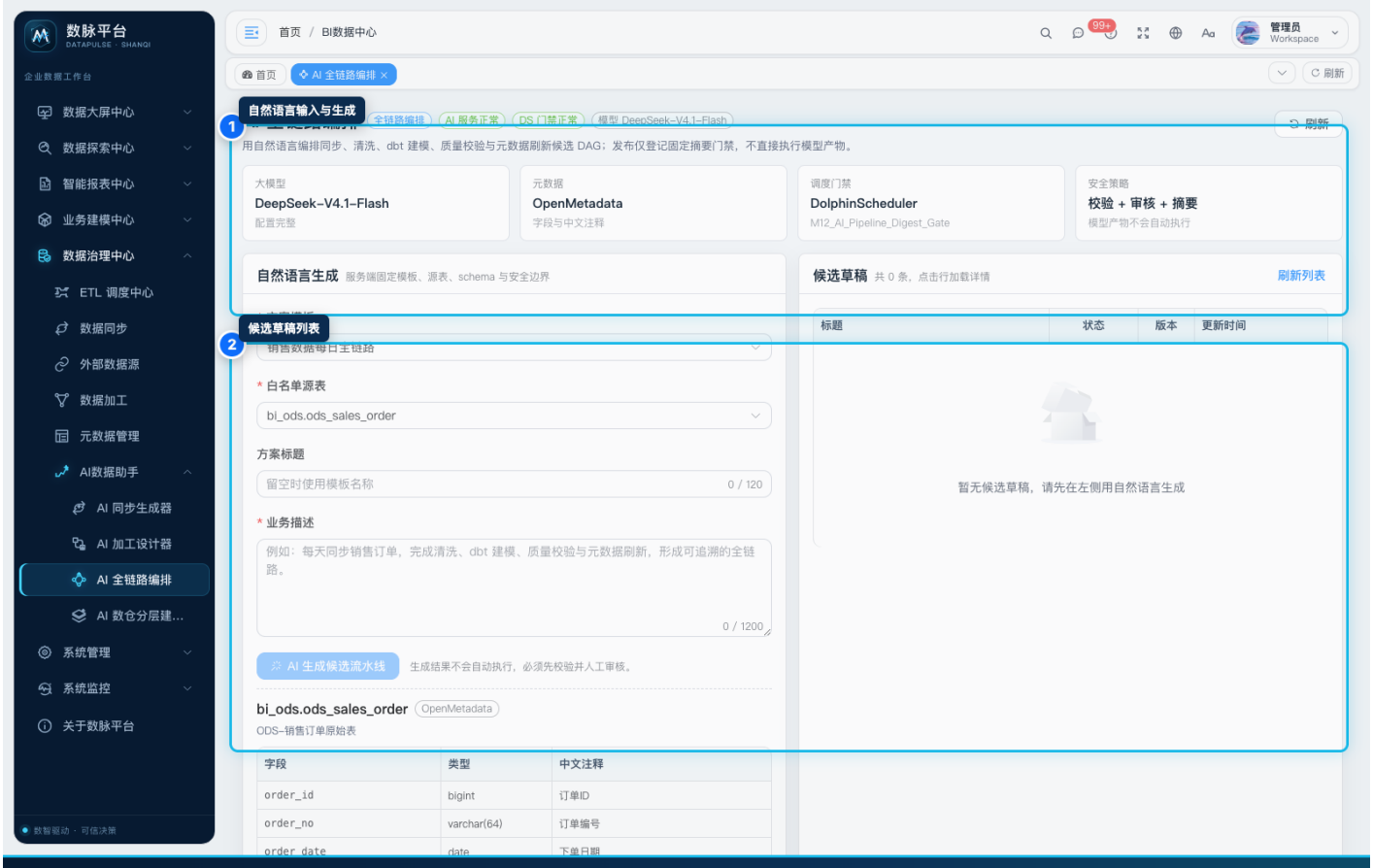
AI □□□□

□□□

AI □□□□□□ “□□□□□ ”□□□□□□□□□□□□□□□□□□□□ dbt □□□□□□□□□□□□□□□ DAG □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□



