

# 虚谷实时数据服务平台V2.0-数据开发域用户手册

实时数据服务平台用户指南

数据开发域用户手册

数据集成、模型设计、流程开发与运行管理

|        |                                  |
|--------|----------------------------------|
| 适用产品版本 | 正式版                              |
| 手册版本   | V1.0                             |
| 编写日期   | 2026-06-12                       |
| 适用角色   | 数据集成工程师、数据开发工程师、数仓建模人员、审批人员、运维人员 |

## 修订记录

| 版本   | 日期         | 修订内容 |
|------|------------|------|
| V1.0 | 2026-06-12 | 初次编写 |

## 数据开发域功能概述

数据开发域提供文件汇聚、离线与实时同步、数仓规划、模型设计、工作流开发、发布审批和运行管理能力。用户可以先准备连接与目标数仓，再设计模型和数据处理任务，发布后配置调度或启动实时运行，最后通过工作流实例及告警规则跟踪运行情况。

## 1 数据开发首页

首页集中展示模型、任务定义、运行实例、函数与审批情况，帮助用户了解当前可访问范围内的数据开发工作。

# 1.1 首页总览

通过总览区了解对象数量与状态分布，再进入相应模块开展具体操作。

## 1.1.1 开发概况

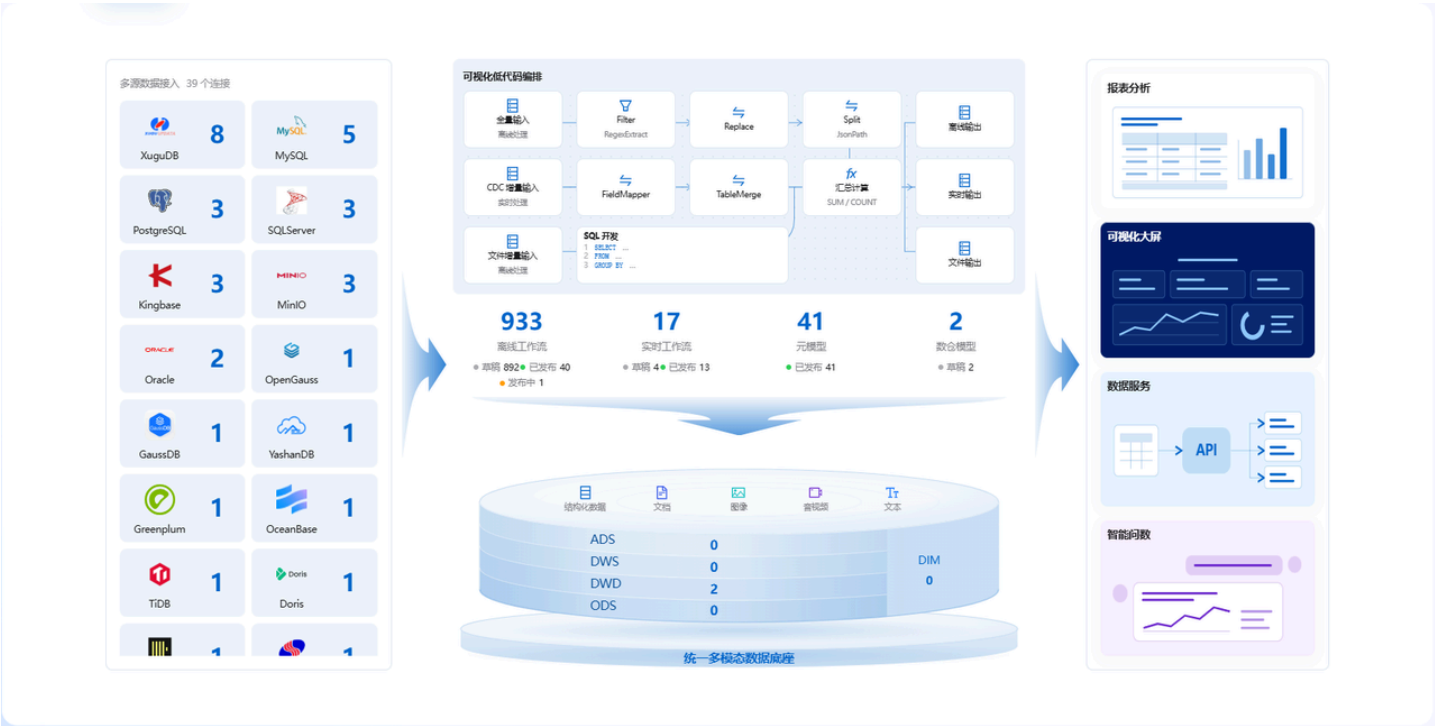


图 1开发概况

说明：

- 进入顶部“数据开发域”，查看首页统计。模型统计区展示元模型、数仓模型及其状态和分层分布。
- 离线、实时任务定义与运行实例分别统计；任务定义代表配置对象，实例代表某一次实际运行，不应将两者数量相加作为任务总数。
- 函数与审批统计用于了解内置能力及审批工作量。图表无数据时先确认访问范围；加载失败时使用页面重试入口重新加载。

# 2 数据集成

数据集成负责将外部数据带入后续处理流程。文件汇聚管理文件对象，离线同步批量读取数据库表，实时同步持续接收数据库变更。连接信息由平台数据源统一维护，各任务使用已授权的数据源，不在任务中重复保存数据库密码。

## 2.1 文件汇聚

文件汇聚以文件服务和路径组织文件，支持浏览、上传、预览及文件管理。文件服务需预先在平台数据源中配置；不同管理操作受用户权限控制。

## 2.1.1 文件列表

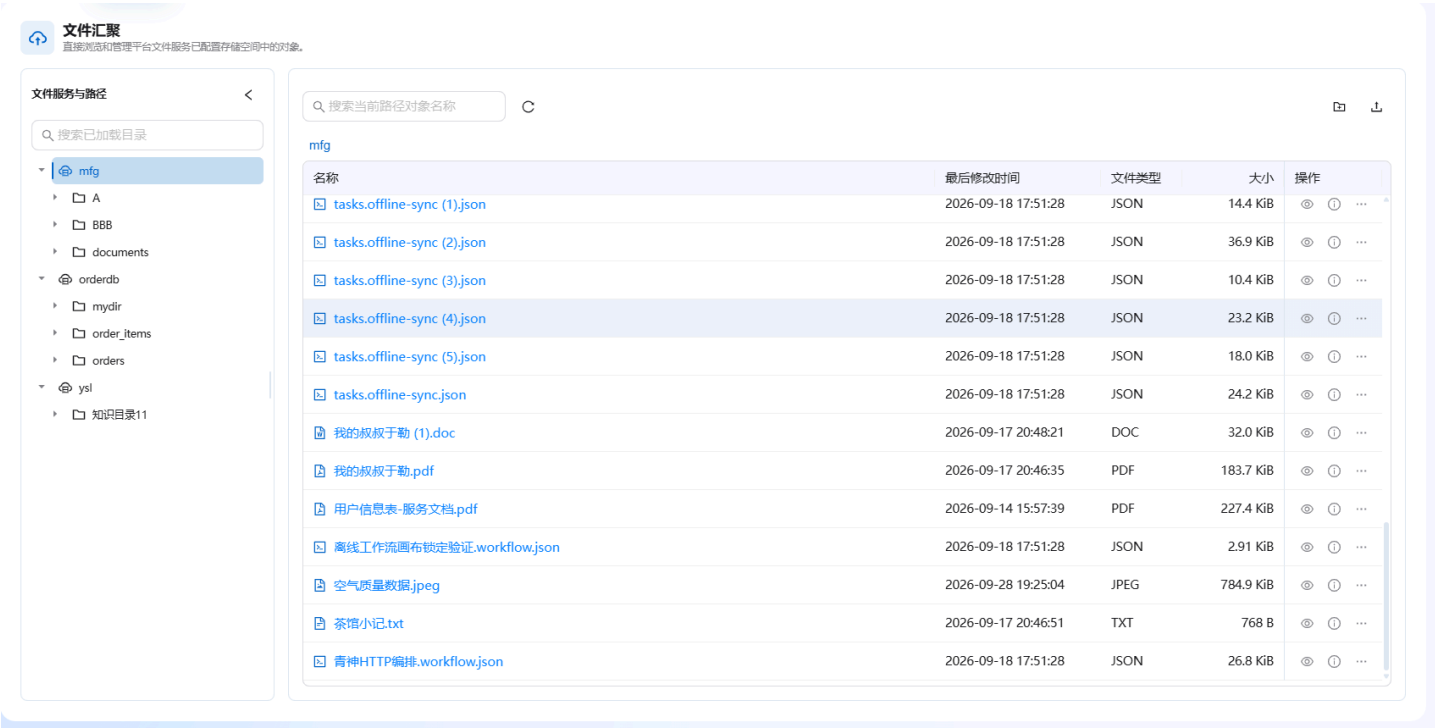


图 2文件列表

说明：

- 从左侧选择文件服务与路径，右侧查看名称、最后修改时间、文件类型、大小和操作。路径表示当前文件所在位置。
- 搜索按名称前缀缩小范围；列表滚动到底部继续加载。切换路径后重新确认当前位置，避免将文件上传到其他目录。

我的叔叔于勒

莫泊桑（法国）

一个白胡子穷老头儿向我们乞讨小钱，我的同伴若瑟夫·达佛朗司竟给了他五法郎的银币。我觉得很奇怪。他于是对我说：“这个穷汉使我回想起一桩故事，这故事，我一直记着不忘，我这就讲给你听。事情是这样的：我的家庭原籍哈佛尔，并不是有钱的人家，也就是刚刚够生活罢了。我父亲做着事，很晚才从办公室回来，挣的钱不多。我有两个姐姐。

我母亲对我们的拮据生活感到非常痛苦。那时家里样样都要节省，有人请吃饭是从来不敢答应的，以免回请；买日用品也是常常买减价的，买拍卖的底货；姐姐的长袍是自己做的，买十五个铜子一米的花边，常常要在价钱上计较半天。

可是每星期日，我们都要衣冠整齐地到海边栈桥上去散步。那时候，一看见从远地方回来的大海船开进港口来，父亲总要说他那句永不变更的话：“唉！如果于勒竟在这只船上，那会叫人多么惊喜呀！”

我的叔叔，父亲的弟弟于勒叔叔，那时候是全家唯一的希望，在这以前则是全家的恐怖。

据说他当初行为不正，糟蹋钱。在穷人家，这是最大的罪恶。在有钱的人家，一个人好玩乐无非算作糊涂荒唐，大家笑嘻嘻地称他一声“花花公子”。在生活困难的人家，一个人要是逼得父母动老本，那就是坏蛋，就是流氓，就是无赖了。于勒叔叔把自己应得的部分遗产吃得一干二净之后，还大大占用了我父亲应得的那一部分。

人们按照当时的惯例，把他送上从哈佛尔到纽约的商船，打发他到美洲去。

我这位于勒叔叔一到那里就写了信回来，说他赚了点儿钱，希望能够赔偿我父亲的损失。这封信在我们家里成了福音。大家分读了又读，拿给所有的人看。

有十年工夫，于勒叔叔再也没有来信，可是父亲的希望却与日俱增。母亲也常常说：“只要这个好心的于勒一回来，我们的境况就不同了。他可真算得一个有办法的人。”

于是每星期日，一看见大轮船喷着黑烟从天边驶过来，父亲总是重复他那句永不变更的话：“唉！如果于勒竟在这只船上，那会叫人多么惊喜呀！”

那时候大家简直好像马上就会看见他挥着手帕喊着：“喂！菲利普！”

对于叔叔回国这桩十拿九稳的事，大家还拟定了上千种计划，甚至计划到要用这位叔叔的钱置一所别墅。我不敢肯定父亲对于这个计划是不是进行了商谈。

我的大姐那时二十八岁，二姐二十六岁。她们老找不着对象，这是全家都十分发愁的事。

终于有一个看中二姐的人上门来了。他是公务员，没有什么钱，但是诚实可靠。我总认为这个青年之所以不再迟疑而下决心求婚，是因为有一天晚上我们给他看了于勒叔叔的信。

我们家赶忙答应了他的请求，并且决定在举行婚礼之后全家到哲尔赛岛去游玩一次。哲尔赛岛是穷人们最理想的游玩的地方。这个小岛是属英国管的。路并不远，乘小轮船渡过海，便到了。因此，一个法国人只要航行两个小时，就可以到一个邻国，看看这个国家的民族，并且研究一下这个不列颠国旗覆盖着的岛上的风俗习惯。

哲尔赛的旅行成了我们的心事，成了我们时时刻刻的渴望和梦想。后来我们终于动身了。我们上了轮船，

图 3对象信息

说明：

- 打开对象信息，查看文件对象键、大小、更新时间、内容类型及服务返回的其他元信息。
- ETag、版本等信息用于识别对象；对象信息与内容预览是不同入口，不用于编辑文件内容。

2.1.2 新建路径

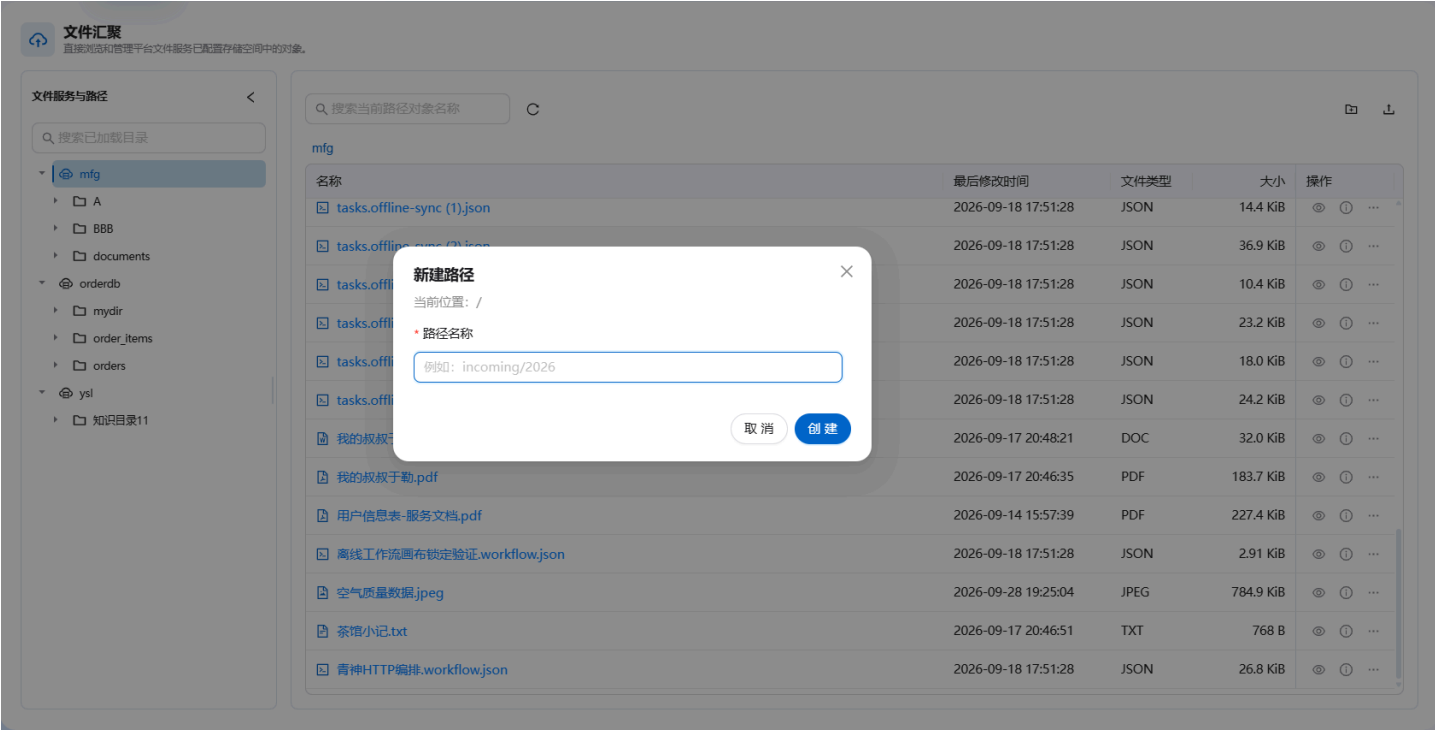


图 4新建路径

说明：

- 在目标文件服务下进入父路径，点击新建路径，填写路径名称并确认。
- 新路径用于组织后续文件；创建后在目录中确认位置，再执行上传。

2.1.3 上传文件

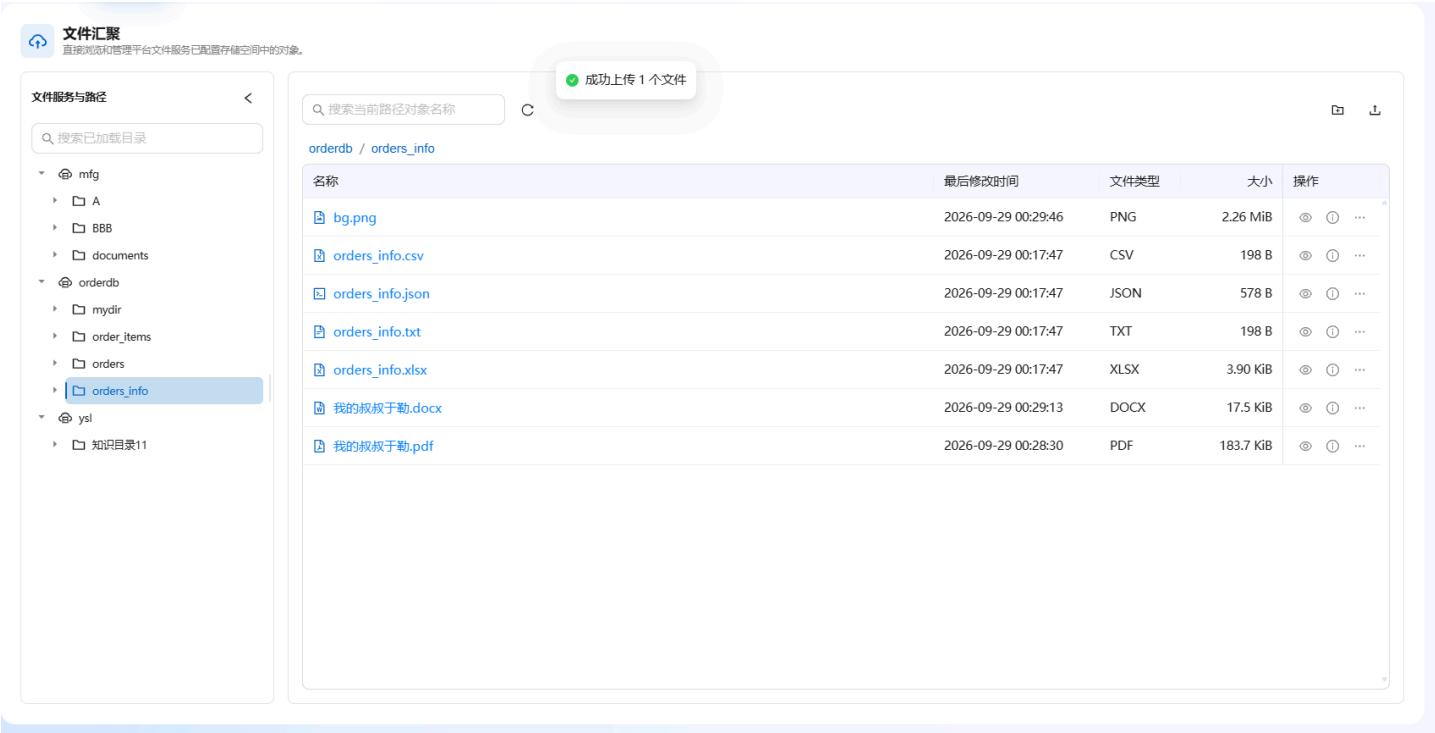


图 5上传文件

- 说明：
- 进入目标路径，点击上传并选择本地文件，查看上传状态。
  - 上传成功后刷新列表，核对文件名和大小。遇到同名提示时先确认是否需要覆盖原文件，不要直接忽略覆盖风险。

2.1.4 预览与下载文件

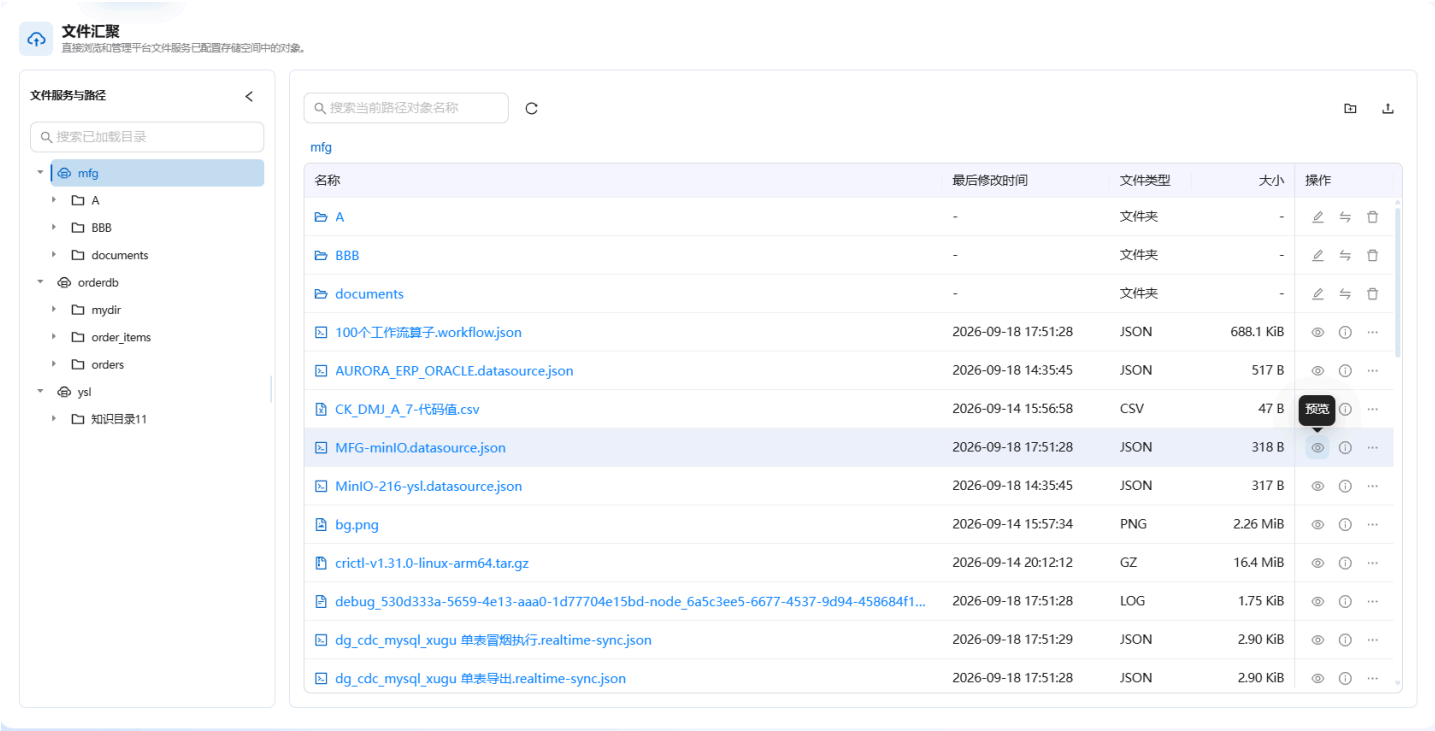


图 6预览文件

- 说明：
- 预览文件
- 点击文件预览入口，在预览抽屉中查看内容。PDF、图片、音频、视频可按相应方式查看或播放。
  - TXT、Markdown、CSV、JSON、Word及Excel等文件使用对应解析预览；解析类文件大小上限为 50 MiB，表格最多预览前1000行、100列。预览不是完整数据导出。
  - 不支持预览或超过限制时，可下载后使用本地应用打开。

下载文件

- 在文件所在行选择下载，将文件保存到本地。
- 下载前核对文件名与路径；含业务敏感信息的文件应按企业规定保存和使用。

2.1.5 整理文件

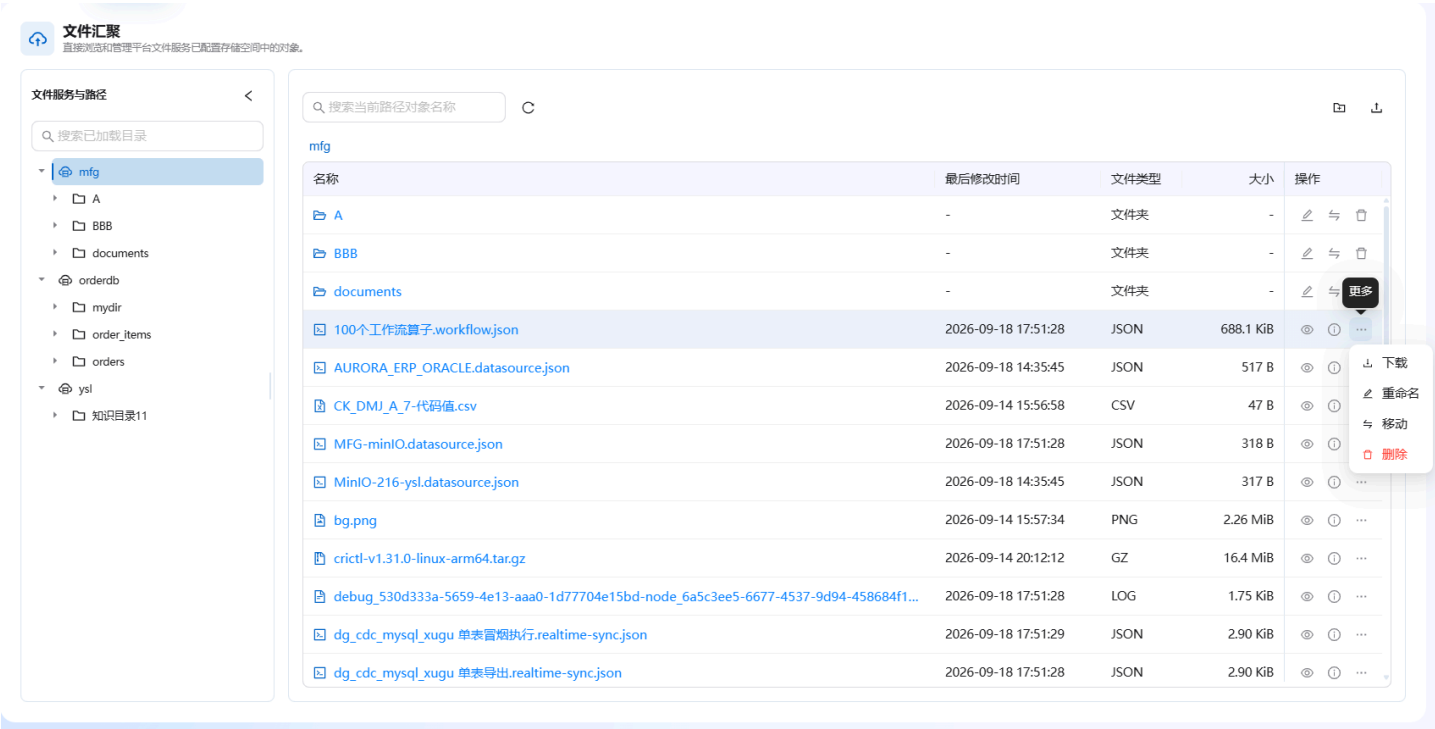


图 7重命名

- 说明：
- 选择重命名，填写新名称并确认。目标名称已存在时，先检查覆盖提示。
  - 操作后确认新旧名称状态。若复制成功但源对象删除失败，可能同时保留两份文件，应依据提示处理，不能只凭新名称出现判断全部完成。

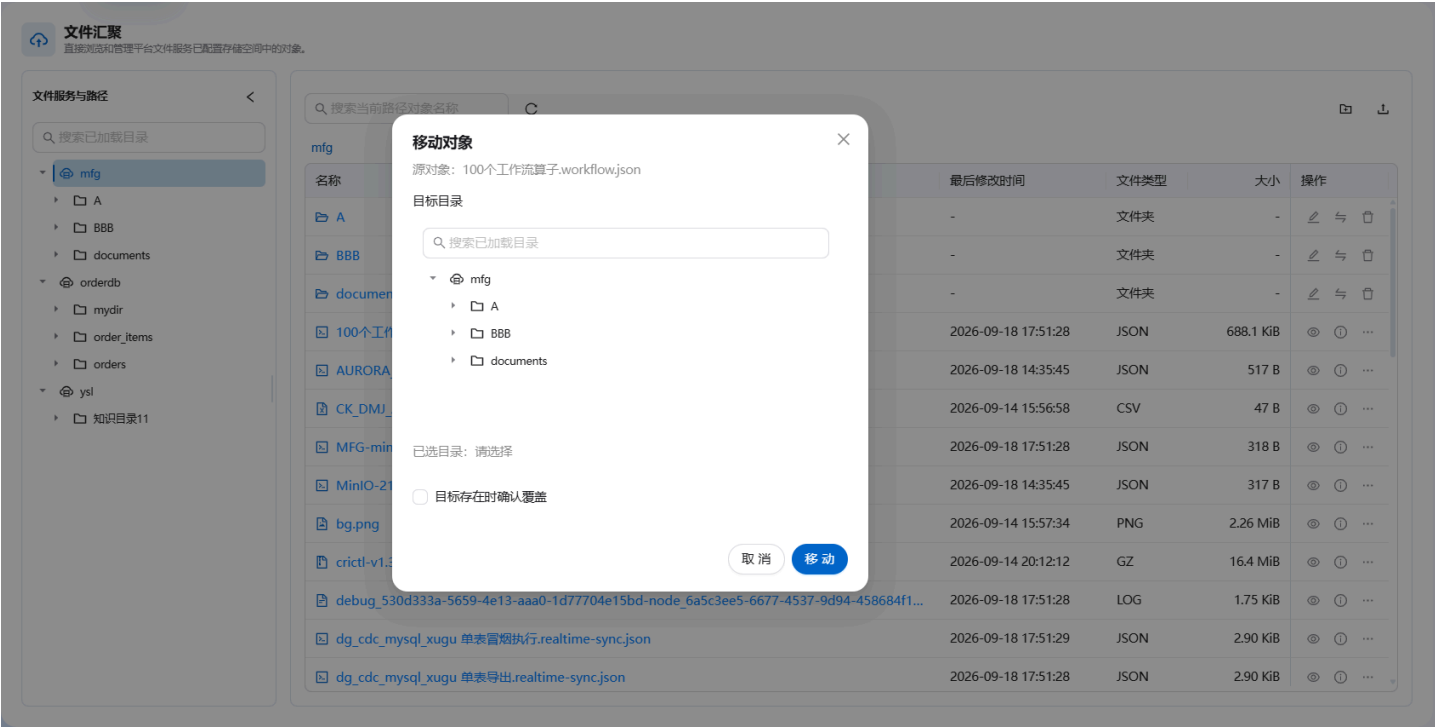


图 8移动文件

- 说明：
- 选择移动并指定目标路径，确认目标位置后提交。

- 目标存在同名文件时需确认覆盖；完成后分别核对目标文件和源路径。移动过程中出现部分失败时，先检查实际对象再重试。

### 2.1.6 删除文件

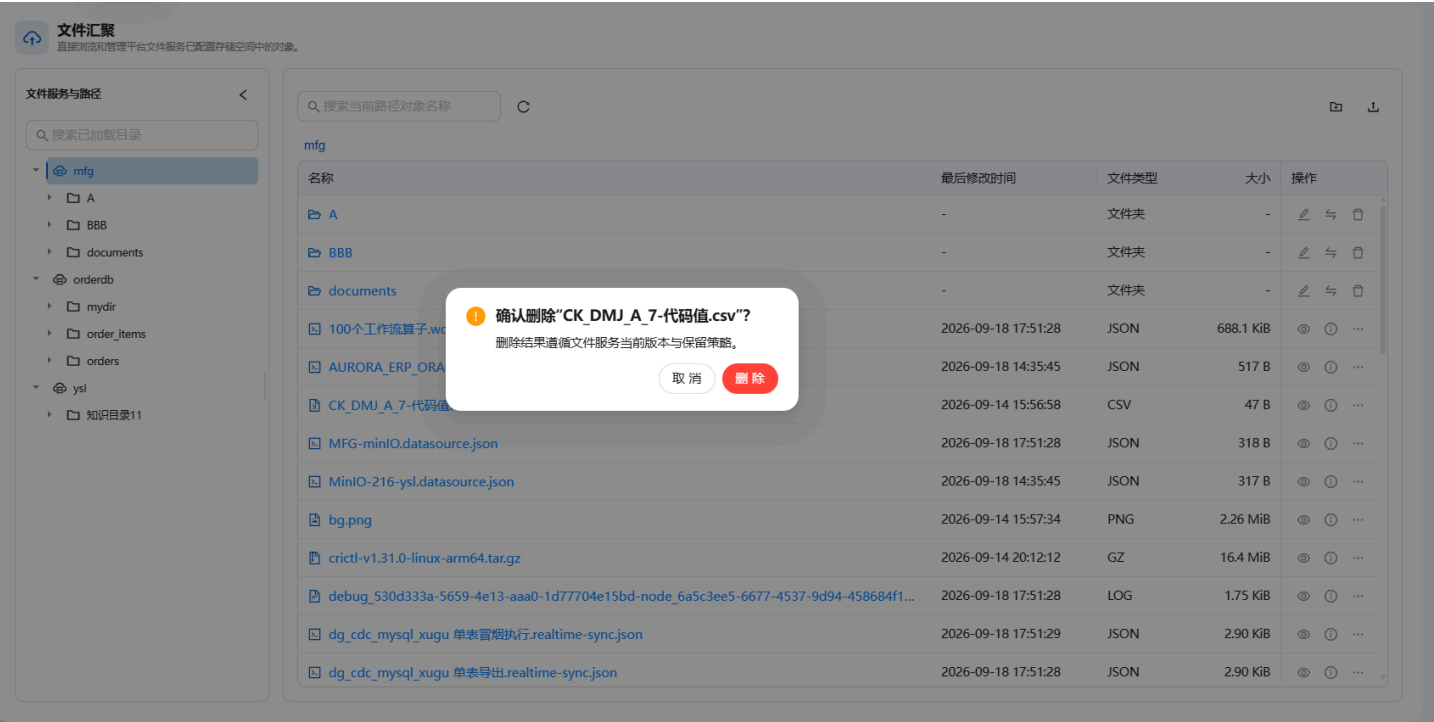


图 9删除文件

说明：

- 在目标文件行选择删除，确认对象与路径后提交。
- 删除会影响文件服务中的实际对象，不只是隐藏列表记录。需要保留的文件应先备份，再执行删除。

## 2.2 离线同步

离线同步面向数据库表的批量搬运。当前任务按全量方式读取来源表，可同时配置多张表，分别指定目标表与字段映射。保存、发布和运行是三个独立动作。

### 2.2.1 离线同步任务列表

离线同步

管理离线同步任务草稿及其版本发布版本。

任务目录

Q 搜索目录或同步任务

全部 82

dg\_sync\_to\_xugu 16

dg\_mysql\_xugu\_sync\_sync0...

dg\_sync\_oracle\_xugu 同步

dg\_sync\_dameng\_xugu 同步

dg\_mysql\_xugu\_版本切换验...

dg\_sync\_kingbase (mysql...

dg\_sync\_kingbase (pg) \_x...

dg\_sync\_kingbase (sqlserv...

dg\_sync\_mysql\_xugu 同步

dg\_sync\_opengauss\_xugu ...

dg\_sync\_postgresql\_xugu ...

dg\_sync\_postgresql\_xugu ...

dg\_sync\_sqlserver\_xugu 同...

dg\_sync\_sqlserver\_xugu 同...

dg\_sync\_xugu\_xugu 同步

dg\_sync\_mysql\_xugu\_参数...

mysql\_user\_basic\_info\_sync...

dg\_sync\_fro\_xugu 9

dg\_sync\_xugu\_dameng 同步

Q 搜索任务名称

来源数据源

目标数据源

创建时间起点

创建时间终点

筛选状态

C

| 任务名称                | 描述               | 来源类型   | 来源                 | 目标类型   | 目标               | 状态  | 操作       |
|---------------------|------------------|--------|--------------------|--------|------------------|-----|----------|
| mysql_user_basic... | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 已发布 | 编辑 删除 更多 |
| 1111                | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 草稿  | 编辑 删除 更多 |
| aaaaa               | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 草稿  | 编辑 删除 更多 |
| 222                 | -                | XuguDB | 数据同步数仓DW4...       | MySQL  | mysql数据源dg_sy... | 草稿  | 编辑 删除 更多 |
| dg_sync_xugu_m...   | -                | XuguDB | mysql数据源dg_sy...   | MySQL  | mysql数据源dg_sy... | 已发布 | 编辑 删除 更多 |
| dg_sync_mysql_x...  | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 已发布 | 编辑 删除 更多 |
| dg_sync_mysql_x...  | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 已发布 | 编辑 删除 更多 |
| dg_sync_mysql_x...  | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 已发布 | 编辑 删除 更多 |
| 1dg_mysql_xugu...   | dg_mysql_xugu... | MySQL  | mysql_mfg_cdc_test | XuguDB | 数据同步数仓DW4...     | 草稿  | 编辑 删除 更多 |
| 离线工作流和离线...         | -                | MySQL  | mysql数据源dg_sy...   | XuguDB | 数据同步数仓DW4...     | 已发布 | 编辑 删除 更多 |

共 82 个任务 < 1 2 3 4 5 ... 9 > 10 条/页

图 10离线同步列表

- 说明：
- 进入“离线同步”，按名称、来源、目标及页面提供的状态条件查找任务。查看名称、连接、状态及版本信息。
  - 点击名称查看配置，使用行操作编辑、发布、导出或删除；新建与导入位于列表工具栏。发布中的任务通过进度入口跟踪处理结果。

2.2.2 新建离线同步任务

新建离线同步

1 数据源 2 同步配置 3 表配置

\* 任务名称

0 / 64

负责人

选择负责人 (非必填)

\* 来源数据源

请选择来源数据源

XuguDB >

Oracle >

MySQL >

Dameng >

PostgreSQL >

(非必填)

\* 目标数据源

请选择目标数据源

0 / 500

保存草稿

下一步

新建离线同步

1 数据源

2 同步配置

3 表配置

任务名称

0 / 64

负责人

选择负责人 (非必填)

来源数据源

请选择来源数据源

目标数据源

请选择目标数据源

Oracle >

MySQL >

Dameng >

PostgreSQL >

SQLServer >

描述

请输入任务描述 (非必填)

0 / 500

保存草稿

下一步

图 11新建任务

说明：

- 点击新建，填写任务名称，例如“物流订单首次入仓”，分别选择来源、目标的数据源类型与实例，可指定负责人。
- 离线同步通过平台统一维护的数据源连接，支持 Xugu、Oracle、Mysql、达梦、PostgreSQL、SQLServer、OpenGauss、人大金仓等多种类型数据源，配置来源端与目标端，实现已支持的集中式存储（如Mysql、PostgreSQL等）和分布式存储（如Xugu、达梦等）之间的批量数据采集与传输。不同存储类型采用对应的读取、写入配置。
- 创建任务时，选择来源数据源、采集对象及目标数据源，配置字段映射、写入方式和执行参数。来源端与目标端可属于不同存储环境，用于跨存储的数据复制或迁移准备。
- 选择的数据源应已启用、连接可用且当前用户有权访问。点击下一步进入同步配置；更换来源后应重新检查表和字段映射。
- 完成配置后，可按平台提供的执行及调度方式运行任务，并在任务记录中检查执行状态、读取量、写入量及异常信息。

← 编辑离线同步

来源 MySQL / mysql\_mfg\_cdc\_test → 目标 XuguDB / 数据同步数仓DW4SYNC

数据源

2 同步配置

3 表配置

任务配置

并行度

4

每线程读取限速 (行/秒)

500

每线程读取限速 (字节/秒)

不限速

引擎作业重试次数

3

引擎作业重试间隔 (秒)

3

来源读取配置

分片行数

8096

JDBC 抓取行数

1000

连接检查超时 (秒)

30

目标写入配置

目标表结构处理

目标表不存在时创建

目标已有数据处理

追加数据

表前缀

mysql\_

表后缀

\_version\_sync\_001

启用 Upsert

☒

写入批大小

1000

连接检查超时 (秒)

30

批量写入重试次数

0

上一步

保存草稿

下一步

图 12同步配置

#### 说明：

- 在同步配置中设置并行度，默认4；每线程读取行数与字节上限用于限速，0表示不限制。
- 来源分片行数默认8096，JDBC抓取行数默认1000，连接检查超时默认30秒。分片控制拆分粒度，抓取行数控制一次从数据库获取的数据量，两者用途不同。
- 当前离线快捷同步为全量读取。重复运行是否产生重复数据，取决于目标已有数据处理与Upsert设置。
- 根据目标表是否已存在选择结构处理策略；自动创建时核对表名前后缀。重建表可能删除原结构和数据，操作前必须确认影响。
- 已有数据处理可按需要选择追加、清空后写入、自定义SQL或有数据时报错。追加不会自动去重，清空后写入也不等于整个任务具备自动回滚。
- 启用Upsert时，后续需配置目标匹配键。写入批大小默认1000、连接检查超时默认30秒、批量写入重试次数默认0。
- COPY或XA等选项仅在支持的目标连接中出现；按对应说明配置，不应将单个连接的事务能力理解为所有表、所有任务的统一事务。
- 进入表配置，点击添加来源表，在弹窗中按表名或注释查找，核对Schema和已添加状态后勾选。支持跨页选择，单个任务最多50张表。
- 确认后每张来源表对应一行映射。检查是否选中了业务Schema中的表，避免选择平台运行表。

← 编辑离线同步

来源 Oracle / oracle数据源dg\_sync → 目标 XuguDB / 数据同步数仓DW4SYNC

✓ 数据源

✓ 同步配置

3 表配置

Q 搜索来源表

已添加1 张表 · 单任务最多50 张

| <input type="checkbox"/> | 来源表                            | 目标表                                                    | Upsert | 目标端唯一键  | 源端分片字段 | 字段结构 | 操作 |
|--------------------------|--------------------------------|--------------------------------------------------------|--------|---------|--------|------|----|
| <input type="checkbox"/> | USER_INFO<br>用户信息表: 台成会员档案及... | USER_INFO<br>ORACLE_USER_INFO_MULTI_TYPE_0001 · 将复用已有表 | 开启     | USER_ID | 请选择    | 配置字段 |    |

共 1 张表 / 1 页 < 1 > 10 条/页

上一步

保存草稿

完成

图 13保存任务

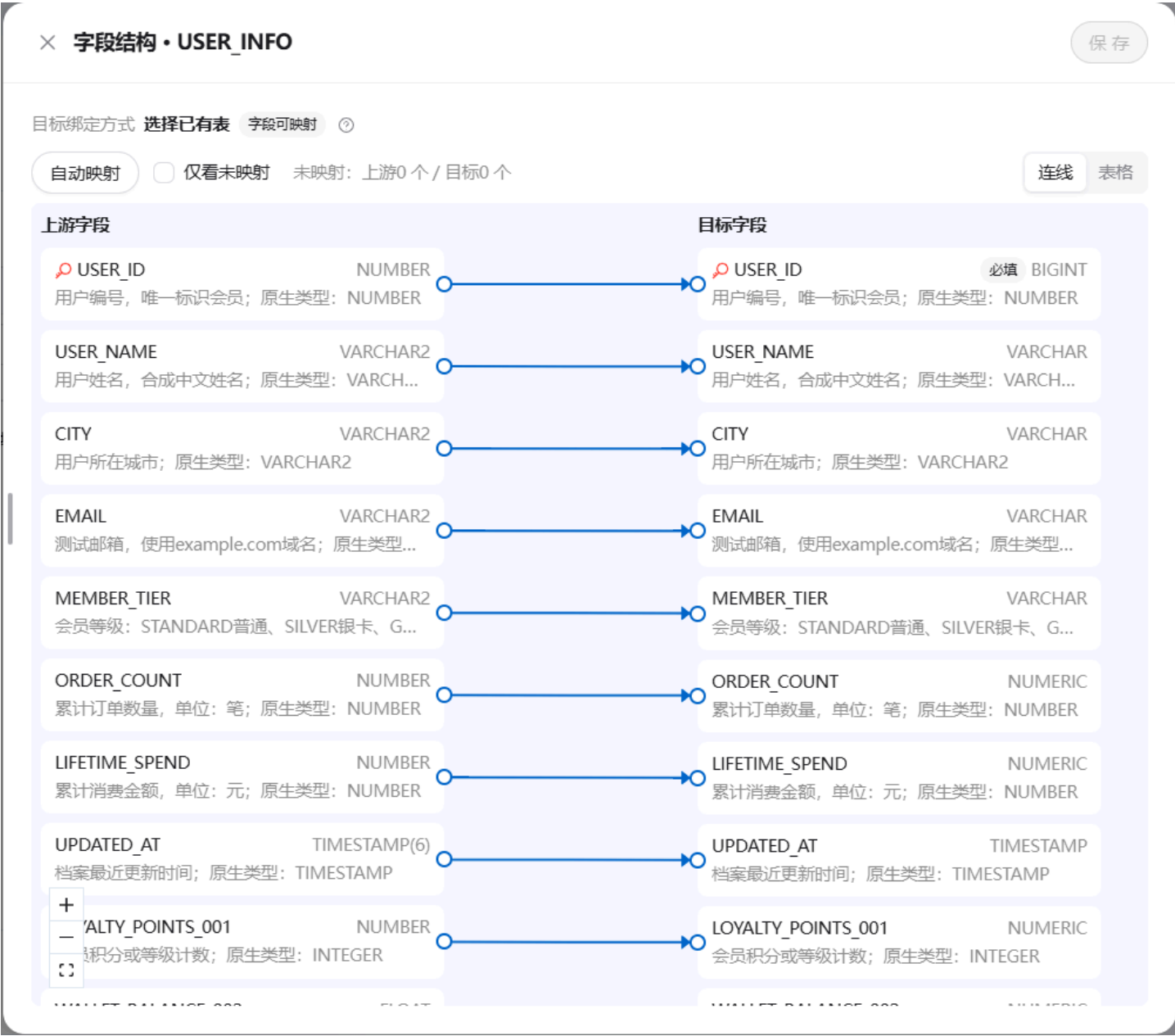


图 14 字段映射配置

- 说明：
- 逐行选择新建目标表或选择已有表。新建方式核对生成的名称；已有表方式选择实际目标表。
  - 同一任务内避免多张来源表重复指向同一目标表。表结构策略和逐表目标方式应一致。
  - 启用Upsert时，为目标表选择能够唯一识别记录的匹配键，可使用组合键。例如运单表可使用真实唯一的运单ID。
  - 分片字段用于拆分来源读取，可选择支持的数值、字符或日期字段；它不等于业务唯一键。请结合数据分布选择，避免大量数据集中在一个取值。
  - 在表配置中使用行删除，或勾选后批量移除。
  - 该操作移除任务中的表配置，不会直接删除来源或目标数据库中的物理表。移除后保存任务使配置生效。
  - 配置尚未完成时使用保存草稿；完成各步骤后点击完成保存任务并返回列表。

- 保存不会自动发布，也不会开始同步。需要运行时，还需发布并到生产工作流执行。
- 点击查看字段。自动建表方式展示同名字段及自动类型映射；使用已有表时，逐项选择对应目标字段。
- 检查来源字段、目标字段、类型和注释。已有表映射应一对一，至少保留一个有效映射，写入所需的匹配键等字段必须映射完整。
- 点击抽屉中的保存保留映射；保存后仍可继续检查字段，再关闭抽屉。类型或必填校验未通过时，先修正配置后完成任务。

### 2.2.3 编辑离线同步任务

← 编辑离线同步

1 数据源

2 同步配置

3 表配置

\* 任务名称

dg\_sync\_oracle\_xugu 同步(多字段类型)01

31 / 64

负责人

选择负责人 (非必填)

\* 来源数据源

Oracle / oracle数据源dg\_sync

\* 目标数据源

XuguDB / 数据同步数仓DW4SYNC

描述

请输入任务描述 (非必填)

0 / 500

保存草稿

下一步

图 14编辑任务

说明：

- 在列表中选择编辑，依次检查数据源、同步配置与表配置，修改后保存。
- 已发布任务的草稿修改不直接改变现有生产版本；需要发布新版本后，生产运行才使用新的配置。

### 2.2.4 离线任务详情



图 15任务详情

说明：

查看离线任务详情

- 点击任务名称打开详情，通过页签查看保存的任务配置。
- 离线任务详情—数据源
- 查看任务名称、负责人及来源和目标连接，确认数据流向。

×

dg\_sync\_oracle\_xugu 同步(多字段类型)01 • 任务详情

数据源

同步配置

表配置 (1)

同步策略

同步方式: 全量同步

任务配置

并行度: 4

每线程行数限速: 不限速

每线程字节限速: 不限速

来源读取配置

分片行数: 8096

JDBC 抓取行数: 1000

来源连接检查超时: 30 秒

目标写入配置

目标表结构处理: 目标表不存在时创建

表前缀: ORACLE\_

表后缀: \_MULTI\_TYPE\_0001

目标已有数据处理: 追加数据

写入批大小: 1000

连接检查超时 (秒) : 30

批量写入重试次数: 0

启用 Upsert: 启用

图 16同步配置

说明：

- 检查读取、写入与性能策略，重点核对全量读取、已有数据处理和Upsert。

数据源 同步配置 表配置 (50)

| 来源表        | 目标表             | 写入策略 | 唯一键      | 分片读取 | 字段结构               |
|------------|-----------------|------|----------|------|--------------------|
| actor_0001 | 新建   actor_0001 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0002 | 新建   actor_0002 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0003 | 新建   actor_0003 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0004 | 新建   actor_0004 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0005 | 新建   actor_0005 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0006 | 新建   actor_0006 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0007 | 新建   actor_0007 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0008 | 新建   actor_0008 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0009 | 新建   actor_0009 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |
| actor_0010 | 新建   actor_0010 | 追加   | actor_id | 自动推断 | <a href="#">查看</a> |

图 17表配置

说明：

- 核对来源表、目标表、匹配键、分片字段；通过查看字段检查字段映射。

2.2.5 发布与版本管理

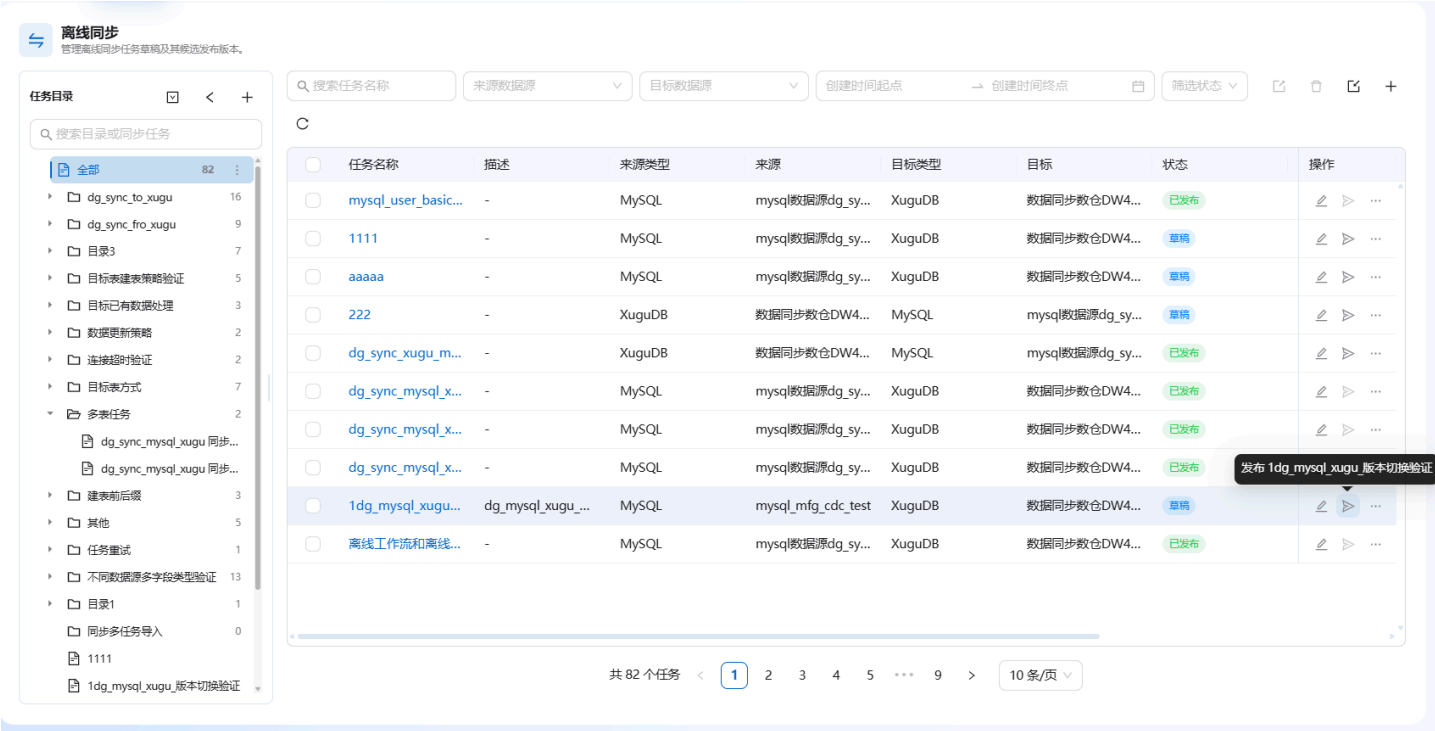


图 18发布任务

- 说明：
- 在列表选择发布，确认待发布配置后提交。需要审批时，到审批管理跟踪；免审时查看发布处理结果。
  - 发布成功后形成生产版本。到生产工作流检查版本，再设置定时计划或手动运行，不要将保存成功当作发布成功。

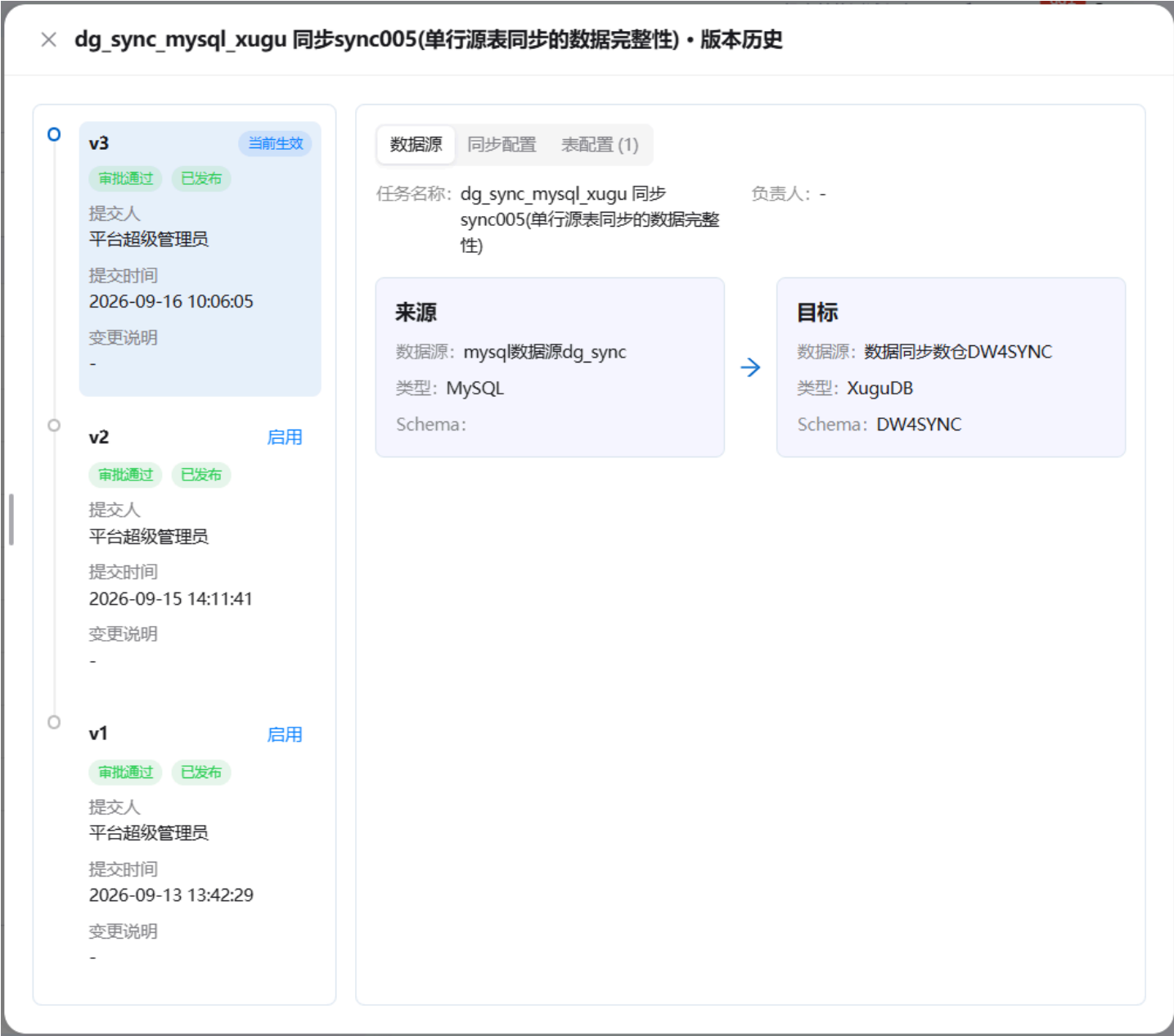


图 19任务版本

说明：

查看离线任务版本

- 点击版本入口打开版本历史，选择版本查看对应保存信息。
- 比对当前任务与生产版本时，以版本号和配置内容为准；修改草稿后不要假定生产版本已更新。

启用历史离线版本

- 在版本历史中选择已发布且不是当前生产版本的版本，使用启用入口并确认。
- 操作后核对当前生产版本。切换版本不撤销此前已写入的数据，执行前确认旧配置仍适用。

2.2.6 导入与导出

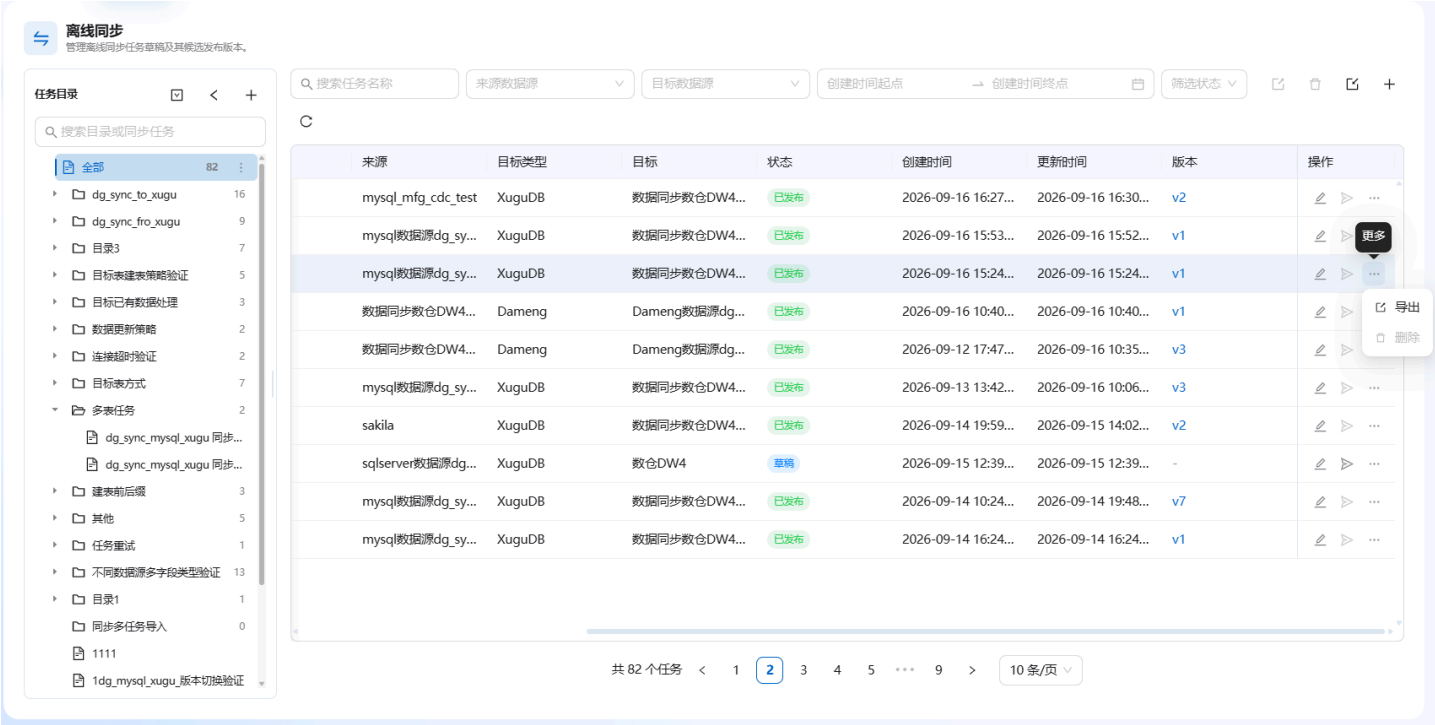


图 20导入导出

- 说明：
- 导入离线任务
- 点击导入，第一步选择配置文件，第二步映射本环境来源和目标数据源，第三步核对导入信息并确认。
  - 导入创建独立草稿；完成后逐项检查表、字段、写入策略，再发布运行。
- 导出离线任务
- 选择导出，保存任务配置文件。导出内容以已保存配置为准，未保存修改应先保存。
  - 配置文件不包含可直接迁移的连接凭据；转移到其他环境后仍需重新映射可访问的数据源。

2.2.7 删除离线任务

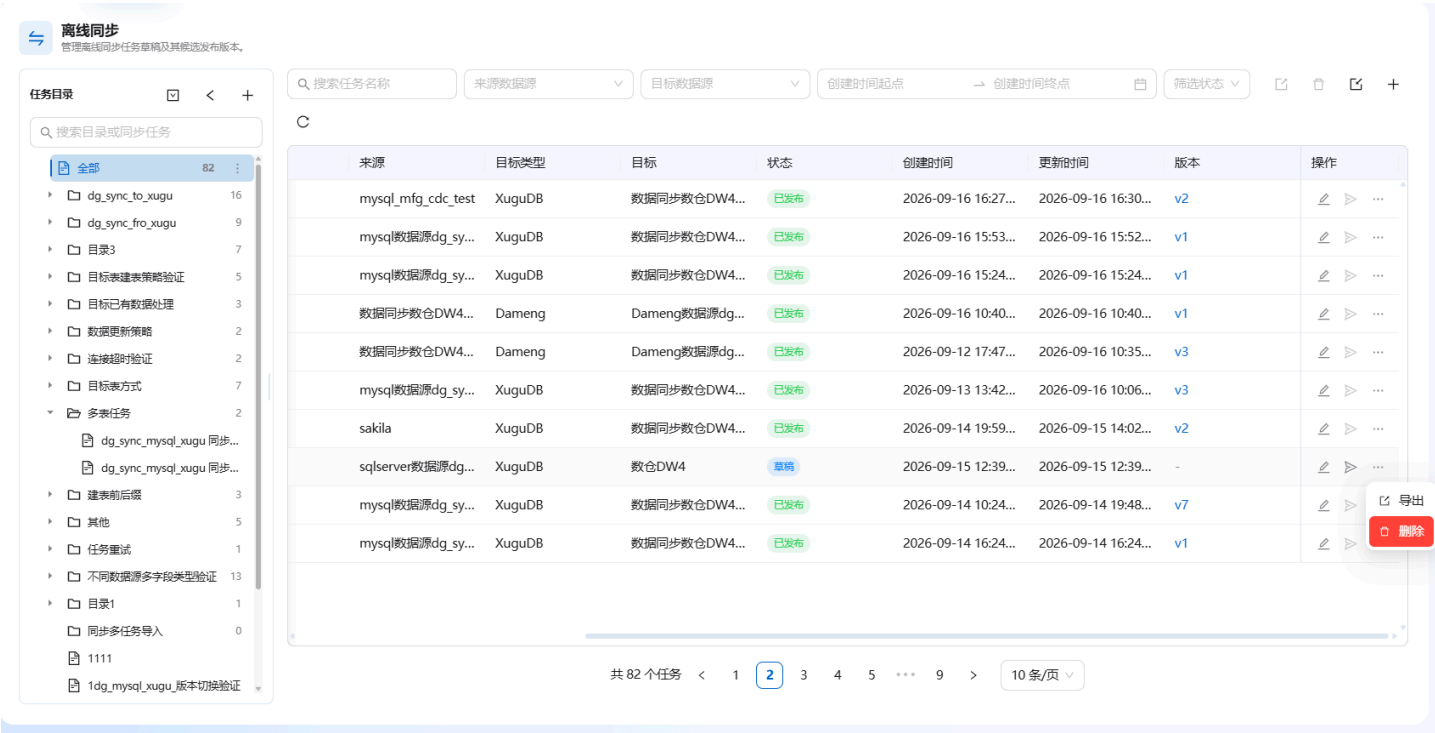


图 21删除任务

- 说明：
- 在列表选择删除并确认目标任务。已发布任务应先停止后续调度，并确保没有仍在运行的实例。
  - 删除或退役任务定义不等于删除历史执行记录，也不会自动清理已经写入的业务数据。

## 2.3 实时同步

实时同步使用数据库变更捕获持续同步新增、修改和删除。它包含数据源、同步配置、表配置、CDC与恢复四个步骤，首次启动与故障恢复需结合来源日志条件和恢复点配置。

### 2.3.1 实时同步任务列表

实时同步

管理实时同步任务草稿及其候选发布版本。

任务目录

搜索目录或同步任务

全部

55

基础任务管理执行

4

dg\_cdc\_mysql\_xugu 单表直...

dg\_cdc\_oracle\_xugu 单表直...

dg\_cdc\_postgresql\_xugu 单...

dg\_cdc\_sqlserver\_xugu 单...

