表的数据内容进行查看。
- 支持开启地图模式，并实现表格和地图的双向联动。
- 支持对物理表字段进行排序，支持根据条件筛选。



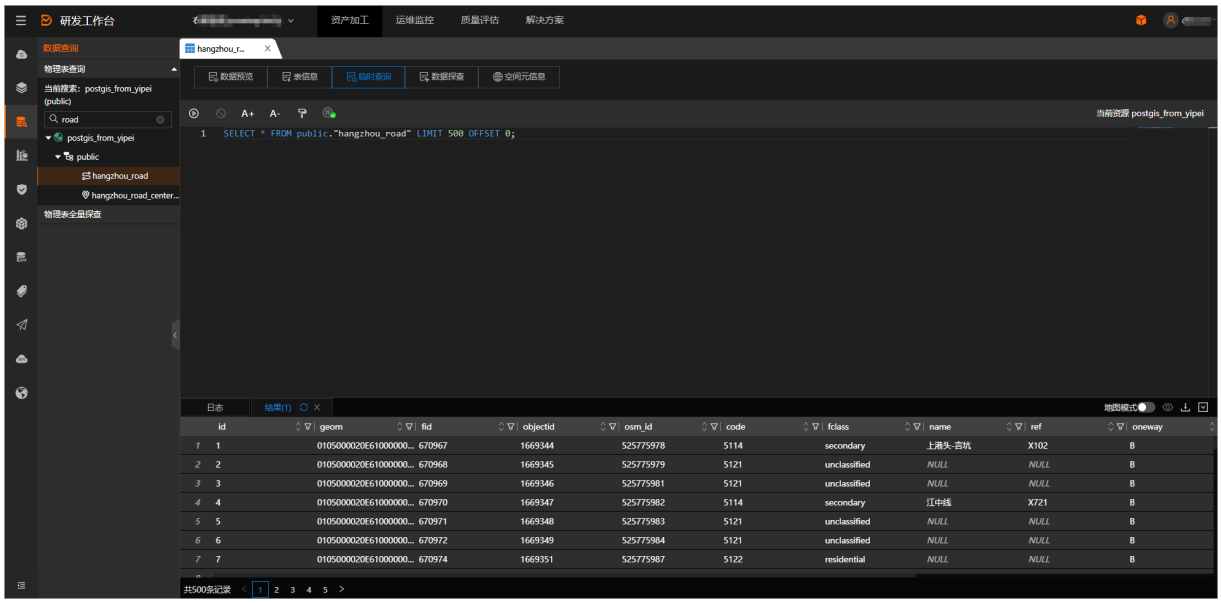
表信息查看

支持不同云计算资源的表信息查看，包括物理表基本信息、表结构和物理表DDL。



临时查询

- 支持通过SQL语句查询预览不同云计算资源的数据资源，可为开发人员和数据应用人员提供全面直观的数据呈现。
- 查询结果支持开启地图模式，并实现表格和地图的双向联动。

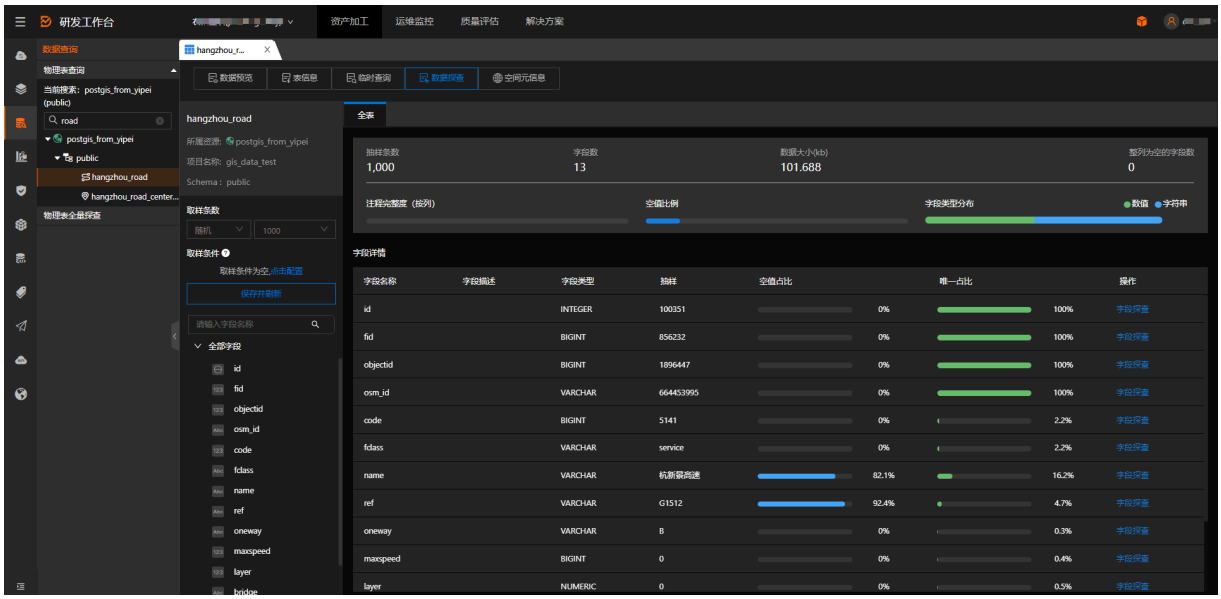


数据探查

支持通过内置的规则及算法对物理表进行探查，快速对物理表整体数据情况、字段情况进行抽样检测。以图表的形式展示物理表数据整体情况，为数据质量规则期望值的设定和数据开发提供输入和依据。

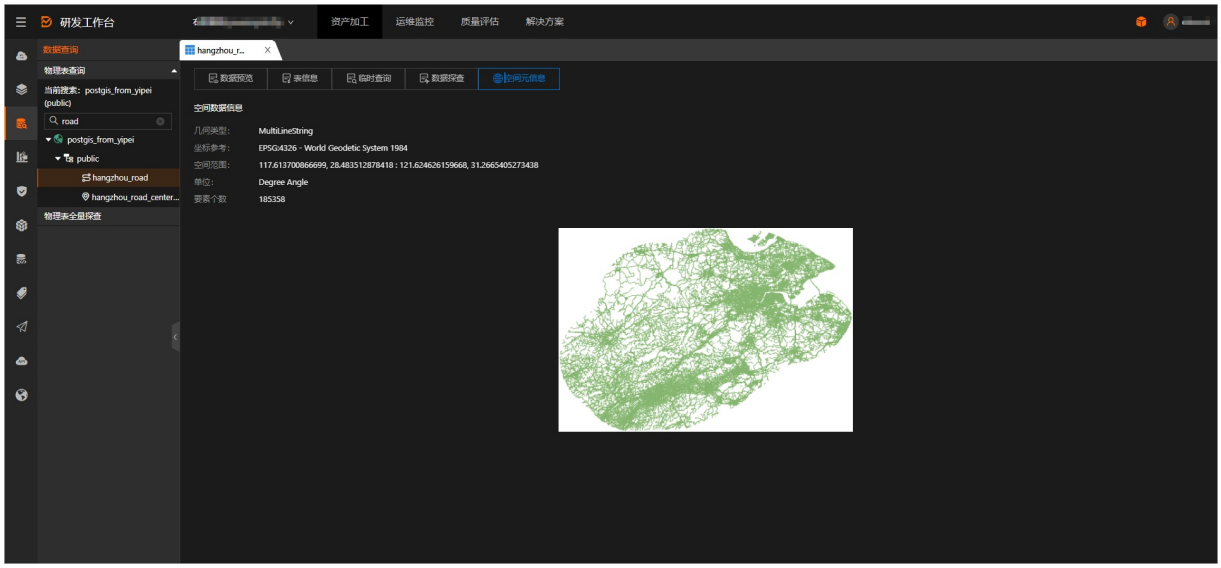
- 支持展示完整表结构、注释完整度、空值率、字段类型分布等表数据分布情况。
- 支持自定义抽样探查规则，支持输入条件组合对标数据进行筛选取样。

- 支持通过内置算法对字段进行分析，包括唯一值、偏斜度、字段频次、异常分布等。以图表形式对分析结果进行展示，并支持对统计图进行下载。



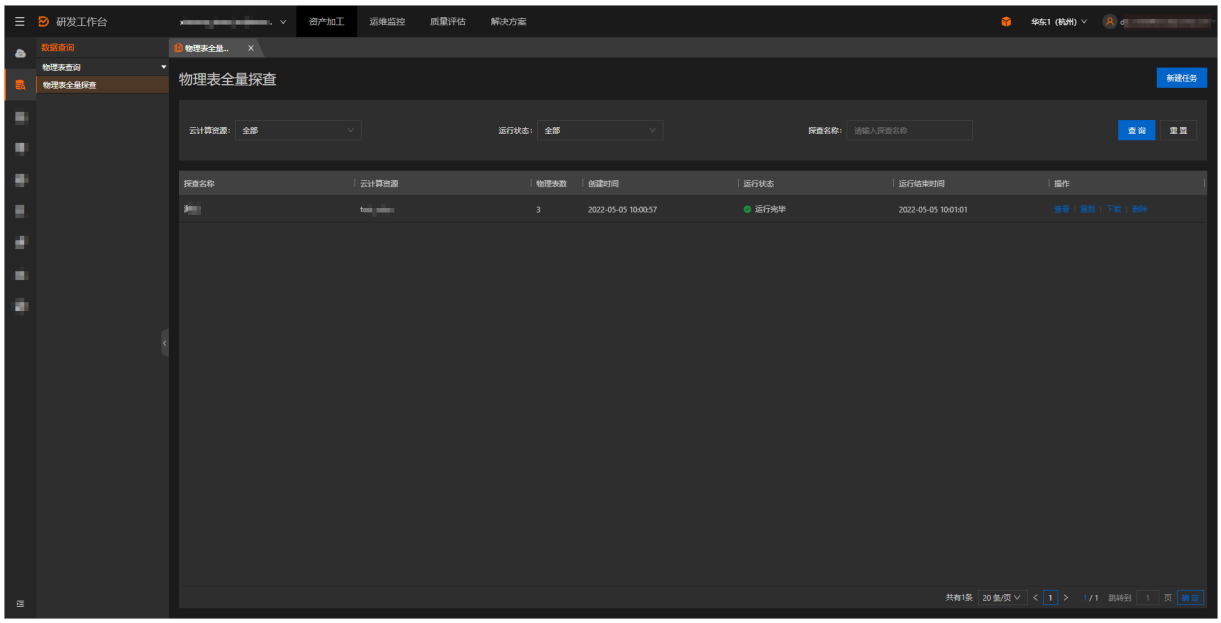
空间元信息

支持查看空间数据基本信息和空间数据缩略图，基本信息包括几何类型、坐标参考、空间范围、单位、要素个数等。



物理表全量探查

支持新建离线任务针对选中计算资源下物理表的全部字段进行全量探查，包括字段类型、唯一值、空值占比、最大最小值等，探查结果支持以Excel报告形式进行下载。

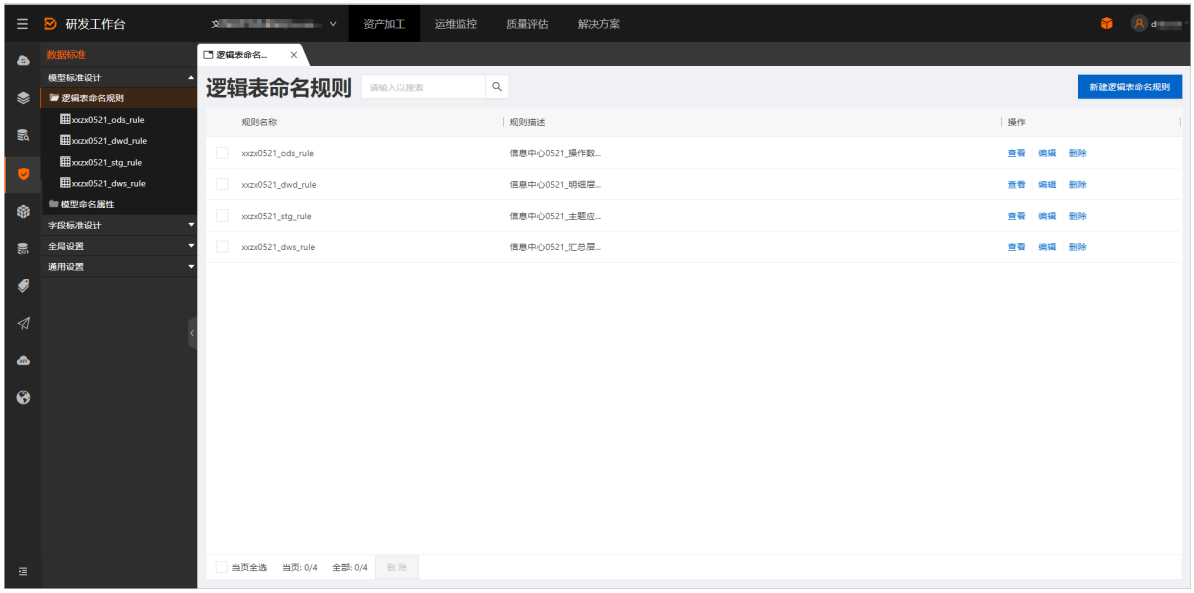


6.3.1.4. 数据标准

数据标准为用户提供模型和字段标准设计和管理的能力。通过模型层级、数据域等属性，对逻辑表的命名进行标准化设计；通过设计标准数据元、数据指标、维度和维度属性，对关键业务对象以及属性和值域进行定义，制定并管理平台遵循的统一数据标准，管控数据的一致性和数据质量。

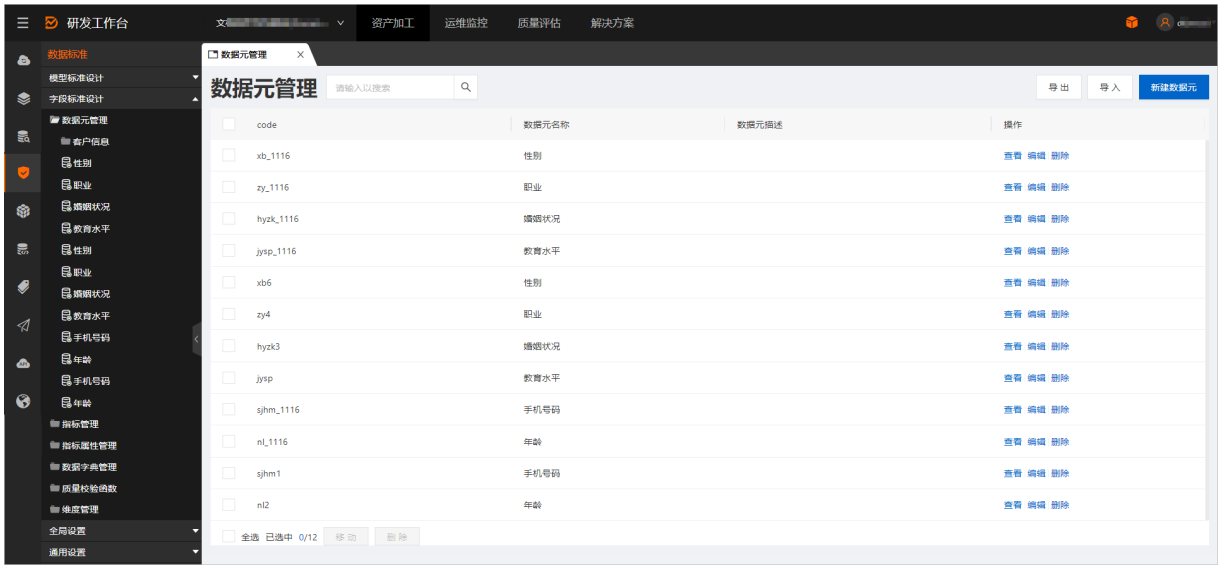
模型标准设计

- **逻辑表命名规则**：支持对逻辑表命名标准的规范设计，可基于模型命名属性，对模型命名规则进行配置，规范统一后续模型设计的命名标准。
- **逻辑表命名属性**：提供对逻辑表命名组成中的一些常用属性进行自定义配置功能，如刷新频率、增量、统计周期。



字段标准设计

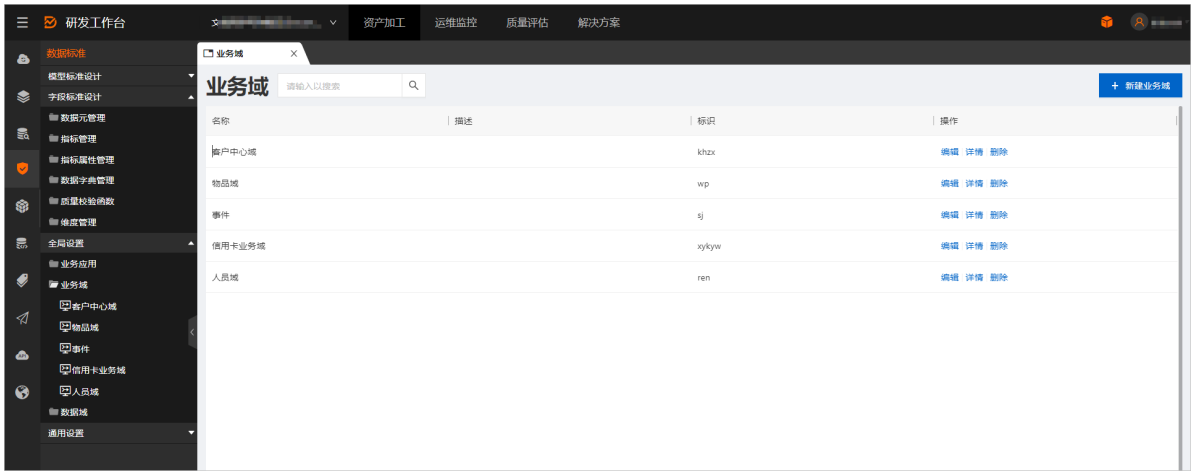
- **数据元管理**：提供对数据元的标准定义和统一管理，通过分析业务流程，抽象关键业务对象、业务对象属性，配置数据元的类型、长度、质量校验函数和引用数据字典，为逻辑表标准化设计及数据治理过程中的质量监控提供规则依据；支持数据元导入、导出，方便行业经验积累和跨项目复制。
- **指标管理**：支持指标类型数据元的定义和管理，可通过选择指标属性的方式，快速新建指标数据元，供逻辑表设计时进行使用。
- **指标属性管理**：支持指标定义的相关属性管理，包含时间修饰、业务修饰、度量属性，帮助快速新建指标数据元。
- **数据字典管理**：支持创建并维护逻辑表中使用的数据字典，方便用户在定义数据元或维度属性时进行数据字典引用，支持数据字典物理化、数据字典导入、导出，方便数据字典使用与跨项目复制。
- **质量校验函数**：支持质量检测函数自定义，方便用户在定义数据元或维度属性时进行质量规则引用，以支持后续对字段的质量规则的快速配置和校验。
- **维度管理**：支持对维度表的统一管理，通过分析业务流程、抽象关键业务对象、业务对象属性，配置维度属性的字段类型、长度、质量校验函数和引用数据字典，为模型标准化设计及数据治理过程中的质量检测提供依据，并提供导入、导出、删除、编辑等功能。



全局设置

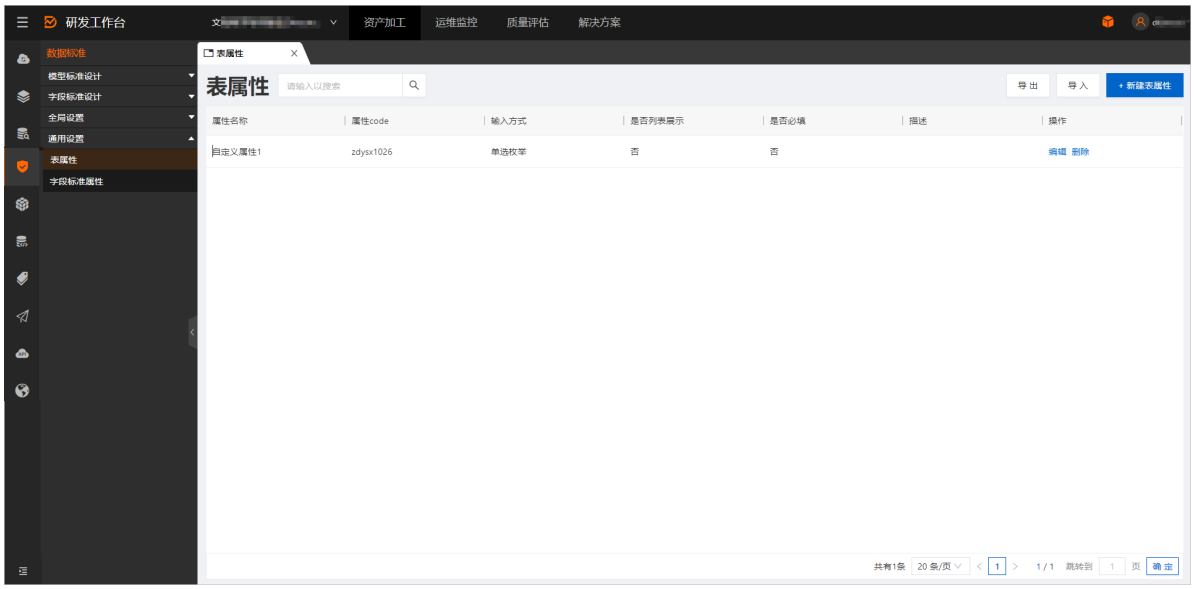
- **业务应用**：对数据模型支撑的应用系统信息进行创建、管理。
- **业务域**：对数据模型来源于哪些系统信息进行创建、管理。

- **数据域**：提供对数据域的定义与管理，供逻辑表设计时进行使用。



通用设置

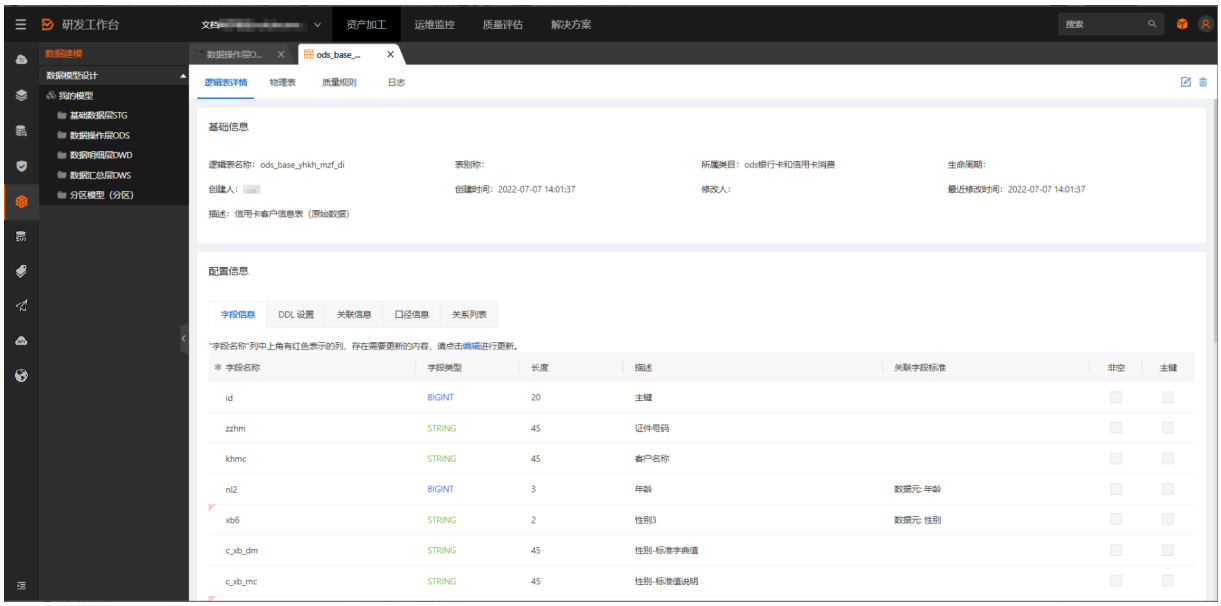
- **表属性**：对逻辑表的属性进行配置、管理。
- **字段标准属性**：对字段标准的属性进行配置、管理。



6.3.1.5. 数据建模

数据建模支持ER模型和维度模型两种数据建模方法，帮助用户完成数据模型、逻辑表的设计和管理。用户可对数据模型的层次结构进行管理，并对逻辑表字段进行设计。字段设计需要在数据标准的约束下，通过模型设计过程中引用标准数据元定义，设置逻辑表相关标准质量规则，保证不同的逻辑表设计人员按照统一口径进行操作。

在数据标准的约束下，保证逻辑表的可持续的维护性及可读性。在数据模型管理模块对逻辑表进行层次结构设计、逻辑表定义，引用标准元素定义，保证不同的逻辑表设计人员按照统一口径进行操作，除此以外还支持：逻辑表质量配置、DDL设置、逻辑表物理化、表关系配置等功能。



- **逻辑表层级管理**：支持对逻辑表层级进行新建，并提供编辑、删除的功能。
- **新建逻辑表**：支持通过可视化的形式创建逻辑表的基本信息、字段信息、关联信息，提供DDL语句新建、引用现有表字段、引用标准库字段、手动添加字段等方式快速创建逻辑表字段及字段信息功能。
- **云计算资源扫描**：支持将云计算资源的已有物理表Meta信息扫描到逻辑表相应层级下，并提供通过正则表达式筛选需要扫描的物理表范围。
- **Excel导入**：支持将Excel里的逻辑表表名、字段信息、所属层级、关联信息导入数据模型管理，便于客户快速搭建数据仓库。
- **Excel导出**：支持将数据模型设计模块中的逻辑表及信息以Excel的形式导出，并提供按全量模型与按查询条件的方式进行导出。
- **DDL设置**：支持根据逻辑表，自动生成并保存不同云计算资源类型的DDL语句。
- **物理化**：支持将已存在的逻辑表物理化到数据库，并对物理化执行结果信息进行展示。
- **表关系配置**：支持建立多个表之间的关联关系，并提供关联关系编辑功能。
- **质量规则配置**：支持可视化与自定义SQL方式制定表级与字段级质量规则，并提供手工创建逻辑表字段时通过关联数据元自动生成质量规则的功能。
- **一致性**：支持当前逻辑表和已物化物理表的字段一致性对比，支持当前逻辑表与数据开发所引用的节点中模型的字段一致性对比。

6.3.1.6. 数据开发

数据开发为用户提供一站式计算节点开发能力，通过对数据加工流程的开发、部署、调试等环节的一体化管理，数据开发实现数据加工工作流编排、加工逻辑的复用，大幅提高数据开发效率。

数据开发帮助用户优化智能系统的架构，提高系统的工程化水平，让用户能够集中精力处理计算逻辑问题，提高开发效率。数据开发除了在系统层面上集中管理计算节点，降低计算节点和系统其他组件的耦合之外，也对计算节点内部的实现提供了一组规范。在实际的业务系统中，数据来源多种多样，不同数据对数据处理的时延和数据量的要求不同，这就需要综合多种不同的平台，包括批量、流式环境等。进行数据处理时，上一个平台计算完毕后把数据传递给下一个平台进行计算，多个平台互相配合来完成任务。数据开发定义了各平台间的数据流转规范，以及各平台计算节点执行顺序的方法，从而让用户能够更加专注于计算逻辑的开发，不必关注底层的技术细节。

数据开发规范了各个计算节点的接口，在一定程度上提高了计算节点复用的能力，降低了多人协作开发的难度。提供了跨多种存储/计算平台的计算节点工作流程配置和快速部署上线的能力。因此，用户在构建数据应用时，只需从解决实际某一业务场景出发，组合多种类型的计算节点来配置计算节点工作流程，从而快速解决业务工作流并依托数据开发的跨平台多异构任务的一键部署和上线能力，极大的提高了工程效率。

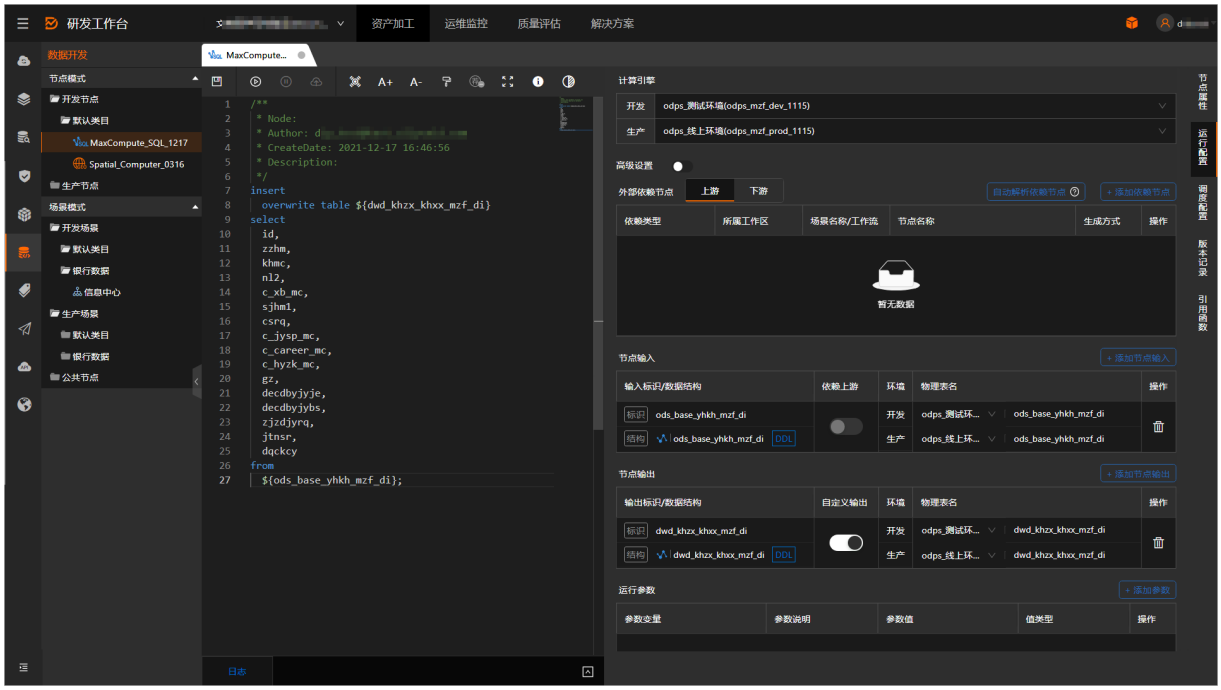
节点模式

节点模式任务主要适用于计算节点相对独立，整个流程由单个节点组成的这类简单数据处理场景，支持灵活的节点开发模式，支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理。支持批量离线计算等多种类型的计算节点。支持同一租户内，不同工作组中的任务依赖和统一调度。

开发节点

支持节点任务类目的新建、编辑和删除以及支持离线节点注册，对节点脚本、自定义函数、节点输入、节点输出、节点参数进行配置和管理以及支持配置节点的依赖关系以及支持节点任务的部署、运行和上线。

支持空间数据节点注册，对空间数据的数据管理、数据导入、数据导出、矢量分析、数据计算、数据筛选、坐标处理和栅格分析等多个算子的直接应用，并对图层输入、图层输出参数进行配置和管理。

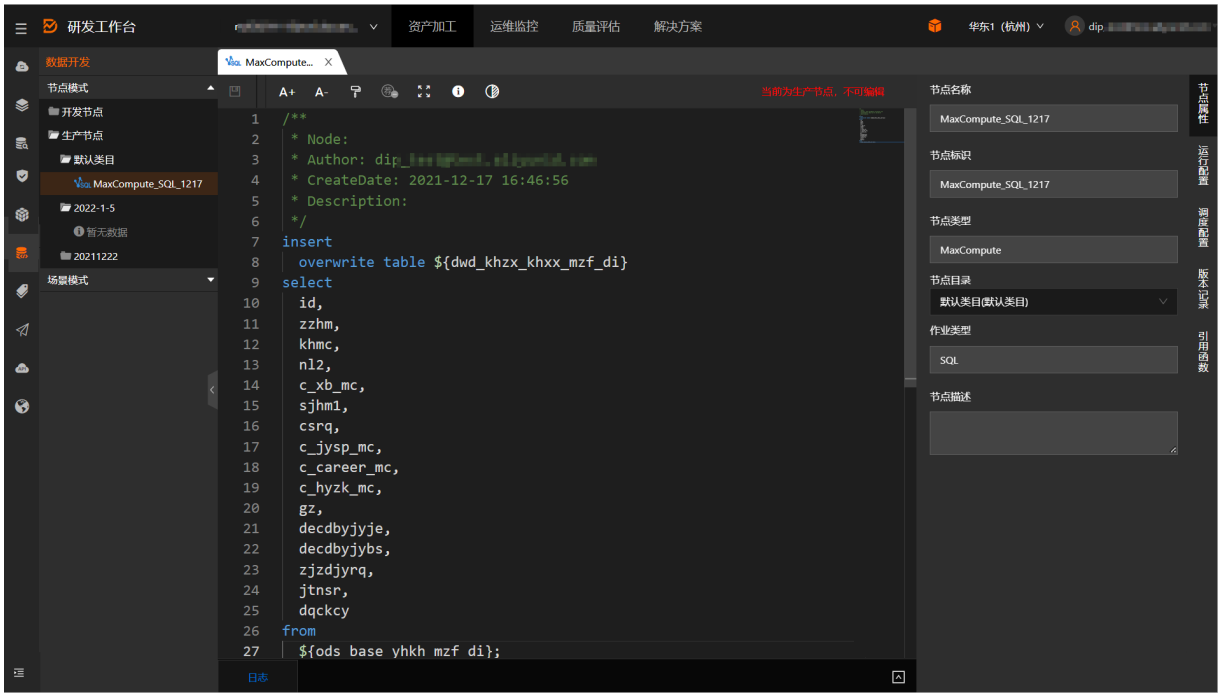


- 支持节点任务类目的新建、编辑和删除。
- 支持节点任务的新建、编辑、删除。
- 支持离线节点注册，对节点脚本、自定义函数、节点输入、节点输出、节点参数进行配置和管理。
- 支持空间数据节点注册，对空间数据的数据管理、数据导入、数据导出、矢量分析、数据计算、数据筛选、坐标处理和栅格分析等多个算子的直接应用，并对图层输入、图层输出参数进行配置和管理。
- 支持根据SQL语言，自动解析离线节点的输入输出。

- 支持输入输出选择任意结构，方便操作。
- 支持配置节点的依赖关系。
- 支持节点任务的部署、运行和上线。

生产节点

支持生产环境中节点任务的查看和下线，对节点的算法信息、计算资源、存储资源、调度周期入、输出等详细信息以及部署信息的查看。

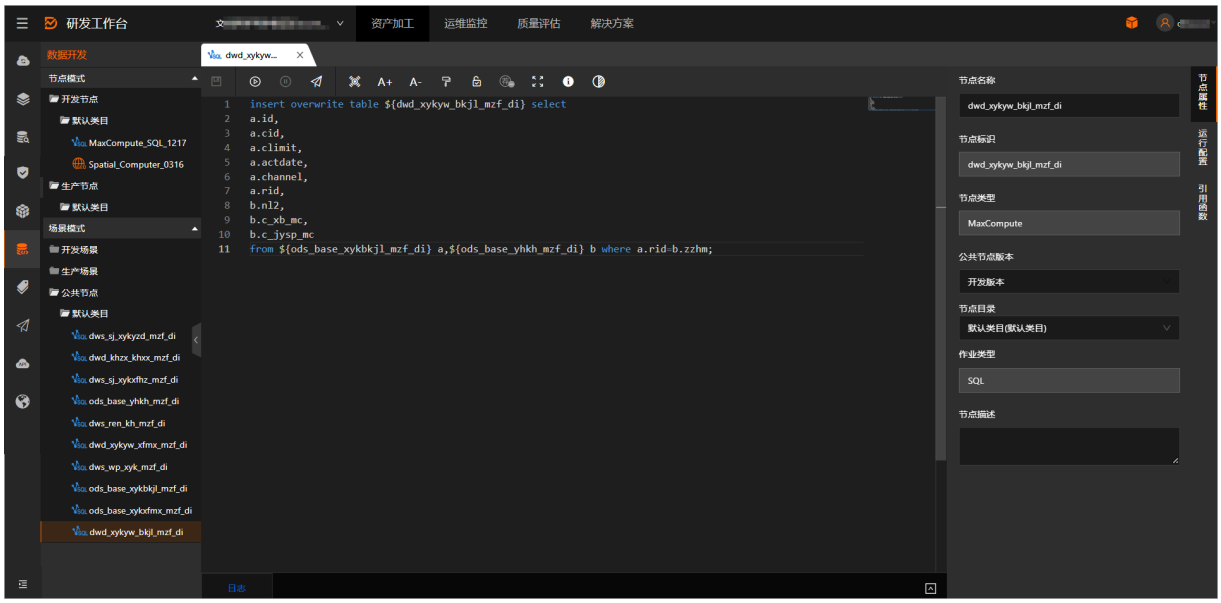


场景模式

场景模式适用于算法相对比较复杂，算法流程由多个节点，在一个场景内的混合编排组成的节点工作流场景，支持可视化的自定义任务流的场景开发模式，支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理。提供数据加工、算法服务任务流程开发、编排与调试、上线、部署、维护等功能。支持流式计算、批量离线计算等多种类型的计算节点在一个工作流中统一编排。支持同一租户内，不同工作组、不同场景间的全局任务依赖和统一调度。提供可视化操作界面，通过托拉拽的方式连接计算节点迅速实现数据加工流程编辑，大幅提升数据开发工作效率。

公共节点

支持公共节点类目管理。支持批量、流式等多种类型的公共节点，对公共计算节点脚本、公共计算节点输入、公共计算节点输出、公共计算节点参数等进行配置和管理。支持公共节点的版本管理。



- 支持公共节点注册，对公共节点脚本、自定义函数、公共节点输入、公共节点输出、公共节点参数进行配置和管理。
- 支持根据SQL语言，自动解析离线节点的输入输出。
- 支持输入输出选择任意结构，方便操作。
- 支持公共节点版本管理，用户在开发环境对公共节点进行修改和调优，调优完成后，发布公共节点并指定版本号，将算法发布到生产环境。
- 支持公共节点的废弃，避免旧版本公共节点被其他使用者重新使用。
- 支持查看公共/计算节点关联的场景，帮助用户清晰的了解公共节点的使用情况。
- 支持批量、流式等多种类型的公共节点。

开发场景

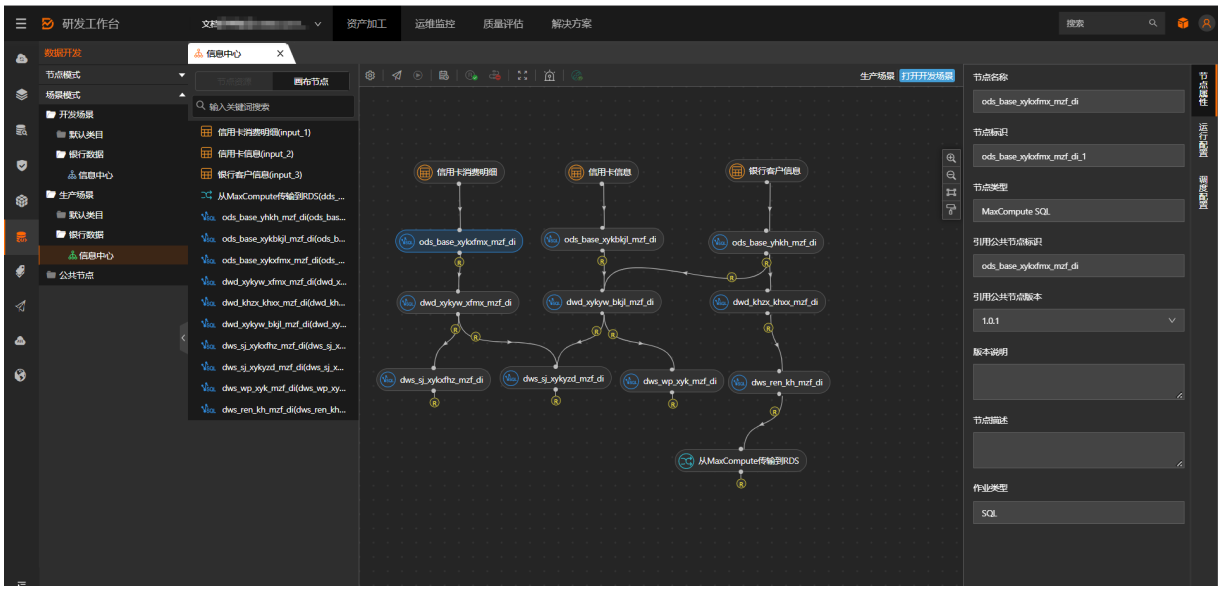
支持开发场景类目管理。支持开发场景引用公共节点。支持数据处理流程编排（多个计算节点组合成一个任务流），用户可灵活的通过可视化方式自定义任务流。支持批量、流式等多种类型的节点在一个场景内的混合编排。支持开发环境中场景的部署、运行和上线，并将开发环境中调试通过的场景发布到生产环境。



- 支持新建、编辑和删除开发场景目录。
- 支持新建、编辑、删除开发场景。
- 支持在开发场景中引用公共节点，并配置离线计算节点的调度信息。
- 支持在开发场景中新建计算节点，配置计算节点的各类参数及离线计算节点的调度信息。
- 支持在开发场景中新建系统节点，并配置系统节点的配置信息。
- 支持通过可视化方式配置上下游计算节点的依赖关系，并配置输入、输出参数的对应关系。
- 支持根据配置的输入信息，自动解析外部依赖节点信息。
- 支持输入和输出参数选择任意结构。
- 支持开发场景的部署、运行和上线。
- 支持自动生成数据质量检查节点（当计算节点的输出参数引用的逻辑表配置了质量规则时）。

生产场景

支持生产环境中场景的运行和下线，对各节点的算法信息、计算资源、存储资源、调度周期入、输出等详细信息的查看。



6.3.1.7. 智能标签

智能标签产品从大数据应用落地点出发，提供了将治理后的数据以业务化视角进行建模、查看、管理及使用，并提供业务衍生标签的自定义功能，为上层应用提供统一的标签数据目录和标签调用接口，沉淀上层应用制作的模型标签，实现高价值标签共享复用，形成标签运营生态。

智能标签在相应行业应用解决方案的结合下，能够让不是很熟悉大数据应用系统开发的程序员也能够快速搭建大数据应用，从而实现大数据价值的快速落地，智能标签帮助数据开发人员进行数据表的实体化，加速应用开发，沉淀业务模型的输入。比如在订单中，卖家能够理解的内容是成交笔数、成交金额、成交日期等实际现实对象，而非一张张表，如订单表。需要智能标签能面向零售行业，提供实体化的对象，如买家这个实体，这个实体上，需要赋予性别、年龄等标签信息，同时和商品建立关系，这个被赋予实际标签的实体，才是卖家能够理解的对象，基于这个对象才能去做分析。

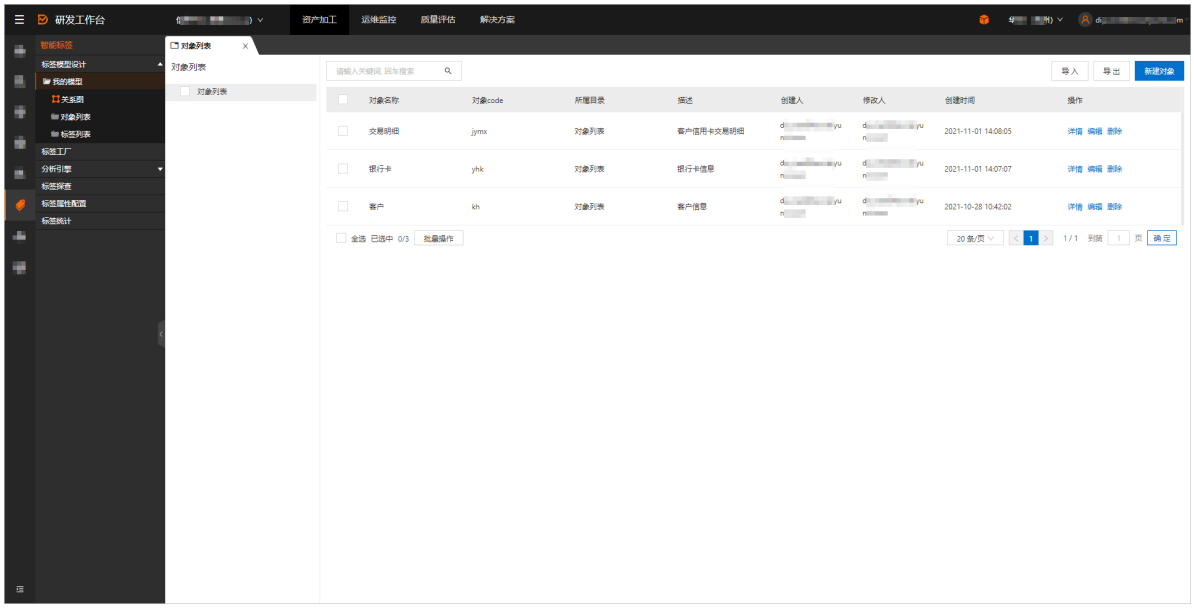
智能标签帮助构建这样的对象，提供给分析型应用实际的内容，同时让应用开发者也直观了解需要去获取的数据对象，直接使用实体化后的对象及标签进行加工、衍生、调用。

智能标签是以标签中心为基础，建立跨多个云计算资源之上的统一逻辑模型，开发者可以在“标签”这种逻辑模型，通过接口的方式进行快速的应用搭建。这种方式的好处：一在于屏蔽掉应用开发人员对于下层多个计算存储资源的深入理解与复杂的系统对接工作，二在于通过数据服务的形式透出也有助于IT部门对数据使用的管理，避免资源的重复和冗余。

标签模型设计

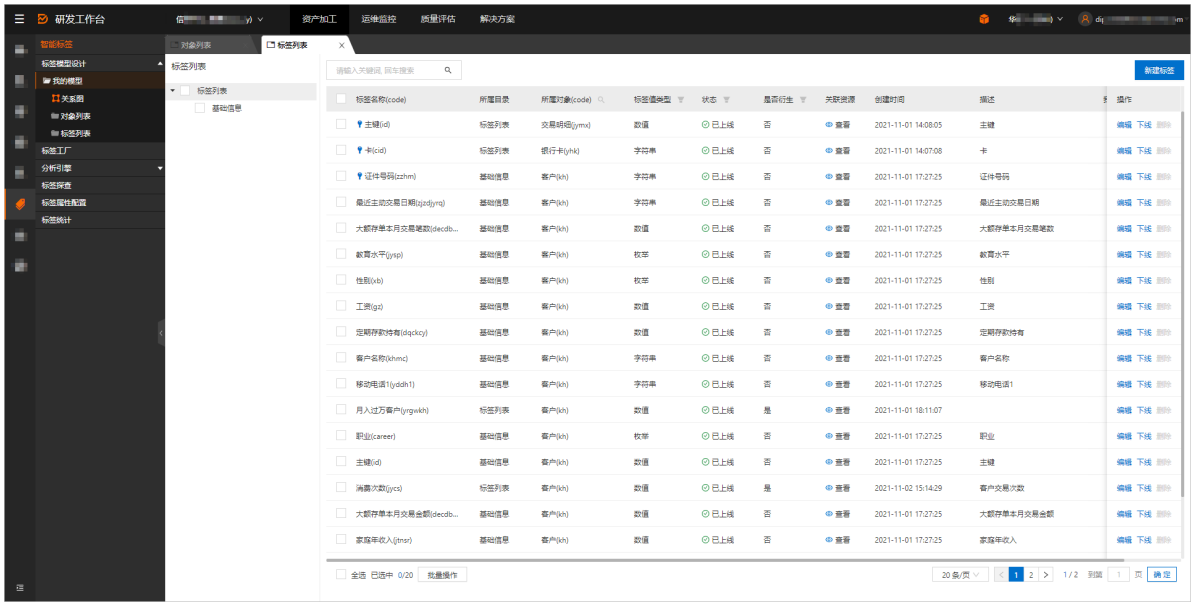
- 标签模型基于OLT（对象、关系、标签）模型来抽象物理数据，构建语义层，为业务人员“看懂数据”从而“使用数据”提供基础，支持对象关系的多级类目管理。
  - 对象定义：包括实体类（无时间维）和事件类（有时间维），描述一个客观存在实体或事件；从数据角度看，是一个多种数据描述同一实体或事件的集合。
  - 关系定义：关系建立在及两个对象上的关系。如用户和购买商品事件的关系，用户和学习课程事件的关系等。
  - 标签定义：标签是挂接在对象，对对象进行描述的属性的集合。
- 对象管理

- 新建对象及标签：支持跨存储类型的数据源整合，低门槛构建“宽表”，形成业务可以理解的对象及其原生标签信息。
- 对象导出/导入：支持将对象导出/导入，包含对象相关的标签、计划、标签API，便于在其他系统或工作组中快速部署。



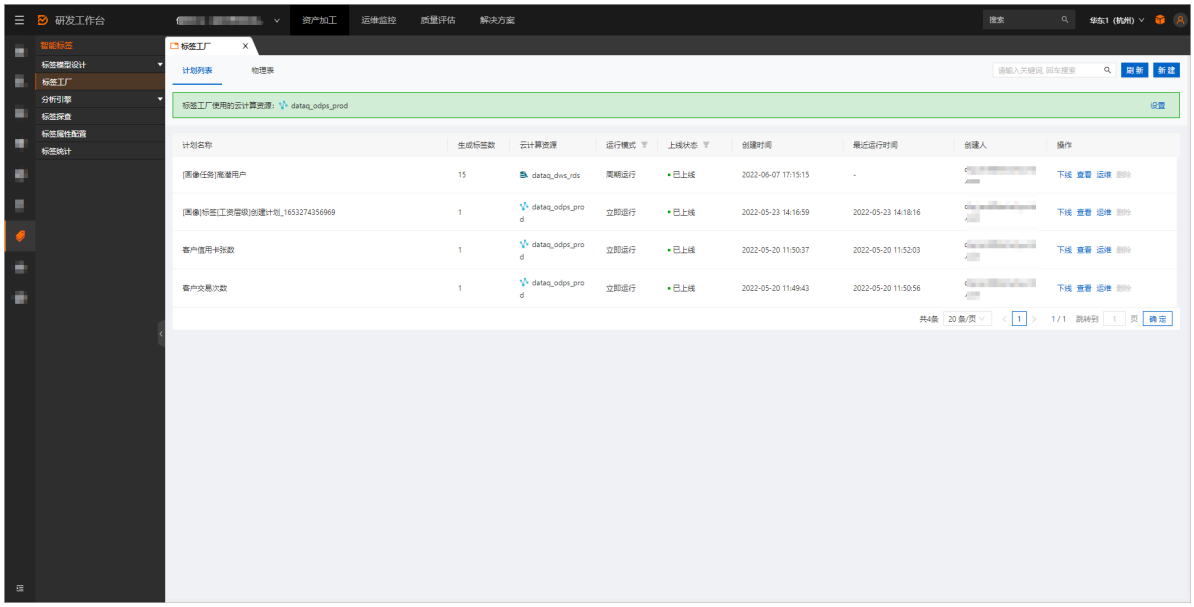
- 标签管理

  支持对标签进行编辑、上线/下线、公开/撤回；实现标签的生命周期管理。



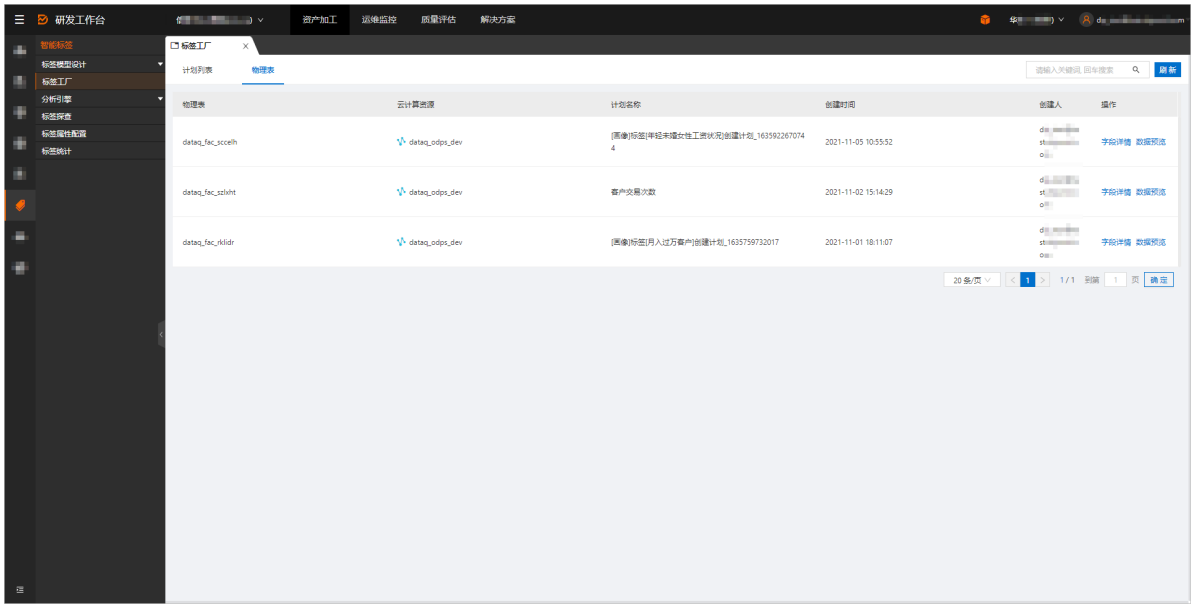
标签工厂

- 计划列表管理
  - 支持已有标签加工生成衍生标签功能。
  - 支持文本TQL编写标签方案，支持衍生标签配置关联到实体或关系上，支持标签方案生成任务一次性调度和周期调度功能。
  - 支持标签方案生成任务的运维管理功能，并提供对计划任务的上线、下线、查看、运维、删除等功能。



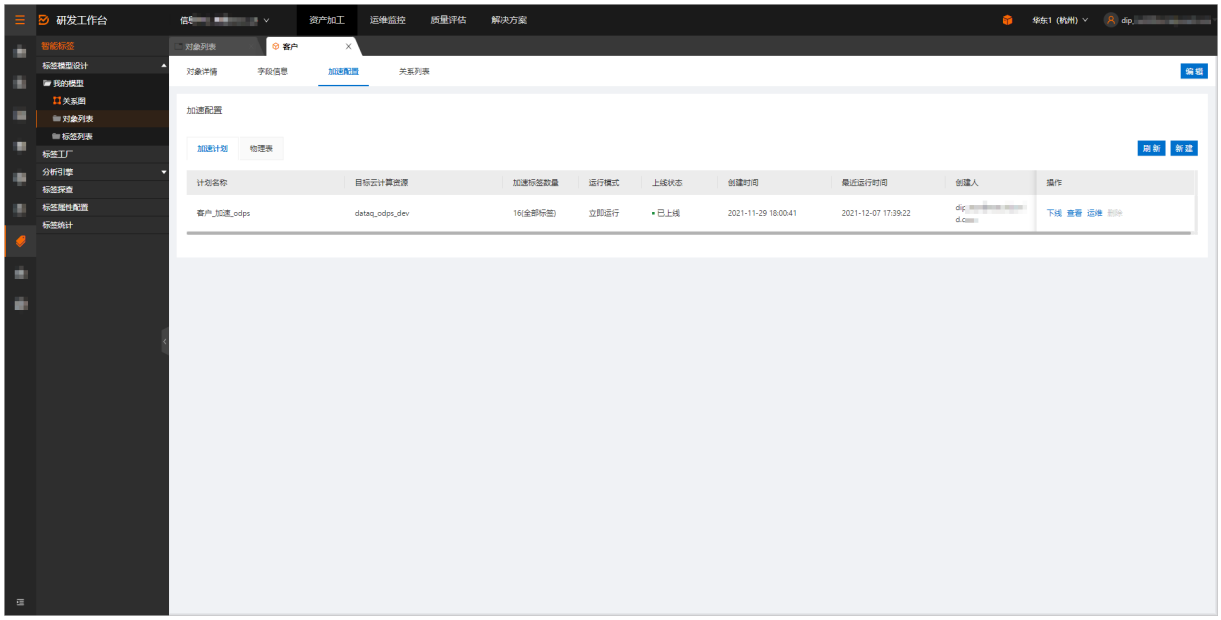
- 物理表查看

  支持查看云计算资源下的物理表信息，并提供刷新、字段详情查看、数据预览等功能。



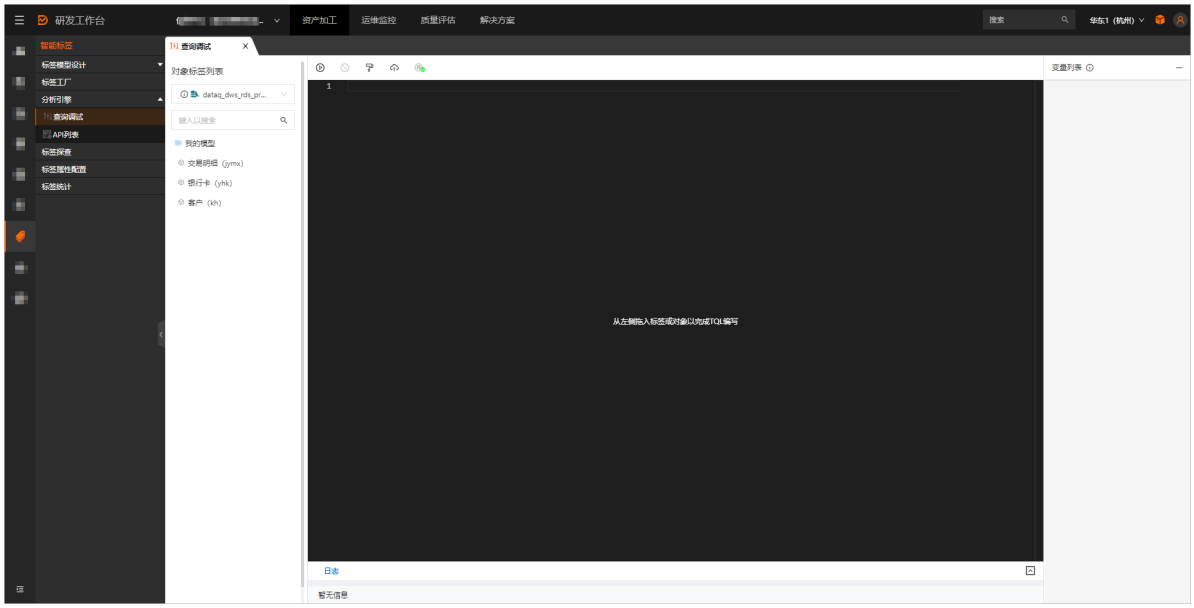
标签加速

支持将构建的标签模型统一加速至分析型数据库形成大宽表，以提供后续分析场景使用。



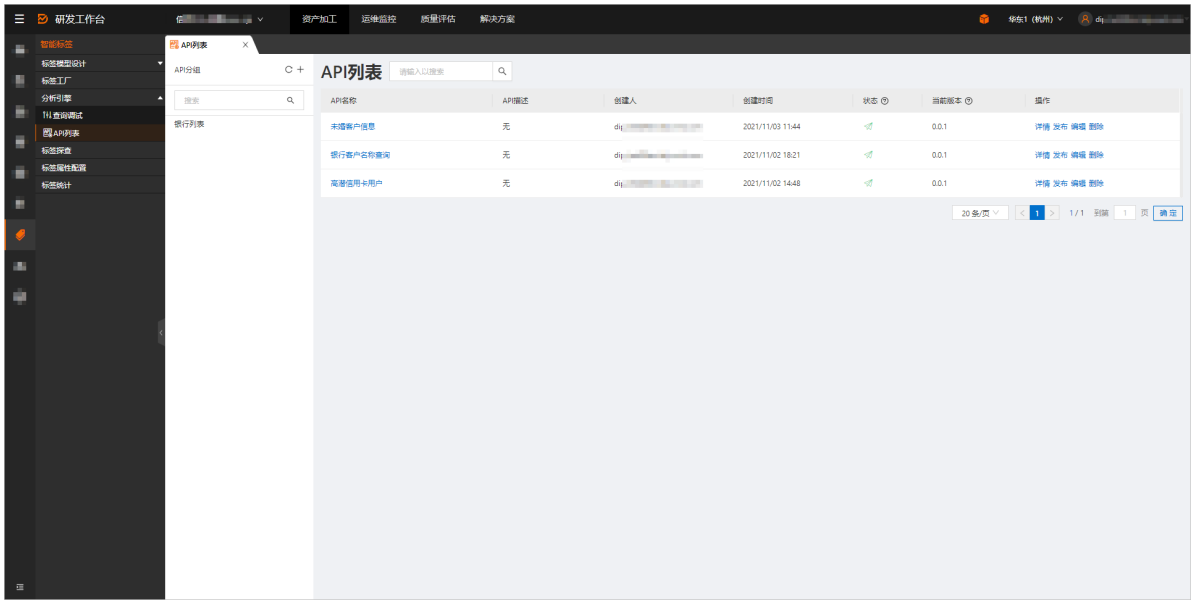
分析引擎

- 查询调试
  - 支持实体关系上的标签和获得授权标签的查询功能。
  - 支持实体关系使用TQL（Tag Query Language）语言的关联查询。
  - 支持查询结果的列表查看和JSON格式查看。
  - 支持可视化的页面调试，展示TQL的每个执行流程；每一流程的时间性能消耗。
  - 支持多种云计算资源中的标签查询。
  - 支持分析查询结果生成API，并提供在分析查询中配置参数。



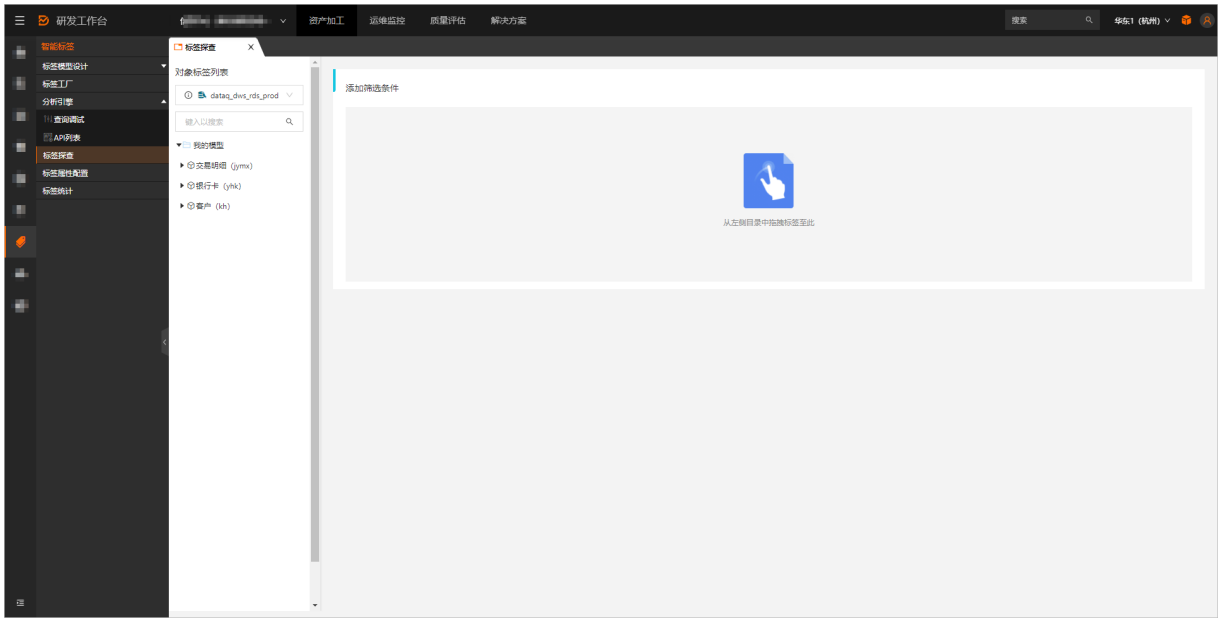
- API列表

支持API的类目管理、列表管理、详情查看、调试和版本管理等功能。



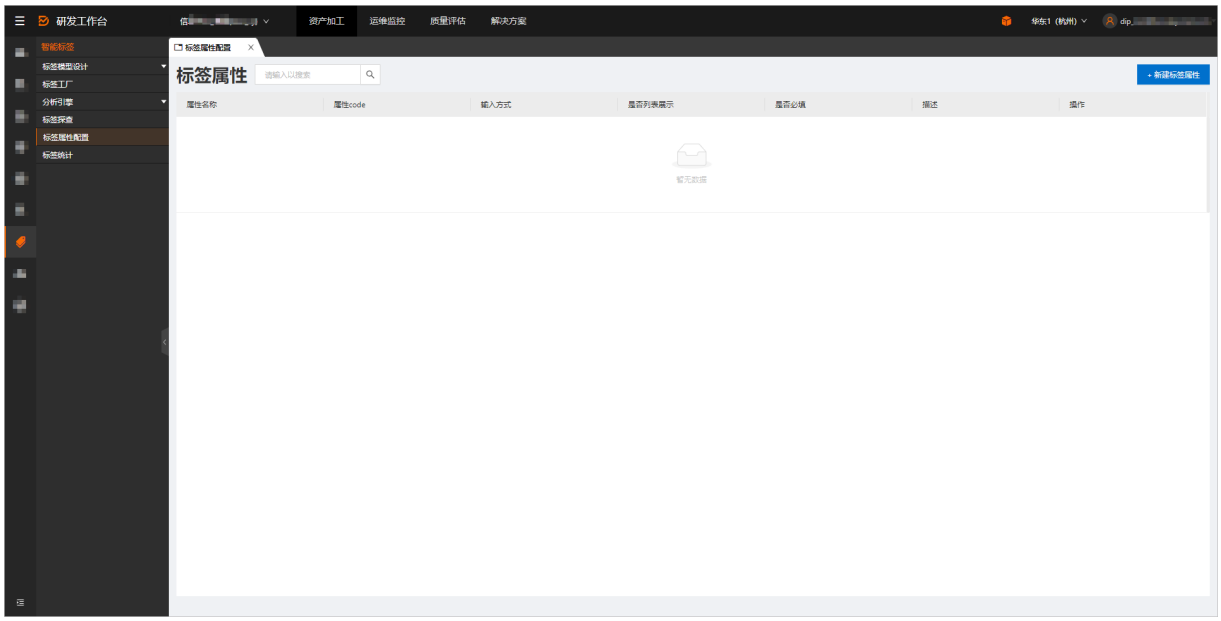
标签探查

支持查询所有实体的标签及标签内容数据，支持多标签的内容筛选能力。



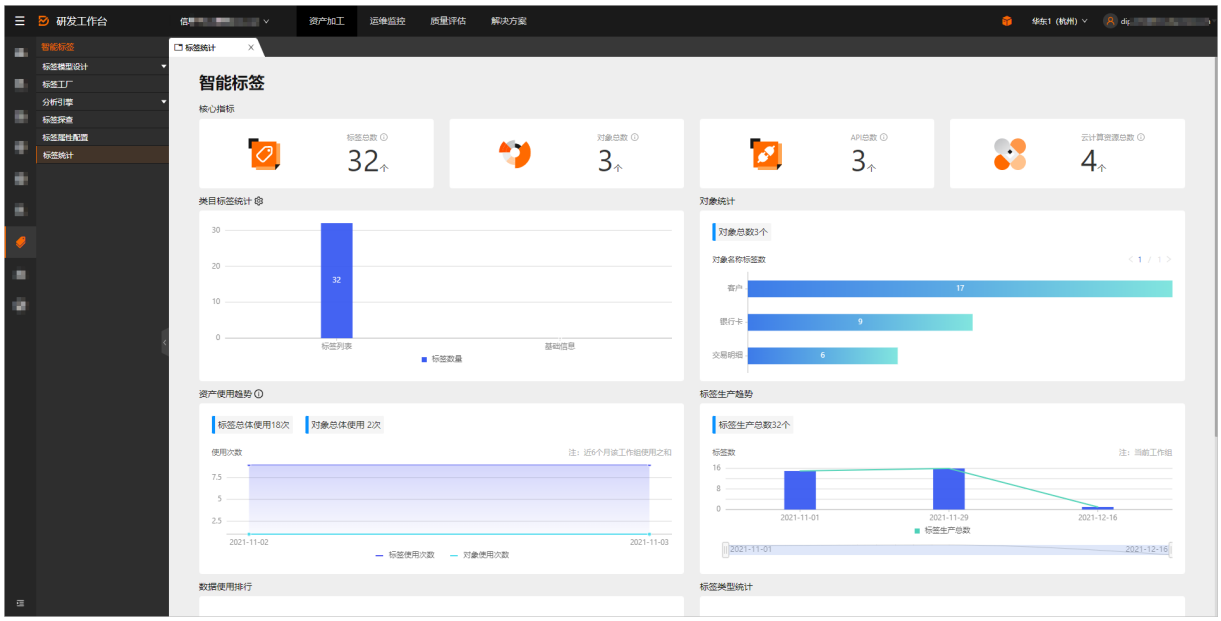
标签属性配置

支持为标签提供易于识别的属性和属性值，便于画像分析时能够提供方便快捷的进行数据筛选，提高了画像分析的效率。



标签使用统计

支持标签生产和使用情况统计，包含标签类目统计、标签使用、对象使用统计等。

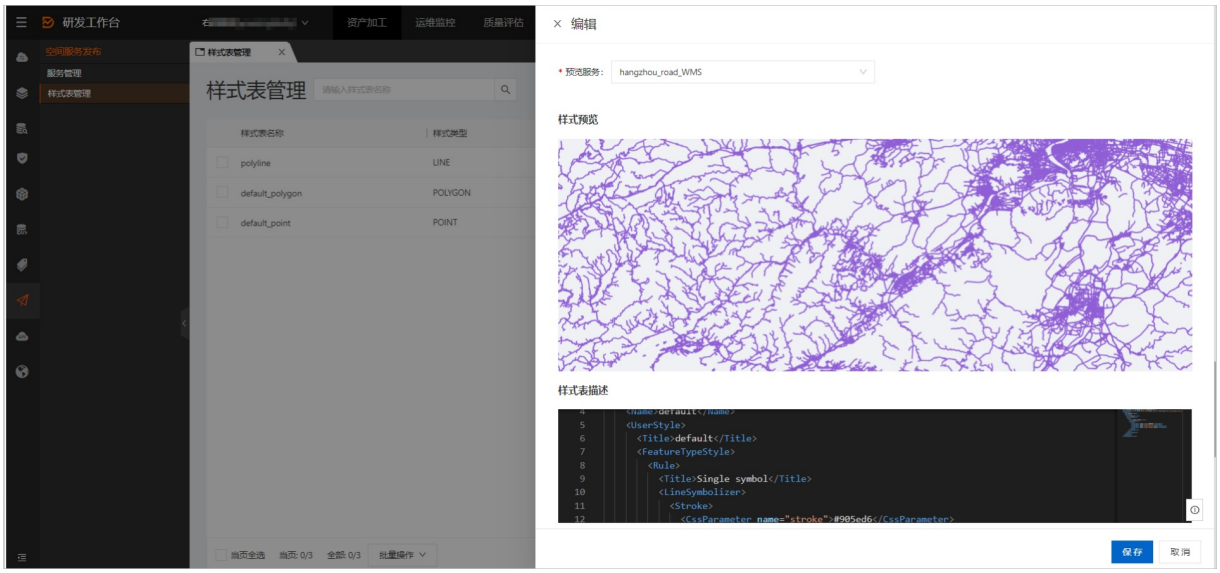


6.3.1.8. 空间服务

通过空间服务发布将空间数管中的各类型空间数据发布成行业标准的空间服务，支持发布OGC WMTS、WMS、WFS服务；TMS（MVT）矢量瓦片服务；S3M、3DTiles等标准的三维服务。

样式表管理

支持通过SLD文件配置地图服务的样式风格，通过上传SLD文件方式添加空间数据的样式表，并支持点、线、面的空间数据样式。



服务管理

支持把空间数据发布为服务，选择需要的数据、设置发布的服务类型、以及坐标系后发布为服务，包括地图服务和瓦片服务。



6.3.1.9. 统一服务

在数据中台建设过程中，搭建核心处理平台，并根据相关标准要求，将源数据或治理后的数据，按照“全面、高效、统筹”的共享原则，“低延时”、“精准化”的传递给用户，方便更多应用消费数据，实现数据价值。协同管控各种组织机构，为智慧化运行提供支撑，实现跨层级、跨部门、跨行业、跨区域信息即时共享，做到高可靠、低延时。同时，数据被应用的使用情况评估也作为数据中台日常运营的重要指标项。提供更多数据生产需求，完成数据从生产、使用、到评估的闭环。

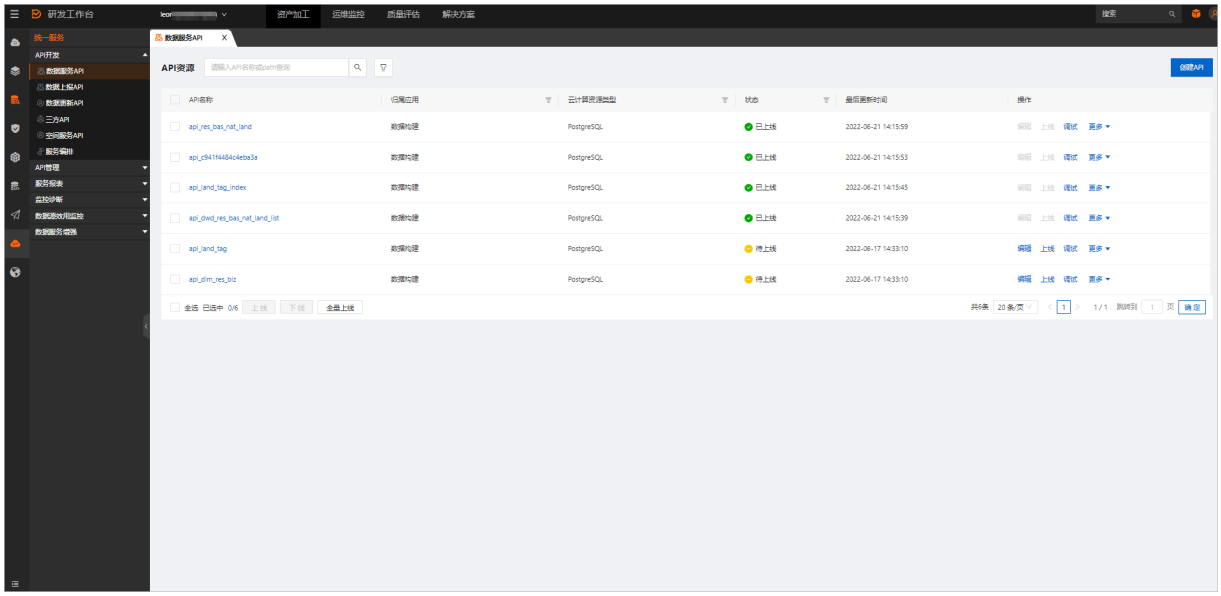
统一服务提供了如下数据服务能力。

- 数据服务API的源头为数据表，可以是单表也可以是多表，请求后最终以SQL的形式到数据库查询，并返回符合条件的数据，将数据组装成报文格式返回。目前数据服务API提供单数据服务API，SQL转API。数据源可以是OTS、ApsaraDB for RDS和AnalyticDB MySQL等数据存储中的数据表。
- 流式数据服务的源头为DataHub或RocketMQ的Topic，服务封装为DataHub Topic的读取API，返回Topic中的数据。数据源可以是DataHub和RocketMQ。
- 批量离线数据服务的源头为MaxCompute的数据表，服务封装为根据分区获取分区内全部数据的API。
- 非结构化数据服务的源头为OSS，服务封装为根据文件名称返回文件数据流。

统一服务为用户提供了数据中台建设过程中的数据服务化能力，面向数据开发者提供覆盖各个加工阶段统一体验的、便捷的数据查询转服务、服务管理、服务运维能力；面向数据资产管理者提供服务的统计分析、服务用量统计分析、热门数据统计分析能力，实现数据中台建设后半场“数据应用”的有效落地，支撑数据智能应用的高效开发。

API开发

支持数据服务API、三方API、数据上报API、数据更新API、服务编排API和空间服务API。支持创建、编辑和调试API，支持的API类型包括实时查询类、离线批量类、流式数据类、非结构化数据类。



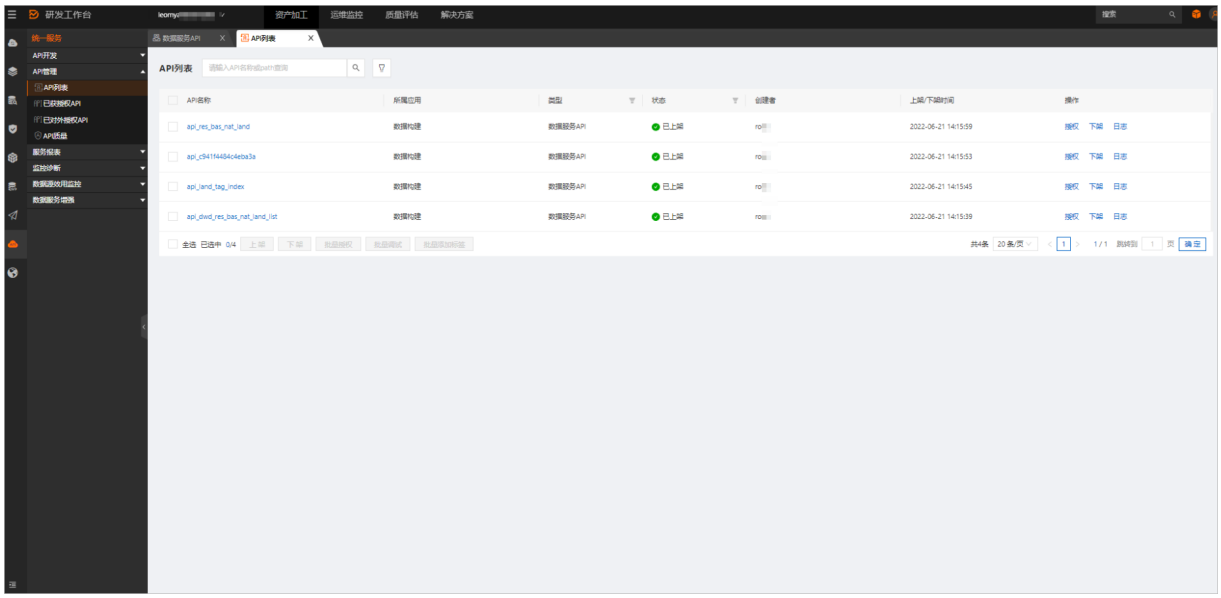
- 数据服务API：将数据库中的表的通过向导模式和SQL模式配置生成API，方便开发者申请使用或调用。目前支持的云计算资源有ApsaraDB for RDS、AnalyticDB MySQL、MaxCompute、DataHub、Table Store、Object Storage Service、AnalyticDB PostgreSQL、RocketMQ、Hologres、Presto、PolarDB等。
- 三方API：将已有API注册到统一服务，通过注册API的相关信息，包括提供商URL、ContentType、请求方式、请求参数、返回参数、返回数据样例等信息将三方API注册到统一服务管理。
- 数据上报API：将数据写入API，用户调用API时将数据上报到服务端，目前支持ApsaraDB for RDS、DataHub和AnalyticDB PostgreSQL作为数据上报目标端。
- 数据更新API：指通过调试或调用创建的数据更新API，修改数据库中的表数据，目前支持的云计算资源有ApsaraDB for RDS和AnalyticDB PostgreSQL。
- 服务编排API：按照业务逻辑，灵活的实现跨计算资源的多个API服务编排，提供拖拽式、可视化的API编排能力。支持按照业务逻辑，以串行、并行和分支等结构编排多个API为工作流。目前支持的服务节点有数据服务API和三方API。
- 空间服务API：将空间数据管理中的各类型空间数据发布成行业标准的空间服务API，支持各类型空间数据的汇聚接入、地图化浏览、空间计算、服务代理。

- API的生命周期管理：API支持数据服务API、三方API、数据上报API、空间服务API和数据更新API。支持API的创建、编辑和调试，支持的API类型包括实时查询类、离线批量类、流式数据类、非结构化数据类。
- API权限管理：包含已获授权API和已对外授权API。

API管理

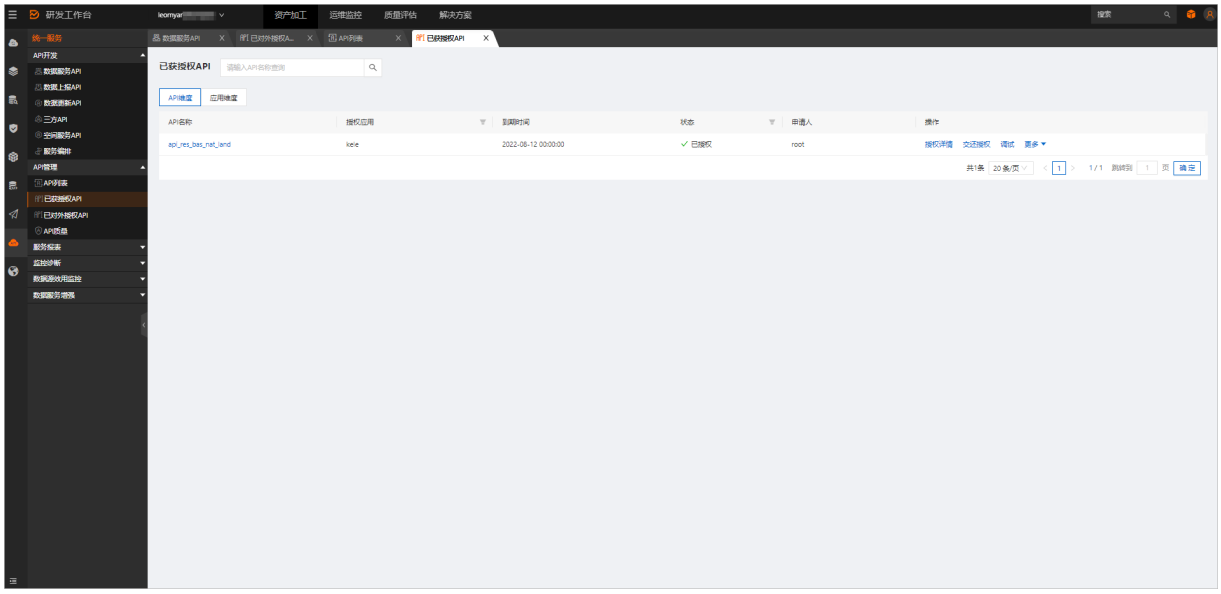
对API做共享管理，支持对已上线的数据API做上架、下架、授权等操作，同时触发审批流程，审批通过后API会在数据资产中心API栏目上架，其他工作组及用户可以申请API使用。API下架时会触发冻结操作，下架后自动收回已对外授权的API的使用权限。

API列表



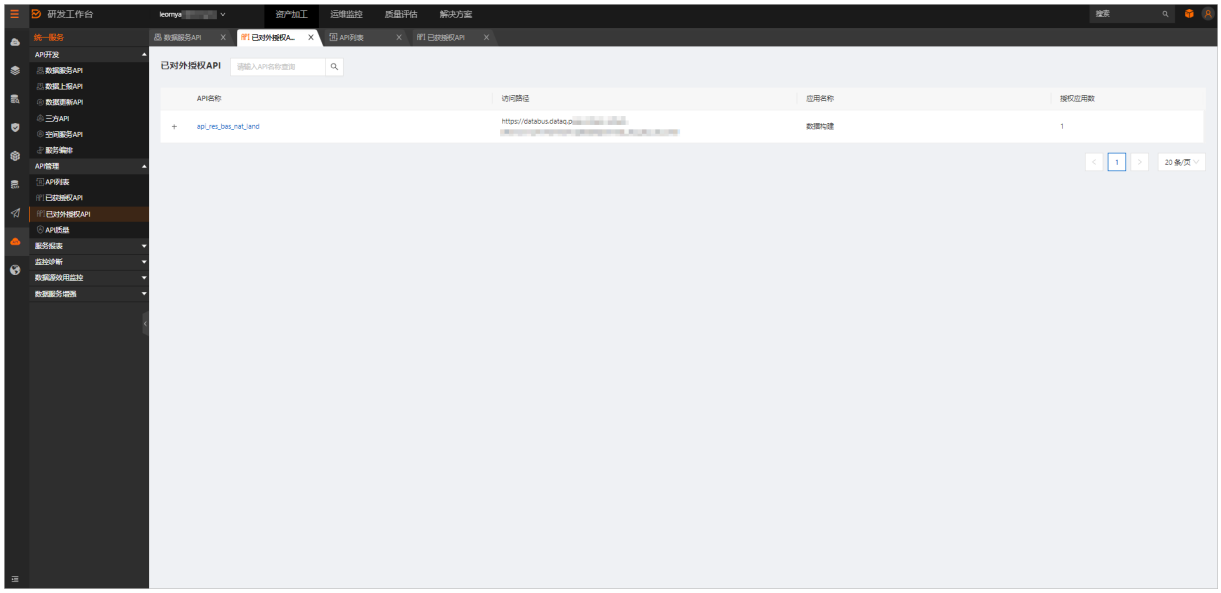
已获授权API

支持对已获授权API列表进行查看，调试API、下载API使用文档、下载密钥SDK和归还权限。支持查看单个API的授权详情，包括：授权时间、访问路径、AppKey和AppSecret。



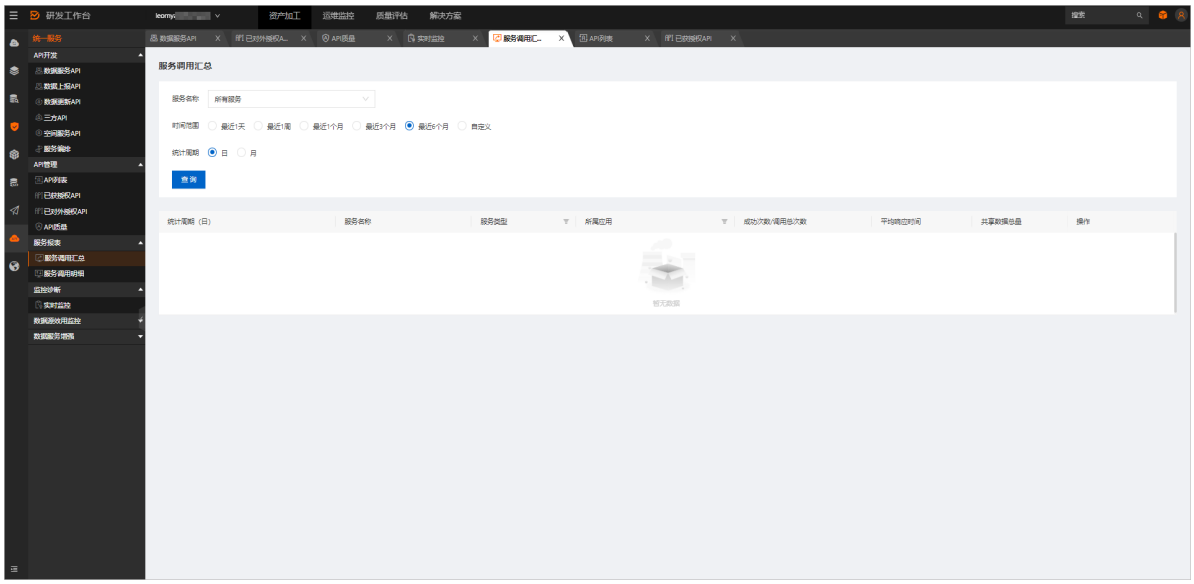
已对外授权API

支持已对外授权API列表的查看，支持查看授权的AK信息、暂停授权、回收授权等操作。

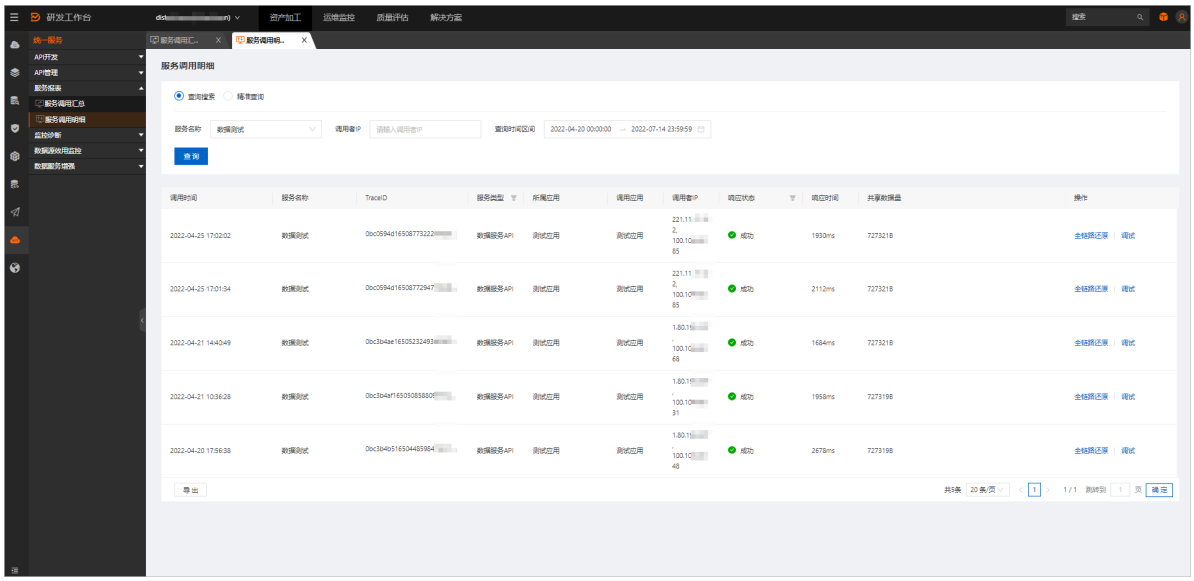


服务报表

- 服务调用汇总支持用户在最近1天、最近1周、最近3个月、最近6个月和自定义等时间范围下，按照15分钟、小时、日、月和天等统计周期统计调用汇总信息，包括总调用次数、平均响应时间、成功调用次数。



- 服务调用明细支持单次调用的明细信息查看，包括调用时间、服务名称、服务类型、响应状态、响应时间等。



实时监控

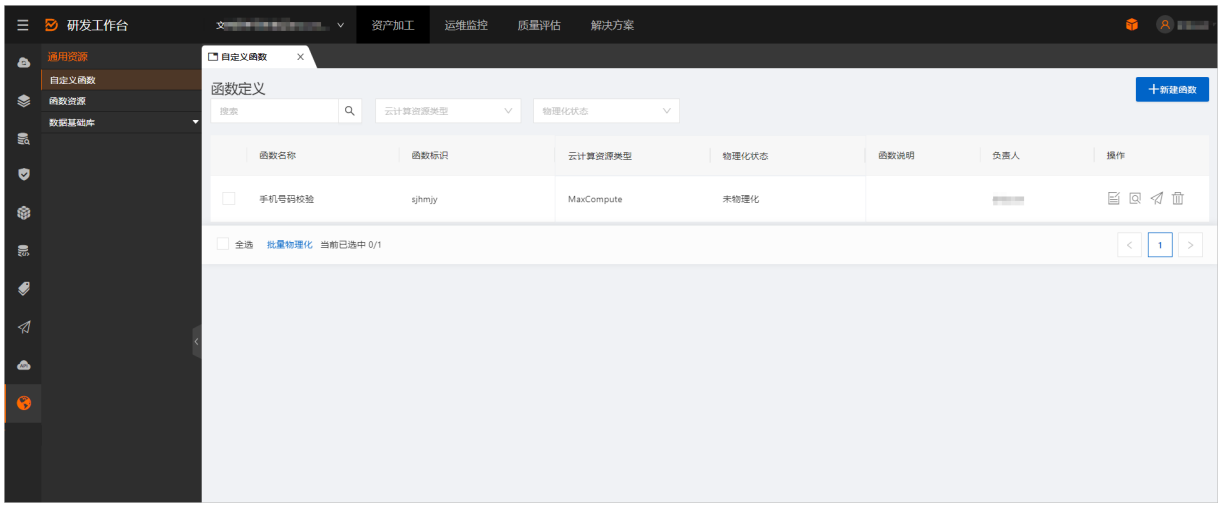
监控最近3小时、12小时和24小时内服务的实时调用趋势及异常趋势。支持查看监控时间点上的服务概览。

6.3.1.10. 通用资源

通用资源为用户提供了自定义函数、函数资源、数据基础库的能力，帮助用户在数据开发等过程中灵活调用各类自定义函数和基础数据。

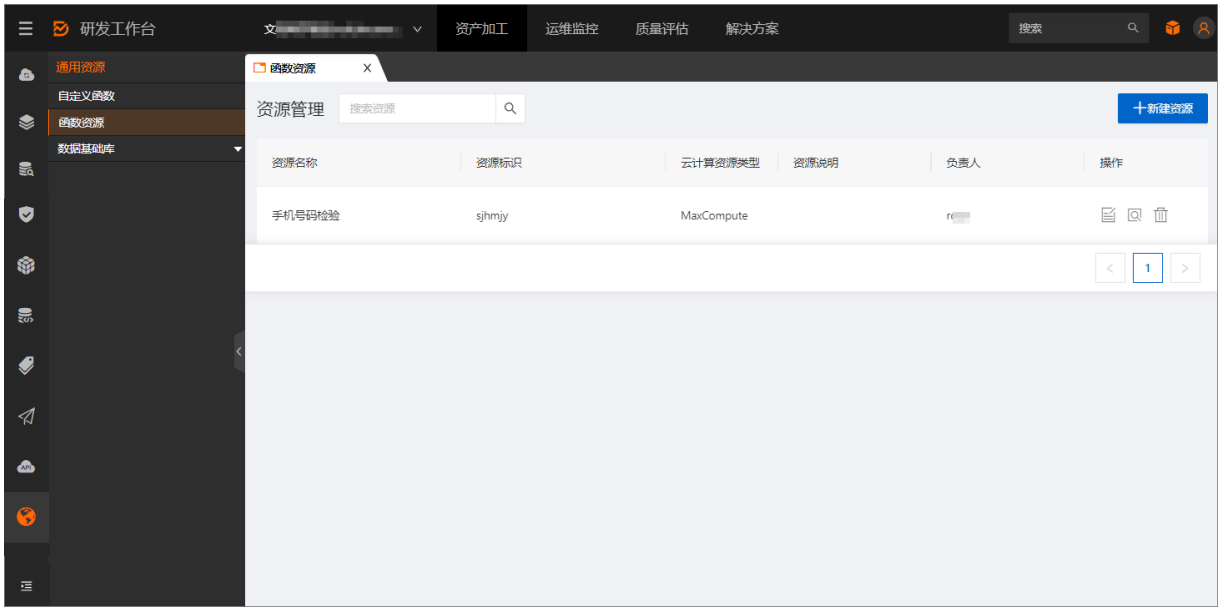
自定义函数

支持对ETL函数自定义和管理，并与公共计算节点相关联，便于快速对数据进行ETL处理。



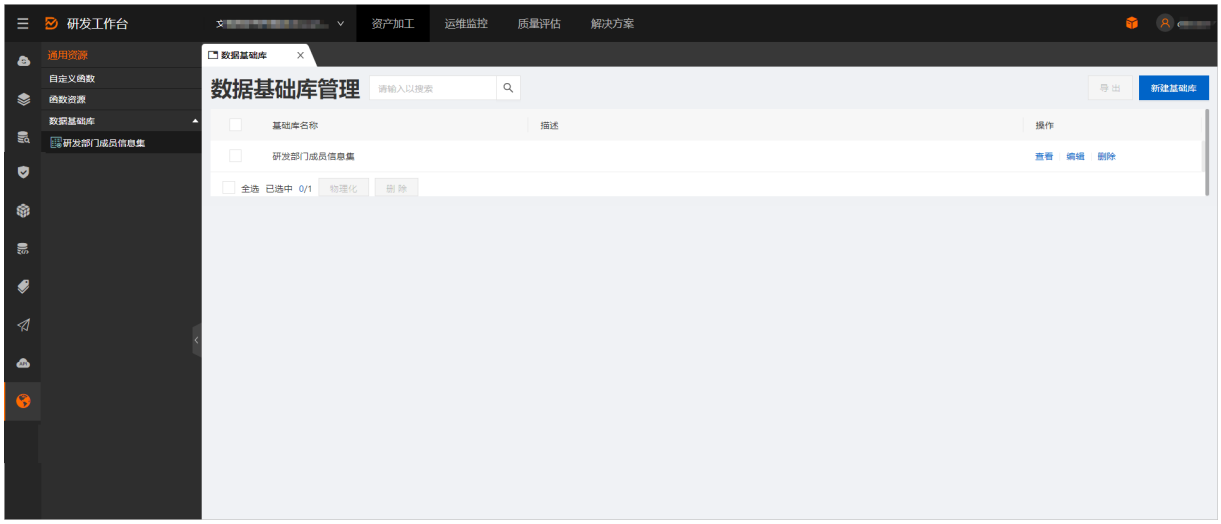
函数资源

支持对自定义函数中所使用的代码包的新建和管理。



数据基础库

支持标准化基础数据的管理维护，以便在模型加工过程中使用行业通用化标准化数据，进行后续数据加工和内容参照。



6.3.1.11. 运维监控

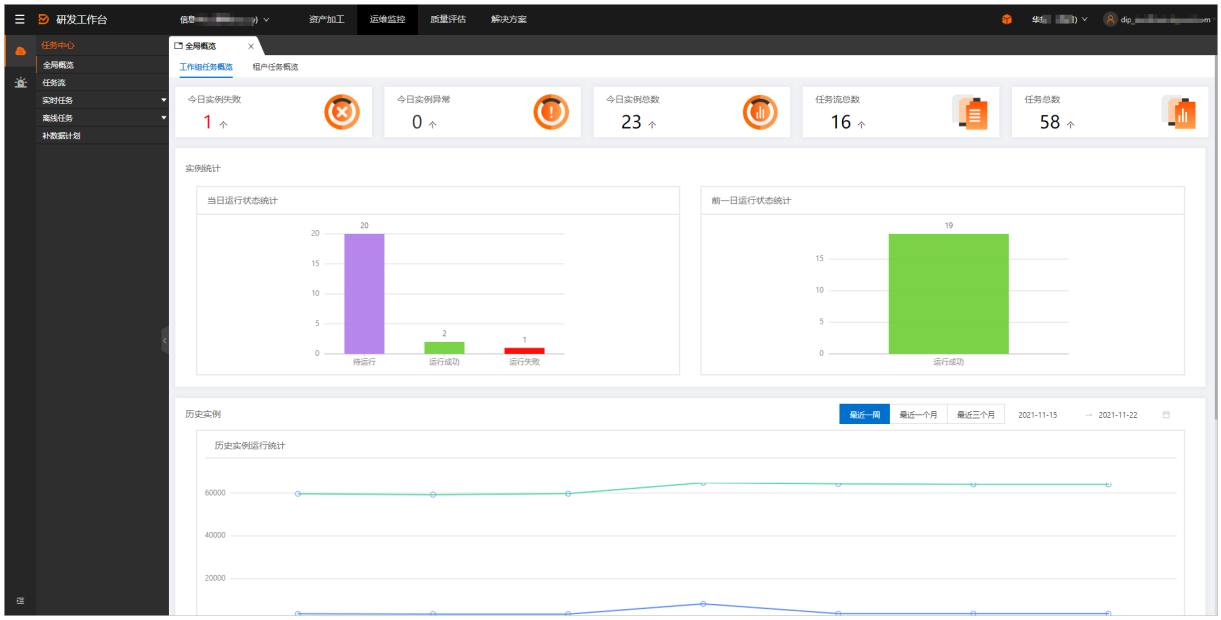
6.3.1.11.1. 任务中心

任务中心对数据资源平台中的数据同步、数据开发、标签加工等任务和实例进行运维管理。

数据资源平台各模块发起的任务，如数据同步、数据开发、标签加工等，都会发到对应的任务流，不同的任务来源通过业务类型区分。根据业务流程，任务流中包含了不同类型的节点，如MaxCompute离线任务、Flink流式任务等。不同类型的节点会生成各类任务，主要包括离线同步任务、离线计算任务、实时同步任务、实时计算任务等。对于离线同步任务、离线计算任务，每次运行时会生成该次运行周期的实例。

