

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：V1.0.1

文档版本：20220523

 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 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

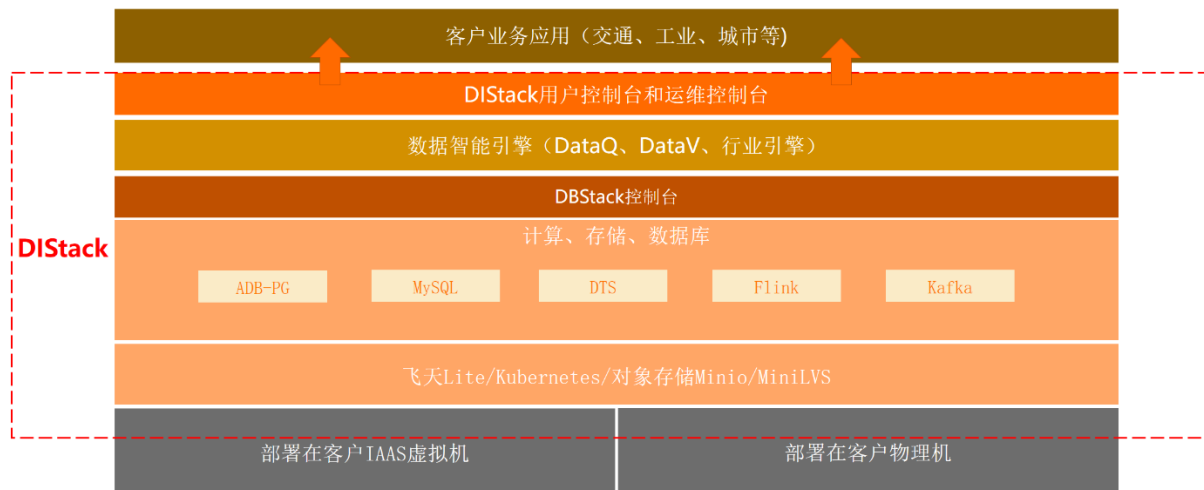
目录

1.DIStack整体介绍	06
2.产品优势	07
3.产品架构	08
4.功能特性	09
5.应用场景	23
6.数据资源平台产品	25
6.1. 产品优势	26
6.1.1. 核心能力	26
6.1.2. 用户价值	30
6.2. 功能特性	31
6.2.1. 研发工作台	31
6.2.1.1. 数据同步	31
6.2.1.2. 数据查询	33
6.2.1.3. 数据标准	36
6.2.1.4. 数据建模	37
6.2.1.5. 数据开发	38
6.2.1.6. 统一服务	42
6.2.1.7. 通用资源	47
6.2.1.8. 运维监控	48
6.2.1.8.1. 任务中心	48
6.2.1.8.2. 全景监控	51
6.2.1.9. 质量评估	58
6.2.1.10. 解决方案	60
6.2.2. 运营工作台	61
6.2.3. 数据资产中心	66
6.2.4. 我的资产	69

6.2.5. 系统设置	70
6.3. 应用场景	74
6.4. 基本概念	76
6.5. 使用限制	78
7.交通云控平台产品	80
7.1. 功能特性	81
7.2. 应用场景	82

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

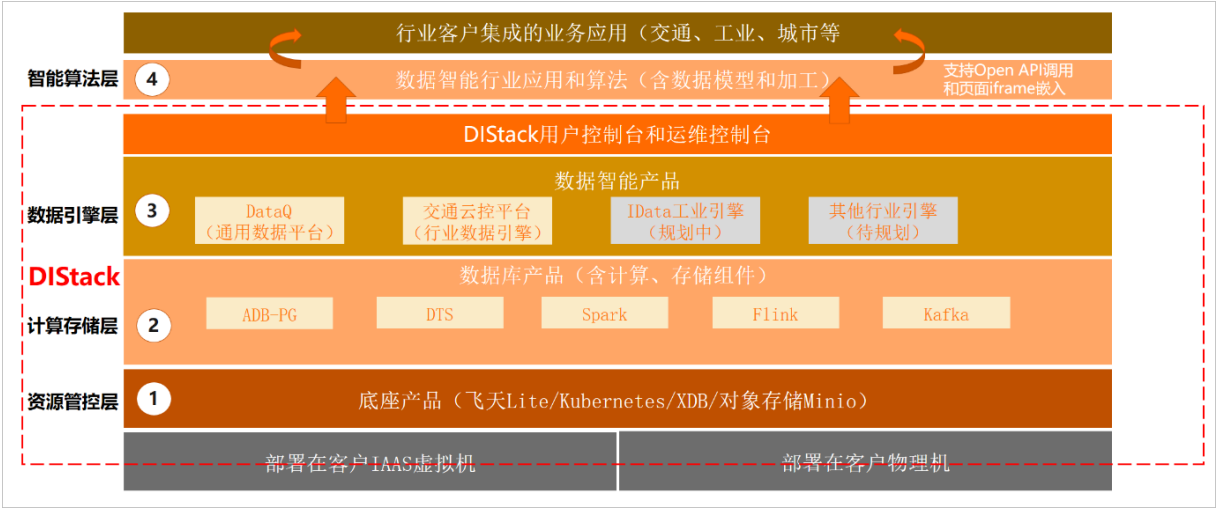
- 支持软硬一体化部署，开箱即用；
- 稳定的云平台计算和存储引擎，调优的通用和行业数据引擎，加速行业数据智能应用一体化开发；
- 内置行业解决方案，贴行业应用一体化输出，支持算法工程师快速定制客户数据需求，提升对客户业务服务能力。

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Kubernetes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使用户自动化运维。
	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。

一级模块	外部数据管理 二级模块	功能描述
		数据管理
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。	
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。	

一级模块	二级模块	功能描述
存储过程	支持存储过程。	
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、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。

	通用资源	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。
		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Oracle 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行数据自动探查，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。
	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。
数据资源平台- 画像分析	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。

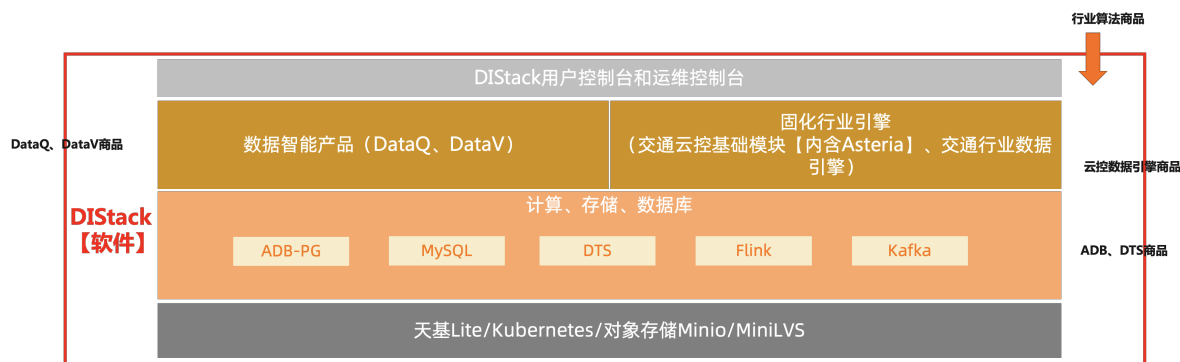
	群体管理	/	支持群体列表管理：支持对群体信息进行新建、删除、更新、状态变更、查看、编辑、查看不同类型群体个数和占比等操作。新建群体支持标签圈群、导入本地文件和群体计算三种方式。支持群体运维：可发布静态和动态群体，并提供在数据库一次性和周期落表的任务运维管理。支持群体计算：通过图形化的计算方式可实现两个群体交并差计算得到新的群体，并提供对群体结果支持查看详细群体、发布API、静态、动态和实时群体。
	群体画像	/	提供对群体或个体进行多维度分析，及群体对比分析，个体对比分析。支持保存分析结果和下载分析报表。

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、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.数据资源平台产品

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、数据探查、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的使用并探索数据，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：

终端	大屏	PC Web	移动钉钉端	
场景应用	领导驾驶舱 智慧选址	土地价值评估	数据孪生沙盘 一码管地	一宅一码
前端能力 (依托DataV)	GIS前端			应用场景区编辑器
	地图场景搭建	二维前端	三维前端	可视化组件库 画布编辑 业务模型编辑
空间数据中台 (依托DataQ)	空间资产管理			
	空间单元 数据资产 地图资产 指标资产 服务资产 分析场景资产			
	空间资产编目	空间资产查找发现	空间资产血缘关联	空间资产共享
	多源数据同步	空间数据查询	空间数据标准与质量	空间数据建模与融合
	结构化空间数据	元数据查询	数据标准定义	地理编码
	结构化空间数据	图形查询	空间质量规则	投影坐标转化
	结构化空间数据	统计查询	空间质量报告	数据网格化
			空间位置关联	Java/Python开发
				字段级数据血缘
				元数据服务
平台底座	「飞天计算平台」 离线处理 MaxCompute 实时计算 PostgreSQL OSS RDS		「ON ADP」 离线处理 Hadoop Hive 实时计算 Spark ADB for PG NAS MySQL	
数据来源	「按业务类型」 社会经济数据 现状数据 规划数据 管理数据		「按数据类型」 矢量 影像 地形 倾斜摄影 激光点云 精模	
			「按数据格式」 GIS数据 BIM数据 三维数据 业务数据	

● 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。
- **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；面向数据资产管理提供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。
- **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，支持对多个数据源端进行展现订阅；可定期自动订阅数据表、标签、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，以可视化的方式体现数据表、标签、API等数据资产的总量与增量的指标信息。
 - **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**：是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。

6.1. 产品优势

6.1.1. 核心能力

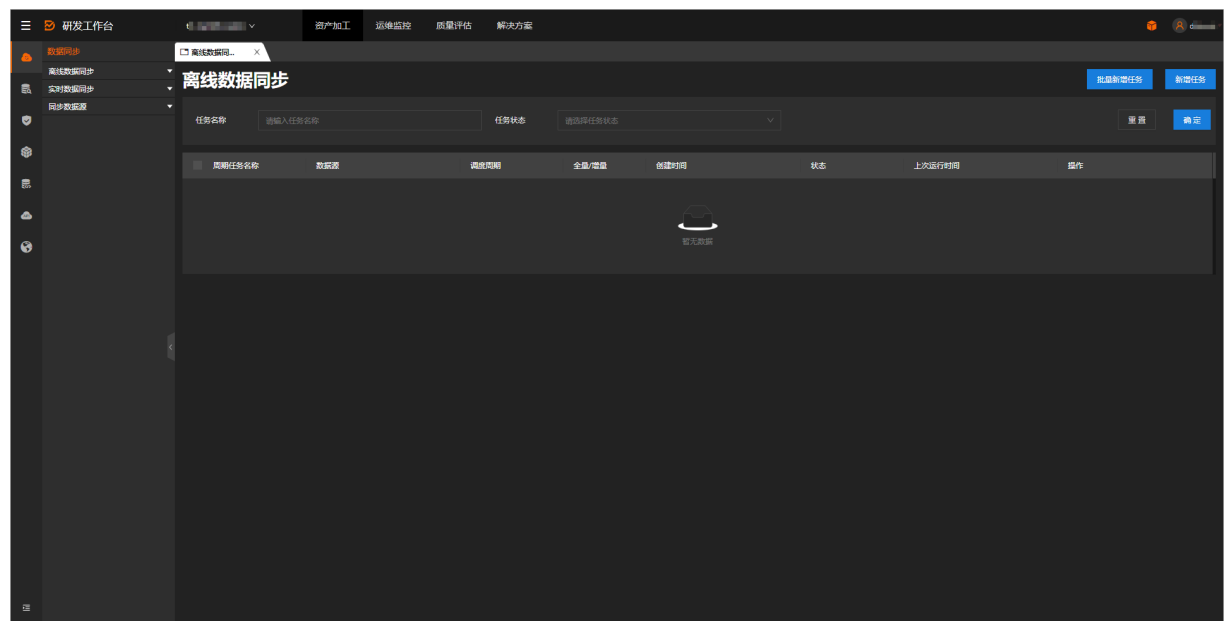
数据资源平台为用户提供了一站式数据资产定义、生产、管理与服务平台，提供企业级数据资产构建能力和一致性使用体验，助力客户快速构建数据智能平台，实现数据资源统一管理，挖掘潜在规律，优化业务决策，让大数据真正的驱动客户业务。本文主要介绍数据资源平台的核心能力。

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

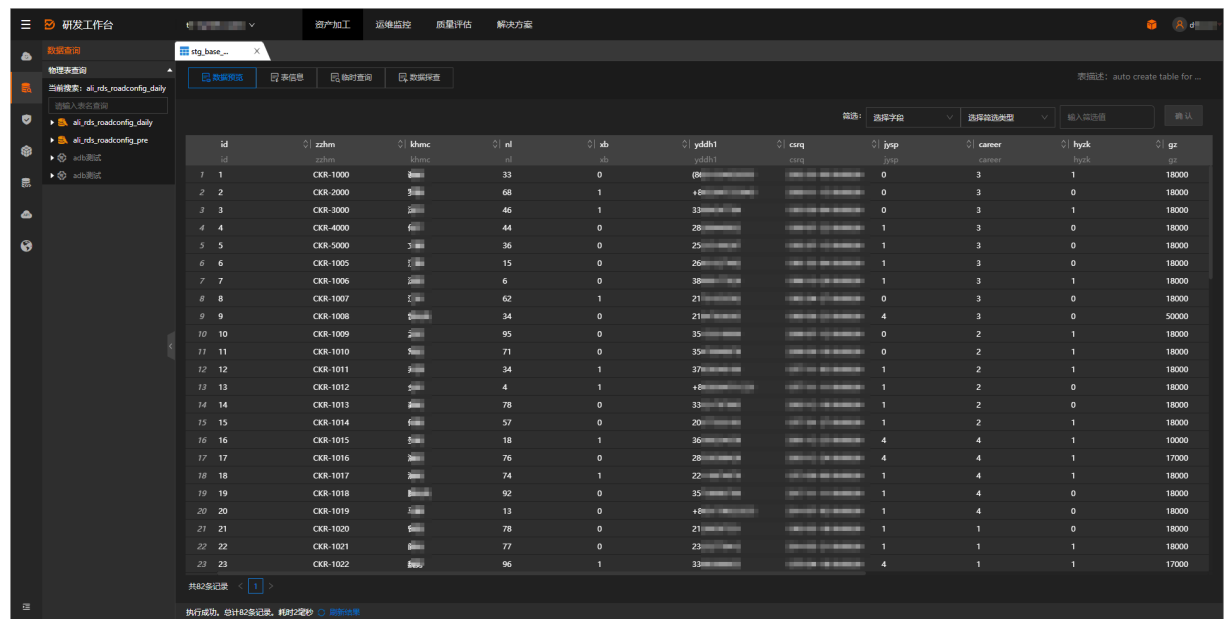
- 支持离线/实时数据传输。

- 周期同步任务调度统一管理。



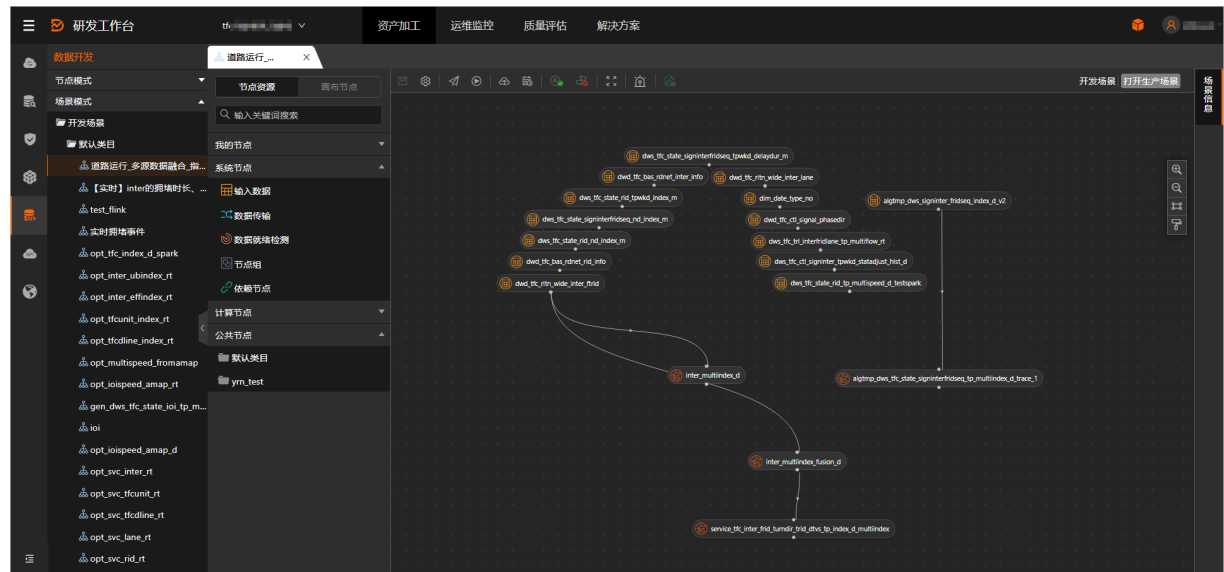
自动灵活的数据探查

- 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。



强大的数据加工能力

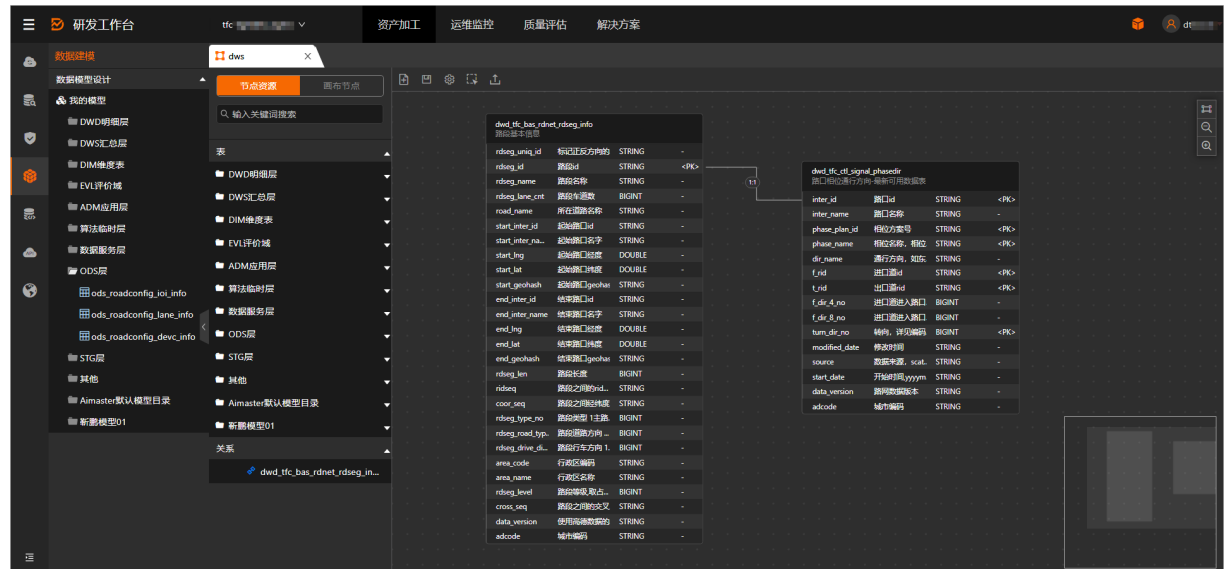
- 离线、实时、在线多类型计算节点。
- 不同类型计算节点在同一个任务流中混合编排调度。



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、API服务资产的目录搜索和详情展现。
- 字段粒度的端到端数据血缘。

统一数据服务封装

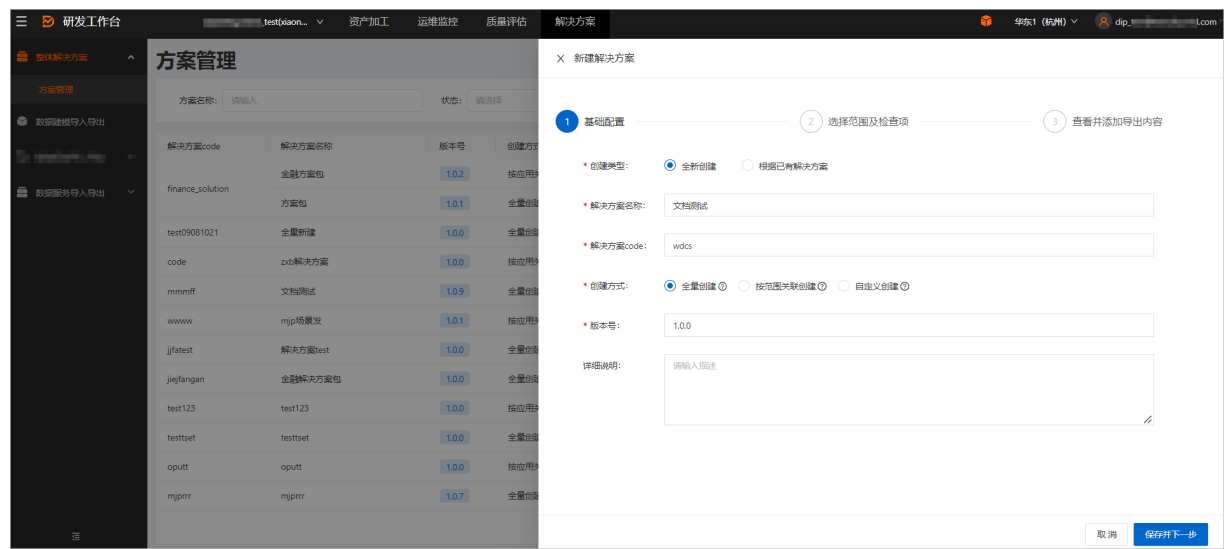
- 实时、离线、半结构化、非结构化、消息队列的数据服务封装。
- 支持向导模式和SQL模式等的服务封装。
- 简单身份认证和加密签名身份认证模式。
- 数据服务调用流量管控。

API名称	归属应用	云资源类型	状态	最后更新时间	操作
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-13 14:46:13	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:55:51	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:37:26	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:35:05	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:34:34	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:33:17	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:33:03	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-12 16:32:06	编辑 上线 调试 更多
交管_道路使用信息	交管服务	AnalyticDB PostgreSQL	已上线	2022-01-11 21:49:43	编辑 上线 调试 更多

