# 数据质量介绍

## 了解数据质量

FiData数据质量模块中质量规则的建立保障了数据工厂中流入流出的数据都能按照规则运行，对不符合规则的数据从而进行定性以及定量的清洗，最大程度保证数据准确性。

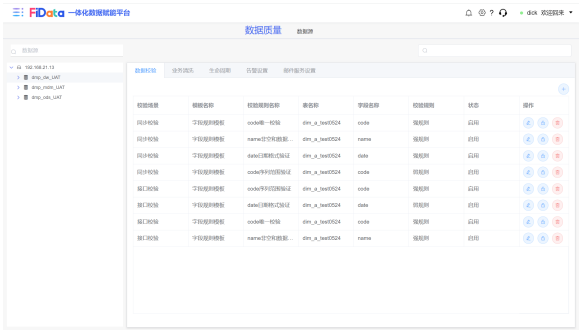

- 了解数据质量
  - 1. 概述
  - 2. 界面概览
  - 3. 创建数据源
  - 4. 配置数据质量规则
    - 4.1 界面描述
    - 4.2 配置数据校验
    - 4.3 检验场景与模板


## 1. 概述

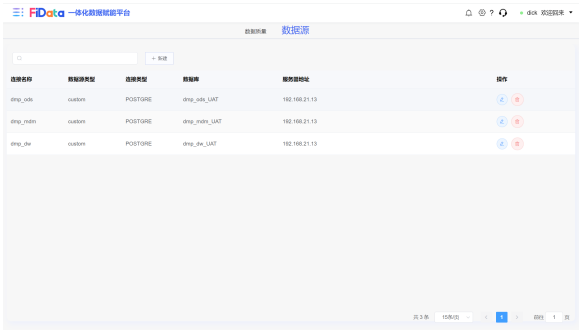
FiData数据质量模块中质量规则的建立保障了数据工厂中流入流出的数据都能按照规则运行，对不符合规则的数据从而进行定性以及定量的清洗，最大程度保证数据准确性。