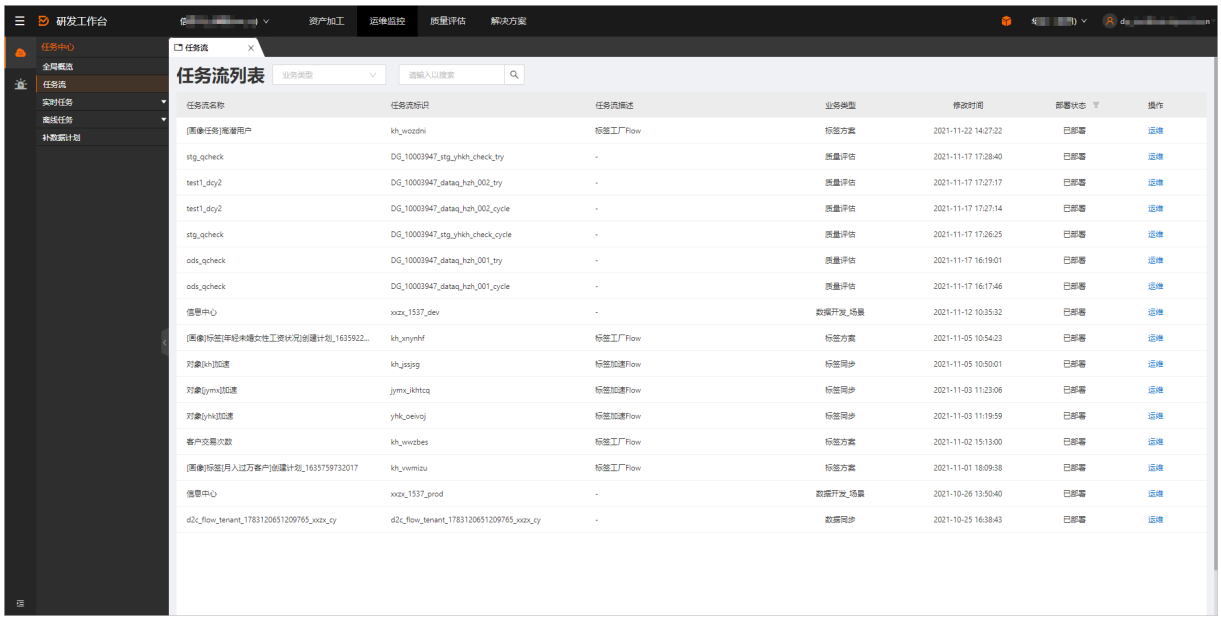
全局概览

支持查看任务流、任务、任务实例的统计信息，并提供以图形化的方式展示历史实例运行信息。



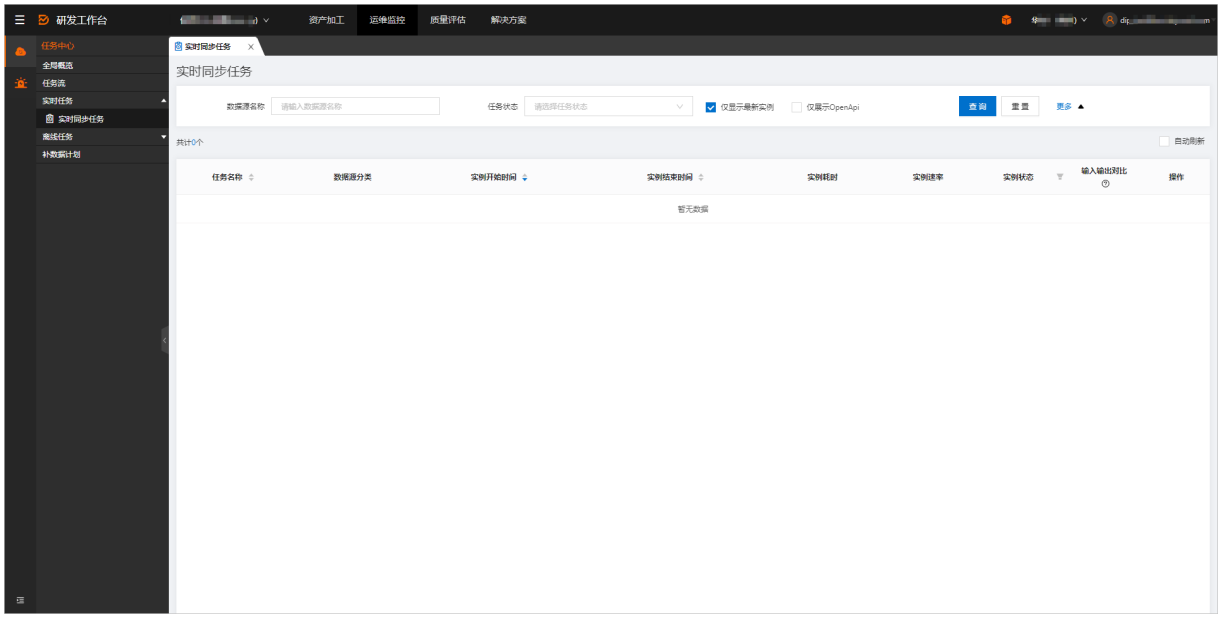
任务流

支持查看任务流的属性信息，包括任务流中包含的任务节点、任务节点信息、任务节点的最新实例等。支持查看任务流的部署、运行日志信息。



实时任务

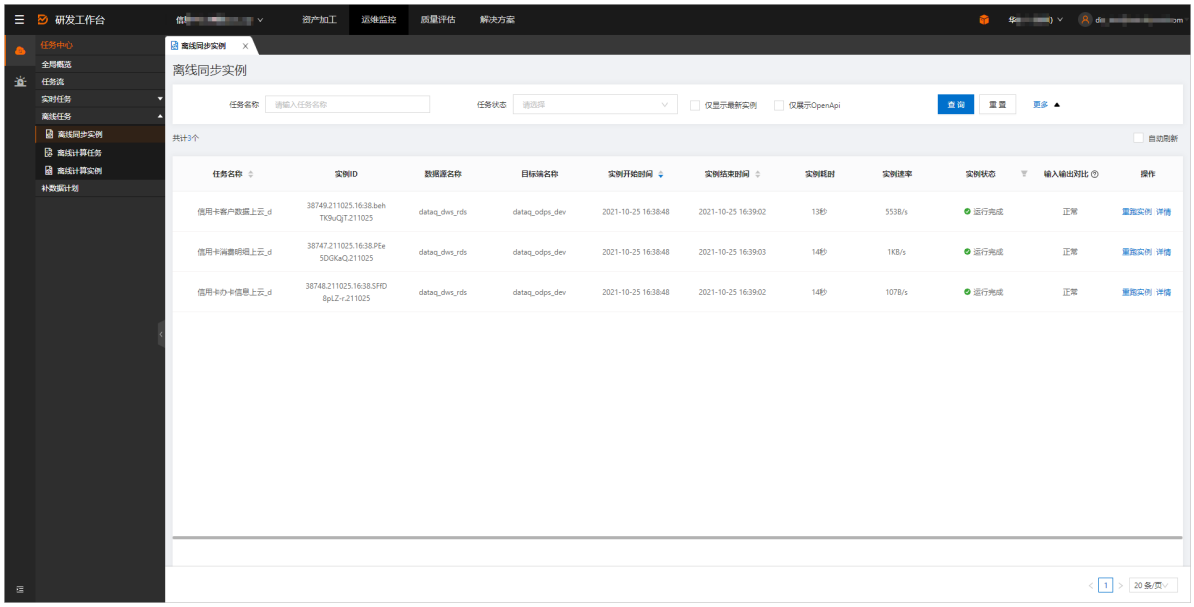
支持以列表的方式展示实时任务信息，并提供当前实例属性与运行日志信息。



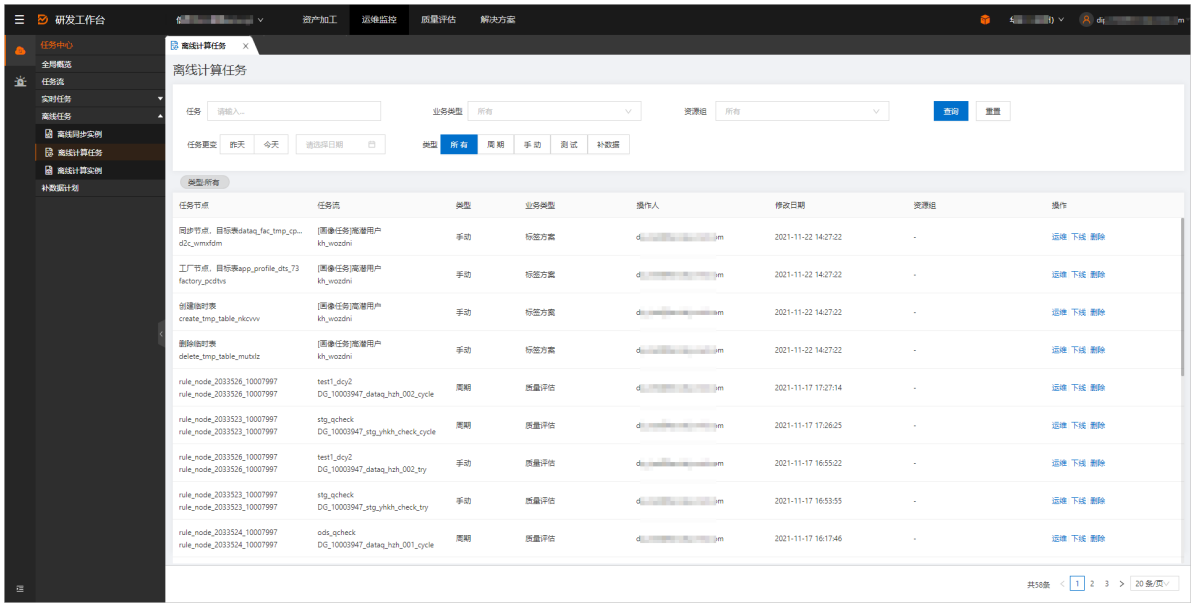
离线任务

支持查看离线计算任务的属性信息，查看离线同步实例、离线计算实例的属性与运行日志。

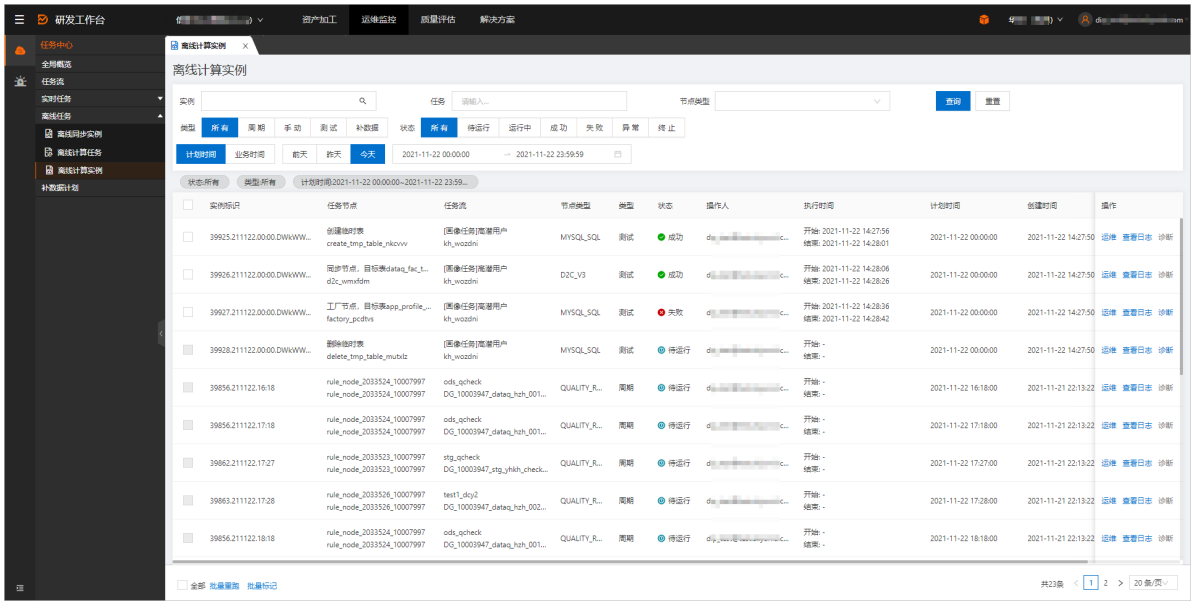
- **离线同步实例**：支持以列表的方式展示离线同步实例信息，并提供离线同步实例属性与运行日志查看；可进行实例重跑的功能。



- **离线计算任务**：支持以列表的方式展示离线计算任务信息，并提供运维、下线、删除的功能。



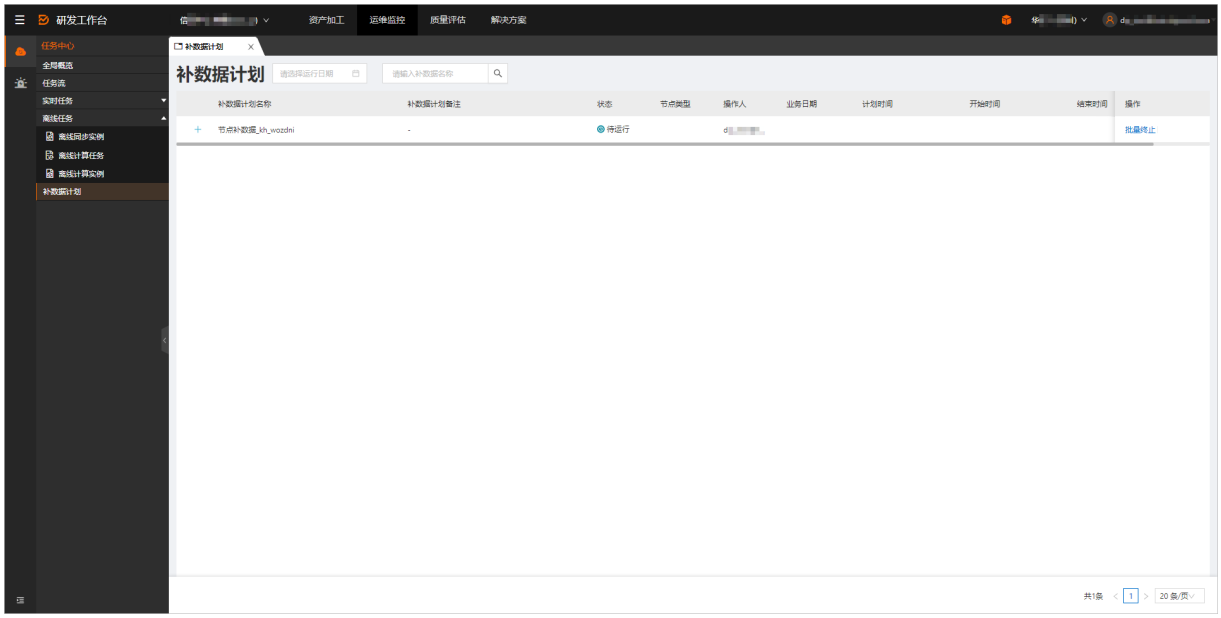
- **离线计算实例**：支持离线计算实例信息，并提供运维与日志的查看。



补数据计划

支持任务节点的补数据操作。

- **新建补数据计划**：支持在任务流和离线计算任务中补数据。
- **终止补数据计划**：支持补数据计划运行后进行终止操作。
- **查看补数据计划**：支持查看补数据的运行状态、节点类型、业务日期，以及查看对应节点的运维信息，包括实例信息、任务代码、实例日志。



6.3.1.11.2. 全景监控

全景监控支持业务链路配置监控场景，通过自定义监控链路和告警消息配置，可快速实现从数据生产到业务应用的全链路编排及溯源，实现全局监控和问题快速锁定。

全景监控提供丰富的告警信息输出方式，包括：钉钉消息、标准协议等方式，出现问题时按照问题的紧急程度，选择告警方式，快速通知相关人员，及时处理问题，保障系统和数据的稳定运行。同时，全景监控提供统一的监控接入标准，全方位对接各种应用系统监控信息，实现覆盖所有系统、全流程的监控管理。

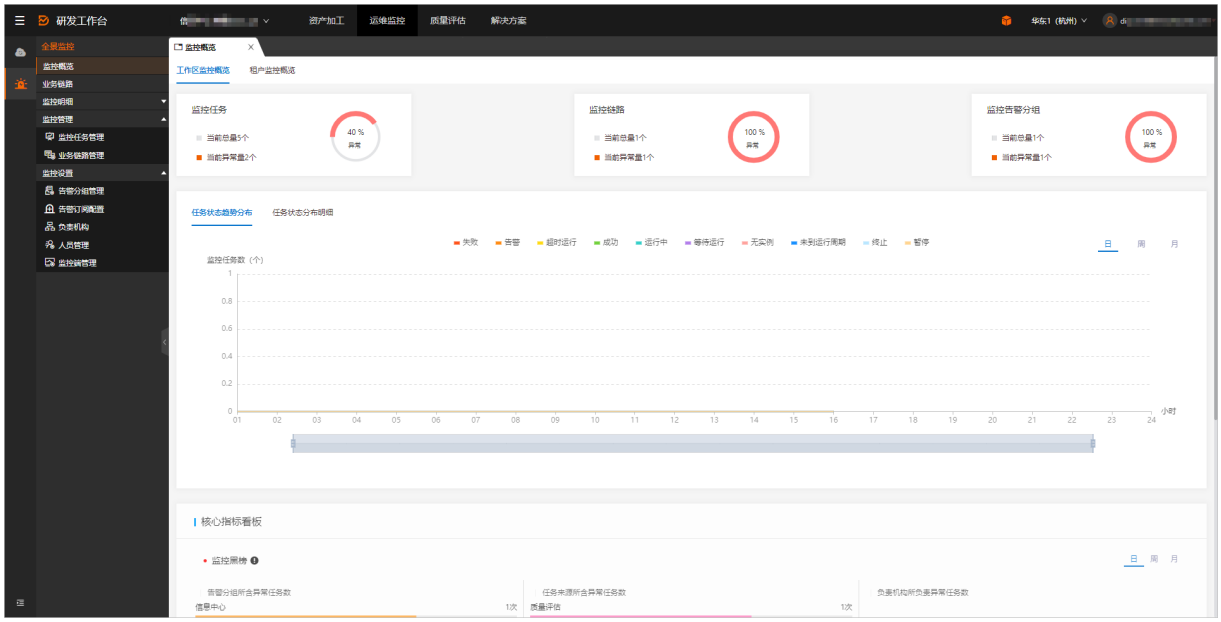
数据全链路保障是指通过采集数据汇聚、数据融合、数据分析挖掘等阶段的各项任务状态信息，对异常状态进行预警和处置，实现对各任务的实时监控和管理。

从应用视角提供平台内跨系统的从数据生产到业务应用的全链路监控能力，其提供的监控端可将多系统的状态信息进行上报汇总和集中推送。提供自定义监控链路配置能力，快速实现从数据生产到数据应用的全链路编排及溯源和监控节点的稳定性分析。例如，当指标出问题时，加工这个指标的全局链路上，涉及系统的接入、数据的加工、API的转换、系统中的算法加工，到最终指标的呈现，到底哪个链路出问题需要立即定位。系统需向各方应用提供统一接入标准，实现全局监控和问题快速锁定。

通过业务场景组织监控链路，包含业务场景名称、业务场景描述、负责人，以及业务场景的正常/异常状态及异常告警个数等统计指标，方便运维人员从全局快速锁定出现问题的业务场景及严重程度。业务场景中包含该场景涉及的业务节点、任务节点，以及这些节点在逻辑上的上下游关系，运维人员可通过上下游关系清晰的了解各任务间的影响和依赖关系。

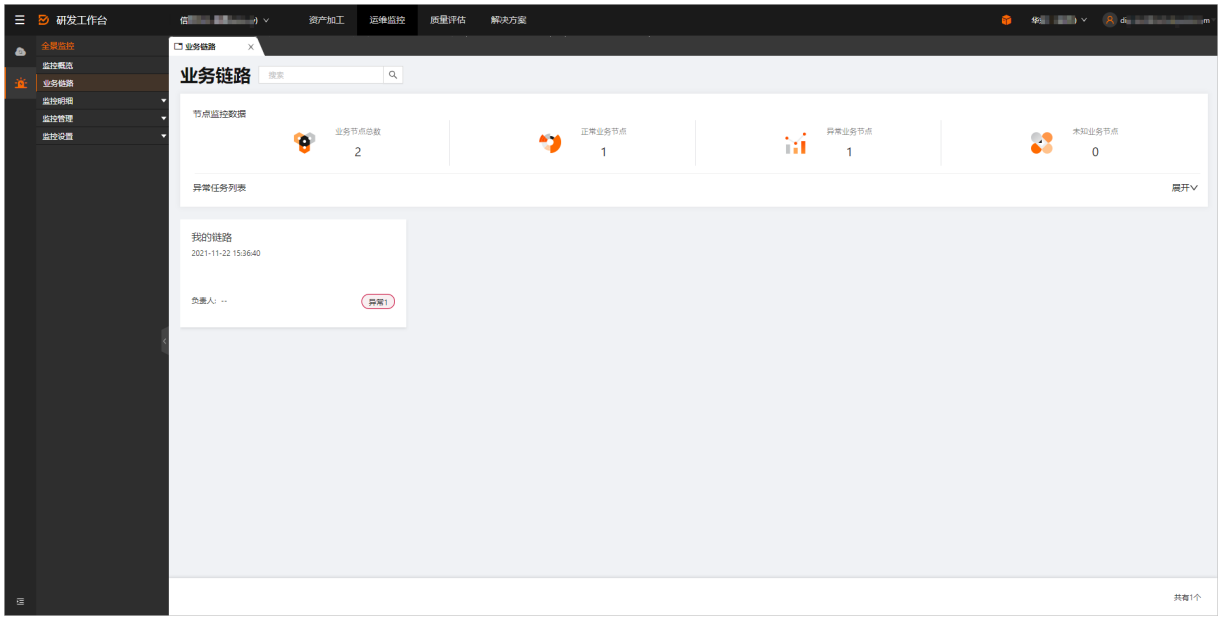
监控概览

提供监控任务、监控链路、监控告警分组等统计信息概览。



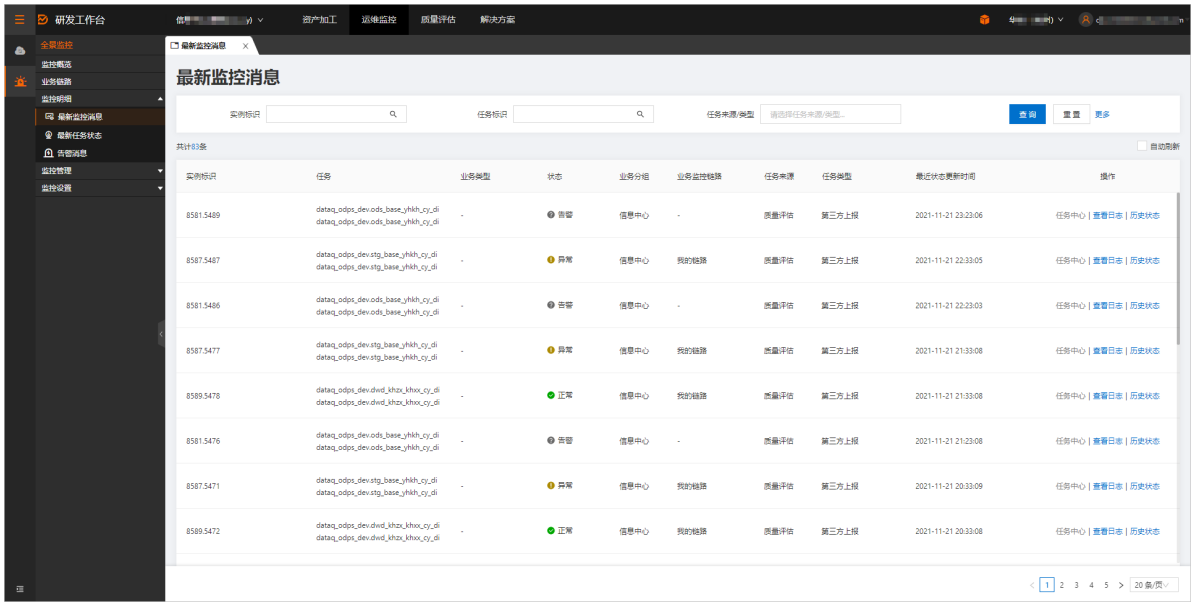
业务链路

支持以业务数据链路为视角进行监控，可查看当前业务链路中节点的运行情况、告警信息。

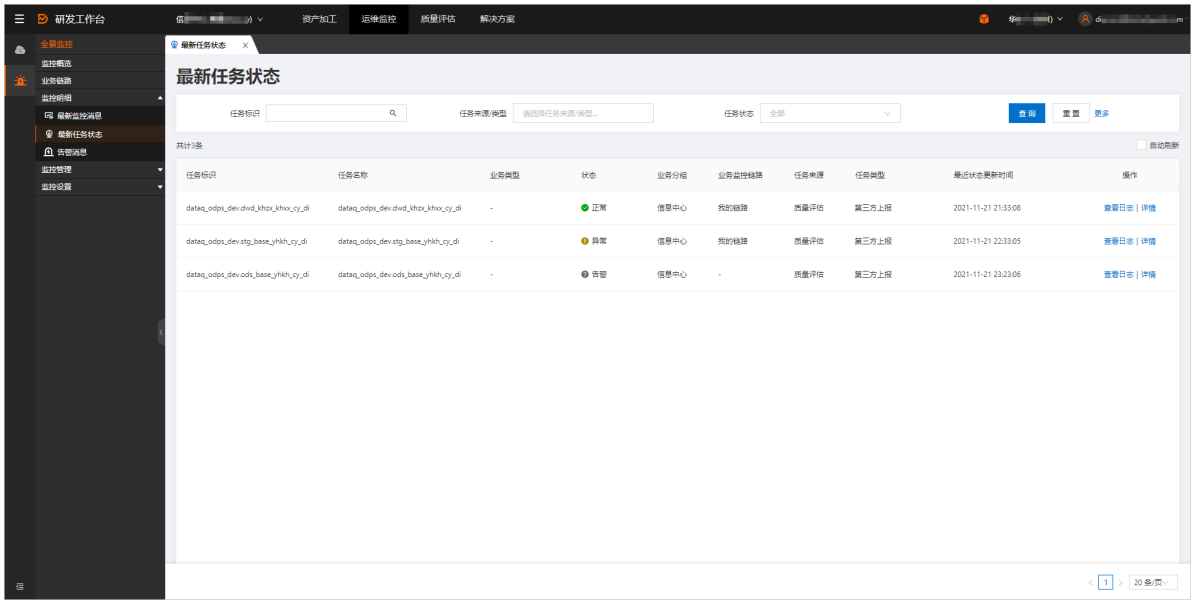


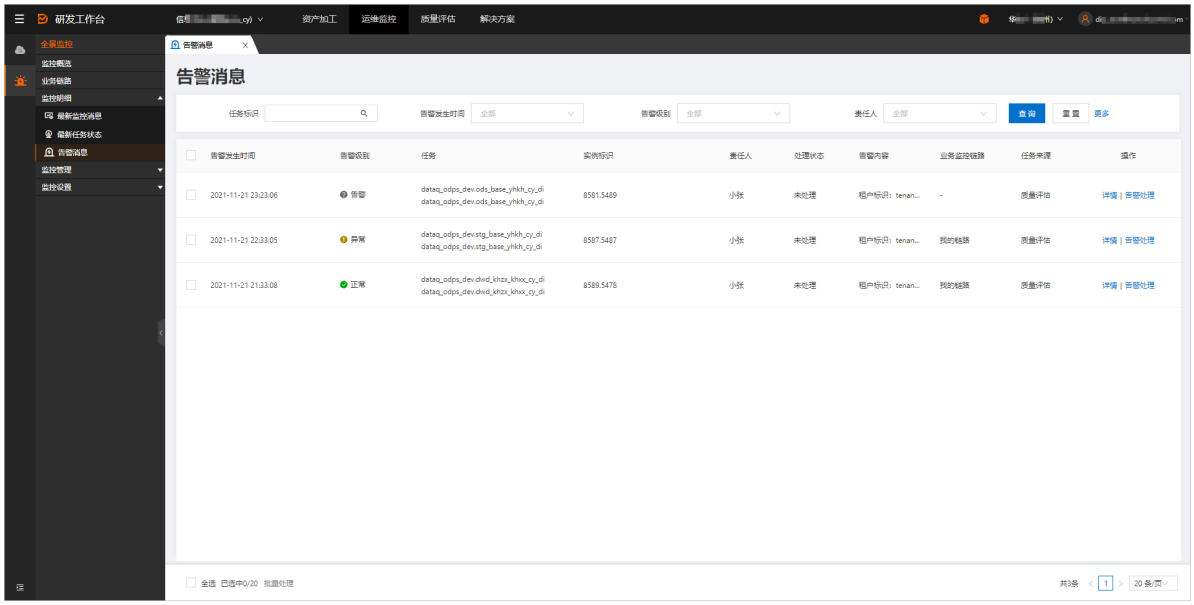
监控明细

- 支持查看任务节点监控结果，可细节掌握每个监控节点的最新状态、任务日志和历史状态分布情况。



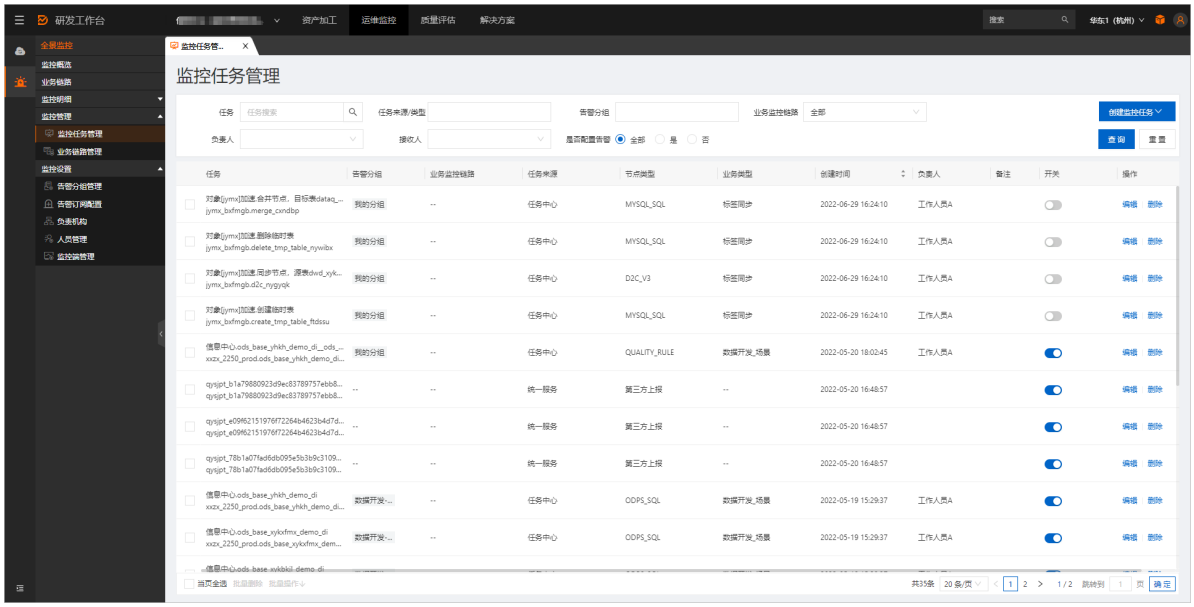
- 支持查看监控任务节点的最近一次告警信息及历史告警信息，并支持手动设置告警处理暂停时间状态。





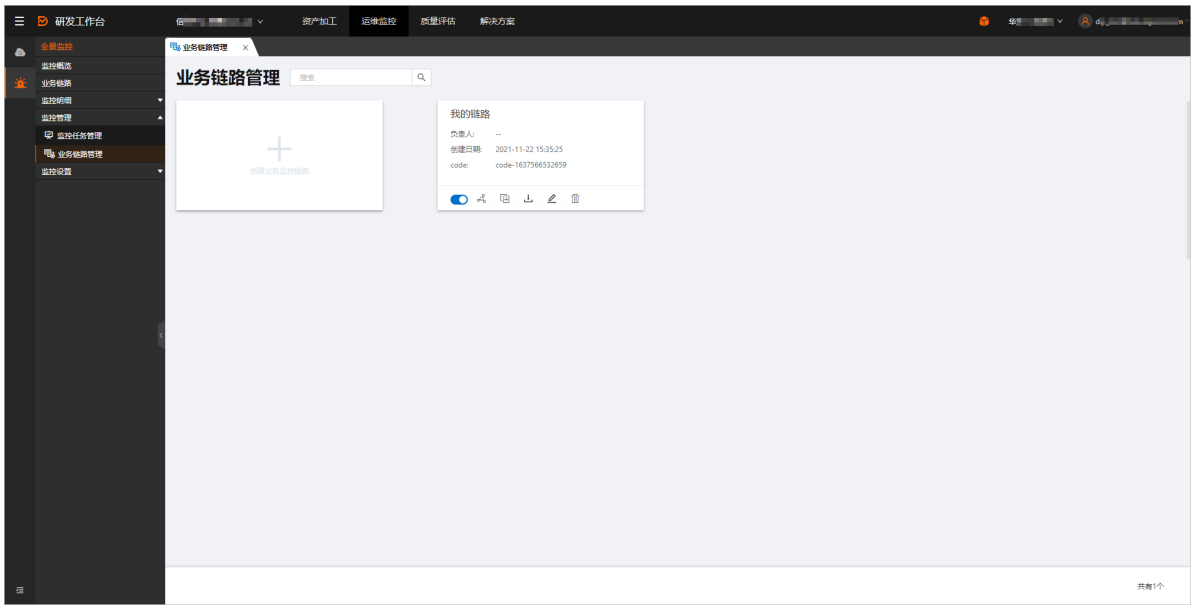
监控管理

- 监控任务管理



- 支持监控任务管理，针对不同来源的任务进行统一监控任务的配置管理。
- 支持通过批量创建、手动创建的方式完成监控任务创建。

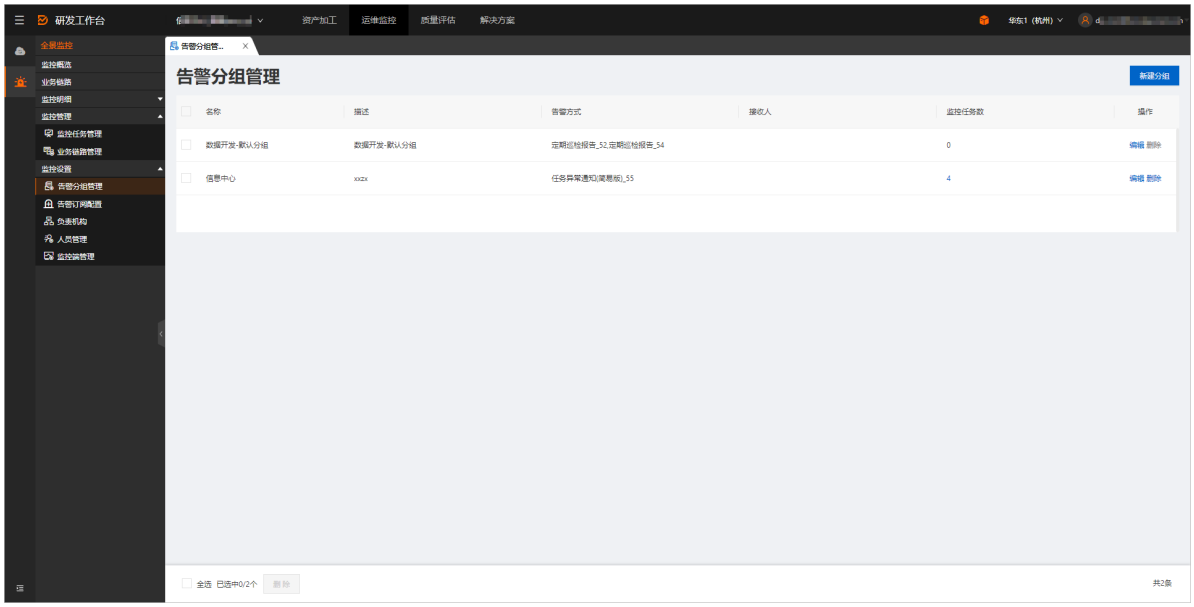
- 业务链路管理

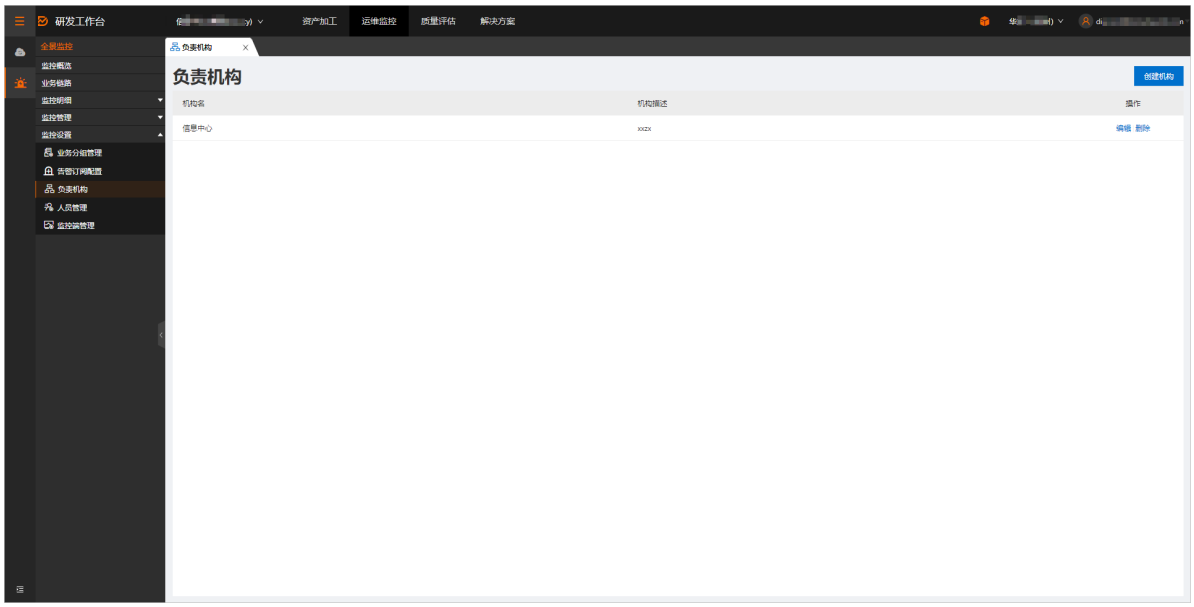


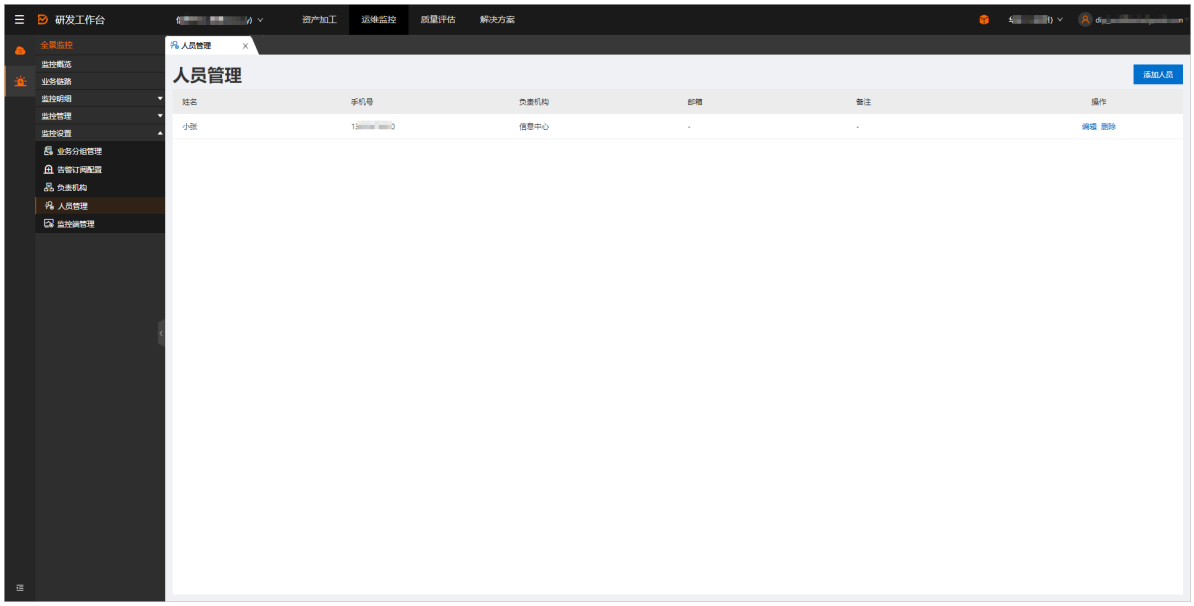
- 支持业务链路管理，按照业务视角自定义配置业务监控链路，可快速开启或关闭业务链路。
- 支持业务链路的导入、导出。

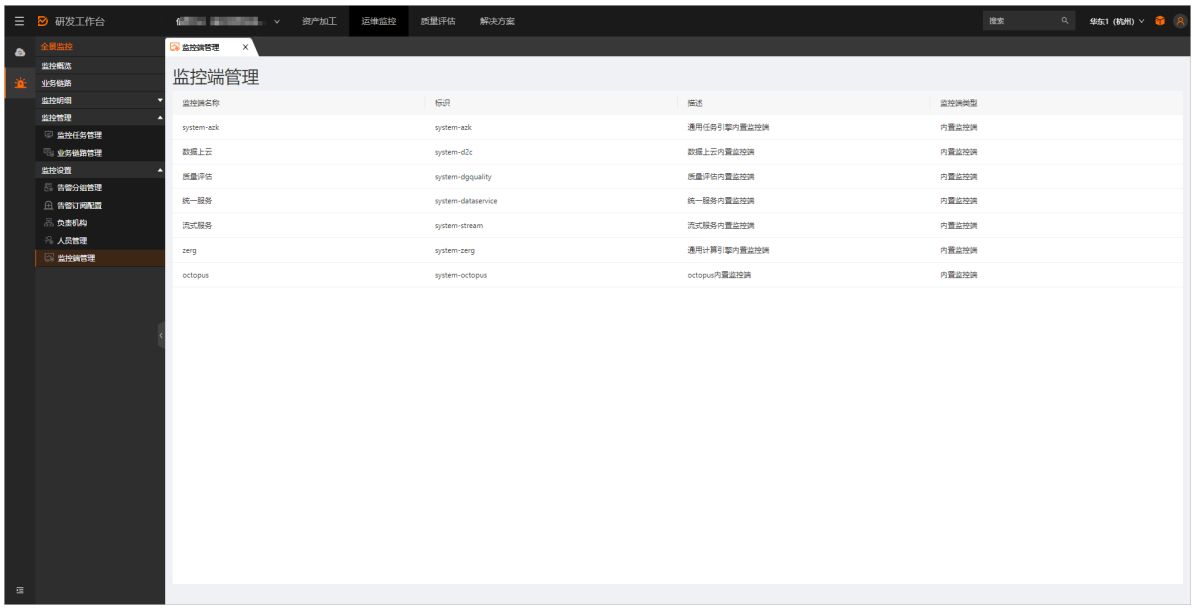
监控设置

- 支持告警分组管理、负责机构、人员管理、监控端管理。

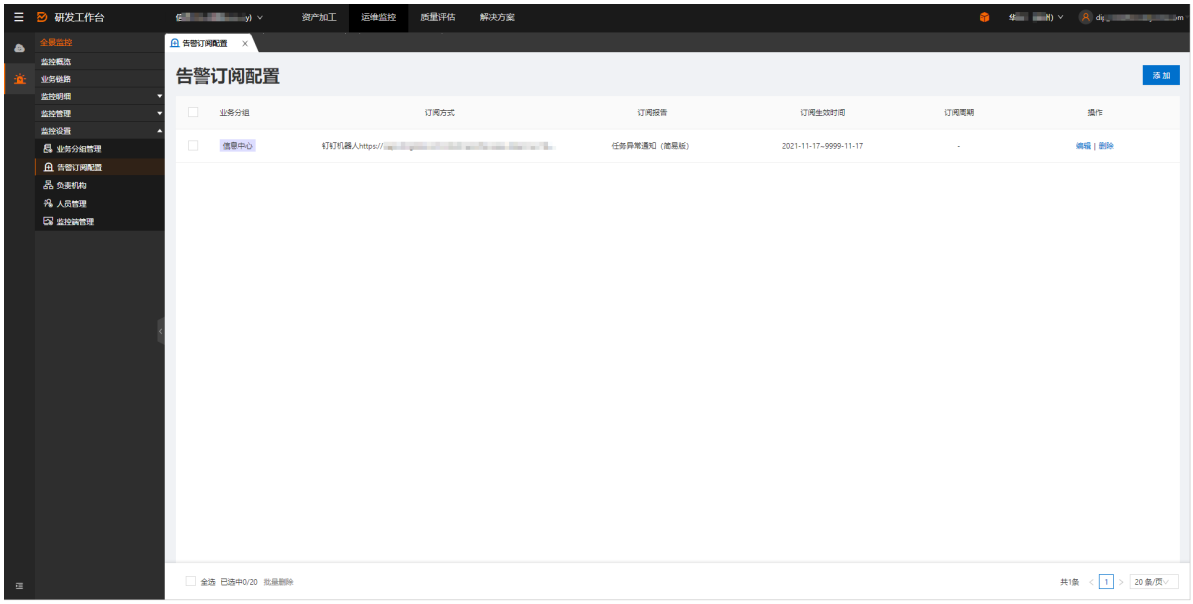





研发工作台
资产加工
运维监控
质量评估
解决方案
人员管理
添加人员
姓名
手机号
负责机构
邮箱
备注
操作
小张
信息中心
编辑 删除


研发工作台
资产加工
运维监控
质量评估
解决方案
监控端管理
监控端名称
标识
描述
监控端类型
system-ack
system-d2c
system-dqquality
system-dataservice
system-stream
system-zerg
system-octopus
内置监控端

- 支持通过钉钉机器人、系统内置标准协议等方式订阅告警消息。



6.3.1.12. 质量评估

质量评估通过用户自定义质量检查计划，根据用户设置的数据质量规则执行面向不同计算资源的质量检查任务，通过数据质量监控报告展现系统整体数据质量概览及多维度细分数据的质量情况。

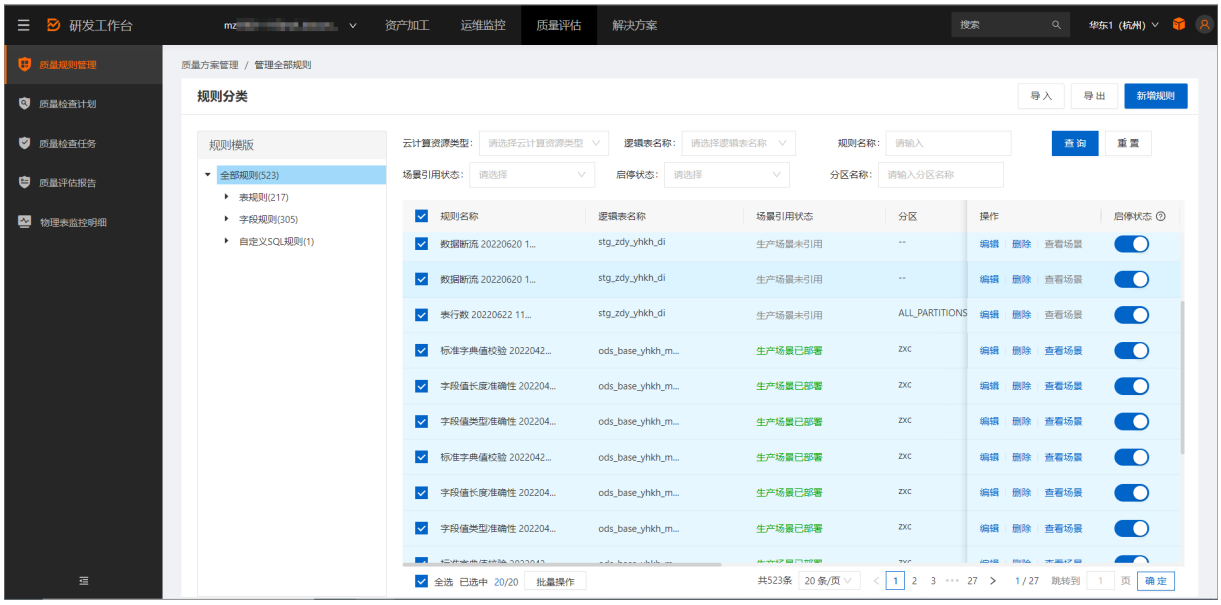
质量评估通过完整性、准确性、一致性、唯一性、时效性、规范性等六个维度构建数据质量指标体系，来实现对数据质量的量化评估。

- 完整性：数据完整性是指一个数据集的特定字段或属性都被赋予了数值。
- 准确性：数据准确性是指数据准确的反映其所建模的“真实世界”的实体。
- 一致性：数据一致性是指确保两个数据集之间或一个数据集内部不同字段间数值或信息的一致。
- 唯一性：数据唯一性是指在一个数据集中，没有重复的记录或信息。
- 时效性：数据时效性是指数据集按时、按量的更新程度。
- 规范性：数据规范性是指数据集对数据标准的满足程度。

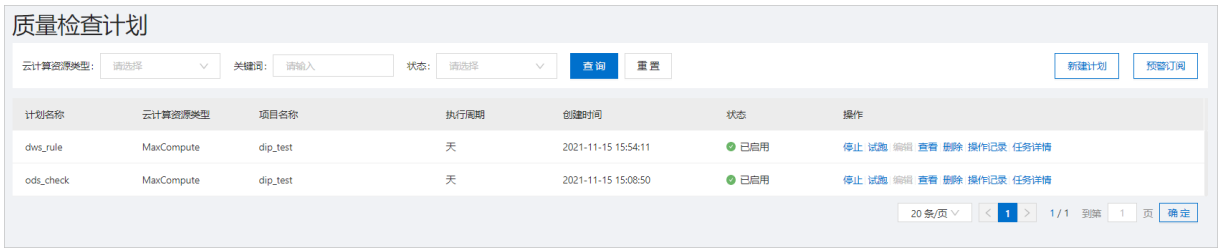
质量规则管理

质量规则管理页面，查看数仓中配置的全部规则，包括表规则、字段规则、自定义规则。

- 支持对所有的物理表统一配置的规则。
- 支持根据资源类型、逻辑表名称、规则名称、启停状态对规则进行筛选和查询。
- 支持对规则的删除、启停操作，删除和启停后规则的状态需要在数仓同步进行更新。
- 支持对全部规则、和选中规则进行导入和导出。
- 支持一键更新生产场景中的质量规则。

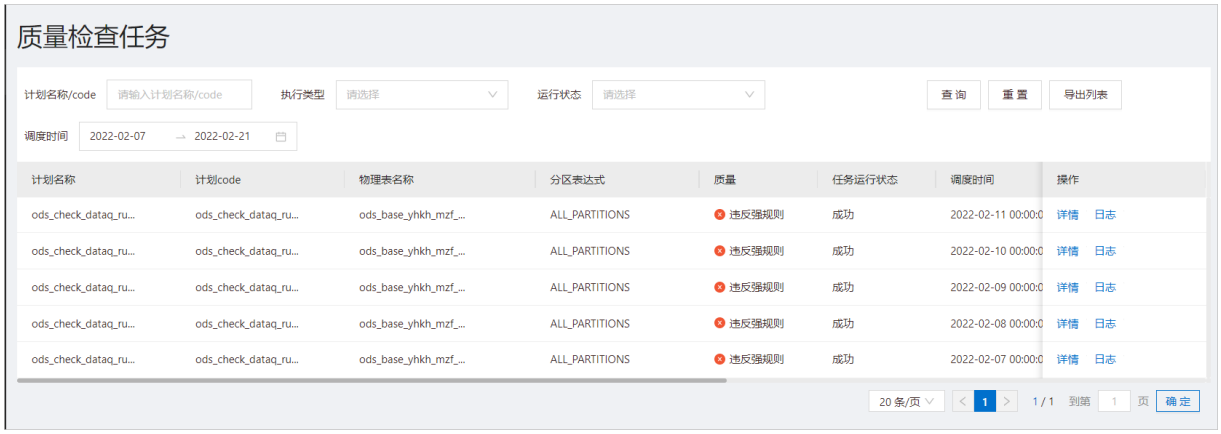


质量检测计划管理



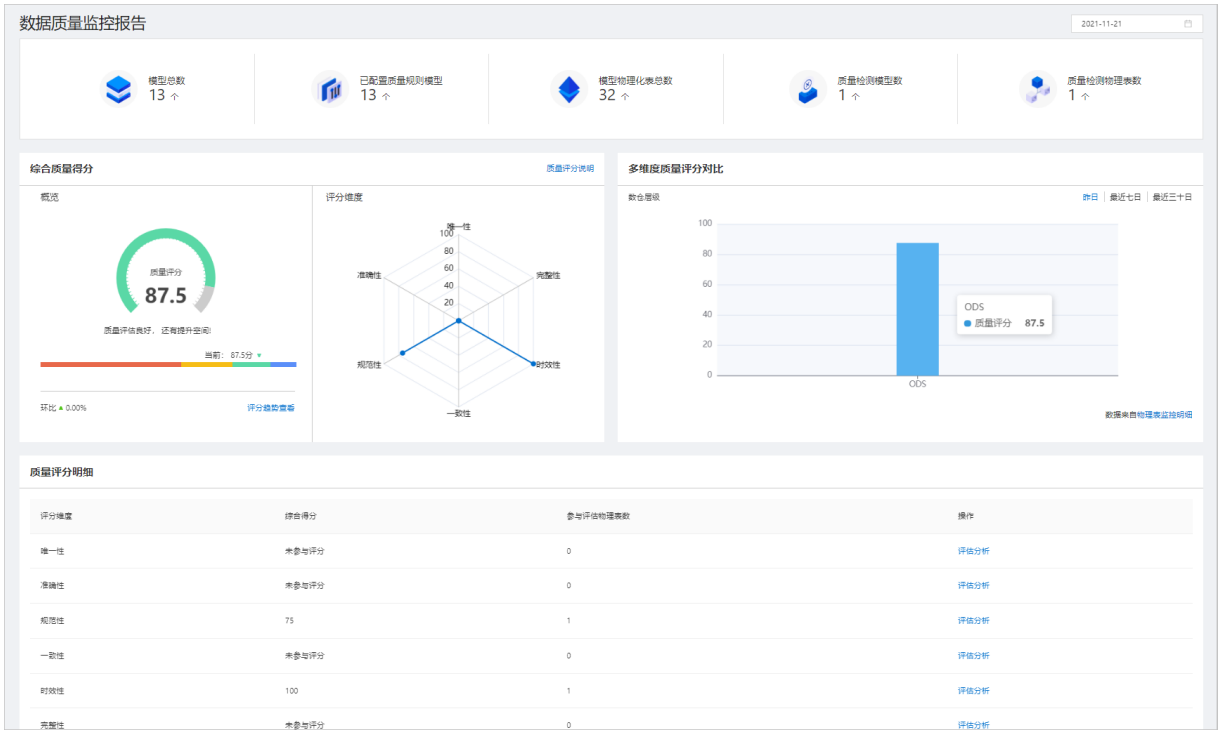
- 支持对质量检测计划的创建，可批量选择需要周期执行且已配置好质量规则的数据模型，配置周期检测任务和调度，进行持续质量检测。
- 支持对计划任务的管理、启动、停止和试跑。
- 支持对物理表质量结果进行告警订阅，在全景监控模块进行预警（依赖全景监控）。

质量检测任务管理



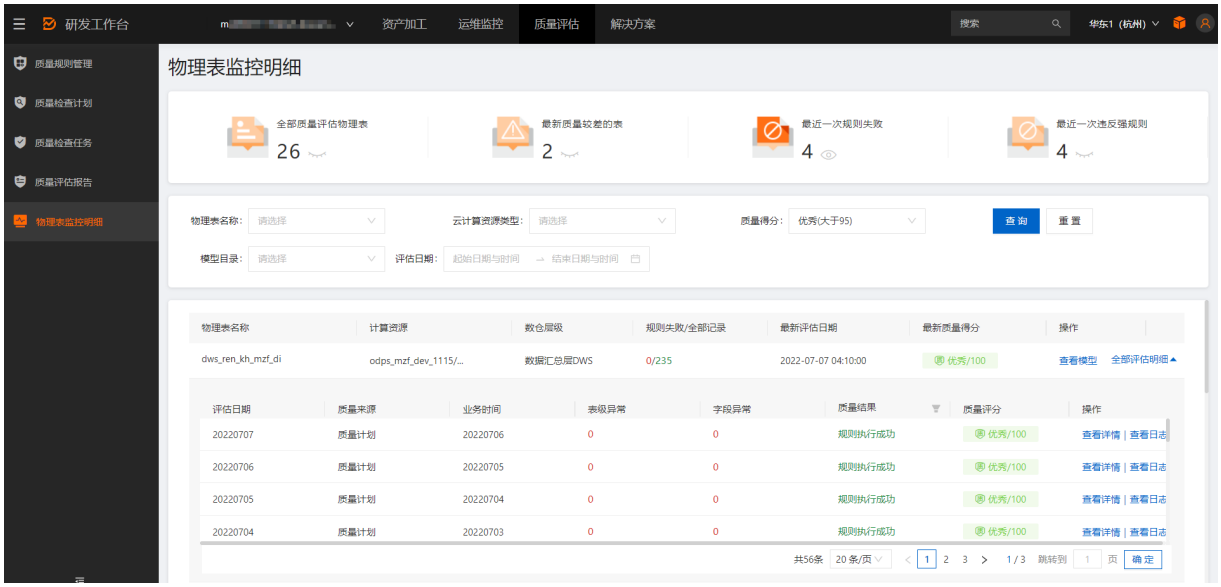
- 支持对质量检测任务运行实例进行查询和管理。
- 支持查看任务实例运行结果详情、运行日志进行查看。
- 支持对质量结果任务运行实例明细进行导出。

质量评估报告



- 支持根据系统数据质量监测运行结果以天的维度自动生成系统整体质量监控报告。
- 支持根据数据唯一性、准确性、规范性、一致性、时效性和完整性进行多维度评估分析。

物理表监控明细



- 支持以物理表为维度，查看物理表在数据开发和质量计划检测中最新的数据质量结果。
- 支持对物理表历史质量实例运行记录进行查看。

6.3.1.13. 解决方案

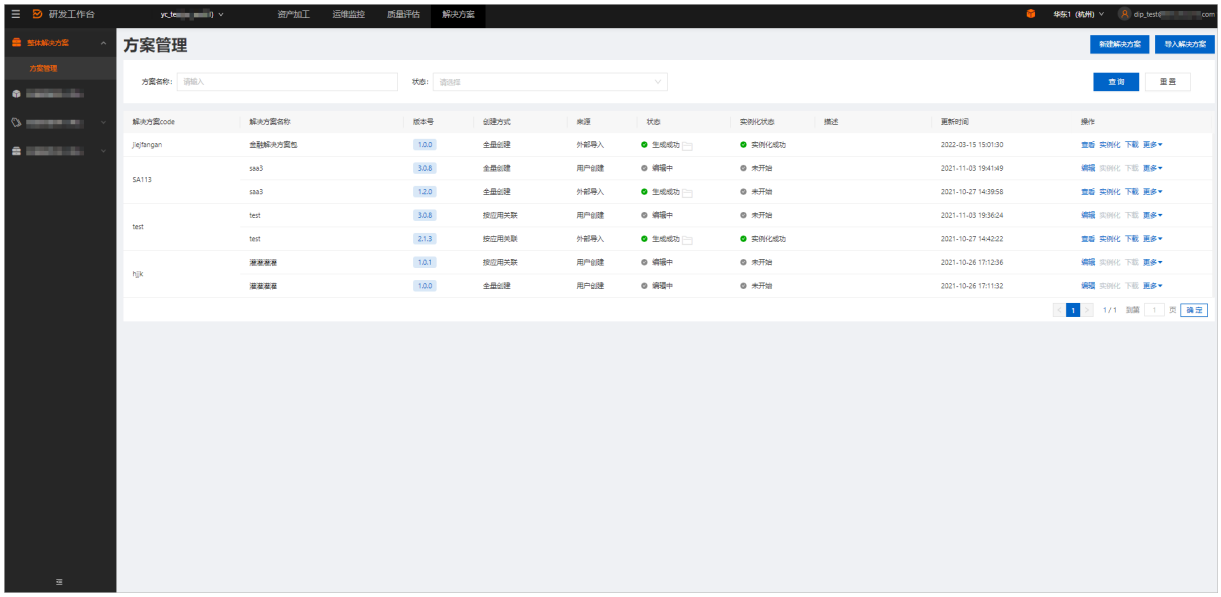
数据资源平台中的数据标准、数据模型、数据加工场景、数据服务API、云计算资源配置等资产，经过验证最终可沉淀为一个特定的解决方案。支持通过不同的创建方式，灵活的选择解决方案中包含的内容。依托数据资源平台中管理的大量元数据，确保解决方案的完整性、一致性和可用性。

根据数据血缘关系和不同的创建方式，自动选择关联的数据资产，保证解决方案的完整性。 同时，通过自动检测待生成的解决方案中是否存在异常（如表一致性、数据质量异常等），保证解决方案的一致性和可用性。对于需要在多项目部署的解决方案，支持以文件形式导出解决方案，并在项目中通过导入到本地的方式，快速实现解决方案部署、发布，以完成数据治理工作环境的初始化、跨平台的资产发布及迁移。

解决方案帮助用户沉淀行业知识，快速完成行业支持的复制，提升数据资产上线的效率和质量，加速数据业务价值的转化。

整体导入导出

将工作组中已沉淀的数据标准、数据模型、数据加工场景、数据服务API配置、云计算资源配置等数据资产创建解决方案后，进行下载、导入以及实例化。



- 支持将工作组内已加工完成的数据标准、数据模型、数据加工场景、数据服务API配置等信息，进行版本化解决方案封装。
- 支持全量、按范围关联两种方式灵活创建解决方案，并可对解决方案内容的逻辑一致性、物理一致性进行检查。
- 支持将解决方案包下载后，导入并实例化至工作组，完成数据治理工作环境的初始化、跨平台的资产发布及迁移。

6.3.2. 资产运营工作台

运营工作台主要面向租户资产管理员，为用户提供数据资产编目、数据资产上线、数据资产下线等能力。通过获取资产元数据信息，并同步到数据资产目录完成数据资产注册。

数据资产管理涉及三种用户角色，包括资产使用方、租户资产管理员、工作组管理员。

- 资产使用方：主要进行数据查找和使用申请，目标是找到需要使用的数据资产，进行权限获取以完成后续业务使用。
- 资产管理员：对所有可对外服务的高价值数据资产进行管理维护与编目上线，进行资产对外服务的生命周期维护，同时针对申请使用进行过程审核。
- 工作组管理员：对该角色负责管理的工作组下数据资产进行使用侧的管理审核。

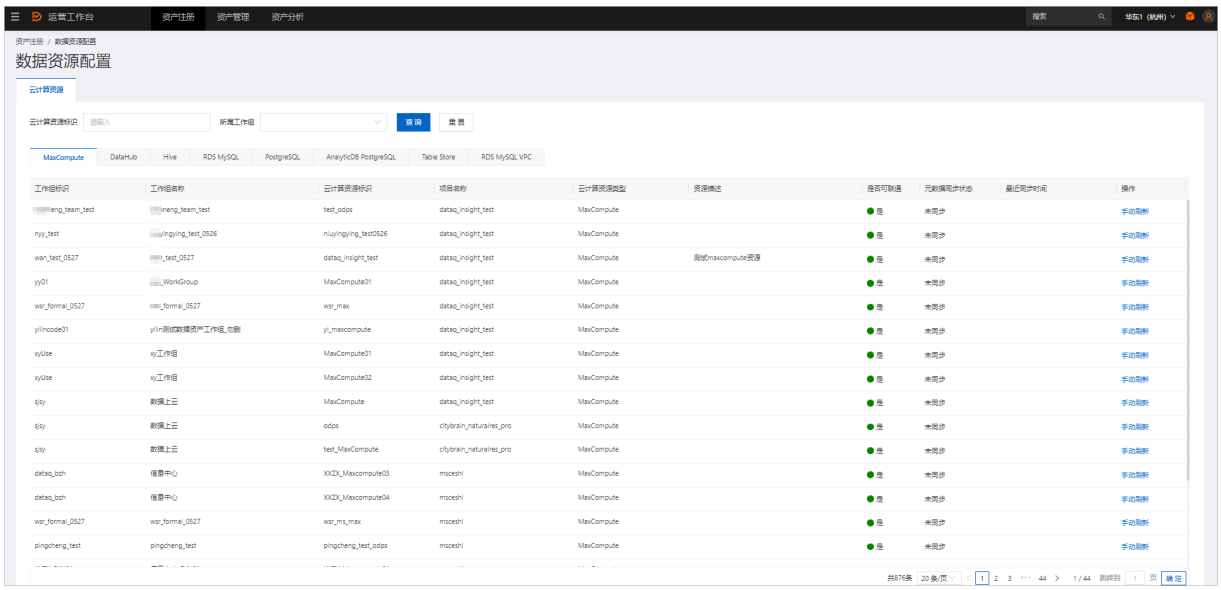
数据资产管理体系主要包括以下几个方面的内容。

- 数据资产编目，数据管理者对平台数据资产进行梳理，确定其共享属性，并按照技术标准，进行数据类目定义。
- 数据资产上线，将已经注册管理的高质量数据资产资源发布，供数据使用者查询和使用。
- 数据资产申请，数据使用者向数据管理方提出使用申请，说明申请方式和申请理由。由租户资产管理员进行资产使用的审核处理。
- 数据资产服务，根据数据资源的性质和特点，选择采用数据表、接口方式提供资产服务。
- 数据资产下线，根据数据的实际情况不再共享数据资产。

资产注册

资产注册功能是数据资源平台资产管理的基础，可快速锁定待管理数据资产的范围。针对支持管理的数据表类型，实现资产信息注册自动化。在工作组中注册存储计算资源后，系统自动执行元数据信息扫描，并完成数据资产发现，为后续资产运营管理持续提供动态更新的资产对象。同时自动完成已上架API资产上报，作为资产管理中待编目维护的数据内容。

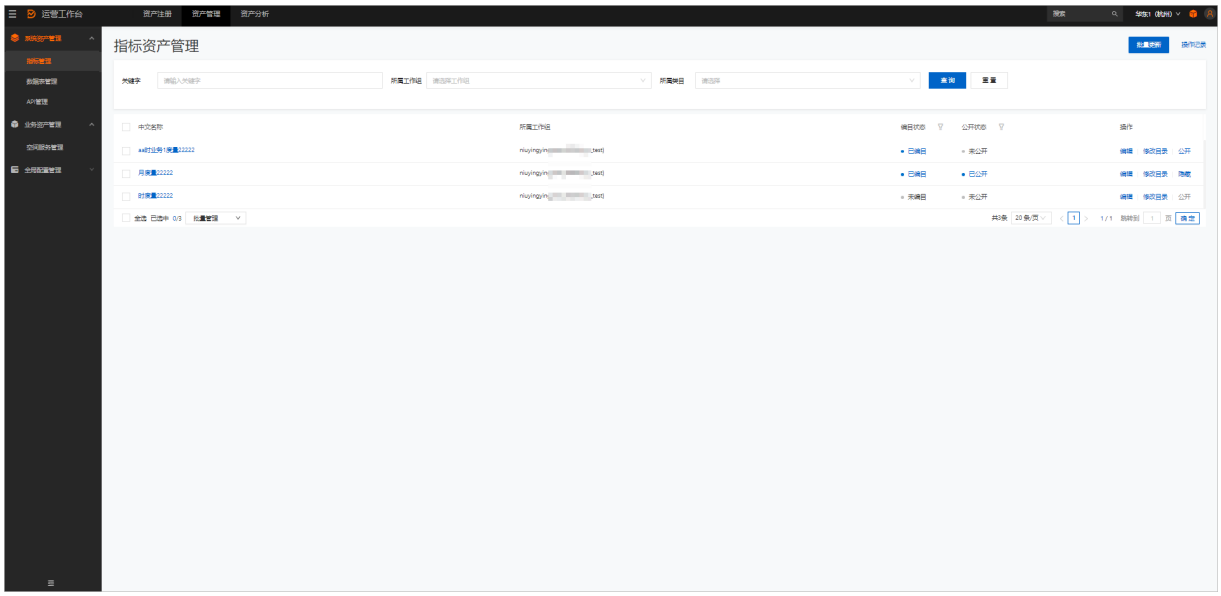




- 支持平台侧需管理的数据资产自动化注册功能，可选择需要管理的数据资产来源作为管理输入。
- 支持对MaxCompute、RDS MySQL、PostgreSQL、DataHub、Hive、AnalyticDB PostgreSQL和Table Store数据表资源进行展现；可通过手动刷新进行元数据同步。

资产管理

系统资产管理：提供对表资产、API资产和指标资产可视化展示，通过编目、公开为资产共享做准备。



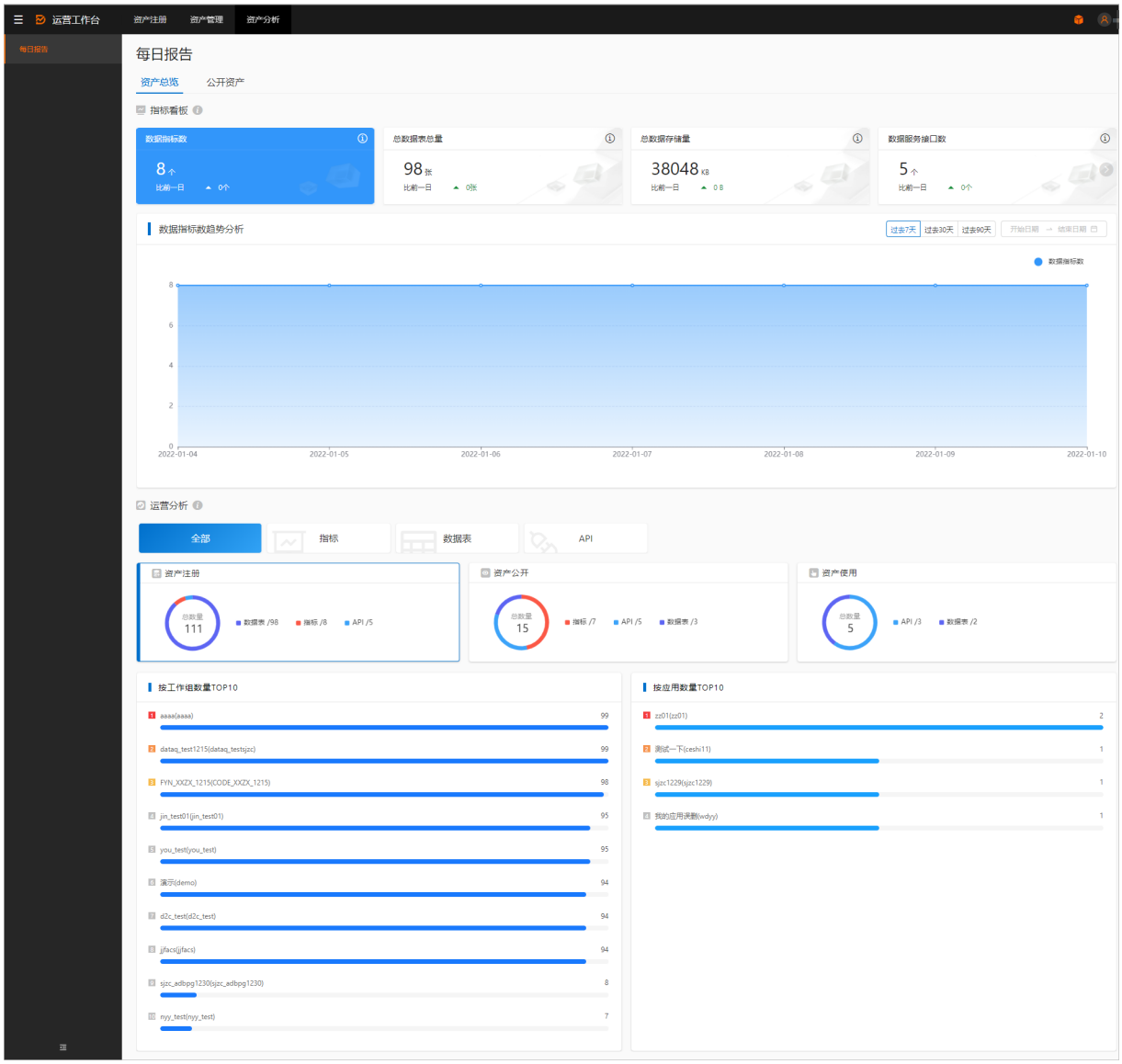
- 数据表资产管理：支持对数据表资产进行编目，并提供查询、公开、隐藏等功能。
- API资产管理：支持对数据服务API、三方API、上报API、数据更新API和服务编排API资产进行编目，并提供查询、公开、隐藏等功能。
- 指标资产管理：支持对指标资产进行编目，并提供查询、公开、隐藏等功能。

业务资产管理：支持对空间服务资产进行编目，并提供查询、公开、隐藏等功能。

全局配置管理：提供自定义目录配置功能，支持设置多级目录，以供数据资产进行后续编目。

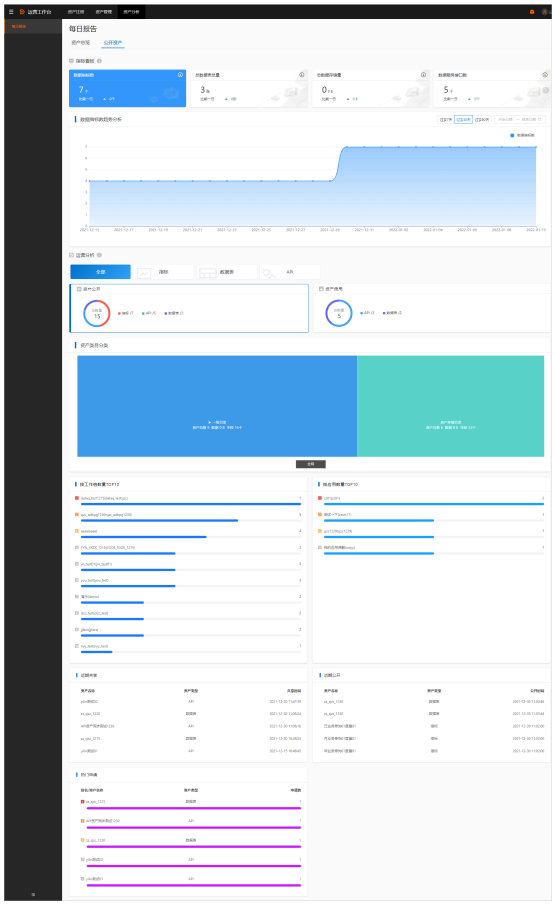
资产分析

资产总览：包括指标看板、数据指标数趋势分析和运营分析。



- 指标看板：展示未公开及已公开的资产，对全租户下已注册的资产进行指标统计，可了解各类资产的量级统计和时间趋势情况。
- 数据指标数趋势分析：展示未公开及已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示未公开及已公开的资产，对全租户下已注册的资产，围绕注册资产的运营阶段（注册-公开-使用），了解各阶段下的多维度统计分析结果。

公开资产：包括指标看板、数据指标数趋势分析和运营分析。



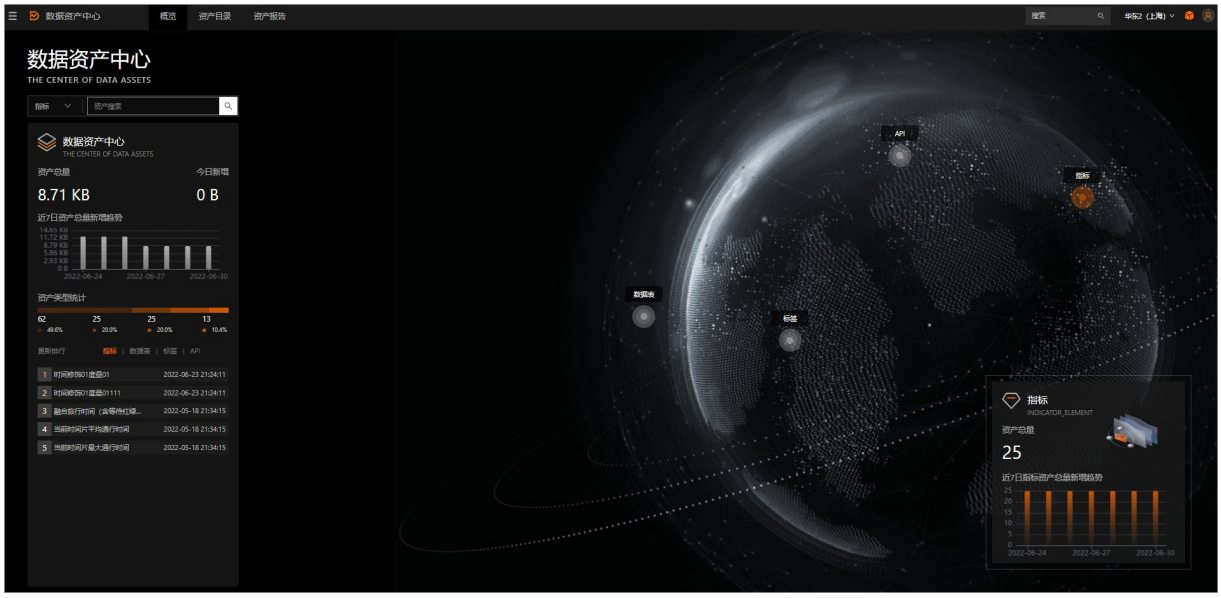
- 指标看板：展示已公开的资产，对全租户下已公开到资产中心的资产进行指标统计，了解各类资产的量级统计和时间趋势情况。
- 数据指标数趋势分析：展示已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示已公开的资产，对全租户下已公开到资产中心的各类资产，围绕公开资产的运营阶段（公开-使用），了解各阶段下的多维度统计分析结果。

6.3.3. 数据资产中心

数据资产中心帮助用户完成数据资产申请、数据资产服务、数据资产共享。通过统一的数据资产目录，沉淀数据供给方生产的各类型数据，帮助资产运营管理者进行高价值、可共享的数据资源的注册、编目分类和公开运营，帮助数据使用者快速了解和获取业务需要的高质量数据资产，促进数据资产发挥更大的业务价值。同时，通过数据资产全局血缘，了解数据资产的上下游关联。

概览

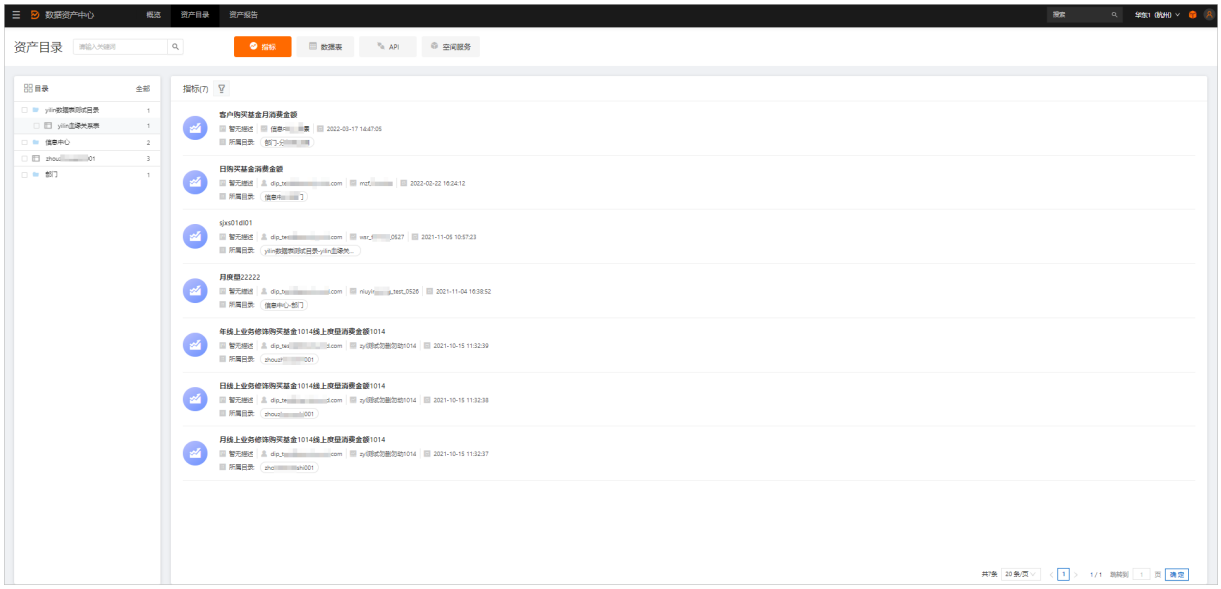
对数据资产的全局及增量指标进行统计和展示，帮助用户快速了解平台管理数据资产的总体情况。



- 支持数据表、API和指标资产类型，同时，支持精确或模糊搜索数据资产。
- 支持数据资产总量、今日新增等全局指标展示，并以趋势图的形式展示增量信息。
- 支持数据表、API和指标资产独立展示增量信息。

资产目录

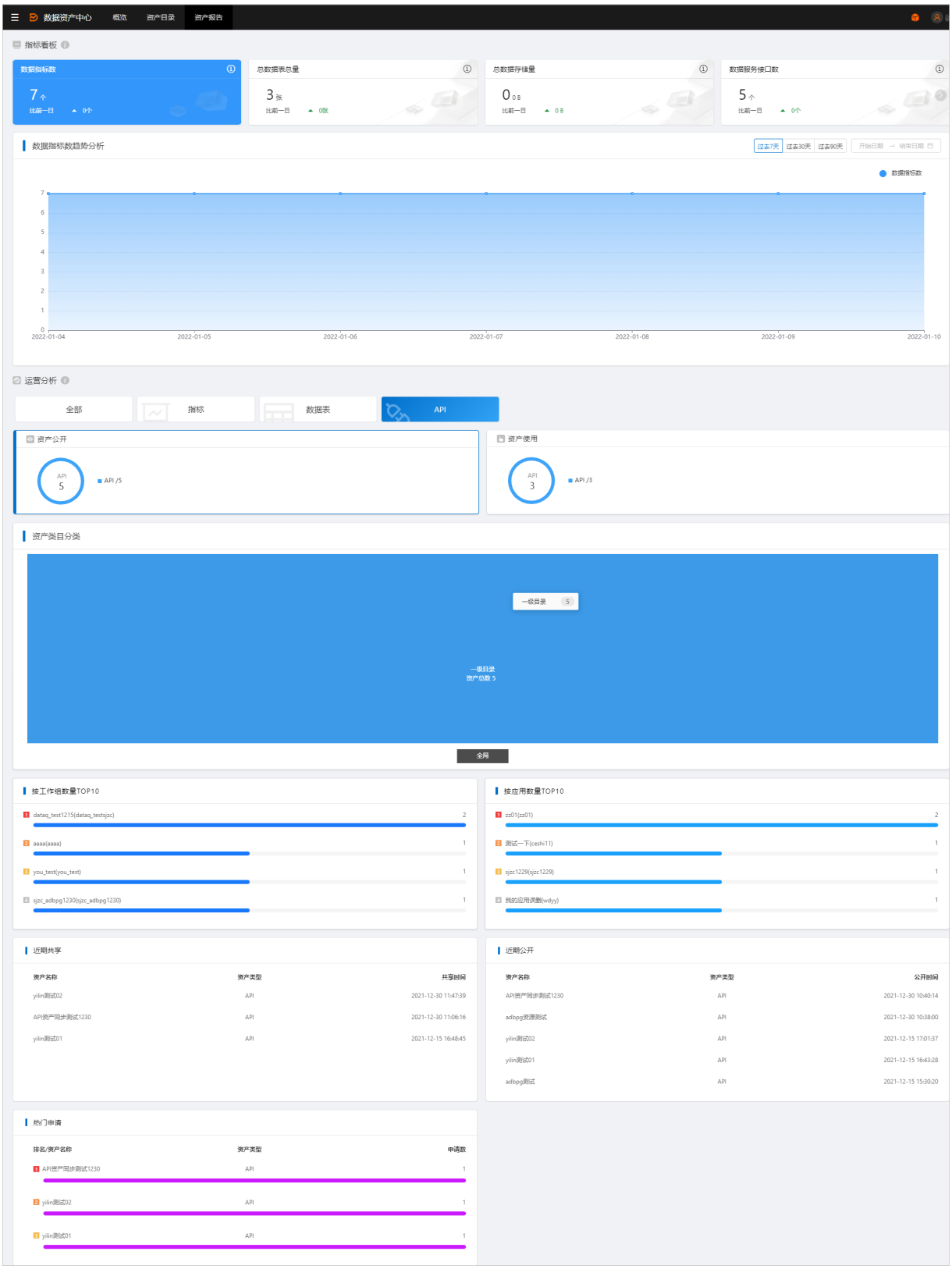
对公开共享的数据资产，可在资源目录进行搜索、目录列表查看，并通过资产详情，了解数据资产的业务价值，促进数据资产的共享和申请使用。



- 支持展示公开的数据表、API和指标资产目录，并提供通过关键词搜索对应资产。
- 支持展示数据表资产基本信息、扩展信息、字段信息、数据预览、血缘关系和数据质量等信息，并提供申请数据表使用权限的功能。
- 支持展示API基本信息、扩展信息、接口信息、血缘信息，并提供申请API使用权限的功能。
- 支持展示指标基本信息、扩展信息、概览信息和关联信息等信息。
- 支持展示空间服务基本信息、扩展信息、概览信息、数据预览和用户评价等。

资产报告

展示指标看板、数据指标数趋势分析和运营分析。



- 指标看板：展示已公开的资产，对全租户下已公开到资产中心的资产进行指标统计，了解各类资产的量级统计和时间趋势情况。包括数据指标数、总数据表总量、总数据存储量、数据服务接口数和各趋势分析图。

- 数据指标数趋势分析：展示已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示已公开的资产，对全租户下已公开到资产中心的各类资产，围绕公开资产的运营阶段（公开-使用），了解各阶段下的多维度统计分析结果。包括指标、数据表和API各类型资产的资产公开数据、资产类目分类和近期公开资产等信息。

6.3.4. 画像分析

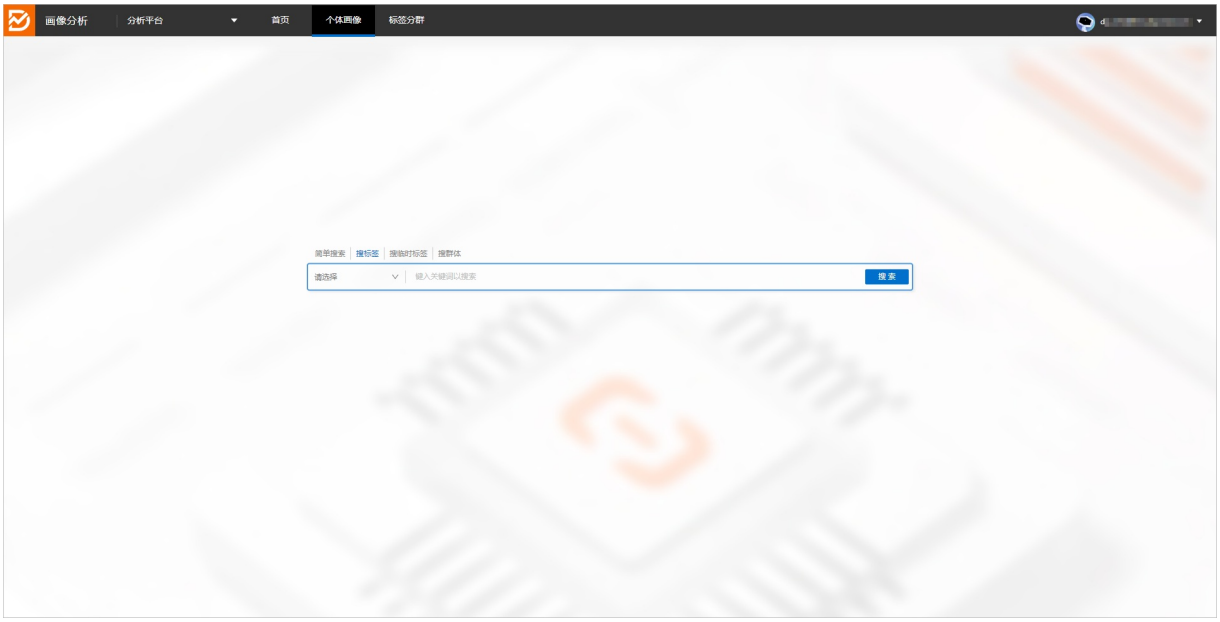
画像分析是对象数据的形象化展示，即将对象的信息画像化，也就是用一系列的画像简洁明了地刻画一个群体或一个个体的特征，是在一连串数据上构建起来的对象模型。对象画像是描述目标对象、分析与聚焦的有效利器，为业务人员提供了无需技术基础，零代码的大数据分析能力。

画像分析系统可基于智能标签系统为数据分析人员、业务人员提供一站式群体分析平台。涵盖画像分析的全链路，主要包括基于标签的群体圈选、群体计算、多状态群体发布等，支持用户建立单个体档案，从而实现用户对于目标群体的精准定位。

画像分析中的人群支持通过将现有标签进行鼠标点选组合即可生成新的标签组合，并可给该标签组合命名，该标签组合在前端保存为标签计算的逻辑，在大数据环境下提供快速计算，不同的新标签在页面可以快速嵌套使用，创建更贴近业务的标签组合。

个体画像

- 提供按群体、个体标签、个体临时标签、关联对象进行搜索的功能，可快速找到目标个体或群体。
- 查询结果支持查看群体和个体详情，群体查询结果支持群体分析。

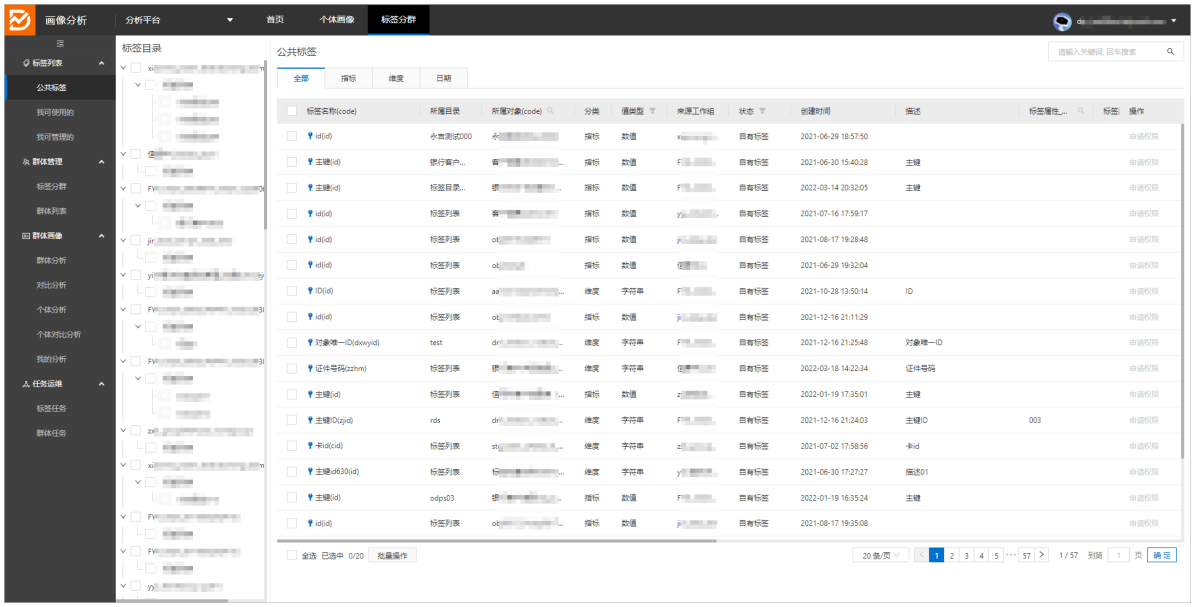


标签分群

标签列表

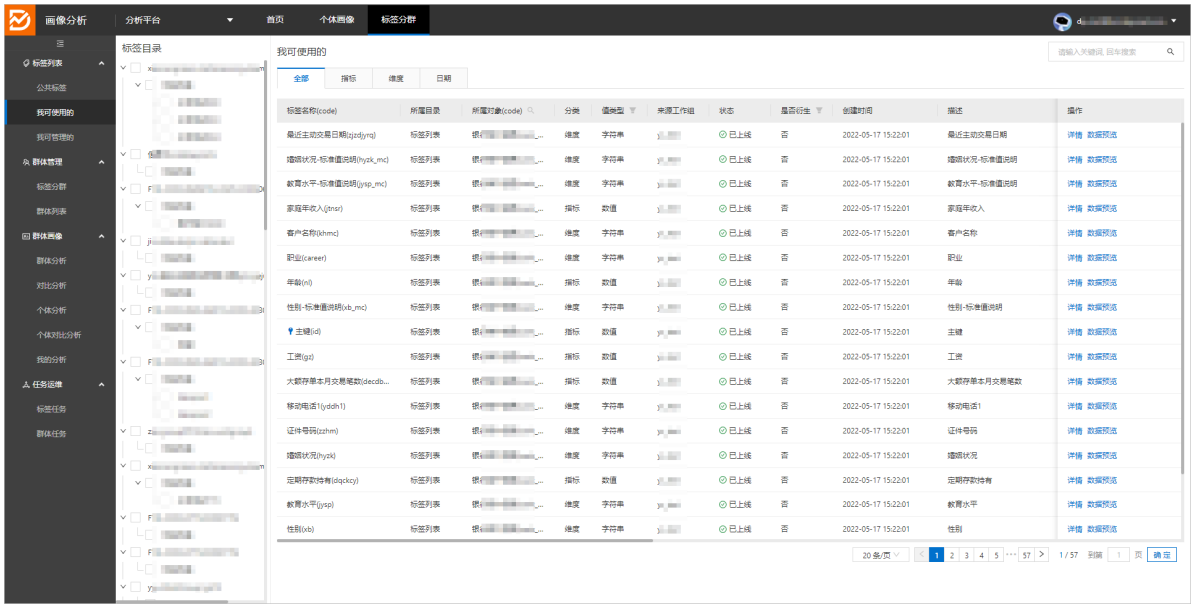
- **公共标签**

支持查看该租户下所有工作组的标签。



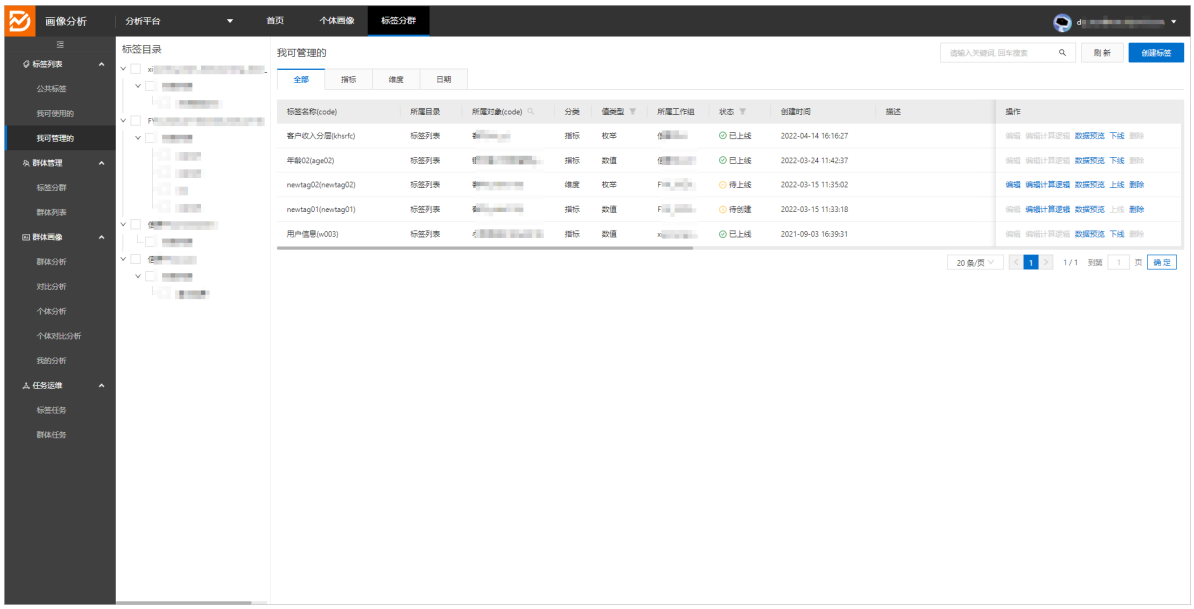
- **我可使用的**

  支持查看该租户下有使用权限的已上线标签，并可进行标签详情查看、数据预览操作。



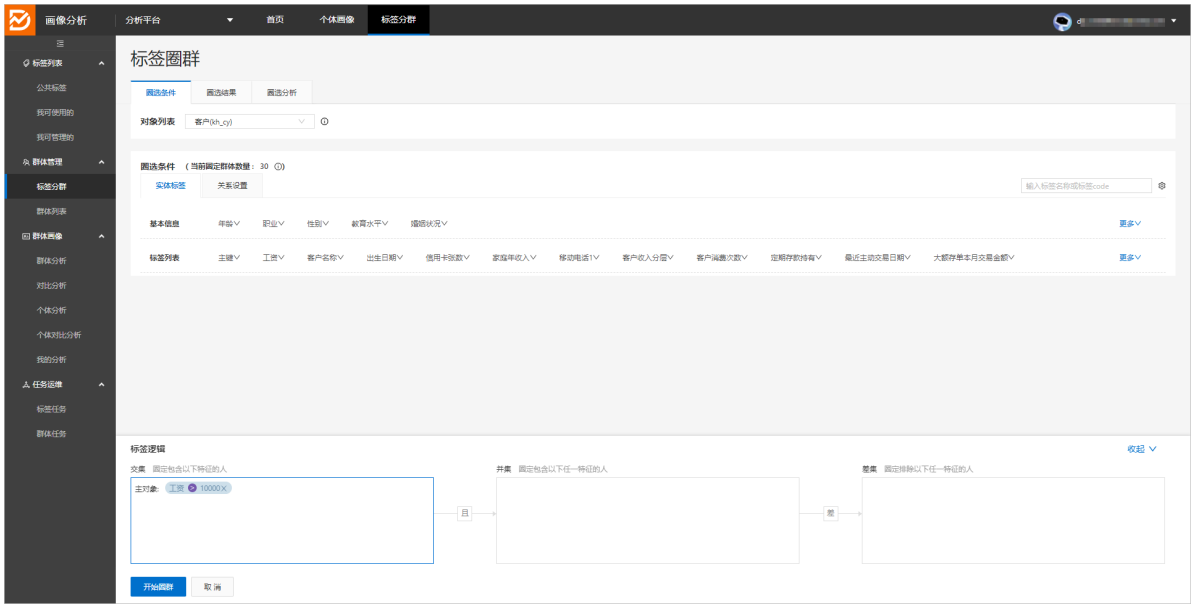
- **我可管理的**

支持通过拖拉拽方式编辑计算逻辑创建枚举型、数值统计型标签，并对衍生标签进行上线、下线方便业务用户沉淀业务经验。



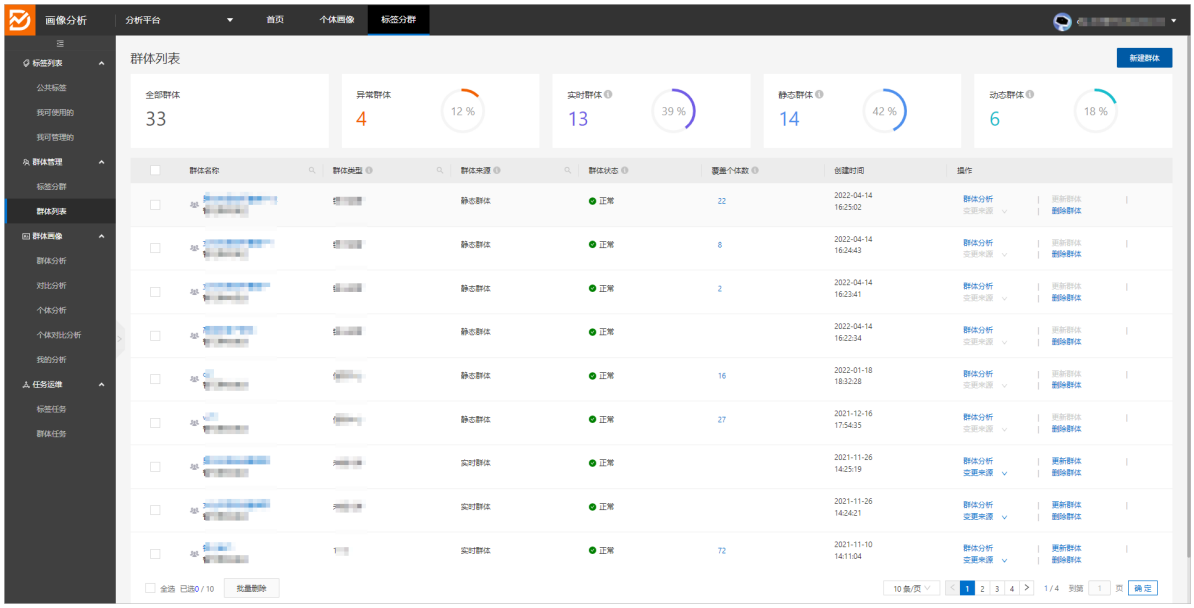
群体管理

- **标签分群**
  - 提供对智能标签里面设定好的对象上的标签设置筛选条件圈定目标群体功能。
  - 支持在标签圈群页面对单个对象或对象标签模型的对象间进行标签的交并差集组合计算生成实时、静态和动态群体。
  - 群体圈定结果支持重新圈定、发布群体、发布API、下载到本地、设置圈选结果展示标签、分析等功能。



- **群体列表**

支持群体统一管理。支持对群体信息进行新建、删除、更新、状态变更、查看、编辑、查看不同类型群体个数和占比等操作。新建群体支持标签圈群、导入本地文件和群体计算三种方式。可实现单个个体多维信息展示，并支持临时打标签能力，具有该临时标签的个体，可以在智能搜索模块被搜索；支持导入Excel格式的群体列表到画像中，进行群体的分析。导入前提，该Excel所含有的标签和值，已经在数据库中存在。支持用户在详情页面中，查看群体详情和个体详情信息。



- **群体计算**

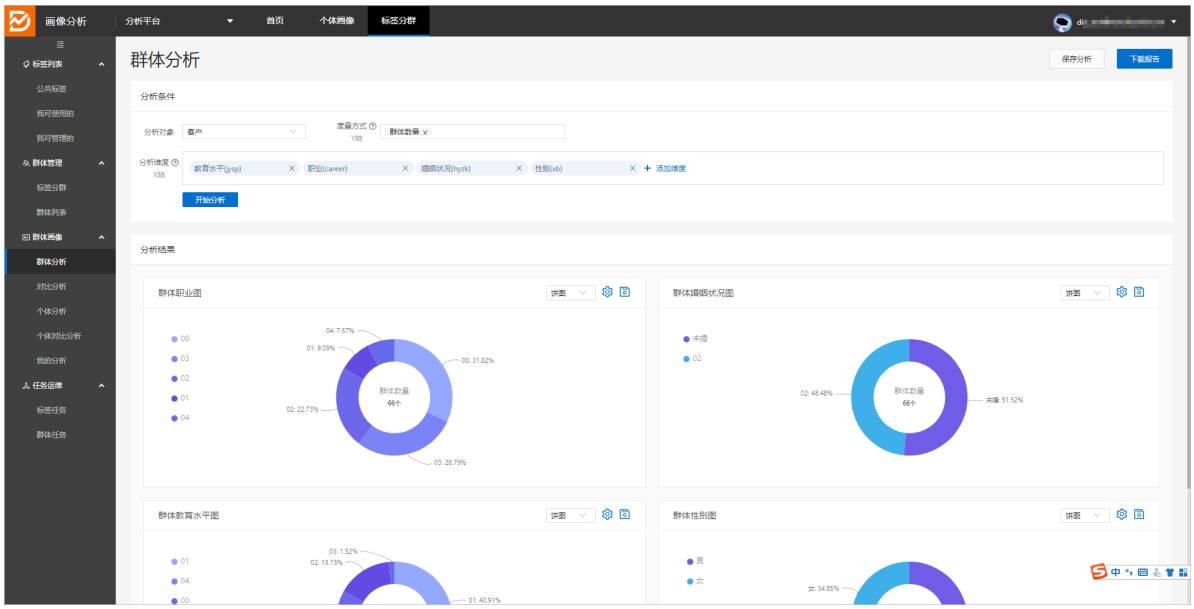
通过图形化的计算方式可实现两个群体交并差计算得到新的群体，并提供对群体结果支持查看详细群体、发布API、静态、动态和实时群体。



群体画像

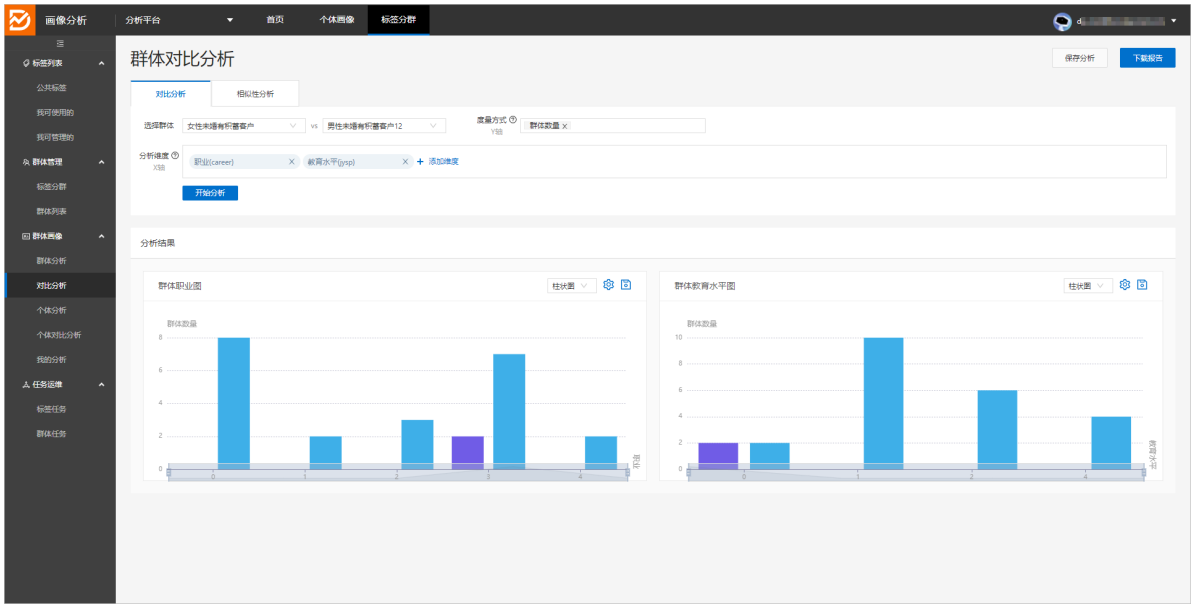
- **群体分析**

支持单群体或对象按单个度量方式的多个维度分析，并可以对分析结果保存为分析模版和下载分析报告。分析结果可生成柱状图、折线图、面积图和饼图，图表可支持的聚合方式包括计数、平均值、求和、最大值和最小值。



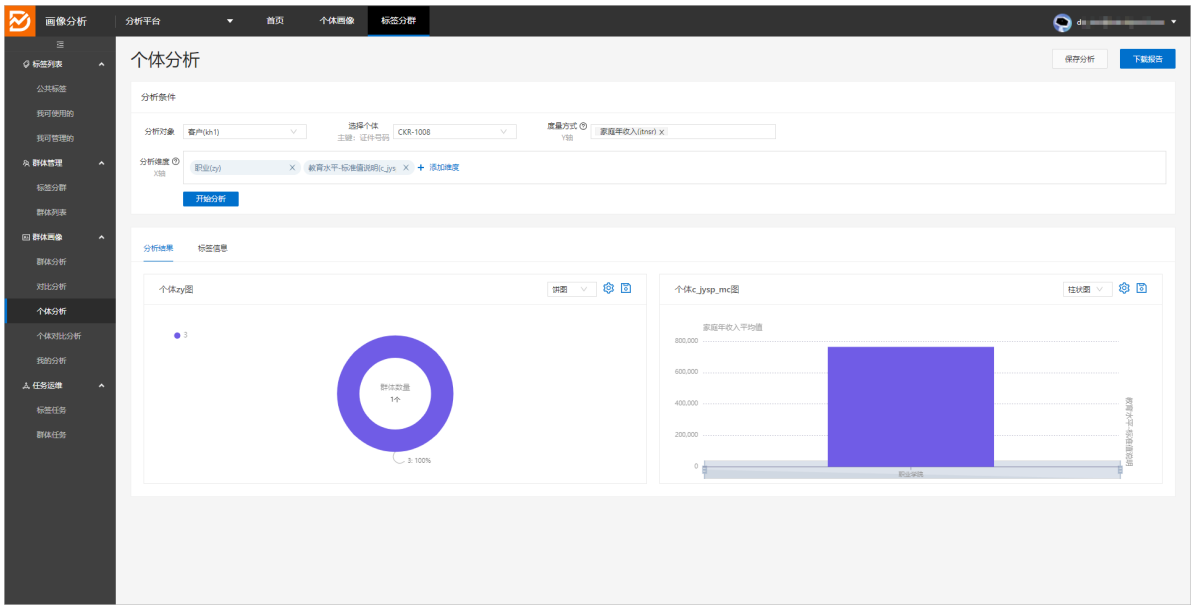
- **对比分析**

支持两个群体按单个度量方式的多个维度分析和相似性分析。两个群体按单个度量方式的多个维度分析，并可以对分析结果保存为分析模版和下载分析报告。分析结果可生成柱状图、折线图和雷达图，图表可支持的聚合方式包括计数、平均值、求和、最大值和最小值。相似性分析支持查看两个群体的群体相似度、标签覆盖人数排行榜和相似个体列表。