基础任务管理执行导入

0

基础任务管理执行5

4

dg\_cdc\_mysql\_xugu 单表执...

dg\_cdc\_oracle\_xugu 单表执...

dg\_cdc\_postgresql\_xugu 单...

dg\_cdc\_sqlserver\_xugu 单...

多表同步

4

dg\_cdc\_mysql\_xugu\_dw4...

dg\_cdc\_oracle\_xugu 多表执...

dg\_cdc\_oracle\_xugu 多表执...

dg\_cdc\_sqlserver\_xugu 多...

Oracle数据源实时同步验证

3

oracle\_xugu\_单表实时任务\_02

Oracle实时同步单表全量加...

搜索任务名称

来源数据源

目标数据源

创建时间起点

创建时间终点

筛选状态

| 任务名称               | 描述             | 来源类型   | 来源                 | 目标类型   | 目标           | 状态  | 操作                                                       |
|--------------------|----------------|--------|--------------------|--------|--------------|-----|----------------------------------------------------------|
| Mysql全量加增量...      | -              | MySQL  | mysql_mfg_cdc_test | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| Oracle实时同步单...     | -              | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| Oracle实时同步单...     | -              | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-Xugu大字段-1...    | UTF8字符集复验; ... | MySQL  | mysql_mfg_cdc_test | XuguDB | 数仓DW4        | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| Mysql大字段表同...      | -              | MySQL  | mysql_mfg_cdc_test | XuguDB | 数仓DW4        | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-cdc-0922-200... | 隔离真数据验证; ...   | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-cdc-0922-200... | 隔离真数据验证; ...   | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-cdc-0922-200... | 隔离真数据验证; ...   | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-cdc-0922-200... | 隔离真数据验证; ...   | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |
| qa-cdc-0922-200... | 隔离真数据验证; ...   | Oracle | ORACLE_MFG_CD...   | XuguDB | 数据同步数仓DW4... | 已发布 | <a href="#">编辑</a> <a href="#">发布</a> <a href="#">更多</a> |

共 55 个任务 < 1 2 3 4 5 6 > 10 条/页

图 22实时同步任务列表

- 说明：
- 进入实时同步，通过搜索与筛选定位任务，查看状态、来源目标和版本。
  - 点击名称查看详情；编辑、发布、导入、导出、删除分别通过相应入口完成。实际启动和停止在生产 workflows 中操作。

### 2.3.2 新建实时同步任务

新建实时同步

1 数据源

2 同步配置

3 表配置

4 CDC 与恢复

任务名称

负责人

0 / 64

选择负责人 (非必填)

描述

请输入任务描述 (非必填)

0 / 500

来源数据源

目标数据源

请选择支持 CDC 的来源数据源

请选择目标数据源

Oracle

MySQL

PostgreSQL

SQLServer

保存草稿

下一步

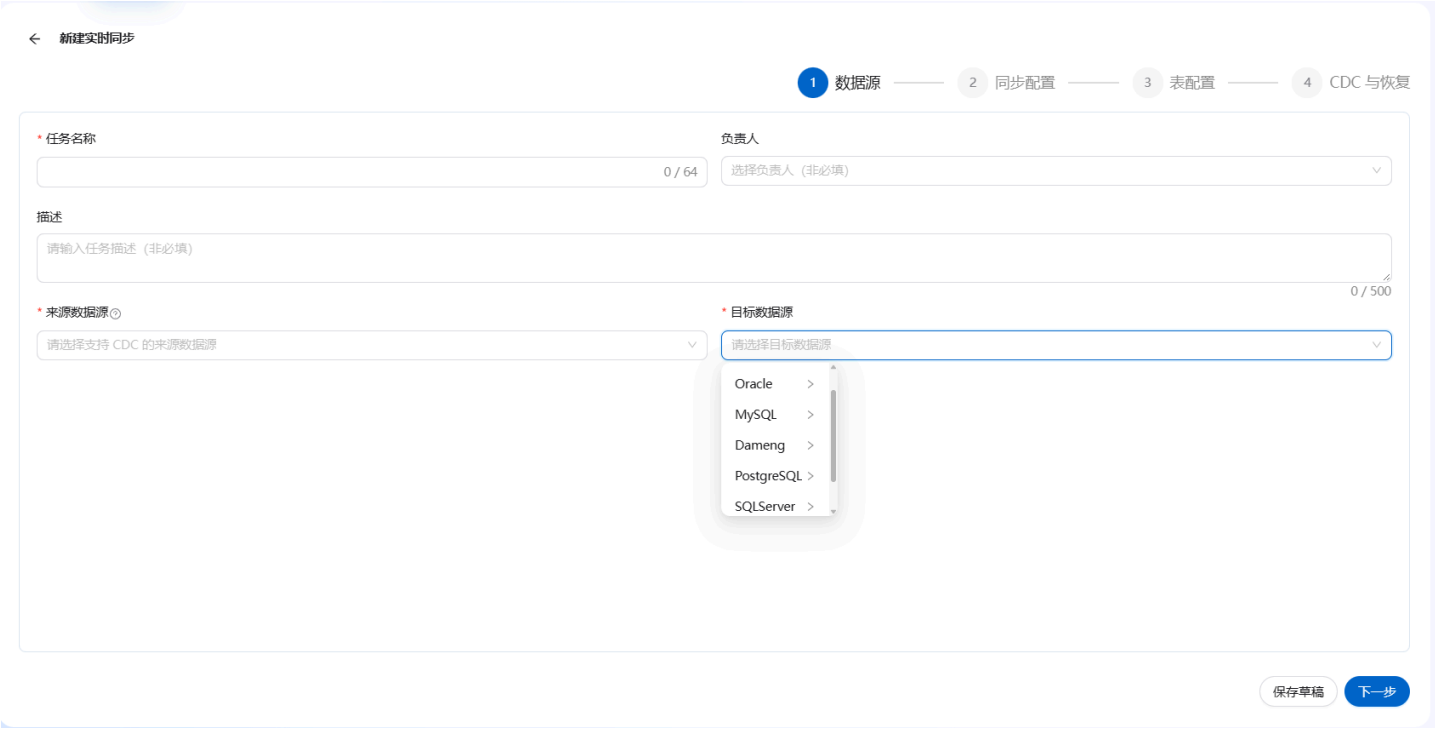


图 23数据源选择

- 说明：
- 点击新建，填写名称，选择受支持的数据库来源与目标实例，按需要指定负责人。
  - 实时来源支持MySQL、Oracle、SQL Server和PostgreSQL。按来源要求准备日志、权限与保留时间，使用前置检查确认条件；平台不会代替数据库管理员开启日志或修改授权。



图 24同步配置

说明：

- 在同步配置设置并行度、读取限速、快照读取参数及连接参数。Checkpoint超时用于限制检查点完成时间，作业重试次数和间隔用于运行恢复。
- 结构变更同步仅在相应能力支持时开启；目标选择自动建表或使用已有表，核对前后缀、Upsert、写入批大小及批量重试。
- 实时同步按变更类型处理数据。关闭Upsert不应理解成只做追加、忽略更新和删除。

新建实时同步

来源 Oracle / ORACLE\_MFG\_CDC\_TEST → 目标 XuguDB / 数据同步数仓DW4CDC1

数据源同步配置表配置CDC与恢复

Q 搜索来源表

全部 10 检测通过 10 检测不通过 0

· 单任务最多50 张

| <input type="checkbox"/> | 来源表                                      | CDC 就绪 | 目标表                                                          | Upsert | 目标主键唯一键 | 字段结构 | 操作 |
|--------------------------|------------------------------------------|--------|--------------------------------------------------------------|--------|---------|------|----|
| <input type="checkbox"/> | BOM_HEADERS<br>BOM头：产品物料清单表头，记...        | 已就绪    | BOM_HEADERS<br>ORACLE_CDC_BOM_HEADERS · 将新建表                 | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | BOM_ITEMS<br>BOM明细：产品物料清单明细，...          | 已就绪    | BOM_ITEMS<br>ORACLE_CDC_BOM_ITEMS · 将新建表                     | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | CUSTOMERS<br>客户：客户主数据，记录汽车零...           | 已就绪    | CUSTOMERS<br>ORACLE_CDC_CUSTOMERS · 将新建表                     | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | CUSTOMER_COMPLAINTS<br>客户投诉：客户投诉记录，关联... | 已就绪    | CUSTOMER_COMPLAINTS<br>ORACLE_CDC_CUSTOMER_COMPLAINTS · 将新建表 | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | DISPATCH_ORDERS<br>派工单：生产派工记录，分配工...     | 已就绪    | DISPATCH_ORDERS<br>ORACLE_CDC_DISPATCH_ORDERS · 将新建表         | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | EMPLOYEES<br>员工：员工主数据，记录所属年...           | 已就绪    | EMPLOYEES<br>ORACLE_CDC_EMPLOYEES · 将新建表                     | 开启     | ID ×    | 查看字段 |    |
| <input type="checkbox"/> | EQUIPMENT<br>设备：设备台账，记录安装工位...           | 已就绪    | EQUIPMENT<br>ORACLE_CDC_EQUIPMENT · 将新建表                     | 开启     | ID ×    | 查看字段 |    |

上一步

保存草稿

下一步

目标绑定方式 **新建目标表** 自动映射 ⓘ

| 来源字段                                                                                                              |              | 目标字段                                                                                                 |
|-------------------------------------------------------------------------------------------------------------------|--------------|------------------------------------------------------------------------------------------------------|
|  ID<br>记录主键，由数据生成脚本分配，非数据库自增列    | NUMBER       |  ID<br>按目标数据库规则自动映射 |
|  CODE<br>记录唯一业务编号；INC-前缀表示增量测试记录 | NVARCHAR2    | CODE<br>按目标数据库规则自动映射                                                                                 |
| STATUS<br>记录业务状态；基准值ACTIVE=有效；增量DRAFT=草稿...                                                                       | NVARCHAR2    | STATUS<br>按目标数据库规则自动映射                                                                               |
| REVISION<br>记录修订次数；初始为0，增量UPDATE时加1                                                                               | NUMBER       | REVISION<br>按目标数据库规则自动映射                                                                             |
| CREATED_AT<br>记录创建时间，字段不包含时区信息                                                                                    | TIMESTAMP(6) | CREATED_AT<br>按目标数据库规则自动映射                                                                           |
| UPDATED_AT<br>记录最后修改时间，字段不包含时区信息                                                                                  | TIMESTAMP(6) | UPDATED_AT<br>按目标数据库规则自动映射                                                                           |
| PRODUCT_ID<br>产品主键，外键关联products.id                                                                                | NUMBER       | PRODUCT_ID<br>按目标数据库规则自动映射                                                                           |
| BOM_VERSION<br>BOM版本号                                                                                             | NVARCHAR2    | BOM_VERSION<br>按目标数据库规则自动映射                                                                          |
| EFFECTIVE_ON<br>版本生效日期                                                                                            | TIMESTAMP(6) | EFFECTIVE_ON<br>按目标数据库规则自动映射                                                                         |

图 25表配置

- 说明：
- 添加需要订阅的来源表，逐行设置目标表与唯一键。实时每张表均需有效的唯一识别字段。
  - 通过查看字段配置来源与目标映射，覆盖目标主键及必需字段；已有表应先检查结构兼容性。实时表配置不使用离线读取的分片字段。

← 编辑实时同步

来源 Oracle / ORACLE\_MFG\_CDC\_TEST → 目标 XuguDB / 数据同步数仓DW4CDC1

✓ 数据源 — ✓ 同步配置 — ✓ 表配置 — 4 CDC 与恢复

启动配置 ⓘ

\* 首次启动方式

同步现有数据并持续同步

同步现有数据并持续同步

先同步当前全部数据，再持续同步新增、修改和删除

仅同步启动后的变更

不处理现有数据，仅同步任务启动后的变更

30

\* 恢复策略

最近成功检查点

上一步

保存草稿

完成

图 26CDC恢复

## 说明：

### 配置CDC与恢复

- 在CDC与恢复步骤选择来源支持的首次启动方式：先全量后增量读取快照，再接收变化；从最新位置开始不会补齐此前历史数据，同时支持同步现有数据并持续同步（全量+增量）和 仅同步启动后的变更（增量）同步。
- 设置Checkpoint间隔，用于定期保存处理进度。恢复时选择最近成功检查点，或在存在可用候选时选择指定恢复点。
- 恢复点与来源日志需同时可用。不要在日志已清理后假定仍可从旧位置完整恢复；确认无法恢复时再制定重新初始化方案。

### 保存实时任务草稿

- 填写各步骤后保存草稿，或点击完成保存任务。
- 保存只保留配置，不会自动启动CDC。表映射或恢复设置有错误时，按提示修正后再次保存。

## 2.3.3 编辑实时同步任务

← 编辑实时同步

1 数据源

2 同步配置

3 表配置

4 CDC 与恢复

\* 任务名称

Oracle实时同步全量加增量多表17 / 64

负责人

选择负责人 (非必填)

描述

请输入任务描述 (非必填)

0 / 500

\* 来源数据源

Oracle / ORACLE\_MFG\_CDC\_TEST

数据库级已就绪

\* 目标数据源

XuguDB / 数据同步数仓DW4CDC1

保存草稿

下一步

图 27编辑任务

说明：

- 从列表进入编辑，修改需要调整的配置并保存。
- 已运行任务的配置变更需重新发布；发布涉及运行升级时，阅读恢复影响提示，确认可用恢复点后执行，避免直接将旧进度用于不兼容配置。

2.3.4 实时任务详情

数据源 同步配置 表配置 (10) CDC 与恢复

任务名称: Oracle实时同步全量加增量多表

负责人: -

来源

数据源: ORACLE\_MFG\_CDC\_TEST  
类型: Oracle  
数据库: xe  
Schema: MFG\_CDC\_TEST



目标

数据源: 数据同步数仓DW4CDC1  
类型: XuguDB  
数据库: DEMODW  
Schema: DW4CDC1

图 28数据源

说明:

查看实时任务详情

- 打开任务详情，分别查看四个页签。

实时任务详情—数据源

- 确认来源、目标与负责人。

数据源 同步配置 表配置 (10) CDC 与恢复

任务配置

|                         |                   |
|-------------------------|-------------------|
| 同步列结构变更：禁用              | 并行度：4             |
| 每线程读取限速（行/秒）：不限速        | 每线程读取限速（字节/秒）：不限速 |
| Checkpoint 超时（毫秒）：30000 | 引擎作业重试次数：3        |
| 引擎作业重试间隔（秒）：3           |                   |

来源读取配置

|                |                  |
|----------------|------------------|
| 快照分片大小（行）：8096 | 快照单次抓取上限（行）：1024 |
| 连接池大小：20       | 连接超时（毫秒）：30000   |
| 连接重试次数：3       |                  |

目标写入配置

|                   |                                |
|-------------------|--------------------------------|
| 目标表结构处理：自动创建不存在的表 | 写入前目标数据处理：保留目标表数据（APPEND_DATA） |
| 表前缀：ORACLE_CDC_   | 启用 Upsert：启用                   |
| 写入批次大小（条）：1000    | 连接检查超时（秒）：30                   |
| 批量写入重试次数：0        |                                |

图 29同步配置

说明：

- 查看读取、写入、重试和结构变更策略。

数据源 同步配置 表配置 (10) CDC 与恢复

| 来源表                  | 目标表                  | 写入策略    | 唯一键 | 字段结构               |
|----------------------|----------------------|---------|-----|--------------------|
| BOM_HEADERS          | 新建   BOM_HEADERS     | 更新插入并删除 | ID  | <a href="#">查看</a> |
| BOM_ITEMS            | 新建   BOM_ITEMS       | 更新插入并删除 | ID  | <a href="#">查看</a> |
| CUSTOMERS            | 新建   CUSTOMERS       | 更新插入并删除 | ID  | <a href="#">查看</a> |
| CUSTOMER_COMPLAI...  | 新建   CUSTOMER_CO...  | 更新插入并删除 | ID  | <a href="#">查看</a> |
| DISPATCH_ORDERS      | 新建   DISPATCH_ORD... | 更新插入并删除 | ID  | <a href="#">查看</a> |
| EMPLOYEES            | 新建   EMPLOYEES       | 更新插入并删除 | ID  | <a href="#">查看</a> |
| EQUIPMENT            | 新建   EQUIPMENT       | 更新插入并删除 | ID  | <a href="#">查看</a> |
| GOODS_RECEIPTS       | 新建   GOODS_RECEIPTS  | 更新插入并删除 | ID  | <a href="#">查看</a> |
| GOODS_RECEIPT_ITEMS  | 新建   GOODS_RECEIP... | 更新插入并删除 | ID  | <a href="#">查看</a> |
| INCOMING_INSPECTI... | 新建   INCOMING_INS... | 更新插入并删除 | ID  | <a href="#">查看</a> |

共 10 条 / 1 页 < 1 > 10 条/页 ▾

图 30表配置

说明：

- 查看订阅表、目标表、唯一键及字段映射。

数据源 同步配置 表配置 (10) CDC 与恢复

启动配置

首次启动方式: 同步现有数据并持续同步

检查点与恢复

检查点间隔: 30 秒

恢复策略: 最近成功检查点

图 31CDC恢复

说明:

- 查看首次启动方式、Checkpoint设置和恢复策略。

## 2.3.5 发布与版本管理

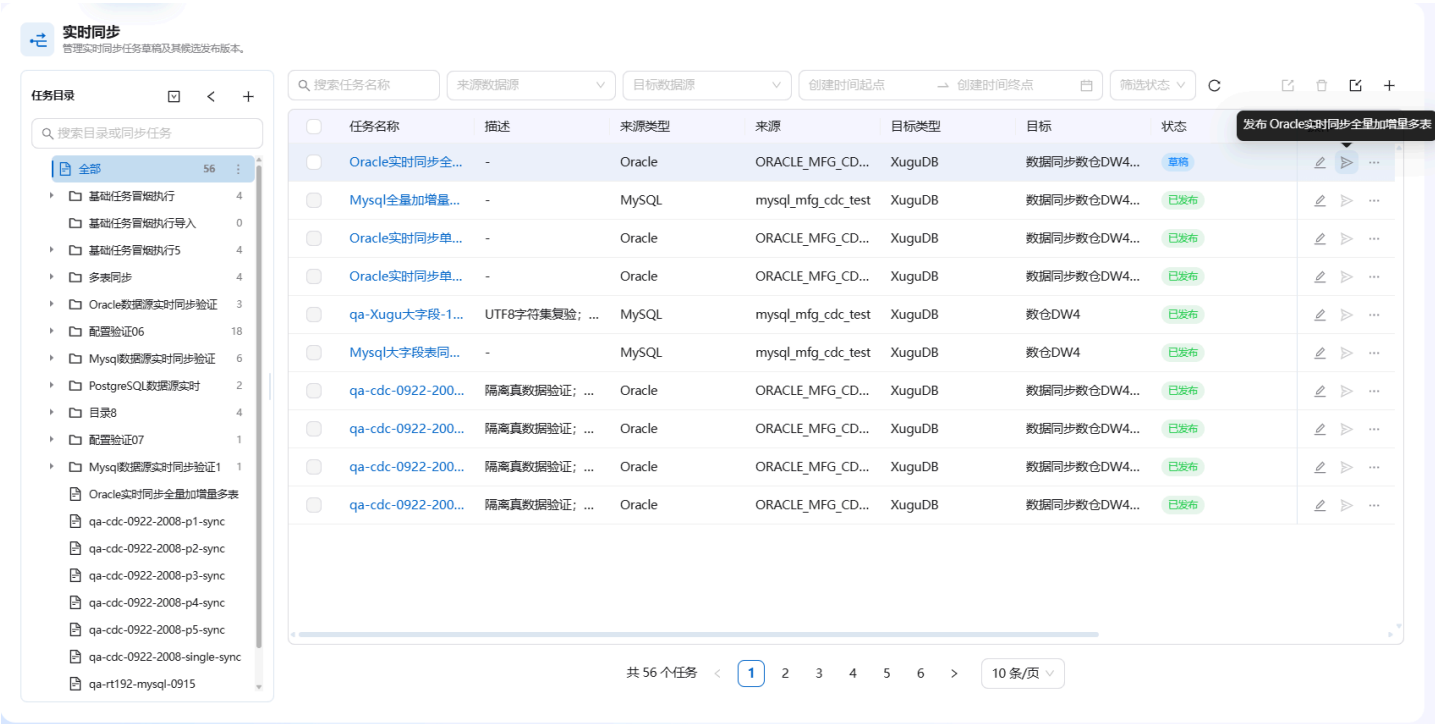


图 32发布任务

说明：

- 从列表提交发布，需审批时跟踪审批结果；发布处理完成后核对生产版本。
- 首次运行需到生产工作流启动，发布成功本身不代表已经开始同步。

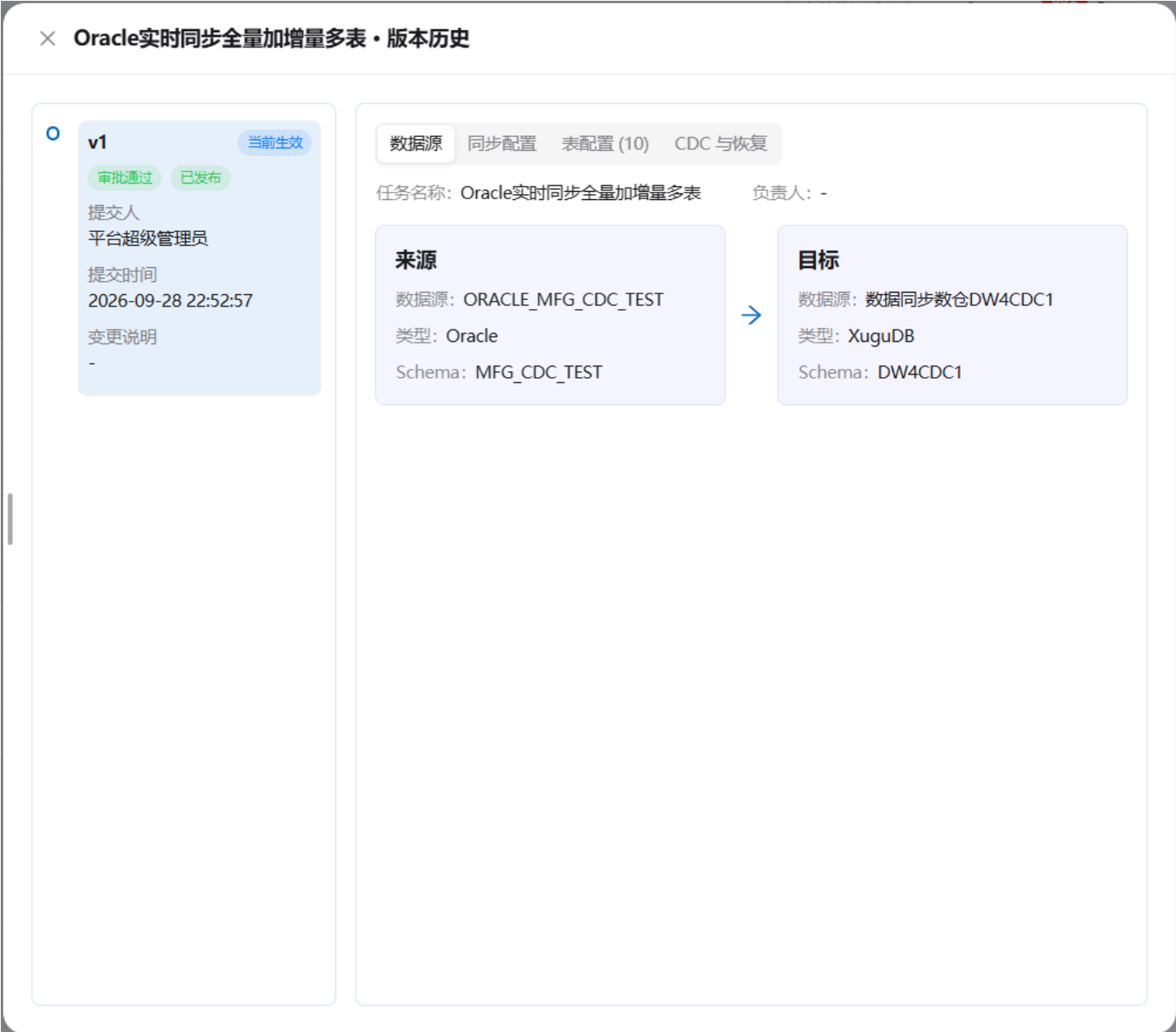


图 33版本历史

说明：

查看实时任务版本

- 进入版本历史，选择需要查看的版本，核对配置及发布时间信息。
- 运行问题定位时记录实际运行版本，避免使用新草稿解释旧实例。

启用历史实时版本

- 在版本历史选择允许启用的已发布版本，确认版本切换提示后提交。
- 核对生产版本和运行状态；涉及恢复时，确认旧版本与恢复点兼容，不把版本切换等同于业务数据回滚。

2.3.6 导入与导出

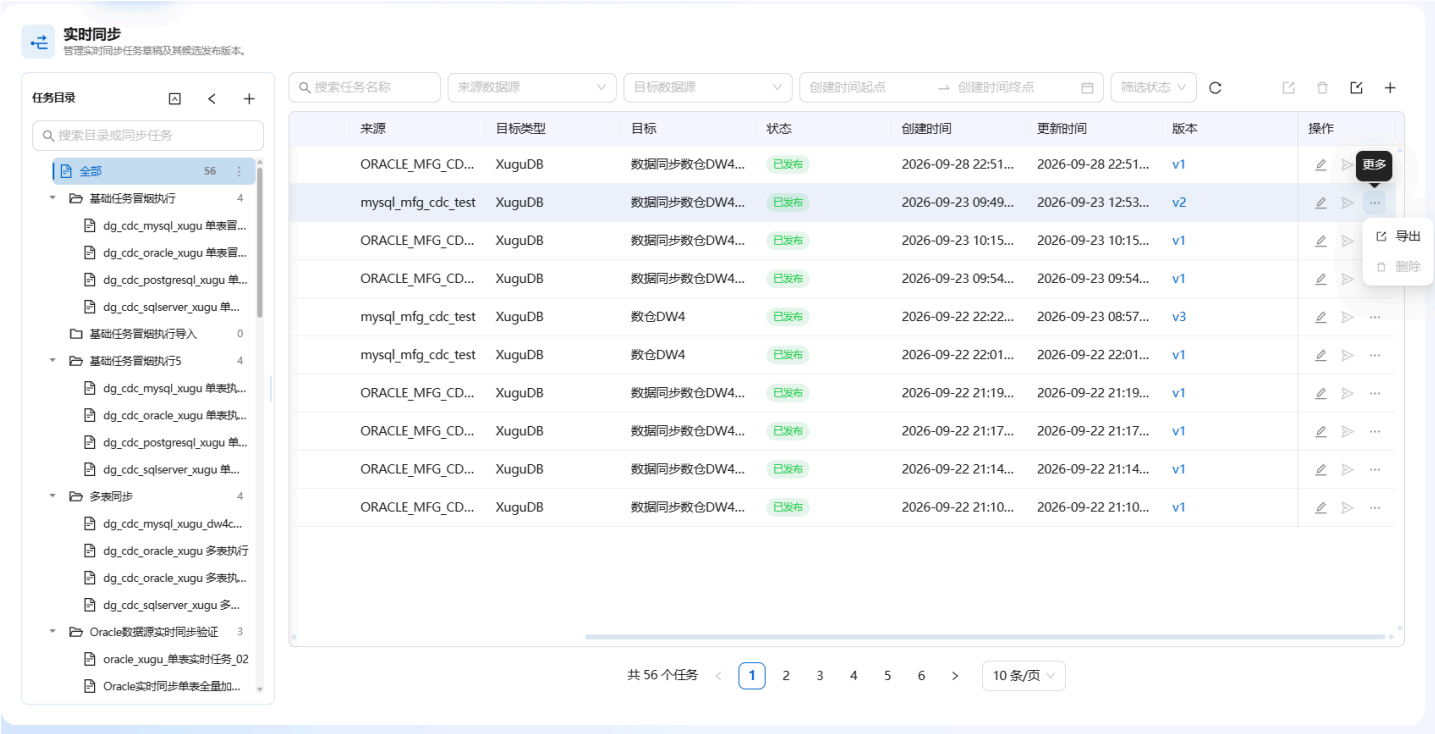


图 34导入导出

- 说明：
- 导入实时任务
- 选择配置文件，映射当前环境数据源，确认导入为独立任务。
  - 导入后重新检查CDC前置条件、表结构、唯一键与首次启动策略，不要直接套用其他环境的恢复位置。
- 导出实时任务
- 从行操作导出已保存配置，并妥善保管配置文件。
  - 配置导出不等于导出数据库日志或可跨环境使用的恢复状态。

2.3.7 删除实时任务

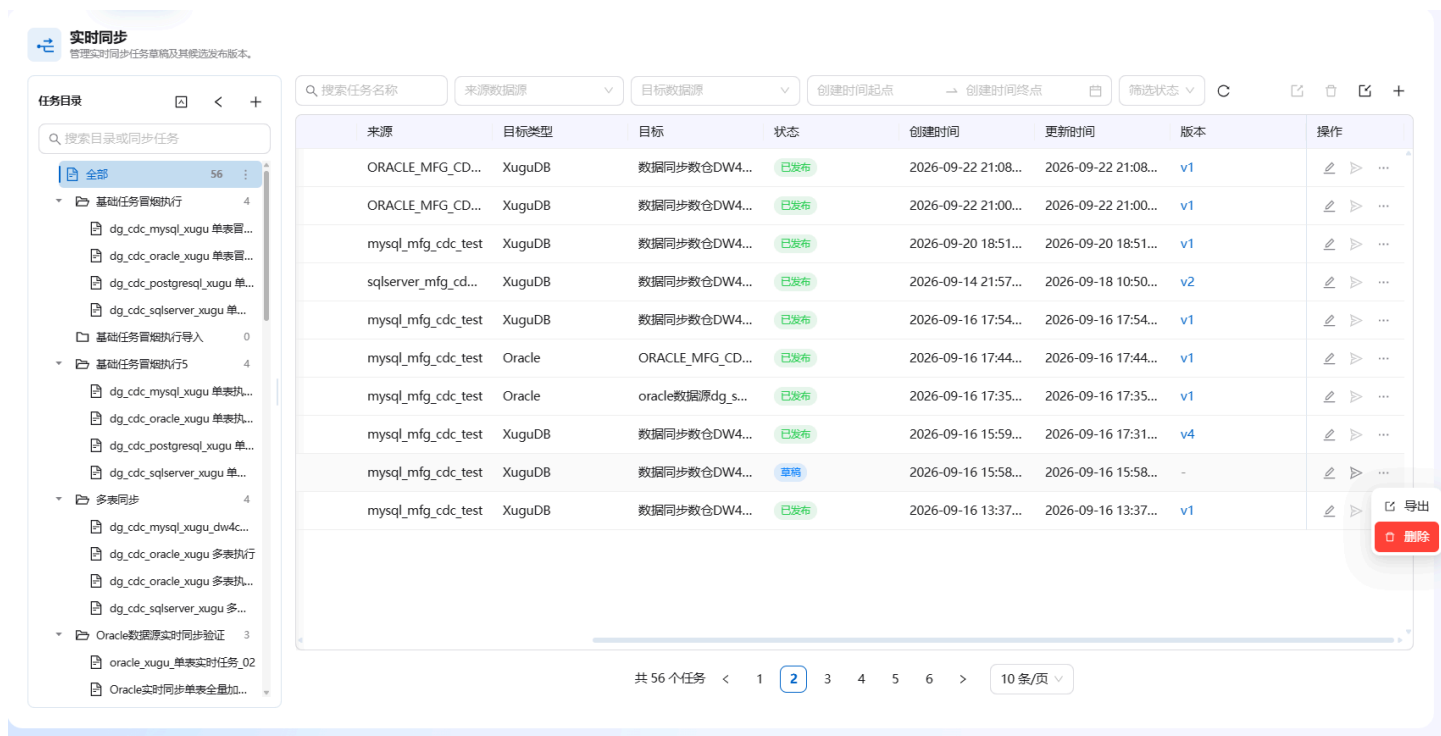


图 35删除任务

说明：

- 删除前在生产 workflow 停止运行，确认没有活动实例，再返回任务列表删除。
- 删除配置不会撤销已同步的数据；需要数据清理时，应另行确认目标范围并操作。

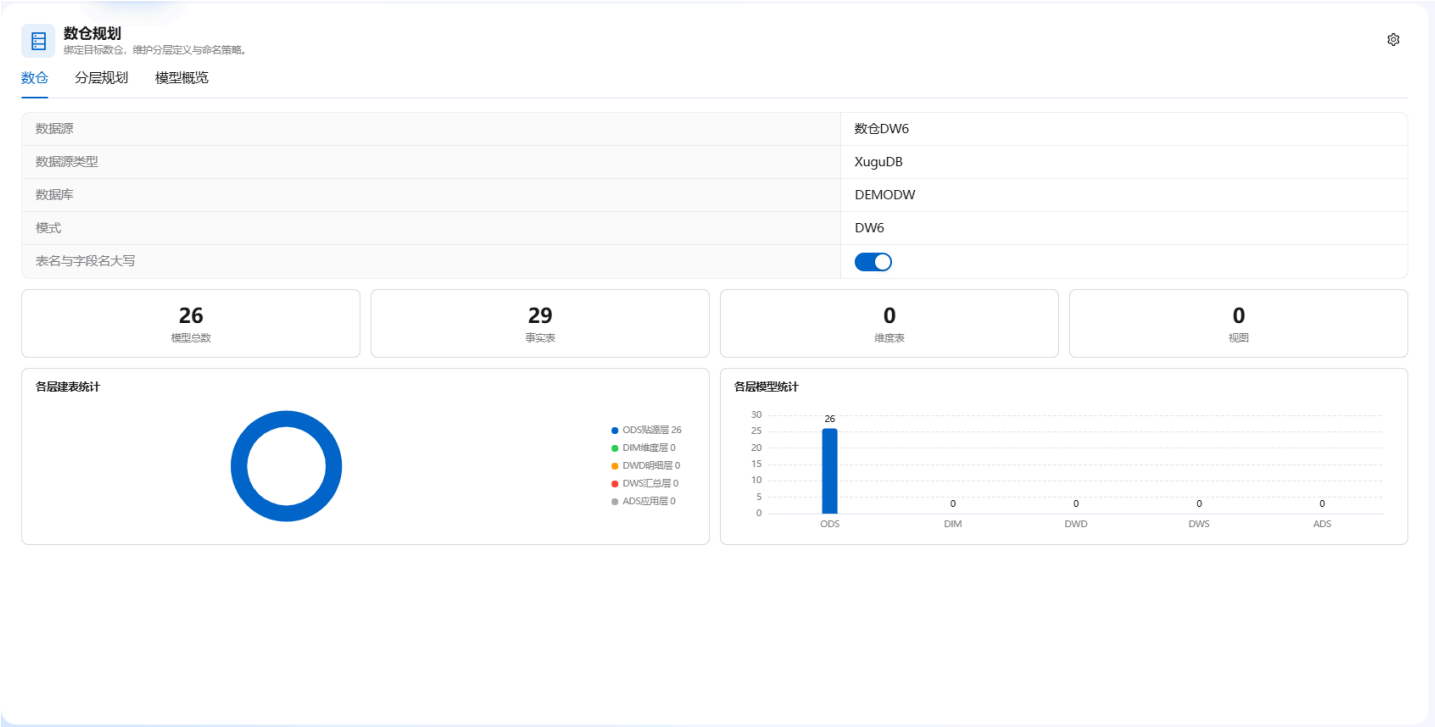
## 3 建模与开发

建模与开发将目标数仓、模型结构和执行逻辑连接起来。数仓规划确定承载位置与分层，模型定义字段和约束，数据开发通过工作流组织处理步骤。模型发布、物理建表、数据写入分别完成，不应将其中一个动作视为全部完成。

### 3.1 数仓规划

数仓规划维护绑定数仓与分层结构，并提供模型概览和字段类型映射策略，为数仓建模提供统一的组织方式。

#### 3.1.1 数仓信息



数仓规划

绑定目标数仓，维护分层定义与命名规范。

数仓

分层规划

模型概览

| 分层编码 | 分层名称 | 排序 | 前缀   | 模型数 | 类型   | 操作                        |
|------|------|----|------|-----|------|---------------------------|
| ODS  | 贴源层  | 1  | ods_ | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DIM  | 维度层  | 2  | dim_ | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DWD  | 明细层  | 3  | dwd_ | 2   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DWS  | 汇总层  | 4  | dws_ | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| ADS  | 应用层  | 5  | ads_ | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |

图 37查看分层规划

说明：

- 切换分层规划，查看分层编码、名称、模型数量及内置或自定义标识。
- ODS用于贴源数据，DIM组织维度，DWD组织明细，DWS组织汇总，ADS组织面向应用的结果；实际模型仍需在各层创建和发布。

数仓规划

绑定目标数仓，维护分层定义与命名规范。

数仓

分层规划

模型概览

| 分层编码 | 分层名称 | 模型数 | 类型   | 操作                        |
|------|------|-----|------|---------------------------|
| ODS  | 贴源层  | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DIM  | 维度层  | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DWD  | 明细层  | 2   | 内置分层 | <div>✎</div> <div>🗑</div> |
| DWS  | 汇总层  | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |
| ADS  | 应用层  | 0   | 内置分层 | <div>✎</div> <div>🗑</div> |

新增分层

\* 分层编码

\* 分层名称

\* 排序

前缀

取消

创建数仓分层

图 38新增分层

说明：

- 点击新增，填写分层编码和名称。编码使用3至16位大写字母或下划线，首尾不用下划线；名称不超过32字。

- 保存后在数仓建模中确认新增分层。分层名称用于业务识别，编码还影响相应模型前缀。

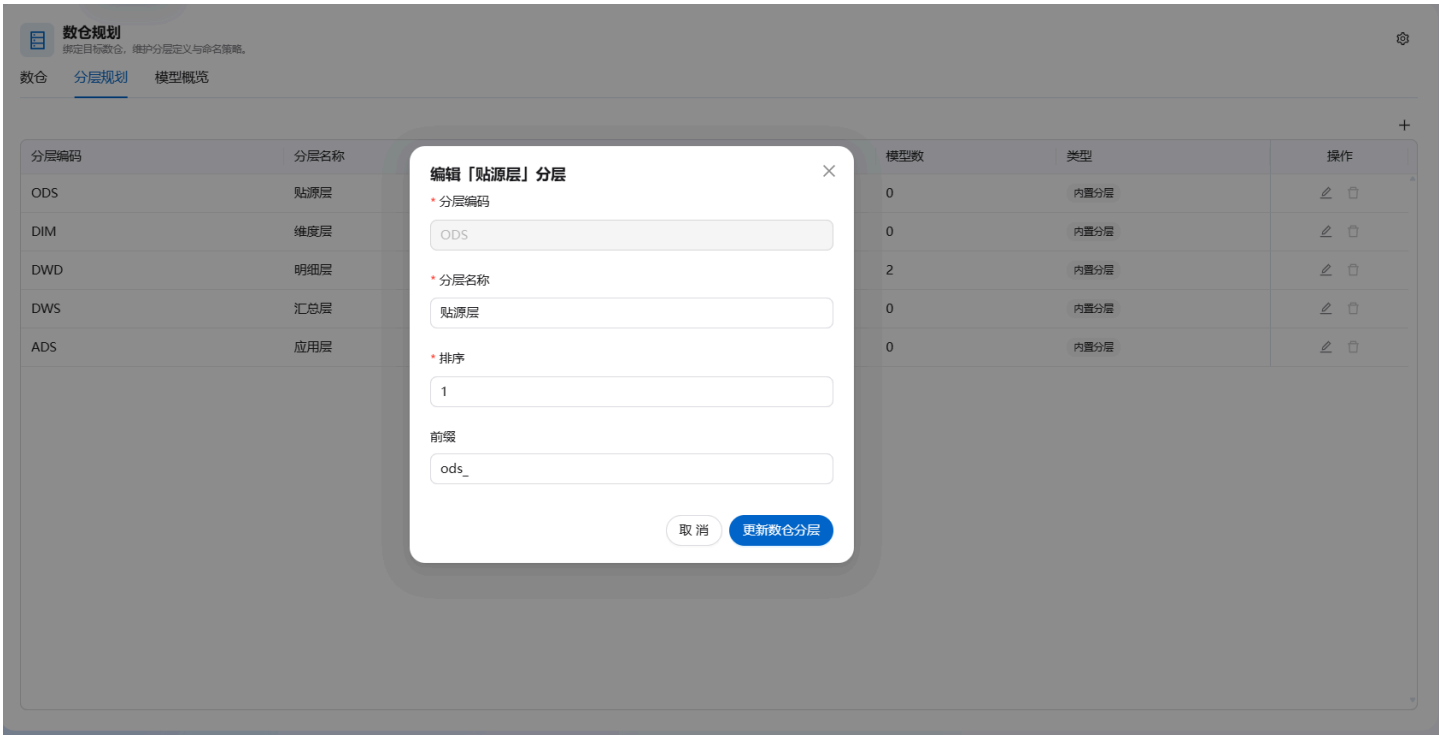


图 39编辑分层

说明：

- 选择需要编辑的分层，修改允许编辑的字段后保存。
- 内置分层或已有模型使用的分层，其编码不可随意调整；根据页面限制维护名称。

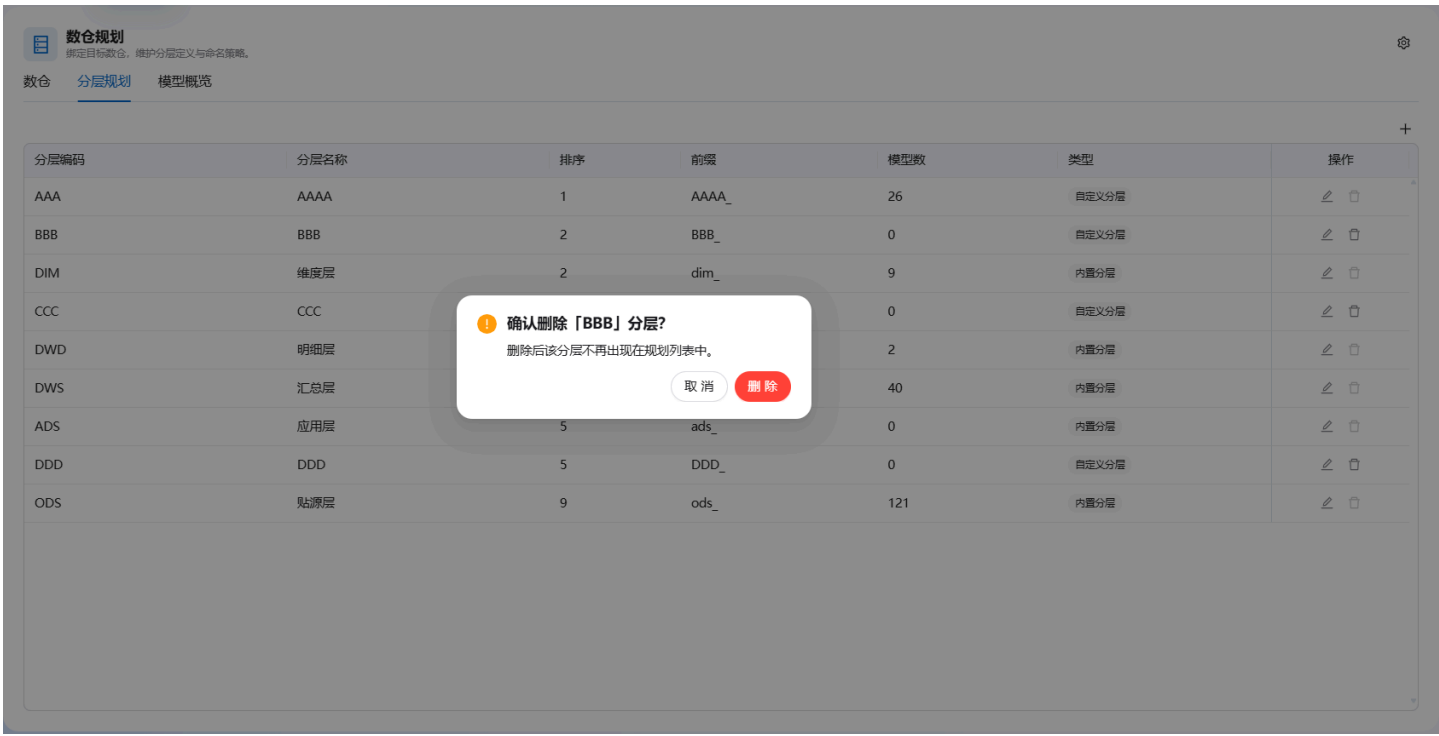


图 40删除分层

说明：

- 选择未被使用的自定义分层并确认删除。

- 内置分层或已经包含模型的分层不可直接删除，先核对模型归属，避免破坏分层组织。

### 3.1.3 模型概览



图 41模型概览

说明：

- 切换模型概览，查看当前规划下模型的整体分布，使用缩放等画布操作调整阅读范围。
- 该页用于查看；模型属性与关系维护请进入元模型或数仓建模相应编辑入口。

### 3.1.4 字段类型映射

×

字段类型映射策略

新增

MySQL

Q 搜索源字段/虚谷字段类型

| 源字段类型           | 虚谷字段类型                       | 来源   | 说明     | 操作 |
|-----------------|------------------------------|------|--------|----|
| BIGINT          | BIGINT                       | 平台内置 | 平台内置映射 |    |
| BIGINT UNSIGNED | DECIMAL(20,0)                | 平台内置 | 平台内置映射 |    |
| BINARY          | BINARY({length})             | 平台内置 | 平台内置映射 |    |
| BIT             | BOOLEAN                      | 平台内置 | 平台内置映射 |    |
| BLOB            | BLOB                         | 平台内置 | 平台内置映射 |    |
| BOOL            | BOOLEAN                      | 平台内置 | 平台内置映射 |    |
| BOOLEAN         | BOOLEAN                      | 平台内置 | 平台内置映射 |    |
| CHAR            | CHAR({length})               | 平台内置 | 平台内置映射 |    |
| DATE            | DATE                         | 平台内置 | 平台内置映射 |    |
| DATETIME        | TIMESTAMP                    | 平台内置 | 平台内置映射 |    |
| DECIMAL         | DECIMAL({precision},{scale}) | 平台内置 | 平台内置映射 |    |
| DOUBLE          | DOUBLE                       | 平台内置 | 平台内置映射 |    |
| DOUBLE UNSIGNED | DOUBLE                       | 平台内置 | 平台内置映射 |    |
| ENUM            | CLOB                         | 平台内置 | 平台内置映射 |    |
| FLOAT           | FLOAT                        | 平台内置 | 平台内置映射 |    |

图 42查看字段类型映射

- 说明：
- 打开字段类型映射策略，选择来源数据库类型并搜索字段类型。列表展示源类型、虚谷目标类型、来源及说明。
  - 内置规则提供基础映射，租户规则用于补充或覆盖。此处的模型类型映射与快捷同步连接器的自动建表策略应分别配置和检查。

**新增类型映射**

\* 源数据库

MySQL

\* 源字段类型

例如 JSON 或 VARCHAR

\* 虚谷字段类型

例如 TEXT 或 VARCHAR({length})

说明

0 / 100

取消 保存

图 43新增映射规则

说明：

- 点击新增，填写来源数据库、源字段类型、虚谷字段类型和说明。
- 需要继承长度的目标类型可按界面支持的占位格式填写，例如VARCHAR({length})。保存后确认规则来源与优先使用结果。
- 选择可编辑的租户规则，修改映射类型或说明后保存。
- 内置规则不直接修改；调整规则后，重新检查受影响模型的字段类型，避免只改规则而忽略已有模型。
- 在租户规则行选择删除并确认。
- 删除覆盖规则后应重新检查可用的内置映射，尤其关注字符长度、数值精度等约束。

## 3.2 元模型

以模型列表管理结构，以编辑页维护属性，以概览画布查看和维护模型关系。元模型用于维护可复用的结构定义，不使用数仓分层筛选。

### 3.2.1 元模型列表

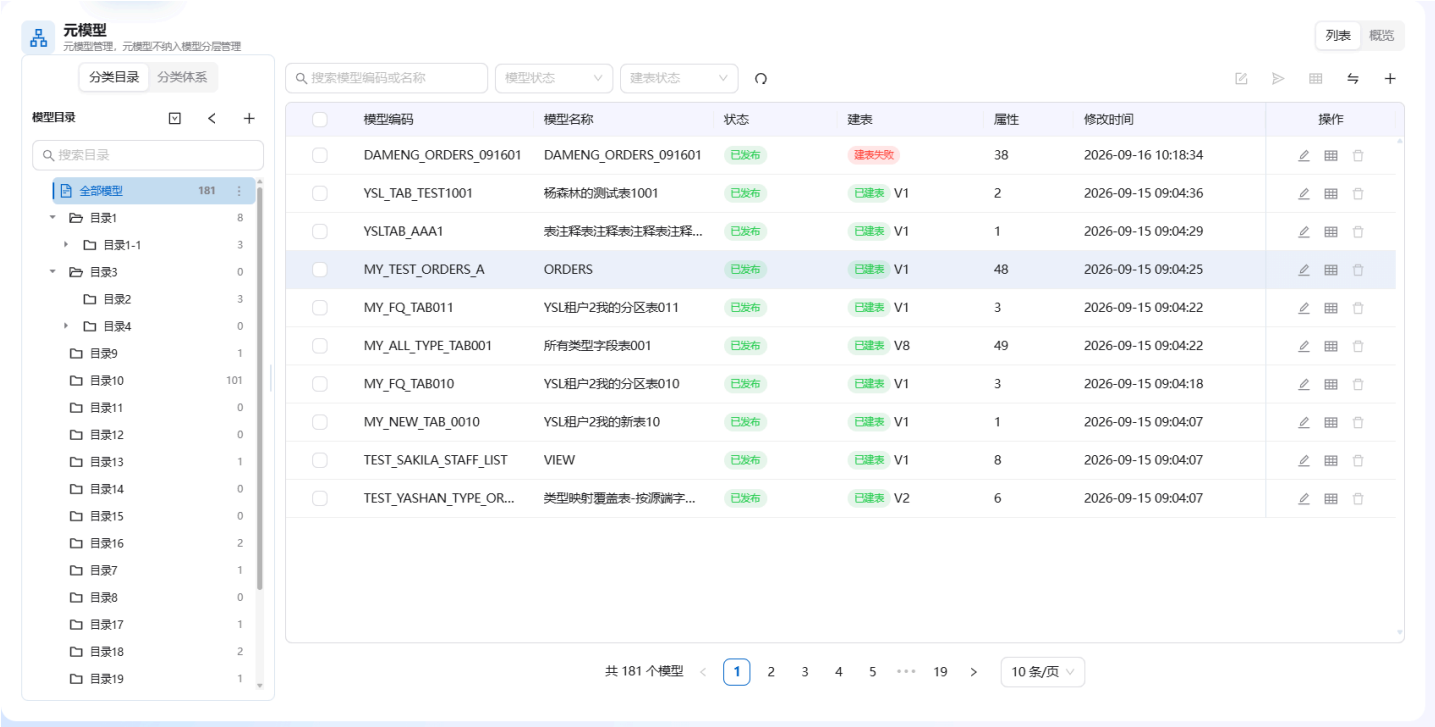


图 44元模型列表

- 说明：
- 左侧通过分类目录或分类体系限定范围，右侧按模型代码或名称查找；分类体系用于按业务分类查找，不等于存储目录。
  - 列表展示模型代码、名称、状态与版本、建表状态、属性数量、修改时间。点击版本查看历史，行操作进入修改、建表或删除。

3.2.2 模型目录管理

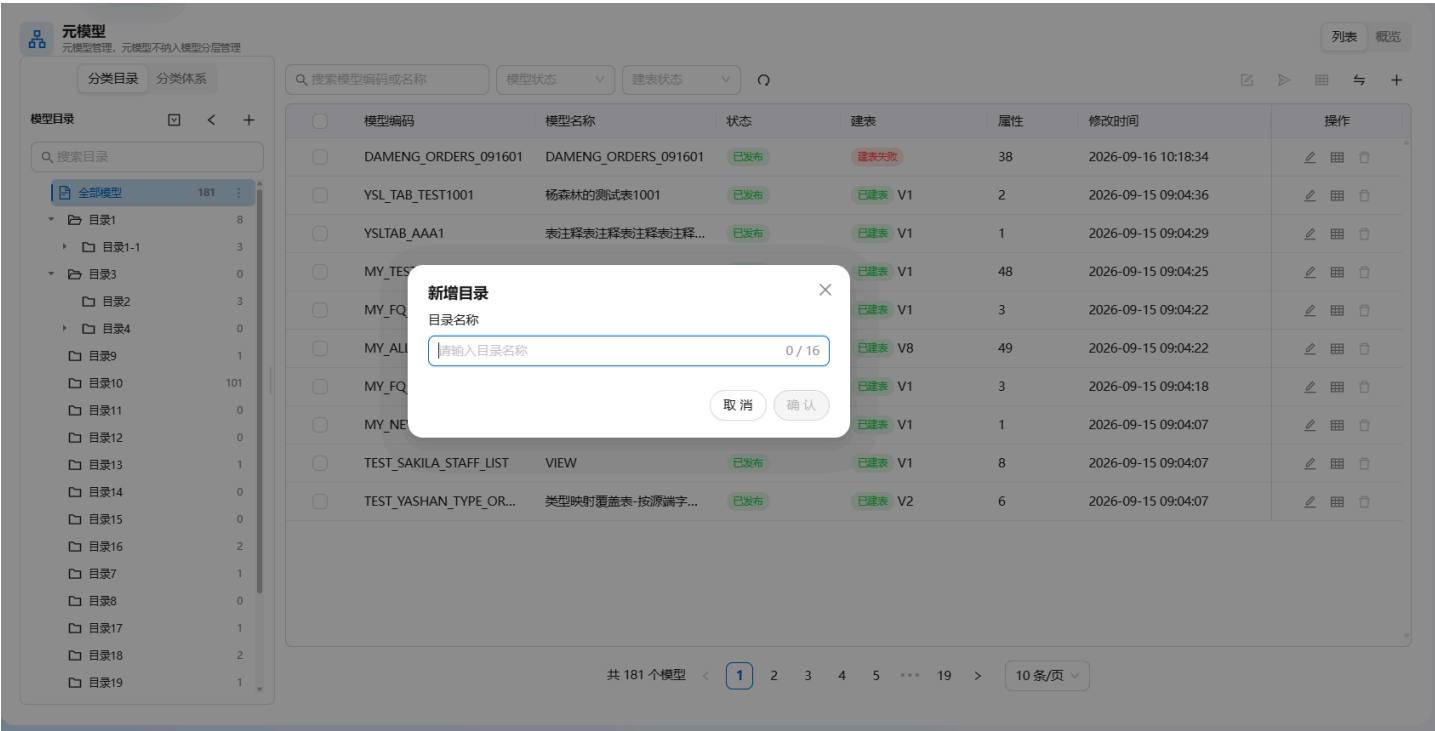


图 45新建模型目录

说明：

- 在分类目录中选择父目录，使用新增入口填写目录名称并保存。
- 目录用于组织模型，例如“物流履约”；它不会自动创建数据库Schema。

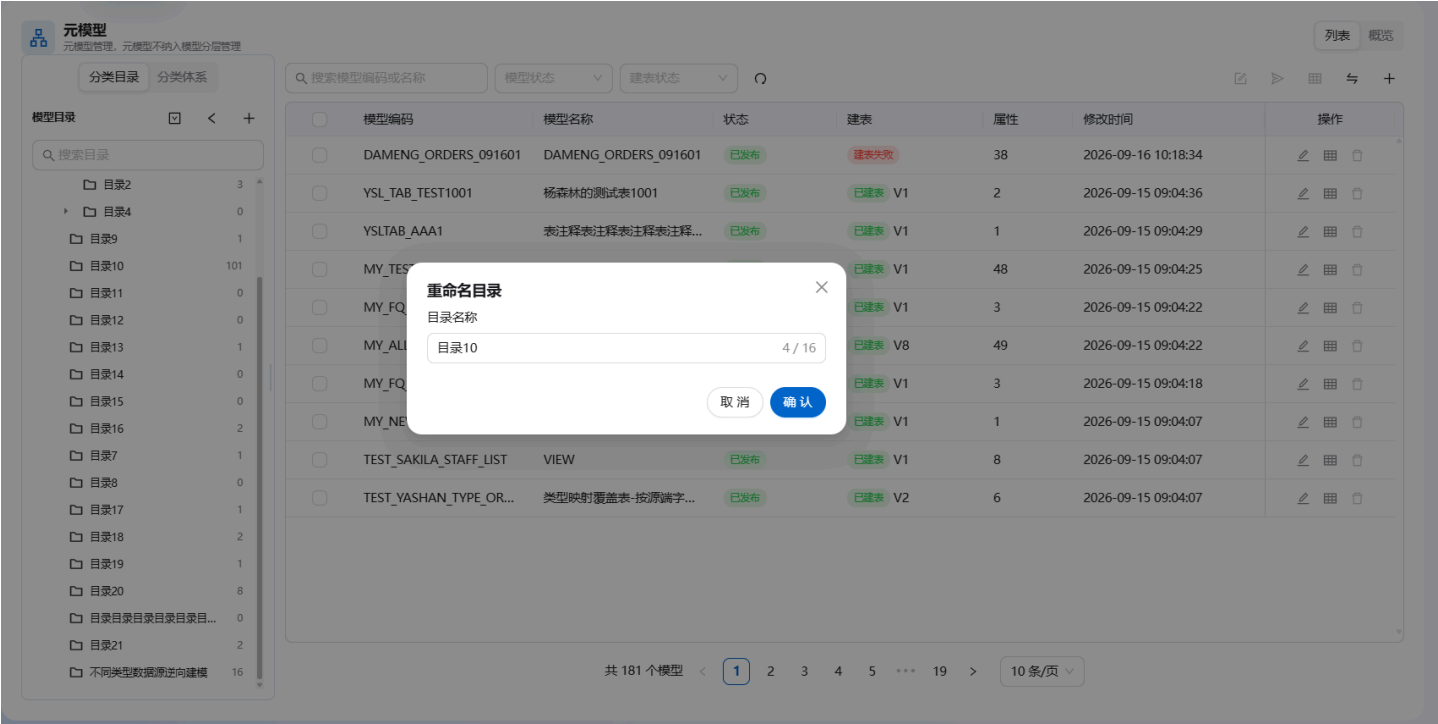


图 46重命名模型目录

说明：

- 在目录操作菜单选择重命名，填写名称并确认。
- 重命名只调整模型组织目录，不修改模型编码或物理表名。

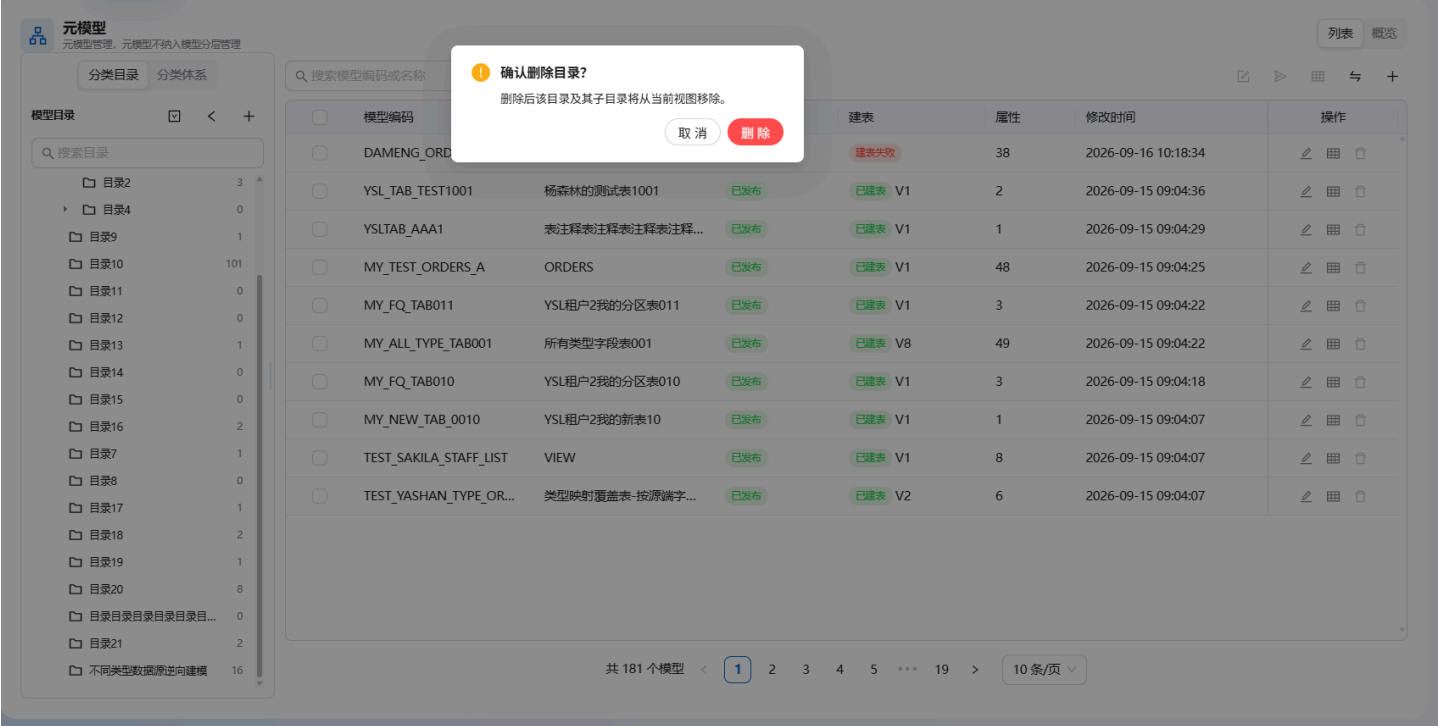


图 47删除模型目录

说明：

- 选择需要删除的目录，核对其中模型及下级目录后按提示确认。
- 存在关联内容时先处理内容归属；不要将删除目录作为删除数据库表的方法。

### 3.2.3 新建元模型

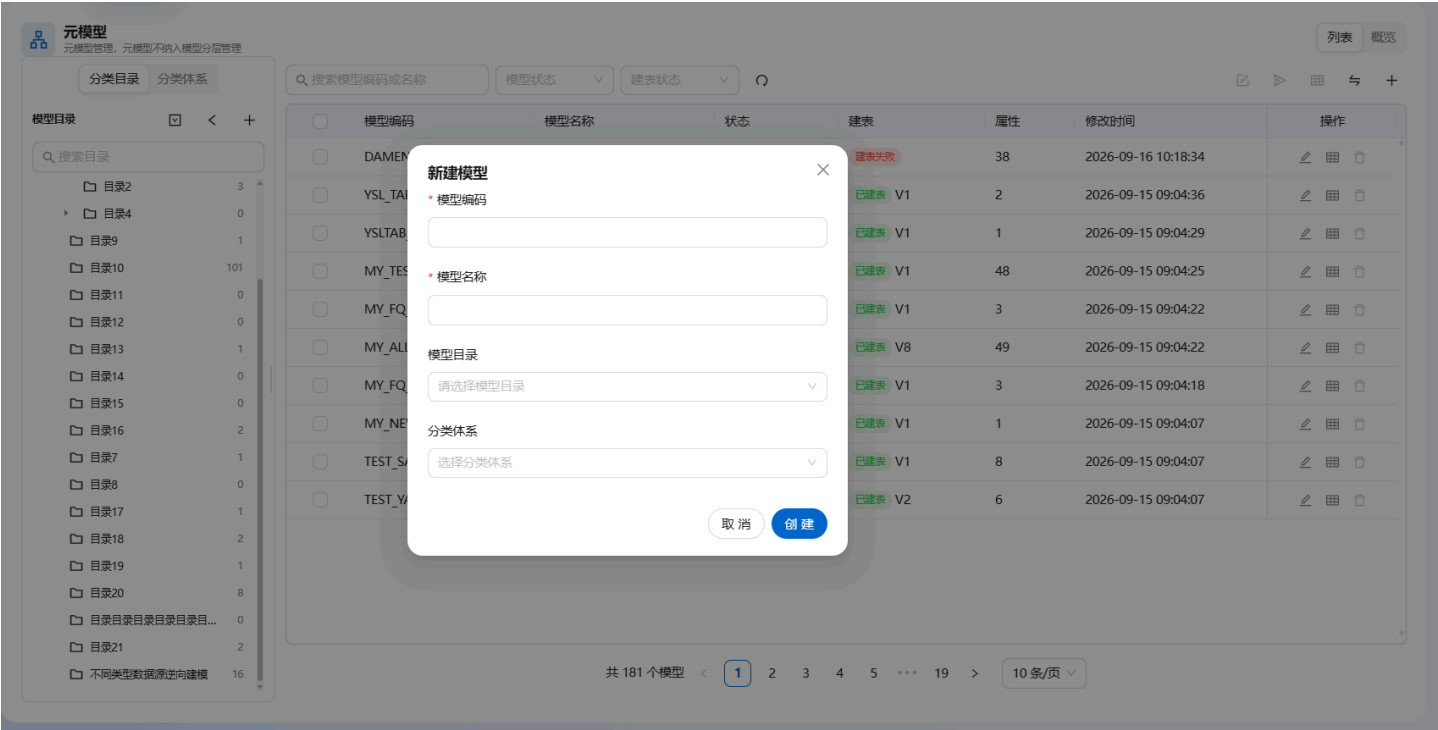


图 48新建元模型

说明：

- 点击新建，填写模型编码、模型名称并选择模型目录。编码以字母开头，使用字母、数字和下划线，长度3至128位；名称不超过128字。
- 使用能够表达对象含义的编码，例如logistics\_order。
- 确认创建后进入模型编辑页，继续补充属性；新建模型不等于数据库已经建表。

