

操作手册

TDMP 数据脱敏系统操作手册

2019 年 04 月

目 录

1. 平台操作.....	6
1.1 登陆界面.....	6
1.2 主页.....	6
1.3 资料库配置.....	7
1.4 用户.....	7
1.4.1. 新增用户.....	8
1.4.2. 激活用户.....	9
1.4.3. 修改用户.....	10
1.4.4. 重置密码.....	10
1.4.5. 解锁用户.....	11
1.4.6. 锁定用户.....	11
1.4.7. 删除用户.....	12
1.5 权限.....	12
1.5.1. 角色.....	12
1.5.1.1. 新增角色.....	13
1.5.1.2. 修改角色.....	13
1.5.1.3. 删除角色.....	14
1.5.2. 用户组.....	14
1.5.2.1. 新增用户组.....	14
1.5.2.2. 删除用户组.....	15
1.5.2.3. 查看用户组详情.....	15
1.6 服务器配置.....	15
1.6.1. 数据库服务器.....	15
1.6.1.1. 新增.....	16
1.6.1.2. 修改.....	17

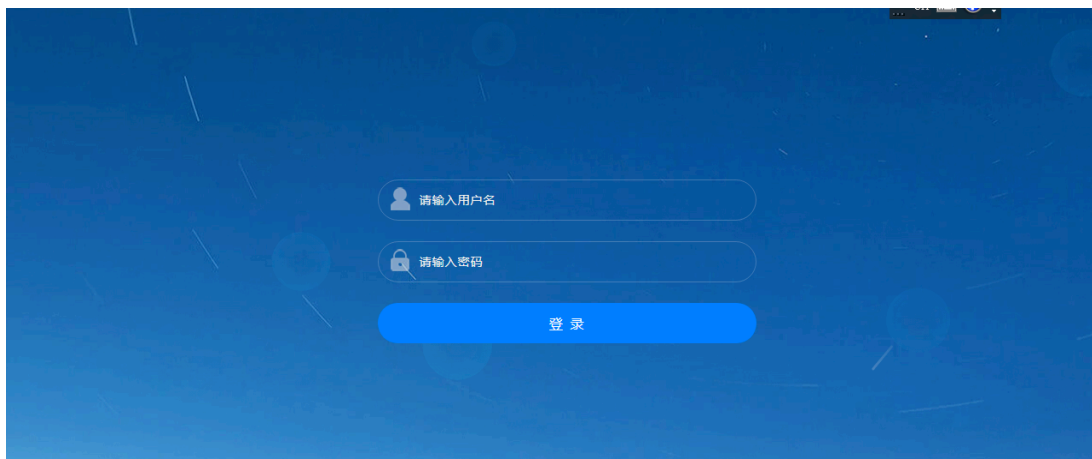
1.6.1.3. 删除	17
1.6.2. 结构化文本服务器.....	18
1.6.2.1. 新增	18
1.6.2.2. 修改	19
1.6.2.3. 删除	19
1.6.3. 远程结构化文本服务器.....	19
1.6.3.1. 新增	20
1.6.3.2. 修改	21
1.6.3.3. 删除	21
2. 规则管理.....	22
2.1 敏感数据发现规则	22
2.2 数据脱敏规则	23
2.3 脱敏策略	24
3. workflow配置.....	26
3.1 敏感元数据管理	26
3.1.1. 新增数据元数据管理.....	26
3.1.1.1. 脱敏产品模式数据元数据.....	26
3.1.1.1.1. 常规配置模式.....	27
3.1.1.1.2. 模板导入模式.....	28
3.1.2. 修改敏感元范围.....	31
3.1.3. 修改分析结果.....	31
3.1.4. 删除敏感数据元.....	31
3.2 数据集配置	32
3.2.1 新增数据集	32
3.2.2 删除数据集	33
3.3 任务配置	34
3.3.1 新增任务配置	34
3.3.1.1 脱敏产品内存方式.....	34
3.3.1.2 脱敏产品文件落地方式	36

- 3.3.2 打开任务配置 38
- 3.3.3 关闭 workflow 38
- 3.3.4 修改任务配置 38
- 3.3.5 删除 workflow 38
- 4. 任务管理 39**
 - 4.1 UDF 监控 39
 - 4.2 脱敏产品监控 41
 - 4.3 SFTP/FTP 远程文本脱敏监控 44
 - 4.4 SFTP/FTP 源文件格式化文本->目标数据库 (GP) 监控 45
 - 4.5 定时调度 46
- 5. 日志 47**
- 6. 报表 48**

1.平台操作

1.1 登陆界面

通过火狐或谷歌浏览器访问平台地址（http://172.22.246.171:8080/opt3）登录脱敏产品测试数据管理平台页面，如下图：



1.2 主页

登录平台，进入主页界面。如下图：所示，为管理员用户登录后看到的管理菜单。

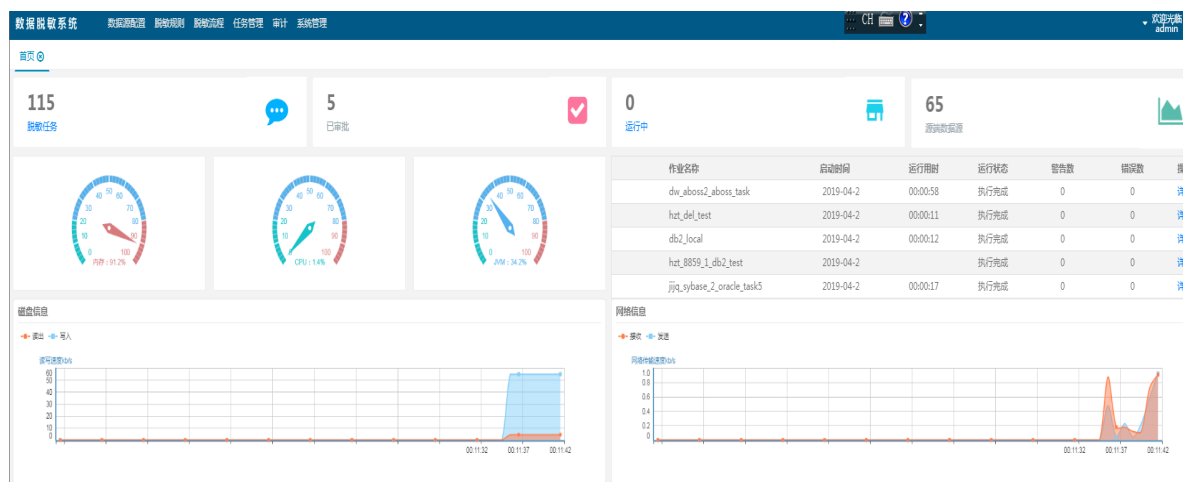


图 1.2.1 主页

在主页界面右上角修改密码、编辑个人资料（用户退出），如下图：

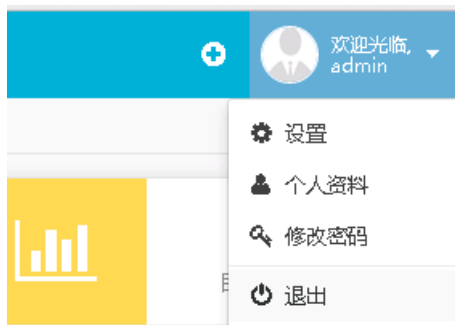


图 1.2.2 登录用户信息

1.3 资料库配置

资料库配置是系统管理子菜单,脱敏产品 3.0 初始化之前需要配置资料库信息,作为脱敏产品 3.0 的元数据参数库。文件配置路径:

/root/脱敏产品 3.0/webapps/opt3/config/jdbc.properties 文件中,填写相应的 IP 地址,数据库名称,登录端口号,数据库用户名与密码。

登录页面点击资料库配置,能看到 jdbc.properties 中的配置信息如下图:

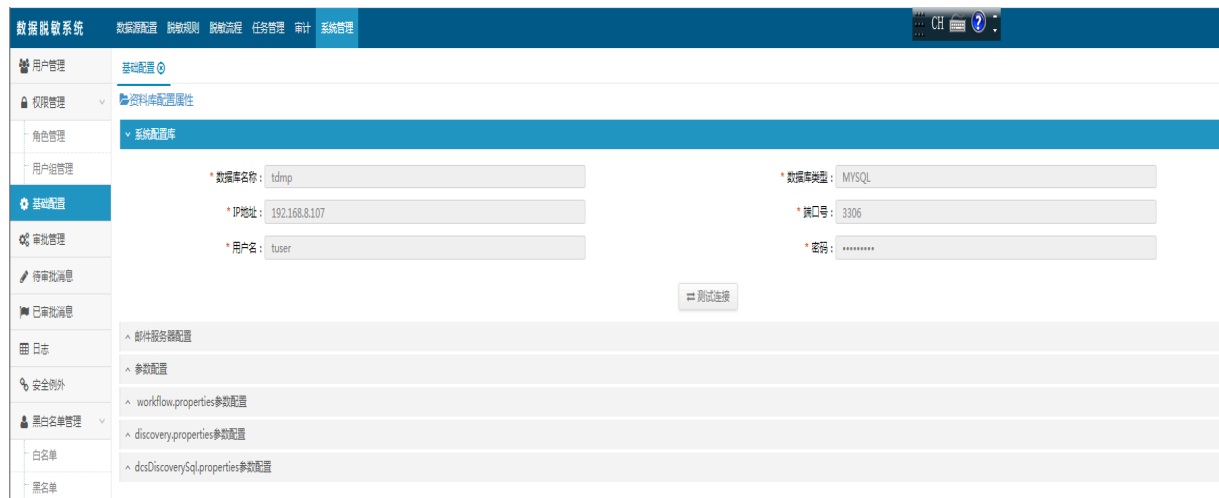


图 1.3 资料库配置

1.4 用户

用户管理的遵循以下管理原则,从上而下,由超级管理员,到管理员,再到操作员、审计员的管理顺序;不可越级管理,即超级管理员只能管理管理员角色的用户,管理员管理操作员和审计员的角色。



1.4.1. 新增用户

新增用户需填写用户名，用户描述信息，选择角色和所属域，填写密码和联系邮箱、联系电话。

注：所属域决定了用户可以访问服务器配置中的哪些数据库，如果当前用户和服务器配置中数据库不再同一个域，那么用户无权访问和操作该服务器中数据库。

点击右上角新增用户如下图：



新增用户

* 用户名: 长度: 3-30, 只包含中文、字母、下划线和数字, 不能以数字开头

* 密码: 长度: 6-20位, 由数字、字母、[特殊字符] 组成

* 确认密码:

* 密码有效天数: (范围: 1-1000)

手机:

邮箱:

* 角色:

* 用户组:

备注: 剩余字数: 100字

图 1.4.1 新增用户

1.4.2. 激活用户

新建的用户, 用户状态处于初始化状态, 需激活该用户, 退出系统重新登录。

用户密码修改

提示: 初始化, 修改密码

原密码 *

新密码 * 长度: 6-20位, 由数字、字母、[特殊字符] 组成

确认密码 *

图 1.4.2 激活用户

1.4.3. 修改用户

选中需修改的用户，点击右侧选择修改用户，修改用户的各种信息，修改完毕，点击完成即可 如下图：

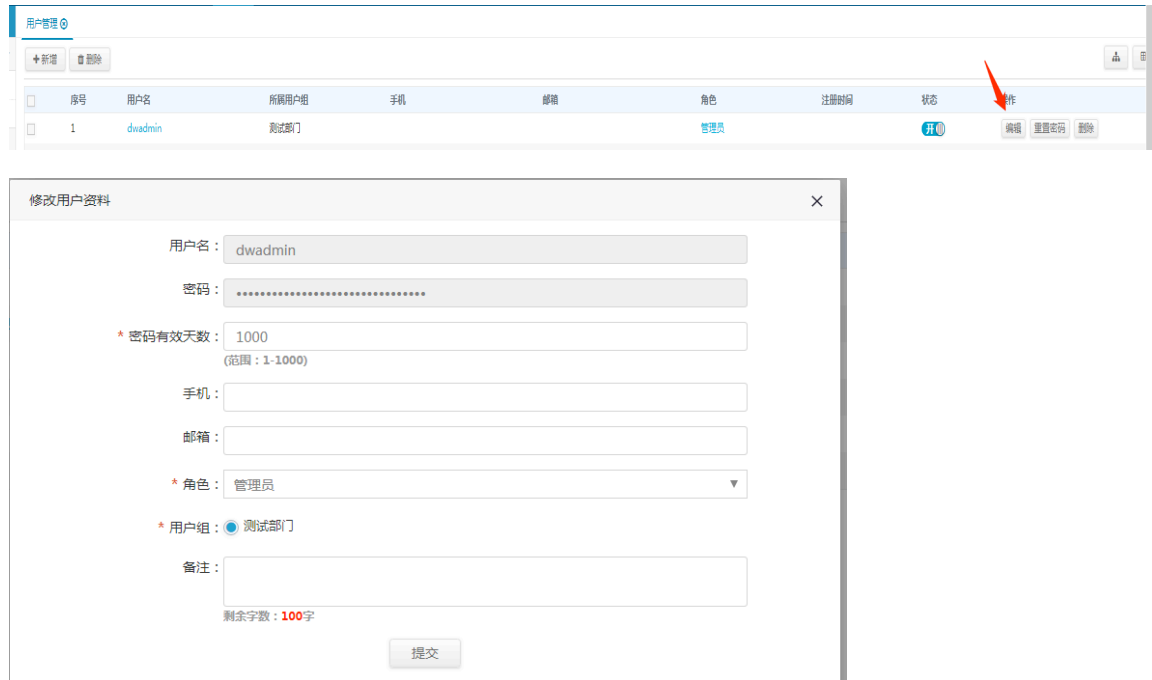
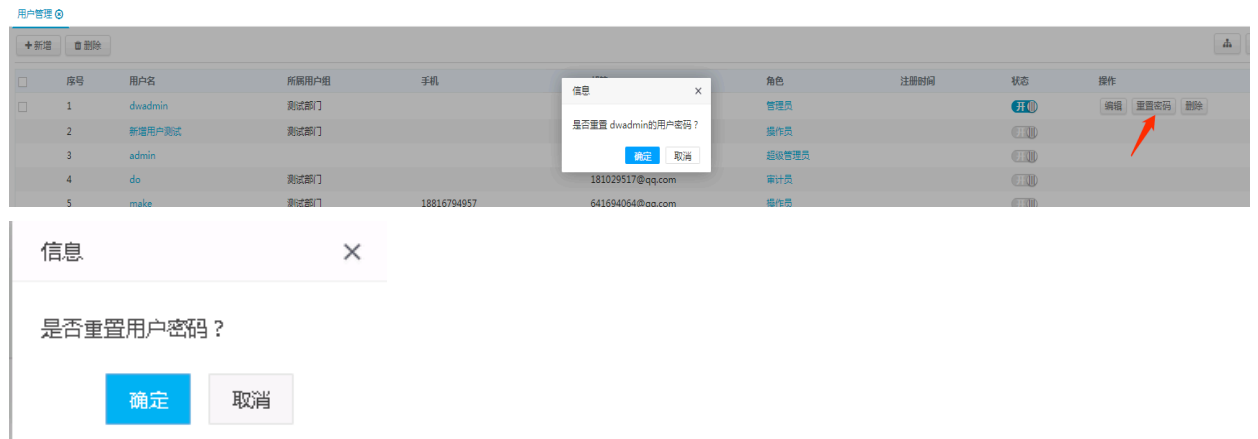