行业知识敏捷交付

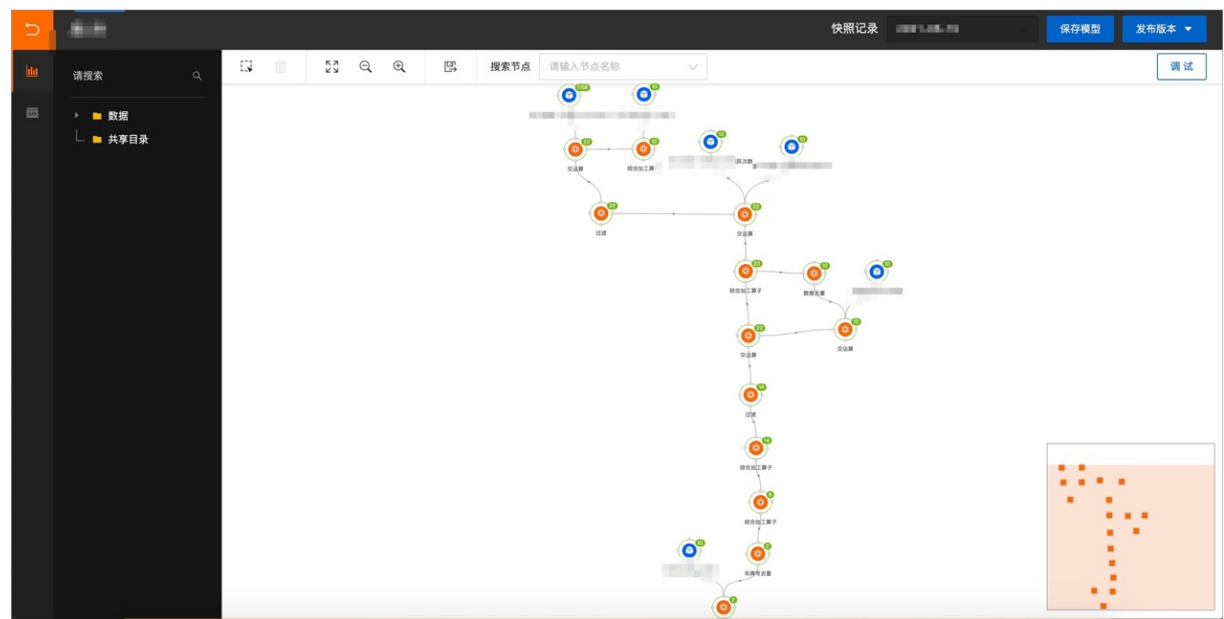
- 工作组下API服务、模式、节点、场景等全类型资产构建解决方案。

- 通过全局数据血缘关联资产灵活构建解决方案。



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。



6.1.2. 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。

6.2. 功能特性

6.2.1. 研发工作台

6.2.1.1. 数据同步

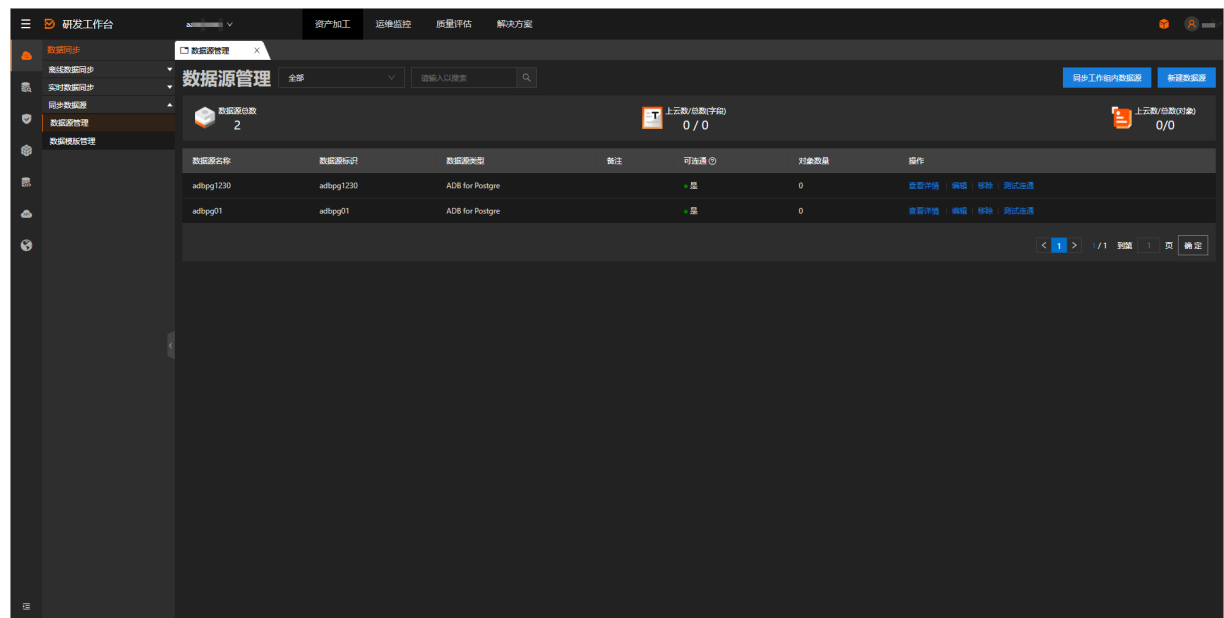
数据同步提供批量、实时多源异构数据的便捷同步或接入能力，系统提供完善的数据接入配置、任务运行监控等功能，有效保障数据接入的稳定性和可控性，满足各类平台、数据源及应用系统间的数据汇聚需求。

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。

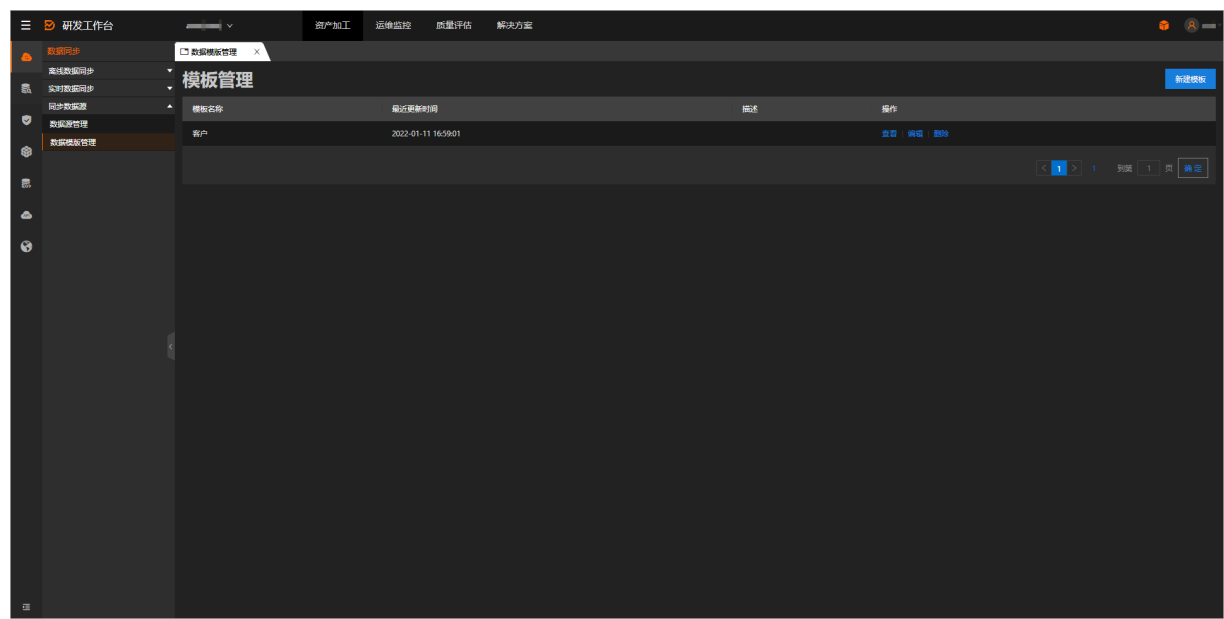
数据源管理

- 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

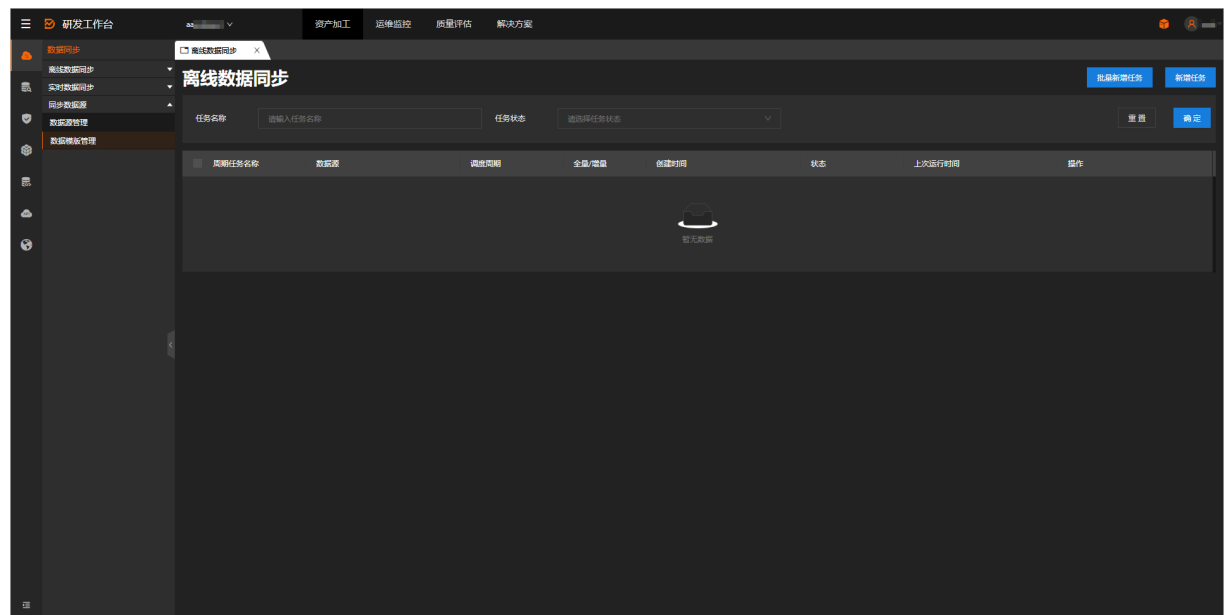
- 针对半结构化、非结构化数据，支持创建数据模板，支持自定义数据字段及字段类型，并提供编辑、删除等功能。
- 支持在离线数据同步、实时数据同步任务中引用已添加的数据模板，根据数据模板定义的数据结构执行数据同步任务。



离线数据同步

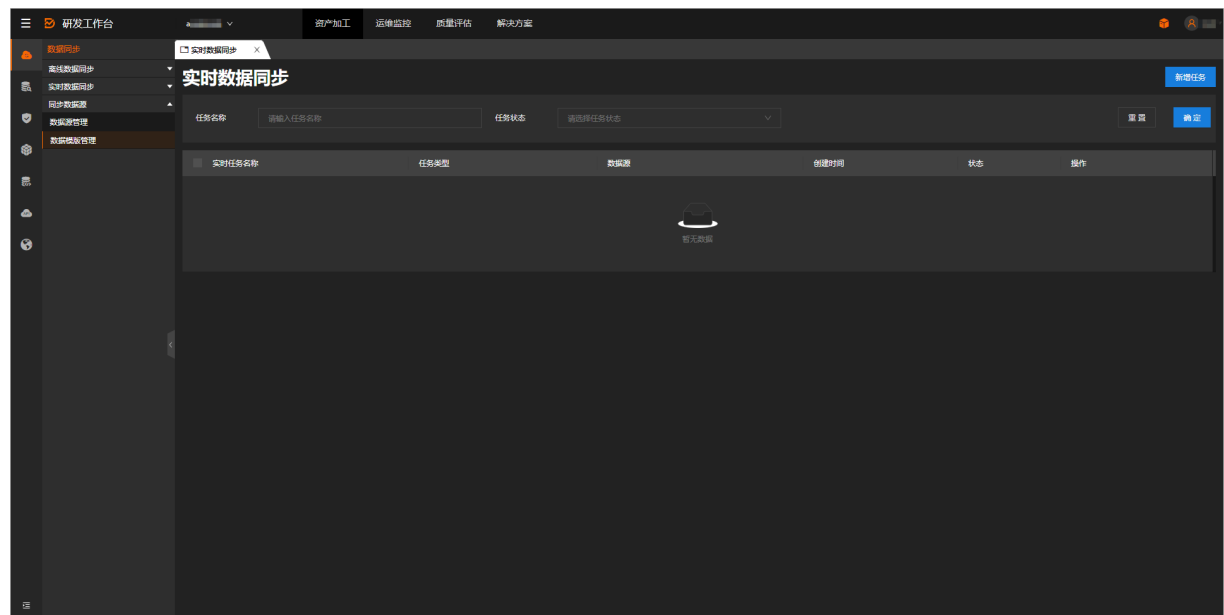
- 支持单表离线任务、批量离线任务的新建。

- 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。
- 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、Rabbit MQ、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。

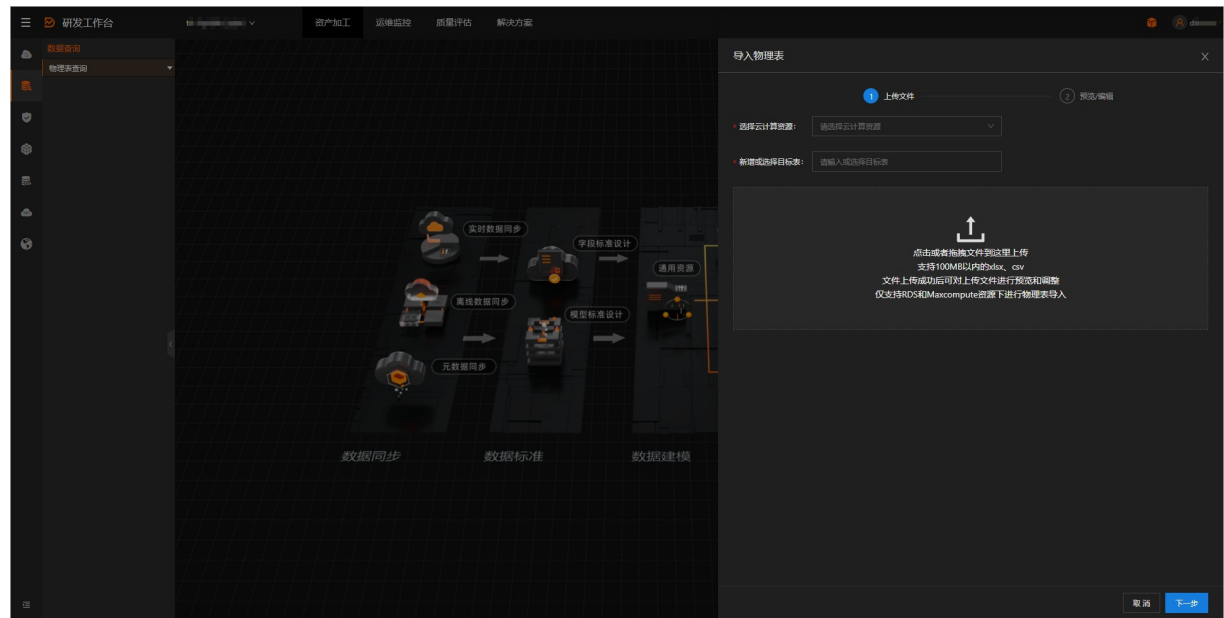


6.2.1.2. 数据查询

数据探查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。

物理表导入

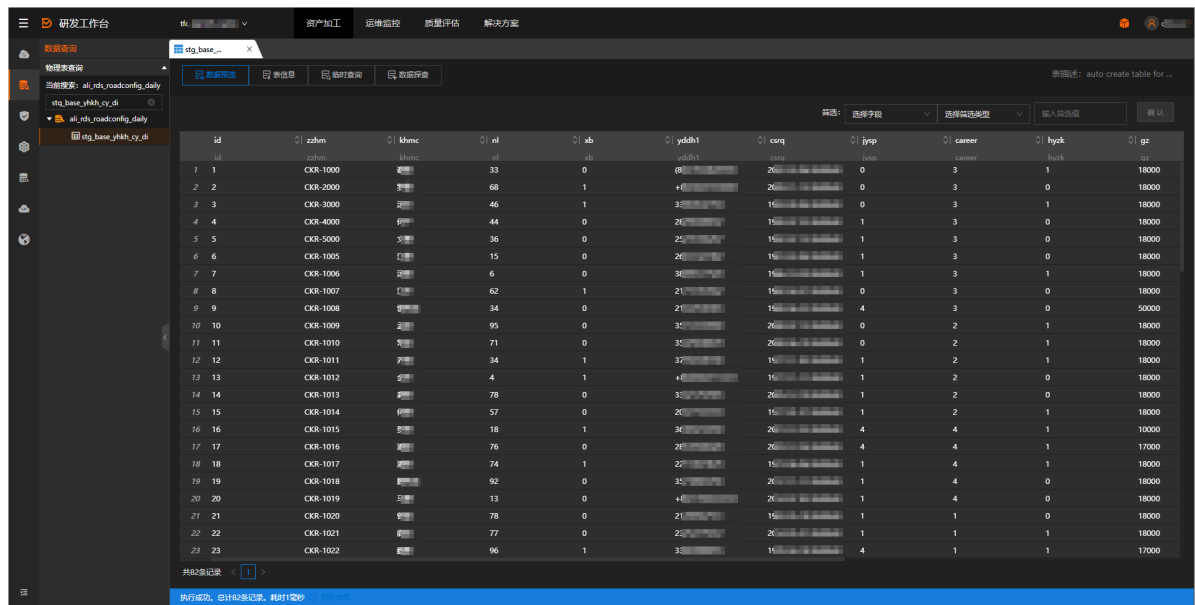
支持100 MB以内的XLSX、CSV文件快捷向RDS MySQL和MaxCompute资源进行导入，支持对要导入的文件进行预览和编辑。



数据预览

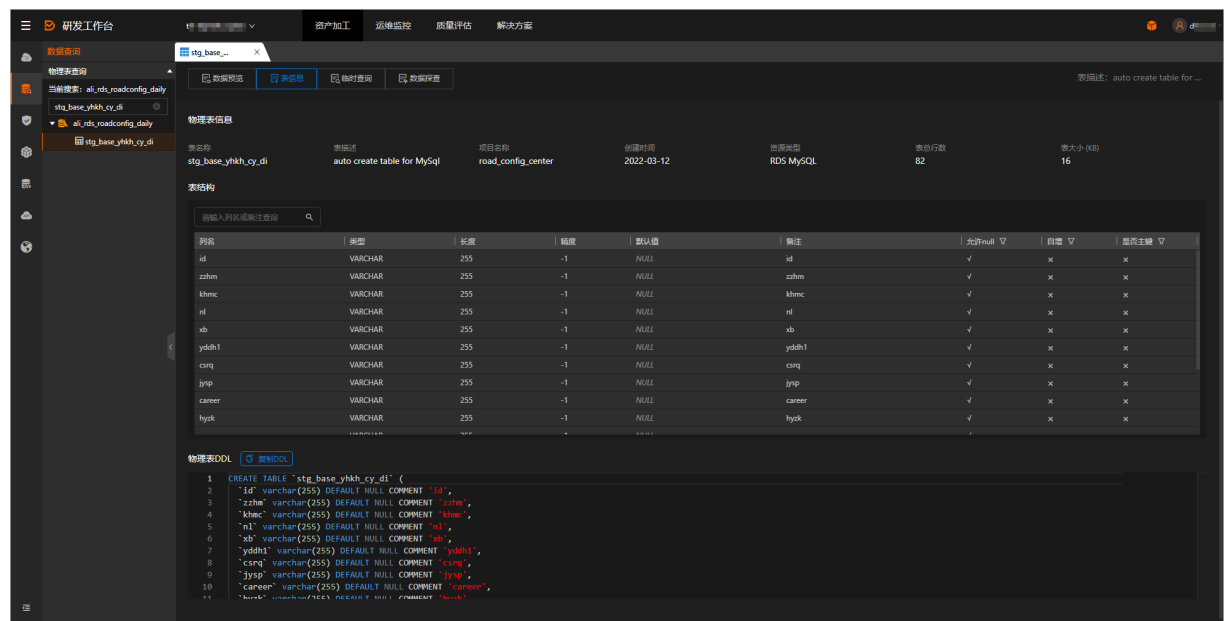
支持不同云计算资源数据的快速预览，可为开发人员和数据应用人员提供全面直观的数据呈现。

- 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。



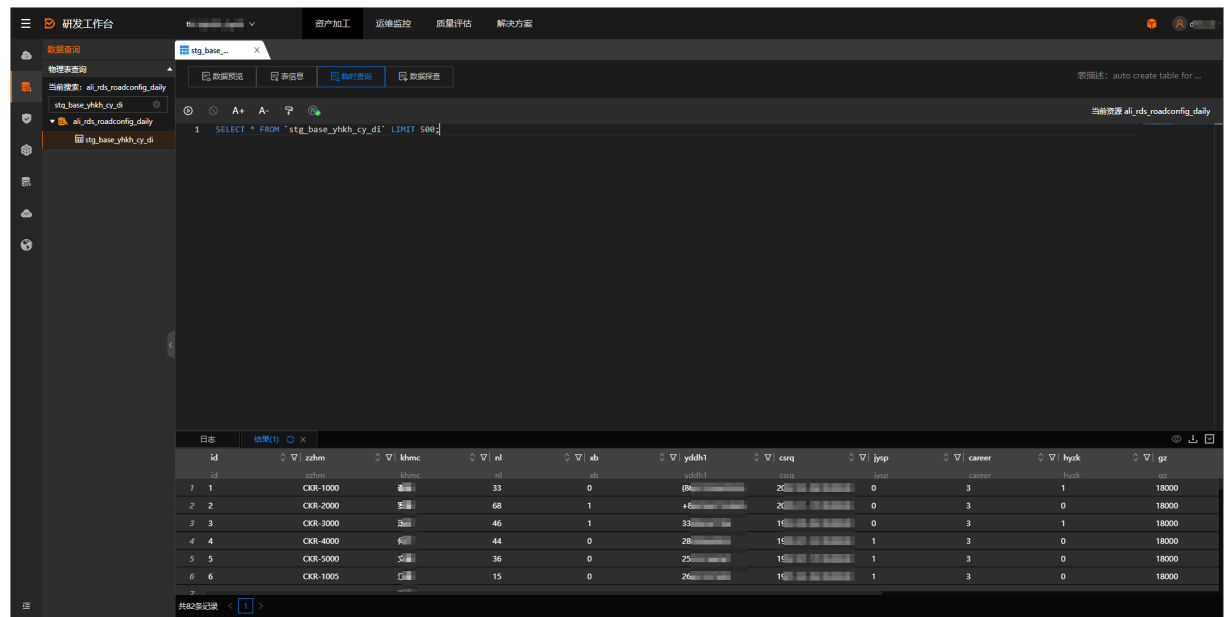
表信息查看

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

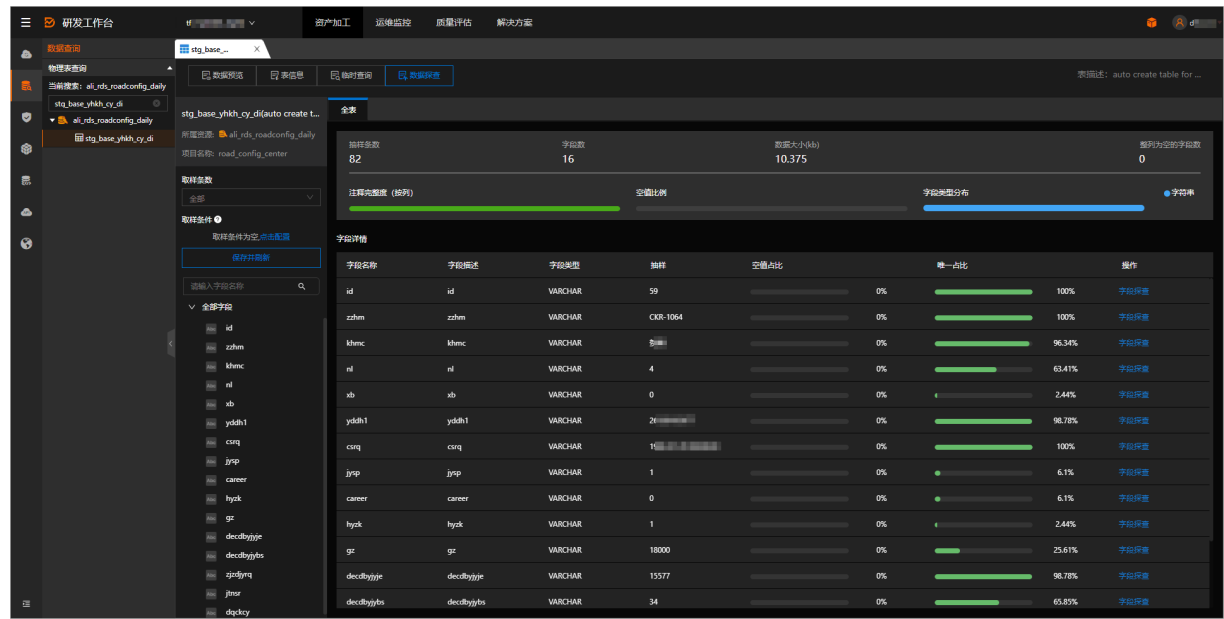
支持通过SQL语句查询预览不同云计算资源的数据资源，可为开发人员和数据应用人员提供全面直观的数据呈现。