### 3.2.4 编辑元模型

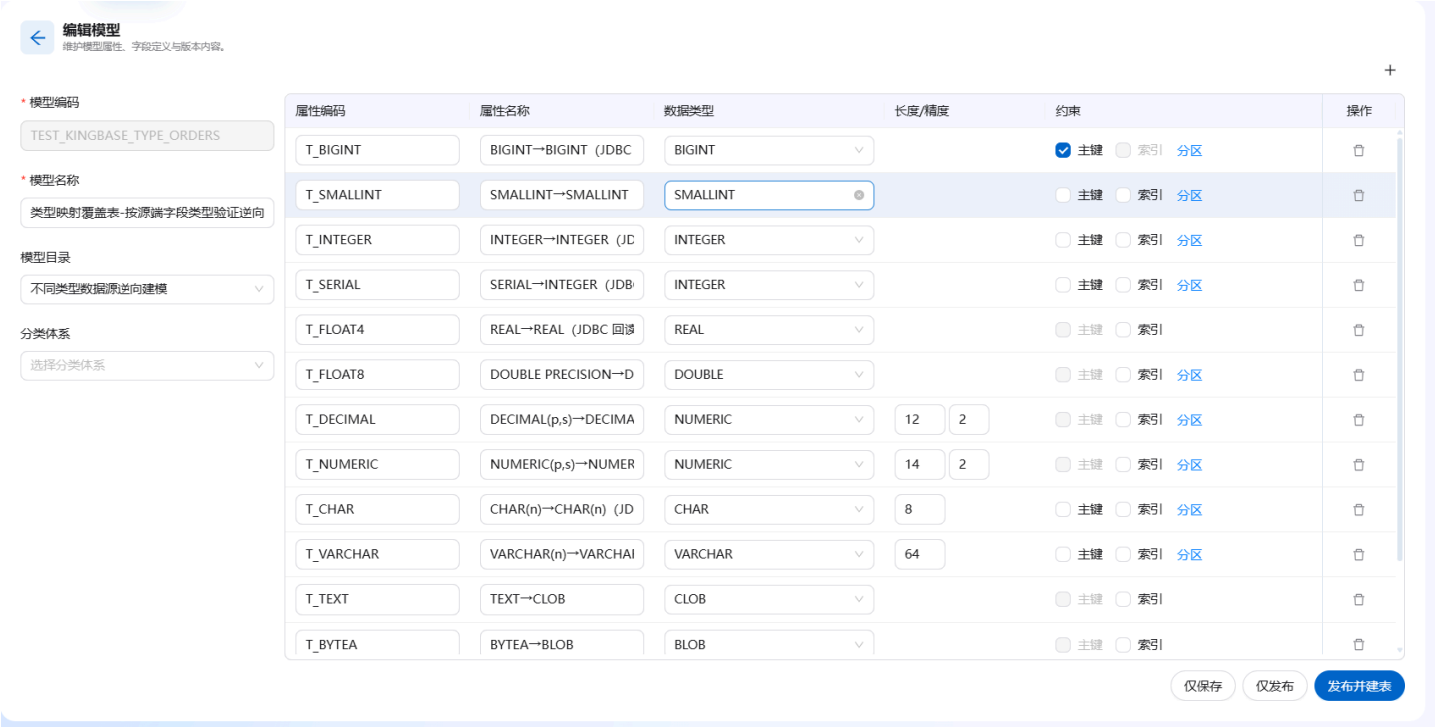


图 49编辑元模型

说明：

编辑模型基本信息

- 从模型列表选择修改，左侧维护模型代码、名称、目录及分类体系归属。
- 右侧维护属性，底部选择仅保存、仅发布或发布并建表。发布中的对象应等待处理结束再编辑。

新增模型属性

- 在属性表点击添加属性，填写属性代码、名称、数据类型及所需长度或精度。
- 例如订单ID使用整数类型，运单号使用字符类型；属性代码不能重复，要求长度或精度的类型应填写有效正数。模型至少保留一个属性。

删除模型属性

- 在目标属性行点击删除并确认，核对是否影响主键、关联关系或下游使用。
- 删除后保存模型；若随后执行建表变更，可能涉及物理字段删除和数据损失，应先备份并确认影响。

仅保存模型

- 完成编辑后点击仅保存，保留当前模型内容。
- 该动作不发布新生产版本，也不创建或修改物理表。保存后在列表确认状态。
- 在属性行调整类型与约束，按需要指定主键。单字段主键可根据类型选择不自动生成、数据库自增、平台UUID或平台雪花ID。
- 复合主键不使用单字段生成方式；非主键字段的唯一与索引选项按界面互斥规则设置。修改后检查关联关系与现有物理表的兼容性。

3.2.5 模型概览与关系管理



图 50查看概览

说明：

- 切换概览视图，在画布查看模型及属性，使用缩放和定位调整查看范围。
- 模型间连线表达结构关联，不代表同步任务已经创建或数据已完成关联计算。

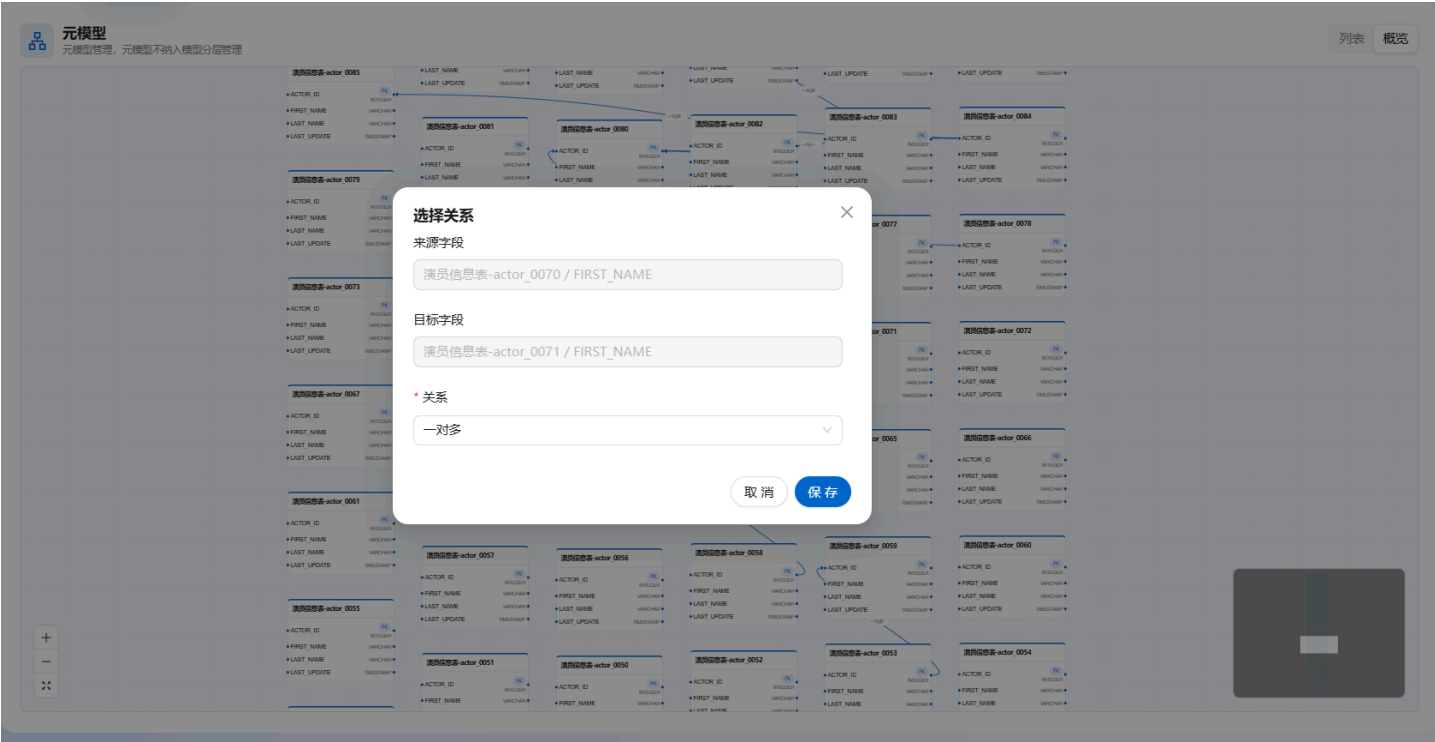


图 51新增模型关系

说明：

- 在概览画布从来源属性连接至目标属性，选择一对一、一对多、多对一或多对多关系并保存。
- 例如承运商主数据与运单可表达一对多关系；请选择能够关联的实际属性，并确认方向。
- 点击概览画布中的已有关系连线，在关系弹窗中调整关系基数并保存。
- 核对来源模型、目标模型和关联属性，确认关系方向仍符合业务含义。
- 点击已有关系连线，在关系弹窗中选择删除并确认。
- 删除关系只取消模型间的关联定义，不删除两侧模型或业务记录。

3.2.6 发布与版本管理

← 编辑模型

维护模型属性、字段定义与版本内容。

模型编码

MY\_ALL\_TYPE\_TAB001

模型名称

所有类型字段表001

模型目录

目录21

分类体系

人力资源域/员工

| 属性编码 | 属性名称 | 数据类型           | 长度/精度 | 约束                                                                         | 操作 |
|------|------|----------------|-------|----------------------------------------------------------------------------|----|
| A01  | A01  | INTEGER        |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A02  | A02  | BINARY         |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A03  | A03  | BINARY_DOUBLE  |       | <input checked="" type="checkbox"/> 主键 <input type="checkbox"/> 索引         | 🗑  |
| A04  | A04  | BINARY_FLOAT   |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引                    | 🗑  |
| A05  | A05  | BINARY_INTEGER |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A06  | A06  | BIGINT         |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A07  | A07  | BIT            |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A08  | A08  | BLOB           |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引                    | 🗑  |
| A09  | A09  | BOOLEAN        |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A10  | A10  | BIGINT         |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A11  | A11  | BIGINT         |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |
| A12  | A12  | CHAR           | 32    | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> | 🗑  |

仅保存

仅发布

发布并建表

图 52模型发布

说明：

仅发布模型

- 点击仅发布，确认发布信息并提交；需要审批时跟踪审批结果。
- 发布成功形成模型版本，但不代表物理表已经创建。可通过建表状态区分模型发布与物理落地。

发布并建表

- 选择发布并建表，阅读预检结果和结构变化提示，确认目标后提交。
- 存在已有表时重点检查字段删除、类型调整等风险，必要时先备份。需要审批的变更待审批及后续执行完成后，分别核对版本和建表状态。

× 所有类型字段表001·版本历史

V8 已发布

2026-09-12 14:03:39

发布人：平台超级管理员

V7 已发布

2026-09-12 14:03:23

发布人：平台超级管理员

V6 已发布

2026-09-12 14:03:16

发布人：平台超级管理员

V5 已发布

2026-09-12 14:02:08

发布人：平台超级管理员

V4 已发布

2026-09-12 14:01:50

发布人：平台超级管理员

V3 已发布

2026-09-12 14:01:37

发布人：平台超级管理员

V2 已发布

2026-09-11 14:13:28

发布人：平台超级管理员

V1 已发布

2026-09-11 14:12:37

发布人：平台超级管理员

V8 版本内容

2026-09-12 14:03:39

基准版本

>

对比

| 属性编码 | 属性名称 | 数据类型          | 长度 | 主键 | 唯一 | 变化 |
|------|------|---------------|----|----|----|----|
| A01  | A01  | INTEGER       | -  | -  | -  | -  |
| A02  | A02  | BINARY        | -  | -  | -  | -  |
| A03  | A03  | BINARY_DO...  | -  | -  | -  | -  |
| A04  | A04  | BINARY_FL...  | -  | -  | -  | -  |
| A05  | A05  | BINARY_INT... | -  | -  | -  | -  |
| A06  | A06  | BIGINT        | -  | -  | -  | -  |
| A07  | A07  | BIT           | -  | -  | -  | -  |
| A08  | A08  | BLOB          | -  | -  | -  | -  |
| A09  | A09  | BOOLEAN       | -  | -  | -  | -  |
| A10  | A10  | BIGINT        | -  | -  | -  | -  |
| A11  | A11  | BIGINT        | -  | -  | -  | -  |
| A12  | A12  | CHAR          | 32 | -  | -  | -  |
| A13  | A13  | CHARACTER     | 32 | -  | -  | -  |
| A14  | A14  | CLOB          | -  | -  | -  | -  |
| A15  | A15  | DATE          | -  | -  | -  | -  |

图 53 模型版本

说明：

- 点击版本链接打开历史抽屉，选择版本查看其属性快照。
- 通过版本号与时间判断所查看的结构，不把当前草稿属性当作历史版本属性。
- 在版本历史使用对比版本，选择基准版本，查看新增、修改或删除的属性差异。
- 对比用于理解变更，不自动将模型还原为旧版本。

### 3.2.7 执行建表

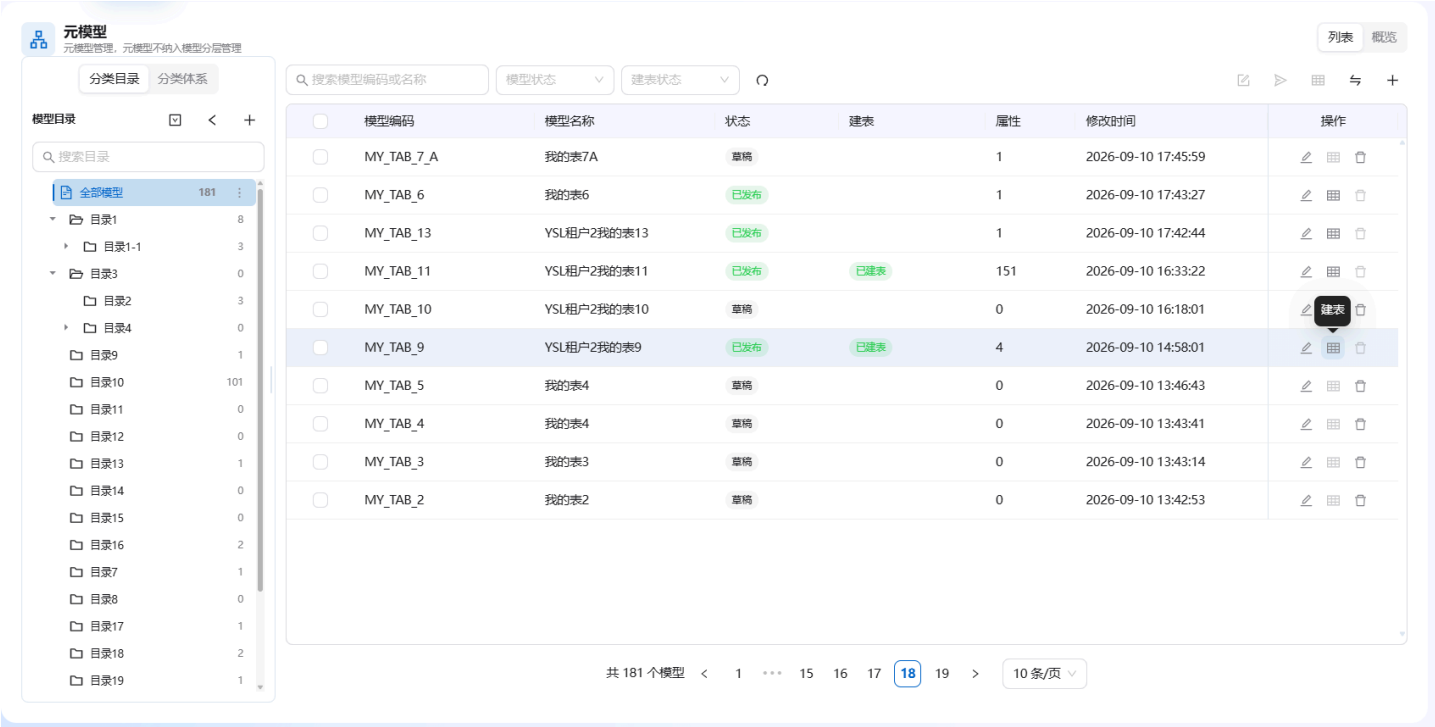


图 54建表执行

- 说明：
- 在已发布模型行使用建表入口，依据已发布版本检查物理表预检结果，确认后执行。
  - 建表处理中查看状态；失败时查看提示并修正原因后重试。该动作处理表结构，不会同步来源数据。

3.2.8 逆向生成模型

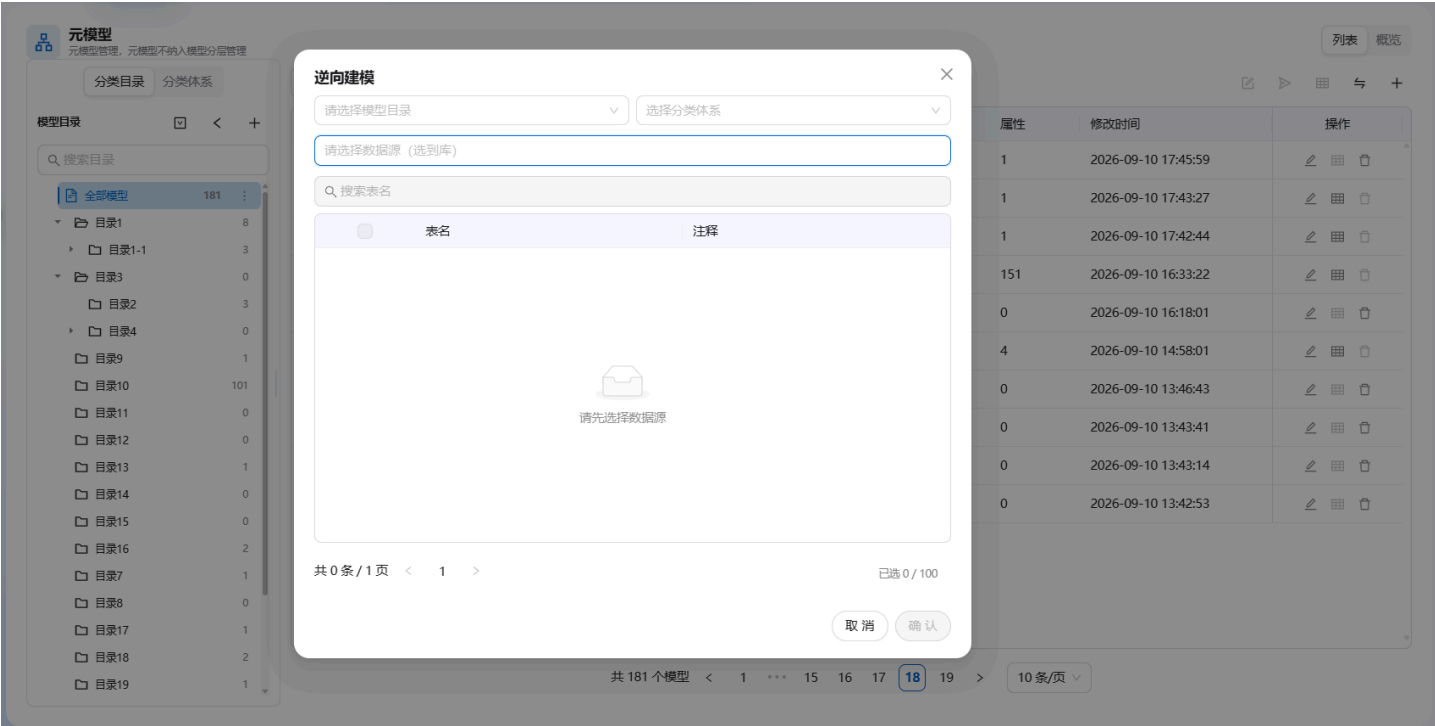


图 55逆向生成模型

说明：

- 打开逆向入口，选择数据源与数据库，查找并勾选真实表；可跨页选择。确认当前目录后生成。
- 查看成功与失败数量及失败明细。部分成功时先到列表确认已生成模型，仅重新处理失败对象，避免重复生成成功项。

### 3.2.9 删除元模型

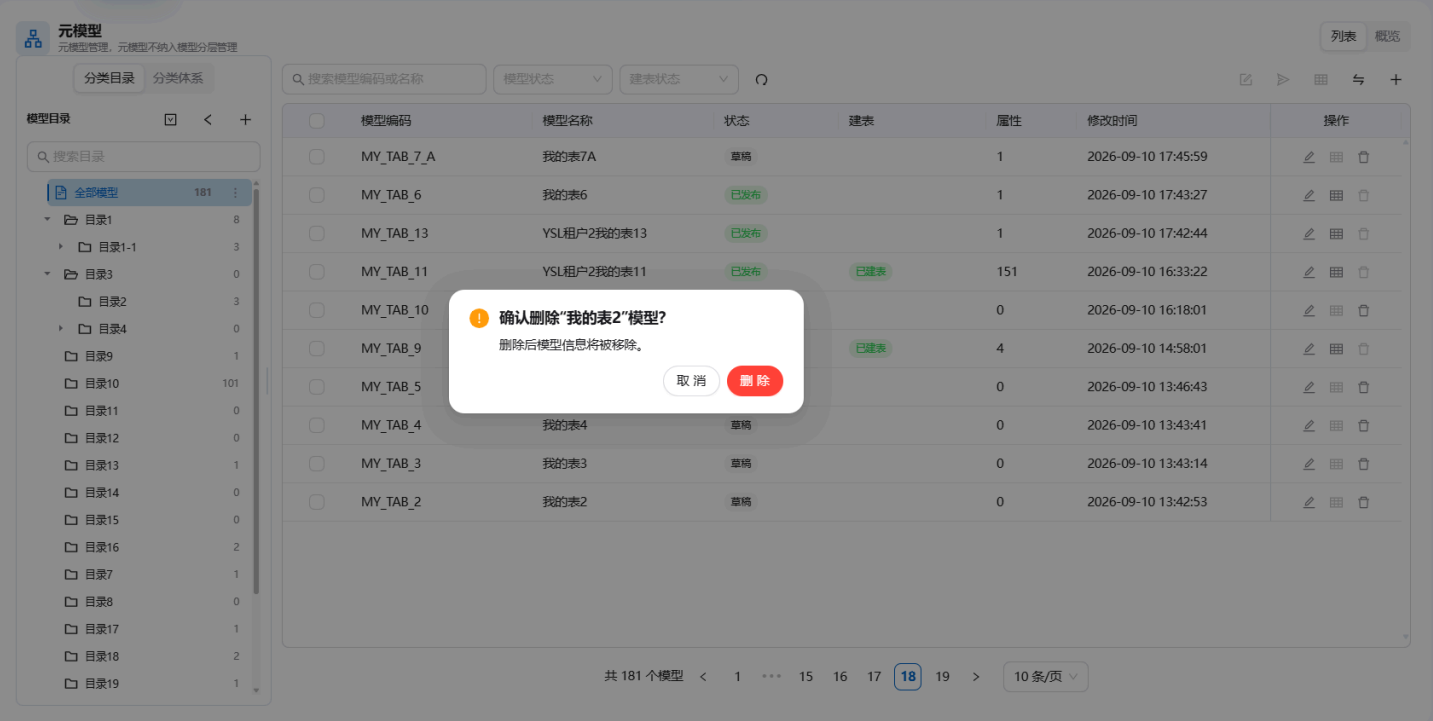


图 56删除模型

说明：

- 从列表选择删除，检查提示后确认。已有版本的模型不能直接删除。
- 模型定义管理与物理数据库管理不同；不要假定删除模型会清空数据库，也不要通过删除物理表绕过模型状态限制。

## 3.3 数仓建模

以模型列表管理结构，以编辑页维护属性，以概览画布查看和维护模型关系。数仓模型按已配置的数仓分层组织。

### 3.3.1 数仓模型列表

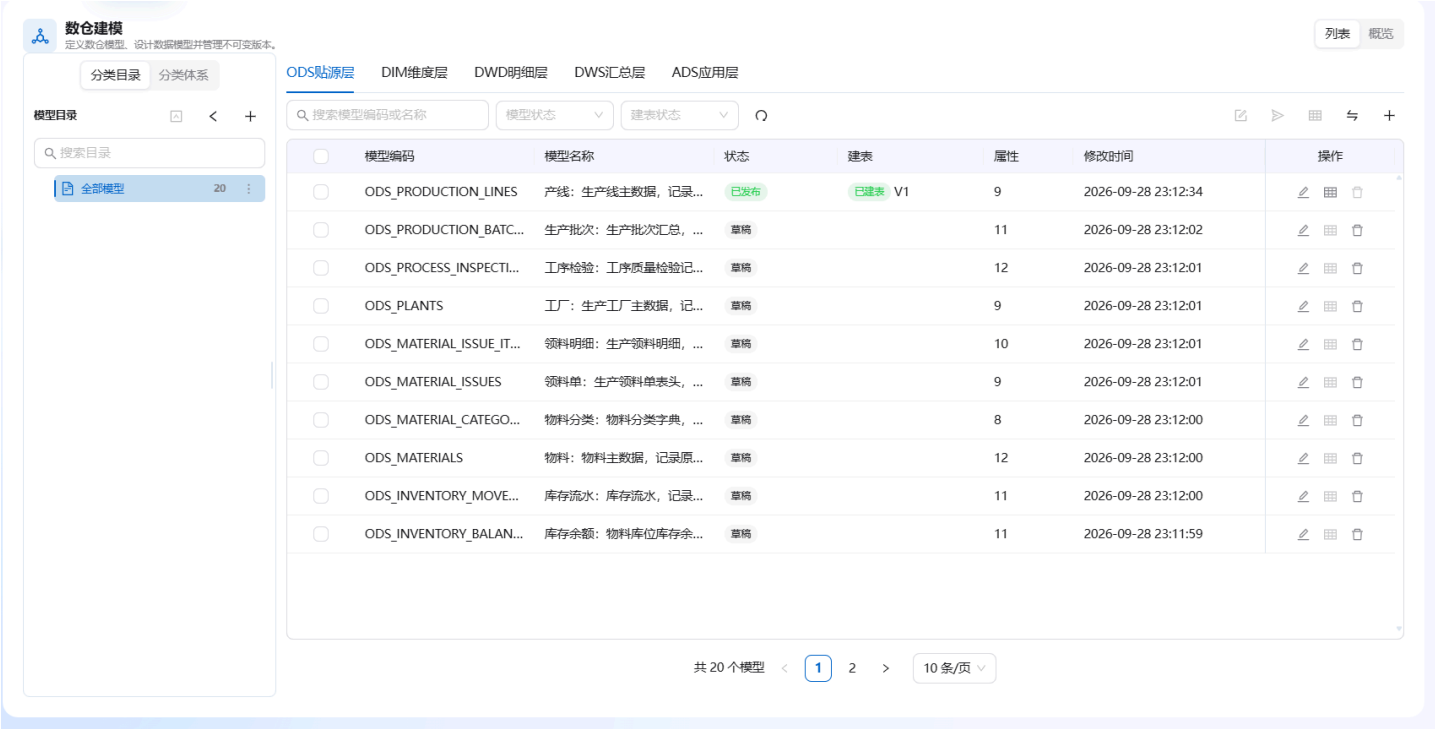


图 57查看数仓建模列表

- 说明：
- 左侧通过分类目录或分类体系限定范围，右侧按模型代码或名称查找；先选择ODS、DIM等分层。
  - 列表展示模型代码、名称、状态与版本、建表状态、属性数量、修改时间。点击版本查看历史，行操作进入修改、建表或删除。

3.3.2 模型目录管理

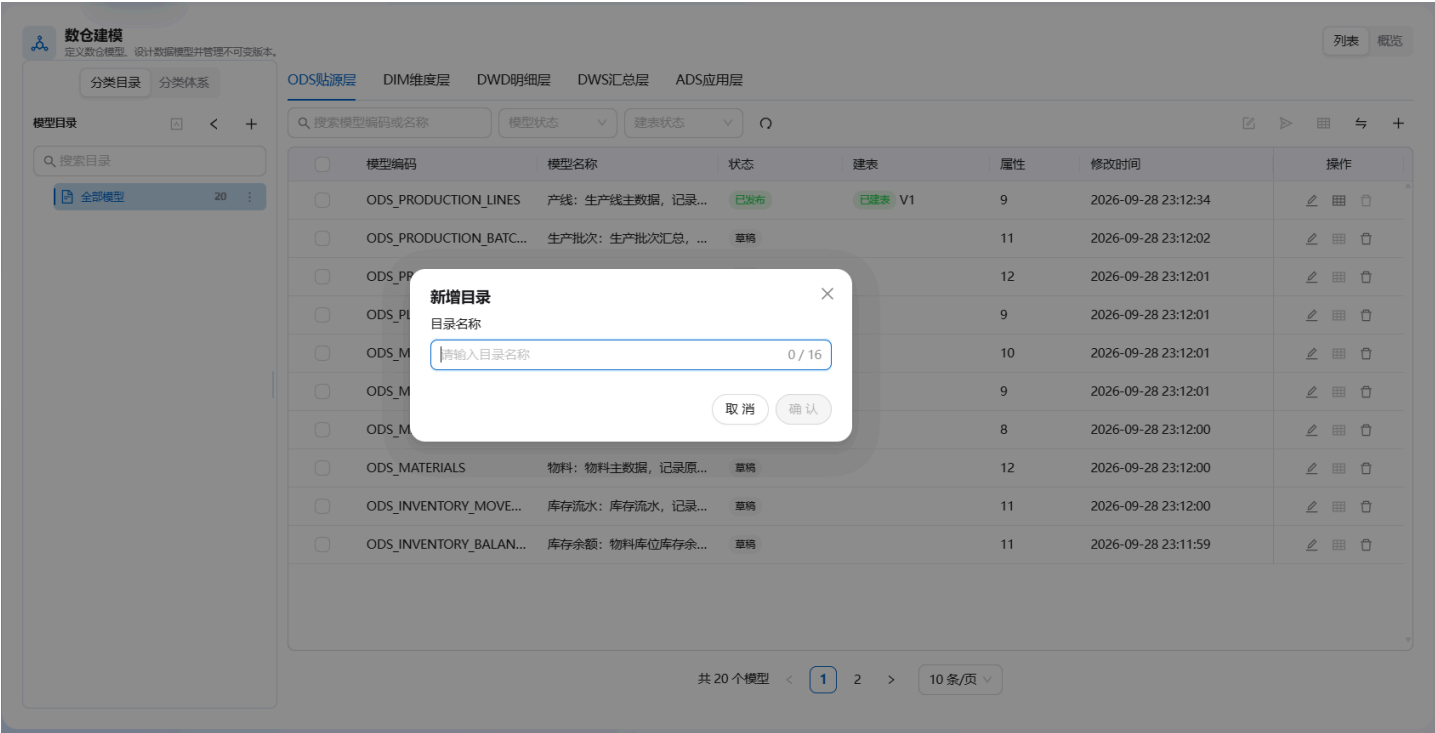


图 58新建目录

- 说明：

- 在分类目录中选择父目录，使用新增入口填写目录名称并保存。
- 目录用于组织模型，例如“物流履约”；它不会自动创建数据库Schema。

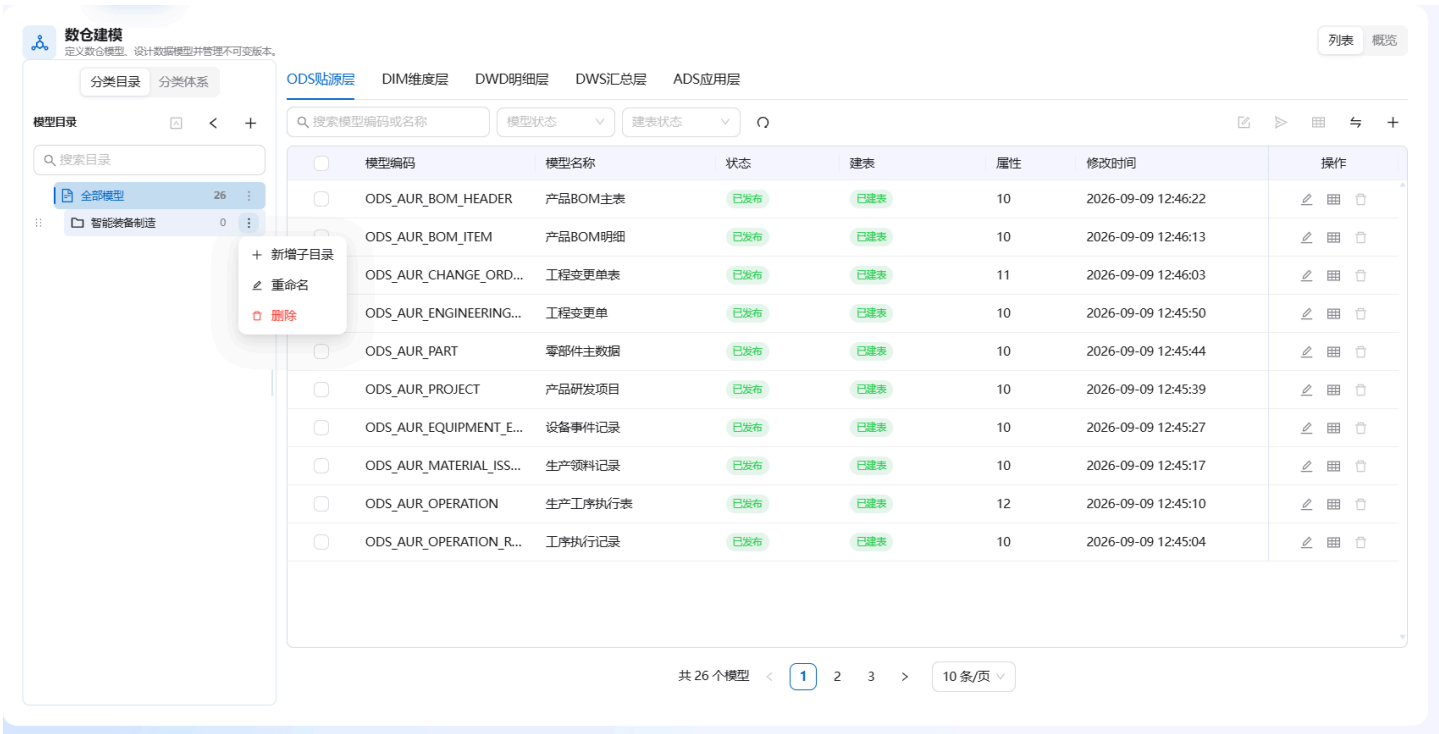


图 59重命名目录

说明：

- 在目录操作菜单选择重命名，填写名称并确认。
- 重命名只调整模型组织目录，不修改模型编码或物理表名。

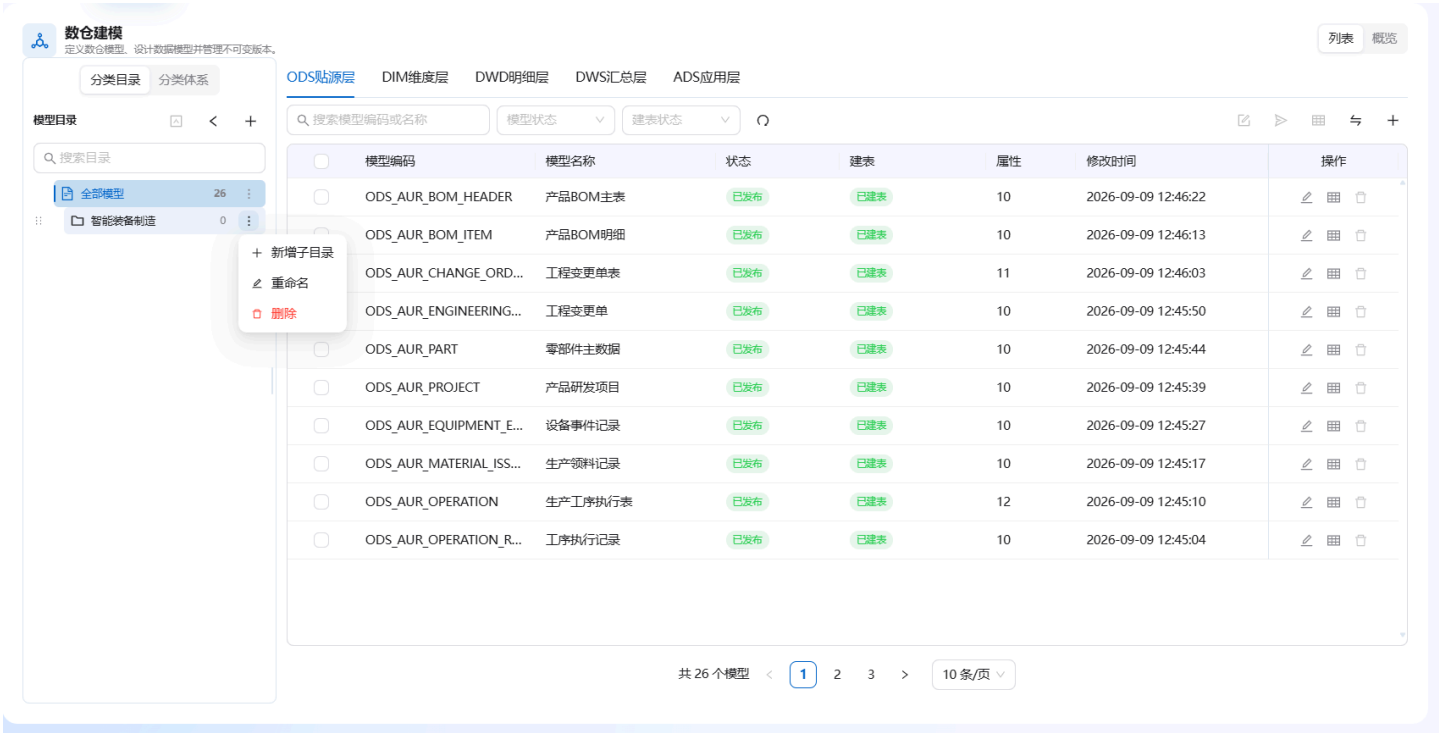


图 60删除目录

说明：

- 选择需要删除的目录，核对其中模型及下级目录后按提示确认。
- 存在关联内容时先处理内容归属；不要将删除目录作为删除数据库表的方法。

### 3.3.3 新建数仓模型

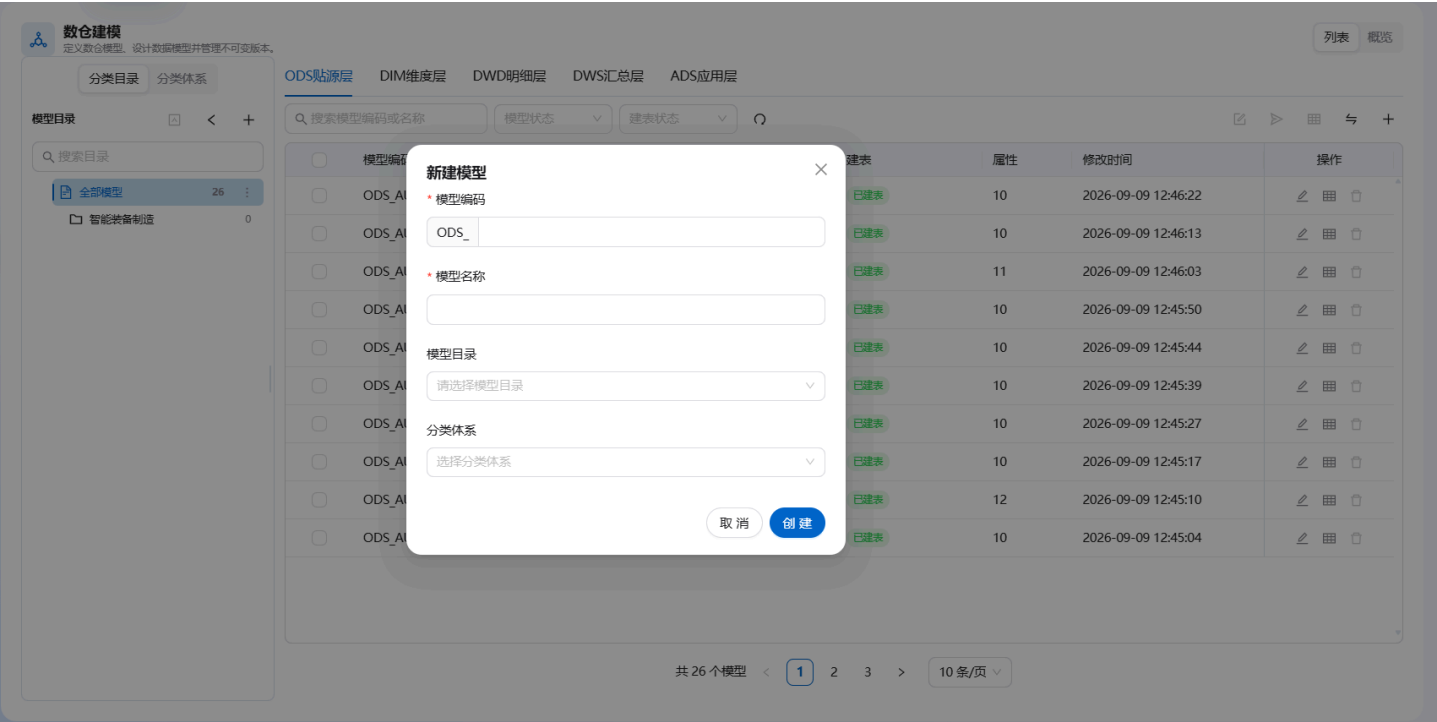


图 61新建模型

- 说明：
- 点击新建，填写模型编码、模型名称并选择模型目录。编码以字母开头，使用字母、数字和下划线，长度3至128位；名称不超过128字。
  - 模型编码前显示当前分层前缀，只填写主体部分，例如logistics\_order。
  - 确认创建后进入模型编辑页，继续补充属性；新建模型不等于数据库已经建表。

### 3.3.4 编辑数仓模型

编辑模型

维护模型属性、字段定义与版本内容。

模型编码

ODS\_ PROCESS\_INSPECTIONS

模型名称

工序检验：工序质量检验记录，记录报工

模型目录

请选择模型目录

分类体系

选择分类体系

属性编码

属性名称

数据类型

长度/精度

约束

操作

ID

记录主键，由数据生成脚

DECIMAL

19

0

☒ 主键 ☐ 索引 [分区](#)

CODE

记录唯一业务编号；INC-I

NVARCHAR2

160

☐ 主键 ☐ 索引 [分区](#)

STATUS

记录业务状态；基准值AC

NVARCHAR2

160

☐ 主键 ☐ 索引 [分区](#)

REVISION

记录修订次数；初始为0，

DECIMAL

19

0

☐ 主键 ☐ 索引 [分区](#)

CREATED\_AT

记录创建时间，字段不包

TIMESTAMP

☐ 主键 ☐ 索引 [分区](#)

UPDATED\_AT

记录最后修改时间，字段

TIMESTAMP

☐ 主键 ☐ 索引 [分区](#)

REPORT\_ID

生产报工主键，外键关联

DECIMAL

19

0

☐ 主键 ☐ 索引 [分区](#)

INSPECTED\_QTY

本次检验产品数量，单位：

DECIMAL

19

0

☐ 主键 ☐ 索引 [分区](#)

GOOD\_QTY

合格产品数量，单位：件

DECIMAL

19

0

☐ 主键 ☐ 索引 [分区](#)

REJECT\_QTY

不合格产品数量，单位：个

DECIMAL

19

0

☐ 主键 ☐ 索引 [分区](#)

DIMENSION\_MM

关键尺寸实测值，单位：mm

DECIMAL

18

2

☐ 主键 ☐ 索引 [分区](#)

RESULT

工序检验结论：PASS=合格

NVARCHAR2

160

☐ 主键 ☐ 索引 [分区](#)

仅保存

仅发布

发布并建表

说明：

- 从模型列表选择修改，左侧维护模型代码、名称、目录及分类体系归属。
- 右侧维护属性，底部选择仅保存、仅发布或发布并建表。发布中的对象应等待处理结束再编辑。

- 在属性表点击添加属性，填写属性代码、名称、数据类型及所需长度或精度。
- 例如订单ID使用整数类型，运单号使用字符类型；属性代码不能重复，要求长度或精度的类型应填写有效正数。模型至少保留一个属性。

- 在目标属性行点击删除并确认，核对是否影响主键、关联关系或下游使用。
- 删除后保存模型；若随后执行建表变更，可能涉及物理字段删除和数据损失，应先备份并确认影响。

- 完成编辑后点击仅保存，保留当前模型内容。
- 该动作不发布新生产版本，也不创建或修改物理表。保存后在列表确认状态。
- 在属性行调整类型与约束，按需要指定主键。单字段主键可根据类型选择不自动生成、数据库自增、平台UUID或平台雪花ID。
- 复合主键不使用单字段生成方式；非主键字段的唯一与索引选项按界面互斥规则设置。修改后检查关联关系与现有物理表的兼容性。

3.3.5 模型概览与关系管理

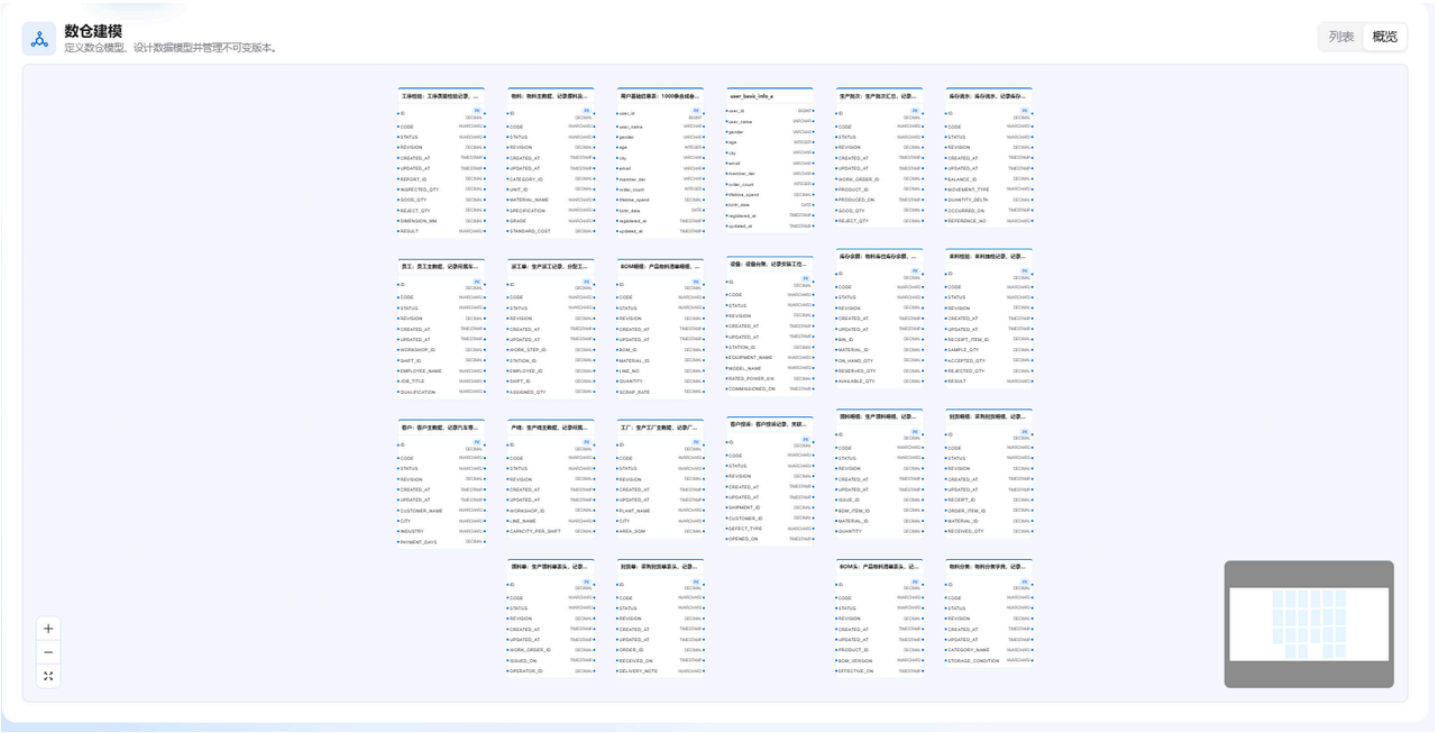


图 63查看模型概览

说明：

- 切换概览视图，在画布查看模型及属性，使用缩放和定位调整查看范围。
- 模型间连线表达结构关联，不代表同步任务已经创建或数据已完成关联计算。
- 在概览画布从来源属性连接至目标属性，选择一对一、一对多、多对一或多对多关系并保存。
- 例如承运商主数据与运单可表达一对多关系；请选择能够关联的实际属性，并确认方向。
- 点击概览画布中的已有关系连线，在关系弹窗中调整关系基数并保存。
- 核对来源模型、目标模型和关联属性，确认关系方向仍符合业务含义。
- 点击已有关系连线，在关系弹窗中选择删除并确认。
- 删除关系只取消模型间的关联定义，不删除两侧模型或业务记录。

3.3.6 发布与版本管理

编辑模型

维护模型属性、字段定义与版本内容。

模型编码

ODS\_ PRODUCTION\_LINES

模型名称

产线：生产线主数据，记录所属车间及单

模型目录

请选择模型目录

分类体系

选择分类体系

| 属性编码               | 属性名称           | 数据类型      | 长度/精度 | 约束                                                                                    | 操作 |
|--------------------|----------------|-----------|-------|---------------------------------------------------------------------------------------|----|
| ID                 | 记录主键，由数据生成脚    | DECIMAL   | 190   | <input checked="" type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a> |    |
| CODE               | 记录唯一业务编号；INC-I | NVARCHAR2 | 160   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| STATUS             | 记录业务状态；基准值AC   | NVARCHAR2 | 160   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| REVISION           | 记录修订次数；初始为0。   | DECIMAL   | 190   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| CREATED_AT         | 记录创建时间，字段不包    | TIMESTAMP |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| UPDATED_AT         | 记录最后修改时间，字段    | TIMESTAMP |       | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| WORKSHOP_ID        | 车间主键，外键关联work  | DECIMAL   | 190   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| LINE_NAME          | 生产线名称          | NVARCHAR2 | 160   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |
| CAPACITY_PER_SHIFT | 每班标准产能，单位：件/   | DECIMAL   | 190   | <input type="checkbox"/> 主键 <input type="checkbox"/> 索引 <a href="#">分区</a>            |    |

仅保存

仅发布

发布并建表

图 64模型发布

说明：

仅发布模型

- 点击仅发布，确认发布信息并提交；需要审批时跟踪审批结果。
- 发布成功形成模型版本，但不代表物理表已经创建。可通过建表状态区分模型发布与物理落地。

发布并建表

- 选择发布并建表，阅读预检结果和结构变化提示，确认目标后提交。
- 存在已有表时重点检查字段删除、类型调整等风险，必要时先备份。需要审批的变更待审批及后续执行完成后，分别核对版本和建表状态。



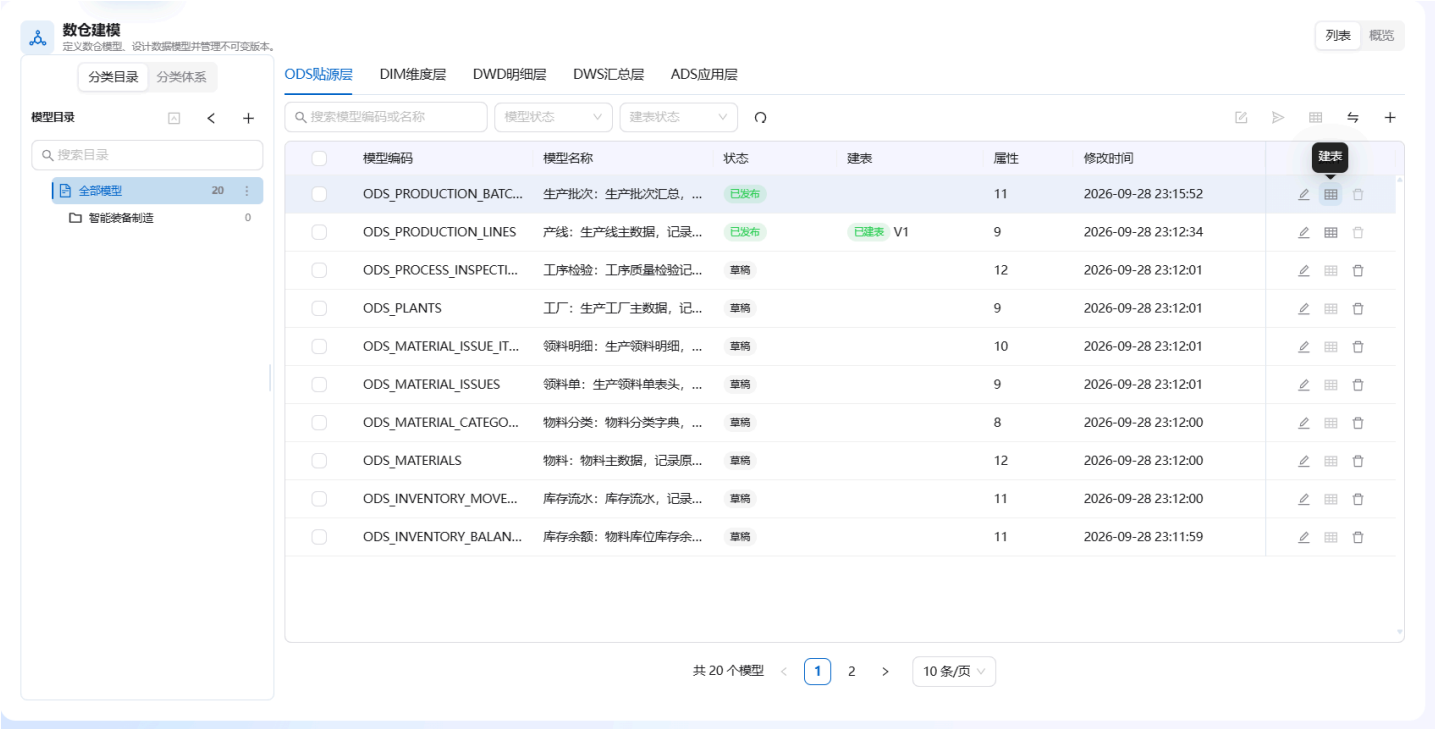


图 66执行建表

- 说明:
- 在已发布模型行使用建表入口, 依据已发布版本检查物理表预检结果, 确认后执行。
  - 建表处理中查看状态; 失败时查看提示并修正原因后重试。该动作处理表结构, 不会同步来源数据。

3.3.8 逆向生成模型

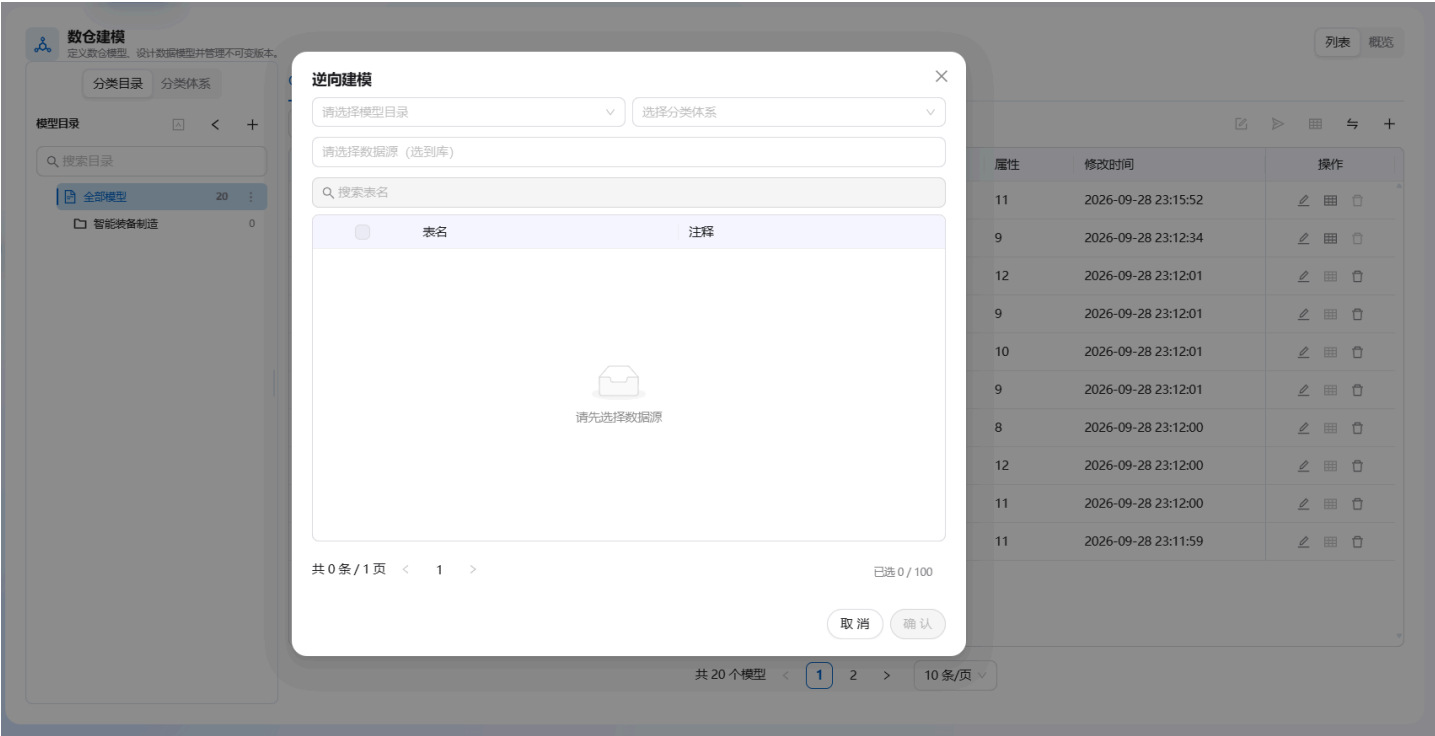


图 67逆向生成模型

说明:

- 打开逆向入口，选择数据源与数据库，查找并勾选真实表；可跨页选择。确认当前目录、分层与前缀后生成。
- 查看成功与失败数量及失败明细。部分成功时先到列表确认已生成模型，仅重新处理失败对象，避免重复生成成功项。

