



DATABLAU · DAM

元数据管理

产品模块用户手册

版本 9.1.2

目录

- 01 元数据管理
 - 02 快速开始
 - 03 从元数据对象查看基础血缘
 - 04 配置元数据采集
 - 05 查询和维护表、字段元数据
 - 06 元数据对象维护
 - 07 采集自定义元模型数据
 - 08 认识元模型与预置模型
 - 09 设计自定义元模型
 - 10 导入和验证自定义元模型数据
 - 11 模型变更与元数据复核
 - 12 管理报表与文件元数据
 - 13 配置数据源采集账号权限
 - 14 将字段关联到数据标准
-

元数据管理

元数据管理用于登记、采集、组织和查询数据库对象、报表、文件、API 及企业自定义对象等元数据，并提供从元数据对象出发的基础血缘查看。

“元数据”是描述数据资源本身的信息，例如表和字段的名称、结构、来源、含义和责任归属；采集或编辑这些描述不等于读取、修改表中的业务记录。需要管理企业特有对象的结构时，再阅读[元模型](#)。

当前模块按三个工作区组织入口：

工作区	页面	主要用途
元数据	元数据概览、数据集、报表、文件	查询已经登记或采集的对象
采集管理	应用系统、数据源、采集管理、元模型	准备来源并运行采集
血缘管理	血缘空间、预处理管理、命名空间管理	管理元数据侧的基础血缘关联

元数据概览、采集管理和各查询入口解决的问题不同：先准备来源与插件，再选择采集类型并运行任务，最后按对象类型到“数据集”“报表”或“文件”查看结果。采集时选择了“文件”，不代表结果一定出现在“文件”栏目；数据字典、DDL 和 API 元数据会被解析成结构化对象。

先判断采集对象和查询位置

要接入的内容	采集入口	采集结果主要在哪里查询
关系型数据库、非关系型数据库或数据库平台	数据源	“数据集”：数据源、表、字段、视图、存储过程、函数等
FineBI、FineReport、SmartBI 等受支持的 BI 报表	报表	“报表”
DAM 新建或导入、SQLink“报表管理”导入的自定义报表	报表模块手工维护	“报表”
DDM 导出的物理 / 逻辑数据字典、DDL 文件	文件	解析后进入“数据集”，原始文件不作为“文件”栏目中的资产展示
文件存储服务器中的文件	文件	“文件”
API 元数据或其他按元模型定义的对象	文件、数据源或元模型对应方案	结构化对象进入“数据集”

如果不确定入口，先阅读[配置元数据采集](#)中的采集决策表，再执行任务。

数据模型的设计、比对和同步请参考 DDM 产品用户手册。模型发生变更后，应在 DAM 中重新采集并复核元数据、标准映射和血缘关系。

阅读路线

第一次使用先完成[快速开始](#)。日常任务按下表进入。

目标	操作指南
准备系统与连接	应用系统与数据源
接入数据库、报表、文件或元模型对象	配置采集
为数据库采集准备最小权限账号	数据源采集账号权限
查找、编辑、导出和批量更新对象	表与字段元数据
维护对象、报表、文件和模型变更后的元数据	元数据对象维护 与 模型变更复核
统一字段含义和值域	关联数据标准
管理报表与文件对象	报表与文件
从对象查看上下游	基础血缘
扩展企业特有资产结构	设计元模型
理解预置模型和自定义模型	元模型概念
按自定义结构录入对象	采集自定义元模型数据
检查模板导入与血缘同步	元模型导入

查看元数据概览

概览按系统、采集、标准、规则和报表等维度展示统计，可在编辑模式调整组件与布局。统计依赖后台计算任务，不应以刚保存后的概览数值替代采集结果核对。

常用口径包括：接入率为已有数据源的应用系统数占全部应用系统数；标准覆盖率为已关联标准字段数占全部字段数；标准核标率为完全映射字段数占已关联标准字段数。应用系统下的数据源统计只计入已采集元数据的来源。

与 SQLink 的边界

元数据管理中的血缘用于从元数据对象查看上下游关系。“血缘管理”License 不等同于 SQLINK License；血缘地图、应用场景、专业解析和影响分析等完整能力参见[数据血缘](#)。

元数据管理快速开始

以“采集一个数据源、核对一张表、为一个字段关联标准”为起点，完成从技术结构到业务含义的第一次整理。开始前准备应用系统、数据源账号及可访问的标准。

1. 准备来源

按[应用系统](#)与[数据源](#)建立来源并测试连接。初次选择少量已知表，便于人工核对。

2. 配置并执行采集

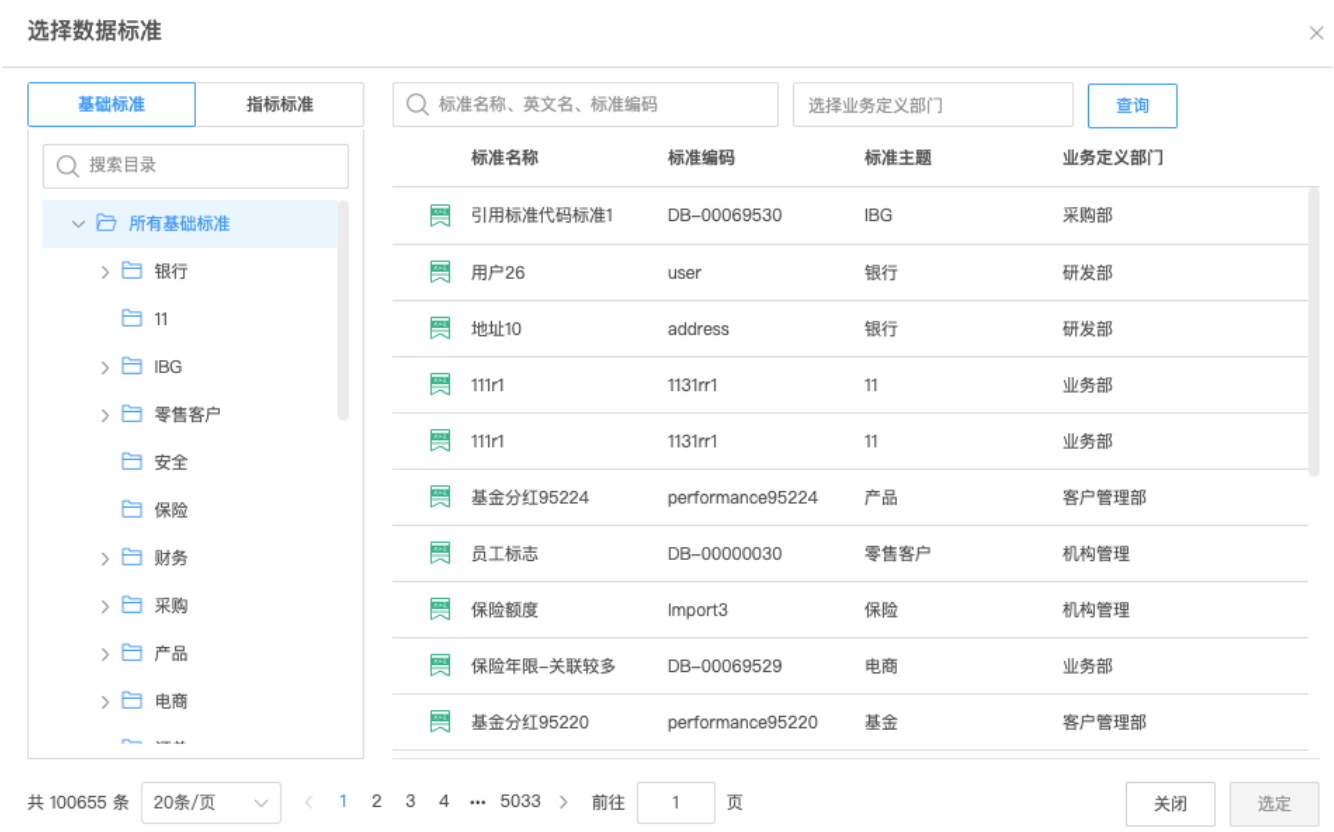
按[配置采集](#)选择数据源和 Schema，检查采集范围并创建任务。查看执行历史确认本次结果；保存采集配置不等于采集已完成。

