数据源与画像

功能定位
数据源与画像用于登记只读数据库连接、同步元数据目录、查看表结构和执行脱敏数据画像。它回答 “ 有哪些数据、字段质量如何、能否支撑分析 ”。

图 6-4 只读数据源、表目录与画像任务

适用角色
数据管理员、数据架构师和 ETL 管理员。

页面入口
业务建模中心 → 数据源与画像（ /bi/bm/source）。

操作步骤
1. 选择项目，点击 “新增数据源 ”。
2. 填写数据源名称、类型、主机端口、数据库、 Schema 和只读账号；口令加密保存。
3. 点击 “测试连接 ”，确认网络、账号和权限。
4. 执行 “同步目录 ”，读取表、字段、类型和注释；该操作只读元数据。
5. 选择一张或多张表执行 “画像 ”，查看行数、空值率、唯一值、最值、分布和格式异常。
6. 查看任务记录和最近验证时间；同步和画像同一数据源同时只允许一个任务。
7. 不再使用的数据源先确认下游引用，再按流程删除。

输入与输出
输入是只读连接和表范围；输出是数据源状态、表目录、字段画像和脱敏统计。系统不保存原始样本行。

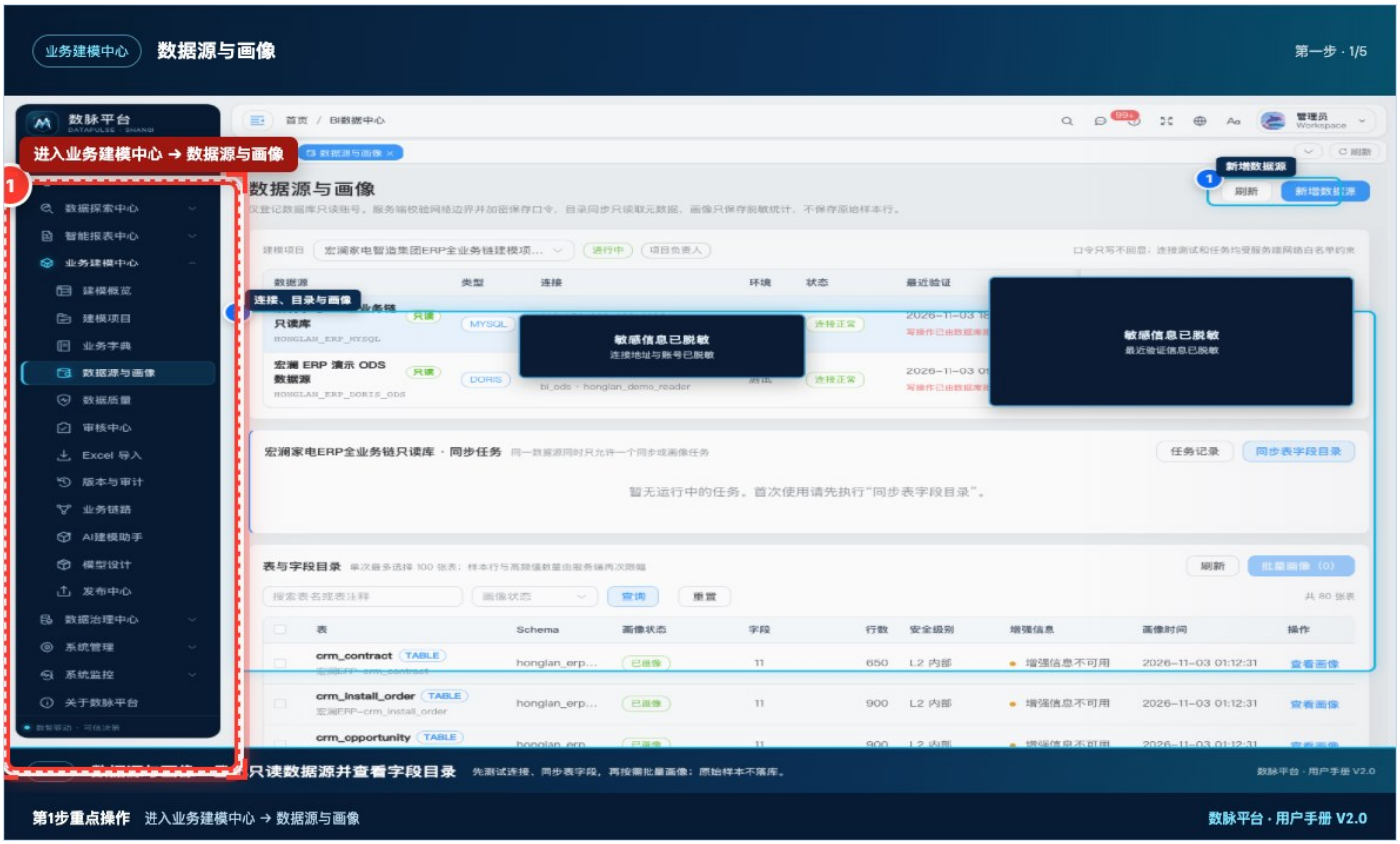
宏澜家电示例
接入 ERP 只读库后同步 80 张表，对 sales_order、 sales_order_item、 inventory_stock 和 finance_voucher 做画像，发现订单金额空值率、客户编码重复率和日期越界问题，作为质量规则设计依据。