□ *7-8* □□□□□□□□□□□□□□

□□□

□□□□□□ ETL □□□□□□□□□□□□□□□□□□□□□□

页面入口
数据治理中心 → AI 数据助手 → AI 全链路编排（ /bi/etl/ai/whale）。

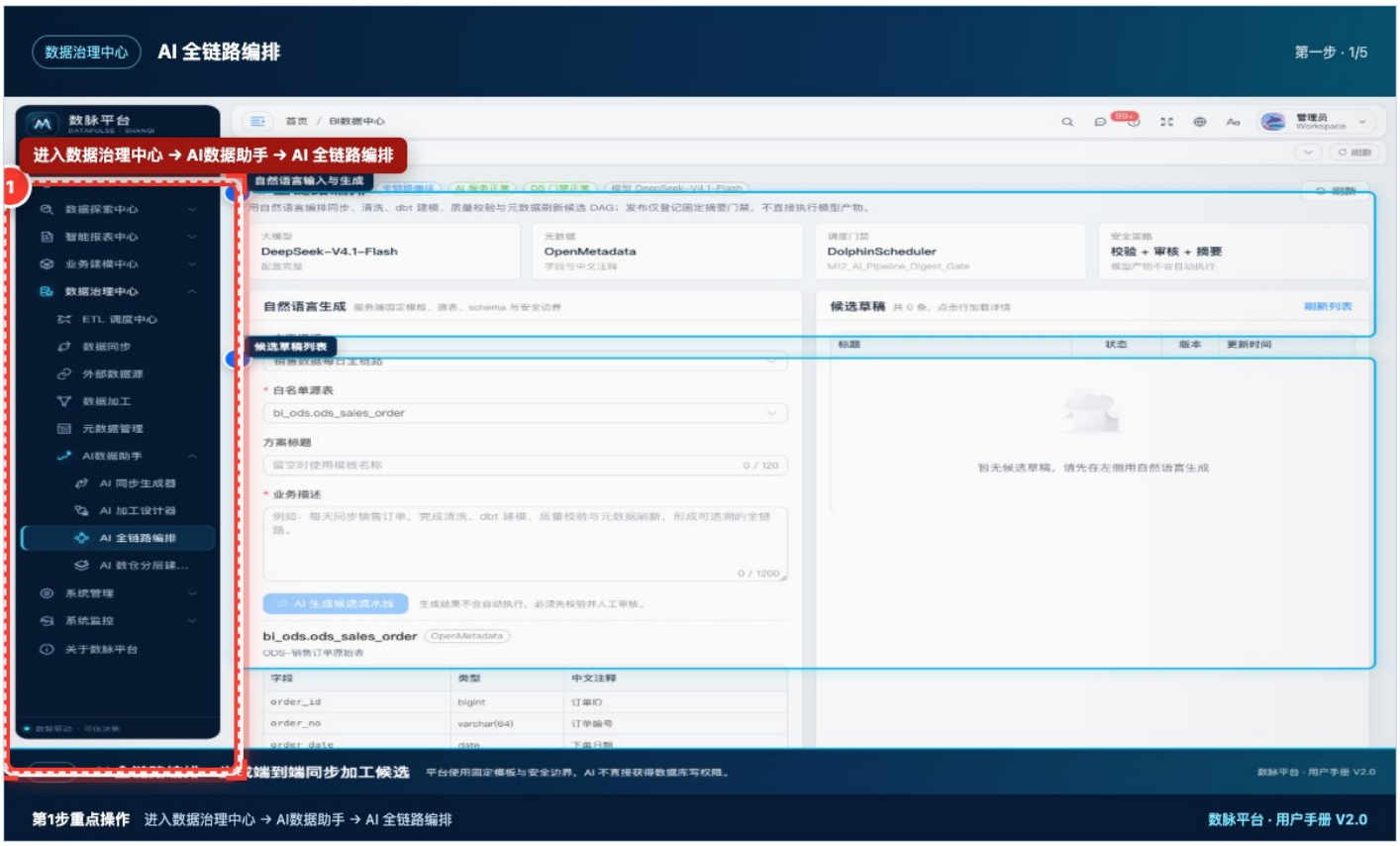
操作步骤
1. 选择全链路模板，例如 “销售数据每日主链路 ”。
2. 选择白名单源表，填写业务描述：同步频率、目标分层、清洗规则、质量要求和刷新范围。
3. 生成候选 DAG，查看节点类型、目标表、依赖边和关键参数。
4. 检查链路是否包含 “源表就绪 → 同步 → DWD 清洗 → DWS 汇总 → 质量测试 → 元数据刷新 ” 的必要阶段。
5. 对每个节点检查失败策略、重试次数、超时和依赖条件，避免一个节点失败后下游仍继续。
6. 执行确定性校验，处理循环依赖、缺失字段、越权表和非白名单组件。
7. 提交审核并登记发布；发布后仍需在调度中心按正常流程执行和监控。