3. 检查表与字段

进入“数据集”，按来源和名称查找目标表。按[元数据维护](#)核对字段、类型、长度及业务属性，补充中文名和描述。

4. 关联标准

打开一个业务含义明确的字段，按[关联标准](#)选择适用标准及代码。回到表详情检查关系，再从标准引用信息反向确认。



图示展示选择标准窗口。先核对候选名称、编码和业务含义，再选定；图中对象仅用于说明界面。单击图片可查看原图。

5. 查看关系与安排更新

若已准备血缘解析结果，从表或字段查看[基础血缘](#)。若尚无关系，继续准备血缘来源，不以空图判断没有上下游。

确认首次采集范围和属性正确后，再设置周期调度。

完成标志

- 目标来源下能查到正确的表和字段。
- 结构与源系统一致，业务描述能解释对象用途。
- 字段引用的标准语义准确，关联可双向核对。
- 执行记录可追溯，后续采集周期符合要求。

下一步可按[数据标准](#)管理统一口径，或按[数据质量](#)验证实际数据。

从元数据对象查看基础血缘

元数据血缘用于从表、字段或报表出发了解来源与去向。能够显示哪些关系，取决于元数据采集、血缘文件解析和对象匹配结果，不能由表名或字段名自动推定。

查看操作

1. 在数据集或报表中定位目标对象。
2. 打开详情中的血缘关系或血缘分析区域。
3. 核对当前对象所属来源及名称，再查看上下游节点和连线。
4. 按业务问题沿相关路径查看：追溯来源时关注上游，评估变更时关注下游。
5. 对关键节点回到元数据详情，核对物理对象及业务含义。

字段血缘与表级血缘的粒度不同；只看到表间关系时，不能据此声称已确认某字段的计算来源。

没有关系时

先检查目标对象是否已采集、相关血缘文件是否解析成功、命名空间和 Schema 是否匹配，再检查资源同步情况。图为空可能是数据准备不足，不代表没有实际业务依赖。

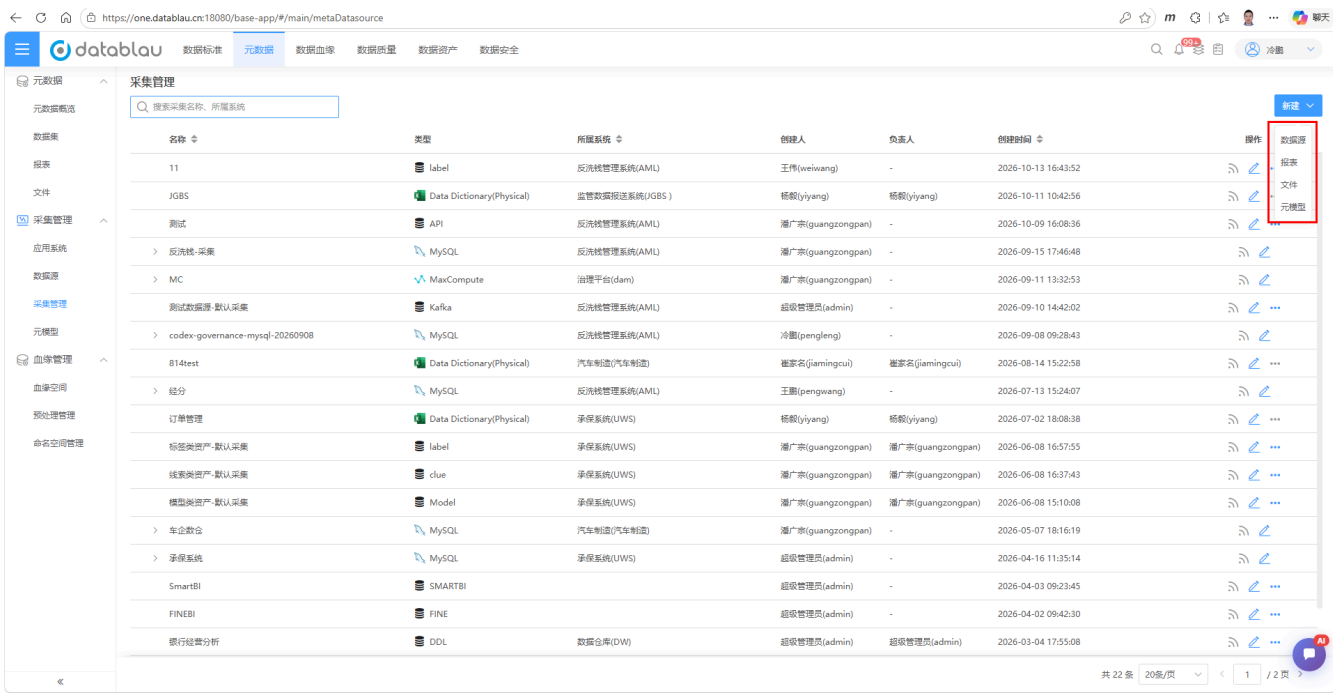
与 SQLink 的区别

元数据中的血缘管理及从元数据查看的全链路血缘，属于元数据血缘范围。独立“数据血缘”栏目中的血缘地图、应用场景等属于 SQLink，需要对应能力。

血缘采集、空间、命名空间与预处理涉及管理员维护，参见[血缘系统管理](#)。本页不将 SQLink 专业分析入口描述为元数据默认功能。

配置元数据采集

采集管理定义从什么来源、以什么方式、以什么范围获取元数据。当前新建采集分为“数据源”“报表”“文件”“元模型”四大类。建立连接、安装插件、配置采集和执行任务是不同步骤；仅保存连接或采集配置不会产生可查询的元数据。