## 2. 界面概览

入口

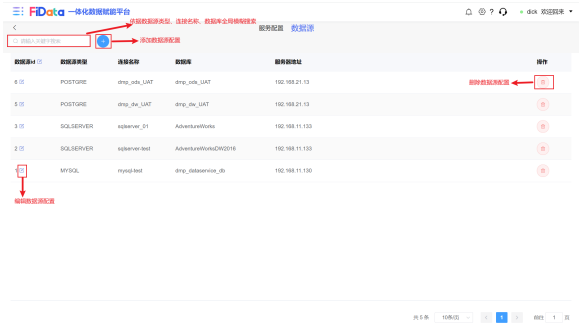



数据源设计页面

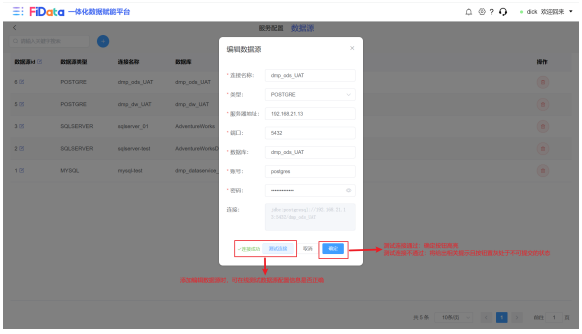



## 3. 创建数据源

数据源列表

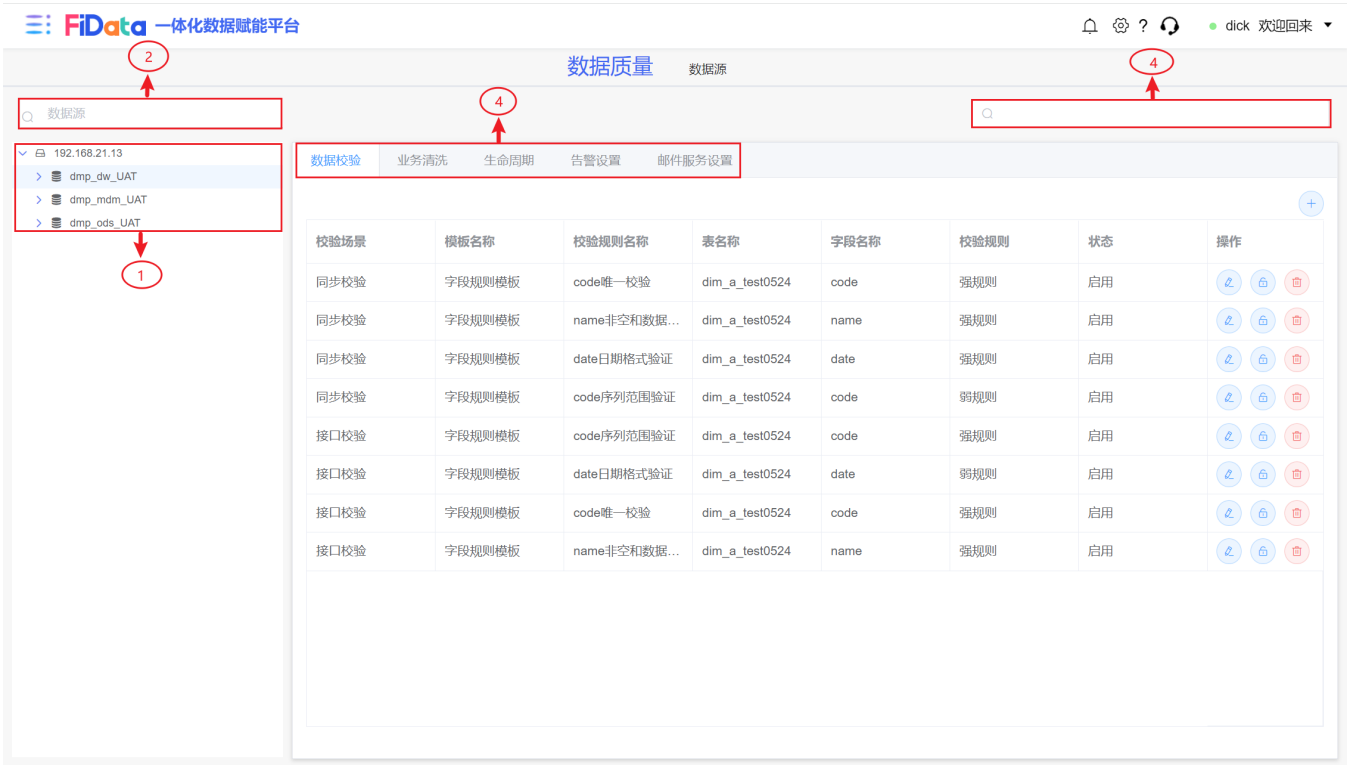



数据源新增/编辑



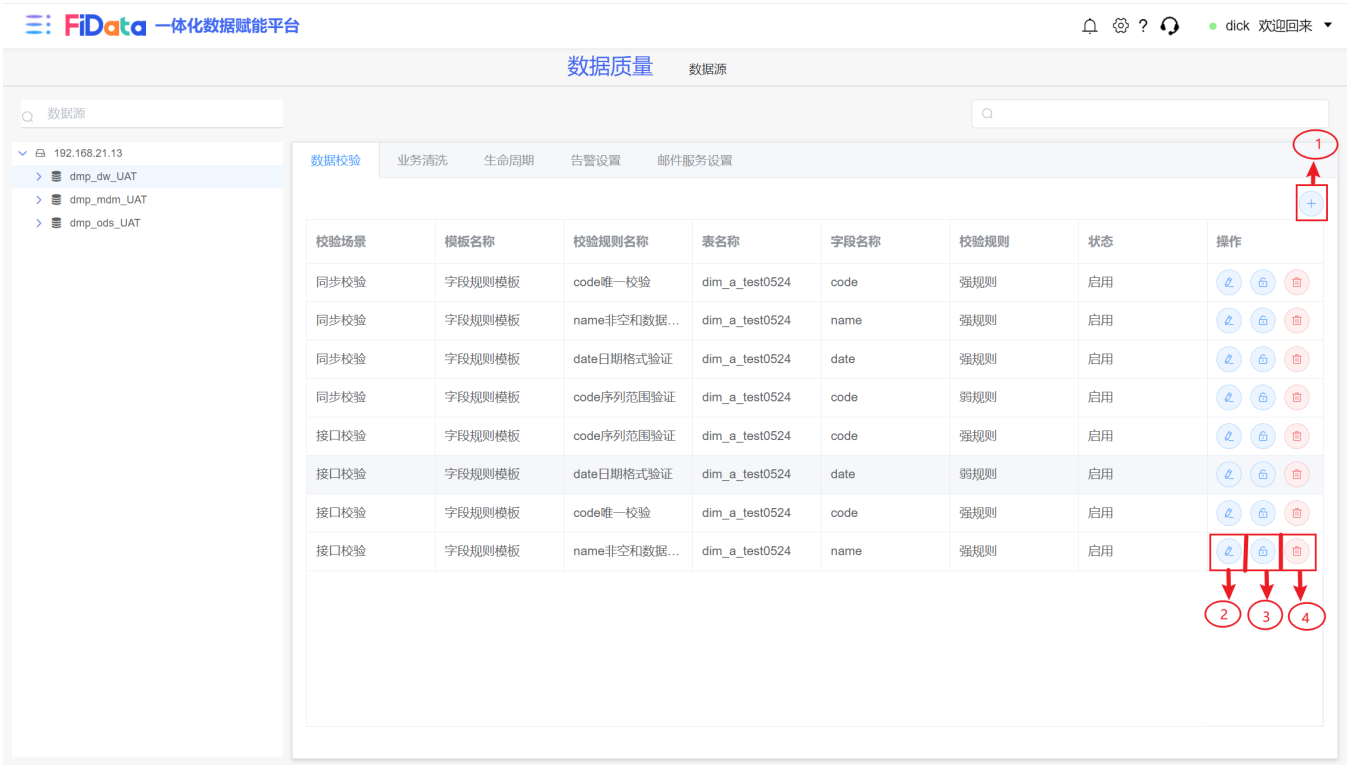

## 4. 配置数据质量规则

### 4.1 界面描述

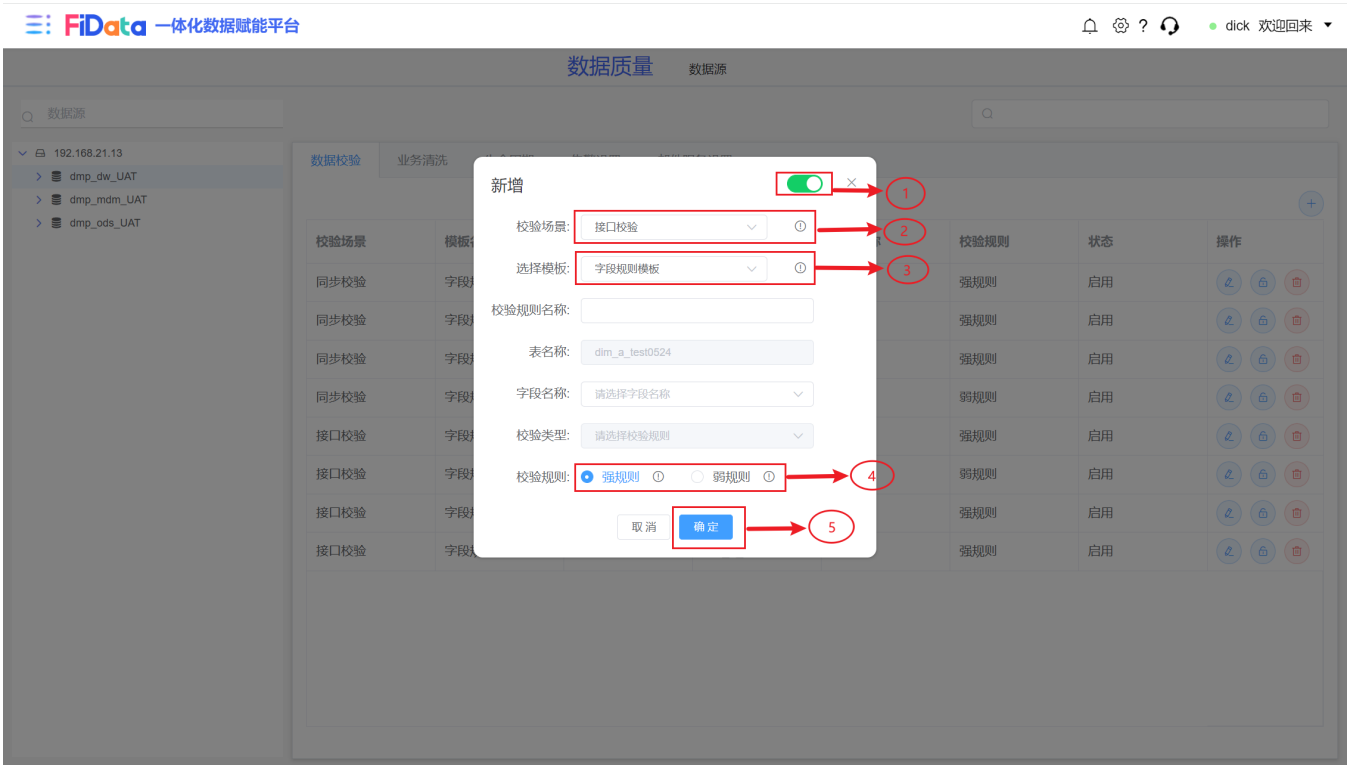



1. 数据源元数据Tree。
2. 搜索数据源节点。
3. 全局搜索规则。
4. a. 数据校验：配置表、字段的校验规则。
   b. 业务清洗：配置表数据的清洗规则。
   c. 回收周期：配置表的回收规则。
   d. 告警设置：配置通知方式，邮件/站内