图 1.4.3 修改用户

1.4.4. 重置密码

修改所创建用户密码操作，进入修改用户选择重置密码，此时会弹出一个临时密码，重新用此用户登录输入临时密码，进行修改密码操作，点击完成即可修改密码成功，如下图：





1.4.5. 解锁用户

用户登录密码输入错误超过相应的次数，此时账号就会被锁定，需要对此用户解锁根据提示进行修改。



对于角色为操作员和审计员的用户

图 1.4.5 解锁用户

1.4.6. 锁定用户

锁定账号，选择状态关闭如下图：



图 1.4.6 锁定用户

1.4.7. 删除用户

删除用户，右侧操作中选择删除用户即可：如下图：

序号	用户名	所属用户组	手机	邮箱	角色	注册时间	状态	操作
1	dwadmin	测试部门			管理员		关	编辑 重置密码 删除
2	新增用户测试	测试部门			操作员		开	
3	admin				超级管理员		开	
4	do	测试部门		181029517@qq.com	审计员		开	



图 1.4.7 删除用户

1.5 权限

用户权限板块用于限制用户功能模块使用，分为角色和域两部分。

1.5.1. 角色

角色主要包含用户对哪个功能模块具有操作权限的限定。

系统默认内置四个角色权限，超级管理员、管理员、操作员、审计员。

超级管理员拥有系统最高权限，能操作系统所有功能。而管理员、操作员和审计员只有部分功能权限。

序号	角色名称	描述
1	审计员	系统内置角色。策略查询、工作流查询、任务查询、审计、用户查询、审批查询、日志
2	操作员	系统内置角色。策略查询、工作流配置、任务管理、审批查询、资源监控、消息查询、日志
3	管理员	系统内置角色。具有除权限管理外的所有功能模块权限
4	超级管理员	系统内置角色。具有所有功能模块权限

1.5.1.1. 新增角色

点击右上角选择新增角色，根据需求选择相应的操作功能模块如下图：

角色名称： 长度2-30，只包含中文、字母、下划线和数字，不能以数字开头

角色描述： 剩余字数：200字

权限模板：

权限分配：

- 数据源配置**
 - 文件 查询 新增 编辑 删除
 - 关系型数据库 查询 新增 编辑 删除
 - 大数据 查询 新增 编辑 删除
- 规则管理**
 - 敏感发现规则 查询 新增 编辑 删除
 - 脱敏算法 查询 新增 编辑 删除
 - 脱敏策略 查询 新增 编辑 删除
- 脱敏流程**
 - 一键脱敏
 - 高级配置
 - 敏感元配置 查询 新增 编辑 删除 审核
 - 数据集配置 查询 新增 编辑 删除
 - 任务配置 查询 新增 编辑 删除
- 任务管理**
 - 监控执行 查询 执行
 - 定时调度 查询 新增 修改 删除
 - 任务报表 查询
 - 任务配置 查询 新增 编辑 删除

图 1.5.1.1 新增角色

1.5.1.2. 修改角色

选中角色，点击编辑角色，如下图：

序号	角色名称	描述	操作
1	monitor		<input type="button" value="编辑"/> <input type="button" value="删除"/>

可以对角色的权限

角色名称：

角色描述： 剩余字数：200字

权限分配：

- 数据源配置**
 - 文件 查询 新增 编辑 删除
 - 关系型数据库 查询 新增 编辑 删除
 - 大数据 查询 新增 编辑 删除
- 规则管理**
 - 敏感发现规则 查询 新增 编辑 删除
 - 脱敏算法 查询 新增 编辑 删除
 - 脱敏策略 查询 新增 编辑 删除

图 1.5.1.2 修改角色

1.5.1.3. 删除角色

选中用户，选择删除角色如或在对应的角色后面，点击删除标识，如下图:

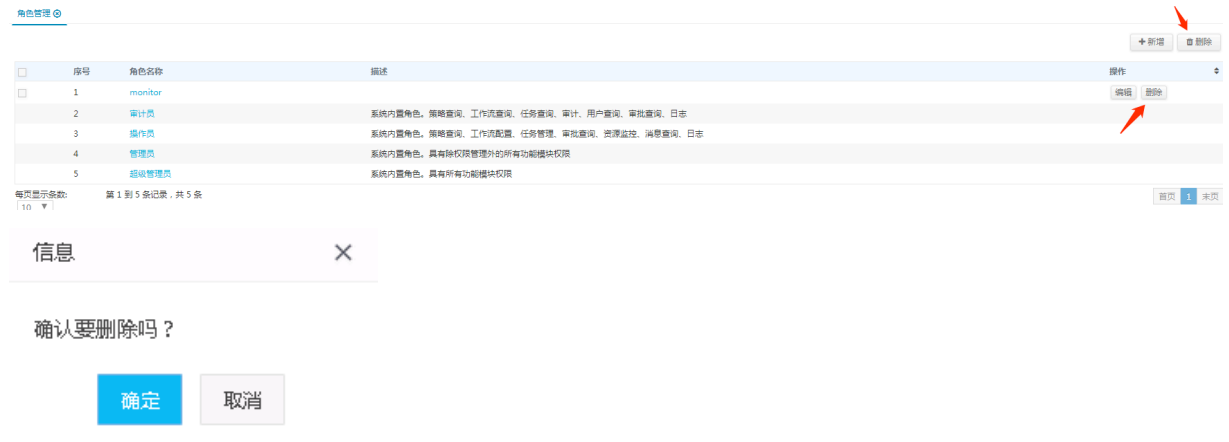


图 1.5.1.3 删角色

1.5.2. 用户组

用户组是对用户操作需要脱敏数据库的范围限定。点击用户组子节点，可以看到用户组相关信息，一个用户组可以包含多个数据库别名，一个数据库别名只能属于一个用户组。

1.5.2.1. 新增用户组

选择新增用户组如下图:

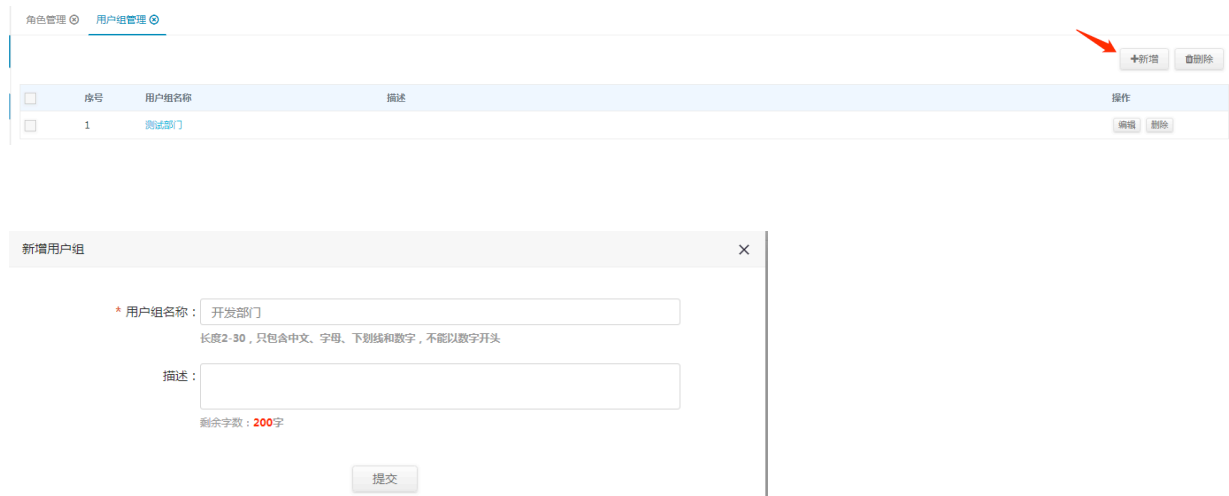


图 1.5.2.1 新增用户组

1.5.2.2. 删除用户组

选中用户组，点击右键选择删除用户组，如下图：

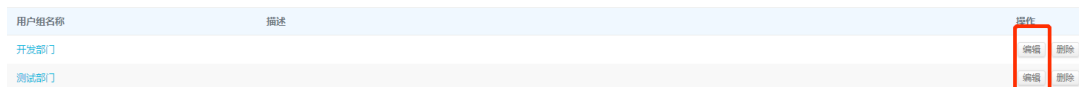


图 1.5.2.2 删除用户组

1.5.2.3. 查看用户组详情

选中用户组，点击用户组名称，如下图：

<input type="checkbox"/>	序号	用户组名称	描述
<input type="checkbox"/>	1	开发部门	
<input type="checkbox"/>	2	测试部门	

图 1.5.2.3 查看用户组详情

1.6 服务器配置

服务器有三种数据类型：数据库，结构化文本，远程结构化文本，根据不同的需求可以选择相应的配置

1.6.1. 数据库服务器

数据库服务器是脱敏数据库为主的一种类型，配置信息包括：

数据库别名：可以根据实际的情况自拟定名称

描述：填写相应的数据库相关信息可选择填写

数据库模型：分为脱敏产品与 UDF 两种

用途：分三类：源，目标，BOTH

所属域：根据之前创建的域选择即可

数据库类型：可选择 oracle,db2,as400,mysql,sqlserver,sybase,greenplum

字符集：根据实际数据库的字符集类型，选择相应的数据库字符集

IP 地址：根据实际情况填写相应的 IP 地址

端口号：根据实际数据库的端口号填写

数据库名：根据实际的数据库名称填写

用户名：根据实际的数据库登录用户填写

密码：根据实际的数据库登录用户填写相应的密码

注释：1、UDF 是调用数据库自定义函数，适用于源库和目标库是同一个库的场景。UDF 后台处理逻辑是直接更新表，类似于：

```
UPDATE tablename SET columnname=function(columnname,param);
```

2、脱敏产品是传统的 ETL 处理流程，包含两种方式，分别是内存方式和落地文件方式，内存方式适用于源库和目标库是不同库的场景。从源库提取数据、经过数据脱敏处理再插入到目标测试库。落地方式即适用于不同库之间，也适用于源库和目标库是同一个库的情况。