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。

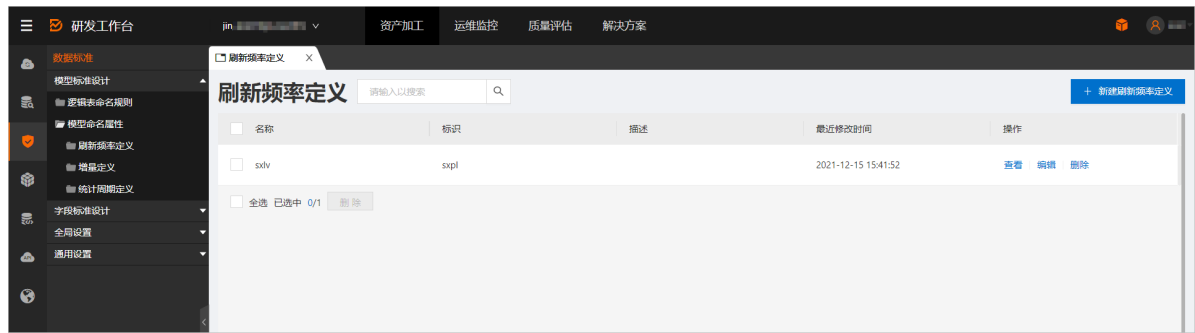


6.2.1.3. 数据标准

数据标准为用户提供模型和字段标准设计和管理的工具。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

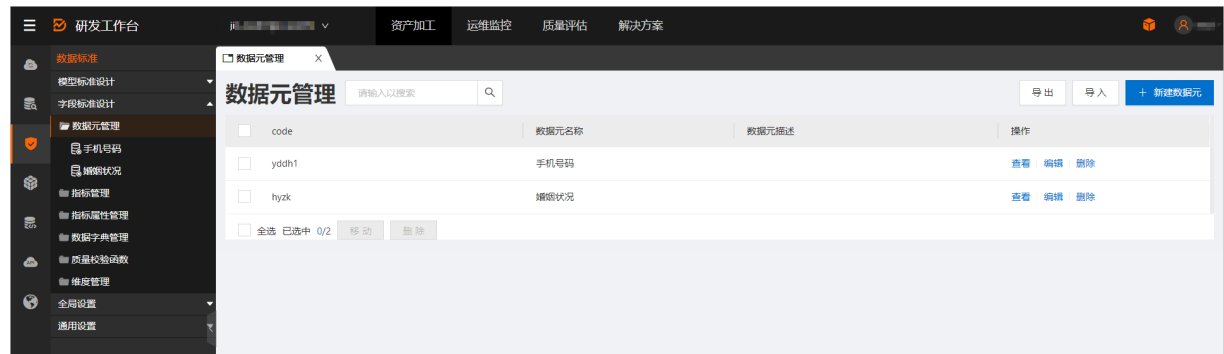
- **逻辑表命名规则：**支持对逻辑表命名标准的规范设计，可基于模型命名属性，对模型命名规则进行配置，规范统一后续模型设计的命名标准。
- **逻辑表命名属性：**提供对逻辑表命名组成中的一些常用属性进行自定义配置功能，如刷新频率、增量、统计周期。



字段标准设计

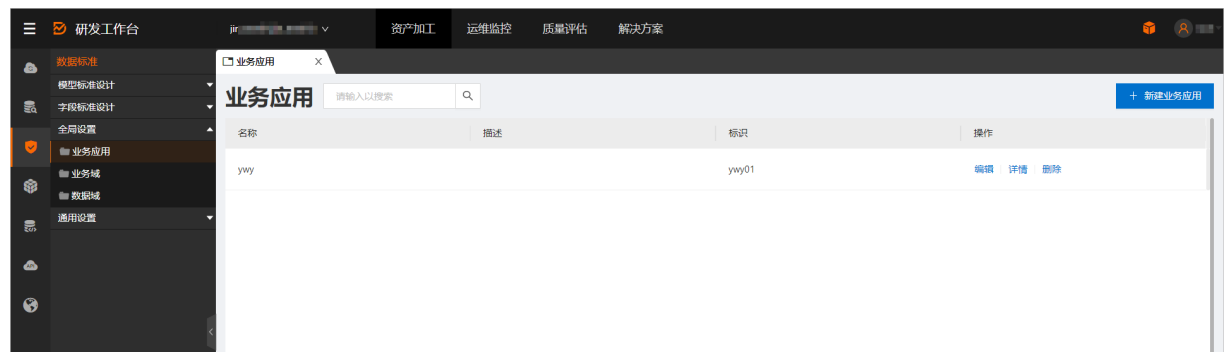
- **数据元管理：**提供对数据元的标准定义和统一管理，通过分析业务流程，抽象关键业务对象、业务对象属性，配置数据元的类型、长度、质量校验函数和引用数据字典，为逻辑表标准化设计及数据治理过程中的质量监控提供规则依据；支持数据元导入、导出，方便行业经验积累和跨项目复制。
- **指标管理：**支持指标类型数据元的定义和管理，可通过选择指标属性的方式，快速新建指标数据元，供逻辑表设计时使用。
- **指标属性管理：**支持指标定义的相关属性管理，包含时间修饰、业务修饰、度量属性，帮助快速新建指标数据元。

- **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。



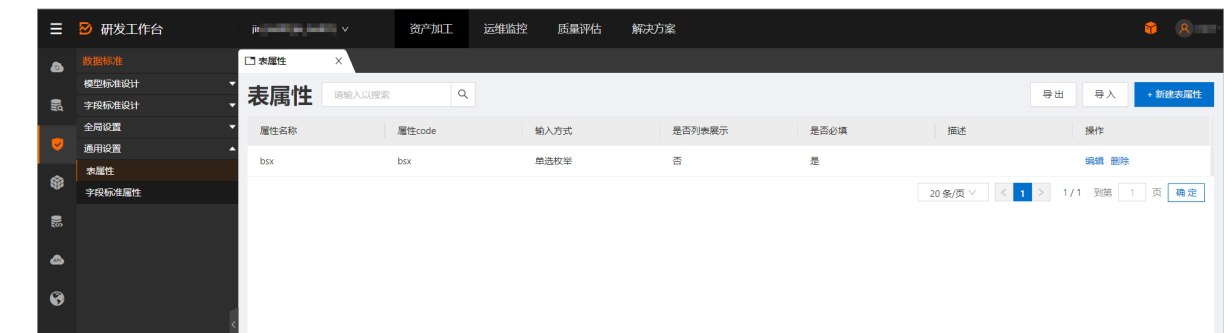
全局设置

- **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

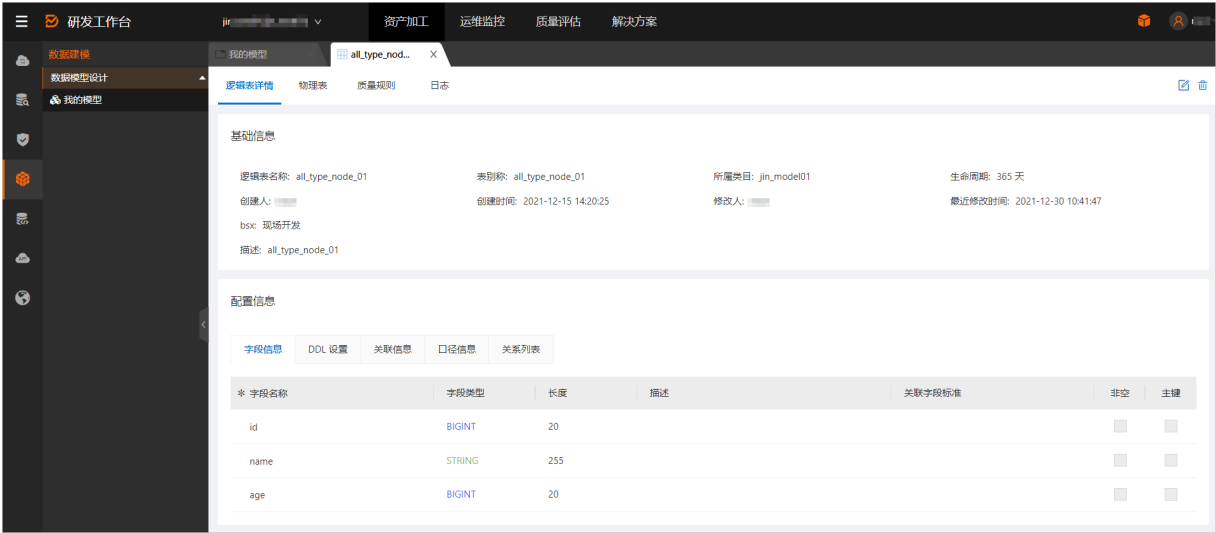
- **表属性**：对逻辑表的属性进行配置、管理。
- **字段标准属性**：对字段标准的属性进行配置、管理。



6.2.1.4. 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。

6.2.1.5. 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 workflow 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。

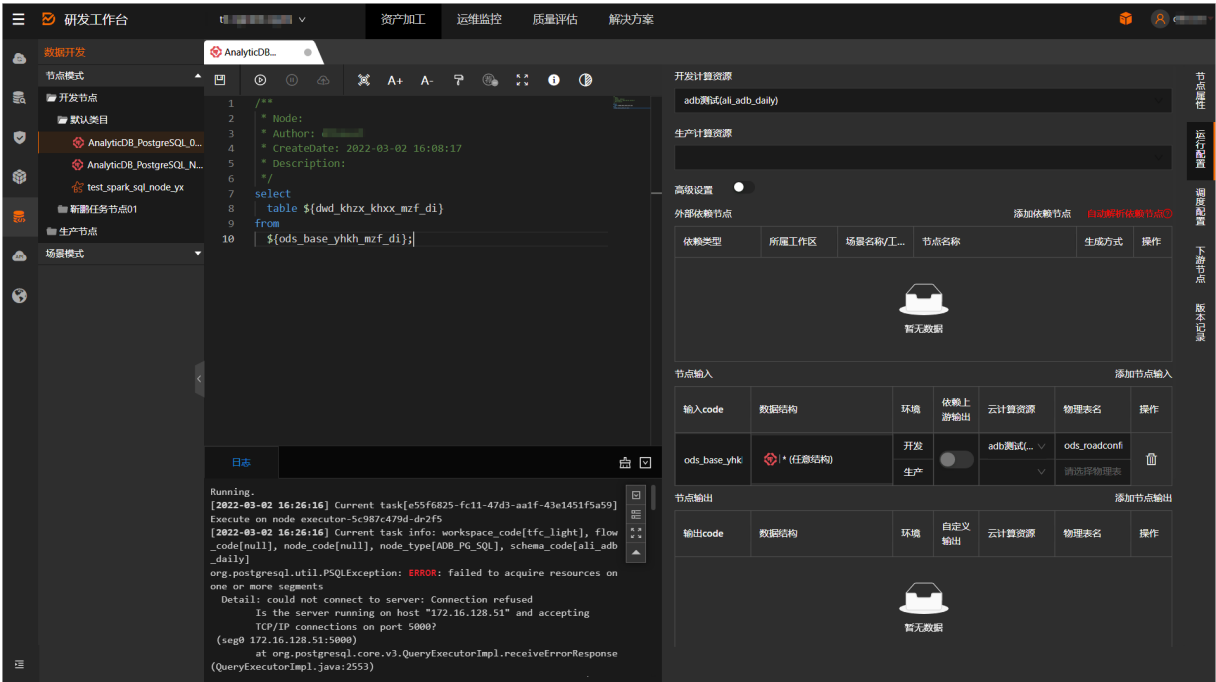
数据开发规范了各个计算节点的接口，在一定程度上提高了计算节点复用的能力，降低了多人协作开发的难度。提供了跨多种存储/计算平台的计算节点工作流程配置和快速部署上线的能力。因此，用户在构建数据应用时，只需从解决实际某一业务场景出发，组合多种类型的计算节点来配置计算节点工作流程，从而快速解决业务工作流并依托数据开发的跨平台多异构任务的一键部署和上线能力，极大的提高了工程效率。

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

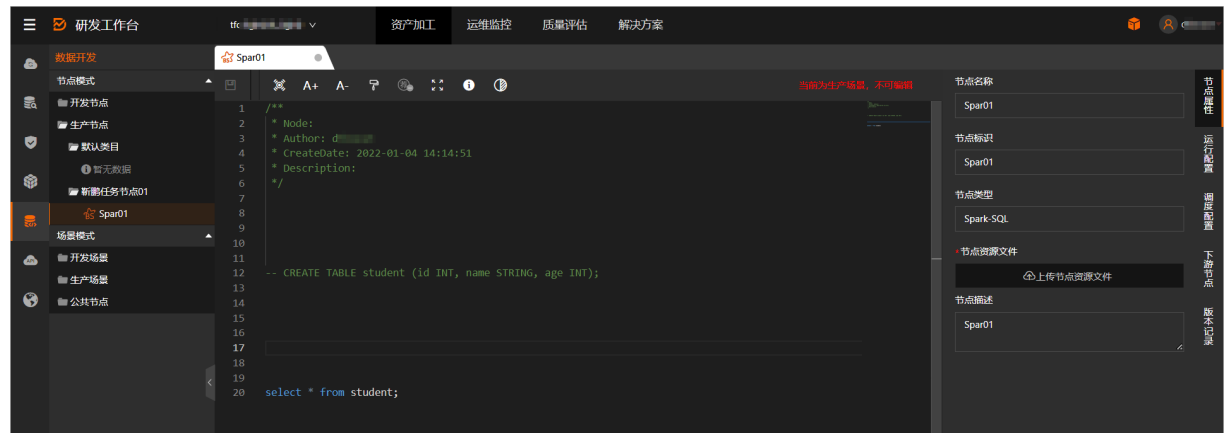
支持节点任务类目的新建、编辑和删除以及支持离线节点注册，对节点脚本、自定义函数、节点输入、节点输出、节点参数进行配置和管理以及支持配置节点的依赖关系以及支持节点任务的部署、运行和上线。



- 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。

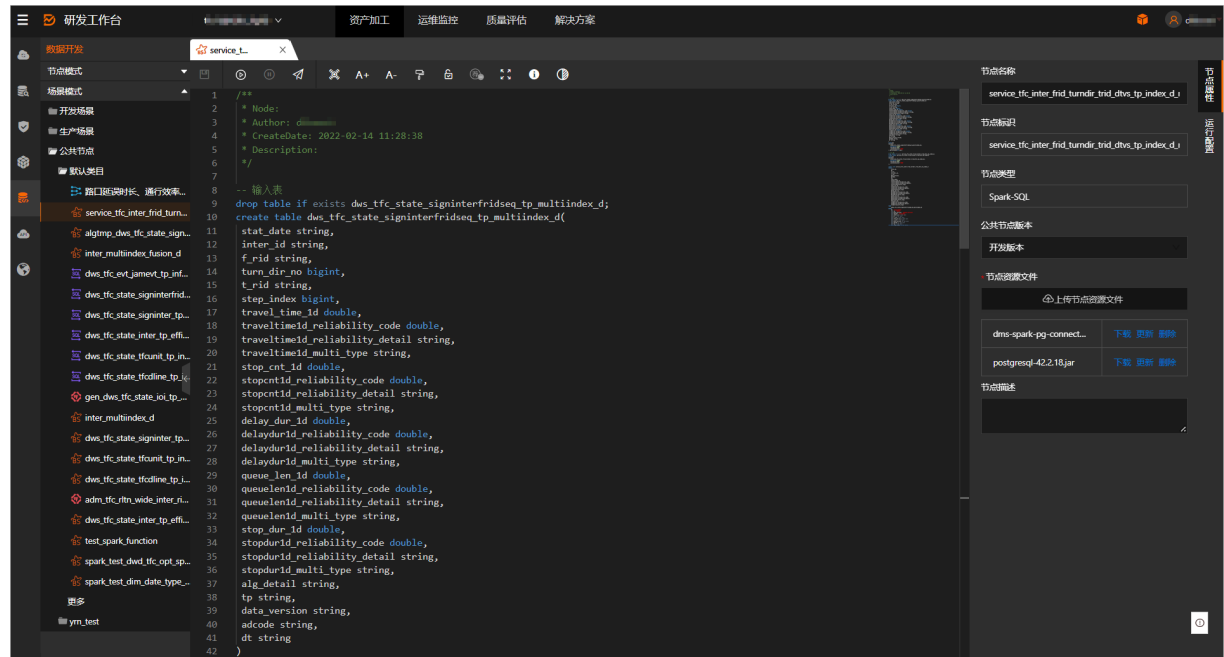


场景模式

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，通过拖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。

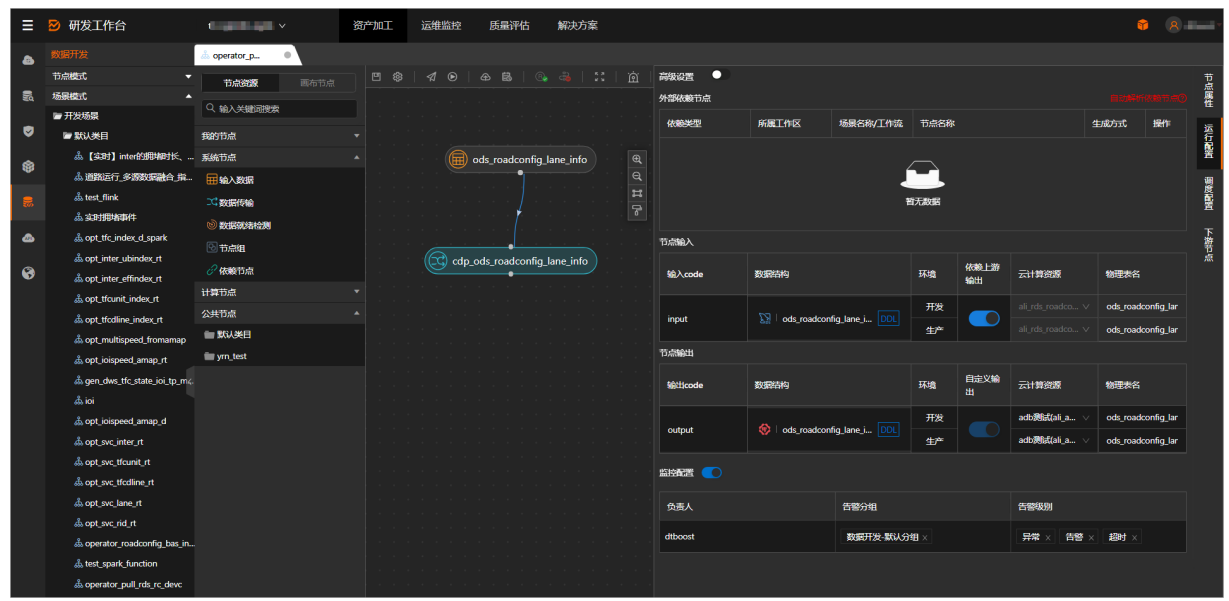


- 支持公共节点注册，对公共节点脚本、自定义函数、公共节点输入、公共节点输出、公共节点参数进行配置和管理。
- 支持根据SQL语言，自动解析离线节点的输入输出。

- 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

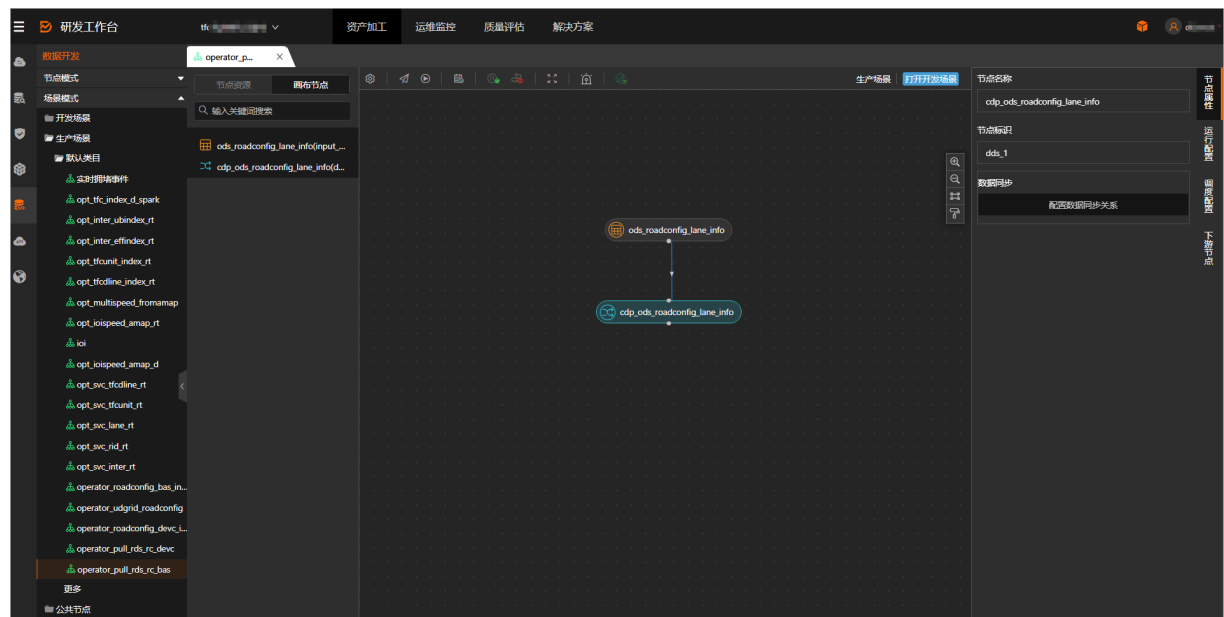
支持开发场景类目管理。支持开发场景引用公共节点。支持数据处理流程编排（多个计算节点组合成一个任务流），用户可灵活的通过可视化方式自定义任务流。支持批量、流式等多种类型的节点在一个场景内的混合编排。支持开发环境中场景的部署、运行和上线，并将开发环境中调试通过的场景发布到生产环境。



- 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。



6.2.1.6. 统一服务

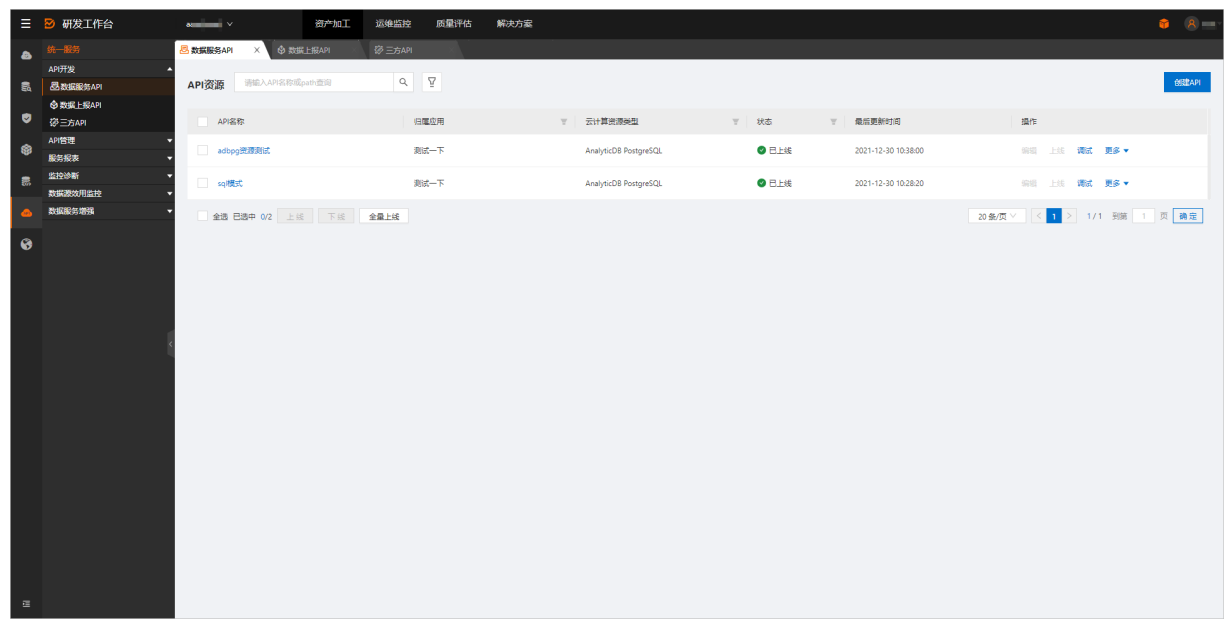
在数据中台建设过程中，需要搭建核心处理平台，并按相关标准要求，将源数据或治理处理后的数据，按照“全面、高效、统筹”的共享原则，按需低延时、精准化的传递给用户，方便更多应用消费数据，实现数据的价值。协同管控各种组织机构，为智慧化运行提供支撑，实现跨层级、跨部门、跨行业、跨区域信息即时共享，做到高可靠、低延时。同时数据被应用的使用情况评估也作为数据中台日常运营的重要指标项，提出更多数据生产需求，完成数据从生产到使用到评估的闭环。

统一服务提供了如下数据服务能力。

- 表转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。目前表转API提供三种形式注册方式，分别为单表转API，sql转API。数据源可以是OTS、RDS，MySQL等数据存储中的数据表。
- 流式数据服务的源头为DataHub或RabbitMQ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、RabbitMQ。
- 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；面向数据资产管理提供服务统计分析、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和数据上报API等三种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。

?