截图中的名称、人员和数量来自核验环境，仅用于说明入口位置。

选择采集路径

先判断最终要管理的对象，再选择采集入口。不要只根据材料是不是一个文件来判断结果位置。

实际来源或目标对象	采集入口与方式	主要前置条件	采集后检查位置
关系型数据库	数据源	应用系统、连接、驱动、数据源插件、逆向插件、数据库与 Schema 范围	“数据集”中的表、字段、视图、存储过程、函数等
HBase、MongoDB 等非关系型数据库或数据库平台	数据源	与该类型匹配的连接、驱动及插件；可采集对象以插件能力为准	“数据集”中的对应技术对象
自定义数据源	数据源	已发布的元模型、匹配的数据源插件和逆向插件、模板或适配程序	“数据集”中的元模型对象及其层级关系
FineBI、FineReport、SmartBI 等 BI 工具	报表	产品支持的报表类型、服务地址、认证信息和报表目录	“报表”中的目录、报表、信息项及关联对象
自定义报表	DAM“报表”新建 / 导入，或 SQLink“报表管理”导入	报表模板、对象编码、所属系统及关联信息	统一在 DAM“报表”查询
DDM 物理数据字典	文件 → Data Dictionary (Physical)	DDM 导出的物理数据字典及正确的模板结构	解析后的表、字段等进入“数据集”
DDM 逻辑数据字典	文件 → Data Dictionary (Logical)	DDM 导出的逻辑数据字典及正确的模板结构	解析后的实体、属性等进入“数据集”
DDL 文件	文件 → “解析 DDL”	与脚本方言匹配的解析类型、命名空间和文件	解析后的表、字段、视图、存储过程、函数等进入“数据集”
API 元数据	文件 → “API 元数据”	内置 REST API 元模型；JSON 文件或 HTTP API 访问参数	API 服务、接口、请求 / 响应参数等进入“数据集”
文件服务器中的文件资产	文件 → “文件存储服务”	服务器访问方式、路径、认证和采集范围	仅此类文件主要在“文件”栏目查询和下载
其他企业自定义对象	元模型，或文件 → “元模型文件”	先完成元模型设计、发布，再按对应模板或插件采集	“数据集”中的元模型实例及父子、关联关系

同一份文件可能只是采集载体，也可能本身就是要管理的资产。数据字典、DDL、API JSON 和元模型模板用于生成结构化元数据；只有通过“文件存储服务”采集的文件，才按文件资产进入“文件”栏目。

检查插件前置条件

插件管理虽然位于“系统管理 → 数据源管理”，却直接决定采集类型是否可选、连接是否可建立以及对象能否被识别。创建采集前由管理员检查：

插件类型	作用	缺失或不匹配时的典型表现
数据源插件	提供数据库或自定义来源的类型、连接及相关接入能力	新建数据源时没有目标类型，或无法建立连接
逆向插件	把来源结构转换成 DAM 可管理的元数据对象	连接可用但不能正确识别表、字段或自定义对象
血缘解析插件	解析 SQL、ETL、BI 或其他文件中的加工关系	元数据对象已采集，但血缘文件无法解析或关系不完整

三类插件用途不同，不能用数据库驱动替代插件，也不能用血缘解析插件替代元数据逆向采集。上传、覆盖更新和验证方法见[驱动与插件](#)。

准备连接和采集账号

先维护[应用系统与数据源](#)。确认所需数据库类型、驱动和插件已经配置，再按[数据源采集账号权限](#)准备专用账号。连接测试通过只说明账号能够登录；还需要用少量已知对象验证系统目录、Schema 和对象定义是否能够完整逆向。

采集数据源元数据

“数据源”不仅指传统关系型数据库，也可包括 HBase、MongoDB 等非关系型数据库，以及主要面向元模型的自定义来源。页面可选择类型由已安装插件决定，不同类型能够采集的对象也不同。

1. 进入“元数据管理 → 采集管理”，新建数据源采集。
2. 填写采集管理名称，按应用系统和数据源筛选并选择来源。
3. 单击“测试”检查连通性，通过后进入下一步。
4. 按需求选择数据库、Schema 及采集明细，检查连接、采样、采集设置和白名单。
5. 确定后，系统按所选 Schema 建立模型及默认采集任务。
6. 按业务要求设置采集任务的执行周期，查看执行记录。
7. 进入“数据集”，选择该来源，核对数据源、表、字段、存储过程、函数或元模型对象。

首次采集建议限定少量已知对象。对照源系统核对象数量、名称、数据类型、长度和关系，再扩大范围。这是实施建议，不是产品对象数量限制。

配置项	操作前的检查
采集管理名称	应能区分用途和来源，避免多个任务只使用相同默认名称
应用系统与数据源	核对实际所属系统及环境，连接测试通过后再设置范围
数据库、Schema	勾选本次需要的对象范围，不默认选择全部
连接、采样与采集设置	按实际需要配置；能采集结构并不表示允许读取业务数据
白名单	核对需要的表是否被范围设置排除，漏采时与任务日志一起检查
执行周期	先核对首次结果，再安排更新频率，并检查任务实际触发记录

采集报表元数据

标准报表采集适用于页面已支持的 BI 工具，例如 FineBI、FineReport 和 SmartBI。在采集管理中新建“报表”，选择工具类型，填写对应服务连接信息及报表目录，测试后确定。不同工具的连接字段不同，应使用所选类型的表单，不将一种工具的 URI、认证方式套用于另一种工具。

无法通过标准连接采集的报表属于自定义报表，可在 DAM“报表”中手工新建或按模板导入，也可通过 SQLink“报表管理”导入。标准采集与自定义导入的报表最终都在 DAM“报表”栏目统一查询；手工填写报表属性不会自动产生完整的库表血缘。

采集后到[报表管理](#)核对报表目录、属性、信息项及关联库表。连接成功不等于所有报表都有完整库表关系。

采集文件类来源

“文件”采集包含多种用途。新建采集后，应根据材料和目标对象选择具体类型。

←

↺

🏠

https://one.datablau.cn:18080/base-app/#/main/metaDatasource

☰

datablau

数据标准

元数据

数据血缘

数据质量

数据资产

数据安全

🔍 元数据

元数据概览

数据集

报表

文件

采集管理

应用系统

数据源

采集管理

元模型

血缘管理

血缘空间

预处理管理

命名空间管理

↶ 返回 | 采集管理 / 添加采集

* 采集管理名称

请输入采集管理名称

* 所属系统

请选择

* 数据库

文件存储服务器

* 命名空间

Data Dictionary(Physical)

* 文件路径

Data Dictionary(Logical)

描述

元模型文件

文件存储服务器

解析DDL

API元数据

🔔

☐ 使用用户名密码

测试

物理和逻辑数据字典

Data Dictionary (Physical) 和 Data Dictionary (Logical) 用于接收 DDM 导出的数据字典。物理数据字典通常还原数据库表、字段等技术结构；逻辑数据字典用于接收逻辑实体、属性等模型内容。上传前保持模板列、对象标识和父子层级不变。完成后在“数据集”查询解析出的对象，不在“文件”栏目查找 Excel 原文件。