3，用途有三种类型：源、目标、BOTH

如果数据库模型为 UDF 时，请选择 BOTH；

如果数据库模型为脱敏产品时，根据实际情况设置用途

1.6.1.1. 新增

根据实际情况填写相应的参数配置，点击测试连接，点击下一步，点击完成即可。

如下图：



新增数据库数据源

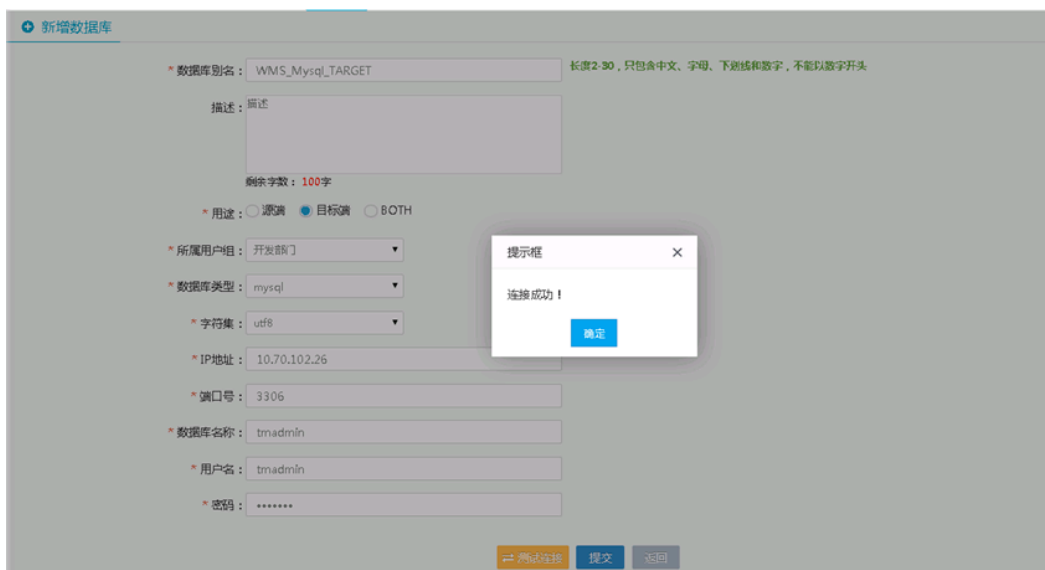


图 1.6.1.1 新增数据库服务器

1.6.1.2. 修改

选中服务器，点击数据源名称后面的编辑操作，如下图：



图 1.6.1.2 修改数据库服务器

1.6.1.3. 删除

选中服务器，点击数据源删除操作，如下图：

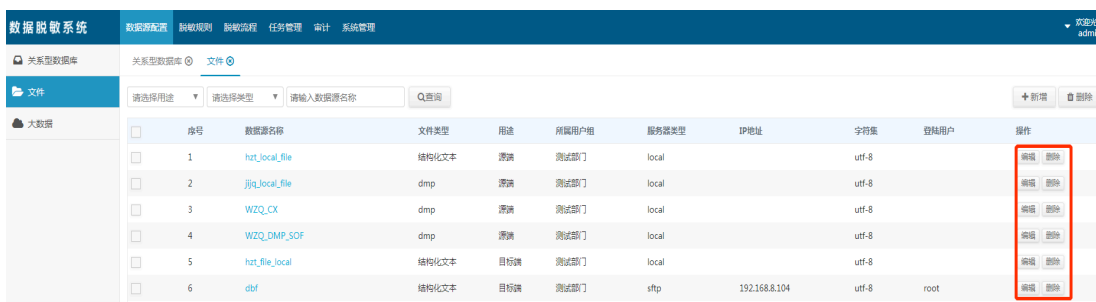


图 1.6.1.3 删除数据库服务器

1.6.2. 结构化文本服务器

结构化文本：是纯文本文件存放在本地服务器目录，以脱敏纯文类型的一种类型，其中配置信息包括：

模板别名：根据实际情况自拟定。

描述：填写数据库文件的描述信息，可选择填写

用途：分三类：源，目标，BOTH

所属域：根据实际情况选择相应的域

字符集：根据实际情况选择相应的文本字符集

文件路径：纯文本存放的绝对路径

字段分隔符：根据文本的实际情况填写

行分隔符：根据实际情况填写

逃逸符：根据实际情况填写

括号符：根据实际情况填写

1.6.2.1. 新增

新增结构化文本服务器，需根据实际情况填写相应的参数，如下图：

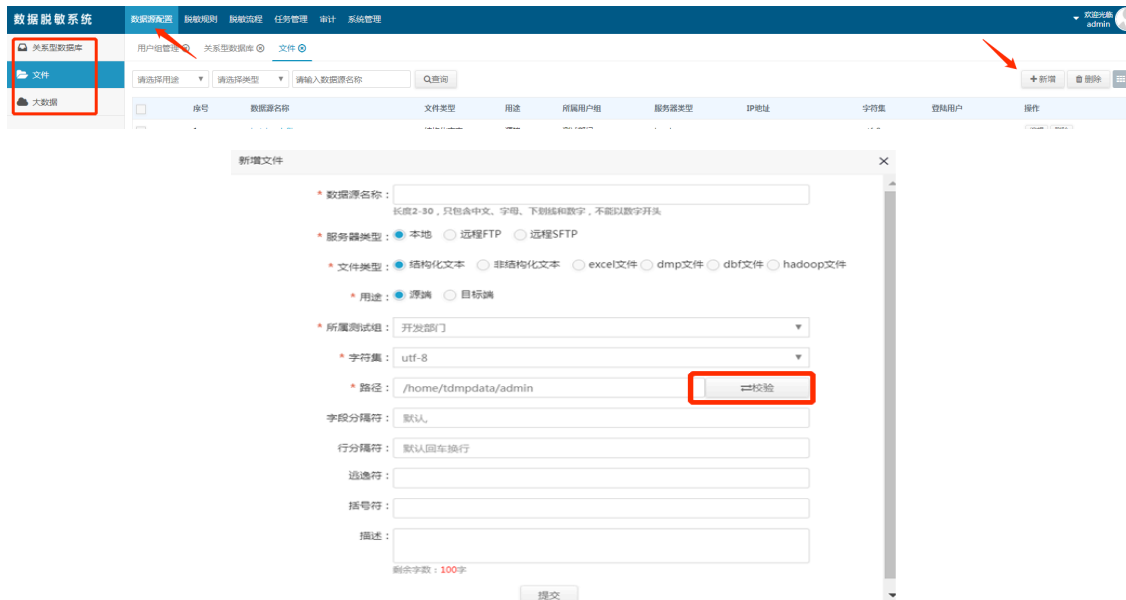


图 1.6.2.1 新增结构化文本服务器

1.6.2.2. 修改

选中服务器，点击右键选择修改，如下图：

序号	数据源名称	文件类型	用途	所属用户组	服务器类型	IP地址	字符集	登陆用户	操作
1	hzt_local_file	结构化文本	源端	测试部门	local		utf-8		编辑 删除
2	jjjq_local_file	dmp	源端	测试部门	local		utf-8		编辑 删除
3	WZQ_CX	dmp	源端	测试部门	local		utf-8		编辑 删除
4	WZQ_DMP_SOF	dmp	源端	测试部门	local		utf-8		编辑 删除
5	hzt_file_local	结构化文本	目标端	测试部门	local		utf-8		编辑 删除
6	dbf	结构化文本	目标端	测试部门	sftp	192.168.8.104	utf-8	root	编辑 删除



图 1.6.2.2 修改结构化文本服务器

1.6.2.3. 删除

选中服务器，点击对应数据源选择删除操作，如下图：

序号	数据源名称	文件类型	用途	所属用户组	服务器类型	IP地址	字符集	登陆用户	操作
1	hzt_local_file	结构化文本	源端	测试部门	local		utf-8		编辑 删除
2	jjjq_local_file	dmp	源端	测试部门	local		utf-8		编辑 删除
3	WZQ_CX	dmp	源端	测试部门	local		utf-8		编辑 删除
4	WZQ_DMP_SOF	dmp	源端	测试部门	local		utf-8		编辑 删除
5	hzt_file_local	结构化文本	目标端	测试部门	local		utf-8		编辑 删除

图 1.6.2.3 删除结构化文本服务器

1.6.3. 远程结构化文本服务器

远程结构化文本：是访问远程服务器存放纯文本目录，以脱敏纯文类型的一种类型，其中配置信息包括：

远程服务器别名：根据实际情况自拟定

描述：根据实际情况描述文件说明

用途：分为源，目标，BOTH

远程服务器类型：分 sftp 与 ftp（其中开放端口分别为 22/21）

所属域：根据实际情况选择填写

字符集：根据实际情况选择填写

IP：根据实际情况选择填写

端口号：根据选择远程服务器类型选择端口号

用户名：为操作系统用户即可

密码：根据提供的操作系统用户密码填写

远程文件路径：填写远程文件存放的绝对路径

字段分隔符：根据文本实际情况填写

行分隔符：根据文本实际情况填写

逃逸符：根据实际情况填写

括号符：根据实际情况填写

1.6.3.1. 新增

点击文件，右上角选择新增，根据实际情况填写相应的参数配置如下图：



* 数据源名称: dbf

* 服务器类型: 本地 远程FTP 远程SFTP

* 文件类型: 结构化文本 非结构化文本 excel文件 dmp文件 dbf文件 hadoop文件

* 用途: 源端 目标端

* 所属测试组: 测试部门

* 字符集: utf-8

* IP地址: 192.168.8.104

* 端口号: 22

* 用户名: root

* 密码: password

* 路径: /root/dbfdir/

* 字段分隔符: .

* 行分隔符: 默认回车换行

逃逸符:

括号符: *

描述:

剩余字数: 100字

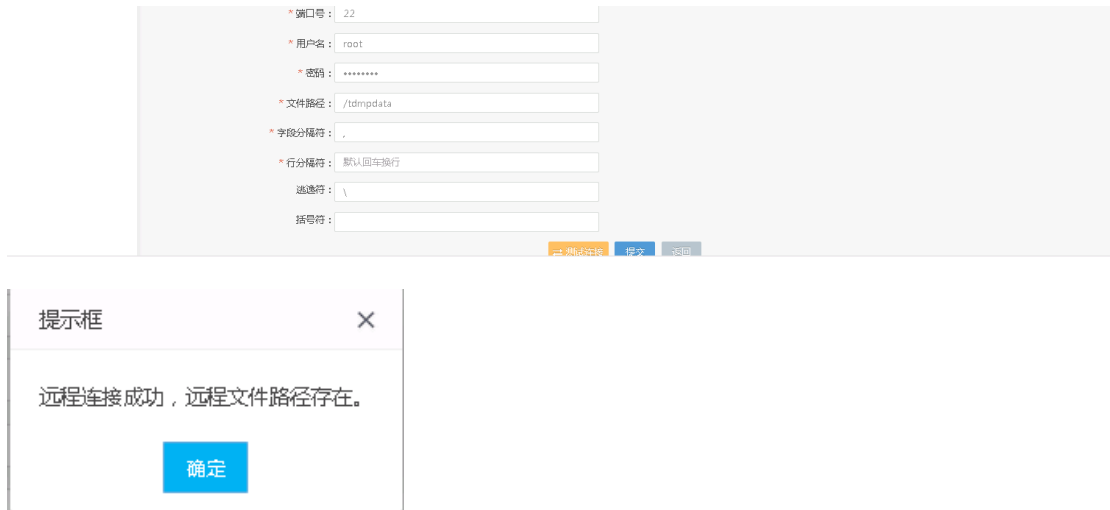


图 1.6.3.1 新增远程结构化文本服务器

1.6.3.2. 修改

选中服务器，选择编辑，如下图：



图 1.6.3.2 修改远程结构化文本服务器

1.6.3.3. 删除

选中服务器，选择删除，如下图：



图 1.6.3.3 删除远程结构化文本服务器

2.规则管理

2.1 敏感数据发现规则

数据发现规则内置 JAVA 正则表达式和 JAVA 校验算法，目前内置有如下正则表达式：信用卡号、地址、名称、日期、电话号码、纳税人识别号、组织机构代码、网址、英文、身份证号、证件号码、邮箱、金额、军官证、工商注册号、护照号、港澳通行证、股票、统一社会信用代码、网址、车架号、车牌号等。

新增数据发现规则： 点击敏感规则选择新增填写相应的参数，如下图：

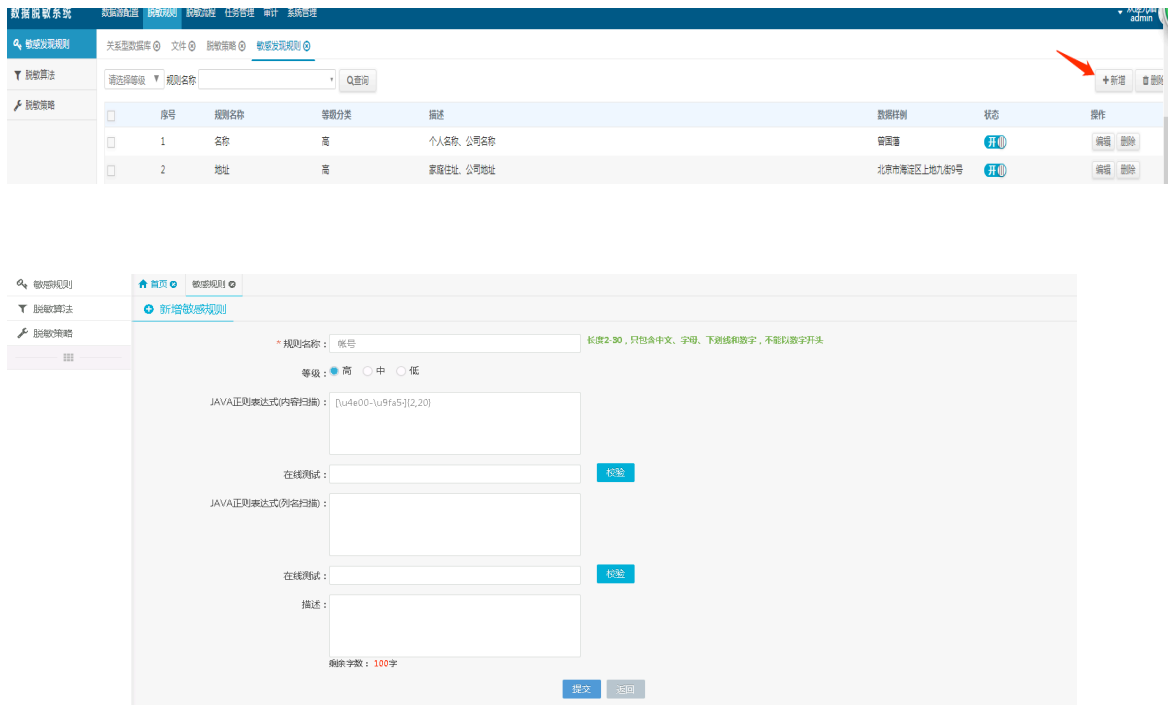


图 2.1.1 新增敏感数据发现规则

修改数据发现规则： 选中数敏感发现规则名称，点击右侧修改，修改相应的参数如下图：



图 2.1.2 修改敏感数据发现规则

删除敏感数据发现规则： 选中敏感是数据发现规则，右侧选择删除如下图：



图 2.1.3 删除敏感数据发现规则

2.2 数据脱敏规则

系统内置有证件号码、名称、地址、电话号码、邮箱、身份证号、信用卡号、组织机构代码、日期、金额、通用账号、截取&替换、英文、保留&脱敏、军官证、工商注册号、护照号、港澳通行证、股票、统一社会信用代码、网址、车架号、车牌号等数据脱敏规则，满足于企业日常数据脱敏的各种需求。



图 2.2.1 数据脱敏规则展示

数据脱敏规则样本测试:

选择脱敏方法：根据不同的敏感字段配置不同的脱敏规则，一般分为随机和 hash 算法。本公司为了保持数据一致性，选择 hash 算法。

样本值：输入原始样本值比，如“张三”

