外部数据源

功能定位
外部数据源用于接入不适合直接连库的第三方系统，例如电商平台、企业微信、开放 API、SaaS 订单接口。系统通过受控 HTTP 出口读取分页或游标数据，写入指定 Doris
目标表，并记录批次、重试、审计和脱敏原因。

图 7-4 外部源配置、同步作业、运行实例与安全审计

适用角色
ETL 管理员、集成开发人员、数据运维人员和审计人员。

页面入口

数据治理中心 → 外部数据源（`/bi/etl/external`）。

页面结构

页签	作用
外部源	源名称、类型、接口地址、目标表、鉴权方式、状态和编辑入口
同步作业	作业名称、外部源、全量/增量模式、自动重试和最近运行
运行实例	批次号、作业、状态、读写行数、DS 实例和重试
安全审计	事件、结果、操作人、脱敏原因和来源IP

操作步骤

1. 点击“新建外部源”，选择模板：通用 REST 分页、企业微信游标或开放 API。
2. 填写源名称、接口地址、请求方法、目标表和分页参数；接口地址只允许服务端出口白名单中的域名。
3. 选择鉴权方式（如 HEADER 或 QUERY），凭据使用 `env:` 或文件引用，页面只显示“已配置”，不回显明文。
4. 保存后测试连通性，确认返回结构、分页字段、游标路径和错误码。
5. 新建同步作业，设置全量或增量、最大重试次数和启停状态。
6. 执行后查看运行实例的读取行数、写入行数和失败原因；重试保持批次号。
7. 在安全审计中核对每次配置、执行和脱敏事件。