DDL 文件

选择“解析 DDL”，并使用与 DDL 方言匹配的数据库或解析类型。系统从建表、视图、存储过程和函数等语句中提取结构对象。没有数据库直连权限时可以使用此方式，但它只反映上传脚本中的定义；脚本是否完整、是否与生产环境一致，需要由提供方确认。结果进入“数据集”。

API 元数据

API 元数据建立在 REST API 元模型上，可通过 JSON 文件或 HTTP API 访问方式采集。采集前核对 API 服务、接口、请求参数、响应参数等对象与内置元模型的映射。结果作为元模型对象进入“数据集”，不是普通文件下载项。

文件存储服务器

选择“文件存储服务器”时，平台采集的是服务器目录中的文件对象。配置服务器访问、路径和范围后执行任务，结果在“文件”栏目查询并按权限查看或下载。这是文件采集类型中与“文件”查询栏目直接对应的路径。

采集元模型对象

元模型先定义“管理哪些对象、对象有什么属性、对象之间如何关联”，采集再把实际对象写入这个结构。元模型适用于 API、消息队列、对象存储、企业专有业务对象等预置数据库模型无法表达的内容，不是数据库采集的通用别名。

采集前先完成[元模型概念](#)和[元模型设计](#)中的设计与发布。根据交付方案选择自定义数据源插件、DAM 模板导入，或“文件 → 元模型文件”路径；不同路径的根对象、对象标识、命名空间和模板规则不能混用。详细操作见[采集自定义元模型数据](#)和[元模型导入](#)。

运行与结果检查

- 在任务执行记录检查本次状态及日志。
- 按本页决策表进入“数据集”“报表”或“文件”，不要只在一个栏目检查所有结果。
- 对数据源、数据字典、DDL、API 和元模型采集，在“数据集”检查来源、对象类型、Schema / 命名空间与数量。
- 对报表采集或导入，在“报表”检查目录、报表、信息项及库表关联。
- 对文件存储服务器采集，在“文件”检查文件名、路径、来源和可下载性。
- 随机选取表、字段和视图，与源系统比较关键属性。
- 查看是否存在漏采、采集失败或白名单排除的对象。
- 需要周期采集时，再配置[任务调度](#)。

常见问题

现象	排查顺序
测试连接失败	网络、连接参数、凭据、登录权限和驱动
测试成功但没有数据	是否执行采集；系统目录权限、Schema、白名单、采集范围及执行日志
采集对象少于预期	源账号可见范围、对象定义权限、筛选条件和失败对象
报表没有关联库表	检查来源是否提供关联信息、是否已采集相关数据库元数据
文件采集成功但“文件”栏目没有记录	先确认采集类型；数据字典、DDL、API 和元模型文件的结果应到“数据集”查询
新建采集时没有目标数据库类型	检查对应数据源插件、逆向插件和驱动是否已安装并适配当前版本
元模型对象层级错误或重复	检查模型是否已发布、模板中的对象标识与父子关系，以及是否混用了不同采集路径的规则
定时任务没有运行	检查启用状态、周期和时间模板交集

查询和维护表、字段元数据

“数据集”集中展示采集后形成的结构化元数据。它不仅包含关系型和非关系型数据源中的对象，也包含由数据字典、DDL、API 文件及自定义元模型生成的对象。业务人员可查看定义，数据管家可维护业务属性，技术人员可检查结构和变更。

哪些对象在数据集查询

来源	数据集中常见的对象
关系型数据库	数据源、数据库、Schema、表、字段、视图、存储过程、函数、索引及约束等
HBase、MongoDB 等非关系型来源	由对应插件识别的库、集合、字段或其他技术对象
DDM 物理 / 逻辑数据字典、DDL 文件	解析后形成的表、字段、实体、属性、视图、存储过程、函数等
API 元数据	API 服务、接口、请求参数、响应参数及其关系
自定义元模型	元模型定义的对象、属性及父子或关联关系

标准采集和手工导入的报表到“报表”查询；只有通过文件存储服务器采集的文件到“文件”查询。详情见[配置元数据采集](#)。

当前数据集页面提供“按技术部门”视图、全部元数据与全部数据源筛选，列表可按英文名称、中文名称、资产类型、业务系统、Schema、技术部门和创建时间阅读对象。对象列表中的操作入口用于打开详情；页底显示当前查询结果数量并支持分页。

The screenshot displays the Datablau metadata management interface. The left sidebar contains navigation menus for '元数据' (Metadata), '数据血缘' (Data Lineage), '数据质量' (Data Quality), '数据资产' (Data Assets), and '数据安全' (Data Security). The '元数据' menu is expanded, showing sub-menus like '元数据概览', '数据集' (Data Sets), '报表', '文件', '采集管理', '应用系统', '数据源', '采集管理', '元模型', '血缘管理', '血缘空间', '预处理管理', and '命名空间管理'. The '数据集' menu is selected, and the '全部数据集' (All Data Sets) option is chosen. The main area shows a list of data objects with columns: '英文名称' (English Name), '中文名称' (Chinese Name), '资产类型' (Asset Type), '业务系统' (Business System), 'Schema', '技术部门' (Technical Department), and '创建时间' (Creation Time). The list includes objects like DIM_PRODUCTS, FACT_SALES_CUSTOME, DIM_EMPLOYEES, FACT_SALE_CONTRACT, FACT_SALES, FACT_SALES_MONTH, DIM_CUSTOMERS, DIM_DATE, KCAB_LRG_AMT_APTM, ODS_P_CUST_TEL, and DIM_PRODUCTS. The bottom right shows pagination: '共 1737 条' (Total 1737 items), '20条/页' (20 items per page), and a page number '1'.

英文名称	中文名称	资产类型	业务系统	Schema	技术部门	创建时间
DIM_PRODUCTS	产品维表	表	数据湖 (ODS)	d3_dw	信息管理...	2025-06-30 17:19:18
FACT_SALES_CUSTOME	大宽表	表	数据湖 (ODS)	d3_dm	信息管理...	2025-06-30 17:19:18
DIM_EMPLOYEES	员工维表中文名	表	数据湖 (ODS)	d3_dw	信息管理...	2025-06-30 17:19:18
FACT_SALE_CONTRACT	销售合同事实表	表	数据湖 (ODS)	d3_dm	信息管理...	2025-06-30 17:19:18
FACT_SALES	销售事实表	表	数据湖 (ODS)	d3_dw	信息管理...	2025-06-30 17:19:18
FACT_SALES_MONTH	销售月度汇总表	表	数据湖 (ODS)	d3_dm	信息管理...	2025-06-30 17:19:18
DIM_CUSTOMERS	客户维表	表	数据湖 (ODS)	d3_dw	信息管理...	2025-06-30 17:19:18
DIM_DATE	时间维表	表	数据湖 (ODS)	d3_dw	信息管理...	2025-06-30 17:19:18
KCAB_LRG_AMT_APTM	大额现金预约表	表	数据湖 (ODS)	AML	信息管理...	2025-06-30 17:38:04
ODS_P_CUST_TEL	客户电话关系表	表	数据湖 (ODS)	AML	信息管理...	2025-06-30 17:38:04
DIM_PRODUCTS	产品维表	表	数据仓库 (DW)	d3_dw	信息管理...	2025-06-30 18:19:12