安全边界
只允许登记只读账号；生产写账号不得接入。
连接地址和凭据不得出现在帮助文档、截图和前端日志。
画像只保存聚合统计，不保存个人手机号、地址等原始样本。

客户应用演示：登记宏澜 ERP 只读数据源

业务情境： 需要把 ERP 中的订单、客户和产品表接入数据治理流程。

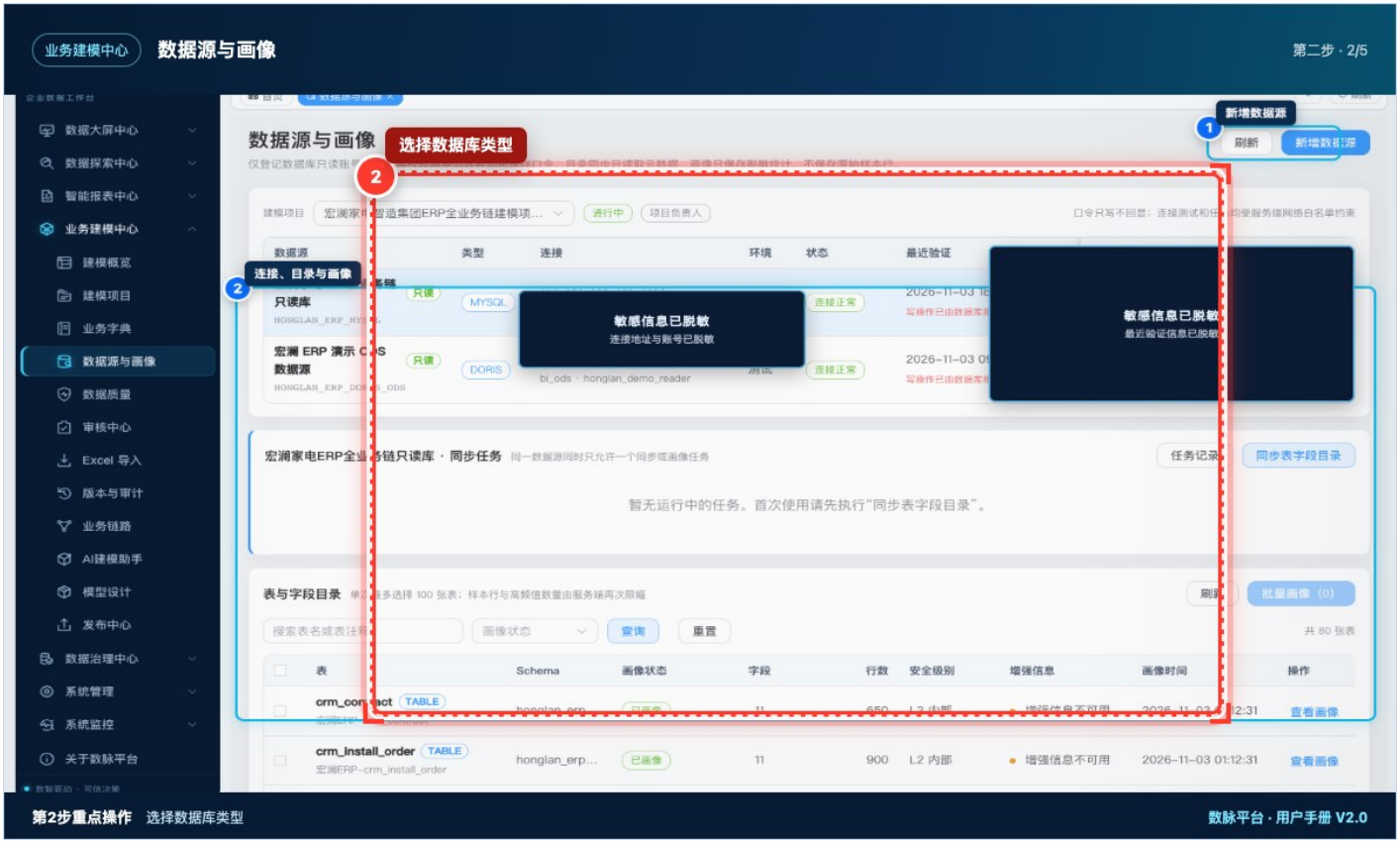
第一步：进入 “业务建模中心 → 数据源与画像 ”