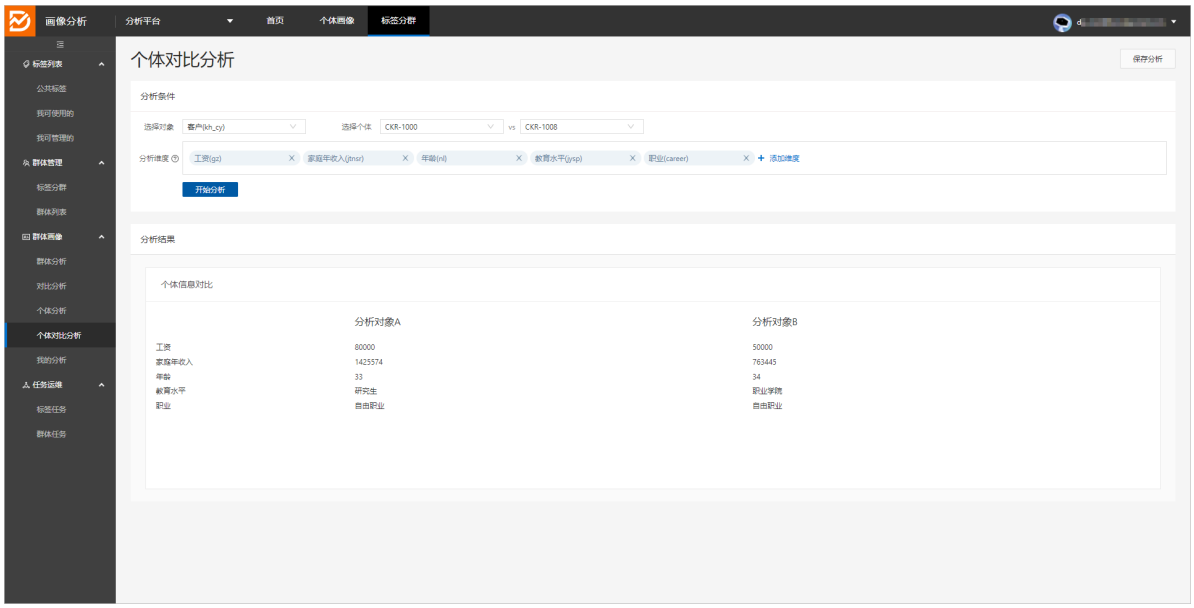
- **个体分析**

支持以对象为基础、以对象主键为核心个体，通过添加一个对比字段为度量方式，添加一个或多个字段为维度对单个个体进行分析，并提供保存、下载分析报告。

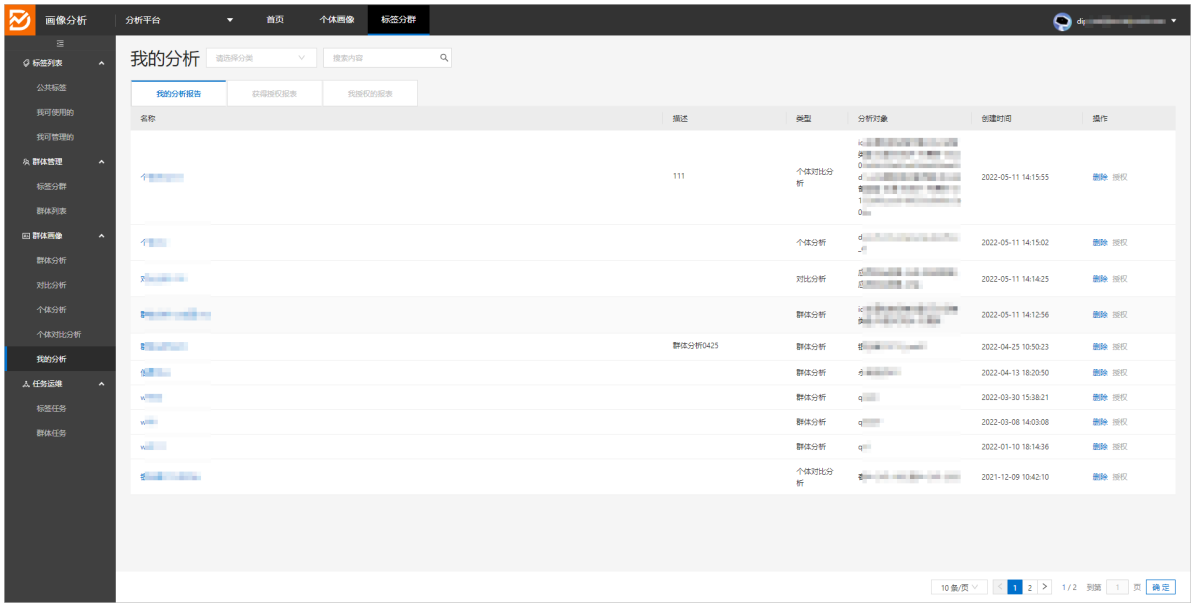


- **个体对比分析**

  支持两个体间的多维对比分析，并提供对分析结果保存为分析模版。



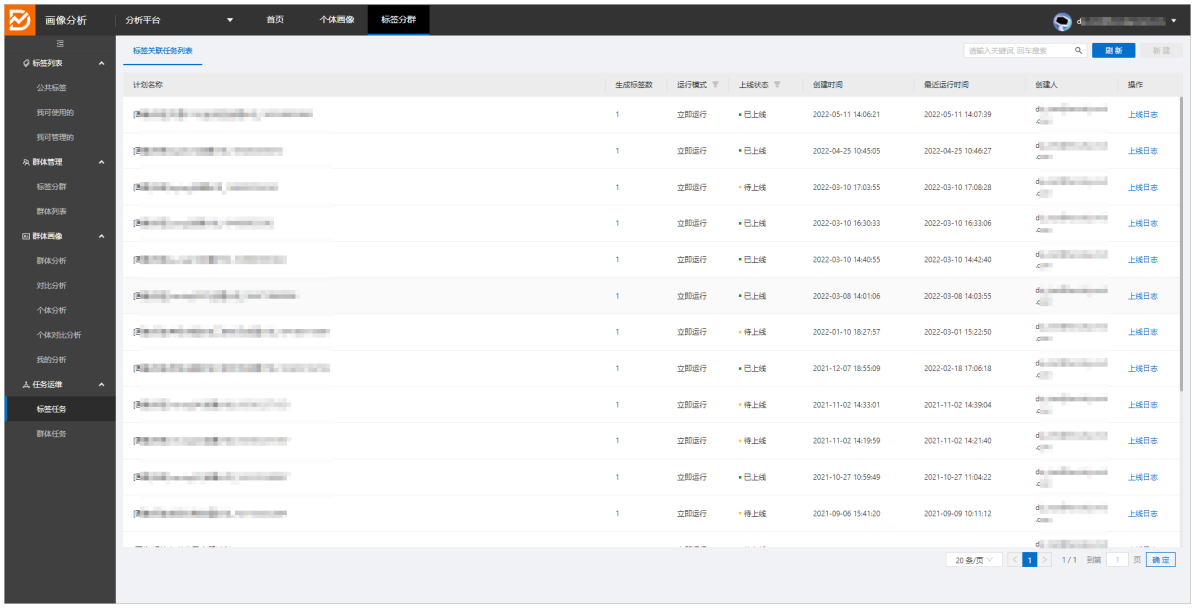
- **我的分析**
  - 支持对查看本工作组下历史保存的分析结果，并提供删除功能。
  - 支持查看其他工作组赋予查看权限的分析报表。
  - 支持赋予其他工作组查看本工作组下分析报告的权限，并且支持权限的撤回。



任务运维

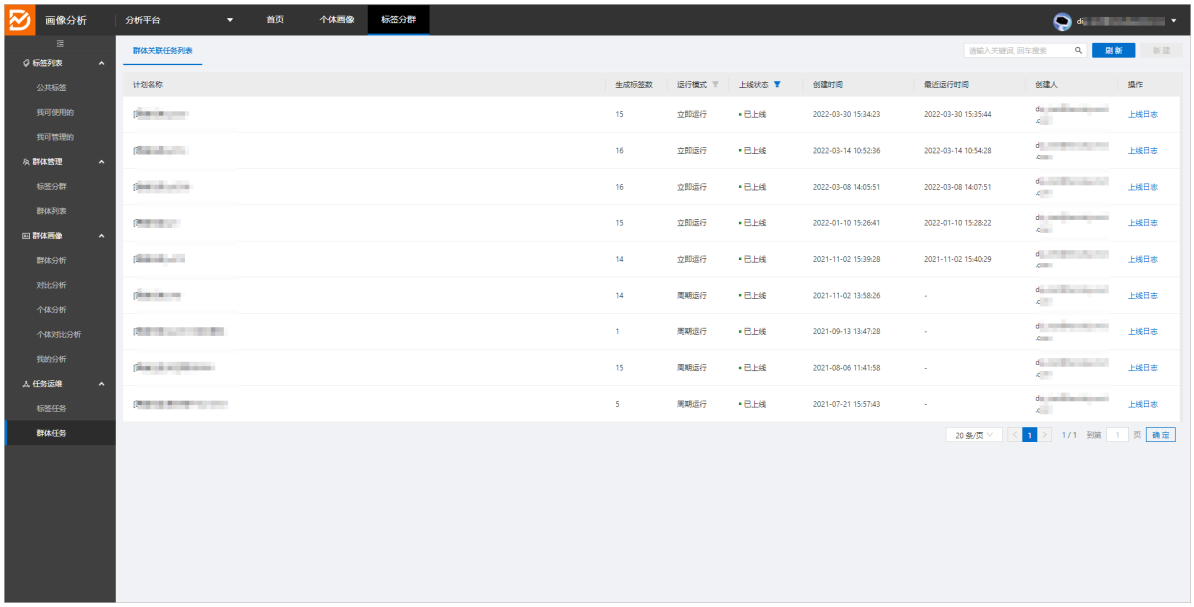
- **标签任务**

  支持创建标签任务的任务运维管理。



- **群体任务**

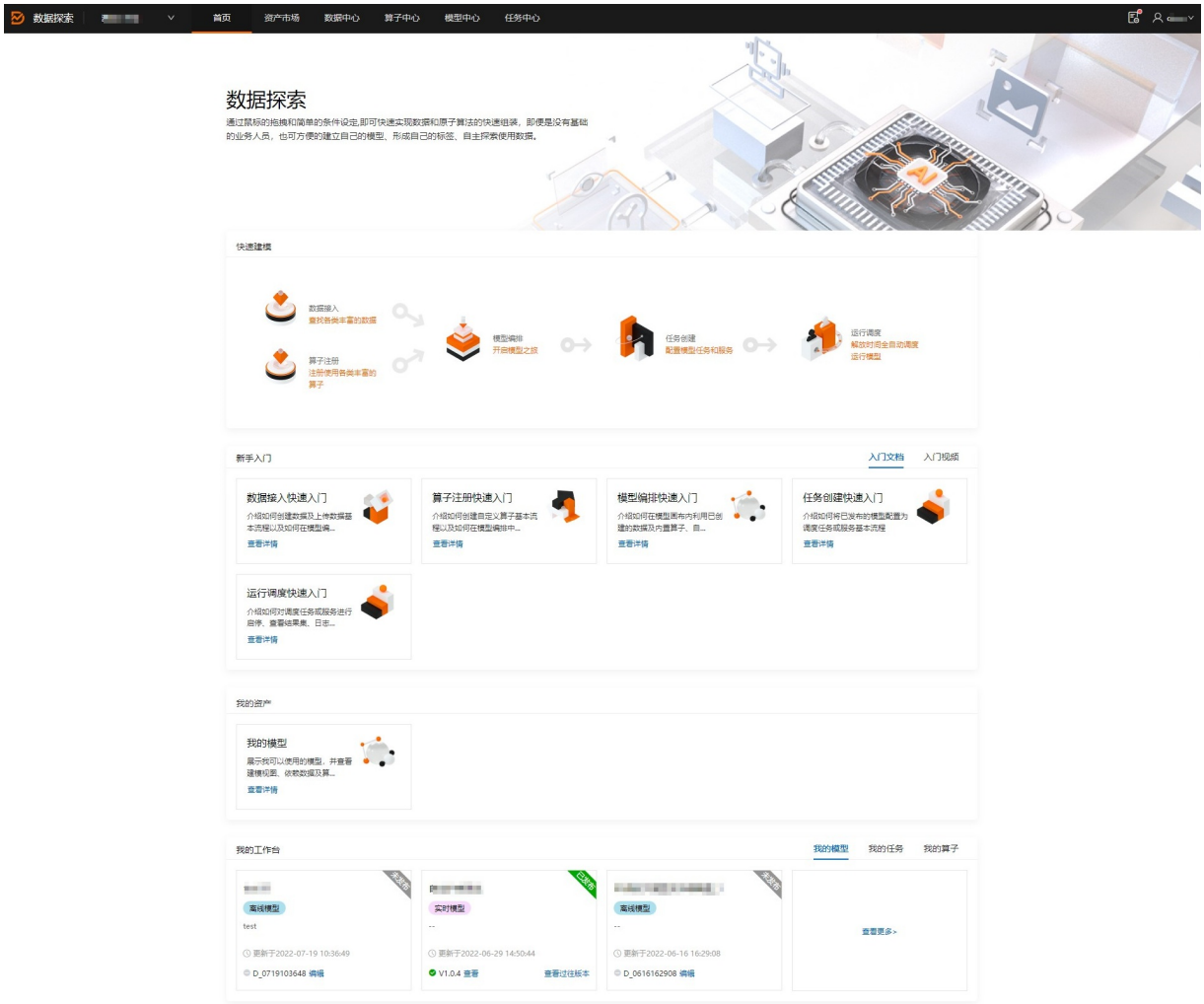
支持发布静态和动态群体，并提供在数据库一次性和周期落表的任务运维管理。



6.3.5. 数据探索

6.3.5.1. 首页

首页主要对系统中的模型、算子、任务和用户情况统计分析，包括模型、算子总量统计、模型、算子增量统计、模型任务量统计，用户总数等概览信息，并针对普通用户提供用户使用指引、入门指南和工作台。



6.3.5.2. 数据中心

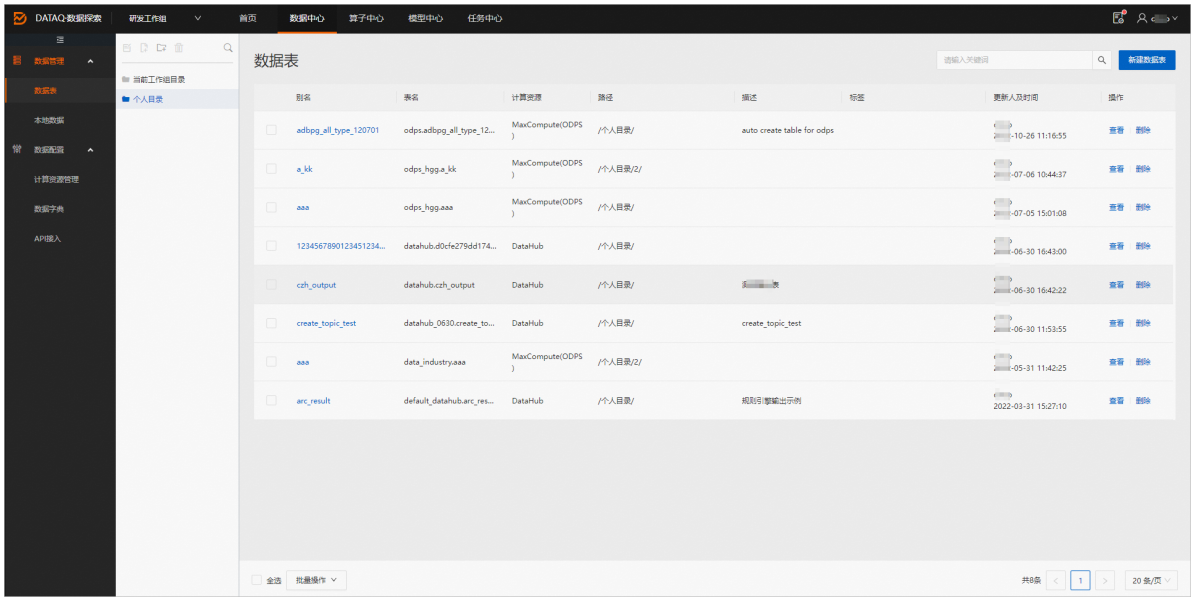
数据中心包含数据管理、数据配置。

数据管理

数据管理模块展示并管理用户添加的数据。主要功能包括数据表引入、本地数据导入。

数据表

1. 支持用户从已注册的的计算资源类型中添加数据表。



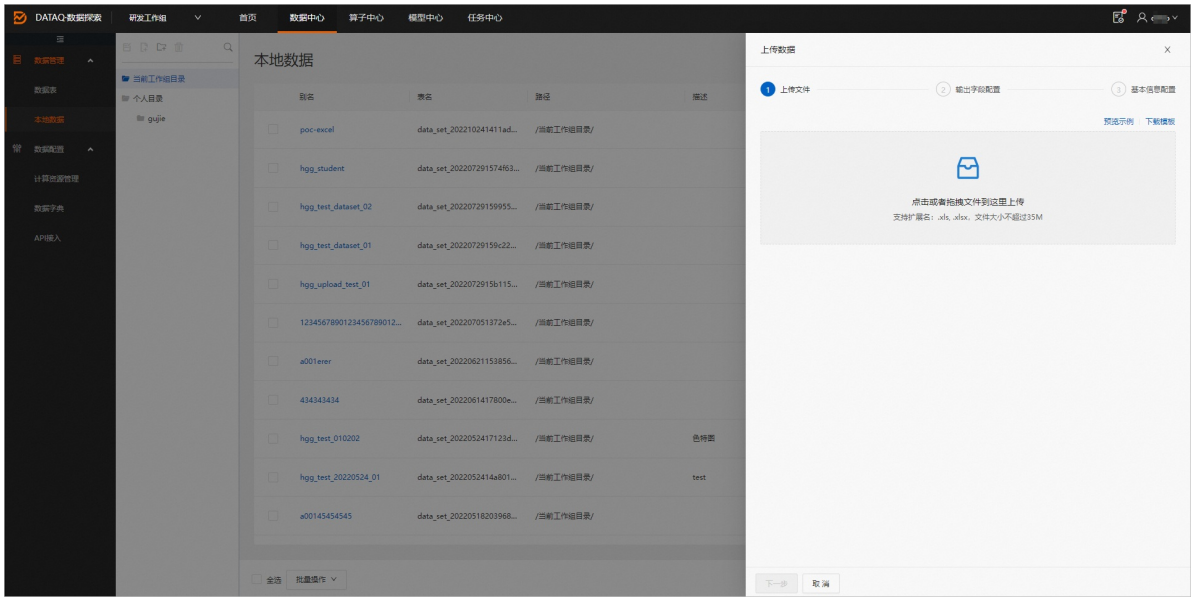
2. 支持基于已添加的数据查看表信息、表内展示的字段。



3. 支持将已添加的源表数据进行删除。

本地数据

1. 支持用户通过上传本地文件方式将数据集导入至系统。



2. 支持上传本地文件更新原数据集，支持增量导入和全量导入。
3. 支持将上传的数据集下载到本地。

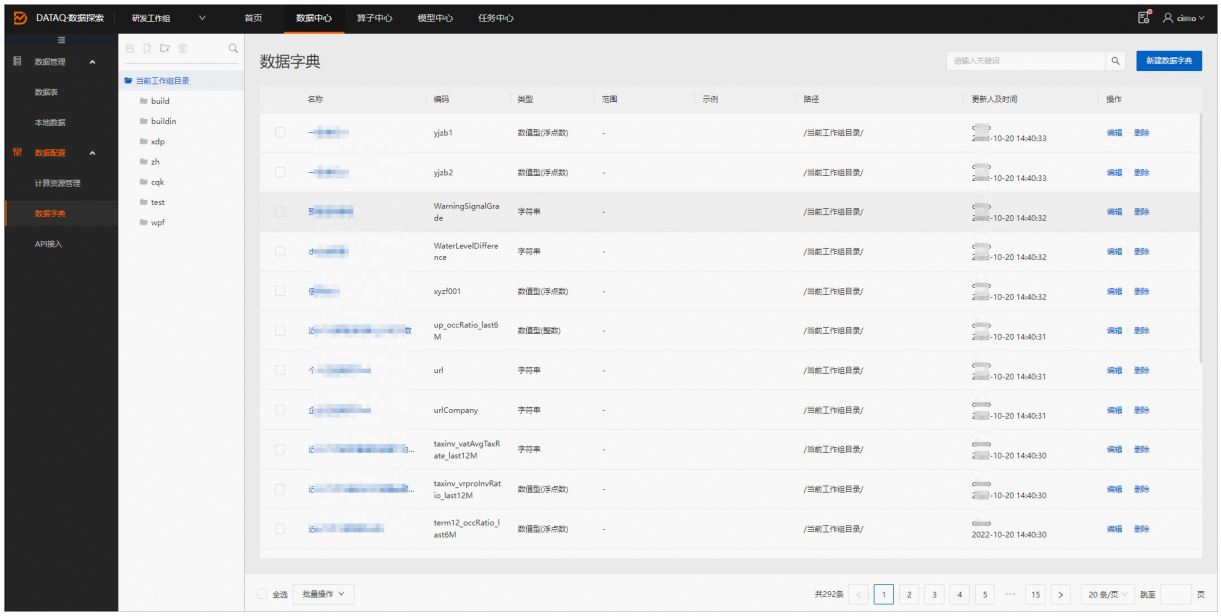
数据配置

计算资源管理

支持用户在系统内添加云计算资源，支持的计算资源类型有：ODPS、hive等。

数据字典

数据字典管理支持创建并维护模型中使用的数据字典，制定并管理平台遵循的统一数据标准，帮助平台管理者和数据管理者管控治理后数据的一致性和数据质量。



API接入

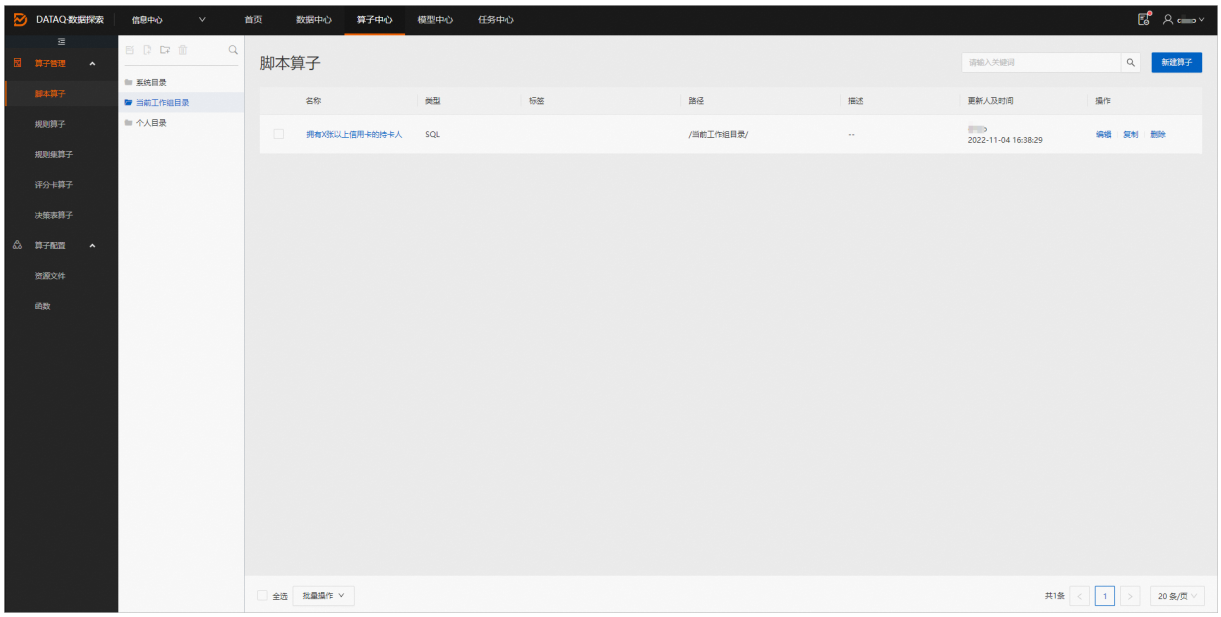
通过API参数配置、API调试、返回示例等对目标数据表进行API调用。

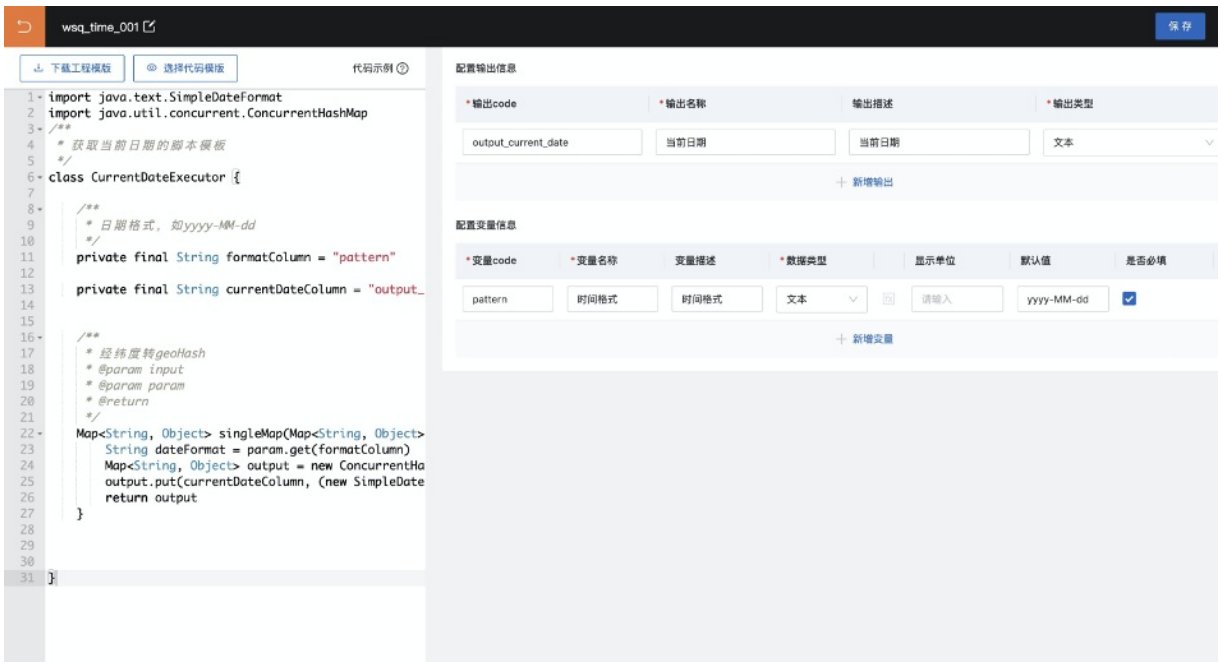
6.3.5.3. 算子中心

算子中心包含算子管理、算子配置。

算子管理

算子管理模块展示并管理用户添加的脚本算子，通过模型生成的模型算子，用于在构建模型时使用。



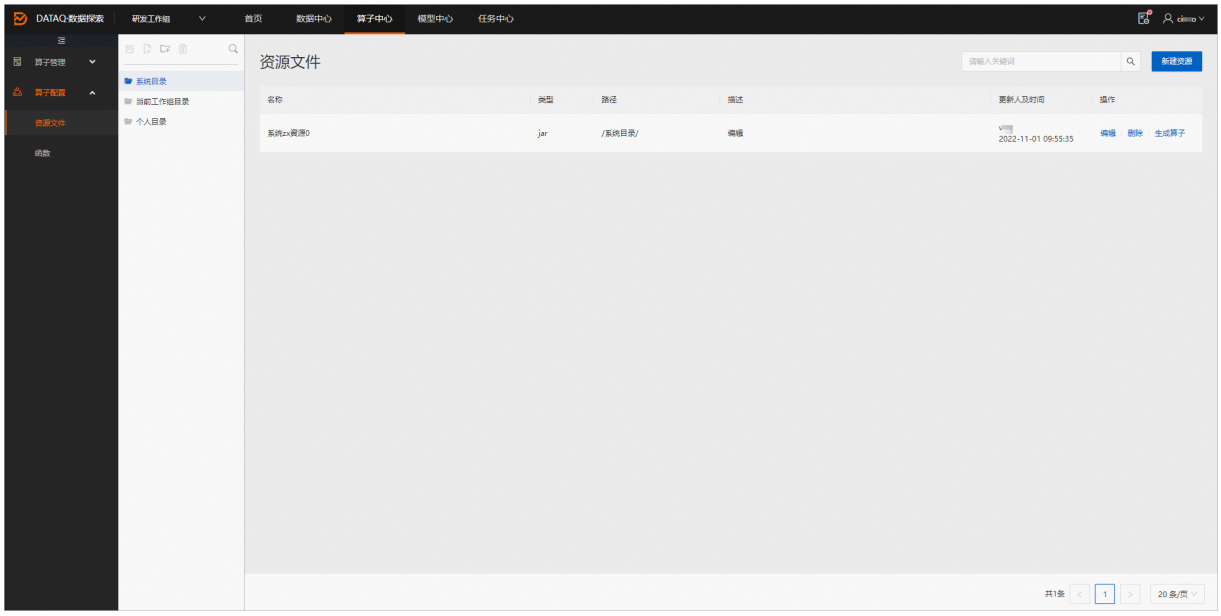


算子配置

算子配置模块展示并管理用户上传的资源文件及函数信息、包含了资源文件和函数。

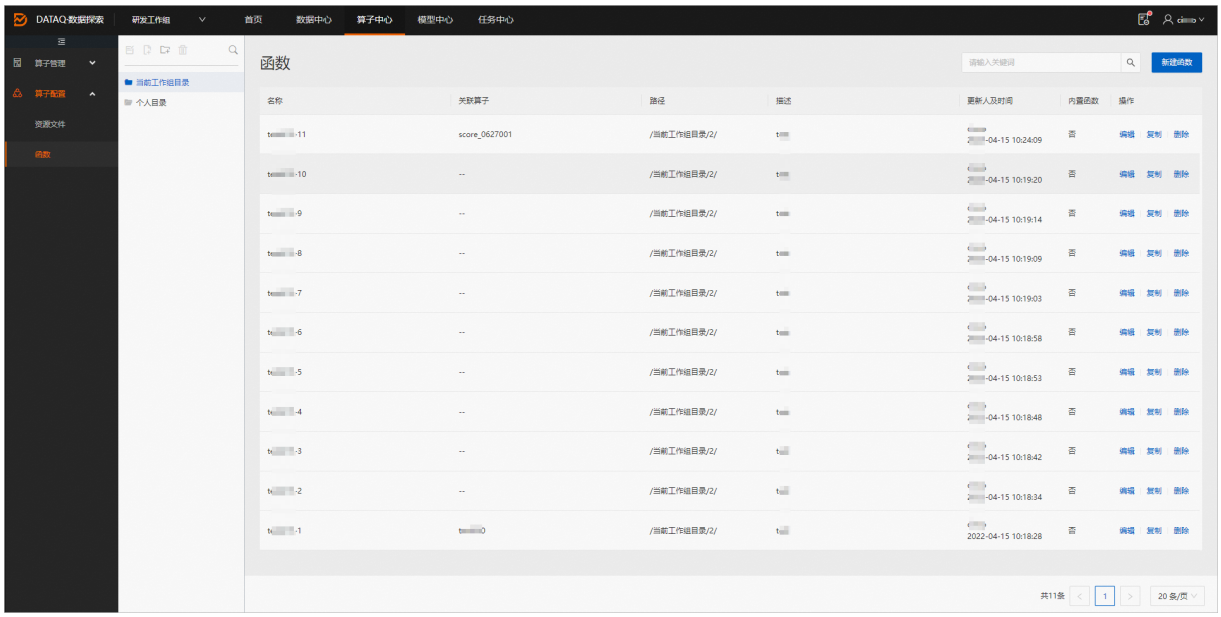
资源文件

支持用户在系统内上传本地资源文件及OSS空间配置，根据上传的资源包可生成对应类型的脚本算子。



函数

支持用户在系统内自定义函数及函数规则逻辑、根据自定义的函数系统默认读取展示该函数的入参和出参。

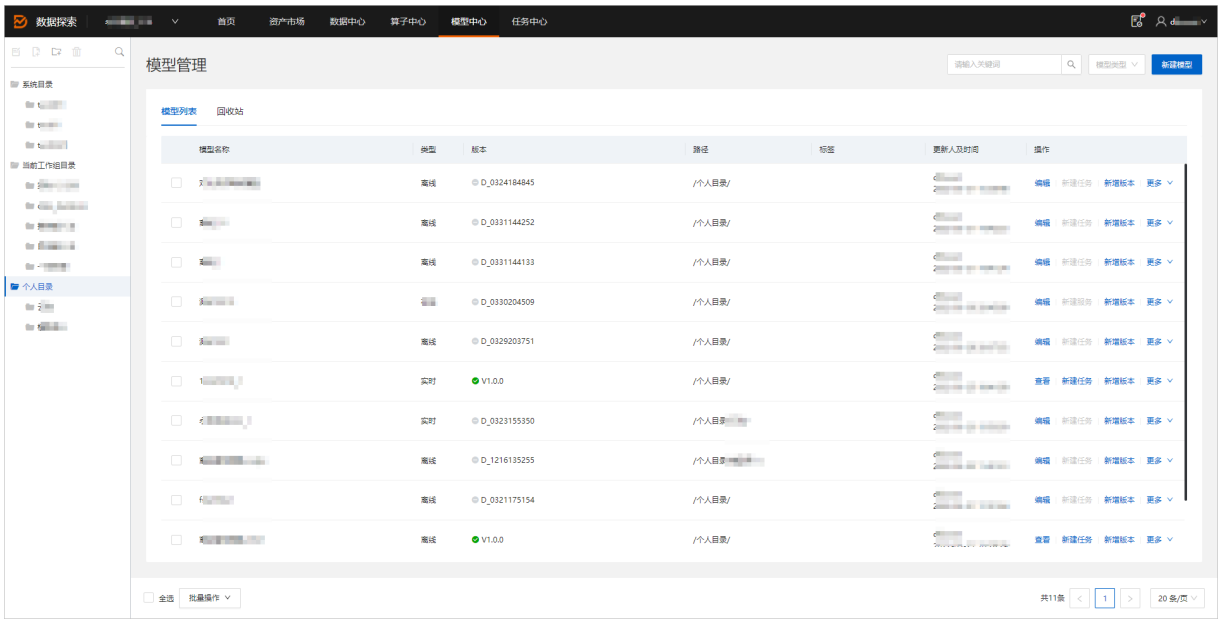


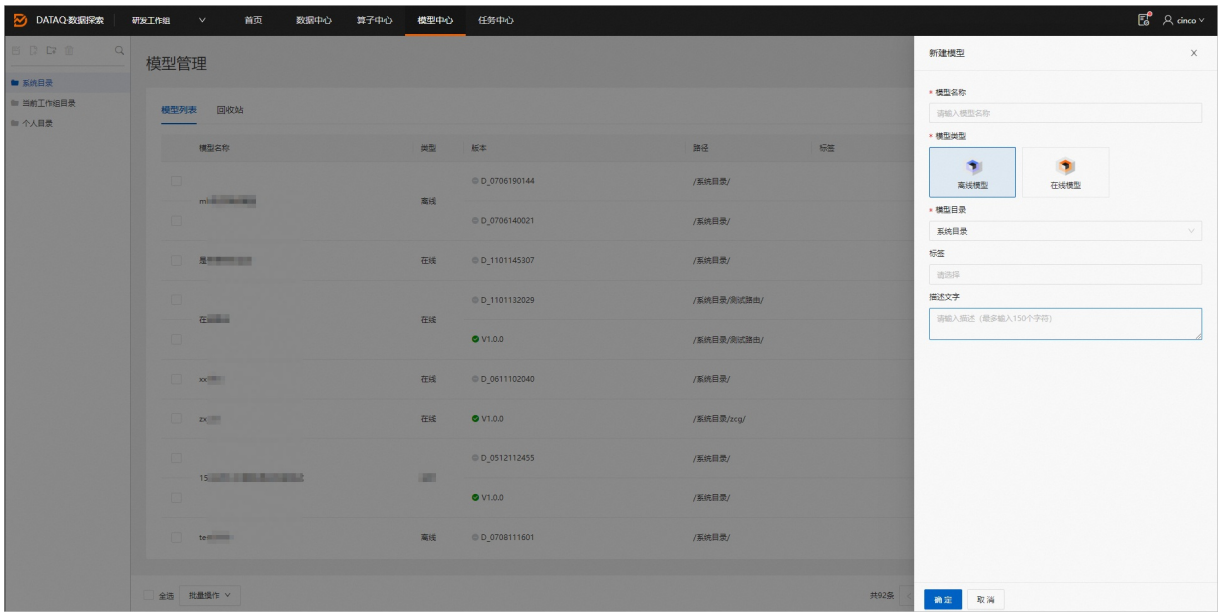
6.3.5.4. 模型中心

模型中心提供业务模型的可视化组装、编排能力，提供拖拽式智能算子的面向业务场景的模型Flow DAG的编辑与配置。通过该模块可以将业务需求转换为业务模型进而通过底层引擎的运行产生业务效果，主要包含以下功能：模型管理、模型编排、模型发版、模型运行调试等功能。

模型管理

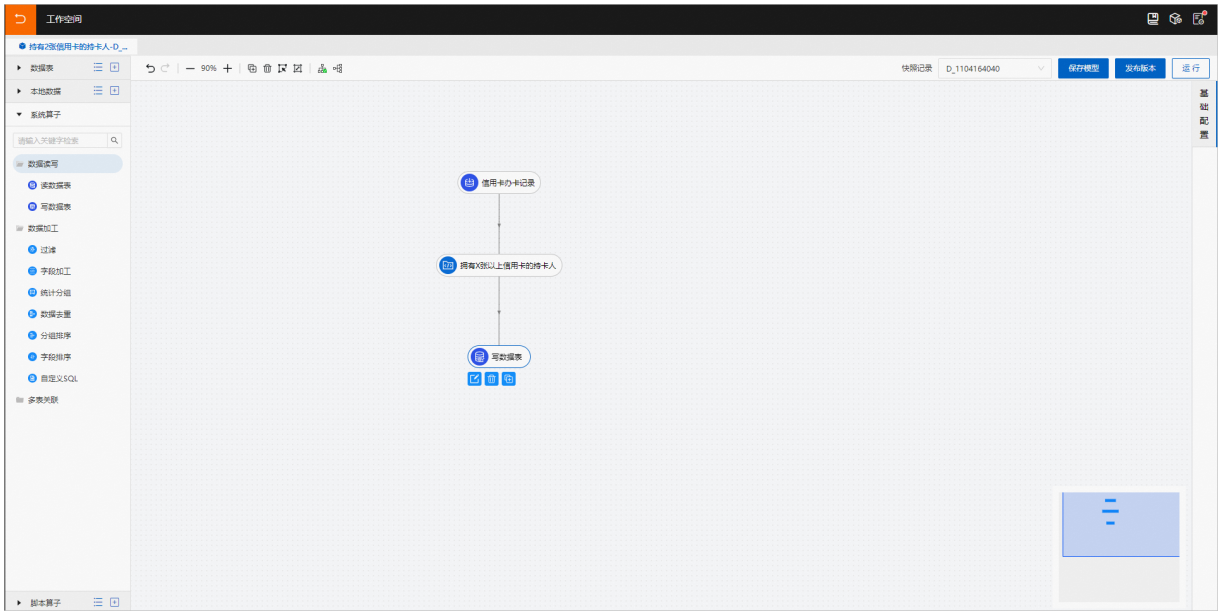
模型管理模块展示并管理用户添加的离线、实时场景的模型、支持对模型进行编排建模、编辑、新建任务、新建服务、版本管理、模型复制、归档。

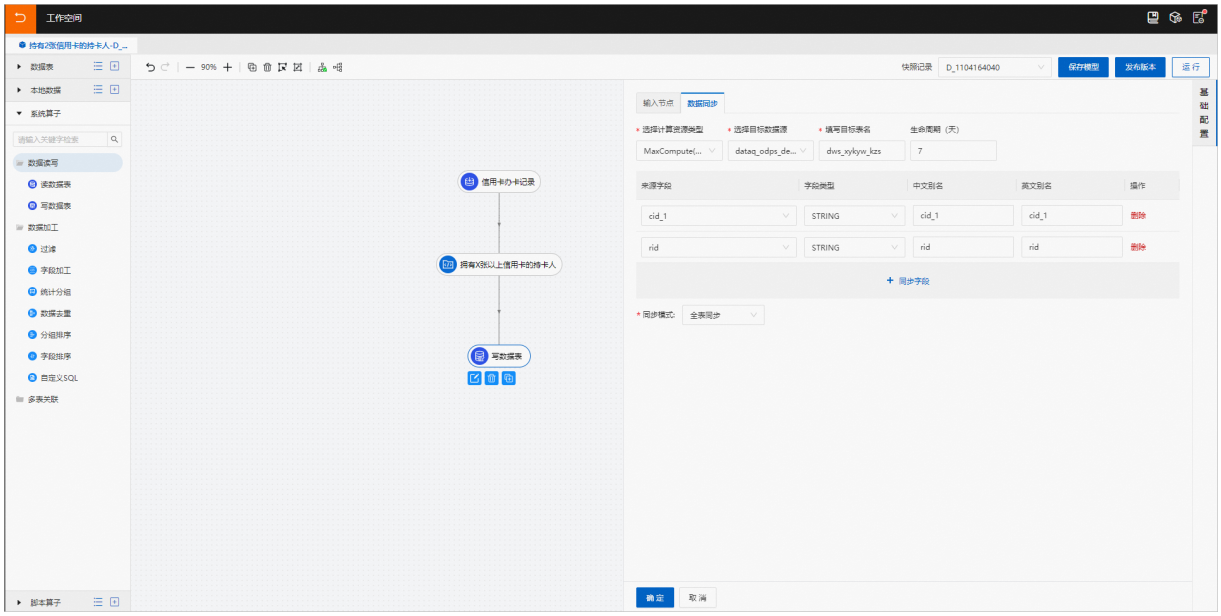




画布编排

模型编排模块提供了丰富的系统内置算子，通过托拉拽的形式与添加的数据进行碰撞产出符合需求的业务模型，在模型画布内支持运行、发布版本、保存、搜索节点等基本功能；支持对运行节点的算子配置条件规则；算子节点支持查看结果集，运行日志、重命名、执行到此处等功能。





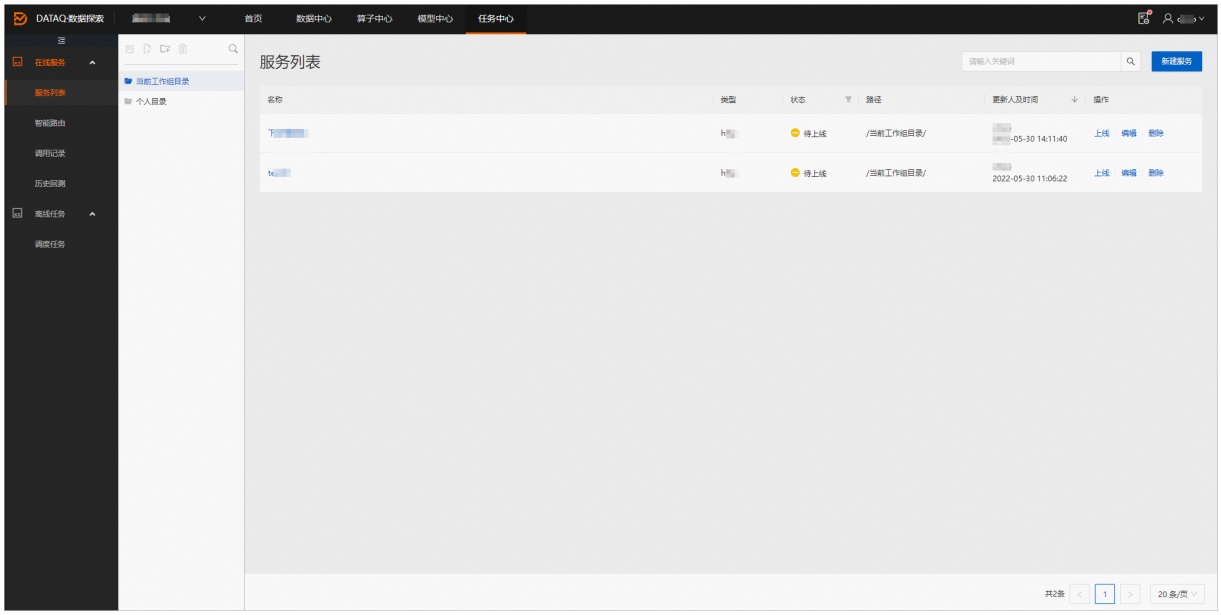
6.3.5.5. 任务中心

任务中心主要是对离线、在线模型自动化调度运行进行任务式管理，每次任务调度时，都将按照模型设置好的参数和数据自动进行运行，通过任务状态监控单次任务的健康情况。

在线服务

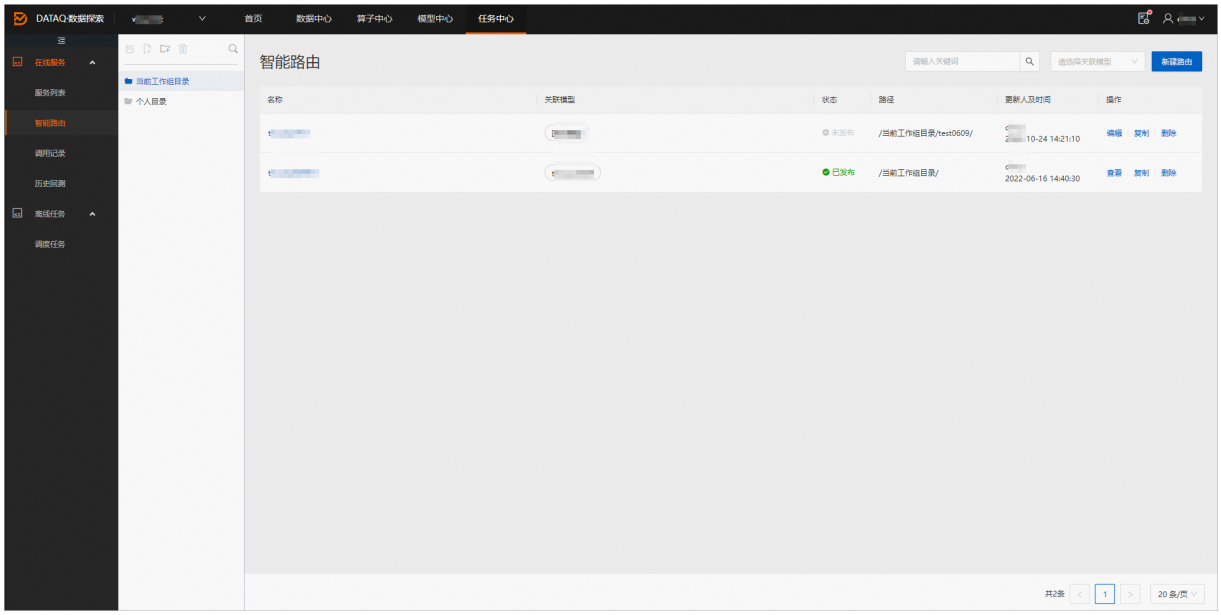
在线服务模块展示并管理用户添加的服务，已创建的服务支持进行上线、下线、查看调用记录等功能。

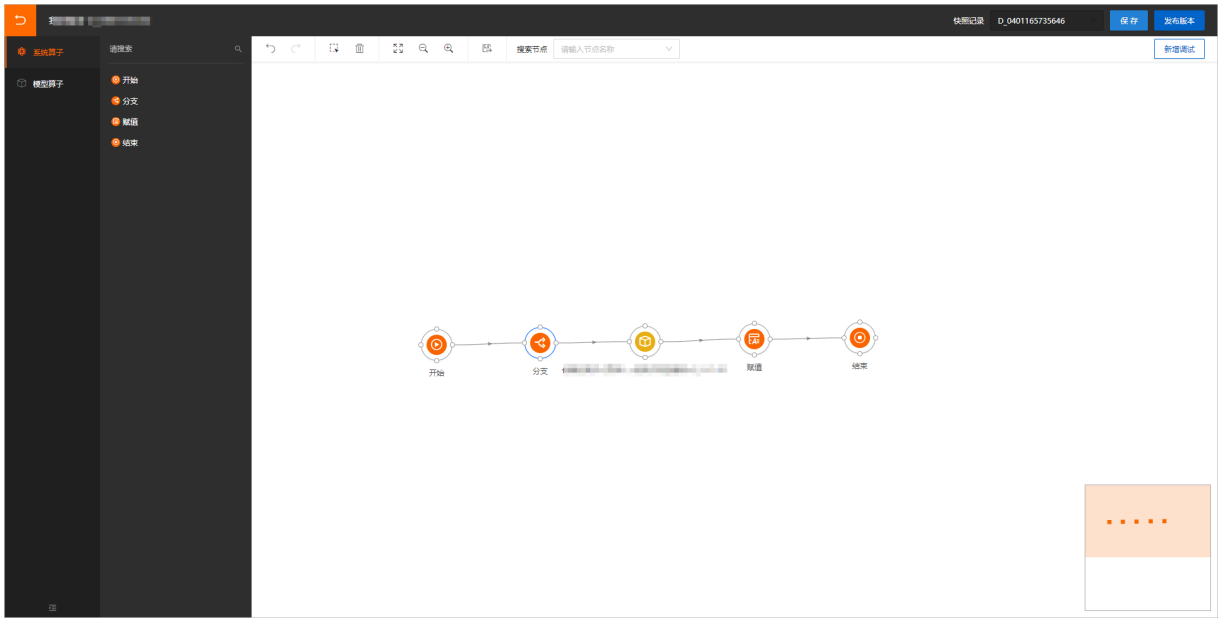
服务列表



智能路由

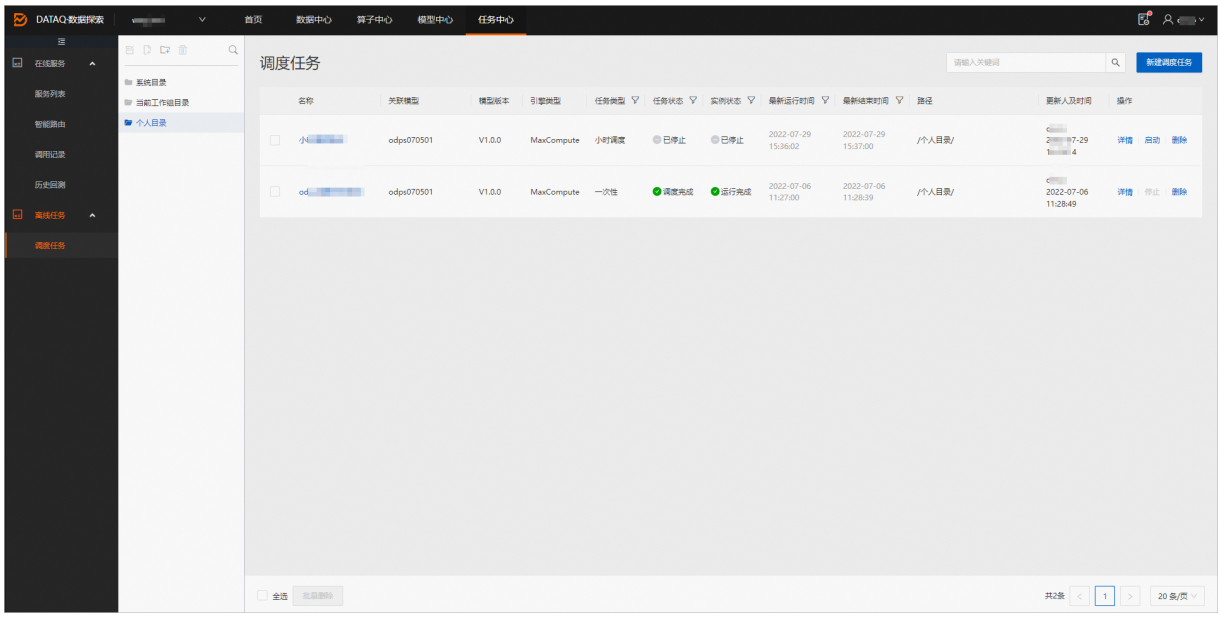
智能理由模块展示并管理用户添加的智能理由，在添加智能路由画布编排页面，编辑路由条件和版本号。


DATAQ·数据探索
首页
数据中心
算子中心
模型中心
任务中心
在线服务
服务列表
智能路由
调用记录
历史回溯
离线任务
调度任务
当前工作组目录
个人目录
智能路由
请输入关键词
请选择关联模型
新建路由
名称
关联模型
状态
路径
更新人及时间
操作
未发布
/当前工作组目录/test0609/
编辑 复制 删除
已发布
/当前工作组目录/
2022-06-16 14:40:30
查看 复制 删除
共2条
1
20 条/页


快照记录
D_0401165735646
保存
发布版本
系统算子
模型算子
请搜索
开始
分支
赋值
结束
搜索节点
请输入节点名称
新增调试
开始
分支
赋值
结束

离线任务

离线任务管理主要是对离线模型自动化调度运行进行任务式管理。每次任务调度时，都将按照模型设置好的参数和数据自动运行，每次任务都是一个实例，通过实例状态监控最近一次任务的健康情况（是否运行成功）。

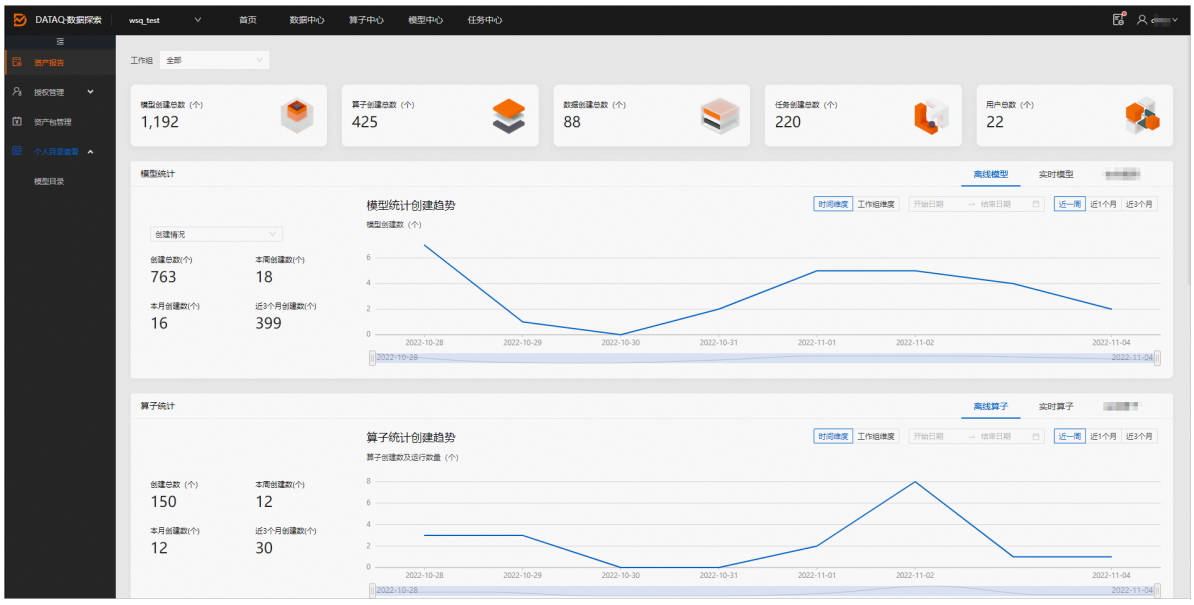


6.3.5.6. 用户中心

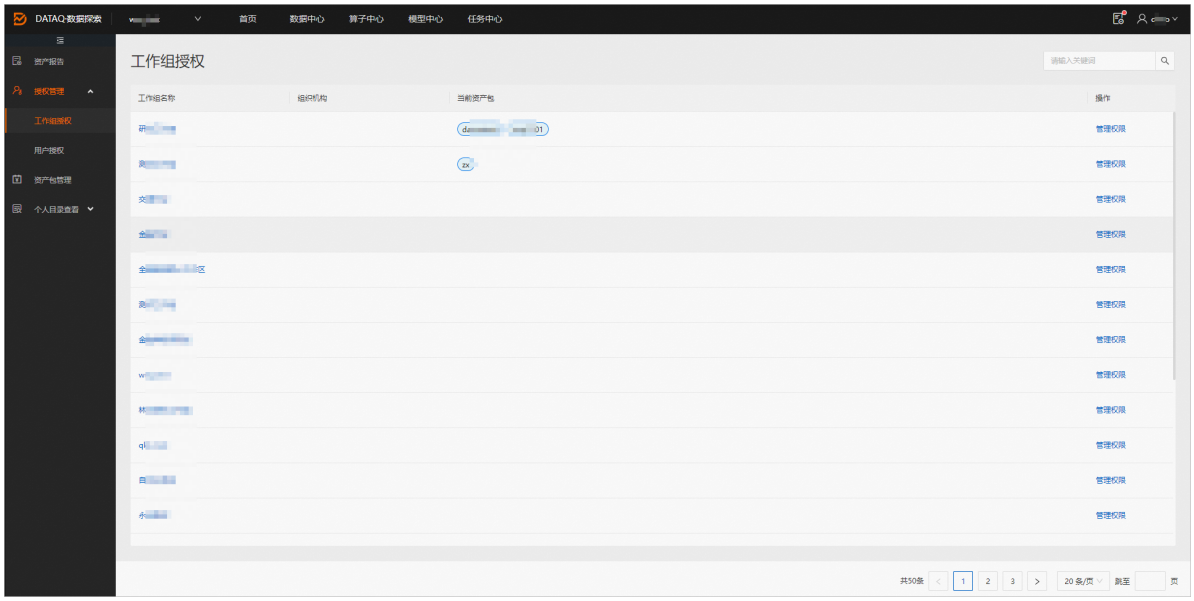
用户中心提供系统管理员视角的相关功能，包括资产中心、系统配置、标签列表、帮助中心。

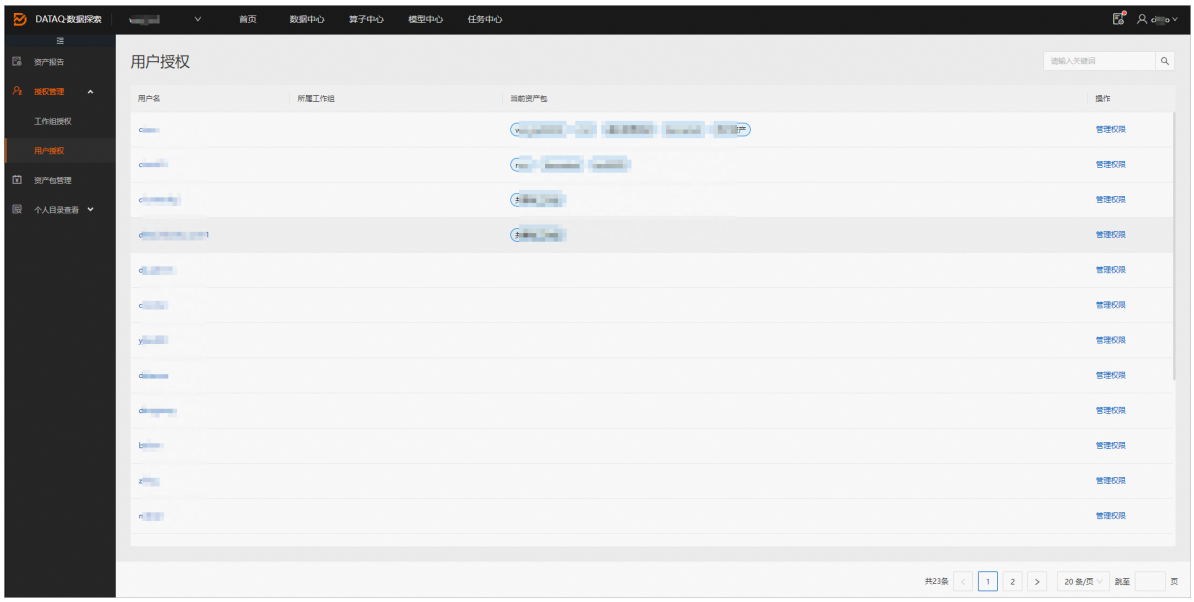
资产中心

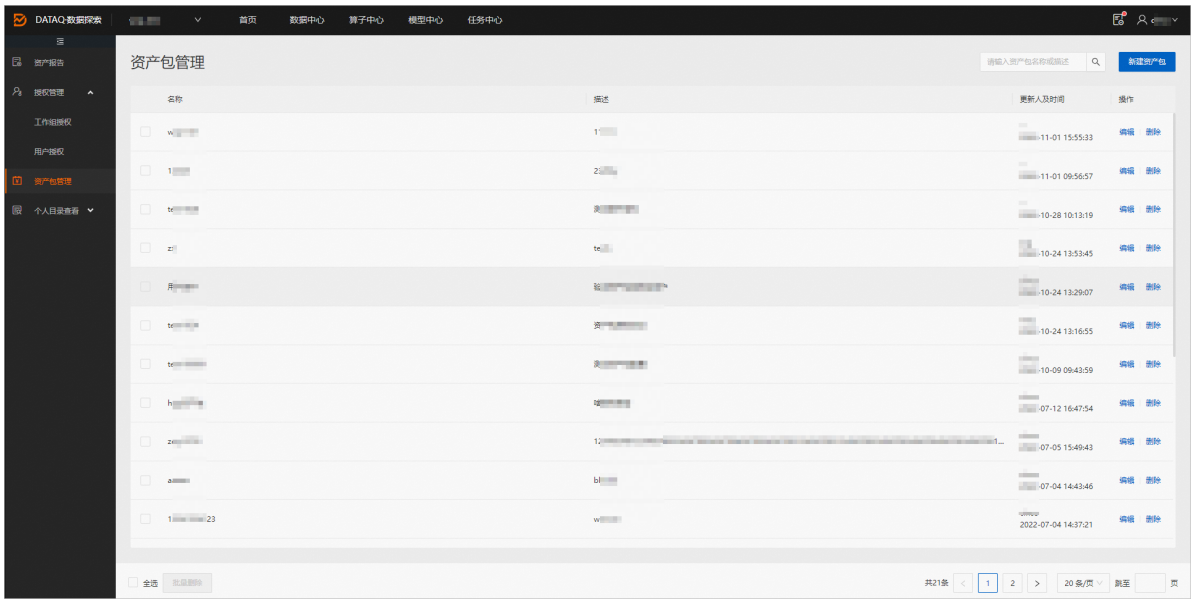
- 支持按工作组查看模型、算子、数据、任务、用户等资产的个数统计和趋势统计报告。



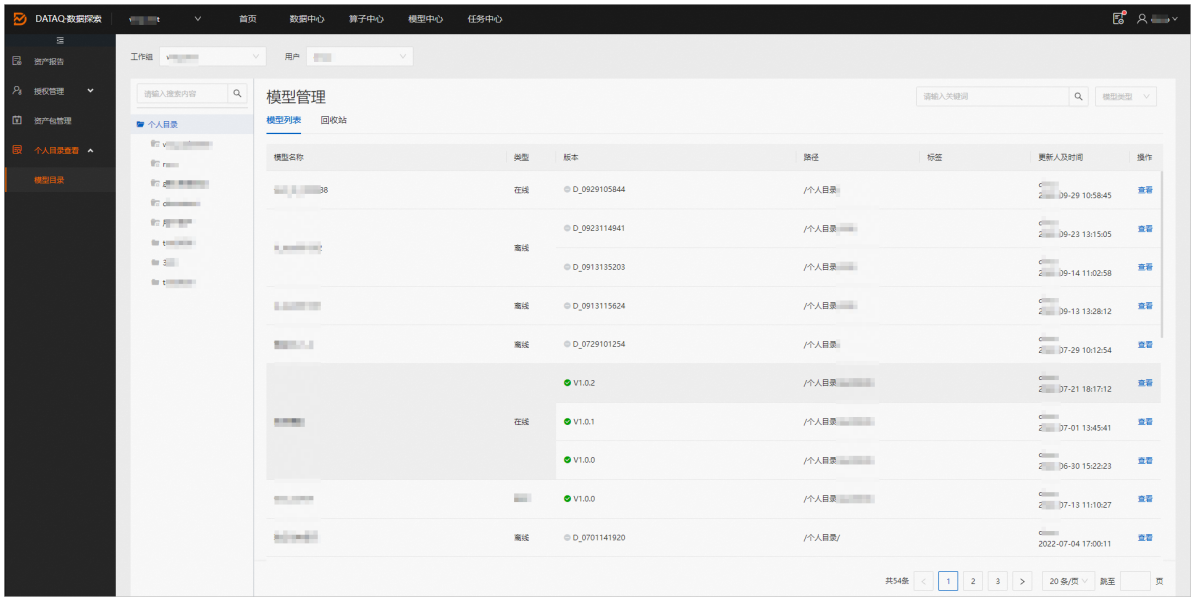
- 支持对数算子和模型按用户或工作组新建资产包，为工作组或用户添加对应的资产包，即拥有该资源下的使用权限。


DATAQ·数据探索
首页
数据中心
算子中心
模型中心
任务中心
资产报告
授权管理
工作组授权
用户授权
资产包管理
个人目录查看
工作组授权
工作组名称
组织机构
当前资产包
操作
管理权限


DATAQ·数据探索
首页
数据中心
算子中心
模型中心
任务中心
资产报告
授权管理
工作组授权
用户授权
资产包管理
个人目录查看
用户授权
用户名
所属工作组
当前资产包
操作
管理权限

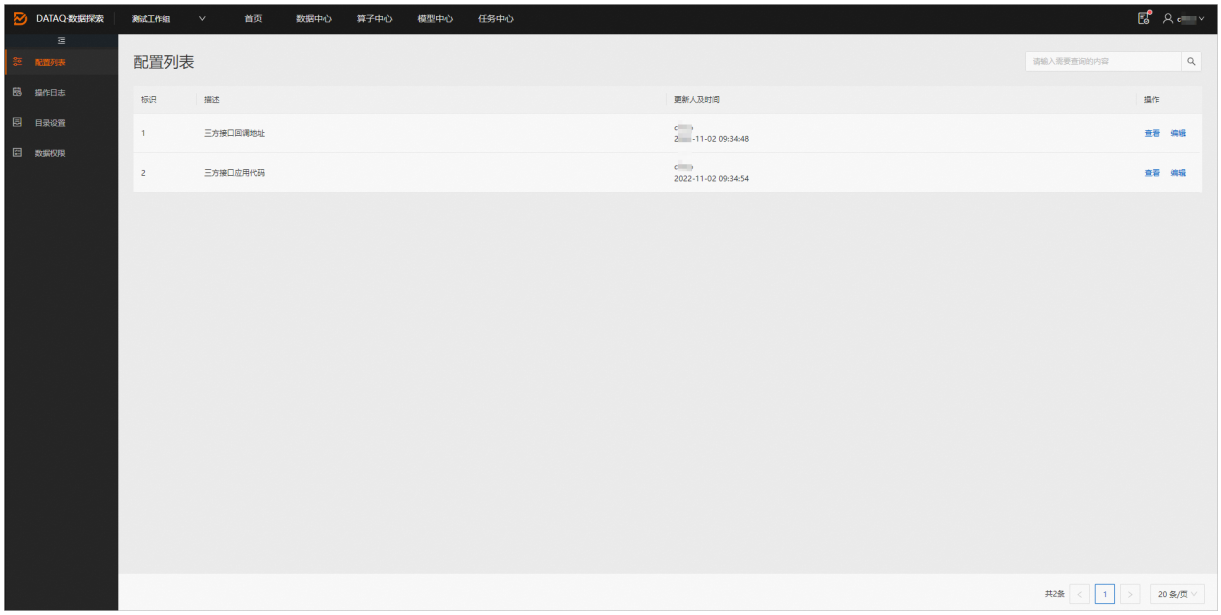

DATAQ·数据探索
首页
数据中心
算子中心
模型中心
任务中心
资产报告
授权管理
工作组授权
用户授权
资产包管理
个人目录查看
资产包管理
新建资产包
名称
描述
更新人及时间
操作
编辑
删除

- 支持租户管理员查看其他用户个人目录下的模型（按工作组区分），并支持查看模型详情。



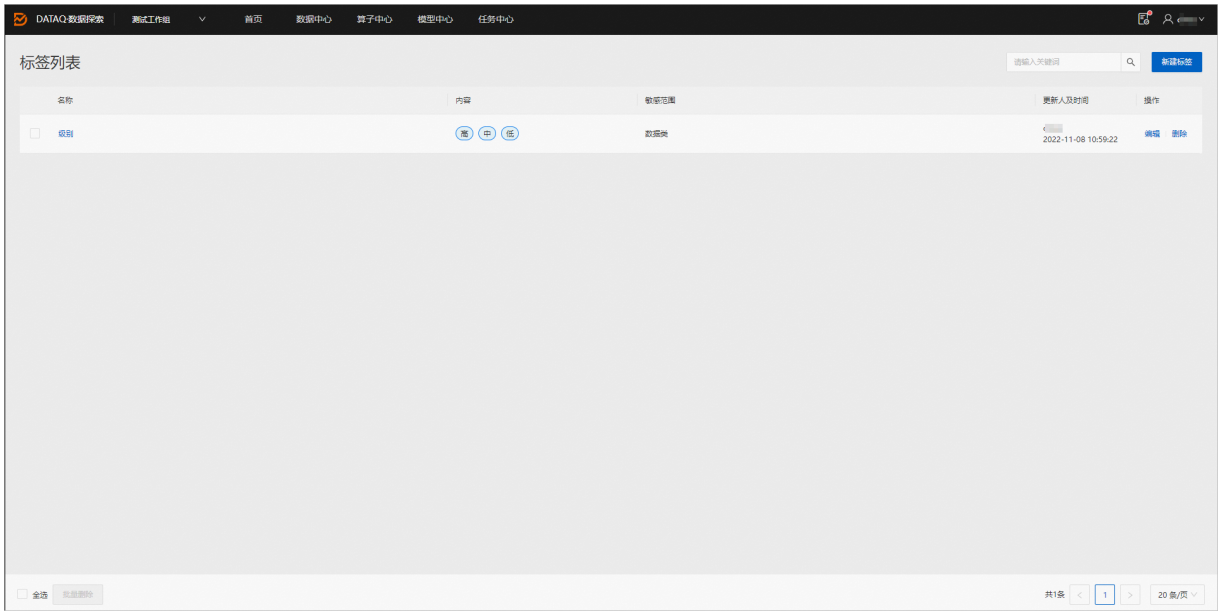
系统配置

支持对系统进行配置、查看操作日志、系统目录设置和数据权限设置。



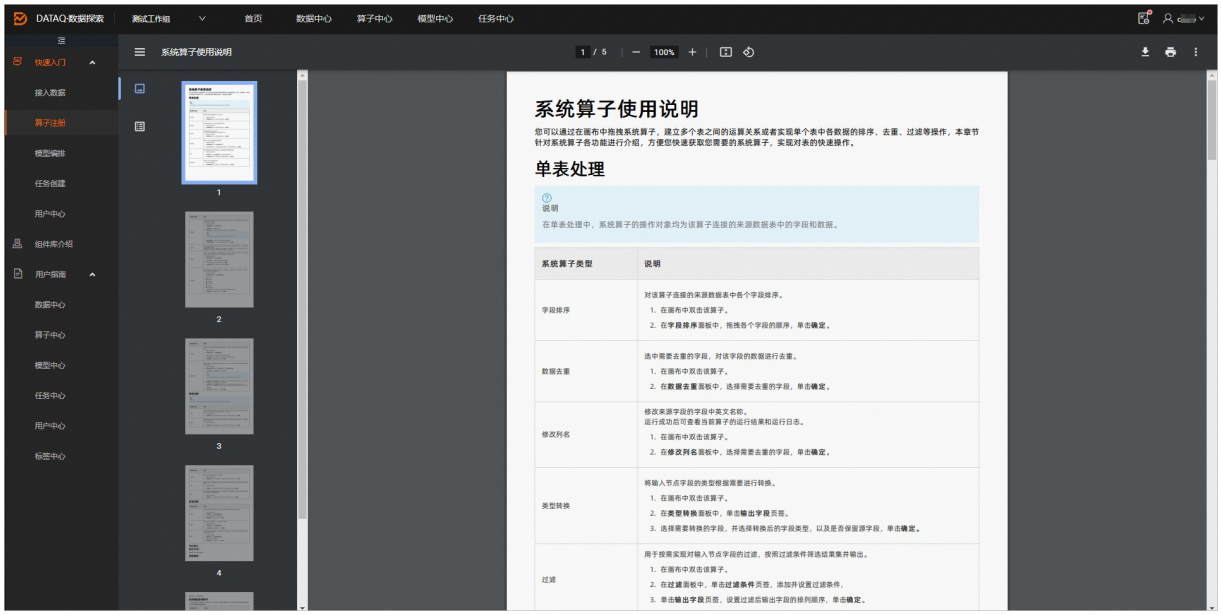
标签列表

支持新建标签，在进行数据、算子、模型管理时，通过绑定标签来描述其属性信息。



帮助中心

支持查看数据探索的帮助文档，包括快速入门、用户指南和模型组件库介绍。



6.3.6. 我的资产

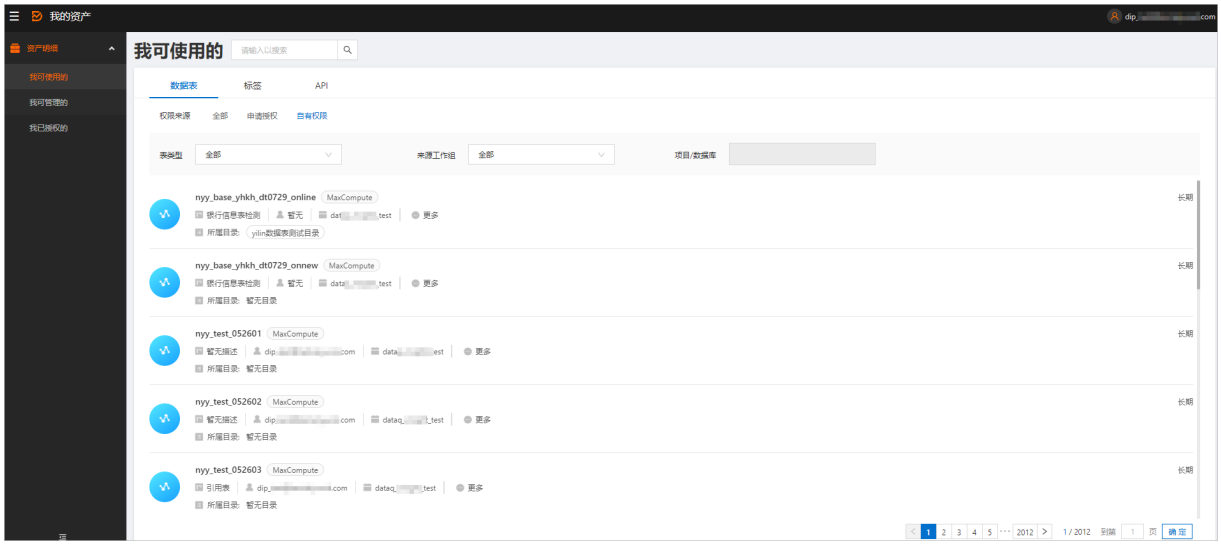
通过我的资产清晰掌握可使用、可管理、已授权的数据表、标签和数据服务API数据资产相关信息。

通过不同的权限来源，用户可获取数据表、标签、API等类型数据资产的使用权限。我的资产以用户的视角展示自有权限或申请授权成功的数据表、标签、数据服务API数据资产，并提供数据资产详情查看等功能，帮助用户清晰的了解在数据加工、数据分析等场景下可使用的数据资产情况。

具备特定角色（如工作组管理员）的用户对一定范围的数据资产有管理权限（如工作组管理员对本工作组的数据资产有管理权限）。我的资产展示用户有管理权限的数据表、标签、数据服务API等数据资产，并提供资产详情查看等功能，帮助用户清晰了解管理数据资产和已授权的数据资产情况。

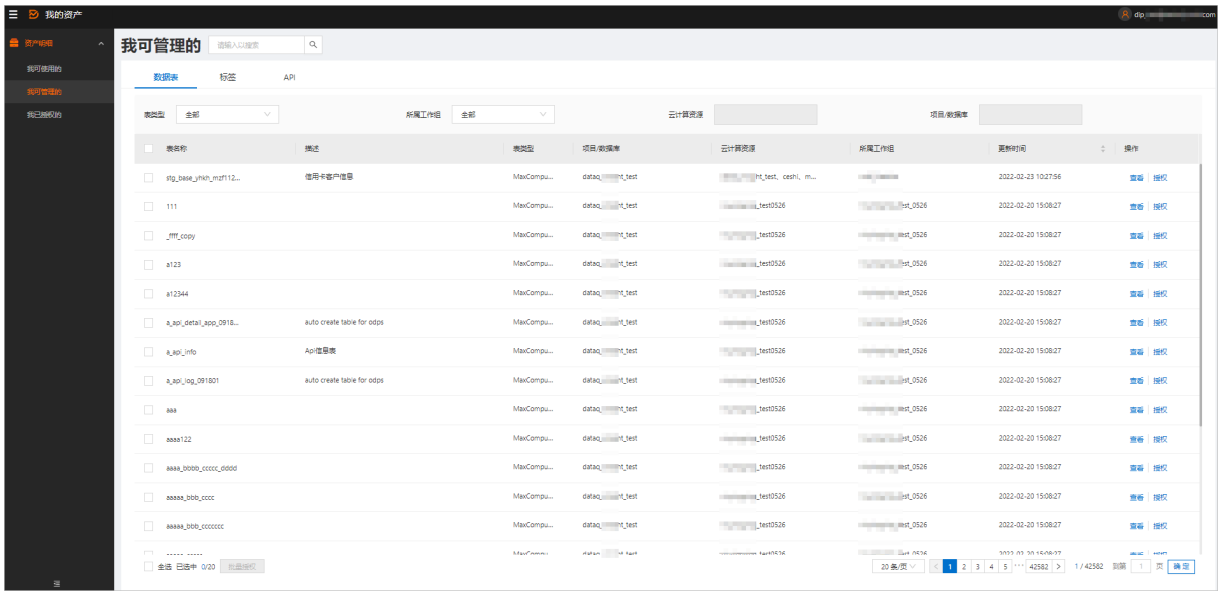
我可使用的

支持展示自有权限或申请授权成功的数据表、标签、数据服务API资产，并提供资产详情查看等功能。



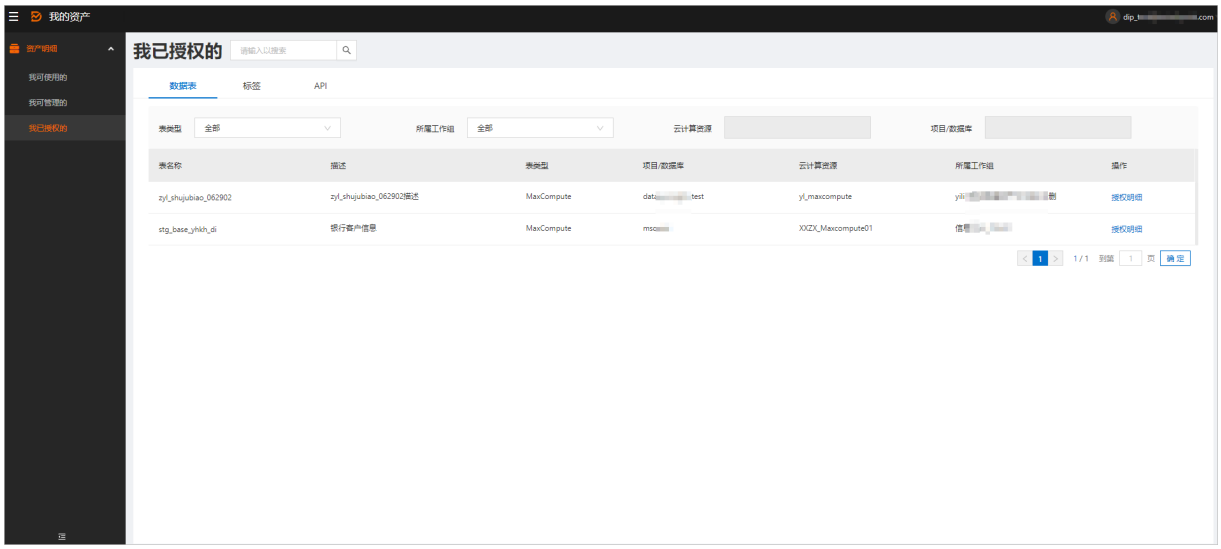
我可管理的

支持展示有管理权限的数据表、标签、数据服务API数据资产，并提供资产详情查看等功能。



我已授权的

支持对已授权的数据表、标签、数据服务API等数据资产进行列表展示，并提供查看授权明细与停止授权等功能。

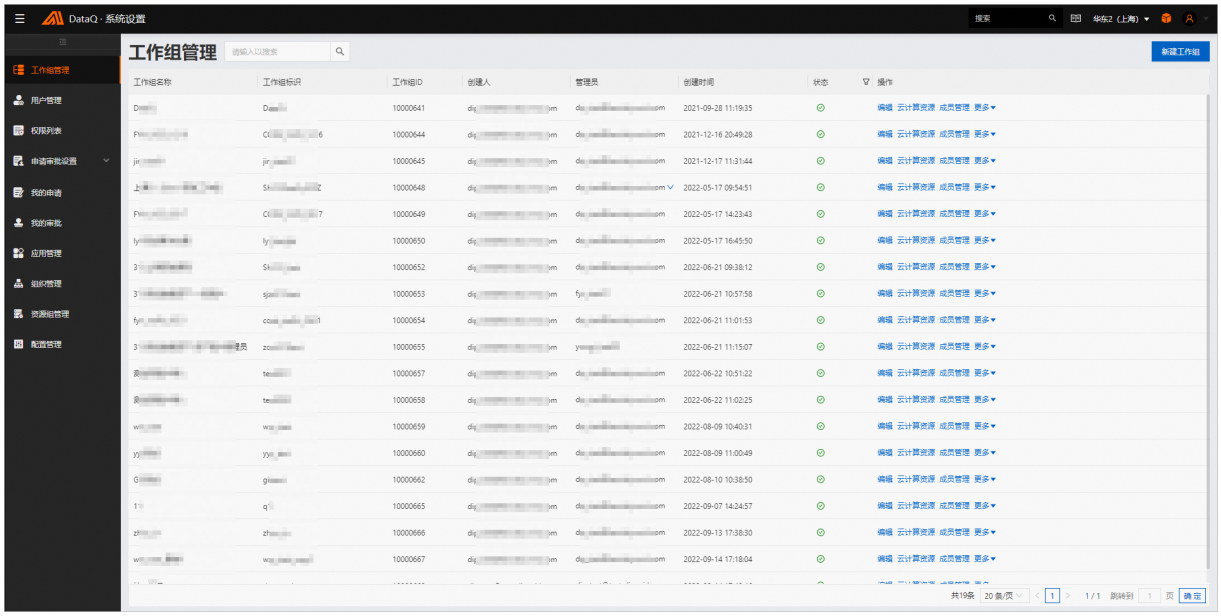


6.3.7. 系统设置

系统设置为用户提供了工作组、云计算资源、成员的创建及管理等能力，为数据资源平台的运行、使用提供了必要的基础保障。

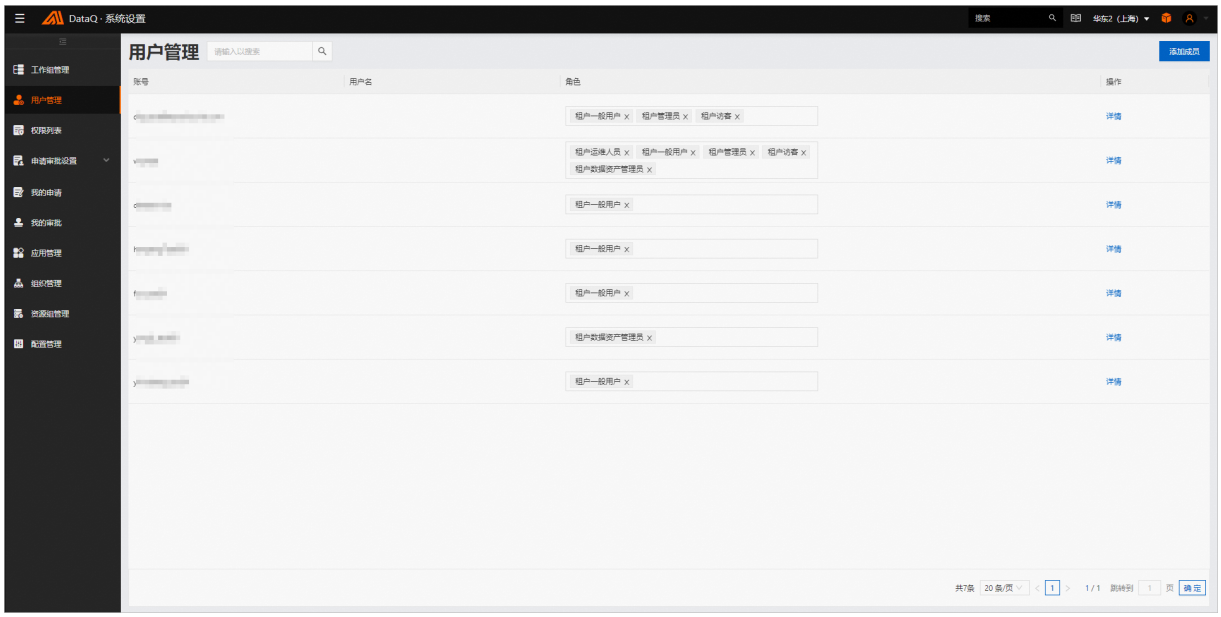
工作组管理

- **工作组管理：**支持工作组的创建、管理，以达到各工作组之间空间隔离的效果，并提供编辑、禁用、复制、删除等功能。
- **云计算资源管理：**支持多种类型的云计算资源的新建管理，并提供复制、编辑、检测、删除等功能。
- **成员管理：**支持对账号的相应工作组管理员、工作组开发者、工作组分析员进行配置，并支持移除成员、详情查看等功能。



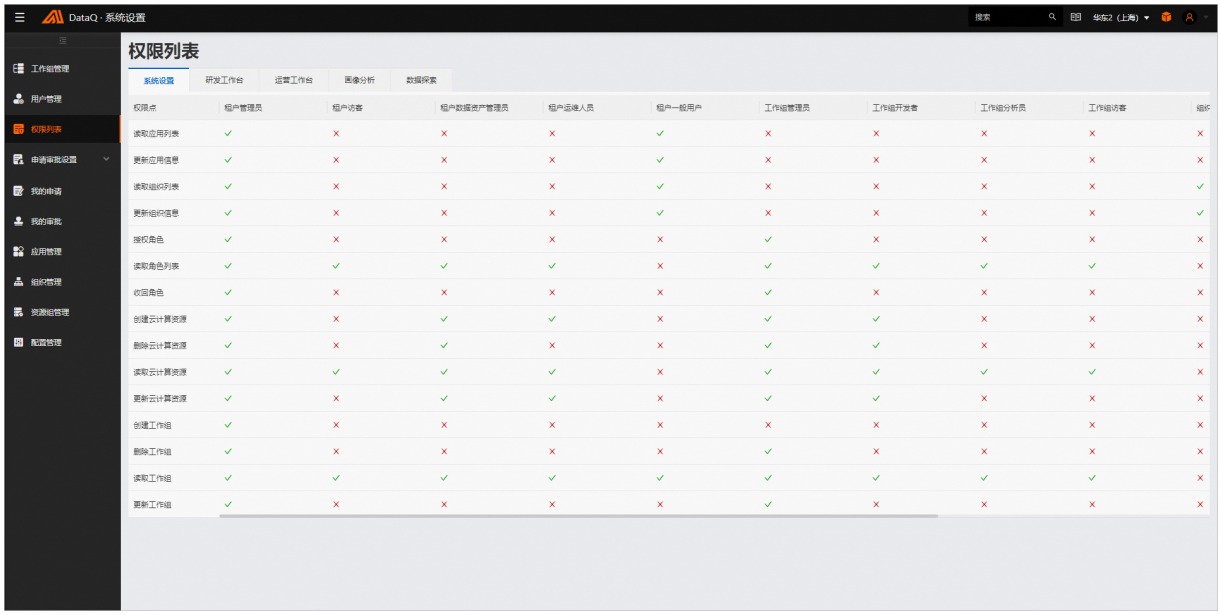
用户管理

支持对人员账号相应的租户管理员、租户访客、租户资产管理员、租户运维、租户一般用户角色等进行添加和管理。



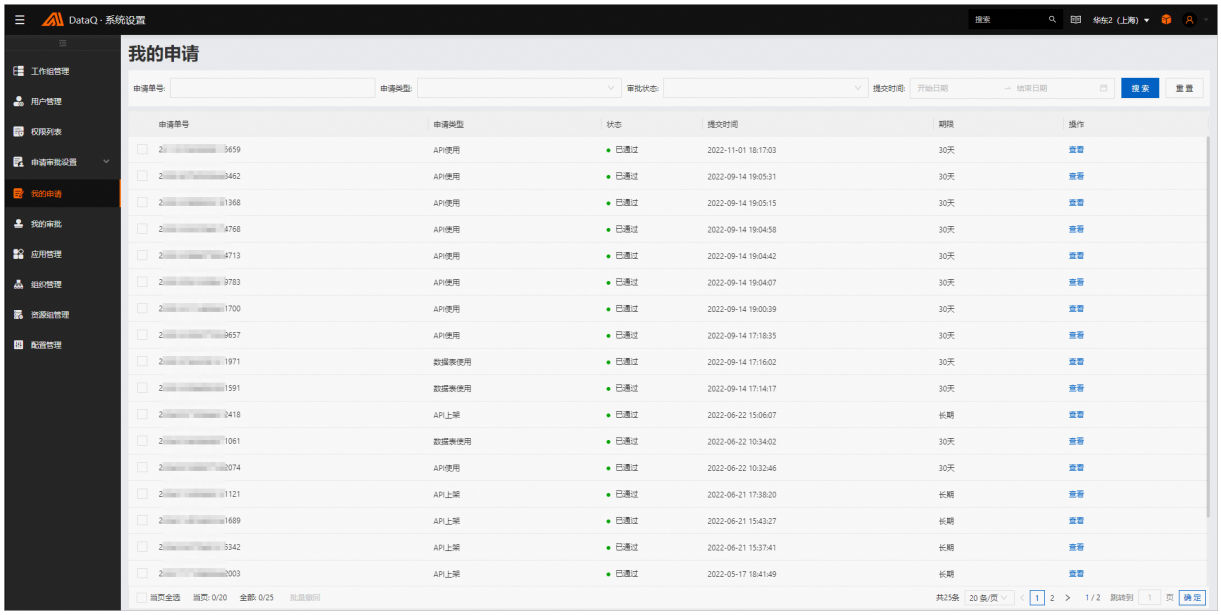
权限列表

支持对系统设置、 数据同步、 数据标准、 数据建模、 智能标签、 统一服务、 通用资源、 质量评估、 任务运维、 画像分析、 运营工作台、 数据资产中心的角色及权限点内容的展示。



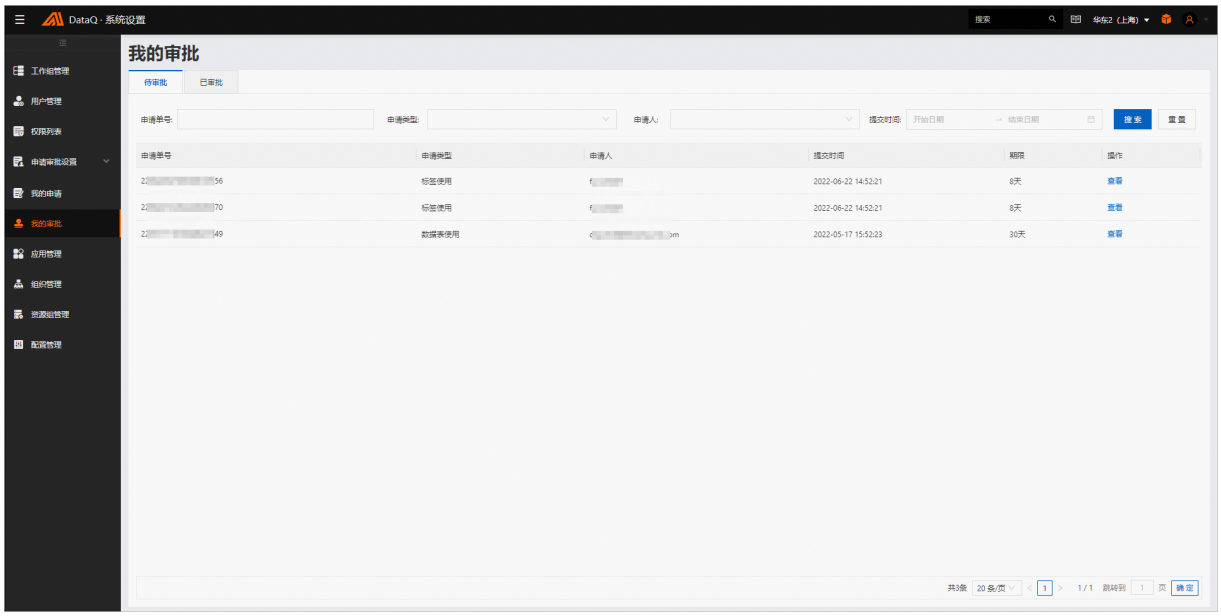
我的申请

支持查看已经提交的申请单，包括审批中和审批完成的。用户可以查看自己申请单的详细信息和审批结果，也可以按审批状态筛选申请单和撤回申请。



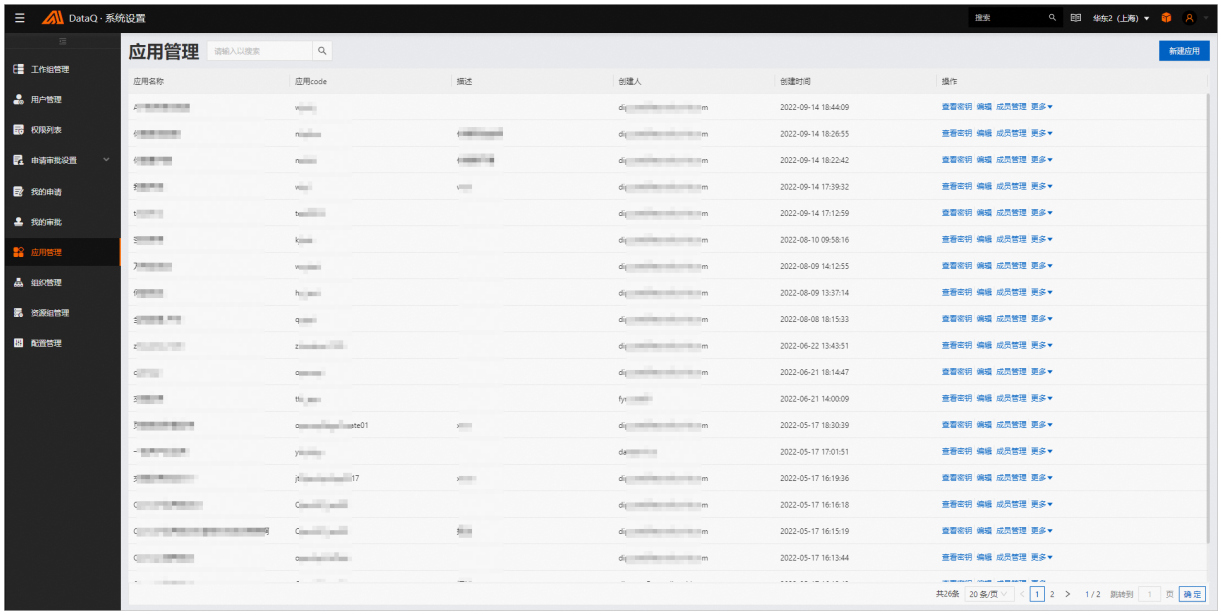
我的审批

展示需要当前用户待操作的申请单。可以查看申请单的详情，包括申请的数据资源信息、申请理由、申请使用时长、审批结果等。支持操作员对申请单的审批动作，支持通过、拒绝等功能。支持查看已审批和未审批的申请单。



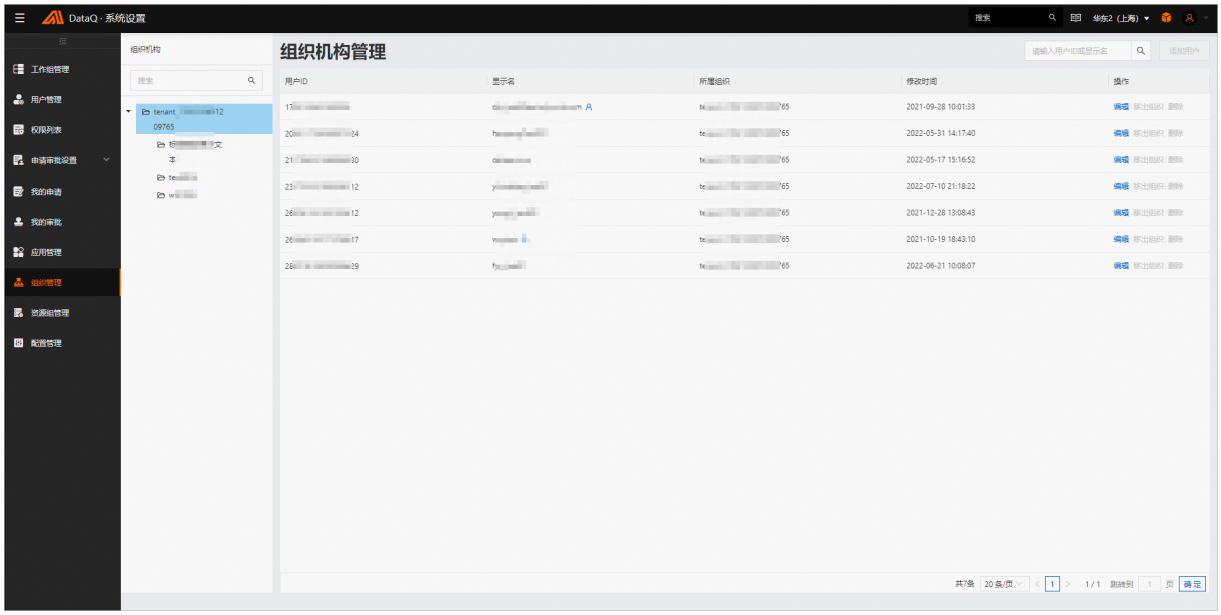
应用管理

支持应用创建、管理功能，支持应用的成员管理功能，提供应用密钥管理功能、OpenAPI授权功能。



组织管理

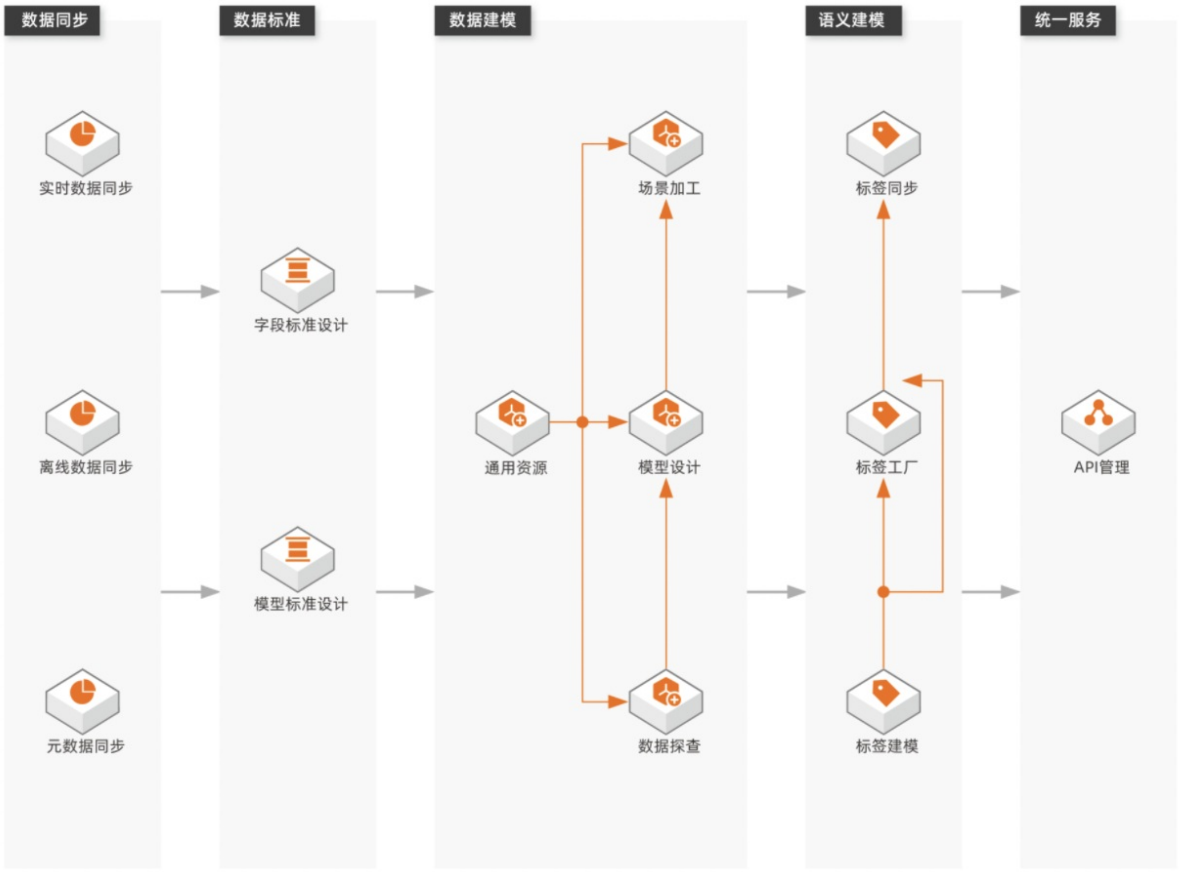
支持树形组织的创建、管理，支持组织内成员的添加、管理功能。



6.4. 应用场景

本文为您介绍数据资源平台的应用场景。

构建数据中台，行业领域模型快速沉淀



场景：快速构建数据仓库，有效治理数据质量，实现政企客户各部门数据的业务协同和共享。

痛点：以政务服务为例，政务服务部门众多，业务系统复杂、流程长，数据来源多且更新频率高；数据口径标准、数据准确性难以保障，传统政务服务需要老百姓多次递交多份纸质材料提供给不同服务部门，服务效率低。

解决方案：

- 数据同步：通过数据同步功能，将不同业务系统数据汇聚到统一的存储计算引擎，实现数据的初步融合。
- 数据标准：管理数据标准和构建数据模型，将数据标准贯彻到数据质量分析、保障及检查的全过程中，将散乱的多源异构数据加工成标准、干净的数据资产，确保数据的完整性、一致性、准确性和可用性。
- 数据建模：通过数据建模模块提供的各种数据开发工具，实现数据的清洗、加工和转换。
- 资产管理：通过资产运营功能实现资产的注册、编目、上架，在资产中心方便各个部门的使用人员搜索数据资源并申请数据权限。

价值：

- 数据标准一致：沉淀行业数据设计规范，保障数据质量。
- 高效资产管理：快捷数据资产搜索，全360度资产盘点，高效管理数据资产。

面向业务沉淀数据资产，客户画像分析助力精准营销



场景：传统金融行业客户寻求数字化转型，从产品销售导向逐渐转变为面向客户金融需求的服务导向，需要全面了解客户需求进行更精准的营销推广。

痛点：

- 运营人员获取高净值客群强依赖IT部门，相似指标IT重复开发，无法有效沉淀数据资产，提高开发效率。
- 多部门数据隔离，无法获得全面客户画像，重复营销成本高。
- 各网点数据无法进行行列权限隔离，安全风险高，数据共享难。

解决方案：

- 智能标签：通过智能标签实现面向业务的数据内容编排，搭建主题标签库，实现标签的闭环运营，沉淀业务经验。
- 标签加速配置：通过创建标签加速计划将各种不同数据源的标签数据进行同步，提升数据运维效率。
- 权限控制：依托智能标签及画像分析的工作组，实现信息中心数据面向各网点安全有序开放。