[查看数据集列表原图](#)。截图中的对象名称和统计数量来自核验环境，仅用于说明页面字段和定位方法。

查找对象

1. 进入“元数据管理 → 数据集”，选择应用系统或数据源。
2. 在搜索框输入表、字段、视图、函数、存储过程、API 或元模型对象名称。
3. 使用类型筛选及页面提供的其他条件缩小结果。
4. 打开对象，核对来源、Schema、物理名称，避免选中不同系统中的同名对象。

如果对象较多，先选择技术部门或数据源，再使用列表搜索；最后通过业务系统、Schema 和物理名称三项交叉确认。分页数量和总条数是当前筛选结果，不代表平台全部元数据。

查看表与字段

表详情提供字段信息、采样数据及血缘关系等内容。打开字段名称进入字段详情；业务描述与技术属性应结合查看，不要用中文名代替物理字段标识。

需要补充属性时，使用对应区域的编辑入口，保存后重新查看。维护前明确哪些属性来自源系统、哪些由治理人员补充；后续采集产生变化时，应结合变更历史检查。

查看视图、存储过程和函数的 SQL 语句受角色权限控制。对象可见并不表示 SQL 正文一定可见。

查看采样数据

在表或视图详情中使用采样功能。采样读取实际数据，需具备相应权限。采样失败与没有数据是两种结果：出现数据库错误时先检查错误内容、查询权限和结构读取；“暂无数据”不能单凭字面判断为连接故障。

采样用于理解字段，不等于全量数据质量检查，也不代表数据市场已完成数据交付。

Hive 采样先读取表结构，再读取指定行数的样本。结构读取失败时，应检查“采样失败”提示及数据库报错；样本查询成功但返回空结果时，页面显示“暂无数据”。排查时分别检查结构读取权限和样本查询结果，不把两种情况都记录成“表中没有数据”。

比较变更历史

在“变更历史”选择两个历史记录进行比较，检查表属性、字段及索引差异。记录新增、删除和属性变化，再结合[基础血缘](#)判断相关上下游对象。

这里的历史记录是业务对象变更历史，与产品软件版本无关。

查看对象关系

在“知识图谱”查看对象与标签、字段、应用系统等资源的关联。建立关系时选择正确的资源和关系用途；知识关联不能直接等同于数据加工血缘。

导出与批量更新

从数据源采集技术元数据后，可使用导出和更新流程批量补充业务语义。常见维护内容包括逻辑名称或中文名、业务描述、标签、责任信息，以及字段关联的数据标准等。批量更新用于补齐 DAM 中的治理属性，不应改写数据库中的物理结构。

内容	更新前检查
逻辑名称或中文名	与业务术语一致，同名对象能通过数据源、Schema 和物理名区分
描述	说明业务含义、使用范围和必要限制，不复制没有解释作用的物理名
标签	使用已批准的标签及分类口径，不在模板中临时创造同义标签
数据标准关联	标准已发布，字段含义、类型和值域与标准一致
责任信息和其他业务属性	责任人、部门和维护边界已经确认

1. 在数据源右侧“...”选择“导出元数据”，或从元数据列表进入导出。
2. 下载全部对象，或勾选表后选择“下载选中的元数据”。超出单个 Excel 可容纳范围时，系统可能拆分文件并以压缩包提供。
3. 以导出文件为基础维护表级和字段级业务语义，保留对象唯一标识、物理名称、数据源、Schema、模板列和工作表结构。
4. 回到数据源“...”选择“更新元数据”，选择修改后的 Excel 数据字典。
5. 查看导入结果和失败明细，确认空值没有覆盖已有属性，标准编码、标签等引用能够正确解析。
6. 完成后逐项核对修改对象；先用少量记录验证，再按同一模板批量更新。

导出结果反映当前筛选和权限范围，不能将局部导出当作整个环境的备份。再次采集可能更新技术属性；业务语义是否保留或发生冲突，应结合产品配置、变更历史和导入结果复核。

检查与排查

现象	检查内容
搜不到对象	系统/数据源选择、类型筛选、采集结果及权限
SQL 不可见	角色是否允许查看语句
采样失败	数据库报错、账号读取权限、连接与结构
更新后未变化	是否选对来源和对象、文件定位信息是否保留、导入结果
对象名相同但内容不同	对照应用系统、数据源、Schema 和物理名称

相关文档：[关联数据标准](#)、[配置采集](#)。

元数据对象维护

元数据管理将数据库对象、元模型对象、报表和文件纳入相应的查询、维护和关系分析入口。采集任务负责获取来源信息，人工维护负责补齐业务属性，两者应区分来源。

按对象进入正确的查询栏目

对象	查询栏目
数据源、表、字段、视图、存储过程、函数等数据库对象	数据集
数据字典或 DDL 解析出的结构对象	数据集
API 和其他元模型对象	数据集
标准采集或自定义导入的报表	报表
文件存储服务器采集的文件	文件

查询与批量维护

在“元数据 → 数据集”使用技术部门、元数据类型、数据源、英文名称、中文名称、资产类型、业务系统、Schema 和创建时间等条件定位对象。打开对象详情后，先确认数据源、库、Schema、表名和采集时间，再编辑业务描述或责任属性。

批量更新前导出当前筛选结果，保留对象唯一标识。可在导出模板中补充逻辑名称、描述、标签、责任信息和数据标准关联等业务语义，再通过“更新元数据”导入。模板只更新允许修改的列；不要改变定位列或用空值覆盖采集得到的技术属性。详细步骤见[查询和维护表](#)、[字段元数据](#)。

模型变更后的元数据复核

模型设计、比对和同步的操作请参考 DDM 产品用户手册。源模型变更后，元数据维护人员应在 DAM 中检查采集对象与业务关联是否仍然准确。

复核时应：

1. 取得已批准的模型变更清单，确认数据源和 Schema。
2. 对受影响范围重新执行元数据采集，并对照清单核对新增、删除和属性修改。
3. 检查标准映射、血缘关系和数据质量规则是否引用受影响字段。
4. 记录未更新的采集对象和失效关系，交给相应责任人处理。

详细步骤见[模型变更与元数据复核](#)。

报表与文件元数据

报表可以通过 BI 工具标准采集，也可以在 DAM“报表”或 SQLink“报表管理”中导入、自定义登记；这些报表统一在 DAM“报表”维护名称、所属系统、负责人和业务说明。文件栏目只记录文件存储服务器采集到的文件或目录对象。导入前检查模板列和对象编码，导出时记录筛选条件。

查看报表或文件时，若需要分析上下游关系，先确认对象已采集并绑定到正确的应用系统；未绑定的对象可能只能看到自身属性，不能得到完整关系。

结果检查

- 查询结果中的对象总数与筛选条件一致。
- 技术属性的采集时间更新，业务属性保留维护记录。
- 标准映射、资产注册和基础血缘仍可从详情页跳转。
- 删除或下线对象前，已检查资产、规则和血缘引用。

