元数据管理

功能定位
元数据管理把 Doris 表、字段、中文注释、 dbt
模型、质量结果和上下游血缘集中到一个可检索目录。它回答三个问题：平台上有什么数据、字段是什么意思、这个指标从哪里来又流向哪里。

图 7-5 治理通道诊断、 Doris 元数据目录与字段血缘入口

适用角色
数据架构师、数据管理员、指标管理员、开发人员和审计人员。

页面入口

数据治理中心 → 元数据管理（`/bi/etl/metadata`）。

页面结构

区域	说明
治理通道诊断	元数据服务、OM 授权、dbt 环境、dbt 产物、Doris 采集、质量同步、ETL 调度服务
元数据目录	仅展示服务端白名单 Schema，可搜索表名或注释
字段与注释	查看字段类型、中文名、描述、敏感级别和是否可用于分析
数据血缘	查看表级和字段级上下游关系
dbt 模型	查看模型、依赖、最近运行和质量结果
质量与运行	查看质量测试、运行状态和失败原因
调度实例	查看与元数据刷新相关的工作流实例

操作步骤

1. 先查看治理通道诊断。红色状态表示该能力不可用，需管理员处理。
2. 在元数据目录选择 Schema，输入表名或中文注释查询。
3. 点击“字段”查看字段清单、类型和注释，确认口径后再用于建模或分析。
4. 点击“血缘”查看上游来源、下游模型和看板引用，评估变更影响。
5. 需要更新目录时，由管理员执行“元数据采集”，选择 Doris、dbt、质量结果或全量刷新。
6. 需要验证 dbt 时，使用“dbt 命令”执行 debug、run、test 或 docs generate。
7. 需要周期性刷新时，通过“执行受控调度”登记固定工作流，并核对调度实例结果。