输入与输出
输入是源表、业务描述、分层目标和模板；输出是端到端 DAG
、节点参数、依赖关系、校验结果、审核版本和调度门禁。

宏澜家电示例
输入 “
每天同步销售订单，清洗后按区域品类汇总，执行金额非负和主键唯一测试，刷新元数据，供经营看板使用 ”。
AI 生成五段式 DAG；数据架构师补充 “质量失败阻断下游 ”的条件后审核发布，避免脏数据进入 ADS。

客户应用演示：生成 “

销售数据每日主链路 ”
业务情境： 需要把同步、清洗、汇总、质量和刷新串成一条每日 DAG。

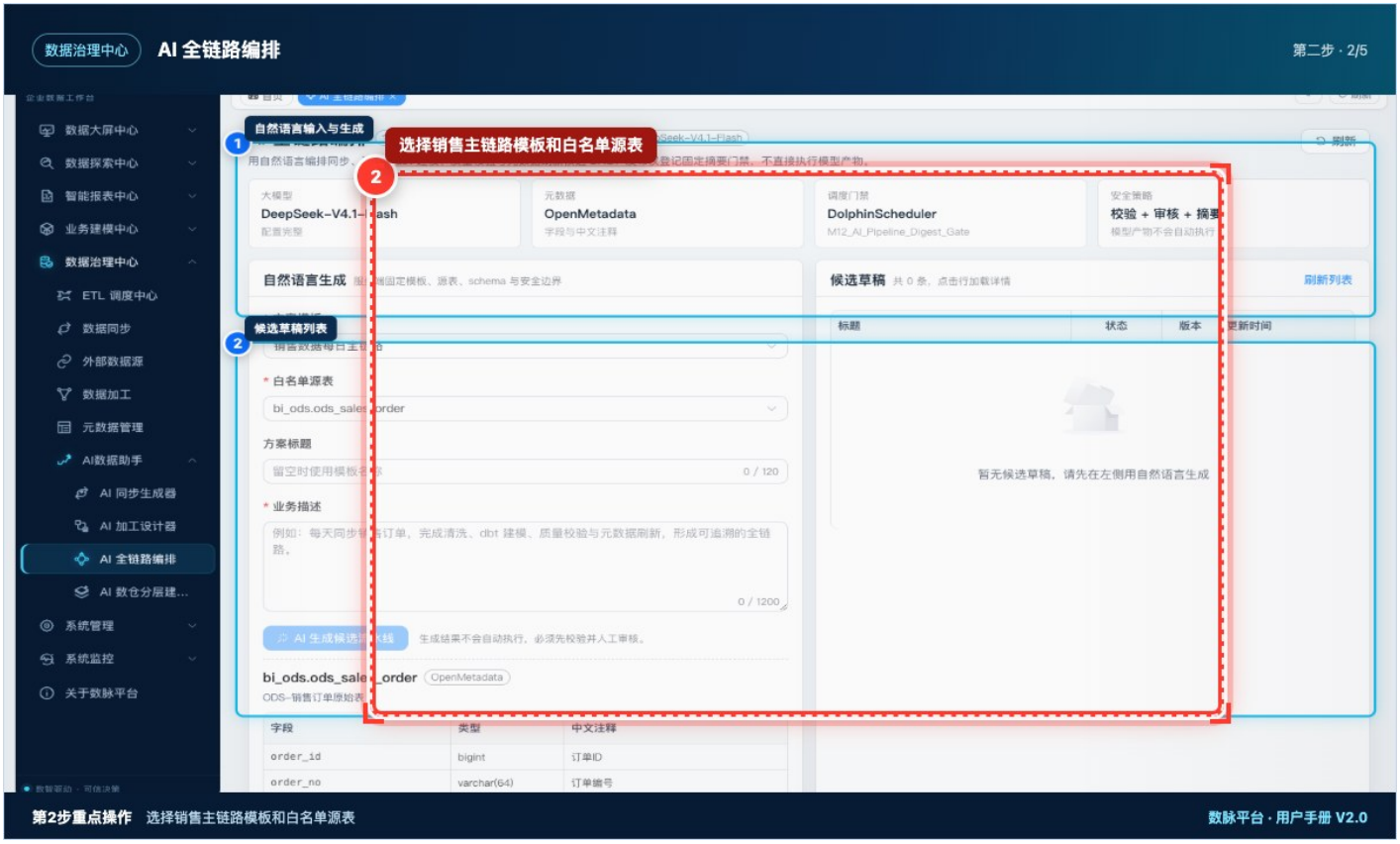


第一步：进入 “数据治理中心 → AI数据助手 → AI 全链路编排 ”

步骤图 *42-01* 第 *1*步：进入 “数据治理中心 → *AI*数据助手 → *AI* 全链路编排 ”。红框标出进入数据治理中心 → *AI* 数据助手 → *AI* 全链路编排。