价值：

- 语义化的数据集市，大大降低数据获取和分析门槛，解耦IT人员和运营人员的工作，保证高频运营需求，提升运营效率。
- 高净值客群数据获取时间降低至小时级，灵活的圈群和分析工具，赋能精准营销；由手工Excel表获取目标用户列表的方式升级为系统接口实时获取。
- 通过交叉运营，利用不同部门之间的数据资产，降低拉新成本，提升拉新成功率。
- 实现仅信息中心IT人员可查看，到多网点运营及业务人员独立查看，实现行内数据普惠。

6.5. 基本概念

本文介绍数据资源平台的基本概念。

系统设置

租户：在数据资源平台产品中，租户为最高级权限隔离的命名空间，一个租户下可以有若干账号，不同账号可以分配不用的角色，数据可以在租户内共享，不同租户之间数据完全隔离。

工作组：数据资源平台中用于数据隔离的命名空间，通常为一个用户组、一个项目或一个应用。工作组下包含云计算资源、工作组成员。工作组内的数据通常为工作组内全部成员可见，研发工作台的数据权限以工作组为单元相互隔离。

数据同步

数据源：是指数据库应用程序所使用的数据库或者数据库服务器，在数据资源平台的数据同步中可以将配置好的数据源端数据同步到目标端。

多源异构数据：指的是多种不同结构的数据源。

数据模板：用于对非结构化数据的解释，当创建数据同步任务时，且源端数据类型为消息中间件时，通过数据模板对其进行解释。

Oracle CDC：实时数据同步模式，对源表进行INSERT、UPDATE或DELETE等操作的同时可以提取数据，变化的数据被保存在数据库的变化表中，通过订阅变化数据的方式实现数据的秒级同步。

MySQL Binlog：实时数据同步模式，通过同步并解析记录所有数据库表结构变更（例如CREATE、ALTER TABLE...）以及表数据修改（INSERT、UPDATE、DELETE...）的二进制日志，实现数据的秒级同步。

SQLServer CDC：实时数据同步模式，开启CDC的源表在插入、更新和删除活动时会插入数据到日志表中。CDC通过捕获进程将变更数据捕获到变更表中，通过CDC提供的查询函数，实现数据的秒级同步。

数据探查

内置算法：内置多种数据统计分析算法，根据不同字段类型采用不同的计算统计规则。

数据标准

数据标准：是对含义相同但字段名称不同的数据进行统一规范管理的数据准则，数据标准可定义逻辑表的命名规范、字段的取值范围、度量单位等内容。

数据元：表示一个数据的最小信息项，包含基本的描述、值域范围和约束规则。

生命周期：指的是表的生命周期，指表（分区）数据从最后一次更新的时间算起，在经过指定的时间后没有变动，则此表（分区）将被数据库（例如：MaxCompute）自动回收。这个指定的时间就是生命周期，生命周期的单位一般是天。

元素的长度：数据元的长度。

固定值：是指数据质量运行的一行一列结果与期望的固定值进行比较，根据比较条件判断质量结果执行是否通过。

趋势：指的是质量规则设定的模式，分为固定值、固定值和固定值比较，以及固定值波动率比较。

数据建模

数据模型：是数据特征的抽象，它从抽象层次上描述了系统的静态特征、动态行为和约束条件，为数据库系统的信息表示与操作提供一个抽象的框架，数据资源平台支持逻辑模型和物理模型。

- **逻辑模型：**反映的是系统分析设计人员对数据存储的观点，是对概念数据模型进一步的分解和细化；数据资源平台中可通过手动、DDL语句、EXCEL导入、云计算资源扫描的方式来实现逻辑模型创建。
- **物理模型：**是对真实数据库的描述；数据资源平台中通过将已创建好的逻辑表物理化到数据库来实现物理模型的创建。

- **质量规则：**是指在特定业务环境下，用户定义数据符合使用目的一组定量或定性的规定要求；在数据资源平台中，可以通过对表级与字段级的数据质量任务运行，真实反映数据接入的唯一性、准确性、规范性、一致性、时效性、完整性。

数据开发

场景：场景定义数据输出的方式，一个场景包含一种或多种数据输出的方式。

计算资源：可以执行计算节点的资源。使用“场景加工”时并不会单独执行一个计算节点任务，需要把计算节点组织成流程执行，所以需要计算资源支持流程的批量执行。

存储资源：可以存储数据的资源，包括离线存储、在线存储。

计算节点：场景加工中的计算节点由五要素组成：计算资源、代码、输入数据、输出数据和参数。代码可在计算资源上执行（代码是否需要编译由计算资源决定）；输入数据、输出数据由逻辑表定义；参数用于控制计算节点的执行。在使计算用节点配置任务流程时，流程中连接上下游的连接线表示计算节点的输入和输出数据，而节点参数需要通过单独的界面进行配置。

节点任务流程：通过有向无环的流程图（DAG）描述，定义了一段处理逻辑。节点任务流程中所有计算节点的输入、输出数据都是具体的数据表、文件和数据流。

开发场景与生产场景：每个工作区都有一个开发场景和一个生产场景，它们是两套独立的资源配置。数据开发工程师只能在开发环境中注册计算节点，配置节点任务流程，并进行开发。开发通过后要经过上线操作才能把开发环境中的节点任务流程发布到生产环境中。任何人都不能编辑生产环境中的节点流程，但可以修改参数，所有修改都会记录在日志中。

智能标签

物理表：是指具体某个数据源中的一张表。

主表：在数据库中建立的表格，其中存在主键（primary key）用于与其它表相关联，并且作为在主表中的唯一性标识，例如：银行客户信息表（证件号码、姓名、出生日期、电话号码、卡号），信用卡消费信息表（卡号、卡额度、办卡渠道），其中信用卡消费信息表中的卡号为银行客户信息表的外键，其与信用卡消费信息表中卡号属性相关联，因此，信用卡消费信息表为主表，银行客户信息表为从表。

加速表：是指加速到目标云计算资源而生成的物理表。

标签模型：是围绕实体（Object）、关系（Link）、标签（Tag）三大元素对分布在不同数据库中的数据进行网络化的建模方式。

- 实体用于描述某个客观的对象，如设备人员地址等，对应到物理数据表上一般就是属性表，有一个主键来代表每一个对象，剩下的每一列就是标签即描述对象的属性。
- 关系是表示对象和对象之间的联系、事件、行为，一般对应到物理数据表上一般就是事实流水表。

标签：是指对用户这一对象的一个最小描述单元，代表着所描述对象某一个具体的客观事实的抽象表达，如属性（教育水平：职业高中、本科、专科、研究生），行为（成交金额、收藏次数、位置定位），或者是兴趣（对于多个关键词的偏好度），是一种以业务视角出发的数据建模方法，标签既可能是数值、也可能是枚举值，也可以是多个Key-Value组织的列，还可能是多字段组成的事实表（如对象、时间、谓语、宾语）。

标签工厂：是指在数据资源平台中对衍生标签进行生产、加工的功能模块。

值码表：数据库相关字段编码值的含义，例如：性别字段代码表的值【M】代表 【男】，【F】代表【女】。

任务运维

任务流：由数据同步节点、离线计算任务节点、流式计算任务节点、自定义计算节点等组成的数据加工或算法服务工作流。

业务类型：提交任务流的来源，包括数据同步、数据开发、标签加工、标签同步等。

任务类型：提交任务流的来源，包括数据同步、数据开发、标签加工、标签同步等。离线同步任务、离线计算任务的任务类型，主要分为手动、周期、补数据等。

节点类型：任务流中包含的同步节点、MaxCompute_SQL节、Flink_Vvp节点、自定义计算节点等在不同计算资源上运行的计算任务节点。

补数据计划：通常情况下，对于离线周期任务，当前序任务失败或任务第一次上线，需要通过补数据计划同步或处理当前业务日期之前数据。

画像分析

群体：分析对象上满足一定条件的单体的集合。

群体计算：同一个实体的群体和群体之间的交集、差集和并集计算，生成新群体。

资产运营

资产目录：是对数据资源平台元数据的有序组织，是记录数据体系的保障。数据类目是目录信息与服务、保障与支撑所组成的一个整体。针对资产的合理组织，需要对资产进行对应分类，如部门类目、主题类目、行业类目，可进行该类目下资产数据进行搜索、展现以及权限申请。

数据资产：数据资源平台中存在大量的数据表、标签、API等各类数据资产，数据管理者通过数据汇聚、数据治理、数据分析后，需要对整个平台数据进行统一管控，了解平台的核心数据资产，提供对应的数据资产管理规范。

统一服务

应用：是调用API服务时的身份。每个应用有一组AppKey和AppSecret，可以理解为账号密码，调用API的时候需要将AppKey做参数传入，AppSecret用于签名计算，计网关会校验这对密钥对您进行身份认证。

API：应用程序编程接口，是一些预先定义的函数，目的是提供应用程序与开发人员基于某软件或硬件得以访问一组例程的能力。

数据API：API的源头为数据表，可以是单表也可以是多表，请求后最终以类SQL的形式到数据库查询并返回符合条件的数据，将数据组装成报文格式返回。

应用API：API的源头为数据表，可以是单表也可以是多表，请求后最终以类SQL的形式到数据库查询并返回符合条件的数据，将数据组装成报文格式返回。

6.6. 使用限制

本文为您介绍数据资源平台产品各模块的使用限制。

系统设置

目前角色包括：租户管理员、租户访客、租户数据资产管理员、租户运维人员、工作组管理员、工作组开发者、工作组分析员、工作组访客、租户一般用户。

资产管理

系统目前支持：

- **数据表：**MaxCompute、DataHub、Hive云计算资源类型。
- **标签：**RDS MySQL、MaxCompute云计算资源类型。
- **API：**RDS MySQL、MaxCompute 、PostgreSQL、Table Store、DataHub、Object Storage Service、AnalyticDB PostgreSQL、Hologres云计算资源类型。

数据同步

- 目前系统支持的任务链路：

链路类型	数据源/目标端	Oracle	SQLSever	Postgres	MySQL	ADB PG	MaxCompute	DataHub	Hive	OSS	Kafka	HBase(Phoenix)	TableStore	Hologres(不支持关键字)	ADB MySQL	Neo4j(4.3)	ES	Greenplum	MINIO	POSTGIS PG	GANOS PG
离线链路	Oracle(10g, 11g)	✔	✔	✔	✔	✔	✔	✔	✔												
	SQLSever(2008, 2014, 2016, 2017)	✔	✔	✔	✔	✔	✔	✔	✔												
	Postgres(9.4)	✔	✔	✔	✔		✔		✔												
	MySQL(5.6, 5.7)	✔	✔	✔	✔	✔	✔	✔	✔						✔						
	HDFS(Hadoop 2.7.1, 2.5.0, 2.6.0)								✔												
	Hive(1.1.1, 1.2.2, Hadoop 2.7.1, 2.5.0, 2.6.0)				✔(API)		✔					✔(API)									
	FTP									✔											
	MaxCompute			✔	✔	✔	✔						✔	✔	✔	✔	✔			✔	
	ADB PG				✔	✔	✔								✔						
	MINIO					✔												✔			
	KUDU						✔														
	SDXPLUS PG																			✔	✔
	ARCSDE ORACLE																			✔	✔
	POSTGIS PG																			✔	✔
	GANOS PG																			✔	✔
	OSS			✔	✔	✔									✔						
实时链路	Oracle(10g, 11g)							✔			✔										
	SQLSever(2008, 2017)							✔			✔										
	MySQL(5.6, 5.7)							✔			✔										
	PUSH API																				
	Kafka (2.3, 2.4, 2.5)			✔		✔		✔											✔		
	RabbitMQ							✔													
	PostgresqlCDC			✔		✔															
	DM							✔													

> 说明
> - “√”代表支持，其他则为不支持。
> - 目标端支持版本与源端一致。

- 目前系统支持的批量任务链路：

链路类型	数据源/目标端	Oracle	SQLSever	Postgres	MySQL	ADB PG	MaxCompute	DataHub	ADB Mysql
离线链路	Oracle(10g, 11g)	✔	✔	✔	✔		✔		
离线链路	SQLSever(2008, 2014, 2016, 2017)	✔	✔	✔	✔		✔		
离线链路	MySQL(5.6, 5.7)	✔	✔	✔	✔	✔	✔		✔
离线链路	ADB PG				✔	✔	✔		✔
实时链路	Oracle CDC/LogMiner (10g, 11g)							✔	
实时链路	MySQL binlog(5.6, 5.7)				✔			✔	

> 说明
> - “√”代表支持，其他则为不支持。
> - 目标端支持版本与源端一致。

数据查询和自助分析

系统目前支持：MaxCompute、RDS MySQL、AnalyticDB PostgreSQL、Hologres、Hive云计算资源类型预览和探查。

逻辑表物理化

系统目前支持：MaxCompute、DataHub、Hive、ADB for PostgreSQL云计算资源类型。

场景

- **计算节点：**目前系统支持MaxCompute、ZergStandalon、ZergService、Hive、Blink、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。
- **读写数据：**目前系统支持MaxCompute、Hive、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。
- **数据同步：**目前系统支持MaxCompute、ApsaraDB for RDS、Hive、Table Store、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。

智能标签

目前系统支持：RDS MYSQL、AnalyticDB PostgreSQL（包括2.0和3.0版本）、MaxCompute、Hologres、AnalyticDB MySQL、PostgreSQL云计算资源类型。

画像分析

目前系统支持：RDS MYSQL、AnalyticDB PostgreSQL（包括2.0和3.0版本）、MaxCompute、Hologres、AnalyticDB MySQL、PostgreSQL云计算资源类型。

统一服务

目前系统支持：RDS MySQL、MaxCompute 、PostgreSQL、Table Store、DataHub、Object Storage Service、AnalyticDB PostgreSQL、Hologres 云计算资源类型。

7.DataV数据可视化 产品简介

7.1. 什么是DataV数据可视化

DataV数据可视化平台旨在让更多的人看到数据可视化的魅力，帮助非专业的工程师通过图形化的界面轻松搭建专业水准的可视化应用，满足会议展览、业务监控、风险预警、地理信息分析等多种业务的展示需求。此外，DataV推出高性能三维渲染组件 DataV.GL、天猫双十一大屏同款可视化工具等将三位渲染能力引入地理场景，借助 GPU 算力实现海量数据渲染，提供低成本、可复用的三维可视化方案，适用于智慧城市、智慧交通、安全监控、商业智能等场景。

背景信息

相比于传统图表与数据仪表盘，如今的数据可视化致力于用更生动、友好的形式，即时呈现隐藏在瞬息万变且庞杂数据背后的业务洞察。无论在零售、物流、电力、水利、环保、还是交通领域，通过交互式实时数据可视化大屏来帮助业务人员发现并诊断业务问题，越来越成为大数据解决方案中不可或缺的一环。

创造一个可视化应用，您可能会面临如下几个问题：

- 对于数据可视化的设计无从下手，团队内的设计师对于复杂数据的展现经验不足。
- 对于数据可视化的实现比较困难，设计出来的很多图表与特效开发耗时耗力。
- 对于非传统报表数据类型（如时空数据、关系数据）的分析展现，缺乏相关的组件或工具支持。
- 对于在大屏幕上的展示，总会遇到分辨率适配的种种问题。
- 对于大数据的分析展示，现有工具无法很好支撑。

为什么选择DataV数据可视化

使用DataV数据可视化，您可以：

- 开发天猫双11、阿里云城市大脑同款数据大屏。

  DataV提供了满足各类场景和人群需求的可视化大屏工具，开箱即用，效果惊艳。

- 开发工业级的数据可视化项目。

  DataV可以面向软件开发商和开发者，提供更灵活的项目权限管理方案、自定义组件开发工具和强大的交互配置能力。

- 使用高性能的三维渲染引擎。

  DataV将游戏级三维渲染的能力引入地理场景，借助GPU计算能力实现海量数据渲染，提供低成本、可复用的三维数据可视化方案，适用于智慧城市、智慧交通、安全监控、商业智能等场景。

- 使用海量的炫酷图表组件。

  DataV支持绘制各类基础图表，接入ECharts等第三方图表库，即便没有设计师，也可搭建出高水准的可视化应用。

- 搭建专业级地理信息可视化大屏。

  DataV支持绘制地理轨迹、地理飞线、热力分布、地域区块、3D地球等效果，支持地理数据多层叠加。

- 使用行业模板创建大屏。

  DataV定制了数十种行业数据模板，用户简单修改即可使用，业务全景一目了然。

- 将多种数据源接入DataV大屏中进行展示。

  DataV支持接入包括分析型数据库、关系型数据库、本地CSV上传等数据源，支持动态请求。

- 轻松搭建可视化大屏。

DataV提供图形化编辑页面，使用拖拽的方式即可完成样式和数据配置，无需编程就能轻松搭建数据大屏。

- 灵活部署和发布可视化大屏。

  DataV可视化大屏可以适配非常规拼接大屏，支持加密发布。

7.2. 产品优势

本文介绍DataV产品的优势。

“一站式”开箱即用数据可视化解决方案

DataV数据可视化平台致力于提供“一站式”开箱即用数据可视化解决方案，帮助非专业的数据工程师通过图形化的界面轻松搭建专业水准的可视化应用。

专业级大数据可视化

专精于地理信息与业务数据融合的可视化，提供丰富的行业模版和图表组件，处理与展示百万级的复杂数据。

图形化编辑界面

拖拉拽的图形化配置方式，通过简单拖拽即可完成样式和数据配置，无需编程就能轻松搭建大屏。

智能主题配色

数据可视化可以通过智能主题功能，无需专业的大屏设计师，也能对大屏进行合理的配色，快速解决在设计大屏时遇到的配色困难的问题。

一键美化样式

数据可视化可以通过一键美化功能，快速调整可视化应用的布局，并通过内置样式丰富可视化应用的内容，快速解决在设计可视化应用时遇到的整体样式配置困难的问题。

工具箱滤镜功能

DataV工具箱提供了滤镜配置的功能。通过滤镜配置，可以对大屏中组件的色相、饱和度、亮度、对比度以及透明度等颜色属性进行配置。

移动端创建生成

DataV移动端专用编辑器，专业的移动端可视化体验。一键PC转移动端，快速获得移动端可视化应用，帮助您随时在手机查看数据应用。

多种数据源支持

支持多种数据源接入，充分发挥阿里云大数据计算的能力，提供多种图表组件支撑多种数据类型的展示，接入包括阿里云分析型数据库、关系型数据库、本地CSV上传和在线API等，支持动态请求。

专业级地理信息可视化大屏

DataV支持绘制地理轨迹、地理飞线、热力分布、地域区块、3D地球等效果，支持地理数据多层叠加。

支持组件二次开发

平台提供二次开发环境，开发者根据规范可以将自己开发的Javascript组件接入到DataV产品中，二次开发组件也可以拖拽布局与数据接入。

支持蓝图编辑器交互

蓝图编辑器一般称为visual programming或者flow based programming 。通过蓝图编辑器功能，您可以使用可视化连线的方式，定义图层与图层之间的交互行为。使得用户可以自由管理大屏中多个组件之间的交互关系。

蓝图编辑器的优势：

- 蓝图编辑器区别于回调ID，可以保证交互和数据的实时性和同步性。
- 蓝图编辑器支持数据请求合并和数据分发的功能。
- 蓝图编辑器可模块化拆分，专注单个的交互链路，不需要考虑代码的整理和规范，只需要专注于业务规则和交互需求即可。
- 蓝图编辑器支持自研组件和第三方组件。该功能对开发者型用户较为便利。用户自行开发的第三方组件，必须是在遵循开发者规范的前提下，才能在DataV中的蓝图编辑器页面无缝衔接使用。

灵活部署和安全发布

针对拼接大屏端的展示做了分辨率优化，同时创建的可视化应用能够发布分享，作为用户对外数据业务展示的窗口。适配非常规拼接大屏，支持本地部署，加密发布时通过密码/Token的方式进行访问权限控制，保护数据隐私安全。

专业团队解决技术难题

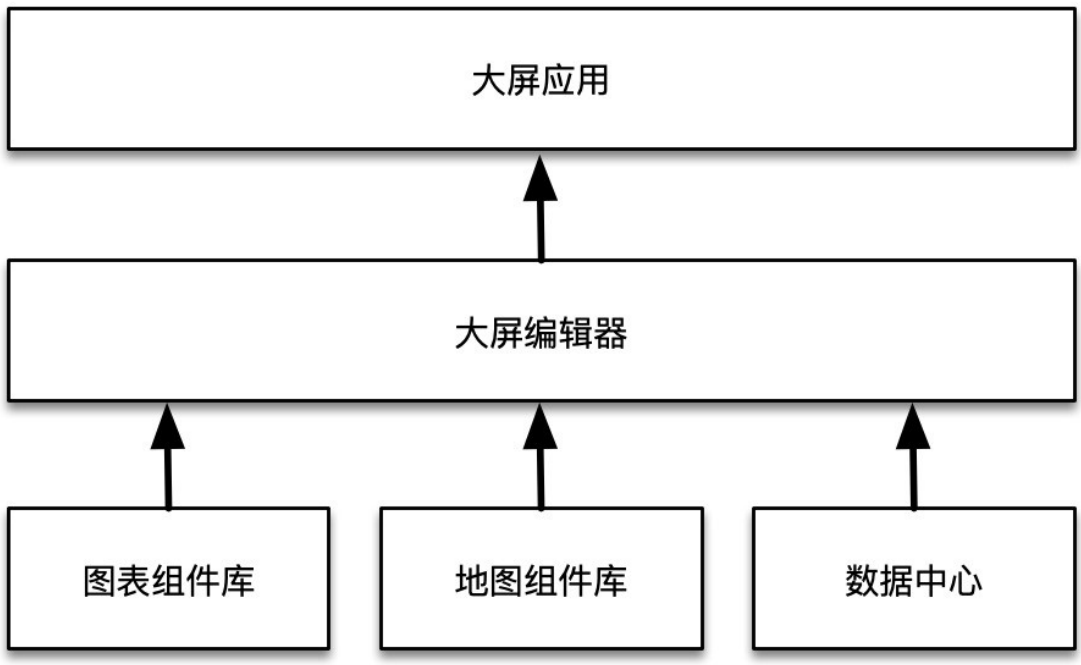
专业团队帮助您解决大屏设计和搭建、数据对接、交互配置等一系列技术难题。

7.3. 产品架构

7.3.1. 功能架构

DataV数据可视化平台是以日常图表组件库d.chart，以及地理相关组件库d.map为基础，专业数据可视化模板设计为视觉框架，依托cube前端框架搭建起来的Web服务。

通过DataV数据可视化平台，您可以使用已有组件，低成本还原设计师产出的原型图，快速完成多种数据源的绑定。并通过最终的可视化配置调整系统完成所见即所得的视觉调试，最终产出数据大屏。

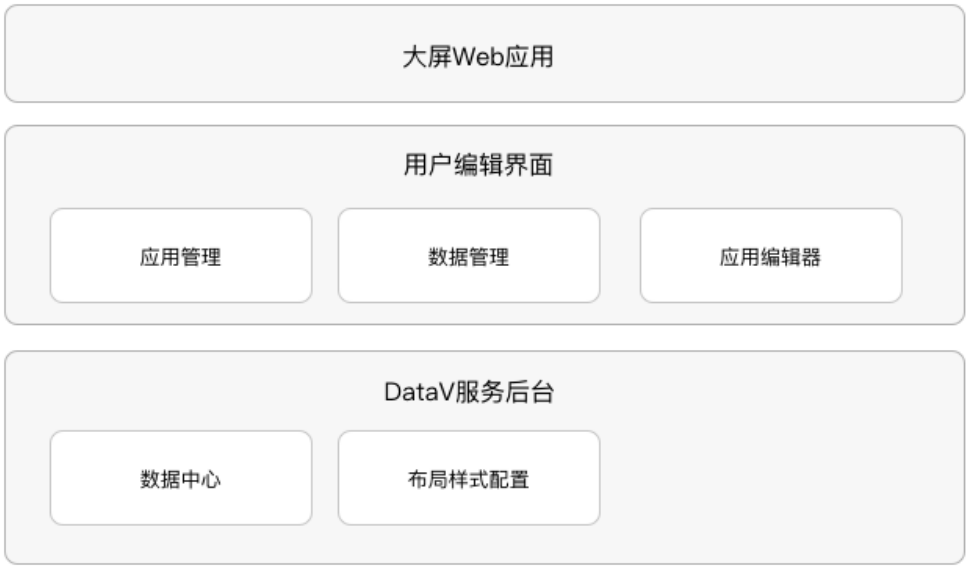


- 组件库：多种图表组件，支撑多种数据类型的分析展示。除了针对业务展示优化过的常规图表外，还能够绘制包括海量数据的地理轨迹、地理飞线、热力分布、地域区块、3D 地图、3D 地球，实现地理数据的多层叠加。此外还有拓扑关系、树图等异形图表供您自由搭配。
- 数据源：多种数据源接入，充分发挥阿里云大数据计算的能力。能够接入常见的分析型数据库和关系型数据库，支持本地CSV上传及动态请求。满足您各类大数据实时计算、监控的需求，充分发挥大数据计算的能力。
- Web服务：提供多种业务模块级别而非图表组件的工具，所见即所得的配置方式。图形化的搭建工具，无需专业编程人员也可快速实现。多分辨率适配与灵活的发布方式，满足您不同场合下的使用。创建的可视化应用能够发布分享，所有用户都可以访问到应用，作为您对外数据业务展示的窗口。

7.3.2. 系统架构

DataV数据可视化平台包括数据API、Web客户端、DataV服务端等多部分组成，采用B/S结构设计，核心内容是数据展现部分，支持绘制各类基础图表，接入EChart等第三方图表库，支持专业级地理信息可视化，配套丰富的行业模版，最终实现数据可视化平台的灵活部署与适配，并呈现丰富的数据展现形式。

数据可视化平台的架构如下。



7.4. 功能特性

创建可视化应用

创建可视化应用时，可以选择直接使用模板创建，也可以使用空白画布从零开始设计。

编辑可视化应用

编辑可视化应用，可以让您快速修改可视化应用的内容和配置。

复制可视化应用

支持复制可视化应用，复制后，会生成一个名为“xxx_副本”的应用（复制功能只会将应用复制给自己）。复制的应用继承原应用的配置和数据，可稍作修改，生成一个符合要求的大屏。

删除可视化应用

支持整体删除可视化应用，删除后应用不可恢复。

拷贝可视化应用

拷贝功能是指将“xxx”应用，通过用户识别码拷贝给其他用户，实现大屏协作与共享（拷贝功能是将应用拷贝给其他用户，区别于复制可视化应用）。

重命名可视化应用

支持修改可视化应用的名称，名称可以重复。

预览可视化应用

应用创建完成后，使用预览功能可预览应用，展示动态效果。

发布可视化应用

可视化应用发布功能提供了四种方式：公开分享、密码访问、TOKEN验证免登和IP白名单访问分享。公开分享的应用通过URL链接即可进行公开访问；密码访问功能，通过URL访问时需要输入密码才能进行；TOKEN验证免登功能，通常用于权限体系集成时用此方式实现。设置启用IP白名单后，发布后的可视化应用仅允许在设置范围内的白名单IP下访问。

- 公开分享：通过URL链接，即可访问大屏应用。
- 密码访问：可设置大屏访问密码，通过URL访问大屏时，需要输入密码。
- TOKEN验证免登：需要进行TOKEN验证，通常用来实现权限体系集成。
- IP白名单访问分享：设置启用IP白名单后，发布后的可视化应用仅允许在设置范围内的白名单IP下访问。

设计可视化应用

可视化模板

数据可视化的设计难点在于，如何能在简单的一页之内读懂数据之间的层次与关联，这就关系到色彩、布局、图表的综合运用。DataV提供多种精美的模板，对应指挥中心、地理分析、实时监控、汇报展示等多种场景常见设计，直接套用使可视化作品显现出高设计水准。



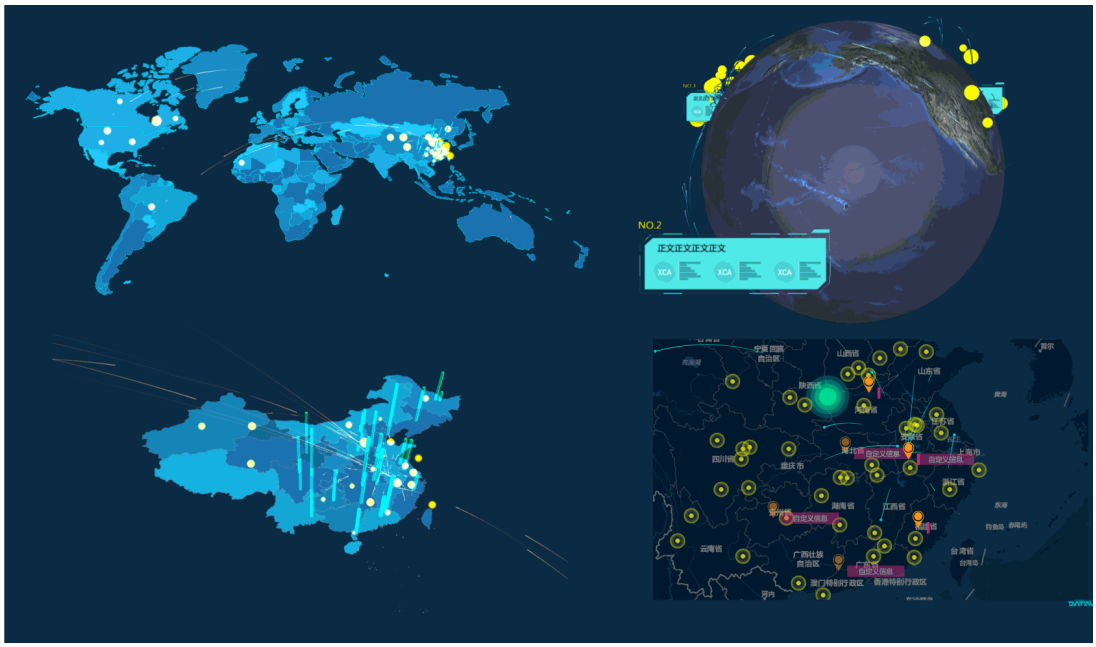
组件种类

DataV数据可视化平台在常规组件上再进行了设计，优化过的条、饼、柱、线等传统图在视觉呈现上更加新颖、多样，相比一般的图表更适合在大屏幕上呈现。组件包括常规图表、地图、指标、关系网络、文字以及其他更多第三方组件，还能绘制出包括海量数据的地理轨迹、地理飞线、热力分布、地域区块、3D 地图、3D 地球，以及地理数据的多层叠加、拓扑关系、树图等异形图表组件，此外还提供ECharts和AntV-G2等第三方开源图表库。

常规图表种类



地图图表种类



组件设置

任何一个组件都可设置其配置、数据和交互。配置设置因组件不同有差异化，可设置全局样式、图表位置等外观上的显示；数据默认使用静态数据，可设置数据源和具体数值；交互数值即参数变量，可以用于控制组件之间参数的传递，从而达到组件之间数据交互的目的。

> 说明
>
> - 交互事件中value绑定到变量，可以用于控制组件之间参数的传递，从而达到组件之间数据交互的目的。
> - 部分组件可以作为主动回调方，任意组件都可以作为被回调方，被能够绑定交互事件的组件回调。可以作为主动回调的组件有：Tab组件、时间轴、轮播列表、数字翻牌器、搜索框、地图类组件的区域热力层、行政下钻热力层、线热力层、散点层、图标点层、多系列散点层。

组件异常提示

在缺少数据、数据连接失败、数据获取错误等情况下会发生数据异常，此时在数据设置页面会出现提示，该组件不能正常工作。

连接数据源

数据可视化平台目前支持接入的数据源有：数据库类、文件类、API类和数据代理等。

- 数据库类

  数据可视化平台支持数据库类数据源包括RDS for MySQL、RDS for PostgreSQL、RDS for SQLServer、TableStore、兼容MySQL数据库等。

- 文件类

  数据可视化平台支持的文件类数据源有两种：CSV文件和静态JSON。

- API类

  数据可视化平台支持API类数据源包括POP API、阿里云API网关和自定义HTTP API。

- 数据代理

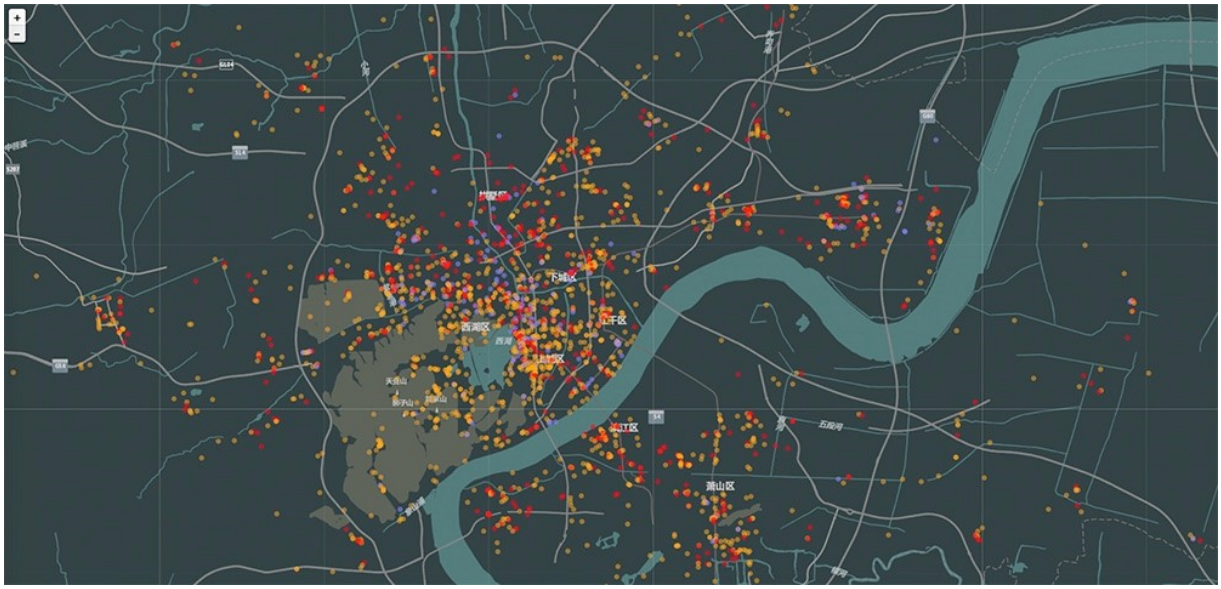
  数据代理服务提供开源数据库代理服务，支持在ECS上进行部署。通过数据代理服务，可降低数据库暴露IP带来的风险。

7.5. 应用场景

DataV数据可视化平台是大屏应用的开发工具和运行环境，平台提供多种布局、组件，能够在模板的基础上通过配置的方式实现可视化前端应用。并能够和各类常见数据库连接，实现多种数据可视化展示场景的需求。

智慧城市-美食大屏

DataV数据可视化平台创建美食大屏，通过美食大屏可以直接观察到菜系的分布。



单击地图某个点还能显示该餐厅细节。



汇报展示-业绩大屏

经营状况实时数据可以通过大屏直接动态展示，可实时监控当日交易支付金额，成交会员分析、热销商品排行等品牌全域核心数据。



8.交通云控平台产品

交通云控平台是阿里云面向交通行业打造的，可提供地图、数据、智能算法、 云边协同、控制下发、视觉渲染六大能力的大数据开放平台。

产品介绍

交通云控平台是基于“数据驱动+人工智能”为核心的云计算技术，充分利用交通行业海量数据，构建的具备多源数据融合能力、即时大数据计算能力、多种人工智能算法协同计算分析能力的，可提供准确数据、智能算法方案、信息发布通道等业务支撑的数据智能中台产品。

交通云控平台为交通行业应用的开发提供图、数、智、控、云、视等能力，交通行业开发者可基于平台能力，结合不同场景的具体业务需求构建相应的智能应用。如对城市交通、高速运营、大型枢纽调度等业务场景进行智慧化改造，用创新的、智能的应用和优化方案服务于交通管理者和交通出行者。



开放能力

交通云控平台的开放性主要体现在以下四个层面：

1. **大数据计算能力开放**

   交通云控平台提供一整套支撑开发的数据计算能力与计算架构，提供安全、高可用、弹性的高性能计算平台。

   - Blink
   - MaxCompute
   - DataHub
   - Apsara RDS
   - AnalyticDB
   - ECS
   - ...

2. **多源融合数据开放**

   交通云控平台针对交通行业各领域（交管、高速、运输等）提供行业数据融合模型，用户只需根据数据清单接入数据，即可生成标准的行业数据模型及指标的质量评价，降低开发者使用数据时的沟通和理解成本。

3. **算法能力开放**

   交通云控平台在数百亿节点数万亿边的超大规模网络上，处理EB级别数据，通过模糊认知反演算法，发现复杂场景背后的超时、超距弱关联。同时，平台将算法以引擎（如智能优化、仿真推演、深度分析等引擎）的方式对外提供。

4. **业务市场开放**

   交通云控平台与开发者共享客户资源。开发者基于交通云控平台开发的软件，可以进驻相关领域的行业云市场向行业客户进行销售。

   - 效果推广：共享客户资源，实现合作共赢。
   - 品牌传播：进驻相关领域的云市场（如交警行业市场），向行业客户进行销售。

产品架构

交通云控平台的产品框架如下图所示，将各类交通数据，例如路侧设备数据、交警数据、视频结构化数据，按照行业数据模型进行数据清洗，并融合多元数据进行计算，形成统一路网、设施设备规整数据，并在此基础上形成不同模型的融合交通参数，服务于不同的业务应用。

核心优势

交通云控平台旨在帮助交通从业者打造新一代智能交通系统，助力交通治理由“感而不知、感而略知”进化到“感而全知”，推动交通业务由“信息化优势”转化为“决策优势”，进而达到“行动优势”。

- 全息数字化还原：互联网数据和设备数据融合，交通评价更精准
- 大规模实时仿真：城市级仿真推演时间从几小时缩短到几分钟
- 智能优化调度方案：6年信号优化算法沉淀、5000+高速控制策略
- 全域多端交互协同：云边端一体化协同，海量边缘节点云端轻松运维

8.1. 功能特性

交通云控平台为交通行业应用的开发提供地图、数据、智能、控制、视觉和计算六大能力，旨在帮助交通从业者打造新一代智能交通系统，助力交通治理由“感而不知、感而略知”进化为“感而全知”，推动交通业务由“信息化优势”转化为“决策优势”，进而达到“行动优势”。

地图

为交通应用提供地图展示、查询、修改的能力。

- 数字化路网：将物理路网还原为统一编码、可计算的数字化路网。可进行路网查询、修改道路实体信息，以及将相关设备挂载于路网上，使设备数据准确的参与融合计算。
- 地图引擎：提供图层渲染、空间搜索等地图引擎专业能力。同时数据与数字化路网同步联动，满足用户定制需求。

数据

多源数据融合更精准、完善的开发工具更高效。

- 评价指标：基于互联网数据和设备数据秒级融合，还原城市的全量出行轨迹，从而计算出更精准的交通运行评价指标。

- 数据开发工具：一套覆盖数据采集、加工、治理全流程的数据开发工具，高效管控数据的产出和质量。

智能

算法结合业务知识，沉淀多领域的交通AI算法、决策方案。

- 视觉分析服务：提供对人、车辆的特征检索、交通参数识别、交通事件识别、时空轨迹还原等视觉分析算法服务。
- 交通智能优化：提供对路口、道路、运输工具在多场景下的调度控制策略，如路口的信号优化方案、高速的流控方案等。
- 实时仿真预演：提供多分辨率实时仿真服务，自带仿真数据，城市仿真推演分钟级别完成。

控制

提供互联网C端、交通设备端控制渠道。

- 互联网出行服务：提供向互联网应用（如导航APP）推送出行消息的服务。
- 路侧统一控制：提供向信号机、诱导屏、RSU、匝道、雾灯等设备下发方案的能力。
- 智能车联网：提供将路侧设备实时计算结果推送给斑马汽车、千寻车载屏等车载智能设备的能力。

视觉

3D视觉渲染服务、交通视觉组件，助力应用研发更快速。

- 实时渲染引擎：支持2D/3D实时渲染、支持GIS地理特性、产品简单易用，非专业工程师也可轻松搭建专业水准的可视化应用。
- 交通视觉组件：提供二三维渠化图、时序图、蔓延分析图、流量溯源图等多种交通视觉组件，帮助前端开发者快速构建应用。

计算

提供云端、边端的计算能力，云边协同一体化管理。

- 云端计算：云控平台支持公共云、专有云、混合云、一体机四种形态的部署方式，同时支持开源底座，兼顾不同类型客户对算力、安全、价格的要求。
- 边缘计算：云控平台的边缘计算产品Neuro，提供视频分析和信号优化方案产出和下发的能力，同时云端自动纳管，运维无忧。

8.2. 应用场景

为您说明交通云控平台在四大交通领域的主要应用场景和案例。

城市交通管理

核心场景

- 路网数字还原：标精/高精矢量地图路网还原，提供路档查询和编辑工具，交通运行事件融合精细认知。
- 车路协同控制：提供全局协同信号优化、可变车道控制等路侧控制方案，以及车内红绿灯、行人/事件预警和车速诱导等车载控制方案。
- 出行诱导服务：通过诱导屏/情报板等路侧设备发布交通事件、图/文诱导路径和建议车速。同时，车内互联网导航系统可上报交通事件、拥堵路段信息。
- 重点车辆管理：可以结合信用评估，驾驶人安全码管理两客一危重点车辆，路径规划及违法实时预警。

案例说明

- 场景：某城市某商区附近道路常发失衡和拥堵，管理者希望通过信号优化减缓拥堵。

- 外场布设：激光雷达（2个）、雷视一体（24路）、车速诱导屏（2块）
- 解决方案：云（交通云控平台）边（Neuro）协同实现区域信号优化。其中，Neuro部署在目标商圈附近的信号机箱，搭载视频解析和信号优化算法。
- 案例效果：
  - 全时段混合交通流下区域自适应秒级优化推荐最佳行驶车速，交通延误缩短20%。
  - 高峰时段单路口每小时可多通行150辆车。

高速运营管理

核心场景

- 路网数字还原：百米精度路网还原，动态信息还原，交通气象预警，三急一速隐患点分析。
- 车路协同控制：10+车路协同场景库，路侧控制和车载控制联动，提升通行效率，降低事故率。
- 应急事件推演：评估事件影响，推荐交通管控策略，通过出行平台发布，并对策略效果仿真评估。
- 精准在线计费：实时还原车辆路径，收费稽核。

案例说明

- 场景：某省希望建立智慧高速综合管理云平台，
- 解决方案：基于公共云部署交通云控平台，基于平台能力进行开发，实现600+公里路网数字还原，1400路视频公共云实时分析。
- 案例效果：
  - 实现全路网交通运行状态实时监测，视频事件检测准确率>98%。
  - 实现两客一危车辆在途状态的实时监测，提升了安全监管水平。
  - 智能产品开发周期缩短50%，实现智慧高速产品线从无到有，一年时间拓展3个省份。

公共交通运营

核心场景

- 全量出行还原：职住通勤客流还原、出行方式分析、景区客流分析预测。
- 公交运行评价：线网评价（线网密度/站点覆盖率/可达性）、线路评价（满载率/运行速度/可靠性/首发站准点率）、站点评价（到站时间预测/到站准点率）。
- 接驳优化推荐：动态公交调度、公交线网优化、公交排班优化、地铁运营与检修排班。
- MaaS出行服务：个性化联程方式出行规划、一码通乘服务、出行核验、用户运营管理。

案例说明

- 场景：某城市希望提升公共交通在城市级大型赛会活动时的服务水平和接驳疏散能力，保障城市交通服务正常运行。
- 解决方案：基于交通云控平台提供交通需求预测、运力供给匹配、协同运力调度三大方案。
- 案例效果：
  - 满足赛会期间300万左右观众的交通信息和出行需求。
  - 支撑2400余辆赛会专用车辆的实时调度需求。

海陆空交通枢纽

核心场景

- 场面数字还原：飞机动线轨迹还原、保障节点自动识别、车辆轨迹还原、路段巡检监测、站台作业识别、船号船型识别、船舶轨迹还原、航道密度识别。
- 安全事件感知：红线区域清空与锥桶摆放识别、飞机泊位识别预警、机场围界入侵检测、人员着装识别、车辆越界与停止线识别、施工辅助盯防、站台安全监控、海上动态感知。
- 运行时刻推测：飞机滑行路径规划、可变滑行时间预测、保障节点时间预测、施工持续时间预测、站台安全风险预测、船舶航行/集卡路径规划、桥吊作业时间预测、龙门吊作业事件预测。
- 运力资源调度：机位智能分配、人员智能排班、车辆智能调度、船舶智能避碰、泊位智能分配、集卡车辆智能调度、龙门吊桥吊智能调度。

案例说明

- 场景：某机场飞机滑行轨迹难以实现全链路追踪和还原；保障节点靠人工记录漏报误报情况普遍；安全事件人工识别发现率低且不及时
- 解决方案：基于枢纽机场云控平台能力开发一体机，实现数字化坪一体机方案快速落地部署。
- 案例效果：
  - 保障节点智能识别16+个，准确率95%+。
  - 全天候飞机动线全轨迹还原，准确率95%+。
  - 安全事件识别适配场景10+个，识别准确率96%+。

9.Hdata产品简介

9.1. 什么是医疗健康数智平台

阿里云医疗健康数智平台是医疗健康行业大数据治理和应用服务平台，内置丰富医疗知识图谱，通过医疗文本结构化解析、医疗文书实体识别、智能编码等算法，疾病风险分层等模型支持用户挖掘数据价值。

阿里云医疗健康数智平台，是医疗健康行业大数据治理和应用服务平台，为业内数据管理者，包括卫健委、医保局、医院、医药企业等，提供符合医疗行业数据特性的数据治理和服务全套工具，具备简单易用、生态开放、算法智能、数据安全等特性，帮客户真正的管好数据、用好数据、发挥好数据智能价值

目前医疗健康数智平台提供以下功能：

- 研发工作台
  - **数据同步**：数据同步可实现离线、实时多源异构数据的便捷同步或接入，系统可提供完善的数据接入配置、 数据模板配置、数据同步任务运行监控等功能，有效保障数据接入的稳定性和可控性，满足各类平台、数据源及应用系统间的数据汇聚需求。
  - **数据查询**：提供对云计算资源中物理表的探查，快速了解物理表详情及分布情况。支持面向多种云计算资源中的表及字段进行预览和表数据自动探查，提供表级别、字段级别探查结果图表化展示。
  - **数据标准**：提供逻辑表标准、字段标准（数据元、指标、维度、数据字典等）管理能力。帮助用户通过设计标准数据元素，定义关键业务对象、业务对象属性及值域定义，并规范标准数据字典，制定并管理平台遵循的统一数据标准，帮助平台管理者和数据管理者管控治理后数据的一致性和数据质量。
  - **数据建模**：帮助用户在数据标准的约束下构建数据模型，将数据标准贯彻到数据质量分析、保障及检查的全过程中，将散乱的多源异构数据加工成标准、干净的数据资产，确保数据的完整性、一致性、准确性、可用性，通过客观量化评估指标帮助客户了解数据治理工作进程，指引数据治理工作的螺旋式上升过程。为了完成在云上积累可运营数据资源的目标，提供数据充分融合、数据高质量可用的必要保障。
  - **数据开发**：支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理，同时提供数据加工、算法服务任务流程开发、编排与调试、上线、部署、维护等功能。支持流式计算、批量离线计算等类型的计算节点在一个工作流中统一编排，通过可视化操作界面，通过拖拉拽的方式连接计算节点迅速实现数据加工流程编辑。支持通过空间算子编辑器零代码实现对空间数据的处理，包括地址空间化（依赖达摩院地址标准化产品）、坐标转换、矢量分析和计算、栅格分析和计算等。空间计算节点可作为节点任务统一编排到工作流，实现空间数据和非空间数据的融合治理及治理链路固化。
  - **智能标签**：可将治理后的数据以业务化视角进行建模、查看、管理及使用，并提供业务衍生标签的自定义功能，为上层应用提供统一的标签数据目录和标签调用接口，沉淀上层应用制作的模型标签，实现高价值标签共享复用，形成标签运营生态。
  - **统一服务**：是一款数据中台建设过程中的数据服务化组件，面向数据开发者提供覆盖各个加工阶段统一体验的、便捷的数据查询转服务、服务管理、服务运维能力；面向数据资产管理者提供服务的统计分析、服务用量统计分析能力，实现数据中台建设后半场“数据应用”的有效落地，支撑数据智能应用的高效开发。
  - **运维监控**：对数据资源平台中的数据同步、数据开发、标签加工等任务进行运维管理。按业务链路配置监控场景，支持自定义监控链路和告警消息配置，可快速实现从数据生产到业务应用的全链路编排及溯源。
  - **质量评估**：对数据模型中建立的逻辑表进行质量规则配置和管理，支持根据配置的质量规则在数据开发环节自动生成质量检测节点和设置自定义质量检查计划。提供面向不同计算资源多种类型质量规则，可通过数据质量监控报告展现系统整体数据质量概览，和多维度细分数据的质量情况。
  - **解决方案**：对数据资源平台中已沉淀的数据标准、数据模型、数据加工场景、数据服务API配置、云计算资源配置等数据资产进行导入导出，便于在不同环境中快速搭建数据治理工作的初始化。
- 运营工作台

- **资产注册**：支持平台侧需管理的数据表资产自动化注册功能，可对多个数据源端注册信息进行展示，并通过手动刷新方式同步元数据；同时，可自动上报数据表、标签和API资产的信息。
- **资产目录**：支持按资产类型分别维护资产，帮助资产目录内容提供者对各类资产进行管理维护，并针对资产进行编目分类及发布上下线管控。

- 数据资产中心
  - **资产概览**：提供统一的数据资源视图，以可视化的方式展示数据表、标签、API和指标等数据资产的总量与增量的信息。
  - **资产目录**：对已接入上线的数据资产，可在资源目录对资产进行搜索、目录列表查看及资产详情信息展示。
- 业务工作台
  - **业务看板**：分层分域、规范统一的数据模型；多主题、多类别的资产服务化全局视图；多视角、多维度的一站式质控分析；医疗结构化处理算法唤醒海量医学沉睡数据
  - **数据治理**：标准数据依据按照“标准统一使用、接口统一制定、资源统一管理”的思路，内置卫生健康信息化标准体系，为各级各类医疗卫生健康机构以及相关部门之间实现互联互通、数据共享打下坚实基础。阿里云医疗健康数智平台内置医疗标准数据元、数据值域，为客户做标准化数据管理提供行业基础标准。

    提供数据标准的展示和查看功能。模型设计者可通过研发工作台资产加工模块设计标准数据元素，定义关键业务对象、业务对象属性及值域定义，并规范标准数据字典，制定并管理平台遵循的统一数据标准，帮助平台管理者和数据管理者管控治理后数据的一致性和数据质量。
  - **医疗数据资产**：支持对平台可共享的数据资产进行展示、搜索和权限申请，促进数据资产的共享和使用。
  - **智能业务能力**：医疗业务产品合集。
- **画像分析**：可基于标签的多维度分析，按照业务需求圈定特定群体范围，实现对复杂群体数据的智能化统计、筛选、加工、沉淀，建立单实体360度全方位个性化档案，通过关系图谱展示、可视化线索分析、相似人员智能化推荐等功能实现用户对于目标群体的精准定位。
- **数据探索**：数据探索面向行业客户/业务人员，提供工具内容一体化的业务模型构建平台，实现低代码、可视化构建全场景专家模型，重塑大数据服务创新模式。平台通过将业务数据沉淀为智能数据或智能算子，以可视化拖拉拽和简单图形化条件设定进行模型编排，支持离线、在线、实时全场景，采用简化建模过程、提高模型运行效能、融合智能化算法等方式帮助用户将数据与业务结合起来，不断积累和沉淀专家业务模型，服务于日常事件挖掘、实时预警事件、在线风险识别
- **系统设置**：数据资产平台的基础功能，包含账号管理、系统设置。该功能帮助用户快速、便捷完成用户角色及相应权限等基础配置，并实现必要的系统管理与控制，保障模块正常运转。

9.2. 产品优势

9.2.1. 核心能力

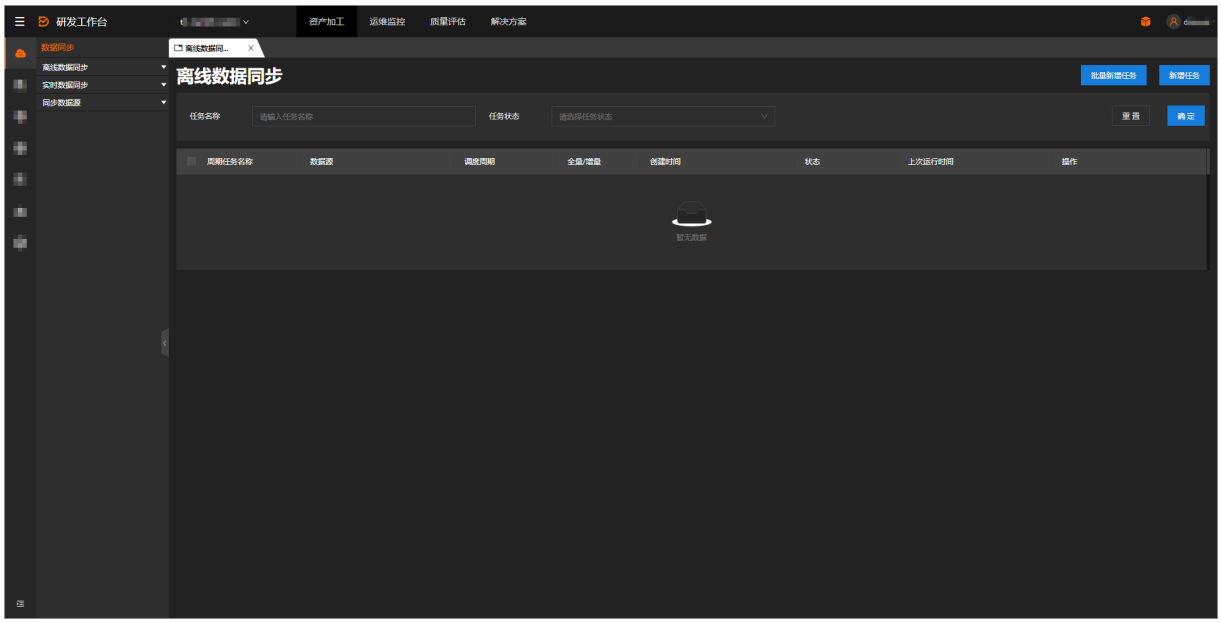
阿里云HData医疗健康数智平台，是医疗健康行业大数据治理和应用服务平台，为业内数据管理者，包括卫健委、医保局、医院、医药企业等，提供符合医疗行业数据特性的数据治理和服务全套工具，具备简单易用、生态开放、算法智能、数据安全等特性，帮客户真正的管好数据、用好数据、发挥好数据智能价值

一站式大数据开发平台，提供一致性开发体验

一站式数据同步

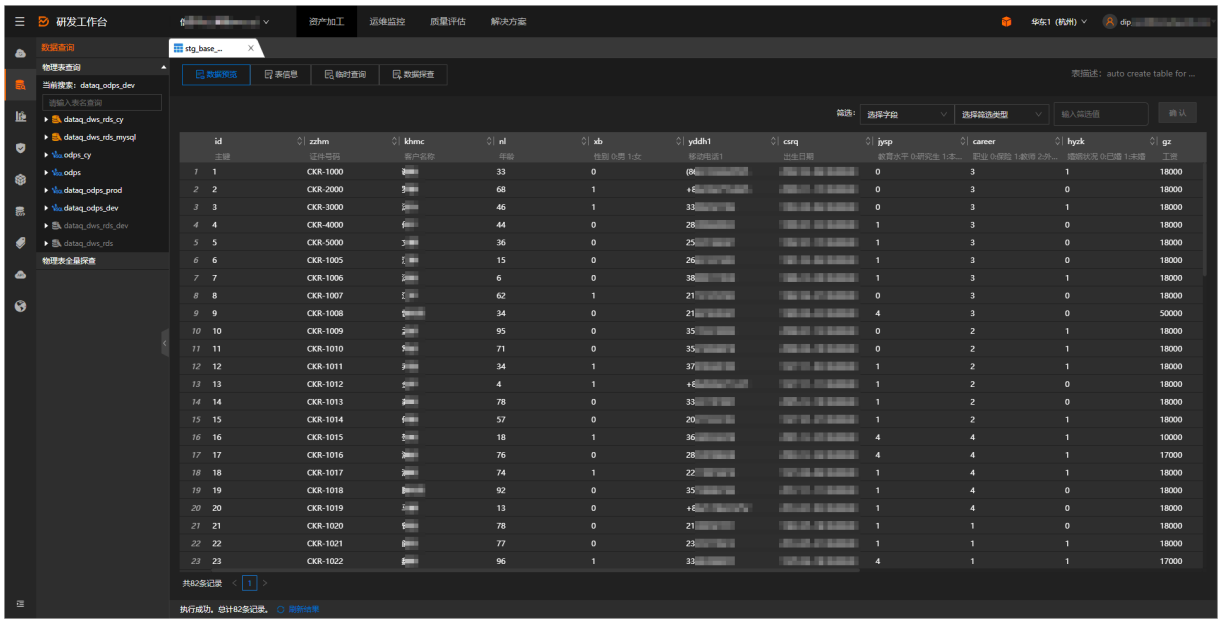
- 支持离线/实时数据传输。

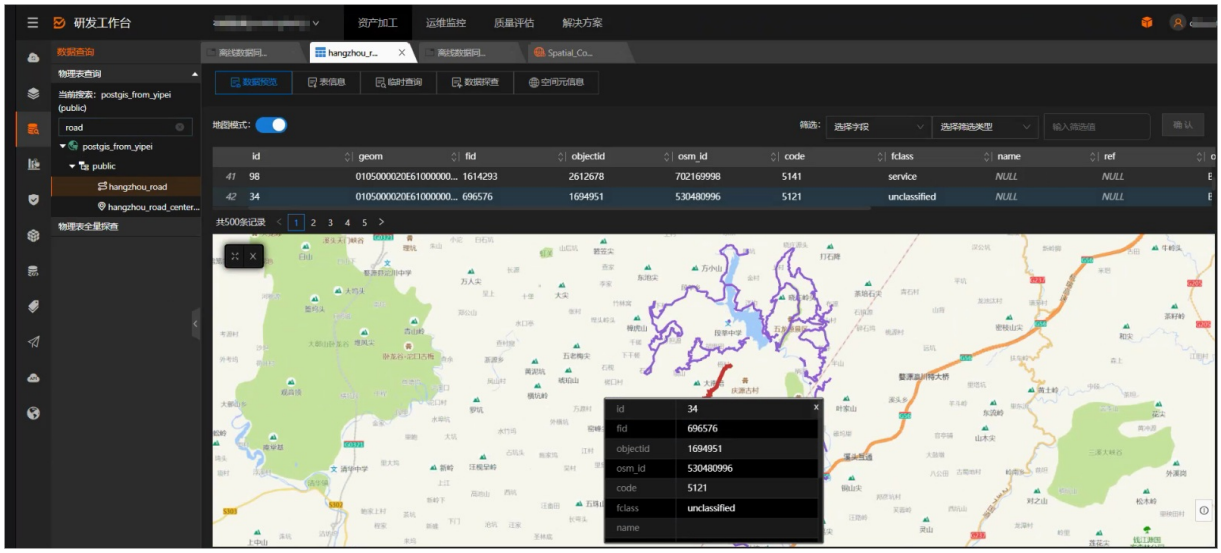
- 周期同步任务调度统一管理。
- 主流空间数据库全面支持同步汇聚（超图SDX+，ArcGIS SDE，PostGIS，Ganos）。



自动灵活地数据探查

- 物理表预览和表数据自动探查。
- 多种探查规则和算法，可快速实现表级别、字段级别探查结果图表化展示。
- 支持空间数据的地图可视化查看。

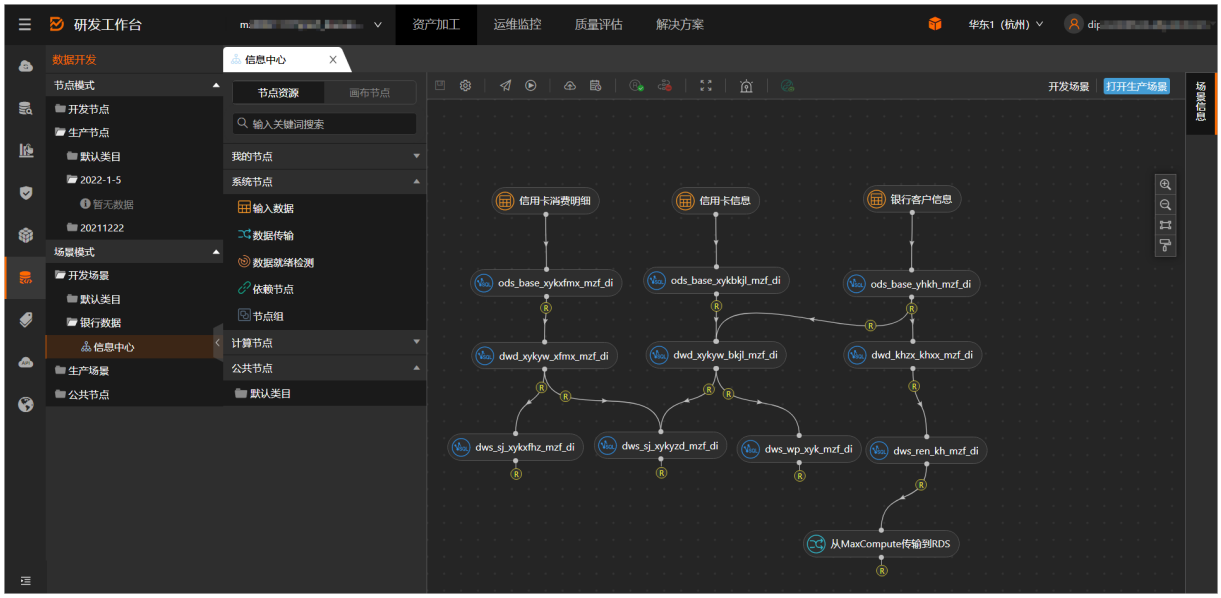




强大的数据加工能力

- 离线、实时、在线多类型计算节点。
- 不同类型计算节点在同一个任务流中混合编排调度。
- 支持空间算子编辑器，零代码针对空间数据进行计算、处理和分析。

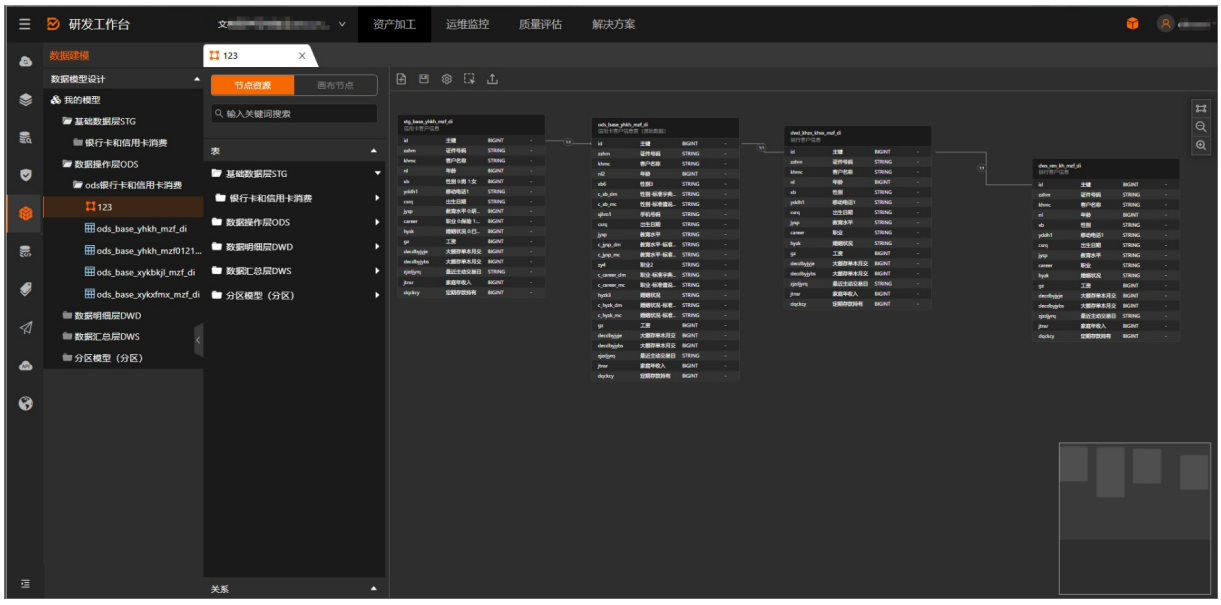


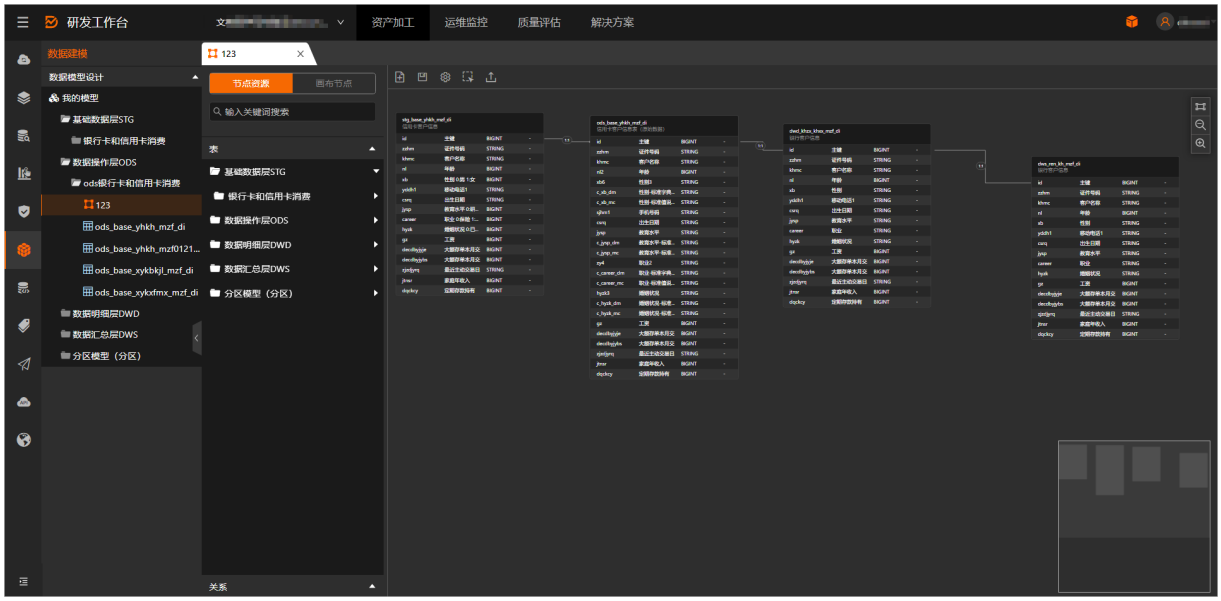


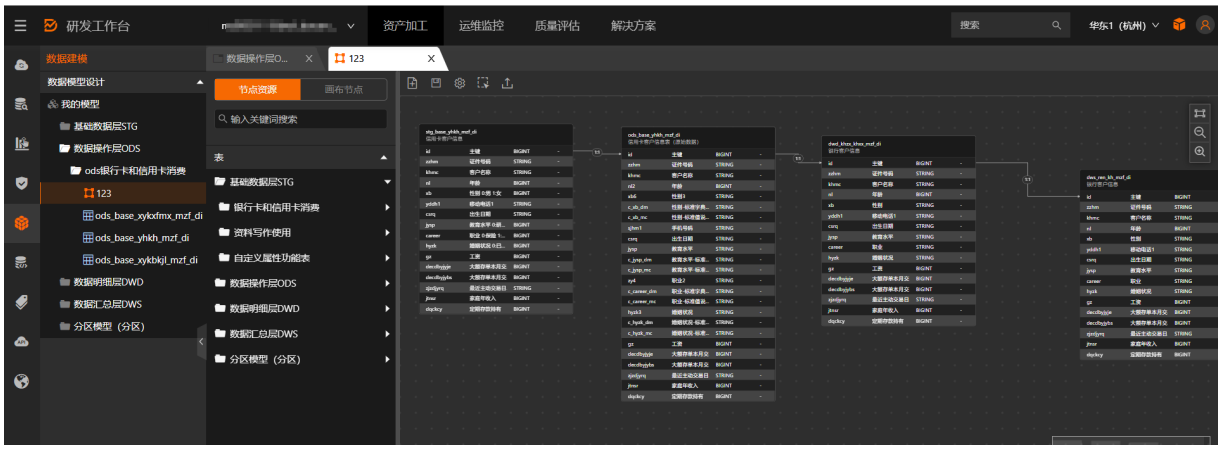
遵循数据标准的数据建模，实现数据质量闭环控制

可视化、符合标准的数据建模

- 可视化ER模型、维度模型设计，自动生成物理表DDL。
- 数据标准贯彻到模型、质量、保障及检查的全过程。
- 自定义表属性、自定义数据元属性等元模型设计。
- 内置6000+标准数据元、700+标准值域、200+标准数据集，满足医疗业务日常标准化治理场景搭建的标准基础







全面、准确的数据质量评估报告

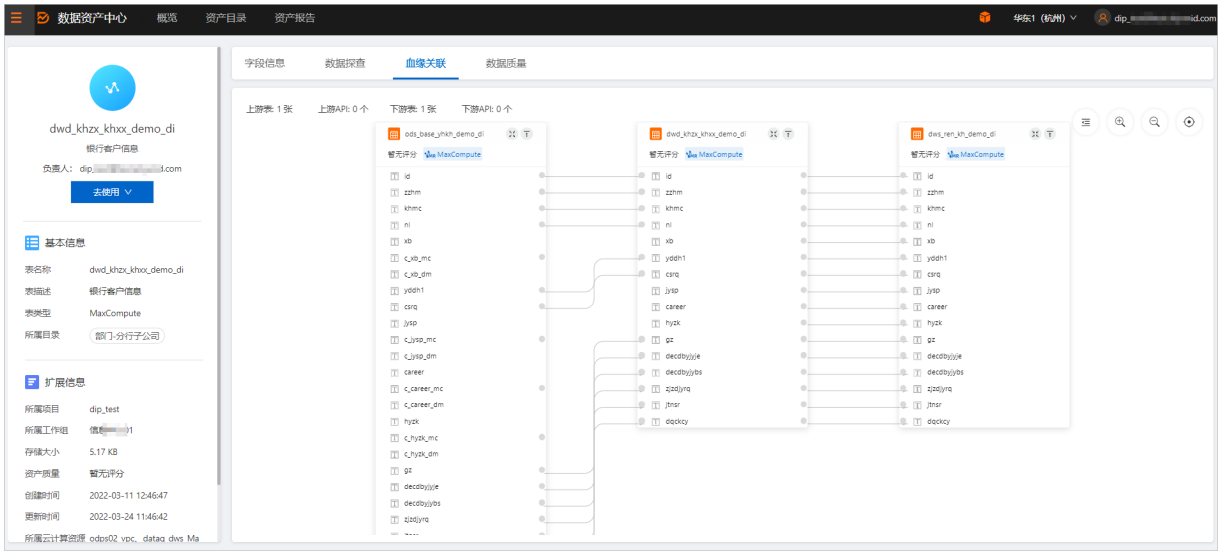
- 多种类型数据源自定义表及字段质量规则配置。
- 支持自定义SQL质量规则。
- 多维度数据质量评估，可视化质量评估报告。



数据资产的管理、服务与敏捷交付，使用数据资产更便捷

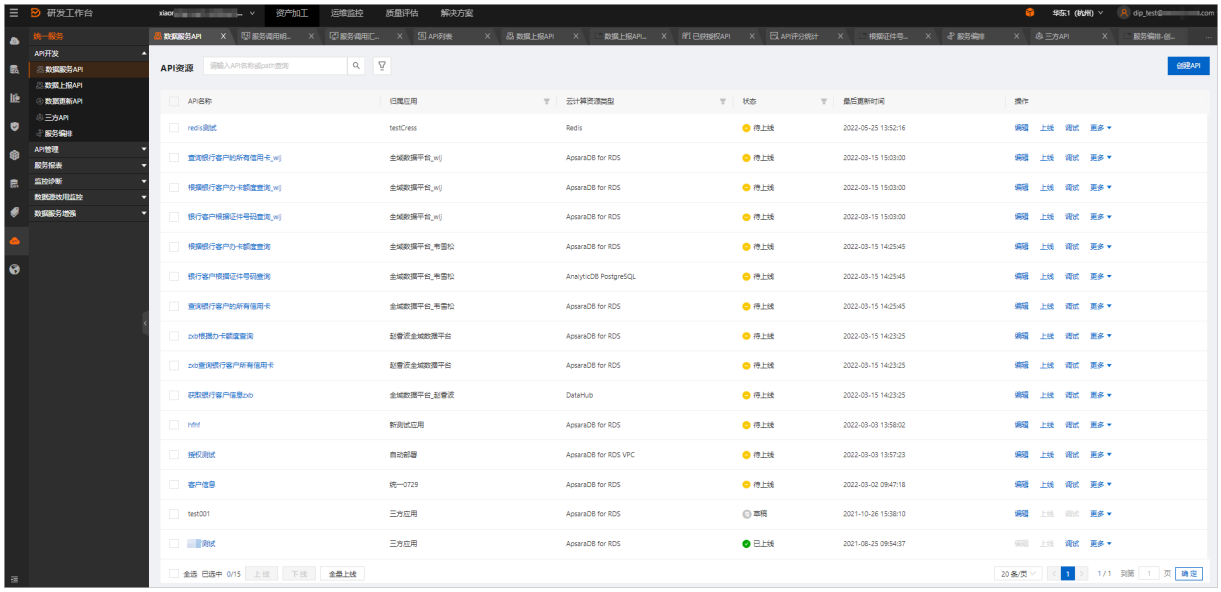
数据资产智能管理

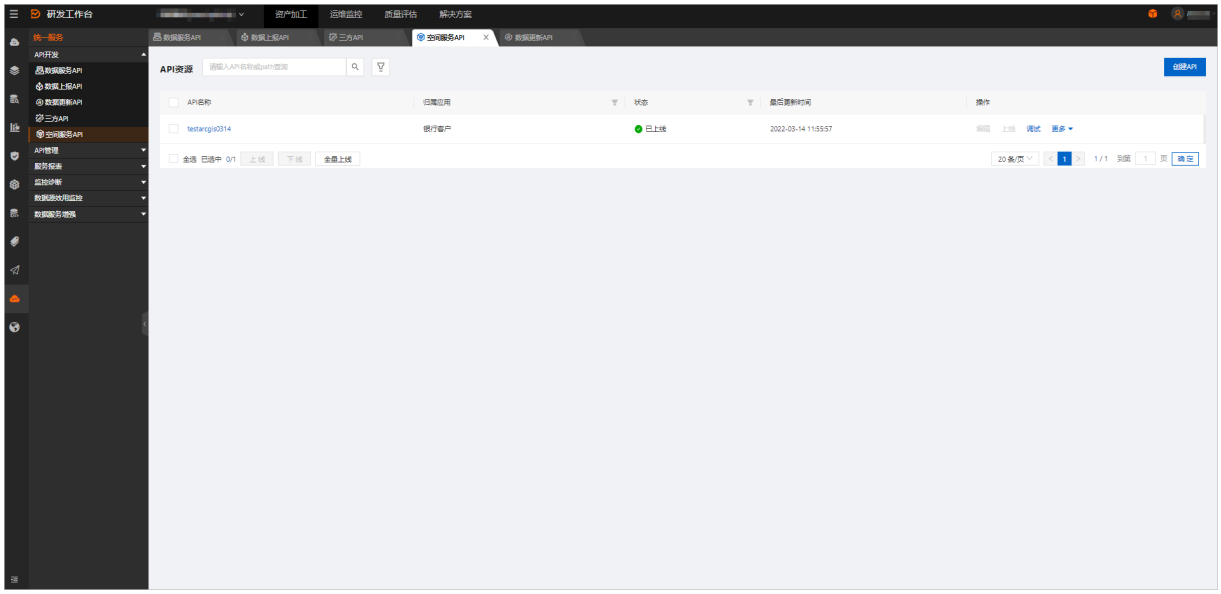
- 表资产、标签资产、API服务资产和指标资产的目录搜素和详情展现。
- 表资产、标签资产、API资产的使用申请和审批管理。

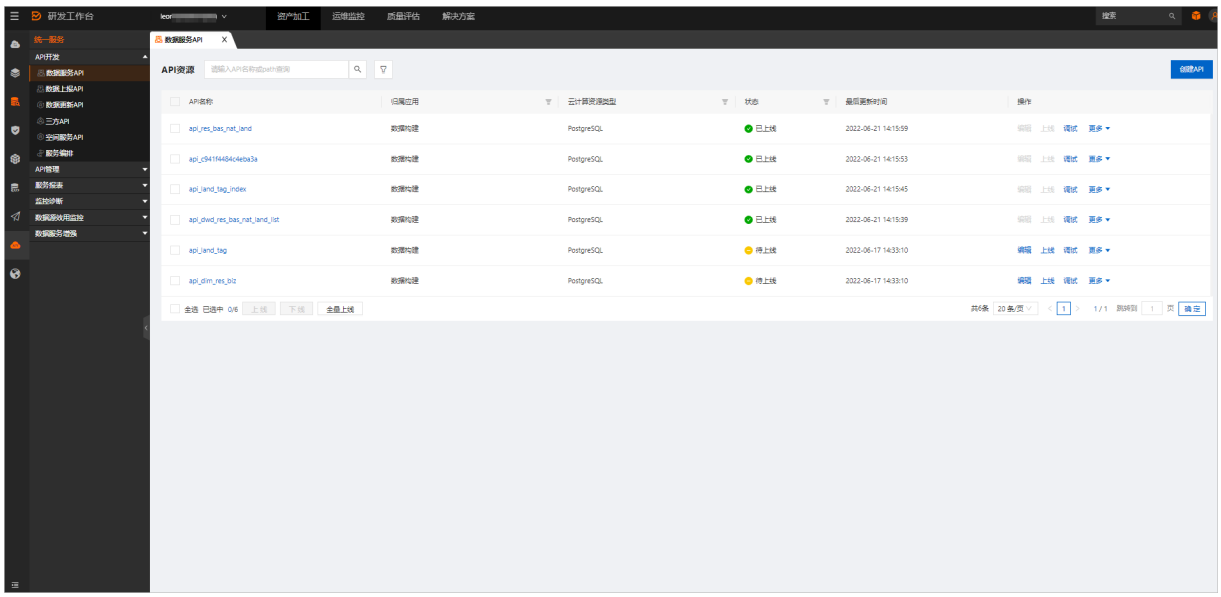


统一数据服务封装

- 实时、离线、半结构化、非结构化、消息队列的数据服务封装。
- 支持向导模式和SQL模式等的服务封装。
- 简单身份认证和加密签名身份认证模式。
- 数据服务调用流量管控。

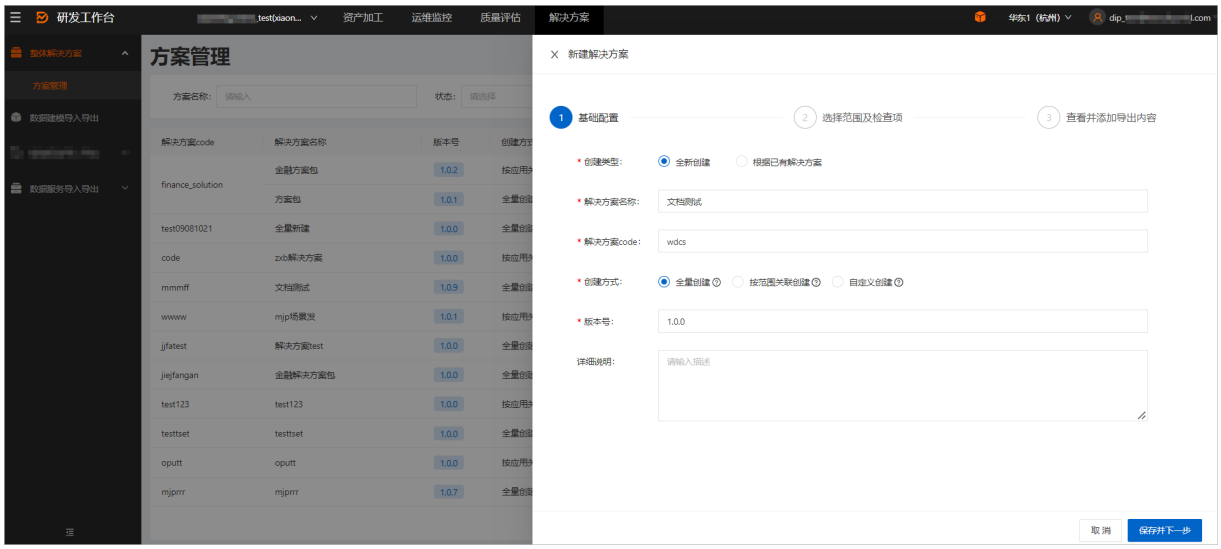






行业知识敏捷交付

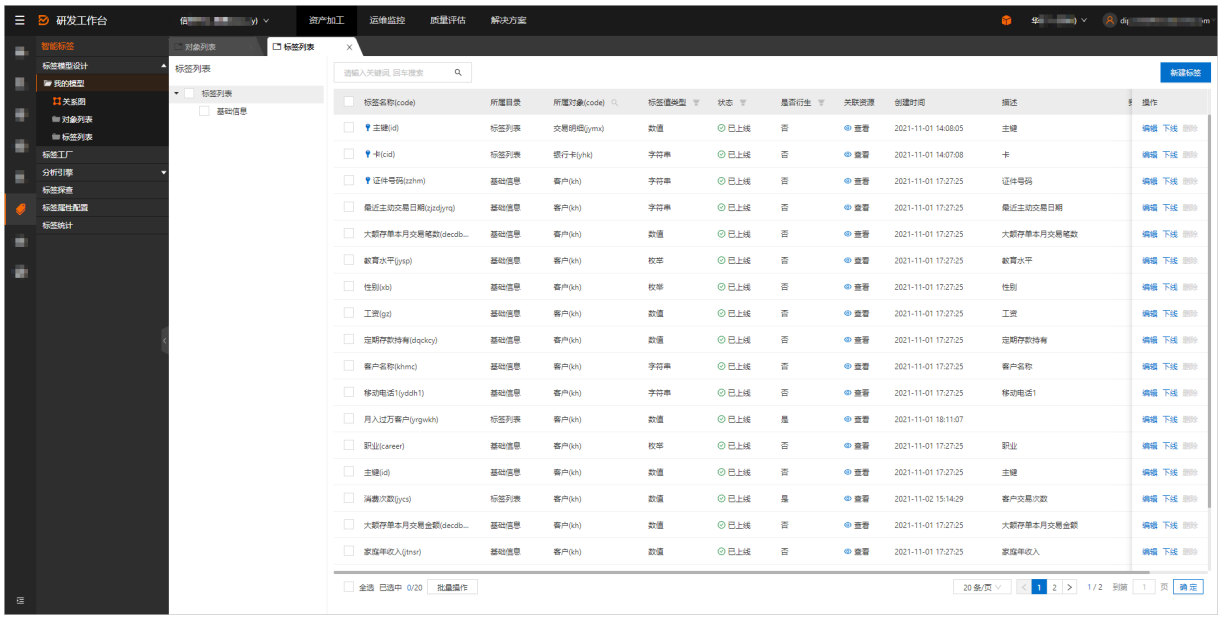
- 工作组下API服务、模式、节点、场景等全类型资产构建解决方案。
- 通过全局数据血缘关联资产灵活构建解决方案。



数据标签化，基于标签的可视化数据分析

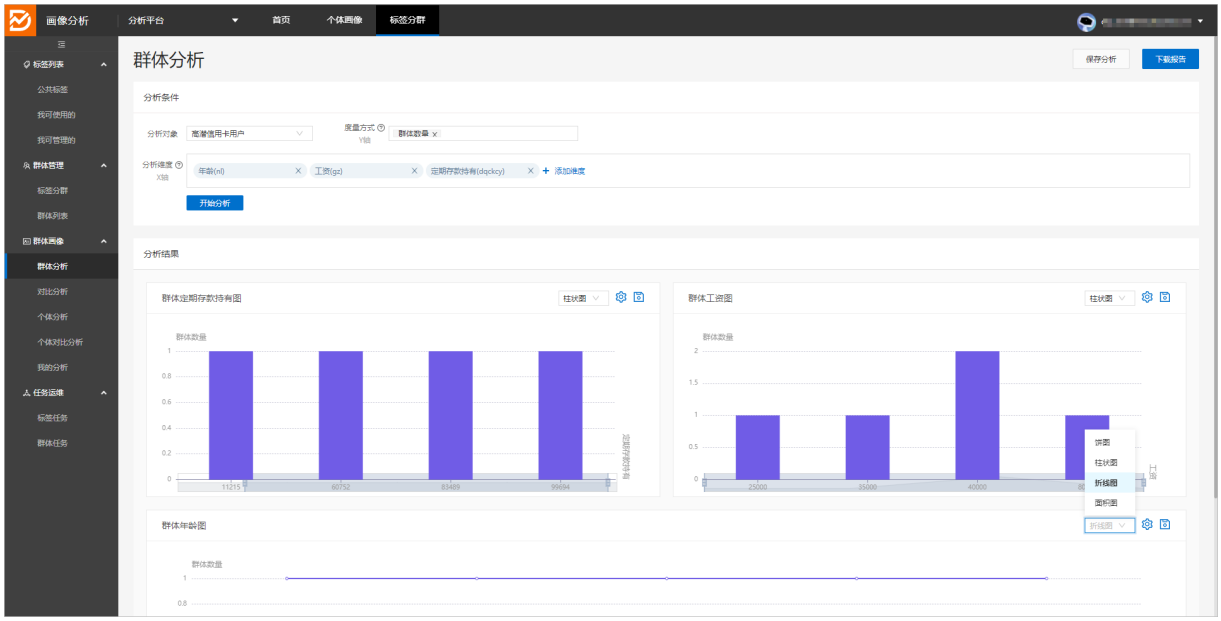
面向业务场景的标签体系构建

- 提供低门槛“宽表”构建能力，帮助上层应用构建业务模型，输出标签库和标签服务。
- 跨存储类型的数据源整合，降低数据同步运维成本。
- 零脚本积累标签画像资产，将数据开发与高频业务运营实现解耦，提高运营效率。



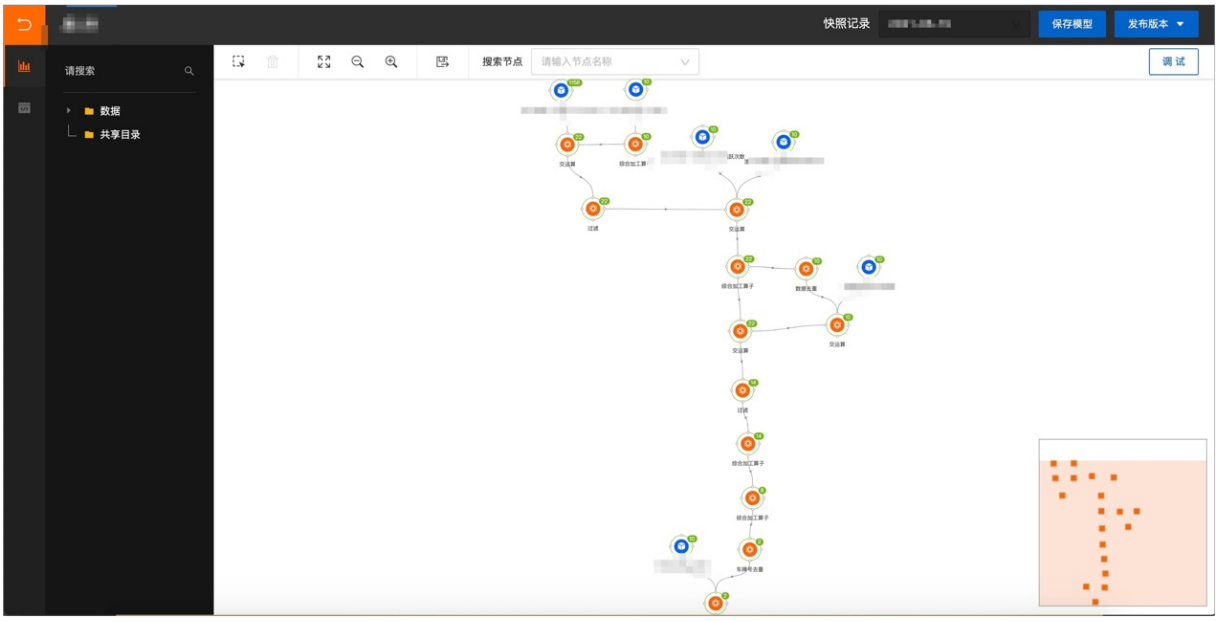
精准运营神器画像分析

- 客群画像分析，通过丰富标签，深入用户洞察，圈定目标群体，并实现精准画像的大数据分析。
- 灵活配置交互式界面：支持多维透视、群体计算等热门分析模型。
- 可沉淀的数据接口，快速集成标签数据、画像报告至各业务系统。



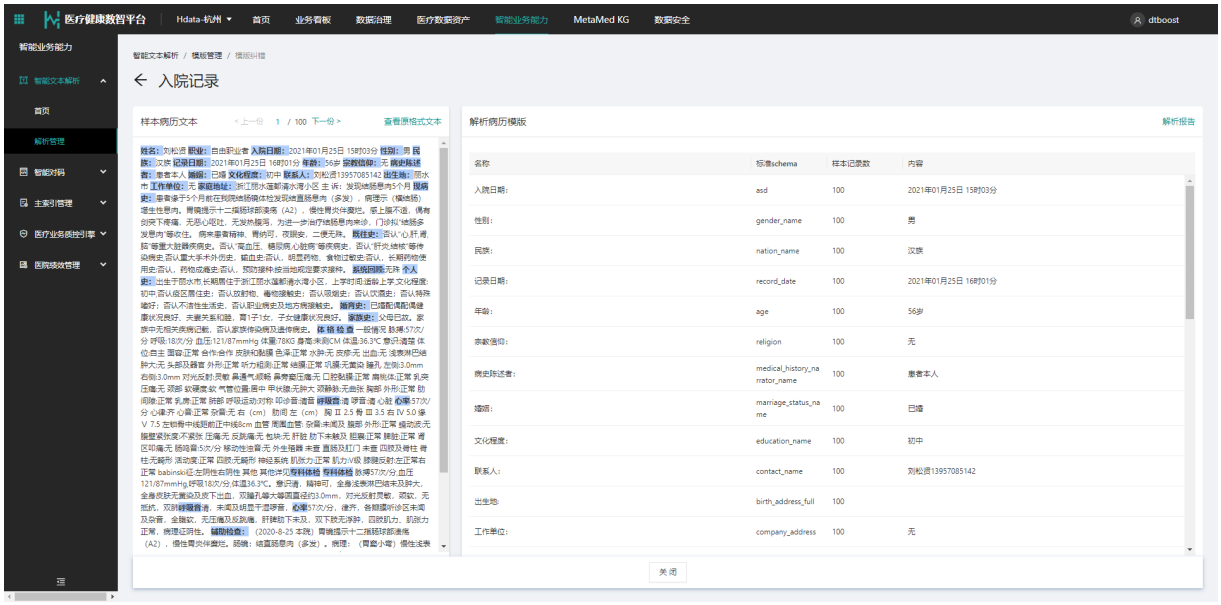
面向业务人员的低代码业务模型构建

- 实现在线、离线、实时模型全计算场景，提供模型编排、发布、服务实例的全生命周期管理能力。
- 实现跨行业通用层逻辑抽象，减少接入新行业/新场景的成本开销。
- 实现工具+内容的业务价值，沉淀行业算子/模型内容库，具备基础算子跨行业输出、业务模型跨行业交付的能力。



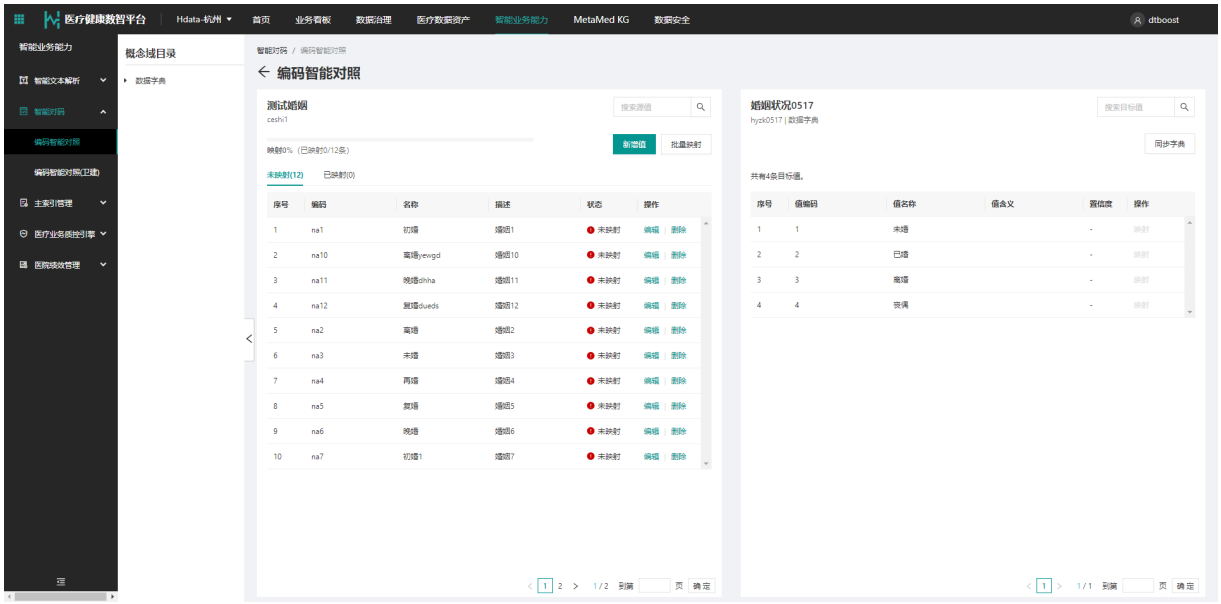
Hdata医疗数据治理智能算法

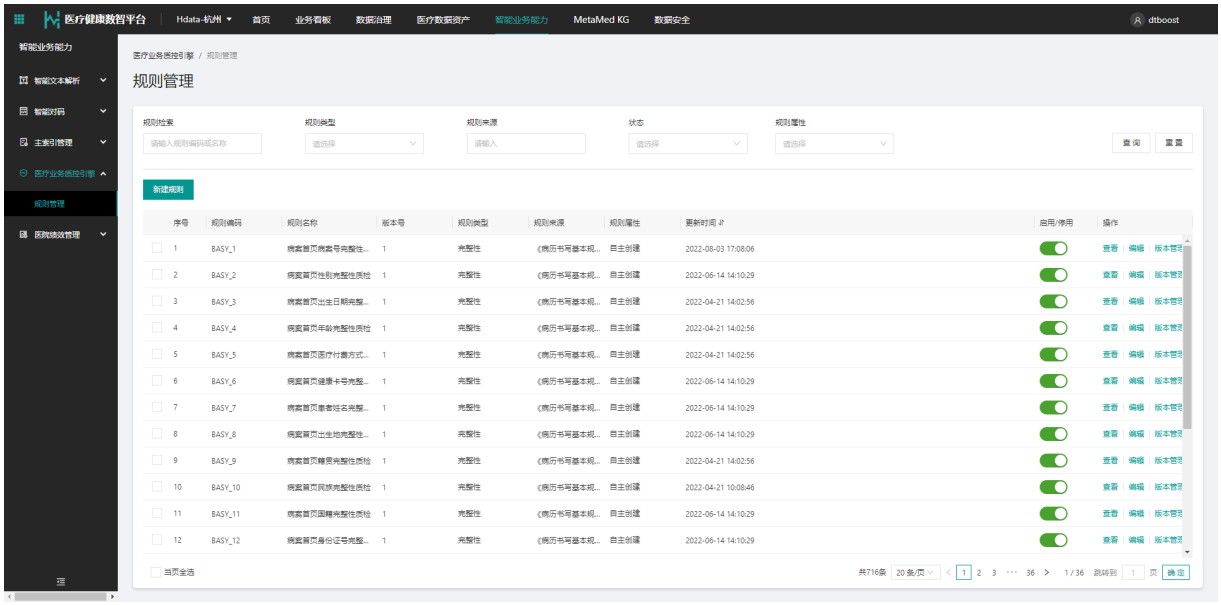
- 医疗文本结构化解析、医疗地址标准化解析、医疗文书实体识别、医疗术语智能归一等智能算法，加速医疗海量数据价值挖掘
- 丰富的医学知识图谱与规则库，包含100万+医学概念、300万+实体关系、3万+规则，支撑上层业务智能化能力



Hdata智能医疗业务引擎

- 提供医疗知识服务，支持合理用药、医学诊断图谱等知识查询和服务化使用，满足智能辅助诊断、医疗业务质控等知识使用场景的便捷使用需求
- 业务质控、绩效评价、患者数字化分层、合理用药、智能辅助诊断等丰富的业务智能引擎支持ISV能力集成





9.2.2. 用户价值

数据资源平台为用户提供了一站式数据资产定义、生产、管理与服务平台，提供企业级数据资产构建能力和一致性使用体验，助力客户快速构建数据智能平台，实现数据资源统一管理，挖掘潜在规律，优化业务决策，让大数据真正的驱动客户业务。 本文主要介绍数据资源平台的用户价值。

- 数据开发更简单：平台提供的各种工具产品能够极大的简化数据开发过程，缩短数据治理周期，降低数据治理成本。通过标准化、精细化、规格化的智能数据生产流程，完成流水线式的数据生产作业，提升数据资源生产效率、消除数据供应品质差异。
- 数据服务更便捷：能够赋予数据以业务价值，让各级用户能够直观的理解数据，并以此为基础向应用输出多样、便捷的数据服务，不断提升数据面向业务的价值。
- 数据应用更智能：提供面向业务人员的无代码业务模型构建能力和数据分析能力，大大降低数据获取和分析门槛，让业务人员可以直接使用数据，积累沉淀业务模型，能够向上层应用提供更加智能的数据。
- 数据资产更清晰：从宏观到微观助力数据管理方全面盘点数据资产，通过智能化全链资产血缘分析，理清战略数据资源，做到让管理者心中有数。
- 数据运营更高效：遵循应用先行、以用带存、由存而通、因通促用的理念，实现城市数据运营，驱动客户业务创新。

9.3. 功能特性

9.3.1. 业务工作台

9.3.1.1. 数据治理

数据治理提供数据标准的新建、审批、发布、展示和查看等功能，包括标准管理、数据标准和数据质量三个模块。

标准数据按照“标准统一使用、接口统一制定、资源统一管理”的思路，内置卫生健康信息化标准体系，为各级各类医疗卫生健康机构以及相关部门之间实现互联互通、数据共享打下坚实基础。阿里云医疗健康数智平台内置医疗标准数据元、数据值域，为客户做标准化数据管理提供行业基础标准。

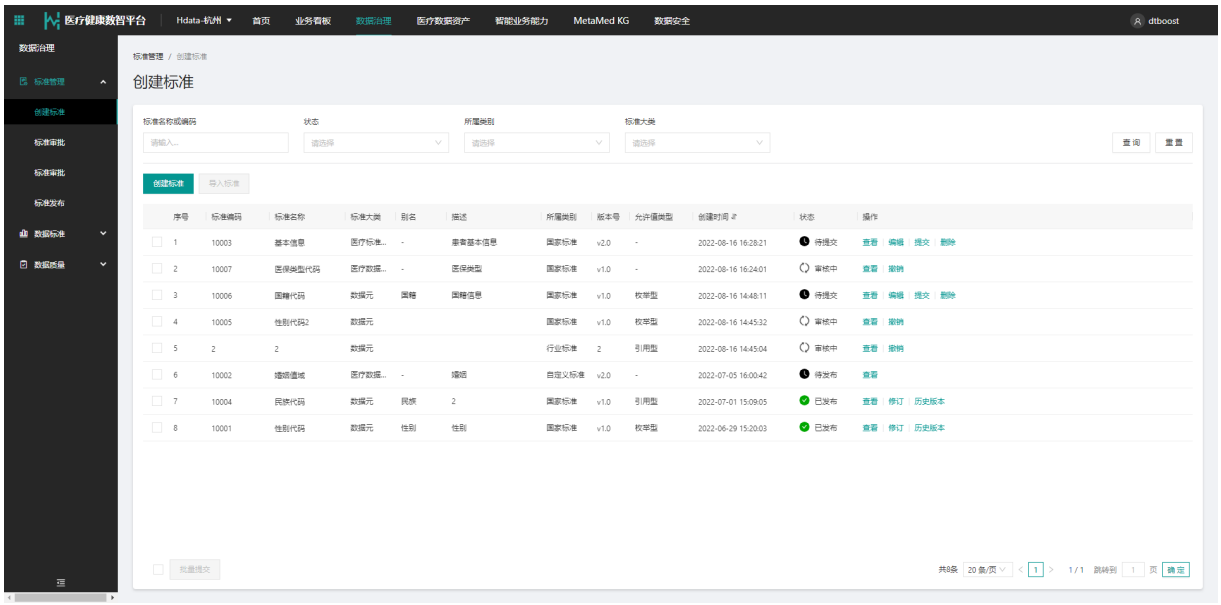
提供数据标准的展示和查看功能。模型设计者可通过研发工作台资产加工模块设计标准数据元素，定义关键业务对象、业务对象属性及值域定义，并规范标准数据字典，制定并管理平台遵循的统一数据标准，帮助平台管理者和数据管理者管控治理后数据的一致性和数据质量。

标准管理

医疗健康数据中台上所使用的标准（包括数据元、值域、数据集）的管理，不仅仅是管理部门参与，平台的机构也有参与标准共建的诉求。

包括标准管理模块，核心提供标准创建、更新、审核、发布模块。

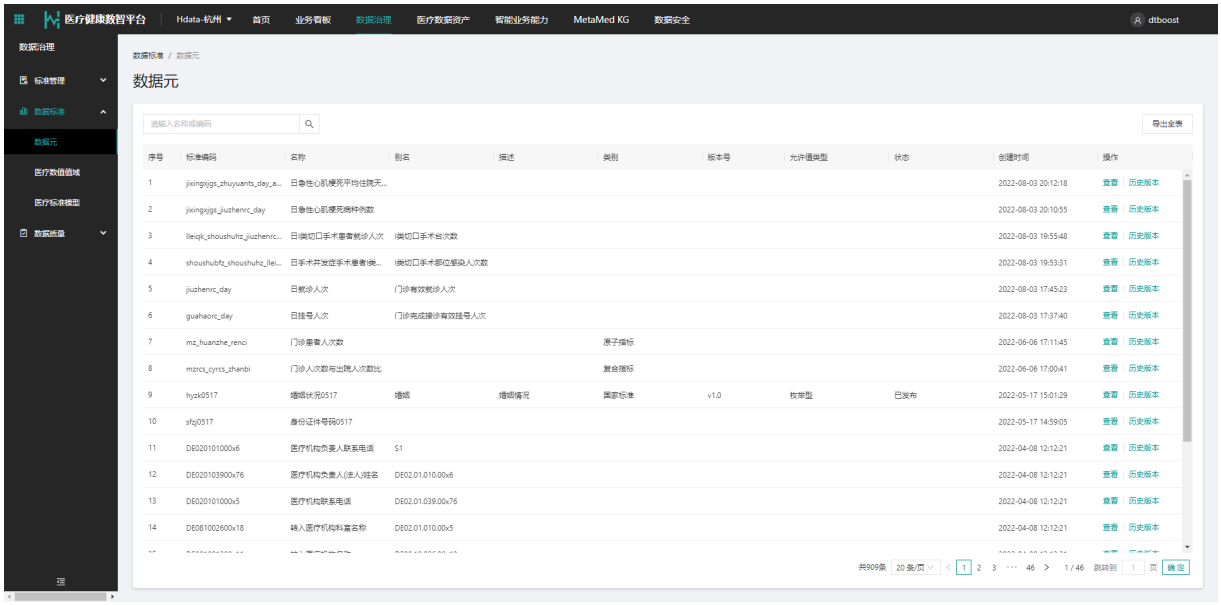
- 支持对标准的创建、编辑、修订、提交审核以及撤销审核，同时支持标准的检索和查询。
- 支持标准审批流的创建与维护。
- 支持标准审批留的分配、审核通过与驳回。
- 支持标准的发布。