输入与输出

输入是 Doris Schema、dbt manifest、质量结果和调度状态；输出是可搜索目录、字段注释、血缘图、dbt 模型清单和采集任务记录。

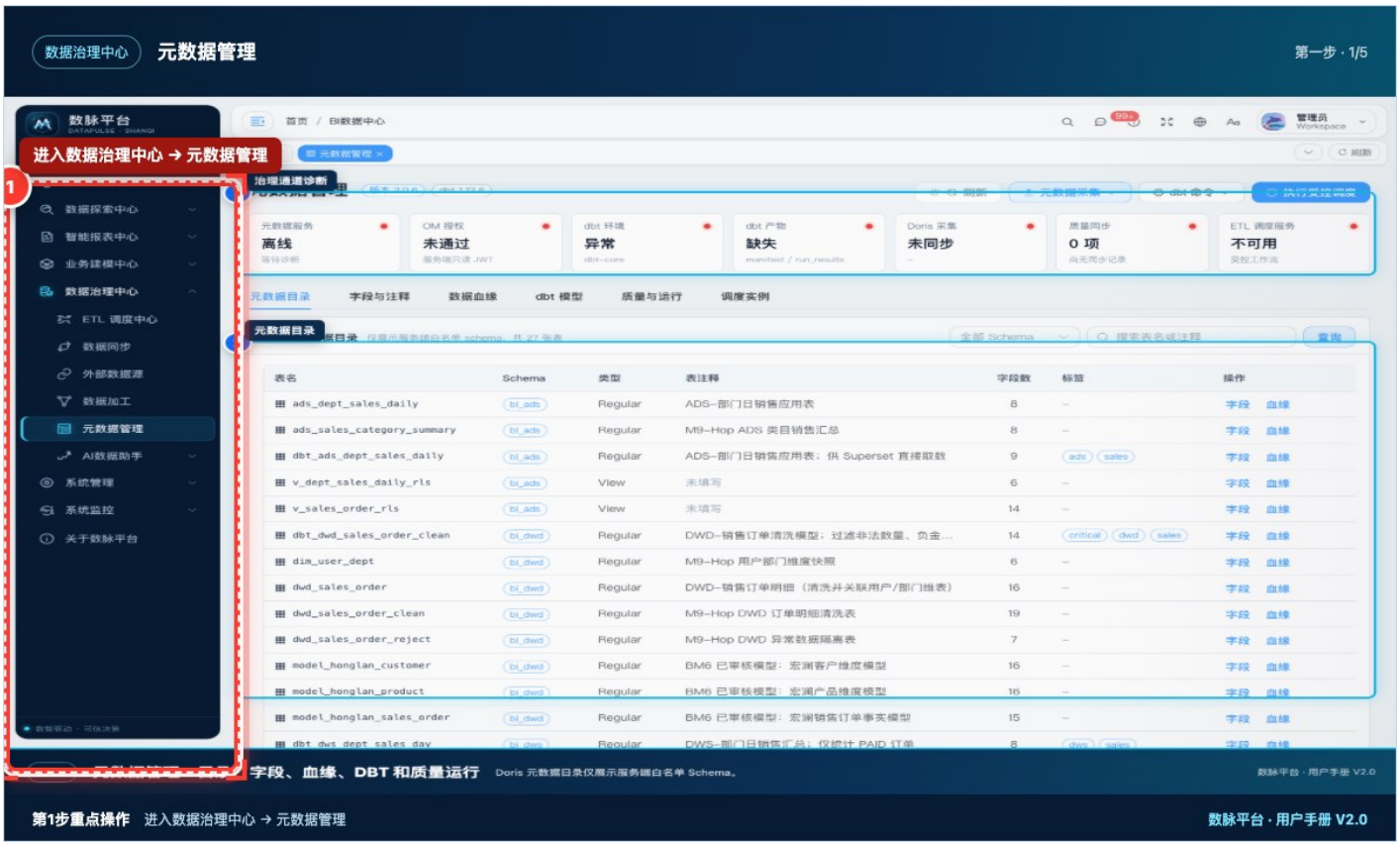
宏澜家电示例

分析师发现 “区域销售额 ”与财务口径不一致，在元数据管理中打开 ads_dept_sales_daily 的血缘：上游是 dws_sales_category_summary 和 dwd_sales_order_clean，再向上到 ERP 订单表。通过字段注释确认 “ 净销售额 ”已扣除退货，从而定位差异来自财务使用含税收入口径。

客户应用演示：追踪 “销售额 ”