进入 “数据治理中心 → AI数据助手 → AI 全链路编排 ”。

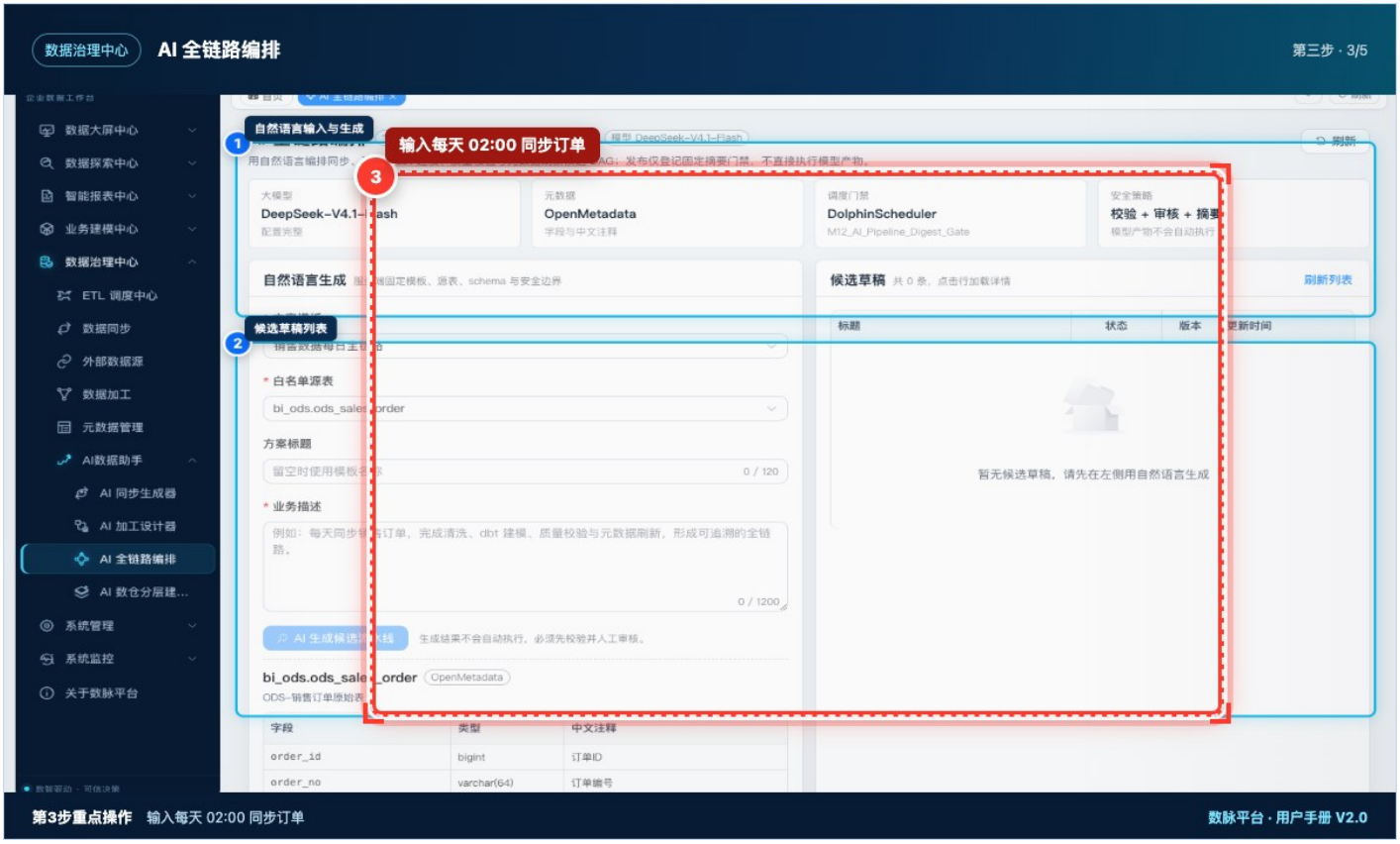
第二步：选择销售主链路模板和白名单源表



步骤图 42-02 第2步：选择销售主链路模板和白名单源表。红框标出选择销售主链路模板和白名单源表。

