

数据开发空间（Ottomi-Forge）社区版功能介绍

端到端闭环：数据开发→调度→BI 分析→大屏呈现→权限管理，全流程一站式贯通。

低门槛易用：拖拽式组件，无需代码基础，降低人才依赖。

高性能引擎：并行计算，轻松应对海量数据，释放数据库压力。

灵活编排：多任务自由组合为作业流，支持条件触发、优先级配置，适应复杂业务场景。

BI 开箱即用：内置 DataEase 可视化引擎，20+图表模板、实时大屏设计器，告别二次开发。

智能运维：任务状态实时跟踪、组件使用统计、失败自动预警，风险拦截。

主要功能

数据源连接

系统支持多种数据源及数据格式：

国外数据库：MySQL、PostgreSQL、SQLserver、Oracle、DB2

国产数据库：DM、GaussDB、KingBase、GBase、StarRocks

NoSQL 数据库：Hive

流式数据：Kafka

服务器终端：SSH

数据库类型

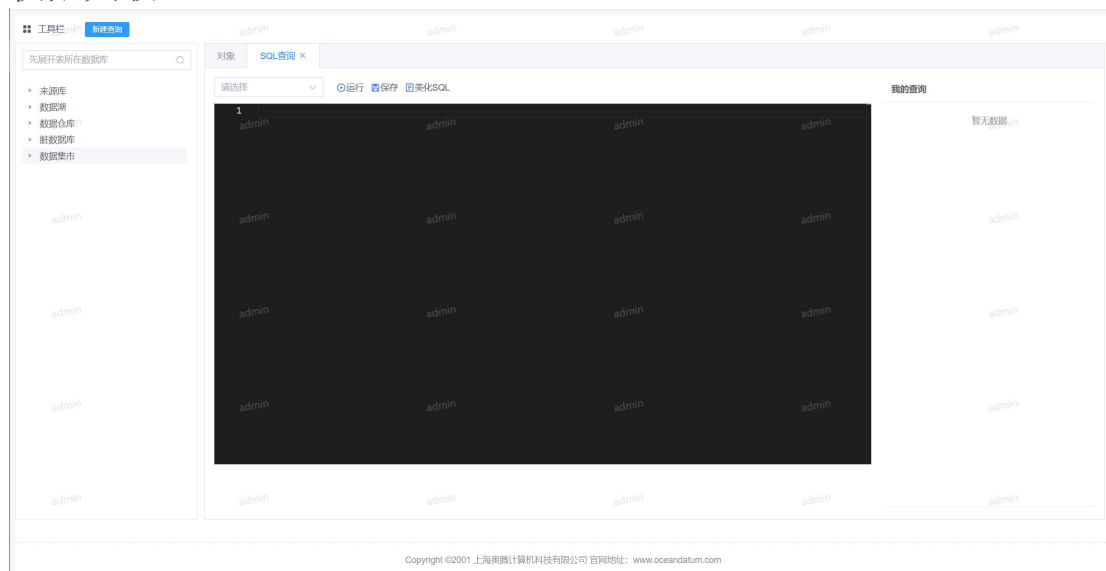
×



取消

资源透视

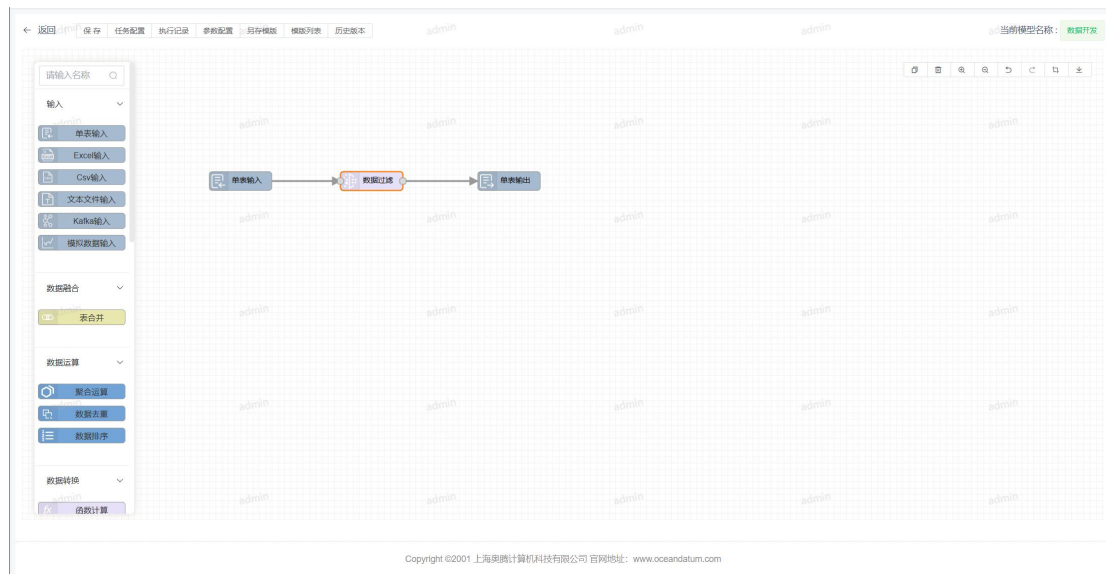