数据标准

数据标准经过创建、审批、发布后，本功能将所有发布成功的标准进行沉淀，给用户提供标准查询、查看的功能。

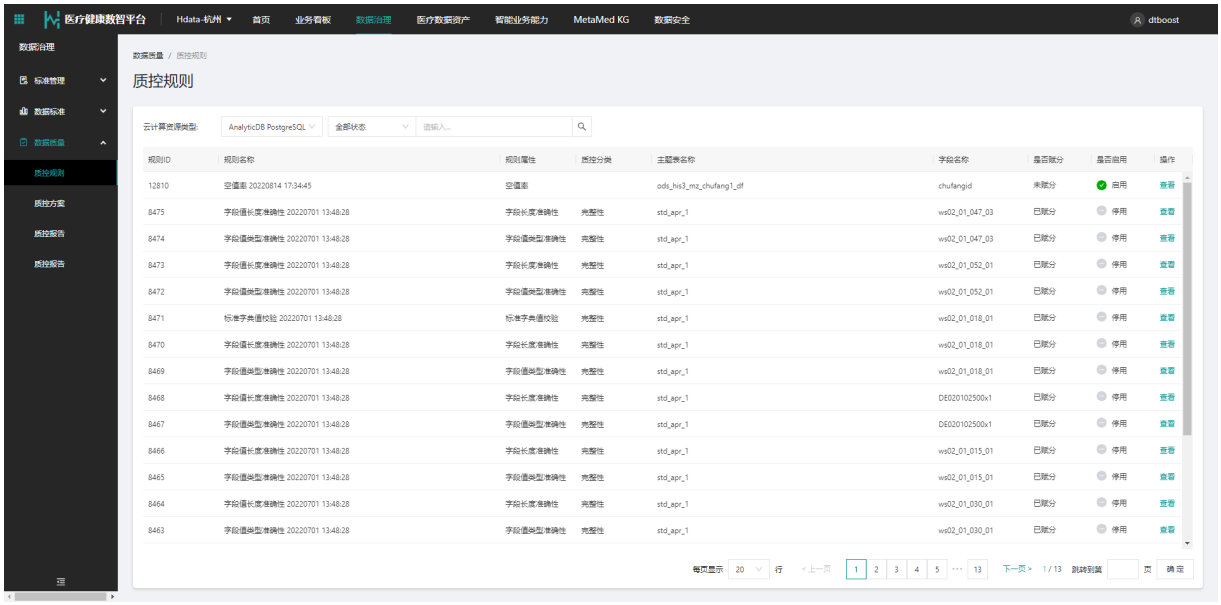
- 支持数据元的查看、导出全表、历史版本查看以及数据元的检索。
- 支持医疗数据值域的查看、导出全表、历史版本查看以及医疗数据值域的检索。
- 支持医疗标准模型的查看、导出全表、历史版本查看以及医疗标准模型的检索。



数据质量

医疗数据质量是医疗智能化应用使用的基础，医疗健康数智平台提供字段级、表级质量规则创建、质量规则分组、业务化的质量方案及计划的创建和监控，并自动化生成结果报告。

- 用户可在研发工作台-数据建模、数据元管理等功能处新建和维护质控规则，支持表级、字段级、自定义SQL数据质量规则。支持MaxCompute、Hive、RDS for MySQL、AnalyticDB PostgreSQL、Kafka、PG存储类型的质量规则。支持根据关联字段标准，自动创建对应质量规则。通过丰富的质量规则，帮助用户全方位的保障数据质量。质控规则页面仅用作数据质控规则配置以及赋分后结果的查看和展示。包括启用/禁用状态的筛选以及规则名称或ID的搜索。
- 质控规则页面仅用作数据质控规则配置以及赋分后结果的查看和展示。包括启用/禁用状态的筛选以及规则名称或ID的搜索。质控方案页面仅做质量方案页面配置结果的展示和查询。
- 质控报告用于统计和展示质控计划执行明细，区分卫健委视角和单个医疗机构视角。统计内容包括已执行次数、执行周期、方案名称、规则数、计划执行状态、质量得分、各类型质控规则整体得分、具体规则得分明细等。

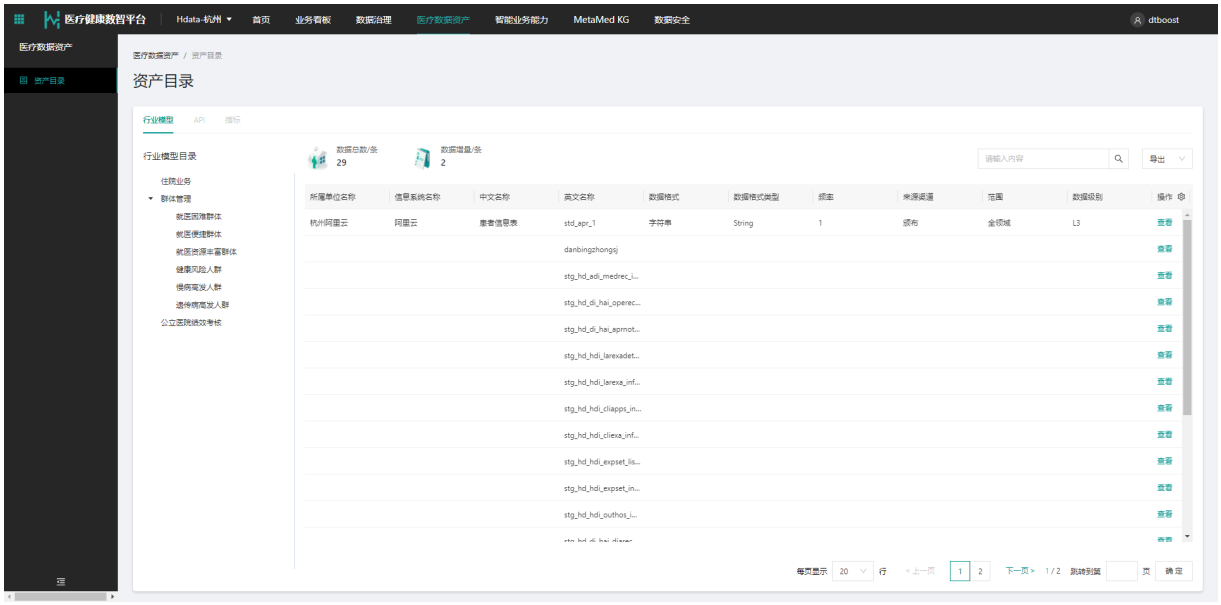


9.3.1.2. 医疗数据资产

医疗数据资产支持对平台可共享的数据资产进行展示、搜索和权限申请，促进数据资产的共享和使用。

资产目录

- 行业模型支持列表展示公开的行业模型资产，用户可通过行业模型类目筛选，或通过关键词搜索，快速定位目标行业模型资产。
- 支持查看医疗行业模型资产的基本信息和字段信息，了解表名称、表描述、数据来源、数据范围、开放共享属性情况、数据安全、字段信息等信息，帮助用户快速了解行业模型资产的基本情况。支持设置表头顺序和展示/隐藏，方便客户个性化展示所关注的信息。
- 支持对有权限的数据表进行数据预览，方便用户了解数据表中存储的内容。



9.3.1.3. 智能业务能力

医疗智能业务应用合集。

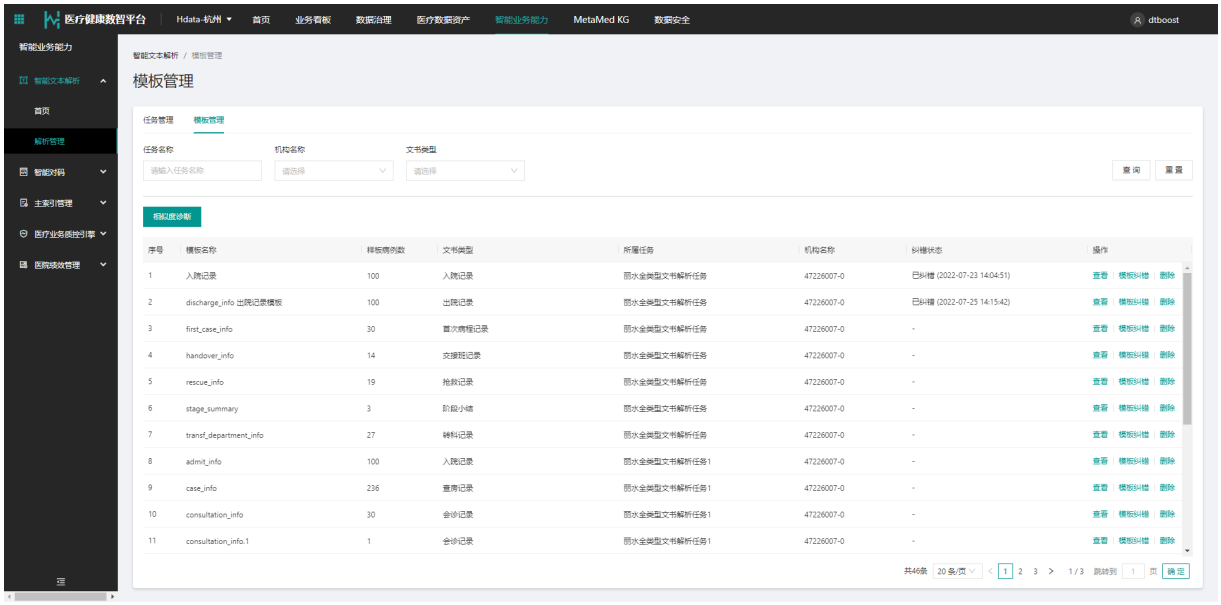
智能文本解析

医疗行业信息化系统使用时间长，部分建设较久的系统数据格式化保存率低，源数据很多以本文的方式存储，给数据分析和结构化带来巨大难度；

HIS这一大块领域没有形成垄断的HIS厂商，同一家医院不同年份都可能采纳不同HIS厂商的产品，再者HIS厂商通常未对医生的书写病历过程提供强制性约束，导致同一类型电子病历文书内部也可能差异很大，实际上，很多医院同一类型文书内部只是大致的统一，内部异质性较强；给同一家医院的时序性分析也带来难度；

文本解析能力是利用NLP能力和医学知识基础，对电子病历等病历文书进行半结构化解析，提取出电子病历文书的模板，以支持下游的实体识别、病历质控等数据分析应用进行。

- 支持解析任务的创建、查看与筛选。解析任务创建支持系统导入与手动上传两种方式。
- 支持解析后的病历模板管理、模板纠错、模板删除。
- 支持模板的相似度诊断。



智能对码

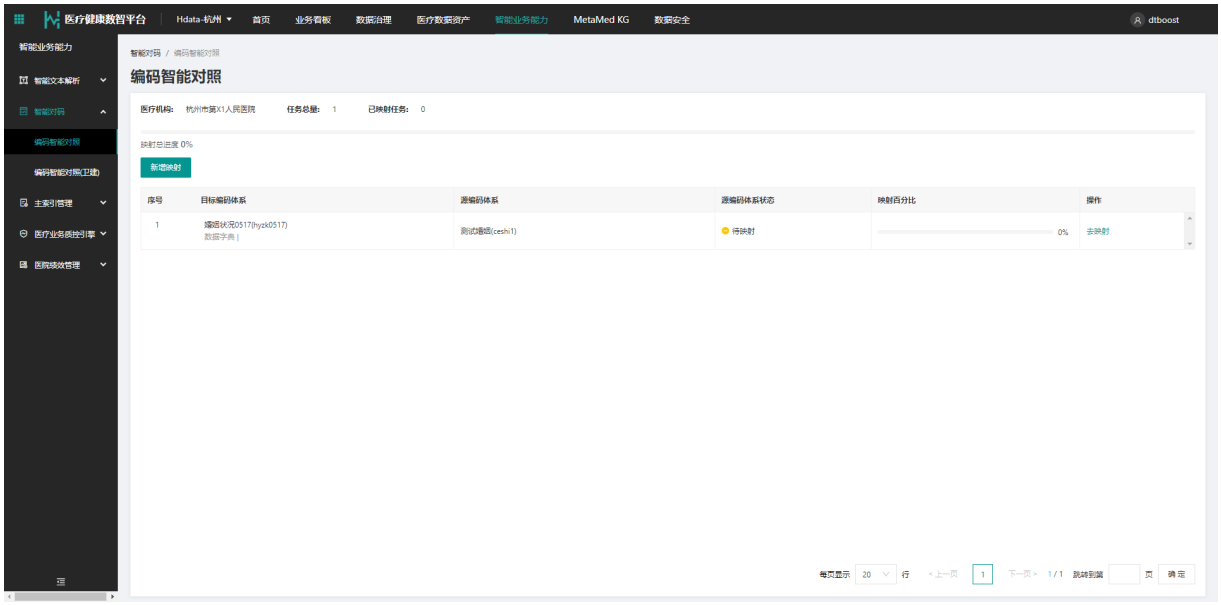
智能对照模块实现目标表与源表的映射，执行映射时系统调用匹配算法匹配目标表与原表字段，并计算出对应的置信度，根据置信度进行批量匹配

本模块由映射任务列表、映射任务详情、智能映射三个部分组成，可实现用户批量映射、手动映射、查看映射进度、查看映射关系、解除映射关系的需求。此模块用户包括卫健委管理员与医疗机构对码员，由于使用人员的角色不同，此模块面对两者提供的功能权限和数据权限也不同，如下所述：

卫健委管理人员：查看映射任务列表、查看映射任务列表详情，管理人员可查看所有的映射任务。

医疗机构对码人员：新建映射任务，查看映射任务列表、查看映射任务列表详情，导入源表数据、进行批量映射、进行手动映射、解除映射。各医院信息科人员仅可查看自己所属医院的映射任务。**需要注意的是：**医院对码人员在创建映射任务时同时也建立了一个源编码体系映射表，此表存储于后台，用于记录源编码体系的名称、编码和值集、待导入值集后进行映射。

- 支持新增和管理编码映射任务。
- 支持批量导入源值，可以进行批量映射和单个新增。
- 支持自动编码匹配和手动编码匹配。
- 支持卫健视角查看下属医疗机构的编码映射进度。



主索引管理

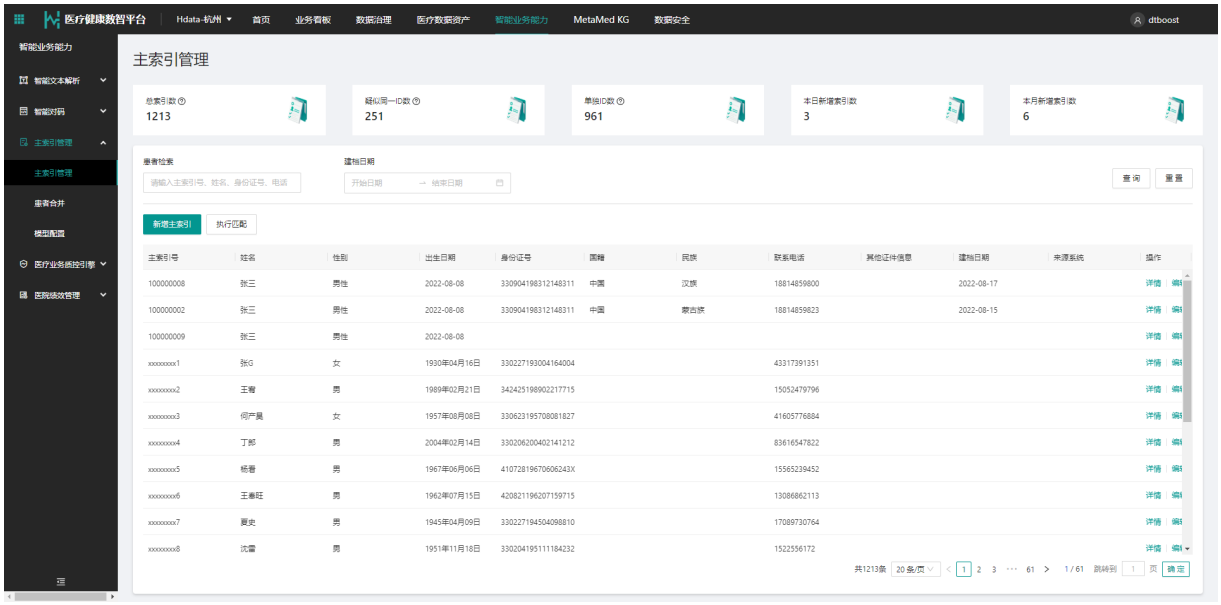
“一人多卡、一人多号”由于用户可能在不同业务系统进行首次登记，且可能随着时间的变化二次登记，比如在急诊用身份证挂号一次，在门诊用社保卡挂号一次，在急诊因为紧急情况，无挂号只保留手机号等就诊一次等，这些不同的记录，实际上是一个患者同人，相关的就诊记录、体征信息、患病信息等，需要打通，在患者360、用药提醒、全面健康档案、健康管理等业务场景下，需要结合各系统的信息做综合分析。

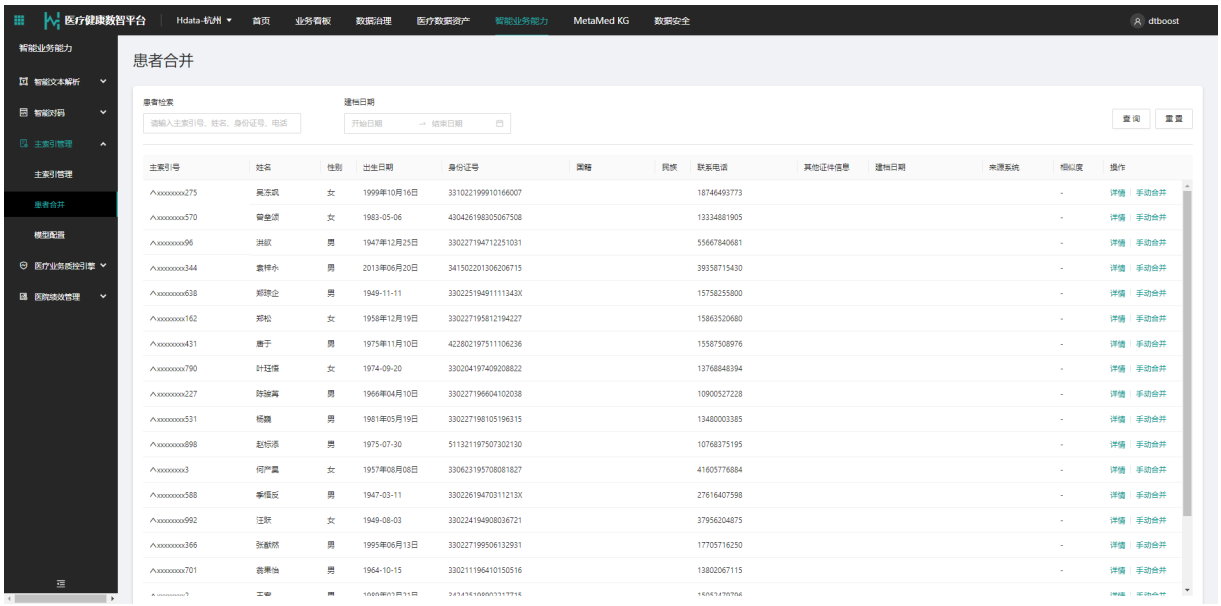
患者主索引：核心提供患者主索引看板、主索引管理、主索引规则配置等能力；内置患者同人ID算法模型，并支持用户对模型做规则配置；支持用户对识别后的索引列表进行合并、拆分等操作。实现索引整合后，以打通不同业务系统中患者相关的就诊记录、体征信息、患病信息等，支持上层患者360、用药提醒、全面健康档案、健康管理等业务场景下使用。

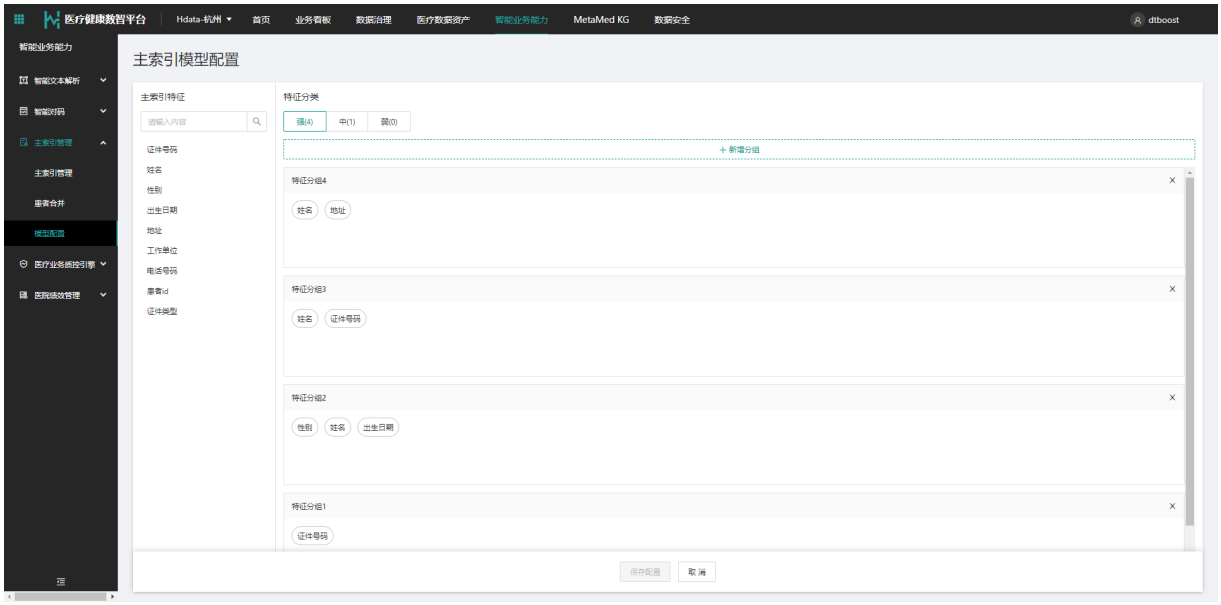
EMPI本质上是一个数据整合系统，它把来自多个不同的系统中的病人标识统一成一个，实现对同一病人仅仅有一个标识对应，同一病人的信息也归并在同一个标识之下。这样就有效地解决了多系统中识别病人身份的问题。

EMPI系统会为每个注册的医院信息系统分配一个Domain ID，用以标识每个医院信息系统。当增加病人信息（或医生的基本信息）的非主数据部分后，医院信息系统则会产生一个Internal ID，与此同时，医院信息系统将发出请求，在EMPI系统中添加相应的病人（或医生）信息的主数据部分。在添加完主数据后，EMPI系统会为其生成一个Global ID，并且建立起Global ID 与 Domain Id、Internal ID 的映射关系，从而实现了 PIX（patient identifier cross-reference） 功能。

- 支持手动新增主索引，也可以通过数据同步完成主索引新增。
- 支持主索引相关数据的统计。
- 支持对主索引信息进行编辑和查看以及检索。
- 支持手动执行主索引合并匹配算法，也设置定时任务执行。
- 主索引合并支持忽略与选择主要主索引信息。
- 支持自主配置患者合并匹配算法的模型。


医疗健康数智平台
Hdata-杭州
首页
业务看板
数据治理
医疗数据资产
智能业务能力
MetaMed KG
数据安全
dtboost
智能业务能力
智能文本解析
智能对码
主索引管理
主索引管理
患者合并
模型配置
医疗业务质控引擎
医院绩效管理
主索引管理
总索引数
1213
疑似同一ID数
251
单独ID数
961
本日新增索引数
3
本月新增索引数
6
患者检索
建档日期
查询
重置
新增主索引
执行匹配
主索引号
姓名
性别
出生日期
身份证号
国籍
民族
联系电话
其他证件信息
建档日期
来源系统
操作

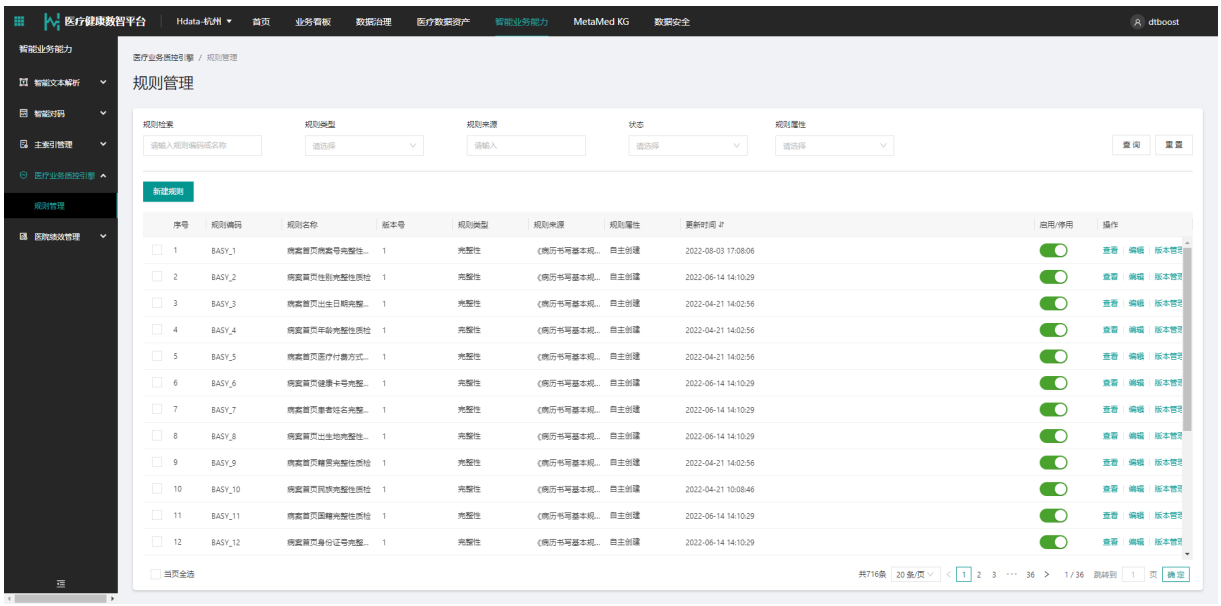

医疗健康数智平台
Hdata-杭州
首页
业务看板
数据治理
医疗数据资产
智能业务能力
MetaMed KG
数据安全
dtboost
智能业务能力
智能文本解析
智能对码
主索引管理
主索引管理
患者合并
模型配置
医疗业务质控引擎
医院绩效管理
患者合并
患者检索
建档日期
查询
重置
主索引号
姓名
性别
出生日期
身份证号
国籍
民族
联系电话
其他证件信息
建档日期
来源系统
相似度
操作



医疗业务质控引擎

医疗业务质检规则的开发与管理。

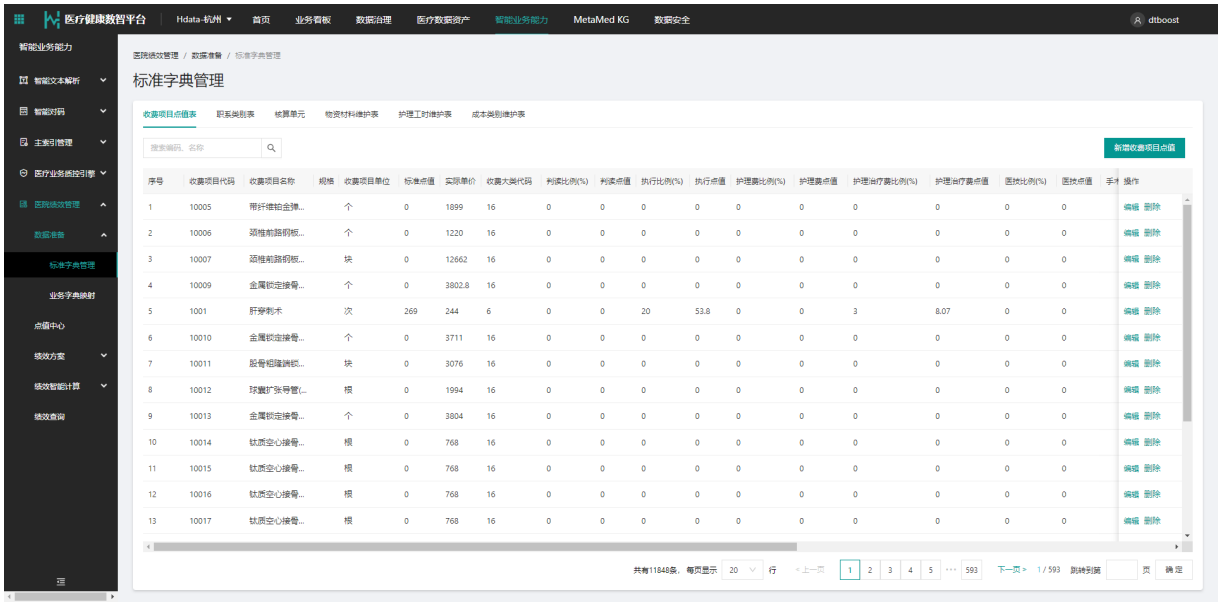
- 支持质检规则的开发、调试、发布以及版本管理。
- 支持质检规则的查看、编辑、启用/停用、检索。

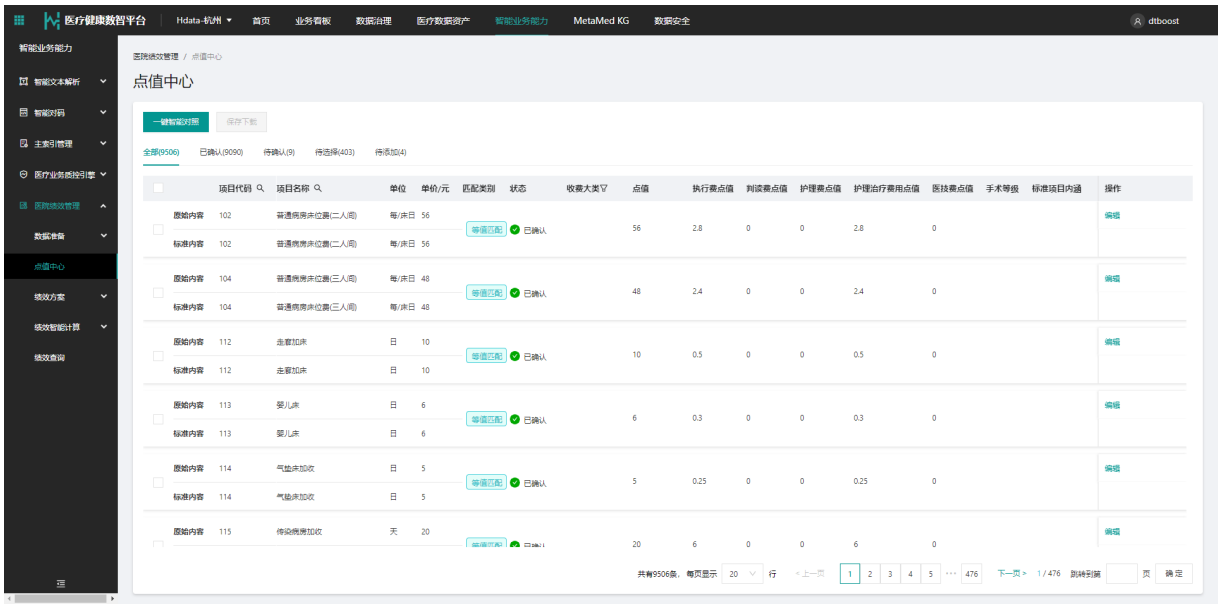


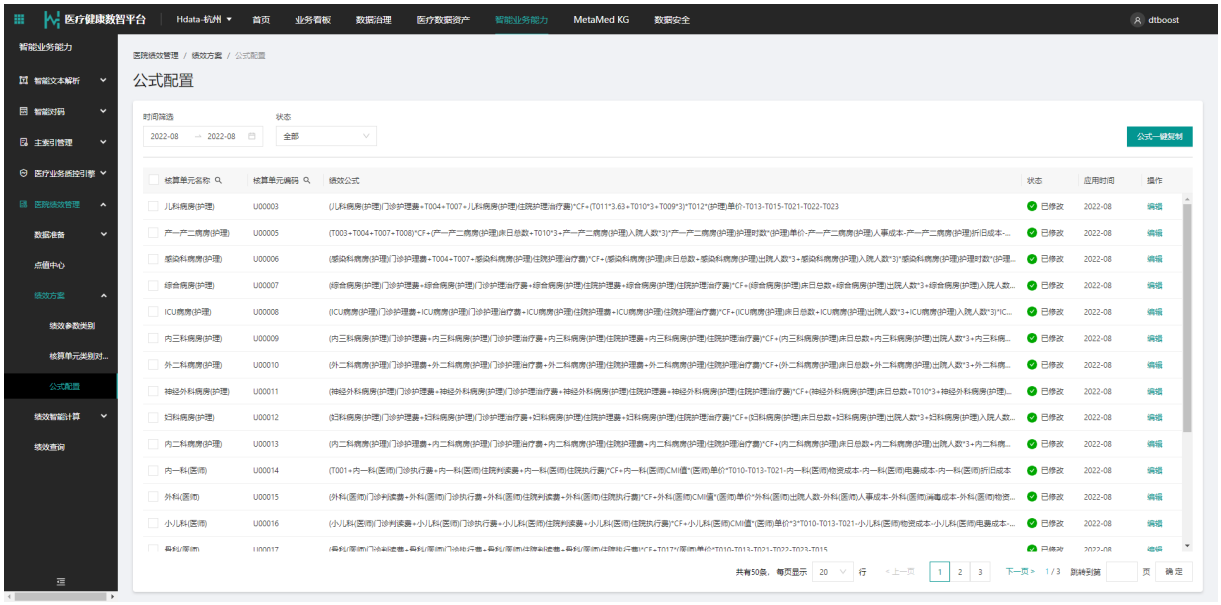
医院绩效管理

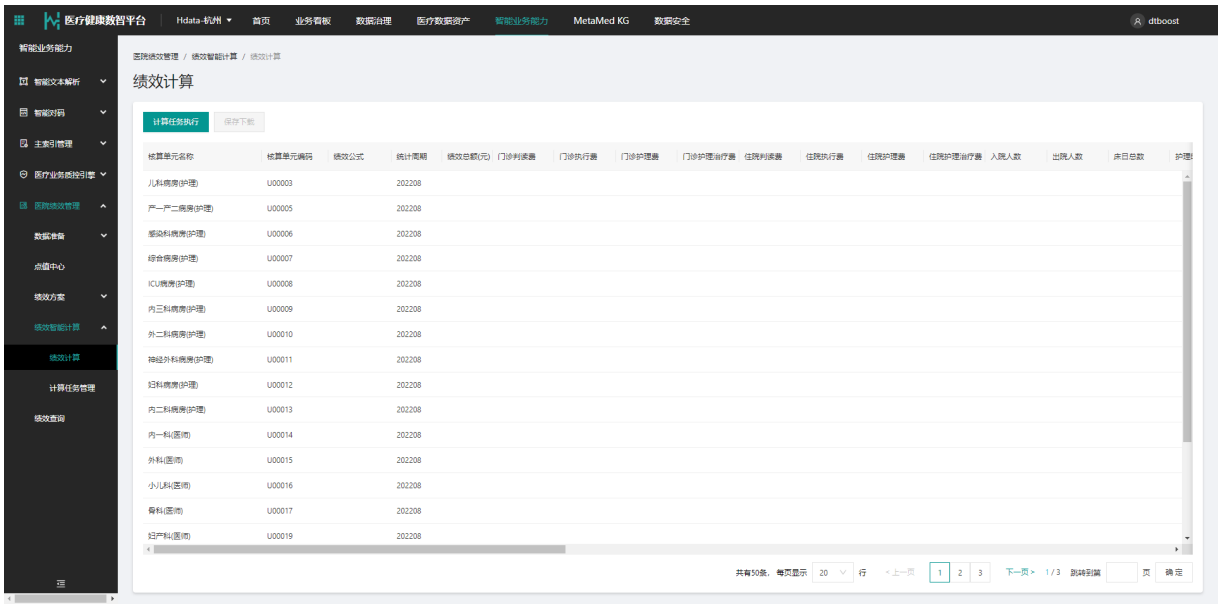
医院绩效管理系统是基于智能数据引擎-医疗产品的数据计算能力、分析能力和数据底盘，构建指标走势实时感知、人工处置、方案评估、诱导策略为一体的绩效管理平台。

- 支持标准字典的维护和业务字典的映射与展示。
- 支持点值中心绩效核算项目的维护，点值的编辑，支持一键智能对照医院原始点值和标准点值。
- 支持绩效方案的设定，包含绩效参数的设定，核算单元参数类别对照，绩效计算公式的设定。
- 支持智能绩效计算和绩效查询。


医疗健康数智平台
Hdata-杭州
首页
业务看板
数据治理
医疗数据资产
智能业务能力
MetaMed KG
数据安全
dtboost
智能业务能力
智能文本解析
智能对码
主索引管理
医疗业务感控引擎
医院绩效管理
数据准备
标准字典管理
业务字典映射
点值中心
绩效方案
绩效智能计算
绩效查询
医院绩效管理 / 数据准备 / 标准字典管理
标准字典管理
收费项目点值表
职系类别表
核算单元
物资材料维护表
护理工时维护表
成本类别维护表
新增收费项目点值
序号
收费项目代码
收费项目名称
规格
收费项目单位
标准点值
实际单价
收费大类代码
判读比例(%)
判读点值
执行比例(%)
执行点值
护理费比例(%)
护理费点值
护理治疗费比例(%)
护理治疗费点值
医技比例(%)
医技点值
手术
操作
编辑 删除
共有11848条，每页显示 20 行
1 2 3 4 5 593
下一页 1 / 593 跳转到第 页 确定


医疗健康数智平台
Hdata-杭州
首页
业务看板
数据治理
医疗数据资产
智能业务能力
MetaMed KG
数据安全
dtboost
医院绩效管理 / 点值中心
点值中心
一键智能对照
保存下载
全部(9506)
已确认(9090)
待确认(9)
待选择(403)
待添加(4)
项目代码
项目名称
单位
单价/元
匹配类别
状态
收费大类
点值
执行费点值
判读费点值
护理费点值
护理治疗费用点值
医技费点值
手术等级
标准项目内涵
操作
原始内容
标准内容
等值匹配
已确认
编辑
共有9506条，每页显示 20 行
1 2 3 4 5 476
下一页 1 / 476 跳转到第 页 确定

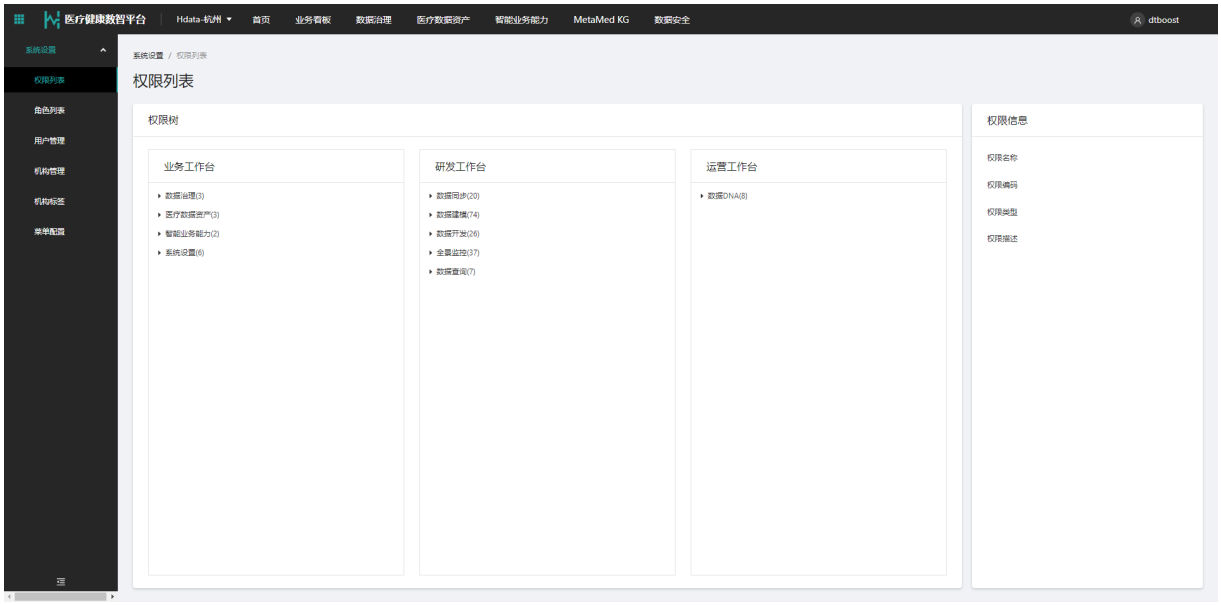




系统设置

对智能业务能力中的各个医疗业务产品进行基础的权限、角色、用户、机构等的配置。

- 支持权限列表的展示、角色和用户的管理。
- 支持机构和机构标签的管理。
- 支持自定义功能菜单配置。



9.3.2. 研发工作台

9.3.2.1. 资产加工

9.3.2.1.1. 数据同步

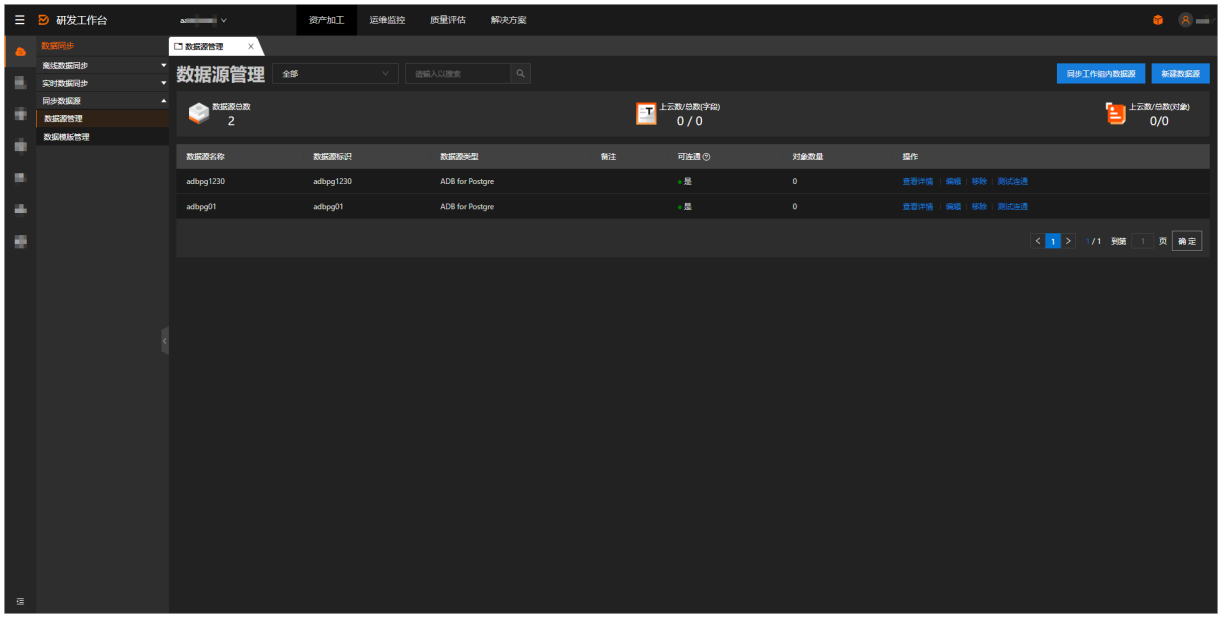
数据同步提供批量、实时多源异构数据的便捷同步或接入能力，系统提供完善的数据接入配置、任务运行监控等功能，有效保障数据接入的稳定性和可控性，满足各类平台、数据源及应用系统间的数据汇聚需求。同时支持对存储在基于ArcGIS、超图等主流GIS平台、开源PostGIS空间数据库中的各类空间数据进行同步汇聚。

通过数据同步的自定义同步任务配置能力，可实现快速创建周期数据同步任务和实时同步任务。用户可选择已经注册的数据源端数据，并选定目标端创建同步任务。系统提供目标端自动建表、源端目标端字段自动映射、批量创建同步任务和引用数据模块进行同步任务配置等多种便捷同步任务创建能力，帮助用户更便捷的完成数据同步配置。

在完成数据同步配置后，数据同步模块可对数据同步任务的执行和运行情况进行全面的监控运维。可对任务总数、线上任务、正运行任务等核心指标进行统计，并可以实现异常任务执行状态的监控，当前实例状态监控。对任务的开始时间、结束时间、耗时、速率、状态和执行日志等信息进行跟踪和展示，并支持周期任务重跑。

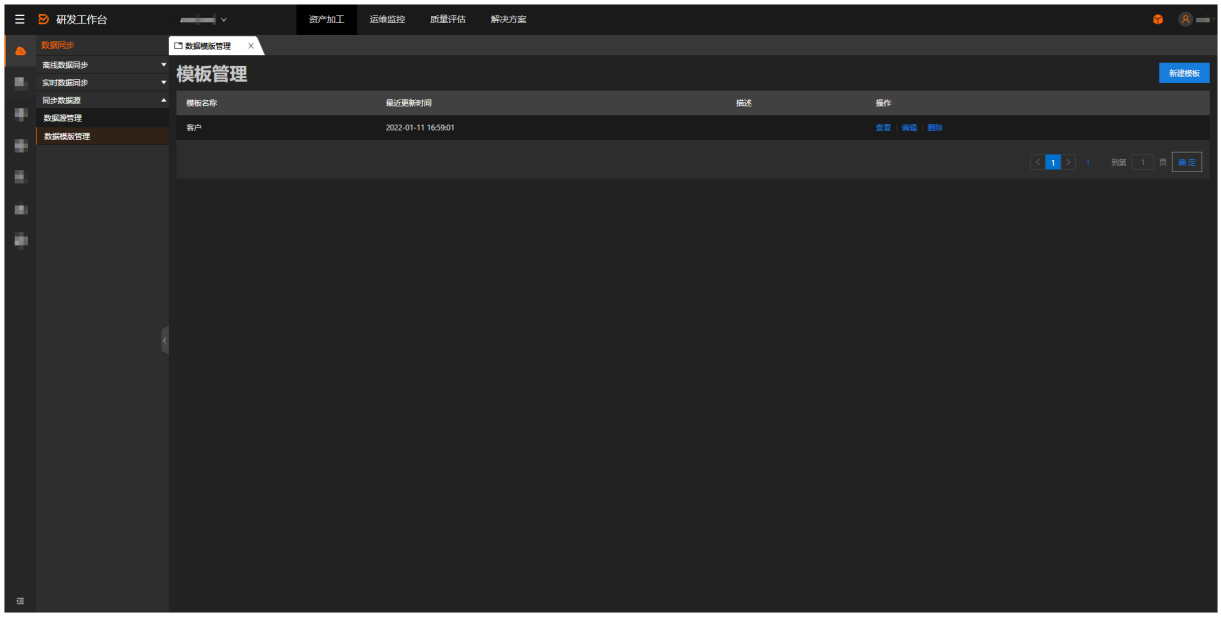
数据源管理

- 支持对数据同步的数据源端和目标端进行统一的注册和管理，支持对关系型数据库、文件、消息队列等多种类型数据源的配置，并可对所填写数据源的连通性进行有效性验证。支持对已配置联通的数据源进行元数据同步及数据对象查看，帮助用户了解元数据资产内容。
- 支持一键同步工作组内已配置的数据资源，快速对工作组已有资源在数据同步任务中进行应用。



数据模板管理

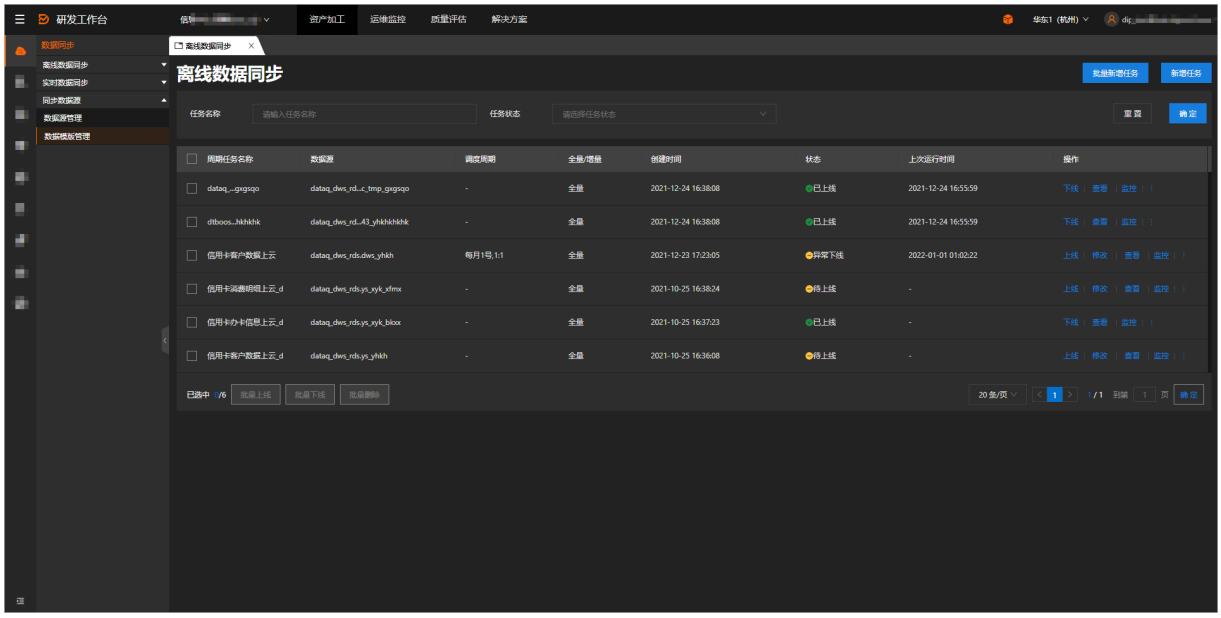
- 针对半结构化、非结构化数据，支持创建数据模板，支持自定义数据字段及字段类型，并提供编辑、删除等功能。
- 支持在离线数据同步、实时数据同步任务中引用已添加的数据模板，根据数据模板定义的数据结构执行数据同步任务。



离线数据同步

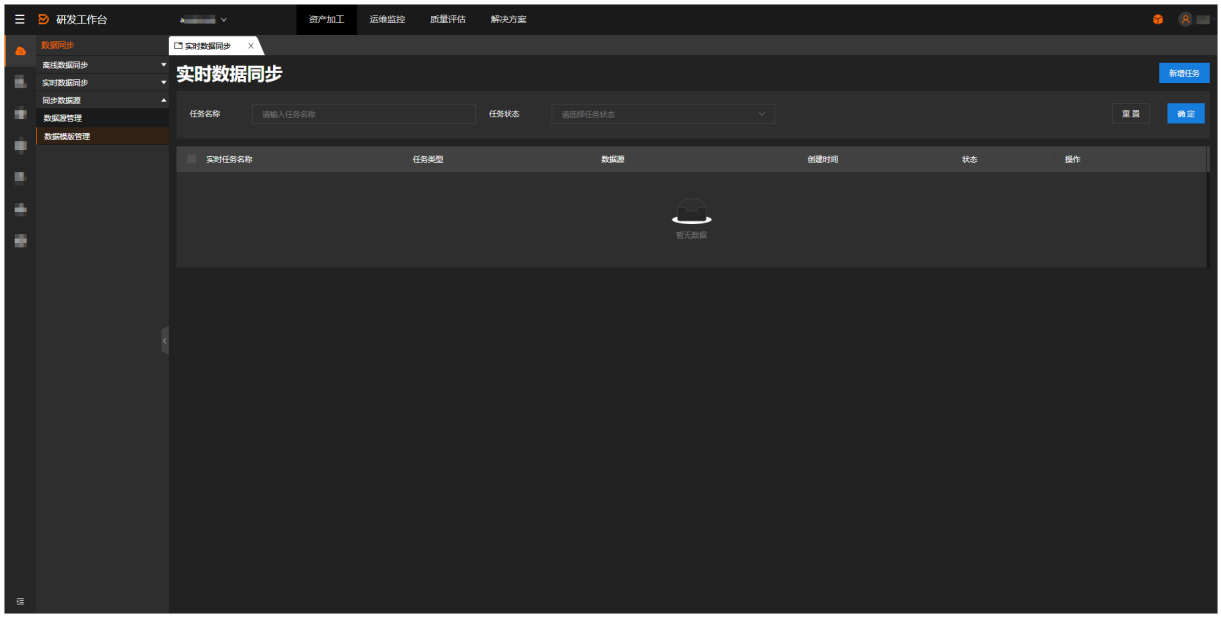
- 支持单表离线任务、批量离线任务的新建。
- 支持自定义选择已经注册的数据源端数据和目标端，支持多种常见离线同步链路，如PostGIS、Ganos、超图SDX和ArcGIS SDE的空间数据同步到PostGIS和Ganos中。
- 提供全量任务自动建表、同名字段自动映射、任务调度配置功能。
- 支持对离线任务实例的运行属性和运行日志进行监看。

- 支持对非增量离线同步实例进行实例重跑。



实时数据同步

- 支持目标端为DataHub的单表实时任务的新建。
- 支持Oracle CDC、MySQL binlog、SQL Server CDC、RabbitMQ、Kafka五种实时数据同步模式。
- 提供完善的实时数据接入配置等功能，有效保障实时任务接入的稳定性和可控性。



9.3.2.1.2. 数据查询

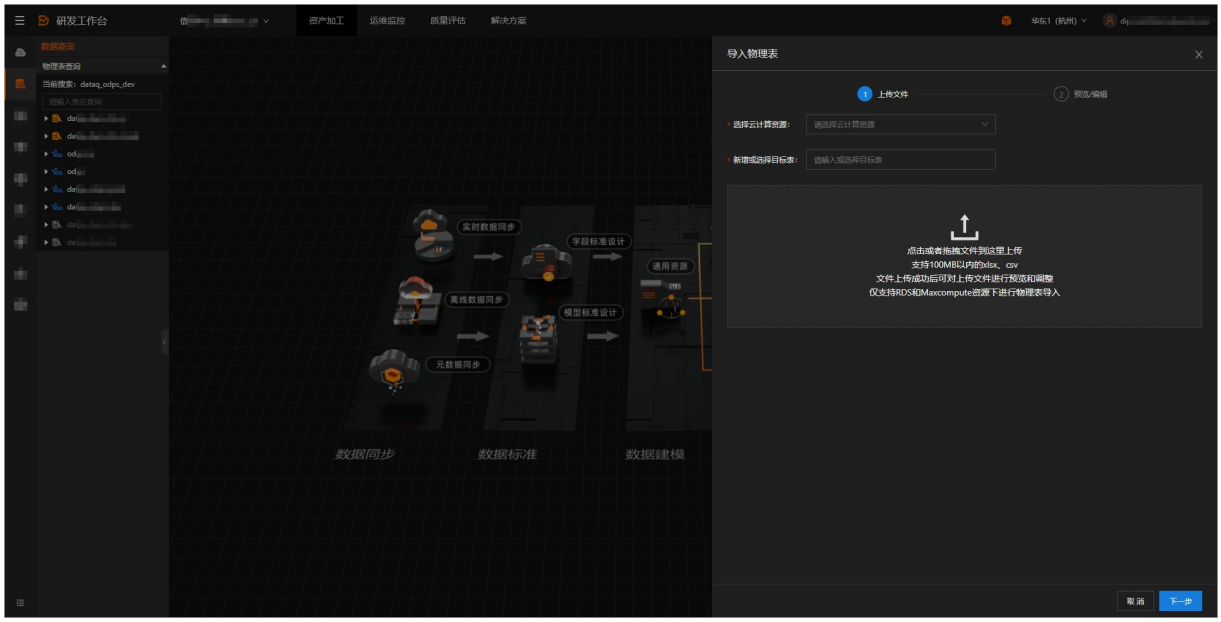
数据查询为用户提供对云计算资源中物理表基本情况的概览能力，帮助用户直观了解物理表基础信息、字段内容分布等情况，建立对数据表的宏观理解，为建立数据标准、定义数据模型、定义数据加工逻辑等工作提供基础依据。

如果数据表格存储的是空间数据，支持在表格的基础上开启地图模式，并实现表格和地图的双向联动。还能查看空间的元数据系统，以及支持SQL的方式进行查询浏览。同时支持对汇聚同步后的空间数据进行地图可视化

浏览，支持地图、属性联动，支持对空间元信息（坐标系、空间范围等）进行全方位查看。

物理表导入

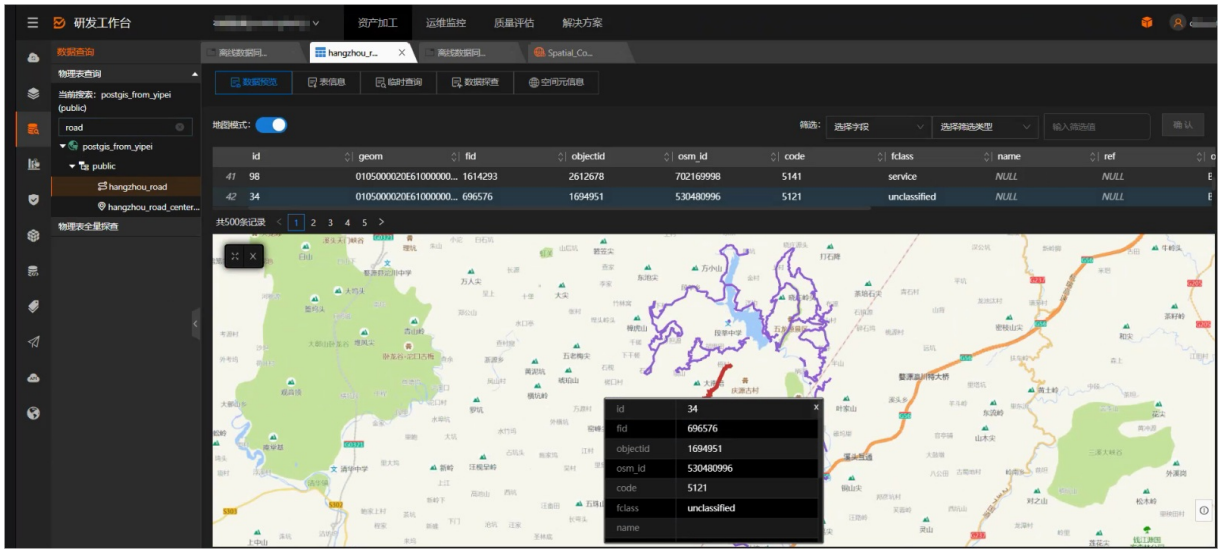
支持100 MB以内的XLSX、CSV文件快捷向RDS MySQL和MaxCompute资源进行导入，支持对要导入的文件进行预览和编辑。



数据预览

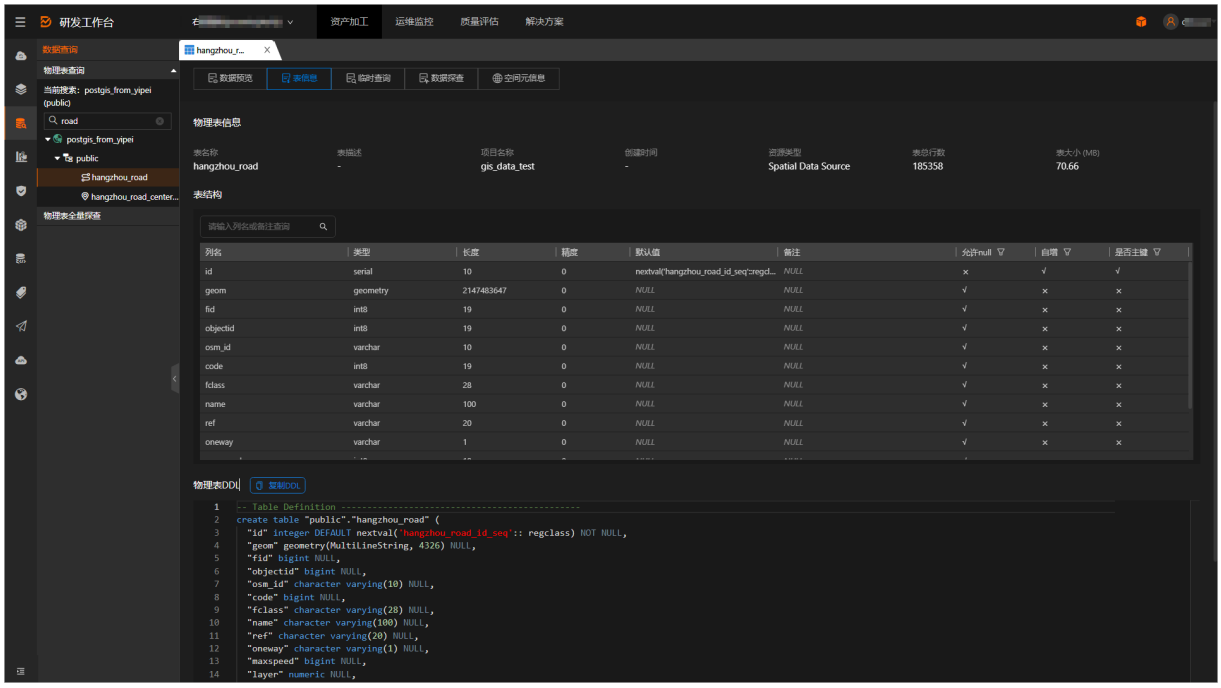
支持不同云计算资源数据的快速预览，可为开发人员和数据应用人员提供全面直观的数据呈现。

- 支持MaxCompute、RDS MySQL、PostgreSQL、AnalyticDB PostgreSQL、Hologres多种云计算资源进行物理表数据预览，可快速对各类数据资源下的物理表的数据内容进行查看。
- 支持开启地图模式，并实现表格和地图的双向联动。
- 支持对物理表字段进行排序，支持根据条件筛选。



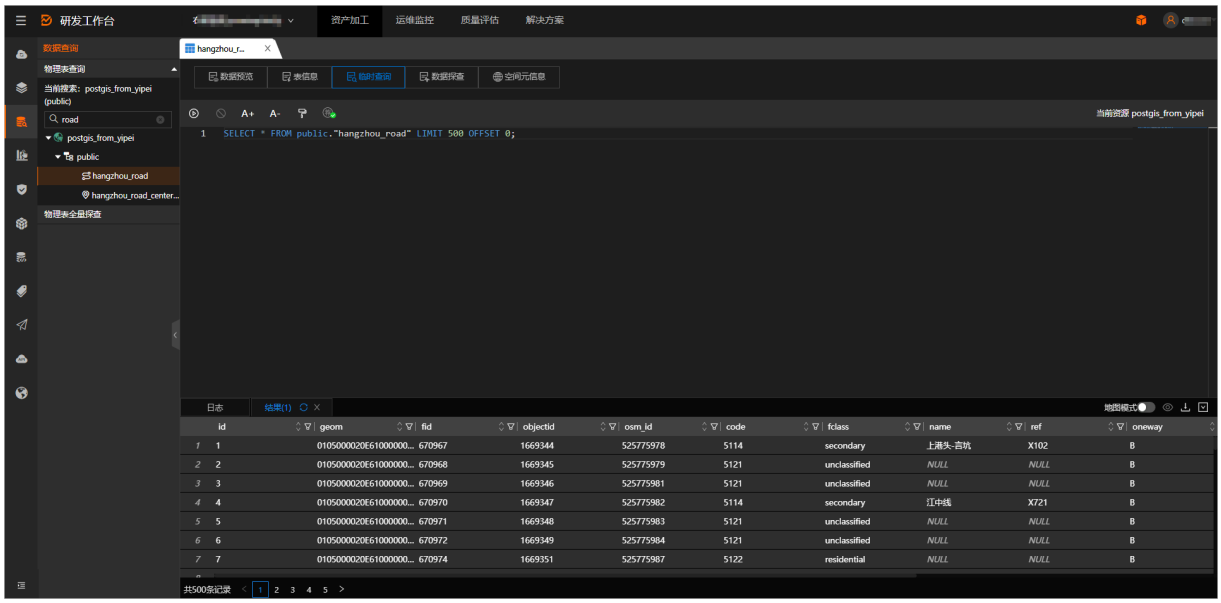
表信息查看

支持不同云计算资源的表信息查看，包括物理表基本信息、表结构和物理表DDL。



临时查询

- 支持通过SQL语句查询预览不同云计算资源的数据资源，可为开发人员和数据应用人员提供全面直观的数据呈现。
- 查询结果支持开启地图模式，并实现表格和地图的双向联动。

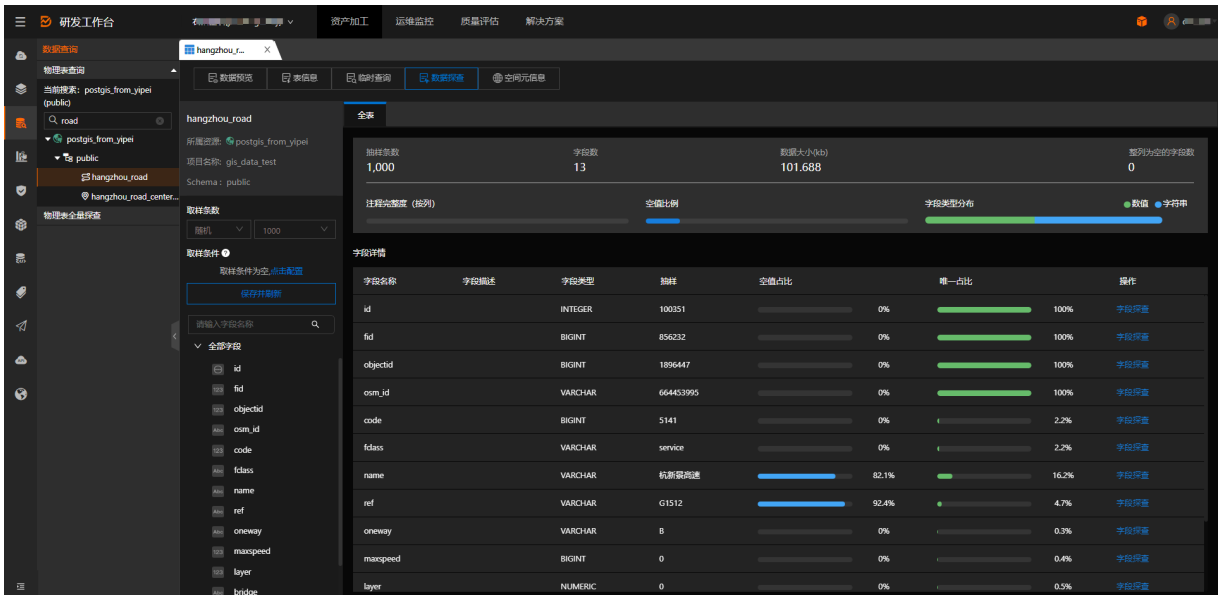


数据探查

支持通过内置的规则及算法对物理表进行探查，快速对物理表整体数据情况、字段情况进行抽样检测。以图表的形式展示物理表数据整体情况，为数据质量规则期望值的设定和数据开发提供输入和依据。

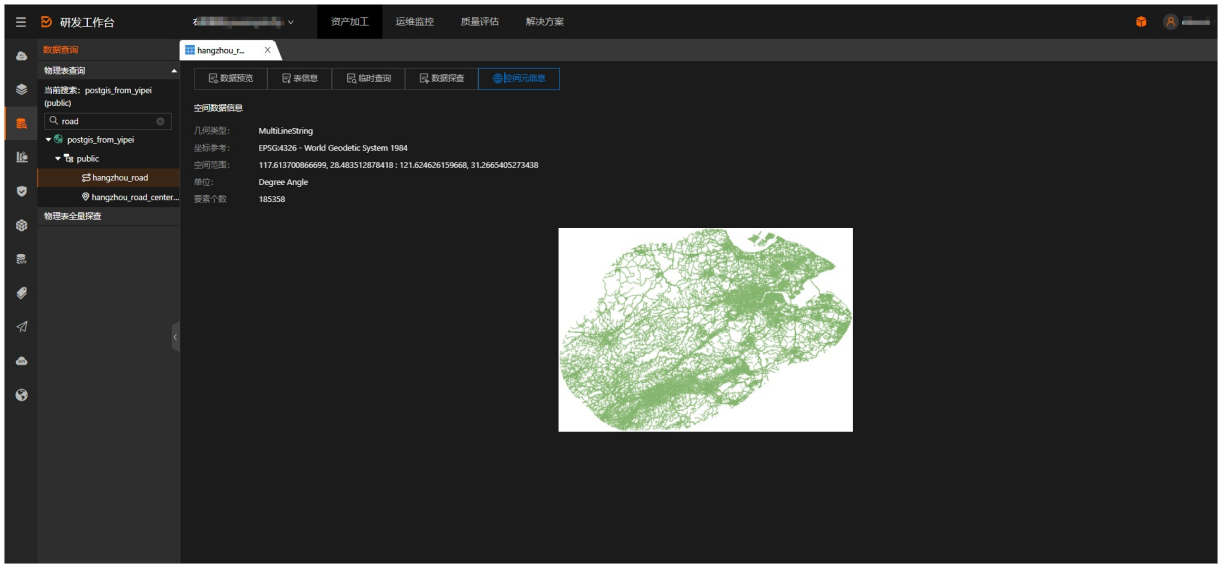
- 支持展示完整表结构、注释完整度、空值率、字段类型分布等表数据分布情况。
- 支持自定义抽样探查规则，支持输入条件组合对标数据进行筛选取样。

- 支持通过内置算法对字段进行分析，包括唯一值、偏斜度、字段频次、异常分布等。以图表形式对分析结果进行展示，并支持对统计图进行下载。



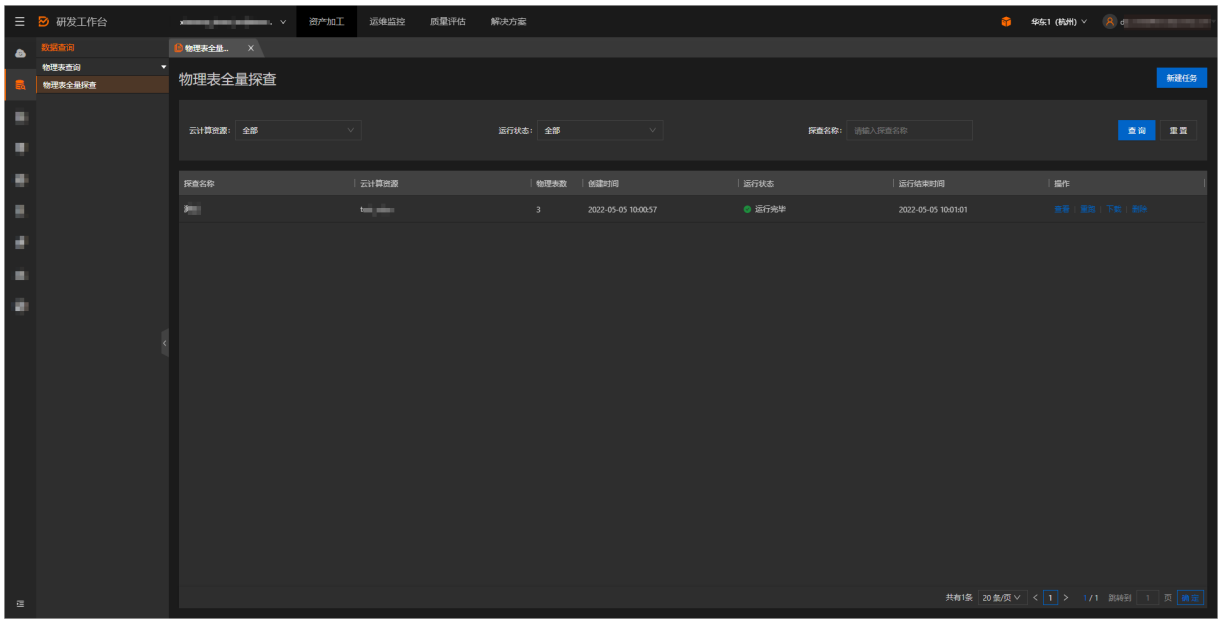
空间元信息

支持查看空间数据基本信息和空间数据缩略图，基本信息包括几何类型、坐标参考、空间范围、单位、要素个数等。



物理表全量探查

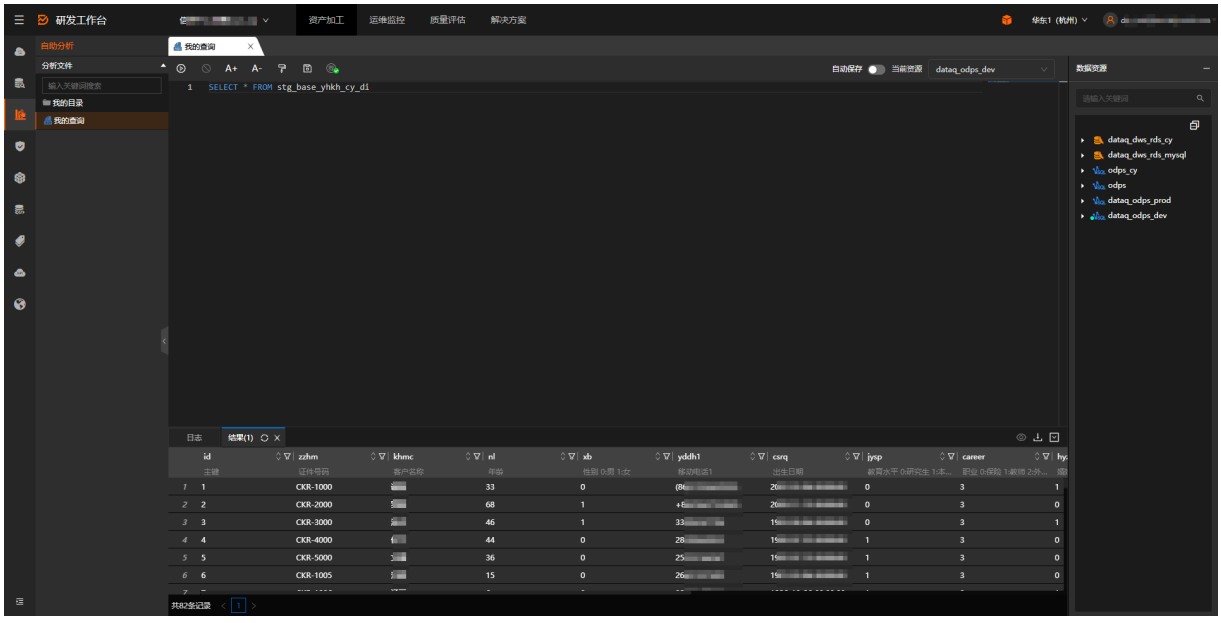
支持新建离线任务针对选中计算资源下物理表的全部字段进行全量探查，包括字段类型、唯一值、空值占比、最大最小值等，探查结果支持以Excel报告形式进行下载。



9.3.2.1.3. 自助分析

自助分析为用户提供对云计算资源中物理表的数据查询能力。

- 持通过SQL语句查询预览不同云计算资源的数据资源，可为开发人员和数据应用人员提供全面直观的数据呈现。
- 支持将查询过程保存，以便下次查阅。

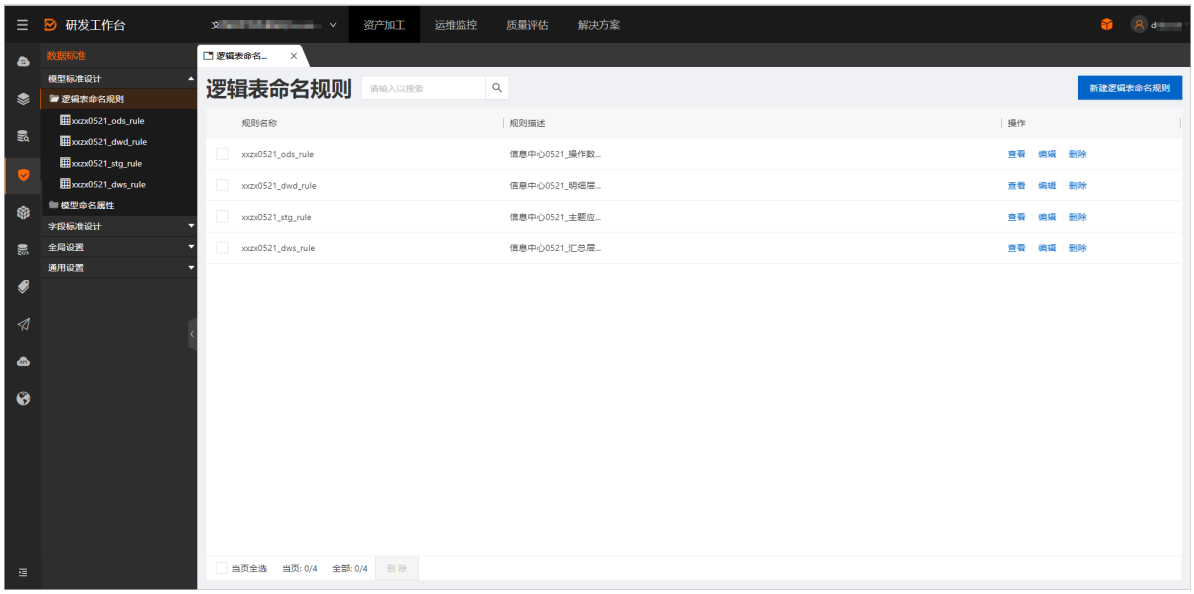


9.3.2.1.4. 数据标准

数据标准为用户提供模型和字段标准设计和管理的能力。通过模型层级、数据域等属性，对逻辑表的命名进行标准化设计；通过设计标准数据元、数据指标、维度和维度属性，对关键业务对象以及属性和值域进行定义，制定并管理平台遵循的统一数据标准，管控数据的一致性和数据质量。

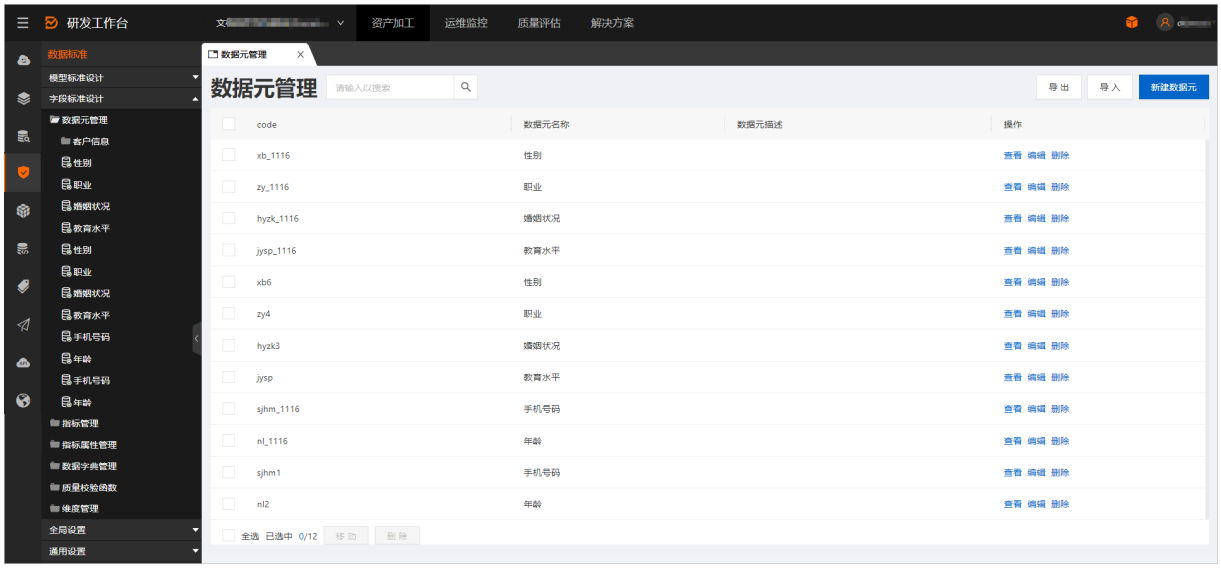
模型标准设计

- **逻辑表命名规则**：支持对逻辑表命名标准的规范设计，可基于模型命名属性，对模型命名规则进行配置，规范统一后续模型设计的命名标准。
- **逻辑表命名属性**：提供对逻辑表命名组成中的一些常用属性进行自定义配置功能，如刷新频率、增量、统计周期。



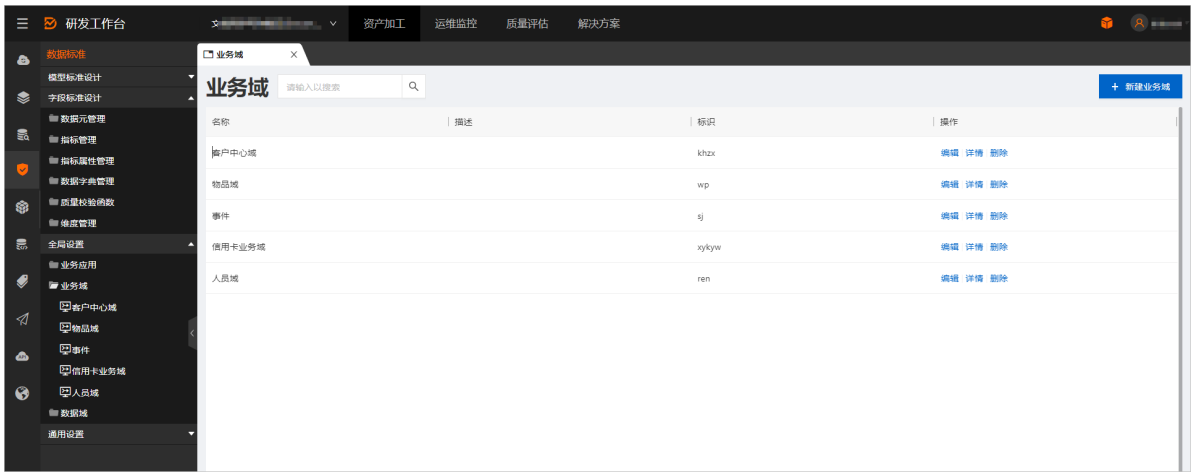
字段标准设计

- **数据元管理**：提供对数据元的标准定义和统一管理，通过分析业务流程，抽象关键业务对象、业务对象属性，配置数据元的类型、长度、质量校验函数和引用数据字典，为逻辑表标准化设计及数据治理过程中的质量监控提供规则依据；支持数据元导入、导出，方便行业经验积累和跨项目复制。
- **指标管理**：支持指标类型数据元的定义和管理，可通过选择指标属性的方式，快速新建指标数据元，供逻辑表设计时进行使用。
- **指标属性管理**：支持指标定义的相关属性管理，包含时间修饰、业务修饰、度量属性，帮助快速新建指标数据元。
- **数据字典管理**：支持创建并维护逻辑表中使用的数据字典，方便用户在定义数据元或维度属性时进行数据字典引用，支持数据字典物理化、数据字典导入、导出，方便数据字典使用与跨项目复制。
- **质量校验函数**：支持质量检测函数自定义，方便用户在定义数据元或维度属性时进行质量规则引用，以支持后续对字段的质量规则的快速配置和校验。
- **维度管理**：支持对维度表的统一管理，通过分析业务流程、抽象关键业务对象、业务对象属性，配置维度属性的字段类型、长度、质量校验函数和引用数据字典，为模型标准化设计及数据治理过程中的质量检测提供依据，并提供导入、导出、删除、编辑等功能。



全局设置

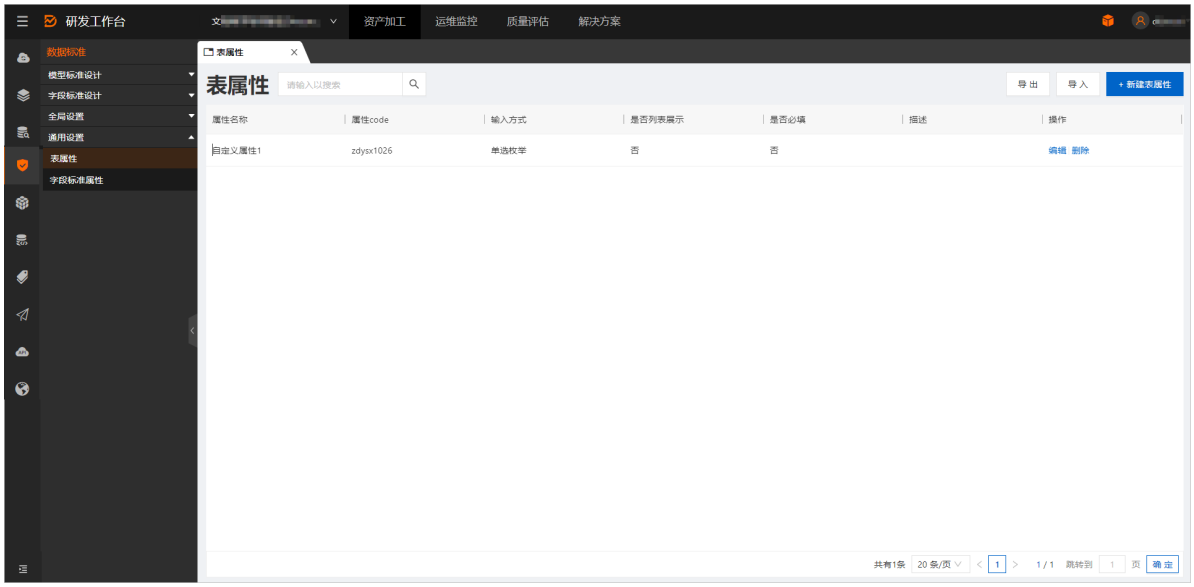
- **业务应用**：对数据模型支撑的应用系统信息进行创建、管理。
- **业务域**：对数据模型来源于哪些系统信息进行创建、管理。
- **数据域**：提供对数据域的定义与管理，供逻辑表设计时进行使用。



通用设置

- **表属性**：对逻辑表的属性进行配置、管理。

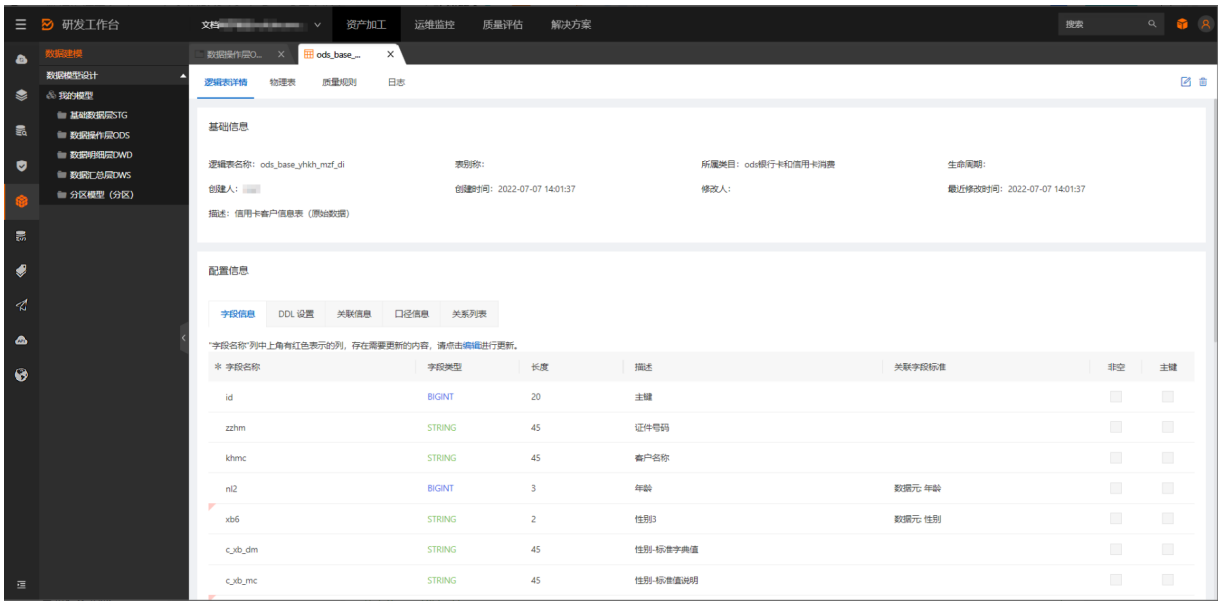
- **字段标准属性**：对字段标准的属性进行配置、管理。



9.3.2.1.5. 数据建模

数据建模支持ER模型和维度模型两种数据建模方法，帮助用户完成数据模型、逻辑表的设计和管理。用户可对数据模型的层次结构进行管理，并对逻辑表字段进行设计。字段设计需要在数据标准的约束下，通过模型设计过程中引用标准数据元定义，设置逻辑表相关标准质量规则，保证不同的逻辑表设计人员按照统一口径进行操作。

在数据标准的约束下，保证逻辑表的可持续的维护性及可读性。在数据模型管理模块对逻辑表进行层次结构设计、逻辑表定义，引用标准元素定义，保证不同的逻辑表设计人员按照统一口径进行操作，除此以外还支持：逻辑表质量配置、DDL设置、逻辑表物理化、表关系配置等功能。



- **逻辑表层级管理**：支持对逻辑表层级进行新建，并提供编辑、删除的功能。
- **新建逻辑表**：支持通过可视化的形式创建逻辑表的基本信息、字段信息、关联信息，提供DDL语句新建、引用现有表字段、引用标准库字段、手动添加字段等方式快速创建逻辑表字段及字段信息功能。

- **云计算资源扫描：** 支持将云计算资源的已有物理表Meta信息扫描到逻辑表相应层级下，并提供通过正则表达式筛选需要扫描的物理表范围。
- **Excel导入：** 支持将Excel里的逻辑表表名、字段信息、所属层级、关联信息导入数据模型管理，便于客户快速搭建数据仓库。
- **Excel导出：** 支持将数据模型设计模块中的逻辑表及信息以Excel的形式导出，并提供按全量模型与按查询条件的方式进行导出。
- **DDL设置：** 支持根据逻辑表，自动生成并保存不同云计算资源类型的DDL语句。
- **物理化：** 支持将已存在的逻辑表物理化到数据库，并对物理化执行结果信息进行展示。
- **表关系配置：** 支持建立多个表之间的关联关系，并提供关联关系编辑功能。
- **质量规则配置：** 支持可视化与自定义SQL方式制定表级与字段级质量规则，并提供手工创建逻辑表字段时通过关联数据元自动生成质量规则的功能。
- **一致性：** 支持当前逻辑表和已物化物理表的字段一致性对比，支持当前逻辑表与数据开发所引用的节点中模型的字段一致性对比。

9.3.2.1.6. 数据开发

数据开发为用户提供一站式计算节点开发能力，通过对数据加工流程的开发、部署、调试等环节的一体化管理，数据开发实现数据加工工作流编排、加工逻辑的复用，大幅提高数据开发效率。

数据开发帮助用户优化智能系统的架构，提高系统的工程化水平，让用户能够集中精力处理计算逻辑问题，提高开发效率。数据开发除了在系统层面上集中管理计算节点，降低计算节点和系统其他组件的耦合之外，也对计算节点内部的实现提供了一组规范。在实际的业务系统中，数据来源多种多样，不同数据对数据处理的时延和数据量的要求不同，这就需要综合多种不同的平台，包括批量、流式环境等。进行数据处理时，上一个平台计算完毕后把数据传递给下一个平台进行计算，多个平台互相配合来完成任务。数据开发定义了各平台间的数据流转规范，以及各平台计算节点执行顺序的方法，从而让用户能够更加专注于计算逻辑的开发，不必关注底层的技术细节。

数据开发规范了各个计算节点的接口，在一定程度上提高了计算节点复用的能力，降低了多人协作开发的难度。提供了跨多种存储/计算平台的计算节点工作流程配置和快速部署上线的能力。因此，用户在构建数据应用时，只需从解决实际某一业务场景出发，组合多种类型的计算节点来配置计算节点工作流程，从而快速解决业务工作流并依托数据开发的跨平台多异构任务的一键部署和上线能力，极大的提高了工程效率。

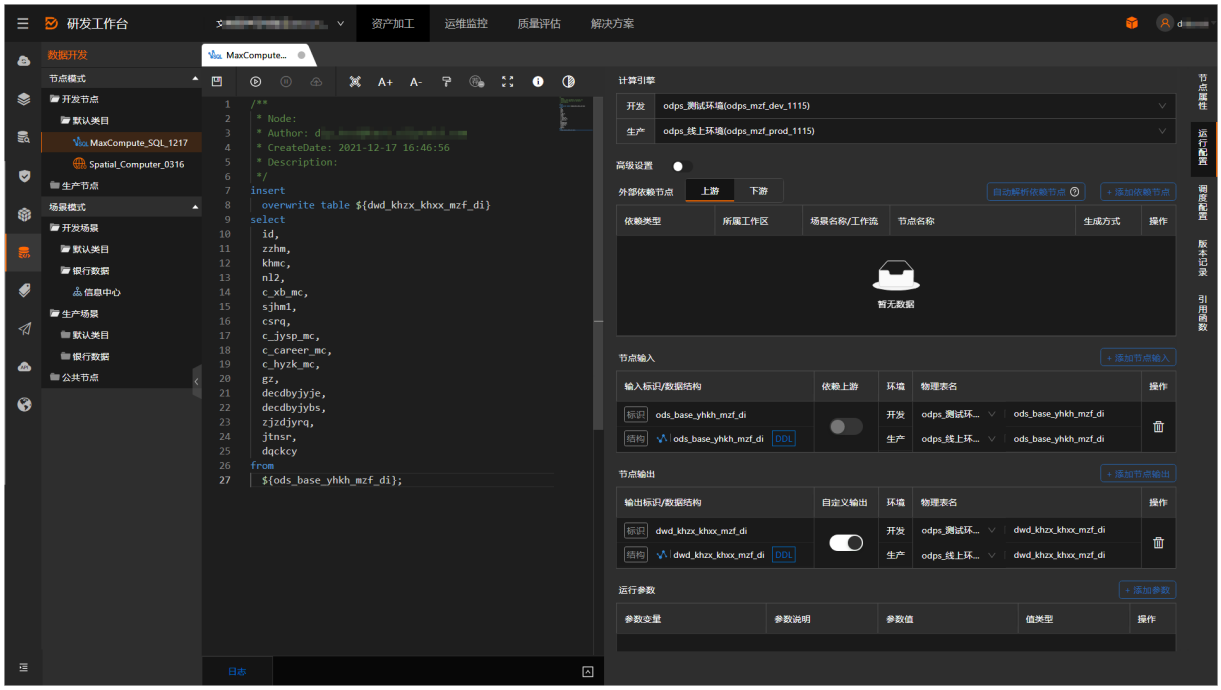
节点模式

节点模式任务主要适用于计算节点相对独立，整个流程由单个节点组成的这类简单数据处理场景，支持灵活的节点开发模式，支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理。支持批量离线计算等多种类型的计算节点。支持同一租户内，不同工作组中的任务依赖和统一调度。

开发节点

支持节点任务类目的新建、编辑和删除以及支持离线节点注册，对节点脚本、自定义函数、节点输入、节点输出、节点参数进行配置和管理以及支持配置节点的依赖关系以及支持节点任务的部署、运行和上线。

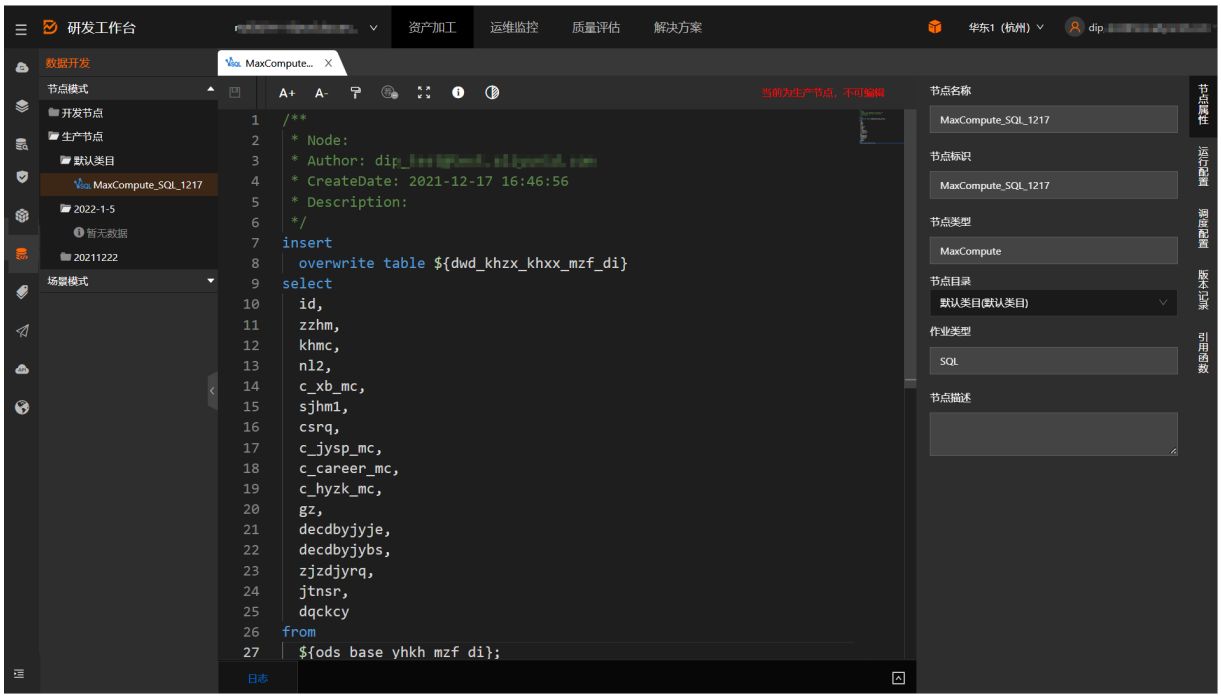
支持空间数据节点注册，对空间数据的数据管理、数据导入、数据导出、矢量分析、数据计算、数据筛选、坐标处理和栅格分析等多个算子的直接应用，并对图层输入、图层输出参数进行配置和管理。



- 支持节点任务类目的新建、编辑和删除。
- 支持节点任务的新建、编辑、删除。
- 支持离线节点注册，对节点脚本、自定义函数、节点输入、节点输出、节点参数进行配置和管理。
- 支持空间数据节点注册，对空间数据的数据管理、数据导入、数据导出、矢量分析、数据计算、数据筛选、坐标处理和栅格分析等多个算子的直接应用，并对图层输入、图层输出参数进行配置和管理。
- 支持根据SQL语言，自动解析离线节点的输入输出。
- 支持输入输出选择任意结构，方便操作。
- 支持配置节点的依赖关系。
- 支持节点任务的部署、运行和上线。

生产节点

支持生产环境中节点任务的查看和下线，对节点的算法信息、计算资源、存储资源、调度周期入、输出等详细信息以及部署信息的查看。

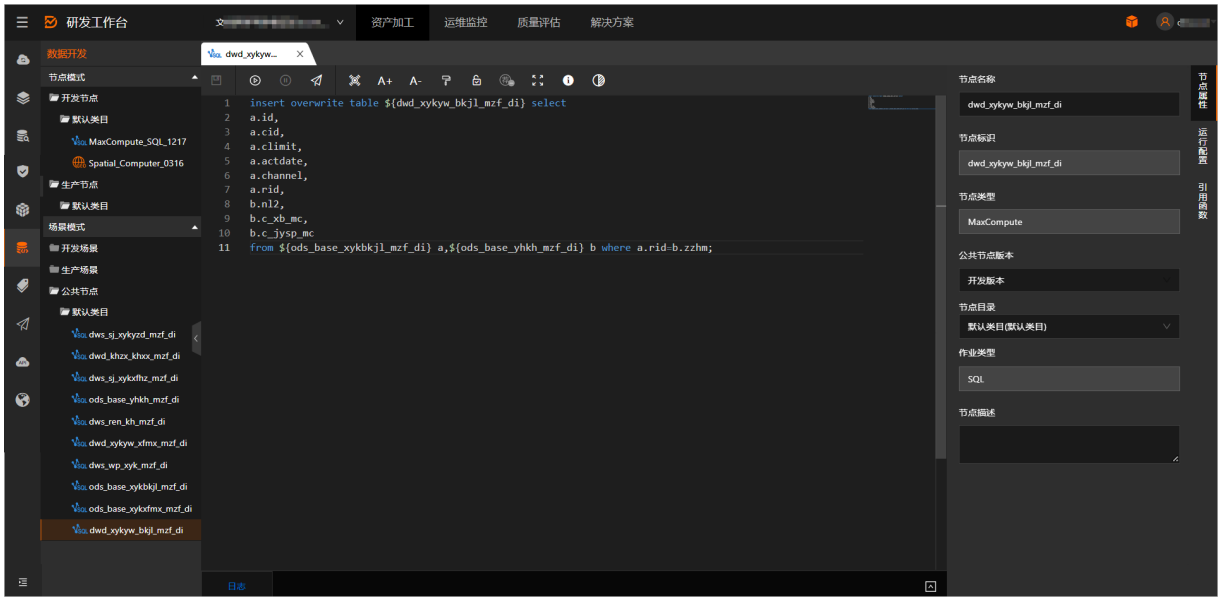


场景模式

场景模式适用于算法相对比较复杂，算法流程由多个节点，在一个场景内的混合编排组成的节点工作流场景，支持可视化的自定义任务流的场景开发模式，支持对计算节点中的脚本、自定义函数、节点输入、节点输出、参数等进行配置和管理。提供数据加工、算法服务任务流程开发、编排与调试、上线、部署、维护等功能。支持流式计算、批量离线计算等多种类型的计算节点在一个工作流中统一编排。支持同一租户内，不同工作组、不同场景间的全局任务依赖和统一调度。提供可视化操作界面，通过托拉拽的方式连接计算节点迅速实现数据加工流程编辑，大幅提升数据开发工作效率。

公共节点

支持公共节点类目管理。支持批量、流式等多种类型的公共节点，对公共计算节点脚本、公共计算节点输入、公共计算节点输出、公共计算节点参数等进行配置和管理。支持公共节点的版本管理。



- 支持公共节点注册，对公共节点脚本、自定义函数、公共节点输入、公共节点输出、公共节点参数进行配置和管理。
- 支持根据SQL语言，自动解析离线节点的输入输出。
- 支持输入输出选择任意结构，方便操作。
- 支持公共节点版本管理，用户在开发环境对公共节点进行修改和调优，调优完成后，发布公共节点并指定版本号，将算法发布到生产环境。
- 支持公共节点的废弃，避免旧版本公共节点被其他使用者重新使用。
- 支持查看公共/计算节点关联的场景，帮助用户清晰的了解公共节点的使用情况。
- 支持批量、流式等多种类型的公共节点。

开发场景

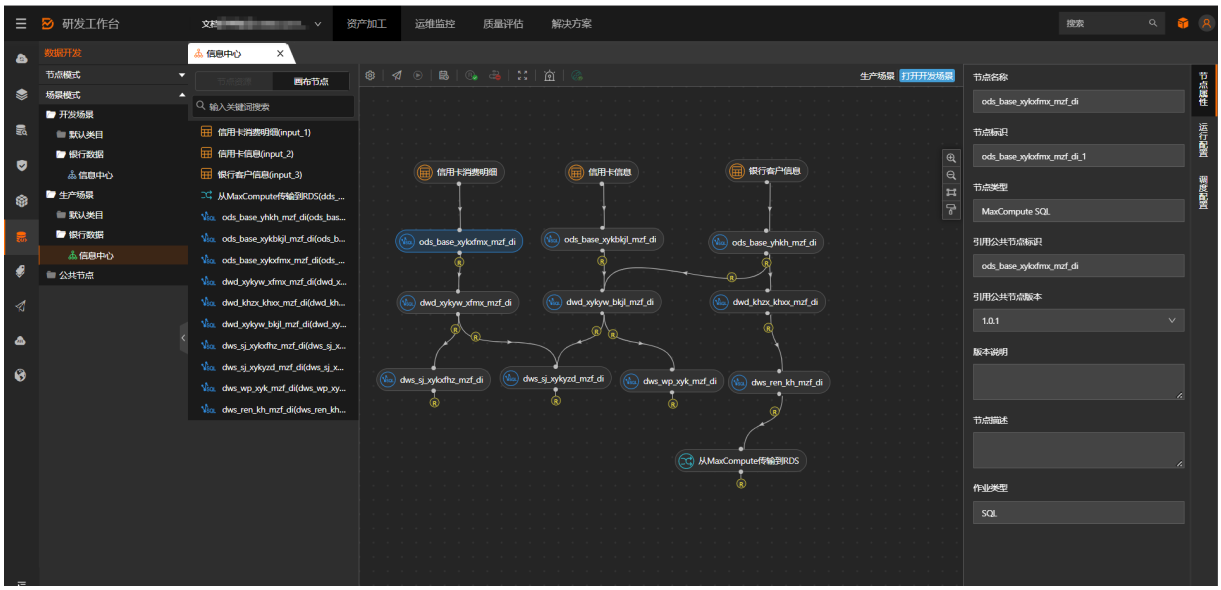
支持开发场景类目管理。支持开发场景引用公共节点。支持数据处理流程编排（多个计算节点组合成一个任务流），用户可灵活的通过可视化方式自定义任务流。支持批量、流式等多种类型的节点在一个场景内的混合编排。支持开发环境中场景的部署、运行和上线，并将开发环境中调试通过的场景发布到生产环境。



- 支持新建、编辑和删除开发场景目录。
- 支持新建、编辑、删除开发场景。
- 支持在开发场景中引用公共节点，并配置离线计算节点的调度信息。
- 支持在开发场景中新建计算节点，配置计算节点的各类参数及离线计算节点的调度信息。
- 支持在开发场景中新建系统节点，并配置系统节点的配置信息。
- 支持通过可视化方式配置上下游计算节点的依赖关系，并配置输入、输出参数的对应关系。
- 支持根据配置的输入信息，自动解析外部依赖节点信息。
- 支持输入和输出参数选择任意结构。
- 支持开发场景的部署、运行和上线。
- 支持自动生成数据质量检查节点（当计算节点的输出参数引用的逻辑表配置了质量规则时）。