### 3.3.9 删除数仓模型

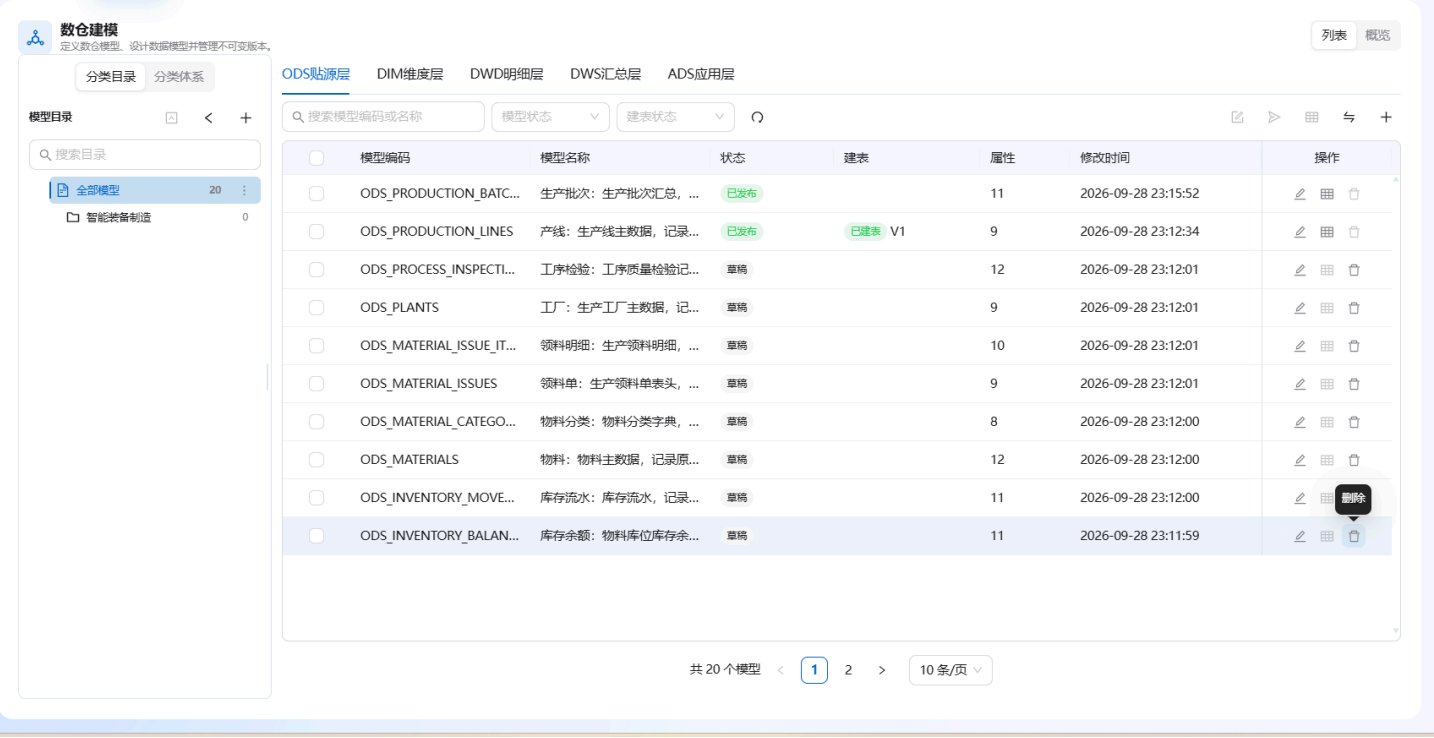


图 68删除模型

- 说明：
- 从列表选择删除，检查提示后确认。已有版本的模型不能直接删除。
  - 模型定义管理与物理数据库管理不同；不要假定删除模型会清空数据库，也不要通过删除物理表绕过模型状态限制。

### 3.3.10 代码预览

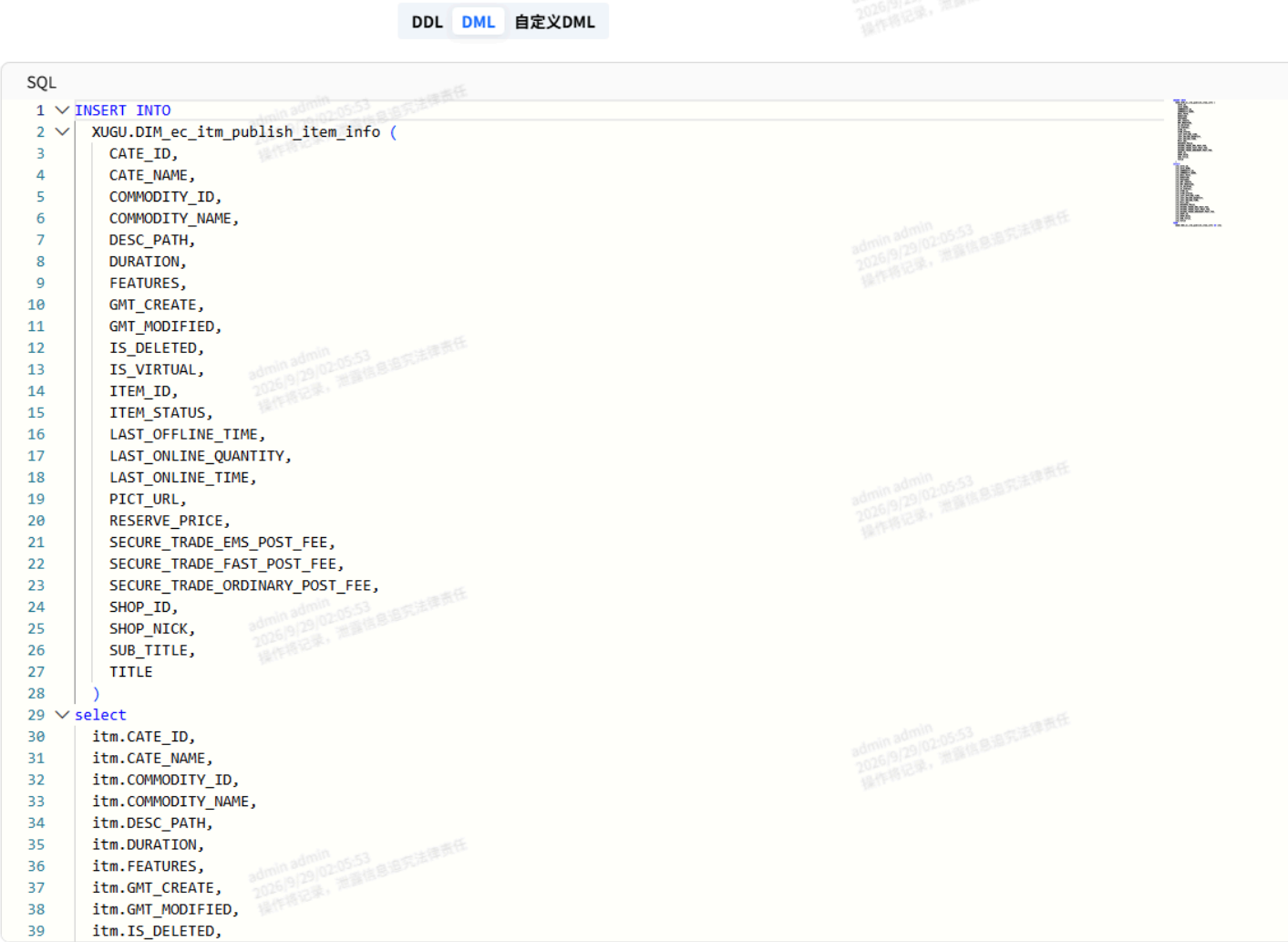


图 69模型代码预览

说明：

- 完成模型配置后，可通过“代码预览”查看模型对应的 DDL、DML 和自定义 DML，核对表结构及数据处理逻辑。
- DDL**：查看模型对应的表结构定义 SQL，核对表名称、字段名称、字段类型等信息，确认表结构是否符合建模要求。
- DML**：查看根据建模配置生成的数据写入 SQL，核对目标表、来源表及字段映射关系。图示采用 `INSERT INTO ... SELECT ...` 语句，将来源表中的指定字段写入目标表。
- 自定义DML**：查看模型配置的自定义数据处理 SQL，核对个性化的数据转换与加工逻辑。
- 代码区域提供语法高亮、行号和代码折叠功能，便于阅读与检查。核对完成后，单击窗口右上角的关闭按钮，返回模型编辑页面。

3.4 数据开发

数据开发通过工作流编排SQL、脚本、接口调用、条件分支和数据管道。工作流定义处理步骤，管道进一步定义输入、转换、输出。编辑与生产运行分离，保存后可调试，发布后交由运维管理。

### 3.4.1 工作流目录

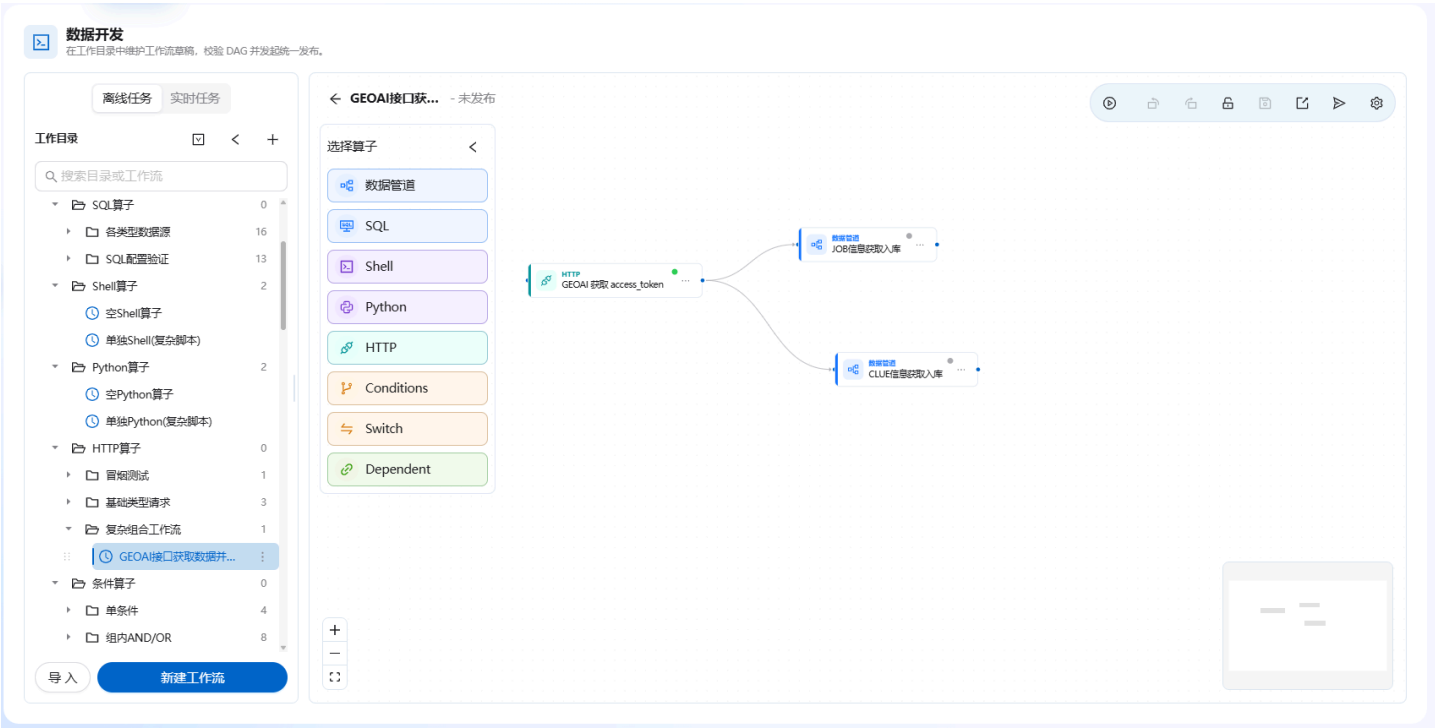


图 69工作流目录

#### 说明：

#### 查看工作流目录

- 进入数据开发，左侧查看统一目录及离线、实时工作流，点击工作流进入画布。
- 离线与实时通过对象类型区分；使用目录组织项目，例如“物流履约分析”。

#### 移动工作流目录

- 在目录树拖动目录到允许的目标位置，调整父目录或顺序。
- 移动后确认目录层级；目录层级不改变工作流的节点执行关系。

#### 移动工作流

- 在目录树将工作流移到目标目录，确认移动后的归属。
- 移动用于重新组织对象，不改变数据源位置，也不执行数据搬迁。

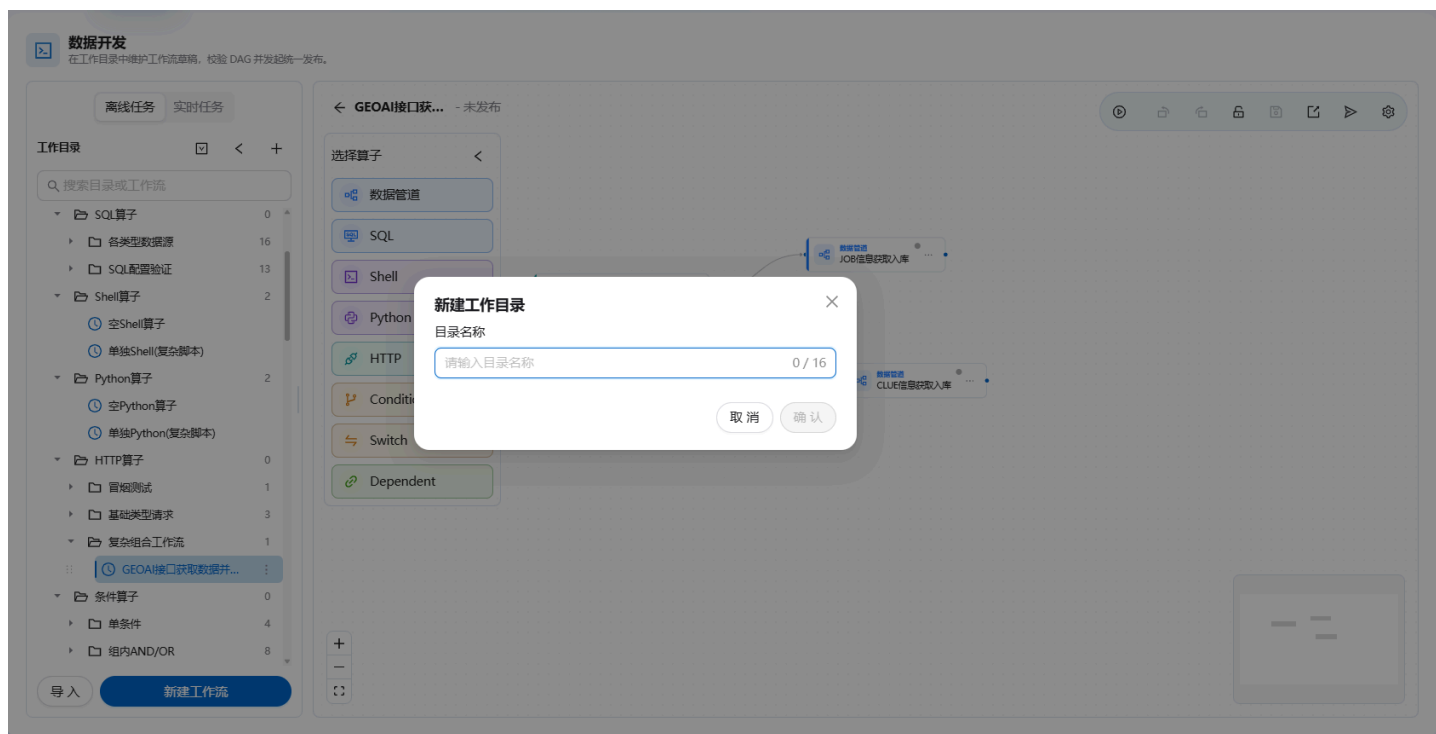


图 70新建目录

说明：

- 在目录区域选择新建目录，填写名称及所属位置后确认。
- 目录用于组织工作流，不是任务之间的执行依赖。

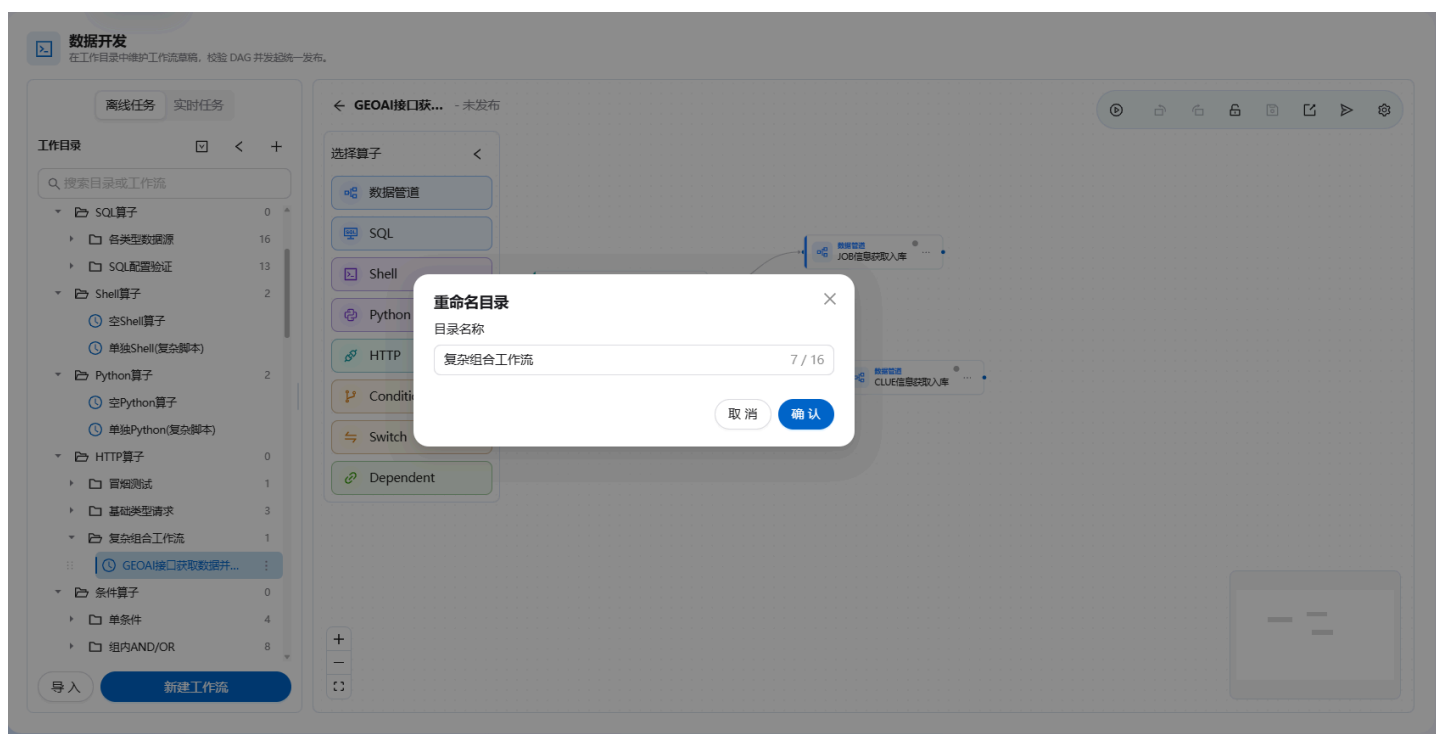


图 71重命名

说明：

- 在目录菜单选择重命名，修改名称后保存。
- 名称调整后重新核对目录位置，不影响工作流内部节点连线。

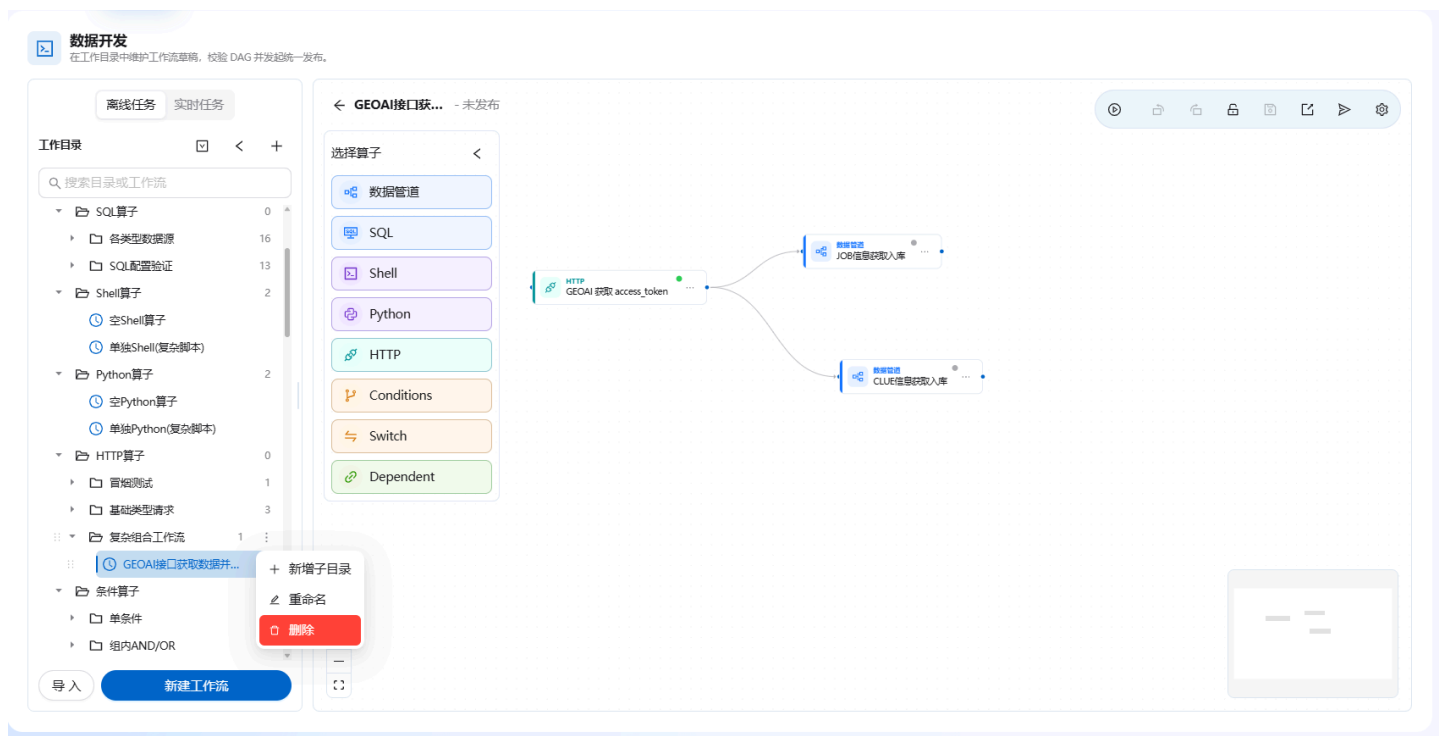


图 72删除目录

### 说明：

- 在目录操作菜单选择删除，先核对包含的工作流和子目录。
- 按提示处理关联内容后删除，避免把仍需运行的工作流一并误处理。

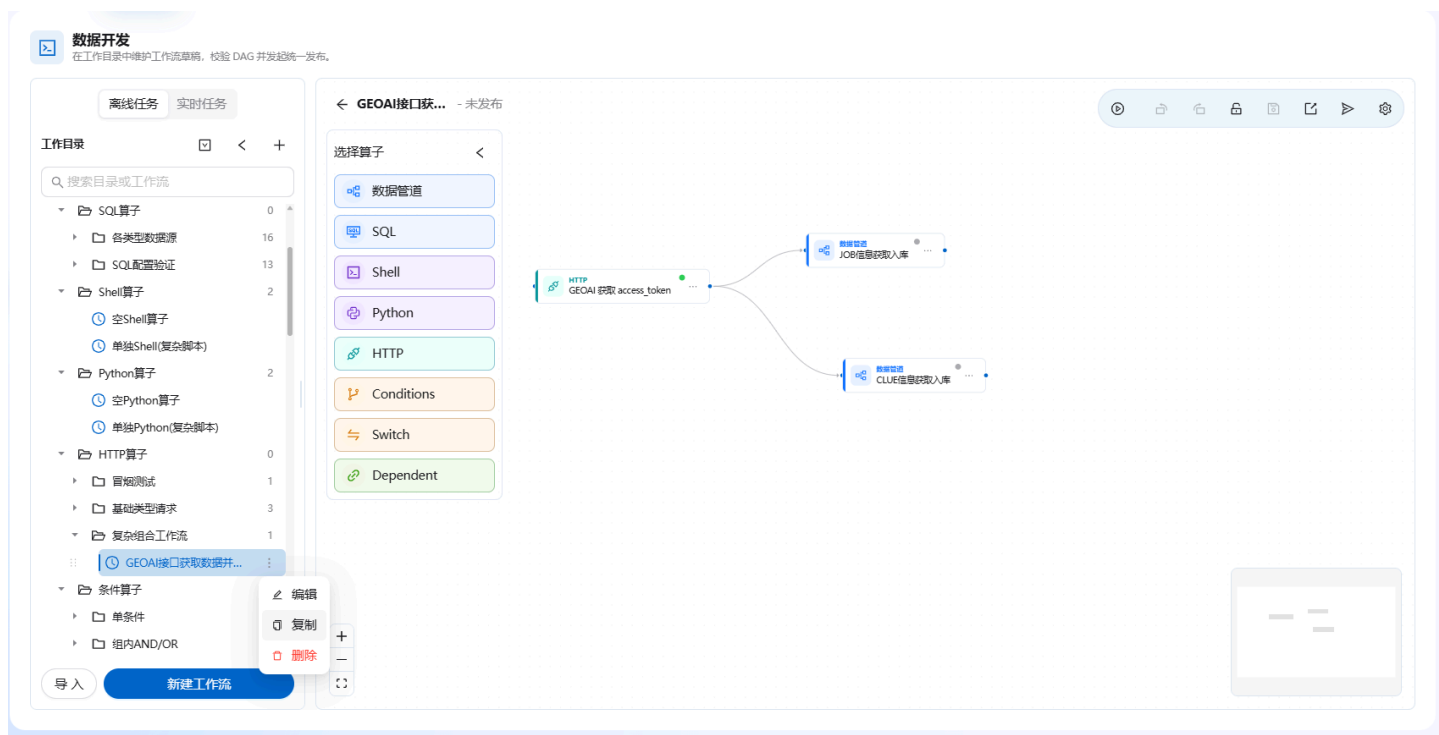


图 73复制工作流

### 说明：

- 在工作流菜单选择复制，按提示完成新对象创建。

- 打开副本核对名称、连接、参数与输出位置，再保存或发布；不要让副本误写原任务正在使用的目标。

### 3.4.2 新建 workflow

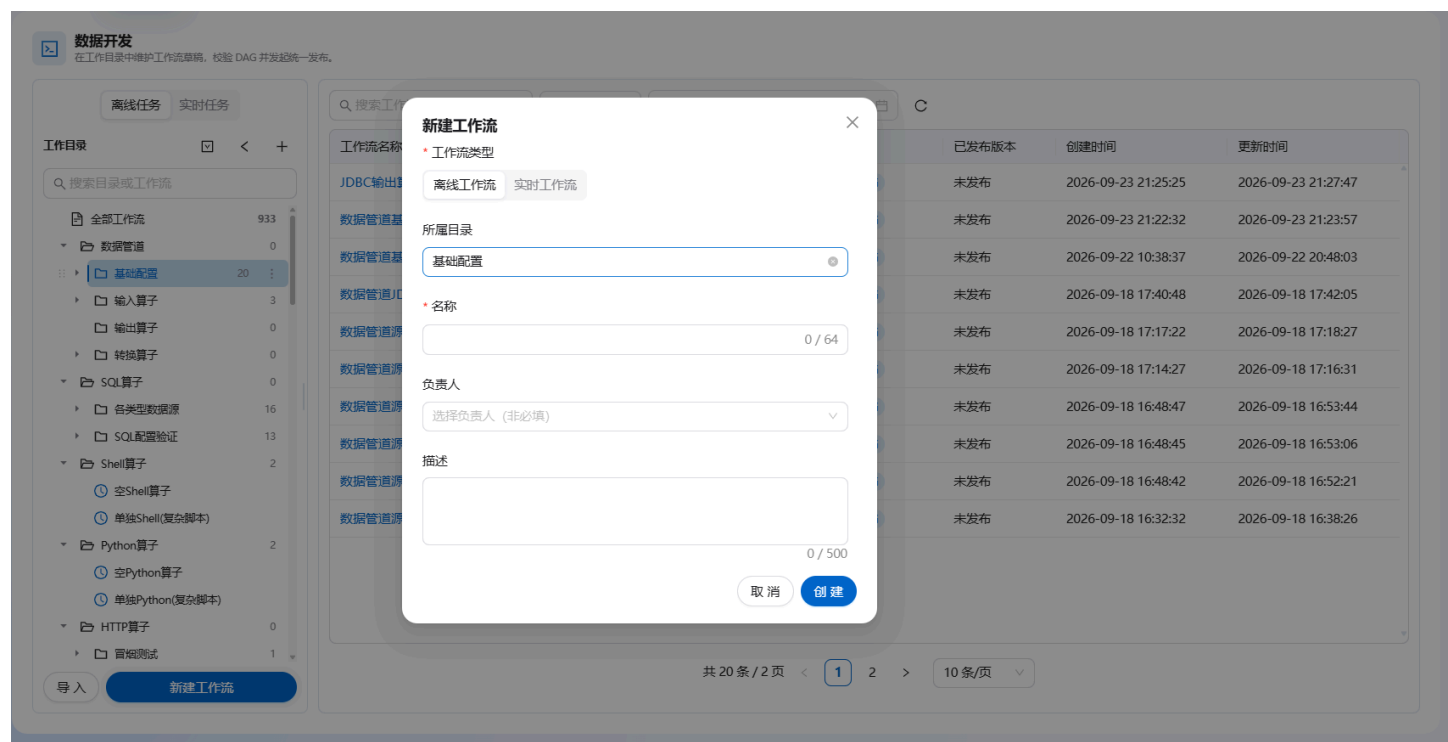


图 74新建工作流

- 说明：
- 点击新建工作流，选择离线或实时类型，填写名称、目录、负责人和描述。名称最多32字，描述最多500字。
  - 类型在创建后不可直接切换。离线可组织多种任务节点；实时工作流使用数据管道节点处理持续数据流。

### 3.4.3 编辑与编排工作流

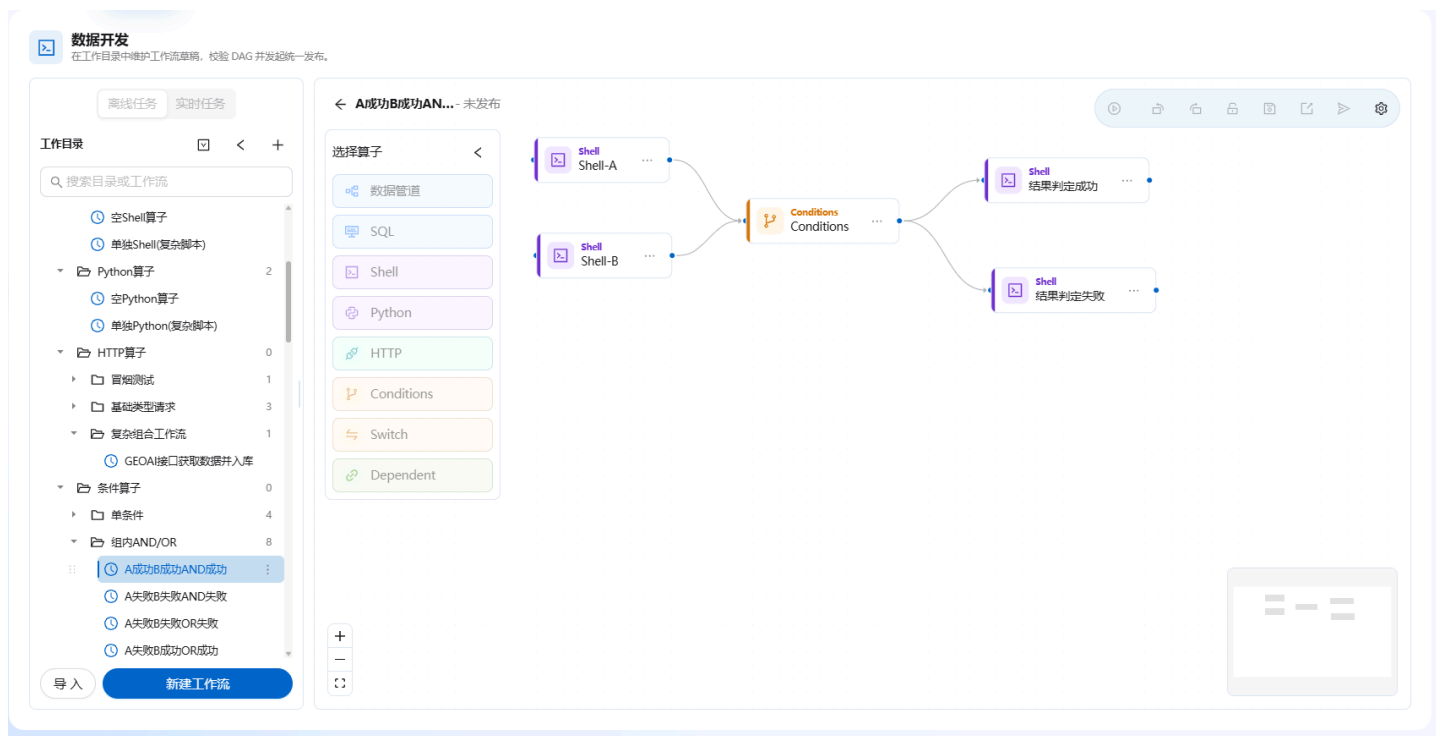


图 75画布操作

## 说明：

### 编排工作流

- 从节点库添加节点，通过连线定义先后依赖，调整位置使流程清晰。工具栏提供保存、配置、导出、发布与锁定等入口。
- 通过节点更多菜单打开配置；数据管道使用进入入口打开二层画布。选择节点不等于打开配置。
- 使用撤销、重做及缩放辅助编辑。删除节点会影响相连依赖，保存前检查起止关系与条件分支。

### 保存工作流

- 点击保存，检查名称、节点必填配置、连线与参数提示，修正错误后再次保存。
- 保存保留草稿，不替换生产版本。离开前确认已保存，避免只保留浏览器内修改。

### 锁定工作流

- 使用锁定入口保护当前工作流编辑状态。
- 被其他用户锁定时按只读状态查看，不强行覆盖对方修改；需要调整时与持锁人协作。

### 解锁工作流

- 本人锁定的工作流可使用解锁入口恢复编辑。
- 解锁后重新确认内容版本，再修改和保存，避免覆盖他人已完成的调整。

×

workflows

保存

基本信息

workflows参数

\* 名称

A成功B成功AND成功11 / 64

workflows类型 ⓘ

离线 workflows

负责人

选择负责人 (非必填) ▾

描述

0 / 500

图 76 workflows配置

说明：

- 打开 workflows配置的基本信息，查看或维护名称、负责人、描述和组织信息。
- 配置完成后保存，使 workflows画布和基本信息保持一致。

×

× workflows配置

保存

基本信息

工作流参数

参数名

值来源

值

amount

自定义

99

✕

+

图 77工作流参数

- 说明：
- 切换工作流参数，新增参数键并设置来源和值。参数键应唯一，节点可使用\${参数键}引用。
  - 常量用于固定配置，系统调度时间表达式用于按运行时间生成值。例如业务日期应由参数明确传递，不能假定补数时间自动成为SQL过滤条件。



图 78实时启动与恢复

说明：

- 实时 workflows配置中切换“启动与恢复”，设置首次启动与检查点恢复选项。
- 根据任务已有的可用恢复点选择恢复策略；新建 workflows尚无历史恢复点时，先完成首次启动。

×

数据管道 · JOB信息获取入库

保存

基础配置

参数配置

运行控制

| 参数名          | 方向 | 数据类型    | 值来源 ? | 值      |                           |
|--------------|----|---------|-------|--------|---------------------------|
| access_token | IN | VARCHAR | 上游出参  | GEO... | <div>✖</div> <div>+</div> |

图 79参数配置

说明：

- 打开节点配置的参数配置页签，维护输入、输出参数及其取值来源。输入可来自常量、工作流参数或上游输出。

- 使用函数选择入口为支持的参数构造表达式；敏感值按界面掩码显示。输出参数需与实际执行结果对应，不能只定义名称而不产生结果。

×

数据管道 · JOB信息获取入库

保存

基础配置

参数配置

运行控制

运行标志

运行

延时执行时间 (分钟)

0

超时配置

超时告警

任务优先级

失败重试

次数

0

间隔 (分钟)

0

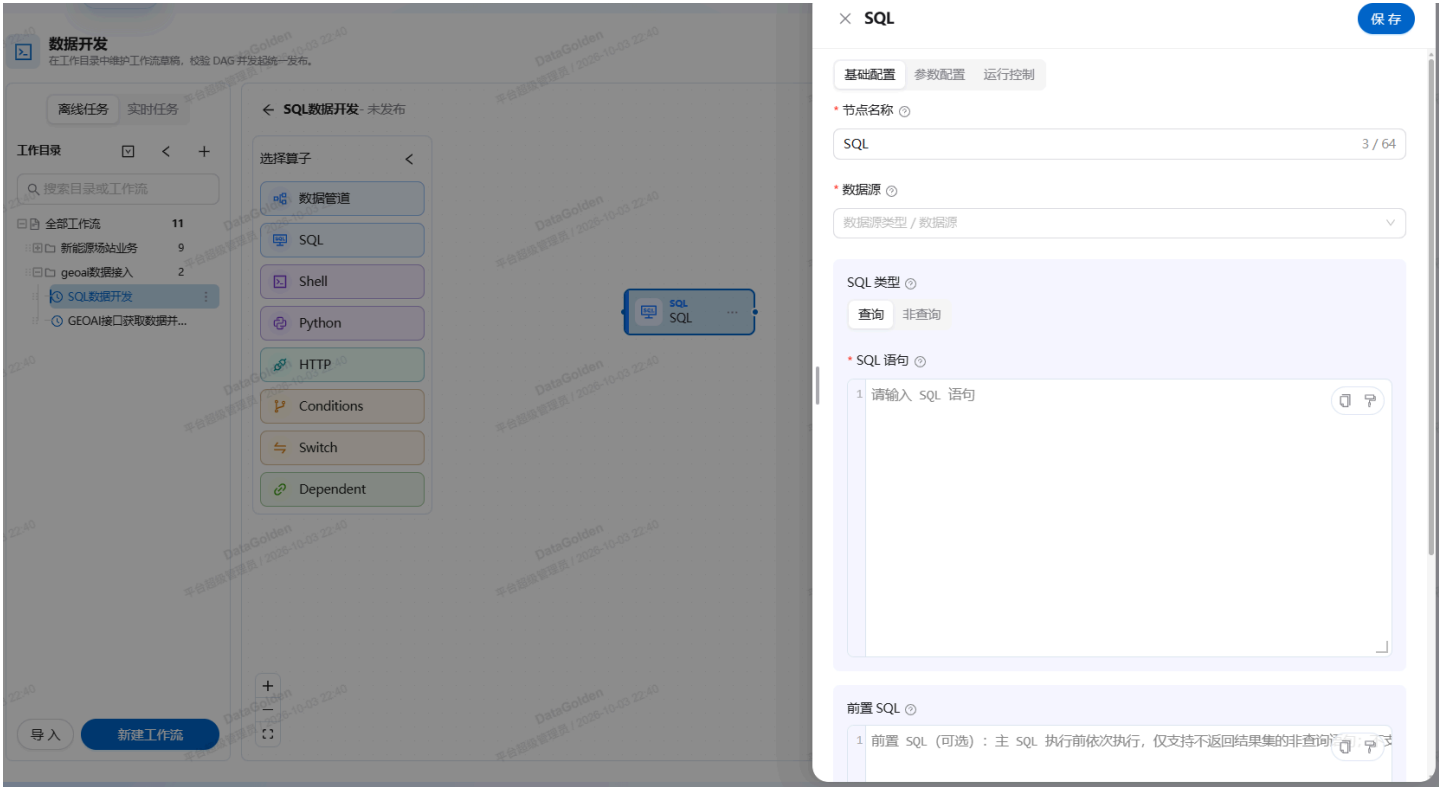
图 80运行控制

说明：

- 切换节点运行控制，设置任务优先级、失败重试次数与间隔、超时及相应处理方式。
- 重试可能再次执行外部写入，应保证脚本、SQL或接口可安全重复执行。超时告警与超时失败的处理含义不同。

### 3.4.4 离线 workflow 节点配置

#### SQL 算子



×

SQL

保存

基础配置

参数配置

运行控制

\* 节点名称

SQL

3 / 64

\* 数据源

SQLServer / sqlserver数据源dg\_sync\_1433 (192.168.31.218:1433/dg\_sync\_260906)

SQL 类型

查询

非查询

\* SQL 语句

1 SELECT \* FROM user\_basic\_info

前置 SQL

1 前置 SQL (可选) : 主 SQL 执行前依次执行, 仅支持不返回结果集的非查询语句

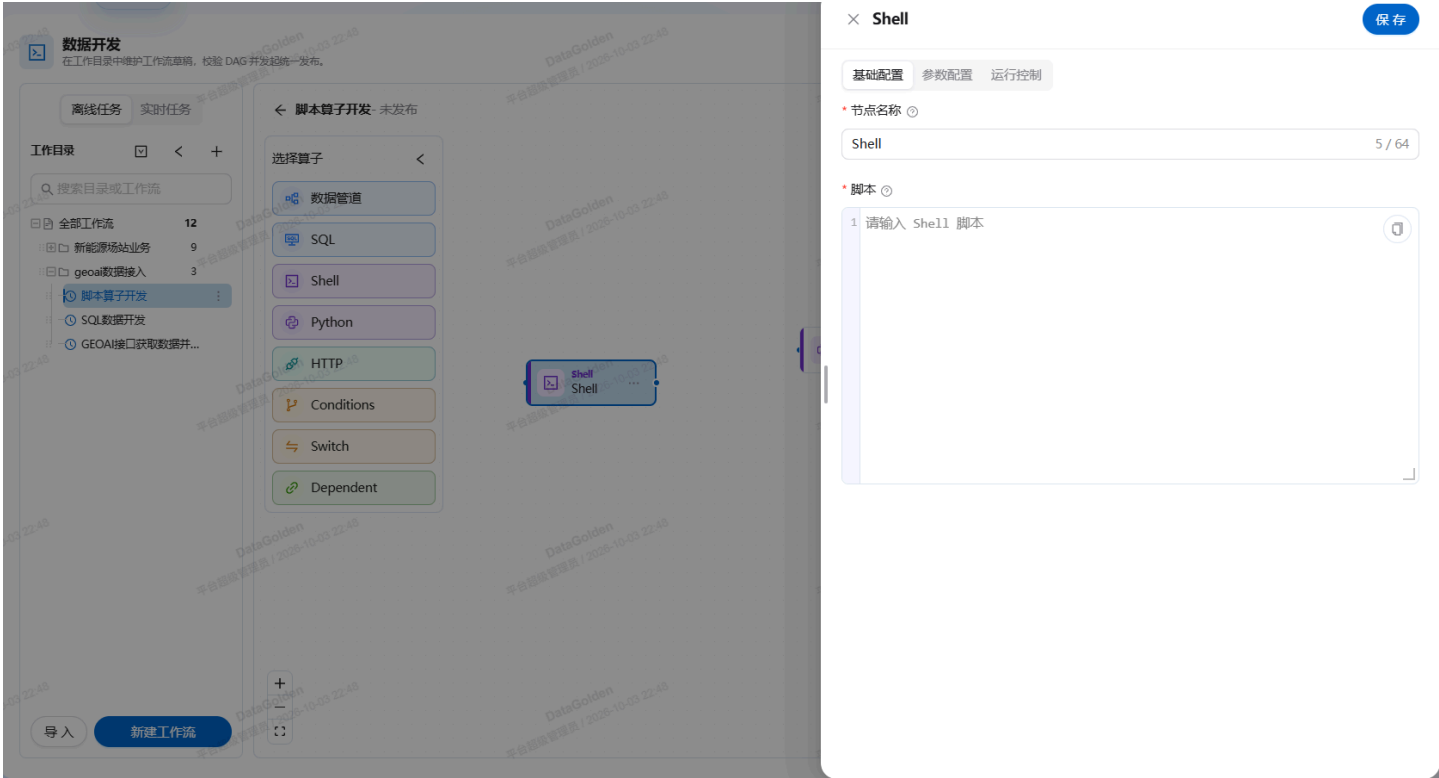
图 81配置SQL节点

说明:

- 支持虚谷、Oracle、Mysql、达梦、人大金仓、PostgreSQL、SQLServer、ClickHouse、Doris、OceanBase、GaussDB等数据源。

- 在节点基础配置选择数据源，选择查询或非查询类型，编写SQL；根据需要填写前置、后置SQL和结果展示行数。
- 查询用于返回结果，非查询用于建表或写入等操作。预览展示行数只控制展示，不代表SQL一定只处理这些行。
- 汇总数据可在目标库使用SQL聚合订单，但执行前应检查目标表、参数和写入重复风险。

## Shell算子



基础配置

参数配置

运行控制

\* 节点名称 ?

Shell

5 / 64

\* 脚本 ?

```
1 # =====
2 # 脚本自测工具 - 验证自身功能的完整性
3 # 用法: ./self_test.sh
4 # =====
5
6 # ===== 颜色定义 =====
7 RED='\033[0;31m'
8 GREEN='\033[0;32m'
9 YELLOW='\033[1;33m'
10 BLUE='\033[0;34m'
11 NC='\033[0m'
12
13 PASSED=0
14 FAILED=0
15 TOTAL=0
16
17 # ===== 断言函数 =====
18 assert_eq() {
19     local expected="$1"
```

图 82配置Shell节点

说明:

- 在基础配置填写Shell脚本，按需在参数配置中接收工作流或上游参数。
- 脚本运行于执行环境，不是当前浏览器所在电脑；使用环境中真实存在且有权限的路径和命令。保存后使用测试资源验证输出与退出状态。

- 在基础配置编写Python脚本，查看平台提供的只读运行环境信息。
- 依赖包和文件需在执行环境可用，不能假定本机安装的依赖会随脚本上传。通过节点参数传递输入并检查实际输出。

## Python算子

The screenshot displays the DataGolden platform's workflow editor. On the left, a sidebar shows a project named '数据开发' (Data Development) with a search bar and a list of workflows. The main workspace is titled '脚本算子开发 - 未发布' (Script Operator Development - Unpublished). A '选择算子' (Select Operator) panel is open, listing various operators: 数据管道 (Data Pipeline), SQL, Shell, Python, HTTP, Conditions, Switch, and Dependent. The 'Python' operator is selected and added to the workflow. On the right, a configuration panel for the 'Python' operator is shown. It has tabs for '基础配置' (Basic Configuration), '参数配置' (Parameter Configuration), and '运行控制' (Execution Control). The '基础配置' tab is active, showing the node name 'Python' (6/64 characters), the Python runtime environment 'Python 3.144 (平台默认)' (Platform Default), and a script editor with the placeholder text '请输入 Python 脚本' (Please enter Python script).

基础配置

参数配置

运行控制

\* 节点名称 ?

Python

6 / 64

Python 运行环境 ?

Python 3.14.4 (平台默认)

\* 脚本 ?

```
1 #!/usr/bin/env python3
2 # =====
3 # Python 脚本自测工具 - 验证自身功能的完整性
4 # 用法: python3 self_test.py
5 # =====
6
7 import os
8 import sys
9 import re
10 import subprocess
11 import tempfile
12 import time
13 from datetime import datetime
14
15 # ===== 颜色定义 =====
16 class Colors:
17     RED = '\033[0;31m'
18     GREEN = '\033[0;32m'
```

图 83配置Python节点

说明:

- 在基础配置编写Python脚本，查看平台提供的只读运行环境信息。
- 依赖包和文件需在执行环境可用，不能假定本机安装的依赖会随脚本上传。通过节点参数传递输入并检查实际输出。

×

HTTP • GEOAI 获取 access\_token

保存

基础配置参数配置运行控制

\* 节点名称 ⓘ

GEOAI 获取 access\_token21 / 64

\* 方法

POST

\* URL

http://192.168.31.210:8001/gateway/geoai-auth-service/oauth/token

连接超时 (毫秒) ⓘ

60000

ParamsAuthHeadersBody

| 参数名           | 值                                 |    |                |
|---------------|-----------------------------------|----|----------------|
| Content-Type  | application/x-www-form-urlencoded | fx | <div>🗑️+</div> |
| Authorization | 函数 · concat                       | fx | <div>🗑️+</div> |

出参

在线探查

图 84HTTP节点

说明：

- 在基础配置填写完整URL和请求方法。HTTP节点用于调用接口，与管道内分页读取并入库的HTTP输入用途不同。

ParamsAuthHeadersBody

参数名

值

Key

值

fx

🗑️

+

出参

在线探查

图 85HTTP节点 Params

说明：

- 配置请求查询参数，核对键名、值和参数引用。

ParamsAuthHeadersBody

Type

No Auth

▼

出参

在线探查

图 86HTTP节点—Auth

说明：

- 按接口要求填写认证设置，避免把密码写入URL或普通说明。

ParamsAuthHeadersBody

参数名

值

Content-Type

application/x-www-form-urlencoded

fx

🗑️

+

Authorization

函数 · concat

fx

🗑️

+

图 87HTTP节点—Headers

说明：

- 填写请求头，例如内容类型或业务要求的头字段。

ParamsAuthHeadersBody

表单参数JSON

| 参数名        | 值                    |                 |
|------------|----------------------|-----------------|
| grant_type | client_credentialsfx | <div>🗑️ +</div> |

出参

在线探查

图 88HTTP节点—Body

说明：

- 选择接口需要的请求体形式；表单与JSON按界面规则选择，不同时提交相互冲突的内容。
- 在具备接口访问条件时探查响应，从响应结构选择需要的输出，或配置JSONPath、XPath等提取表达式。
- 下游引用的是提取后的输出参数。探查也会实际调用接口，对有写入作用的请求应使用测试地址。

条件算子

×

Conditions

保存

基础配置

参数配置

运行控制

\* 节点名称 ?

Conditions10 / 64

状态 ?

成功

\* 分支流转

结果判定成功

状态 ?

失败

\* 分支流转

结果判定失败

添加前置检查条件

且

且

+

+

Shell-A

成功

+

Shell-B

成功

+

图 89配置Conditions节点

- 说明：
- 先连接上下游，再在基础配置选择AND或OR，设置上游任务成功、失败等条件。
  - 分别指定条件成立与不成立的后续节点，两个分支应按页面要求区分。该节点判断执行状态，不直接筛选订单数据行。

## 分支算子

×

Switch

保存

基础配置

参数配置

运行控制

\* 节点名称 ?

Switch6 / 64

条件分支 ?

\${amount}> 100

Shell-B1

+

默认分支 ?

Shell-D

图 90配置Switch节点

说明：

- 依次填写判断条件和对应后续节点，设置唯一的默认分支。

- 按规则顺序检查匹配结果；默认分支用于没有条件命中的情况，避免遗漏未匹配路径。

依赖算子

×

Dependent

保存

基础配置

参数配置

运行控制

\* 节点名称

Dependent9 / 64

添加依赖

\* 依赖工作流

被依赖工作流A成功 (v1)

▼

\* 时间周期

日今天

传递依赖输出参数

且

\* 依赖工作流

被依赖工作流B失败 (v1)

▼

\* 时间周期

日今天

传递依赖输出参数

或

\* 依赖工作流

▼

\* 时间周期

传递依赖输出参数

+

\* 依赖工作流

▼

\* 时间周期

传递依赖输出参数

图 91配置Dependent节点

说明：

- 选择已发布工作流，并指定依赖全部任务或单个任务，设置当天、前一天、前周或前月等依赖周期。
- 配置检查间隔，以及依赖不满足时失败或继续等待和等待时长。依赖判断的是外部执行结果，不会自动补建缺失的来源数据。

### 3.4.5 调试工作流

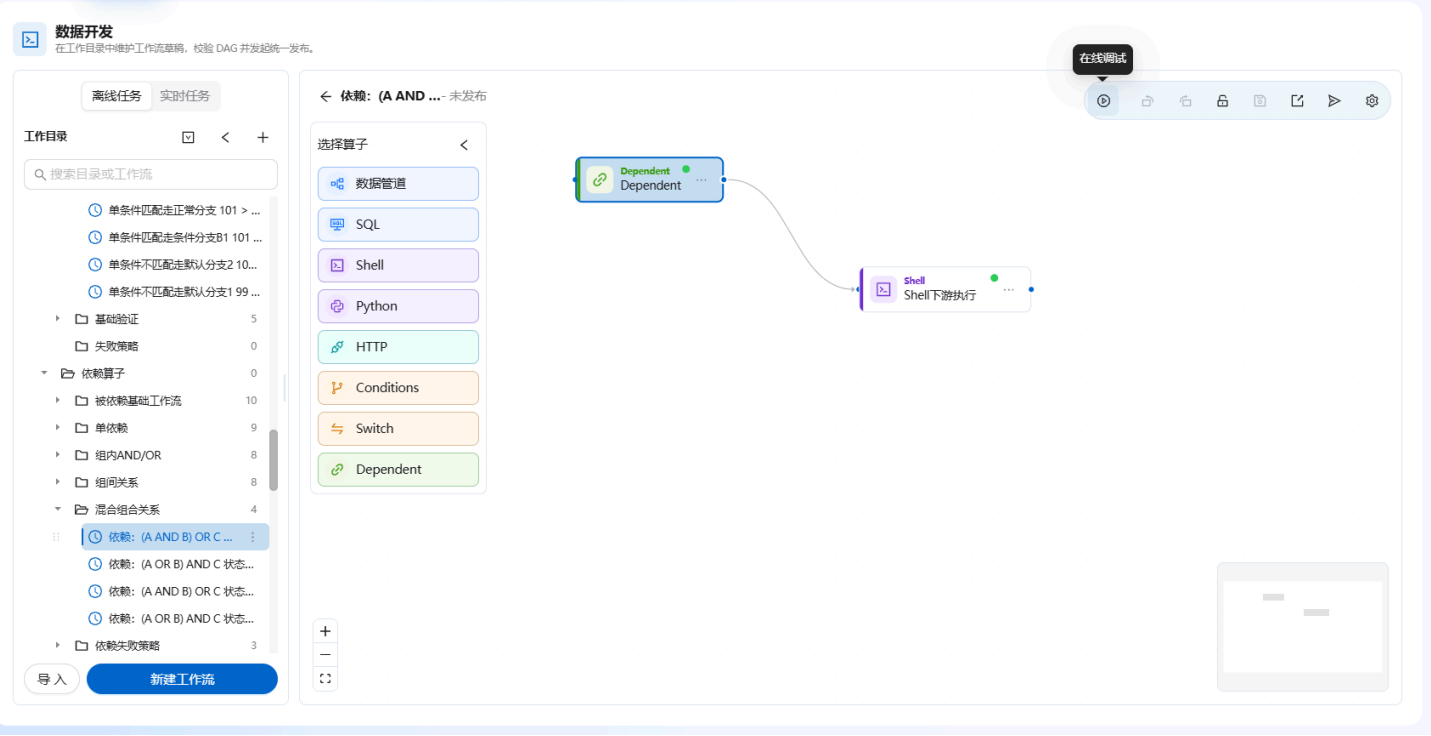


图 92工作流调试

#### 说明：

##### 调试整个工作流

- 保存配置后点击调试，启动整个工作流的实际执行。
- 调试会访问真实数据源并可能写入数据库或调用外部接口，应使用测试资源。调试不是静态语法检查，也不是发布的替代动作。

##### 调试当前节点

- 在目标节点更多菜单选择调试当前节点，确认该节点需要的输入已准备后执行。
- 该动作聚焦单个节点；依赖的上游输出未产生时，应改用调试到此或先准备输入。

##### 调试到当前节点

- 在节点菜单选择调试到此，执行目标节点及其所需上游，不执行目标之后的下游。
- 适合逐段验证物流数据读取和清洗；运行前仍需检查上游可能产生的写入。

##### 查看调试结果

- 打开调试结果查看运行信息、节点状态和错误提示。

停止调试

- 在调试结果顶部使用停止入口，等待运行状态更新。
- 关闭抽屉或离开页面不会自动停止调试。停止也不会自动撤销此前已提交的外部写入。

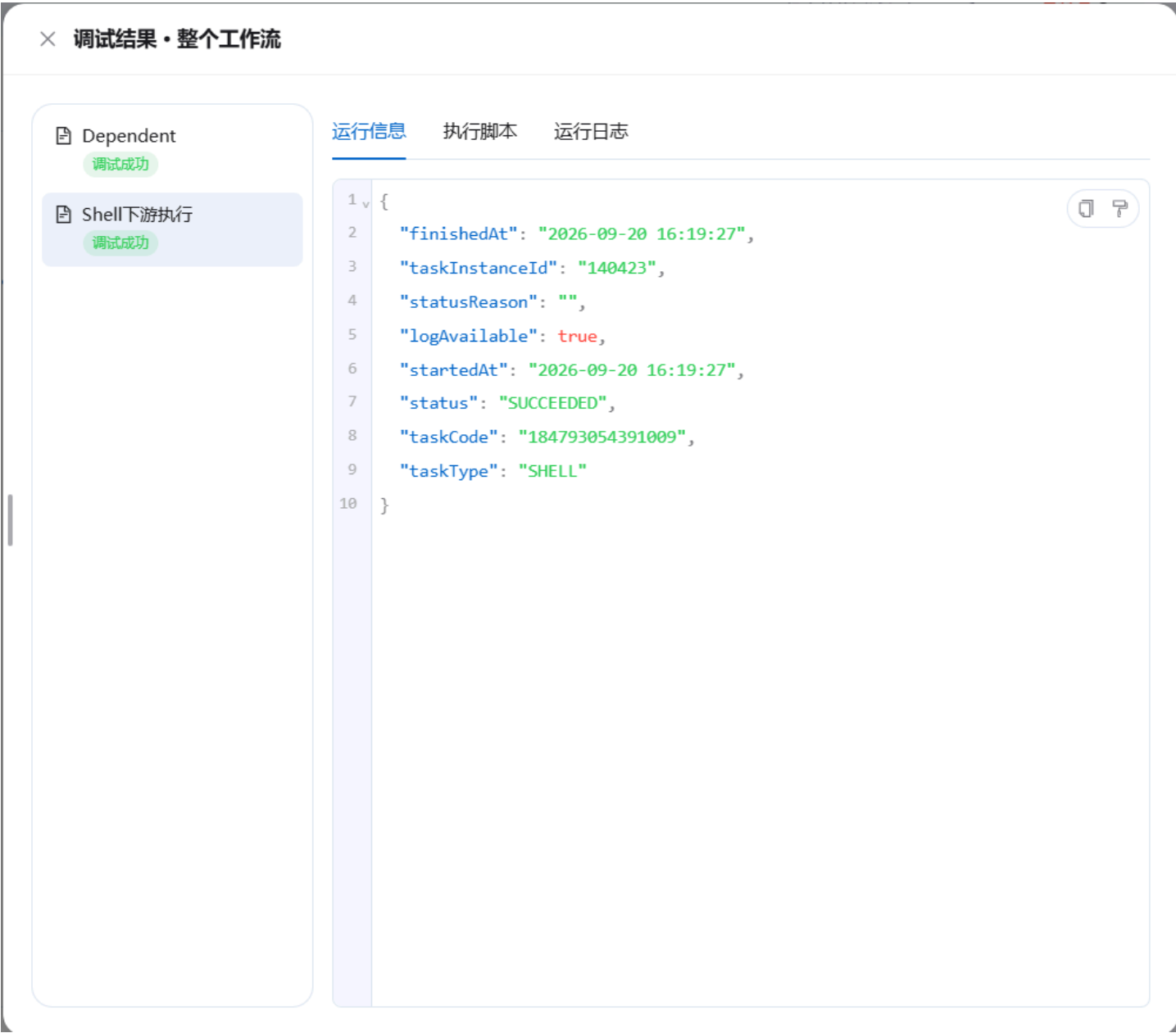


图 93运行信息

说明：

- 按节点类型查看实际执行脚本或配置，核对参数替换与目标位置。

数据管道  
调试成功

运行信息 执行配置 运行日志

调度日志 作业日志

在已加载日志中搜索

已加载 283 / 共 283 行

↺

🔍

📄

⬇️

⬆️

⬇️

|     |                         |      |                                                                              |
|-----|-------------------------|------|------------------------------------------------------------------------------|
| 259 | 2026-09-28 19:25:42,366 | INFO | [o.a.s.c.z.h.HikariDataSource] [BlockingWorker-TaskGroupLocation[jobId=      |
| 260 | 2026-09-28 19:25:42,366 | INFO | [o.a.s.c.z.h.HikariDataSource] [BlockingWorker-TaskGroupLocation[jobId=      |
| 261 | 2026-09-28 19:25:42,366 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 262 | 2026-09-28 19:25:42,366 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 263 | 2026-09-28 19:25:42,410 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 264 | 2026-09-28 19:25:42,415 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 265 | 2026-09-28 19:25:42,415 | INFO | [JdbcMultiTableResourceManager] [BlockingWorker-TaskGroupLocation[jc         |
| 266 | 2026-09-28 19:25:42,415 | INFO | [o.a.s.c.z.h.HikariDataSource] [BlockingWorker-TaskGroupLocation[jobId=      |
| 267 | 2026-09-28 19:25:42,415 | INFO | [o.a.s.c.z.h.HikariDataSource] [BlockingWorker-TaskGroupLocation[jobId=      |
| 268 | 2026-09-28 19:25:42,415 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 269 | 2026-09-28 19:25:43,052 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 270 | 2026-09-28 19:25:43,054 | INFO | [o.a.s.e.s.TaskExecutionService] [hz.main.seaTunnel.task.thread-31556] - [10 |
| 271 | 2026-09-28 19:25:43,055 | INFO | [a.s.e.s.s.DefaultSlotService] [hz.main.generic-operation.thread-19] - recei |
| 272 | 2026-09-28 19:25:43,527 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 273 | 2026-09-28 19:25:43,527 | INFO | [o.a.s.e.s.TaskExecutionService] [hz.main.seaTunnel.task.thread-31556] - [10 |
| 274 | 2026-09-28 19:25:43,658 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 275 | 2026-09-28 19:25:43,658 | INFO | [o.a.s.e.s.TaskExecutionService] [hz.main.seaTunnel.task.thread-31556] - [10 |
| 276 | 2026-09-28 19:25:43,659 | INFO | [a.s.e.s.s.DefaultSlotService] [hz.main.generic-operation.thread-28] - recei |
| 277 | 2026-09-28 19:25:43,745 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 278 | 2026-09-28 19:25:43,745 | INFO | [o.a.s.e.s.TaskExecutionService] [hz.main.seaTunnel.task.thread-31556] - [10 |
| 279 | 2026-09-28 19:25:43,746 | INFO | [a.s.e.s.s.DefaultSlotService] [hz.main.generic-operation.thread-15] - recei |
| 280 | 2026-09-28 19:25:43,766 | INFO | [o.a.s.e.s.TaskExecutionService] [BlockingWorker-TaskGroupLocation[jobId=    |
| 281 | 2026-09-28 19:25:43,766 | INFO | [o.a.s.e.s.TaskExecutionService] [hz.main.seaTunnel.task.thread-31556] - [10 |
| 282 | 2026-09-28 19:25:43,766 | INFO | [a.s.e.s.s.DefaultSlotService] [hz.main.generic-operation.thread-17] - recei |
| 283 | 2026-09-28 19:25:44,429 | INFO | [a.s.e.s.s.DefaultSlotService] [hz.main.generic-operation.thread-38] - recei |