数据查询工具，直连数据库数据实时刷新。可以查询接入的源端目标端的数据。提供 SQL 引擎查询数据。支持一键美化 sql。页面数据查询排序查看。本菜单开放给管理员查看数据，因而不与分类分级权限关联。



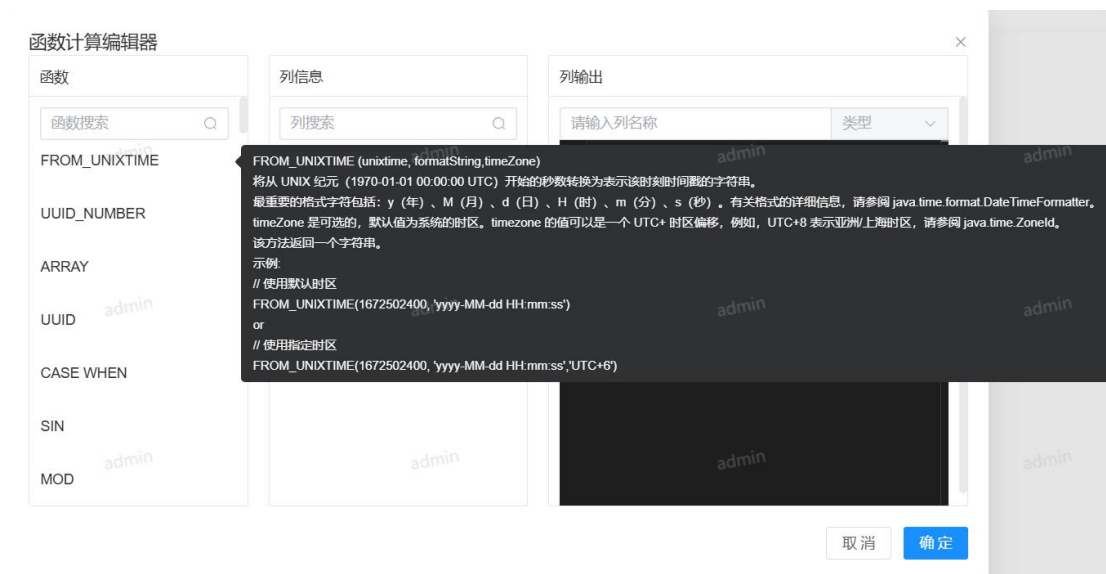
数据开发

提供可视化开发环境，拖拽式组件，无需代码基础，便于理解数

据流和处理过程，加强数据治理和监控能力，满足多样化需求。



此外，系统具有一站式建模、低门槛建模、高性能计算等优点，内置多种组件和数十种计算函数



单表输入组件

单表输入主要用于对单一表进行数据处理，包括数据聚合、过滤、替换、添加字段等操作。通过可视化的界面简化操作步骤，用户可以根据实际业务需求灵活筛选和编辑数据。

单表数据聚合处理：支持按照指定字段或多个字段进行聚合操作。

单表数据过滤处理：支持条件表达式两种方式对数据进行过滤筛选。

单表数据替换处理：支持字符串替换，并允许编写 SQL 语句进行数据替换。

单表添加新字段处理：支持生成唯一 ID、当前时间等字段，使用函数公式生成新的字段列。

文件输入组件

文件输入应对大文件、大数据量且需要灵活清洗的场景，支持多种文件类型（如 Excel、文本文件、CSV 等），并提供自定义字段分隔符、文件编码转换等功能，确保不同类型文件能够正确解析和输出。