说明

数据API即表转API。

- **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和Object Storage Service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**：将数据写入的API，用户调用API时将数据上报到服务端，目前支持ApsaraDB for RDS、DataHub作为数据上报目标端。
- **API的生命周期管理**：API支持数据API、三方应用API、数据上报API三种。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。数据API也即表转API。
- **API权限管理**：包含API共享、已获授权API和已对外授权API。

上架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。

已获授权API

支持对已获授权API列表进行查看，支持评价API、调试API、下载API使用文档、调用SDK。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

支持查看通过数据资产中心已完成对外授权审批流程的API列表。支持已对外授权API列表的查看，支持查看授权的AK信息、暂停授权等操作。

API管理

API列表

研发工作台

数据服务API

API列表

API列表

API名称	所属应用	类型	状态	创建者	上架/下架时间	操作
交管_车牌日期信息	交管服务	数据API	待上架	dtb...		授权 上架 日志
交管_车牌BaseInfo	交管服务	数据API	已上架	dtb...	2022-01-12 16:55:51	授权 下架 日志
交管_车牌最新实时指标	交管服务	数据API	已上架	dtb...	2022-01-12 16:37:26	授权 下架 日志
交管_车牌路口设备相关信息	交管服务	数据API	已上架	dtb...	2022-01-12 16:35:05	授权 下架 日志
交管_车牌信息ioi	交管服务	数据API	已上架	dtb...	2022-01-12 16:34:34	授权 下架 日志
交管_车牌路段rdseg基本信息	交管服务	数据API	已上架	dtb...	2022-01-12 16:33:17	授权 下架 日志
交管_车牌和ID指定日期分时间片...	交管服务	数据API	已上架	dtb...	2022-01-12 16:33:03	授权 下架 日志
交管_车牌对应的信息点	交管服务	数据API	已上架	dtb...	2022-01-12 16:32:06	授权 下架 日志
交管_车牌分方向道路droad基本...	交管服务	数据API	待上架	dtb...		授权 上架 日志

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。

已获授权API

研发工作台

数据服务API

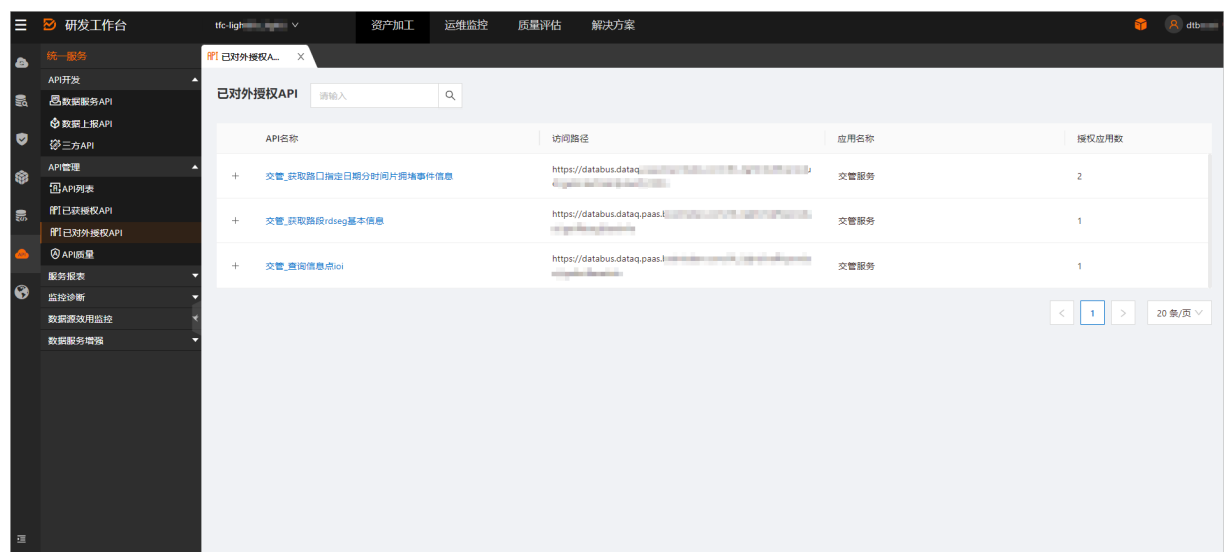
已获授权API

已获授权API

API名称	授权应用	到期时间	状态	申请人	我的评价	操作
交管_车牌信息ioi	解决方案应用xy_01	2022-02-11 00:00:00	已暂停	dtb...	无	授权详情 API评价 调试 更多
交管_车牌路段rdseg基本...	数据资产	2022-02-11 00:00:00	已暂停	dtb...	无	授权详情 API评价 调试 更多
交管_车牌路口指定日期...	数据资产	2038-01-01 00:00:00	已授权	dtb...	无	授权详情 API评价 调试 更多
交管_车牌路口指定日期...	解决方案应用xy_01	2038-01-01 00:00:00	已授权	dtb...	无	授权详情 API评价 调试 更多

支持对已获授权API列表进行查看，支持评价API、调试API、下载API使用文档、调用SDK。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。

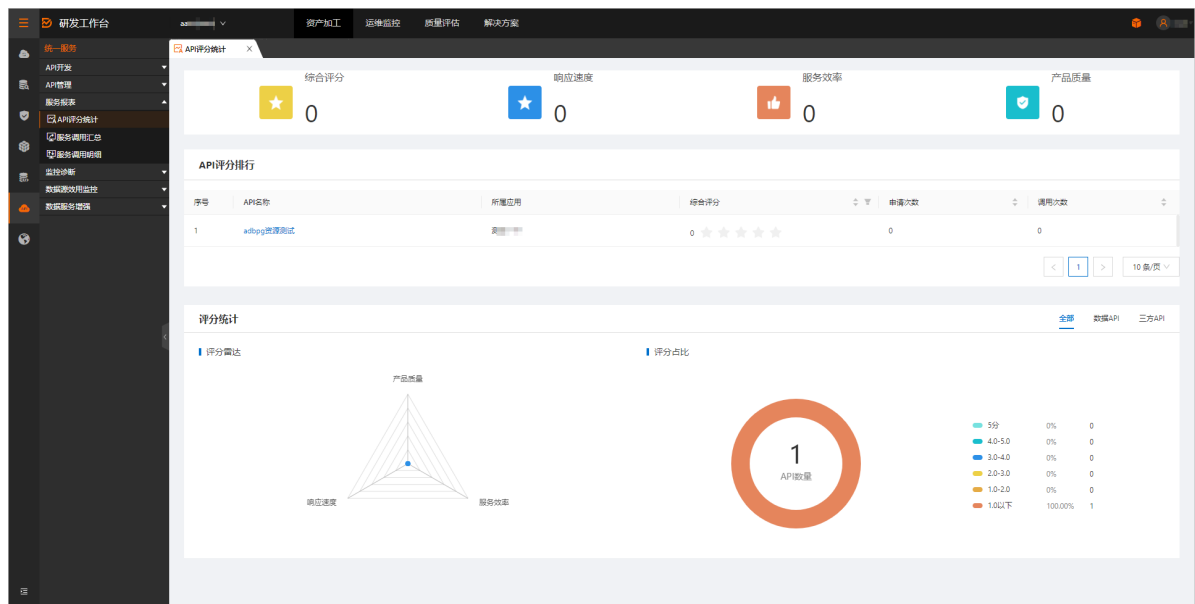
已对外授权API



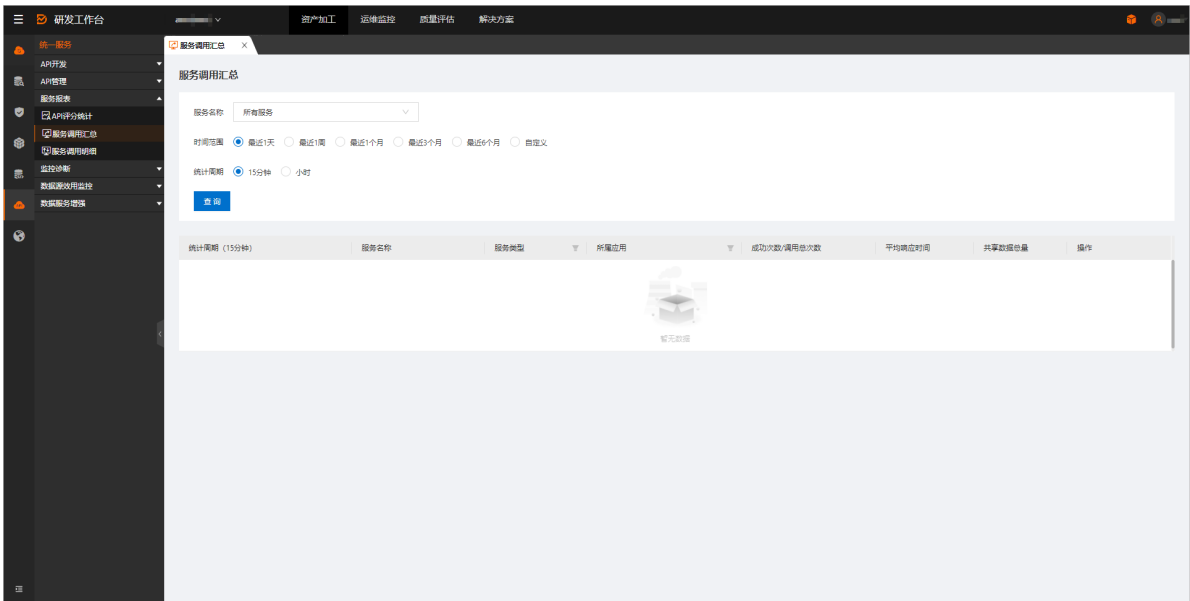
支持查看通过数据资产中心已完成对外授权审批流程的API列表。支持已对外授权API列表的查看，支持查看授权的AK信息、暂停授权等操作。

服务报表

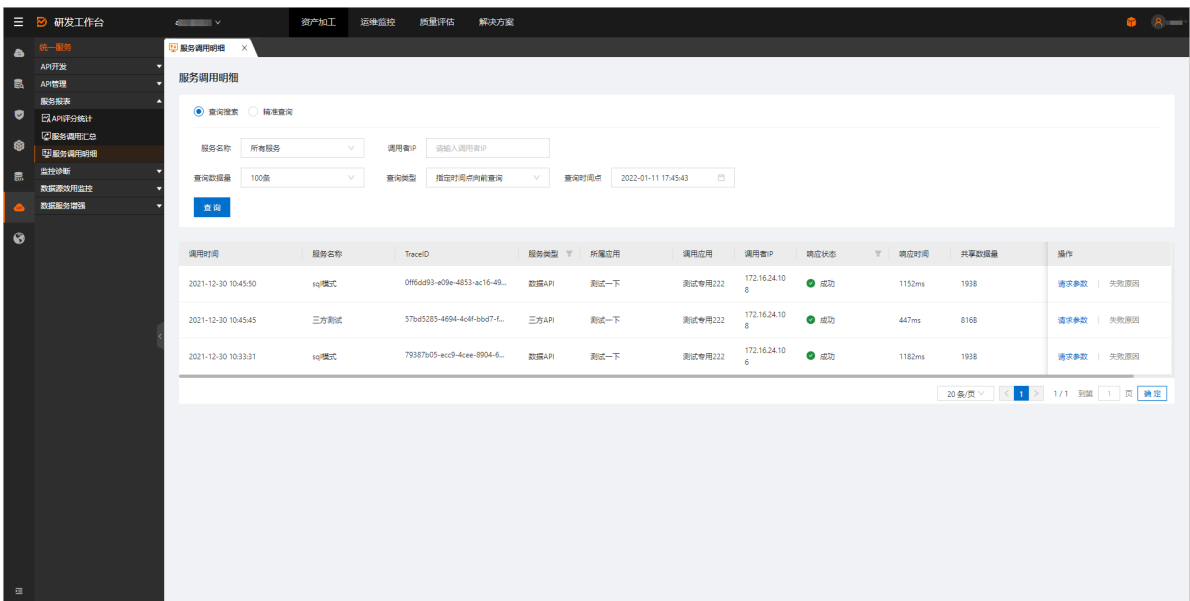
- API被调用后，支持对API的响应速度、服务效率、产品质量的进行评分统计，并将统计结果以报表形式展示。



- 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。

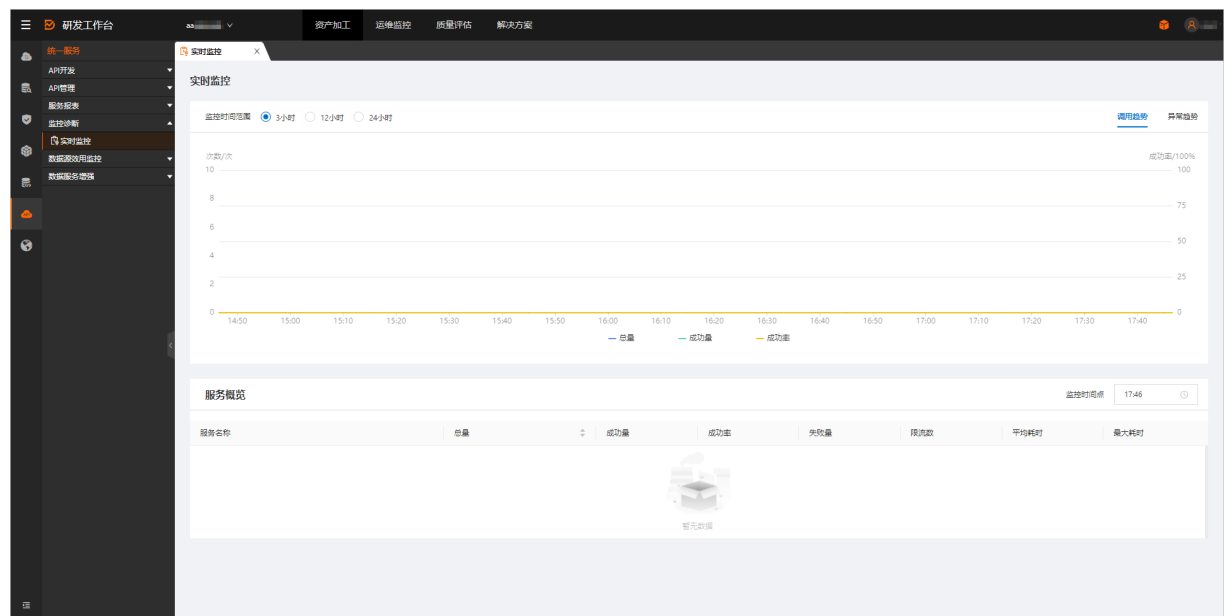


- 服务调用明细支持单次调用的明细信息查看，包括：调用时间、TraceID、响应状态、响应时间、请求参数、失败原因。



实时监控

监控最近24小时内服务的实时调用趋势及异常调用。支持查看监控时间点上的服务调用概览和服务调用明细。



6.2.1.7. 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。

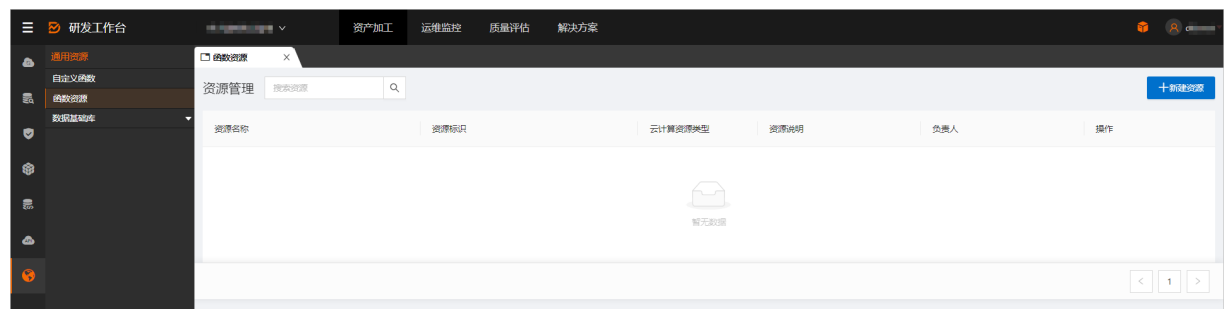
自定义函数

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。



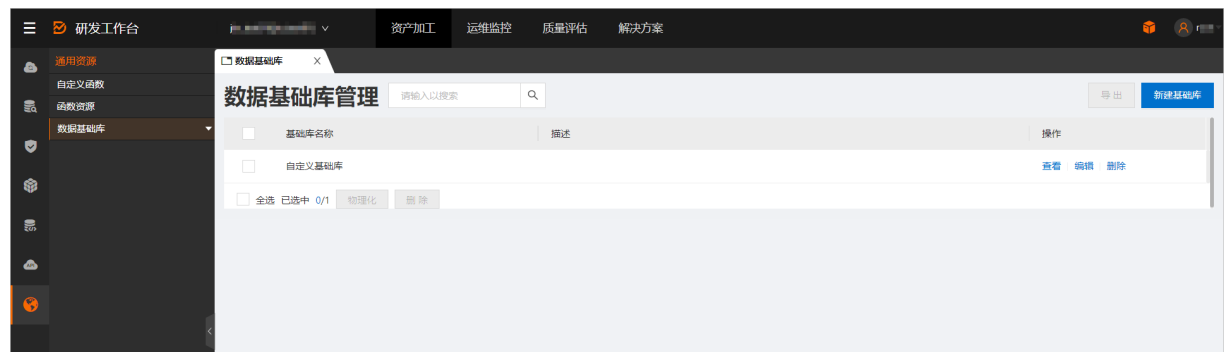
函数资源

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。



6.2.1.8. 运维监控

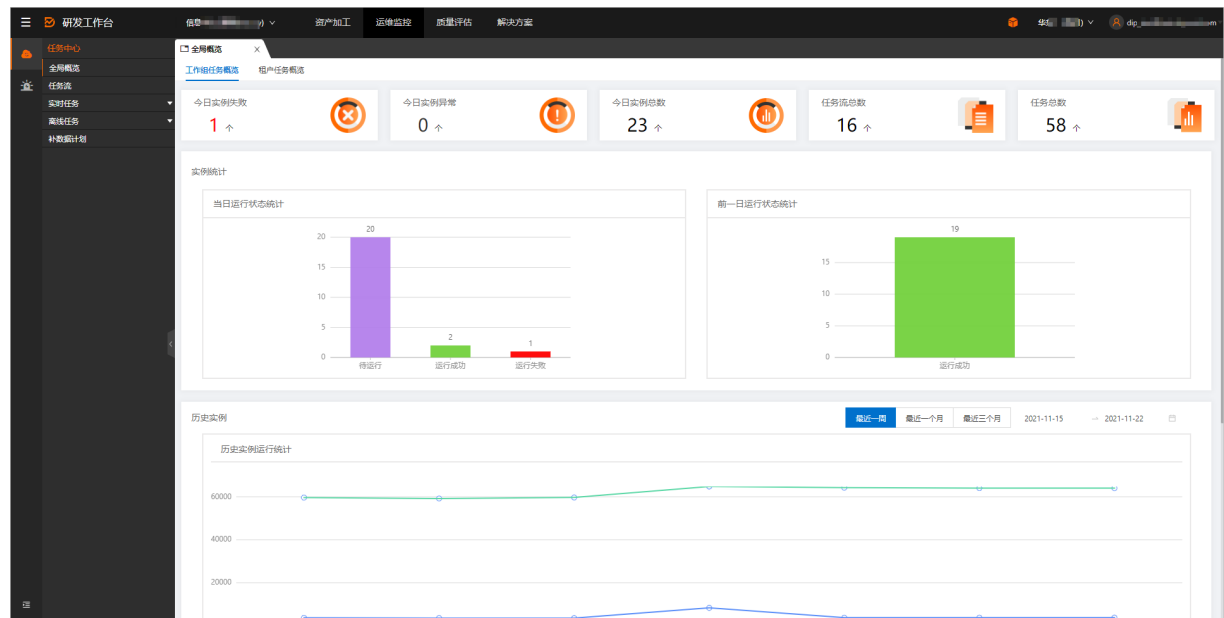
6.2.1.8.1. 任务中心

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，每次运行时会产生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。

研发工作台

任务中心

全局概览

任务流

实时任务

离线任务

补数据计划

任务流列表

业务类型

请输入以搜索

任务流名称	任务流标识	任务流描述	业务类型	修改时间	部署状态	操作
(画像任务)批量用户	kh_wxzdnh	标签工厂Flow	标签方案	2021-11-22 14:27:22	已部署	运维
stg_qcheck	DG_10003947_stg_ykhk_check_try	-	质量评估	2021-11-17 17:28:40	已部署	运维
test1_doy2	DG_10003947_dataq_hsh_002_try	-	质量评估	2021-11-17 17:27:17	已部署	运维
test1_doy2	DG_10003947_dataq_hsh_002_cycle	-	质量评估	2021-11-17 17:27:14	已部署	运维
stg_qcheck	DG_10003947_stg_ykhk_check_cycle	-	质量评估	2021-11-17 17:26:25	已部署	运维
odt_qcheck	DG_10003947_dataq_hsh_001_try	-	质量评估	2021-11-17 16:19:01	已部署	运维
odt_qcheck	DG_10003947_dataq_hsh_001_cycle	-	质量评估	2021-11-17 16:17:46	已部署	运维
信息中心	xxcx_1537_dev	-	数据开发_场景	2021-11-12 10:35:32	已部署	运维
(画像)标签(年龄和婚恋女性工资状况)创建计划_1635922...	kh_wxyzhf	标签工厂Flow	标签方案	2021-11-05 10:54:23	已部署	运维
对象(hn)加速	kh_jsgjg	标签加速Flow	标签同步	2021-11-05 10:50:01	已部署	运维
对象(jymm)加速	jymm_khbcq	标签加速Flow	标签同步	2021-11-03 11:23:06	已部署	运维
对象(ykh)加速	ykh_peicjg	标签加速Flow	标签同步	2021-11-03 11:19:59	已部署	运维
客户交易次数	kh_wxwzbes	标签工厂Flow	标签方案	2021-11-02 15:13:00	已部署	运维
(画像)标签(月收入过万客户)创建计划_1635759732017	kh_wxmzizu	标签工厂Flow	标签方案	2021-11-01 18:09:38	已部署	运维
信息中心	xxcx_1537_prod	-	数据开发_场景	2021-10-26 13:50:40	已部署	运维
d2c_flow_tenant_1783120651209765_xxcx_cy	d2c_flow_tenant_1783120651209765_xxcx_cy	-	数据同步	2021-10-25 16:38:43	已部署	运维