图 94 运行日志

说明：

- 查看日志中的执行过程与错误，不把界面返回成功提示当作数据处理全部成功；结合节点终态和目标结果判断。

### 3.4.6 发布与版本管理

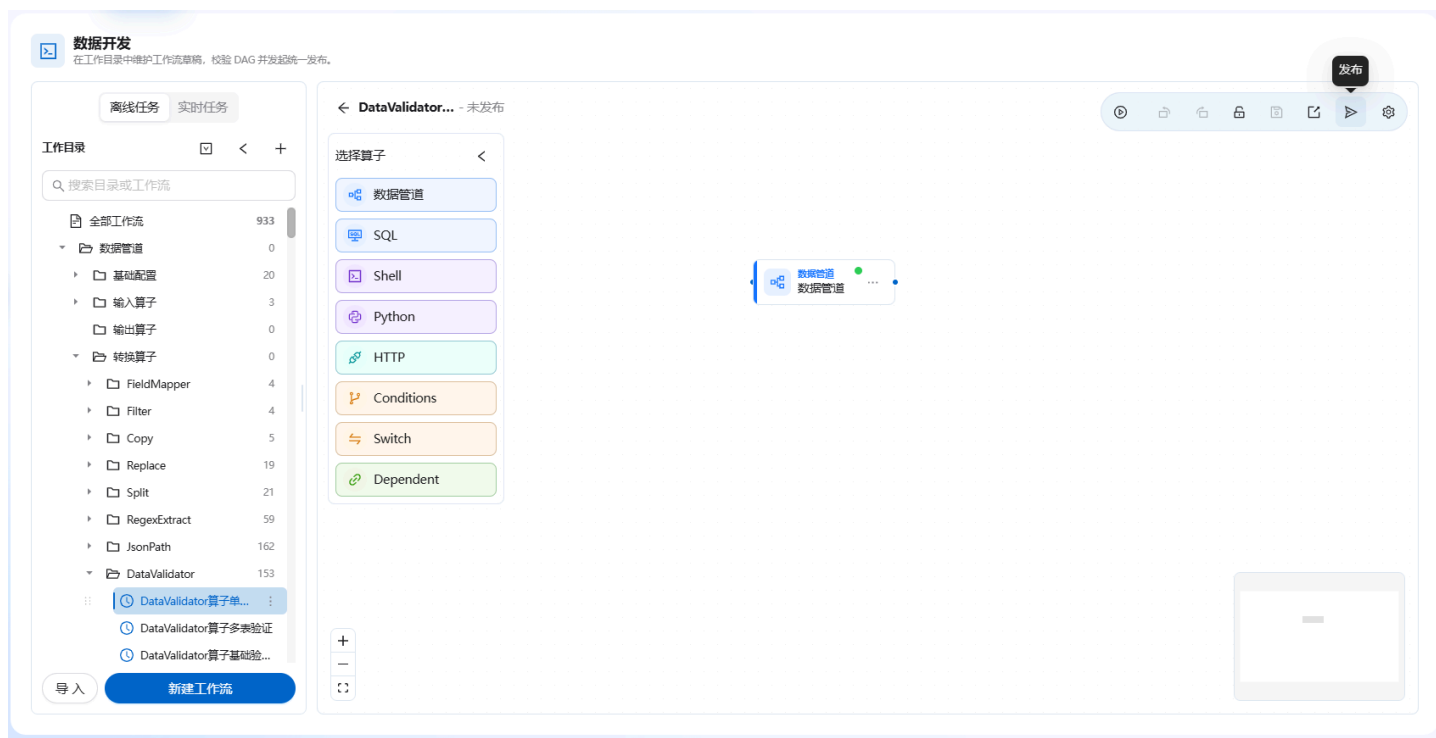


图 95发布工作流

#### 说明：

- 保存后点击发布，检查待发布内容并提交，按需等待审批。
- 发布成功后到生产工作流核对版本，再配置调度或启动。在线调试成功不代表工作流已经发布。

○

v2

当前生效

审批通过 已发布

提交人  
平台超级管理员

提交时间  
2026-09-20 14:48:28

变更说明  
提交工作流开发版本

○

v1

启用

审批通过 已发布

提交人  
平台超级管理员

提交时间  
2026-09-20 14:20:44

变更说明  
提交工作流开发版本

基本信息 工作流参数 节点配置

名称: 被依赖工作流A Group成功

工作流类型: 离线工作流

负责人: user04006

描述: —

图 96版本历史

#### 说明：

##### 查看工作流版本

- 在工作流配置中进入版本页签，查看已发布版本及当前版本信息。
- 问题定位时使用实例对应的生产版本；新草稿和历史生产版本的参数、连线可能不同。

##### 启用历史工作流版本

- 在版本历史中选择可启用的已发布版本，确认后启用。
- 完成后到生产工作流核对当前版本；该操作不撤销历史实例产生的外部写入，也不代替数据修复。

### 3.4.7 导入与导出

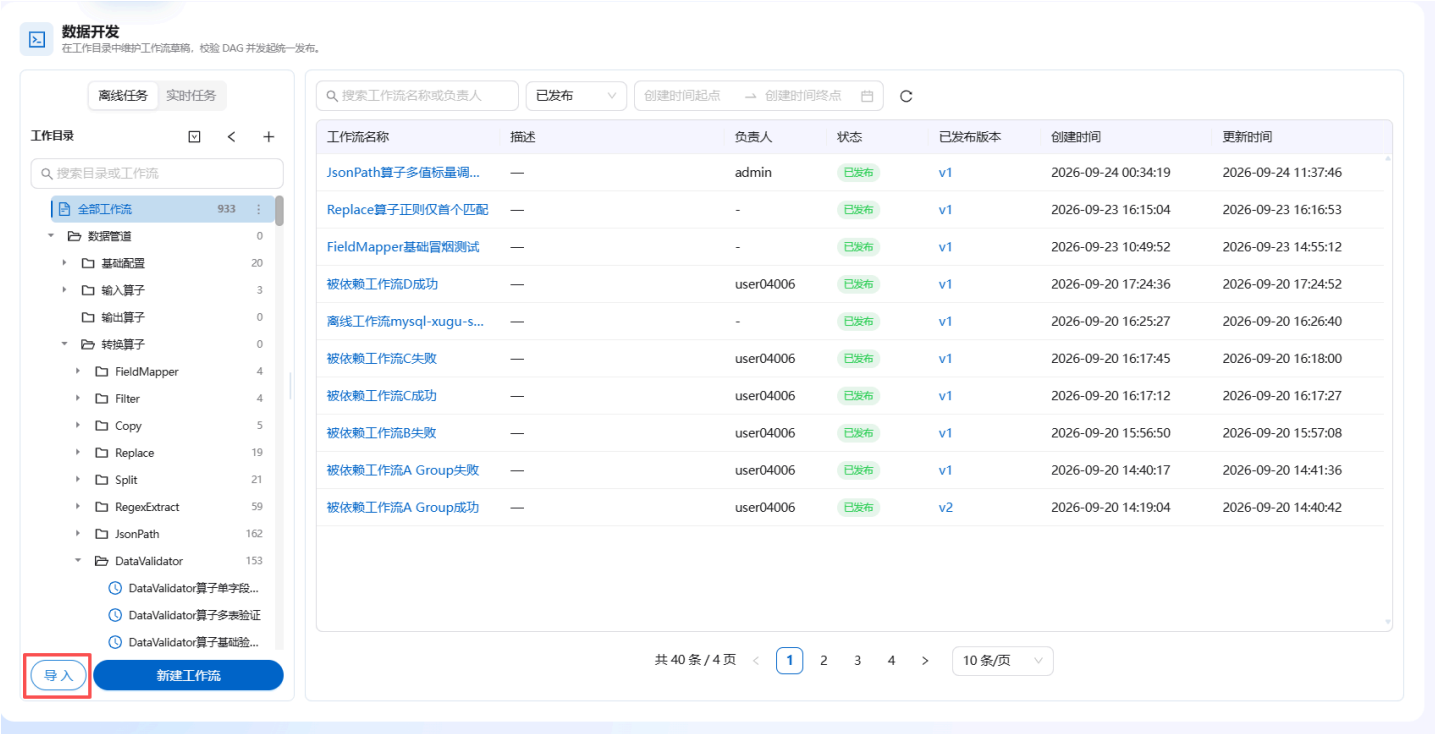


图 97导入导出

- 说明：
- 导入工作流
- 使用导入入口选择配置文件，填写新工作流名称与所属目录，并完成本环境连接检查。
  - 导入生成新的工作流对象；发布前检查参数、数据源、子工作流及输出位置，不直接沿用其他环境的资源。
- 导出工作流
- 点击导出，保存已保存工作流的配置文件。
  - 先保存需要保留的修改；配置导出不包含环境凭据或真实业务数据。

### 3.4.8 删除工作流

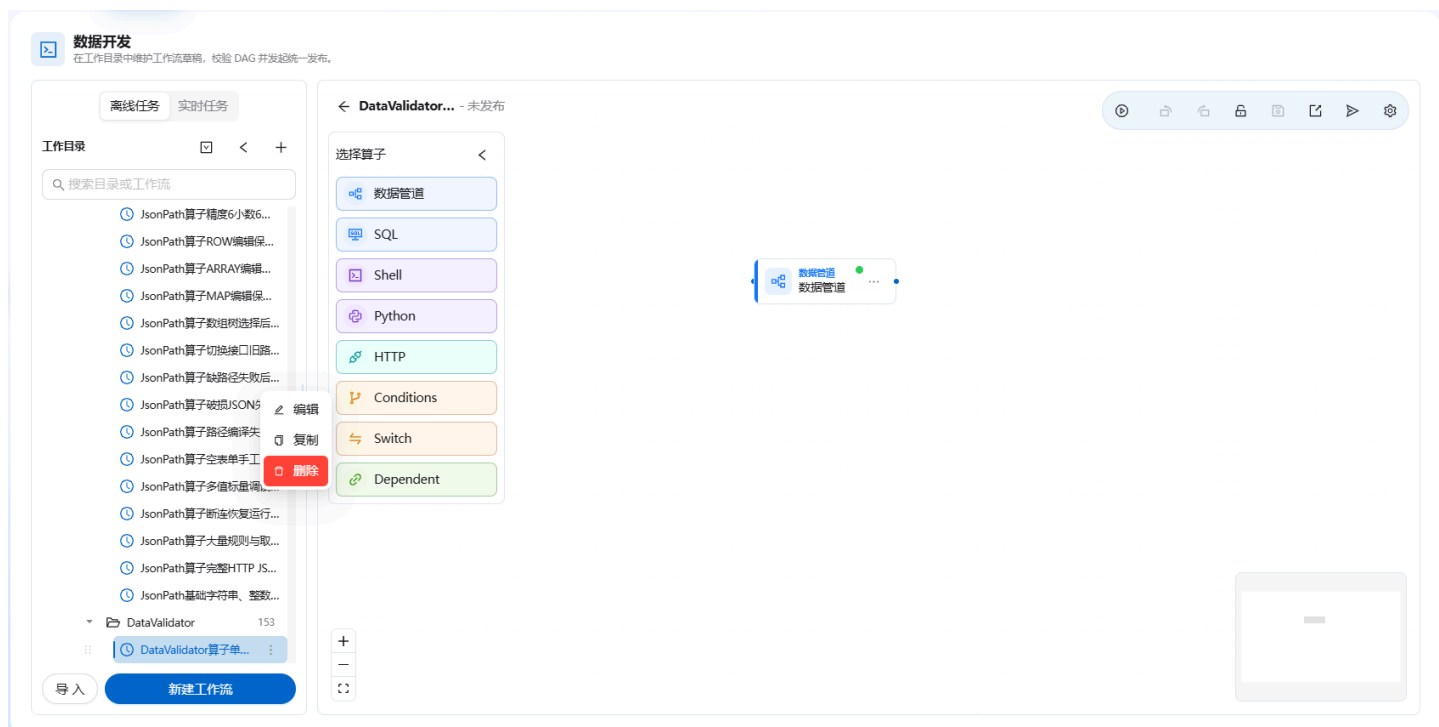


图 98删除工作流

#### 说明：

- 在工作流操作入口选择删除，确认对象。已发布对象先停用后续调度并确认无活动实例。
- 历史实例保留供查询；删除定义不负责清理输出表或撤销历史写入。

## 3.5 数据管道

数据管道在工作流节点内部编排输入、转换和输出算子。先连线提供上游字段，再设置转换和目标映射。离线与实时模式显示的算子不同；配置随所属工作流保存和发布。

### 3.5.1 管道画布

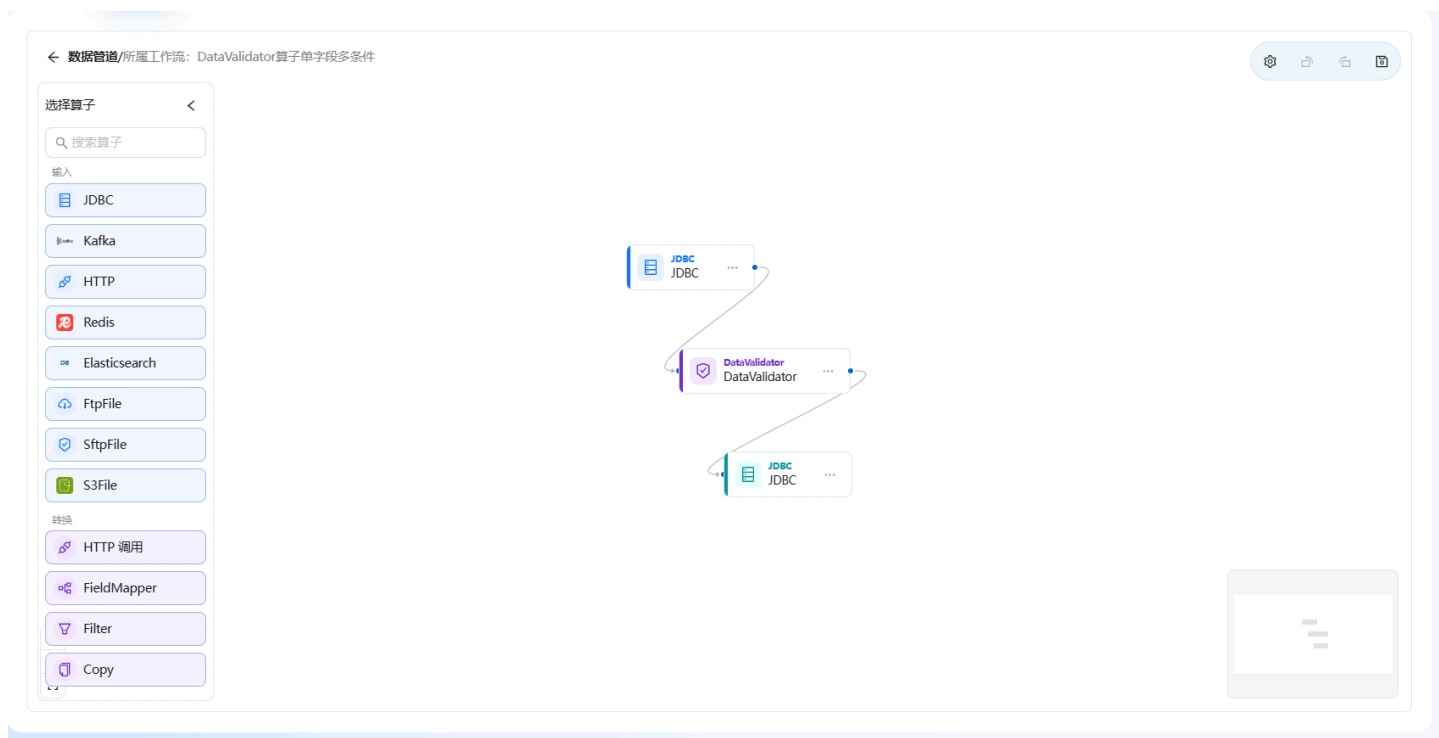


图 99编辑管道画布

- 说明：
- 通过数据管道节点的进入入口打开二层画布，从输入、转换、输出分类添加算子并连接。
  - 选择算子打开配置，逐项检查必填字段。上游变更后重新核对下游字段；完成后返回工作流并保存，管道不单独发布。

3.5.2 离线 workflow 管道算子配置

3.5.2.1 输入算子配置

JDBC输入

×

JDBC

保存

基础配置

读取配置

\* 算子名称

JDBC4 / 64

\* 数据源

MySQL / mysql数据源dg\_sync (192.168.31.213:3306/dg\_sync\_260906)

读取方式

按表读取

Q 搜索已选来源表

已添加1 张表 · 单输入最多 50 张 +

| 来源表                       | 分片字段 | 操作 |
|---------------------------|------|----|
| tab_user_info_id<br>用户信息表 | id   |    |

共 1 条 / 1 页

1

>

10 条/页

图 100配置JDBC输入

- 说明：

- 基础配置中选择数据库数据源和来源表。查询SQL可留空以读取表字段，也可填写SELECT限定字段、WHERE过滤行，或使用数据库支持的查询。
- 来源表用于确定表身份和元数据；填写查询SQL后以查询结果作为下游输入，注意字段名与类型。
- JDBC输入来源数据源通过平台统一维护的数据源连接，支持 Xugu、Oracle、Mysql、达梦、PostgreSQL、SQLServer、OpenGauss、人大金仓等多种类型数据源，实现已支持的集中式存储（如Mysql、PostgreSQL等）和分布式存储（如Xugu、达梦等）之间的批量数据采集与传输。不同存储类型采用对应的读取配置。

Kafka输入

×

Kafka

保存

基础配置

字段配置

运行控制

\* 算子名称

Kafka5 / 64

\* 数据源 ⓘ

Kafka / orderdb-kafka-216 (192.168.31.216:9092)▼

\* Topic ⓘ

orderdb-orders▼

消费组 ⓘ

dg.pipeline.37006260150ddc2b4d874fdb1c2a471

图 102配置Kafka输入

说明：

- 基础配置选择Kafka数据源与单个Topic。消费组由系统生成并在任务内保持稳定。

×

Kafka

保存

基础配置

字段配置

运行控制

\* 消息格式

JSON

输出字段

探查消息结构

已读取1 条消息，识别19 个顶层字段

消息结构

{ 19 项

c\_array : array<string>

c\_bool : boolean

c\_double : double

c\_int64 : bigint

c\_object : { 2 项

batch\_id : string

sequence : int

字段名

数据类型

order\_id

BIGINT

+

order\_no

STRING

+

phone

STRING

+

status

STRING

+

total\_amount

DOUBLE

+

图 103Kafka输入一字段配置

- 说明：
- 选择JSON或Text。JSON可探查消息结构并生成输出字段，也可手工维护字段；Text按需要设置分隔符。
  - 输出字段应与实际消息结构一致；探查样例不能保证所有消息均符合该结构。

×

Kafka

保存

基础配置

字段配置

运行控制

\* 起始位置

最早位置

Checkpoint 完成后提交 Kafka 位点

拉取超时 (毫秒)

10000

图 104 Kafka输入—运行控制

说明：

- 选择消费组已提交位置、最早位置或最新位置，设置拉取超时。启用Checkpoint提交位点时，在检查点完成后提交消费进度。
- 从最新位置开始不会读取此前所有积压消息；重跑前确认消费组位置和恢复策略。

HTTP输入

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

\* 节点名称

获取无人机飞行任务JOB数据14 / 64

\* 方法

POST

\* URL

http://192.168.31.210:8001/gateway/geoai-api-service/v1/job/list

Params

Auth

Headers

Body

参数名

值

Key

值

fx

🗑

+

出参

在线探查

{ 5 项

code : string

msg : string

data : { 2 项

total : bigint

records : [ 11 项 设为数据记录

id : bigint

jobId : string

☒ 分页读取

页码字段

起始页码

每页记录数

总页数

图 105 配置HTTP输入

说明：

- 基础配置填写URL、请求方法、请求头、参数和请求体。接口分页时启用页码分页，设置页码字段、起始页、每页条数及可用的总量信息。

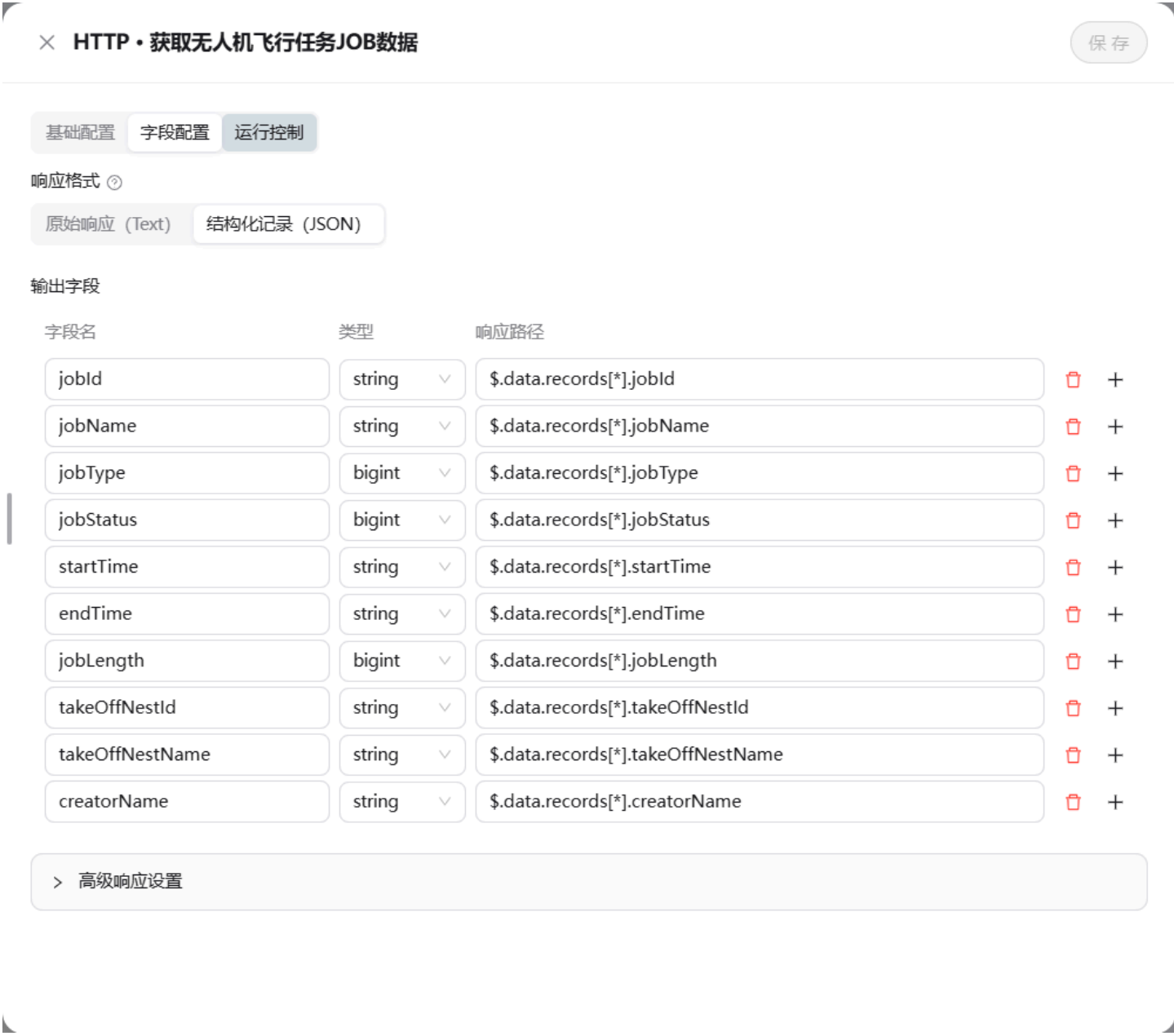


图 106 HTTP输入—字段配置

说明：

- 选择原始文本或结构化JSON解析；使用响应探查确定内容路径、数组位置和输出字段。
- 提取订单列表时选取实际列表数组，不把整个响应包装对象误作一条订单。

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

最大重试次数

3

重试间隔基准 (毫秒) ?

100

重试间隔上限 (毫秒) ?

10000

连接超时 (毫秒)

12000

读取超时 (毫秒)

60000

图 107 HTTP输入—运行控制

说明：

- 配置请求重试、重试间隔上限、连接和读取超时。接口限流或失败时结合返回信息调整，避免高并发持续重试。

Redis输入

×

Redis

保存

基础配置

字段配置

运行控制

\* 算子名称

Redis5 / 64

\* 数据源

Redis / Redis数据源dg\_sync\_6380 (192.168.31.216:6380/0)

\* 键或匹配规则

dg\_sync\_260906:user:260906009

\* 数据类型

String

图 108 配置Redis输入

说明：

- 基础配置选择Redis数据源、数据库编号、键范围和数据类型。键与类型应符合实际存储方式。

×

Redis

保存

基础配置

字段配置 (1)

运行控制

\* 值格式 ⓘ

JSON

\* 数据 Schema ⓘ

探查字段

已根据 1 条样本生成，请检查后保存。

1 {

2 "fields": {

3 "user\_id": "STRING",

4 "user\_name": "STRING",

5 "gender": "STRING",

6 "age": "STRING",

7 "city": "STRING",

8 "email": "STRING",

9 "member\_tier": "STRING",

10 "order\_count": "STRING",

11 "lifetime\_spend": "STRING",

请配置非空的 fields 对象，字段名和类型不能为空。

图 109 Redis输入—字段配置

说明：

- 选择JSON或Text并维护输出字段，确保读取值能按选定格式解析。

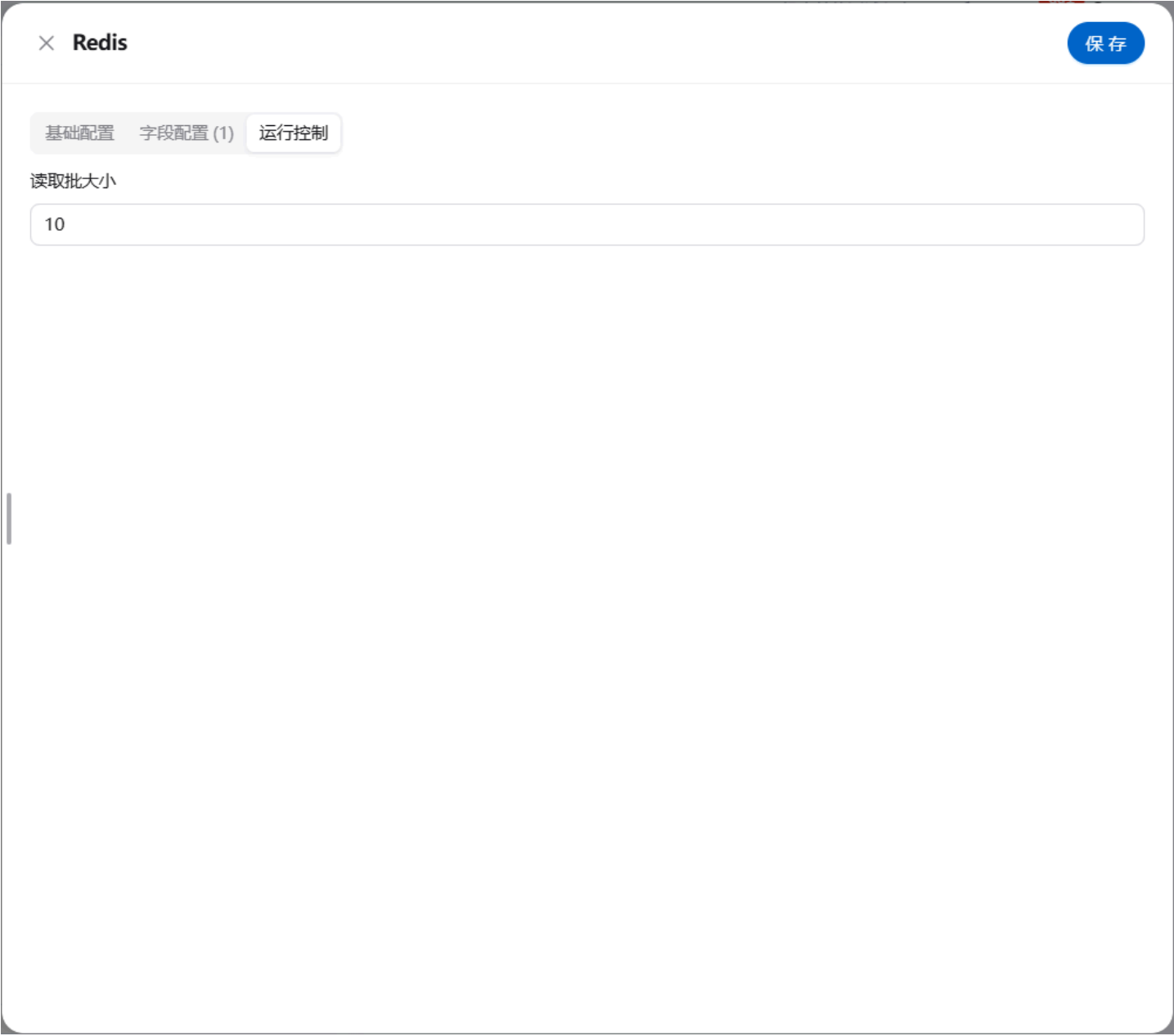


图 110 Redis输入—运行控制

说明：

- 设置批量读取大小，结合键数量与服务负载控制读取压力。

Elasticsearch输入

×

Elasticsearch

保存

基础配置

运行控制

\* 算子名称

Elasticsearch13 / 64

\* 数据源

数据源类型 / 数据源

\* 索引

请先选择数据源

查询 DSL

1{"match\_all":{}}

读取字段

搜索并选择字段，留空读取全部字段

图 111 配置Elasticsearch输入

说明：

- 基础配置选择数据源、索引并填写查询DSL，限定需要读取的文档。

×

Elasticsearch

保存

基础配置

运行控制

滚动批大小 ⓘ

100

滚动上下文时长

1m

图 112 Elasticsearch输入—运行控制

说明：

- 设置滚动读取大小及滚动上下文保持时间，结合查询规模调整；查询条件应先确认不会误读无关业务索引。

FtpFile输入

×

FtpFile

保存

基础配置

字段配置

\* 算子名称

FtpFile7 / 64

\* 数据源 ⓘ

FTP / FTP-216 (192.168.31.216:21//)▼

\* 文件路径

/data/

递归扫描子目录

☒

图 113 配置FtpFile输入

说明：

- 基础配置选择对应文件数据源，填写文件路径，并选择是否递归扫描子目录。原文件传输可选择全部文件或仅新增、变化文件，并按文件大小与修改时间或内容判断变化。

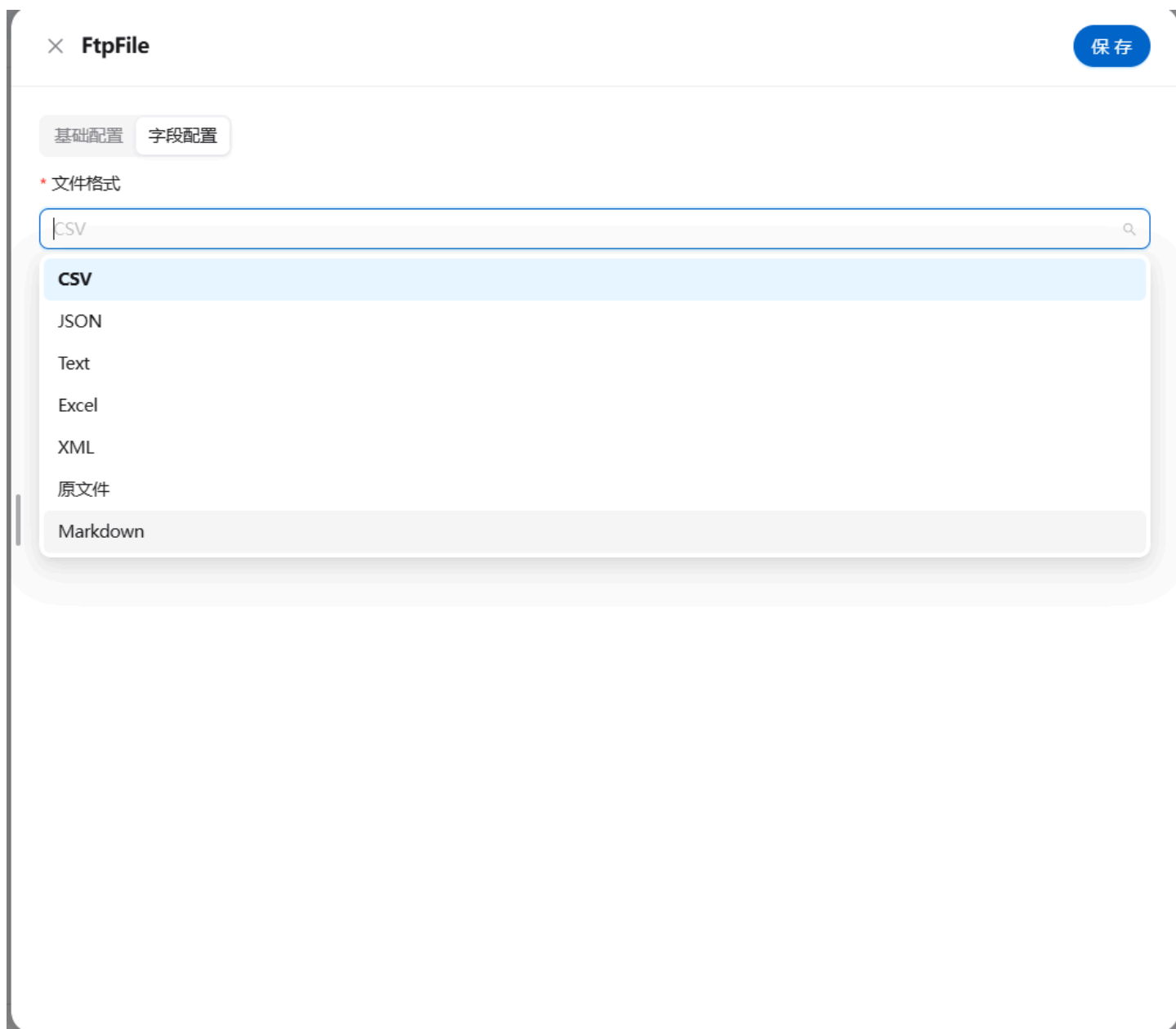


图 114 FtpFile输入一字段配置

#### 说明：

- FtpFile 输入算子用于采集 FTP 文件服务器中的文件，支持根据文件类型和业务需求，将文件内容解析为数据后入库，或将原文件采集至平台统一纳管，支持CSV、JSON、Text、Excel、XML、原文件或Markdown等适用格式。
- 对于 CSV、Excel 等结构化文件，选择对应的文件格式并配置解析规则，将文件内容解析为数据记录，通过下游输出算子写入目标数据库，适用于业务明细、统计报表等数据的采集。
- 对于 JSON、XML 等半结构化文件，选择对应的文件格式并配置解析规则，提取所需字段，通过下游输出算子写入目标数据库，适用于日志、接口结果等数据的采集。
- 对于 PNG 等图片及其他需要保留原始内容的文件，选择“原文件”格式，通过后续流程将其他文件服务器中的文件统一采集至平台纳管，适用于图片、文档等文件资源的集中管理。

#### SftpFile输入

×

SftpFile

保存

基础配置

字段配置

\* 算子名称

SftpFile8 / 64

\* 数据源 ?

SFTP / SFTP-216 (192.168.31.216:2222//)

\* 文件路径

/data

递归扫描子目录

☒

图 115 配置SftpFile输入

说明：

- 基础配置选择对应文件数据源，填写文件路径，并选择是否递归扫描子目录。原文件传输可选择全部文件或仅新增、变化文件，并按文件大小与修改时间或内容判断变化。

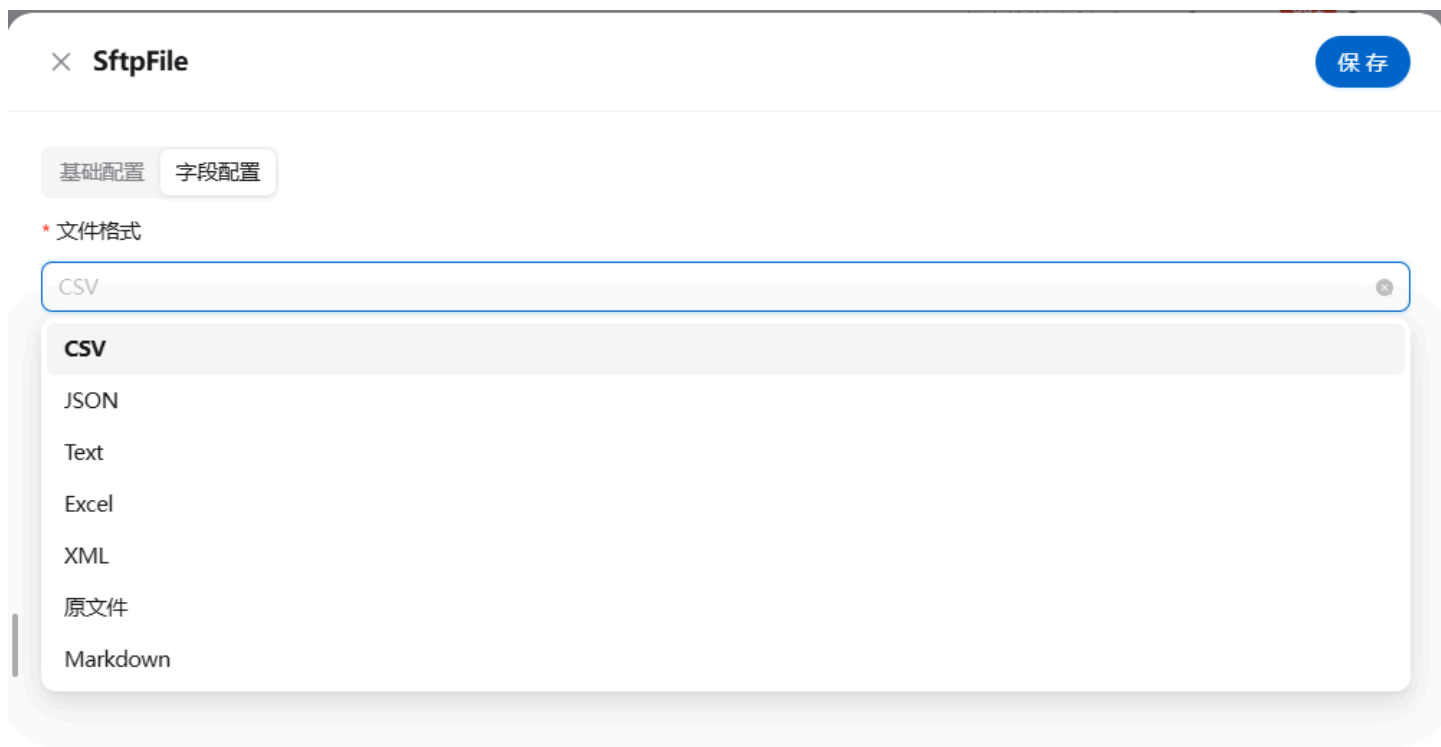


图 116 SftpFile输入—字段配置

说明：

- SftpFile 输入算子用于采集 SFTP 文件服务器中的文件，支持根据文件类型和业务需求，将文件内容解析为数据后入库，或将原文件采集至平台统一纳管，支持CSV、JSON、Text、Excel、XML、原文件或Markdown等适用格式。
- 对于 CSV、Excel 等结构化文件，选择对应的文件格式并配置解析规则，将文件内容解析为数据记录，通过下游输出算子写入目标数据库，适用于业务明细、统计报表等数据的采集。
- 对于 JSON、XML 等半结构化文件，选择对应的文件格式并配置解析规则，提取所需字段，通过下游输出算子写入目标数据库，适用于日志、接口结果等数据的采集。
- 对于 PNG 等图片及其他需要保留原始内容的文件，选择“原文件”格式，通过后续流程将其他文件服务器中的文件统一采集至平台纳管，适用于图片、文档等文件资源的集中管理。

S3File输入

×

S3File

保存

基础配置

字段配置

\* 算子名称

S3File6 / 64

\* 数据源 

MinIO / orderdb (192.168.31.216:9000/orderdb) ▾

\* 文件路径

/orders\_info

递归扫描子目录

☒

图 117 配置S3File输入

说明：

- 基础配置选择对应文件数据源，填写文件路径，并选择是否递归扫描子目录。S3File使用S3兼容对象存储数据源。

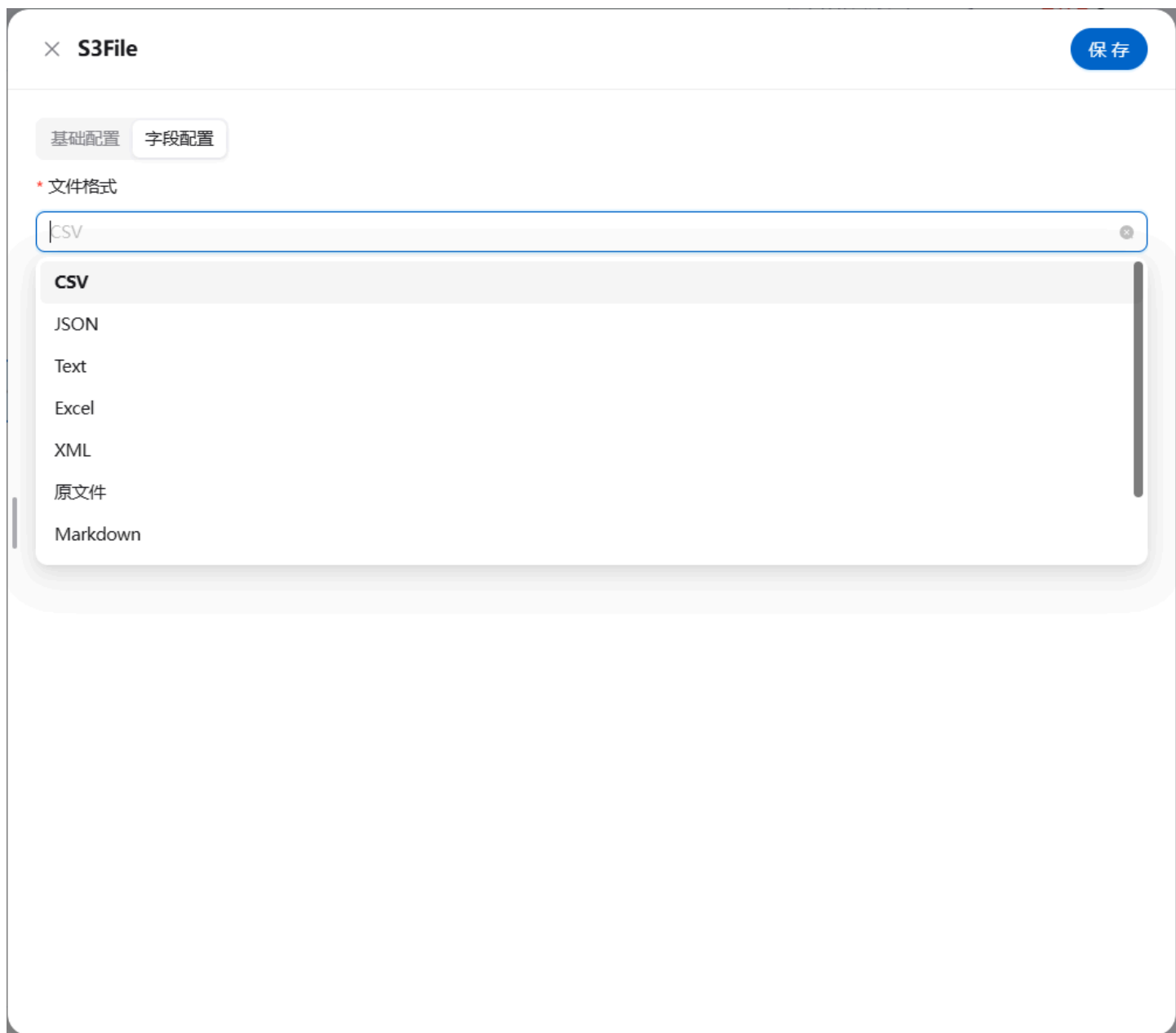


图 118 S3File输入—字段配置

说明：

- 选择CSV、JSON、Text、Excel、XML、原文件或Markdown等适用格式；S3File还支持Parquet、ORC。按格式填写分隔符和字符编码。
- 原文件传输保留文件内容，不把文件解码为订单字段；结构化转换应选择可解析的对应格式。

### 3.5.2.2 转换算子配置

#### FieldMapper

×

FieldMapper

①

保存

基础配置

字段配置

\* 算子名称

FieldMapper

11 / 64

图 119 转换算子通用基础配置

说明：

- 基础配置中确认算子名称和使用说明，连接直接上游后配置字段规则。

×

FieldMapper

①

保存

基础配置

字段配置

|   | 启用                                  | 源字段         | 类型                | 输出字段      |
|---|-------------------------------------|-------------|-------------------|-----------|
| ⋮ | <input checked="" type="checkbox"/> | actor_id    | SMALLINT UNSIGNED | actorId   |
| ⋮ | <input checked="" type="checkbox"/> | last_name   | VARCHAR           | lastName  |
| ⋮ | <input checked="" type="checkbox"/> | first_name  | VARCHAR           | firstName |
| ⋮ | <input type="checkbox"/>            | last_update | TIMESTAMP         |           |

> 生成配置预览

图 120 FieldMapper算子—字段配置

说明：

- 选择输出字段，调整顺序并填写重命名映射。
- 例如将order\_no输出为order\_number；被移除的字段不再提供给下游。

Filter

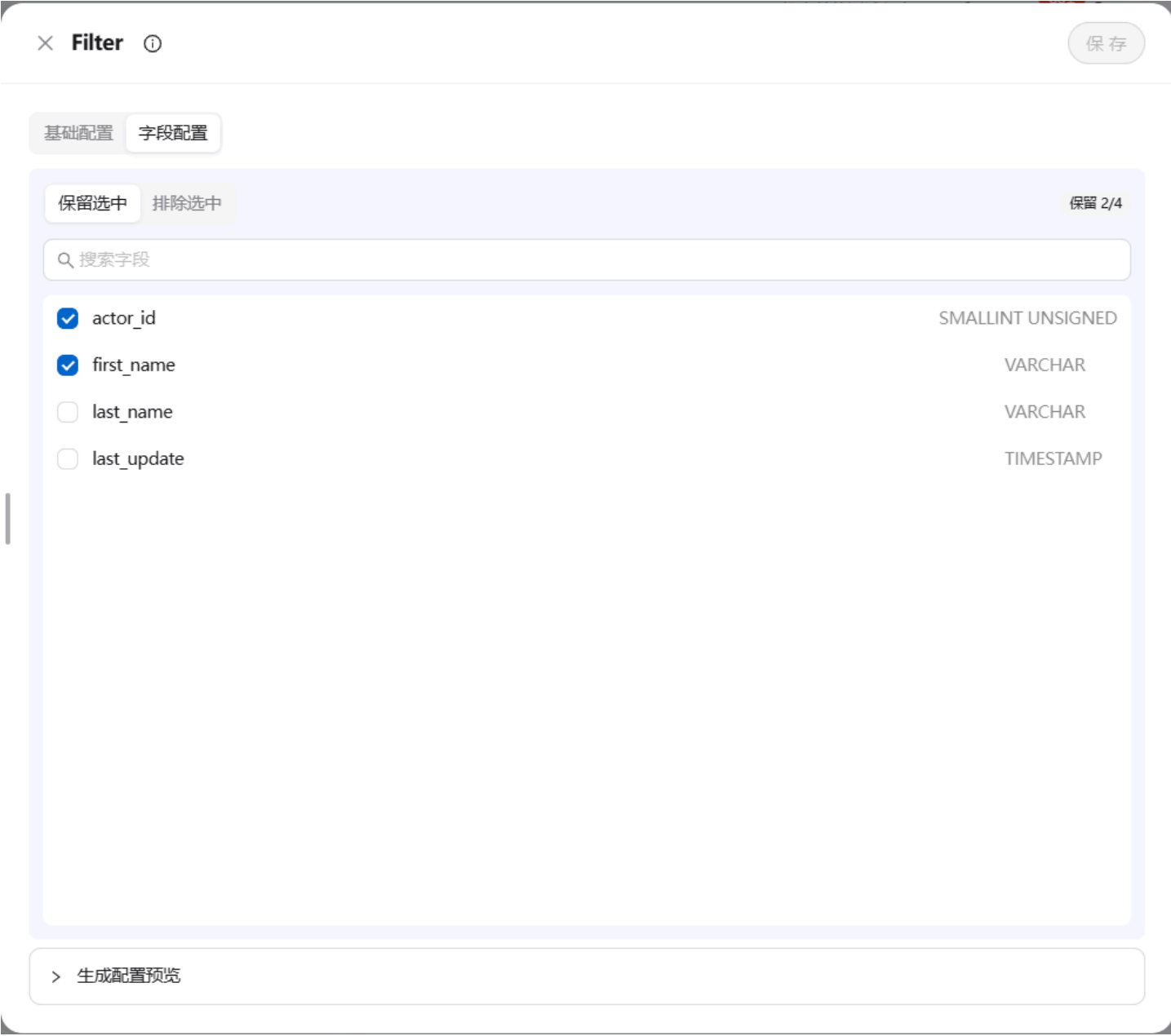


图 121 Filter算子—字段配置

说明：

- 选择要保留或排除的字段。
- 这里过滤的是字段，不是数据行；按订单金额过滤行应使用SQL转换。

Copy

×

Copy

ⓘ

保存

基础配置

字段配置

\*复制规则 ⓘ

源字段

输出字段

actor\_id · SMALLINT UNSIGNED

actor\_id01

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id02

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id03

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id04

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id05

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id06

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id07

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id08

✕

+

first\_name · VARCHAR

first\_name01

✕

+

first\_name · VARCHAR

first\_name02

✕

+

first\_name · VARCHAR

first\_name03

✕

+

first\_name · VARCHAR

first\_name04

✕

+

...

...

✕

+

> 生成配置预览

图 122 Copy算子一字段配置

说明：

- 配置源字段与一个或多个目标字段名称。
- 例如将waybill\_no复制为waybill\_no\_backup；源字段仍保留，新增字段与源值相同。

Replace

×

Replace

ⓘ

保存

基础配置

字段配置

\* 处理字段 ⓘ

txt · TEXT

▼

\* 匹配内容 ⓘ

abc

替换内容 ⓘ

X

使用正则表达式 ⓘ

☐

仅替换首个匹配 ⓘ

☐

> 生成配置预览

图 123 Replace算子—字段配置

#### 说明：

- 选择文本字段，填写匹配内容与替换内容，可启用正则或仅替换首个匹配。
- 例如将地址中的指定分隔符统一为逗号；替换内容留空表示删除匹配内容。

#### Split

×

Split

①

保存

基础配置

字段配置

\* 拆分字段 ②

raw\_value · TEXT

▼

\* 分隔符 ②

,

\* 输出字段 ②

part\_one × part\_two × part\_three ×

▼

> 生成配置预览

图 124 Split算子一字段配置

说明：

- 选择待拆分字段、分隔符及按顺序排列的输出字段。
- 例如将“华东|上海”拆为region\_name、city\_name，输出顺序应与原值一致。

RegexExtract

×

RegexExtract

①

保存

基础配置

字段配置

\* 源字段 ②

raw\_value · TEXT

▼

\* 正则表达式 ②

^([^\@]+)\.([^\@]+)\.([^\@]+)\$

\* 输出字段 ②

extract\_one × extract\_two × extract\_three ×

▼

默认值 ②

▼

> 生成配置预览

图 125 RegexExtract算子一字段配置

#### 说明：

- 填写源字段、正则捕获组、输出字段和需要的默认值。
- 捕获组与输出字段逐一对应；未匹配的记录按配置使用默认值。

#### JsonPath

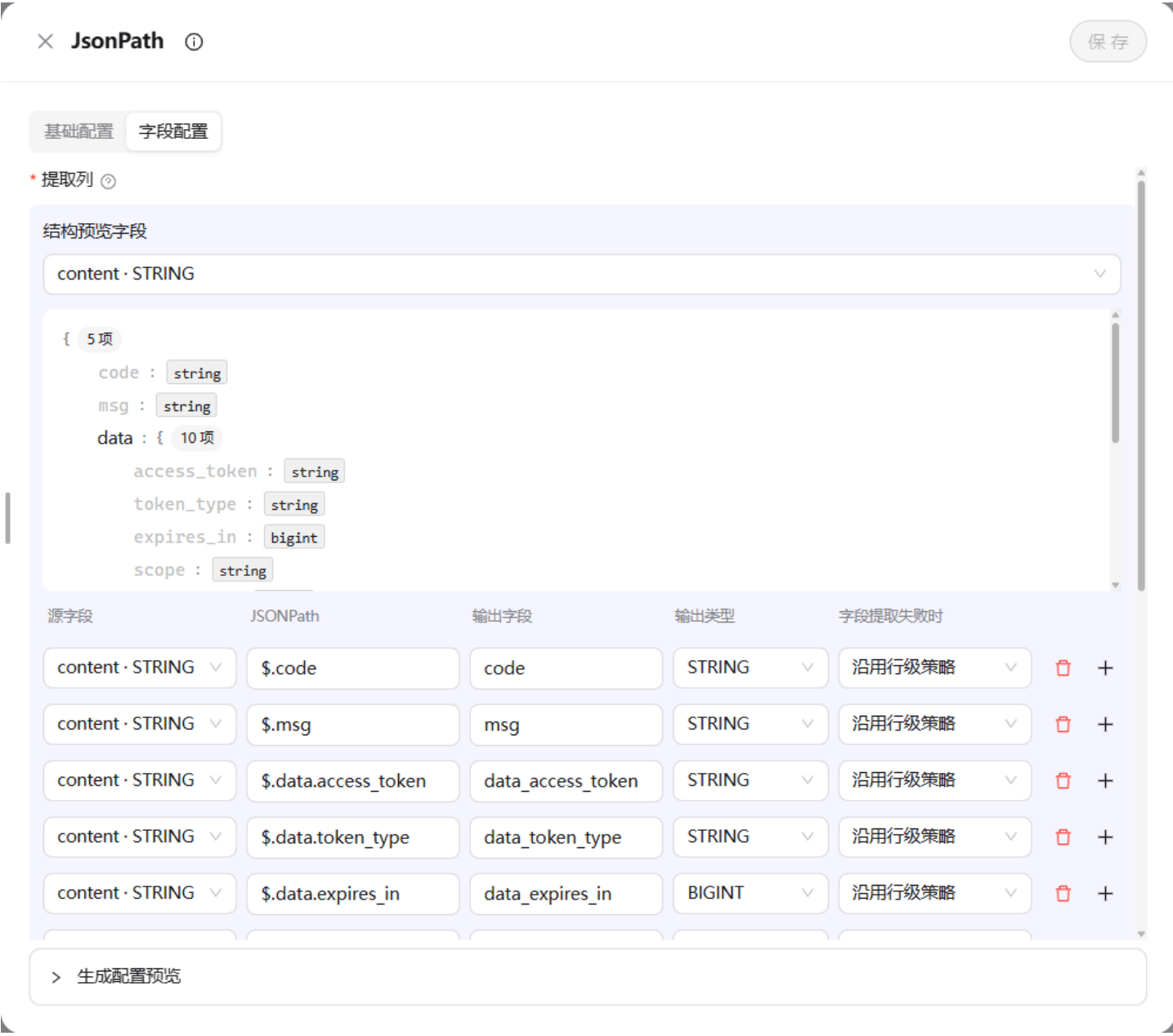


图 126 JsonPath算子一字段配置

说明：

- 逐列设置源字段、JSONPath、目标名称和类型，并选择错误时失败或跳过。
- 例如从payload中的\$.order.id提取order\_id；跳过错误行会影响最终记录数量。

SQL

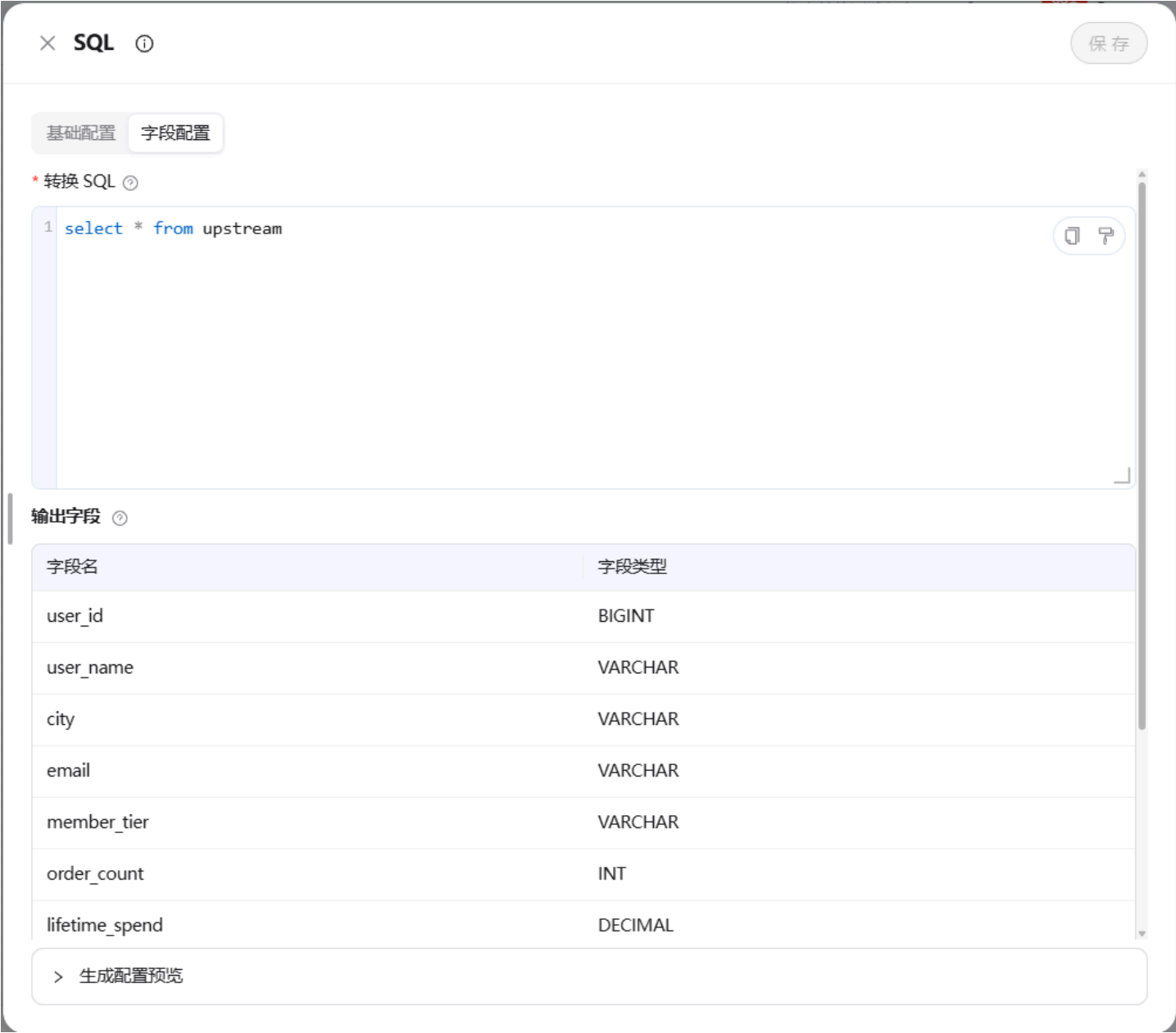


图 127 SQL转换算子—字段配置

- 说明：
- 填写针对直接上游的SELECT，可派生字段、转换类型和按条件过滤数据行；多表输入可设置作用表正则及单表覆盖规则。
  - 此转换不支持多表关联或聚合。需要物流日汇总或多表JOIN时，可在 workflow SQL 节点中使用数据库支持的SQL完成。

Metadata

×

Metadata

①

保存

基础配置

字段配置

\* 元数据字段

②

元数据类型

输出字段

Database

mysql\_database

✕

+

Table

mysql\_table

✕

+

RowKind

mysql\_rowkind

✕

+

> 生成配置预览

图 128 Metadata算子—字段配置

- 说明：
- 配置来源数据库、表名、事件类型、事件时间等元信息到普通字段的映射。
  - 例如输出source\_table追踪数据来源；元信息可用性取决于上游，不会自动生成来源未提供的真实事件时间。

RowKindExtractor

×

RowKindExtractor ①

保存

基础配置

字段配置

输出字段 ②

row\_kind\_multit

输出格式 ②

完整格式

> 生成配置预览

图 129 RowKindExtractor算子一字段配置

说明：

- 指定保存原变更类型的输出字段，并选择短格式或完整格式。
- 该算子同时把所有记录转为INSERT追加流；原删除事件可记为DELETE，但下游不再按删除事件执行。普通批量数据无需使用。

FilterRowKind

×

FilterRowKind • FILTER\_ROW\_KIND

①

保存

基础配置

字段配置

保留事件

排除事件

INSERT ×

▼

> 生成配置预览

图 130 FilterRowKind算子—字段配置

说明：

- 在保留与排除方式中二选一，选择INSERT、UPDATE\_BEFORE、UPDATE\_AFTER或DELETE。
- 例如排除DELETE后删除事件不传递下游，目标可能保留原记录，应符合业务要求。

TableFilter

×

TableFilter

ⓘ

保存

基础配置

字段配置

数据库匹配 ⓘ

\*

Schema 匹配 ⓘ

\*

表名匹配 ⓘ

actor\_\*

匹配模式 ⓘ

保留匹配表

> 生成配置预览

图 131 TableFilter算子一字段配置

说明：

- 填写数据库、Schema、表名正则，并选择保留或排除匹配表。
- 该操作限制输入表范围，不是对单表字段值进行过滤。

TableRename

×

TableRename • TABLE\_RENAME

①

保存

基础配置

字段配置

大小写转换 ②

小写

名称前缀 ②

名称后缀 ②

正则替换 ②

匹配表达式

替换内容

例如: ^ods\_(.\*)\$

例如: dwd\_\$1

🗑️

+

> 生成配置预览

图 132 TableRename算子一字段配置

说明：

- 配置表名大小写转换、前后缀或正则替换。
- 例如增加ods\_前缀形成目标表身份；不会直接重命名来源数据库中的物理表。

TableMerge

×

TableMerge • TABLE\_MERGE

①

保存

基础配置

字段配置

\* 来源表正则 ②

.\*

\* 目标表 ②

orders\_all

> 生成配置预览

图 133 TableMerge算子—字段配置

#### 说明：

- 填写来源表匹配正则与合并后的目标数据库、Schema、表名。
- 将结构兼容的分表汇入同一表身份，不执行多表JOIN；目标字段结构和唯一键需能够容纳合并后的数据。

#### HTTP调用

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

\* 节点名称

获取无人机飞行任务JOB数据14 / 64

\* 方法

POST

\* URL

http://192.168.31.210:8001/gateway/geoai-api-service/v1/job/list

Params

Auth

Headers

Body

参数名

值

Key

值

fx

🗑

+

出参

在线探查

{ 5 项

code : string

msg : string

data : { 2 项

total : bigint

records : [ 11 项 设为数据记录

id : bigint

jobId : string

☒ 分页读取

页码字段

起始页码

每页记录数

总页数

图 105 配置HTTP调用

说明：

- 基础配置填写URL、请求方法、请求头、参数和请求体。接口分页时启用页码分页，设置页码字段、起始页、每页条数及可用的总量信息。

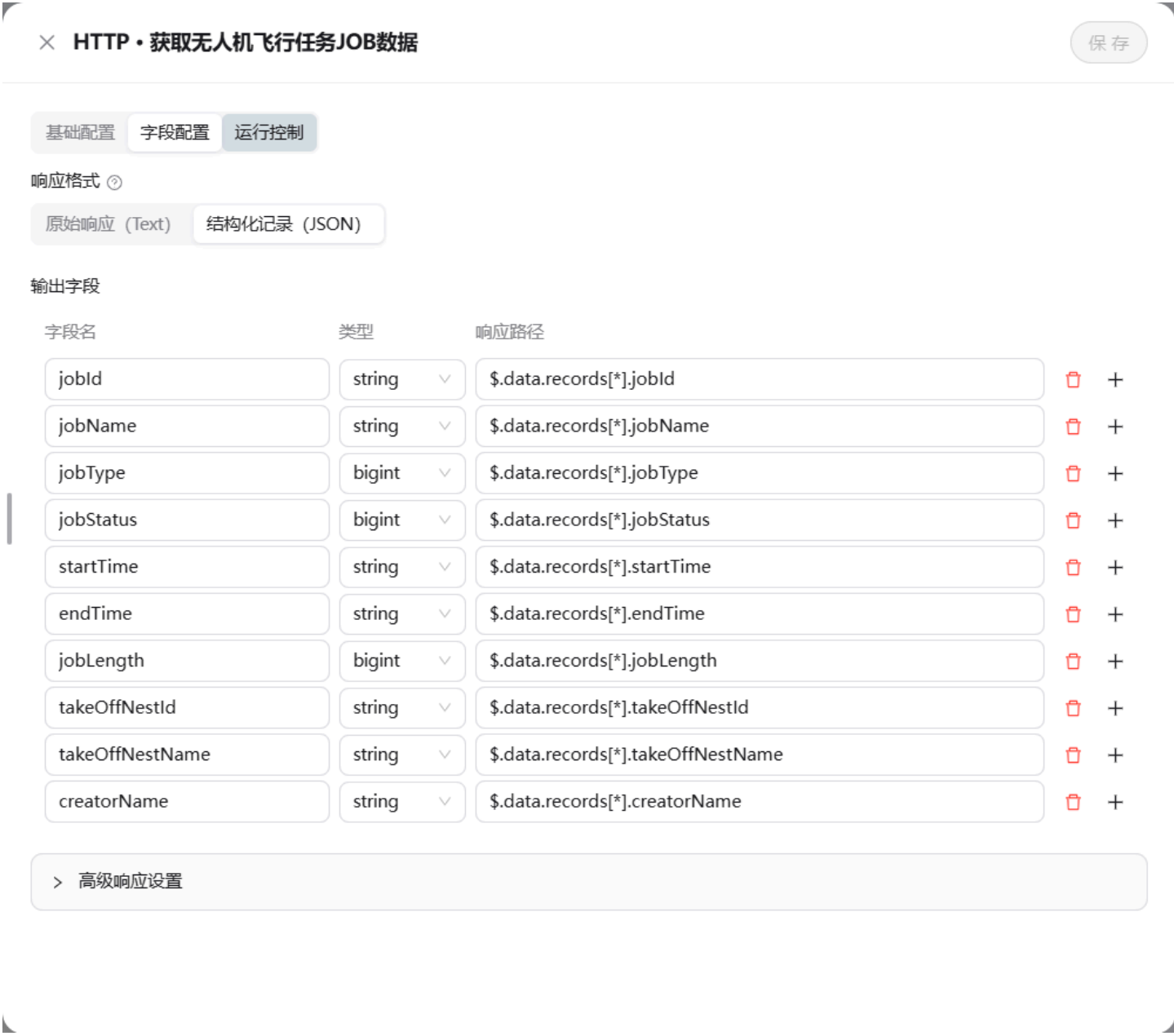


图 106 HTTP调用—字段配置

说明：

- 选择原始文本或结构化JSON解析；使用响应探查确定内容路径、数组位置和输出字段。
- 提取订单列表时选取实际列表数组，不把整个响应包装对象误作一条订单。

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

最大重试次数

3

重试间隔基准 (毫秒) ?