输入与输出

输入是接口地址、鉴权引用、分页规则、字段映射和目标表；输出是写入 Doris 的外部数据、批次记录、审计事件和脱敏摘要。

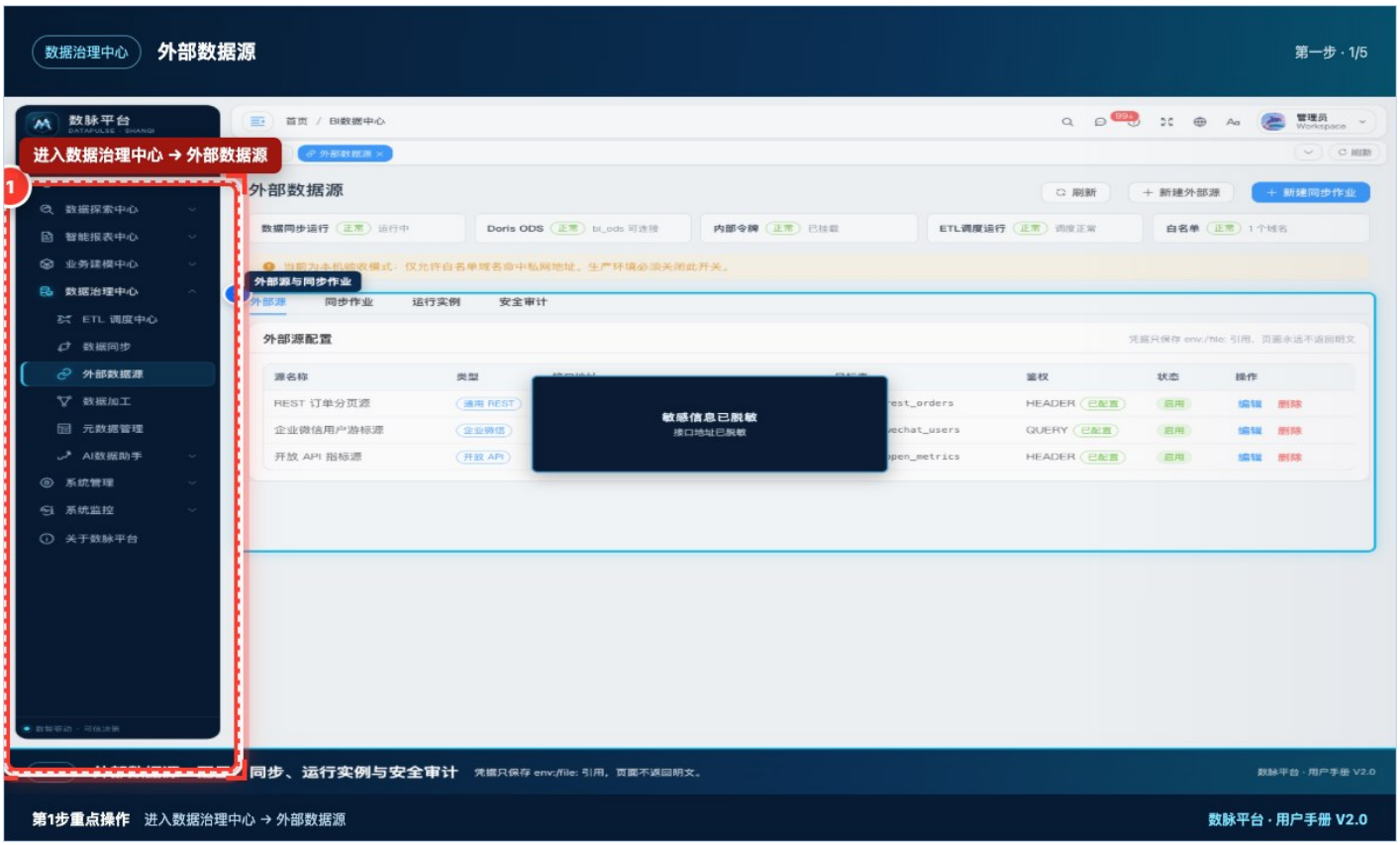
宏澜家电示例

接入某电商平台的订单接口：使用 REST 分页，按更新时间增量拉取，订单号作为幂等键写入