采集自定义元模型数据

自定义元模型采集分为 DAM 路径和 SQLink 路径。两者都把实际对象映射到模型结构，但数据源、根节点和血缘准备方式不同。

DAM 路径：插件与模板

安装插件

1. 由管理员进入“系统管理 → 插件管理”。
2. 安装与自定义模型匹配的数据源插件和逆向插件。
3. 检查插件状态为已启用，再进入元数据数据源配置。

示例资料中的 METAMODELDEMO 适合模板验证；生产环境是否使用该插件或开发专用插件，应由交付人员确认。

定义数据源和采集任务

1. 在“元数据 → 数据源”新建自定义类型数据源。
2. 关联应用系统并选择对应插件，保存数据源。
3. 按标准采集流程创建采集任务。该任务主要建立数据源与模型的映射，采集名称通常作为 Schema 标识。

生成并填写 Excel 模板

1. 从数据源详情下载依据模型结构生成的模板。
2. 保留根节点 Sheet 中系统预填的 ID 和 Name；Name 对应 Schema，不要随意改名。
3. 在子对象 Sheet 填写对象数据及父对象 ID、Name。
4. 更深层对象按模板要求填写父子层级，保证同一对象在所有 Sheet 中使用一致标识。
5. 保存后进入“元数据 → 数据集”，选择对应数据源并执行“导入元数据”。

导入日志中的缺失父 ID、重复 Name、空必填属性和层级不一致应逐条修正。首次导入先使用少量对象，确认树状展示和属性分组后再批量导入。

SQLink 路径：文件采集与同步

1. 在 SQLink 的采集管理中选择“文件”类型，配置自定义元模型采集。
2. 填写血缘解析所需的命名空间；SQLink 路径的根节点 ID 可以为空，子对象可通过 Name 表达层级。
3. 运行采集任务，检查任务状态和对象数量。

4. 在 SQLink 数据目录按数据源筛选对象，确认对象已同步到 DAM 数据集。

不能把 DAM 模板要求的根节点 ID 规则套用于 SQLink 文件采集，也不能把 SQLink 的空根节点填写方式套回 DAM Excel 模板。

展示和增量更新

验证顺序建议为：根对象 → 子对象 → 属性分组 → 关联关系 → 数据集筛选 → 资产目录注册。增量更新只提交发生变化的对象，保留原有对象标识，避免重复创建。

为自定义对象建立血缘

1. 进入 SQLink 血缘空间，新建 Excel 血缘任务。
2. 下载关系模板，填写源对象、目标对象、命名空间和 Schema；支持多对多关系。
3. 从元数据对象的血缘分析页面关闭中文显示，可查看对象使用的命名空间和 Schema。
4. 导入并执行解析，回到元数据对象检查上下游关系。
5. 图为空时，检查元模型对象是否已同步到血缘空间的数据库资源，必要时手工同步。

自定义元模型对象的关系定义和血缘关系是两类信息：前者表达模型内的业务关联，后者表达实际数据处理链路，不能互相替代。

认识元模型与预置模型

元模型定义“平台要管理什么对象、对象有哪些属性、对象之间如何关联”；元数据是按照这些定义保存的实际对象。元模型是结构设计，不是某一张表或一次采集任务。

元模型的四个组成部分

组成	作用	示例
模型	定义一类资产的整体结构	数据库模型、接口模型、数据产品模型
对象	定义需要管理的实体类型	数据库、数据表、接口、数据产品
属性	描述对象的特征	名称、定义、负责人、请求地址
关系	表达包含、引用、依赖或业务关联	数据库包含表、指标来源于字段

对象是用户在数据集和资产目录中看到的管理单元；属性决定详情页展示内容；关系决定树状导航、对象跳转和后续分析能否成立。

预置元模型

平台对通用技术资产提供预置模型，并配套采集、展示和治理页面。常见预置模型包括：

- 通用数据库元模型：数据库、Schema、表、字段、视图、存储过程、函数、索引及主外键关系。
- REST API 元模型：服务、接口、请求参数、响应参数和接口关系。

使用预置模型时，用户不需要重新设计对象和属性，也不需要把数据库字段逐一映射到自定义模型。具体数据库或接口支持范围以[兼容性说明](#)为准。

预置模型还是自定义模型

需求	推荐路径	原因
管理数据库、数据仓库和常规表字段	预置数据库模型	已有专用采集和对象展示
管理 OpenAPI/REST 接口	预置 API 模型	可复用接口结构与采集方式
管理监管报送、设备、数据产品等企业对象	自定义模型	对象结构和业务属性具有组织特性
只是给现有表补充业务描述	不新建模型	直接维护元数据属性或使用标准映射

只有当预置模型无法表达对象边界时才创建自定义模型。模型一旦被采集任务、资产目录或接口引用，修改核心对象和属性会影响后续数据，必须先做影响评估。

元模型与 DAM、SQLink 的关系

DAM 和 SQLink 使用同一类元模型思想，但采集和展示入口不同：DAM 自定义模型通常通过插件和 Excel 模板导入；SQLink 自定义模型更偏向文件采集和血缘关联。两条路径创建的对象不能简单套用对方的根节点、命名空间或 Schema 填写规则。

下一步阅读[设计自定义元模型](#)和[采集自定义元模型数据](#)。

设计自定义元模型

元模型定义对象的结构，元数据是按该结构记录的实际内容。数据库和 REST API 等已有预置模型的资产优先使用预置模型；企业特有对象再考虑自定义。

四类要素

要素	含义	设计时应回答的问题
模型	某类资产的整体结构	管理什么范围
对象	具体实体类型	有哪些对象层级
属性	对象特征	每类对象记录什么信息
关系	层级、引用或业务关联	对象之间如何连接

设计前画出对象和关系，确定标识、名称、描述等公共属性。不要为了未来可能需求一次加入大量无用途字段。

创建模型

1. 进入“元模型”，创建新模型。
2. 填写代码、名称和描述。代码用于识别与引用，应稳定且具有业务含义。
3. 进入模型编辑页面。ID、GUID 由系统生成，不支持人工修改。
4. 核对模型数据区与模型结构区，随后建立对象、属性和关系。

模型编辑中的修改实时保存。生产模型调整前应先评估采集、模板和现有数据的影响，不要把关闭页面视为取消修改。

基本信息怎么填写

字段	用途	建议
代码	系统识别模型的稳定标识	使用有业务含义的英文编码，创建后尽量不改
名称/别名	页面展示和业务识别	中文名称表达管理对象，避免使用项目临时简称
描述	说明模型目标和范围	写清适用系统、责任团队和不包含的对象
ID、GUID	平台内部标识	由系统生成，不手工维护