步骤图 26-01 第1步：进入 “业务建模中心 → 数据源与画像 ”。红框标出进入业务建模中心 → 数据源与画像。

进入 “业务建模中心 → 数据源与画像 ”，点击 “新增数据源 ”。

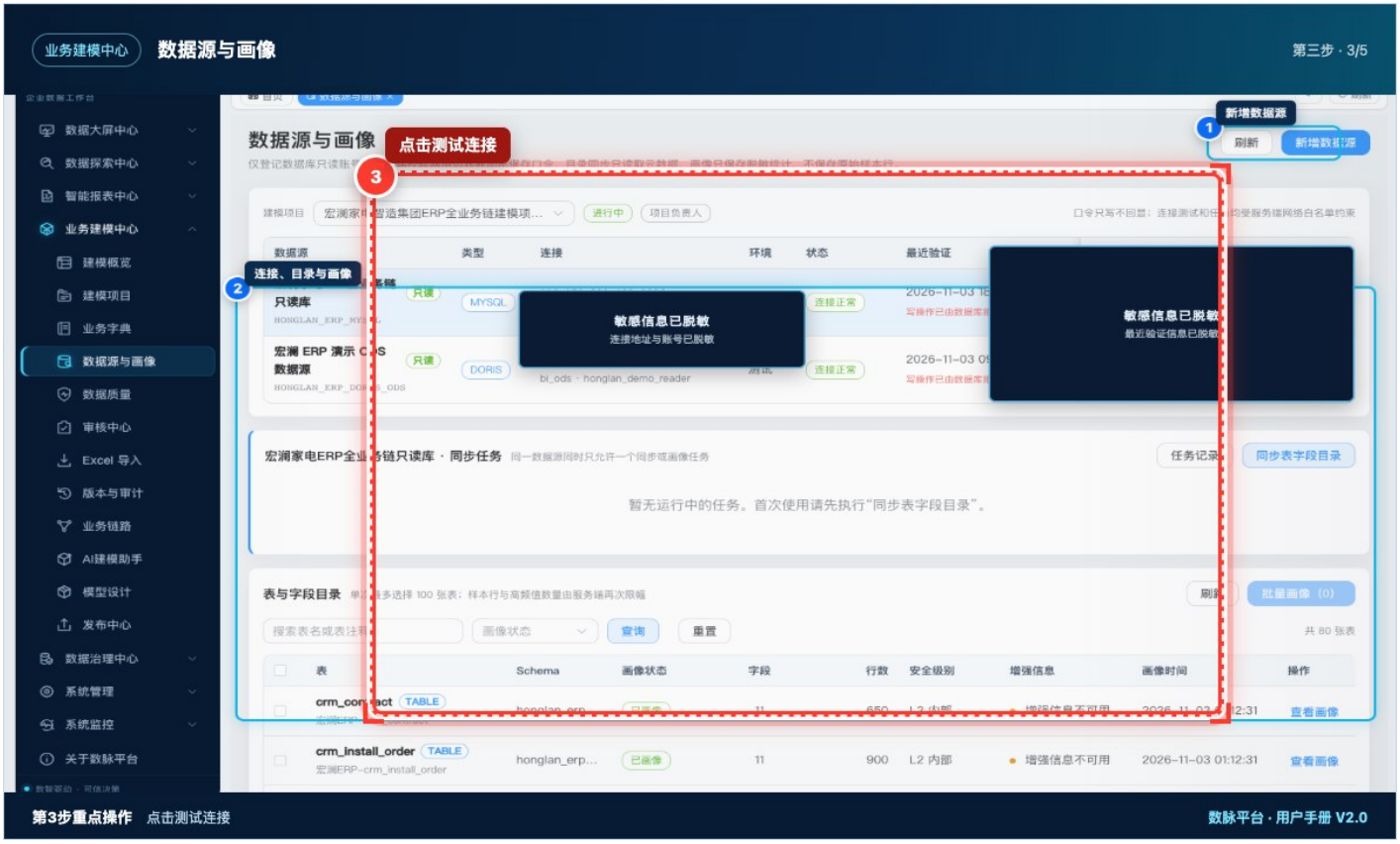