ods_ecommerce_order。凭据由环境变量注入，页面只显示 “HEADER 已配置 ”
。若平台限流，作业按指数退避重试并保留原批次号。

客户应用演示：接入宏澜 ERP

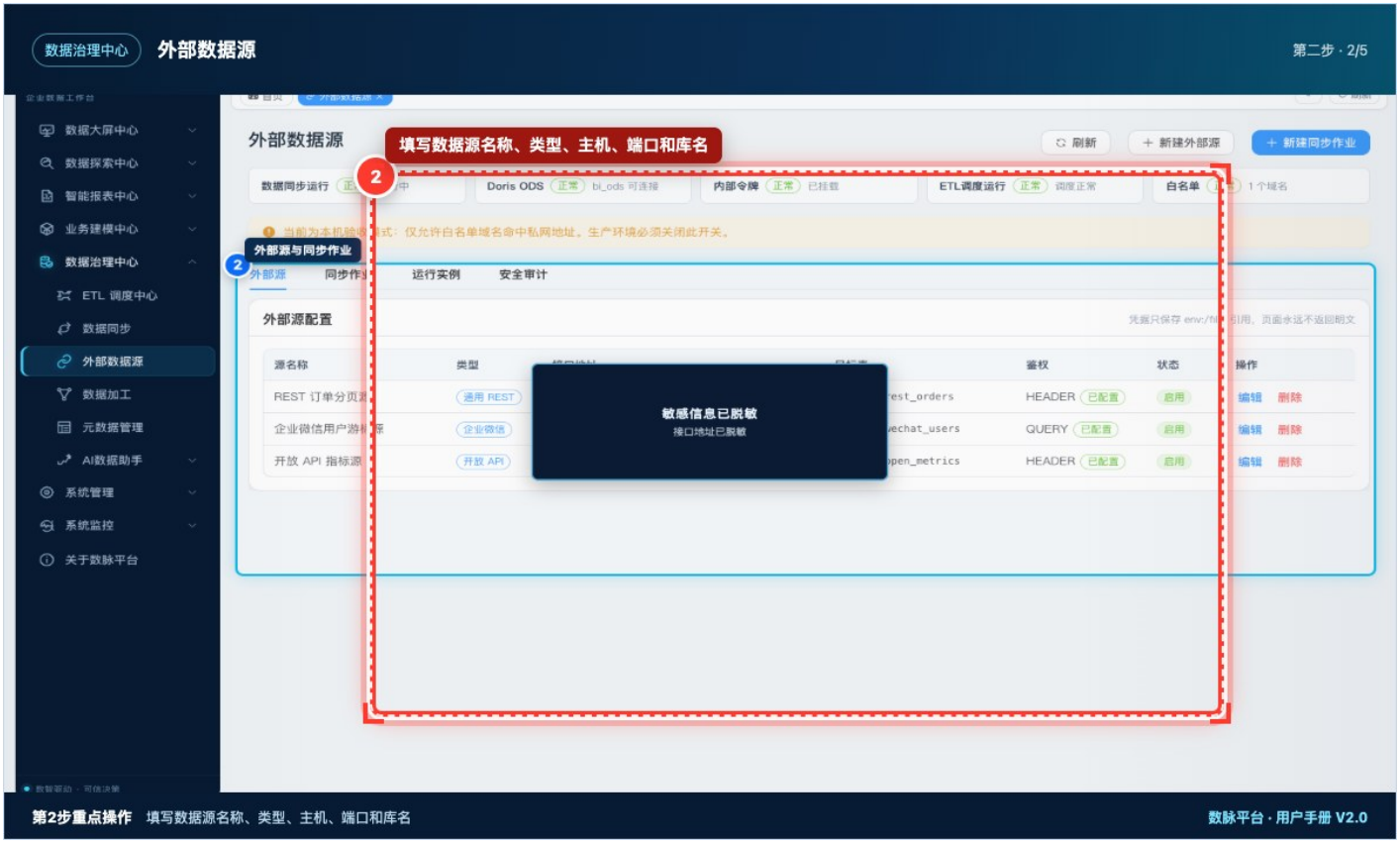
只读数据库
业务情境： 数据管理员需要把外部 ERP 作为后续同步和建模的数据源。