参数配置：根据实际情况填写相应的参数配置

如下图：

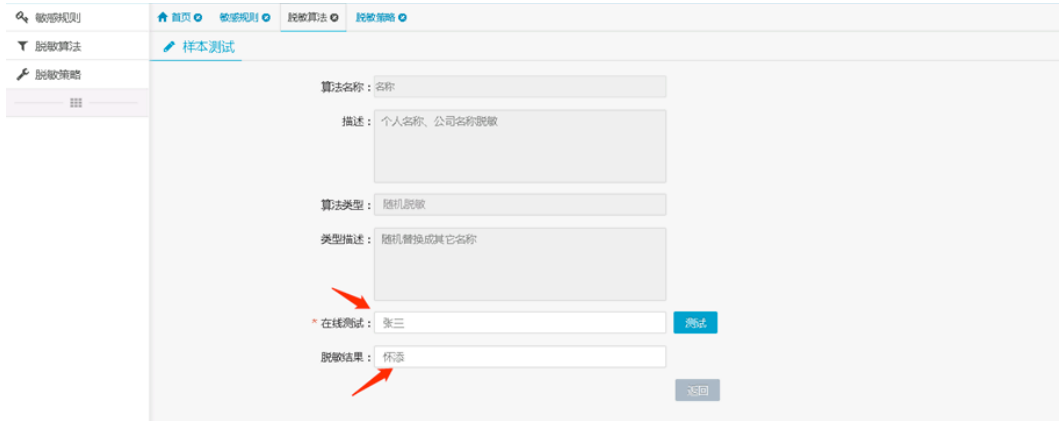


图 2.2.2 数据脱敏规则样本测试

注意：数据脱敏规则是只读板块，不可以新增、修改和删除

2.3 脱敏策略

脱敏策略板块：主要用于配置数据发现规则和数据脱敏规则。根据实际情况选择相应的数据发现规则，数据脱敏规则以及参数配置。

新增脱敏策略：点击脱敏策略，选择新增，如下图：

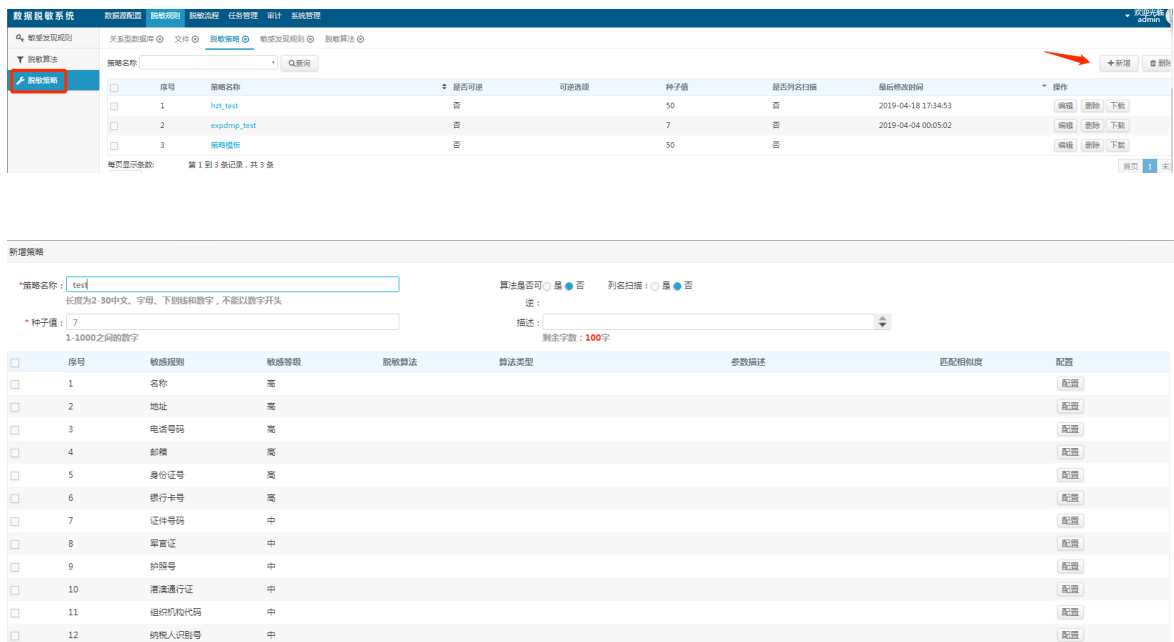


图 2.3.1 脱敏策略新增页面

配置策略：根据实际情况选择相应的数据发现规则之前打勾，点击配置根据实际情况填写相应的参数如下图：

<input type="checkbox"/>	序号	敏感规则	敏感等级	脱敏算法	算法类型	参数描述	匹配相似度	配置
<input checked="" type="checkbox"/>	1	名称	高					
<input type="checkbox"/>	2	地址	高					
<input type="checkbox"/>	4	邮箱	高					
<input type="checkbox"/>	12	证件号码	中					

* 策略名称: CBS_POLICY 长度2-30,只包含中文、字母、下划线和数字,不能以数字开头

算法是否可逆: 是 否

列名扫描: 是 否

* 种子值: 2 (1-1000之间的数字)

描述:

剩余字数: 100字

<input type="checkbox"/>	序号	敏感规则	敏感等级	脱敏算法	算法类型	参数	匹配相似度	配置
<input checked="" type="checkbox"/>	1	名称	高	名称	唯一仿真脱敏	原名称保留前0位	30	
<input checked="" type="checkbox"/>	2	地址	高	地址	唯一仿真脱敏	原地址保留前0位	30	
<input checked="" type="checkbox"/>	3	电话号码	高	电话号码	唯一仿真脱敏	原电话号码保留前0位	30	
<input checked="" type="checkbox"/>	4	邮箱	高	邮箱	唯一仿真脱敏	不保留域名	30	
<input checked="" type="checkbox"/>	5	证件号码	中	证件号码	唯一仿真脱敏	组织机构号保留前0位身份证不保留生日	30	

修改脱敏策略

敏感规则:

脱敏算法:

算法类型:

匹配相似度:

保留前几位: 范围0-6

图 2.3.2 脱敏策略配置页面

脱敏策略展示: 配置完成之点击保存, 如下图:

序号	策略名称	是否可逆	可逆选项	种子值	是否列名扫描	最后修改时间	操作
1	hzt_test	否		50	否	2019-04-18 17:34:53	<input type="button" value="编辑"/> <input type="button" value="删除"/> <input type="button" value="下载"/>
2	expdmp_test	否		7	否	2019-04-04 00:05:02	<input type="button" value="编辑"/> <input type="button" value="删除"/> <input type="button" value="下载"/>
3	策略模板	否		50	否		<input type="button" value="编辑"/> <input type="button" value="删除"/> <input type="button" value="下载"/>

每页显示条数: 1-20 第 1 到 3 条记录, 共 3 条 首页 1

敏感规则	敏感等级	脱敏算法	算法类型	算法描述	参数	匹配相似度
名称	高	名称	唯一仿真脱敏	脱敏成其它名称, 保证每次脱敏的结果不变	原名称保留前0位	30
地址	高	地址	唯一仿真脱敏	脱敏成其它地址, 保证每次脱敏的结果不变	原地址保留前0位	30
电话号码	高	电话号码	唯一仿真脱敏	脱敏成相同长度电话号码, 保证每次脱敏的结果不变	原电话号码保留前0位	30
邮箱	高	邮箱	唯一仿真脱敏	脱敏成其它邮箱, 保证每次脱敏的结果不变	不保留域名	30
证件号码	中	证件号码	唯一仿真脱敏	脱敏成相同格式证件号码, 保证每次脱敏的结果不变	组织机构号保留前0位身份证不保留生日	30

图 2.3.3 脱敏策略展示

修改脱敏策略: 选中脱敏策略, 点击选择修改如下图:

序号	策略名称	是否可逆	可逆选项	种子值	是否列名扫描	最后修改时间	操作
1	hzt_test	否		50	否	2019-04-18 17:34:53	编辑 删除 下载
2	expdmp_test	否		7	否	2019-04-04 00:05:02	编辑 删除 下载
3	策略模板	否		50	否		编辑 删除 下载

图 2.3.4 修改脱敏策略

删除脱敏策略：选中脱敏策略，选择删除如下图：

序号	策略名称	是否可逆	可逆选项	种子值	是否列名扫描	最后修改时间	操作
1	hzt_test	否		50	否	2019-04-18 17:34:53	编辑 删除 下载
2	expdmp_test	否		7	否	2019-04-04 00:05:02	编辑 删除 下载
3	策略模板	否		50	否		编辑 删除 下载

图 2.3.5 删除脱敏策略

3. workflows配置

3.1 敏感元数据管理

3.1.1. 新增数据元数据管理

3.1.1.1. 脱敏产品模式数据元数据

数据元数据配置参数包括：

子集名称：可自拟定名称

描述：可以自拟定填写

数据源：选择之前创建的数据库服务器别名

数据类型：自动根据数据源填写

数据库类型：自动根据数据源填写

敏感信息文本导入：分否与是

注释：敏感信息文本导入“否”是指：直接从数据库中选择相应的表自动敏感发现敏感字段

敏感信息文本导入“是”是指：提前根据提供的敏感信息文本模板填写好，相应的模式名（或用户名），


表名，字段名，字段备注，敏感发现规则，SQL 条件，抽取限制条数，*等相关信息

3.1.1.1.1. 常规配置模式

1, 此步骤敏感信息文本导入选择“否”的操作步骤
选择需要做数据脱敏处理的数据源如下图:



序号	敏感元数据名称	数据源	数据源描述	脱敏策略	日期	状态	操作
1	WQZQ_CHECK_190425172941_MD	WQZQ_CHECK	local_dmp	策略脱敏	部分	已审核	修改范围 修改结果 删除
2	WQZQ_CK_190425172942_MD	WQZQ_CK	local_dmp	策略脱敏	部分	已审核	修改范围 修改结果 删除
3	cg_db2_src_77_190425151532_MD	cg_db2_src_77	db2_192.168.8.77_db2test	策略脱敏	部分	未审核	修改范围 修改结果 删除
4	ht_local_file_190425145316_MD	ht_local_file	local_structured	策略脱敏	部分	已审核	修改范围 修改结果 删除
5	ht_local_file_190425145246_MD	ht_local_file	local_structured	策略脱敏	部分	已审核	修改范围 修改结果 删除



数据脱敏系统 数据源配置 脱敏规则 脱敏流程 任务管理 审计 系统管理

敏感元配置

1 新增敏感元数据 2 确定数据范围

数据源类型: 关系型数据库 文件 大数据

* 数据源: dw_aboss2_aboss

* 敏感元数据名称: dw_aboss2_aboss_190424225050_MD

数据源描述: local_dmp_root/datamask_dmp_src/aboss2/aboss.dmp

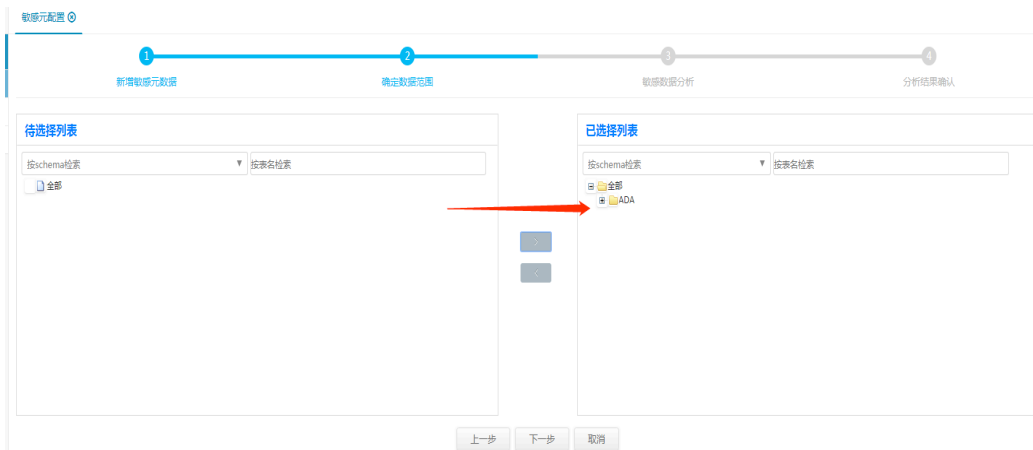
是否含表头: 否

描述:
剩余字数: 100字

下一步 取消

图 3.1.1.1.1 新增敏感元数据

选择需要脱敏的表放到右侧如下图:



敏感元配置

1 新增敏感元数据 2 确定数据范围 3 敏感数据分析 4 分析结果确认

待选择列表

按schema检索 按表名检索

全部

已选择列表

按schema检索 按表名检索

全部 ADA

上一步 下一步 取消

图 3.1.1.1.1 新增敏感元数据

选择脱敏策略，点击数据扫描如下图：

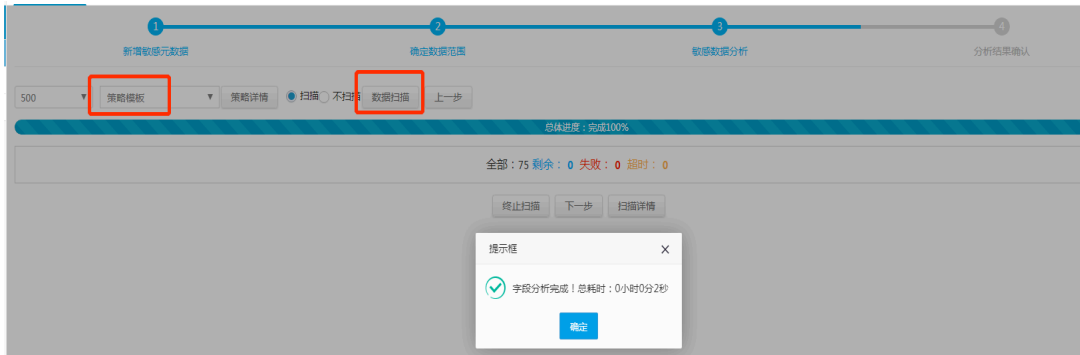


图 3.1.1.1.1 新增敏感元数据

手动确定自动敏感发现字段是否正确，可自主选择敏感发现规则，确认无误，点击分析结果确认，保存如下图：

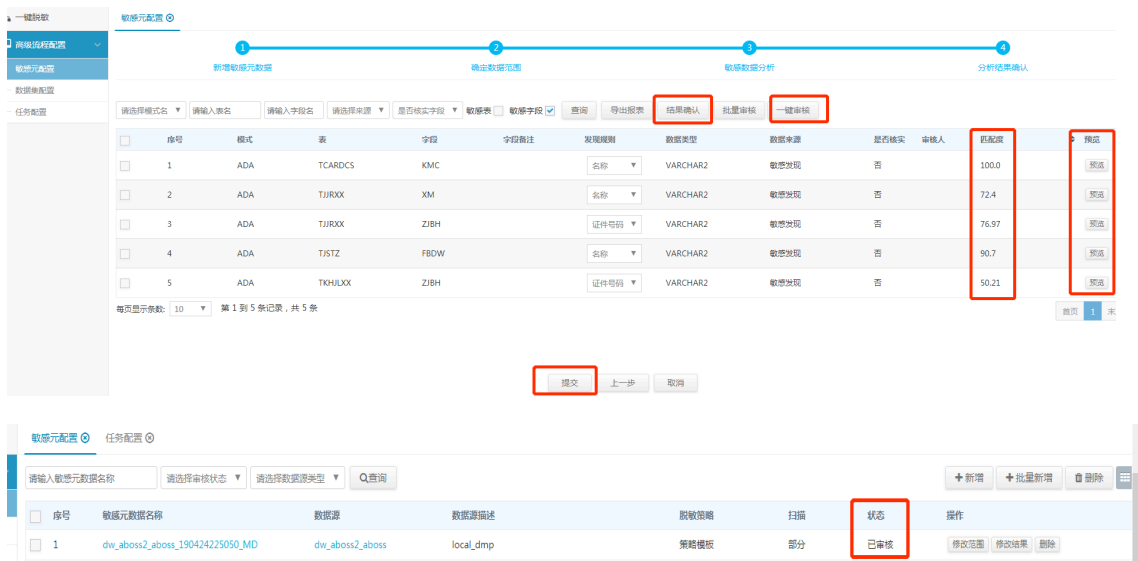


图 3.1.1.1.1 新增敏感元数据

3.1.1.1.2. 模板导入模式

以下是敏感信息文本导入选择“是”的操作步骤
创建数据元，敏感信息文本导入选择“是”如下图：



图 3.1.1.1.2 新增敏感元数据

文件模板为：是以.CSV 文件为后缀如下图：

	A	B	C	D	E	F
1	模式名(或表名)		字段名	字段备注	敏感发现规则	*
2	CQTEST	CQ_INDEX_TEST01	CUSTNO			*
3	CQTEST	CQ_INDEX_TEST01	IDNO		证件号码	*
4	CQTEST	CQ_INDEX_TEST01	ENT_SHORT_NAME			*
5	CQTEST	CQ_INDEX_TEST01	TELNO		电话号码	*
6	CQTEST	CQ_INDEX_TEST01	MOBILENO		电话号码	*
7	CQTEST	CQ_INDEX_TEST01	ADDRESS		地址	*
8	CQTEST	CQ_INDEX_TEST01	ENT_CUSTNAME		名称	*
9	CQTEST	CQ_INDEX_TEST01	PERSON_CUSTNAME		名称	*

图 3.1.1.1.2 新增敏感元数据

根据提示上传文件，点击浏览选择文件，点击上传文件，点击完成如下图：





图 3.1.1.1.2 新增敏感元数据

选择相应的脱敏策略，点击下一步，如下图：



图 3.1.1.1.2 新增敏感元数据

根据上传的.csv 敏感字段文件确认敏感字段如下图：



图 3.1.1.1.2 新增敏感元数据

点击提交如下图：



图 3.1.1.1.2 新增敏感元数据

3.1.2. 修改敏感元范围

针对以上创建的敏感数据元类型，选择需要修改的敏感数据元，点击右侧修改，修改相应的参数，点击提交如下图：

序号	敏感元数据名称	数据源	数据源描述	脱敏策略	扫描	状态	操作
1	WZQ_CHECK_190425172941_MD	WZQ_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
2	WZQ_CX_190425172429_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
3	cq_db2_src_77_190425151522_MD	cq_db2_src_77	db2_192.168.8.77_db2test	策略模板	部分	未审核	修改范围 修改结果 删除
4	hzt_local_file_190425145356_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
5	hzt_local_file_190425145249_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
6	hzt_local_file_190425144831_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
7	hzt_local_file_190425144154_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
8	WZQ_DMP_TARGET_CHECK_190425114204_MD	WZQ_DMP_TARGET_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
9	WZQ_CX_190425102649_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
10	hzt_local_file_190425100128_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除

图 3.1.2 修改敏感数据元

3.1.3. 修改分析结果

选中需要修改的数据子集，点击右键选择修改分析结果，修改敏感字段敏感发现规则，修改完成点击分析结果确认，如下图：

序号	敏感元数据名称	数据源	数据源描述	脱敏策略	扫描	状态	操作
1	WZQ_CHECK_190425172941_MD	WZQ_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
2	WZQ_CX_190425172429_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
3	cq_db2_src_77_190425151522_MD	cq_db2_src_77	db2_192.168.8.77_db2test	策略模板	部分	未审核	修改范围 修改结果 删除
4	hzt_local_file_190425145356_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
5	hzt_local_file_190425145249_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
6	hzt_local_file_190425144831_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
7	hzt_local_file_190425144154_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
8	WZQ_DMP_TARGET_CHECK_190425114204_MD	WZQ_DMP_TARGET_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
9	WZQ_CX_190425102649_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
10	hzt_local_file_190425100128_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除

图 3.1.4 修改分析结果

3.1.4. 删除敏感数据元

选中需要删除的敏感数据元，点击右侧选择删除敏感数据元如下图：

序号	敏感元数据名称	数据源	数据源描述	脱敏策略	扫描	状态	操作
1	WZQ_CHECK_190425172941_MD	WZQ_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
2	WZQ_CX_190425172429_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
3	cq_db2_src_77_190425151522_MD	cq_db2_src_77	db2_192.168.8.77_db2test	策略模板	部分	未审核	修改范围 修改结果 删除
4	hzt_local_file_190425145356_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
5	hzt_local_file_190425145249_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
6	hzt_local_file_190425144831_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
7	hzt_local_file_190425144154_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除
8	WZQ_DMP_TARGET_CHECK_190425114204_MD	WZQ_DMP_TARGET_CHECK	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
9	WZQ_CX_190425102649_MD	WZQ_CX	local_dmp	策略模板	部分	已审核	修改范围 修改结果 删除
10	hzt_local_file_190425100128_MD	hzt_local_file	local_structured	策略模板	部分	已审核	修改范围 修改结果 删除

图 3.1.5 删除敏感数据元

3.2 数据集配置

数据集板块是对一个需要做数据脱敏处理的数据库的子集范围的确定。可以选择需要脱敏的表范围，以及每个表的 Where 条件。

点击 workflows 配置--数据集配置，右上角可以选择新增和删除数据集配置

3.2.1 新增数据集

点击右上角新增

序号	数据集名称	敏感元数据	数据源描述	策略策略	创建时间	操作
1	cwd_mysql_47_190408092019_MD_DS	cwd_mysql_47_190408092019_MD	mysql_192.168.8.47_idmp	策略模板	2019-04-08 09:20:45	编辑 删除
2	WZQ_MYSQL_104_190403130414_MD_DS	WZQ_MYSQL_104_190403130414_MD	mysql_192.168.8.104_test	策略模板	2019-04-11 17:49:45	编辑 删除
3	WZQ_ORCL_40_190411145028_MD_DS	WZQ_ORCL_40_190411145028_MD	oracle_192.168.8.40_orcl	策略模板	2019-04-11 18:15:01	编辑 删除
4	cwd_mysql_47_190408091941_MD_DS	cwd_mysql_47_190408091941_MD	mysql_192.168.8.47_idmp	策略模板	2019-04-12 17:57:24	编辑 删除
5	WZQ_ORCL_40_190415151307_MD_DS	WZQ_ORCL_40_190415151307_MD	oracle_192.168.8.40_orcl	策略模板	2019-04-15 15:13:41	编辑 删除

图 3.2.1 新增数据集

选择敏感元数据，点击下一步

图 3.2.1 新增数据集

选择敏感数据范围--点击移向右边--然后点击下一步--where 条件配置--提交

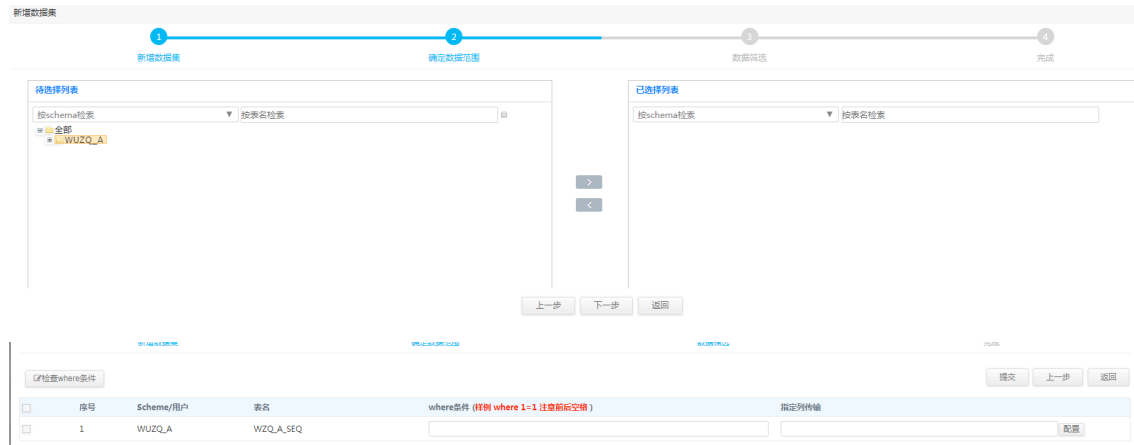


图 3. 2. 1 新增数据集

3.2.2 删除数据集

选择要删除的数据集--点击右上角删除或者在右侧操作中直接删除

序号	数据集名称	敏感元数据	数据源描述	脱敏策略	创建时间	操作
1	cwd_mysql_47_190408092019_MD_DS	cwd_mysql_47_190408092019_MD	mysql_192.168.8.47_tdmp	策略模板	2019-04-08 09:20:45	编辑 删除
2	WZQ_MYSQL_104_190403130414_MD_DS	WZQ_MYSQL_104_190403130414_MD	mysql_192.168.8.104_test	策略模板	2019-04-11 17:49:45	编辑 删除
3	WZQ_ORCL_40_190411145028_MD_DS	WZQ_ORCL_40_190411145028_MD	oracle_192.168.8.40_orcl	策略模板	2019-04-11 18:15:01	编辑 删除
4	cwd_mysql_47_190408091941_MD_DS	cwd_mysql_47_190408091941_MD	mysql_192.168.8.47_tdmp	策略模板	2019-04-12 17:57:24	编辑 删除
5	WZQ_ORCL_40_190415151307_MD_DS	WZQ_ORCL_40_190415151307_MD	oracle_192.168.8.40_orcl	策略模板	2019-04-15 15:13:41	编辑 删除