第二步：选择数据库类型



步骤图 26-02 第2步：选择数据库类型。红框标出选择数据库类型。

选择数据库类型，填写主机、端口、库名和只读账号的环境变量引用。

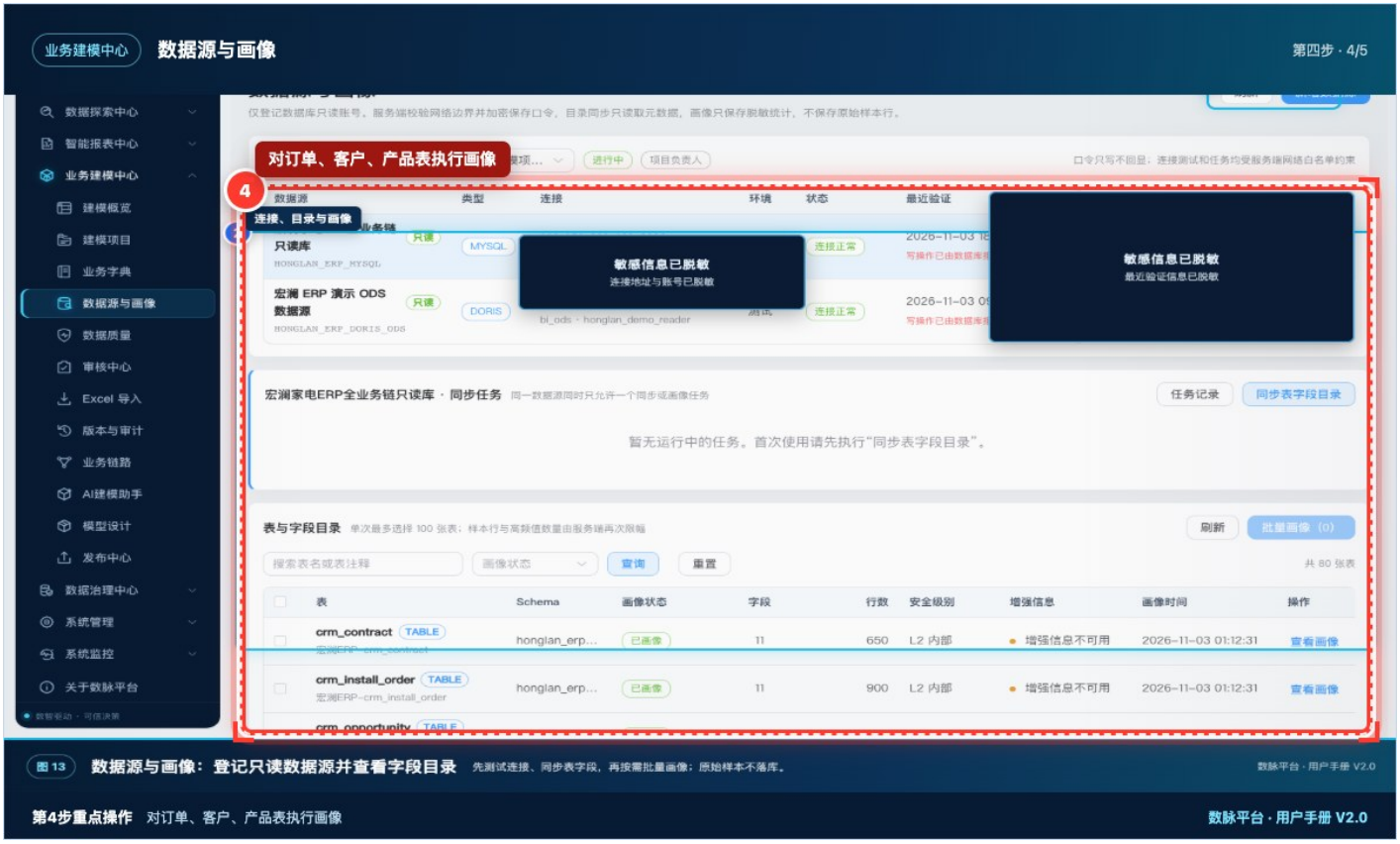
第三步：点击 “测试连接 ”