第一步：进入 “数据治理中心 → 外部数据源 ”



步骤图 38-01 第1步：进入 “数据治理中心 → 外部数据源 ”。红框标出进入数据治理中心 → 外部数据源。
进入 “数据治理中心 → 外部数据源 ”，点击新增。

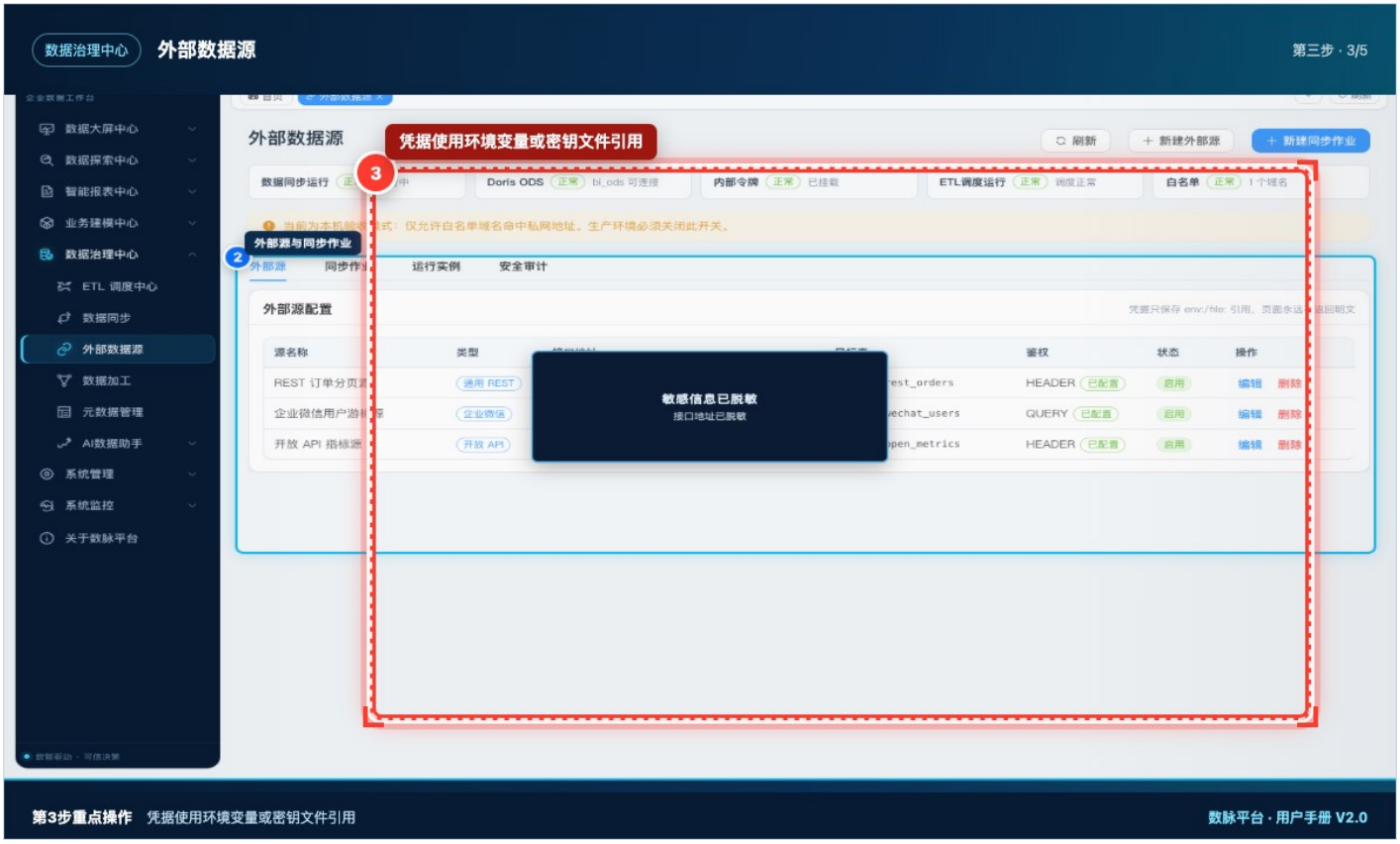
第二步：填写数据源名称、类型、主机、端口
和库名



步骤图 38-02 第2
步：填写数据源名称、类型、主机、端口和库名。红框标出填写数据源名称、类型、主机、端口和库名。

填写数据源名称、类型、主机、端口和库名。

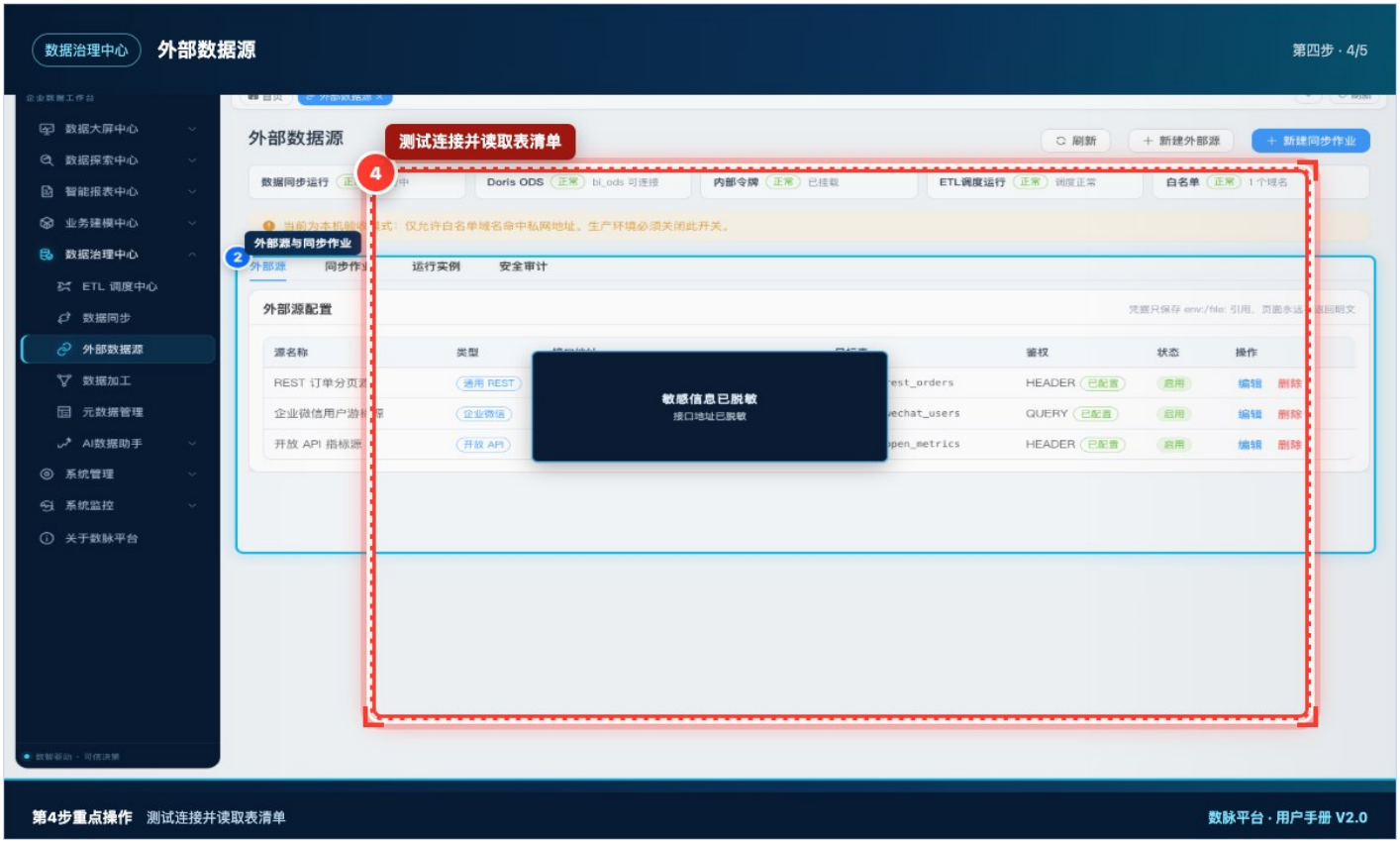
第三步：凭据使用环境变量或密钥文件引用



步骤图 38-03 第3步：凭据使用环境变量或密钥文件引用。红框标出凭据使用环境变量或密钥文件引用。

凭据使用环境变量或密钥文件引用，不填写明文到页面说明中。

第四步：测试连接并读取表清单



步骤图 38-04 第4步：测试连接并读取表清单。红框标出测试连接并读取表清单。

测试连接并读取表清单，确认只读账号权限边界。

第五步：保存后执行一次连接诊断



步骤图 38-05 第5步：保存后执行一次连接诊断。红框标出保存后执行一次连接诊断。

保存后执行一次连接诊断，并把数据源负责人和用途登记完整。

最终结果与效果
外部数据源登记成功，后续同步作业可按白名单引用；页面不回显密码，操作进入审计日志。