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的。

研发工作台

任务中心

全局概览

任务流

实时任务

离线任务

补数据计划

实时同步任务

数据源名称

请输入数据源名称

任务状态

请选择任务状态

☒ 仅显示最新实例

☐ 仅显示OpenApi

查询

重置

更多

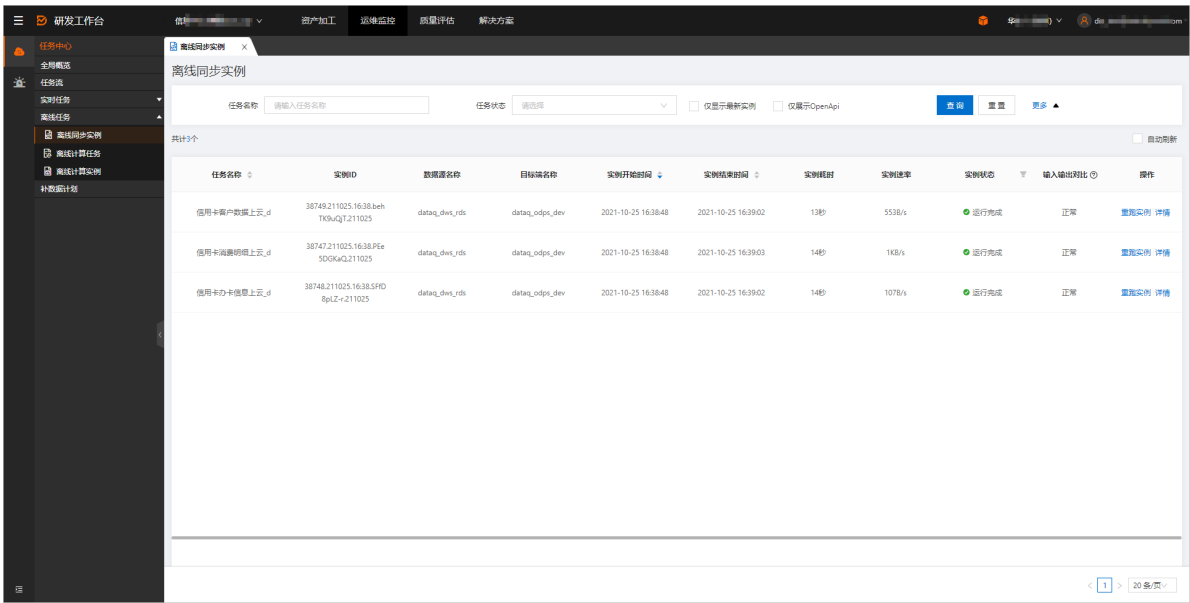
共计0个

任务名称	数据源分类	实例开始时间	实例结束时间	实例耗时	实例速率	实例状态	输入输出对比	操作
暂无数据								

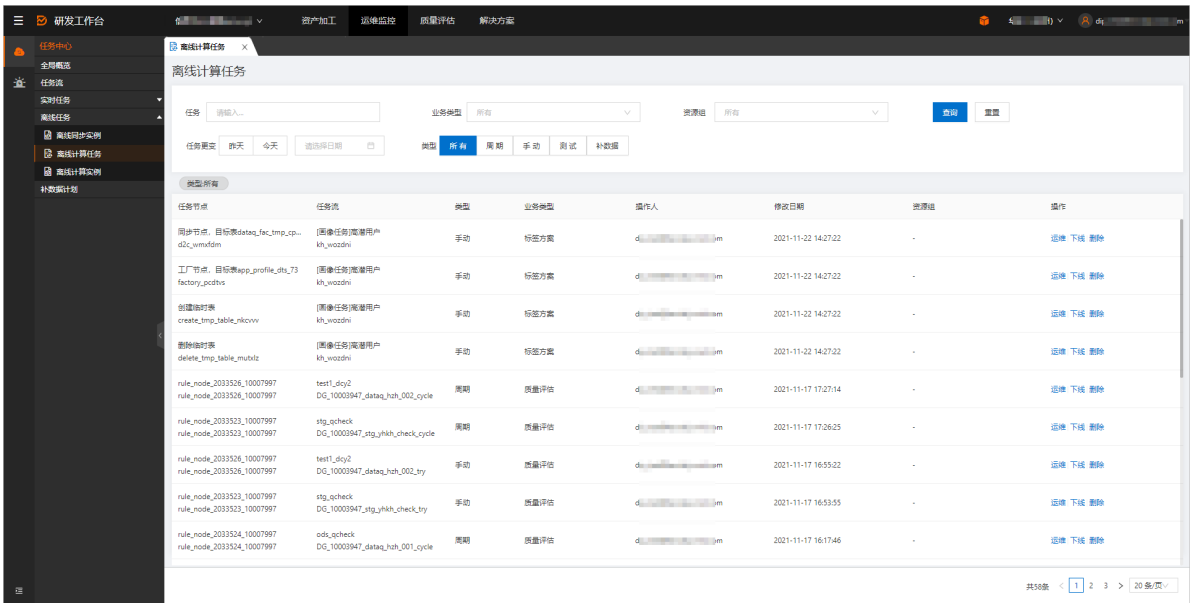
离线任务

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。

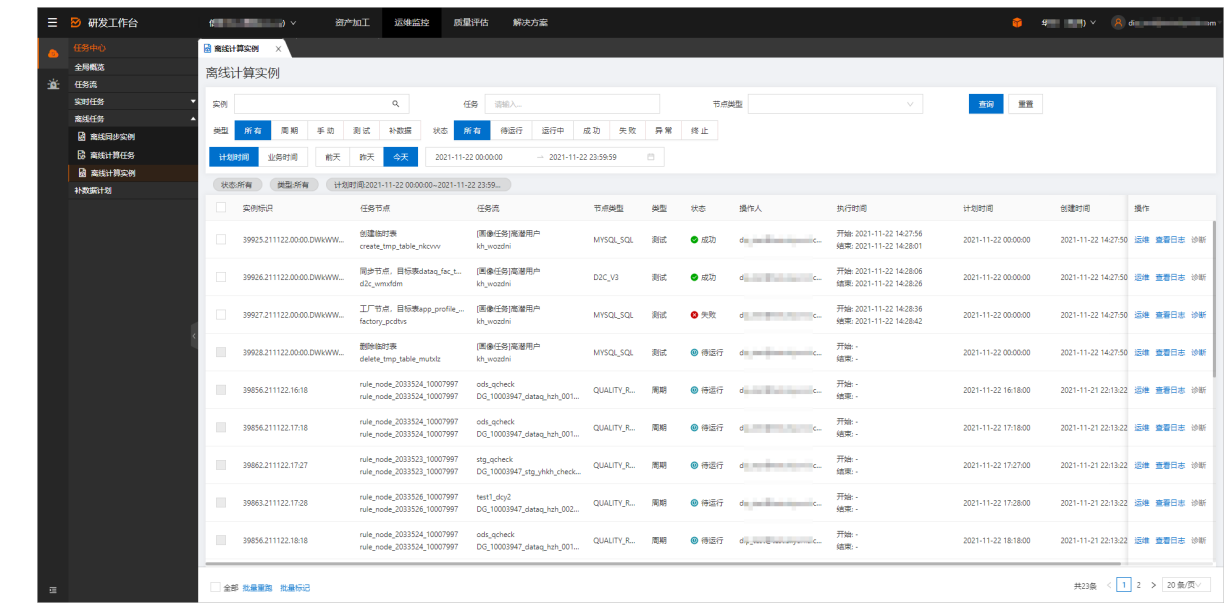
- **离线同步实例：**支持以列表的方式展示离线同步实例信息，并提供离线同步实例属性与运行日志查看；可进行实例重跑的功能。



- **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。



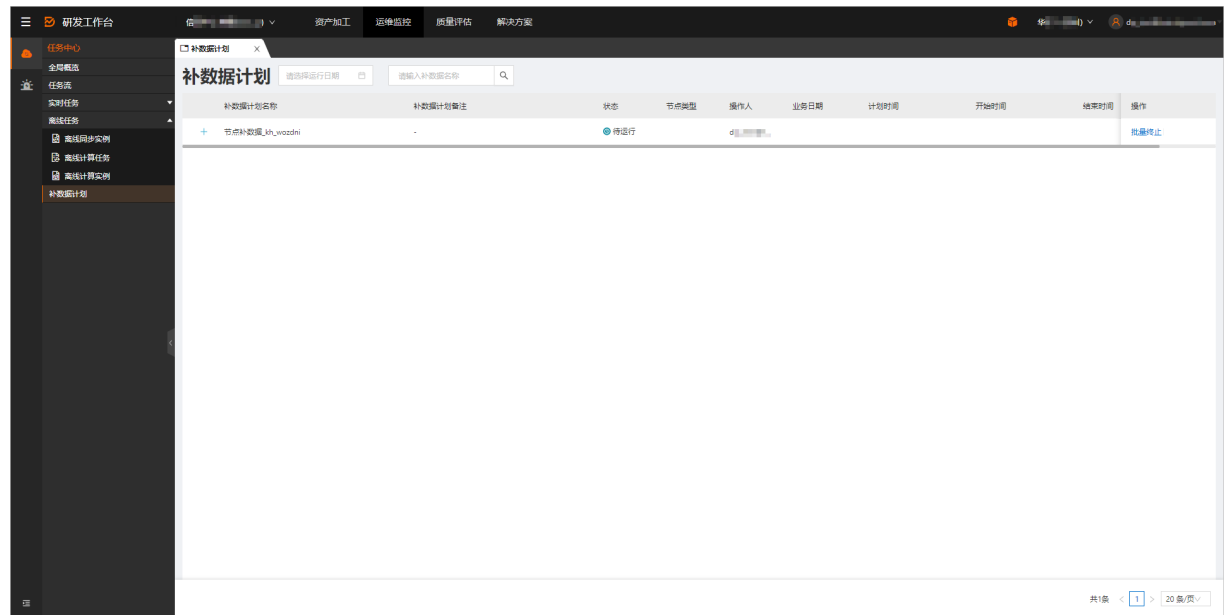
- 离线计算实例：支持离线计算实例信息，并提供运维与日志的查看。



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。



6.2.1.8.2. 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。

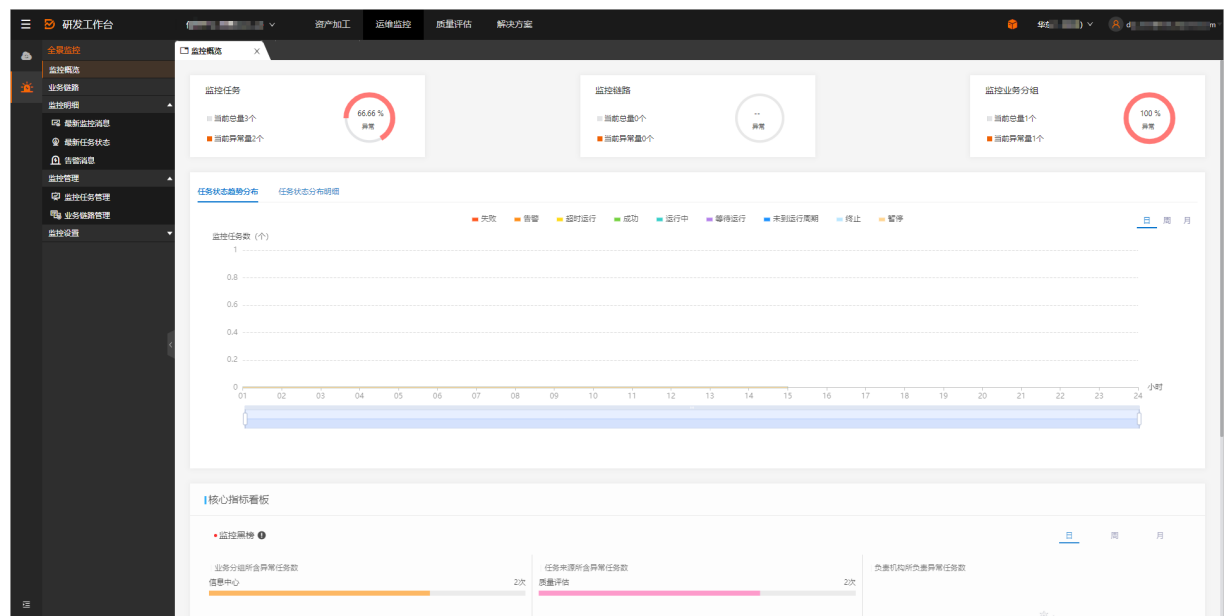
数据全链路保障是指通过采集数据汇聚、数据融合、数据分析挖掘等阶段的各项任务状态信息，对异常状态进行预警和处置，实现对各任务的实时监控和管理。