一个模型代码只对应一套结构。若新旧结构必须并行，先建立独立模型或评估迁移策略，不要直接覆盖生产模型。

组装模型结构

1. 在“对象”集合中建立业务对象，填写名称、别名、描述及展示图标。
2. 如需作为资产目录中的资产类型，按设计设置是否允许注册到数据资产。
3. 在“属性”集合统一定义属性。同义属性尽量复用，避免各对象重复创建。
4. 为属性设置显示控件、分组名、显示顺序号；未分组属性进入默认分组。
5. 回到对象，通过“关联属性”选择已有属性。
6. 使用“关联对象”建立层级和业务关系。
7. 编辑模型根节点，关联公共基础属性，再关联业务对象。

根节点关联不可省略。 只创建对象集合与属性集合，没有把它们组装到根节点，可能导致模板不能生成、导入失败或展示不完整。



图示标出业务对象、基础属性与根节点的关联位置；对象名称按实际模型设计填写。

五步设计检查

第一步：创建对象集合

先列出真正需要管理的实体类型，例如 Agent、Goal、Knowledge、Memory 和 Model。对象不是一条数据记录，而是详情页和权限边界的类型。需要作为数据资产发布的对象，才开启“允许注册到数据资产”。

第二步：创建属性集合

在属性集合中统一定义属性，重复含义只建立一次。为属性设置控件类型、分组名和显示顺序；日期、枚举和长文本应选择适合的控件，避免全部使用普通字符。

第三步：关联属性

回到对象集合，使用“关联属性”选择已经定义的属性。关联不是复制属性，后续修改公共属性可能影响多个对象，修改前检查引用对象。

第四步：建立对象关系

使用“关联对象”表达父子、引用、依赖和业务关联。关系名称应说明方向和语义；不要把“可能相关”直接建成强依赖关系，避免搜索和影响分析产生误导。

第五步：组装根节点

根节点是模型实例的入口。先关联公共属性（例如 ID、名称、逻辑名称、定义），再关联业务对象集合。完成后用模板预览检查根节点、对象层级和属性分组。

设计原则

- 业务驱动：从要管理的业务对象出发，而不是从数据库表结构出发。
- 简洁：先保留支撑任务的核心属性，后续按真实需求扩展。
- 稳定：核心对象、属性编码和关系方向投入使用后尽量不变。
- 可扩展：预留扩展属性和新对象的边界，但不要预先创建无用途字段。
- 统一命名：对象、属性和关系采用同一命名规范，并在描述中记录业务定义。

发布前验收

用少量样例完成以下检查：

1. 模板可以生成，根节点 ID/Name 规则正确。
2. 对象树能按预期展开，父子关系没有循环。
3. 详情页的属性分组和顺序符合业务阅读习惯。
4. 可注册资产的对象能被目录类型选择，其他对象不会误出现在资产注册入口。
5. 采集、导入和血缘关系使用的对象标识一致。

验证设计

在少量示例对象上验证模板生成、父子关系、属性分组及展示顺序。检查名称是否唯一、层级是否符合业务、属性是否被错误重复定义。确认后按[元模型导入](#)建立实际数据。

对象关系只表达所定义的关联；需要加工血缘时，应另行提供源目标关系，不能假设所有对象关系自动成为血缘。

导入和验证自定义元模型数据

本页介绍 DAM 的 Excel 模板导入路径，并说明与 SQLink 文件采集的差异。开始前完成[元模型设计](#)，尤其是根节点组装。

准备 DAM 插件与数据源

1. 由管理员在插件管理中安装产品提供的元模型数据源和逆向插件。基础模板方案使用 METAMODELDEMO 插件套件。
2. 检查插件启用状态。此方案适合初期验证；接入其他来源时应使用适配的插件。
3. 在“元数据 → 数据源”新建自定义类型数据源，关联应用系统并选择元模型插件。
4. 按标准采集流程创建元模型采集。该模板方案的采集配置用于建立管理结构，不会直接从外部读取业务元数据；采集名称作为 Schema 标识。

不要把示范插件表单中的无效连接项当作真实外部连接配置。

填写并导入 Excel

1. 从数据源详情下载依据元模型结构生成的 Excel 模板。
2. 保留首个根节点 Sheet 的系统预填 ID 和 Name；Name 对应 Schema，不可任意修改。
3. 在子对象 Sheet 填写对象数据以及父对象 ID、Name。
4. 对更深层对象按模板要求填写层级名称，保证父子引用一致。
5. 保存后进入“元数据 → 数据集”，选择对应数据源，执行“导入元数据”。
6. 检查导入日志中的失败项，修正缺失父 ID、重复 Name 等错误。

如 AI_AGENT.DIG 这样的层级名称仅为格式示例，实际名称应由所采用的根节点和父子层级组成。

核对结果

在数据源下查看实例列表，进入根对象检查基础属性和子对象，沿层级导航检查属性分组及关联。增量更新时只准备需要变更的内容，并在导入后核对合并结果，先用少量对象确认规则。

SQLink 路径的区别

SQLink 在采集管理中使用“文件”类型进行元模型采集，需要填写命名空间。该路径不使用 DAM 示范插件，根节点 ID 可为空，子对象可以通过 Name 描述根节点；不能把这种填写方式套用到 DAM 的根节点模板。

运行采集后，在 SQLink 数据目录筛选对应数据源查看对象，并核对同步到 DAM 数据集的结果。

建立自定义血缘

1. 在 SQLink 血缘空间新建 Excel 血缘任务，下载关系模板。
2. 按模板填写源、目标、命名空间和 Schema；可定义多对多关系。
3. 如不确定对象路径，可在对象血缘分析中关闭中文显示，查看命名空间与 Schema。
4. 导入并执行解析，在元数据对象上检查血缘。
5. 图为空时检查元模型元数据是否同步至血缘空间的数据库资源，必要时手工同步。

相关文档：[SQLink 管理](#)、[基础血缘](#)。

模型变更与元数据复核

模型比对用于把 DDM 中维护的数据模型与 DAM 从数据库采集到的实际结构进行比较，识别表、字段、视图及属性差异，并生成比对结果。典型场景是先在 DDM 将数据库结构逆向成模型，再在 DDM 维护模型元数据，最后通过 DAM 的模型比对确认差异并按批准范围更新。

DDM 负责数据库逆向、模型设计和模型版本管理；DAM 负责采集实际数据库结构、发起模型比对以及复核治理关联。两端比较的对象、版本和同步方向必须明确，不能把模型设计稿未经审核直接覆盖生产元数据。

建立可比较的两个基线

1. 在 DAM 为目标数据库建立数据源并执行元数据采集，确认数据库、Schema、表和字段范围完整。
2. 在 DDM Client 使用[逆向数据库](#)读取同一数据库和 Schema，生成物理数据模型。
3. 将模型保存到约定的 DDM 模型库目录并建立可识别版本，记录其对应的数据库环境和逆向时间。
4. 在 DDM 中按设计要求维护中文名、描述、数据标准、字段定义及其他模型元数据，完成模型评审并发布需要参与比对的版本。
5. 在开始比较前确认 DAM 采集时间与 DDM 模型版本，避免拿不同环境或不同时点的数据作结论。