100

重试间隔上限 (毫秒) ?

10000

连接超时 (毫秒)

12000

读取超时 (毫秒)

60000

图 107 HTTP调用—运行控制

说明：

- 配置请求重试、重试间隔上限、连接和读取超时。接口限流或失败时结合返回信息调整，避免高并发持续重试。

LLM

×

LLM

ⓘ

保存

\* 算子名称

LLM

3 / 64

\* 供应商

DEEPSEEK

\* 接口地址

https://api.deepseek.com/chat/completions

API Key

.....

🔒

已保护；重新输入可替换

\* 模型

点击查询并选择，或手动输入模型名称

deepseek-flash

deepseek-v4-pro

自定义

\* 提示词

\* 输入字段

图 108 配置LLM

说明：

- 支持通用OpenAI协议，常见OPENAI、DEEPSEEK、DOUBAO、KIMIAI、ZHIPU、Ollama等多种协议模型，能够使用LLM算子借助调用大模型实现数据智能化处理。

### 3.5.2.3 输出算子配置

#### JDBC输出

×

JDBC • JDBC\_SINK

保存

基础配置

写入配置

表配置

\* 算子名称

JDBC\_SINK9 / 64

\* 目标数据源

XuguDB / 数据同步数仓DW4FLOW (192.168.31.213:5138/DEMOW)▼

图 134 配置JDBC输出

说明：

- 基础配置选择目标数据源与表，设置表结构和已有数据处理策略。自动生成或自定义建表SQL应与目标库类型一致。
- 启用Upsert时配置目标主键；清空、自定义SQL等可能改变既有数据，执行前确认范围。
- JDBC输出目标数据源通过平台统一维护的数据源连接，支持 Xugu、Oracle、Mysql、达梦、PostgreSQL、SQLServer、OpenGauss、人大金仓等多种类型数据源，实现已支持的集中式存储（如Mysql、PostgreSQL等）和分布式存储（如Xugu、达梦等）之间的批量数据采集与传输。不同存储类型采用对应的写入配置。

×

JDBC • JDBC\_SINK

保存

基础配置

写入配置

表配置

写入方式

自动生成

\* 目标表结构处理

表不存在时报错

\* 目标已有数据处理

追加数据

启用 Upsert

写入批大小

1000

目标连接检查超时 (秒)

30

批量写入重试次数

0

图 135 JDBC输出—写入配置

- 说明：
- 使用表格或画布映射来源与目标字段，检查字段类型及必需字段。由目标或平台生成的ID无需映射为上游字段。

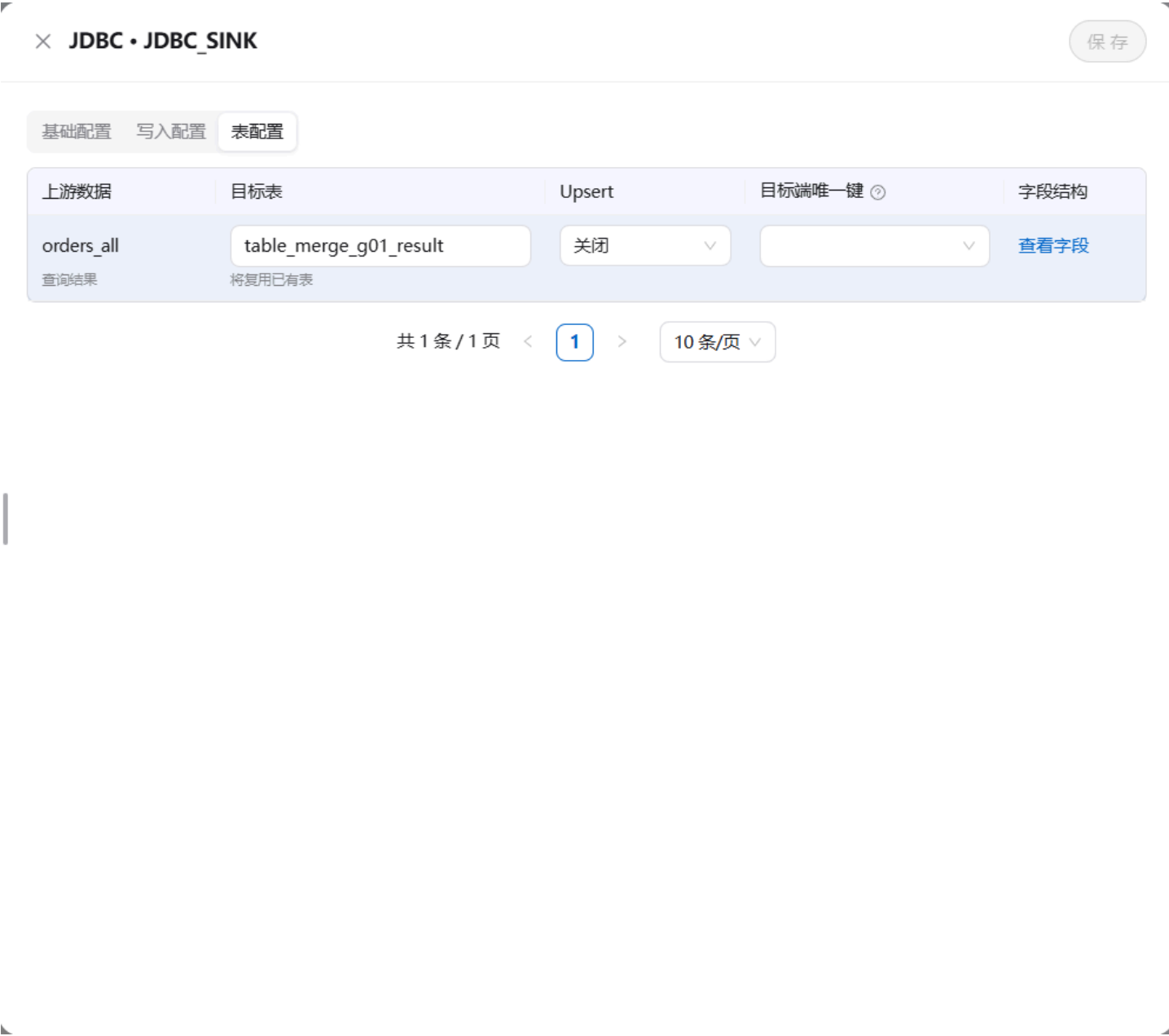


图 136 JDBC输出一表配置

- 说明：
- 设置目标表、更新策略和目标端唯一键，并配置字段映射。

Kafka输出

×

Kafka

保存

基础配置

字段配置

运行控制

\* 算子名称

Kafka5 / 64

\* 数据源 ⓘ

Kafka / orderdb-kafka-216 (192.168.31.216:9092)

\* Topic

orderdb-orders

分区键字段

id ×

图 137 配置Kafka输出

说明：

- 基础配置选择Kafka数据源、Topic与分区键，确认下游消费者使用的Topic。

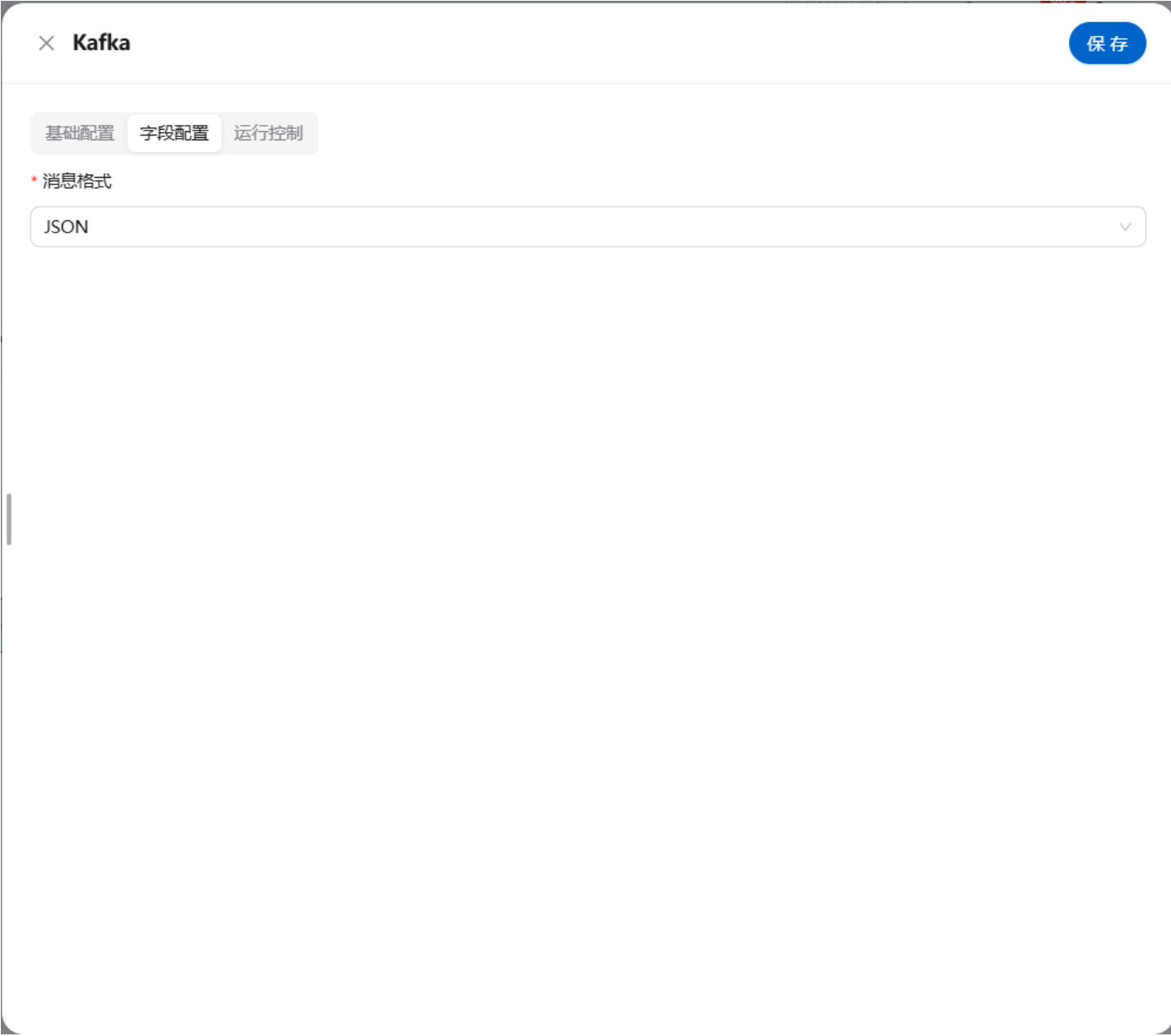


图 138 Kafka输出一字段配置

说明：

- 选择消息格式，并按所选格式配置分隔符等参数。输出结构应与消费者约定一致。

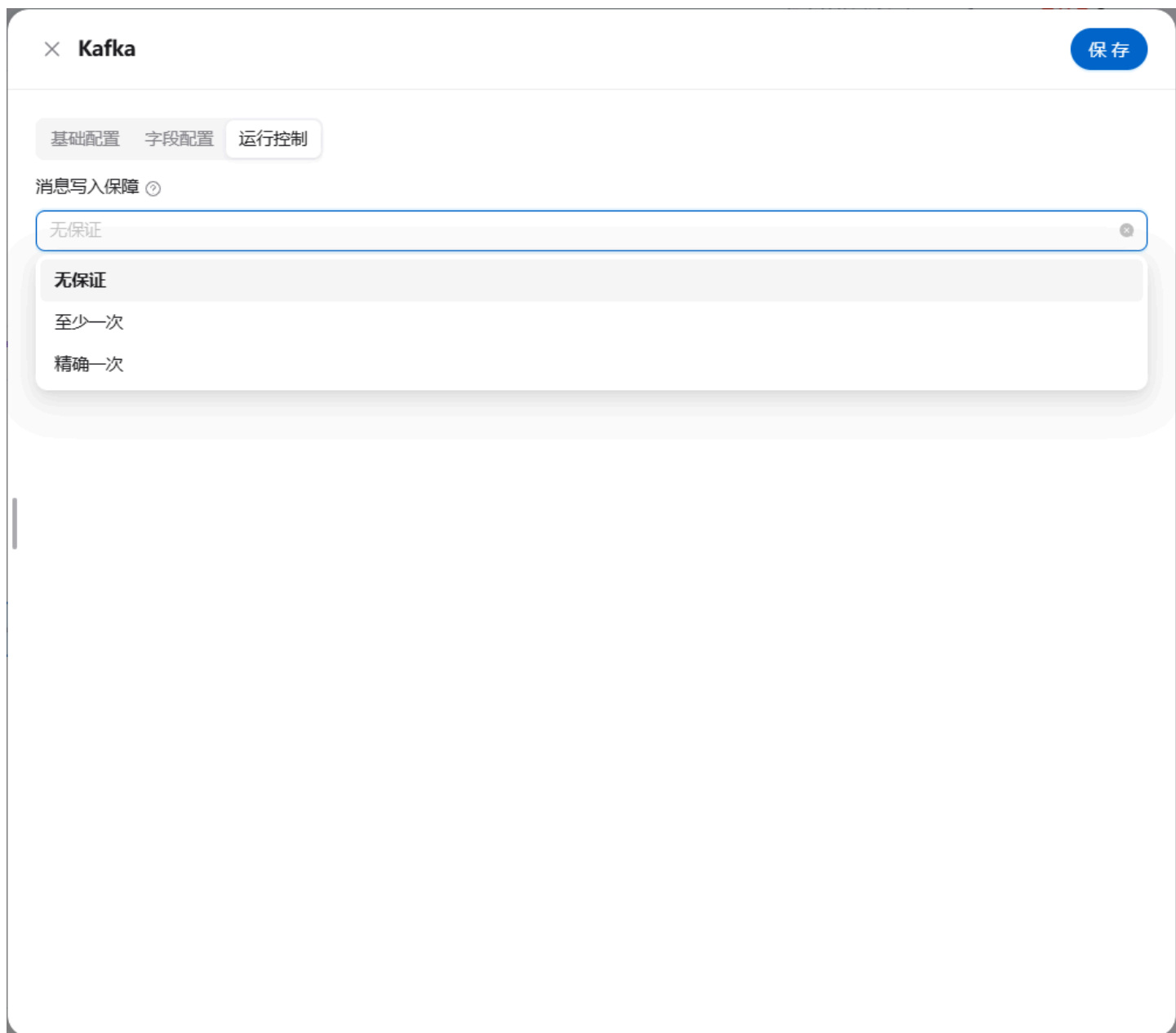


图 139 Kafka输出—运行控制

说明：

- 选择交付保证。使用精确一次时按要求填写唯一的事务前缀，并结合Checkpoint设置。
- 该选项约束对应Kafka写入，不保证工作流其他外部接口同时具备相同事务能力。

HTTP输出

×

HTTP

保存

基础配置

运行控制

\* 算子名称

HTTP4 / 64

\* 数据源

数据源类型 / 数据源

\* Webhook URL

https://api.example.com/orders

请求头

| 参数名 | 值 |
|-----|---|
| Key | 值 |

数组批量发送

☐

请求批大小

1

图 140 配置HTTP输出

- 说明：
- 基础配置选择HTTP数据源或填写Webhook地址、请求头，设置是否数组批量发送及批大小。
  - 按接口接收约定发送JSON；写入接口需考虑幂等键，避免重试重复创建物流单据。

×

HTTP

保存

基础配置

运行控制

最大重试次数

3

连接超时 (毫秒)

12000

读取超时 (毫秒)

60000

图 141 HTTP输出—运行控制

- 说明：
- 配置最大重试次数、连接和读取超时。请求超时不一定表示服务端未处理，应结合接口幂等机制核对结果。

Redis输出

×

Redis

保存

基础配置

字段配置

运行控制

\* 算子名称

Redis5 / 64

\* 数据源 

Redis / Redis数据源dg\_sync\_6380 (192.168.31.216:6380/0) 

\* 键表达式

orders:\${id}

\* 数据类型

String

过期秒数 

-1

图 142 配置Redis输出

说明：

- 基础配置选择数据源、键表达式、数据类型及过期时间；过期时间为-1表示不自动过期。

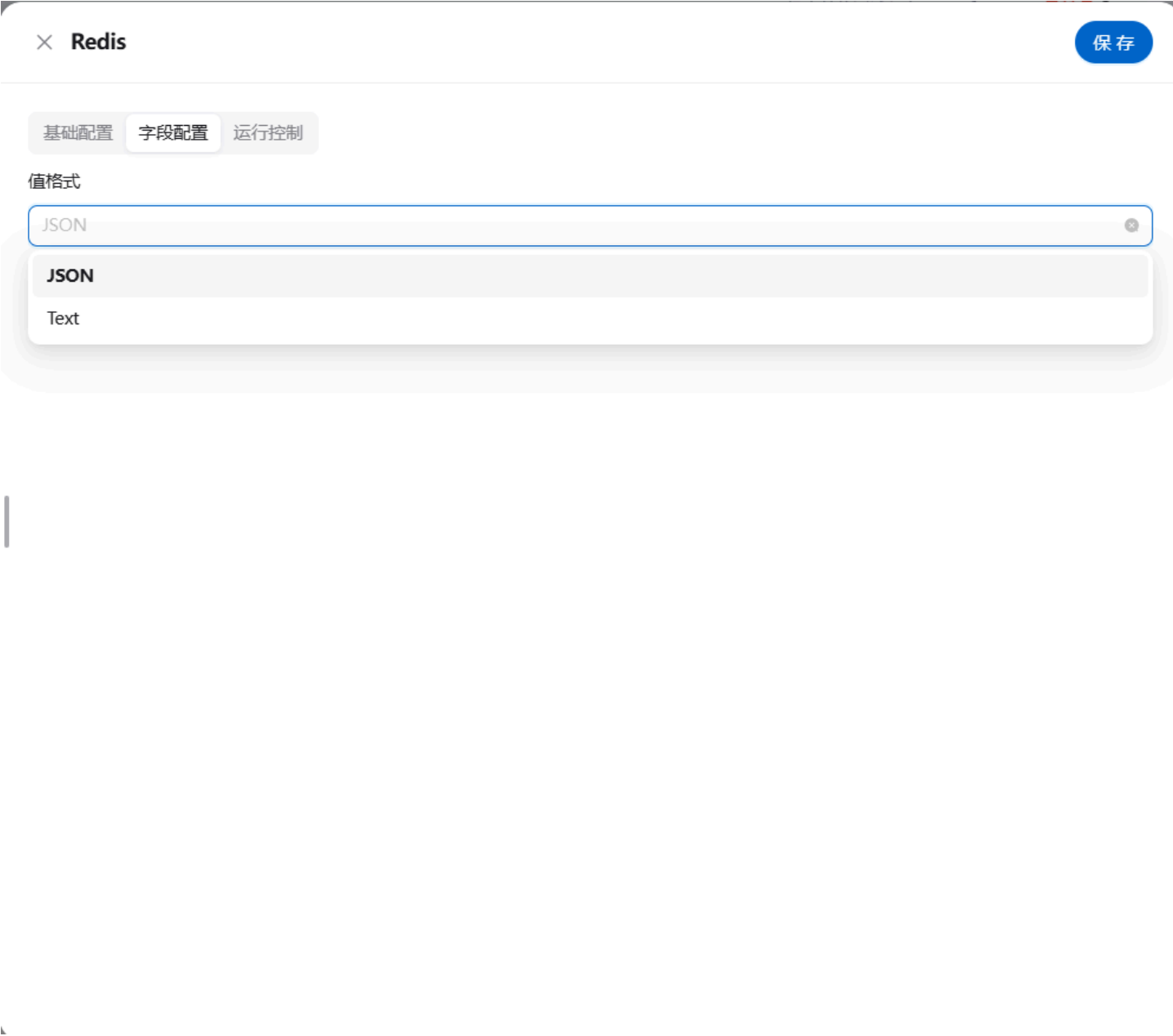


图 143 Redis输出一字段配置

说明：

- 选择JSON或Text等输出格式，检查生成值与后续读取方式一致。

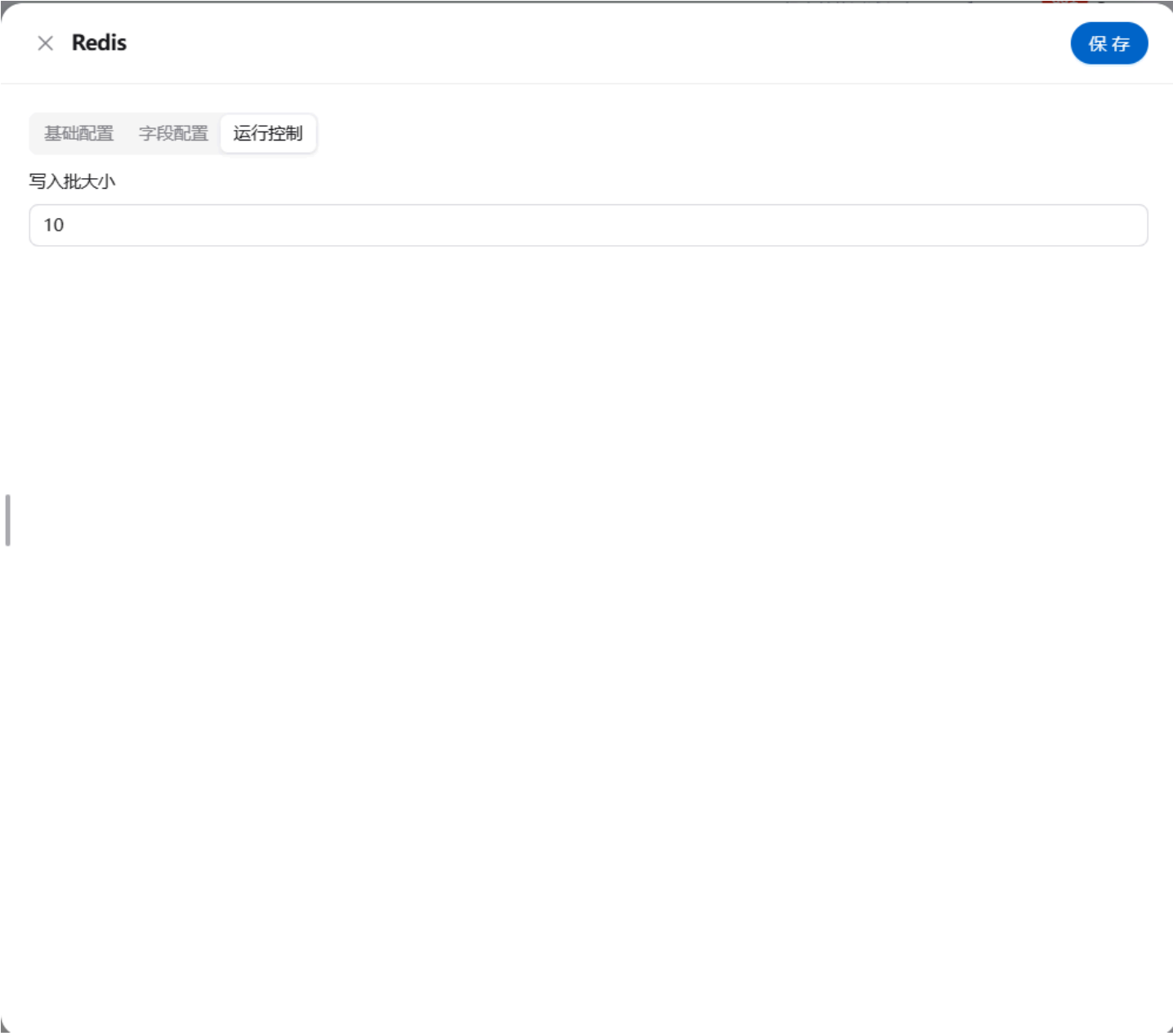


图 144 Redis输出—运行控制

说明：

- 配置批量写入大小。键表达式应避免不同订单意外使用同一个键而互相覆盖。

Elasticsearch输出

×

Elasticsearch

保存

基础配置

运行控制

\* 算子名称

Elasticsearch13 / 64

\* 数据源 

数据源类型 / 数据源

\* 索引

orders

文档主键字段

请先选择目标数据源

索引处理

索引不存在时创建

已有数据处理

追加数据

图 145 配置Elasticsearch输出

说明：

- 基础配置选择数据源、索引及文档主键，设置索引结构和已有数据处理。
- 文档主键用于识别目标文档，应结合来源唯一字段配置，避免覆盖无关文档。

×

Elasticsearch

保存

基础配置

运行控制

批量写入条数

10

最大重试次数

3

图 146 Elasticsearch输出—运行控制

说明：

- 设置批大小与重试次数，结合目标索引写入能力控制压力。

FtpFile输出

×

FtpFile

保存

基础配置

字段配置

\* 算子名称

FtpFile7 / 64

\* 数据源 ⓘ

FTP / FTP-216 (192.168.31.216:21//)▼

\* 输出路径

/data

中转目录 ⓘ

/tmp/datagolden-transfer

自定义文件名

☐

文件名表达式 ⓘ

part-\${now}-\${uuid}

图 147 配置FtpFile输出

说明：

- 基础配置选择数据源与输出路径，按需启用自定义文件名并填写表达式；未启用时使用默认命名。中转目录需在执行环境可写。
- 重复运行前核对路径和文件命名，避免覆盖需保留的历史文件。

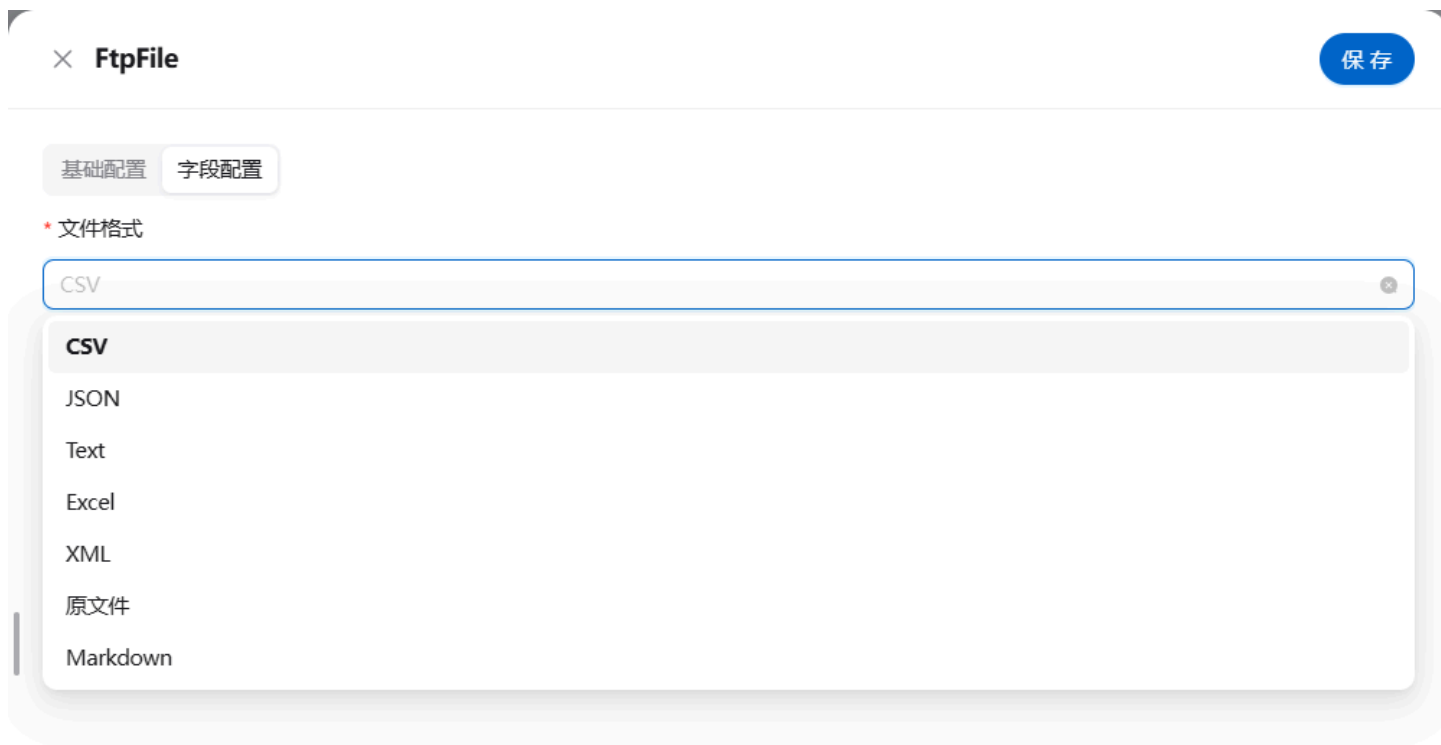


图 148 FtpFile输出一字段配置

说明：

- 选择输出格式及对应分隔符、字符编码，确保接收方能够按约定读取。原文件方式不进行结构化字段计算。

SftpFile输出

×

SftpFile

保存

基础配置

字段配置

\* 算子名称

SftpFile8 / 64

\* 数据源

SFTP / SFTP-216 (192.168.31.216:2222//)

\* 输出路径

/data

中转目录

/tmp/datagolden-transfer

自定义文件名

文件名表达式

part-\${now}-\${uuid}

图 149 配置SftpFile输出

说明：

- 基础配置选择数据源与输出路径，按需启用自定义文件名并填写表达式；未启用时使用默认命名。中转目录需在执行环境可写。
- 重复运行前核对路径和文件命名，避免覆盖需保留的历史文件。

×

SftpFile

保存

基础配置

字段配置

\* 文件格式

CSV

字段分隔符 

,

字符编码

UTF-8

图 150 SftpFile输出一字段配置

- 说明：
- 选择输出格式及对应分隔符、字符编码，确保接收方能够按约定读取。原文件方式不进行结构化字段计算。

×

S3File

保存

基础配置

字段配置

\* 算子名称

S3File6 / 64

\* 数据源 ?

MinIO / orderdb (192.168.31.216:9000/orderdb)▼

\* 输出路径

order\_info

自定义文件名

☐

文件名表达式 ?

part-\${now}-\${uuid}

图 151 配置S3File输出

说明：

- 基础配置选择数据源与输出路径，按需启用自定义文件名并填写表达式；未启用时使用默认命名。
- 重复运行前核对路径和文件命名，避免覆盖需保留的历史文件。

×

S3File

保存

基础配置

字段配置

\* 文件格式

CSV

字段分隔符 ?

,

字符编码

UTF-8

图 152 S3File输出一字段配置

说明：

- 选择输出格式及对应分隔符、字符编码，确保接收方能够按约定读取。原文件方式不进行结构化字段计算。

Console输出

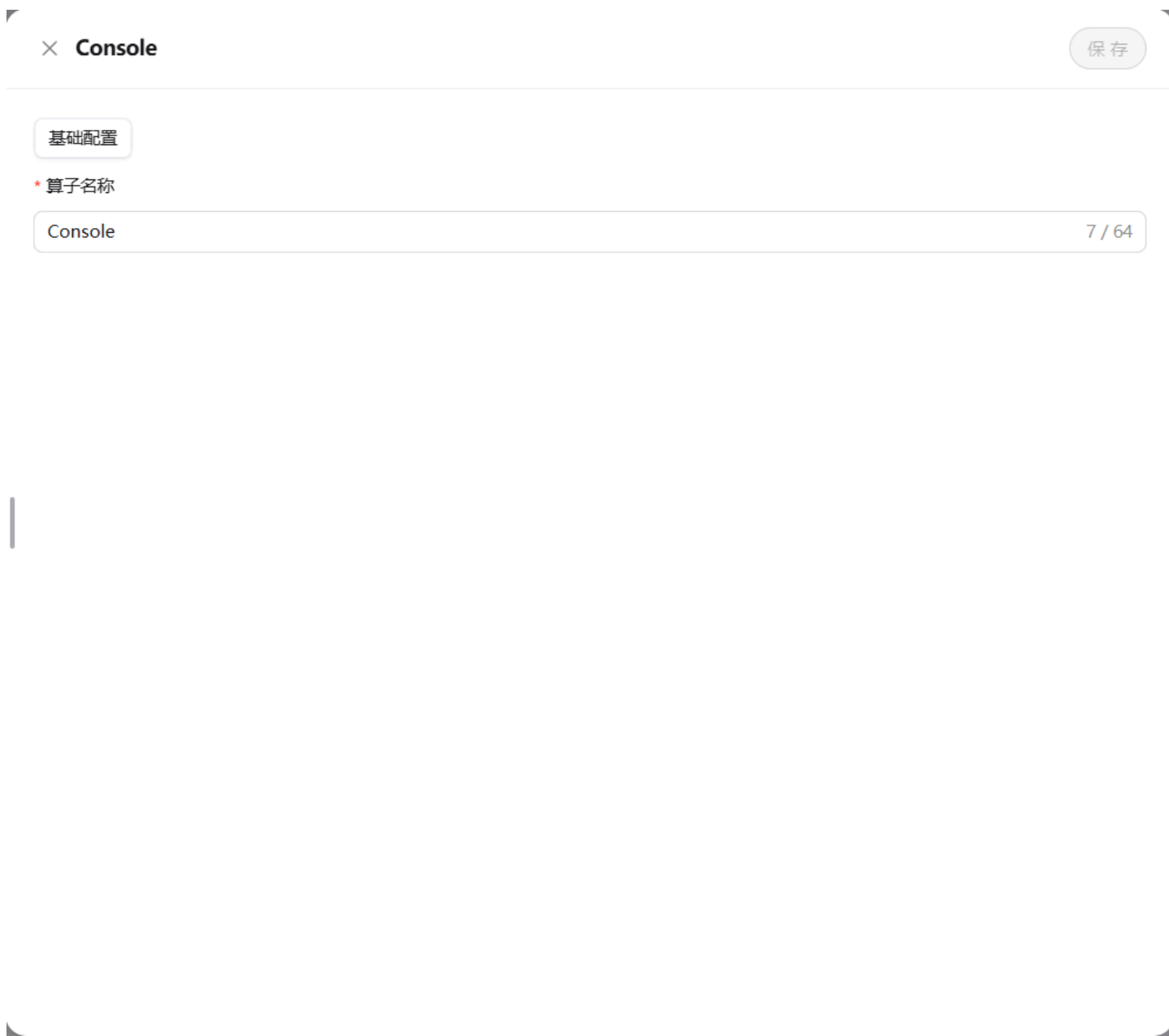


图 153 配置Console输出

说明：

- 添加Console输出连接上游，用于查看处理结果和辅助调试。
- 控制输出的数据范围，避免在日志中暴露敏感业务数据。Console输出不等于数据已写入业务目标库。

#### 3.5.2.4 自定义 API 调用与参数传递

画布工作流支持将自定义 API 调用作为任务节点纳入流程编排。用户可配置接口地址、请求方式、请求头和请求参数，通过节点连线设置执行顺序及上下游依赖，实现 API 调用与其他数据处理任务的协同执行。

数据开发

在工作目录中维护工作流草稿，校验 DAG 开发起统一发布。

离线任务实时任务

工作目录

搜索目录或工作流

单独立Python(复杂脚本)

HTTP算子5

冒烟测试1

基础类型请求3

复杂组合工作流1

GEOAI接口获取数据并...26

条件算子26

单条件4

组内AND/OR8

组间AND/OR8

混合嵌套4

基础验证2

分支算子20

分支判断基础场景9

参数验证6

基础验证5

失败策略0

导入新建工作流

选择算子

数据管道

SQL

Shell

Python

HTTP

Conditions

Switch

Dependent

HTTP GEOAI 获取 access\_token

数据管道 JOB信息获取入库

数据管道 CLUE信息获取入库

← JOB信息获取入库/所属工作流: GEOAI接口获取数据并入库 副本

选择算子

搜索算子

输入

JDBC

Kafka

HTTP

Redis

Elasticsearch

FtpFile

SftpFile

S3File

转换

HTTP 调用

FieldMapper

Filter

Copy

HTTP 获取无人机飞行任务JOB数据

HTTP 调用 获取无人机飞行任务JOB明...

JDBC 无人机飞行任务数据入库

API调用支持参数定义与传递：可通过定义全局变量，实现HTTP算子之间的参数值传递；在HTTP算子工作流中，可定义局部的工作流参数，实现工作流内部对变量参数使用的需求

新增变量

变量名称 请输入名称

更新时间

| 变量名称     | 编码       | 状态   | 变量类型 | 变量值 | 变量格式 | 偏移量 | 更新时间  |
|----------|----------|------|------|-----|------|-----|-------|
| test_oa  | test_oa  | ● 上线 | 普通变量 | 01  | -    | -   | 2026- |
| test1    | test1    |      |      |     |      |     | 2026- |
| DATA_... | DATA_... |      |      |     |      |     | 2025- |

新增全局变量

\* 变量名称:

请输入名称0 / 30

\* 编码:

请输入编码0 / 30

\* 所属目录:

test1

\* 变量类型:

☒ 普通变量

☐ 日期时间

\* 变量值:

请输入变量值0 / 512

描述:

请输入0 / 100

取消

确定

× workflow配置

保存

基本信息

workflow参数

| 参数名           | 值来源 | 值                              |   |   |
|---------------|-----|--------------------------------|---|---|
| client_id     | 自定义 | b2dcbf8528ae4c8ba1b3be296d234e | 🗑 | + |
| client_secret | 自定义 | MGRjMzhjODk1N453NDQ4MzgzMj     | 🗑 | + |

配置参数值

函数

concat

01

常量

Bearer

02

当前算子入参

access\_token

取消

应用参数值

3.5.3 实时 workflow 管道算子配置

3.5.2.1 输入算子配置

JDBC 输入

JDBC • JDBC1

保存

基础配置

读取配置

\* 算子名称

JDBC1

5 / 64

\* 数据源

MySQL / mysql\_mfg\_cdc\_test (192.168.31.213:3306/mfg\_cdc\_test)

数据库级已就绪

Q 搜索已选来源表

已添加3 张表 · 单输入最多 50 张

| 来源表                                                                                                    | 操作          |
|--------------------------------------------------------------------------------------------------------|-------------|
| <div><div>purchase_returns</div><div>采购退货：采购退货记录，记录关联到货明细、退货数量、原因和日期。汽车零部件制造业仿真CDC测试数据。</div></div>    | <div></div> |
| <div><div>purchase_orders</div><div>采购订单：采购订单表头，记录供应商、下单日期、交期、币种和明细合计金额。汽车零部件制造业仿真CDC测试数据。</div></div> | <div></div> |
| <div><div>purchase_order_items</div><div>采购明细：采购订单明细，记录订购物料、数量、单价及行金额。汽车零部件制造业仿真CDC测试数据。</div></div>   | <div></div> |

共 3 条 / 1 页

1

10 条/页

×

JDBC • JDBC1

保存

基础配置

读取配置

\* 首次启动方式 ⓘ

全量后增量：先同步现有数据，再持续读取变更

快照分片大小（行） ⓘ

8096

快照单次抓取上限（行） ⓘ

1024

连接超时（毫秒） ⓘ

30000

连接池大小 ⓘ

20

连接重试次数 ⓘ

3

图 100配置JDBC输入

- 说明：
- 基础配置中选择数据库数据源和来源表。查询SQL可留空以读取表字段，也可填写SELECT限定字段、WHERE过滤行，或使用数据库支持的查询。
  - 来源表用于确定表身份和元数据；填写查询SQL后以查询结果作为下游输入，注意字段名与类型。

Kafka输入

×

Kafka

保存

基础配置

字段配置

运行控制

\* 算子名称

Kafka

5 / 64

\* 数据源

Kafka / orderdb-kafka-216 (192.168.31.216:9092)

▼

\* Topic

orderdb-orders

▼

消费组

dg.pipeline.37006260150ddc2b4d874fdb1c2a471

图 102配置Kafka输入

说明：

- 基础配置选择Kafka数据源与单个Topic。消费组由系统生成并在任务内保持稳定。

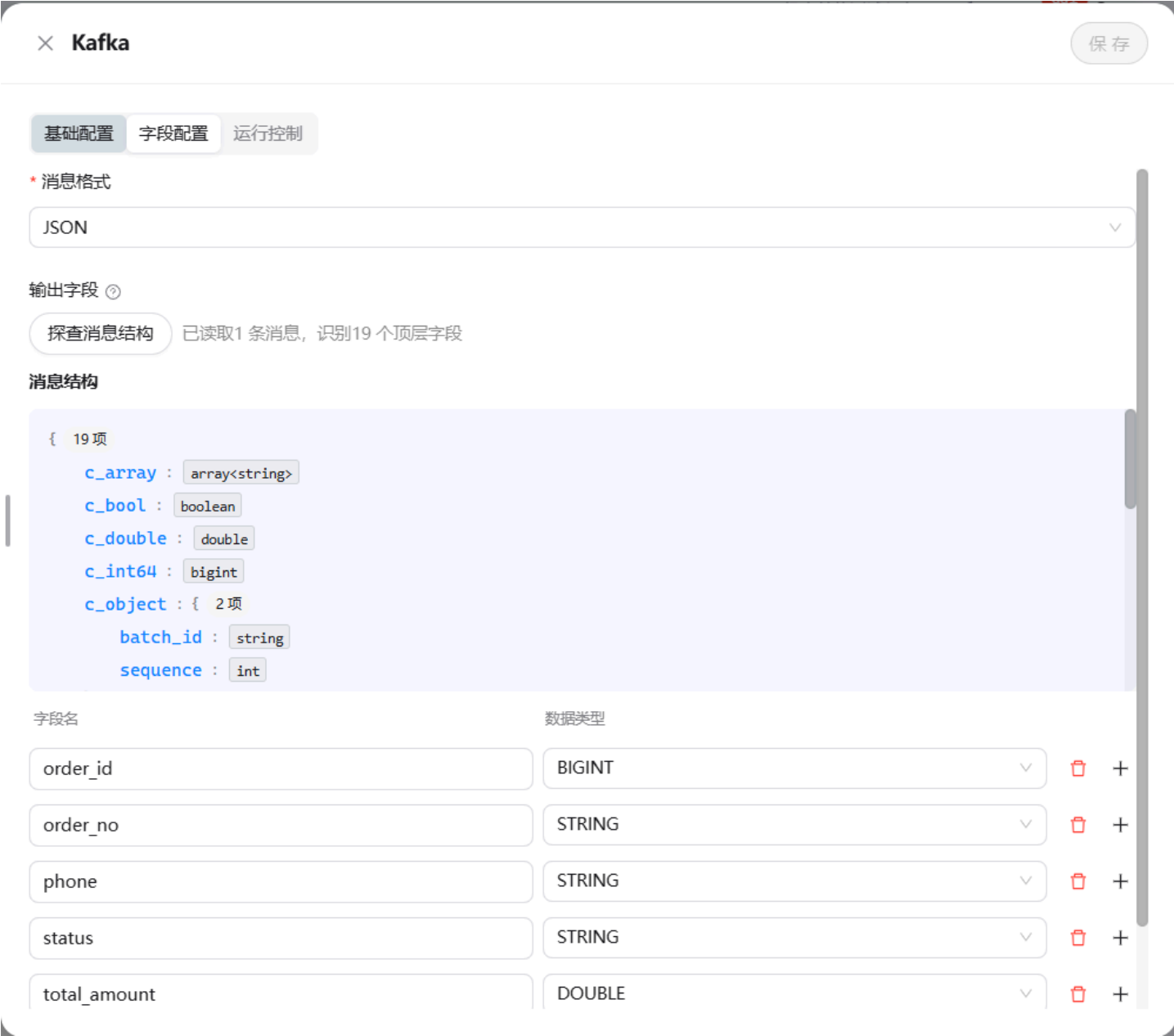


图 103Kafka输入一字段配置

- 说明：
- 选择JSON或Text。JSON可探查消息结构并生成输出字段，也可手工维护字段；Text按需要设置分隔符。
  - 输出字段应与实际消息结构一致；探查样例不能保证所有消息均符合该结构。

×

Kafka

保存

基础配置

字段配置

运行控制

\* 起始位置

最早位置

Checkpoint 完成后提交 Kafka 位点

拉取超时 (毫秒)

10000

图 104 Kafka输入—运行控制

说明：

- 选择消费组已提交位置、最早位置或最新位置，设置拉取超时。启用Checkpoint提交位点时，在检查点完成后提交消费进度。
- 从最新位置开始不会读取此前所有积压消息；重跑前确认消费组位置和恢复策略。

### 3.5.2.2 转换算子配置

#### FieldMapper

×

FieldMapper

①

保存

基础配置

字段配置

\* 算子名称

FieldMapper

11 / 64

图 119 转换算子通用基础配置

说明：

- 基础配置中确认算子名称和使用说明，连接直接上游后配置字段规则。

×

FieldMapper

①

保存

基础配置

字段配置

|   | 启用                                  | 源字段         | 类型                | 输出字段      |
|---|-------------------------------------|-------------|-------------------|-----------|
| ⋮ | <input checked="" type="checkbox"/> | actor_id    | SMALLINT UNSIGNED | actorId   |
| ⋮ | <input checked="" type="checkbox"/> | last_name   | VARCHAR           | lastName  |
| ⋮ | <input checked="" type="checkbox"/> | first_name  | VARCHAR           | firstName |
| ⋮ | <input type="checkbox"/>            | last_update | TIMESTAMP         |           |

> 生成配置预览

图 120 FieldMapper算子—字段配置

说明：

- 选择输出字段，调整顺序并填写重命名映射。
- 例如将order\_no输出为order\_number；被移除的字段不再提供给下游。

Filter

×

Filter

ⓘ

保存

基础配置

字段配置

保留选中

排除选中

保留 2/4

Q 搜索字段

☒ actor\_id

SMALLINT UNSIGNED

☒ first\_name

VARCHAR

☐ last\_name

VARCHAR

☐ last\_update

TIMESTAMP

> 生成配置预览

图 121 Filter算子—字段配置

说明：

- 选择要保留或排除的字段。
- 这里过滤的是字段，不是数据行；按订单金额过滤行应使用SQL转换。

Copy

×

Copy ⓘ

保存

基础配置

字段配置

\* 复制规则 ⓘ

源字段

输出字段

actor\_id · SMALLINT UNSIGNED

actor\_id01

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id02

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id03

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id04

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id05

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id06

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id07

✕

+

actor\_id · SMALLINT UNSIGNED

actor\_id08

✕

+

first\_name · VARCHAR

first\_name01

✕

+

first\_name · VARCHAR

first\_name02

✕

+

first\_name · VARCHAR

first\_name03

✕

+

first\_name · VARCHAR

first\_name04

✕

+

...

...

✕

+

> 生成配置预览

图 122 Copy算子一字段配置

说明：

- 配置源字段与一个或多个目标字段名称。
- 例如将waybill\_no复制为waybill\_no\_backup；源字段仍保留，新增字段与源值相同。

Replace

×

Replace

ⓘ

保存

基础配置

字段配置

\* 处理字段 ⓘ

txt · TEXT

▼

\* 匹配内容 ⓘ

abc

替换内容 ⓘ

X

使用正则表达式 ⓘ

☐

仅替换首个匹配 ⓘ

☐

> 生成配置预览

图 123 Replace算子—字段配置

#### 说明：

- 选择文本字段，填写匹配内容与替换内容，可启用正则或仅替换首个匹配。
- 例如将地址中的指定分隔符统一为逗号；替换内容留空表示删除匹配内容。

#### Split

×

Split

①

保存

基础配置

字段配置

\* 拆分字段 ②

raw\_value · TEXT

▼

\* 分隔符 ②

,

\* 输出字段 ②

part\_one × part\_two × part\_three ×

▼

> 生成配置预览

图 124 Split算子一字段配置

说明：

- 选择待拆分字段、分隔符及按顺序排列的输出字段。
- 例如将“华东|上海”拆为region\_name、city\_name，输出顺序应与原值一致。

RegexExtract

×

RegexExtract

①

保存

基础配置

字段配置

\* 源字段 ②

raw\_value · TEXT

▼

\* 正则表达式 ②

^([^\@]+)\@([^\.]+)\.(.+)\$

\* 输出字段 ②

extract\_one × extract\_two × extract\_three ×

▼

默认值 ②

▼

> 生成配置预览

图 125 RegexExtract算子一字段配置

#### 说明：

- 填写源字段、正则捕获组、输出字段和需要的默认值。
- 捕获组与输出字段逐一对应；未匹配的记录按配置使用默认值。

#### JsonPath

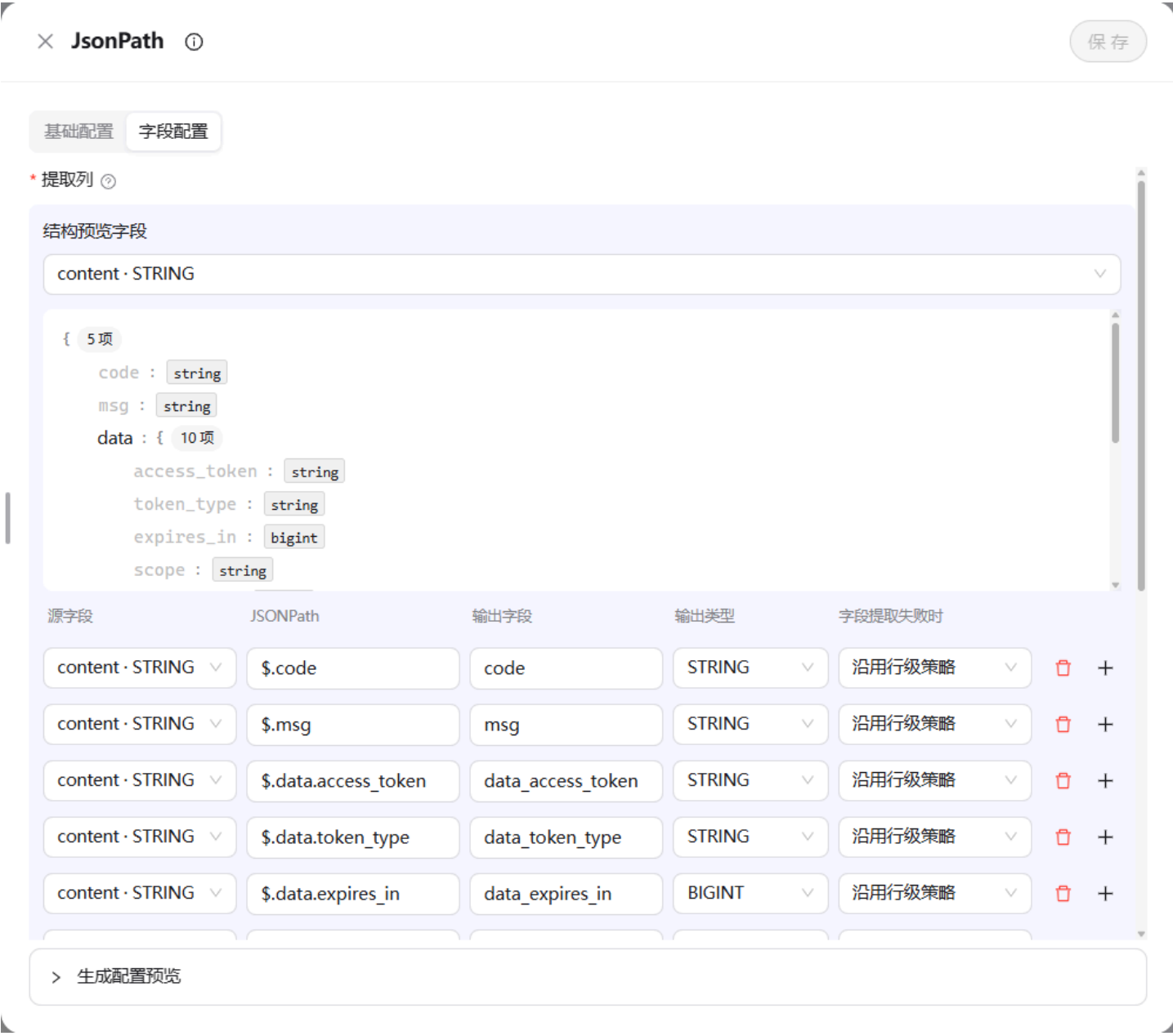


图 126 JsonPath算子一字段配置

说明：

- 逐列设置源字段、JSONPath、目标名称和类型，并选择错误时失败或跳过。
- 例如从payload中的\$.order.id提取order\_id；跳过错误行会影响最终记录数量。

SQL

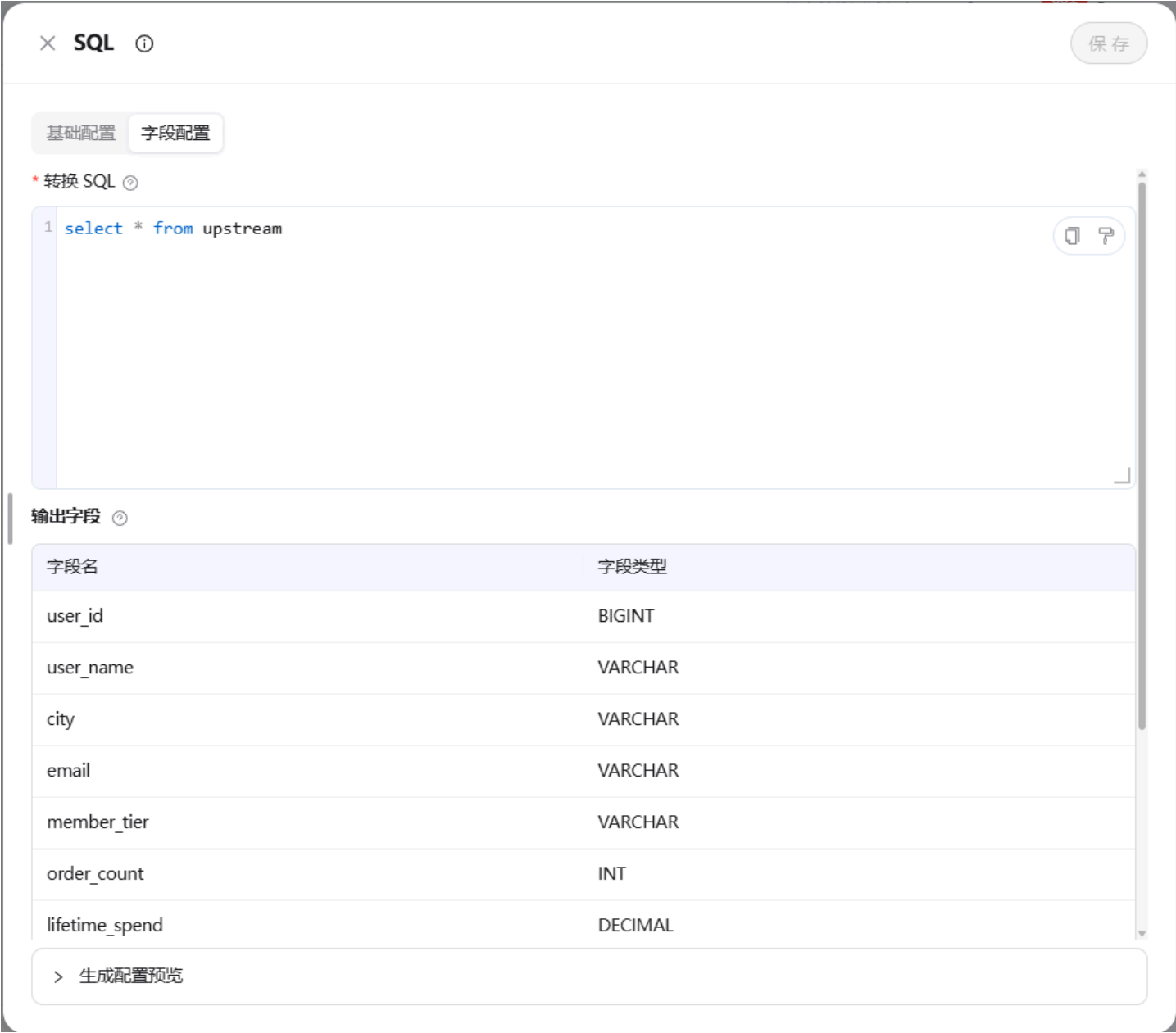


图 127 SQL转换算子—字段配置

- 说明：
- 填写针对直接上游的SELECT，可派生字段、转换类型和按条件过滤数据行；多表输入可设置作用表正则及单表覆盖规则。
  - 此转换不支持多表关联或聚合。需要物流日汇总或多表JOIN时，可在 workflow SQL 节点中使用数据库支持的SQL完成。

Metadata

×

Metadata

①

保存

基础配置

字段配置

\* 元数据字段

?

元数据类型

输出字段

Database

mysql\_database

✕

+

Table

mysql\_table

✕

+

RowKind

mysql\_rowkind

✕

+

>

生成配置预览

图 128 Metadata算子—字段配置

- 说明：
- 配置来源数据库、表名、事件类型、事件时间等元信息到普通字段的映射。
  - 例如输出source\_table追踪数据来源；元信息可用性取决于上游，不会自动生成来源未提供的真实事件时间。

RowKindExtractor

×

RowKindExtractor ①

保存

基础配置

字段配置

输出字段 ②

row\_kind\_multit

输出格式 ②

完整格式

> 生成配置预览

图 129 RowKindExtractor算子一字段配置

说明：

- 指定保存原变更类型的输出字段，并选择短格式或完整格式。
- 该算子同时把所有记录转为INSERT追加流；原删除事件可记为DELETE，但下游不再按删除事件执行。普通批量数据无需使用。

FilterRowKind

×

FilterRowKind • FILTER\_ROW\_KIND

①

保存

基础配置

字段配置

保留事件

排除事件

INSERT ×

▼

> 生成配置预览

图 130 FilterRowKind算子—字段配置

说明：

- 在保留与排除方式中二选一，选择INSERT、UPDATE\_BEFORE、UPDATE\_AFTER或DELETE。
- 例如排除DELETE后删除事件不传递下游，目标可能保留原记录，应符合业务要求。

TableFilter

×

TableFilter

ⓘ

保存

基础配置

字段配置

数据库匹配 ⓘ

\*

Schema 匹配 ⓘ

\*

表名匹配 ⓘ

actor\_\*

匹配模式 ⓘ

保留匹配表

> 生成配置预览

图 131 TableFilter算子一字段配置

说明：

- 填写数据库、Schema、表名正则，并选择保留或排除匹配表。
- 该操作限制输入表范围，不是对单表字段值进行过滤。

TableRename

×

TableRename • TABLE\_RENAME

①

保存

基础配置

字段配置

大小写转换 ②

小写

名称前缀 ②

名称后缀 ②

正则替换 ②

匹配表达式

替换内容

例如: ^ods\_(.\*)\$

例如: dwd\_\$1

🗑️ +

> 生成配置预览

图 132 TableRename算子一字段配置

说明：

- 配置表名大小写转换、前后缀或正则替换。
- 例如增加ods\_前缀形成目标表身份；不会直接重命名来源数据库中的物理表。

TableMerge

×

TableMerge • TABLE\_MERGE

①

保存

基础配置

字段配置

\* 来源表正则 ②

.\*

\* 目标表 ②

orders\_all

> 生成配置预览

图 133 TableMerge算子—字段配置

#### 说明：

- 填写来源表匹配正则与合并后的目标数据库、Schema、表名。
- 将结构兼容的分表汇入同一表身份，不执行多表JOIN；目标字段结构和唯一键需能够容纳合并后的数据。

#### HTTP调用

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

\* 节点名称 ⓘ

获取无人机飞行任务JOB数据14 / 64

\* 方法

POST

\* URL

http://192.168.31.210:8001/gateway/geoai-api-service/v1/job/list

Params

Auth

Headers

Body

参数名

值

Key

值

fx

🗑

+

出参 ^

在线探查

{ 5 项

code : string

msg : string

data : { 2 项

total : bigint

records : [ 11 项 设为数据记录

id : bigint

jobId : string

☒ 分页读取 ⓘ

页码字段 ⓘ

起始页码

每页记录数 ⓘ

总页数 ⓘ

图 105 配置HTTP调用

说明：

- 基础配置填写URL、请求方法、请求头、参数和请求体。接口分页时启用页码分页，设置页码字段、起始页、每页条数及可用的总量信息。

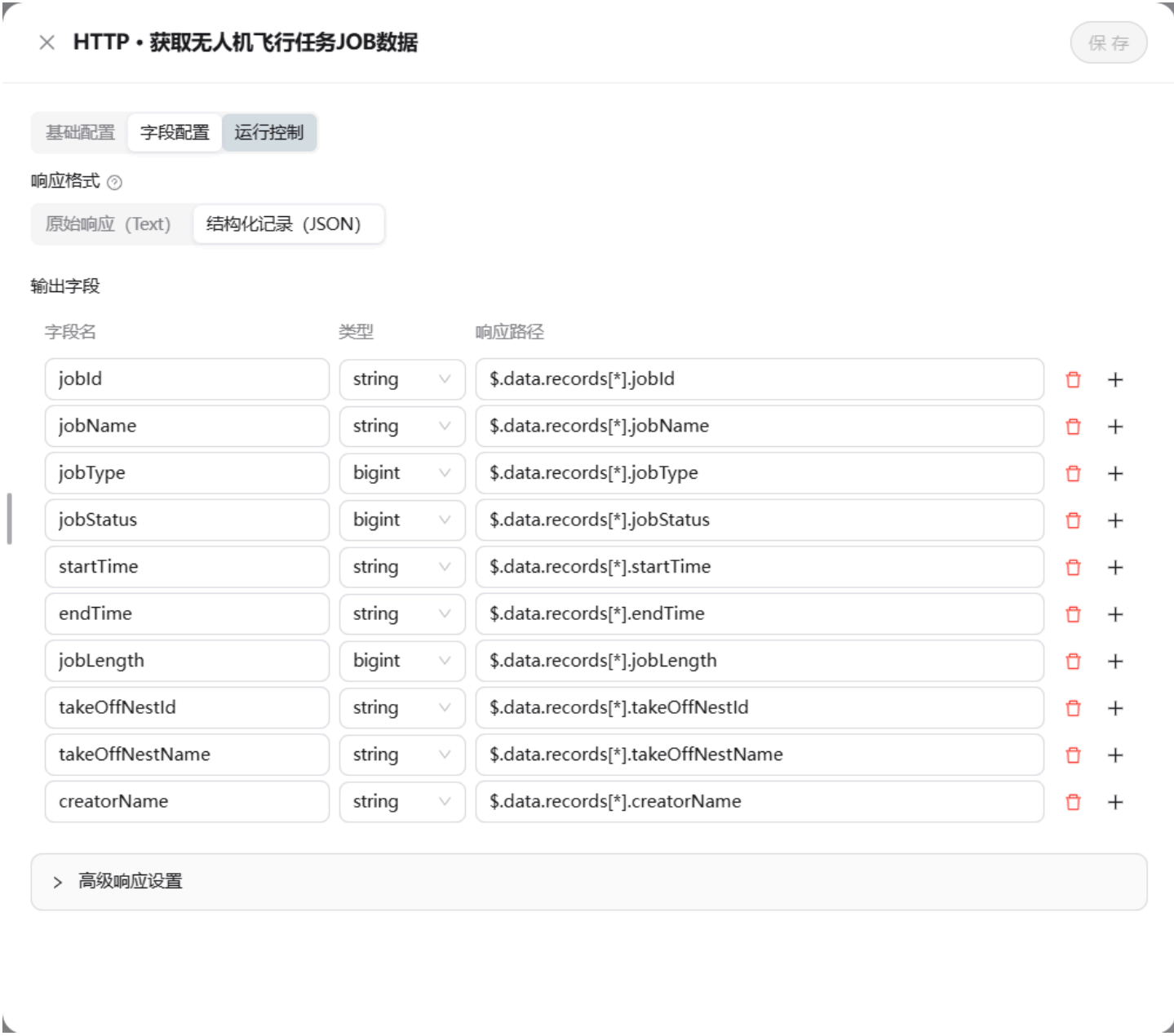


图 106 HTTP调用—字段配置

说明：

- 选择原始文本或结构化JSON解析；使用响应探查确定内容路径、数组位置和输出字段。
- 提取订单列表时选取实际列表数组，不把整个响应包装对象误作一条订单。

×

HTTP • 获取无人机飞行任务JOB数据

保存

基础配置

字段配置

运行控制

最大重试次数

3

重试间隔基准 (毫秒) ?