生产场景

支持生产环境中场景的运行和下线，对各节点的算法信息、计算资源、存储资源、调度周期入、输出等详细信息的查看。



9.3.2.1.7. 智能标签

智能标签产品从大数据应用落地点出发，提供了将治理后的数据以业务化视角进行建模、查看、管理及使用，并提供业务衍生标签的自定义功能，为上层应用提供统一的标签数据目录和标签调用接口，沉淀上层应用制作的模型标签，实现高价值标签共享复用，形成标签运营生态。

智能标签在相应行业应用解决方案的结合下，能够让不是很熟悉大数据应用系统开发的程序员也能够快速搭建大数据应用，从而实现大数据价值的快速落地，智能标签帮助数据开发人员进行数据表的实体化，加速应用开发，沉淀业务模型的输入。比如在订单中，卖家能够理解的内容是成交笔数、成交金额、成交日期等实际现实对象，而非一张张表，如订单表。需要智能标签能面向零售行业，提供实体化的对象，如买家这个实体，这个实体上，需要赋予性别、年龄等标签信息，同时和商品建立关系，这个被赋予实际标签的实体，才是卖家能够理解的对象，基于这个对象才能去做分析。

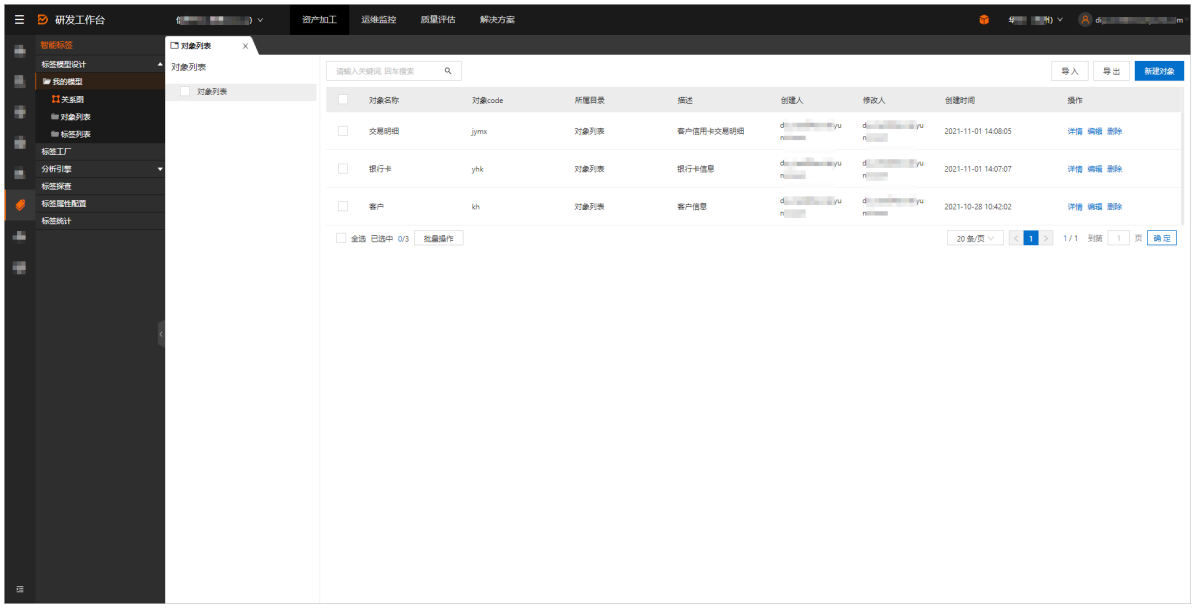
智能标签帮助构建这样的对象，提供给分析型应用实际的内容，同时让应用开发者也直观了解需要去获取的数据对象，直接使用实体化后的对象及标签进行加工、衍生、调用。

智能标签是以标签中心为基础，建立跨多个云计算资源之上的统一逻辑模型，开发者可以在“标签”这种逻辑模型，通过接口的方式进行快速的应用搭建。这种方式的好处：一在于屏蔽掉应用开发人员对于下层多个计算存储资源的深入理解与复杂的系统对接工作，二在于通过数据服务的形式透出也有助于IT部门对数据使用的管理，避免资源的重复和冗余。

标签模型设计

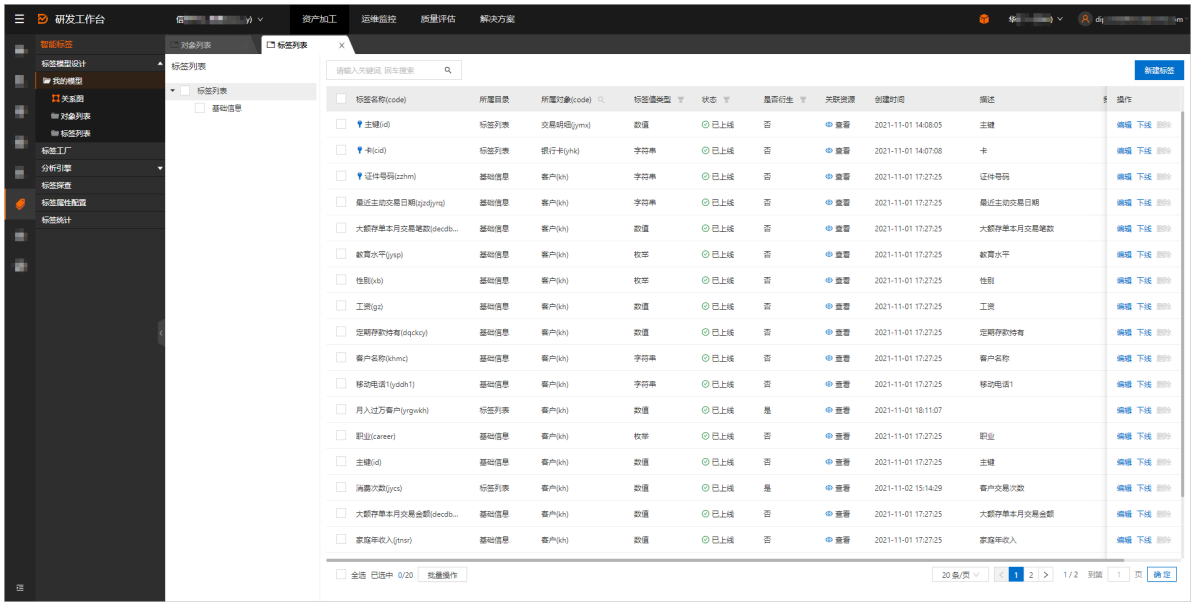
- 标签模型基于OLT（对象、关系、标签）模型来抽象物理数据，构建语义层，为业务人员“看懂数据”从而“使用数据”提供基础，支持对象关系的多级类目管理。
  - 对象定义：包括实体类（无时间维）和事件类（有时间维），描述一个客观存在实体或事件；从数据角度看，是一个多种数据描述同一实体或事件的集合。
  - 关系定义：关系建立在及两个对象上的关系。如用户和购买商品事件的关系，用户和学习课程事件的关系等。
  - 标签定义：标签是挂接在对象，对对象进行描述的属性的集合。
- 对象管理

- 新建对象及标签：支持跨存储类型的数据源整合，低门槛构建“宽表”，形成业务可以理解的对象及其原生标签信息。
- 对象导出/导入：支持将对象导出/导入，包含对象相关的标签、计划、标签API，便于在其他系统或工作组中快速部署。



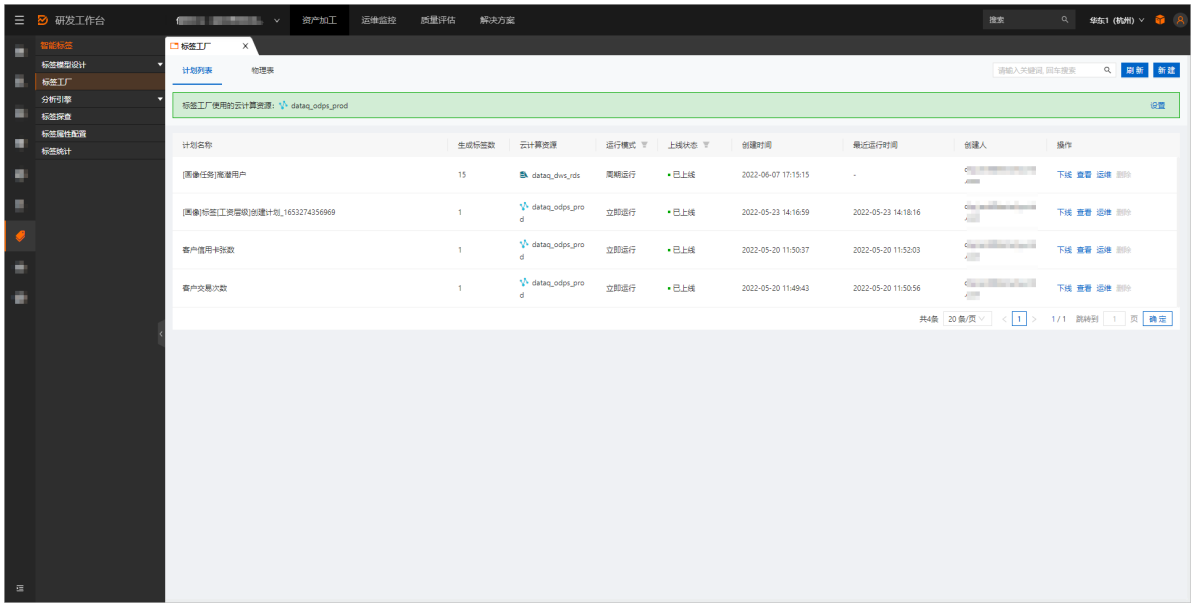
- 标签管理

  支持对标签进行编辑、上线/下线、公开/撤回；实现标签的生命周期管理。



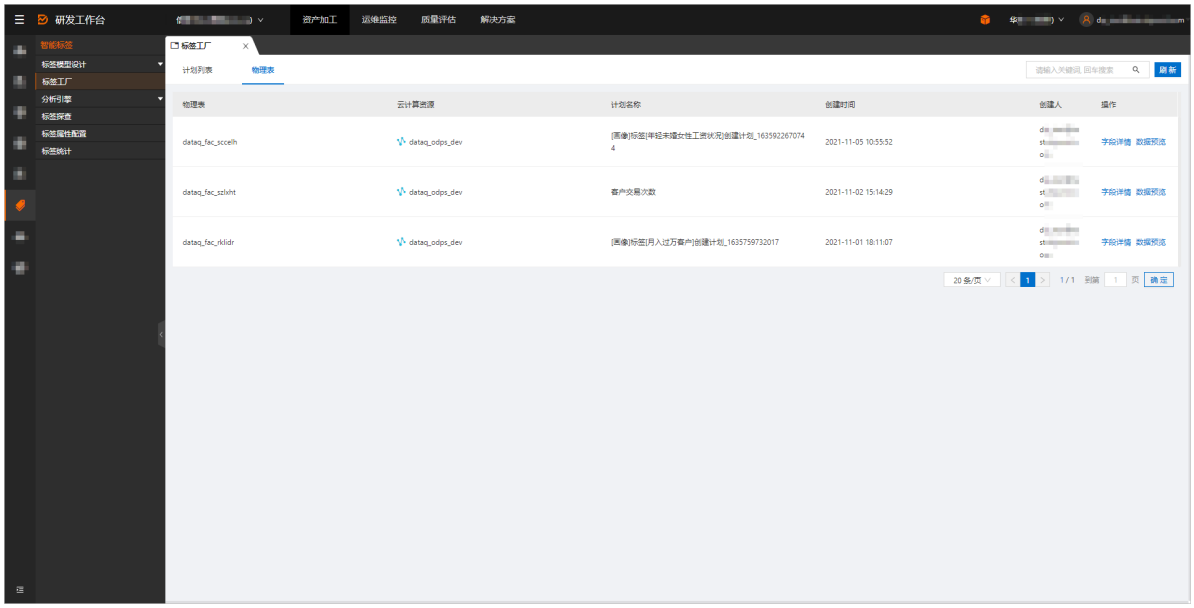
标签工厂

- 计划列表管理
  - 支持已有标签加工生成衍生标签功能。
  - 支持文本TQL编写标签方案，支持衍生标签配置关联到实体或关系上，支持标签方案生成任务一次性调度和周期调度功能。
  - 支持标签方案生成任务的运维管理功能，并提供对计划任务的上线、下线、查看、运维、删除等功能。



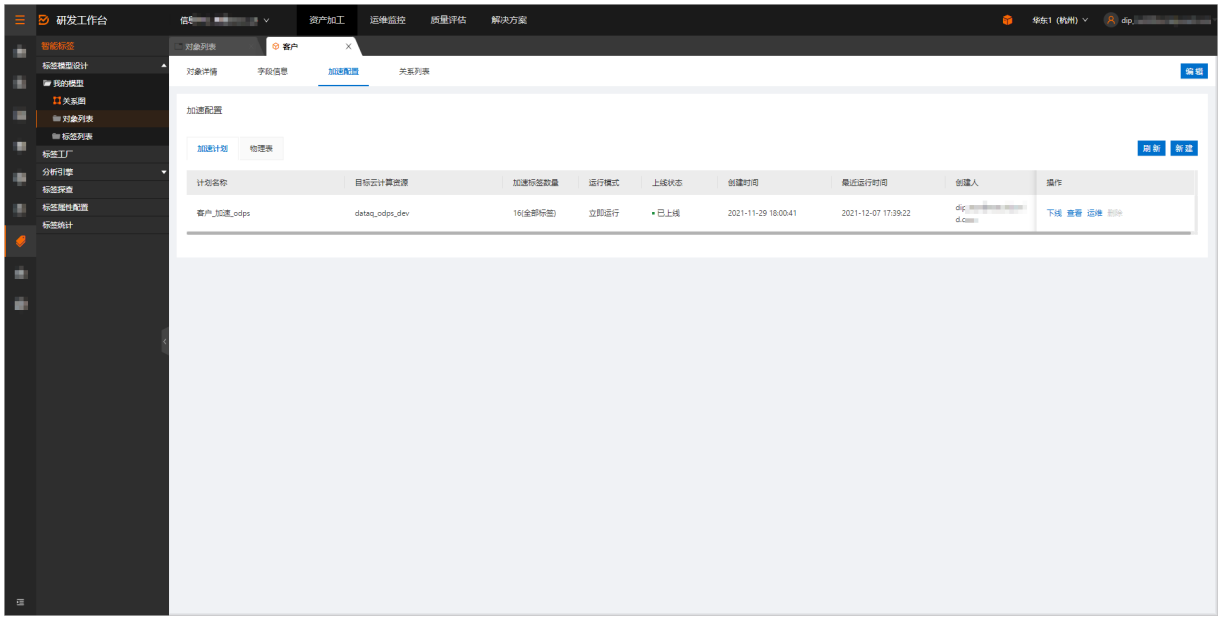
- 物理表查看

  支持查看云计算资源下的物理表信息，并提供刷新、字段详情查看、数据预览等功能。



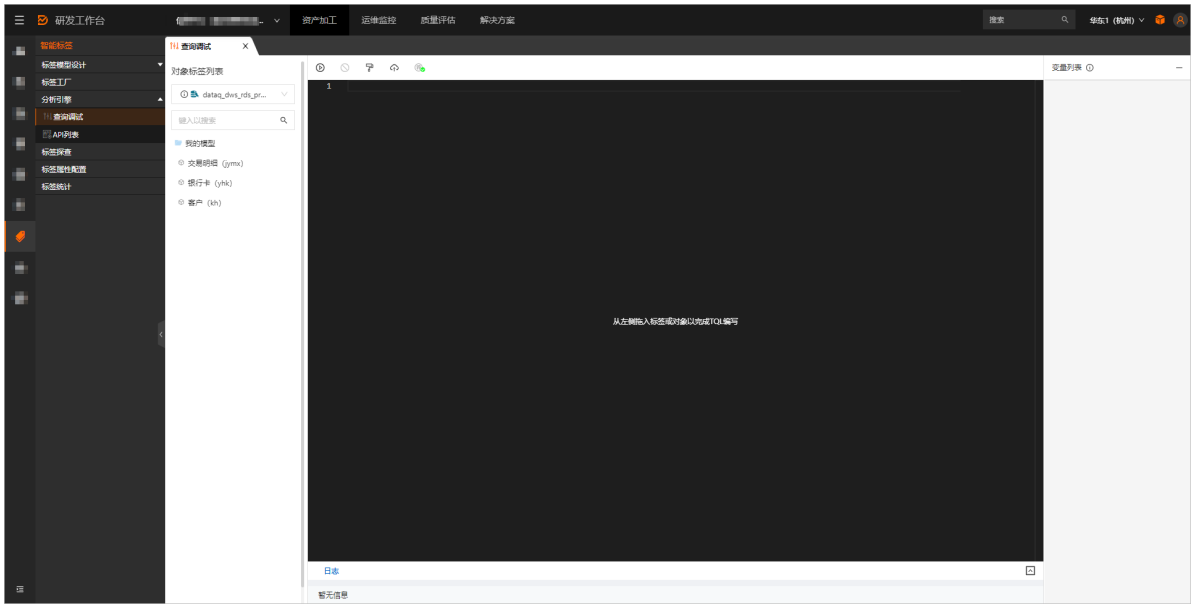
标签加速

支持将构建的标签模型统一加速至分析型数据库形成大宽表，以提供后续分析场景使用。



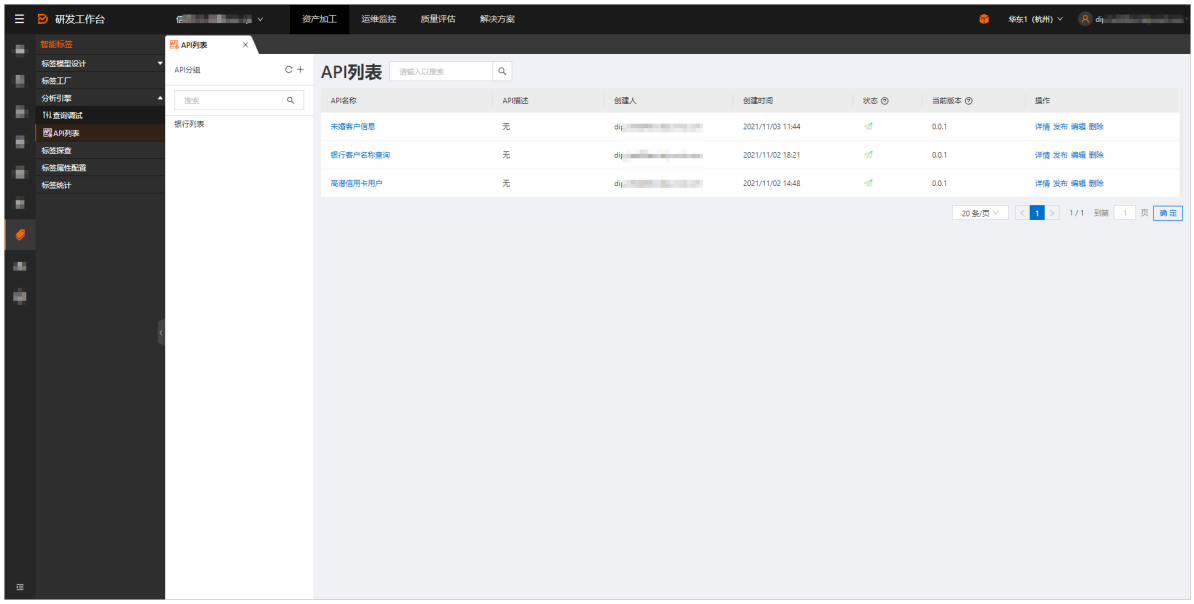
分析引擎

- 查询调试
  - 支持实体关系上的标签和获得授权标签的查询功能。
  - 支持实体关系使用TQL（Tag Query Language）语言的关联查询。
  - 支持查询结果的列表查看和JSON格式查看。
  - 支持可视化的页面调试，展示TQL的每个执行流程；每一流程的时间性能消耗。
  - 支持多种云计算资源中的标签查询。
  - 支持分析查询结果生成API，并提供在分析查询中配置参数。



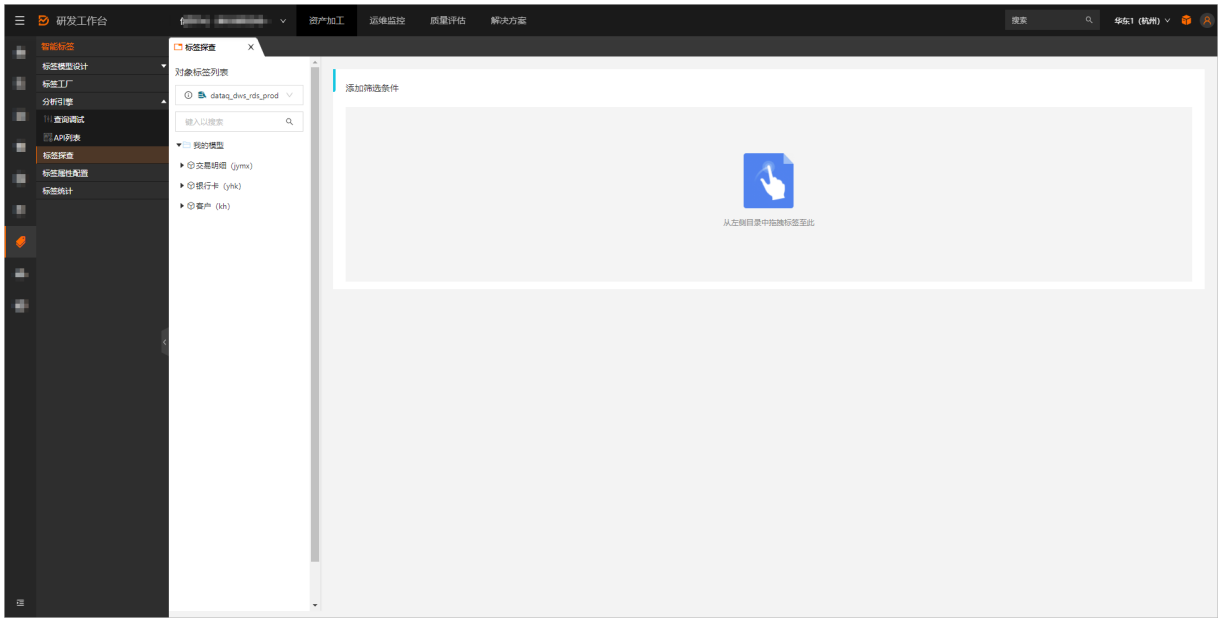
- API列表

支持API的类目管理、列表管理、详情查看、调试和版本管理等功能。



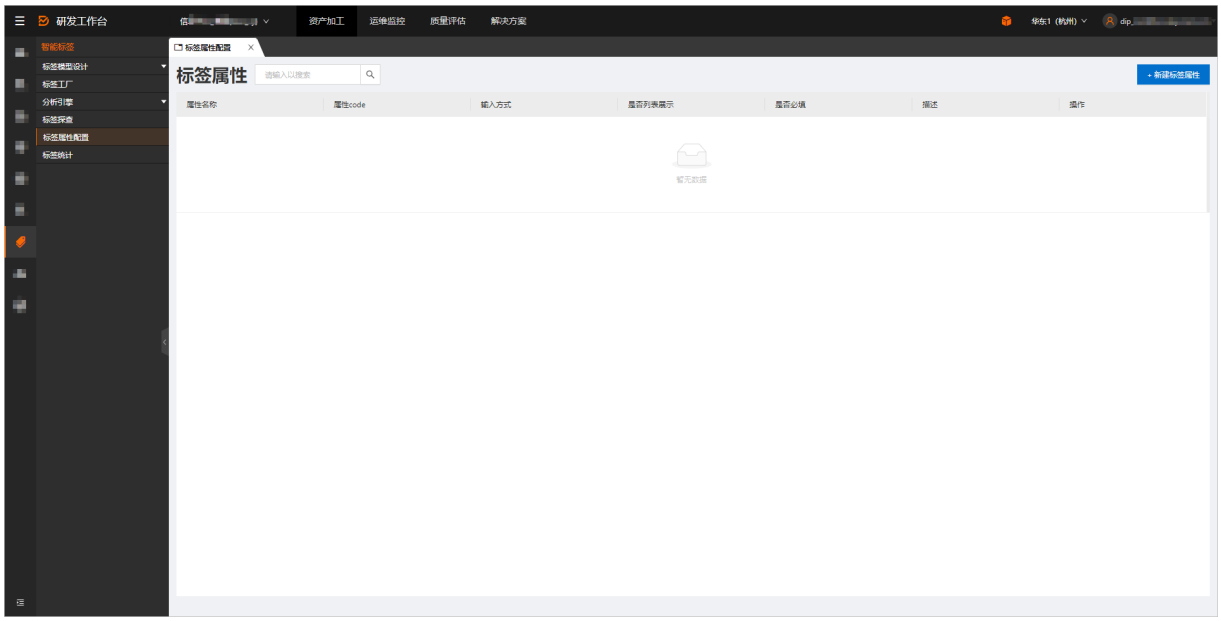
标签探查

支持查询所有实体的标签及标签内容数据，支持多标签的内容筛选能力。



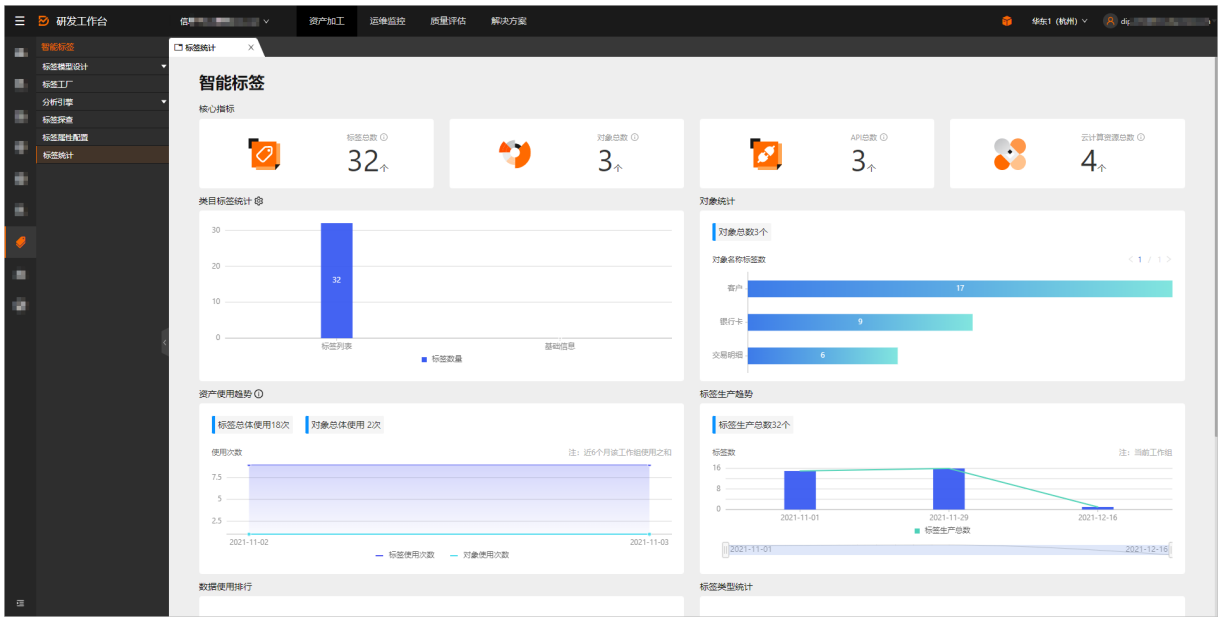
标签属性配置

支持为标签提供易于识别的属性和属性值，便于画像分析时能够提供方便快捷的进行数据筛选，提高了画像分析的效率。



标签使用统计

支持标签生产和使用情况统计，包含标签类目统计、标签使用、对象使用统计等。



9.3.2.1.8. 统一服务

在数据中台建设过程中，搭建核心处理平台，并根据相关标准要求，将源数据或治理后的数据，按照“全面、高效、统筹”的共享原则，“低延时”、“精准化”的传递给用户，方便更多应用消费数据，实现数据价值。协同管控各种组织机构，为智慧化运行提供支撑，实现跨层级、跨部门、跨行业、跨区域信息即时共享，做到高可靠、低延时。同时，数据被应用的使用情况评估也作为数据中台日常运营的重要指标项。提供更多数据生产需求，完成数据从生产、使用、到评估的闭环。

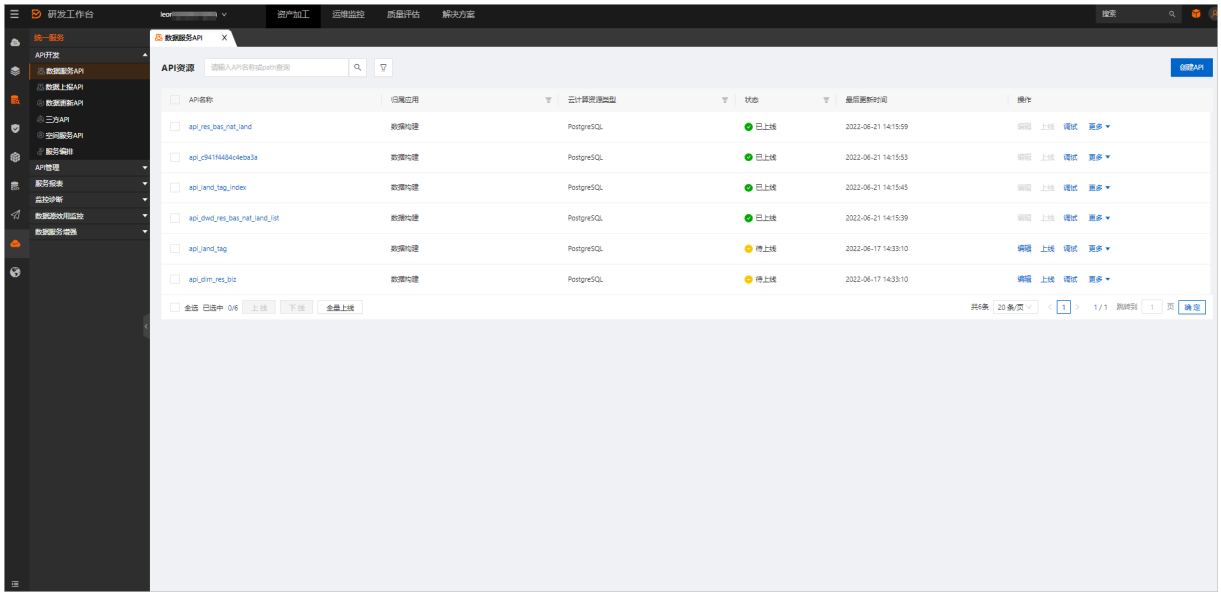
统一服务提供了如下数据服务能力。

- 数据服务API的源头为数据表，可以是单表也可以是多表，请求后最终以SQL的形式到数据库查询，并返回符合条件的数据，将数据组装成报文格式返回。目前数据服务API提供单数据服务API，SQL转API。数据源可以是OTS、ApsaraDB for RDS和AnalyticDB MySQL等数据存储中的数据表。
- 流式数据服务的源头为DataHub或RocketMQ的Topic，服务封装为DataHub Topic的读取API，返回Topic中的数据。数据源可以是DataHub和RocketMQ。
- 批量离线数据服务的源头为MaxCompute的数据表，服务封装为根据分区获取分区内全部数据的API。
- 非结构化数据服务的源头为OSS，服务封装为根据文件名称返回文件数据流。

统一服务为用户提供了数据中台建设过程中的数据服务化能力，面向数据开发者提供覆盖各个加工阶段统一体验的、便捷的数据查询转服务、服务管理、服务运维能力；面向数据资产管理者提供服务的统计分析、服务用量统计分析、热门数据统计分析能力，实现数据中台建设后半场“数据应用”的有效落地，支撑数据智能应用的高效开发。

API开发

支持数据服务API、三方API、数据上报API、数据更新API、服务编排API和空间服务API。支持创建、编辑和调试API，支持的API类型包括实时查询类、离线批量类、流式数据类、非结构化数据类。



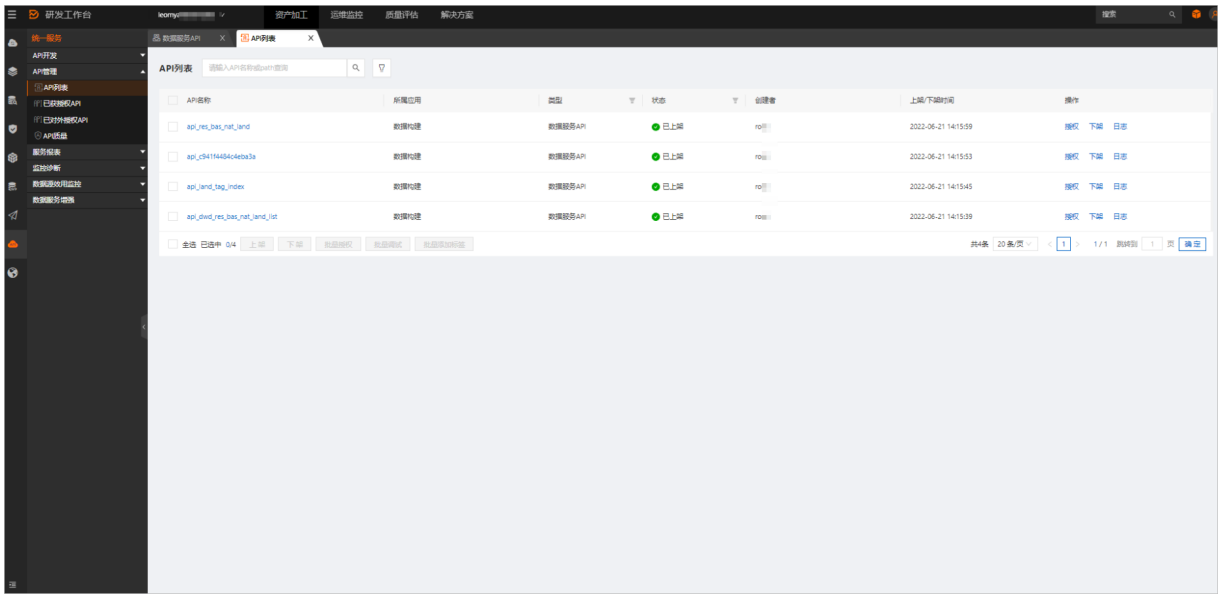
- 数据服务API：将数据库中的表的通过向导模式和SQL模式配置生成API，方便开发者申请使用或调用。目前支持的云计算资源有ApsaraDB for RDS、AnalyticDB MySQL、MaxCompute、DataHub、Table Store、Object Storage Service、AnalyticDB PostgreSQL、RocketMQ、Hologres、Presto、PolarDB等。
- 三方API：将已有API注册到统一服务，通过注册API的相关信息，包括提供商URL、ContentType、请求方式、请求参数、返回参数、返回数据样例等信息将三方API注册到统一服务管理。
- 数据上报API：将数据写入API，用户调用API时将数据上报到服务端，目前支持ApsaraDB for RDS、DataHub和AnalyticDB PostgreSQL作为数据上报目标端。
- 数据更新API：指通过调试或调用创建的数据更新API，修改数据库中的表数据，目前支持的云计算资源有ApsaraDB for RDS和AnalyticDB PostgreSQL。
- 服务编排API：按照业务逻辑，灵活的实现跨计算资源的多个API服务编排，提供拖拽式、可视化的API编排能力。支持按照业务逻辑，以串行、并行和分支等结构编排多个API为工作流。目前支持的服务节点有数据服务API和三方API。
- 空间服务API：将空间数据管理中的各类型空间数据发布成行业标准的空间服务API，支持各类型空间数据的汇聚接入、地图化浏览、空间计算、服务代理。

- API的生命周期管理：API支持数据服务API、三方API、数据上报API、空间服务API和数据更新API。支持API的创建、编辑和调试，支持的API类型包括实时查询类、离线批量类、流式数据类、非结构化数据类。
- API权限管理：包含已获授权API和已对外授权API。

API管理

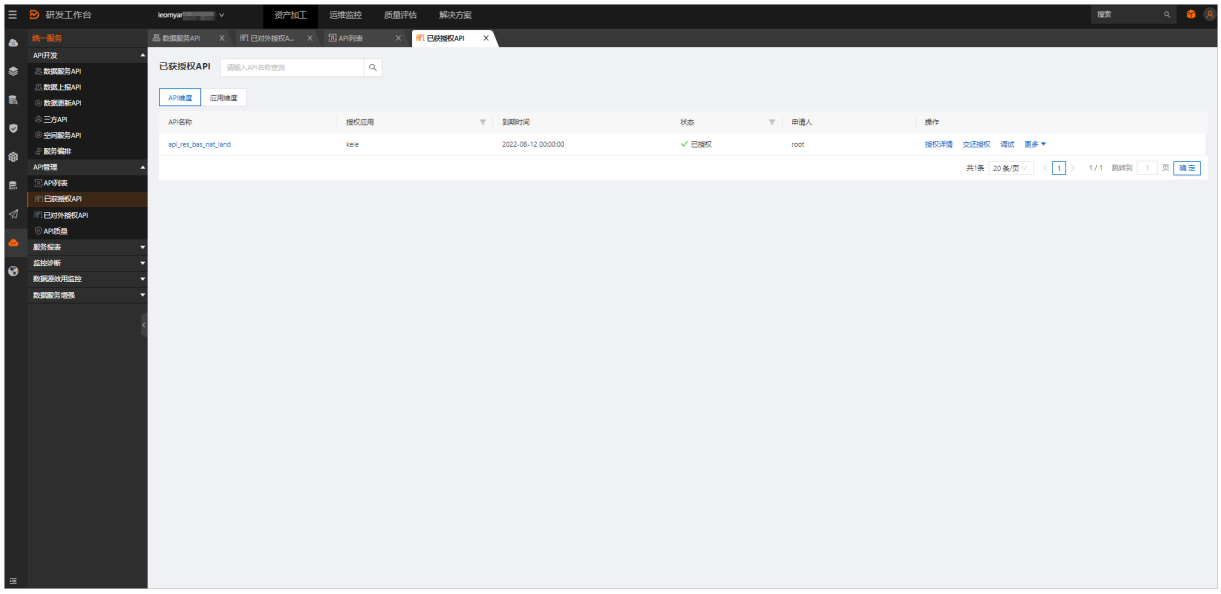
对API做共享管理，支持对已上线的数据API做上架、下架、授权等操作，同时触发审批流程，审批通过后API会在数据资产中心API栏目上架，其他工作组及用户可以申请API使用。API下架时会触发冻结操作，下架后自动收回已对外授权的API的使用权限。

API列表



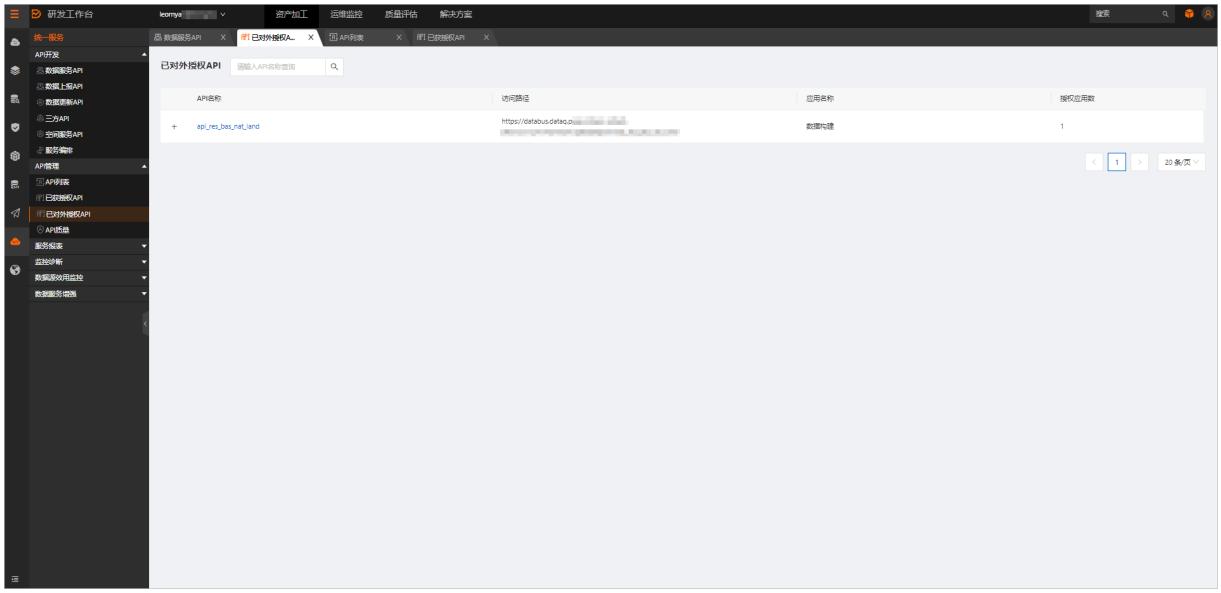
已获授权API

支持对已获授权API列表进行查看，调试API、下载API使用文档、下载密钥SDK和归还权限。支持查看单个API的授权详情，包括：授权时间、访问路径、AppKey和AppSecret。



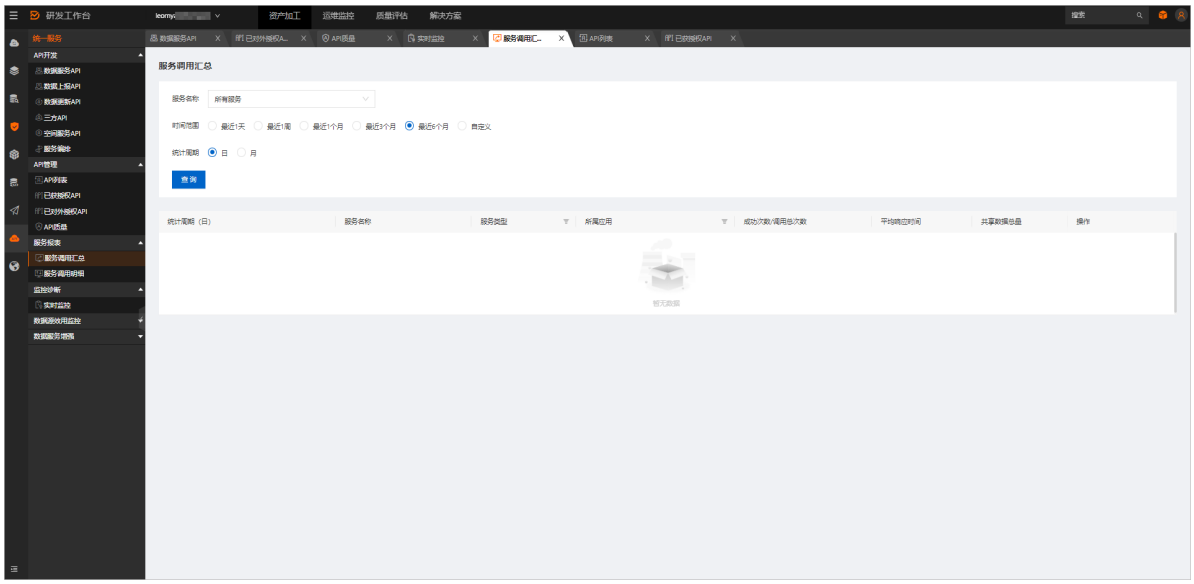
已对外授权API

支持已对外授权API列表的查看，支持查看授权的AK信息、暂停授权、回收授权等操作。

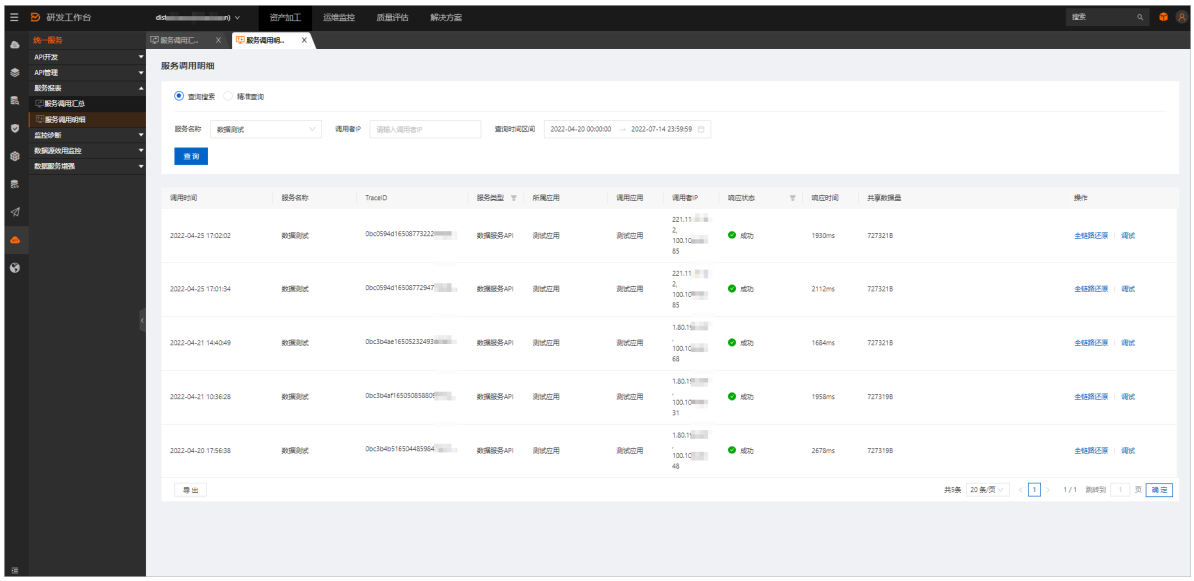


服务报表

- 服务调用汇总支持用户在最近1天、最近1周、最近3个月、最近6个月和自定义等时间范围下，按照15分钟、小时、日、月和天等统计周期统计调用汇总信息，包括总调用次数、平均响应时间、成功调用次数。



- 服务调用明细支持单次调用的明细信息查看，包括调用时间、服务名称、服务类型、响应状态、响应时间等。



实时监控

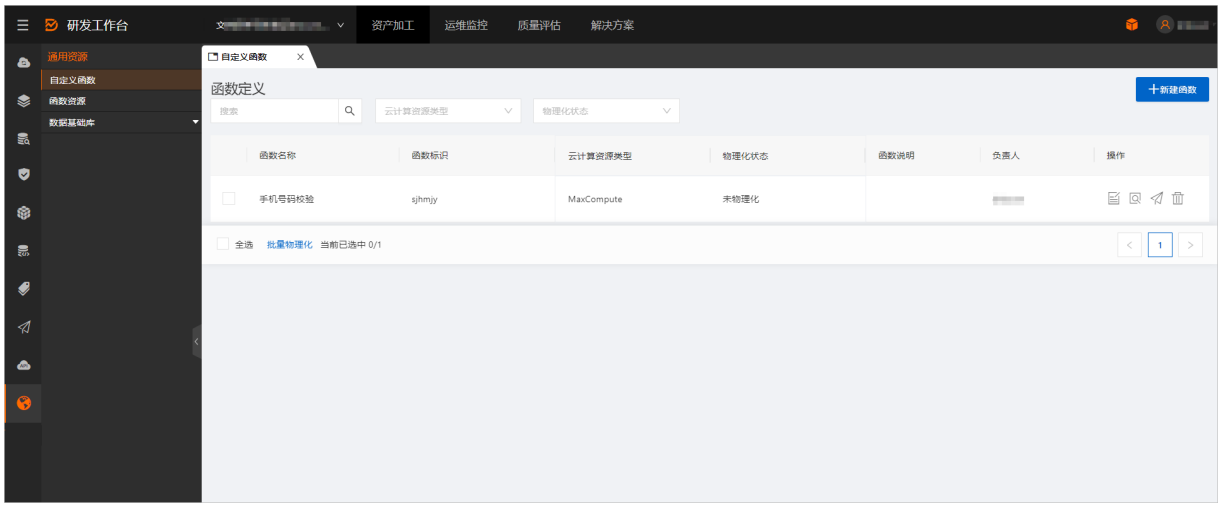
监控最近3小时、12小时和24小时内服务的实时调用趋势及异常趋势。支持查看监控时间点上的服务概览。

9.3.2.1.9. 通用资源

通用资源为用户提供了自定义函数、函数资源、数据基础库的能力，帮助用户在数据开发等过程中灵活调用各类自定义函数和基础数据。

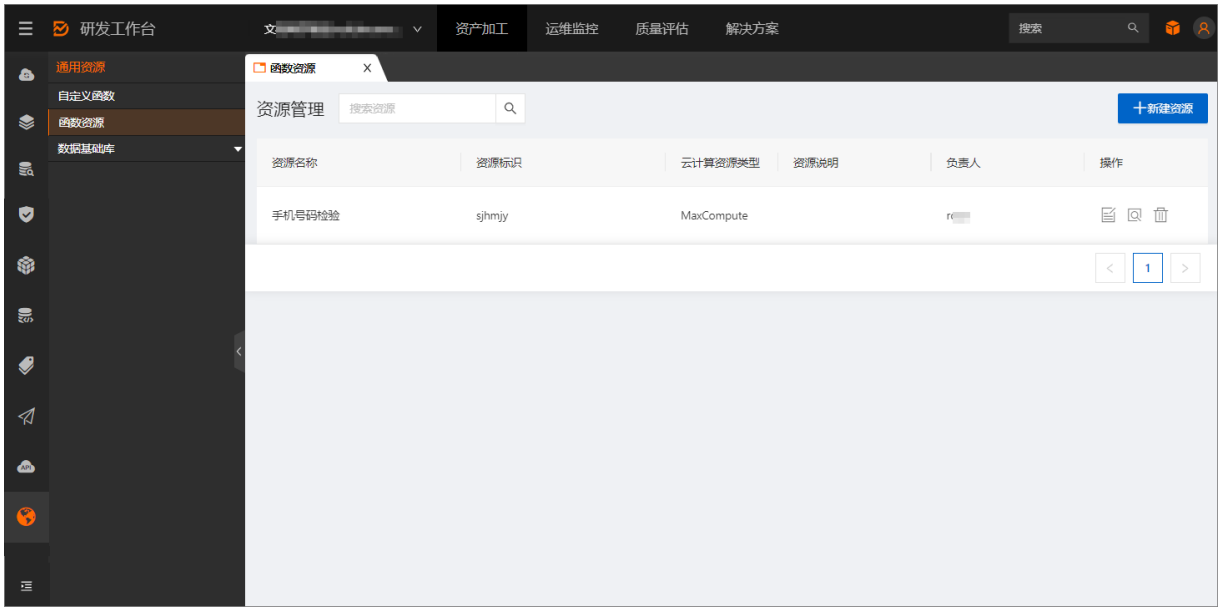
自定义函数

支持对ETL函数自定义和管理，并与公共计算节点相关联，便于快速对数据进行ETL处理。



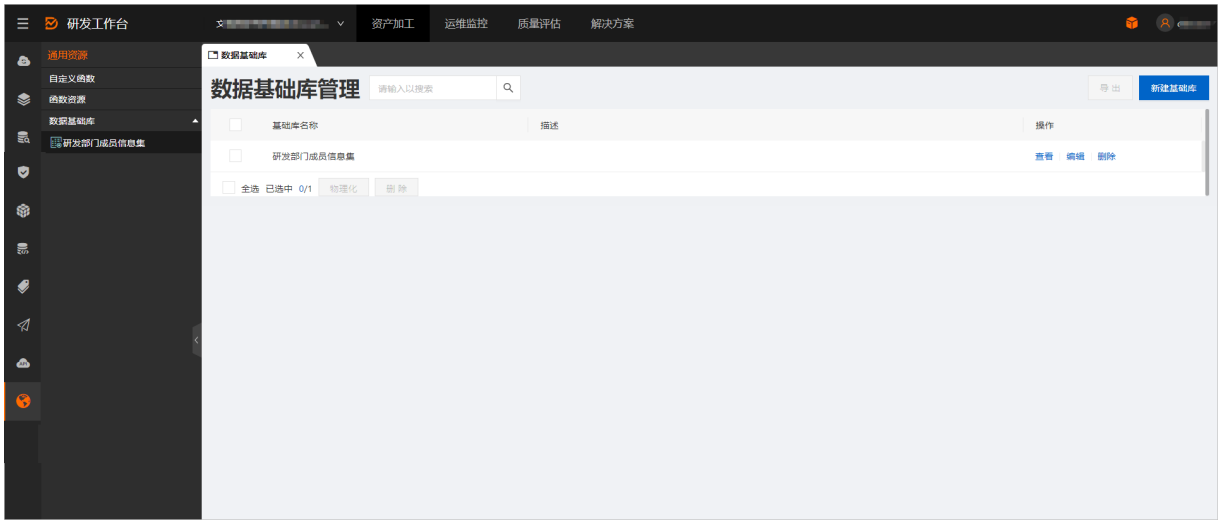
函数资源

支持对自定义函数中所使用的代码包的新建和管理。



数据基础库

支持标准化基础数据的管理维护，以便在模型加工过程中使用行业通用化标准化数据，进行后续数据加工和内容参照。



9.3.2.2. 运维监控

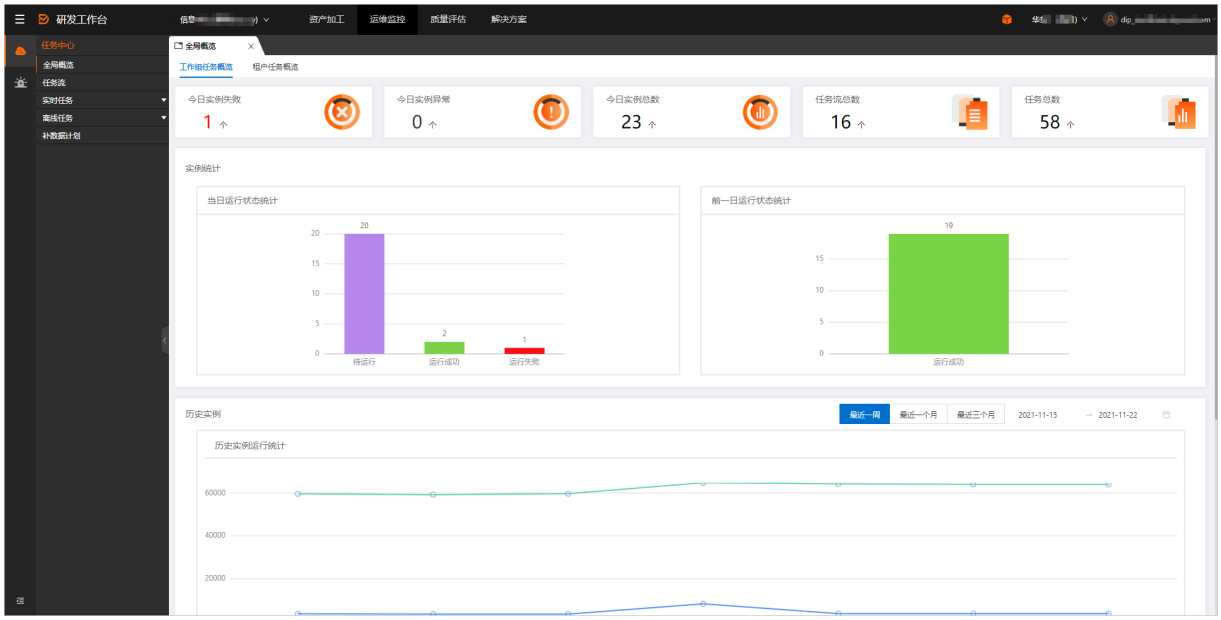
9.3.2.2.1. 任务中心

任务中心对数据资源平台中的数据同步、数据开发、标签加工等任务和实例进行运维管理。

数据资源平台各模块发起的任务，如数据同步、数据开发、标签加工等，都会发到对应的任务流，不同的任务来源通过业务类型区分。根据业务流程，任务流中包含了不同类型的节点，如MaxCompute离线任务、Flink流式任务等。不同类型的节点会生成各类任务，主要包括离线同步任务、离线计算任务、实时同步任务、实时计算任务等。对于离线同步任务、离线计算任务，每次运行时会生成该次运行周期的实例。

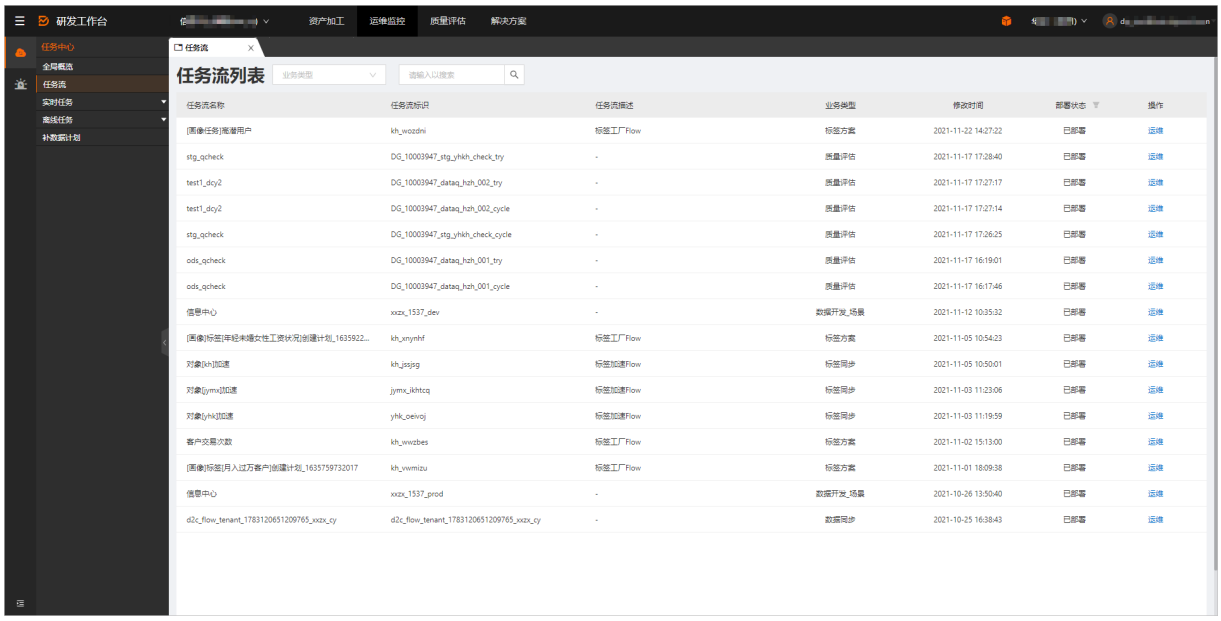
全局概览

支持查看任务流、任务、任务实例的统计信息，并提供以图形化的方式展示历史实例运行信息。



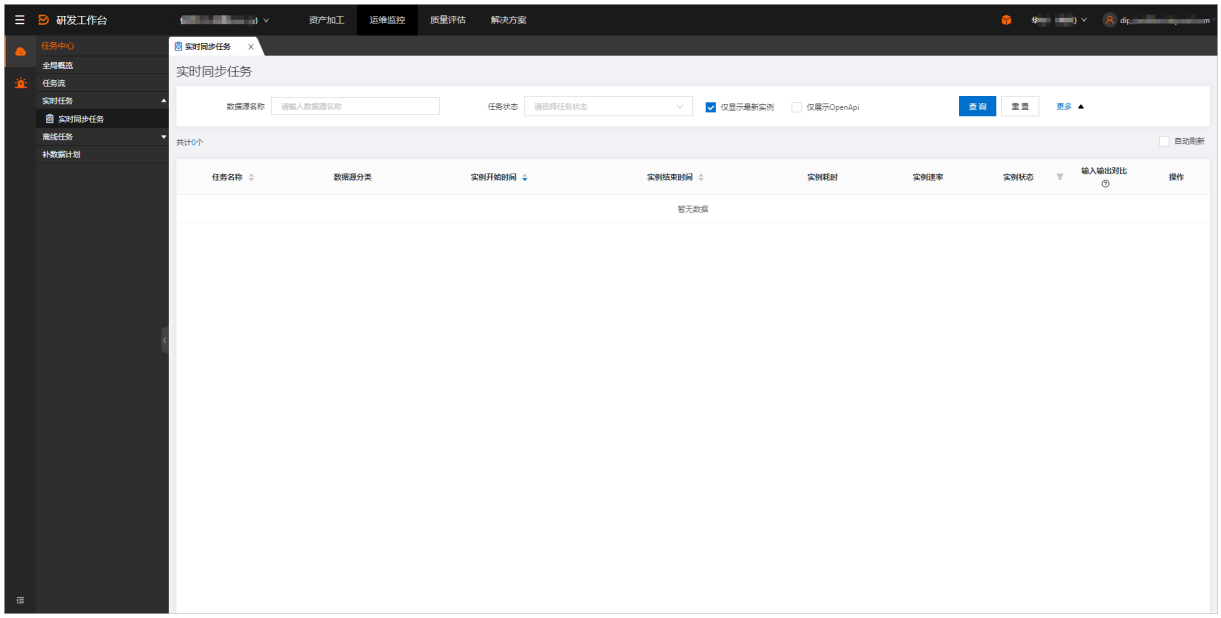
任务流

支持查看任务流的属性信息，包括任务流中包含的任务节点、任务节点信息、任务节点的最新实例等。支持查看任务流的部署、运行日志信息。



实时任务

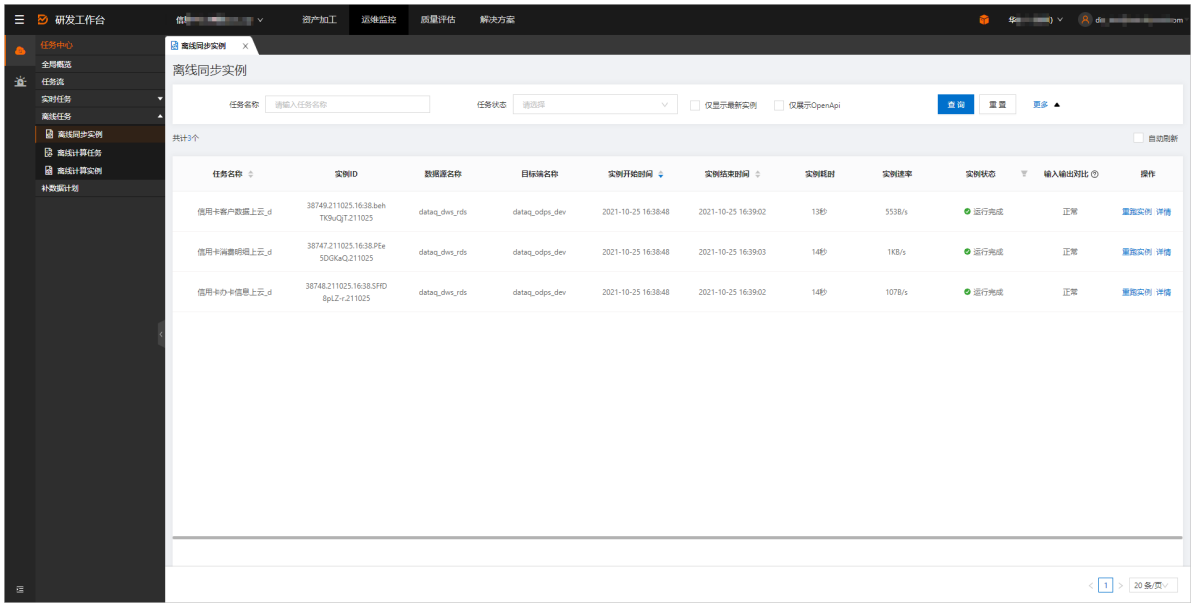
支持以列表的方式展示实时任务信息，并提供当前实例属性与运行日志信息。



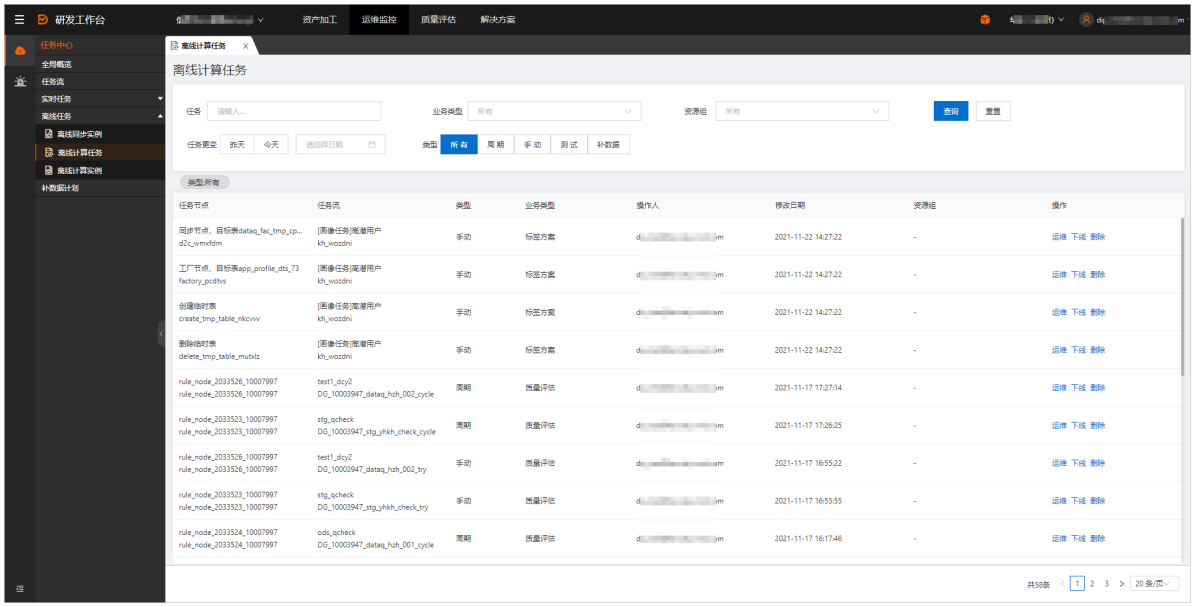
离线任务

支持查看离线计算任务的属性信息，查看离线同步实例、离线计算实例的属性与运行日志。

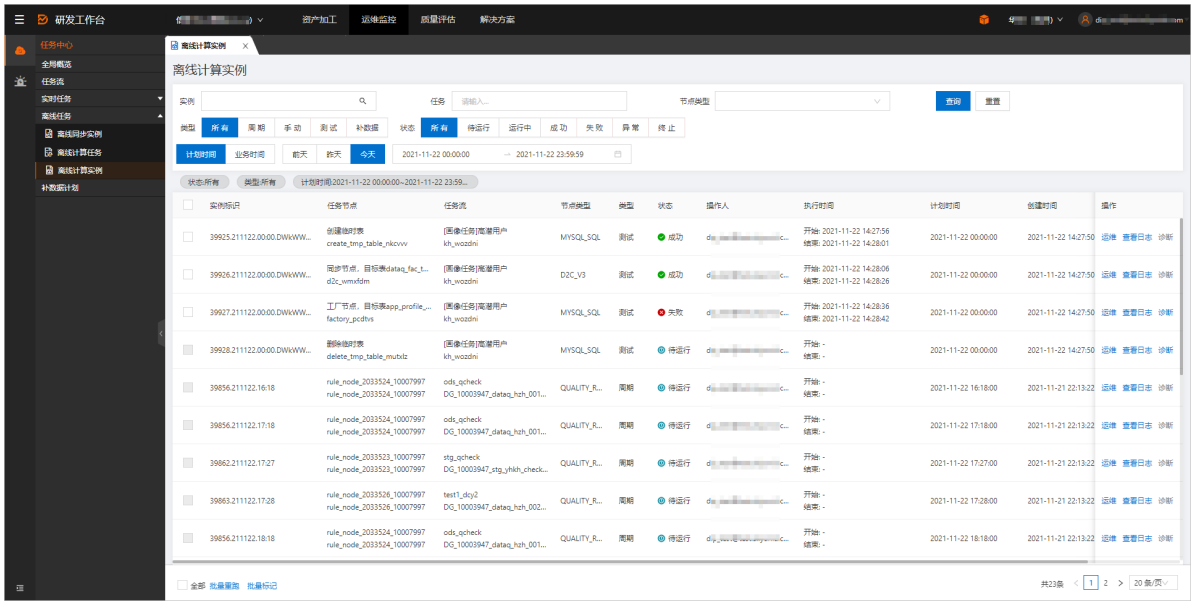
- **离线同步实例**：支持以列表的方式展示离线同步实例信息，并提供离线同步实例属性与运行日志查看；可进行实例重跑的功能。



- **离线计算任务**：支持以列表的方式展示离线计算任务信息，并提供运维、下线、删除的功能。



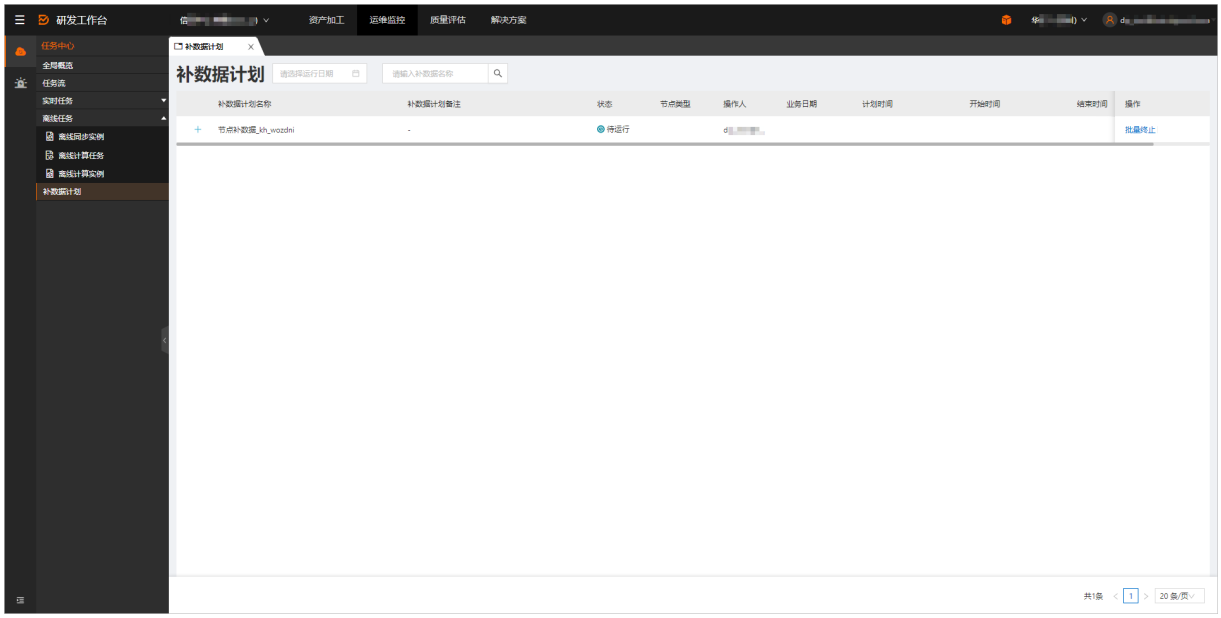
- **离线计算实例**：支持离线计算实例信息，并提供运维与日志的查看。



补数据计划

支持任务节点的补数据操作。

- **新建补数据计划**：支持在任务流和离线计算任务中补数据。
- **终止补数据计划**：支持补数据计划运行后进行终止操作。
- **查看补数据计划**：支持查看补数据的运行状态、节点类型、业务日期，以及查看对应节点的运维信息，包括实例信息、任务代码、实例日志。



9.3.2.2.2. 全景监控

全景监控支持业务链路配置监控场景，通过自定义监控链路和告警消息配置，可快速实现从数据生产到业务应用的全链路编排及溯源，实现全局监控和问题快速锁定。

全景监控提供丰富的告警信息输出方式，包括：钉钉消息、标准协议等方式，出现问题时按照问题的紧急程度，选择告警方式，快速通知相关人员，及时处理问题，保障系统和数据的稳定运行。同时，全景监控提供统一的监控接入标准，全方位对接各种应用系统监控信息，实现覆盖所有系统、全流程的监控管理。

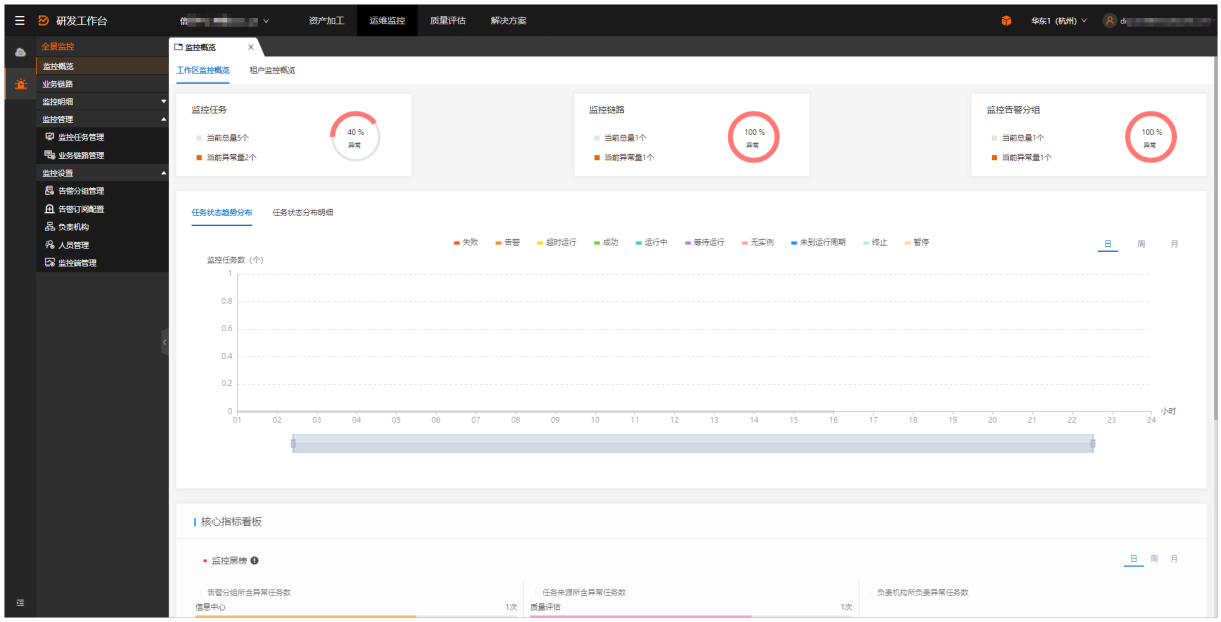
数据全链路保障是指通过采集数据汇聚、数据融合、数据分析挖掘等阶段的各项任务状态信息，对异常状态进行预警和处置，实现对各任务的实时监控和管理。

从应用视角提供平台内跨系统的从数据生产到业务应用的全链路监控能力，其提供的监控端可将多系统的状态信息进行上报汇总和集中推送。提供自定义监控链路配置能力，快速实现从数据生产到数据应用的全链路编排及溯源和监控节点的稳定性分析。例如，当指标出问题时，加工这个指标的全局链路上，涉及系统的接入、数据的加工、API的转换、系统中的算法加工，到最终指标的呈现，到底哪个链路出问题需要立即定位。系统需向各方应用提供统一接入标准，实现全局监控和问题快速锁定。

通过业务场景组织监控链路，包含业务场景名称、业务场景描述、负责人，以及业务场景的正常/异常状态及异常告警个数等统计指标，方便运维人员从全局快速锁定出现问题的业务场景及严重程度。业务场景中包含该场景涉及的业务节点、任务节点，以及这些节点在逻辑上的上下游关系，运维人员可通过上下游关系清晰的了解各任务间的影响和依赖关系。

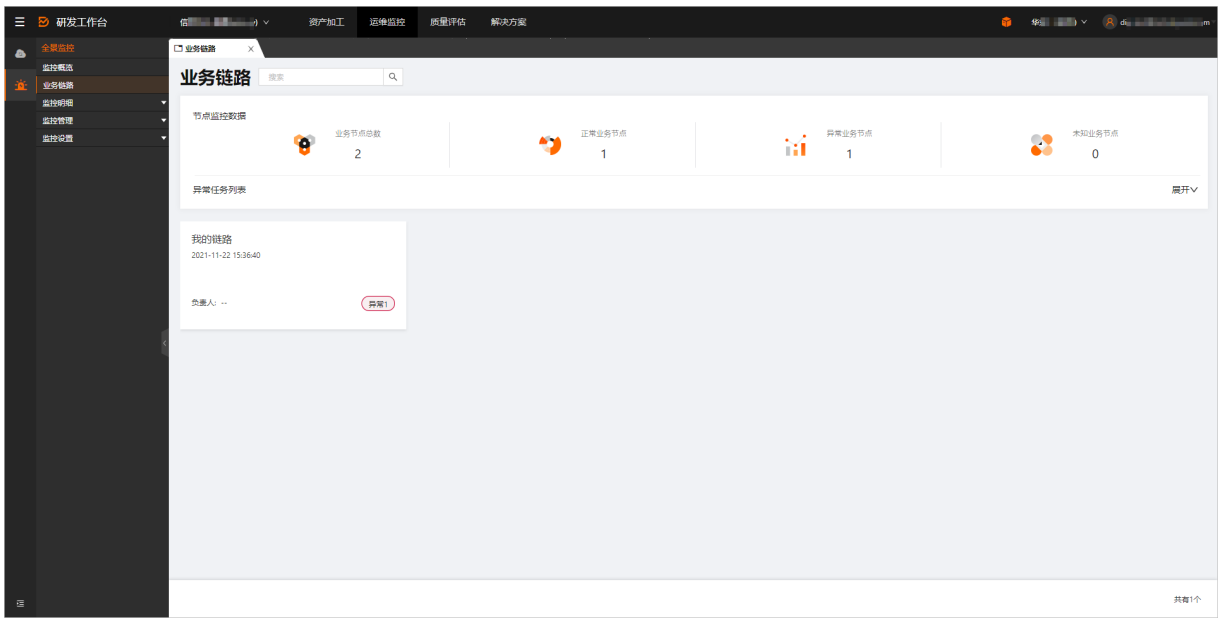
监控概览

提供监控任务、监控链路、监控告警分组等统计信息概览。



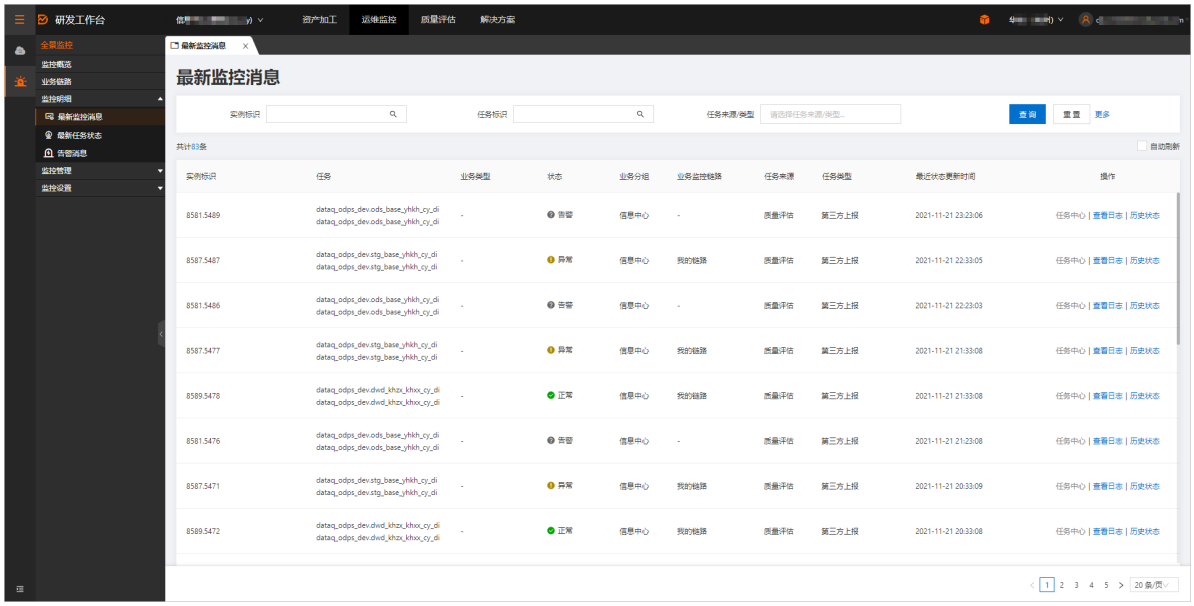
业务链路

支持以业务数据链路为视角进行监控，可查看当前业务链路中节点的运行情况、告警信息。

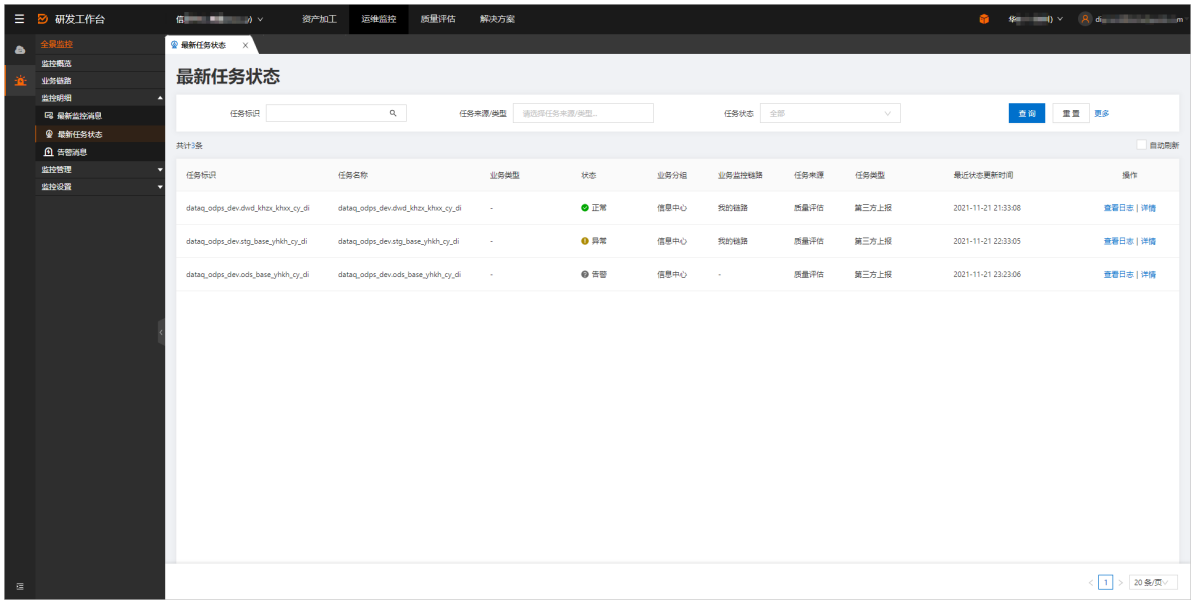


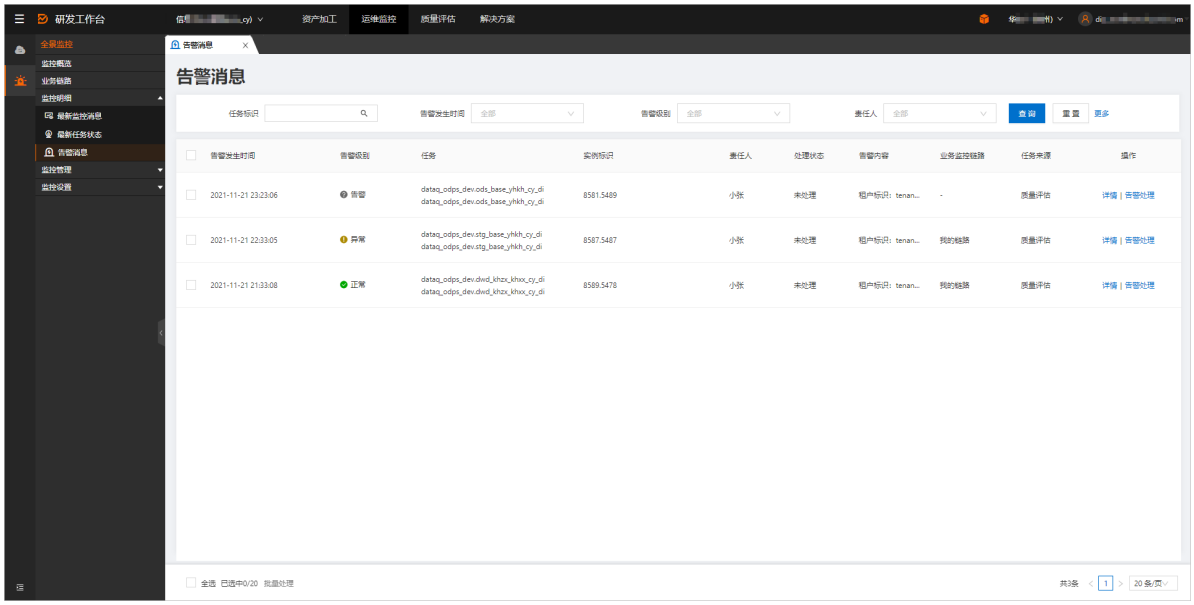
监控明细

- 支持查看任务节点监控结果，可细节掌握每个监控节点的最新状态、任务日志和历史状态分布情况。



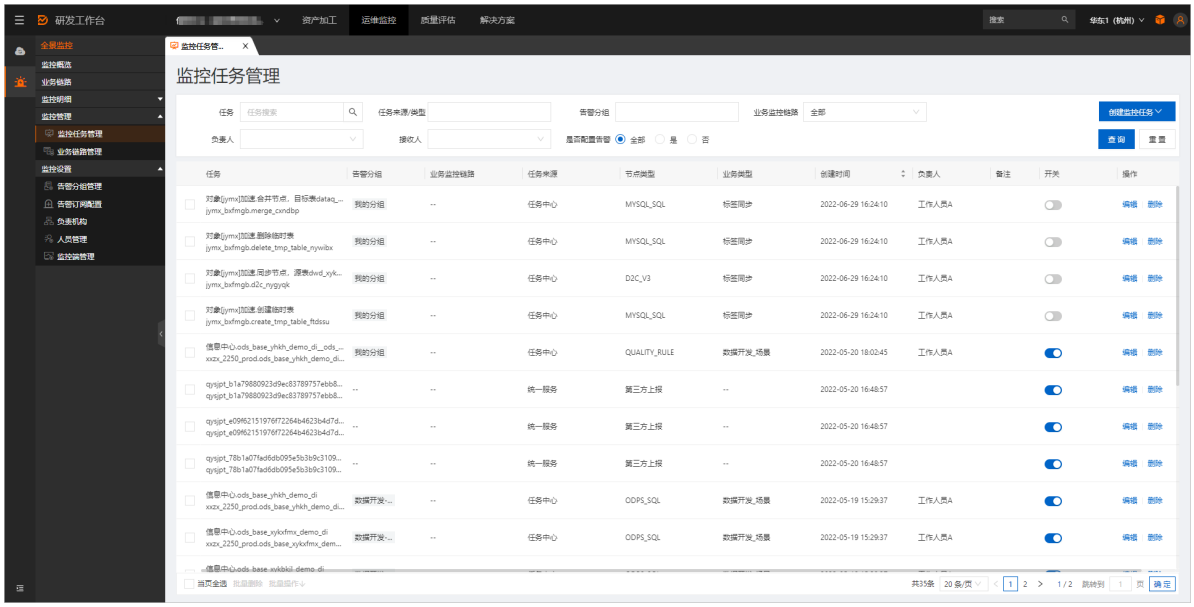
- 支持查看监控任务节点的最近一次告警信息及历史告警信息，并支持手动设置告警处理暂停时间状态。





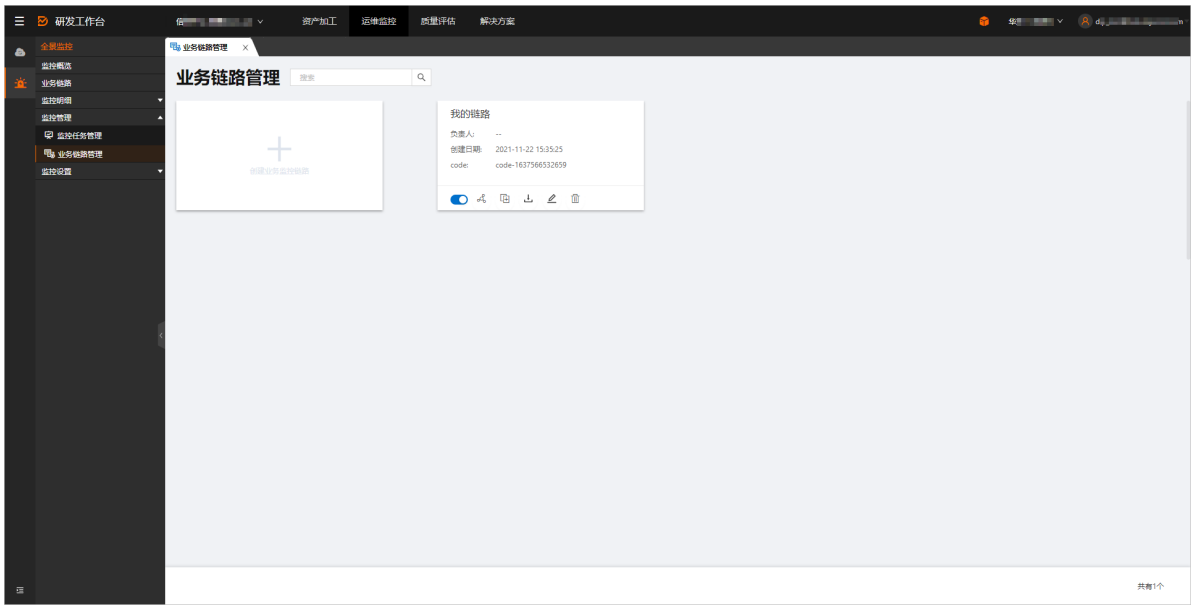
监控管理

- 监控任务管理



- 支持监控任务管理，针对不同来源的任务进行统一监控任务的配置管理。
- 支持通过批量创建、手动创建的方式完成监控任务创建。

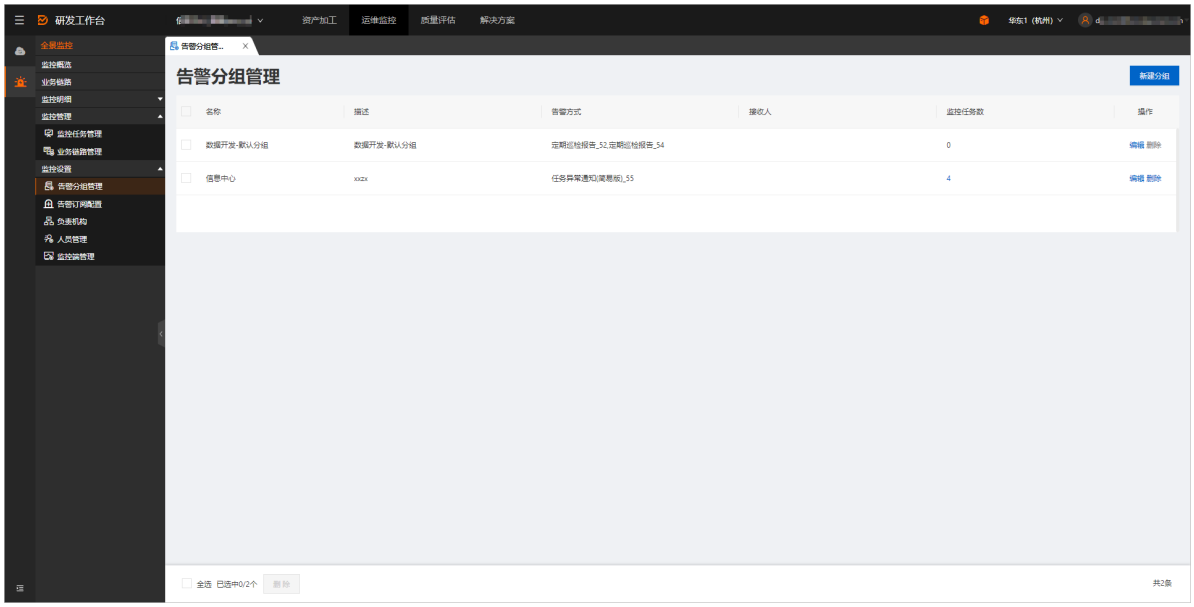
- 业务链路管理

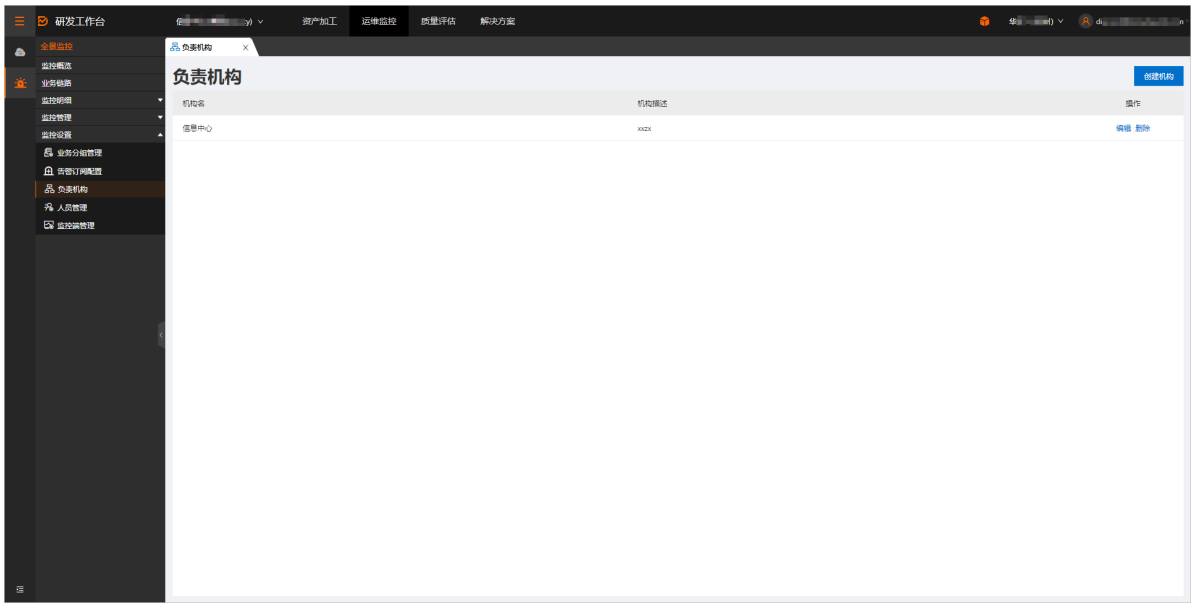


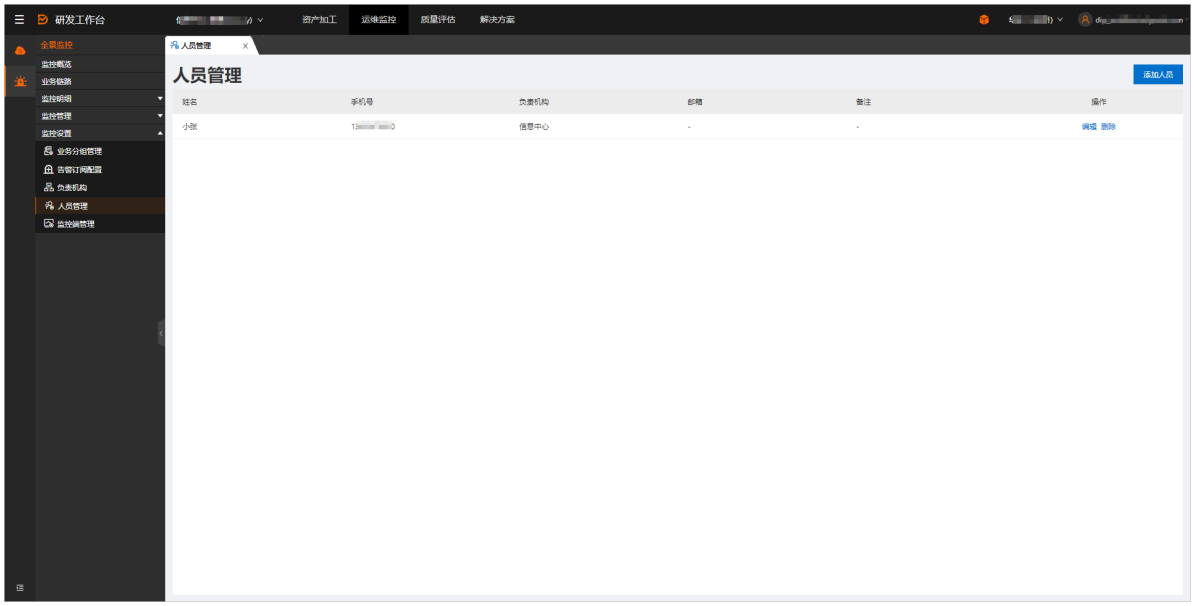
- 支持业务链路管理，按照业务视角自定义配置业务监控链路，可快速开启或关闭业务链路。
- 支持业务链路的导入、导出。

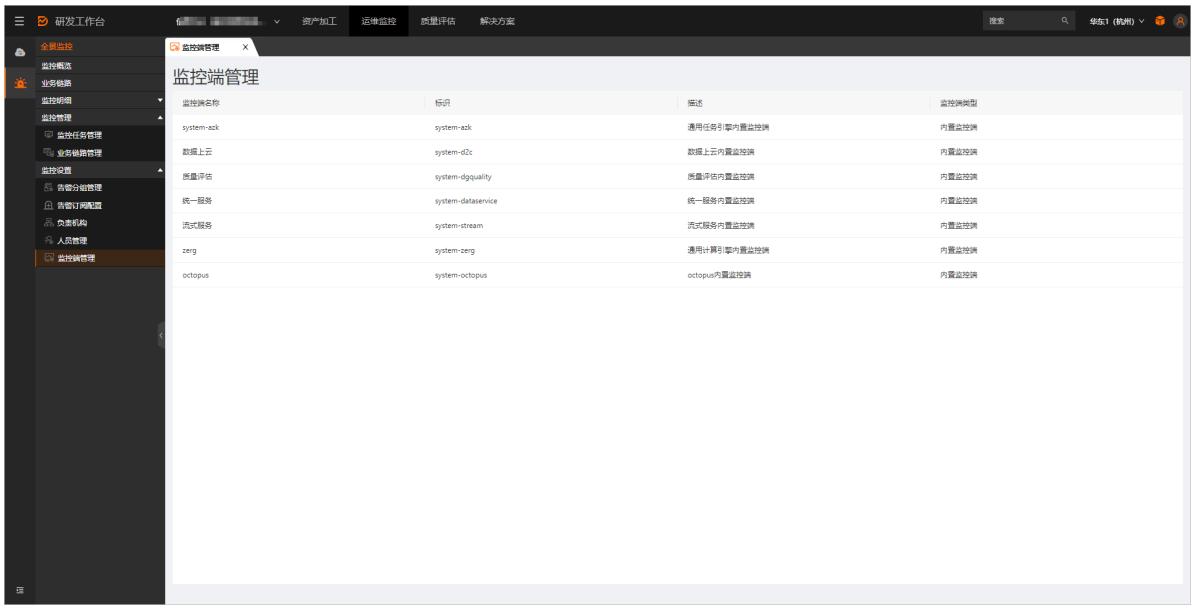
监控设置

- 支持告警分组管理、负责机构、人员管理、监控端管理。

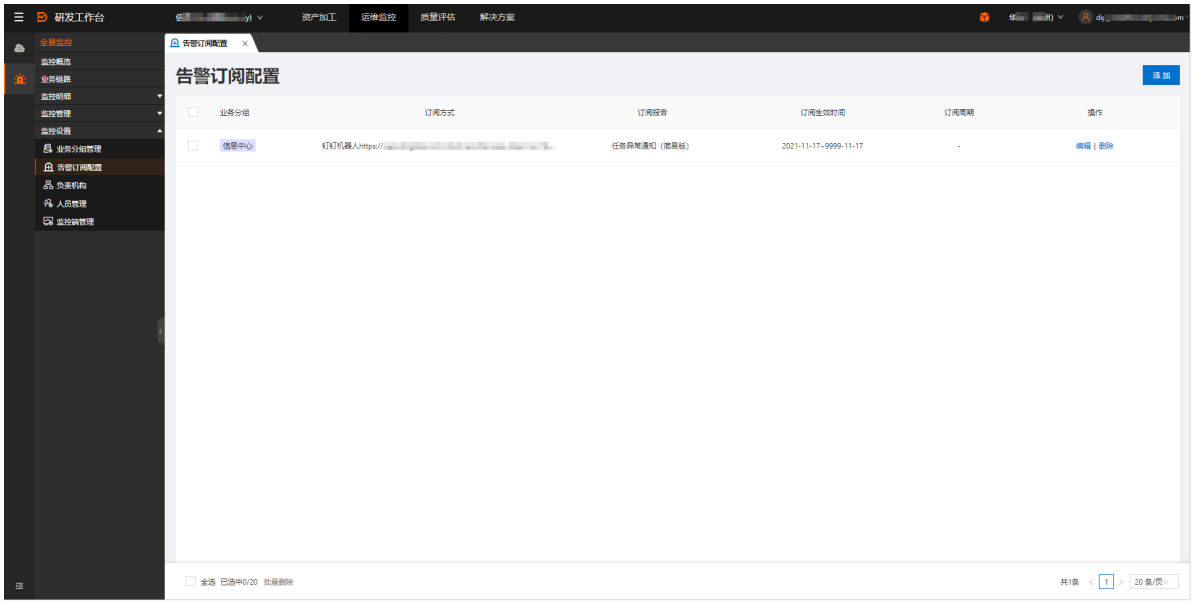





研发工作台
资产加工
运维监控
质量评估
解决方案
全景监控
监控概览
业务链路
监控明细
监控管理
监控设置
业务分组管理
告警订阅配置
负责机构
人员管理
监控端管理
人员管理
添加人员
姓名
手机号
负责机构
邮箱
备注
操作
小张
信息中心
编辑 删除


研发工作台
资产加工
运维监控
质量评估
解决方案
华东1（杭州）
全景监控
监控概览
业务链路
监控明细
监控管理
监控任务管理
业务链路管理
监控设置
告警分组管理
告警订阅配置
负责机构
人员管理
监控端管理
监控端管理
监控端名称
标识
描述
监控端类型
system-ack
system-ack
通用任务引擎内置监控端
内置监控端
数据上云
system-d2c
数据上云内置监控端
内置监控端
质量评估
system-dqquality
质量评估内置监控端
内置监控端
统一服务
system-dataservice
统一服务内置监控端
内置监控端
流式服务
system-stream
流式服务内置监控端
内置监控端
zerg
system-zerg
通用计算引擎内置监控端
内置监控端
octopus
system-octopus
octopus内置监控端
内置监控端

- 支持通过钉钉机器人、系统内置标准协议等方式订阅告警消息。



9.3.2.3. 质量评估

质量评估通过用户自定义质量检查计划，根据用户设置的数据质量规则执行面向不同计算资源的质量检查任务，通过数据质量监控报告展现系统整体数据质量概览及多维度细分数据的质量情况。

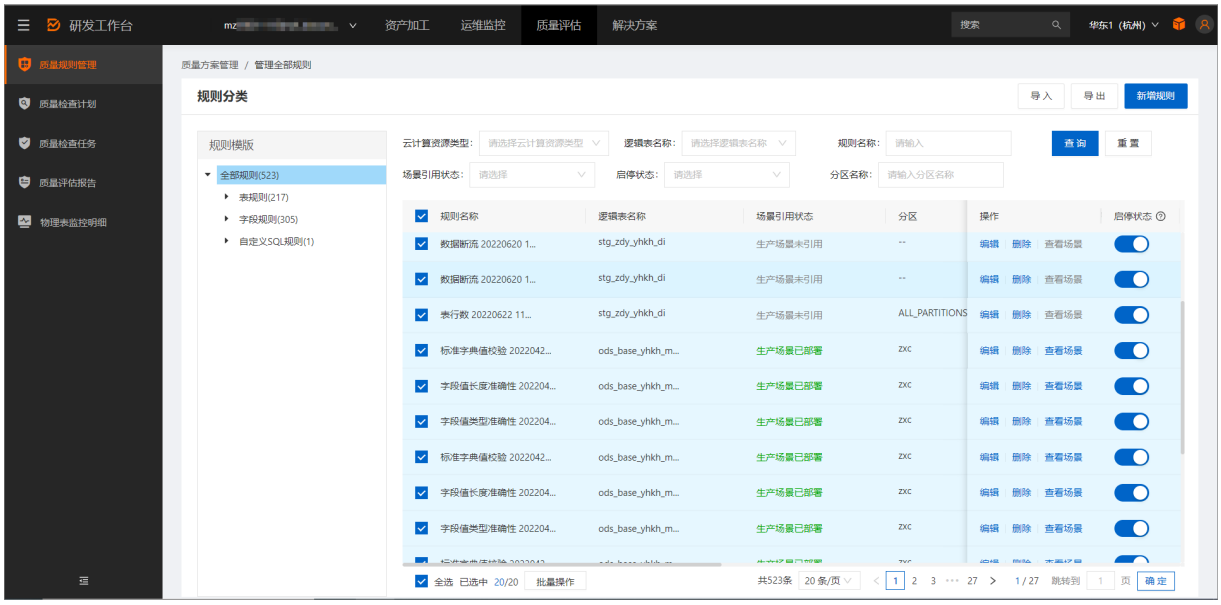
质量评估通过完整性、准确性、一致性、唯一性、时效性、规范性等六个维度构建数据质量指标体系，来实现对数据质量的量化评估。