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。

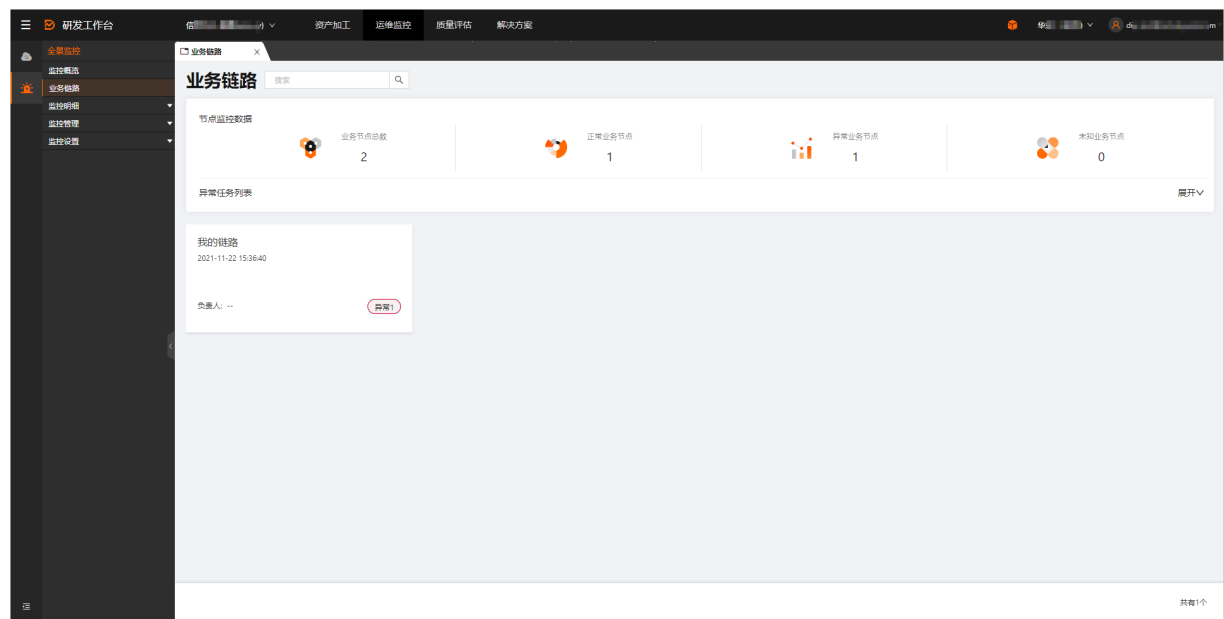
监控概览

提供监控任务、监控链路、监控业务分组等统计信息概览。



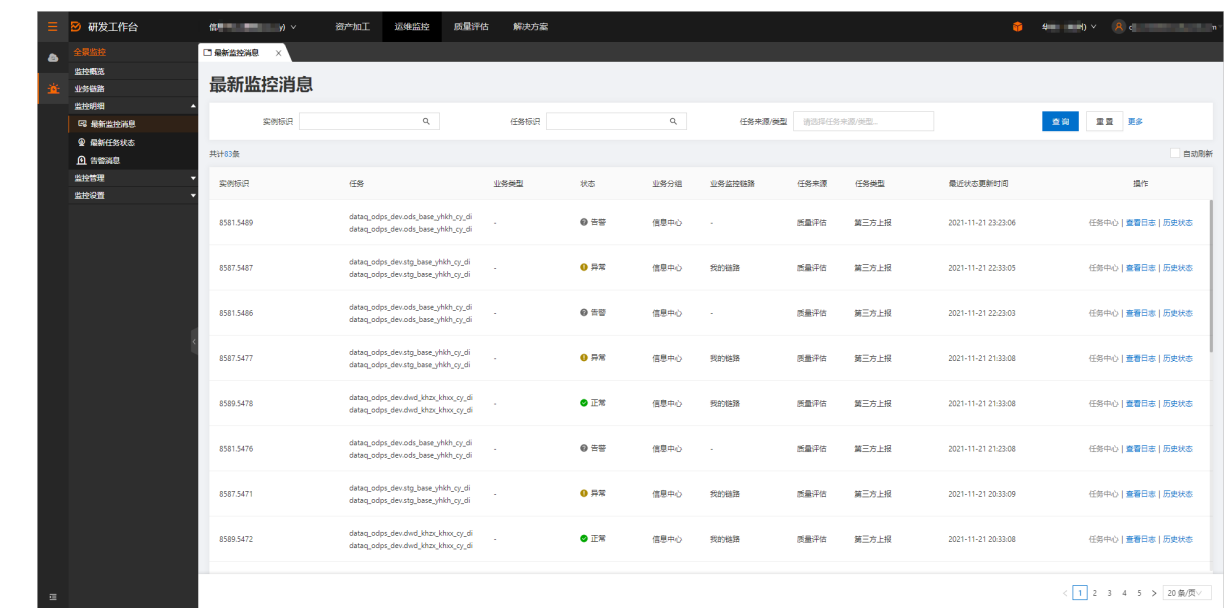
业务链路

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。

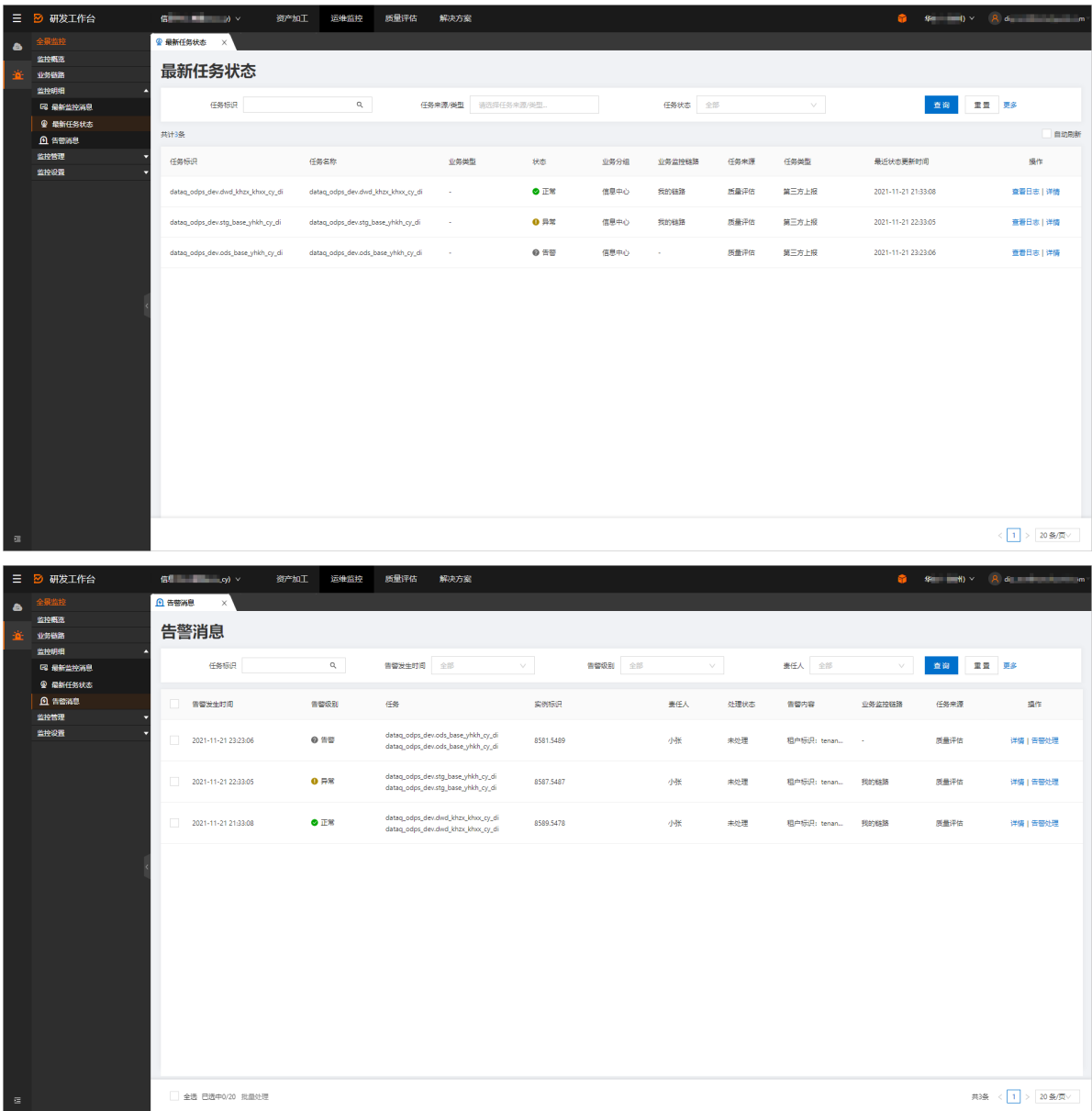


监控明细

- 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。

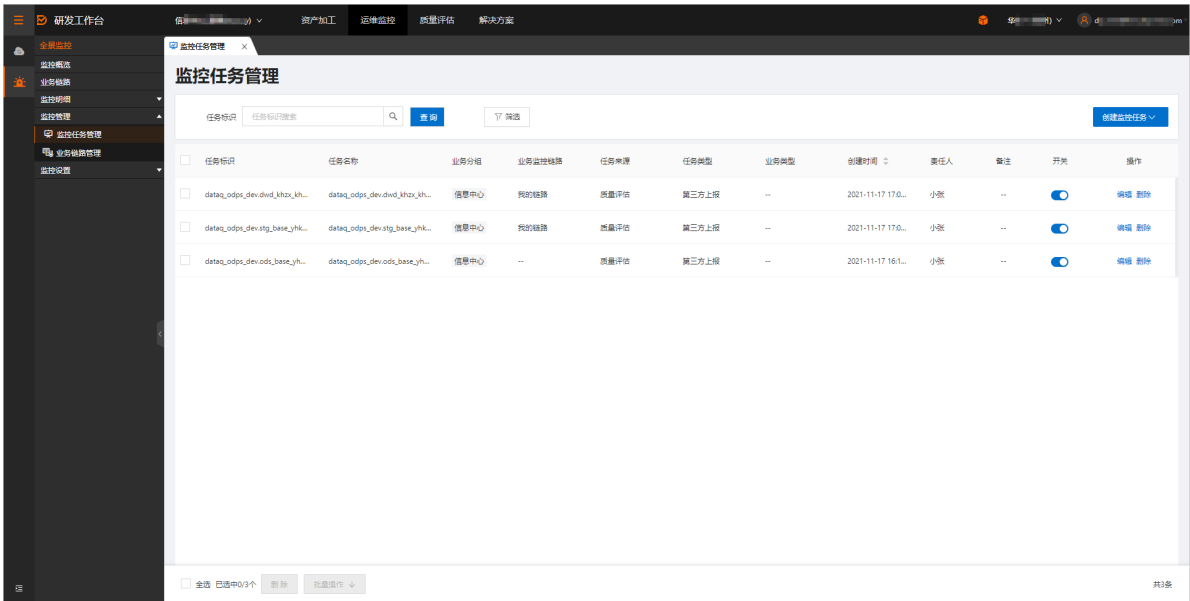


- 支持查看监控任务节点的最近一次告警信息及历史告警信息，并支持手动设置告警处理暂停时间状态。

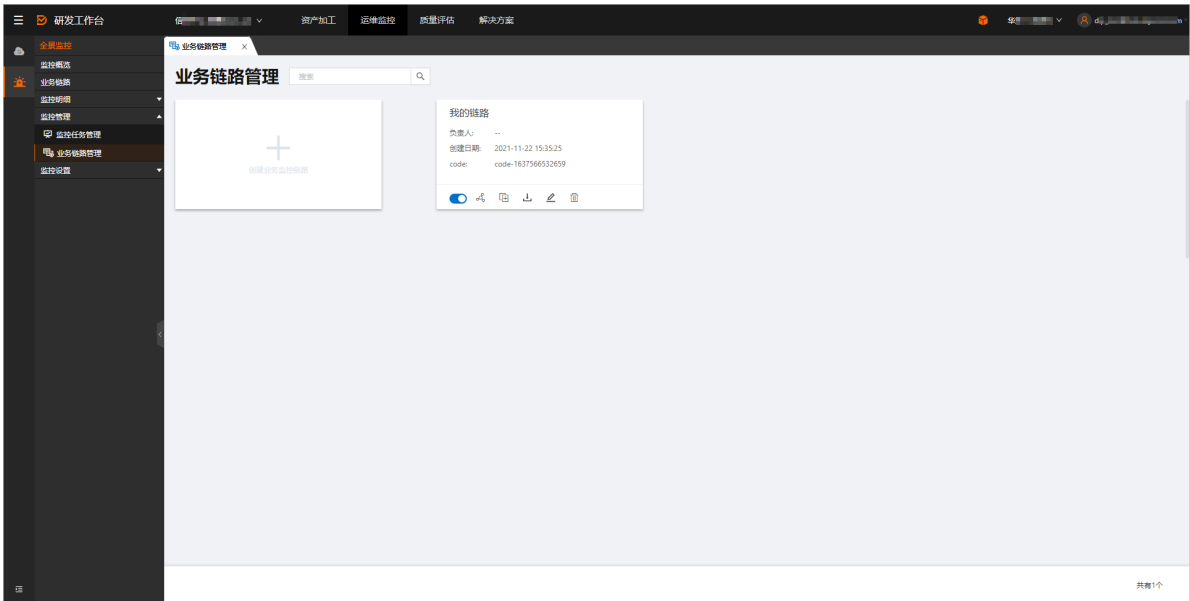


监控管理

- 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。

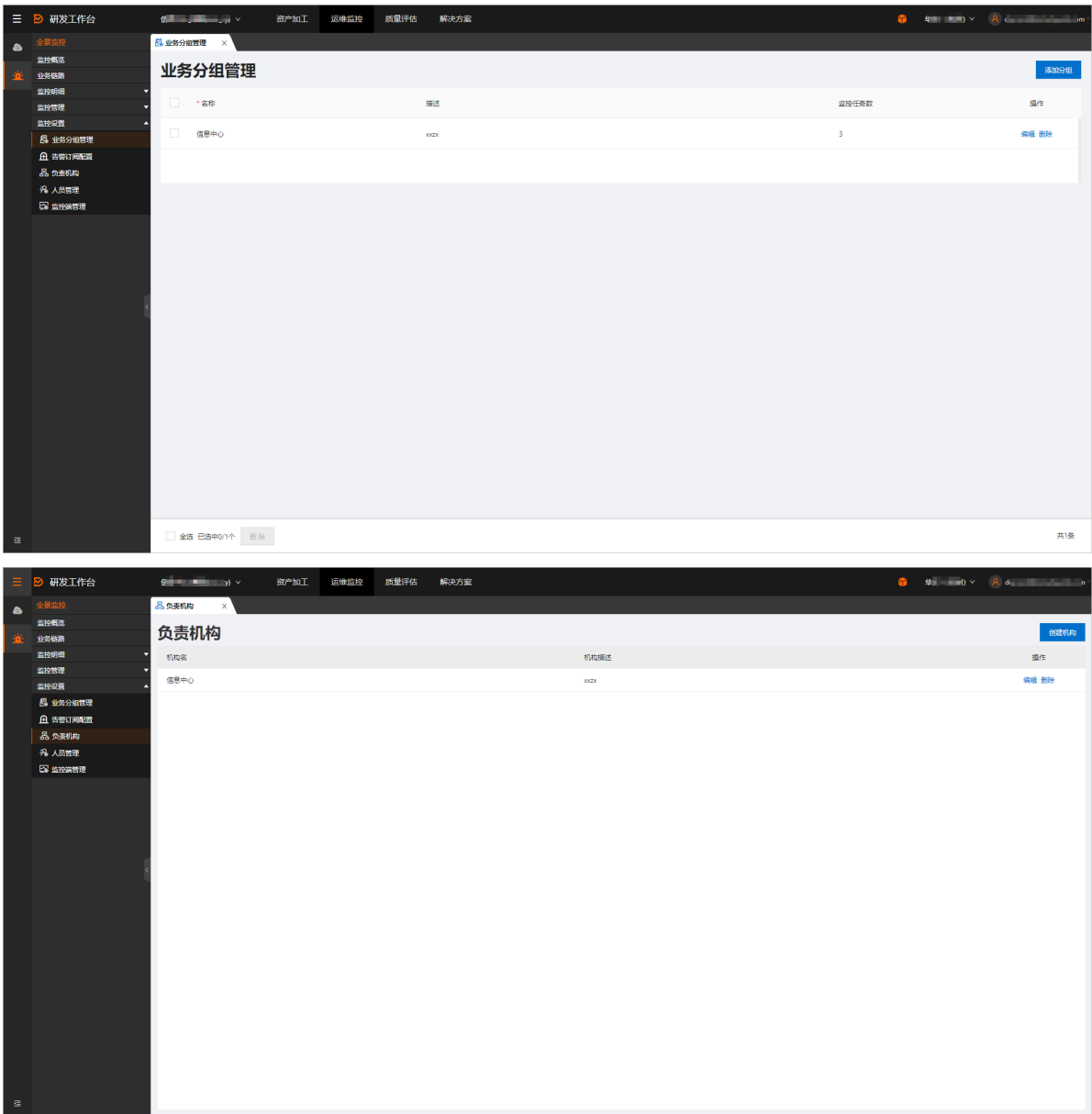


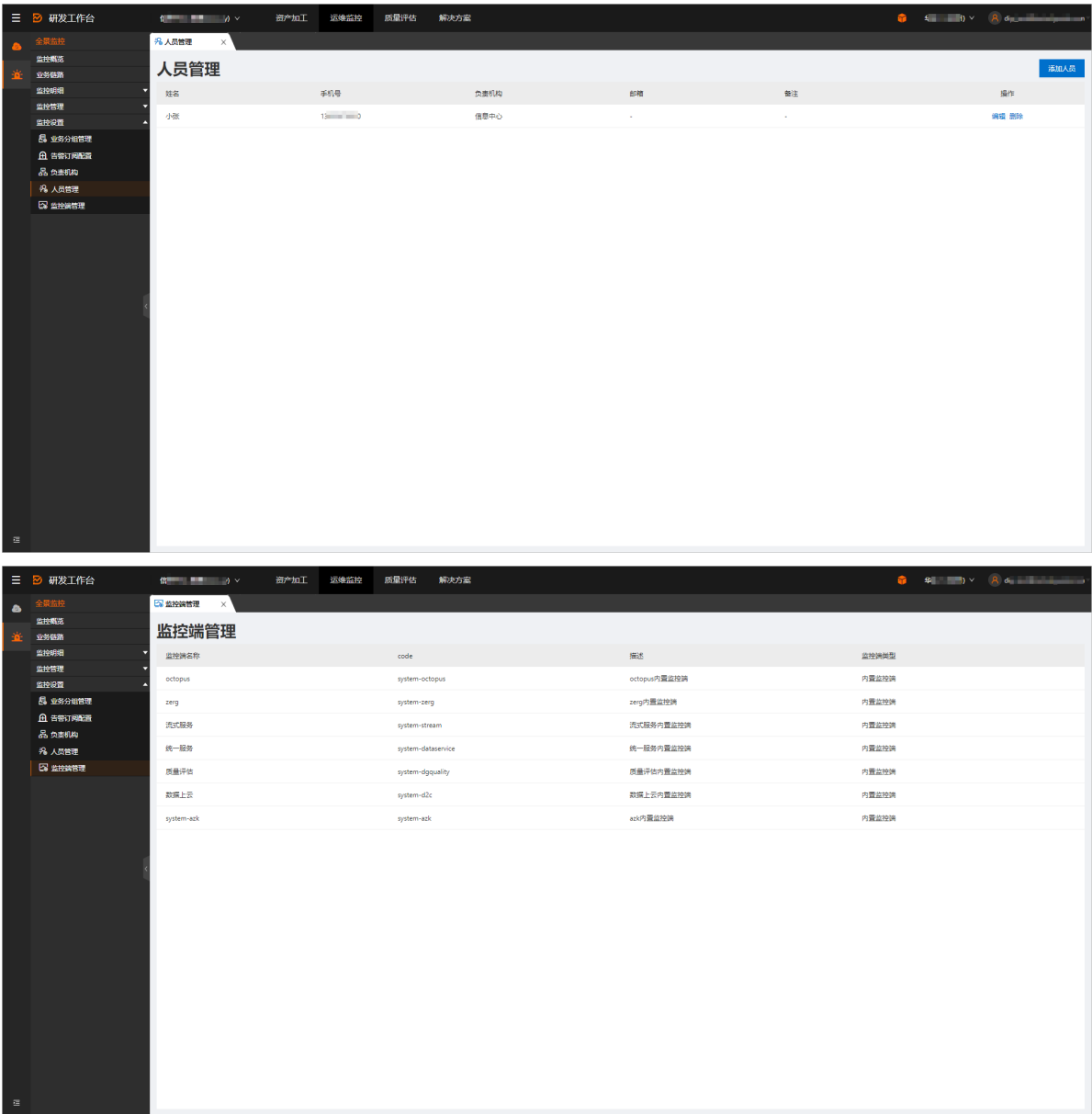
- 业务链路管理
 - 支持业务链路管理，按照业务视角自定义配置业务监控链路，可快速开启或关闭业务链路。
 - 支持业务链路的导入、导出。



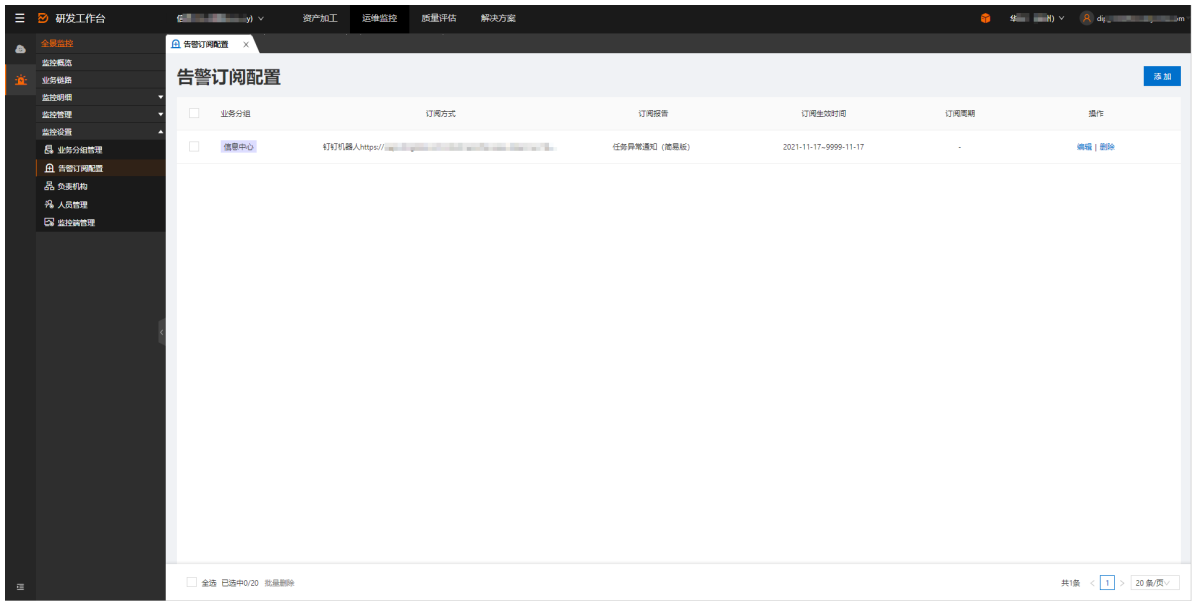
监控设置

- 支持业务分组管理、负责机构、人员管理、监控端管理。





- 支持通过钉钉机器人、系统内置标准协议等方式订阅告警消息。



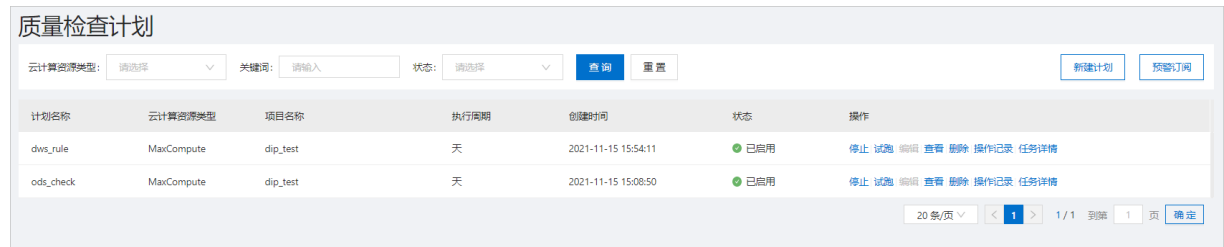
6.2.1.9. 质量评估

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

质量检查任务

计划名称/code

执行类型

运行状态

调度时间

→

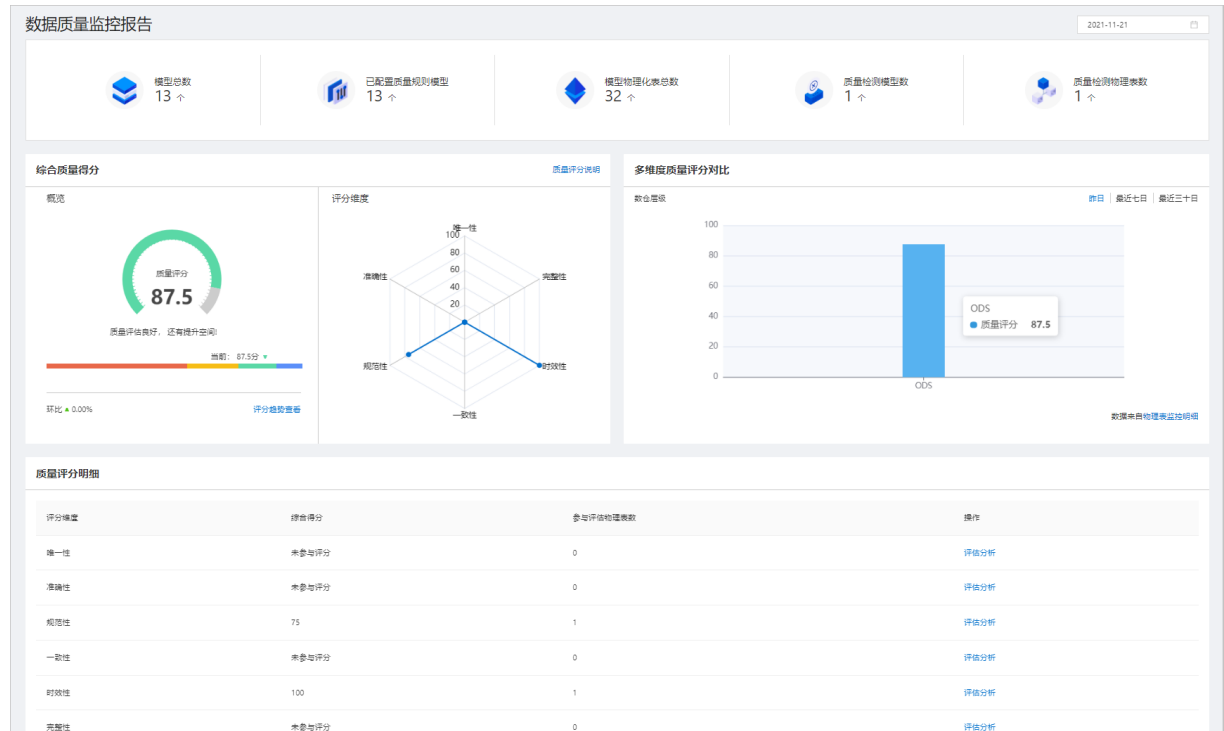
计划名称	计划code	物理表名称	分区表达式	质量	任务运行状态	调度时间	操作
ods_check_dataq_ru...	ods_check_dataq_ru...	ods_base_yhkh_mzf_...	ALL_PARTITIONS	❗ 违反强规则	成功	2022-02-11 00:00:00	详情 日志
ods_check_dataq_ru...	ods_check_dataq_ru...	ods_base_yhkh_mzf_...	ALL_PARTITIONS	❗ 违反强规则	成功	2022-02-10 00:00:00	详情 日志
ods_check_dataq_ru...	ods_check_dataq_ru...	ods_base_yhkh_mzf_...	ALL_PARTITIONS	❗ 违反强规则	成功	2022-02-09 00:00:00	详情 日志
ods_check_dataq_ru...	ods_check_dataq_ru...	ods_base_yhkh_mzf_...	ALL_PARTITIONS	❗ 违反强规则	成功	2022-02-08 00:00:00	详情 日志
ods_check_dataq_ru...	ods_check_dataq_ru...	ods_base_yhkh_mzf_...	ALL_PARTITIONS	❗ 违反强规则	成功	2022-02-07 00:00:00	详情 日志

20条/页

1 / 1 到第 1 页

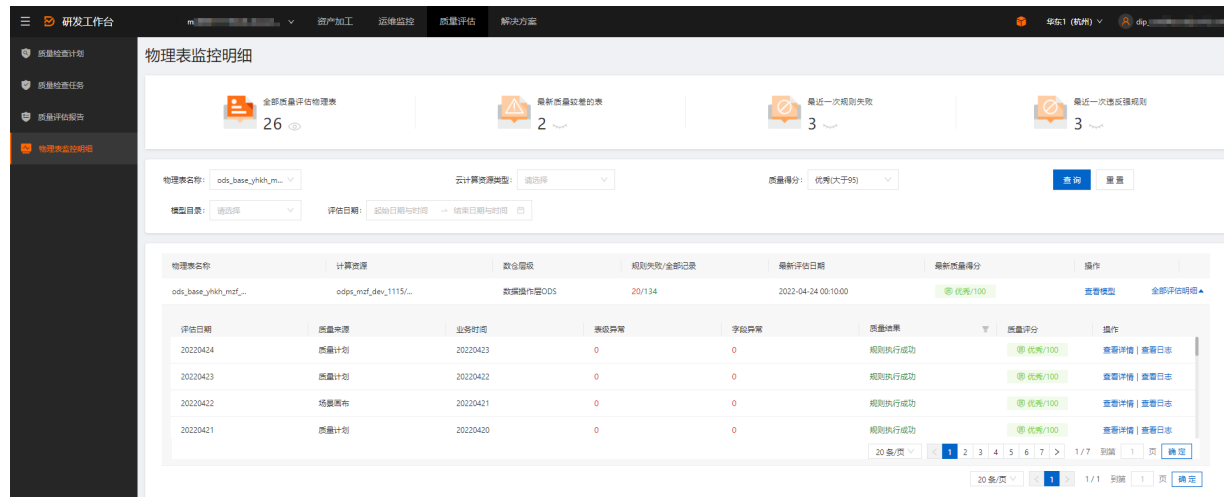
- 支持对质量检测任务运行实例进行查询和管理。
- 支持查看任务实例运行结果详情、运行日志进行查看。
- 支持对质量结果任务运行实例明细进行导出。

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。

6.2.1.10. 解决方案

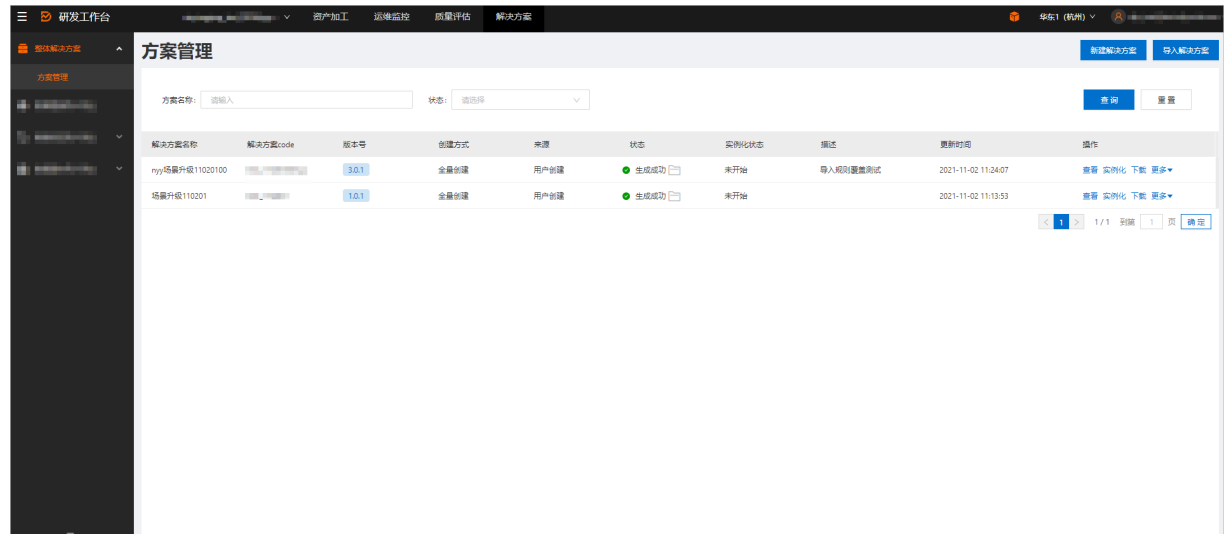
数据资源平台中管理的数据标准、数据模型、数据加工场景、数据服务API、云计算资源配置等数据资产，经过验证最终可沉淀为一个特定问题的解决方案。通过不同的创建方式，灵活的选择解决方案中包含的内容。依托数据资源平台中管理的大量元数据，数据资源平台帮助用户保证解决方案的完整性、一致性和可用性。

根据数据血缘关系，数据资源平台会根据不同的创建方式，自动选择关联的数据资产，保证解决方案的完整性。同时，数据资源平台会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，数据资源平台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数据资源平台的方式，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。

6.2.2. 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。运营工作台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。