操作提示： “
连接失败先检查网络白名单、账号权限和密钥引用；不要把真实连接串写入文档。

当前演示环境说明
截图显示 “外部数据源集成已关闭 ”，接口地址已脱敏，白名单为 0 个域名。该页面展示配置结构和审计方式；实际调用前必须由管理员开启集成并配置允许访问的域名。

安全边界
禁止接入任意公网地址，禁止通过页面探测内网服务。
凭据只保存引用，不返回明文；日志和审计不得记录 Token、Cookie 或完整响应。
目标表必须在白名单内，字段映射必须显式声明。
删除数据源前先停用作业并确认历史批次和下游依赖。

常见问题
保存时提示域名不在白名单。 由管理员在服务端出口策略中登记域名，不要在页面反复换写法绕过。

运行成功但字段为空。 检查 JSON 路径、分页游标和字段映射，先用小页数验证。

凭据显示 “未配置 ”。 环境变量或文件引用没有加载，需管理员检查部署配置，不要把密钥直接填到浏览器。

开发人员参考
接口： /bi/ext/sources、 /jobs、 /runs、 /audits、 /options、 /status
权限： bi:etl:ext
使用受控出口网关，服务端校验目标域名、请求方法、响应大小和字段白名单。
审计只保存事件、结果和脱敏原因；凭据引用不写入数据库明文。

版本号 #2
由 Admin 创建于 2026-11-06 00:23:21 UTC
由 Admin 更新于 2026-11-06 03:53:09 UTC