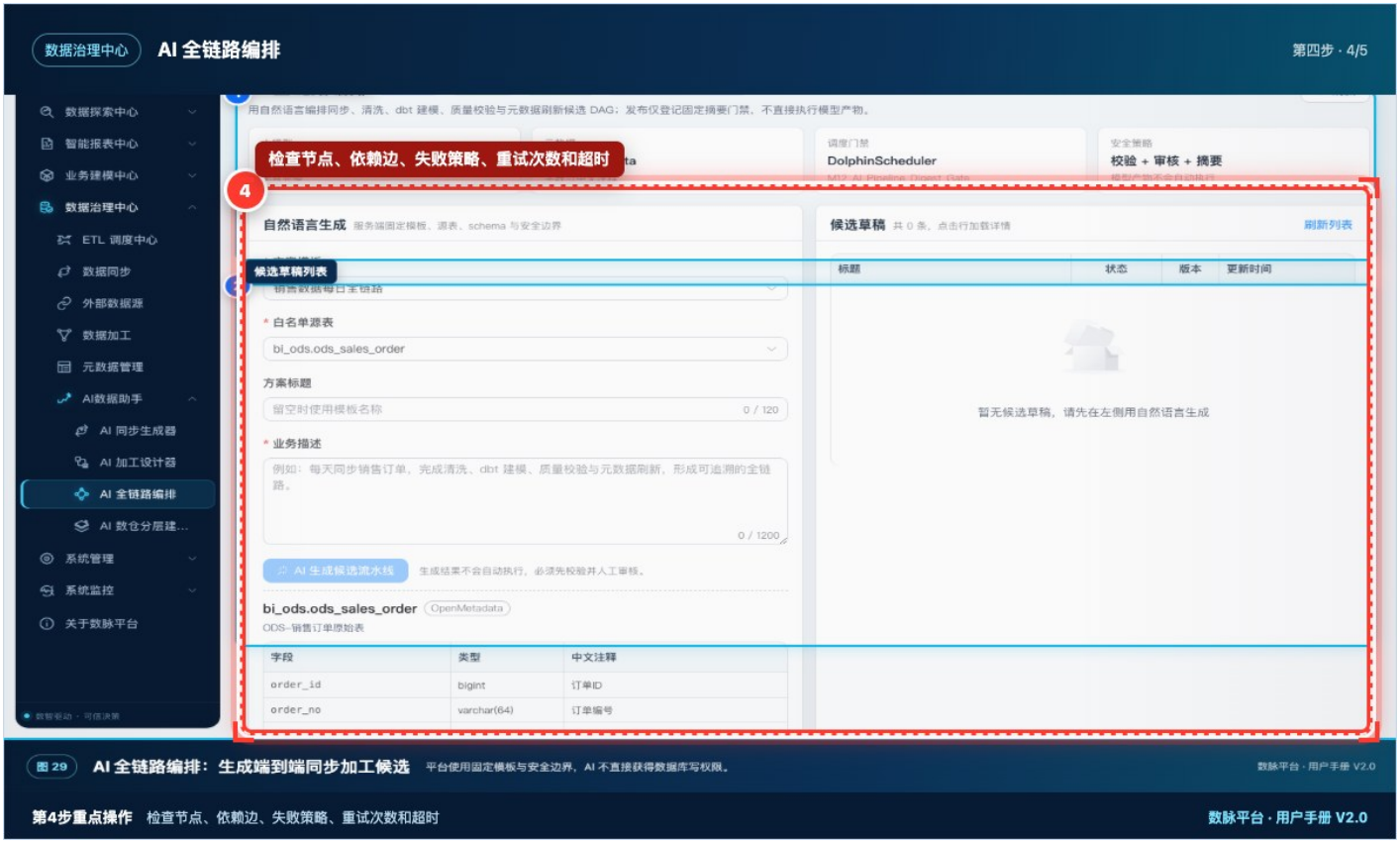
选择销售主链路模板和白名单源表。

第三步：输入 “每天 02:00 同步订单



步骤图 42-03 第3步：输入 “每天 02:00 同步订单。红框标出输入每天 02:00 同步订单。
输入 “每天 02:00 同步订单，清洗后汇总区域品类，执行金额非负和主键唯一测试，再刷新数据集 ”。