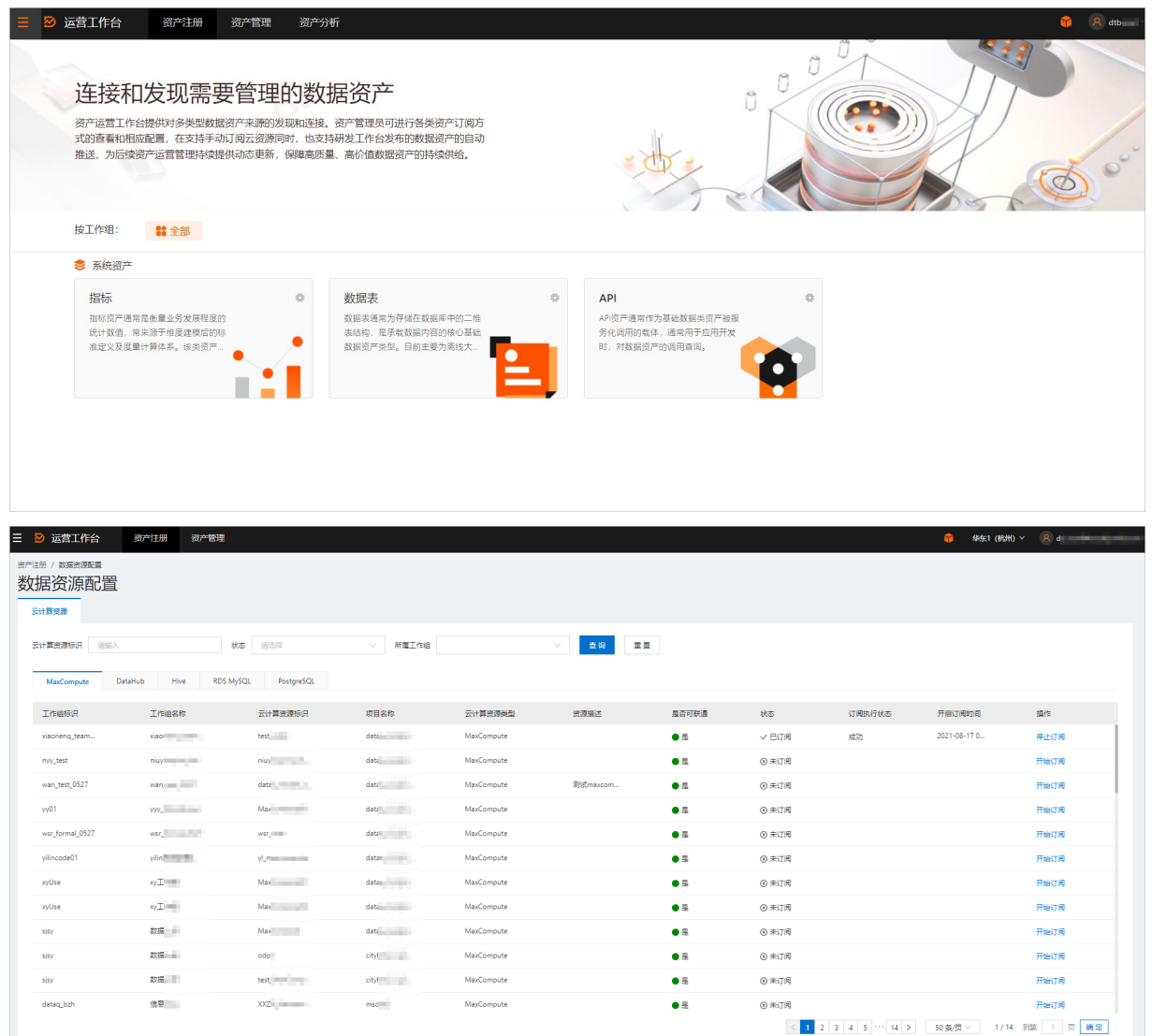
- 资产使用方：主要进行数据查找和使用申请，目标是找到需要使用的数据资产，进行权限获取以完成后续业务使用。
- 资产管理员：对所有可对外服务的高价值数据资产进行管理维护与编目上线，进行资产对外服务的生命周期维护，同时针对申请使用进行过程审核。
- 工作组管理员：对该角色负责管理的工作组下数据资产进行使用侧的管理审核。

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、标签或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

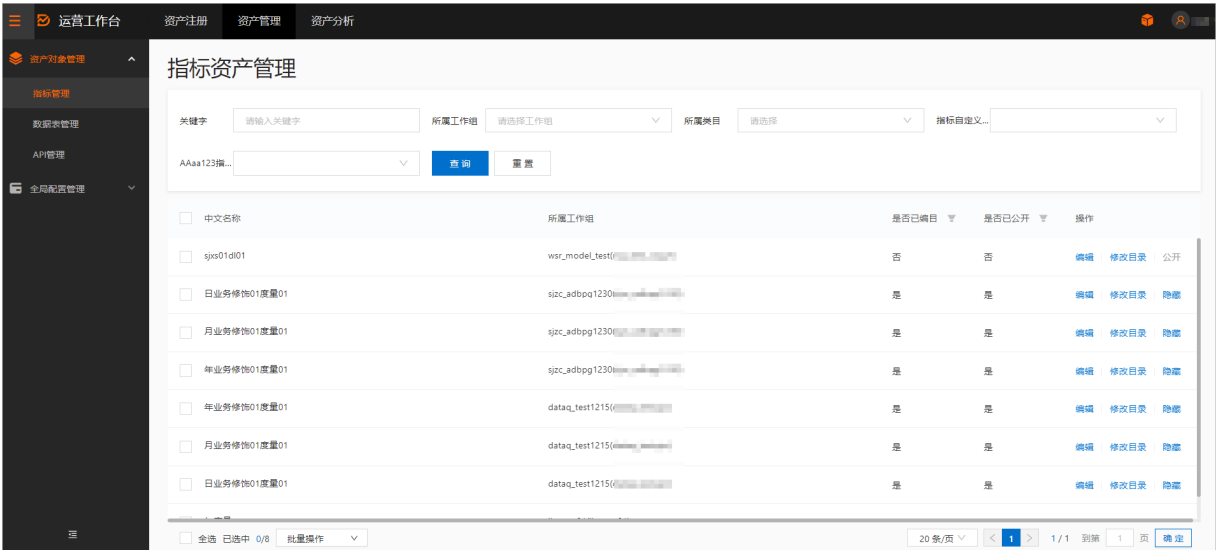
资产注册功能是数据资源平台资产管理的基础，可快速锁定待管理数据资产的范围。通过资产注册实现多类型数据表元数据的订阅，同时自动完成已上架API资产订阅，作为资产管理中待编目维护的数据内容。



- 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和Hive数据表资源进行展现订阅；可自动订阅研发工作台的数据表、API等资产信息。

资产管理

资产对象管理：提供对表资产、API资产和指标资产可视化展示，通过编目、公开为实现资产共享做准备。



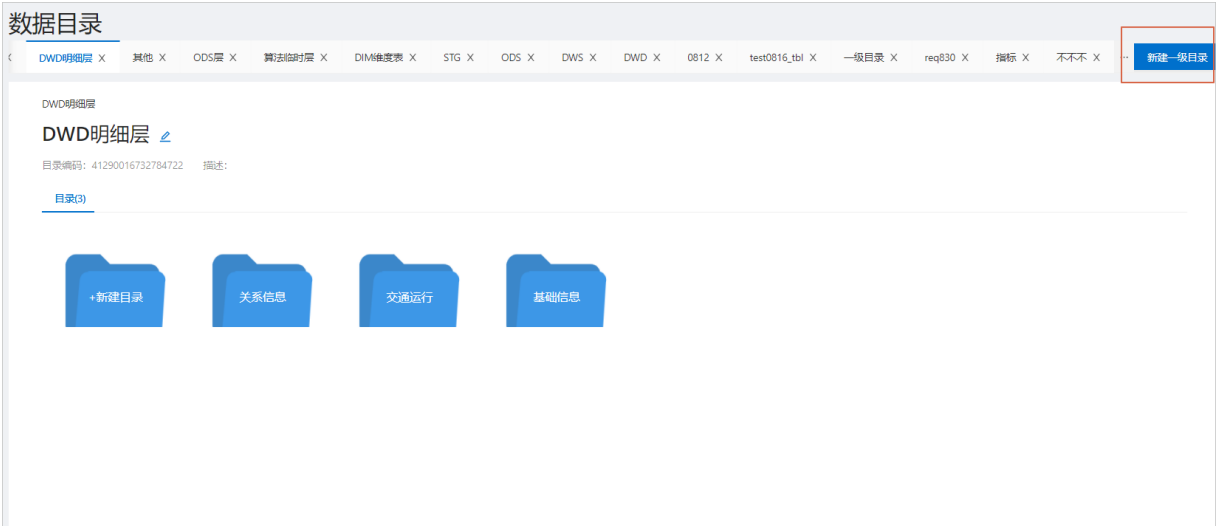
- 表资产管理：支持对数据表资产进行编目，并提供查询、公开、隐藏等功能。
- API资产管理：支持对数据API、三方API和上报API资产进行编目，并提供查询、公开、隐藏等功能。
- 指标资产管理：支持对指标资产进行编目，并提供查询、公开、隐藏等功能。

资产目录管理：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。

全局配置管理

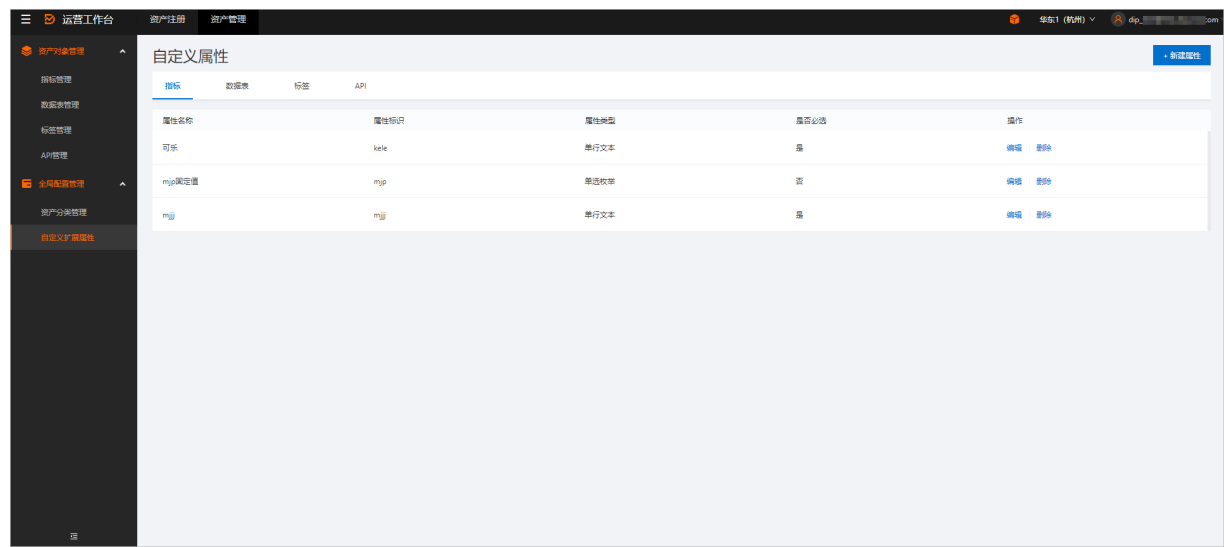
资产分类管理

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。



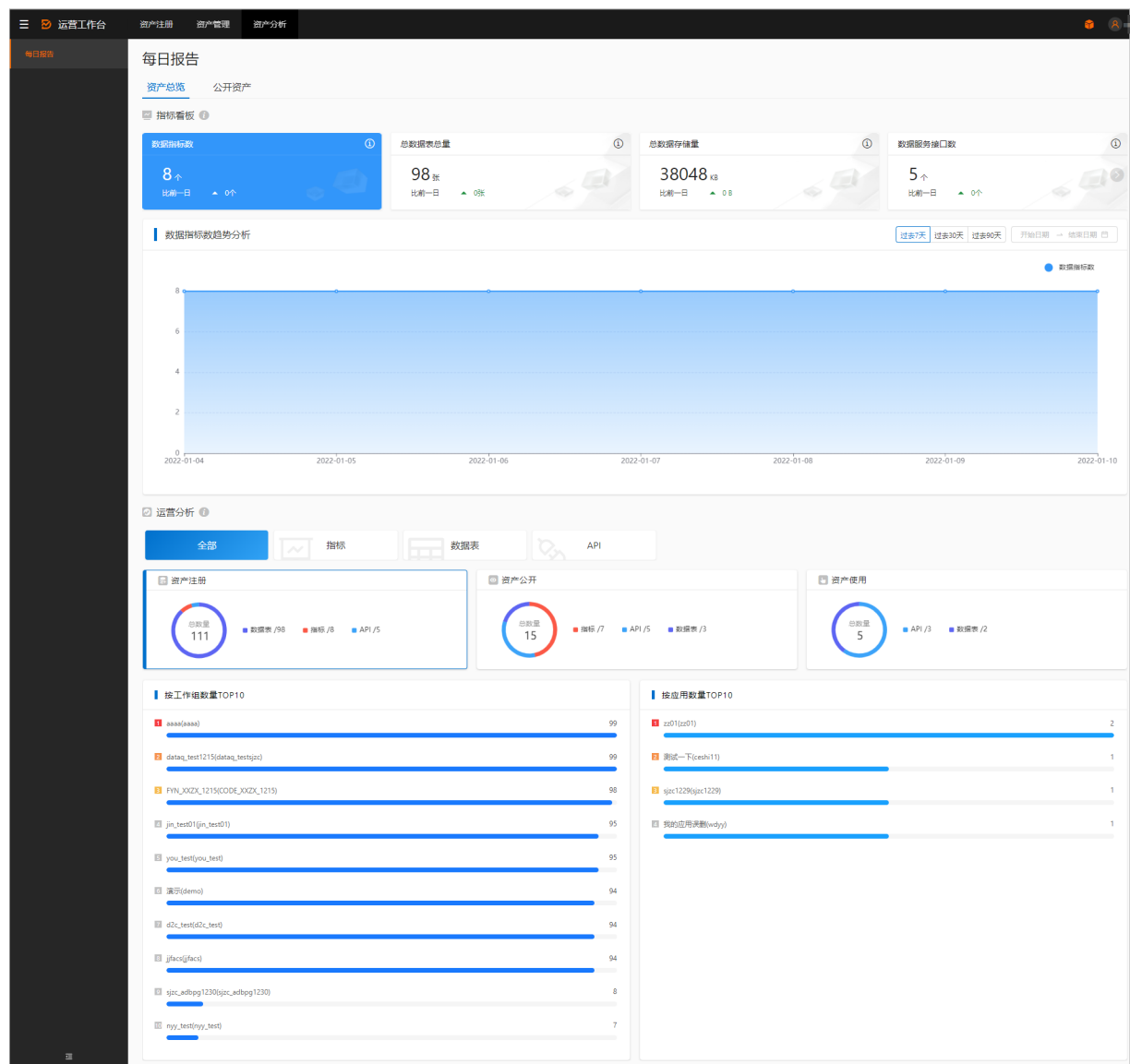
自定义扩展属性

提供对数据资产创建自定义扩展属性，以备在资产对象管理中为数据资产添加易于识别的特色属性值。



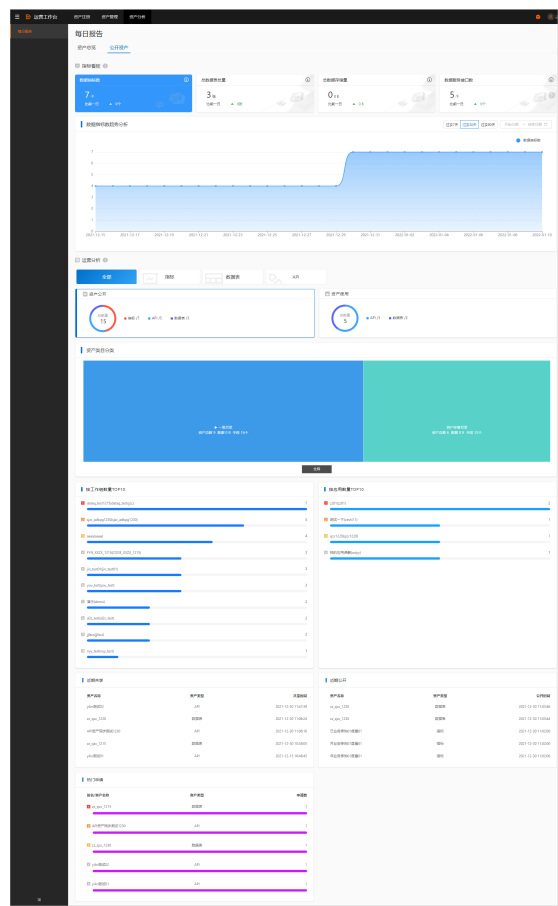
资产分析

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。



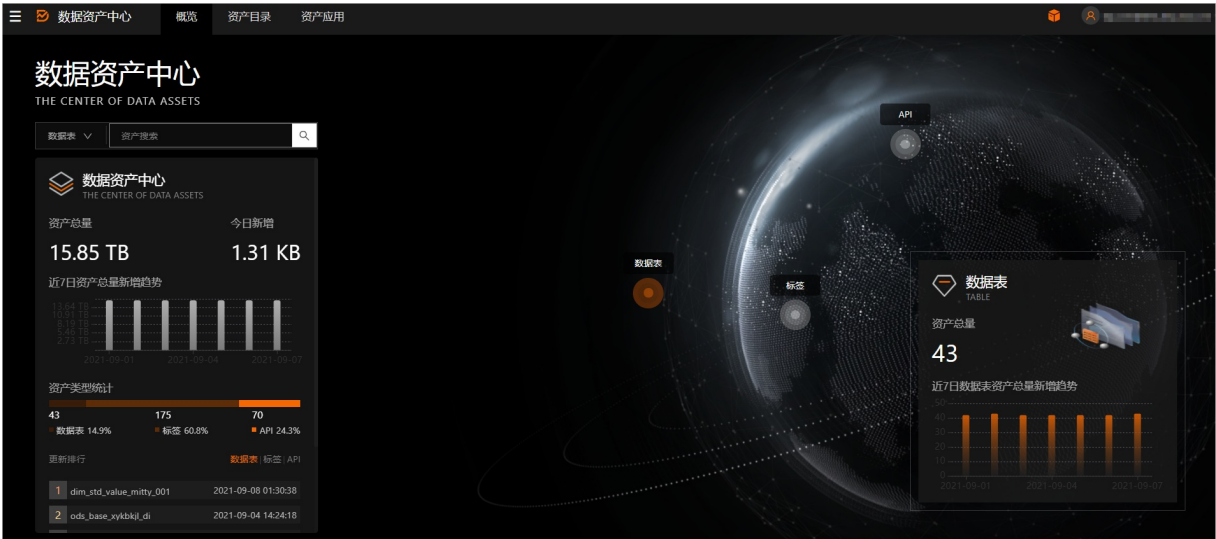
- 指标看板：展示已公开的资产，对全租户下已公开到资产中心的资产进行指标统计，了解各类资产的量级统计和时间趋势情况。
- 数据指标数趋势分析：展示已公开的资产过去7天、过去30天、过去90天及自定义时间范围内的指标趋势图。
- 运营分析：展示已公开的资产，对全租户下已公开到资产中心的各类资产，围绕公开资产的运营阶段（公开-使用），了解各阶段下的多维度统计分析结果。

6.2.3. 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

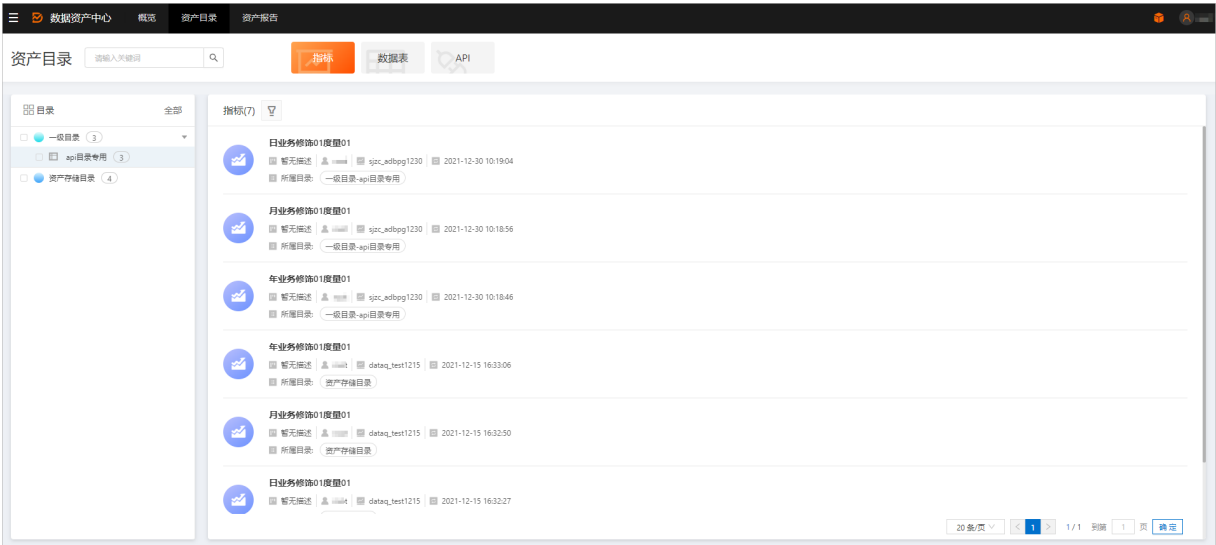
对数据资产的全局及增量指标进行统计、展示，帮助用户快速了解平台管理数据资产的总体情况。



- 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等资产分类型展示指标与增量信息。

资产目录

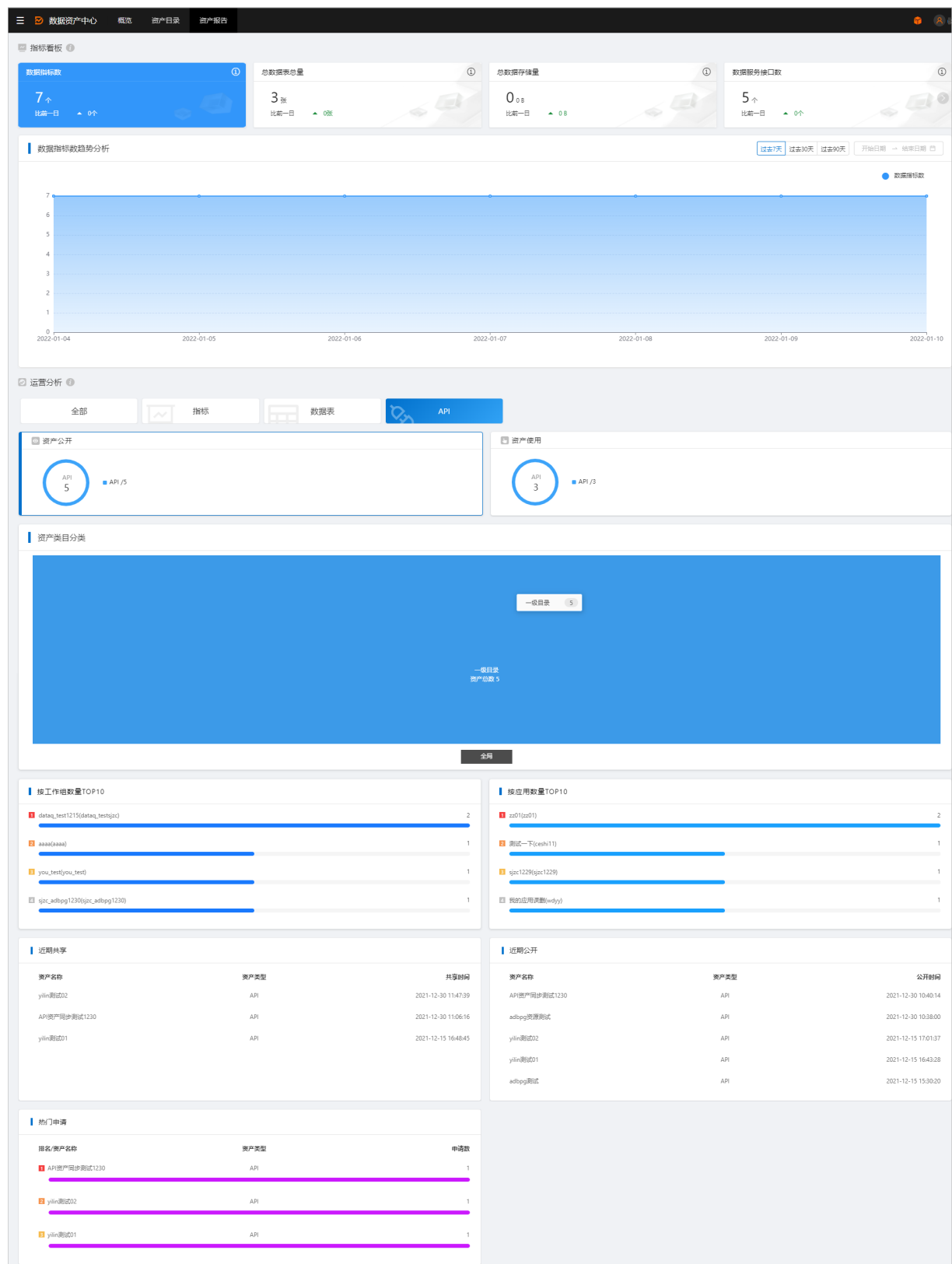
对公开共享的数据资产，可在资源目录进行搜索、目录列表查看，并通过资产详情，了解数据资产的业务价值，促进数据资产的共享和申请使用。