模拟数据组件

在缺乏真实数据的情况下，系统提供模拟数据的能力，支持生成不同类型的模拟数据，满足性能测试需求。模拟的数据量可增加至 100 万，数据值域和范围可自定义。

Kafka 输入组件

Kafka 输入支持读取消息队列中各类非结构化数据，支持消费者组自定义这使得可横向扩展消费者数量应对高并发的数据场景，此外提供多种读取点位进行消费队列中的数据，如消费者组偏移量、历史数据、最新数据和指定时间，默认为当前组的消费偏移量。同时提供多种数据格式用于解构消息队列中非结构化数据支持文本格式、json、debezium_json、canal_json。

多表同步组件

数据分析过程中经常需要将多张表的维度进行相互补充，对数据开展横向扩维，纵向汇聚的需求。多表数据处理算法集主要用于两张输入表之间的碰撞运算，包括交集处理、左连接处理、右连接处理、全连接处理等。

数据运算组件

通过聚合字段（相当于 sql 的 group by），可对字段进行求最大值、最小值、平均值、字段计数或拼接，还可以对字段数据去重、排序处理。

数据转换组件

数据转换是数据处理过程中的重要环节，系统内置了丰富的数据转换算法，包括规则过滤、函数计算、字段过滤、数据过滤、值替换、NULL 值替换、数据脱敏、加密解密等多个功能模块，支持快速可视化数据处理。

➤ 函数计算

系统提供丰富的函数包，分为字符串函数、时间函数、系统函数、数值函数等多个函数包，满足不同业务场景下各类需求，函数计算组件支持将当前函数计算后的数据生成新的字段或只替换当前字段。

➤ 字段过滤

系统提供选择需要的字段输出至下一步，抛弃无用字段。

➤ 数据过滤

系统提供对数据进行筛选过滤，保留符合要求的数据，可以选择对比固定值，也可以选择对比固定列。

➤ 值替换

系统提供将指定字段值替换为设定的值，通常将编码转换为字典值。

➤ NULL 值替换

系统提供仅将字段值为 null 的替换为指定值。

➤ 数据脱敏

系统提供将敏感数据根据脱敏规则静态脱敏，此版本只支持内置身份证号、手机号、中文名、座机号、电子邮件、地址、车牌、密码、银行卡等脱敏规则。

➤ 数据加密解密

系统提供静态的加密解密实现数据存储。数据预览、处理过程更加安全。加密解密的算法包括 ADS，DES，MD5，SHA-1 等。

➤ 字段名称映射

系统提供修改字段名称，如带有特殊字符的字段名称、中文字段名称导致建表失败或者需要统一规范字段名称。

➤ 唯一 ID 生成

系统提供生成唯一值列，支持算法有雪花算法、常用 UUID 算法、UUID 正整数算法、ObjectId 算法。

➤ 字符串拆分

系统提供自主配置拆分依据，自主生成字段列，实现指定信息的自定义拆分。

➤ 大小写转换

系统提供将字段值大写转小写，小写转大写功能。

➤ 字符串拼接

系统提供将同行中多个字段拼接为新的字段，可自定义分隔符。

➤ 字符串剪切

系统提供自主配置截取位置及长度，实现字段值的自定义截取。

➤ 字符串修剪

主要解决了数据不规范的问题，根据业务场景提供多种操作方式：包括特殊字符去除（回车，空格，换行符，制表符的去除），去除指定前缀、后缀符号，也可自定义去除指定符号。

➤ JAVA 高级转换

JAVA 高级转换是对常规算法无法满足场景的有效补充，系统提供了 JAVA 脚本的输入界面，并基于常用的场景封装了多种内置方法，便于用户开展数据处理工作。主要包括列返回，值修改，值获取，值丢弃等。

➤ JSONPath 提取

系统提供将字段值为 JSON 类型的数据提取，生成新的字段。

➤ 列转行

系统提供根据分隔符，将列数据转换为多行。如

id	name	subjects
1	张三	语文、数学

转换后



id	name	subjects
1	张三	语文
1	张三	数学

输出组件

支持将数据处理后的结果保存为实体数据，生成新表供其他运算或对外服务应用。输出类型包括库表、文件、非结构化输出端，支持自定义刷写批量条数和输出前后的 SQL 脚本，增强了输出灵活性。