字段的来源与影响
业务情境： 业务人员质疑销售额口径，需要确认字段来自哪里、被哪些资产使用。

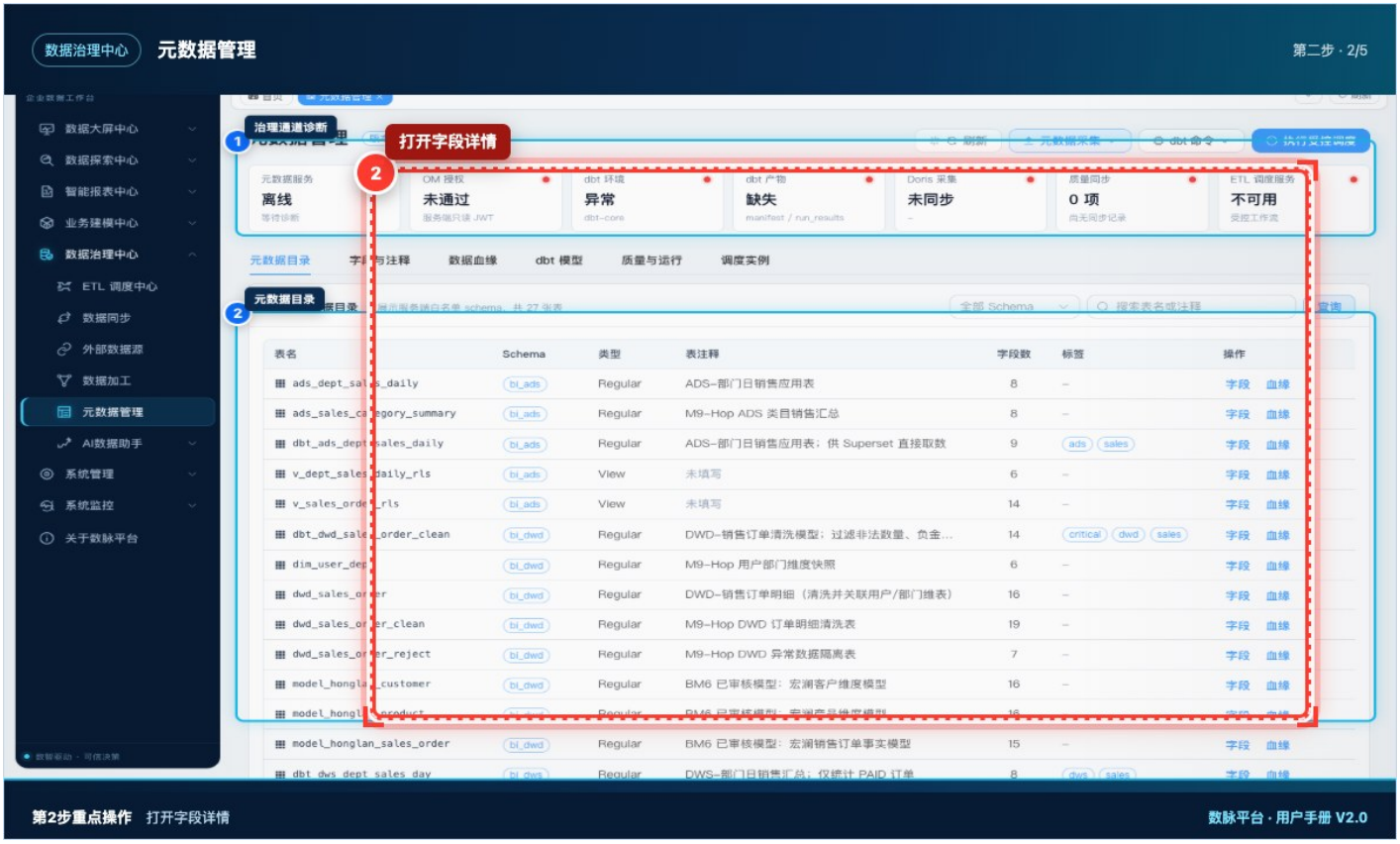
第一步：进入 “数据治理中心 → 元数据管理 ”