- 完整性：数据完整性是指一个数据集的特定字段或属性都被赋予了数值。
- 准确性：数据准确性是指数据准确的反映其所建模的“真实世界”的实体。
- 一致性：数据一致性是指确保两个数据集之间或一个数据集内部不同字段间数值或信息的一致。
- 唯一性：数据唯一性是指在一个数据集中，没有重复的记录或信息。
- 时效性：数据时效性是指数据集按时、按量的更新程度。
- 规范性：数据规范性是指数据集对数据标准的满足程度。

质量规则管理

质量规则管理页面，查看数仓中配置的全部规则，包括表规则、字段规则、自定义规则。

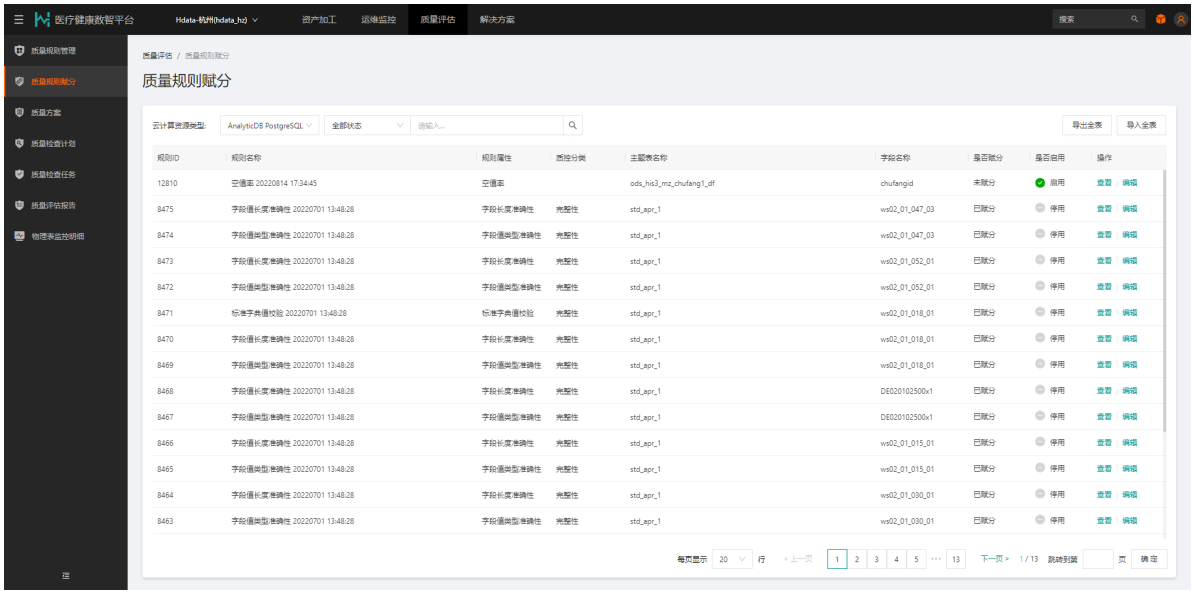
- 支持对所有的物理表统一配置的规则。
- 支持根据资源类型、逻辑表名称、规则名称、启停状态对规则进行筛选和查询。
- 支持对规则的删除、启停操作，删除和启停后规则的状态需要在数仓同步进行更新。
- 支持对全部规则、和选中规则进行导入和导出。
- 支持一键更新生产场景中的质量规则。



质量规则赋分

质量规则赋分页面，查看数仓中配置的全部规则，并对规则进行赋分操作。

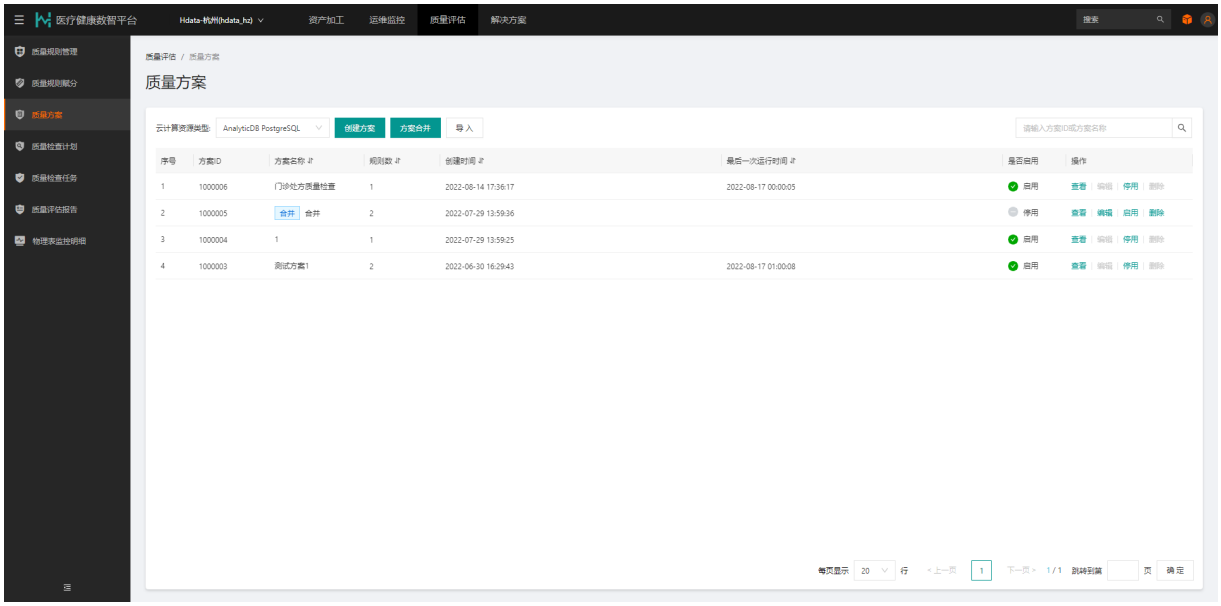
- 支持对规则赋分结果进行配置和编辑，包括配置期望区间、结果值以及对应的分值。
- 支持配置规则的质控类型。
- 支持对全部规则和选中规则进行导入和导出。



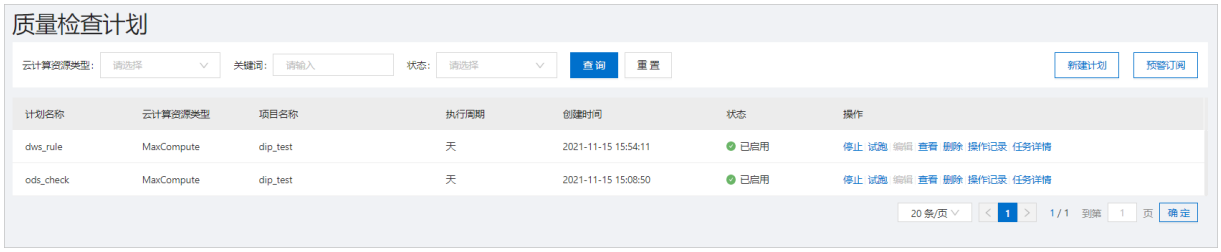
质量方案

质量方案页面，可查看所有维护好的质检规则组合即质量方案，并支持维护质量方案和方案组。

- 支持新建质量方案，将多条规则任意组成质量方案。
- 支持将任意质量方案合并成新的质量方案。
- 支持按照格式批量导入质量方案或合成质量方案。

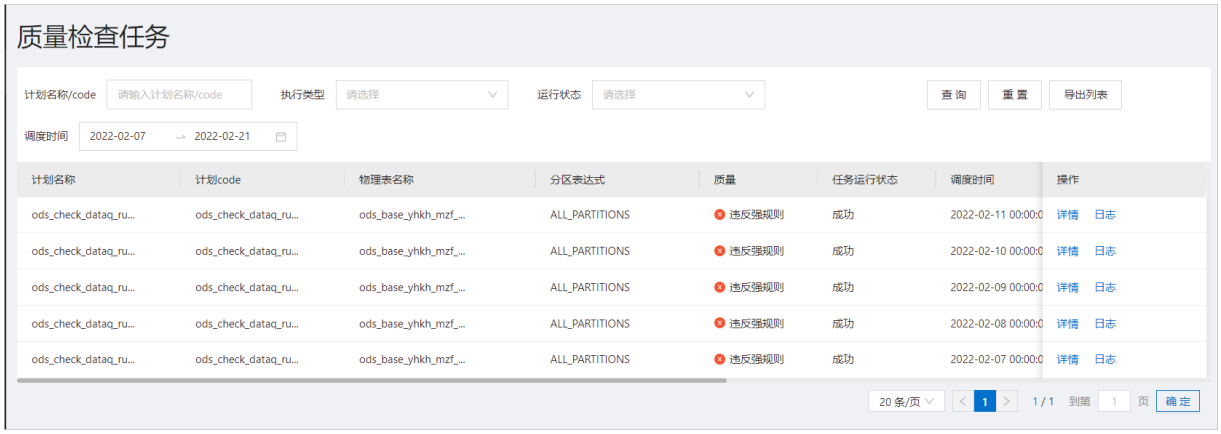


质量检测计划管理



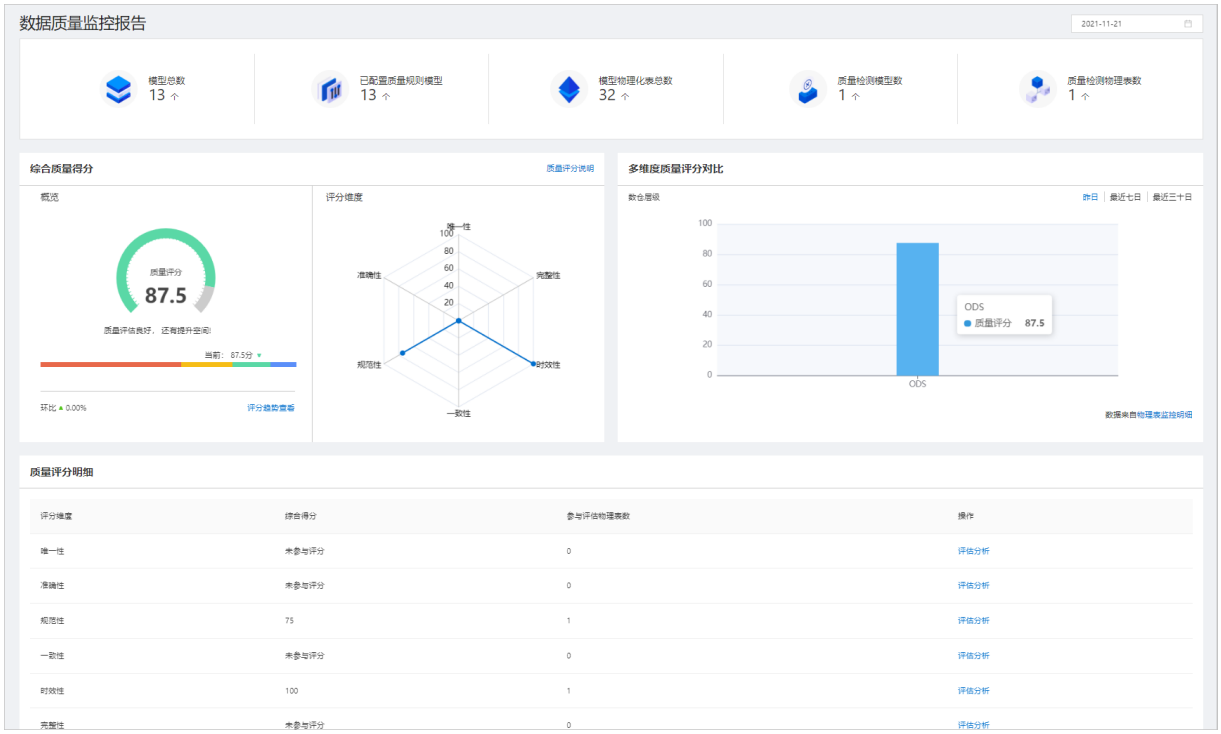
- 支持对质量检测计划的创建，可批量选择需要周期执行且已配置好质量规则的数据模型，配置周期检测任务和调度，进行持续质量检测。
- 支持对计划任务的管理、启动、停止和试跑。
- 支持对物理表质量结果进行告警订阅，在全景监控模块进行预警（依赖全景监控）。

质量检测任务管理



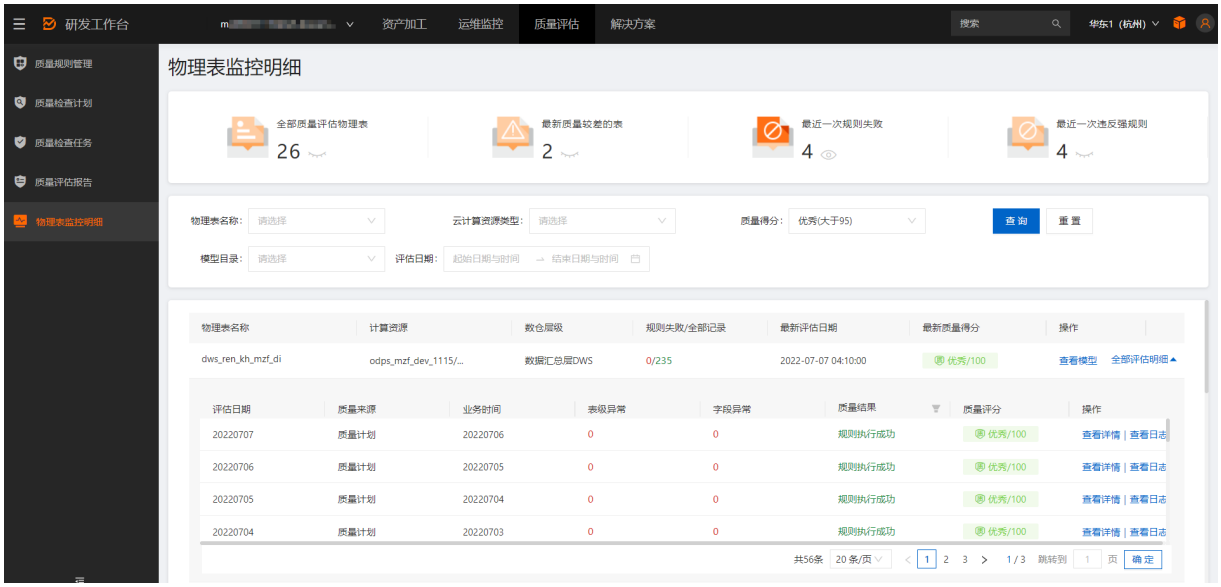
- 支持对质量检测任务运行实例进行查询和管理。
- 支持查看任务实例运行结果详情、运行日志进行查看。
- 支持对质量结果任务运行实例明细进行导出。

质量评估报告



- 支持根据系统数据质量监测运行结果以天的维度自动生成系统整体质量监控报告。
- 支持根据数据唯一性、准确性、规范性、一致性、时效性和完整性进行多维度评估分析。

物理表监控明细



- 支持以物理表为维度，查看物理表在数据开发和质量计划检测中最新的数据质量结果。
- 支持对物理表历史质量实例运行记录进行查看。

9.3.2.4. 解决方案

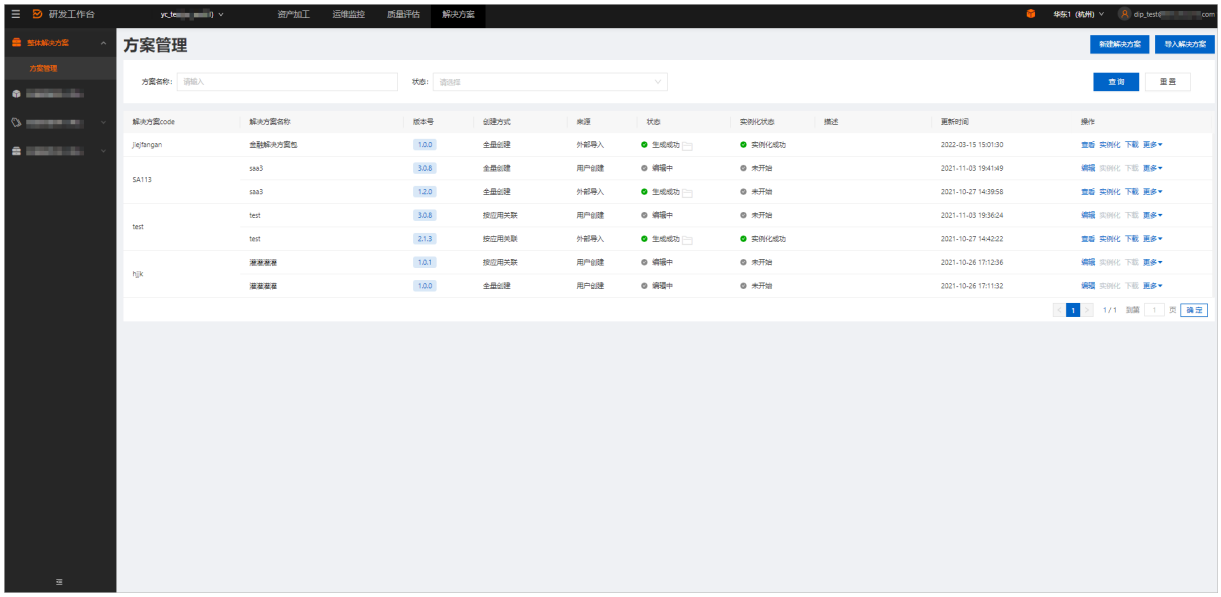
数据资源平台中的数据标准、数据模型、数据加工场景、数据服务API、云计算资源配置等资产，经过验证最终可沉淀为一个特定的解决方案。支持通过不同的创建方式，灵活的选择解决方案中包含的内容。依托数据资源平台中管理的大量元数据，确保解决方案的完整性、一致性和可用性。

根据数据血缘关系和不同的创建方式，自动选择关联的数据资产，保证解决方案的完整性。 同时，通过自动检测待生成的解决方案中是否存在异常（如表一致性、数据质量异常等），保证解决方案的一致性和可用性。对于需要在多项目部署的解决方案，支持以文件形式导出解决方案，并在项目中通过导入到本地的方式，快速实现解决方案部署、发布，以完成数据治理工作环境的初始化、跨平台的资产发布及迁移。

解决方案帮助用户沉淀行业知识，快速完成行业支持的复制，提升数据资产上线的效率和质量，加速数据业务价值的转化。

整体导入导出

将工作组中已沉淀的数据标准、数据模型、数据加工场景、数据服务API配置、云计算资源配置等数据资产创建解决方案后，进行下载、导入以及实例化。



- 支持将工作组内已加工完成的数据标准、数据模型、数据加工场景、数据服务API配置等信息，进行版本化解决方案封装。
- 支持全量、按范围关联两种方式灵活创建解决方案，并可对解决方案内容的逻辑一致性、物理一致性进行检查。
- 支持将解决方案包下载后，导入并实例化至工作组，完成数据治理工作环境的初始化、跨平台的资产发布及迁移。

9.3.3. 运营工作台

运营工作台主要面向租户资产管理员，为用户提供数据资产编目、数据资产上线、数据资产下线等能力。通过获取资产元数据信息，并同步到数据资产目录完成数据资产注册。

数据资产管理涉及三种用户角色，包括资产使用方、租户资产管理员、工作组管理员。

- 资产使用方：主要进行数据查找和使用申请，目标是找到需要使用的数据资产，进行权限获取以完成后续业务使用。
- 资产管理员：对所有可对外服务的高价值数据资产进行管理维护与编目上线，进行资产对外服务的生命周期维护，同时针对申请使用进行过程审核。
- 工作组管理员：对该角色负责管理的工作组下数据资产进行使用侧的管理审核。

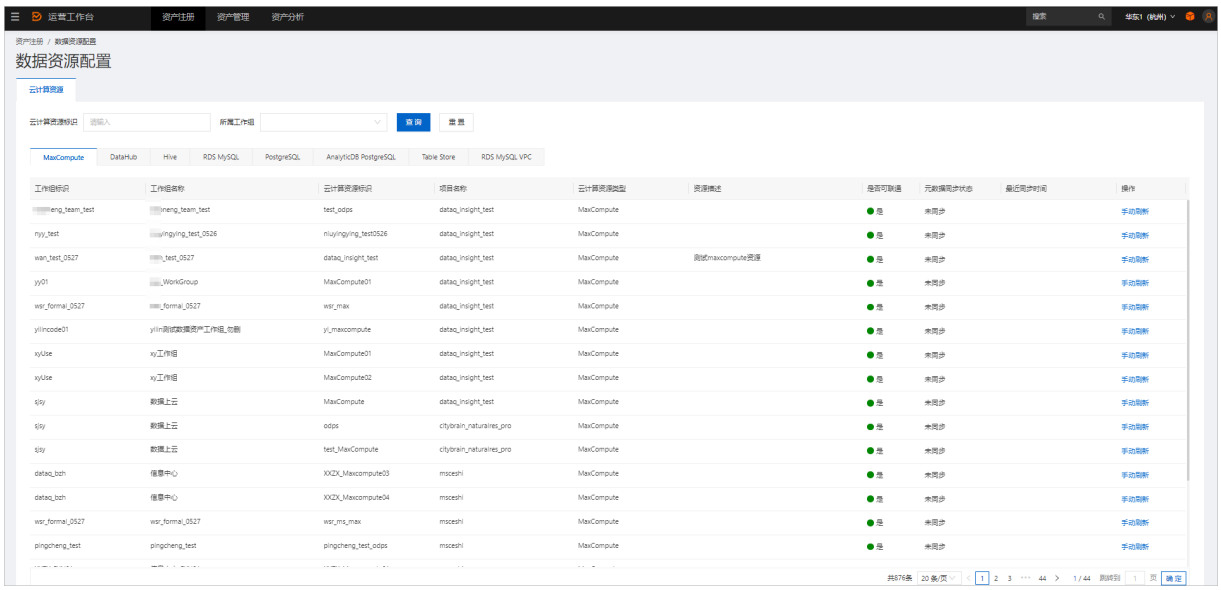
数据资产管理体系主要包括以下几个方面的内容。

- 数据资产编目，数据管理者对平台数据资产进行梳理，确定其共享属性，并按照技术标准，进行数据类目定义。
- 数据资产上线，将已经注册管理的高质量数据资产资源发布，供数据使用者查询和使用。
- 数据资产申请，数据使用者向数据管理方提出使用申请，说明申请方式和申请理由。由租户资产管理员进行资产使用的审核处理。
- 数据资产服务，根据数据资源的性质和特点，选择采用数据表、接口方式提供资产服务。
- 数据资产下线，根据数据的实际情况不再共享数据资产。

资产注册

资产注册功能是数据资源平台资产管理的基础，可快速锁定待管理数据资产的范围。针对支持管理的数据表类型，实现资产信息注册自动化。在工作组中注册存储计算资源后，系统自动执行元数据信息扫描，并完成数据资产发现，为后续资产运营管理持续提供动态更新的资产对象。同时自动完成已上架API资产上报，作为资产管理中待编目维护的数据内容。

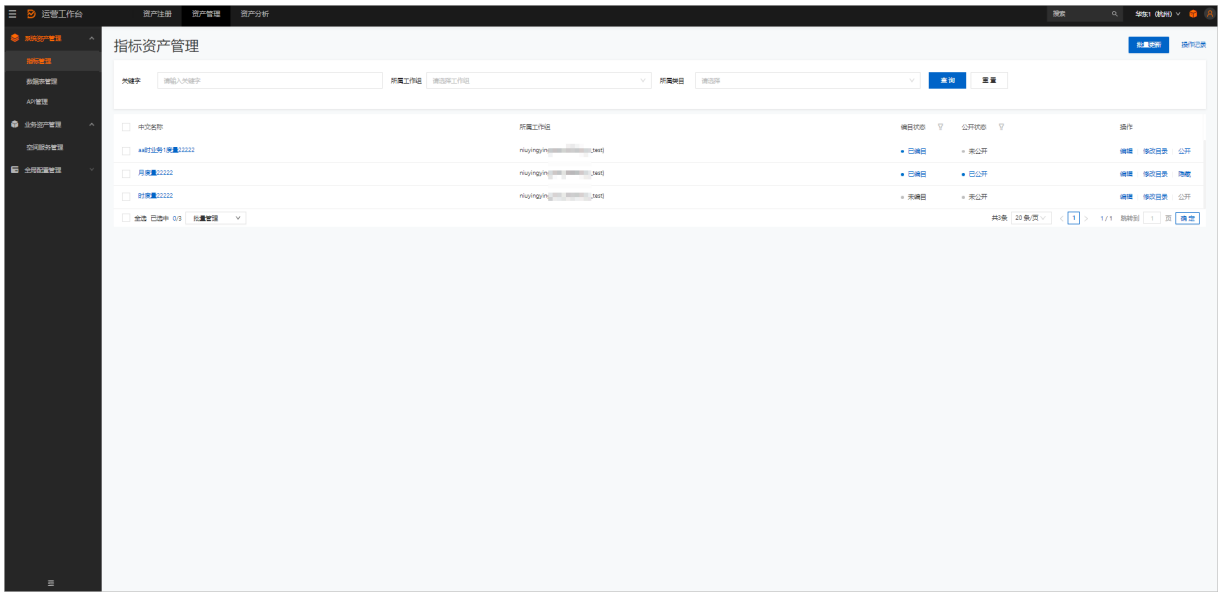




- 支持平台侧需管理的数据资产自动化注册功能，可选择需要管理的数据资产来源作为管理输入。
- 支持对MaxCompute、RDS MySQL、PostgreSQL、DataHub、Hive、AnalyticDB PostgreSQL和Table Store数据表资源进行展现；可通过手动刷新进行元数据同步。

资产管理

系统资产管理：提供对表资产、API资产和指标资产可视化展示，通过编目、公开为资产共享做准备。



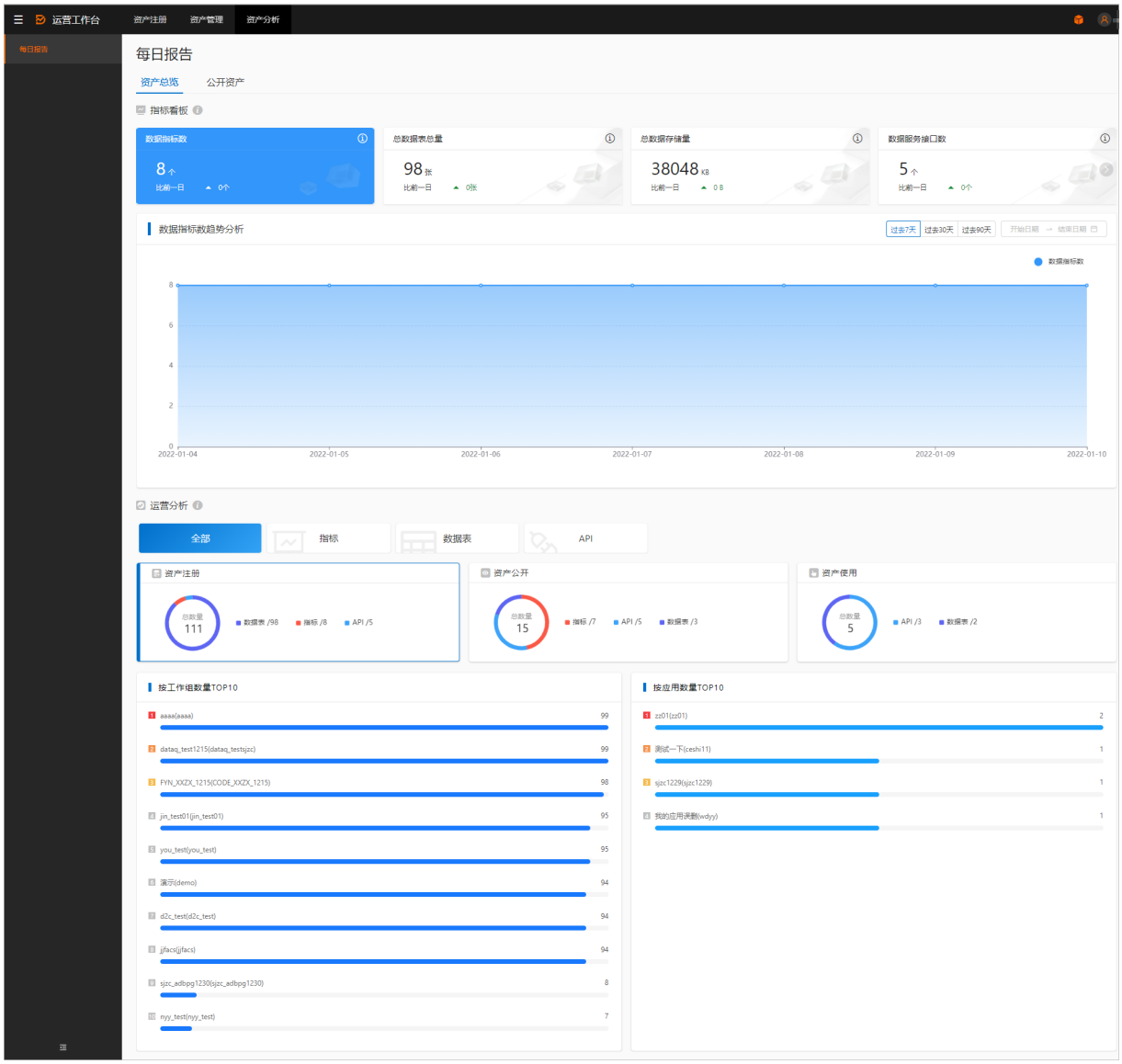
- 数据表资产管理：支持对数据表资产进行编目，并提供查询、公开、隐藏等功能。
- API资产管理：支持对数据服务API、三方API、上报API、数据更新API和服务编排API资产进行编目，并提供查询、公开、隐藏等功能。
- 指标资产管理：支持对指标资产进行编目，并提供查询、公开、隐藏等功能。

业务资产管理：支持对空间服务资产进行编目，并提供查询、公开、隐藏等功能。

全局配置管理：提供自定义目录配置功能，支持设置多级目录，以供数据资产进行后续编目。

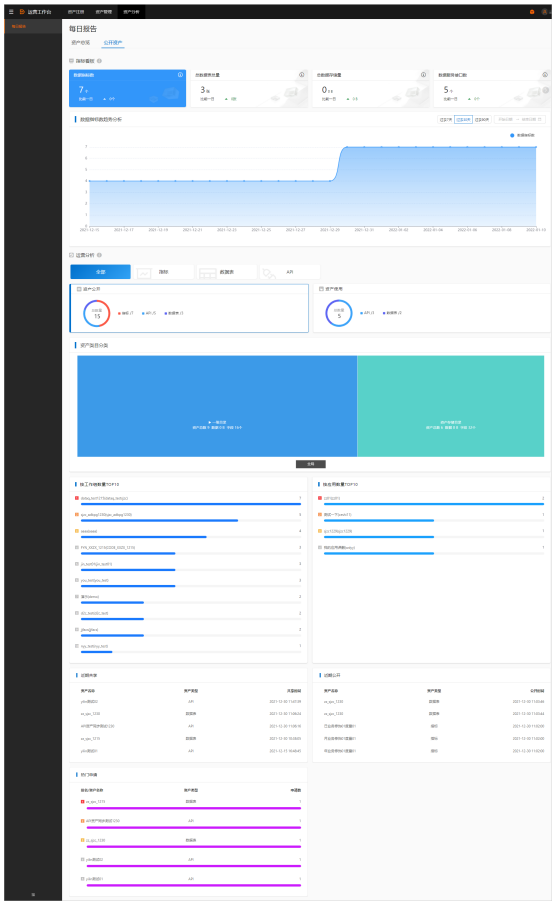
资产分析

资产总览：包括指标看板、数据指标数趋势分析和运营分析。



- 指标看板：展示未公开及已公开的资产，对全租户下已注册的资产进行指标统计，可了解各类资产的量级统计和时间趋势情况。
- 数据指标数趋势分析：展示未公开及已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示未公开及已公开的资产，对全租户下已注册的资产，围绕注册资产的运营阶段（注册-公开-使用），了解各阶段下的多维度统计分析结果。

公开资产：包括指标看板、数据指标数趋势分析和运营分析。



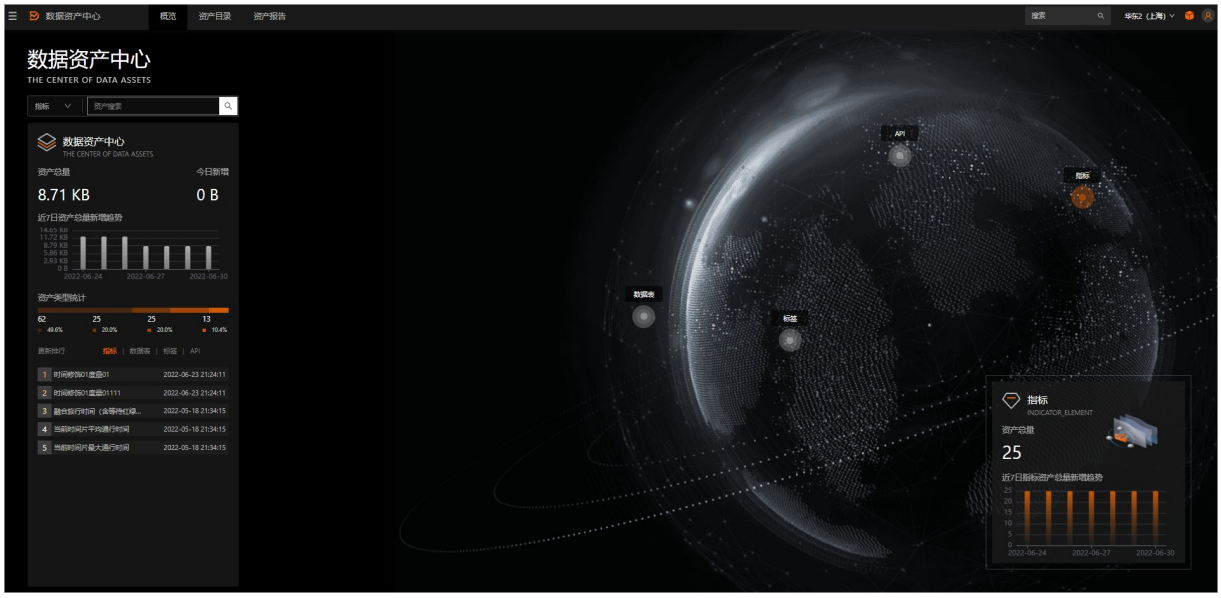
- 指标看板：展示已公开的资产，对全租户下已公开到资产中心的资产进行指标统计，了解各类资产的量级统计和时间趋势情况。
- 数据指标数趋势分析：展示已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示已公开的资产，对全租户下已公开到资产中心的各类资产，围绕公开资产的运营阶段（公开-使用），了解各阶段下的多维度统计分析结果。

9.3.4. 数据资产中心

数据资产中心帮助用户完成数据资产申请、数据资产服务、数据资产共享。通过统一的数据资产目录，沉淀数据供给方生产的各类型数据，帮助资产运营管理者进行高价值、可共享的数据资源的注册、编目分类和公开运营，帮助数据使用者快速了解和获取业务需要的高质量数据资产，促进数据资产发挥更大的业务价值。同时，通过数据资产全局血缘，了解数据资产的上下游关联。

概览

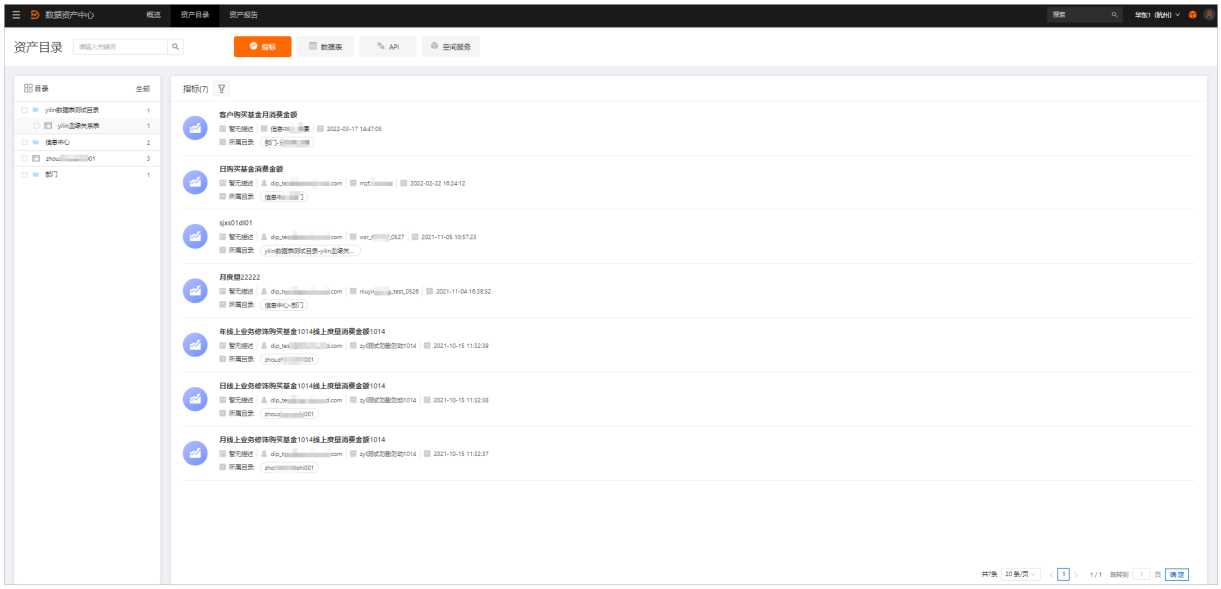
对数据资产的全局及增量指标进行统计和展示，帮助用户快速了解平台管理数据资产的总体情况。



- 支持数据表、API和指标资产类型，同时，支持精确或模糊搜索数据资产。
- 支持数据资产总量、今日新增等全局指标展示，并以趋势图的形式展示增量信息。
- 支持数据表、API和指标资产独立展示增量信息。

资产目录

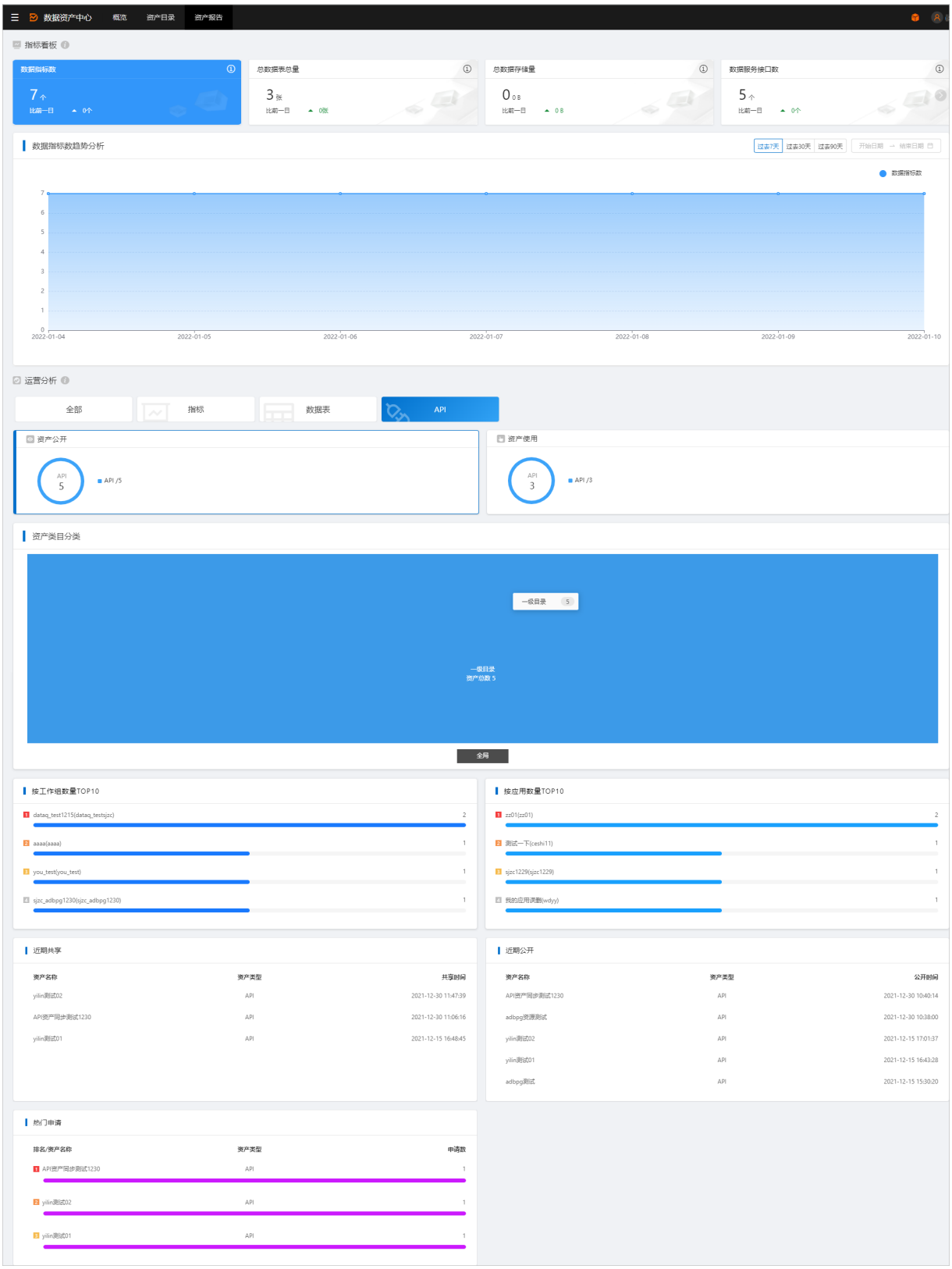
对公开共享的数据资产，可在资源目录进行搜索、目录列表查看，并通过资产详情，了解数据资产的业务价值，促进数据资产的共享和申请使用。



- 支持展示公开的数据表、API和指标资产目录，并提供通过关键词搜索对应资产。
- 支持展示数据表资产基本信息、扩展信息、字段信息、数据预览、血缘关系和数据质量等信息，并提供申请数据表使用权限的功能。
- 支持展示API基本信息、扩展信息、接口信息、血缘信息，并提供申请API使用权限的功能。
- 支持展示指标基本信息、扩展信息、概览信息和关联信息等信息。
- 支持展示空间服务基本信息、扩展信息、概览信息、数据预览和用户评价等。

资产报告

展示指标看板、数据指标数趋势分析和运营分析。



- 指标看板：展示已公开的资产，对全租户下已公开到资产中心的资产进行指标统计，了解各类资产的量级统计和时间趋势情况。包括数据指标数、总数据表总量、总数据存储量、数据服务接口数和各趋势分析图。

- 数据指标数趋势分析：展示已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示已公开的资产，对全租户下已公开到资产中心的各类资产，围绕公开资产的运营阶段（公开-使用），了解各阶段下的多维度统计分析结果。包括指标、数据表和API各类型资产的资产公开数据、资产类目分类和近期公开资产等信息。

9.3.5. 画像分析

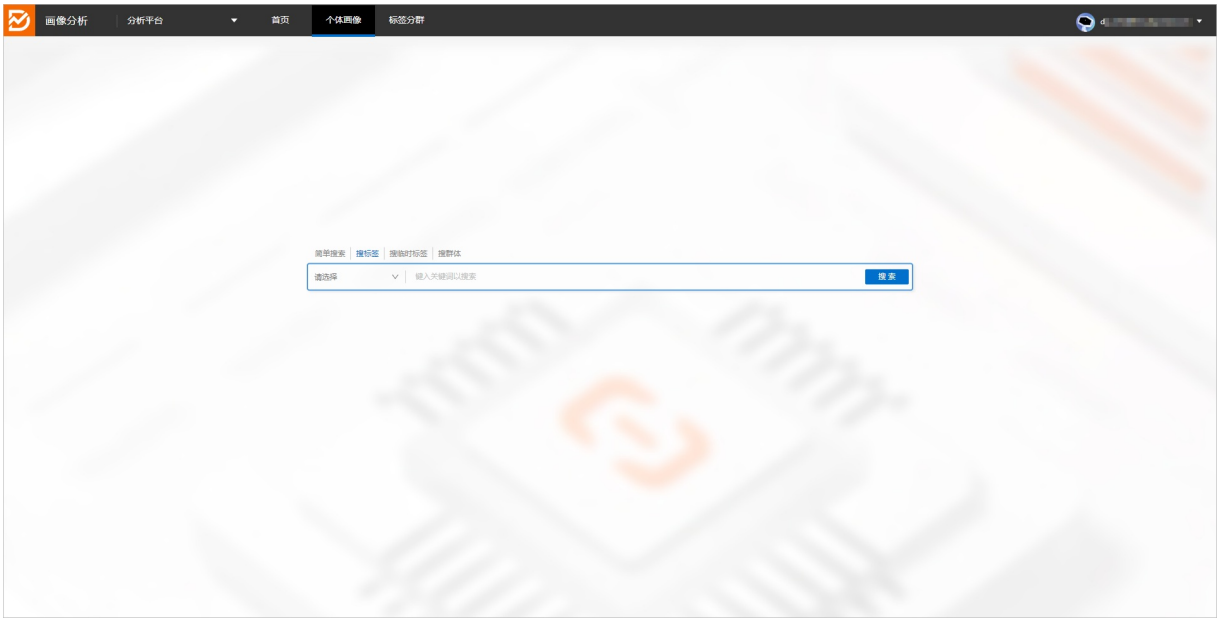
画像分析是对象数据的形象化展示，即将对象的信息画像化，也就是用一系列的画像简洁明了地刻画一个群体或一个个体的特征，是在一连串数据上构建起来的对象模型。对象画像是描述目标对象、分析与聚焦的有效利器，为业务人员提供了无需技术基础，零代码的大数据分析能力。

画像分析系统可基于智能标签系统为数据分析人员、业务人员提供一站式群体分析平台。涵盖画像分析的全链路，主要包括基于标签的群体圈选、群体计算、多状态群体发布等，支持用户建立单个体档案，从而实现用户对于目标群体的精准定位。

画像分析中的人群支持通过将现有标签进行鼠标点选组合即可生成新的标签组合，并可给该标签组合命名，该标签组合在前端保存为标签计算的逻辑，在大数据环境下提供快速计算，不同的新标签在页面可以快速嵌套使用，创建更贴近业务的标签组合。

个体画像

- 提供按群体、个体标签、个体临时标签、关联对象进行搜索的功能，可快速找到目标个体或群体。
- 查询结果支持查看群体和个体详情，群体查询结果支持群体分析。

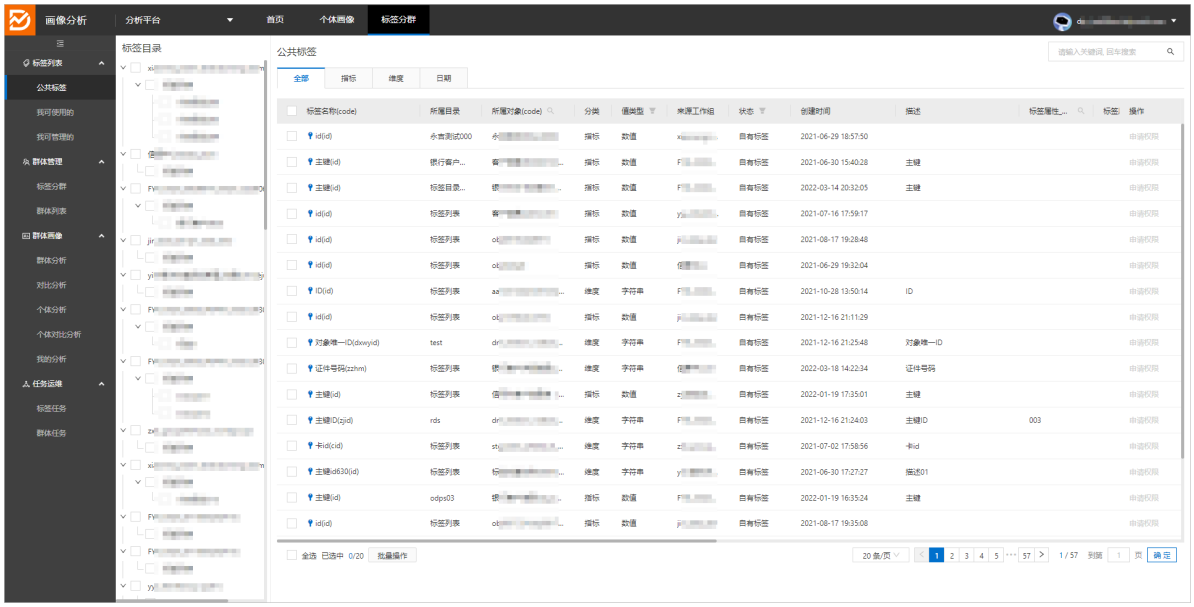


标签分群

标签列表

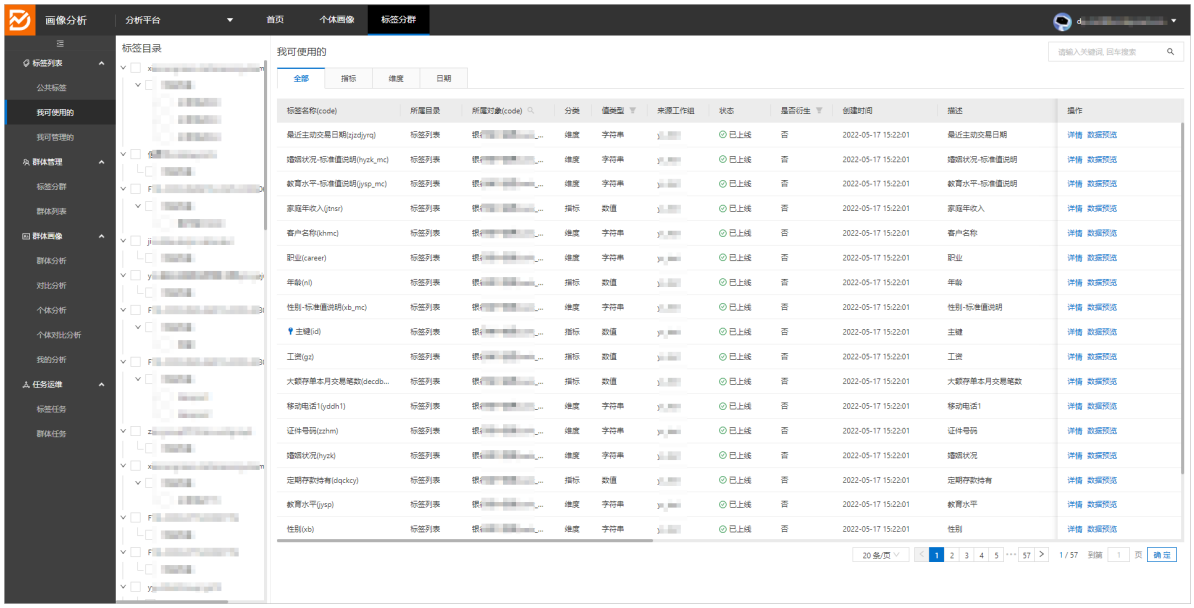
- **公共标签**

支持查看该租户下所有工作组的标签。



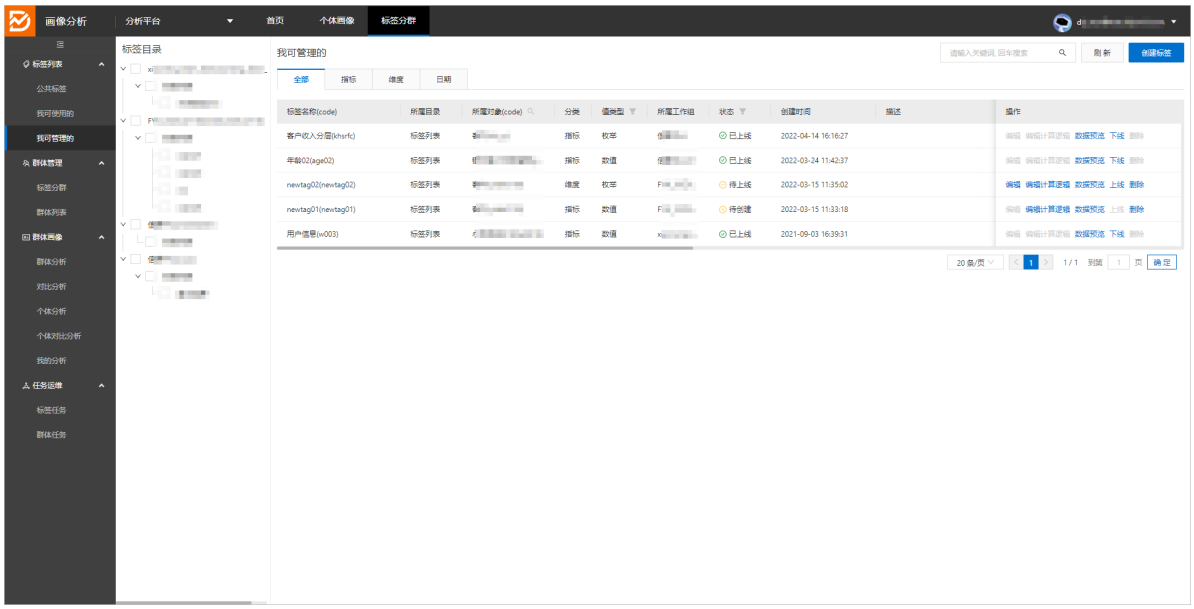
- **我可使用的**

  支持查看该租户下有使用权限的已上线标签，并可进行标签详情查看、数据预览操作。



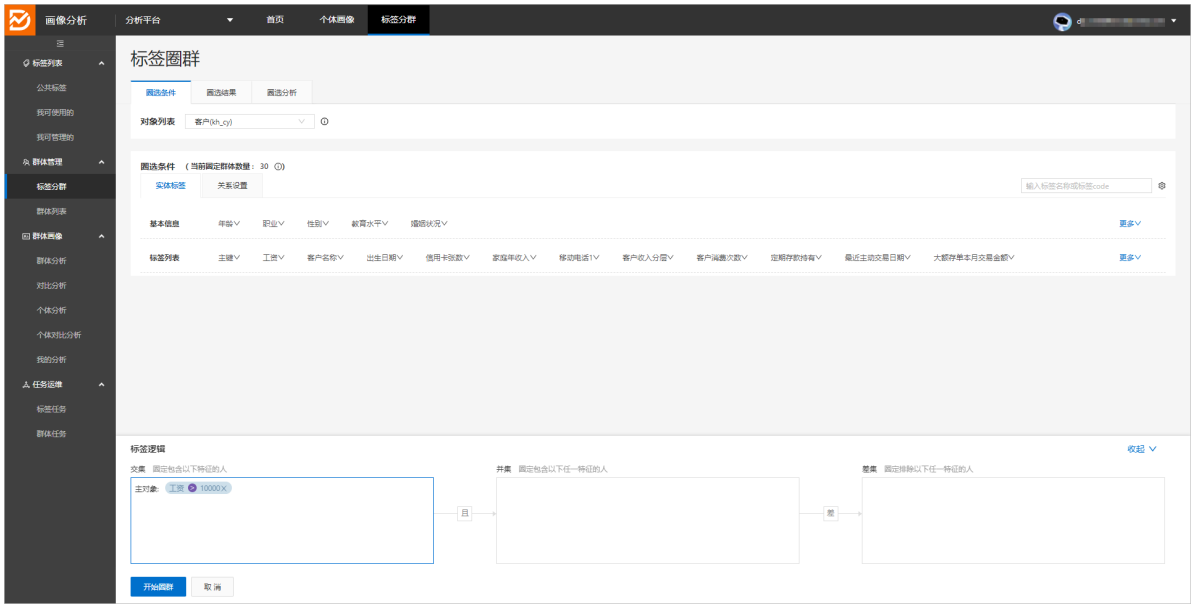
- **我可管理的**

支持通过拖拉拽方式编辑计算逻辑创建枚举型、数值统计型标签，并对衍生标签进行上线、下线方便业务用户沉淀业务经验。



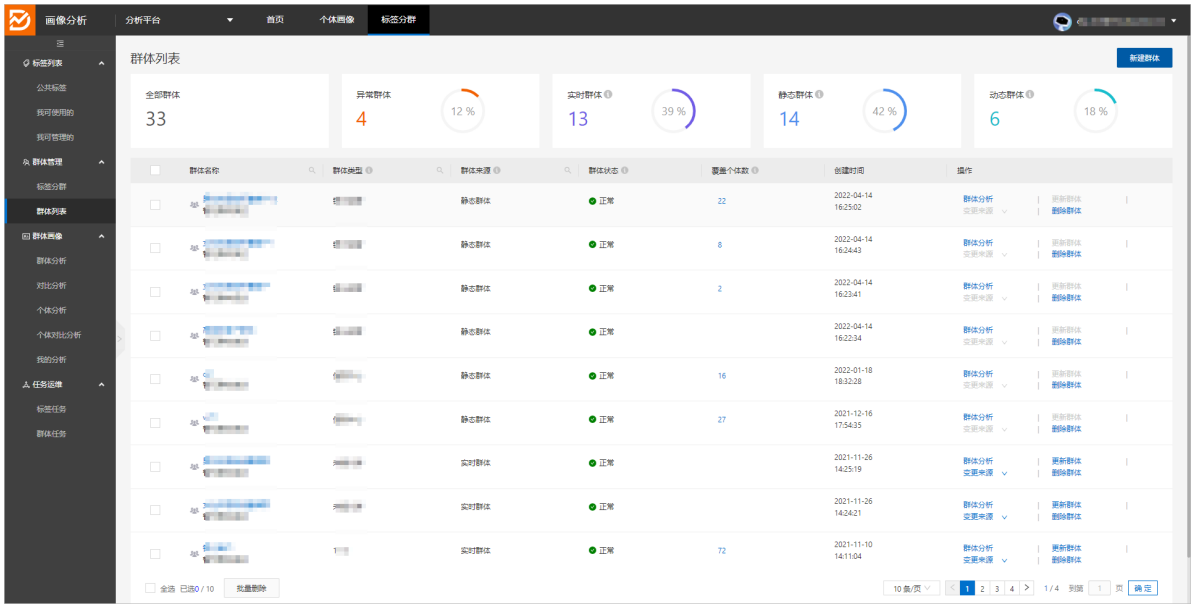
群体管理

- **标签分群**
  - 提供对智能标签里面设定好的对象上的标签设置筛选条件圈定目标群体功能。
  - 支持在标签圈群页面对单个对象或对象标签模型的对象间进行标签的交并差集组合计算生成实时、静态和动态群体。
  - 群体圈定结果支持重新圈定、发布群体、发布API、下载到本地、设置圈选结果展示标签、分析等功能。



- **群体列表**

支持群体统一管理。支持对群体信息进行新建、删除、更新、状态变更、查看、编辑、查看不同类型群体个数和占比等操作。新建群体支持标签圈群、导入本地文件和群体计算三种方式。可实现单个个体多维信息展示，并支持临时打标签能力，具有该临时标签的个体，可以在智能搜索模块被搜索；支持导入Excel格式的群体列表到画像中，进行群体的分析。导入前提，该Excel所含有的标签和值，已经在数据库中存在。支持用户在详情页面中，查看群体详情和个体详情信息。



- **群体计算**

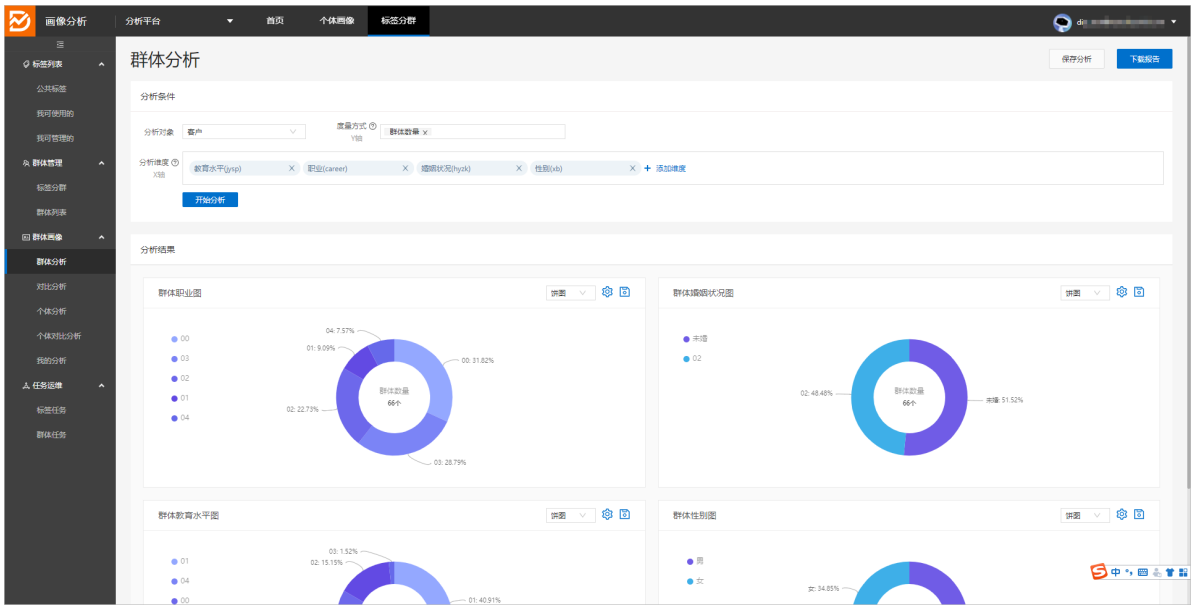
  通过图形化的计算方式可实现两个群体交并差计算得到新的群体，并提供对群体结果支持查看详细群体、发布API、静态、动态和实时群体。



群体画像

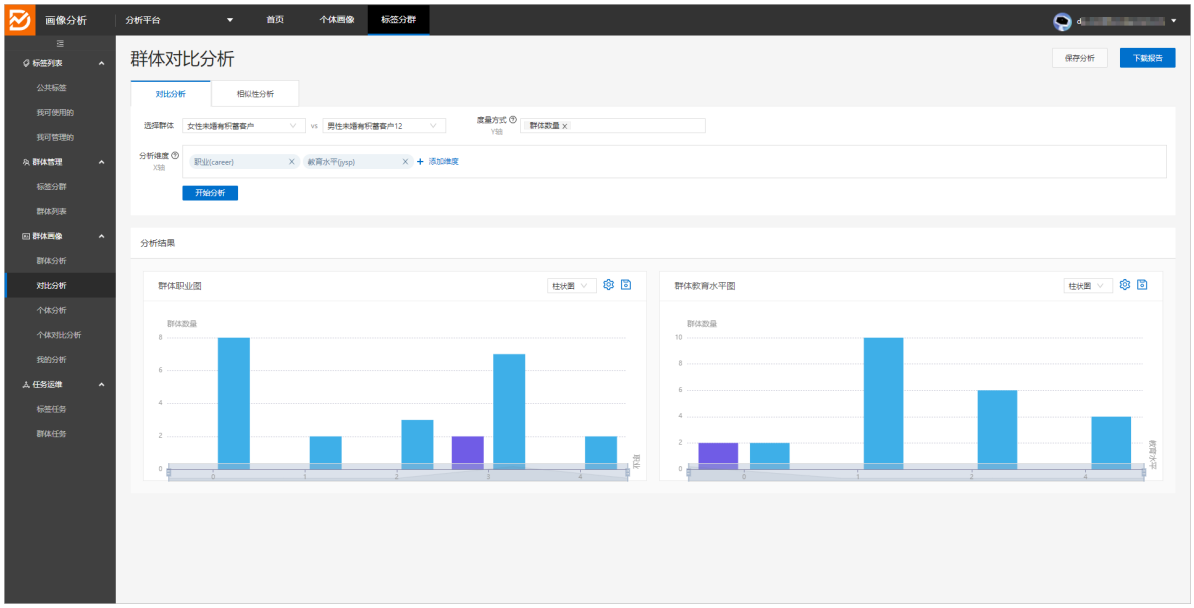
- **群体分析**

支持单群体或对象按单个度量方式的多个维度分析，并可以对分析结果保存为分析模版和下载分析报告。分析结果可生成柱状图、折线图、面积图和饼图，图表可支持的聚合方式包括计数、平均值、求和、最大值和最小值。



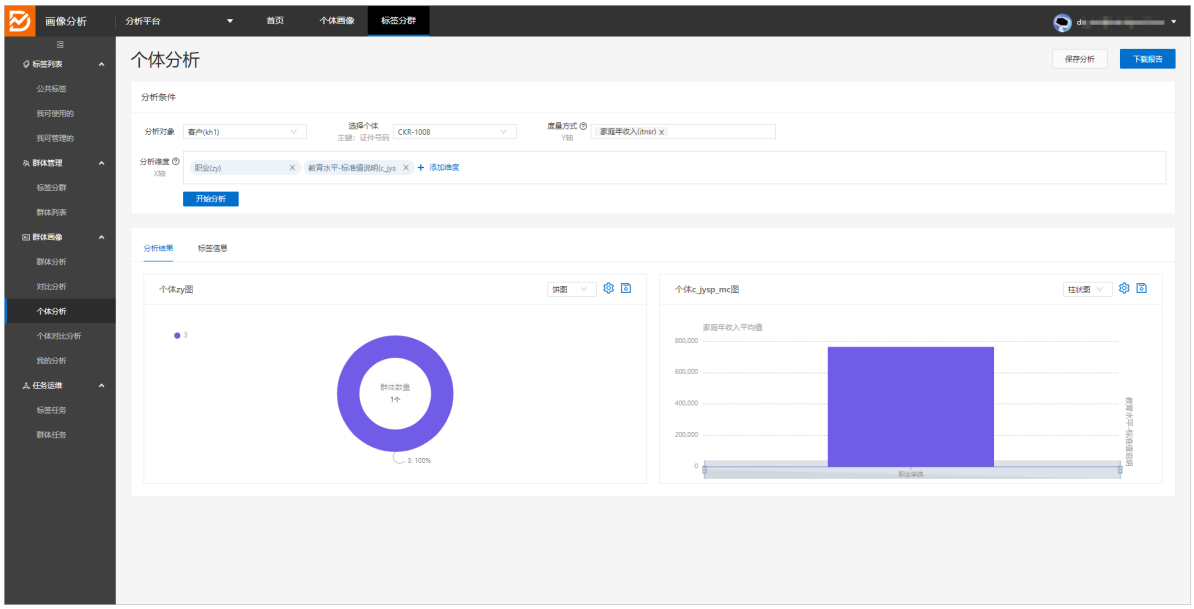
- **对比分析**

支持两个群体按单个度量方式的多个维度分析和相似性分析。两个群体按单个度量方式的多个维度分析，并可以对分析结果保存为分析模版和下载分析报告。分析结果可生成柱状图、折线图和雷达图，图表可支持的聚合方式包括计数、平均值、求和、最大值和最小值。相似性分析支持查看两个群体的群体相似度、标签覆盖人数排行榜和相似个体列表。



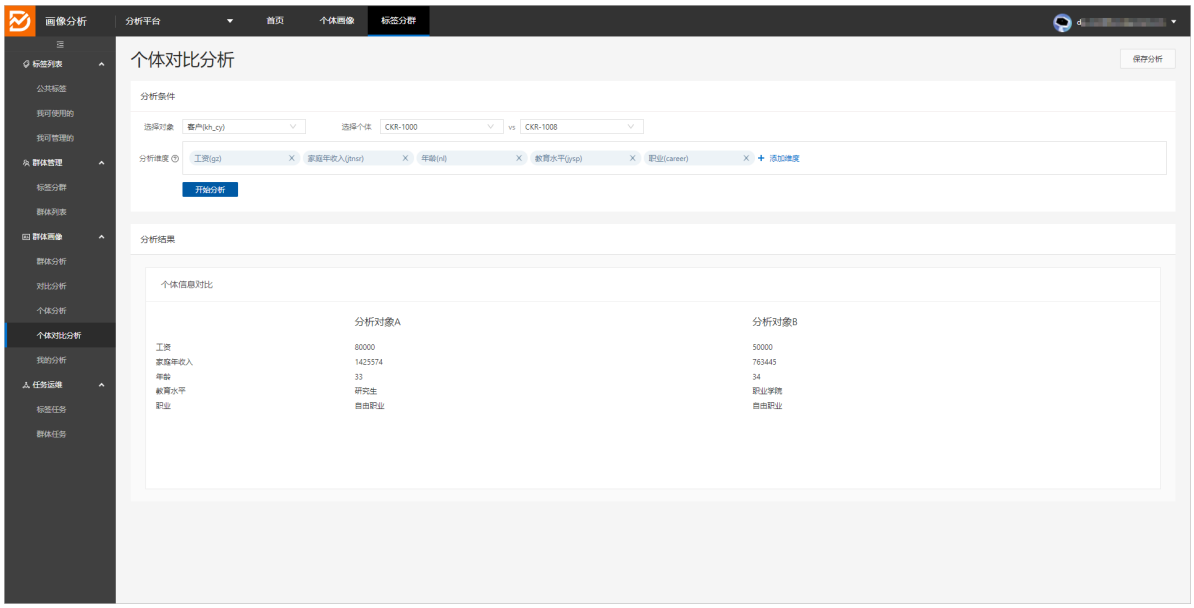
- **个体分析**

支持以对象为基础、以对象主键为核心个体，通过添加一个对比字段为度量方式，添加一个或多个字段为维度对单个个体进行分析，并提供保存、下载分析报告。

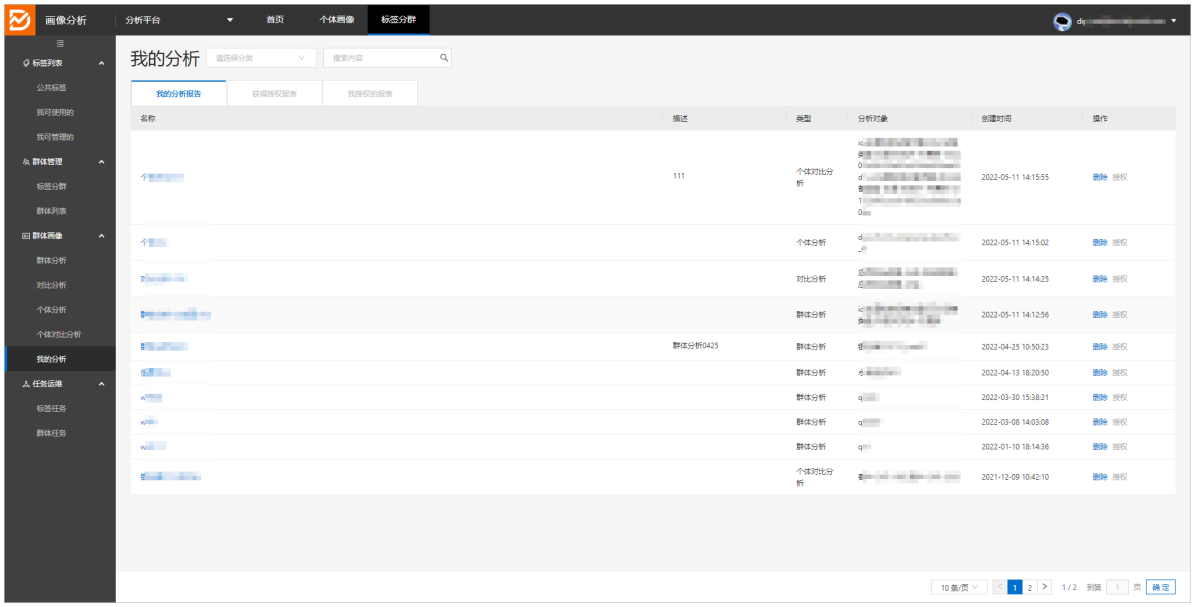


- **个体对比分析**

  支持两个体间的多维对比分析，并提供对分析结果保存为分析模版。



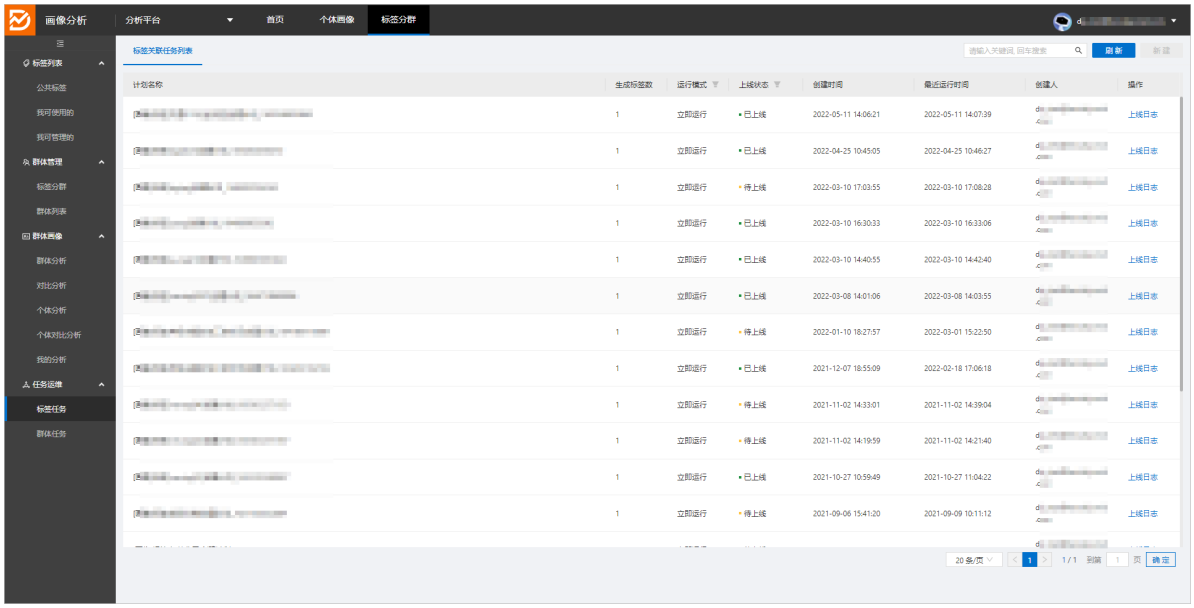
- **我的分析**
  - 支持对查看本工作组下历史保存的分析结果，并提供删除功能。
  - 支持查看其他工作组赋予查看权限的分析报表。
  - 支持赋予其他工作组查看本工作组下分析报告的权限，并且支持权限的撤回。



任务运维

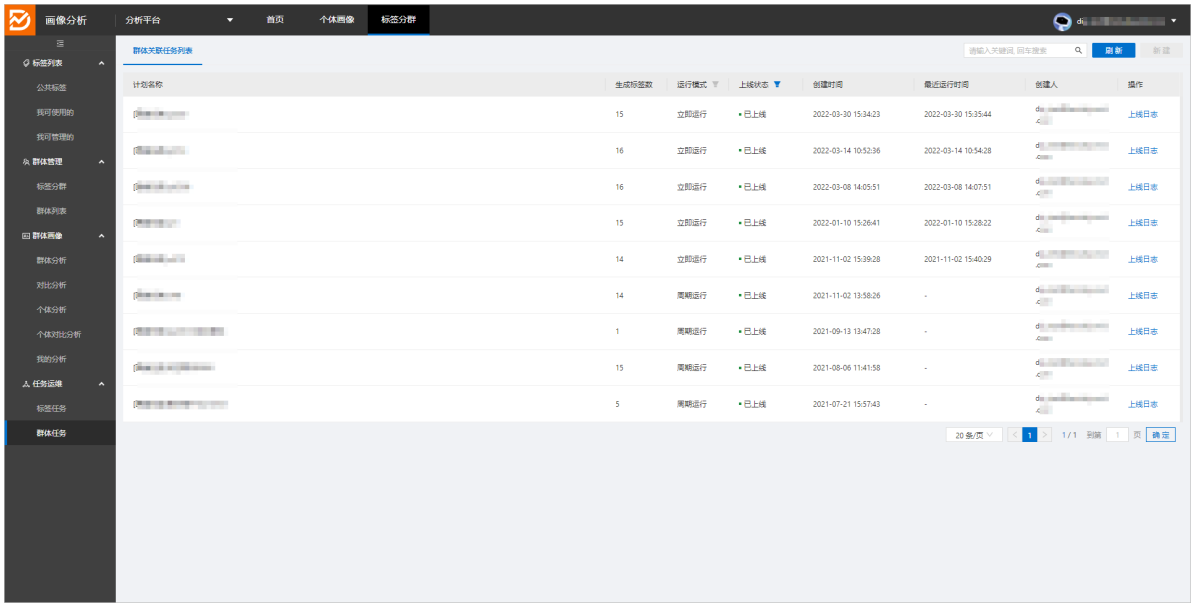
- **标签任务**

  支持创建标签任务的任务运维管理。



- **群体任务**

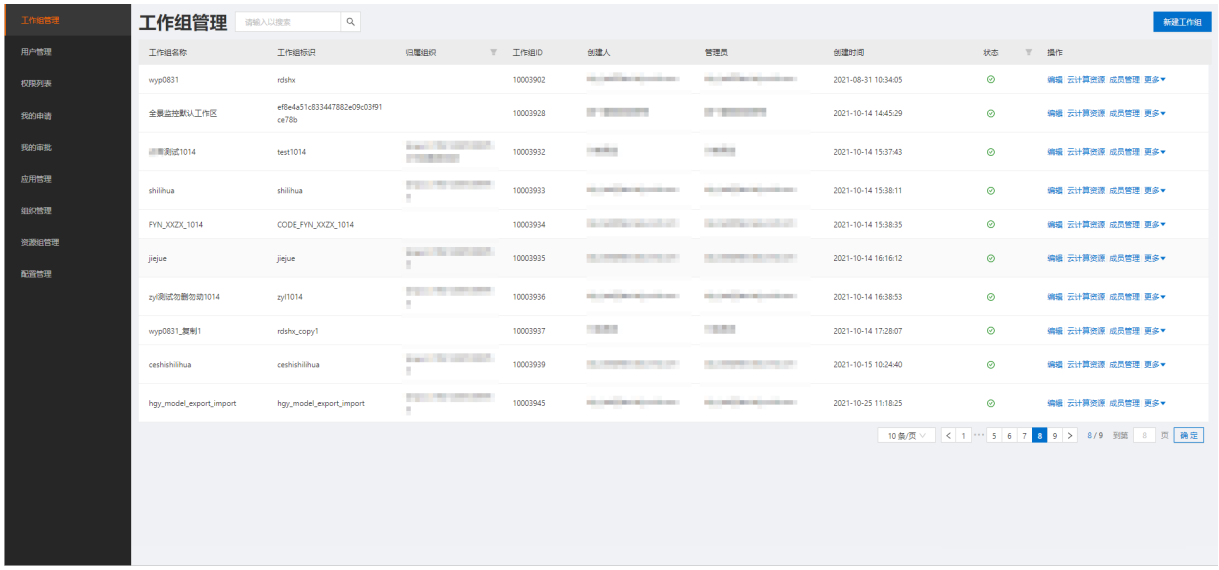
支持发布静态和动态群体，并提供在数据库一次性和周期落表的任务运维管理。



9.3.6. 系统设置

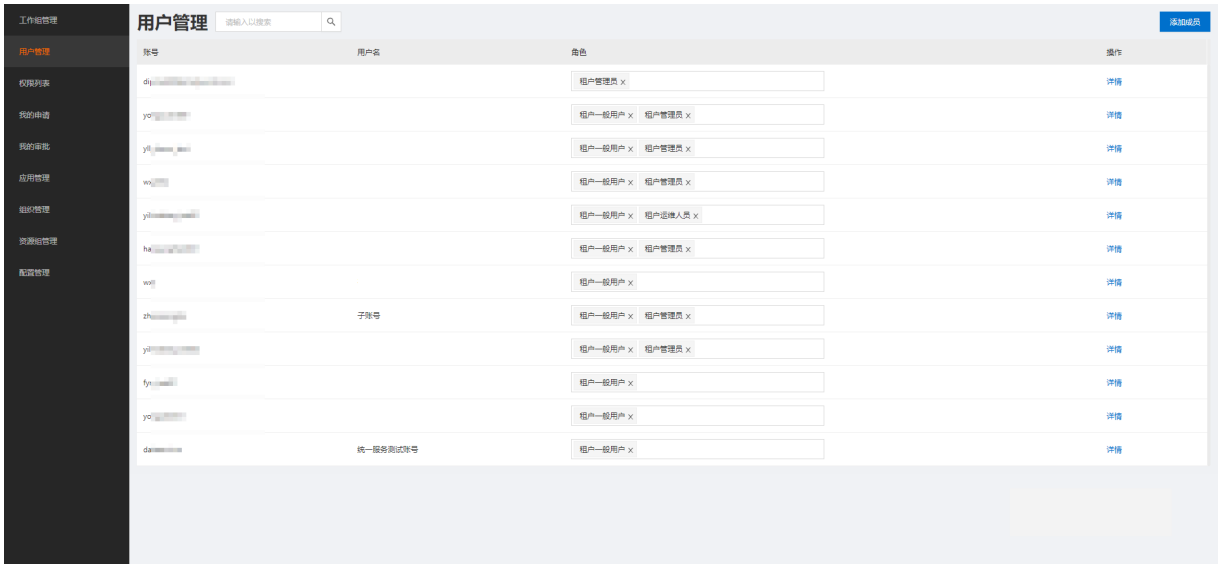
系统设置为用户提供了工作组、云计算资源、成员的创建及管理等能力，为数据资源平台的运行、使用提供了必要的基础保障。

工作组管理



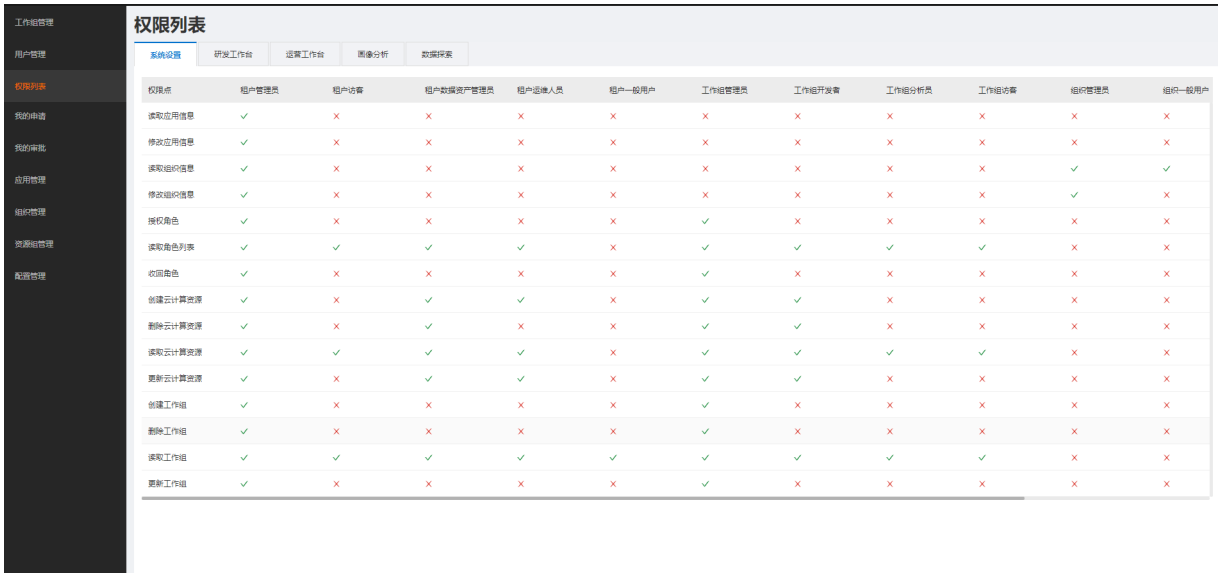
- **工作组管理：**支持工作组的创建、管理，以达到各工作组之间空间隔离的效果，并提供编辑、禁用、复制、删除等功能。
- **云计算资源管理：**支持多种类型的云计算资源的新建管理，并提供复制、编辑、检测、删除等功能。
- **成员管理：**支持对账号的相应工作组管理员、工作组开发者、工作组分析员进行配置，并支持移除成员、详情查看等功能。

用户管理



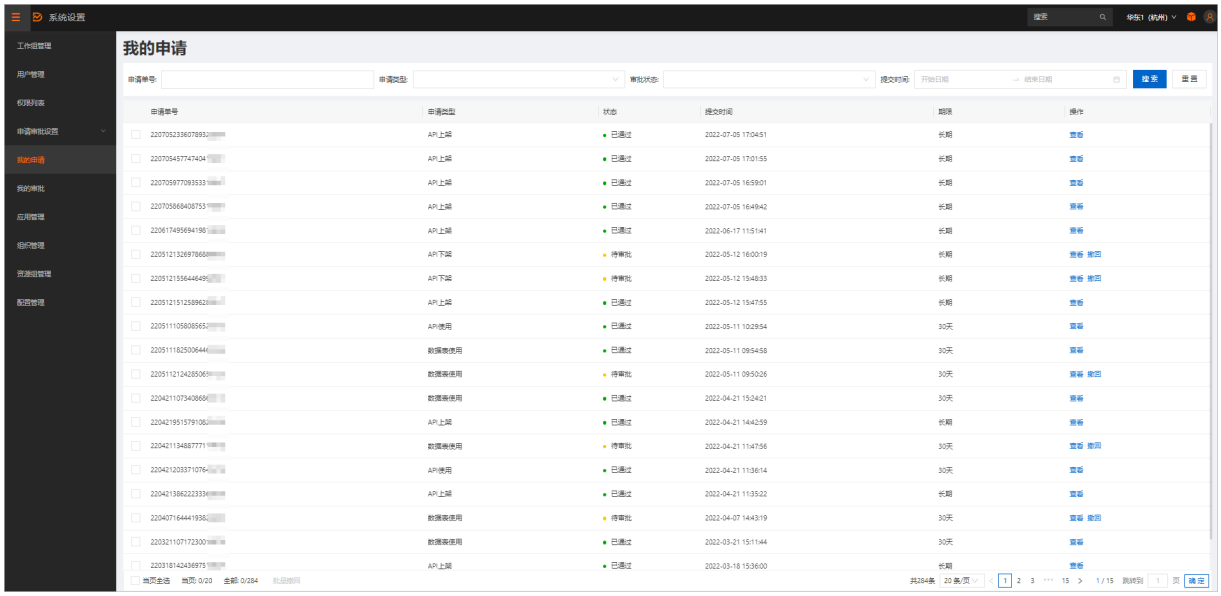
支持对人员账号相应的租户管理员、租户访客、租户资产管理员、租户运维、租户一般用户角色等进行添加和管理。

权限列表



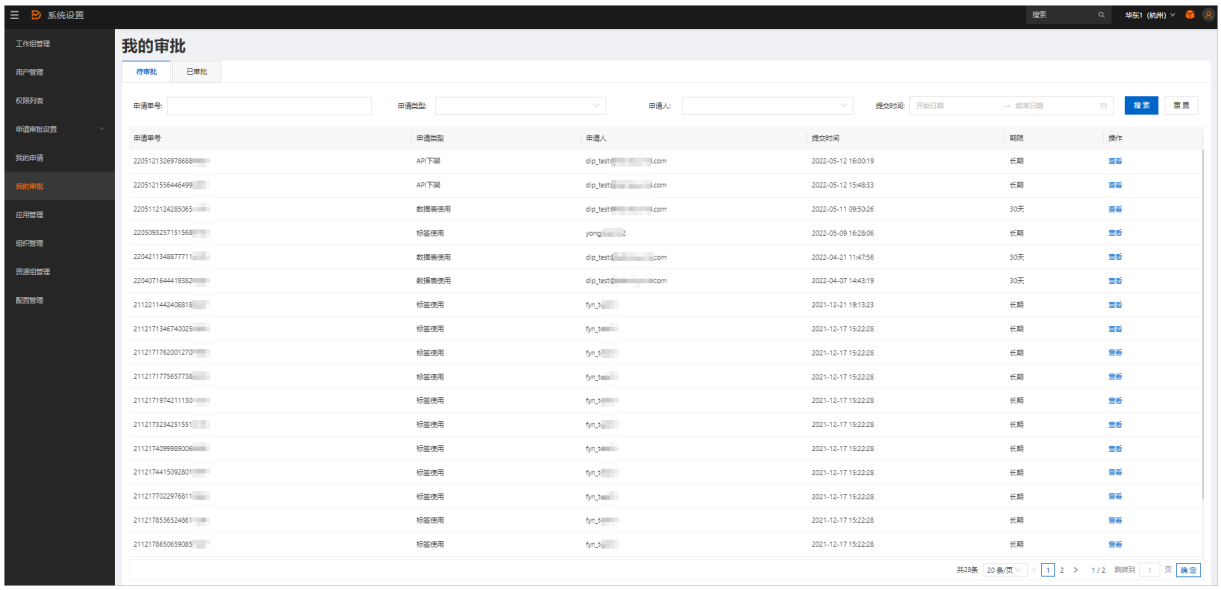
支持对系统设置、 数据同步、 数据标准、 数据建模、 智能标签、 统一服务、 通用资源、 质量评估、 任务运维、 画像分析、 运营工作台、 数据资产中心的角色及权限点内容的展示。

我的申请



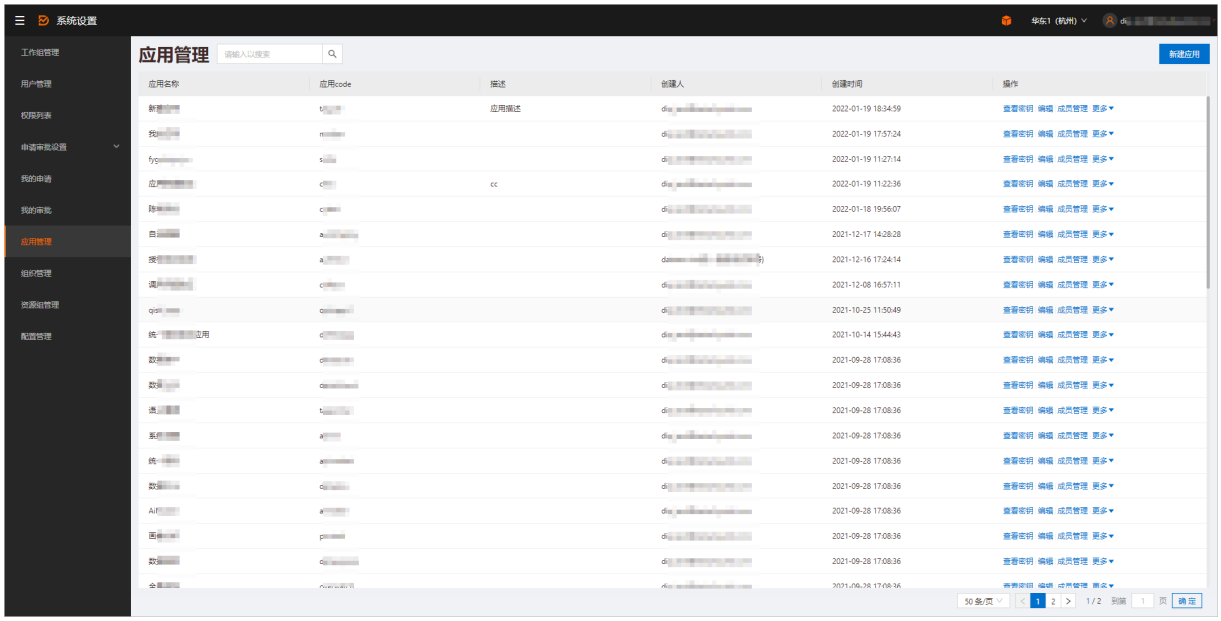
支持查看已经提交的申请单，包括审批中和审批完成的。用户可以查看自己申请单的详细信息和审批结果，也可以按审批状态筛选申请单和撤回申请。

我的审批



展示需要当前用户待操作的申请单。可以查看申请单的详情，包括申请的数据资源信息、申请理由、申请使用时长、审批结果等。支持操作员对申请单的审批动作，支持通过、拒绝等功能。支持查看已审批和未审批的申请单。

应用管理



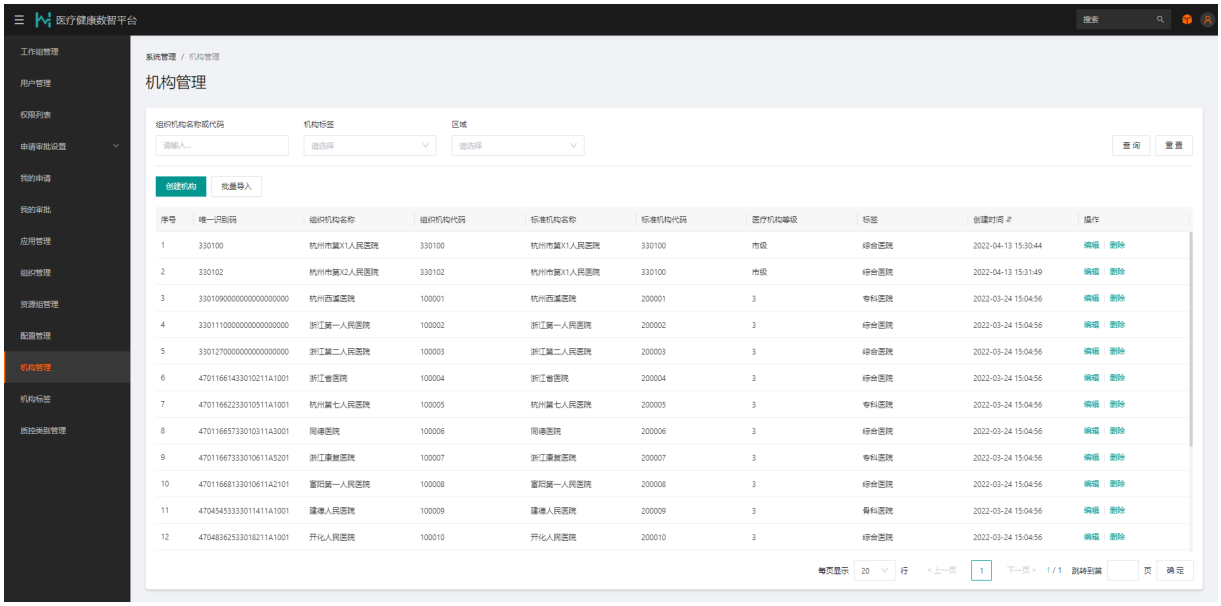
支持应用创建、管理功能，支持应用的成员管理功能，提供应用密钥管理功能、OpenAPI授权功能。

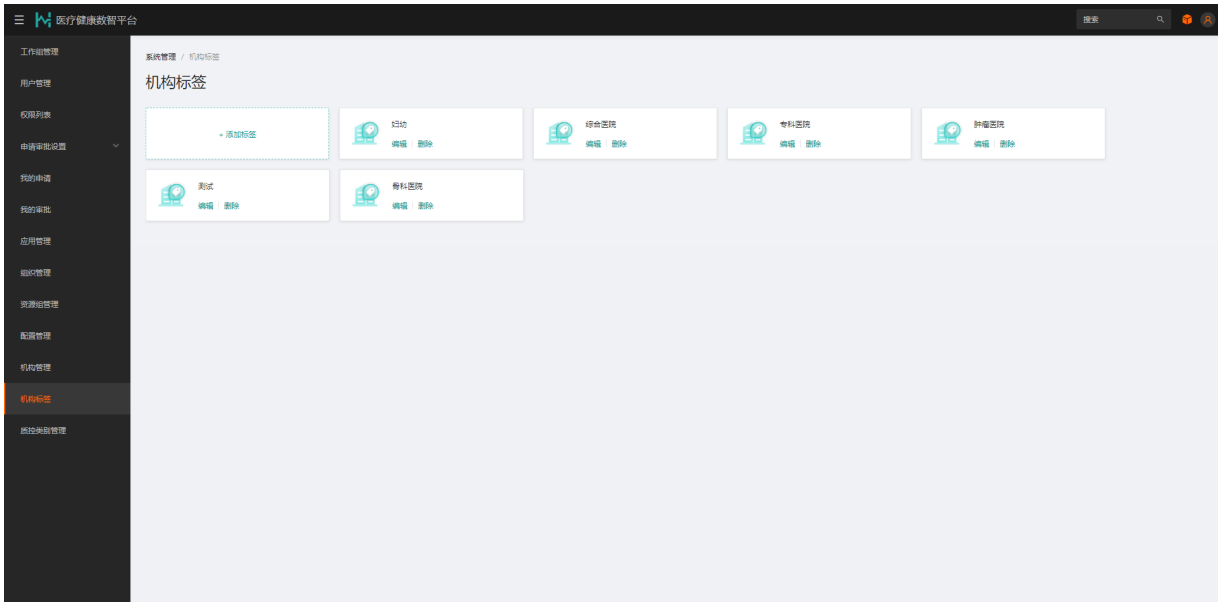
组织管理



支持树形组织的创建、管理，支持组织内成员的添加、管理功能。

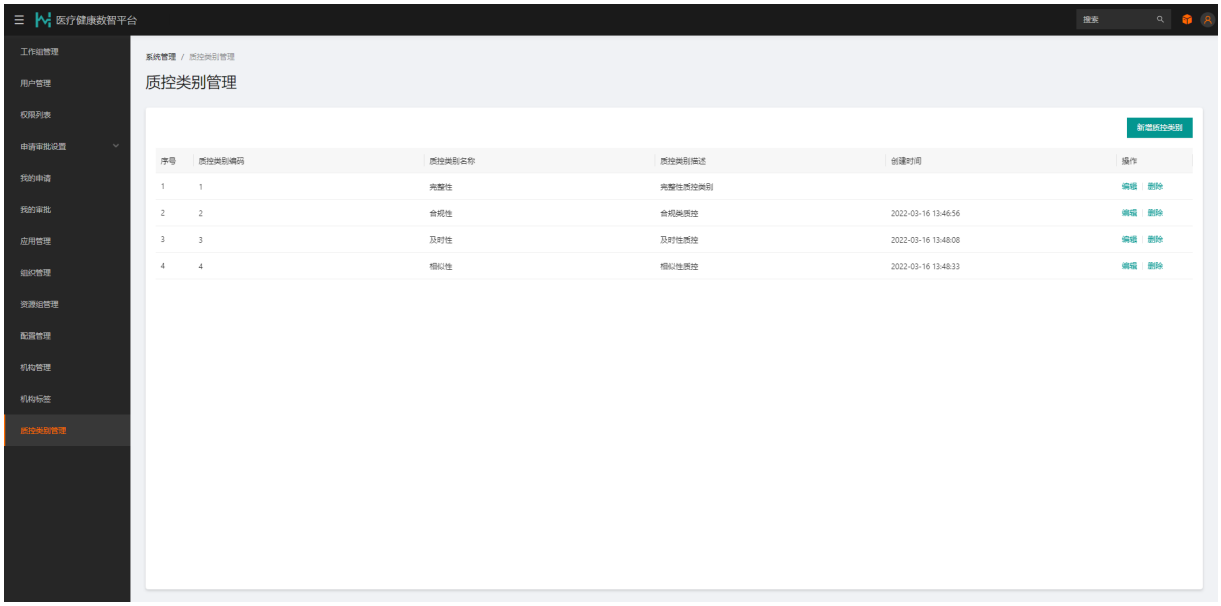
机构管理





- **机构管理**：用于维护业务工作台中使用业务应用的医疗机构，支持机构基本信息的增删改查等功能，支持批量导入机构。
- **机构标签**：用于维护机构的描述标签，可用作后续的机构分类与筛选中。

质控类别管理



质控类别管理，用于维护医疗业务质检规则的质控类别字典。

9.4. 应用场景

阿里云医疗健康数智平台应用场景。

卫健大数据中心及智能化应用解决方案（全民健康信息平台应用）

一站式医疗数据治理、数据资产管理以及数据服务等相关功能，利用大数据技术实现全区域医疗健康信息的互联互通。依托高可用数据在传染病防治、慢性病人群健康管理等领域为卫健委提供智能化管理工具。

医院智慧中台数据解决方案

建立医院标准数据模型、数据元、数据定义，实现医院全量数据集成、持续数据治理，规范医院数据需求服务，助力医院实现“内部驱动、及时响应、自主开发”的数据管理模式。

医保数据中台及智能化应用解决方案

高可用医保数据中台产品及数据治理服务，内嵌各类知识图谱、通用算法和医保行业关算法及医保模拟仿真沙盘。为各业务单元提供快速的数据分析、挖掘、建模支持，输出专业化算法和应用，支撑医保上层各应用系统及应用服务。

9.5. 基本概念

本文介绍数据资源平台的基本概念。

系统设置

租户：在数据资源平台产品中，租户为最高级权限隔离的命名空间，一个租户下可以有若干账号，不同账号可以分配不用的角色，数据可以在租户内共享，不同租户之间数据完全隔离。

工作组：数据资源平台中用于数据隔离的命名空间，通常为一个用户组、一个项目或一个应用。工作组下包含云计算资源、工作组成员。工作组内的数据通常为工作组内全部成员可见，研发工作台的数据权限以工作组为单元相互隔离。

数据同步

数据源：是指数据库应用程序所使用的数据库或者数据库服务器，在数据资源平台的数据同步中可以将配置好的数据源端数据同步到目标端。

多源异构数据：指的是多种不同结构的数据源。

数据模板：用于对非结构化数据的解释，当创建数据同步任务时，且源端数据类型为消息中间件时，通过数据模板对其进行解释。

Oracle CDC：实时数据同步模式，对源表进行INSERT、UPDATE或DELETE等操作的同时可以提取数据，变化的数据被保存在数据库的变化表中，通过订阅变化数据的方式实现数据的秒级同步。

MySQL Binlog：实时数据同步模式，通过同步并解析记录所有数据库表结构变更（例如CREATE、ALTER TABLE...）以及表数据修改（INSERT、UPDATE、DELETE...）的二进制日志，实现数据的秒级同步。

SQLServer CDC：实时数据同步模式，开启CDC的源表在插入、更新和删除活动时会插入数据到日志表中。CDC通过捕获进程将变更数据捕获到变更表中，通过CDC提供的查询函数，实现数据的秒级同步。

数据探查

内置算法：内置多种数据统计分析算法，根据不同字段类型采用不同的计算统计规则。

数据标准

数据标准：是对含义相同但字段名称不同的数据进行统一规范管理的数据准则，数据标准可定义逻辑表的命名规范、字段的取值范围、度量单位等内容。

数据元：表示一个数据的最小信息项，包含基本的描述、值域范围和约束规则。

生命周期：指的是表的生命周期，指表（分区）数据从最后一次更新的时间算起，在经过指定的时间后没有变动，则此表（分区）将被数据库（例如：MaxCompute）自动回收。这个指定的时间就是生命周期，生命周期的单位一般是天。

元素的长度：数据元的长度。

固定值：是指数据质量运行的一行一列结果与期望的固定值进行比较，根据比较条件判断质量结果执行是否通过。

趋势：指的是质量规则设定的模式，分为固定值、固定值和固定值比较，以及固定值波动率比较。

数据建模

数据模型：是数据特征的抽象，它从抽象层次上描述了系统的静态特征、动态行为和约束条件，为数据库系统的信息表示与操作提供一个抽象的框架，数据资源平台支持逻辑模型和物理模型。

- **逻辑模型：**反映的是系统分析设计人员对数据存储的观点，是对概念数据模型进一步的分解和细化；数据资源平台中可通过手动、DDL语句、EXCEL导入、云计算资源扫描的方式来实现逻辑模型创建。
- **物理模型：**是对真实数据库的描述；数据资源平台中通过将已创建好的逻辑表物理化到数据库来实现物理模型的创建。
- **质量规则：**是指在特定业务环境下，用户定义数据符合使用目的一组定量或定性的规定要求；在数据资源平台中，可以通过对表级与字段级的数据质量任务运行，真实反映数据接入的唯一性、准确性、规范性、一致性、时效性、完整性。

数据开发

场景：场景定义数据输出的方式，一个场景包含一种或多种数据输出的方式。

计算资源：可以执行计算节点的资源。使用“场景加工”时并不会单独执行一个计算节点任务，需要把计算节点组织成流程执行，所以需要计算资源支持流程的批量执行。

存储资源：可以存储数据的资源，包括离线存储、在线存储。

计算节点：场景加工中的计算节点由五要素组成：计算资源、代码、输入数据、输出数据和参数。代码可在计算资源上执行（代码是否需要编译由计算资源决定）；输入数据、输出数据由逻辑表定义；参数用于控制计算节点的执行。在使计算用节点配置任务流程时，流程中连接上下游的连接线表示计算节点的输入和输出数据，而节点参数需要通过单独的界面进行配置。