脚本组件

➤ SQL 脚本

SQL 脚本是对常规算法无法满足场景的有效补充，系统提供了 SQL 脚本的输入界面，并基于常用的场景封装了多种自定义函数，便于用户开展数据处理工作。

➤ Python 脚本

系统支持 Python 代码编写管理，用户可在面板上编写 Python 代码并提交至所属的服务器环境上运行，支持定时调度 Python 脚本。

➤ Shell 脚本

系统支持 Shell 代码编写管理，用户可在面板上编写 Shell 代码并提交至所属的服务器环境上运行，支持定时调度 Shell 脚本。

任务管理

系统提供丰富的任务调度和管理，可以随时启停开发任务查看任务日志，设置调度周期、重试次数等参数。

编辑任务

×

* 任务名称 速度测试-单机单节点

* 调度周期 0 0 0 * * ?

每秒流量上限 - + 字节(Byte) v

每秒条数上限 - +

并行度 - 1 + 推荐值: 8

失败重试次数 - 0 + ?

重试间隔(秒) - 3 + ?

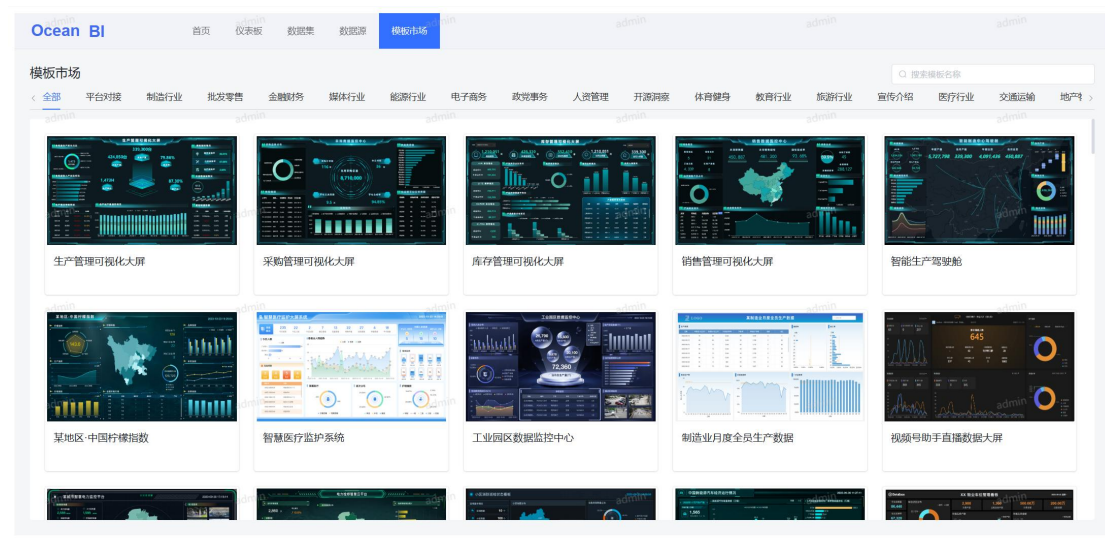
定时执行 ☐

取消

确定

数据分析

内置 DataEase 可视化引擎，图表组件丰富，支持自定义报表布局，包含多种模版、快速生成交互式报表/指挥大屏。



Copyright ©2001 上海奥腾计算机科技有限公司 官网地址: www.oceandatum.com

系统管理

预先设定机构层级关系，配置角色的菜单按钮权限、数据权限（包括：所有部门数据权限、本部门数据权限、本部门及下级部门数据权限、自定义部门数据权限），然后创建用户并关联角色和机构。即可实现用户权限的隔离，同部门用户数据共享、跨部门用户数据隔离、上级部门用户有权查看下级部门的数据。同级部门之间的数据不共享。

添加

×

角色名称

角色码

admin

角色描述

角色权限

请选择

菜单

全部

自定义

本级及下级

本级

本人

任务管理

任务列表

任务日志

系统配置

用户管理

部门管理

角色管理

菜单管理

日志管理

取消

确定

系统能够解决传统数据处理工具中的诸多难题，极大提升数据处理的效率和灵活性，