步骤图 39-01 第1步：进入 “数据治理中心 → 元数据管理 ”。红框标出进入数据治理中心 → 元数据管理。

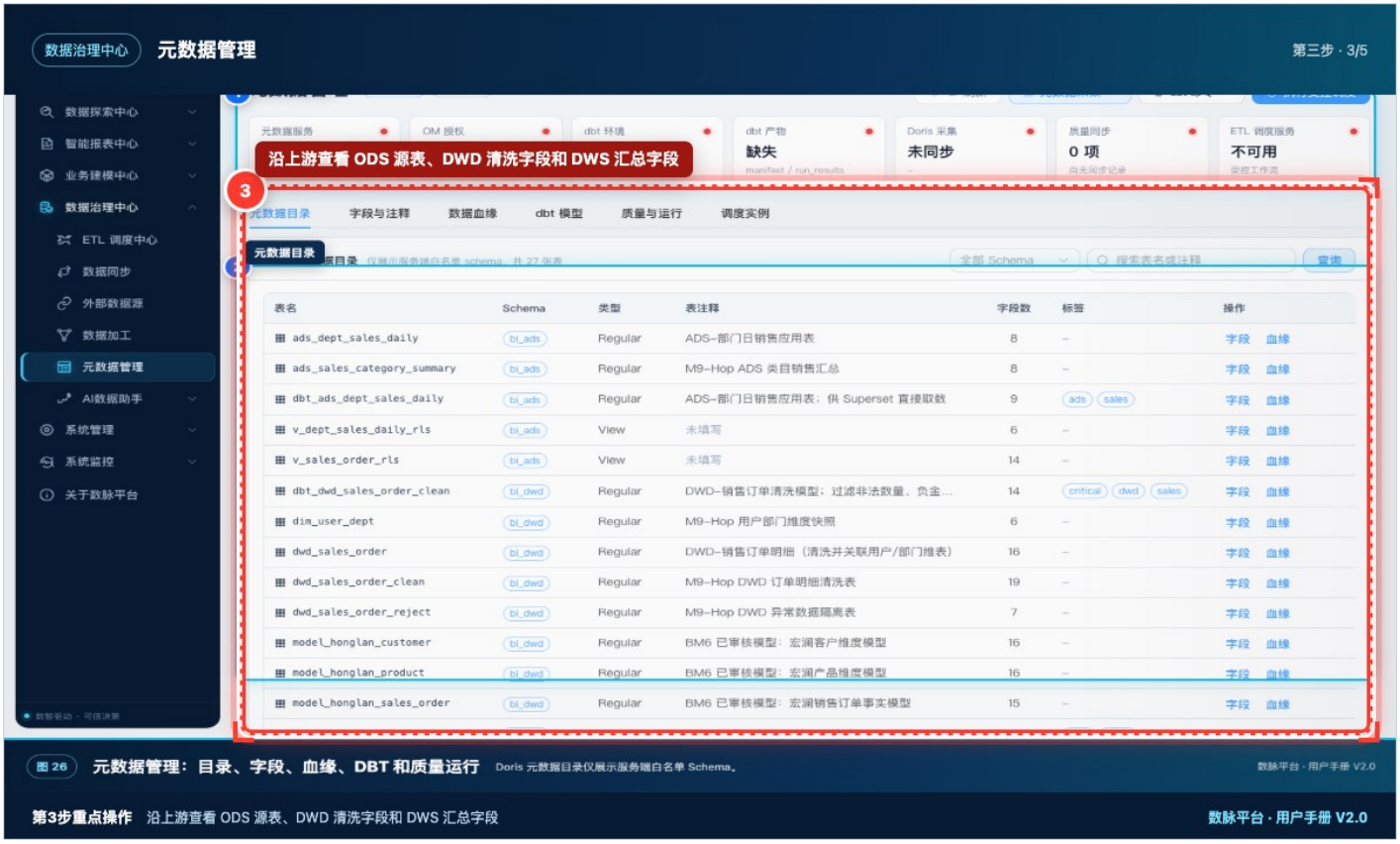
进入 “数据治理中心 → 元数据管理 ”，搜索销售额字段。

第二步：打开字段详情



步骤图 39-02 第2步：打开字段详情。红框标出打开字段详情。
打开字段详情，查看所属表、数据类型、注释和负责人。

第三步：沿上游查看 ODS 源表、 DWD 清洗字段和 DWS 汇总字段

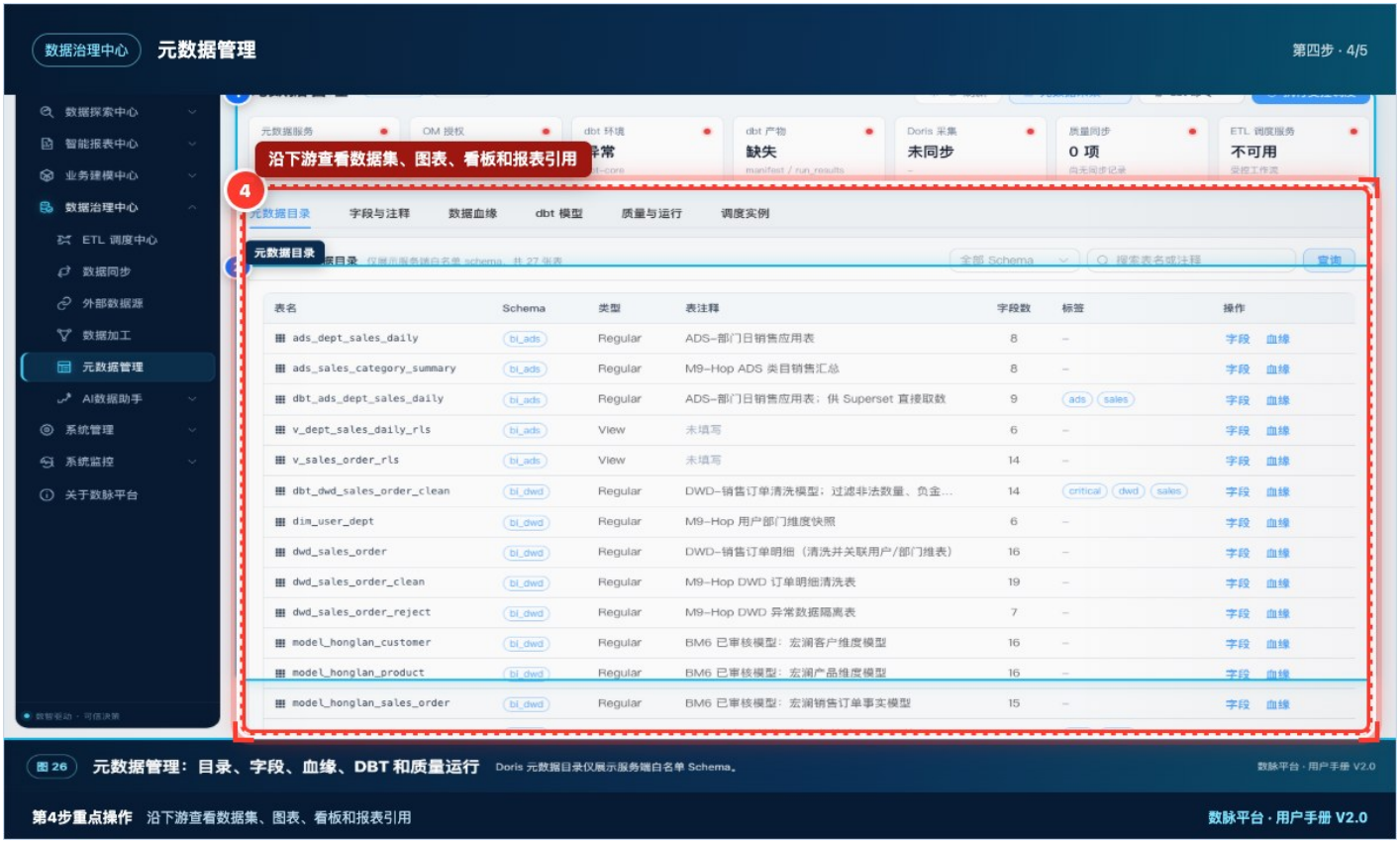


□□□ *39-03 □ 3□□□□□□□ ODS □□□ DWD □□□□□ DWS □□□□□□□□□□□□□□□ ODS □□□ DWD □□□□□ DWS □□□□□*

□□□□□ ODS □□□ DWD □□□□□ DWS □□□□□

□□□□□□□□□□□□□□□□□□□□□