业务建模中心
数据源与画像
第三步 · 3/5
数据大屏中心
数据探索中心
智能报表中心
业务建模中心
建模概览
建模项目
业务字典
数据源与画像
数据质量
审核中心
Excel 导入
版本与审计
业务链路
AI建模助手
模型设计
发布中心
数据治理中心
系统管理
系统监控
关于数脉平台
数据源与画像
点击测试连接
新增数据源
刷新
新增数据源
连接、目录与画像
敏感信息已脱敏
连接地址与账号已脱敏
敏感信息已脱敏
最近验证信息已脱敏
宏澜家电ERP全业务链只读库 · 同步任务
暂无运行中的任务。首次使用请先执行"同步表字段目录"。
表与字段目录
crm_contract TABLE
crm_install_order TABLE
第3步重点操作 点击测试连接
数脉平台 · 用户手册 V2.0

□□□ *26-03 □ 3*□□□□ “□□□□ ”□□□□□□□□□□□□

□□ “□□□□ ”□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□



步骤图 26-04 第4步：对订单、客户、产品表执行画像。红框标出对订单、客户、产品表执行画像。对订单、客户、产品表执行画像，查看行数、空值、枚举和主键候选。

第五步：保存数据源并记录负责人、用途和访问范围



步骤图 26-05 第5
步：保存数据源并记录负责人、用途和访问范围。红框标出保存数据源并记录负责人、用途和访问范围。

保存数据源并记录负责人、用途和访问范围。

最终结果与效果
数据源登记成功，页面显示表数量、字段目录和画像摘要；密码不会在页面回显。

操作提示： “
必须使用只读账号；若连接失败，先检查网络白名单和密钥引用，不要粘贴明文口令。

开发人员参考
接口： /bi/bm/sources、.../tables、.../profile、.../jobs
权限： bi:bm:source:admin、 bi:bm:profile:run
服务端校验网络边界并加密保存口令；目录同步与画像均为受控异步任务。

版本号 #2
由 Admin 创建于 2026-11-06 00:23:13 UTC

Admin 2026-11-06 03:52:55 UTC