图 3. 2. 2 删除数据集

点击确定删除数据集

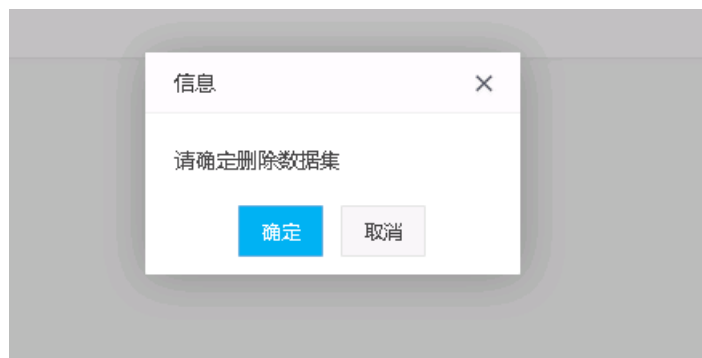


图 3. 2. 2 删除数据集

3.3 任务配置

任务配置模块分为：

源数据库--->目标数据库（ETL）

就地库（udf）

源数据库--->目标格式化文本

源格式化文本--->目标格式化文本

源格式化文本--->目标数据库

SFTP/FTP 源文件格式化文本--->SFTP/FTP 目标格式化文本

3.3.1 新增任务配置

任务配置板块，是对所创建的数据子集和脱敏策略进行任务的定制。

选择数据源，其中数据源就是数据集中创建或修改的数据子集。

3.3.1.1 脱敏产品内存方式

新增脱敏产品内存方式 workflow，右侧选择新增，选择源数据-->目标数据库，下拉菜单选择数据集，如下图：



图 3.3.1.1 新增脱敏产品内存 workflow

填写相应的参数：

任务名称：自拟定名称

描述：自拟定描述

目标数据库：下拉菜单选择目标数据库

并发数：根据实际情况填写，建议 3-5 个

分段提交量：根据实际情况选择 1-5W

SCHEMA 映射：填写相应的 aimSchema

根据实际情况填写参数，点击下一步，如下图：

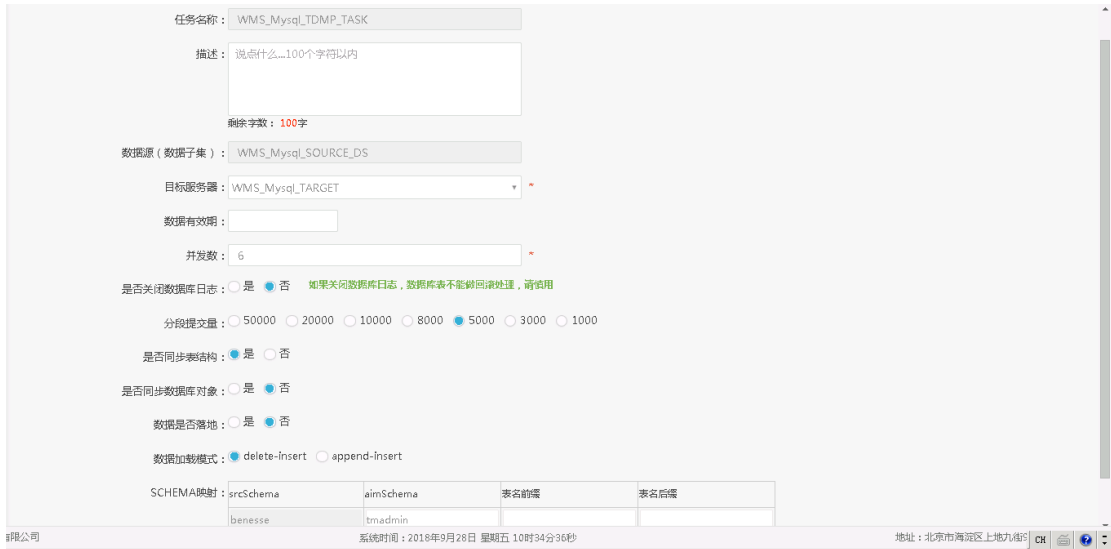


图 3.3.1.1 新增脱敏产品内存 workflow

内存方式参数设置如下：

是否落地：否；

是否同步表结构：根据实际情况选择

是否同步数据库对象：根据需求方提供的信息来选择

根据实际需求填写，点击下一步，如下图：



图 3.3.1.1 新增脱敏产品内存 workflow

每个任务右侧都可以进行修改和删除，开关状态需要打开才可以进行执行和调度

序号	脱敏任务	源端		目标端		脱敏方式	脱敏元数据 (数据集)	创建时间	开关状态	操作
		类型	名称	类型	名称					
1	CBS_ORACLE_EDITEST_TDMP_TASK	db	CBS_ORACLE_SOURCE	db	CBS_ORACLE_EDITEST_TGT	ETL	CBS_ORACLE_SOURCE_DS	2018-09-27 13:22:11	开	
2	WMS_Mysql_TDMP_TASK	db	WMS_Mysql_SOURCE	db	WMS_Mysql_TARGET	ETL	WMS_Mysql_SOURCE_DS	2018-09-27 13:20:10	关	
3	BLS_EDITEST_LOAD_TASK	db	CBS_ORACLE_SOURCE	db	CBS_ORACLE_EDITEST_TGT	ETL	CBS_ORACLE_SOURCE_DS	2018-09-26 16:21:30	开	
4	CBS_ORACLE_PDBCBSR2_TASK	db	CBS_ORACLE_SOURCE	db	CBS_ORACLE_PDBCBSR2_TGT	ETL	CBS_ORACLE_SOURCE_DS	2018-09-26 15:36:53	开	
5	WMS_Mysql_TASK	db	WMS_Mysql_SOURCE	db	WMS_Mysql_TARGET	ETL	WMS_Mysql_SOURCE_DS	2018-09-26 13:11:51	开	
6	CBS_ORACLE_EDITEST_TASK	db	CBS_ORACLE_SOURCE	db	CBS_ORACLE_EDITEST_TGT	ETL	CBS_ORACLE_SOURCE_DS	2018-09-26 11:01:37	开	

每页显示条数: 10 第 1 到 6 条记录, 共 6 条

图 3.3.1.1 新增脱敏产品内存 workflows

3.3.1.2 脱敏产品文件落地方式

新增脱敏产品文件落地方式 workflows，右侧选择新增，选择源数据--》目标数据库，下拉菜单选择数据集，如下图：

新增任务配置

ETL (源-目标)
 UDF (本地库更新)
 文件导入

源

数据源类型: 数据库 文本 大数据

数据集类型: 脱敏元数据 数据集

数据库类型: mysql

数据集: WMS_Mysql_SOURCE_DS

目标

数据源类型: 数据库 文本 大数据

数据库类型: mysql

目标服务器: WMS_Mysql_TARGET

图 3.3.1.2 新增脱敏产品文件落地 workflows

填写相应的参数：

任务名称：自拟定名称

描述：自拟定描述

目标数据库：下拉菜单选择目标数据库

并发数：根据实际情况填写，建议 3-5 个

分段提交量：根据实际情况选择 1-5W

SCHEMA 映射：填写相应的 aimSchema

根据实际情况填写参数，点击下一步，如下图：



图 3.3.1.2 新增脱敏产品文件落地 workflow

落地文件方式可以参照内存方式来配置，不同的是，第二步需做如下配置：

是否落地：“是”

是否同步表结构：根据实际情况选择

是否同步数据库对象：根据需求方提供的信息来选择

数据加载模式：insert 效率比 load 低，load 方式需在服务器装客户端；

落地文件路径：选择空间合适的路径，需校验路径合法性（/脱敏产品 data）；

脱敏成功后是否清空落地文件：一般选“是”，以腾出更多的磁盘空间。

根据实际情况选择相应参数配置，点击下一步如下图：



图 3.3.1.2 新增脱敏产品文件落地 workflow

3.3.2 打开任务配置

点击开关状态选择打开，任务状态为“打开”如下图：

序号	脱敏任务	源端		目标端		脱敏方式	敏感元数据 (数据库)	开关状态	操作
		类型	名称	类型	名称				
1	jjq_sybase_2_oracle_task13	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
2	jjq_sybase_2_oracle_task12	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
3	jjq_sybase_2_oracle_task11	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
4	WZQ_CHECK_TKXGLG	dmp	WZQ_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_CHECK_190425172941_MD	开	编辑 删除 同步 下载
5	db2_db2_synDbObject_test	db	hzt_db2_a	db	hzt_db2_test	ETL	hzt_db2_a_190423111509_MD	开	编辑 删除 同步 下载
6	WZQ_CX_MORT_TTT	dmp	WZQ_CX	dmp	WZQ_DMP_TARGET	ETL	WZQ_CX_190425172429_MD	开	编辑 删除 同步 下载
7	WZQ_CHECK_TASK	dmp	WZQ_DMP_TARGET_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_DMP_TARGET_CHECK_190425114204_MD	开	编辑 删除 同步 下载
8	WZQ_CX_FAST_TASK	dmp	WZQ_CX	dmp	WZQ_DMP_TARGET	ETL	WZQ_CX_190425102649_MD	开	编辑 删除 同步 下载

图 3.3.2 打开工作流

3.3.3 关闭工作流

点击开关状态选择关闭，任务状态为“关闭”如下图：

7	WZQ_CHECK_TASK	dmp	WZQ_DMP_TARGET_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_DMP_TARGET_CHECK_190425114204_MD	开	编辑 删除 同步 下载
8	WZQ_CX_FAST_TASK	dmp	WZQ_CX	dmp	WZQ_DMP_TARGET	ETL	WZQ_CX_190425102649_MD	开	编辑 删除 同步 下载
9	file_db2_hzt	file	hzt_local_file	db	hzt_db2_test	ETL	hzt_local_file_190425100128_MD	开	编辑 删除 同步 下载
10	WZQ_DMP_MORE_FAST	dmp	WZQ_DMP_SOF	dmp	WZQ_DMP_TARGET	ETL	WZQ_DMP_SOF_190425101601_MD	关	编辑 删除 同步 下载

图 3.3.3 关闭工作流

3.3.4 修改任务配置

右侧操作中点击修改，可修改相应的参数配置。注：修改工作流之前一定要先关闭工作流，如下图：

序号	脱敏任务	源端		目标端		脱敏方式	敏感元数据 (数据库)	开关状态	操作
		类型	名称	类型	名称				
1	jjq_sybase_2_oracle_task13	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
2	jjq_sybase_2_oracle_task12	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
3	jjq_sybase_2_oracle_task11	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
4	WZQ_CHECK_TKXGLG	dmp	WZQ_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_CHECK_190425172941_MD	开	编辑 删除 同步 下载
5	db2_db2_synDbObject_test	db	hzt_db2_a	db	hzt_db2_test	ETL	hzt_db2_a_190423111509_MD	开	编辑 删除 同步 下载
6	WZQ_CX_MORT_TTT	dmp	WZQ_CX	dmp	WZQ_DMP_TARGET	ETL	WZQ_CX_190425172429_MD	开	编辑 删除 同步 下载
7	WZQ_CHECK_TASK	dmp	WZQ_DMP_TARGET_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_DMP_TARGET_CHECK_190425114204_MD	开	编辑 删除 同步 下载

图 3.3.4 修改工作流

3.3.5 删除工作流

右侧操作中点击删除，注：删除工作流之前必须先关闭工作流。如下图：

1	jjq_sybase_2_oracle_task13	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
2	jjq_sybase_2_oracle_task12	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
3	jjq_sybase_2_oracle_task11	db	jjq_sybase_70_both	db	jjq_oracle_47_tgt	ETL	jjq_sybase_70_both_190424161444_MD	开	编辑 删除 同步 下载
4	WZQ_CHECK_TKXGLG	dmp	WZQ_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_CHECK_190425172941_MD	开	编辑 删除 同步 下载
5	db2_db2_synDbObject_test	db	hzt_db2_a	db	hzt_db2_test	ETL	hzt_db2_a_190423111509_MD	开	编辑 删除 同步 下载
6	WZQ_CX_MORT_TTT	dmp	WZQ_CX	dmp	WZQ_DMP_TARGET	ETL	WZQ_CX_190425172429_MD	开	编辑 删除 同步 下载
7	WZQ_CHECK_TASK	dmp	WZQ_DMP_TARGET_CHECK	dmp	WZQ_DMP_TARGET	ETL	WZQ_DMP_TARGET_CHECK_190425114204_MD	开	编辑 删除 同步 下载

图 3.3.5 删除工作流

4. 任务管理

任务管理板块不仅用于进行 workflow 任务的调度执行、中止、单步执行，更重要的是通过颜色图标的区分清晰的展示了整个任务的执行情况。任务中每个表的每个阶段完成状态都用未开始、运行中、错误中止、正常结束、告警标识。未开始代表程序未运行，运行中代表程序正在运行处理中，错误中止代表程序运行失败或人为中止，正常结束代表程序正常运行成功，告警代表程序虽然运行完成但有警告信息。

4.1 UDF 监控

执行：点击执行按钮开始执行任务（任务中途终止，再次点击执行可以断点续传）

中止：点击中止按钮停止任务工作

初始化：需重新执行任务之前，点击初始化，之前日志将被清除

通过数据更新图标颜色可以查看运行状态，可查看预估记录数与实际记录数，耗时等信息，如下图：

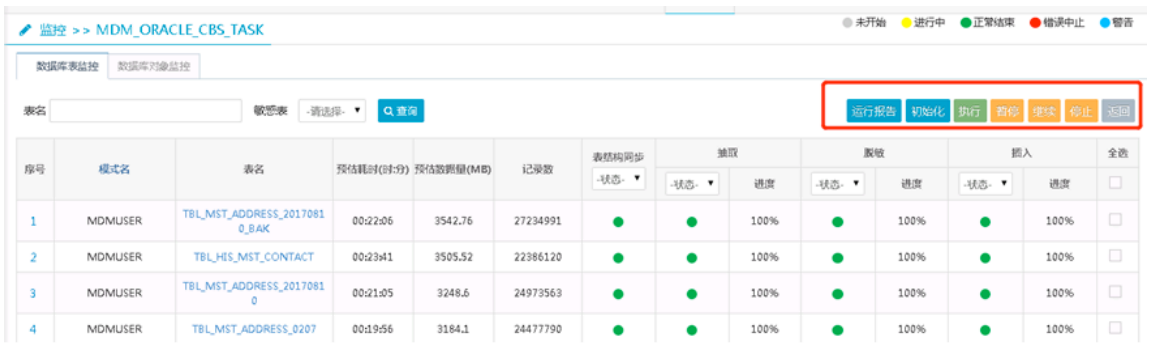


图 4.1.1 UDF 监控页面

单表执行：选择单个表，点击执行即选择单表执行如下图：

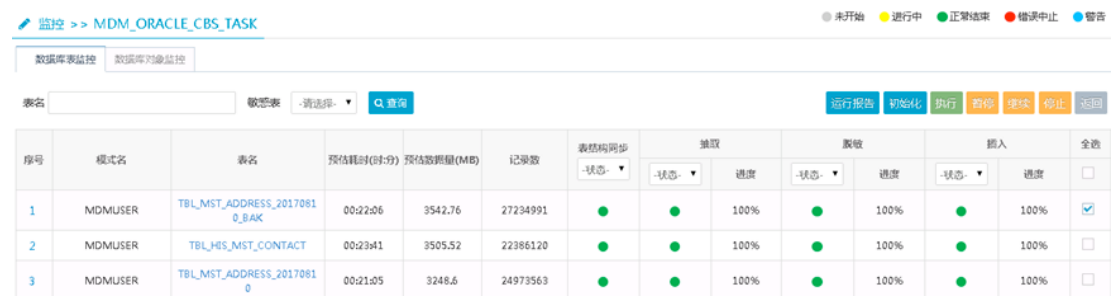


图 4.1.2UDF 单表执行

查看日志：点击序列号查看日志如下图：

序号	模式名	表名	预估耗时(时:分)	预估数据量(MB)	记录数	表结构同步	抽取	脱敏	插入	全选			
1	MDMUSER	TBL_MST_ADDRESS_2017081_0_BAK	00:22:06	3542.76	27234991	●	●	100%	●	100%	●	100%	<input checked="" type="checkbox"/>
2	MDMUSER	TBL_HIS_MST_CONTACT	00:23:41	3505.52	22386120	●	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	MDMUSER	TBL_MST_ADDRESS_2017081_0	00:21:05	3248.6	24973563	●	●	100%	●	100%	●	100%	<input type="checkbox"/>
4	MDMUSER	TBL_MST_ADDRESS_0207	00:19:56	3184.1	24477790	●	●	100%	●	100%	●	100%	<input type="checkbox"/>