- 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。

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、标签、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。

6.2.4. 我的资产

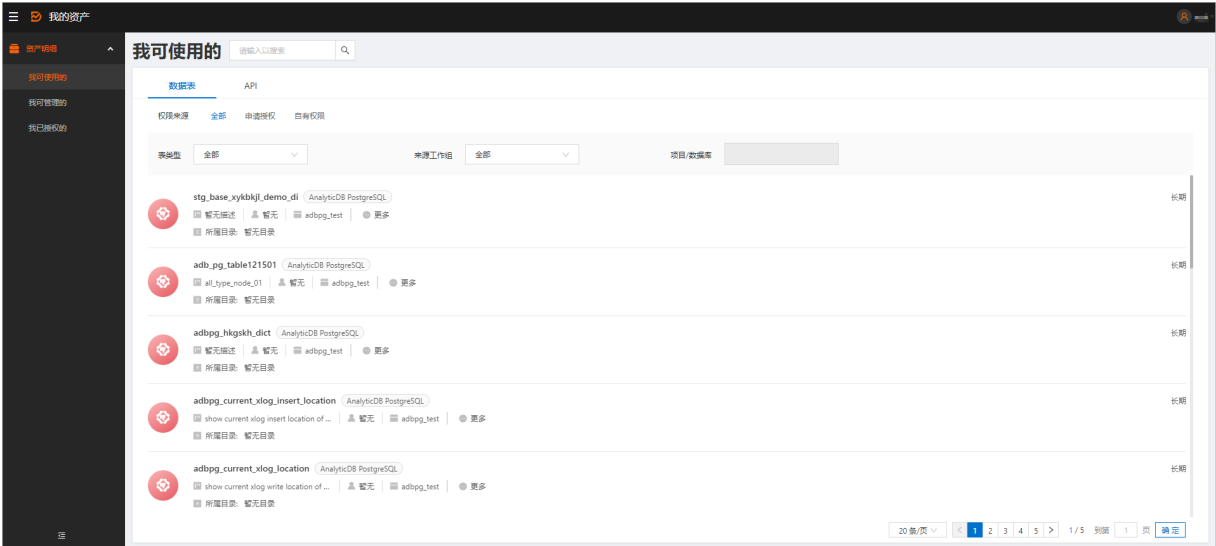
通过我的资产清晰掌握可使用、可管理、已授权的数据表、数据服务API数据资产相关信息。

通过不同的权限来源，用户可获取数据表、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、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、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。

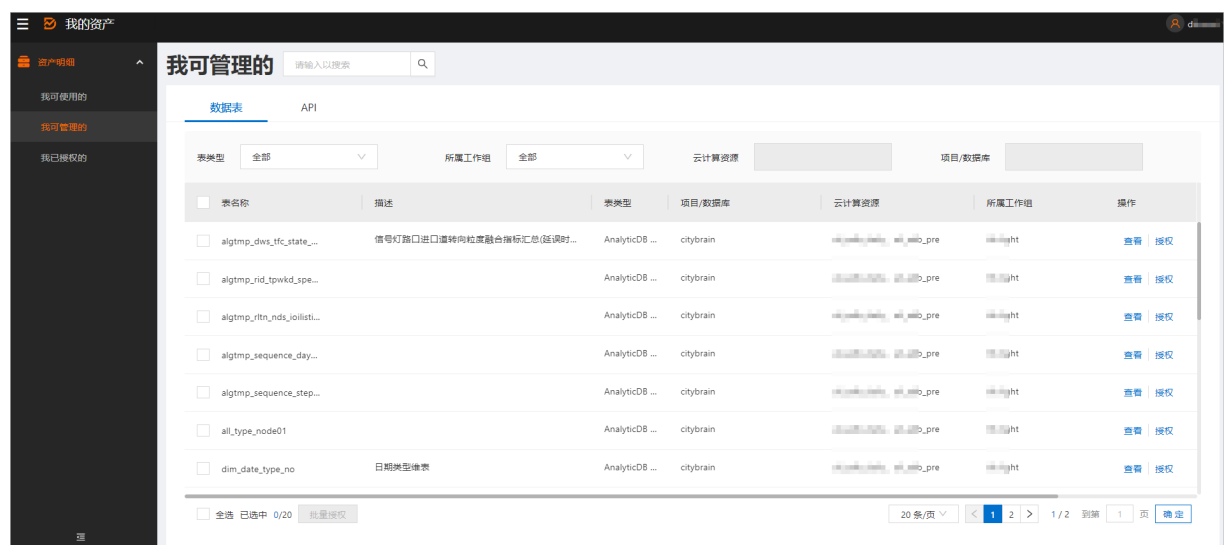
我可使用的

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



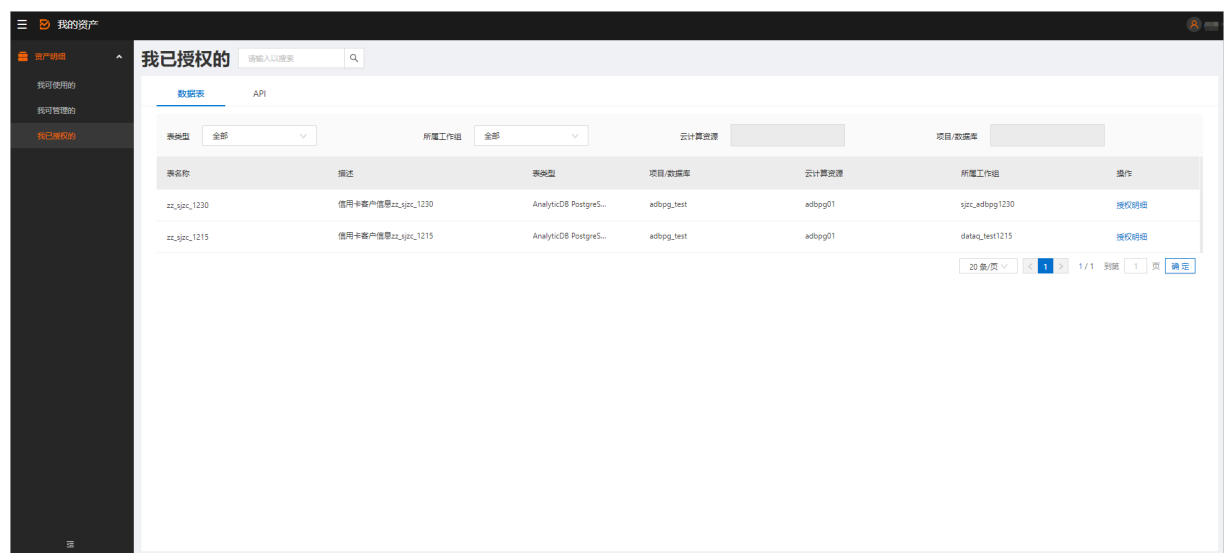
我可管理的

支持展示有管理权限的数据表、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、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。



6.2.5. 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功能，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

[illegible]

- **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：**支持多个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

工作台管理

用户管理

数据列表

我的申请

我的审批

应用管理

组织管理

资源站管理

配置管理

请输入以搜索

Q

添加成员

序号	用户名	角色	操作
01	admin	租户管理员 X	详情
02	user	租户一般用户 X 租户管理员 X	详情
03	user	租户一般用户 X 租户管理员 X	详情
04	user	租户一般用户 X 租户管理员 X	详情
05	user	租户一般用户 X 租户管理员 X	详情
06	user	租户一般用户 X 租户管理员 X	详情
07	user	租户一般用户 X	详情
08	user	租户一般用户 X 租户管理员 X	详情
09	user	租户一般用户 X 租户管理员 X	详情
10	user	租户一般用户 X	详情
11	user	租户一般用户 X	详情
12	user	租户一般用户 X	详情

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、租户运维、租户一般用户角色进行添加、管理。

权限列表

工作流管理	权限列表											
	系统设置	研发工作台	运营工作台	画像分析	数据探索							
权限列表	权限点	租户管理员	租户访客	租户数据资产管理	租户运维人员	租户一般用户	工作流管理员	工作流开发	工作流分析	工作流访客	组织管理员	组织一般用户
我的申请	读取应用信息	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
我的审批	修改应用信息	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
应用管理	读取组织信息	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓
组织管理	修改组织信息	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗
资源管理	授权角色	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗
配置管理	读取角色列表	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✗
	收回角色	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗
	创建云计算资源	✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗
	删除云计算资源	✓	✗	✓	✗	✗	✓	✓	✗	✗	✗	✗
	读取云计算资源	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✗
	更新云计算资源	✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗
	创建工作组	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗
	删除工作组	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗
	读取工作组	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗
	更新工作组	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗

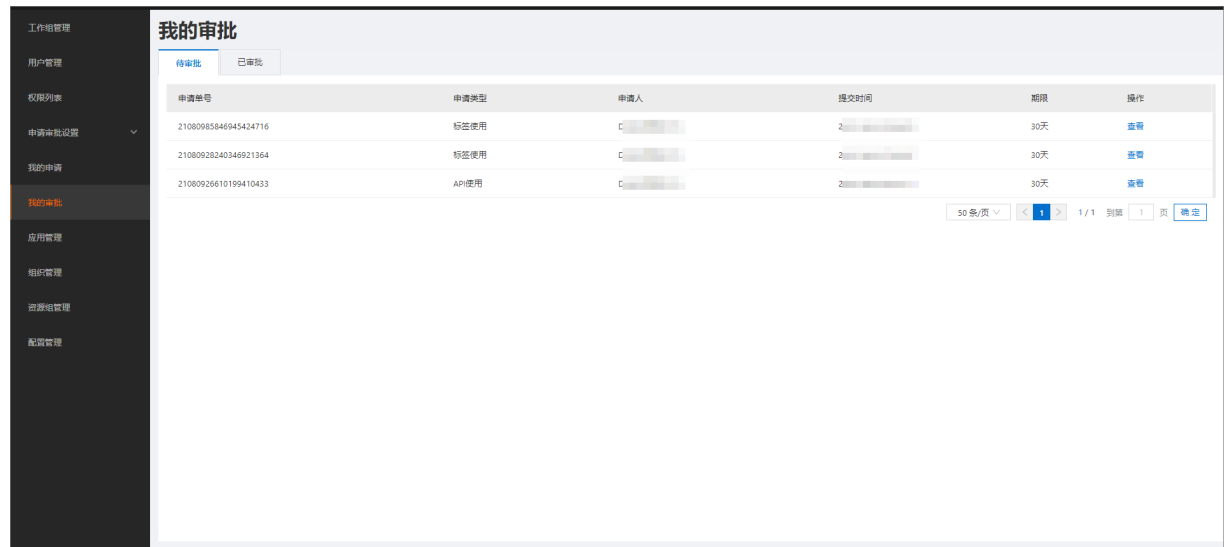
支持对系统管理、数据同步、数据标准、数据建模、智能标签、统一服务、通用资源、质量评估、任务运维、画像分析、运营工作台、数据资产中心的角色及权限点内容的展示。

我的申请

工作流管理	我的申请											
	申请单号	申请类型	状态	提交时间	期限	操作						
权限列表	210142089	API上架	待审批	2021-10-14 16:30:10	长期	查看 撤回						
我的审批	211014843	API上架	待审批	2021-10-14 16:23:29	长期	查看 撤回						
应用管理	210830194	API上架	已通过	2021-08-30 10:26:14	长期	查看						
组织管理	210830172	API使用	已通过	2021-08-30 09:56:48	30天	查看						
资源管理	210827206	API上架	待审批	2021-08-27 10:33:32	长期	查看 撤回						
配置管理	210823340	API上架	待审批	2021-08-25 09:50:55	长期	查看 撤回						
	210823856	API上架	待审批	2021-08-25 09:48:12	长期	查看 撤回						
	210825141	API上架	待审批	2021-08-25 09:44:21	长期	查看 撤回						
	210824624	API上架	待审批	2021-08-24 18:04:42	长期	查看 撤回						
	210811849	API上架	待审批	2021-08-11 17:50:53	长期	查看 撤回						
	210810163	标签使用	已通过	2021-08-10 15:50:19	30天	查看						
	210728342	数据表使用	已通过	2021-07-22 13:17:37	30天	查看						
	210721104	标签使用	已通过	2021-07-21 13:12:37	30天	查看						
	210721222	标签使用	已通过	2021-07-21 13:10:28	30天	查看						
	210721170	API上架	待审批	2021-07-21 11:04:23	长期	查看 撤回						
	210721921	API使用	已通过	2021-07-21 11:02:15	30天	查看						
	210721197	API上架	待审批	2021-07-21 10:44:50	长期	查看 撤回						

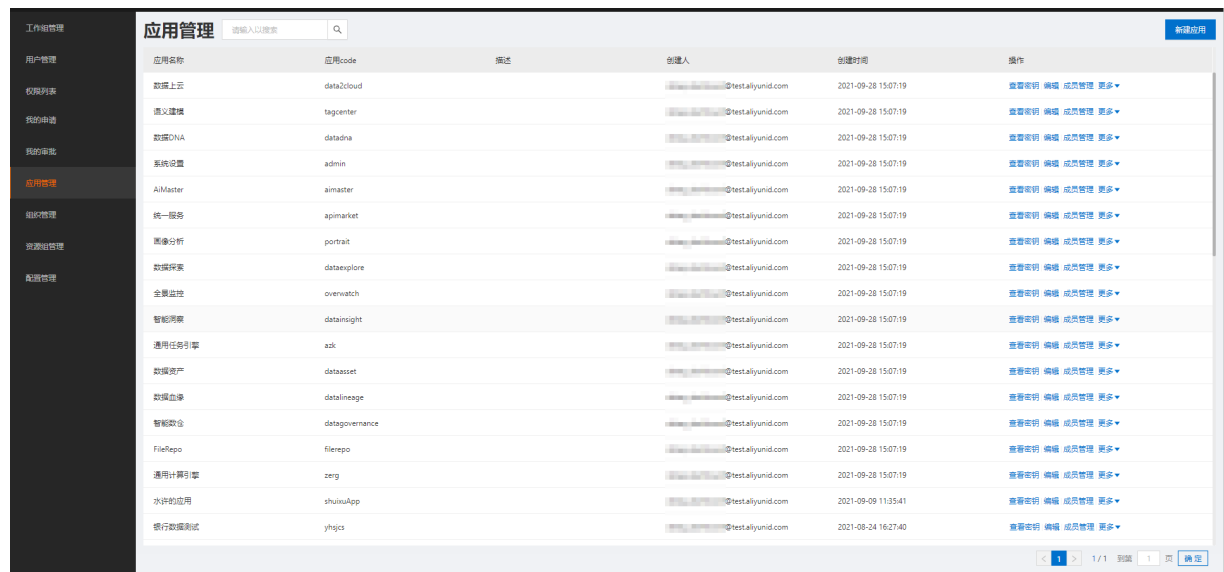
支持查看已经提交的申请单，包括在途和审批完成的。用户可以查看自己申请单的详细信息和审批结果，用户也可以按审批状态筛选申请单和撤回申请。

我的审批



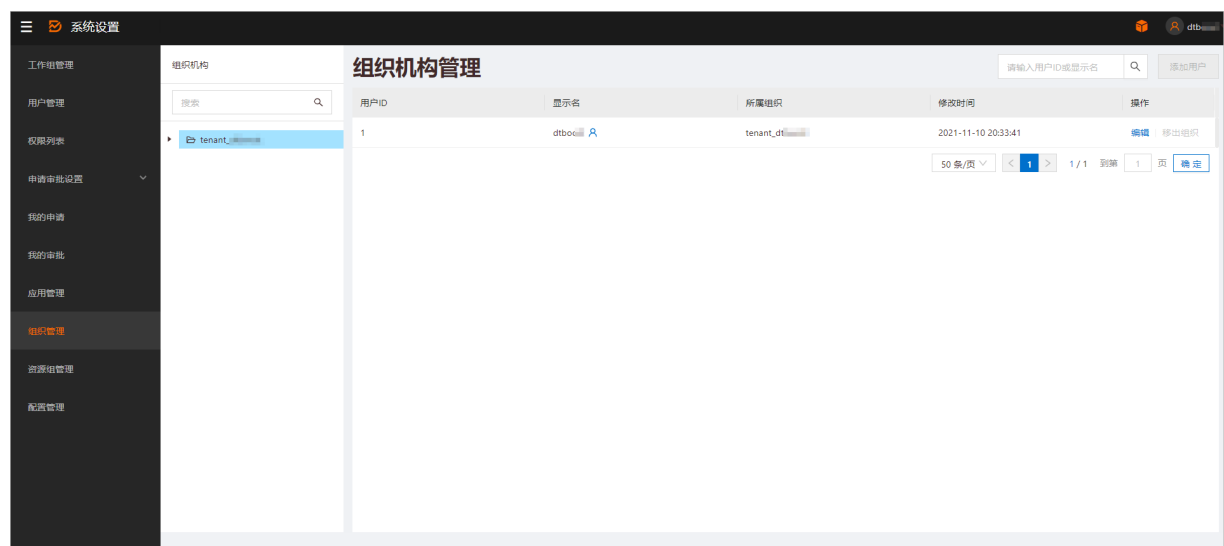
展示需要当前用户待操作的申请单。可以查看申请单的详情，包括申请的数据资源信息、申请理由、申请使用时长、审批结果等。支持操作员对申请单的审批动作，支持通过、拒绝等功能。支持查看已审批和未审批的申请单。

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、Open API 授权功能。

组织管理

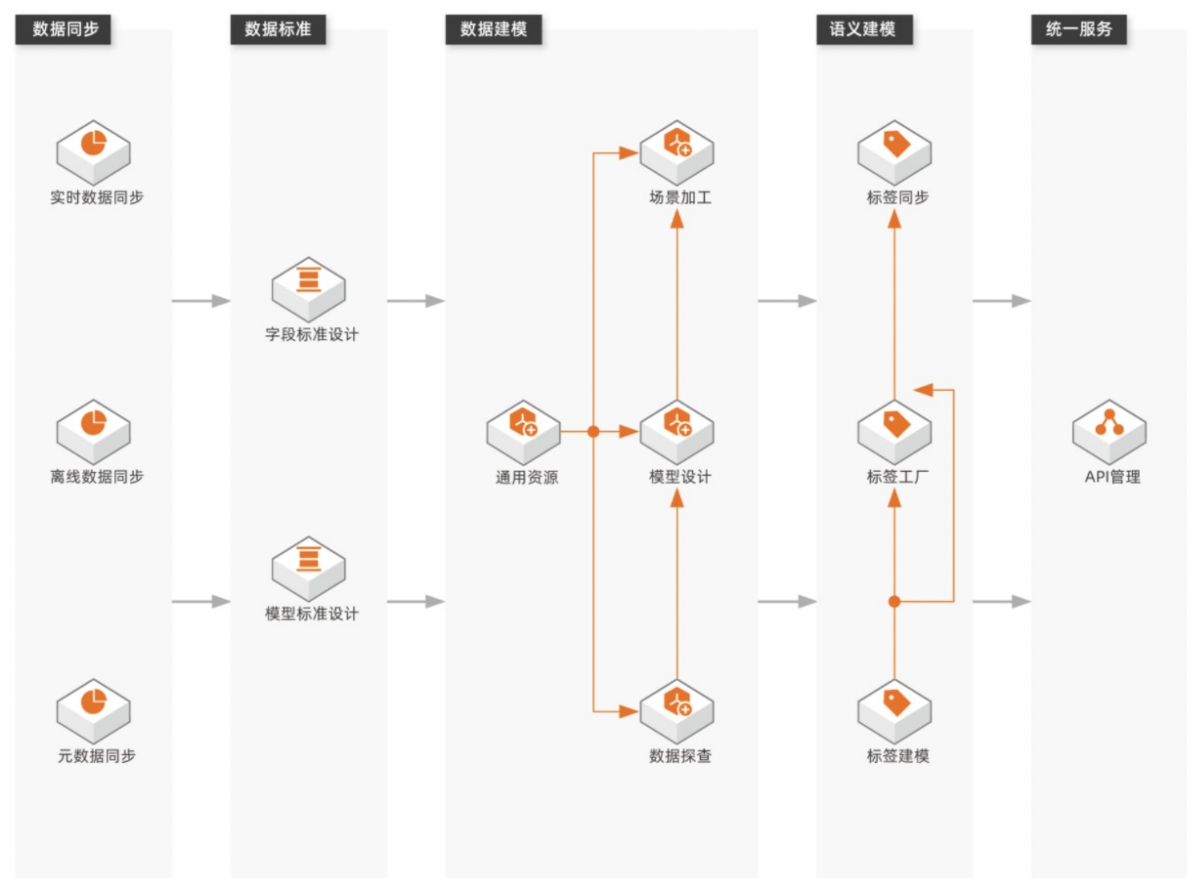


支持树形组织的创建、管理，支持组织内成员的添加、管理功能。

6.3. 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。

6.4. 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同的角色，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。

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 workflow。

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。步任务、离线计离线同步任务、离线计算任务的类型，主要分为手动、周期、补数据等。离线同步任务、离线计算任务的类型，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节点、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。

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、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。

6.5. 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、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。
- **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	数据源类型与版本	目标端												
		Oracle	SQLServer	Postgres	MySQL	ADB PG	MaxCompute	DataHub	Hive	OSS	HBase(Phoenix)	TableStore	Hologres (不支持关键字)	ADB MySQL
离线链路	Oracle(10g, 11g)	✓	✓	✓	✓		✓	✓	✓					
离线链路	SQLServer(2008, 2014, 2016, 2017)	✓	✓	✓	✓		✓	✓	✓					
离线链路	Postgres(9.4)	✓	✓	✓	✓		✓	✓	✓					
离线链路	MySQL(5.6, 5.7)	✓	✓	✓	✓	✓	✓	✓	✓					✓
离线链路	HDFS(Hadoop 2.7.1, 2.5.0, 2.6.0)							✓						
离线链路	Hive(1.1.1, 1.2.2)													
	Hadoop 2.7.1, 2.5.0, 2.6.0				✓(API)		✓				✓(API)			
离线链路	FTP									✓				
离线链路	MaxCompute			✓	✓	✓	✓					✓	✓	✓
离线链路	ADB PG				✓	✓	✓						✓	

② 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	数据源类型与版本	目标端
		DataHub
实时链路	SQLSever(2008, 2017)	√
实时链路	MySQL(5.6, 5.7)	√
实时链路	PUSH API	√
实时链路	Kafka (2.3, 2.4, 2.5)	√
实时链路	RabbitMQ	√

② 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。

数据探查

系统目前支持：MaxCompute、RDS MySQL、AnalyticDB PostgreSQL、Hologres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。

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. 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。

交通云控平台与生态应用的关系



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

7.1. 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。

7.2. 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%+。