节点任务流程：通过有向无环的流程图（DAG）描述，定义了一段处理逻辑。节点任务流程中所有计算节点的输入、输出数据都是具体的数据表、文件和数据流。

开发场景与生产场景：每个工作区都有一个开发场景和一个生产场景，它们是两套独立的资源配置。数据开发工程师只能在开发环境中注册计算节点，配置节点任务流程，并进行开发。开发通过后要经过上线操作才能把开发环境中的节点任务流程发布到生产环境中。任何人都不能编辑生产环境中的节点流程，但可以修改参数，所有修改都会记录在日志中。

智能标签

物理表：是指具体某个数据源中的一张表。

主表：在数据库中建立的表格，其中存在主键（primary key）用于与其它表相关联，并且作为在主表中的唯一性标识，例如：银行客户信息表（证件号码、姓名、出生日期、电话号码、卡号），信用卡消费信息表（卡号、卡额度、办卡渠道），其中信用卡消费信息表中的卡号为银行客户信息表的外键，其与信用卡消费信息表中卡号属性相关联，因此，信用卡消费信息表为主表，银行客户信息表为从表。

加速表：是指加速到目标云计算资源而生成的物理表。

标签模型：是围绕实体（Object）、关系（Link）、标签（Tag）三大元素对分布在不同数据库中的数据进行网络化的建模方式。

- 实体用于描述某个客观的对象，如设备人员地址等，对应到物理数据表上一般就是属性表，有一个主键来代表每一个对象，剩下的每一列就是标签即描述对象的属性。
- 关系是表示对象和对象之间的联系、事件、行为，一般对应到物理数据表上一般就是事实流水表。

标签：是指对用户这一对象的一个最小描述单元，代表着所描述对象某一个具体的客观事实的抽象表达，如属性（教育水平：职业高中、本科、专科、研究生），行为（成交金额、收藏次数、位置定位），或者是兴趣（对于多个关键词的偏好度），是一种以业务视角出发的数据建模方法，标签既可能是数值、也可能是枚举值，也可以是多个Key-Value组织的列，还可能是多字段组成的事实表（如对象、时间、谓语、宾语）。

标签工厂：是指在数据资源平台中对衍生标签进行生产、加工的功能模块。

值码表：数据库相关字段编码值的含义，例如：性别字段代码表的值【M】代表 【男】，【F】代表【女】。

任务运维

任务流：由数据同步节点、离线计算任务节点、流式计算任务节点、自定义计算节点等组成的数据加工或算法服务工作流。

业务类型：提交任务流的来源，包括数据同步、数据开发、标签加工、标签同步等。

任务类型：提交任务流的来源，包括数据同步、数据开发、标签加工、标签同步等。离线同步任务、离线计算任务的任务类型，主要分为手动、周期、补数据等。

节点类型：任务流中包含的同步节点、MaxCompute_SQL节、Flink_Vvp节点、自定义计算节点等在不同计算资源上运行的计算任务节点。

补数据计划：通常情况下，对于离线周期任务，当前序任务失败或任务第一次上线，需要通过补数据计划同步或处理当前业务日期之前数据。

画像分析

群体：分析对象上满足一定条件的单体的集合。

群体计算：同一个实体的群体和群体之间的交集、差集和并集计算，生成新群体。

资产运营

资产目录：是对数据资源平台元数据的有序组织，是记录数据体系的保障。数据类目是目录信息与服务、保障与支撑所组成的一个整体。针对资产的合理组织，需要对资产进行对应分类，如部门类目、主题类目、行业类目，可进行该类目下资产数据进行搜索、展现以及权限申请。

数据资产：数据资源平台中存在大量的数据表、标签、API等各类数据资产，数据管理者通过数据汇聚、数据治理、数据分析后，需要对整个平台数据进行统一管控，了解平台的核心数据资产，提供对应的数据资产管理规范。

统一服务

应用：是调用API服务时的身份。每个应用有一组AppKey和AppSecret，可以理解为账号密码，调用API的时候需要将AppKey做参数传入，AppSecret用于签名计算，计网关会校验这对密钥对您进行身份认证。

API：应用程序编程接口，是一些预先定义的函数，目的是提供应用程序与开发人员基于某软件或硬件得以访问一组例程的能力。

数据API：API的源头为数据表，可以是单表也可以是多表，请求后最终以类SQL的形式到数据库查询并返回符合条件的数据，将数据组装成报文格式返回。

应用API：API的源头为数据表，可以是单表也可以是多表，请求后最终以类SQL的形式到数据库查询并返回符合条件的数据，将数据组装成报文格式返回。

9.6. 使用限制

本文为您介绍数据资源平台产品各模块的使用限制。

系统设置

目前角色包括：租户管理员、租户访客、租户数据资产管理员、租户运维人员、工作组管理员、工作组开发者、工作组分析员、工作组访客、租户一般用户。

资产管理

系统目前支持：

- **数据表：**MaxCompute、DataHub、Hive云计算资源类型。
- **标签：**RDS MySQL、MaxCompute云计算资源类型。
- **API：**RDS MySQL、MaxCompute 、PostgreSQL、Table Store、DataHub、Object Storage Service、AnalyticDB PostgreSQL、Hologres云计算资源类型。

数据同步

- 目前系统支持的任务链路：

链路类型	数据源/目标端	Oracle	SQLSever	Postgres	MySQL	ADB PG	MaxCompute	DataHub	Hive	OSS	Kafka	HBase(Phoenix)	TableStore	Hologres(不支持关键字)	ADB MySQL	Neo4j(4.3)	ES	Greenplum	MINIO	POSTGIS PG	GANOS PG
离线链路	Oracle(10g, 11g)	✔	✔	✔	✔	✔	✔	✔	✔												
	SQLSever(2008, 2014, 2016, 2017)	✔	✔	✔	✔	✔	✔	✔	✔												
	Postgres(9.4)	✔	✔	✔	✔		✔		✔												
	MySQL(5.6, 5.7)	✔	✔	✔	✔	✔	✔	✔	✔						✔						
	HDFS(Hadoop 2.7.1, 2.5.0, 2.6.0)								✔												
	Hive(1.1.1, 1.2.2, Hadoop 2.7.1, 2.5.0, 2.6.0)				✔(API)		✔					✔(API)									
	FTP									✔											
	MaxCompute			✔	✔	✔	✔						✔	✔	✔	✔	✔			✔	
	ADB PG				✔	✔	✔								✔						
	MINIO					✔												✔			
	KUDU						✔														
	SDXPLUS PG																			✔	✔
	ARCSDE ORACLE																			✔	✔
	POSTGIS PG																			✔	✔
	GANOS PG																			✔	✔
	OSS			✔	✔	✔									✔						
实时链路	Oracle(10g, 11g)							✔			✔										
	SQLSever(2008, 2017)							✔			✔										
	MySQL(5.6, 5.7)							✔			✔										
	PUSH API																				
	Kafka (2.3, 2.4, 2.5)			✔		✔		✔											✔		
	RabbitMQ							✔													
	PostgresqlCDC			✔		✔															
	DM							✔													

> 说明
> - “√”代表支持，其他则为不支持。
> - 目标端支持版本与源端一致。

- 目前系统支持的批量任务链路：

链路类型	数据源/目标端	Oracle	SQLSever	Postgres	MySQL	ADB PG	MaxCompute	DataHub	ADB Mysql
离线链路	Oracle(10g, 11g)	✔	✔	✔	✔		✔		
离线链路	SQLSever(2008, 2014, 2016, 2017)	✔	✔	✔	✔		✔		
离线链路	MySQL(5.6, 5.7)	✔	✔	✔	✔	✔	✔		✔
离线链路	ADB PG				✔	✔	✔		✔
实时链路	Oracle CDC/LogMiner (10g, 11g)							✔	
实时链路	MySQL binlog(5.6, 5.7)				✔			✔	

> 说明
> - “√”代表支持，其他则为不支持。
> - 目标端支持版本与源端一致。

数据查询和自助分析

系统目前支持：MaxCompute、RDS MySQL、AnalyticDB PostgreSQL、Hologres、Hive云计算资源类型预览和探查。

逻辑表物理化

系统目前支持：MaxCompute、DataHub、Hive、ADB for PostgreSQL云计算资源类型。

场景

- **计算节点：**目前系统支持MaxCompute、ZergStandalon、ZergService、Hive、Blink、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。
- **读写数据：**目前系统支持MaxCompute、Hive、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。
- **数据同步：**目前系统支持MaxCompute、ApsaraDB for RDS、Hive、Table Store、PostgreSQL、AnalyticDB PostgreSQL、Hologres云计算资源类型。

智能标签

目前系统支持：RDS MYSQL、AnalyticDB PostgreSQL（包括2.0和3.0版本）、MaxCompute、Hologres、AnalyticDB MySQL、PostgreSQL云计算资源类型。

画像分析

目前系统支持：RDS MYSQL、AnalyticDB PostgreSQL（包括2.0和3.0版本）、MaxCompute、Hologres、AnalyticDB MySQL、PostgreSQL云计算资源类型。

统一服务

目前系统支持：RDS MySQL、MaxCompute 、PostgreSQL、Table Store、DataHub、Object Storage Service、AnalyticDB PostgreSQL、Hologres 云计算资源类型。

10.感知数据工厂产品简介

感知数据工厂通过数据治理及轨迹、关系等多种融合大数据分析算法，深度挖掘感知数据价值。包括人/车/场所等多种实体、人车码轨迹关联、落脚点、人人/人车/人地等多种关系、昼伏夜出/疑似无业等多种标签。实现人员的轨迹全关联、关系全掌握、特征全刻画。

10.1. 产品介绍

1.1 产品背景

随着人民生活水平提高，对公共安全、城市治理的要求也不断提高。通过视觉智能和大数据提升警务工作效率成为新时代公安发展的重要课题。在国家大数据发展战略背景下，公安部提出了以视图大数据助推公安工作质量变革、效率变革、动力变革的重要目标。

随着雪亮和天网工程建设，每年都会积累几十亿甚至上百亿的人像数据。但目前视图数据还处在2.0智能化发展初期，仅用于布控报警、轨迹分析和基本的研判。依然面临重点人员管控难、人员客观鲜活画像基本处于空白、全目标融合分析处于空白等问题。离散的人像数据价值有限，缺乏技术手段深度挖掘这座金矿。

阿里云城市安全视图3.0数据化解决方案，实现从离散人像数据聚类归档、到人像聚档数据、车辆数据等视图大数据分析、再到智能化业务应用的全链路升级。

1.2 产品定义

感知数据工厂通过数据治理及轨迹、关系等多种融合大数据分析算法，深度挖掘感知数据价值。包括人/车/场所等多种实体、人车码轨迹关联、落脚点、人人/人车/人地等多种关系、昼伏夜出/疑似无业等多种标签。实现人员的轨迹全关联、关系全掌握、特征全刻画。

感知数据工厂变革了视图数据利用模式，开启了视图大数据时代，带动新一轮感知智能化应用创新与升级。为城市安全、城市治理等业务提供独特的能力支撑。

感知数据工厂提供标准的数据接入格式，对外提供通用的查询、分析和订阅服务。

1.3 用户群体

客户群体：

地市、区县的公安分局，建设方多为科信。

使用客户：

目标用户：科信、isv、交付运维人员。

10.2. 基本概念

实体：客观存在并可相互区别的一类事物，如人员、车辆、场所、组织、设备等。

要素：指客观世界能标识一个实体的属性或属性组合，是构成事物必不可少的因素。如人的身份证号、人脸、人体等。

全息归档率：自动关联合并的轨迹要素ID数/所有轨迹要素ID数*100%，不同质量数据对应值不同。全息归档率衡量人车码关联的覆盖度情况，指标越高越好。

要素关联率：=要素关联对数*2/所有要素ID总数*100%，不同质量数据对应值不同。要素关联率衡量要素间关联的密度，即每个要素与其他要素的平均关联个数。越接近真实情况越好，如城市中一个人关联的平均车辆数量一般在0.3-0.5辆。

伴随分析指标：N次伴随分析查询，m次在top 10结果中有命中伴随人员，伴随成功发现率=m/n*100%。伴随成功发现率衡量伴随分析实战效果，指标越高越好。

聚集分析指标：n次聚集分析查询，m次在top 10结果中有命中聚集人员，聚集成功发现率=m/n*100%。聚集成功发现率衡量聚集分析实战效果，指标越高越好。

采集设备空间重叠率：统计每类采集设备之间在空间上的交集。1个采集设备在1km范围内能找到至少1个其他类型设备则算重叠，重叠率=重叠的设备数/设备总数*100%。重叠率对人车码关联、关系挖掘影响较大。

关系覆盖度：=存在某种关系的实体ID数/所有实体ID数*100%，不同质量数据对应值不同。关系覆盖度衡量某种关系的普遍性情况，越接近真实情况越好，如城市中有朋友（人人关系）的人数量一般比有车（人车）的人数量要多。

关系密度：=某种关系对数*2/所有实体ID总数*100%，不同质量数据对应值不同。关系密度即每个实体拥有某种关系的个数。越接近真实情况越好，如城市中一个人的平均同住所人数一般在2-4人。

关系准确率：某种关系正确的数量/该关系总数量*100%，不同质量数据对应值不同。关系准确率衡量某种关系的精度情况，越高越好。

标签准确率：某种标签正确的数量/该标签总数量*100%，不同质量数据对应值不同。标签准确率衡量某种标签的精度情况，越高越好。

10.3. 应用场景

产品可针对城市范围内人、车、场所等进行大数据分析。分析结果可应用于各类城市安全及城市治理场景。
注：以下场景需由不同SaaS应用实现，底层能力则由产品赋能SaaS应用。

城市安全视觉智能解决方案聚焦在城市安全业务领域，以公安应用场景为主，其他安全相关应用场景为辅。可直接应用在轨迹检索、规律分析、区域排查、客流检测、人物关系分析等共性场景，也可间接应用在相关专业场景，isv也可基于开放的服务接口开发满足专业场景的上层应用。其中相关的专业场景详细描述如下：

3.1 关系分析

3.1.1 场景背景

在城市安全建设进程中，公安已通过广泛的卡口相机基础建设获取了丰富的视频资源和人车结构化数据，充分合理的利用已有的人车数据，分析人员关系，掌握治安安全隐患，排除社会危险因素成为势在必行的趋势。

3.1.2 需求分析

公安急需利用已有的视频数据和结构化能力，自动摸排区域内的陌生人员，出行特征敏感人员，自动分析嫌疑人轨迹、落脚点、自动对嫌疑人人际关系进行梳理。节省民警办案时间，提升民警的办事效率。

3.1.3 功能应对

l 陌生人摸排：通过区域碰撞、首次入城/昼伏夜出等标签摸排特定案件嫌疑人。

l 嫌疑人分析：查看嫌疑人轨迹、落脚点，可进行蹲守或抓捕。

l 关系研判：对嫌疑人关系进行研判，发现同伙及组织。

3.2 团伙分析

3.2.1 场景背景

实际破案或治安管理过程中，公安需要根据部分嫌疑人或重点人员的社交关系进行团伙成员分析或管控，实时、高效的分析社会团伙对公安破案和治安管控具有重要意义。

3.2.2 需求分析

可以快速分析出动态人员的团伙信息；可以根据重点人底库，分析出可能的嫌疑人团伙；可以根据人员底库信息分析家庭等社会组织；

3.2.3 功能应对

l 疑似涉黄/赌/毒/传销/盗窃等团伙；

l 家庭。

3.3 疫情防控

3.3.1 场景背景

当前国内疫情持续不断，暂时做不到完全控制住疫情传染。当区域内疫情发生的时候，第一时间找到密接人员和时空伴随人员成为了防止疫情扩散最重要的一步。传统的人工排查方式，耗时耗力，容易漏查，而且大大增加感染和二次传染风险，急需找到一种实时、远程、自动的疫情防控排查方式。

3.3.2 需求分析

可以通过轨迹碰撞找出与特定人员时空接近的人;可以通过人员关系网找出与特定人员亲密度较高的人。

3.3.3 功能应对

l 时空伴随人员；

l 密接人员排查。

3.4 侦察破案

3.4.1 场景背景

目前公安的视频监控系统，经历了过去10多年的发展，基本实现了数字化和集中化。随着大量新增的摄像头，视频数据得到了极大的丰富，成为公安破案的重要手段。公安系统也基本建设了统一的视频管理平台，对视频文件的查询，保存，使用提供了便捷的手段。

3.4.2 需求分析

经过多年的积累和发展，公安破案已形成一套成熟流程，但每一个环节都还存在诸多问题和痛点，影响破案效率和成功率。主要痛点如下：

l 线索的突破，从0到1，最耗时耗力，目前纯人工排查，无有效手段；

l 线索的追踪，需要大范围搜寻，仍主要靠人工，现有手段难以覆盖全域；

l 警力不足（中美万人拥有警察数量比为13：34），破案率低，小案破案率不20%；线索生成、线索排查、布控抓捕耗费90%时间，成为破案瓶颈。

3.4.3 功能应对

l 视算研判：在全域视频流中，实现对案件线索（图片/特征）的秒级搜索，系统智能推荐符合线索的疑似目标，经民警确认后，复原嫌疑人轨迹。再结合人脸识别、数据碰撞等手段，快速定位出嫌疑人身份和落脚点。

l 监控预警：可根据业务需求选择不同布控规则引擎，对嫌疑人进行多维布控，一旦触网，自动发送告警并触发轨迹与实时位置记录，提升布控及抓捕效率。

l 案件中心：以案件为警务室研判核心，一切的操作都围绕着一起案件展开。一案一档的推出将有助于民警提高对于案件的管理，其中包含案件信息、嫌疑人信息、以及嫌疑人详细的布控信息、轨迹归档、且支持对于案件的报告导出操作。可在视算研判时一键归档到案件中心。

3.5 小区人口管理

3.5.1 场景背景

随着社会的发展，政府对基层的人口管理要求是“一标三实”，公安在进行小区人口管理时存在工作量大、效率低、预算占比高等痛点，面对客户痛点，人口管理工作需要充分利用视频AI及数据AI技术，实时追踪人员轨迹，及时管控，提升工作效率。

3.5.2 需求分析

小区人口管理需要对小区内的人员成分进行管理，包括小区的流动人口、重点人员和进入小区的外来工作人员等。小区人口通过与静态人口库进行档案碰撞，进行关联，实现小区人口实名化。针对小区出入口记录的聚档，可用于小区常住人口管理、流动人口管理、迁入迁出管理。通过人员实体属性、关系、标签等，分析小区人员构成，提高管理水平。

3.5.3 功能应对

l 小区管理：建立需要管理的小区，小区实有人口存放在对应的小区下面，方便针对性管理。在创建管理小区时需要关联小区相应的摄像头，通过前端摄像头采集数据，然后进行分析，可实现动态人口库和静态人口库的关联。

l 小区人口实名化：小区人口通过与静态人口库进行档案碰撞，实现小区人口实名化。不能自动关联成功的小区人口支持手动录入实名信息，实现小区人口实名化。

l 流动人口管理：流动人口的迁入和迁出，根据算法模型进行规则计算。对于动态人口库人员，如若摄像头拍到该人员连续出现或未出现15天，即判定为迁入或迁出人口，并且推荐到系统界面。

10.4. 功能介绍

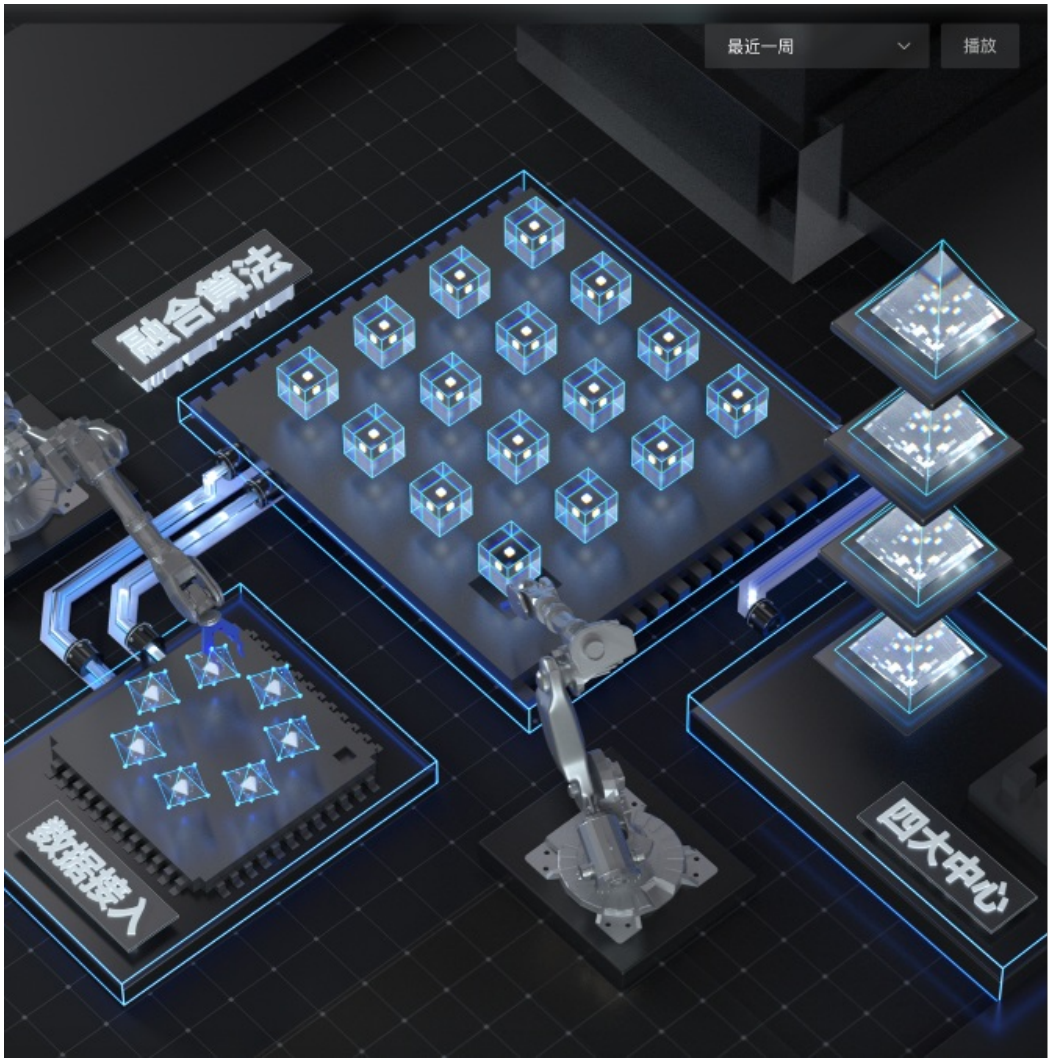
5.1 管理及服务层

5.1.1 资产大盘

资产大盘作为平台的门户，概括展示了数据平台处理数据的过程以及平台数据使用态势。

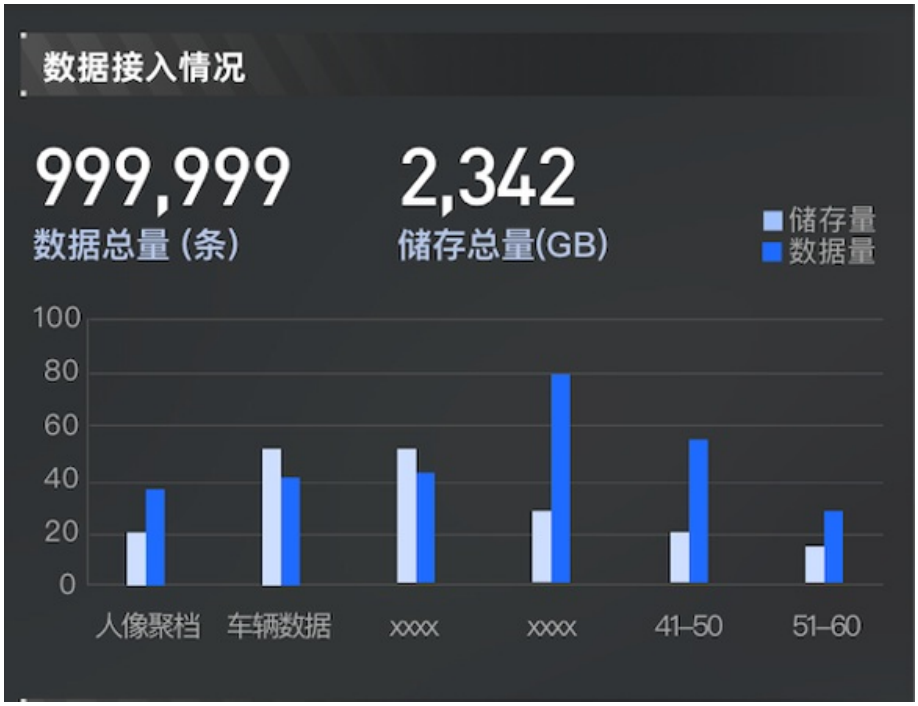
资产大盘包含全链路工厂、数据接入态势、数据加工态势、数据产出态势、数据服务态势、四大中心态势、数据态势总览九大模块。

5.1.1.1 全链路工厂



全链路工厂生动的展示了数据控制台整个平台处理数据的过程，从数据接入、数据加工和数据资产三层间的相互关系呈现数据处理过程。点击每一个节点，都可以展示与该节点相关的加工链路，以及该节点的概况信息。

5.1.1.2 数据接入态势



数据接入层是人像聚档、车辆识别等主要接入的数据类型，通过数据和图像展示不同类型数据接入数量的态势，同时数据支持按最近30天、最近15天、最近一周的维度查看。

5.1.1.3 数据加工态势



数据加工层是实体、轨迹、关系、标签四大类加工算法和数据模型，通过数据和图像展示计算次数和耗时指标，同时数据支持按最近30天、最近15天、最近一周的维度查看。

5.1.1.4 数据产出态势

数据加工层是实体、轨迹、关系、标签四大类加工算法和数据模型，通过数据和图像展示计算次数和耗时指标，同时数据支持按最近30天、最近15天、最近一周的维度查看。

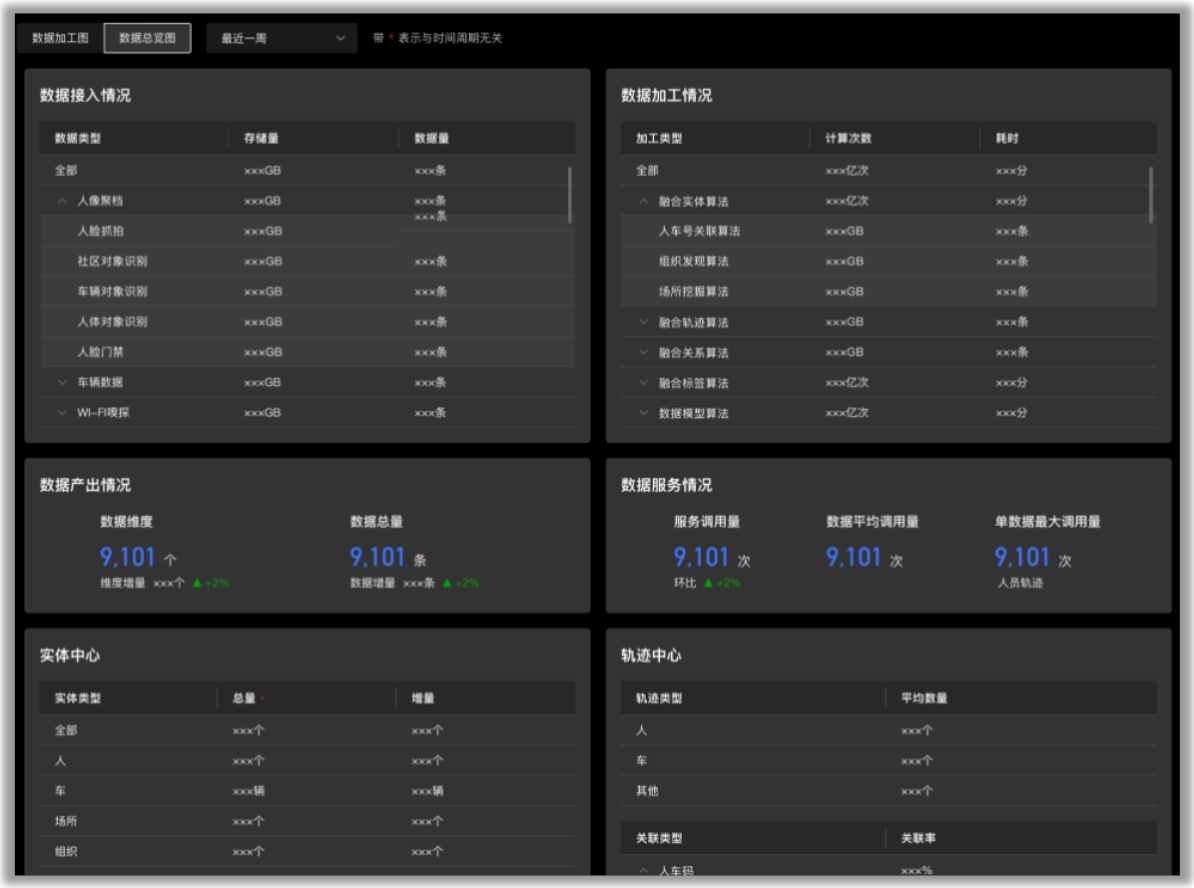
5.1.1.5 数据服务情况

数据服务是对数据加工后服务调用的统计，统计维度包括调用总量、平均调用次数、最大调用量等，同时数据支持按最近30天、最近15天、最近一周的维度查看。

5.1.1.6 四大中心态势

四大中心态势是数据加工分析后的产出指标分析，通过数据和图像分析包括指标有实体中心的数量、关系中心的数量对、轨迹中心的轨迹数量、标签中心的标签数量。同时数据支持按最近30天、最近15天、最近一周的维度查看。

5.1.1.7 数据总览图



数据总览图将数据接入、数据加工、数据产出到数据服务的整体情况进行呈现。可选择近一周、近一月等周期查看。

数据接入情况：按两级分类进行展现，从存储量、数据量两个维度直观了解系统接入的数据类型及数据规模。

数据加工情况：将核心加工算法及数据模型进行展示，从计算次数、计算耗时两个维度直观了解数据加工的运行情况。

数据产出情况：主要从数据维度和数据总量两个方面直观展现整体产出规模。再分别展示实体、轨迹、关系和标签四大中心各自产出资产情况。数据维度指人脸、人体、车牌号、车品牌等可区分的数据类型；数据总量则是每个数据类型的数据量之和。

数据服务情况：从服务调用量、数据平均调用量、单数据最大调用量直观展现数据赋能情况。

5.1.2 资产分类

平台接入数据，形成庞大的大数据资产，为了能够有效的对数据资产进行态势了解、数据关注、趋势关注，平台以实体中心、轨迹中心、关系中心、标签中心四大中心为基础，构建资产分类，可视化平台拥有的数据资产，展现数据资产在归档、轨迹研判、关系研判、标签研判等多方面的能力，为上层提供有力支撑。

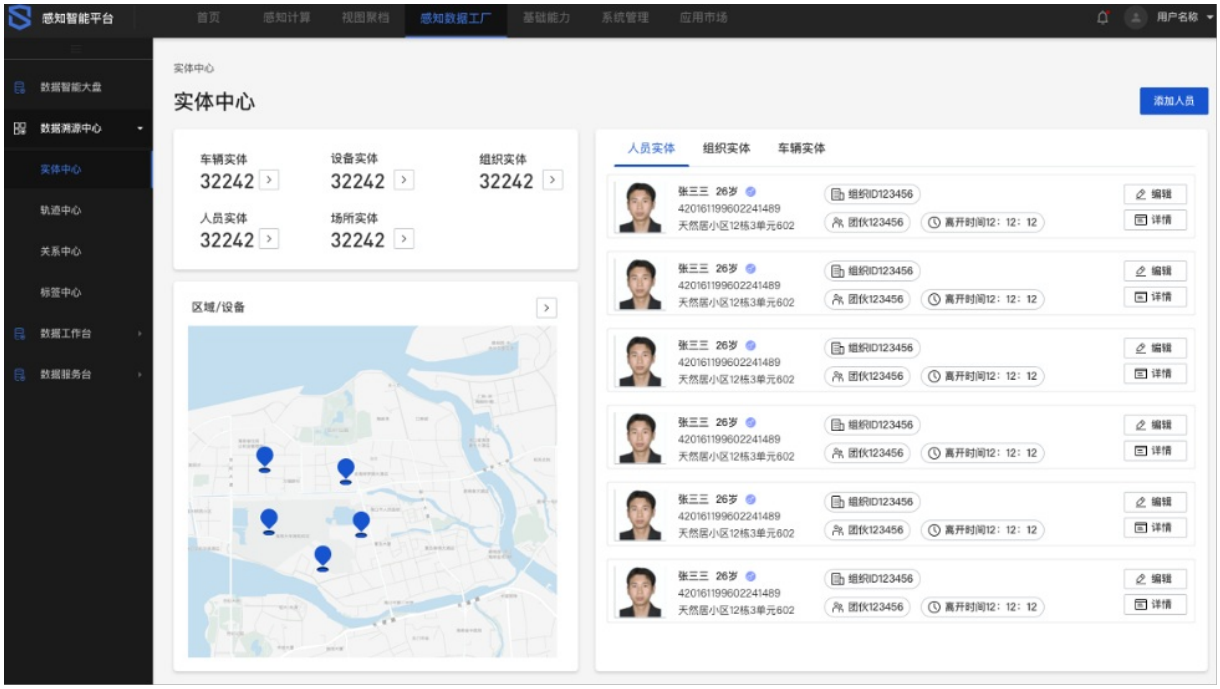
5.1.2.1 实体中心

公安的业务场景主要包括侦查打击、治安防控、指挥调度、活动安保等，业务场景的工作对象可以分为5大类：人、地、事、场所、组织。不同的业务场景下，即针对这5大类工作对象或者称为要素，开展打防管控服。而对这5大类对象，形成实时、鲜活、饱满的画像刻画，是当今时代，大数据警务或警务智能化的一个先决条件。故而我们通过视频大数据，以人、车、场所、设备、组织为主体，通过对视频画面中的时间、地点、要素、关联关系进行深度挖掘，形成实体中心，构建鲜活而丰富的基础信息，辅助下一步的精准治理，是公安各业务警种的通用需求。

5.1.2.1.1 人员实体

1) 人员实体概览

a) 人员实体数据资产



人员实体数据资产展示平台人员实体的数据资产情况，从人员实体数量等多个维度展示人员实体数据资产的情况，对人员实体的数据量有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 人员

默认展示全部人员档案卡片，卡片包含人员封面图（人脸照片）、人员ID、姓名、身份证、居住地信息。人员实体档案按照归档时间倒序排列。点击能够进入“一人一档”。

c) 人员实体图搜

针对人员实体能够上传已有人员图片进行图搜。搜索结果为人员实体档案列表，搜索出来的列表人员档案为上传图片相似度在一定规定值以上。

2) 人员实体详情信息

a) 人员实体图片

人员实体详情展示人员人脸人体各三张图。点击能够放大查看人员的照片详情，能够左右切换图片。

b) 人员实体基础信息

人员实体基础信息来源于增加人员时填写的基础信息或通过人员底库撞库获得的数据。基础数据包含：姓名、身份证、性别、年龄、登记住址、职业

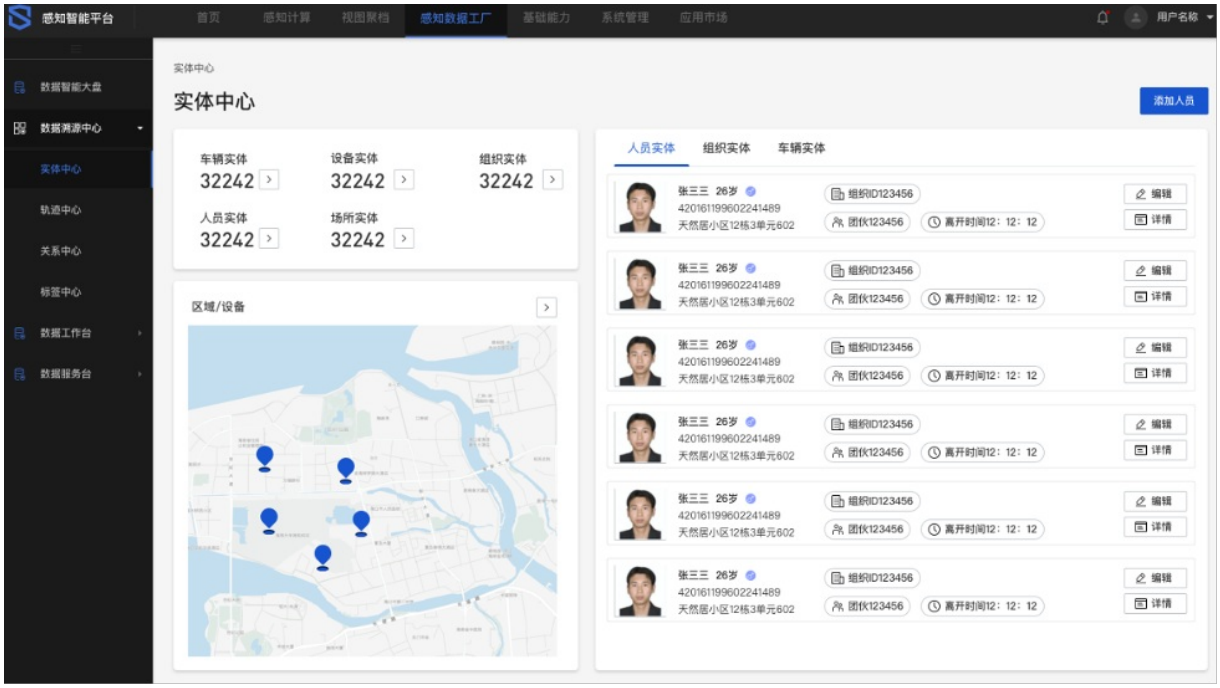
c) 人员实体归档信息

人员实体归档信息来源于视图聚档后，根据轨迹分析分析出来的数据，随着每天的聚档，会进行自动变化。归档信息包含：发型、脸型、体型、身高、是否有胡子、是否有眼镜、建档时间、最早最晚出现时间等。

5.1.2.1.2 车辆实体

1) 车辆实体概览

a) 车辆实体数据资产



展示平台车辆实体的数据资产情况，从车辆实体数量等多个维度展示车辆实体数据资产的情况，对车辆实体的数据量有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 车辆实体数据列表

默认展示全部车辆档案卡片，卡片包含车辆封面图（车辆照片）、车辆ID、车牌号。车辆实体档案按照归档时间倒序排列。

2) 车辆实体详情信息

a) 车辆实体图片

车辆实体详情展示车辆三张照片，点击能够放大查看车辆的照片详情，能够左右切换图片。

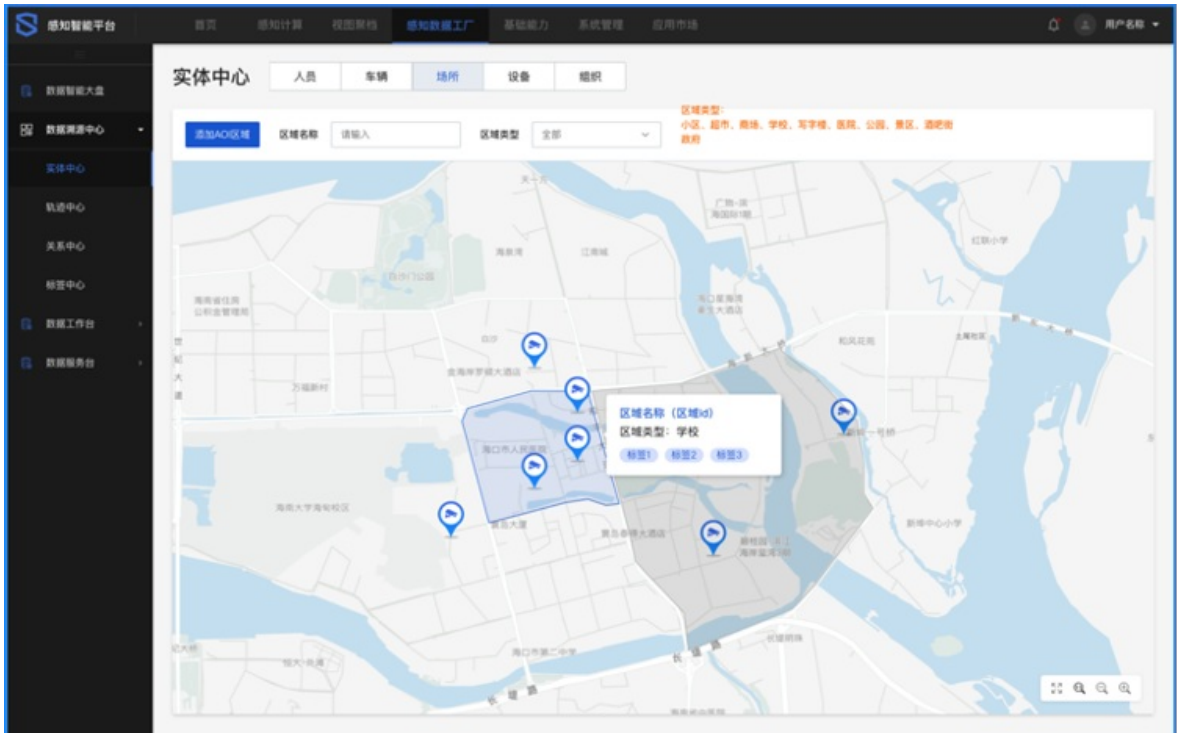
b) 车辆实体归档信息

车辆实体归档信息来源于视图聚档后，根据轨迹分析分析出来的数据，随着每天的聚档，会进行自动变化。归档信息包含：车牌号、车辆ID、车辆颜色、车牌颜色、最早最晚出现时间等。

5.1.2.1.3 场所实体

1) 场所实体概览

a) 场所实体区域地图总览



首次进入时划定关注的AOI区域，后续进入展示全区域地图，同时高亮场所中划定好的AOI区域，以及区域内部的摄像头。

b) 场所实体区域信息查看

针对高亮区域鼠标悬浮时能够查看信息，包括区域名称及区域标签。

2) 场所实体详情信息

a) 场所实体基础信息

展现场所的实体基础信息，包括区域名称、区域类型等。同时会在地图上单独高亮展示该场所。

b) 场所实体标签信息

展现该场所的标签信息，该场所的标签来源有归档后的数据分析结果，也有来源于标签信息维护。针对场所实体的标签信息，能够单独进行维护，去掉不正确的标签，增加未添加的标签。

c) 场所实体重点人员信息

该区域内出现的重点人员按照出现频次进行排序，展示信息：出现频次、性别、重点人员分类。

点击展开查看全部的重点人员信息，点击重点人员进入一人一档详情页。

d) 场所实体区域人员信息

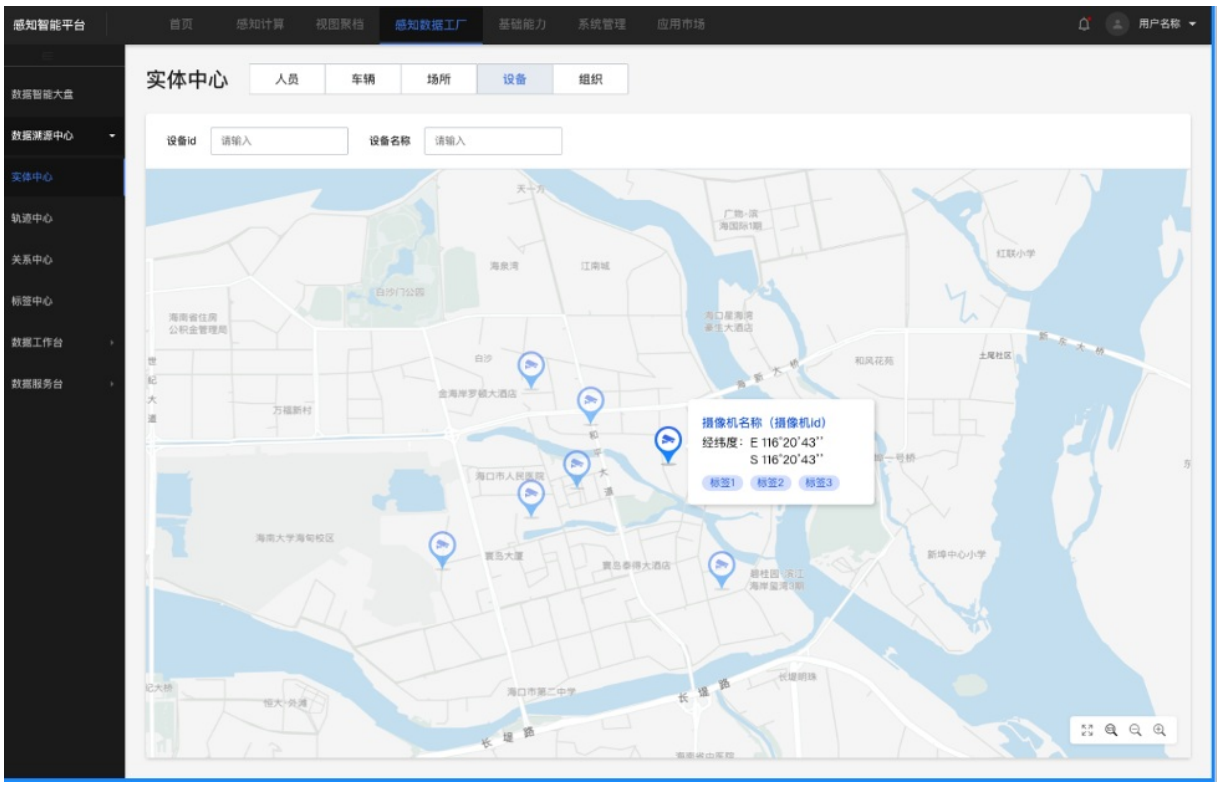
默认展示近一天的出现人员，可以进行时间切换“近一天”、“近一周”、“近一月”。

切换结果按照用户出现频次由高到低排序，鼠标移动至照片展示出现频次性别等相关信息，点击头像进行一人一档详情页。

5.1.2.1.4 设备实体

1) 设备实体概览

a) 设备实体区域地图总览



设备总览支持查看区域内的全部摄像头点位，支持查看点位所属区域，设备状态、能够查看运行中的和停用的设备。。

b) 实体设备信息查看

针对单个摄像头鼠标悬浮时能够查看信息，包括摄像头ID及摄像头名称、摄像头经纬度等，信息来源于摄像头本身的自带信息。

2) 设备详情信息

a) 设备实体基础信息

展示单个摄像头的基本信息，基本信息包括摄像头的ID、名称、所在区域、角度等等信息，按照cpp库中摄像机基本信息来展示。

b) 设备实体标签信息

展示单个摄像头的标签信息，针对设备实体的标签信息，能够单独进行维护，去掉不正确的标签，增加未添加的标签。

c) 多设备关系画像

以本设备为中心，周围200m内的摄像机做关联关系展示，如果关联关系错误可以进行“删除关系”操作。

点击摄像机之间的关联即可展示两者之间的信息

5.1.2.1.5 组织实体

1) 组织实体概览

a) 组织实体数据资产



展示平台组织实体的数据资产情况，从组织实体数量等多个维度展示组织实体数据资产的情况，对组织实体的数据量有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 组织实体团队列表

以单一组织为维度，展示组织团队列表，默认展示一行，能够点击展开。展示信息包括：组织名称、ID、疑似团伙类型、团伙成员人数、发现时间、可信度。同时展现该组织下的团伙成员的照片、成员名字、常住地等。点击展开，能够看到一个团队之间的关系图谱，点击一个人，能够高亮与他有关系的人员。

5.1.2.2 轨迹中心

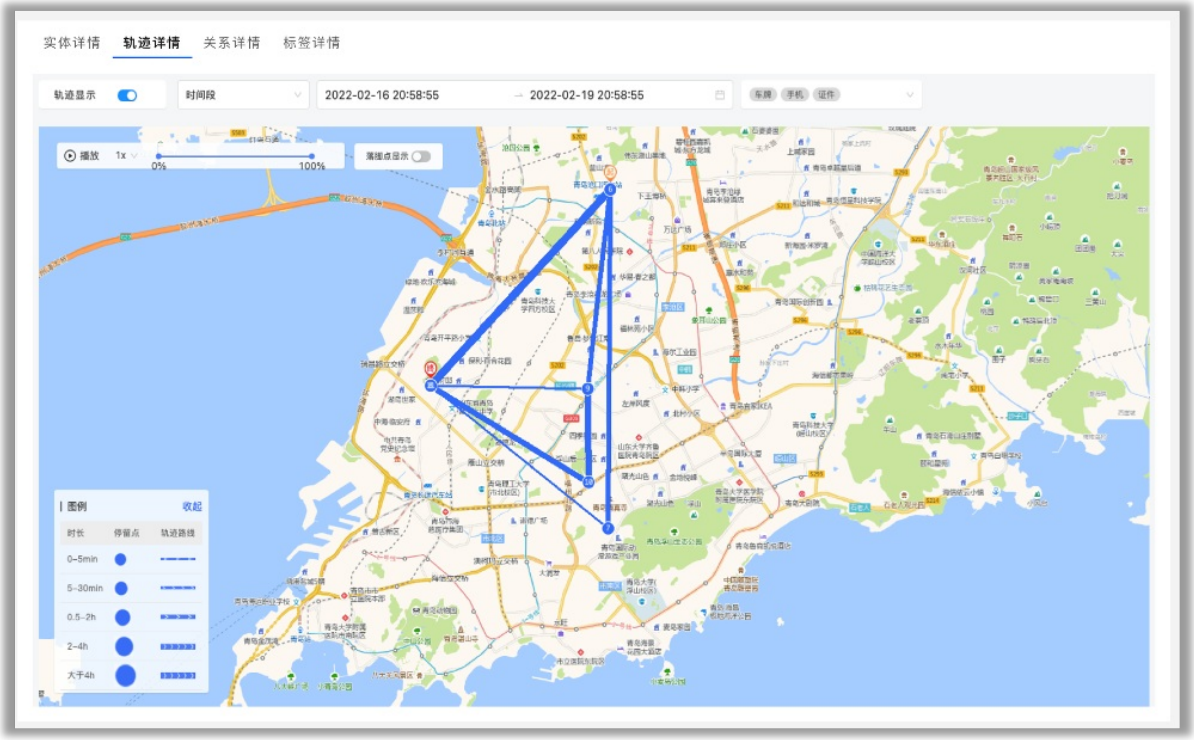
在公安场景中，人员和车辆的轨迹能够展示人员和车辆的动向、落脚点、日常习惯等信息，是作为打击犯罪行为，预警犯罪行为非常重要的辅助功能。基于业务场景的实际需求以及实体中心对实体的数据归档，通过底层算法能力对大数据视频进行自动研判、自动跟踪，产出人员和车辆每日行动轨迹。轨迹中心展示平台的轨迹数据资产，并且使用地图还原实际轨迹。通过全量融合轨迹，大幅度的增加轨迹的点位数量，提升轨迹的密度和准确度，做到精准展示轨迹并且预测轨迹，帮助各类警种进行有效追踪。

5.1.2.2.1 轨迹中心概览

1) 轨迹数据资产

展示平台组轨迹中心的数据资产情况，从轨迹数量、轨迹增量、实体关联度等多个维度展示轨迹中心数据资产的情况，对轨迹中心的使用情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

2) 轨迹信息地图总览



默认使用地图展示当天该地区随机生成的几段轨迹，同时按照先后顺序，自动对轨迹进行播放。鼠标指向某段轨迹能够查看相关信息，展示轨迹的人员、车辆找照片、时间、间隔时长等等。

5.1.2.2.2 轨迹中心详情

1) 实体轨迹查询

能够针对实体进行实体的轨迹查询，主要使用人员实体图片上传查询该人员的轨迹，查询时轨迹按照天生成，能够进行查询时间的修改，最多能够查看7日内的轨迹。

2) 实体轨迹播放

点击播放，能够按发生时间先后顺序进行实体轨迹的播放。可以进行播放、暂停、停止操作，不可进行拖拽。轨迹能够按照0.5、1、2、4、8倍速进行播放。

5.1.2.3 关系中心

人员是公安业务的核心，业务场景是基于人员的行为和关系发展。人员实体在活动中会和多类实体产生多边关系，这类关系是在业务场景研判的关键点之一。 对实体之间同框的次数时长进行算法判断，支撑构建关系中心。关系中心支持人员和人员、人员和车辆、人员和场所等多种实体关系的挖掘，并形成关系图谱，联通多类实体，帮助公安解决寻找关系相关情景。

5.1.2.3.1 人人关系

1) 人人关系概览

a) 人人关系数据资产

展示人人关系的数据资产情况，从人人关系数量、平均关系数量等多个维度展示人人关系数据资产的情况，对人人关系中心的情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 人人关系信息图谱

根据人人关系的次数倒序，形成人人关系图谱。可以随机选择区域内一个人的关系进行图谱拉动展示，展示相关人员的名字、关系、关系次数、年龄等。点击单个人员，与他相关的人员会高亮显示，展示在页面上。

2) 人人关系详情

a) 人人关系标签

展示人人关系的标签，用于分类当前人员的不同关系。标签包括同行、邻居、朋友等。

b) 关系人基础信息

根据关系标签展示和当前人员有标签关系的人员，展示关系人员的信息，包括关系人姓名、关系人ID、关系次数。

c) 人人关系发生信息

展示人人关系发生地的信息，包括发生地、发生时间，同时展示双方同框的照片，图片点击能够放大，能够进行左右切换。

5.1.2.3.2 人车关系

1) 人车关系概览

a) 人车关系数据资产

展示人车关系的数据资产情况，从人车关系数量、平均关系数量等多个维度展示人车关系数据资产的情况，对人车关系中心的情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 人车关系信息图谱

根据人车关系的次数倒序，形成人车关系图谱。图谱形式为人车互相之间的关系为一组，展示人的名字、关系、车牌号、关系次数等。

2) 人车关系详情

a) 人车关系标签

展示人车关系的标签，用于分类当前人员和车的不同关系，关系包括车主、副驾驶等等。

b) 关系车基础信息

根据关系标签展示和当前人员有标签关系的人员，展示关系人员的信息，包括关系人姓名、关系人ID、关系次数。

c) 人车关系发生地信息

展示人车关系发生地的信息，包括发生地、发生时间，展示人与车的同框照片。

5.1.2.3.3 人地关系

1) 人地关系概览

a) 人地关系数据资产

展示人地关系的数据资产情况，从人地关系数量、平均关系数量等多个维度展示人地关系数据资产的情况，对人地关系中心的情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

b) 人地关系信息图谱

根据地点为维度，一个地点中包含一组出现在该地点过的人员，人员信息按照出现次数倒序排序。展示地点的名称、ID、地址等。同时展示人员的名字、性别、出现次数等信息。

2) 人地关系详情

a) 人地关系基础信息

展示人员和地的相关关系。包括人员出现过在哪些地方、近期出现次数、近期一个月出现天数、最近一次出现时间、最后一次出现时间等。

b) 人地关系发生信息

展示人地关系发生地的信息，包括发生地、发生时间，展示人经过地点的照片。

5.1.2.4 标签中心

不同业务场景的五类实体：人、地、事、场所、组织，经过数据接入和数据加工会展现出不同的数据特点，对于数据的共同特点能够形成有效标签，为数据进行标签服务。标签融合形成实体的实体画像，展现实体的特点和形象。实体画像能够帮助公安业务对实体进行有效认知，放大公安在不同业务上的关注点。同时通过标签的不断更新，获取精准的实体画像，为公安业务场景提供支持。

5.1.2.4.1 人员标签

1) 人员标签数据资产

展示人员标签的数据资产情况，从标签包含人员数量等多个维度展示人员标签数据资产的情况，对人员标签情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

2) 人员标签图谱

使用人像环绕的方式展示平台的人员标签能力，点击标签能够进入人员列表，支持查看拥有标签的人员。

3) 人员标签实体人员列表

点击单个人员标签，能够进入该标签下的实体人员列表，列表展示实体人员头像，按照标签标记时间倒序。

5.1.2.4.2 车辆标签

1) 车辆标签数据资产

展示车辆标签的数据资产情况，从标签包含车辆数量等多个维度展示车辆标签数据资产的情况，对车辆标签情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

2) 车辆标签图谱

使用车辆环绕的方式展示平台的车辆标签能力，点击标签能够进入车辆列表，支持查看拥有标签的人员。

5.1.2.4.3 场所标签

1) 场所标签数据资产

展示场所标签的数据资产情况，从标签包含场所数量等多个维度展示场所标签数据资产的情况，对场所标签情况有综合概览认识。同时数据支持按最近30天、最近15天、最近一周的维度查看。

2) 场所标签图谱

使用地图展示该区域，提供区域标签的选择，选择后会在地图上标记拥有相关标签的点位，展示场所标签地点。

5.1.2.5 行为中心

通过行为分析算法得到行为中心。行为中心包含回家、去运动等生活类行为；去上班等工作类行为；去学校等学习类行为。行为分析依赖场所信息。

行为中心主要功能包括人员行为和车辆行为。

5.1.2.5.1 人员行为

人员行为是对人员实体行为作出定义，根据行为的不同进行分类，与视频图形进行碰撞，用于实际人员行为分析。

5.1.2.5.2 车辆行为

对车辆实体行为作出定义，根据行为的不同进行分类，与视频图形进行碰撞，用于实际车辆行为分析。

5.1.3 开放服务

平台通过对外开放的能力，提供多种接口和数据服务，增加开放性，链接多平台。

5.1.3.1 接口服务

平台提供多类接口，其中包括：提供实体、轨迹、关系、标签查询接口。提供区域碰撞接口，根据设定的多个区域、停留时段、离开时段进行碰撞，返回同时满足条件的人。提供伴随分析接口，根据给定的目标及时间段，分析与目标有伴随关系的人。提供熟人聚集分析接口，根据给定的名单及时间段，分析名单中有聚集的人。

5.1.3.2 数据服务

提供实体、轨迹、关系、标签数据同步服务。

5.2 数据资产层

数据资产层包含五大中心：实体中心、轨迹中心、关系中心、行为中心、标签中心。数据资产是对实体全方位画像分析的结果，是产品的核心，其他部分都是围绕丰富数据资产以及更好对外提供数据资产服务展开。

5.2.1 实体中心

实体中心包含人员实体、车辆实体、场所实体和设备实体。分别对各类实体进行构建和基础信息整理，分配实体ID。轨迹、关系、行为、标签基于实体ID进行关联，组成实体完整画像。

5.2.1.1 人员实体

对人像聚档结果以及人脸人体结构化结果进行分析，形成人员实体。

具体包含内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
1	最早出现	最早出现的时间、地点、经纬度（任意关联要素）	是	否	否
2	最后出现	最后出现的时间、地点、经纬度（任意关联要素），动态更新	是	否	否
3	姓名		否	否	是
4	身份证		否	否	是

5	人脸信息	聚档id	是	否	否
6	人体信息	聚档id	是	否	否
7	步态信息	聚档id	是	否	否
8	登记住址		否	否	是
9	性别		是	否	否
10	年龄	具体值，或年龄段（上下限）	是	否	否
11	身高	身高段（上下限）	是	否	否
12	体型	按1400： 0111:特胖；0112:较胖； 0113:中等；0114:较度；0115:特搜； 0119：其他	是	否	否
13	发型	按1400： 1:平头；2:中分；3：偏分； 4：额秃；5：项秃；6：全秃；7：卷发； 8：波浪发；9:麻花辫；10：盘发；11： 披肩；12:长发；13:短发；14:马尾； 15：秃头；99：其他	是	否	否
14	脸型	按1400： 0311:椭圆脸；0312:圆脸； 0313:长方脸；0314:方脸；0315: 倒大脸；0316:三角脸；0317:狭长脸； 0318:菱形脸；0319:畸形脸；0399:其他	是	否	否
15	是否戴眼镜	是，否	是	否	否
16	是否留胡子	是，否	是	否	否
17	人脸照片	质量最好的人脸图片，最多3张（1证件照 +2抓拍图或3抓拍图）	是	否	否
18	人体照片	质量最好的人体正面、侧面、背面抓拍图 各1张	是	否	否

19	步态视频	质量最好的正面、侧面、背面步行视频，各1个	是	否	否

5.2.1.2 车辆实体

对车辆结构化结果进行分析，形成车辆实体。

具体包含内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
20	最早出现	最早出现的时间、地点、经纬度	是	否	否
21	最后出现	最后出现的时间、地点、经纬度，动态更新	是	否	否
22	车辆照片	质量最好的车辆正面、背面抓拍图各1张	是	否	否
23	车牌号	各类机动车号牌编号车牌全部无法识别的以“无车牌”标识，部分未识别的每个字符以半角‘-’代替	是	否	否
24	车牌颜色	按1400：1：黑色，2: 白色，3: 黄色，4：蓝色，5:绿色；6:渐变绿色；7:黄绿双拼；99:其他	是	否	否

25	号牌种类	01:大型汽车号牌；02: 小型汽车号牌；03: 使馆汽车号牌；04: 领馆汽车号牌；05:境外汽车号牌；06:外籍汽车号牌；07:普通摩托车号牌；08:轻便摩托车号牌；09: 使馆摩托车号牌；10: 领馆摩托车号牌；11: 境外摩托车号牌；12:外籍摩托车号牌；13:低速车号牌；14:拖拉机号牌；15:挂车号牌；16: 教练汽车号牌；20:临时入境汽车号牌；21: 临时入境摩托车号牌；22: 临时行驶车牌；23:港警用汽车号牌；24:警用摩托车号牌；25:原农机号牌；26:香港入出境号牌；27:澳门入出境号牌;31:武警号牌；32:军队号牌；51:大蓝牌；52: 单层黄牌；53: 双层黄牌；54: 挂车号牌；55:新使馆号牌；56:黄色学牌；57:普通黑牌；58:白色警牌；59: 单层武警牌；60: 双层武警牌；61: 单层军牌；62:双层军牌；63:兴能源小型车牌；64:新能源大型车牌；65:农用车牌；66: 02式个性号牌；67:民航车牌；68: 港澳出入境车牌；99:其他	是	否	否
26	车辆类型	1:大客车；2:大货车；3:小轿车；4:越野车；5:面包车；6:小货车；7:商务车；8:中型客车；9:皮卡；99：其他	是	否	否
27	特种车辆类型	1:出租车；2:警车；3:救护车；4:混凝土搅拌车；5:洒水车；6:危险品车；7:消防车；8:拖拉机；9:工程车；10:罐车；11:校车；12::普通车；99：其他	是	否	否
28	车辆品牌	见附1-车辆品牌	是	否	否
29	车身颜色	1:黑色；2:白色；3:灰色；4:红色；5:蓝色；6:黄色；7:橙色；8:棕色；9:绿色；10：紫色；11：青色；12:粉色；13:透明；99：其他	是	否	否
30	年检贴状态	0：没有年检贴；1：有年检贴；99：其他	是	否	否

5.2.1.3 设备实体

对相机基础信息及实时状态信息进行分析，形成设备实体。

具体包含内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
31	首次时间	设备最早服役时间	否	否	否
32	末次时间	设备最后上报数据时间	否	否	否
33	设备类型	视频流相机、图片流相机、wifi嗅探、电围、RFID、门禁	否	否	否
34	设备编号	相机为国标码	否	否	否
35	设备名称		否	否	否
36	经纬度		否	否	否
37	安装位置	安装地址	否	否	否
38	设备IP		否	否	否
39	设备MAC	无	否	否	否
40	所属单位名称		否	否	否
41	所属单位代码		否	否	否
42	朝向	相机的视野方向	否	否	否
43	设备状态	是否在线	否	否	否
44	设备厂商信息		否	否	否

5.2.1.4 场所实体

通过AOI构建算法，将高德AOI数据与相机点位数据进行碰撞，自动建立场所与相机的关联，形成场所实体。

具体包含内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
45	建档时间	档案建立的时间	否	是	否
46	场所来源	高德数据/大数据算法/手工添加	否	是	否
47	场所一级分类	按照高德通用数据进行分类（公安场所分类如重点场所等以标签形式）	否	是	否
48	场所二级分类		否	是	否
49	*场所三级分类	机场、火车站、医院、学校、商场、小区、网吧等	否	是	否
50	*场所名称	xxx医院	否	是	否
51	*经纬度	一个点或一个区域	否	是	否
52	*场所面积	场所经纬度区域面积	否	是	否
53	*场所地址	包括对应中文地址和geohash	否	是	否
54	场所照片	最多10张	否	是	否

5.2.2 轨迹中心

轨迹中心主要针对人车轨迹进行分析，提供人车轨迹、轨迹伴随、人车关联、人车落脚点等能力。

5.2.2.1 人车轨迹

人员轨迹来源于人像聚档数据，车辆轨迹来源于车辆结构化数据。系统对人车轨迹重新建表，方便其他模块对轨迹数据的查询和利用。

包含以下数据表及内容：

序号	数据表	内容

1	要素最早最晚分布	要素ID、要素类型、最早最晚抓拍图片、最早最晚相机、最早最晚出现时间、最早最晚出现经纬度、最早最晚出现地点、最早最晚出现场所、最早最晚出现区域
2	档案最早最晚分布	档案ID、档案类型、最早最晚抓拍图片、最早最晚相机、最早最晚出现时间、最早最晚出现经纬度、最早最晚出现地点、最早最晚出现场所、最早最晚出现区域
3	访问场所记录	档案ID、档案类型、要素ID、要素类型、场所ID、场所名称、场所类型、进入时间、进入抓拍图、进入采集设备、离开时间、离开抓拍图、离开采集设备、场所停留时间
4	人脸轨迹	聚档ID、人脸图片ID、设备编号、设备名称、抓拍时间、经纬度、小图坐标、人脸属性等
5	人体轨迹	聚档ID、人体图片ID、设备编号、设备名称、抓拍时间、经纬度、小图坐标、人体属性等
6	车辆轨迹	车牌号、车辆图片ID、设备编号、设备名称、抓拍时间、经纬度、小图坐标、车辆属性等

5.2.2.2 轨迹伴随

基于轨迹可分析出短时间内的显性伴随关系，包括人人伴随、人车伴随以及车车伴随。

轨迹伴随主要内容包含实体1 ID、实体1 类型、实体2 ID、实体2 类型、伴随起始时间、伴随起始地点、伴随结束时间、伴随结束地点。

5.2.2.3 人车关联

经过一段时间的轨迹分析，发现某些人员实体与某些车辆实体经常伴随，满足一定条件后会自动建立人车关联，相当于识别出车辆长期驾驶人。人车关联后，人员轨迹可包含其关联的车辆轨迹，还原人员的完整轨迹。

人车关联主要内容包含人员实体ID、车辆实体ID、关联起始时间、关联结束时间。

5.2.2.4 人车落脚点

基于落脚点特征，通过落脚点算法分析出人车落脚点。包含住所地、工作地、必经点、驻留点、热点。

人车落脚点主要内容包含实体类型、实体ID、落脚点类型、落脚点位置、落脚点标准地址、落脚点经纬度、落脚次数、抓拍图片列表等。

5.2.3 关系中心

通过各种关系算法分析得到关系中心。关系中心包含人人关系、人车关系、人地关系、车车关系。

5.2.3.1 人人关系

人人关系包含朋友、同住所的固定关系及亲密度，以及同驾、同行、同框、同现的临时关系。

主要内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
1	*同住所	经常同行（近3个月超过6天），且住所地相同，且1的其他同住人也是2的同住人。建立后条件不再满足超3个月取消关系	是	否	否
2	*朋友	同时出现达一定程度（近3个月超过3天，且地点不同），且有同行关系，且住所地不同	是	否	否
3	稳固关系亲密度	0-100%	是	否	否
4	稳固关系说明	近3月同现次数、月均同现次数、近3次同现时间和图片、高频同现位置	是	否	否
5	*同驾	同一辆车的主驾副驾。主副驾相关内容都额外依赖超卡数据	是	否	否
6	*同行（伴随）	同时行走、骑车或停留超过一段距离（2km）或时间(0.5h)	是	否	否
7	*同现	出现在同一摄像头，时间间隔很近（5s内）	是	否	否
8	*同框	同一视频画面	是	否	否
9	临时关系记录	每条临时关系：时间（段）、地点、经纬度、图片；每种临时关系次数可按时段统计	是	否	否

5.2.3.2 人车关系

人车关系包含长期乘坐、高频主驾、高频副驾、临时乘坐。

主要内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
10	*长期乘坐	建立了人车关联	是	否	否

11	*高频主驾	列举近1个月最高频（按存在主驾关系的天数之和，同一天多次只算一次）的主驾者，人员ID、姓名和图片，最多5个	是	否	否
12	*高频副驾	列举近1个月最高频（按存在副驾关系的天数之和，同一天多次只算一次）的副驾者，人员ID、姓名和图片，最多5个	是	否	否
13	*临时乘坐	未建立人车关联，机动车时为副驾	是	否	否
14	长期关系置信度	0-100%	是	否	否
15	长期关系建立时间	确定关系的时间	是	否	否
16	长期关系说明	近3月次数、月均次数、近3次时间和图片	是	否	否
17	临时关系记录	可多次：起止时间、起止地点、图片	是	否	否

5.2.3.3 车车关系

车车关系包含同路和同出入。

主要内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
18	*同路（车车伴随）	同时/相邻出现多次连续一定时间或距离	是	否	否
19	*同出入	进出场所时间间隔在3分钟内算一次同进出，当天之内，并且间隔大于半小时	是	是	否

5.2.3.4 人地关系

人地关系包含近期出现次数、近期出现天数、首次/末次出现时间。

主要内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据

20	*近期出现次数	近一个月出现次数，去除5分钟内重复抓拍	是	是	否
21	*近期出现天数	近一个月出现的天数	是	是	否
22	首次出现的时间	第一次出现的时间（不限制时间范围）	是	是	否
23	末次出现的时间	最后一次出现的时间	是	是	否

5.2.4 行为中心

通过行为分析算法得到行为中心。行为中心包含回家、去运动等生活类行为；去上班等工作类行为；去学校等学习类行为。行为分析依赖场所信息。

主要内容如下表：

序号	一级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
1	*去上班	有工作地，且当日去到了工作地，且在工作地有停留（有半小时以上没出现或只出现在工作地）【时间（段）、地点、人物、事件】	是	是	否
2	*回家	有住所地，在住所地出现（当日去之前有别的轨迹），且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
3	*去医院	当日出现在医院，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
4	*去汽车站	当日出现在汽车站，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
5	*去火车站	当日出现在火车站，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否

6	*去机场	当日出现在机场，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
7	*去运动	当日出现在运动场所，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
8	*去网吧	当日出现在网吧，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】	是	是	否
9	*去学校	当日出现在学校，且有停留。（有半小时以上没出现或只出现在该地）【时间（段）、地点、人物、事件】可通过经常去学校判断学生或老师	是	是	否

5.2.5 标签中心

基于各类实体标签算法，分析实体行为标签、关系标签及内容标签，形成标签中心。标签中心包含人员标签、车辆标签、场所标签以及设备标签。

5.2.5.1 人员标签

人员标签内容如下表：

序号	二级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
1	*常住人员	出现天数及频次达到一定程度（近1个月超过10天），打标后超过3个月未达标取消标签	是	否	否
2	*临时人员	偶尔出现（近1个月不超过7天），且每次间隔都较长（间隔3天内出现算一次，每次出现间隔7天以上），且多次出现在交通枢纽；变为常住人口后取消标签	是	否	否
3	*首次出现	历史数据或积累数据1个月之后，针对近7天新出现的人打标；1个月后取消标签	是	否	否

4	是否重点人	根据重点人库获取重点人分类标签 涉X：涉黄、涉赌、涉毒、涉恐、涉传、涉诈、涉黑、涉邪教 其他：前科、信访等	是	否	是
5	*疑似无业	出行不规律（近3个月有7天以上轨迹，且没有一个时间段（1小时）占比超过50%），且无工作地；有工作地后取消标签	是	否	否
6	*昼伏夜出	经常（近1个月超7天）夜间出现白天不出现；3个月不满足取消标签	是	否	否
7	*频繁更换车辆	近1月频繁主驾多个不同车辆（超3个）；3个月不满足取消标签	是	否	否
8	*频繁更换住所地	近1月更换多个住所地（超3个）；3个月不满足取消标签	是	否	否
9	（近X天）特定时段出入小区（视频）	早上7-10点出小区，夜晚17-22点回小区，列举7，15，30天	是	否	否
10	（近X天）本市触网天数（视频）	本市被采集到轨迹，计算触网天数。列举7，15，30天	是	否	否
11	（近X天）是否在本市有活跃轨迹	本市有触网轨迹，计1次，列举7，15，30天	是	否	否
12	*深夜同行异性人数	女性在深夜21点至凌晨5点同行的异性总人数； 同行判定方式一：被同一设备在5秒内抓拍到为1次，累计出现3次及以上视为同行； 同行方式二：一天内同时（前后5秒内）被不同设备（设备相距500米以上）抓拍到视为通行	是	否	否

13	*深夜同行涉x前科人数	在深夜21点至凌晨5点同行的涉黄前科人数； 在深夜21点至凌晨5点同行的涉毒前科人数； 在深夜21点至凌晨5点同行的涉赌前科人数； 在深夜21点至凌晨5点同行的盗窃前科人数； 在深夜21点至凌晨5点同行的涉黑恶前科人数； 在深夜21点至凌晨5点同行的邪教前科人数； 同行定义同2	是	否	是
14	*与涉x前科同行次数	与涉黄前科人员同行的次数； 与涉毒前科人员同行的次数； 与涉赌前科人员同行的次数； 与盗窃前科人员同行的次数； 与涉黑恶前科人员同行的次数； 与邪教前科人员同行的次数； 同行定义同2	是	否	是
15	*经常夜间出现	在近30天内5天及以上且5次及以上在21:00-5:00出现	是	否	否

5.2.5.2 车辆标签

车辆标签内容如下表：

序号	二级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
16	*首次出现	历史数据或积累数据1个月之后，针对近7天新出现的车打标；1个月后取消标签	是	否	否
17	*长期在本地	出现天数及频次达到一定程度（近1个月超过10天）；超过3个月未达标取消标签	是	否	否
18	*临时在本地	偶尔出现一次，且每次停留时间较短（近1个月不超过2次，每次不超过3天），且为外地车牌；变为常住车辆后取消标签	是	否	否
19	*疑似套牌	主驾、车辆属性一致，但车牌有多个	是	否	否

20	*共享汽车	主驾人员频繁更换（1个月超7个）；3个月后不满足则取消标签	是	否	否
21	*昼伏夜出	经常（近1个月超7天）夜间出现白天不出现；3个月不满足取消标签	是	否	否
22	*经常载重点人	经常出现重点人主驾或副驾（1个月超7次）；3个月不满足取消标签	是	否	是
23	*经常夜间出现	在近30天内5天及以上且5次及以上在21:00-5:00出现	是	否	否

5.2.5.3 场所标签

场所标签内容如下表：

序号	二级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
24	*经常出现重点人	近1个月超7天出现过重点人，超7人（不同人）；3个月不满足取消标签	是	是	是
25	*疑似涉X场所	经常出现涉X人员（近1个月超7天出现过涉X人员，超7人）；3个月不满足则取消标签 涉X：涉黄、涉赌、涉毒、涉恐、涉传、涉诈、涉黑、涉邪教、涉前科	是	是	是
26	近7日过人数量	一个人一天出现多次算一次	是	是	否
27	*过人高峰时段	出现人最多的小时时段，top 3	是	是	否
28	近7日过车数量	一个车一天出现多次算一次	是	是	否
29	*过车高峰时段	出现车最多的小时时段，top 3	是	是	否
30	近7日过重点人数量	一个重点人一天出现多次算一次	是	是	否
31	*过重点人高峰时段	出现重点人最多的小时时段，top 3	是	是	是

5.2.5.4 设备标签

设备标签内容如下表：

序号	二级内容	内容描述	依赖人车数据	依赖AOI数据	依赖身份数据
32	近7日感知数据量	相机的结构化量、门禁为刷卡/刷脸量、其他为感知到的ID数量	是	否	否
33	*设备状态异常	数据传输不稳定，或者错误数据	是	否	否

5.3 工具层

工具层包含的数据接入及加工是产品重要组成部分，包含核心的数据处理模型及大数据算法。主要分为数据接入、数据治理和数据加工三个环节。

5.3.1 数据接入

系统需要接入必要的数据进行加工分析，可支持的数据包括人像聚档数据、人车结构化数据、高德地图数据、前端点位数据、重点人数据等。

支持多种数据对接模式：源端接口推送、原始库抽取、消息队列接收。

5.3.1.1 人像聚档数据

人像聚档数据来源于上游聚档结果，可通过上述任意方式进行数据对接，包括已实名档案和未实名档案。人像聚档数据包含人像档案ID、档案基础信息、人脸人体小图及对应的时间经纬度等信息。人像聚档数据是分析人员画像的必要数据。

5.3.1.2 人车结构化数据

人车结构化数据来源于人像/车辆卡口等抓拍的结构化解析结果，可对接视图库或第三方系统获取。人车结构化数据核心信息包括人脸特征（如脸型发型、戴口罩、头盔等属性）、人体衣物体貌特征（如上衣/裤子款式、颜色、长短等）、车辆特征（如车型号、品牌，颜色、车牌及其颜色等）、车辆乘坐人员特征（主副驾人脸）。人车结构化数据是分析人车画像的必要数据。

5.3.1.3 高德地图数据

场所、人地关系、行为等分析依赖高德地图的AOI/POI场所及分类信息、地址标准化能力。在此基础上通过AOI构建算法自动完成场所与前端点位的关联，进而为场所分析、人地关系分析、行为分析等提供支撑。高德地图AOI核心属性包括场所名称、场所地址、场所面积、所属一二级分类、场所经纬度范围等。

高德地图数据采取手动导入方式，支持周期更新。

若无高德地图数据，可以手动建立场所信息以及场所与前端点位的关联关系，通过模版导入系统。但场所数据量大，前端点位变化多，人工维护成本较高。

5.3.1.4 前端点位数据

前端点位数据主要是人脸/人体/车辆等各视频卡口的摄像头信息，核心属性包括摄像头编码、摄像头名称、摄像头安装地址、摄像头经纬度、摄像头IP等基础信息。

前端点位数据可通过上述任意方式进行对接。系统支持从人车结构化数据中提取前端点位核心信息，包括摄像头名称、摄像头编码、摄像头经纬度。

前端点位数据是必要的基础数据，对轨迹分析、人地关系、同行同现关系等分析提供支撑。

5.3.1.5 重点人数据

重点人数据主要是本地重点关注人员（涉黄赌毒黑传销、前科等）身份证号和对应的关注类型以及关注周期。可通过上述任意方式进行对接，用于对重点人员以及疑似重点人员的分析。

5.3.2 数据治理

5.3.2.1 数据架构

数据架构包括数据模型（概念模型、逻辑模型）以及数据间的流转关系，主要对行业数据的分类、分布和流转进行规划、设计，确保感知数据架构能够与现有业务数据架构和应用系统保持一致和融合，避免产生信息孤岛，或者带来重复不必要的数据集成、数据转换。

数据架构图如下：

5.3.2.2 数据标准

系统依赖的感知数据采用视图库1400协议标准来设计数据模型，上层结合业务的实体、轨迹、行为、标签、关系五大中心业务标准；通过标准化的元数据来管理整个感知数据。

5.3.2.3 数据开发

数据开发目的是把分散、零乱、标准不统一的数据整合到一起，为更深入的算法分析及业务应用提供依据。

数据经过数据抽取、清洗去重、规则转换、关联补全之后生成到数据仓库。

数据抽取：从各个不同的数据源抽取到数仓ODS层；

清洗去重：进行数据清洗去重，将脏数据和不完整数据过滤掉，数据去重技术主要有两方面。一方面是指数据文件集合中完全重复上报的数据,只保存唯一的数据单元,从而消除冗余数据；另一方面是按照各个要素具体行为轨迹挑选核心属性字段在业务意义上去重。

规则转换：IDM到IDS层的规则转换过程中，包括数据粒度的转换，在数据仓库中生成按照业务域粒度进行聚合的数据模型，例如实体域，轨迹域等；规则类关系和标签的计算也在此步骤完成。

关联补全：在从ODS到IDM的转换过程，还会进行身份的关联补全。

5.3.2.4 数据质量

数据质量包含两个方面：数据自身的质量和数据的过程质量。

数据本身质量指数据必须准确地反映实际发生的业务，任何业务操作的数据都没有被遗漏，各个字段的实际取值不能跳过其约束条件。在本产品中会通过事前的数据探查和界面指标透出反馈数据本身的质量。

数据的过程质量指在数据治理过程中各环节产出的数据符合预期的结果。在数据治理每个阶段治理链路的产出都会配置质量检测规则对各类数据质量问题的严重程度进行识别、度量、监控、预警和流程熔断。

5.3.3 数据加工

数据加工是指在各类画像分析过程中所用到的大数据算法。核心包含轨迹伴随算法、人车关联算法、人车落脚点算法、关系类算法、AOI构建算法等。

5.3.3.1 轨迹伴随算法

【算法描述】

根据轨迹时空相似性，计算任意两段轨迹序列的显性伴随关系，包括人人伴随、人车伴随以及车车伴随。

【输入】

任意两段轨迹序列，可以是人-人、人-车、车-车轨迹序列。

【输出】

实体1 ID、实体1 类型、实体2 ID、实体2 类型、伴随置信度、伴随起始时间、伴随起始地点、伴随结束时间、伴随结束地点、共现次数。

【方案概述】

轨迹伴随算法目的是判断两个轨迹序列是否存在伴随关系，并给出伴随置信度分值。如下图示意：

现实世界的轨迹数据一般是离散的时空点，利用采集设备对对象的移动路径进行采样得到，本质上反映了对象的移动行为模式。如利用人脸卡口、车辆卡口等设备对人或车在途经重要地点时拍照采样。

受采集设备分布稀疏、设备经纬度误差、数据处理能力有限等影响，简单的时空共现匹配很难获得较好的效果。我们针对数据分布特点，通过合理的时空误差容忍策略、轨迹向量化等算法挖掘潜在的伴随关系，并设计加速算法实现千万量级轨迹序列的两两伴随计算。

5.3.3.2 人车关联算法

【算法描述】

人车关联算法目的是将身份证号和车牌号进行关联，回答“谁在开车”或“谁在车上”的问题联，相当于识别出车辆长期驾驶人或乘车人。人车关联后，人员轨迹可包含其关联的车辆轨迹，还原人员的完整轨迹。

【输入】

指定人和车的长期轨迹序列，确保更高的召回率和准确率，建议时间周期至少6个月。可以采取逐日累加方式，也可针对历史数据进行批量处理。

【输出】

人员实体ID、车辆实体ID、关联起始时间、关联结束时间、关联得分、关联次数、关联地点数、关联天数。

【方案概述】

利用车辆卡口、人脸卡口、超卡等设备采集的人脸轨迹、车辆轨迹数据，通过算法深度分析挖掘人车关联关系。通过轨迹预处理、轨迹伴随、行为规律等算法，挖掘基本人车关联；利用轨迹链接算法提升关联准确率；针对数据稀疏问题，积累长周期数据挖掘提升关联置信度；若有额外的手机等轨迹数据，可用于交叉验证进一步提升人车关联准确率。

自动分析发现某些人员实体与某些车辆实体经常伴随，满足一定条件后会自动建立人车关。

5.3.3.3 落脚点算法

【算法描述】

基于落脚点特征，通过落脚点算法分析出人车落脚点。包含住所地、工作地、必经点、驻留点、热点。

【输入】

指定人或车的长期轨迹序列，时间周期至少1个月。

【输出】

实体类型、实体ID、落脚点类型、落脚点位置、落脚点标准地址、落脚点经纬度、落脚次数、抓拍图片列表等。

【方案概述】

利用聚类算法、驻留点检测算法等，对指定人或车的长期轨迹序列进行时空聚类、驻留点识别，得到一些重要位置点。然后结合住所地、工作地、必经点、驻留点、热点等不同落脚点类型的业务规则定义，计算每一类落脚点。人、车的落脚点分别单独计算。

例如“住所地”类型落脚点的业务规则定义：

5.3.3.4 关系类算法

关系类算法在感知领域中，基于摄像头抓拍的视频数据和其他来源数据，构建人人之间的关系纽带，通过人人关系形成复杂庞大的关系网络结构，在关系网络结构中综合网络结构的拓扑信息和属性信息进行团体发现和团体类型识别，最终产出人人之间的亲密度结果和团体发现的团体结果，为分析人人关系和社群组织提供支撑。

1. 亲密度算法

【算法描述】

在进行关系分析时，亲密度是衡量关系紧密程度的重要指标。基于亲密度算法可实现亲密度的有效计算和动态更新。

【输入】

人人之间原始关系对数据，人员视觉结构化特征，团体发现网络结构和团体结果。

【输出】

人人之间随时间更新的亲密度值。

【方案概述】

基于对现实场景中关系时效性的分析，采用寿命期限阈值对亲密度计算进行实时约束，在寿命期限内亲密度值随时间维度进行叠加积累和有序更新，保证了在有噪数据环境下亲密度计算的稳定性和合理性。

亲密度计算在涵盖人人之间多种类型关系特征的基础上，还融合了社群组织范围下人人之间的关系密度，包括两者的共同邻居情况，是否同属一个团体等网络结构特征，从线到面更加全面丰富地反映人人之间的亲密情况。

2. 团体发现算法

【算法描述】

除了双边关系外，通过团体发现算法还可挖掘多边实体关系即团体。支持对团体的生命周期进行动态更新。

【输入】

人人之间原始关系对数据，人员视觉结构化特征。

【输出】

随时间更新的团体结果序列，包括团体统计分析信息如团体类型，成员数量，成员重要性，成员关联情况等。该团体序列可反映社群组织随时间的演变发展过程，展示团体的生命周期，如团体的壮大，萎缩，分裂，合并，出生和消亡。

【方案概述】

通过增量计算和存量融合的方式，解决了动态网络问题，最终产出可适应于动态发展的动态团体发现结果，同时通过流式对海量的存量数据进行团体发现，仅需少量算力，解决了该场景下常见的冷启动困难的问题。此外结合现实场景中关系边的生卒问题，随时间对关系边进行过期核验和处理，对因关系边死亡引起的团体变化进行实时调整更新。在流式基础上，通过局部网络结构分析，解决了重叠团体问题，可将成员分别归属到其对应的多个团体中。基于对团体在不同时期的跟踪，可展示团体随时间的具体变化情况，揭示整个生命周期，提供团体演化分析。

5.3.3.5 AOI构建算法

1.异常设备发现及纠偏算法

【算法描述】

各类采集设备的坐标质量是本产品各项功能的基础，如果大量设备对应的经纬度发生偏差，将直接影响下游的各类算法、模型效果。例如在轨迹预测中，如果某些设备出现异常，最终下游任务得到结果极有可能南辕北辙。在真实场景中，由于大多数设备信息登记依靠人工，存在大量异常，包括经纬度数值错误、坐标系错误、名称错别字或过于简略等等。本功能通过算法计算，进行异常设备发现及纠偏。

【输入】

设备登记名称、设备登记经纬度、轨迹序列。（至少包含两项）。

【输出】

该设备位置是否异常、置信度、推荐经纬度。

【方案概述】

考虑到单一维度信息无法进行交叉验证，算法需要设备名称、设备登记经纬度、轨迹序列中至少包含两项信息，进而通过多模态信息融合综合判断。包括使用高德地址标准化能力进行设备名称地址标准化、设备经纬度拟地理编码和基于轨迹序列（如人脸轨迹）的设备向量化，并进一步通过卡尔曼滤波或深度回归模型进行异常置信度计算，最终对于异常设备推荐其真实经纬度。

2.设备场所信息关联算法

【算法描述】

采集设备本身无任何业务属性，只是孤立的离散点。通过设备场所信息关联算法赋予采集设备的场所属性（如学校、公司、小区等）是下游任务开展行为分析、轨迹分类、关系判断等能力的必要条件。

【输入】

设备名称、设备经纬度。

【输出】

POI名称、POI类型、AOI名称、AOI类型。

【方案概述】

基于高德的POI、AOI查询接口，并基于计算几何、语义相似度分析、轨迹表示等算法能力实现弹性场所信息关联。尤其是对于可同时匹配到多个POI、AOI的设备，可通过算法进行关联度排序，推荐置信度最高结果。

5.4 性能指标

截止文档发布，产品指标如下：

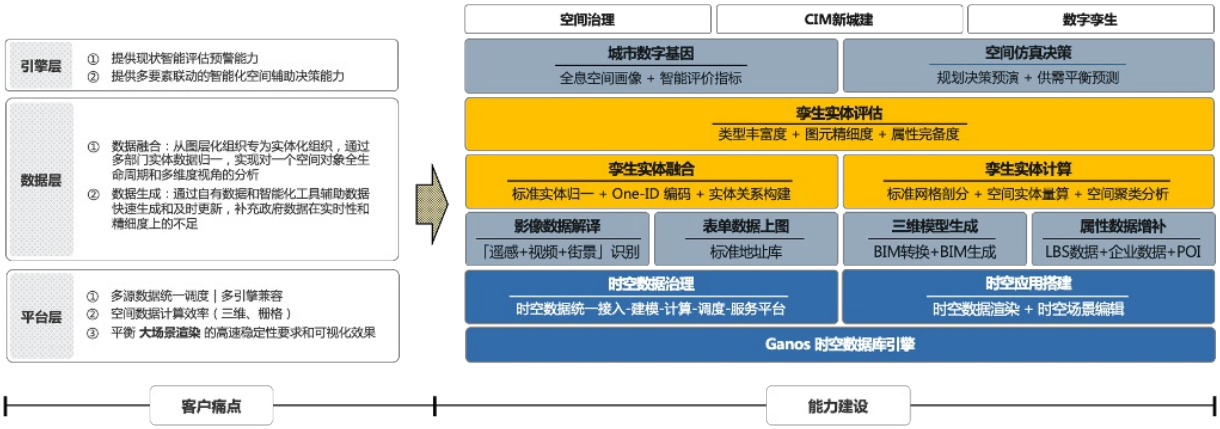
序号	指标项	指标	备注
1	实体数量	实有人口3000万人、实有车辆1000万辆、100万个场所、1000万个组织、100万个设备	临时人员、临时车辆不限
2	人车码关联率、准确率	设备重叠率>20%下，动态关联率>50%、准确率>80%	
3	平均关系准确率	>=75%	
4	平均标签准确率	>=75%	

11.空间治理平台

11.1. 什么是空间治理平台

产品定位

针对 城市时空数据统一底座 的建设需求，基于 云原生 的时空数据技术栈，构建结构化数据和空间数据 在线一体化 治理和应用的平台，以城市智能生成、孪生实体建模等技术构建 全域全要素孪生实体融合模型，以 城市数字基因、空间仿真决策 为精细化空间治理提供量化决策依据，提供城市级孪生数字建设工具。



产品用户

空间治理（CIM）平台主要面向自规、住建、政务、应急等对空间资源有诉求的行业用户，包括政府客户、ISV生态伙伴，用户角色可具体划分为数据开发人员、应用开发人员、业务分析人员。

- **数据开发人员**

平台提供完整的从数据同步、数据建模、数据融合、数据开发到服务发布的数据开发链路工具，并面向各类型业务场景，沉淀CIM、空间治理、灾害、政务等不同领域的数据模型，支持用户一站式快捷搭建自审业务的数据资源体系。

- **应用开发人员**

用户可在平台内进行业务模型、业务算法开发，构建、渲染二三维空间场景，并利用应用资源包进行业务流程、应用界面的快捷定制。

- **业务分析人员**

用户可可查看、管理、运营平台归集、融合的各类数据资产、服务资产及应用资产，并基于智能分析中心提供的全息空间画像、仿真分析引擎等能力进行业务分析。

1.3. 产品价值

- **时空数据研发更简单**

平台提供完善的空间资源行业数据模型，具备城市全域要素的数字化能力，为基于数字孪生城市的各业务场景提供数据底座，同时提供各种工具产品以简化数据开发过程，流水线式的数据治理过程，极大缩短数据治理周期，降 低数据治理成本。

- **业务应用研发更便捷**

平台不仅为业务系统供应商提供稳定数据基座、基础数据服务，同时支持业务算法挖掘、数字孪生业务场景构建、大数据处理，实现快速业务应用系统开发，降低业务研发门槛。

- **业务分析场景更丰富**

平台面向规划、应急、城建、政务等领域，提供业务分析全链条的能力，业务人员不需要做代码开发，利用平台已有能力即可进行的业务场景分析。

- **空间资产管理更清晰**

提供城市数字孪生成果直观展示，全资产监控管理能力，服务成果汇报、现场指挥、领导决策。

11.2. 产品优势

空间治理（CIM）平台具备以下优势：

1、云原生时空数据处理技术栈

适配多种数据库、兼容常见GIS引擎，实现结构化数据和空间数据在线治理

2、智能化的空间数据生产加速器

快速生成标准实体模型数据，实现与政府数据底板基础的融合

3、One- ID的孪生实体融合模型

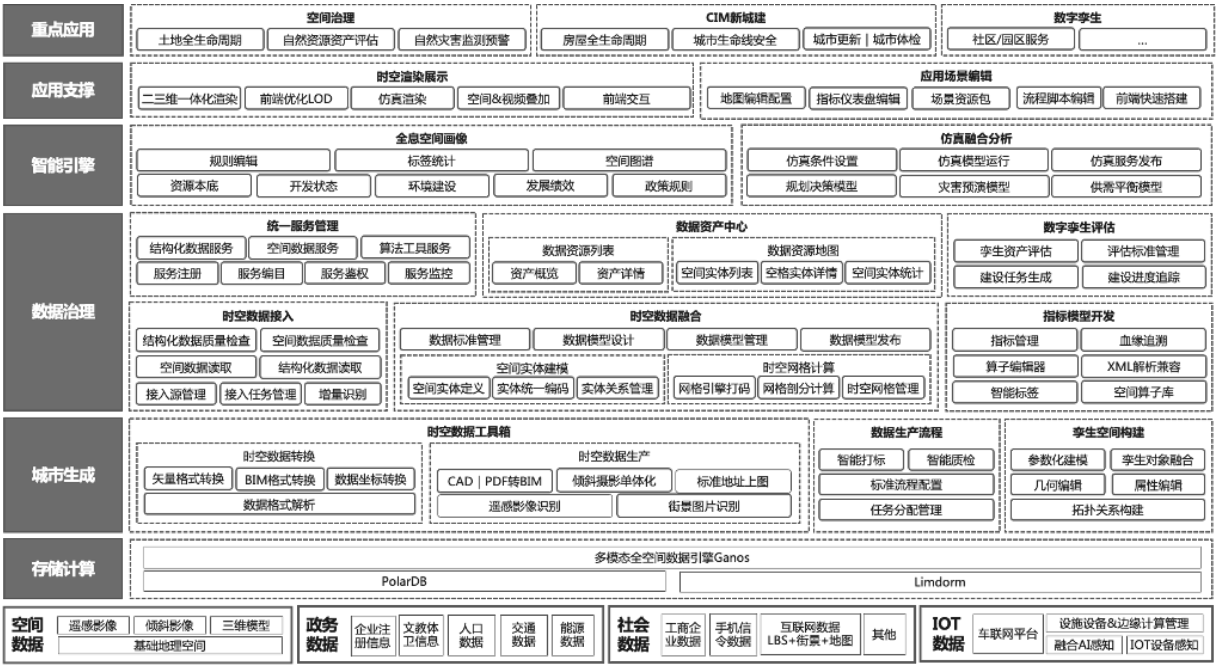
构建全域全要素时空One-ID数据模型，通过统一编码、网格剖分等技术支持多类时空关系建构和检索

4、覆盖评估-预演的智能决策闭环

多维指标刻画、城市体检预警，提供面向空间规划决策的交互式预测和预演模型支撑

特性1 时空计算 云原生的时空数据处理技术栈	特性2 数据生产 智能化的空间数据生产加速器	特性3 数据融合 One- ID的孪生实体融合模型	特性4 数据智能 覆盖评估-预演的智能决策闭环
基于云原生数据库构建大规模、复杂时空存储计算的强健基底，适配多种数据库、兼容常见GIS引擎，实现结构化数据和空间数据在线治理-渲染展现-应用搭建的一体化平台	依托自有高德和标准地址库等数据，以及影像识别、BIM生成等智能化工具，快速生成标准实体模型数据，可实现与政府数据底板基础的融合，提供更新预警和孪生评估等功能	基于全链路实体建模工具，构建全域全要素时空One-ID数据模型，通过统一编码、网格剖分等技术支持多类时空关系建构和检索	构建面向空间对象的多维指标刻画和体检预警能力，提供面向空间规划决策的交互式预测和预演模型支撑
技术要点 ■ **计算高性能** • 全空间多类型数据一库统管，通过计算下沉、多库调度、优化算法等，实现对几何矢量、栅格、点云、轨迹、BIM等空间数据的高效存储和计算 ■ **平台高兼容** • 一库多平台，支持多源时空数据接入、计算、服务和质量管理，支持跨GIS引擎计算链路混编调度及第三方XML文件解析 • 将血缘追溯等能力扩展到空间数据 ■ **渲染高视效** • 与数据库能力协同，支撑大规模城市二三位一体化场景的流畅和高视效渲染	**技术要点** ■ **智能融合生成** • 通过自有和易获取数据，实现地物智能解析、BIM智能生成、动态属性增补等智能生成能力 • 基于数据融合规则与政府数据整合，生成符合标准实体模型的数据集 ■ **时空更新预警** • 识别更新要素并提示比对，协助人工更新核查，变整体更新为局部更新 ■ **数字孪生评估** • 基于国家标准或场景自定义模版，对数据集的数字孪生建设程度进行评价	**技术要点** ■ **全链路实体建模工具** • 实体建模：提供实体定义管理、实体属性管理、实体编码管理、实体关系管理等能力支持实体模型生成 • 空间构建：提供属性编辑、拓扑构建、网格剖分等空间融合标注和计算能力，支持实体关系构建 ■ **全要素标准实体模型** • 全域实体分类 • 统一空间编码 • 标准实体模型	**技术要点** ■ **现状评估预警** • 结合多元数据，从资源本底、区位特征、建成环境、发展绩效、政策规则等多方面建立不同空间单元的自动评估预警指标，作为城市体检和城市更新的量化决策依据 ■ **空间决策预演** • 在交通仿真基础上拓展到交通-用地仿真，为全局综合联动分析奠定基础 • 内置规划决策预演、供需平衡预测等多种智能仿真模型，支持交互式的空间决策预演分析
业务价值 构建云原生技术栈（去GIS化）实现对海量、多源时空数据的高效、一体化处理和资产化管理，满足场景对空间可视化和低代码应用搭建的需求	**业务价值** 为时空数据生产提供智能化解析和精细化补足的能力，支撑政府对城市治理的实时性和精细度要求	**业务价值** 实现以空间实体为核心实现多部门时空数据的全要素融合、全周期追溯，提供一码管空间的数据服务，成为各部门时空数据协同共享的基本能力	**业务价值** 为业务人员提供更精准的现状预警和决策仿真分析能力，支持量化决策会商，降低决策风险和损失

11.3. 产品架构



11.4. 功能特性

空间治理平台由时空数据汇聚、时空数据融合、数据资产管理、智能分析中心等核心模块构成，完整功能构成如下表。

模块	功能	子功能	功能描述
时空数据汇聚	数据接入管理	实时数据同步	支持支持实时任务的新建，并支持一键启动多个上云任务；通过创建实时同步任务，选择已经注册的数据源端空间数据，并选定目标端，实现全量任务自动建表，可批量创建。
		离线数据同步	支持离线任务的新建，并支持一键启动多个上云任务；通过创建离线同步任务，选择已经注册的数据源端空间数据，并选定目标端，实现全量任务自动建表，可批量创建。
		接入数据管理	支持从OSS中或本地以文件导入的方式上传数据，包括各类shapefile、GeoJSON、CSV、PostGIS格式的矢量文件，NetCDF、Image、GeoTIff格式的栅格数据和OSGB、S3M和3DTiles的倾斜摄影文件导入。
	数据标准化	标准化工具	具备一系列空间数据标准化工具，包括坐标、格式、数据生成等工具。
		标准化流程	按照不同类型的数据进行标准化数据治理流程的组织，将相关数据治理工具进行串联，从而灵活地构建在线空间数据生产流水线。

	数据质量检测	质量规则配置	支持空间和非空间质量规则的设定。提供多种空间数据及业务质量检查规则，数据完整性检查，空间类型检查等，并可根据业务需要予以选择。
		质量检查计划	支持对质量检测计划的创建，可批量选择需要周期执行且已配置好质量规则的数据模型，配置周期检测任务和调度，进行持续质量检测。支持对计划任务的管理、启动、停止和试跑。
		质量检查任务	支持对质量检测计划的创建，可选择需要周期执行的带规则数据模型，并配置周期检测计划的时间信息设置。
		质量监控报告	支持根据系统数据质量监测运行结果以天的维度自动生成系统整体质量监控报告。支持根据数据唯一性、准确性、规范性、一致性、时效性和完整性进行多维度评估分析。
时空数据融合	数据标准设计	数据标准查看	可查看数据源、数据字典、数据度量、数据维度等标准。
		数据标准设计	可进行各类型数据标准的新建、设计。
		孪生实体建模	可实现快速创建实体、管理实体目录、浏览、设计和编辑实体模型等功能。实体模型涵盖了实体基本信息、实体空间编码，实体空间表达、实体的业务属性和实体关系ER图等信息。
	孪生空间构建	孪生对象生成	通过对实体数据进行属性、编码、几何、结构、关联关系等方面的系列处理，将数据转换为实体对象，建立与空间实体间的直接映射关系，可灵活进行空间场景构建、行业数据模型构建及支撑上层应用。
		空间语义编辑	支持空间实体的可视化图形编辑及实体键关联关系/挂在关系等语义编辑。
		三维化	可通过空间实体临摹绘制、参数化构建、内置模型自适应的方式进行三维空间生成。
	指标融合计算	指标定义管理	提供空间治理多业务场景下的指标体系，同时将指标与土地、楼房、规划、公服设施、企业等实体关联，在业务维度和空间维度均可查询指标，及其详细信息。
		计算链路编辑	采取算子编辑器，是适合于算子低代码开发的可视化编辑器，实现计算模型的可视化构建和执行，计算过程支持血缘。

		指标链路查看	提供指标计算逻辑说明，包括配图，文字描述与视频讲解，将抽象的数学概念和共识，具象为直观的业务逻辑。
		指标质量溯源	提供指标生成链路溯源与节点质量监控，包括指标字段质量、关联数据表质量、计算实例状态、源数据表质量等信息。
	标准网格引擎	网格剖分计算	利用网格编码技术，为试点区域每一寸空间赋予可计算、易检索的全球唯一空间编码标识。该标识体系具有多尺度、唯一性、立体性、全时空性等特点，为实现建设要素精细化管理提供了基础性技术工具支撑。
		网格检索查询	支持按网格和网格组合的数据穿透式查询检索，利用网格数据“切片”实现数据按需提取和分发交互的能力。
数据资产管理	数据资产中心	数据资源目录	实体、数据、指标、标签、二三维地图服务等资产，可按业务口径分部门别类商家岛数据资源目录。
		数据资源地图	以“实体”为维度为用户提供地图化地探索数据资源全景的能力（多个图层和业务数据“聚合”后形成的实体数据全集），包括：查看实体目录、实体详情；进行实体数据质量查看、数据更新查看；进行实体属性的数值过滤查看及实体数量统计。
		统一服务管理	数据表、标签、指标、地图服务加工完成后，转换为标准的数据服务 API，为应用提供统一的鉴权和数据访问方法，屏蔽了数据的异构，很好的解决的数据安全和访问复杂的问题。
	数字孪生评估	资产评估	数字孪生评估主要内容包括城市数字孪生覆盖度、区域孪生精细度分级、区域物联感知设备覆盖度、区域视频设备覆盖度等。
		任务生成	为用户提供量化制定数字孪生度的总体目标的功能，用户可以分维度、分地区、分行业对数字孪生目标进行管理，并对目标的约束性和引导性进行规定。
		标准管理	用户可查看当前孪生评估的维度、指标和计算模型，并进行维度和指标的增、删、改及模型权重调整等动作。
智能分析中心	全息空间画像	实体标签	对实体从资源本底、建成环境、发展绩效、开发状态、服务水平、区位特征、政策规则等多方面形成标签，可查看标签树，可对标签顺序和层级进行编辑和调整，可查看不同标签实体的基本统计数值及在地图上的分布情况。
		画像分析	基于配置好的规则，可实现土地等实体的统计分析、空间图谱等分析查看。

11.5. 应用场景

- **国土空间**

为国土空间行业提供国土空间治理/实景三维方面应用。

客户案例：浙江省1+11空间治理体系建设

紧密围绕“空间治理”构建数字化支撑体系，直指浙江提高土地利用效率、实现高质量发展、解决资源错配、协同空间规划的阶段性问题

- **城乡建设**

支持CIM新城建建设需求，提供新城建CIM平台

客户案例：宜昌城市大脑数字底座-CIM平台

为宜昌城市大脑提供二三维空间数据底板，支持各数字化应用的空间数据调用需求。

- **政务服务**

为智慧城市建设提供空间底板，提供二三维数据的统一管理

- **孪生仿真**

为水利/水务/园区/地产/应急等多行业数字孪生仿真能力