图

4.1.3 查看日志

数据对比：点击表名查看数据前 10 行原始数据与脱敏数据的对比如下图：

序号	模式名	表名	预估耗时(时:分)	预估数据量(MB)	记录数	表结构同步	抽取	脱敏	插入	全选	
1	MDMUSER	TBL_MST_ADDRESS_2017081_0_BAK	00:22:06	3542.76	27234991	●	●	100%	●	100%	<input checked="" type="checkbox"/>
2	MDMUSER	TBL_HIS_MST_CONTACT	00:23:41	3505.52	22386120	●	●	100%	●	100%	<input type="checkbox"/>
3	MDMUSER	TBL_MST_ADDRESS_2017081_0	00:21:05	3248.6	24973563	●	●	100%	●	100%	<input type="checkbox"/>
4	MDMUSER	TBL_MST_ADDRESS_0207	00:19:56	3184.1	24477790	●	●	100%	●	100%	<input type="checkbox"/>
5	MDMUSER	TBL_MST_ADDRESS	00:24:18	3167.26	28264770	●	●	100%	●	100%	<input type="checkbox"/>

序号	ADDRESSID	ADDRESS	PROVCEID	STREET	ZIP	CONTACTID	ADDRTYPID	CRDT	CRUSR	MDDT	MDUSR	PRMADDRESS
1	1031990024	100号余杭供电局1018	330110	南苑街道	311100	8008223	2	2010-08-03 16:56:57.0	9084	2010-08-03 16:56:57.0		去敏
1	1031990024	河北保定涿州市涿鹿县	330110	南苑街道	311100	8008223	2	2010-08-03 16:56:57.0	9084	2010-08-03 16:56:57.0		去敏
2	1031990025	338号太和商务中心11楼114室	440402	柠溪路	519000	3637681	2	2010-08-03 16:57:00.0	9083	2010-08-03 16:57:00.0		透敏
2	1031990025	河南省开封市兰考县晏庙陈村	440402	柠溪路	519000	3637681	2	2010-08-03 16:57:00.0	9083	2010-08-03 16:57:00.0		透敏
3	1031990026	41号3号楼2单元401室	370205	重庆南路	266000	4083498	2	2010-08-03 16:57:03.0	9089	2011-02-20 09:55:13.0	9027	
3	1031990026	辽宁省盘锦市盘山县	370205	重庆南路	266000	4083498	2	2010-08-03 16:57:03.0	9089	2011-02-20 09:55:13.0	9027	
4	1031990027	北岛新村16号老年康复医院收费处	320602		226000	11292175	2	2010-08-03 16:57:05.0	9095	2012-05-23 15:49:16.0	6934	透敏
4	1031990027	辽宁省葫芦岛市绥中县南京路会大	320602		226000	11292175	2	2010-08-03 16:57:05.0	9095	2012-05-23 15:49:16.0	6934	透敏
5	1031990030	华新开发区船山实验中学	430407		421000	5304734	2	2010-08-03 16:55:56.0	9094	2012-05-15 15:02:05.0	66362	透敏
5	1031990030	重庆市长寿区青春雅舍	430407		421000	5304734	2	2010-08-03 16:55:56.0	9094	2012-05-15 15:02:05.0	66362	透敏
6	1031990033	河南郑州郑州市郑东新区中兴大道天山路9号3单元601	410196		450000	000000000493087	1	2010-08-03 16:57:41.0	1024	2011-01-20 17:15:34.0	67089	
6	1031990033	陕西省宝鸡市金台区南中国雅仕苑	410196		450000	000000000493087	1	2010-08-03 16:57:41.0	1024	2011-01-20 17:15:34.0	67089	
7	1031990034	16号5栋3单元A座	510107	桐梓林北路	610000	4342162	2	2010-08-03 16:57:52.0	67017	2012-12-18 12:42:24.0	9082	透敏
7	1031990034	重庆市酉阳土家	510107	桐梓林北路	610000	4342162	2	2010-08-03 16:57:52.0	67017	2012-12-18 12:42:24.0	9082	透敏
8	1031990003	石岩街道地塘所	440306		518100	8707269	2	2010-08-03 16:58:02.0	9081	2010-08-03 16:58:02.0		
8	1031990003	广西壮族自治区	440306		518100	8707269	2	2010-08-03 16:58:02.0	9081	2010-08-03 16:58:02.0		
9	1031990038	安徽省合肥市蜀山区柏堰新村三期柏苑12栋	340702		244000	8742702	2	2010-08-03 16:58:02.0	9059	2013-09-22 11:06:08.0	68995	
9	1031990038	安徽省合肥市蜀山区柏堰新村三期柏苑12栋	340702		244000	8742702	2	2010-08-03 16:58:02.0	9059	2013-09-22 11:06:08.0	68995	
10	1031990039	江苏省南京市鼓楼区龙江白云园小区100号201	320106		210000	6559806	2	2010-08-03 16:56:38.0	6911	2010-08-03 16:56:38.0		
10	1031990039	广东省汕头市龙湖区金珠坊	320106		210000	6559806	2	2010-08-03 16:56:38.0	6911	2010-08-03 16:56:38.0		

图 4.1.4 查看数据对比

4.2 脱敏产品监控

脱敏产品监控分为：数据表监控，数据库对象监控

数据表监控：

执行：点击执行任务开始（中途中止任务或者报错，再次点击执行可断点续传，断点续传功能适用于脱敏产品内存方式，不适用于脱敏产品落地文件方式）

中止：点击中止任务停止执行

初始化：需重新执行任务之前，点击初始化，之前日志将被清除

通过图标颜色查看每个环节运行状态，查看记录数等信息，如下图：

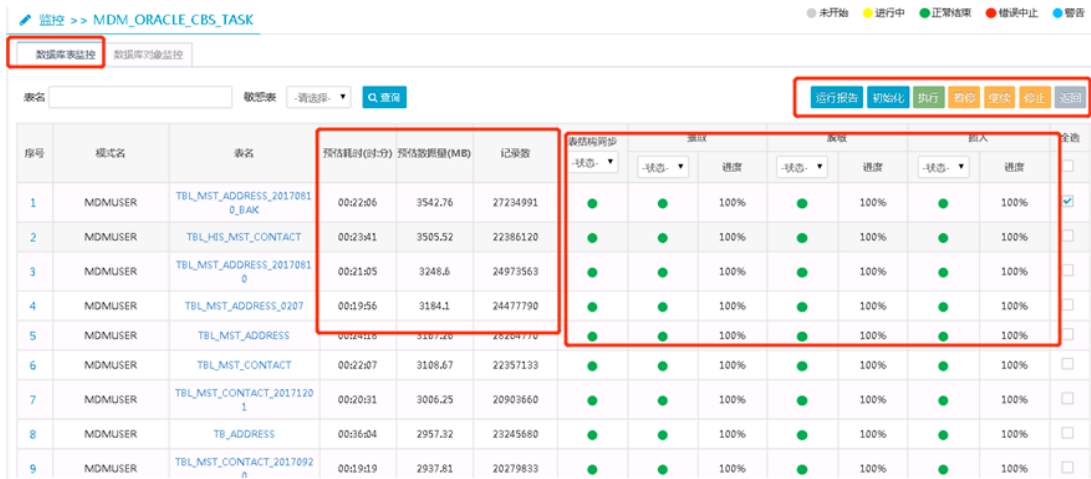


图 4.2.1 脱敏产品监控页面

表结构同步&单表执行：勾选表名点击执行 选择相应的操作事项，如下图：

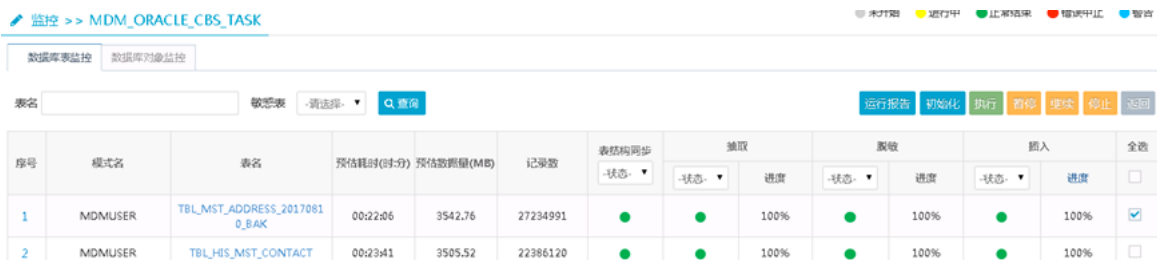


图 4.2.2 脱敏产品表结构同步&单表执行

查看日志：点击序号查看日志如下图：

序号	模式名	表名	预估耗时(时:分)	预估数据量(MB)	记录数	表结构同步	抽取	脱敏	插入	全选
1	MDMUSER	TBL_MST_ADDRESS_2017081_0_BAK	00:22:06	3542.76	27234991	●	● 100%	● 100%	● 100%	<input checked="" type="checkbox"/>
2	MDMUSER	TBL_HIS_MST_CONTACT	00:23:41	3505.52	22386120	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
3	MDMUSER	TBL_MST_ADDRESS_2017081_0	00:21:05	3248.6	24973563	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
4	MDMUSER	TBL_MST_ADDRESS_0207	00:19:56	3184.1	24477790	●	● 100%	● 100%	● 100%	<input type="checkbox"/>

图 4.2.3 脱敏产品查看日志

数据对比：点击表名查看数据前 10 行原始数据与脱敏数据的对比如下图：

序号	模式名	表名	预估耗时(时:分)	预估数据量(MB)	记录数	表结构同步	抽取	脱敏	插入	全选
1	MDMUSER	TBL_MST_ADDRESS_2017081_0_BAK	00:22:06	3542.76	27234991	●	● 100%	● 100%	● 100%	<input checked="" type="checkbox"/>
2	MDMUSER	TBL_HIS_MST_CONTACT	00:23:41	3505.52	22386120	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
3	MDMUSER	TBL_MST_ADDRESS_2017081_0	00:21:05	3248.6	24973563	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
4	MDMUSER	TBL_MST_ADDRESS_0207	00:19:56	3184.1	24477790	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
5	MDMUSER	TBL_MST_ADDRESS	00:24:18	3167.26	28264770	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
6	MDMUSER	TBL_MST_CONTACT	00:22:07	3108.67	22357133	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
7	MDMUSER	TBL_MST_CONTACT_2017120_1	00:20:31	3006.25	20903660	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
8	MDMUSER	TBL_ADDRESS	00:36:04	2957.32	23245680	●	● 100%	● 100%	● 100%	<input type="checkbox"/>
9	MDMUSER	TBL_MST_CONTACT_2017092	00:19:19	2937.81	20279833	●	● 100%	● 100%	● 100%	<input type="checkbox"/>

序号	ADDRESSID	ADDRESS	PROVCEID	STREET	ZIP	CONTACTID	ADDRTPID	CRDT	CRUSR	MDOT	MDUSR	PRMADDRESS
1	1031990024	100号余杭供电局1018	330110	南苑街道	311100	8008223	2	2010-08-03 16:56:57.0	9084	2010-08-03 16:56:57.0		去前
1	1031990024	河北保定涿州市清城县	330110	南苑街道	311100	8008223	2	2010-08-03 16:56:57.0	9084	2010-08-03 16:56:57.0		去前
2	1031990025	338号太和商务中心11楼114室	440402	柠溪路	519000	3637681	2	2010-08-03 16:57:00.0	9083	2010-08-03 16:57:00.0		透前
2	1031990025	河南开封市兰考县晏庙镇陈祥	440402	柠溪路	519000	3637681	2	2010-08-03 16:57:00.0	9083	2010-08-03 16:57:00.0		透前
3	1031990026	41号3号楼2单元401室	370205	霍庆南路	266000	4083498	2	2010-08-03 16:57:03.0	9089	2011-02-20 09:55:13.0	9027	
3	1031990026	辽宁省盘锦市盘山县	370205	霍庆南路	266000	4083498	2	2010-08-03 16:57:03.0	9089	2011-02-20 09:55:13.0	9027	
4	1031990027	北高新村16号老年康乐医院收费处	320602		226000	11292175	2	2010-08-03 16:57:05.0	9095	2012-05-23 15:49:16.0	6934	透前
4	1031990027	辽宁省葫芦岛市绥中县南京路会大	320602		226000	11292175	2	2010-08-03 16:57:05.0	9095	2012-05-23 15:49:16.0	6934	透前
5	1031990030	华新开发区韶山实验中学	430407		421000	5304734	2	2010-08-03 16:55:56.0	9094	2012-05-15 15:02:05.0	66362	透前
5	1031990030	重庆市长寿区青春雅舍	430407		421000	5304734	2	2010-08-03 16:55:56.0	9094	2012-05-15 15:02:05.0	66362	透前
6	1031990033	河南郑州市郑东新区中兴东大街东段9号3单元601	410196		450000	000000000493087	1	2010-08-03 16:57:41.0	1024	2011-01-20 17:15:34.0	67089	
6	1031990033	陕西省宝鸡市金台区南中国雅仕苑	410196		450000	000000000493087	1	2010-08-03 16:57:41.0	1024	2011-01-20 17:15:34.0	67089	
7	1031990034	16号5栋3单元A座	510107	桐梓林北路	610000	4342162	2	2010-08-03 16:57:52.0	67017	2012-12-18 12:42:24.0	9082	透前
7	1031990034	重庆市酉阳土家	510107	桐梓林北路	610000	4342162	2	2010-08-03 16:57:52.0	67017	2012-12-18 12:42:24.0	9082	透前
8	1031990003	石岩街道地粉所	440306		518100	8707269	2	2010-08-03 16:58:02.0	9081	2010-08-03 16:58:02.0		
8	1031990003	广西壮族自治区	440306		518100	8707269	2	2010-08-03 16:58:02.0	9081	2010-08-03 16:58:02.0		
9	1031990038	安徽省蚌埠市淮上区怀远镇怀远镇怀远镇怀远镇	340702		244000	8742702	2	2010-08-03 16:58:02.0	9059	2013-09-22 11:06:08.0	66995	
9	1031990038	安徽省蚌埠市淮上区怀远镇怀远镇怀远镇怀远镇	340702		244000	8742702	2	2010-08-03 16:58:02.0	9059	2013-09-22 11:06:08.0	66995	
10	1031990039	江苏省南京市鼓楼区龙江白云小区100号201	320106		210000	6559806	2	2010-08-03 16:56:38.0	6911	2010-08-03 16:56:38.0		
10	1031990039	广东省汕头市龙湖区金珠坊	320106		210000	6559806	2	2010-08-03 16:56:38.0	6911	2010-08-03 16:56:38.0		