100

重试间隔上限 (毫秒) ?

10000

连接超时 (毫秒)

12000

读取超时 (毫秒)

60000

图 107 HTTP调用—运行控制

说明：

- 配置请求重试、重试间隔上限、连接和读取超时。接口限流或失败时结合返回信息调整，避免高并发持续重试。

LLM

×

LLM

ⓘ

保存

\* 算子名称

LLM

3 / 64

\* 供应商

DEEPSEEK

\* 接口地址

https://api.deepseek.com/chat/completions

API Key

.....

🔑

已保护；重新输入可替换

\* 模型

点击查询并选择，或手动输入模型名称

deepseek-flash

deepseek-v4-pro

自定义

\* 提示词

\* 输入字段

图 108 配置LLM

说明：

- 支持通用OpenAI协议，常见OPENAI、DEEPSEEK、DOUBAO、KIMIAI、ZHIPU、Ollama等多种协议模型，能够使用LLM算子借助调用大模型实现数据智能化处理。

### 3.5.2.3 输出算子配置

#### JDBC输出

×

JDBC • JDBC\_SINK

保存

基础配置

写入配置

表配置

\* 算子名称

JDBC\_SINK9 / 64

\* 目标数据源

XuguDB / 数据同步数仓DW4FLOW (192.168.31.213:5138/DEMOW)▼

图 134 配置JDBC输出

- 说明：
- 基础配置选择目标数据源与表，设置表结构和已有数据处理策略。自动生成或自定义建表SQL应与目标库类型一致。
  - 启用Upsert时配置目标主键；清空、自定义SQL等可能改变既有数据，执行前确认范围。

×

JDBC • JDBC\_SINK

保存

基础配置

写入配置

表配置

写入方式

自动生成

\* 目标表结构处理 ⓘ

目标已有数据处理 ⓘ

表不存在时报错

追加数据

启用 Upsert ⓘ

写入批大小 ⓘ

1000

目标连接检查超时 (秒) ⓘ

批量写入重试次数 ⓘ

30

0

图 135 JDBC输出—写入配置

- 说明：
- 使用表格或画布映射来源与目标字段，检查字段类型及必需字段。由目标或平台生成的ID无需映射为上游字段。

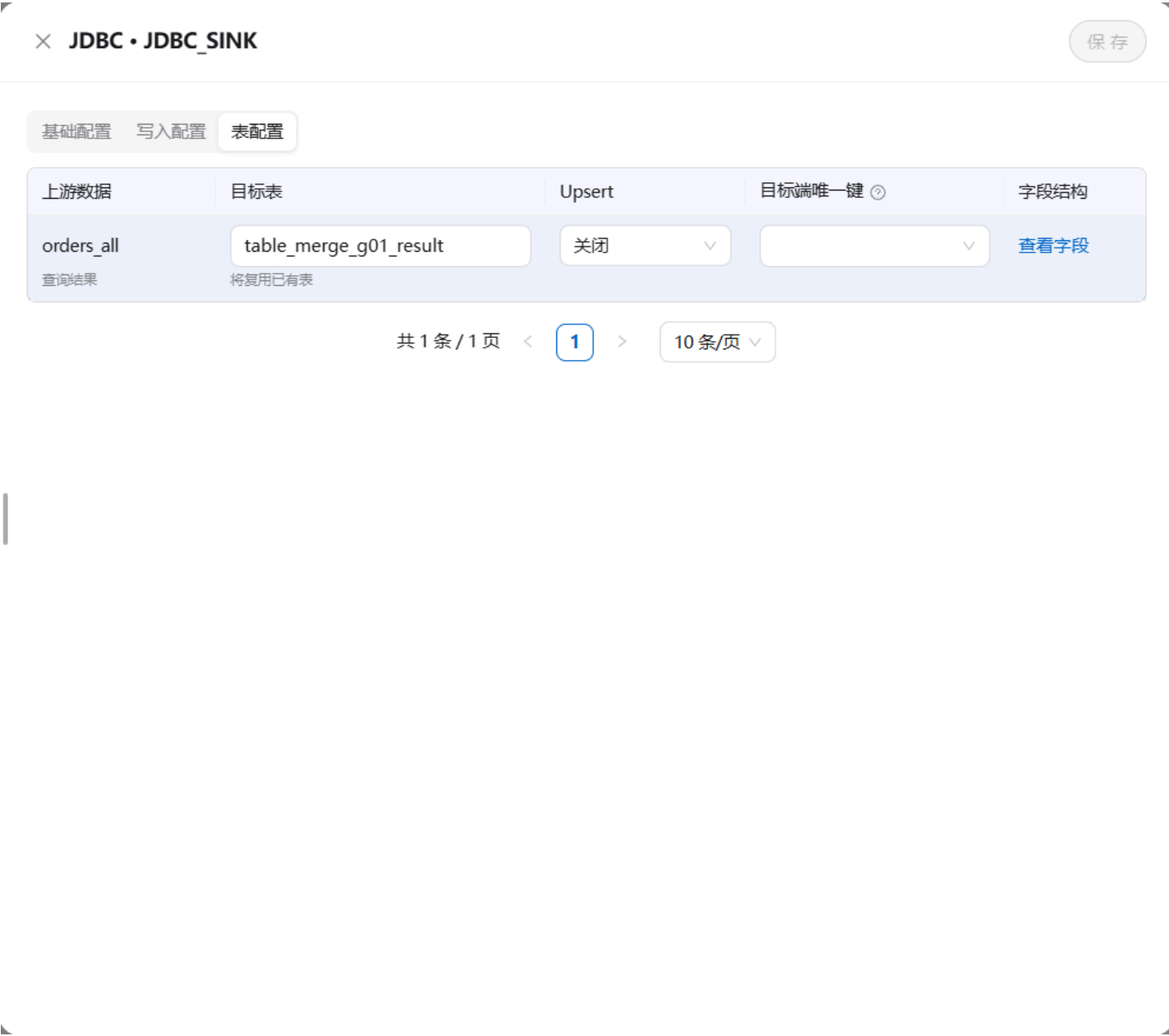


图 136 JDBC输出一表配置

说明：

- 设置目标表、更新策略和目标端唯一键，并配置字段映射。

Kafka输出

×

Kafka

保存

基础配置

字段配置

运行控制

\* 算子名称

Kafka5 / 64

\* 数据源 ⓘ

Kafka / orderdb-kafka-216 (192.168.31.216:9092)

\* Topic

orderdb-orders

分区键字段

id ×

图 137 配置Kafka输出

说明：

- 基础配置选择Kafka数据源、Topic与分区键，确认下游消费者使用的Topic。

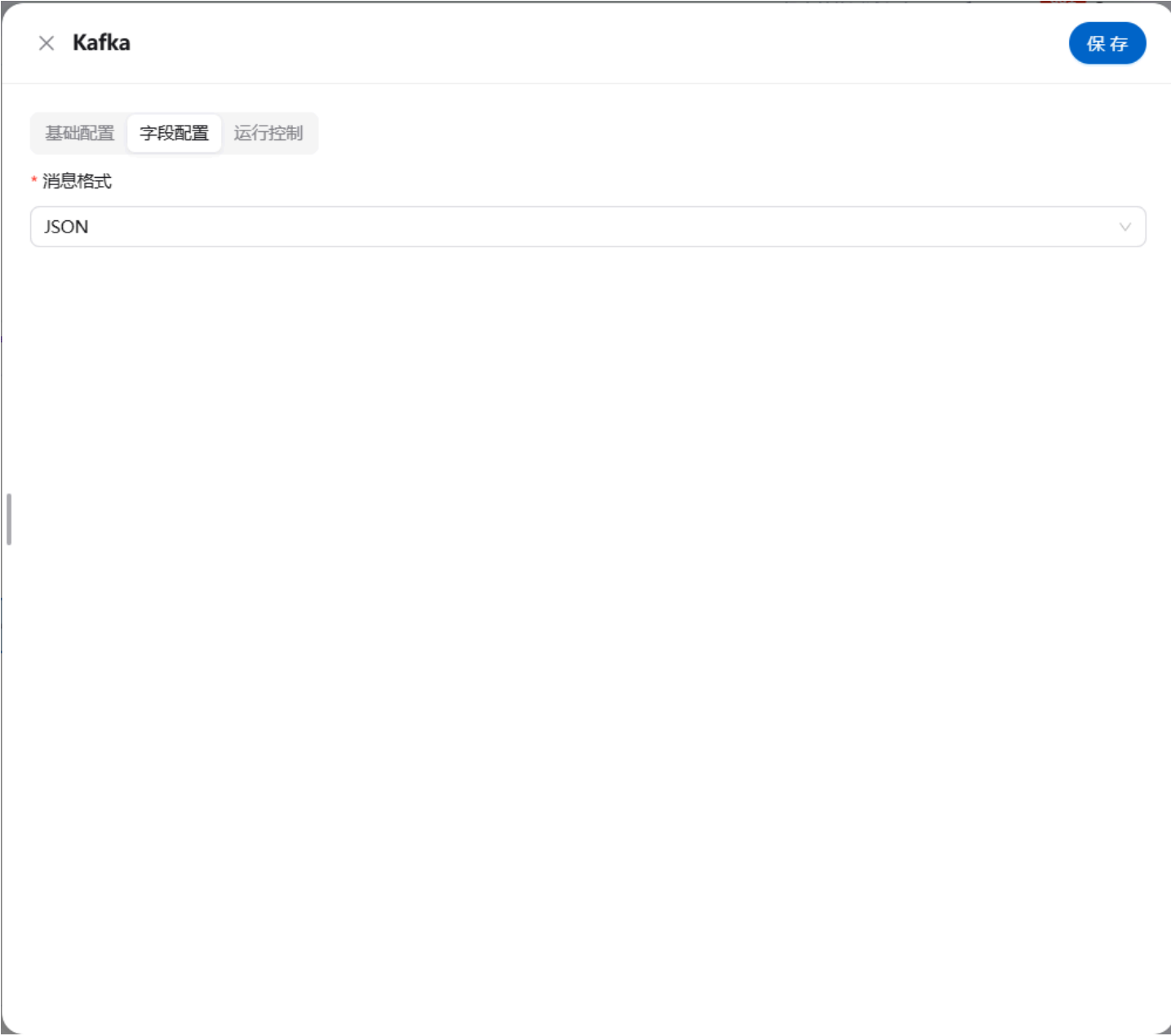


图 138 Kafka输出一字段配置

说明：

- 选择消息格式，并按所选格式配置分隔符等参数。输出结构应与消费者约定一致。

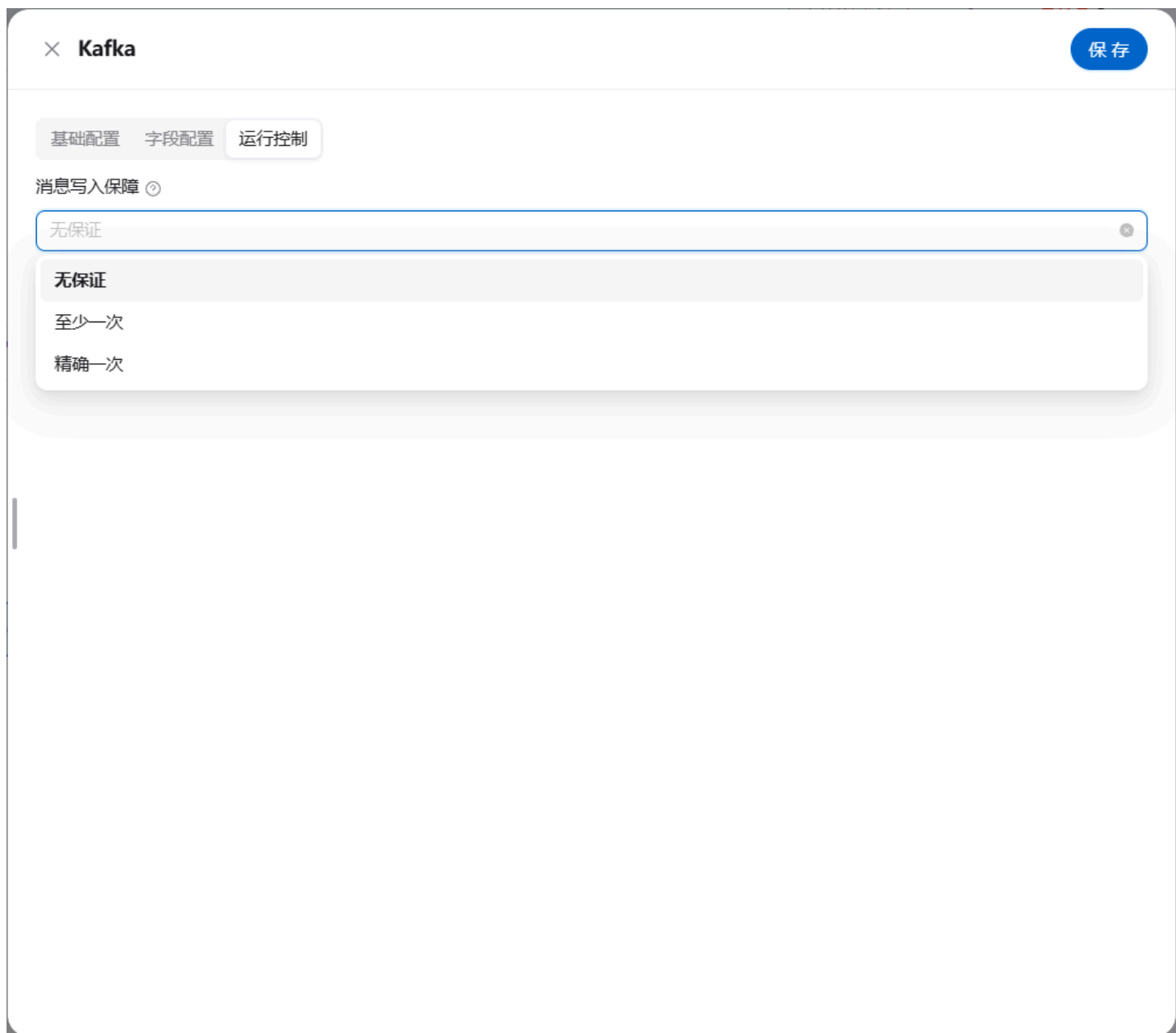


图 139 Kafka输出—运行控制

说明：

- 选择交付保证。使用精确一次时按要求填写唯一的事务前缀，并结合Checkpoint设置。
- 该选项约束对应Kafka写入，不保证工作流其他外部接口同时具备相同事务能力。

HTTP输出

×

HTTP

保存

基础配置

运行控制

\* 算子名称

HTTP4 / 64

\* 数据源

数据源类型 / 数据源

\* Webhook URL

https://api.example.com/orders

请求头

| 参数名 | 值 |
|-----|---|
| Key | 值 |

数组批量发送

请求批大小

1

图 140 配置HTTP输出

- 说明：
- 基础配置选择HTTP数据源或填写Webhook地址、请求头，设置是否数组批量发送及批大小。
  - 按接口接收约定发送JSON；写入接口需考虑幂等键，避免重试重复创建物流单据。

×

HTTP

保存

基础配置

运行控制

最大重试次数

3

连接超时 (毫秒)

12000

读取超时 (毫秒)

60000

图 141 HTTP输出—运行控制

- 说明：
- 配置最大重试次数、连接和读取超时。请求超时不一定表示服务端未处理，应结合接口幂等机制核对结果。

Redis输出

×

Redis

保存

基础配置

字段配置

运行控制

\* 算子名称

Redis5 / 64

\* 数据源 ⓘ

Redis / Redis数据源dg\_sync\_6380 (192.168.31.216:6380/0)▼

\* 键表达式

orders:\${id}

\* 数据类型

String▼

过期秒数 ⓘ

-1

图 142 配置Redis输出

说明：

- 基础配置选择数据源、键表达式、数据类型及过期时间；过期时间为-1表示不自动过期。

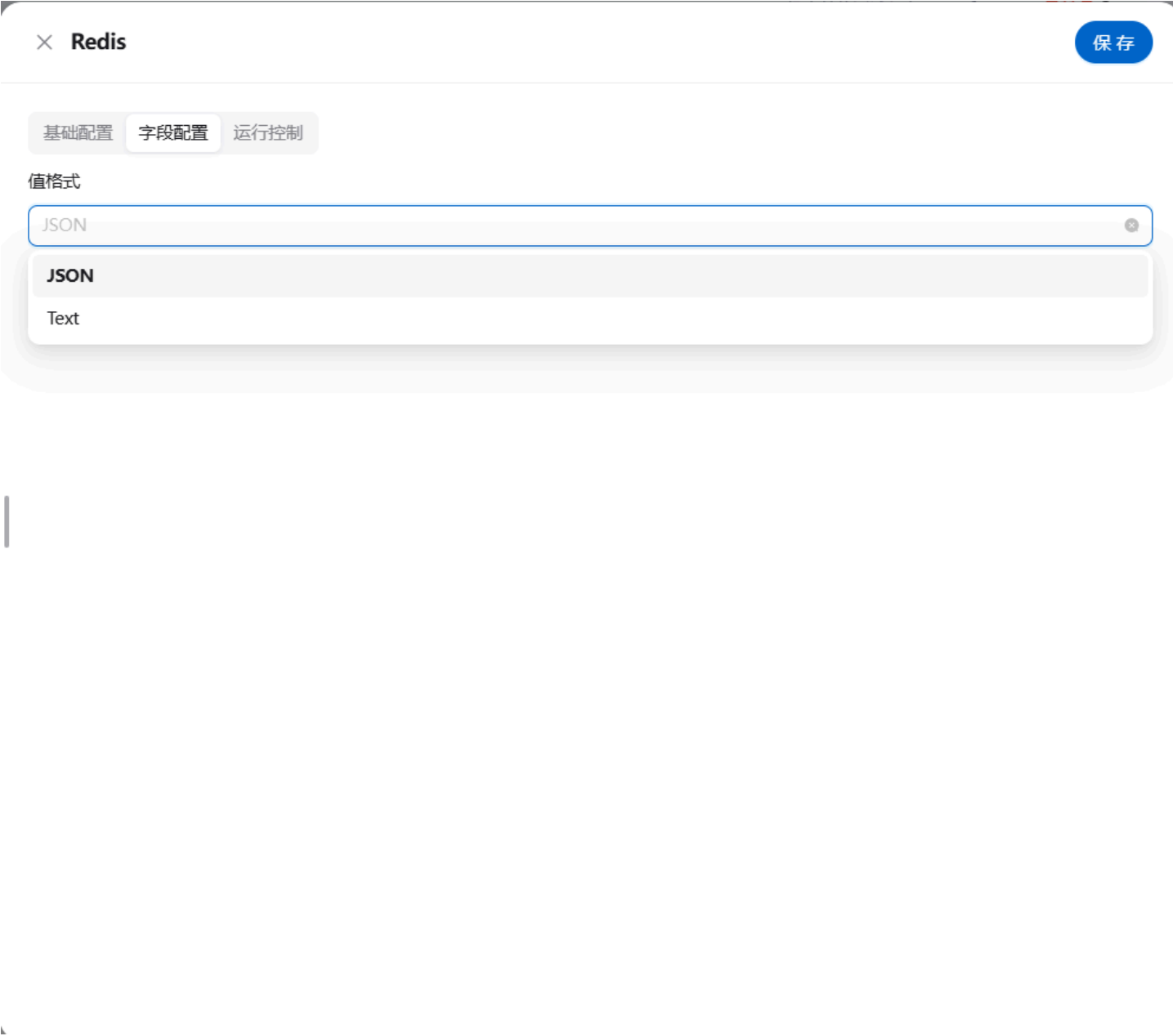


图 143 Redis输出一字段配置

说明：

- 选择JSON或Text等输出格式，检查生成值与后续读取方式一致。

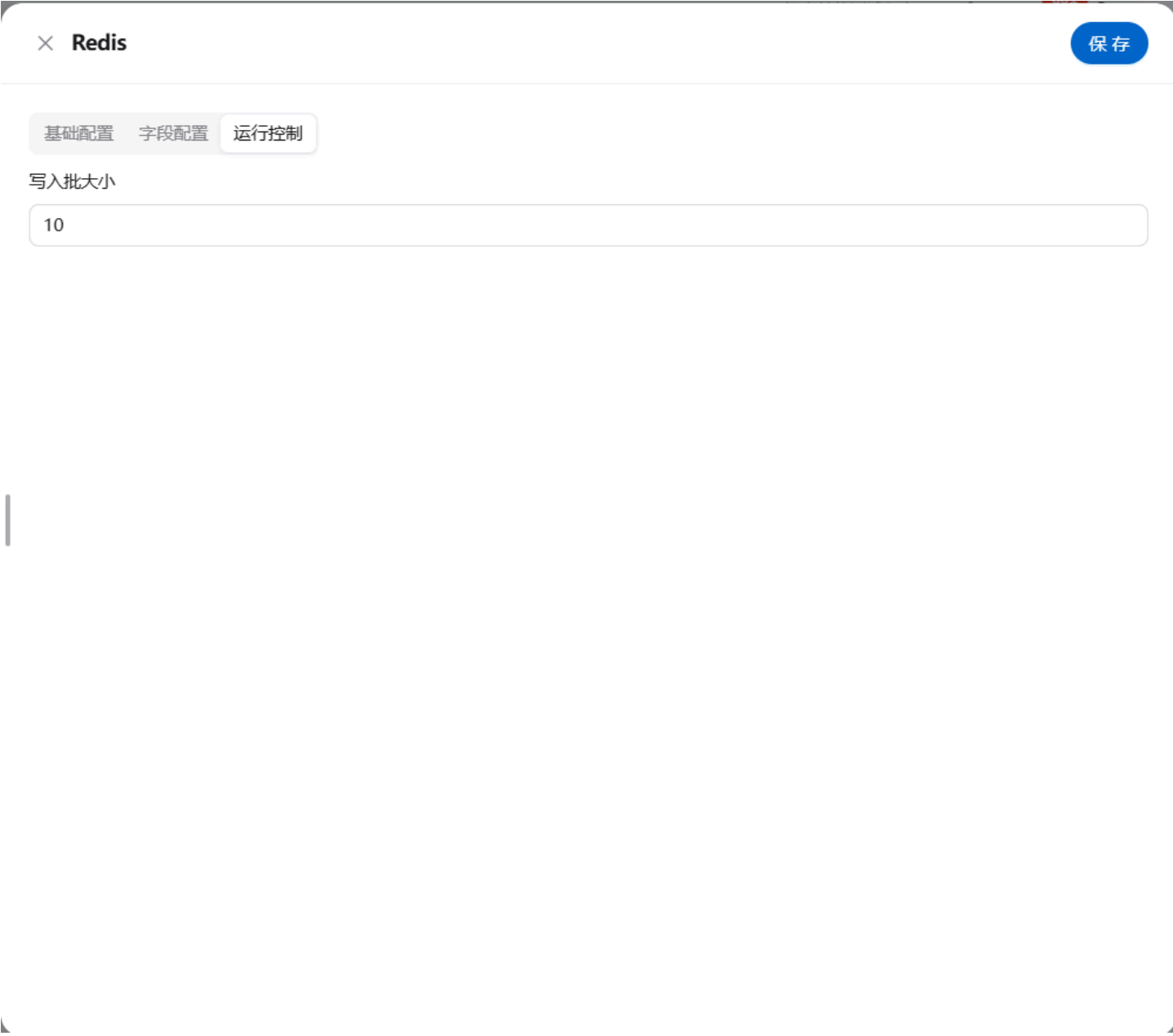


图 144 Redis输出—运行控制

说明：

- 配置批量写入大小。键表达式应避免不同订单意外使用同一个键而互相覆盖。

Elasticsearch输出

×

Elasticsearch

保存

基础配置

运行控制

\* 算子名称

Elasticsearch13 / 64

\* 数据源 

数据源类型 / 数据源

\* 索引

orders

文档主键字段

请选择目标数据源

索引处理

索引不存在时创建

已有数据处理

追加数据

图 145 配置Elasticsearch输出

说明：

- 基础配置选择数据源、索引及文档主键，设置索引结构和已有数据处理。
- 文档主键用于识别目标文档，应结合来源唯一字段配置，避免覆盖无关文档。

×

Elasticsearch

保存

基础配置

运行控制

批量写入条数

10

最大重试次数

3

图 146 Elasticsearch输出—运行控制

说明：

- 设置批大小与重试次数，结合目标索引写入能力控制压力。

Console输出

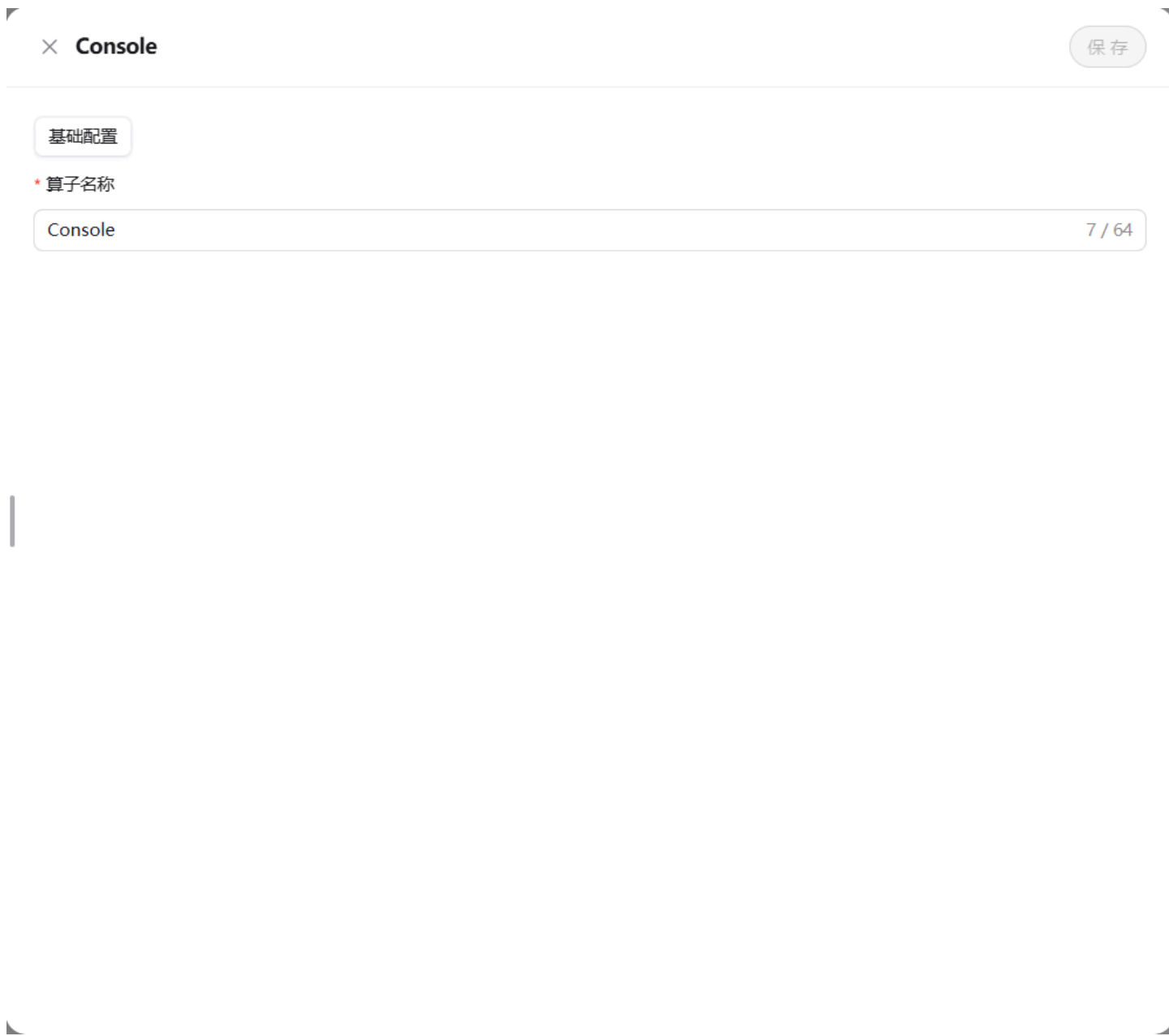


图 153 配置Console输出

说明：

- 添加Console输出连接上游，用于查看处理结果和辅助调试。
- 控制输出的数据范围，避免在日志中暴露敏感业务数据。Console输出不等于数据已写入业务目标库。

## 3.6 函数管理

函数管理集中展示平台提供的内置函数，供开发人员查找参数表达式可用的函数与输入输出约定。

### 3.6.1 函数列表

fx

函数管理

查看节点参数配置可用的平台内置原子函数。

搜索函数名称或标识

C

| 函数名称          | 函数标识          | 类型 | 入参                                | 返回类型    | 状态  | 维护方  |
|---------------|---------------|----|-----------------------------------|---------|-----|------|
| base64        | base64        | 内置 | value:STRING                      | STRING  | 已发布 | 平台内置 |
| concat        | concat        | 内置 | values:STRING                     | STRING  | 已发布 | 平台内置 |
| defaultIfNull | defaultIfNull | 内置 | defaultValue:STRING, value:STRING | STRING  | 已发布 | 平台内置 |
| hmacSha256    | hmacSha256    | 内置 | key:STRING, value:STRING          | STRING  | 已发布 | 平台内置 |
| jsonPath      | jsonPath      | 内置 | json:JSON, path:STRING            | JSON    | 已发布 | 平台内置 |
| jsonStringify | jsonStringify | 内置 | value:STRING                      | STRING  | 已发布 | 平台内置 |
| sha256        | sha256        | 内置 | value:STRING                      | STRING  | 已发布 | 平台内置 |
| timestamp     | timestamp     | 内置 | -                                 | LONG    | 已发布 | 平台内置 |
| toBoolean     | toBoolean     | 内置 | value:STRING                      | BOOLEAN | 已发布 | 平台内置 |
| toLong        | toLong        | 内置 | value:STRING                      | LONG    | 已发布 | 平台内置 |

共 13 条 / 2 页

< 1 2 >

10 条/页

图 154 查看函数列表

- 说明：
- 进入函数管理，使用搜索定位函数，查看函数标识、名称、输入签名和返回类型。可刷新获取最新列表。
  - 根据签名确认参数数量与类型，再在支持的参数表达式入口使用。该列表用于查询内置函数，不提供自定义函数新建、编辑或发布。

## 4 审批中心

审批中心处理发布等操作触发的审批事项，区分待办、已办和本人发起的记录。审批负责授权决策，实际发布或建表仍需在对应对象中确认执行结果。

### 4.1 审批管理

用户依据角色处理待办或跟踪发起事项；具有配置权限的用户可维护场景是否需要审批。审批操作以详情中当前可执行按钮为准。

#### 4.1.1 审批管理列表

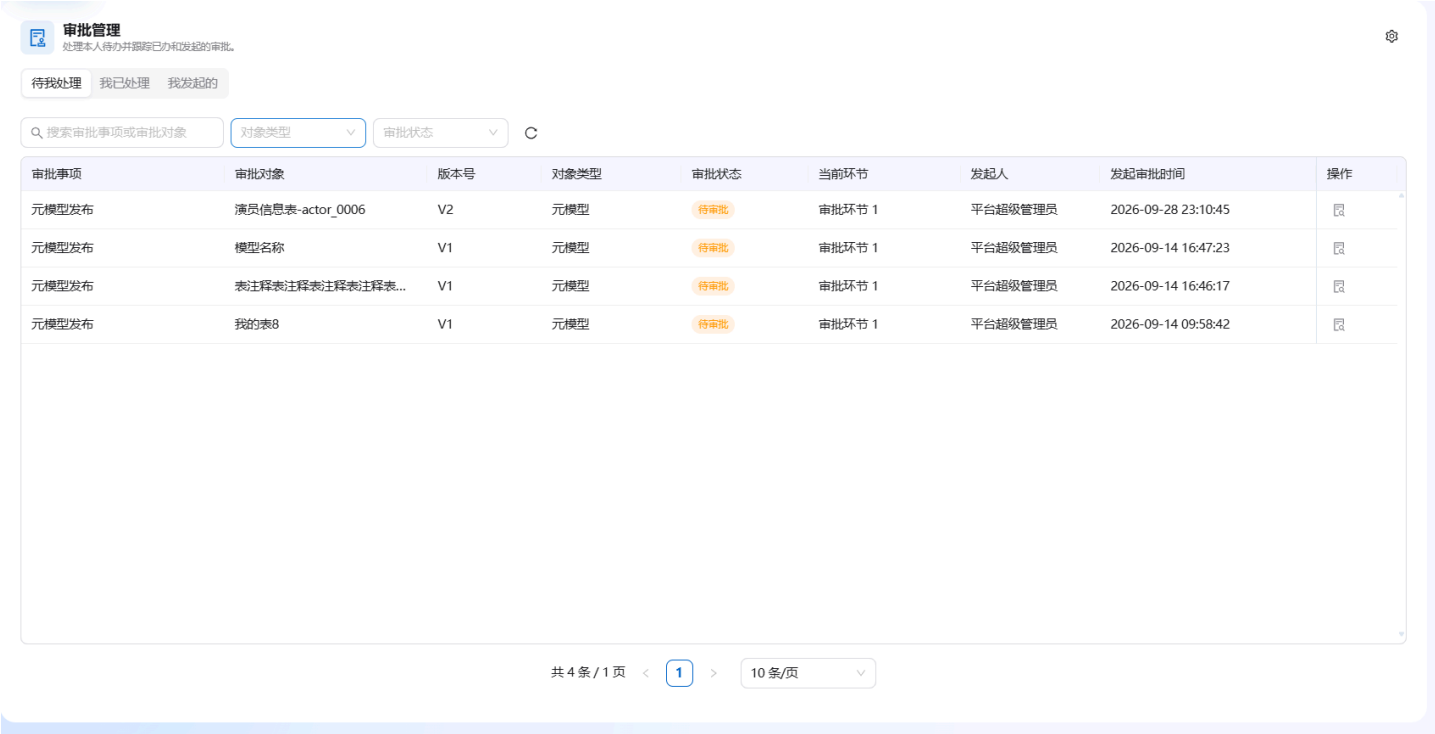


图 155 查看待我处理

说明：

- 切换待我处理，按关键词、类型、状态等条件查找待办，查看对象、版本、环节、发起人和时间。
- 打开记录核对本次变更对象，避免处理同名任务的其他版本。

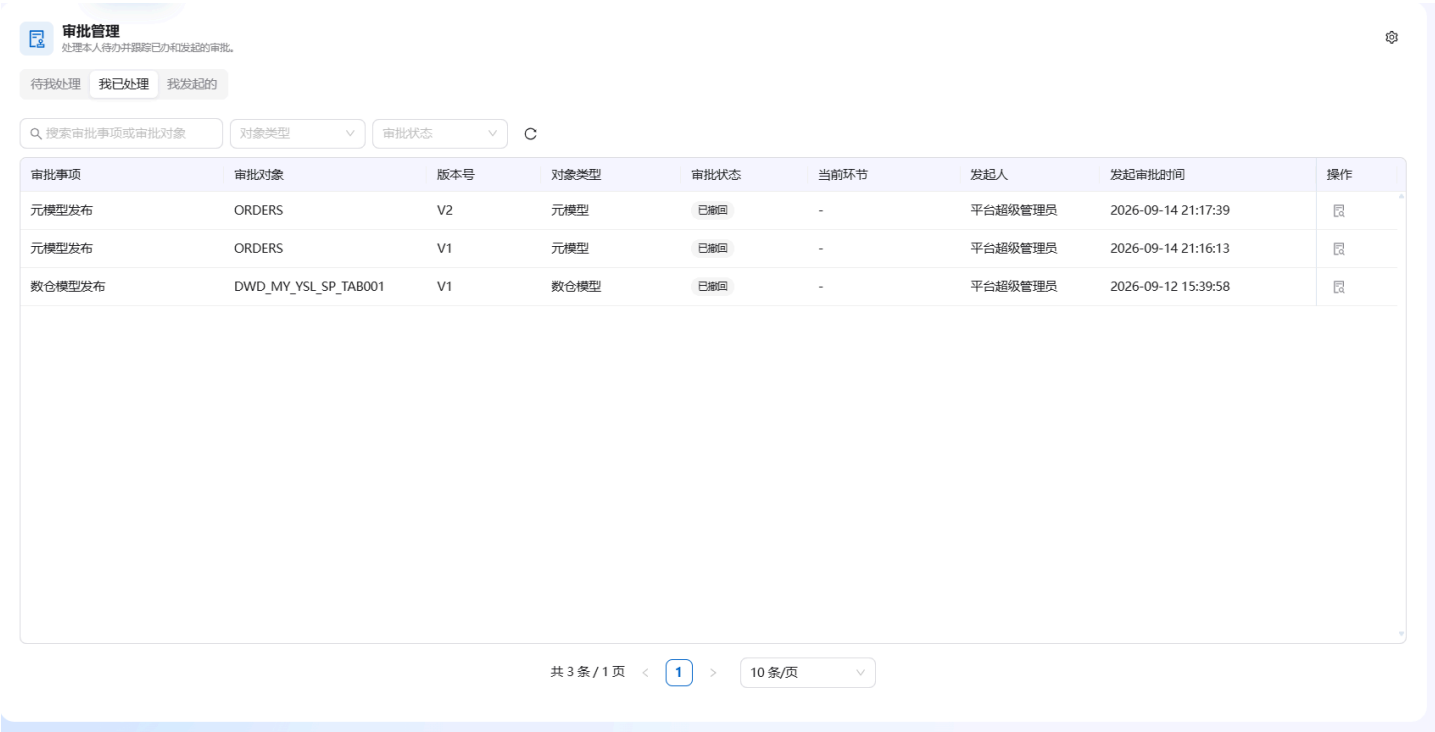


图 156 查看我已处理

说明：

- 切换我已处理，查询本人已经办理的审批记录。
- 打开详情查看意见与流转结果；本人处理完成不代表所有后续审批环节均已结束。

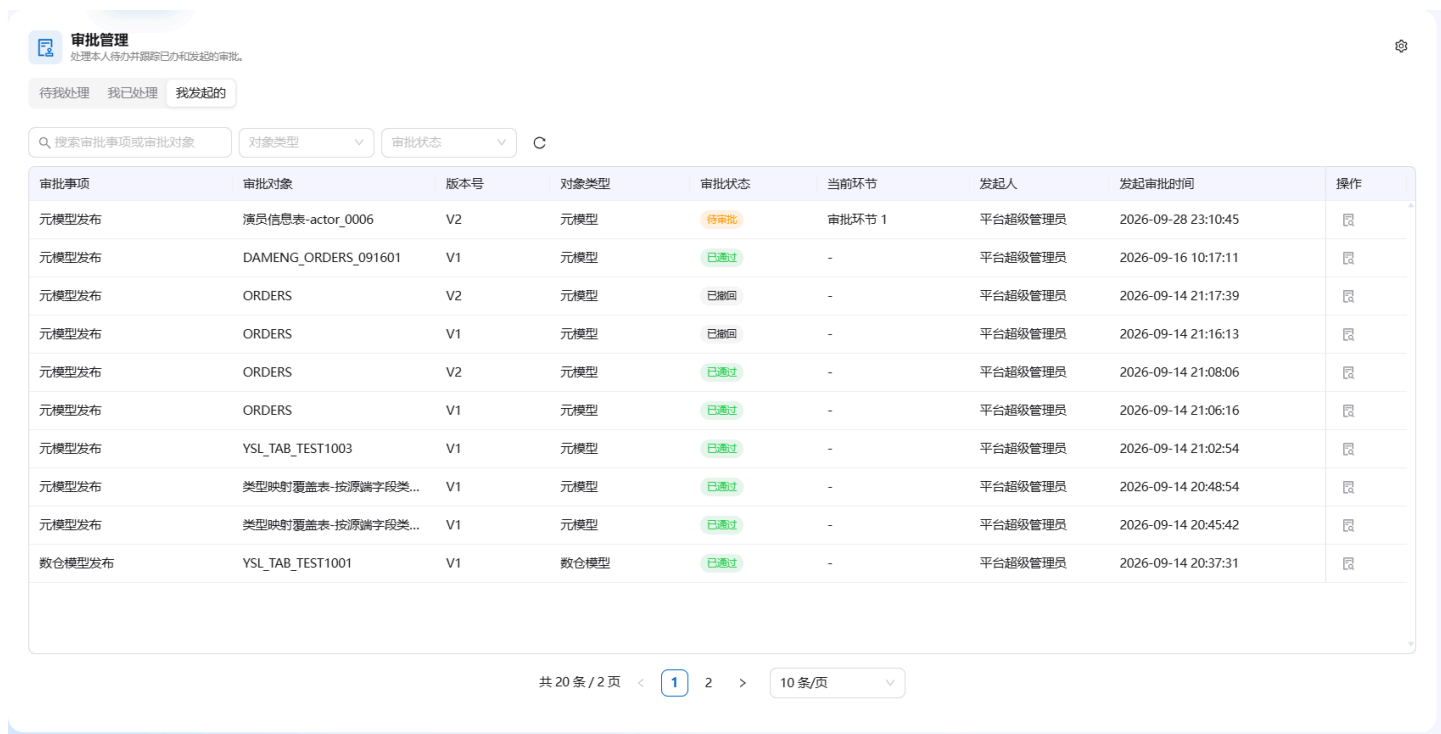


图 157 查看我发起的

说明：

- 切换我发起的，跟踪本人提交事项的当前状态、环节和处理进度。
- 被退回或拒绝后，先根据意见修正原对象，再按允许的入口重新提交。

### 4.1.2 审批详情



图 158 查看审批详情

说明:

- 点击审批事项打开详情，查看对象信息、申请内容、当前状态和处理时间线。
- 通过时间线识别每个处理人及意见；实际发布完成情况应返回任务或模型中检查。

4.1.3 办理审批与撤回申请



图 159 通过审批

说明:

- 在当前可处理事项中选择通过，填写要求的意见并确认。

- 通过表示同意当前申请；有后续节点时继续流转，最终通过后仍需等待发布或建表处理。



图 161 拒绝审批

说明：

- 选择拒绝，填写理由并确认。

- 拒绝前核对申请范围；不要使用拒绝替代任务运行停止，审批与运行控制是不同操作。

×

审批详情

审批事项：元模型发布

审批对象：演员信息表-actor\_0006

版本号：V2

对象类型：元模型

审批状态：

待审批

当前环节：审批环节 1

发起人：平台超级管理员

发起审批时间：2026-09-28 23:10:45

审批时间轴

○

发起审批

平台超级管理员 · 2026-09-28 23:10:45

○

审批环节 1

待处理 · user02001

○

审批环节 2

未开始 · 用户02012

○

结束

未开始 · 待分配

✓

×

转交

↩

图 162 转交审批

说明：

- 选择转交，从可选人员中指定接收人，填写意见后提交。

- 转交后核对当前处理人及流转记录，确保事项交给具有处理资格的人员。



图 163 撤回申请

说明：

- 在本人发起且尚符合撤回条件的事项中选择撤回并确认。

- 已进入不允许撤回的处理阶段时，不再显示可用撤回动作；应与当前处理人协作。

4.1.4 审批配置

×

审批配置

保存

| 审批场景     | 需要审批                                | 更新人     | 更新时间                |
|----------|-------------------------------------|---------|---------------------|
| 元模型发布    | <input checked="" type="checkbox"/> | 平台超级管理员 | 2026-09-14 21:17:11 |
| 数仓模型发布   | <input type="checkbox"/>            | 平台超级管理员 | 2026-09-14 21:17:11 |
| 离线同步发布   | <input type="checkbox"/>            | 平台超级管理员 | 2026-09-14 21:17:11 |
| 实时同步发布   | <input type="checkbox"/>            | 平台超级管理员 | 2026-09-14 21:17:11 |
| 数据开发任务发布 | <input type="checkbox"/>            | 平台超级管理员 | 2026-09-14 21:17:11 |

图 164 配置审批要求

说明：

- 具有权限时点击审批配置，按场景查看是否需要审批及更新信息，调整开关后保存。
- 该配置控制相应操作的审批要求，不用于编辑业务任务内容。变更前确认企业审批制度及影响范围。

5 运维中心

运维中心统一管理已发布工作流的调度与执行，覆盖离线工作流的计划、手动运行和补数，以及实时工作流的启动、停止和恢复。对于已配置的集中式或分布式存储之间的数据采集、复制及迁移工

作流，运维人员可在统一入口查看生产配置、执行实例、节点状态和日志，检查来源读取、数据处理及目标写入情况。

生产 workflow 负责执行计划与启动，workflow 实例记录每次实际运行的参数、配置和执行结果；执行结果通知与节点超时告警用于发现和追踪异常。生产配置、执行实例及通知分别管理，便于区分“准备运行”“正在运行”和“发现异常”。

运维中心统一调度的是已发布的数据处理 workflow。来源、目标及写入方式在对应设计页面配置；用于数据迁移时，执行成功后仍需核验目标数据，并按业务安排完成访问切换及来源数据处置。

## 5.1 生产 workflow

生产 workflow 展示正式发布版本。离线 workflow 可定时、手动或补数运行；实时 workflow 以持续启动和停止管理。修改开发草稿不会直接改写既有生产版本。

### 5.1.1 生产 workflow 列表

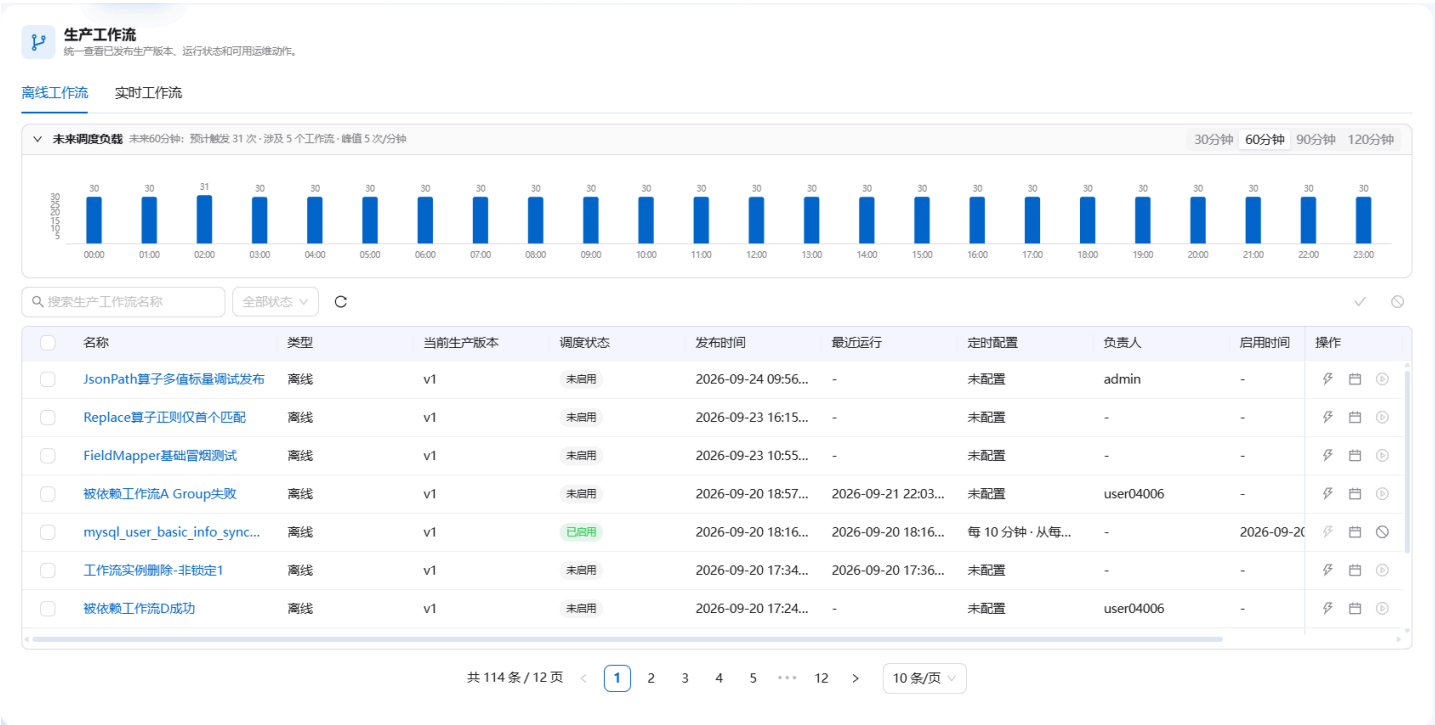


图 165 查看离线生产 workflow

- 说明：
- 切换离线 workflow，按名称与状态查找。查看当前生产版本、启停状态、最近运行、定时配置、负责人和启停时间。
  - 通过行操作配置调度或运行，通过当前页勾选完成允许的批量启停。启用状态不等于此刻存在运行实例。

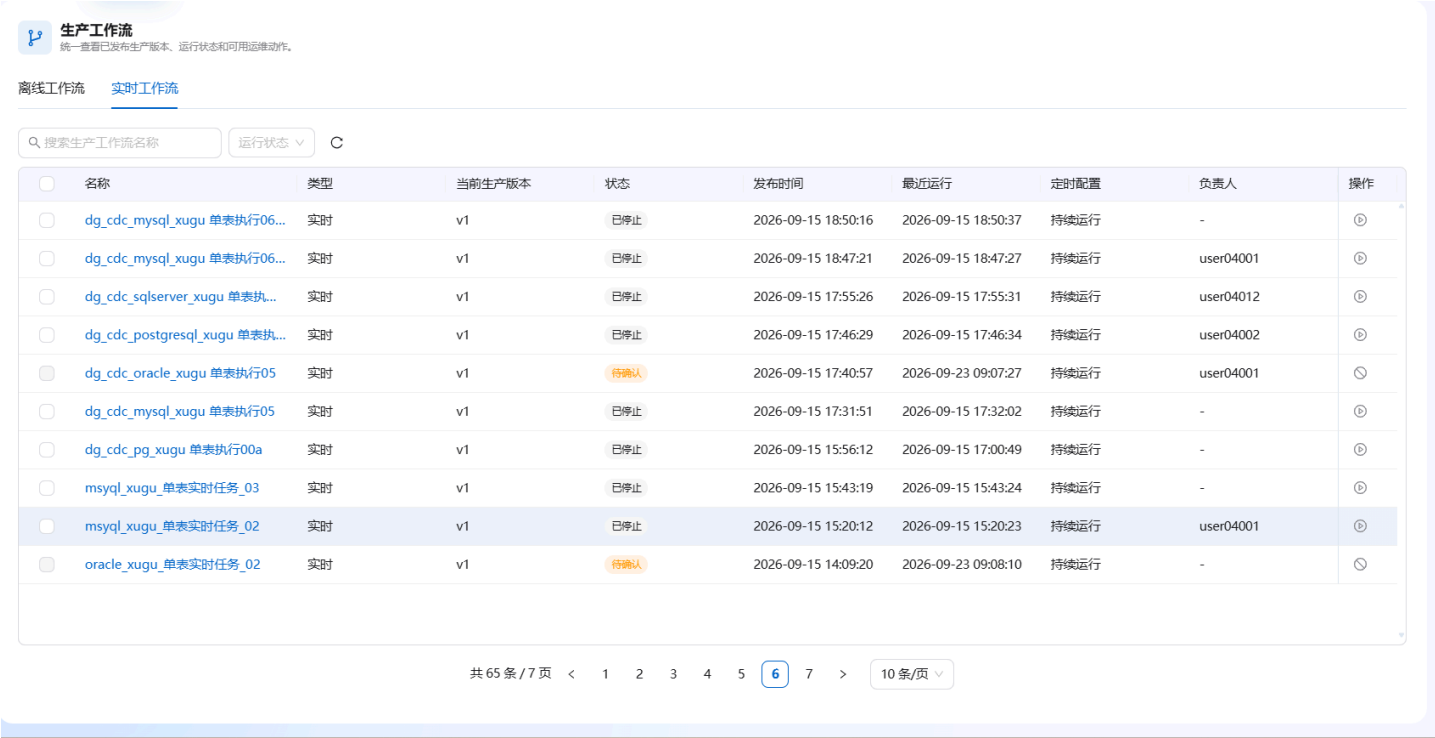


图 166 查看实时生产 workflows

说明：

- 切换实时 workflows，定位已发布任务，查看生产版本与运行状态。
- 通过启动或停止管理持续运行；实际健康状态需进入实时实例查看，而不是只看 workflows 是否已发布。

### 5.1.2 生产 workflows 详情

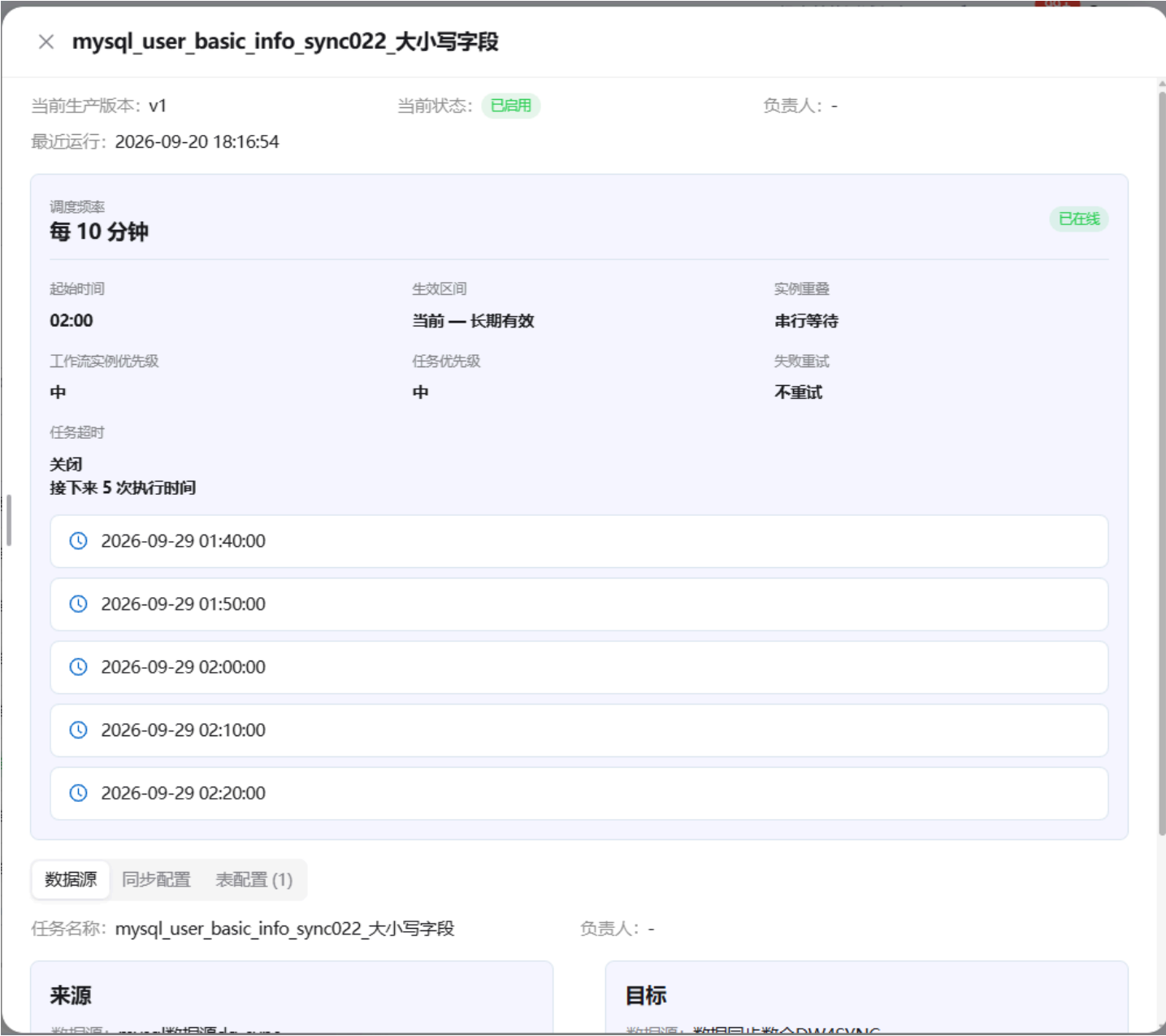


图 167 查看生产 workflow 详情

- 说明:
- 点击名称查看生产配置快照。自定义工作流通过基本信息、编排配置和工作流参数页签核对内容。
  - 查看生产版本节点与连线，确认已发布的是预期流程。
  - 查看生产版本参数定义，运行时传值应符合这些参数约定。
  - 离线工作流查看当前调度计划、有效期、实例策略与任务默认策略。
  - 实时工作流查看启动与恢复配置；快捷同步工作流通过其数据源、同步配置、表配置及CDC与恢复页签核对生产快照。
  - 对跨存储采集、复制及迁移工作流，核对生产版本中的来源数据源、目标数据源、采集对象、字段映射和写入策略。集中式或分布式存储的可用范围，以相应任务支持的连接类型及读写能力为准。
  - 运维页面查看的是已发布配置快照。需要调整来源、目标或处理逻辑时，应返回开发页面修改并发布，再核对生产版本。

5.1.3 调度配置

调度设置 • 被依赖工作流A Group失败

调度修订独立保存，不会改变当前生产版本。

工作流级调度

周期调度

周期

分

时

日

周

月

季度

年

执行频率

每天执行一次

调度时间

02:00

生效日期 (可选)

开始日期 (不限)

→ 结束日期 (不限)

近 5 次执行时间

1. 2026-09-29 02:00

2. 2026-09-30 02:00

3. 2026-10-01 02:00

4. 2026-10-02 02:00

5. 2026-10-03 02:00

实例执行

实例重叠策略

工作流实例优先级

取消

保存

图 168 调度配置完整表单

说明：

配置离线调度周期

- 点击调度配置，选择分钟、小时、日、周、月、年等周期，按界面填写执行时刻或间隔。快捷周期仍需检查实际生成的计划。

- 设置生效区间；未限定结束日期时长期有效。核对平台时区以及接下来5次预计执行时间，确认符合业务要求。
- 保存调度只保存计划，不自动启用未启用 workflow。已启用计划的调整按页面反馈确认是否生效；调度修订不等于发布新的 workflow 业务版本。

配置实例执行策略

- 在调度配置中设置实例优先级与实例重叠策略。串行等待表示上一次尚未完成时，后续实例等待，而不是同时运行。
- 根据任务耗时和数据写入方式安排策略，避免重叠全量写入同一目标表。
- 跨存储复制及迁移任务应结合来源读取压力、目标写入能力和预计耗时配置执行策略。多个任务写入同一目标表或分区时，应检查执行时间和重叠策略，避免重复写入或相互覆盖。
- 调度配置用于安排 workflow 执行，不直接调整底层存储池容量或分配存储资源。

配置任务默认策略

- 设置任务默认优先级、重试次数与间隔、超时时长及处理方式。
- WARN 用于告警提示，失败处理与单纯告警不同。任务重试可能重复执行已经产生外部效果的操作，应结合写入幂等性配置。
- 展开未来调度负载，选择 30、60、90 或 120 分钟窗口，查看预计触发数量与峰值。
- 预测的是计划触发情况，不是实际并发占用；实际任务耗时、排队和失败重试会影响运行负载。