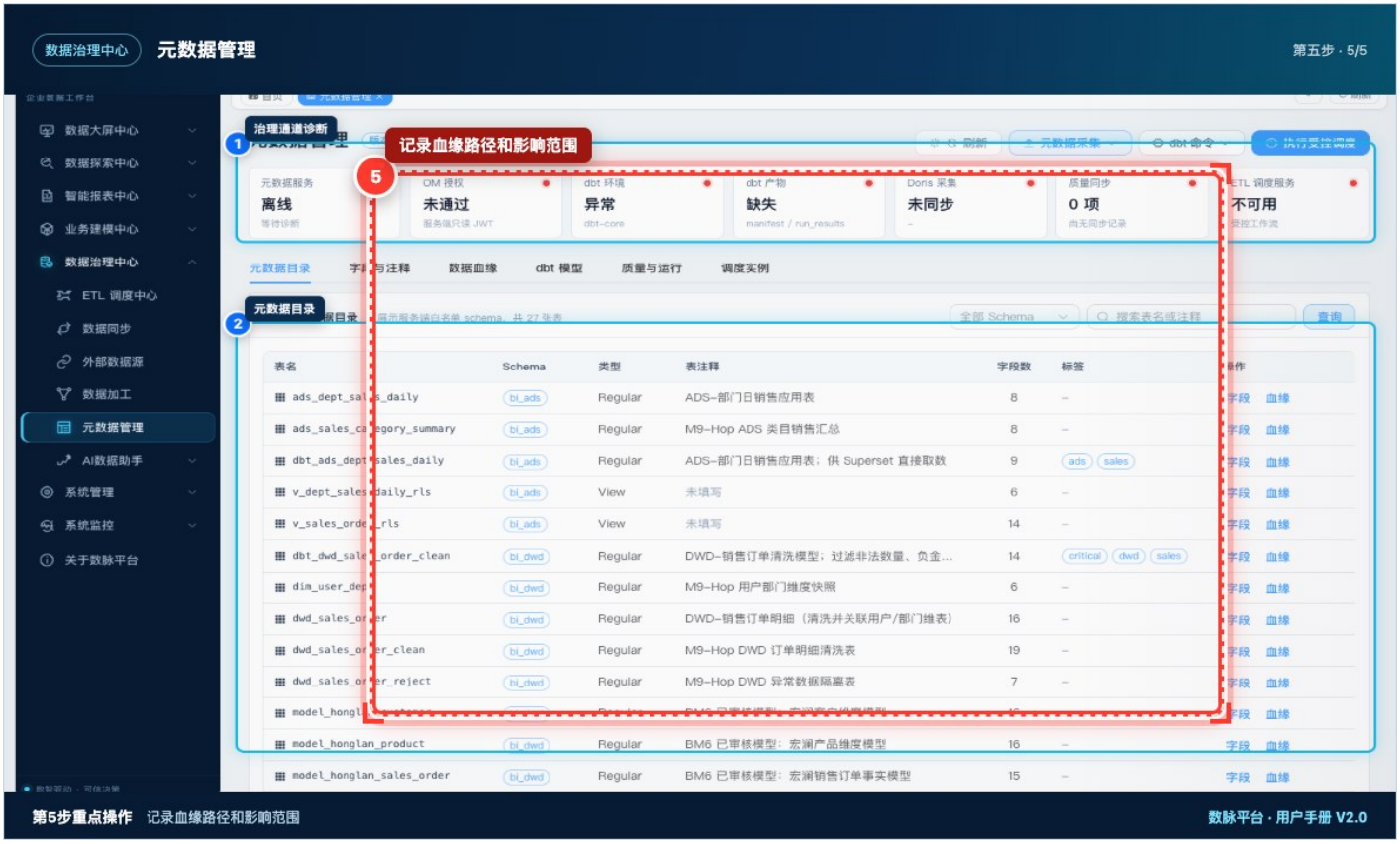
□□□



步骤图 39-04 第4
步：沿下游查看数据集、图表、看板和报表引用。红框标出沿下游查看数据集、图表、看板和报表引用。

沿下游查看数据集、图表、看板和报表引用。

第五步：记录血缘路径和影响范围



步骤图 39-05 第5步：记录血缘路径和影响范围。红框标出记录血缘路径和影响范围。
记录血缘路径和影响范围，如需修改先通知下游负责人。

最终结果与效果
页面显示字段的上游来源、处理过程和下游资产，能够判断一次口径变更会影响哪些报表。

操作提示： 血缘用于定位影响，不替代数据质量校验；修改前仍要抽样对账。 “

当前演示环境说明
截图显示元数据服务离线、 OM 授权未通过、 dbt 环境异常、 dbt 产物缺失、 Doris 采集已同步、质量同步为 0 项、 ETL 调度可用。页面同时展示 27 张白名单表。该状态用于演示诊断，不能据此认定生产血缘完整。

常见问题

搜不到表。 可能不在白名单 Schema，或尚未执行元数据采集。

血缘图只有表、没有字段。 需要 dbt manifest 和字段级依赖解析；先检查 dbt 产物是否生成。

表注释为空。 在模型或建表阶段补充 COMMENT，再重新采集，不要只在前端手工改显示。

开发人员参考
接口： /bi/metadata/status、 /catalog、 /table、 /lineage、 /dbt/models、 /dbt/runs、 /ingest、 /schedule
权限： bi:metadata:view，采集、 dbt 和调度为独立细粒度权限。
仅允许白名单 Schema；采集任务异步执行并记录输入摘要、结果和错误码。
血缘来自 dbt manifest、 Doris 元数据和登记关系的组合，不靠前端猜测。

版本号 #2
由 Admin 创建于 2026-11-06 00:23:22 UTC
由 Admin 更新于 2026-11-06 03:53:10 UTC