图 4.2.4 脱敏产品数据对比

数据库对象监控：

数据库对象监控可以清楚看到表的主键，索引，视图等信息的同步情况，即数据表监控任务全部执行完成以后，数据库对象监控自动开始同步

可以选择单个对象类型点击同步功能如下图：



图 4.2.5 脱敏产品数据对象监控

查看日志：点击序号可以查看日志如下图：



图 4.2.6 脱敏产品数据库对象监控查看日志

查看原库表的 DDL 语句：点击源对象名可以查看 DDL 语句如下图：

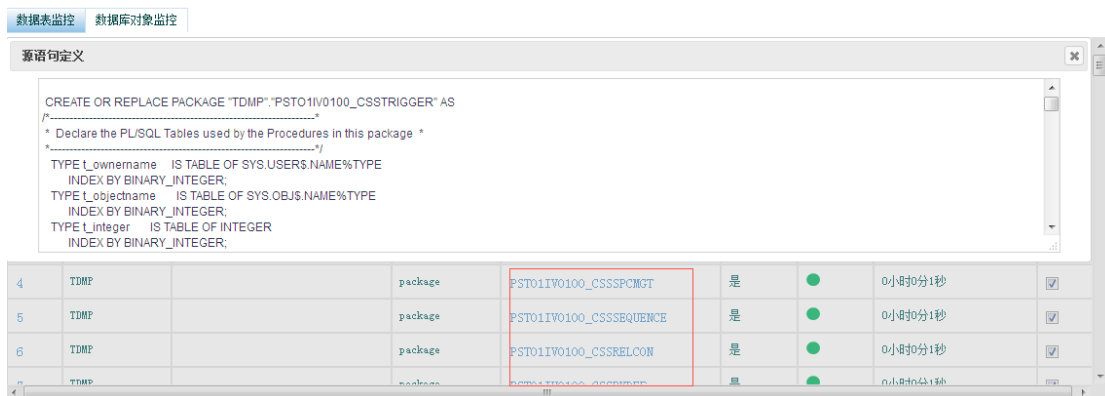


图 4.2.6 脱敏产品数据库对象监控查看 DDL 语句

4.3 SFTP/FTP 远程文本脱敏监控

远程文本脱敏监控页面如下：

执行：点击执行，任务开始

中止：点击中止，任务中断

移动源文件：任务执行完成，源文件会自动或者手动执行移动到指定的目标目录中

初始化：需重新执行任务之前，点击初始化，之前日志将被清除

通过页面可以查看文件运行状态如下图：



图 4.3.1SFTP/FTP 远程文本脱敏监控

单个文件执行：选择文件名点击右键可以单个文件执行如下图：



图 4.3.2SFTP/FTP 远程文本单个文件执行

查看日志：点击序号可以查看日志如下图：



图 4.3.3SFTP/FTP 远程文本查看日志

数据对比：点击文件名可以查看前 10 行源文件与目标文件的脱敏对比如下图：

序号	模式名	文件名	数据下载		数据脱敏		数据上传		全选
			-状态-	进度	-状态-	进度	-状态-	进度	
1	/data/remote_file_test/DB2USER/	DB2USER.optin_details.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
2	/data/remote_file_test/DB2USER/	DB2USER.optin_orders.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	/data/remote_file_test/DB2USER/	DB2USER.optin_sales.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>

全部记录数 3 条记录

图 4.3.4SFTP/FTP 远程文本数据对比

4.4 SFTP/FTP 源文件格式化文本->目标数据库（GP）监控

SFTP/FTP 源文件格式化文本->目标数据库（GP）监控页面如下：

执行：点击执行，任务开始

中止：点击中止，任务中断

移动源文件：任务执行完成，源文件会自动或者手动执行移动到指定的目标目录中

初始化：需重新执行任务之前，点击初始化，之前日志将被清除

通过页面可以查看文件运行状态如下图：

序号	模式名	文件名	数据下载		数据脱敏		数据加载		全选
			-状态-	进度	-状态-	进度	-状态-	进度	
1	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optin_customers_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
2	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optin_orders.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optin_sales_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>

图 4.4.1 SFTP/FTP源文件格式化文本->目标数据库（GP）监控页面

单个执行文件：点击右键选择单个执行文件，如下图：

您的位置: 监控

提示: ● 未开始 ● 进行中 ● 错误中止 ● 正常结束 ● 警告

初始化 执行 中止 移动源文件

序号	模式名	文件名	数据下载		数据脱敏		数据加载		全选
			-状态-	进度	-状态-	进度	-状态-	进度	
1	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_customers_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
2	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_orders.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_sales_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>

图 4.4.2 SFTP/FTPSFTP/FTP 源文件格式化文本->目标数据库 (GP) 单个执行文件

查看日志: 点序号查看日志如下图:

提示: ● 未开始 ● 进行中 ● 错误中止 ● 正常结束 ● 警告

初始化 执行 中止 移动源文件

序号	模式名	文件名	数据下载		数据脱敏		数据加载		全选
			-状态-	进度	-状态-	进度	-状态-	进度	
1	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_customers_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
2	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_orders.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_sales_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>

图 4.4.3 SFTP/FTPSFTP/FTP 源文件格式化文本->目标数据库 (GP) 查看日志

查看数据对比, 点击文件名可查看数据前 10 上的脱敏前后数据对比如下图:

提示: ● 未开始 ● 进行中 ● 错误中止 ● 正常结束 ● 警告

初始化 执行 中止 移动源文件

序号	模式名	文件名	数据下载		数据脱敏		数据加载		全选
			-状态-	进度	-状态-	进度	-状态-	进度	
1	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_customers_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
2	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_orders.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>
3	/data/remote_file_test/TDMP/PUBLIC/	PUBLIC.optim_sales_copy.txt	●	100%	●	100%	●	100%	<input type="checkbox"/>

图 4.4.4 SFTP/FTPSFTP/FTP 源文件格式化文本->目标数据库 (GP) 数据对比

4.5 定时调度

对于已经配置好的任务可以设置一个时间点去进行自动调度

点击任务管理--定时调度--新增



图 4.5.1 新增定时调度

输入作业名称--选择脱敏任务名称--调度时限--执行时间--调度计划后提交

图 4.5.1 新增定时调度

需要生效时打开启用，右侧可以对此作业删除和修改操作

编号	作业名称	脱敏任务名称	调度状态	调度计划	开始时间	结束时间	下次调度时间	启用	操作
1	dw_aboss2_aboss_sched	dw_aboss2_aboss_task	任务过期	每周六执行一次	2019-04-24	2019-04-25	已过期，不再执行	<input checked="" type="checkbox"/>	编辑 删除

图 4.5.1 新增定时调度

5. 日志

日志查询板块会对用户进行所有操作的完整记录，用于审计人员进行审计、查看。日志页面可以按照日志的选择类别进行查看，严重性，以及具体的时间段查看日志信息，具体如下图：

图 5.1 日志查询

日志下载：此下载文件主要是后台整体日志可以下载与删除操作如下图：



图 5.2 日志下载

6. 报表

报表板块记录了每个 workflow 任务的总体运行情况和详细运行情况。

其中脱敏报表能看到任务的总体调度信息，包括任务的总表数、敏感表数、敏感字段数、并发数、操作用户、开始时间、结束时间、总耗时等。通过脱敏报表表可以下钻到详细报表，查看每个表和数据库对象的运行详细信息。用户行为审计报告用于记录用户的日常操作信息。用户登录行为报表记录了用户每次登陆的状态信息。

脱敏报表：按照每个脱敏场景定制了每个场景的报表。如下图：

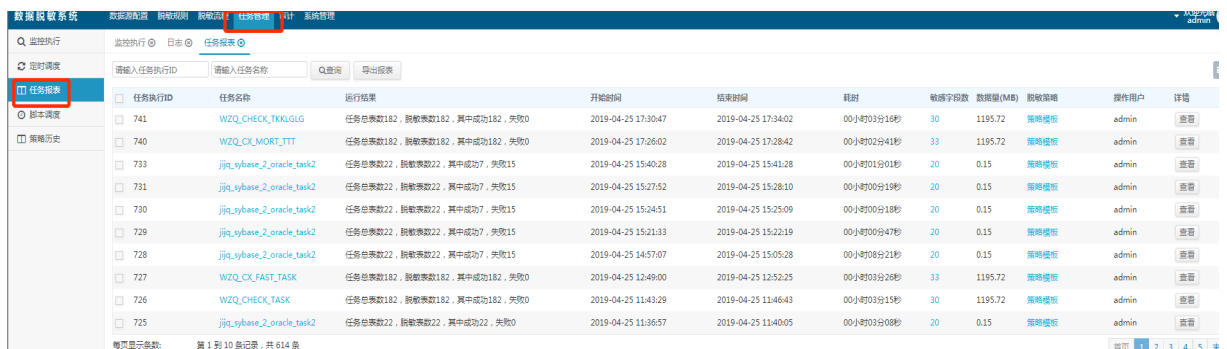


图 6.1 脱敏报表预览

报表检索：输入任务名称，支持模糊查询如下图：

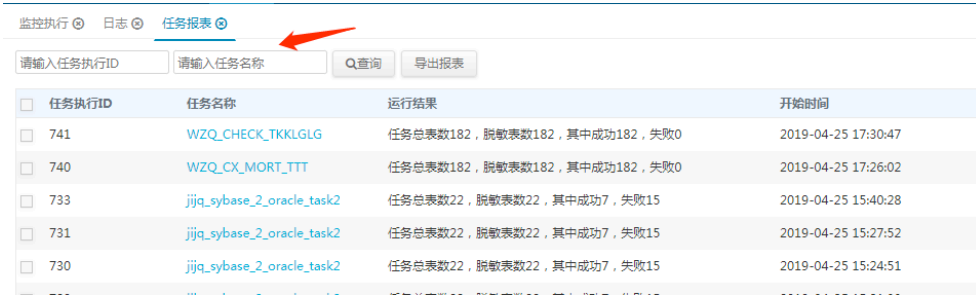


图 6.2 报表检索

导出报表：选择导出报表可以导出某个任务的执行任务信息如下图：



图 6.3 导出报表

查看任务的历史报表：点击任务的历史报表可以查看每次之行的情况如下图：

耗时	敏感字段数	数据量(MB)	脱敏策略	操作用户	详情	
4:02	00小时03分16秒	30	1195.72	策略模板	admin	查看
8:42	00小时02分41秒	33	1195.72	策略模板	admin	查看
1:28	00小时01分01秒	20	0.15	策略模板	admin	查看
8:10	00小时00分19秒	20	0.15	策略模板	admin	查看
5:09	00小时00分18秒	20	0.15	策略模板	admin	查看
2:19	00小时00分47秒	20	0.15	策略模板	admin	查看
5:28	00小时08分21秒	20	0.15	策略模板	admin	查看
2:25	00小时03分26秒	33	1195.72	策略模板	admin	查看
6:43	00小时03分15秒	30	1195.72	策略模板	admin	查看
0:05	00小时03分08秒	20	0.15	策略模板	admin	查看

图 6.4 历史报表

用户行为审计报表：用户所做的操作等信息，也包含检索与导出报表功能如下图：

编号	用户	操作时间	IP地址	操作类别	风险级别	操作状态	详情
1	admin	2019-04-26 00:54:07	192.168.7.146	任务报告检索	低	成功	详情
2	admin	2019-04-26 00:41:59	192.168.7.146	查询任务配置	低	成功	详情
3	admin	2019-04-26 00:41:09	192.168.7.146	查询任务配置	低	成功	详情
4	admin	2019-04-26 00:39:04	192.168.7.146	查询数据表	低	成功	详情
5	admin	2019-04-26 00:39:00	192.168.7.146	增加数据表	中	成功	详情
6	admin	2019-04-26 00:33:43	192.168.7.146	查询数据表	低	成功	详情
7	admin	2019-04-26 00:26:27	192.168.7.146	查询数据元数据	低	成功	详情
8	admin	2019-04-26 00:14:32	192.168.7.146	查询数据规则	低	成功	详情
9	admin	2019-04-26 00:11:23	192.168.7.146	查询数据规则	低	成功	详情

图 6.5 用户行为审计报表

用户登录行为报表：存放用户登录信息，也包含检索与导出报表功能如下图：

编号	用户	操作时间	IP地址	操作类别	风险级别	操作状态	详情
1	admin	2019-04-26 00:54:07	192.168.7.146	任务报告检索	低	成功	详情
2	admin	2019-04-26 00:41:59	192.168.7.146	查询任务配置	低	成功	详情
3	admin	2019-04-26 00:41:09	192.168.7.146	查询任务配置	低	成功	详情
4	admin	2019-04-26 00:39:04	192.168.7.146	查询数据表	低	成功	详情
5	admin	2019-04-26 00:39:00	192.168.7.146	增加数据表	中	成功	详情
6	admin	2019-04-26 00:33:43	192.168.7.146	查询数据表	低	成功	详情
7	admin	2019-04-26 00:26:27	192.168.7.146	查询数据元数据	低	成功	详情
8	admin	2019-04-26 00:14:32	192.168.7.146	查询数据规则	低	成功	详情
9	admin	2019-04-26 00:11:23	192.168.7.146	查询数据规则	低	成功	详情

图 6.6 用户登录行为报表