5.1.4 离线 workflow 启停与运行

生产 workflow

统一查看已发布生产版本、运行状态和可用运维动作。

离线 workflow

实时 workflow

▼ 未来调度负载

未来60分钟：预计触发 31 次，涉及 5 个 workflow，峰值 5 次/分钟

30分钟

60分钟

90分钟

120分钟

Q 搜索生产 workflow 名称

全部状态 ▼

C

✓

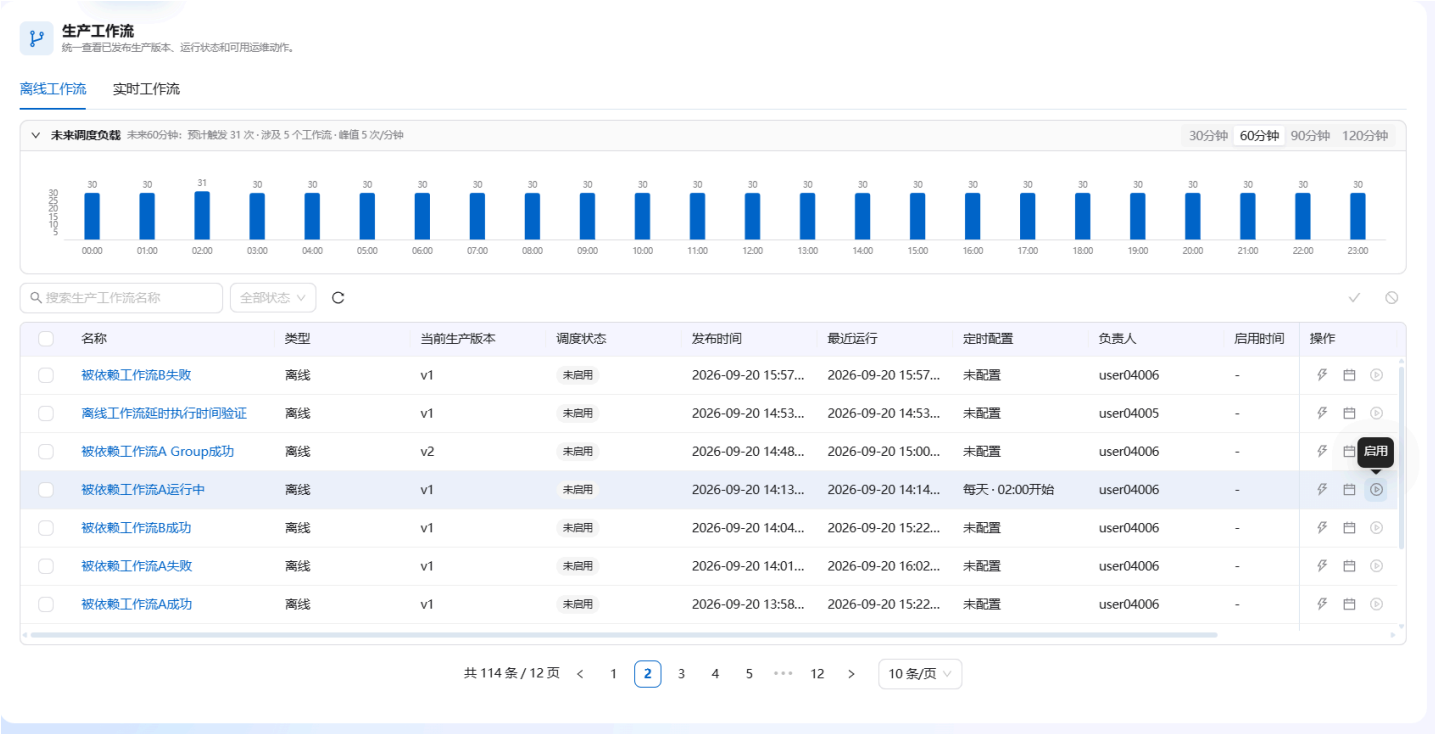
🔍

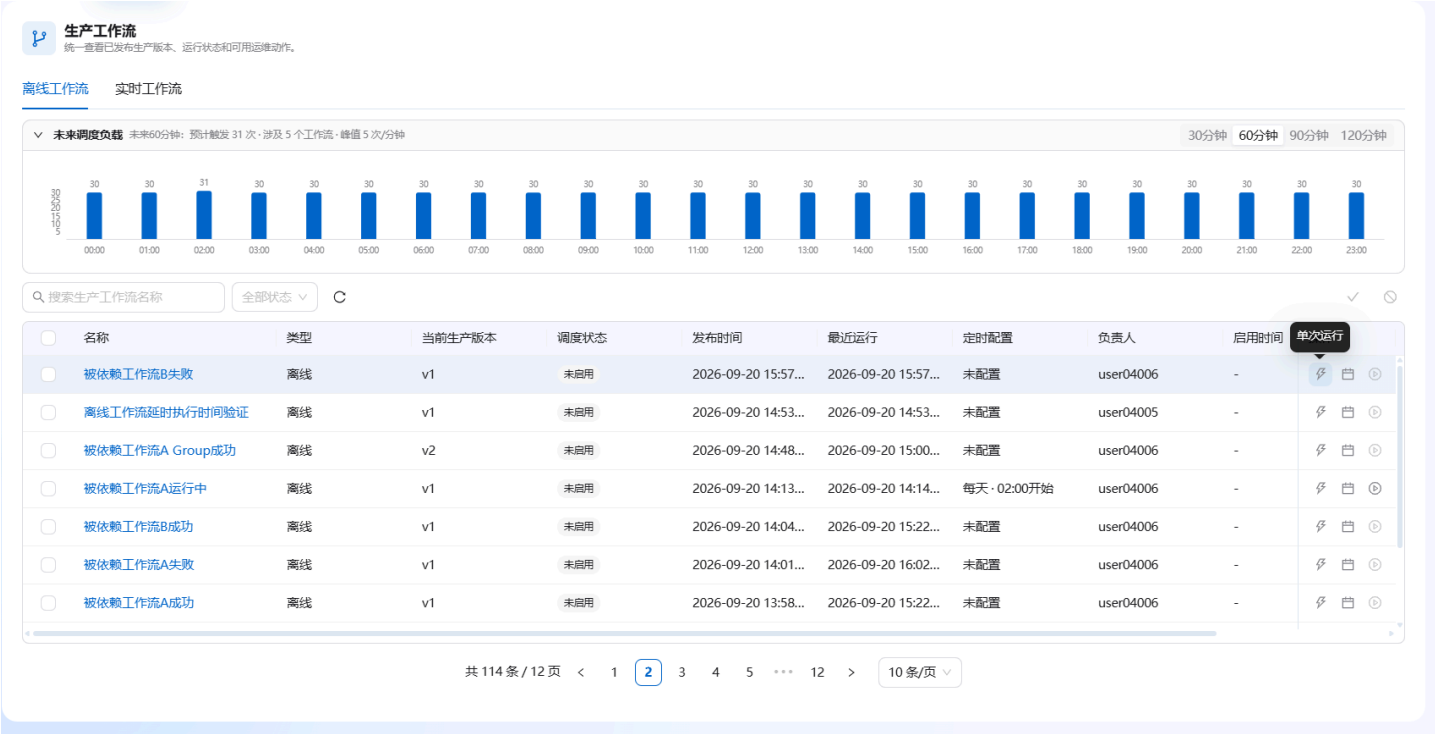
| <input type="checkbox"/>            | 名称                            | 类型 | 当前生产版本 | 调度状态 | 发布时间                | 最近运行                | 定时配置            | 负责人       | 启用时间       | 操作                   |
|-------------------------------------|-------------------------------|----|--------|------|---------------------|---------------------|-----------------|-----------|------------|----------------------|
| <input type="checkbox"/>            | JsonPath算子多值标量调试发布            | 离线 | v1     | 未启用  | 2026-09-24 09:56... | -                   | 未配置             | admin     | -          | <div>🔍 🗑️ ⌛</div>    |
| <input type="checkbox"/>            | Replace算子正则仅首个匹配              | 离线 | v1     | 未启用  | 2026-09-23 16:15... | -                   | 未配置             | -         | -          | <div>🔍 🗑️ ⌛</div>    |
| <input type="checkbox"/>            | FieldMapper基础富烦测试             | 离线 | v1     | 未启用  | 2026-09-23 10:55... | -                   | 未配置             | -         | -          | <div>🔍 🗑️ ⌛</div>    |
| <input type="checkbox"/>            | 被依赖 workflow A Group失败        | 离线 | v1     | 未启用  | 2026-09-20 18:57... | 2026-09-21 22:03... | 未配置             | user04006 | -          | <div>🔍 🗑️ ⌛ 停用</div> |
| <input checked="" type="checkbox"/> | mysql_user_basic_info_sync... | 离线 | v1     | 已启用  | 2026-09-20 18:16... | 2026-09-20 18:16... | 每 10 分钟 · 从每... | -         | 2026-09-20 | <div>🔍 🗑️ ⌛</div>    |
| <input type="checkbox"/>            | workflow实例删除-非锁定1             | 离线 | v1     | 未启用  | 2026-09-20 17:34... | 2026-09-20 17:36... | 未配置             | -         | -          | <div>🔍 🗑️ ⌛</div>    |
| <input type="checkbox"/>            | 被依赖 workflow D成功              | 离线 | v1     | 未启用  | 2026-09-20 17:24... | -                   | 未配置             | user04006 | -          | <div>🔍 🗑️ ⌛</div>    |

共 114 条 / 12 页

< 1 2 3 4 5 ... 12 >

10 条/页 ▼





单次运行

☒ 是否是补数

按 DS 原生调度时间生成实例，时区：Asia/Shanghai。不会自动过滤来源表数据；取数条件由工作任务自身定义。

\* 补数时间

☒ 时间范围

☐ 指定时间列表

开始日期

→ 结束日期

📅

依赖模式

☒ 不执行下游依赖

☐ 执行下游依赖

执行方式

☒ 串行

☐ 并行

执行顺序

☒ 降序

☐ 升序

取消

单次运行

图 170 手动运行离线 workflow 补数

- 在运行配置启用补数，选择历史时间范围或指定时间点；指定时间点列表最多100项。按需要选择依赖范围、串行正序、串行倒序或并行执行。
- 补数使用历史调度时间生成运行，不会自动把SQL加上业务日期过滤。任务应通过参数正确限定补处理范围。
- 全量追加任务补数可能重复写入，应先核对目标处理策略。并行补数还需避免同一表或分区同时被覆盖。

### 5.1.6 实时 workflow 启停

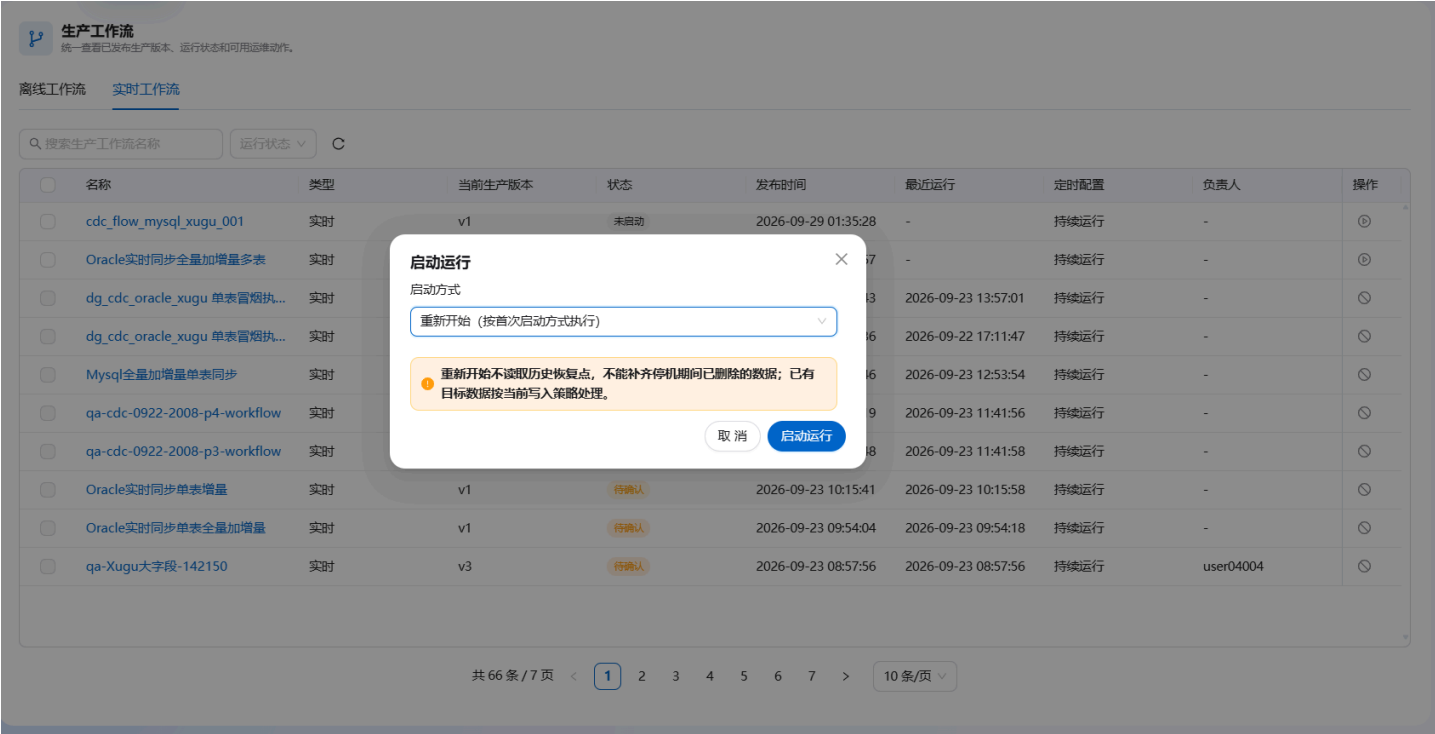


图 171 启动实时 workflow

说明：

- 在没有活动实例时点击启动，核对本本次启动方式、参数与恢复策略并确认。
- 到实时实例查看启动过程、节点状态和Checkpoint，不要只凭请求提交成功判断已稳定运行。

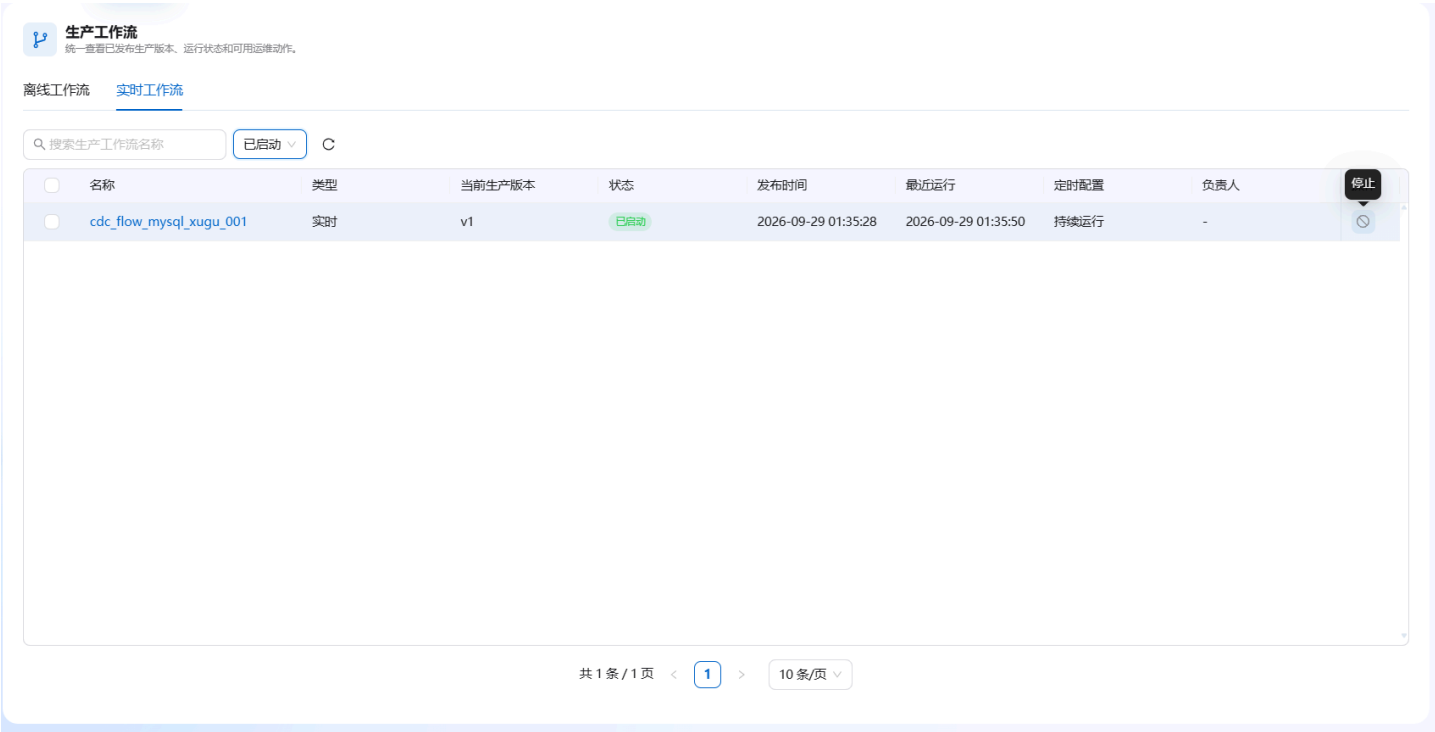


图 172 停止实时 workflow

说明：

- 选择停止，按要求填写原因并提交，等待停止结果和可用恢复信息。

- 停止后如需恢复，应使用实际可用的恢复点。停止请求不等于当前批次或检查点已经完成，应查看实例最终状态。

## 5.2 workflow实例

实例代表一次执行，保留实际参数、执行配置、节点状态与日志。离线区分生产实例和在线调试，实时区提供运行观测及检查点恢复信息。运维动作取决于实例当前状态与用户权限。

### 5.2.1 workflow实例列表

工作流实例

查看生产运行与在线调试固化的版本、参数、节点和运行证据。

离线工作流

实时工作流

生产实例

在线调试

搜索实例或工作流名称

实例状态

触发方式

触发时间起点

→

触发时间终点

清除

刷新

帮助

| <input type="checkbox"/> | 实例名称           | 所属工作流                | 运行版本 | 触发方式 | 状态 | 触发时间                | 实际开始时间              | 等待时长 | 执行时长 | 操作    |
|--------------------------|----------------|----------------------|------|------|----|---------------------|---------------------|------|------|-------|
| <input type="checkbox"/> | 20260928173500 | dg_mysql_xugu_...    | v2   | 调度触发 | 成功 | 2026-09-29 01:35... | 2026-09-29 01:35... | 2 秒  | 40 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928173053 | dg_sync_xugu_my...   | v1   | 手动运行 | 成功 | 2026-09-29 01:30... | 2026-09-29 01:30... | 0 秒  | 9 秒  | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928173000 | mysql_user_basic_... | v1   | 调度触发 | 失败 | 2026-09-29 01:30... | 2026-09-29 01:30... | 2 秒  | 32 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928173000 | dg_mysql_xugu_...    | v2   | 调度触发 | 成功 | 2026-09-29 01:30... | 2026-09-29 01:30... | 2 秒  | 40 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928173000 | dg_sync_mysql_xu...  | v1   | 调度触发 | 成功 | 2026-09-29 01:30... | 2026-09-29 01:30... | 2 秒  | 13 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928173000 | 离线工作流多版本...          | v9   | 调度触发 | 成功 | 2026-09-29 01:30... | 2026-09-29 01:30... | 2 秒  | 3 秒  | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928172500 | dg_mysql_xugu_...    | v2   | 调度触发 | 成功 | 2026-09-29 01:25... | 2026-09-29 01:25... | 1 秒  | 40 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928172000 | 离线工作流多版本...          | v9   | 调度触发 | 成功 | 2026-09-29 01:20... | 2026-09-29 01:20... | 1 秒  | 7 秒  | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928172000 | mysql_user_basic_... | v1   | 调度触发 | 失败 | 2026-09-29 01:20... | 2026-09-29 01:20... | 1 秒  | 43 秒 | 🔍 ⌂ ⋮ |
| <input type="checkbox"/> | 20260928172000 | dg_mysql_xugu_...    | v2   | 调度触发 | 成功 | 2026-09-29 01:20... | 2026-09-29 01:20... | 1 秒  | 40 秒 | 🔍 ⌂ ⋮ |

共 8934 条 / 894 页

1

2

3

4

5

...

894

10 条/页

图 173 离线生产实例列表及批量操作区

说明：

#### 查看离线生产实例

- 切换离线实例中的生产实例区域，按工作流、名称、状态、触发方式等条件查找。
- 记录实例标识、运行版本与触发时间，避免将同一工作流多次执行的日志混在一起。使用刷新更新状态。

#### 批量操作实例

- 在生产实例列表勾选符合相同操作条件的对象，选择允许的批量动作，单批最多100项。
- 提交前逐一核对状态和任务范围；查看每项处理结果，部分失败时仅处理未完成项。

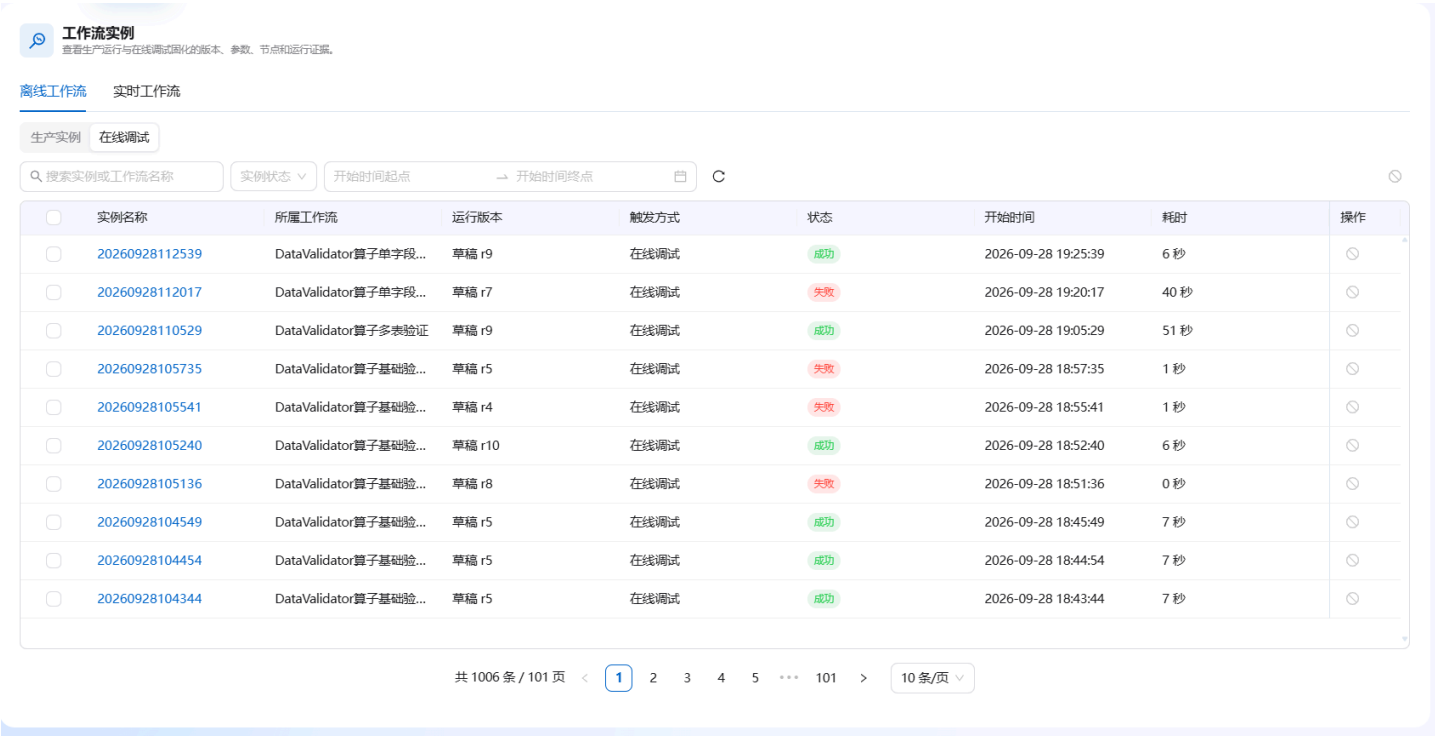


图 174 在线调试历史

说明：

- 切换在线调试区域查看开发过程产生的历史执行。
- 调试记录用于追溯，不通过这里执行生产恢复、重跑等管理动作。

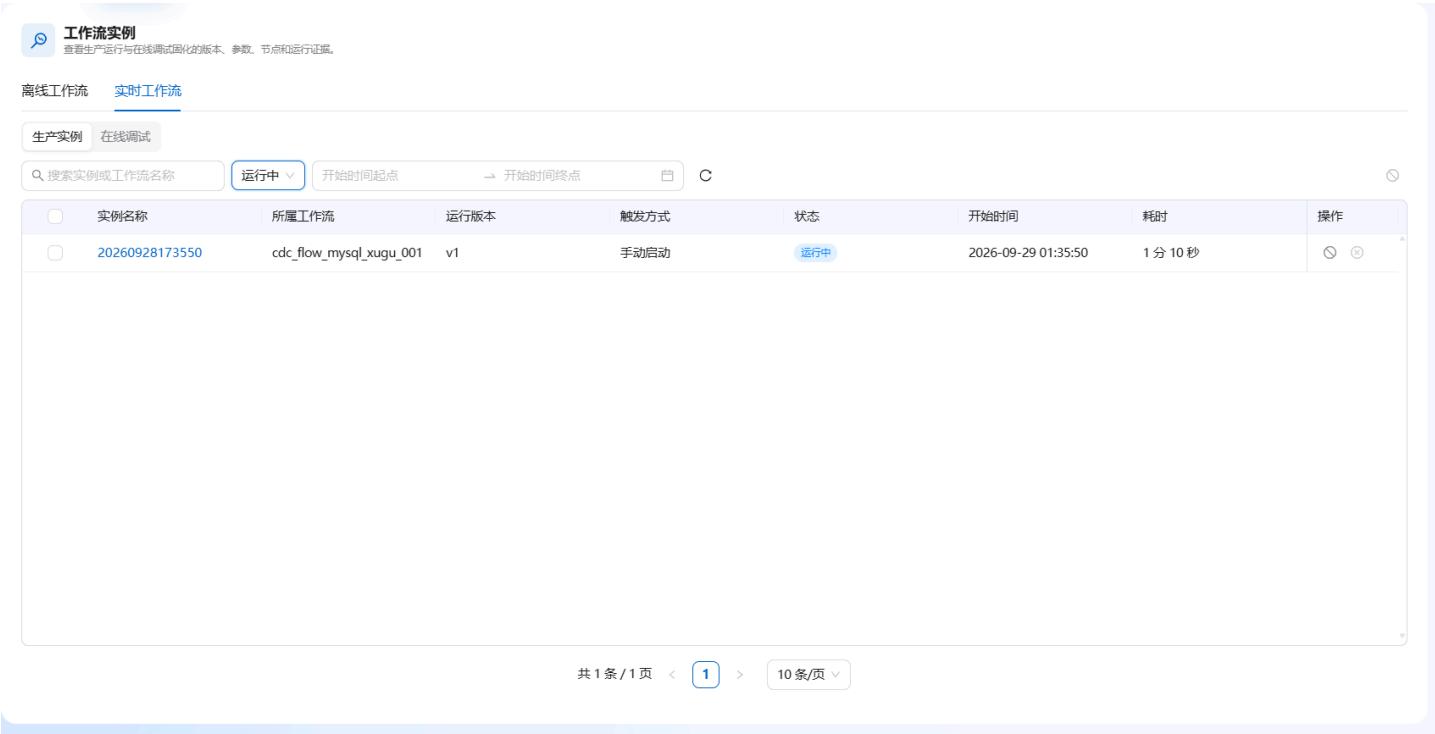
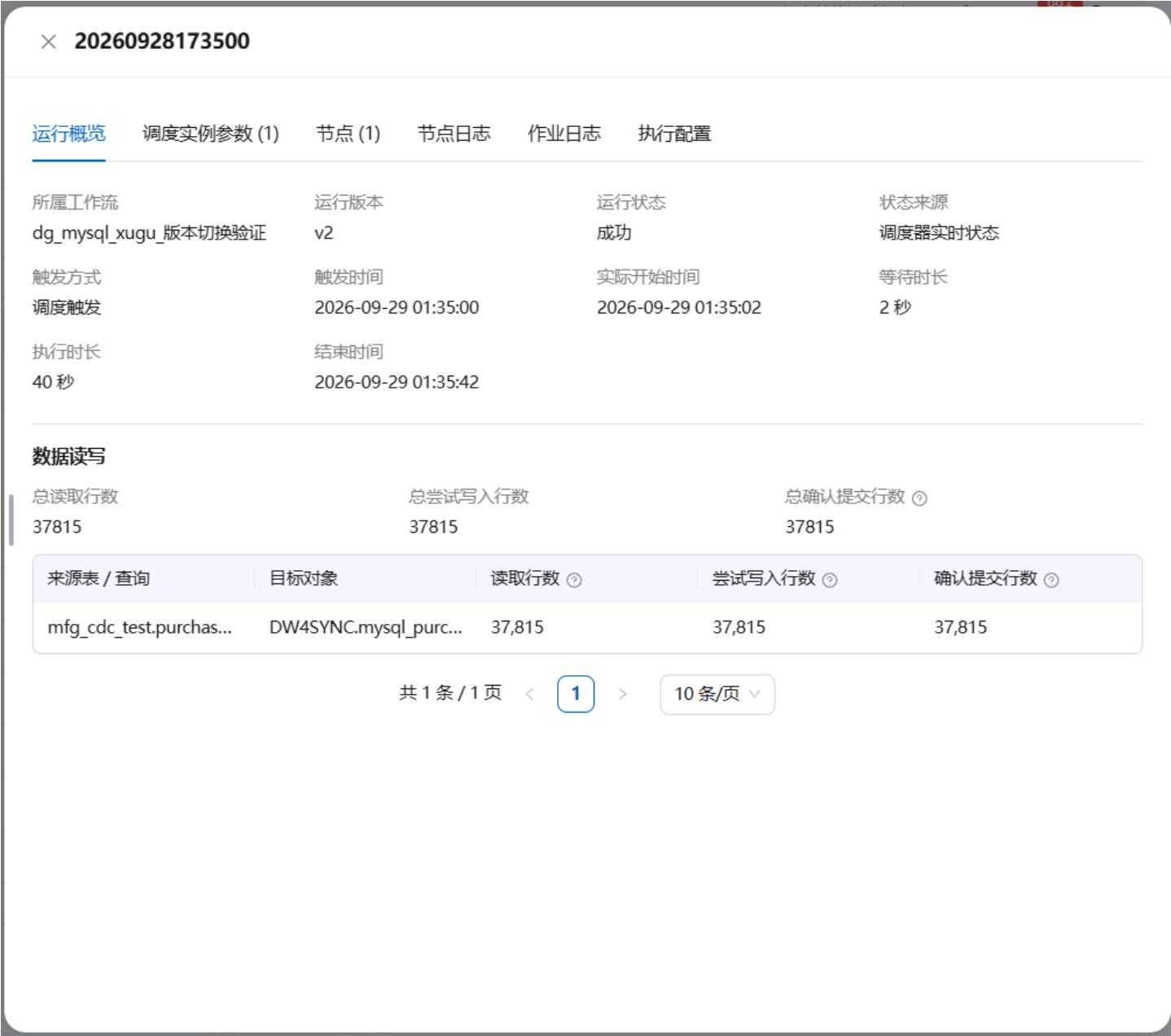


图 175 实时实例列表

说明：

- 切换实时实例，通过工作流、状态和健康等条件查找，打开需要查看的实例。
- 实时运行可能持续较长时间，运行状态与健康状态需结合观测数据判断。

5.2.2 离线实例详情



数据读写

总读取行数

37815

总尝试写入行数

37815

总确认提交行数

37815

来源表 / 查询

目标对象

读取行数

尝试写入行数

确认提交行数

mfg\_cdc\_test.purchas...

DW4SYNC.mysql\_purc...

37,815

37,815

37,815

共 1 条 / 1 页

1

10 条/页

图 176 查看离线运行概览

- 说明：
- 打开离线实例详情的运行概览，查看开始结束时间、耗时与结果状态。
  - 整体失败时继续定位失败节点；整体成功后仍应按业务要求核对目标表数据。
  - 对跨存储采集、复制及迁移实例，结合运行结果、节点日志及目标数据核对处理情况。实例成功表示本次配置流程执行完成，不自动代表来源与目标数据已通过一致性验收。

运行概览 调度实例参数 (1) 节点 (1) 节点日志 作业日志 执行配置

scheduleTime · 来源: 调度引擎: 2026-09-29 01:35:00

- 存在调度参数时切换调度参数，查看本次传入值和业务时间。
- 这是实例实际参数，不是当前开发草稿；补数定位时重点核对历史业务日期。

offline-sync:160d6964-  
15b0-4fdf-b67a-  
972d968a1274\_trigger

成功

```
1 {  
2   "env": {  
3     "job.name": "dg_mysql_xugu_版本切换验证-v2",  
4     "job.mode": "BATCH",  
5     "parallelism": 4,  
6     "job.retry.times": 3,  
7     "job.retry.interval.seconds": 3,  
8     "read_limit.rows_per_second": 500  
9   },  
10  "source": [  
11    {  
12      "url": "jdbc:mysql://192.168.31.213:3306/mfg_cdc_test?useSSL=false&serv  
13      "driver": "com.mysql.cj.jdbc.Driver",  
14      "dialect": "MySQL",  
15      "user": "root",  
16      "password": "*****",  
17      "plugin_name": "DG_JDBC",  
18      "plugin_output": "source_0",  
19      "table_list": [  
20        {  
21          "table_path": "mfg_cdc_test.purchase_orders"  
22        }  
23      ],  
24      "split.size": 8096,
```

图 177 查看离线执行配置

#### 说明:

- 切换执行配置，检查本次运行使用的配置快照。
- 确认来源、目标和写入策略，避免使用后来修改的工作流草稿解释历史运行。
- 核对本次执行的来源、目标、采集范围及写入方式，确认与复制或迁移计划一致。重跑、补数前，应检查目标端已写入的数据，避免追加重复或覆盖非预期范围。

运行概览    调度实例参数 (1)    节点 (1)    节点日志    作业日志    执行配置

| 节点                           | 类型        | 状态 | 状态原因 | 开始时间                | 读取行数  |
|------------------------------|-----------|----|------|---------------------|-------|
| offline-sync:160d6964-15b... | SEATUNNEL | 成功 | —    | 2026-09-29 01:35:02 | 37815 |

图 178 查看离线节点

说明：

- 切换节点，查看各节点执行状态和耗时，定位失败、等待或耗时较长的节点。
- 节点之间的等待可能来自依赖关系，不应直接认定为数据库连接失败。

运行概览 调度实例参数 (1) 节点 (1) 节点日志 作业日志 执行配置

offline-sync:160d6964-15b0-4fdf-b67a-972d968a1274\_trigger

成功

在已加载日志中搜索

已加载 58 / 共 58 行

🔄 📄 📋 ⬇️ ⬆️ ⬇️

```

32
33 2026-09-29 01:35:02.449 INFO - tenantCode :datagolden_40, task dir:/tmp/dolphinscheduler/exec/proces
34 2026-09-29 01:35:02.449 INFO - generate script file:/tmp/dolphinscheduler/exec/process/216689/seatur
35 2026-09-29 01:35:02.449 INFO - SeaTunnel task command: ${SEATUNNEL_HOME}/bin/datagolden-seatur
36 2026-09-29 01:35:02.449 INFO - Final Script Content:
37 =====
38 #!/bin/bash
39 BASEDIR=$(cd `dirname $0`; pwd)
40 cd $BASEDIR
41 ${SEATUNNEL_HOME}/bin/datagolden-seatunnel.sh --config /tmp/dolphinscheduler/exec/process/216689
42 =====
43 2026-09-29 01:35:02.449 INFO - Executing shell command : sudo -u datagolden_40 -i /tmp/dolphinsched
44 2026-09-29 01:35:02.450 INFO - process start, process id is: 1766506
45 2026-09-29 01:35:42.045 INFO - -> HTTP/1.1 200
46 2026-09-29 01:35:42.045 INFO - -> Content-Type: application/json
47 2026-09-29 01:35:42.045 INFO - -> Transfer-Encoding: chunked
48 2026-09-29 01:35:42.045 INFO - -> Date: Mon, 28 Sep 2026 17:35:42 GMT
49 2026-09-29 01:35:42.045 INFO - -> Connection: close
50 2026-09-29 01:35:42.045 INFO - ->
51 2026-09-29 01:35:42.045 INFO - -> 170
52 2026-09-29 01:35:42.045 INFO - -> {"data":{"key":"seatunnel-job","name":"dg_mysql_xugu_版本切换验证
53 2026-09-29 01:35:42.045 INFO - -> 0
54 2026-09-29 01:35:42.045 INFO - ->
55 2026-09-29 01:35:42.046 INFO - process has exited, execute path:/tmp/dolphinscheduler/exec/process/2
56 2026-09-29 01:35:42.185 INFO - 🏁 Finalize Task Instance
57 2026-09-29 01:35:42.185 INFO - FINALIZE_SESSION
58 2026-09-29 01:35:42.185 INFO - Deleted task exec directory: /tmp/dolphinscheduler/exec/process/21668

```

图 179 查看离线节点日志

说明：

- 切换节点日志并选择目标节点，查看执行过程与错误位置。
- 先确认节点和实例对应，再分析连接、SQL、权限或数据问题；分享日志前检查敏感信息。
- 切换作业日志，查看整体提交、调度和作业执行信息。
- 节点日志与作业日志用于不同层次定位；二者时间和状态应结合起来判断。

### 5.2.3 实时实例详情

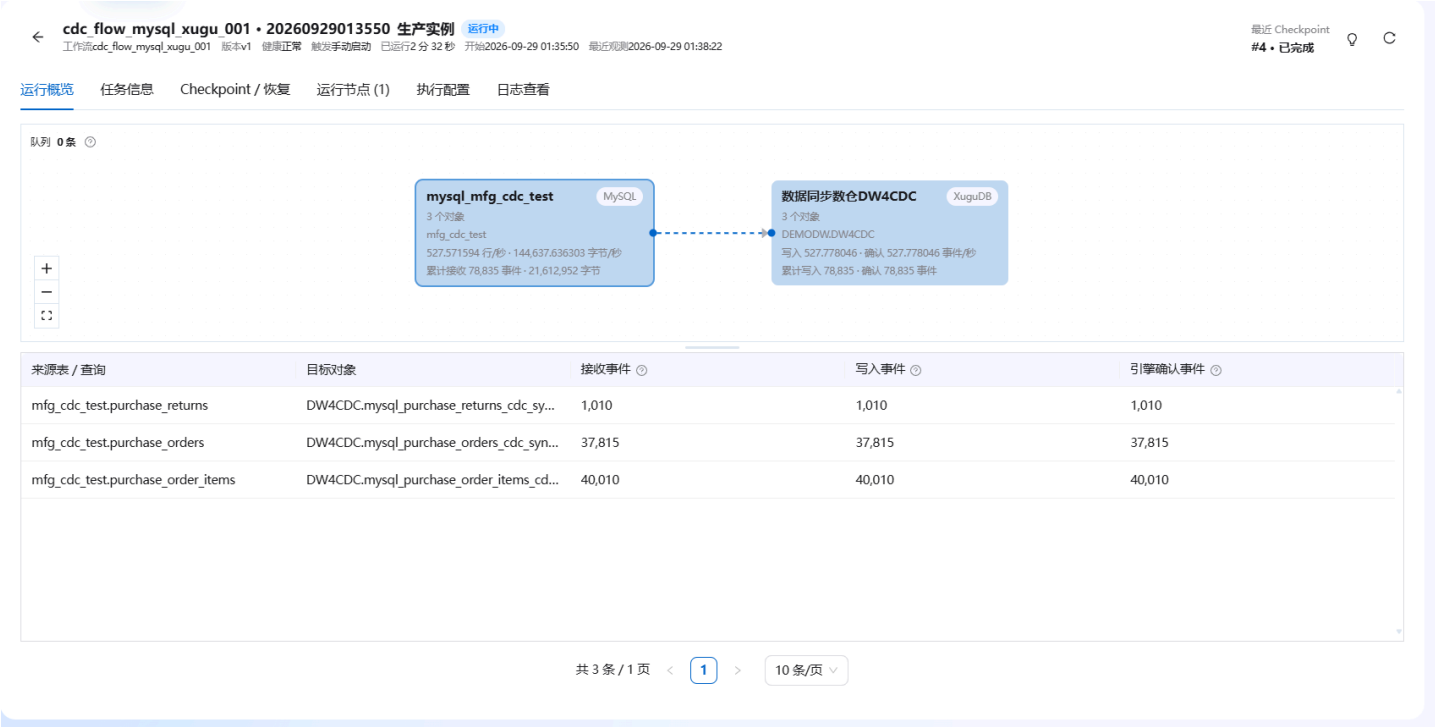


图 180 查看实时运行概览

- 说明：
- 打开实时运行概览，查看拓扑、输入接收、输出写入、检查点确认等指标和节点状态。
  - 输入接收、输出写入与检查点确认统计含义不同，不应把瞬时差值直接视为数据丢失。无观测数据不等于数值为0。
  - 存在调整建议时，通过去调整返回开发配置，修改并发布新版本；建议本身不会直接修改运行参数。
  - 本次具有参数时打开实际参数，核对运行输入值。
  - 排查历史实例时使用此处记录，不使用后来修改的默认值。
  - 对跨存储持续同步实例，结合输入接收、输出写入、节点状态及Checkpoint信息检查运行情况。用于迁移过渡时，应核对初始数据与后续变化数据的衔接情况，确认目标端满足业务要求后再安排访问切换。



图 181 查看实时执行配置

说明：

- 打开执行配置，核对本实例使用的版本与配置快照。
- 升级后检查实例是否使用预期新版本，确认连接、订阅范围与输出策略未误变更。

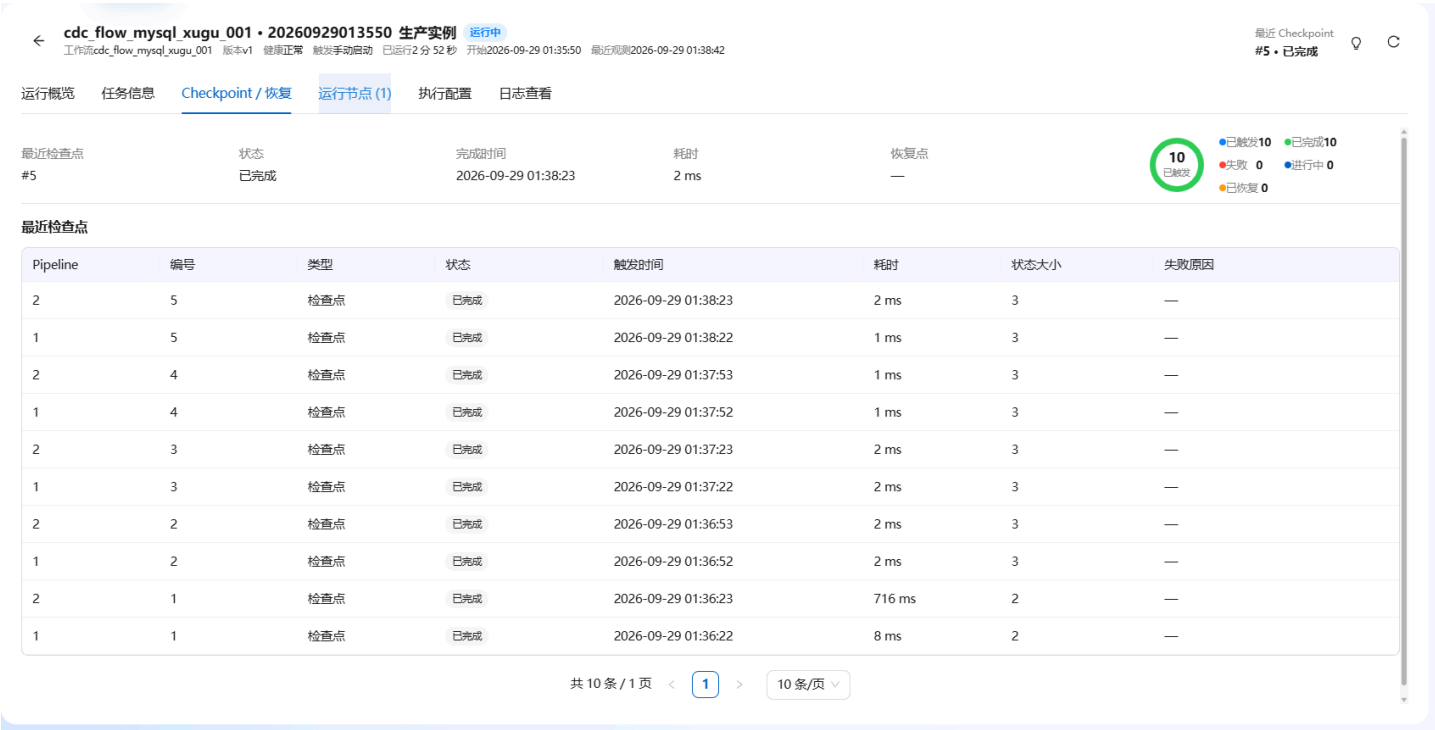


图 182 查看Checkpoint与恢复

说明：

- 打开Checkpoint与恢复，查看检查点成功、失败、耗时和可用恢复点信息。

- 恢复点表示已保存的处理进度；是否可完整恢复还取决于外部日志、消息保留及目标写入条件。不存在成功恢复点时，不能假定自动无损续跑。

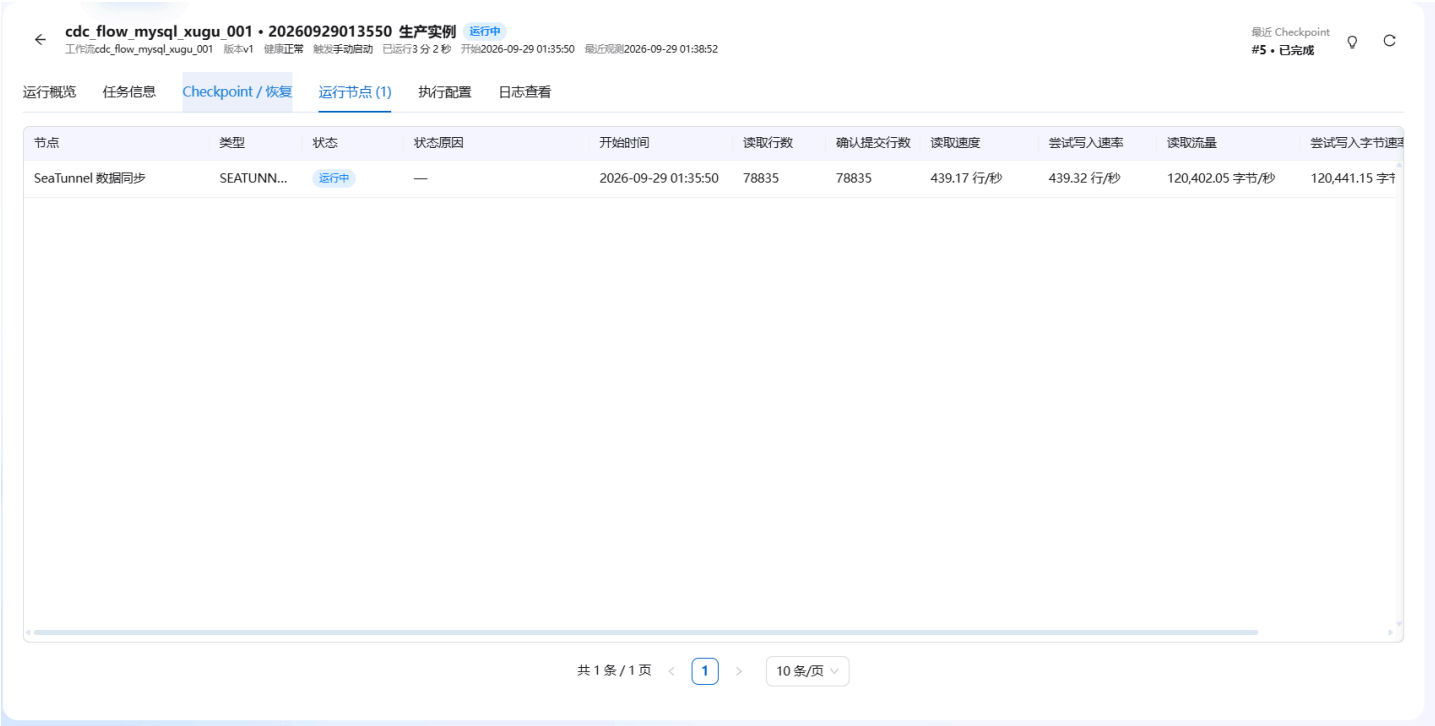


图 183 查看实时运行节点

说明：

- 打开运行节点，查看节点状态和对应运行信息，定位输入、转换或输出环节。
- 持续积压时结合输入、输出速率与Checkpoint情况分析，不只调整并行度。
- 进入查看日志，按运行对象查看错误和执行过程。
- 旧实例若显示已停止采集的历史信息，应作为历史记录阅读，不将其视为当前实时监控。

5.2.4 实例控制与恢复

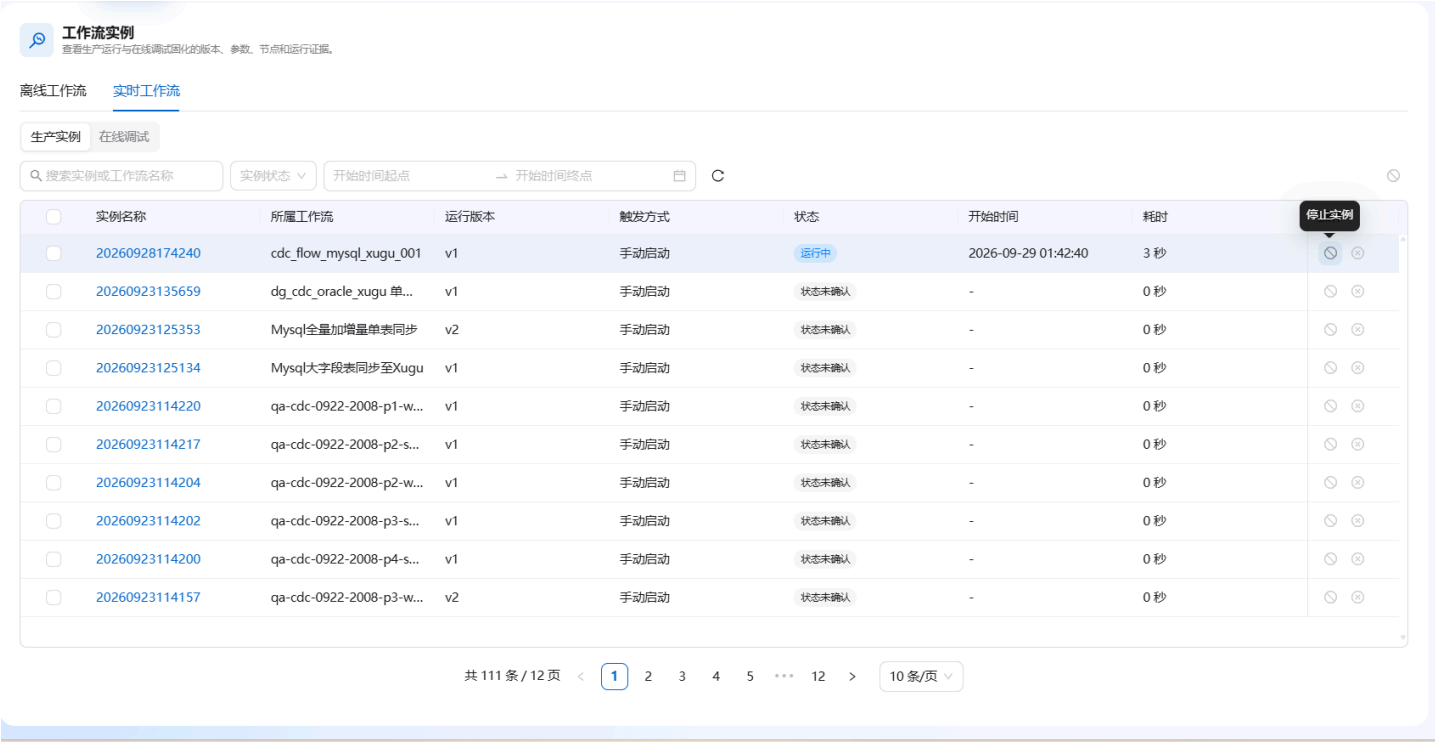


图 184 停止实例

说明：

- 对允许停止的活动实例使用停止，按要求确认原因并等待状态变化。
- 停止只终止后续执行过程，不承诺回滚所有已写入数据；重跑前检查目标部分写入情况。

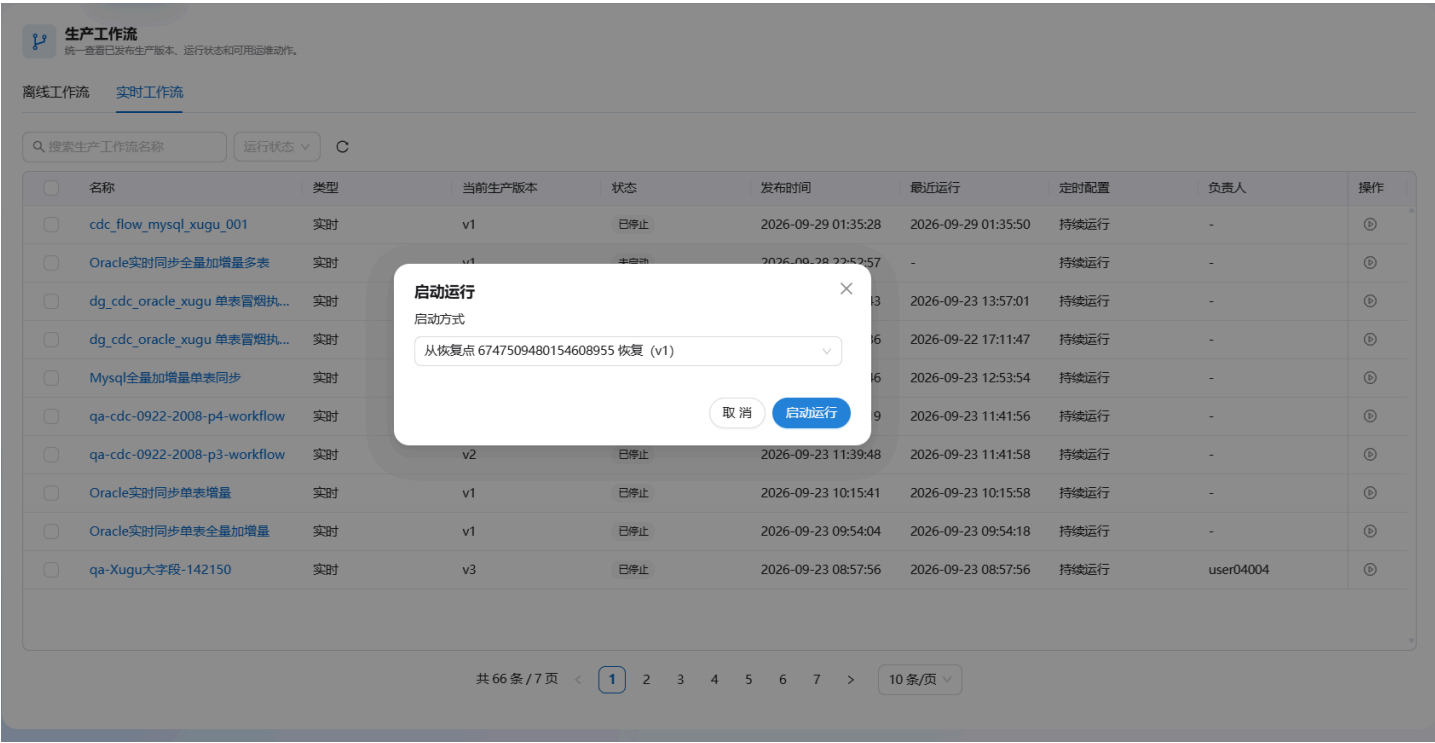


图 185 恢复实例

说明：

- 在允许恢复的实例上选择恢复，核对可用恢复条件后提交。
- 恢复不是任意状态都可用；实时恢复应确认恢复点存在且外部日志仍可访问。

- 跨存储任务恢复前，应核对来源数据或日志是否仍可读取、目标连接是否可用，以及已写入数据与恢复策略是否匹配。恢复不会自动撤销此前产生的目标写入结果。

工作流实例

查看生产运行与在线调试固化的版本、参数、节点和运行证据。

离线工作流 实时工作流

生产实例 在线测试

Q 搜索实例或工作流名称

实例状态

开始时间起点

→ 开始时间终点

🗑

🔍

| <input type="checkbox"/> | 实例名称           | 所属工作流                    | 运行版本 | 触发方式 | 状态    | 开始时间                | 耗时  | <div>强制终止</div> |
|--------------------------|----------------|--------------------------|------|------|-------|---------------------|-----|-----------------|
| <input type="checkbox"/> | 20260928174240 | cdc_flow_mysql_xugu_001  | v1   | 手动启动 | 运行中   | 2026-09-29 01:42:40 | 3 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923135659 | dg_cdc_oracle_xugu_单...  | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923125353 | Mysql全量增量单表同步            | v2   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923125134 | Mysql大字段表同步至Xugu         | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114220 | qa-cdc-0922-2008-p1-w... | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114217 | qa-cdc-0922-2008-p2-s... | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114204 | qa-cdc-0922-2008-p2-w... | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114202 | qa-cdc-0922-2008-p3-s... | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114200 | qa-cdc-0922-2008-p4-s... | v1   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |
| <input type="checkbox"/> | 20260923114157 | qa-cdc-0922-2008-p3-w... | v2   | 手动启动 | 状态未确认 | -                   | 0 秒 | <div>🔄 ⏸</div>  |

共 111 条 / 12 页 < 1 2 3 4 5 ... 12 > 10 条/页

图 187 强制终止实例

- 说明：
- 普通停止无法完成且页面允许时，由具有权限的人员使用强制终止，填写原因并确认风险。
  - 强制终止可能无法完成正常清理或进度保存。执行后核对外部资源与目标数据，不能把它作为常规停止替代。

5.3 工作流告警与通知

工作流告警用于及时了解工作流执行结果和节点超时情况。执行结果通知在工作流运行或调度设置中配置，节点超时告警在节点参数中配置；产生的通知可在平台管理域的告警中心查看。

5.3.1 配置执行结果通知

- 说明：
- 在工作流运行或调度设置中，选择告警类型：
    - 不告警：不发送工作流执行结果通知。
    - 失败告警：工作流执行失败时发送通知。
    - 成功告警：工作流执行成功时发送通知。
    - 全部告警：工作流执行成功或失败时均发送通知。
  - 根据页面提示保存配置；周期调度使用已生效的调度配置，单次运行使用本次运行设置。
  - 工作流通知通过当前租户的默认告警组发送，无需逐个选择接收人。

- 如果页面提示默认通知接入未就绪，请由有权限的管理员按提示初始化或同步调度引擎，确认成功后再操作。
- 保存配置不代表已经产生通知；通知在符合所选条件时由调度引擎触发。

调度设置 · workflow实例删除-非锁定1

任务默认策略

任务优先级

中

任务超时

超时检测

关闭

失败重试

失败重试次数

0

重试间隔 (分钟)

1

告警类型

不发送通知

不发送通知

失败时通知

成功时通知

成功或失败时通知

5.3.2 配置节点超时告警

说明：

- 在支持超时设置的工作流节点中，开启超时检查，按页面标注的单位填写超时时间，并选择超时处理策略。
- 需要收到超时通知时，选择包含告警的策略；是否同时使节点失败，以所选策略为准。
- 节点超时告警与工作流执行结果通知独立。工作流选择“不告警”时，已配置的节点超时告警仍可发送。
- 同一次运行可能先后产生节点超时和工作流失败通知，消息到达顺序不保证与发生顺序一致。

调度设置 · 工作流实例删除-非锁定1



任务默认策略

任务优先级 ?

中



任务超时

超时检测 ?

开启



超时通知策略 ?



告警



失败

任务超时时长 (分钟) ?

60

失败重试

失败重试次数 ?

0

重试间隔 (分钟) ?

1

告警类型

不发送通知



变更说明

取消

保存

### 5.3.3 查看工作流通知

说明：

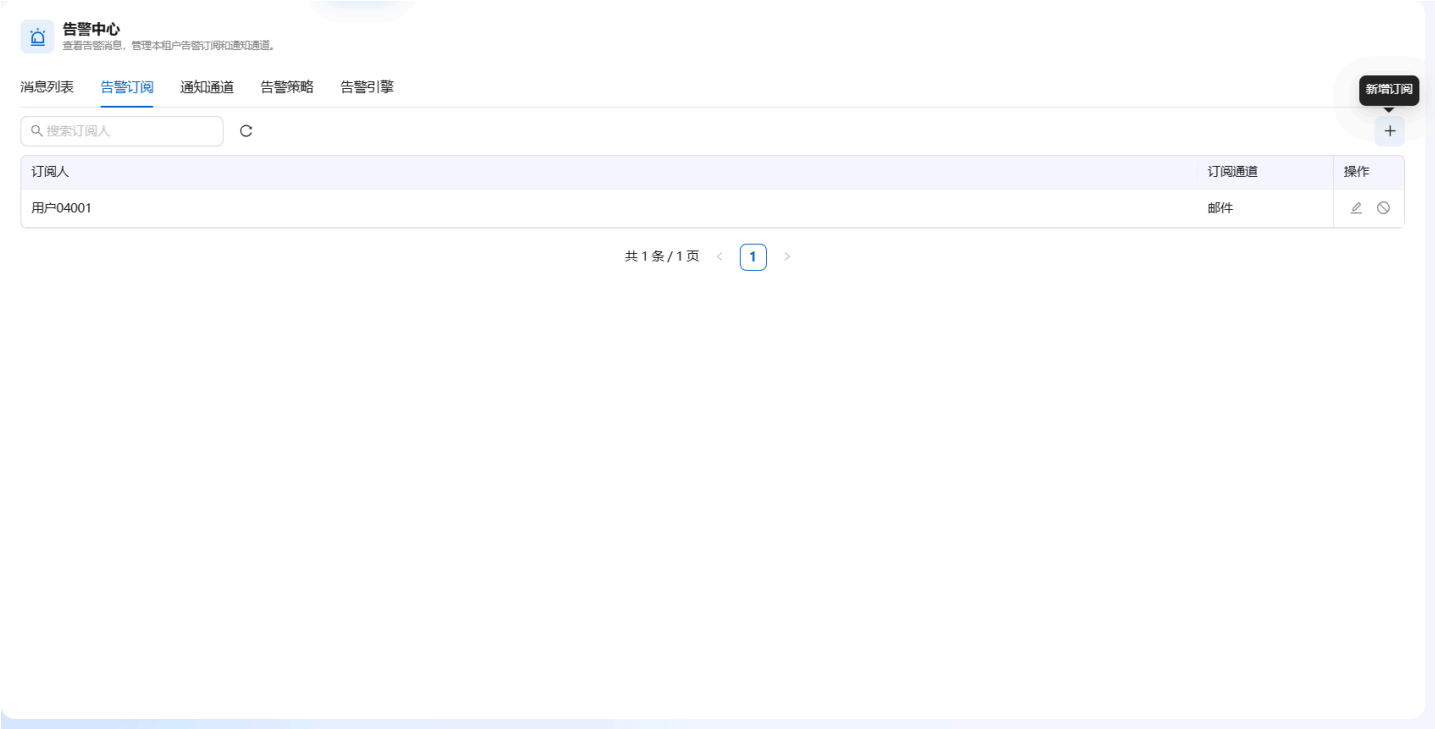
- 进入 平台管理域 → 告警中心 → 消息列表，查看工作流执行成功、执行失败或节点超时通知。
- 根据消息中的工作流和实例信息，进入数据开发域的工作流实例页面，查看运行状态、节点详情和日志，定位问题。
- 站内信按当前租户共享，无需新增告警订阅；各用户的已读状态分别记录。
- 将消息标记为已读仅表示已查看，不会停止工作流，也不代表运行故障已经解决。
- 工作流执行成功通知表示本次执行成功，不自动表示此前失败已恢复或完成处置。
- 收到跨存储采集、复制或迁移任务的失败、超时通知后，根据实例信息检查来源读取、转换处理及目标写入环节。完成异常处理后，应核对后续执行结果和目标数据；消息已读不代表任务恢复或迁移完成。



### 5.3.4 配置邮件和短信订阅

说明：

- 需要额外接收邮件或短信时，进入 平台管理域 → 告警中心 → 通知通道，配置并启用对应通道。
- 进入 告警订阅，点击新增，选择当前租户的订阅人，勾选邮件、短信后保存。
- 订阅人的邮箱和手机号来自用户管理；联系方式缺失或格式不正确时，需先维护用户信息。
- 当前订阅按租户和用户生效，不支持单独选择某个工作流，也不支持仅订阅工作流失败消息。订阅人还可能收到当前租户的质量、机器资源等通知。
- 修改或取消订阅不影响站内信，也不改变工作流的告警类型和超时设置。



### 5.3.5 使用注意事项

- 说明：
- 当前不提供独立的工作流“告警规则”新建、编辑、删除和启停操作。
  - 队列、吞吐、Checkpoint等运行指标可用于查看和诊断，当前不支持通过自定义工作流告警规则配置这些指标的阈值和持续时间。
  - 工作流被停止时，调度引擎可能按失败结果发送通知，应结合实例状态和操作记录判断原因。
  - 调度引擎或通知通道异常可能影响消息送达。未收到通知时，应先查看实例实际状态，再检查告警类型、默认通知接入及通知通道；不能据此判断工作流运行正常。