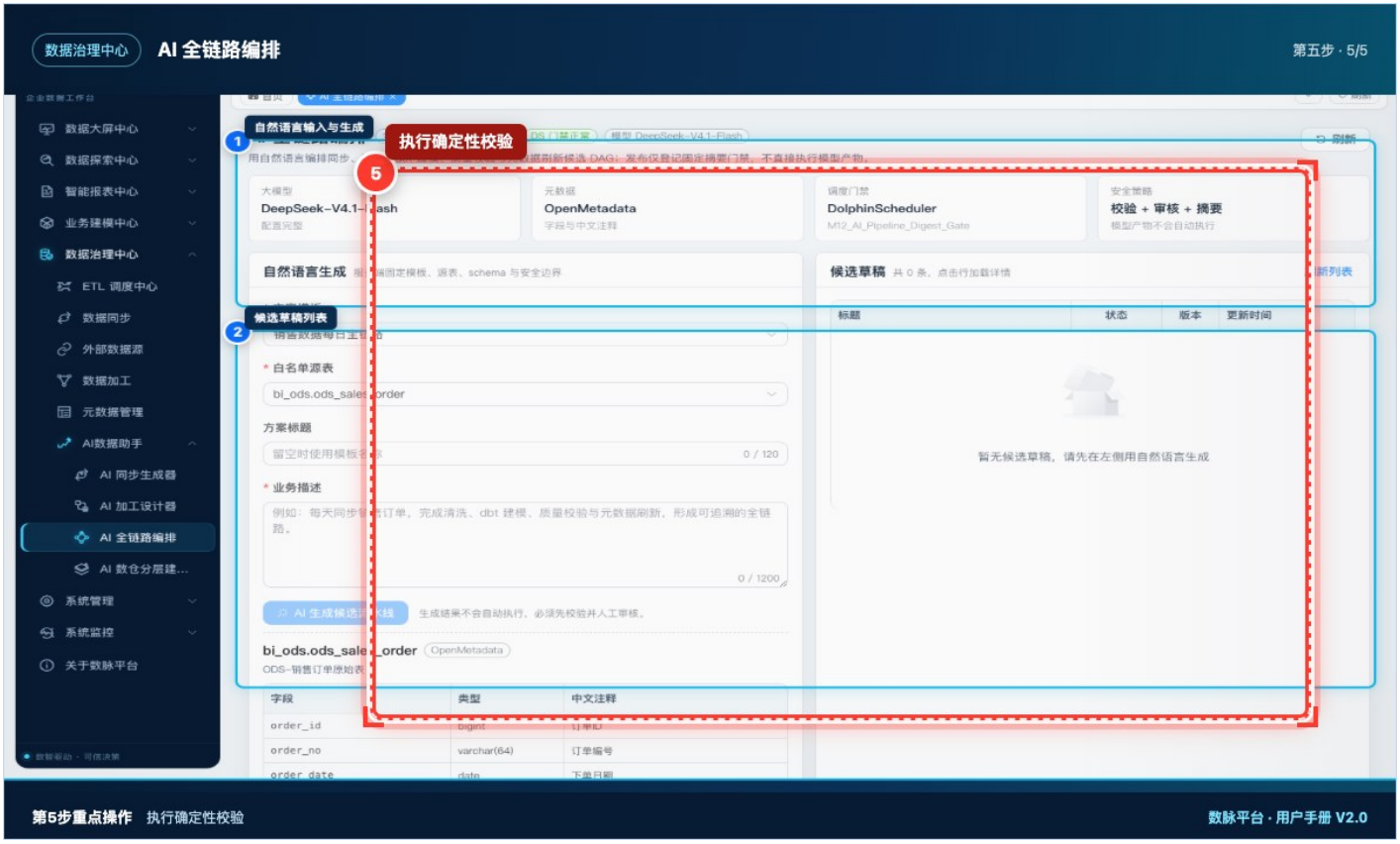
第四步：检查节点、依赖边、失败策略、重试
次数和超时



步骤图 42-04 第4
步：检查节点、依赖边、失败策略、重试次数和超时。红框标出检查节点、依赖边、失败策略、重试次数和超时。

检查节点、依赖边、失败策略、重试次数和超时。

第五步：执行确定性校验



步骤图 42-05 第5步：执行确定性校验。红框标出执行确定性校验。
执行确定性校验，补充 “质量失败阻断下游 ”，提交审核发布。

最终结果与效果
生成端到端 DAG 草稿，包含同步、加工、汇总、质量和刷新阶段；发布只登记调度门禁，不自动运行。

操作提示： 确保链条有源表就绪、失败阻断和批次追踪，避免半成品进入看板。 “

常见问题
DAG 看起来完整但不能发布。 先看校验结果，通常是字段越界、依赖循环或缺少目标表。
全链路是否会自动执行。 不会。发布只登记调度门禁和摘要，实际执行由 DolphinScheduler 负责。某个节点失败如何补数。 从失败节点开始按依赖顺序补跑，避免只补最后一层造成口径断层。

开发人员参考
接口： /bi/ai-pipeline/options?kind=WHALE、 /generate、 /drafts/{id}/validate|review|publish、 /versions/{id}/rollback
权限： bi:etl:ai-whale
服务端对 DAG 做节点白名单、依赖环检测、字段边界和摘要校验。
发布不直接调用执行器，避免 AI 生成内容未经人工确认就改变生产数据。

版本号 #2
由 Admin 创建于 2026-11-06 00:23:24 UTC
由 Admin 更新于 2026-11-06 03:53:14 UTC