   e. 邮件服务器设置：邮件通知的服务器配置

## 4.2 配置数据校验




1. 新增数据校验规则。
2. 编辑数据校验规则。
3. 启用/禁用数据校验规则。
4. 删除数据校验规则。




1. 启用禁用数据校验规则。
2. 选择校验场景。
3. 选择校验模板，不同的场景下对应各自的模板。
4. 强规则：校验不通过熔断数据流，终止后续操作，弱规则：校验不通过不熔断数据流，后续操作任然允许进行。
5. 确定：保存校验规则

### 4.3 检验场景与模板

| 模块 | 应用场景 | 模板名称 | 模板描述 |
| --- | --- | --- | --- |
| 数据校验 | 页面校验（web） | 字段校验模板 | 验证传入字段参数值是否合法，支持长度、非空、唯一、日期格式、序列范围 |
| | | 相似度对比模板 | 验证传入字段参数值的相似度比例，ods中验证 |
| | 同步校验（stg） | 字段校验模板 | 验证数据在stg进入ods之前技术规则验证，确保进入的数据不违反数据库的存储规则，支持唯一、长度、非空、日期格式、序列范围 |
| | 质量报告（ods/mdm/dw） | 业务校验模板 | 自定义验证脚本，验证ods中表数据是否符合技术规则，并将验证结果形成excel报告邮件发出，比如查询年龄大于18周岁的用户 |
| | | 字段聚合波动阀值模板 | 设置字段的聚合属性以及波动阈值，当超过或等于时，发送邮件提醒 |
| | | 表行数波动阀值模板 | 基于更新依据字段设置表行数的波动阈值，当超过或等于时，发送邮件提醒 |
| | | 空表校验模板 | 检查表是否为空，当为空时，发送邮件提醒 |
| | | 表更新校验模板 | 基于更新依据字段检查表数据是否存在更新，当存在时，发送邮件提醒 |
| | | 表血缘断裂校验模板 | 检查表上下游血缘是否断裂（数据管理服务中能查询到此表，否则不可选中此模板），当断裂时，发送邮件提醒 |