DDM 也可直接在[模型比较](#)中比较模型与逆向数据库。需要把模型结果用于 DAM 元数据治理时，再按本页流程在 DAM 建立比对任务。

在 DAM 配置模型比对

1. 进入“元数据 → 数据集”，定位目标数据源，在数据源操作菜单选择“模型比对与同步”。
2. 选择与该数据源对应的 DDM 模型及已批准版本，核对模型库、模型名称、数据库类型和 Schema。
3. 启用模型比较，选择本次参与比较的对象和属性。只比较需要治理的范围，不默认把所有模型属性纳入更新。
4. 执行比较并查看差异，按新增、删除和修改分类检查表、字段、视图、数据类型、长度精度、默认值、主外键及索引。
5. 导出或保存比对报告，记录 DAM 对象、DDM 模型版本、差异类型、处理决定和责任人。

对同名但语义不同的字段，应由业务责任人确认，不能只依据技术名称自动合并。比对结果表示两个基线之间存在差异，不自动证明数据库或模型哪一方正确。

审核并更新元数据

1. 由模型负责人和数据库负责人确认差异原因，区分“数据库已实施但模型未更新”“模型已设计但数据库未实施”“DAM 尚未重新采集”和“比较范围不一致”。

2. 需要把 DDM 中批准的业务属性同步到 DAM 时，启用模型同步并明确属性范围，例如数据标准、标准代码、中文名、描述或自定义属性。
3. 按页面提供的同步策略处理空值和冲突。执行前保留比对报告，先选择少量对象验证同步方向。
4. 若差异属于数据库物理结构变更，应由数据库变更流程实施，再重新执行 DAM 采集；不要用业务属性同步代替数据库 DDL 变更。
5. 更新完成后重新运行比较，确认已处理差异消失，未处理差异仍有明确原因和责任人。

检查关联影响

删除字段、修改类型或调整约束后，检查数据标准映射、质量规则、资产目录、指标映射和数据血缘。发现关联对象失效时，记录责任人和处理结果；不能以采集任务成功或比对任务完成代替业务影响确认。

执行后复核

重新打开受影响对象的详情，检查结构、中文名、描述、标准映射、标签和关联关系。对目标数据源重新执行小范围[元数据采集](#)后再次比较，确认 DAM、DDM 与数据库的状态符合批准结果。模型设计、版本或 DDL 生成问题仍应回到 DDM 操作记录处理。

比对结果检查清单

- DAM 数据源与 DDM 模型指向同一数据库环境和 Schema。
- 比较使用的 DDM 模型版本已经评审，DAM 元数据采集时间符合本次范围。
- 比对报告记录新增、删除、属性修改及不处理原因。
- 同步方向和属性范围经过确认，没有用空值覆盖有效业务语义。
- 数据库物理变更通过数据库发布流程实施，DAM 已重新采集。
- 标准、质量、资产、指标和血缘等下游引用已完成影响复核。

管理报表与文件元数据

报表元数据描述报表、信息项及库表关联；文件元数据描述文件存储服务器采集到的文件对象。两类对象的获取方式和查询范围不同。

查询栏目	会出现的内容	不会因为采集材料是文件就出现在这里的内容
报表	FineBI、FineReport、SmartBI 等标准采集报表，以及 DAM 或 SQLink 手工导入的自定义报表	数据字典、DDL、普通服务器文件
文件	通过“文件存储服务器”采集的文件和目录对象	物理 / 逻辑数据字典、DDL、API 元数据、元模型模板

新建报表

1. 进入元数据管理的“报表”，选择归属目录。
2. 新建报表，填写基础属性及目录信息。
3. 补充信息项和库表信息，建立内容关联。
4. 保存后打开详情，核对目录、属性及关联对象。

需要自动获取受支持 BI 工具的报表时，使用[报表采集](#)，而不是手工重录采集结果。FineBI、FineReport、SmartBI 等标准采集结果与自定义报表最终在本栏目统一管理。

导入报表

1. 从报表导入入口取得模板。
2. 按模板准备属性及关联信息。
3. 选择导入策略，确认是更新还是覆盖，再上传文件。
4. 检查导入结果，并对照原文件核对报表数量和关联库表。

覆盖策略可能改变现有报表内容，先核对同名对象和范围。

维护与导出报表

通过查看进入详情，通过编辑修改属性。列表勾选后可批量删除，删除前核对范围。目录导出的范围为当前目录下的报表，不是全部目录。

如果报表中存在已解析关系，可从报表对象查看[基础血缘](#)。手工填写说明文字不会自动生成完整血缘。

查看与下载文件

本栏目的文件类元数据只通过“文件存储服务器”采集获得。物理 / 逻辑数据字典、DDL、API 元数据和元模型文件虽然从“文件”采集入口上传，但会被解析为结构化对象，应到“数据集”查询。

1. 进入“文件”，通过列表搜索目标文件。
2. 打开详情查看基础属性，需要时维护可编辑属性。
3. 返回列表勾选目标文件，执行文件下载。
4. 核对下载内容与目标文件对应，按组织规则存放和使用。

没有找到文件时，先确认采集类型是“文件存储服务器”，再检查任务是否执行、采集路径和范围是否包含该文件，以及当前用户权限。若使用的是数据字典、DDL、API 元数据或元模型文件类型，请到“数据集”查找解析结果。

配置数据源采集账号权限

元数据采集账号必须能够连接目标数据源，并读取逆向插件所需的系统目录、对象定义和指定 Schema。权限不足时，连接测试仍可能成功，但采集会失败、只返回部分对象，或缺少视图、存储过程、主外键和索引等信息。

数据库授权由数据库管理员执行。优先建立专用只读账号，按目标数据库和 Schema 授予最小权限，不使用业务应用账号，也不把管理员角色作为默认方案。下面的授权示例已按 DAM 元数据抽取和逆向所需能力整理，账号名、数据库名、Schema 名和密码必须替换为当前环境的值。不同数据库版本、认证方式和插件版本可能需要补充系统目录权限，执行前应由 DBA 审核。

区分元数据权限和数据读取权限

使用目的	需要的能力	是否读取业务记录
测试连接	登录目标数据库或服务	否
逆向元数据	读取系统目录、对象定义、Schema、表、字段、视图、过程、函数、约束和索引等	通常不需要读取表中记录
采样或数据探查	对指定表或视图执行受限查询	是
质量检核或其他任务	按任务 SQL 读取目标字段和记录	是

如果项目只需要结构采集，不应因为采样功能存在就默认授予全部业务表查询权限。确需采样时，应限定允许读取的库、Schema、表、行数和账号使用范围。

按数据库类型准备权限

数据库类型	元数据逆向重点	需要额外确认的内容
MySQL	目标库对象定义及引用关系；限定库时按库授权	是否需要视图定义、例程和采样数据
Oracle	建立会话并读取数据字典	是否跨 Schema；是否需要读取表数据
达梦	读取数据字典、视图、序列和目标对象	授权范围是否限定到目标模式
SQL Server	读取目标数据库中的对象定义和依赖信息	不应直接把 db_owner 当作通用最小权限，由 DBA 按版本拆分只读权限
Hive、Impala	访问目标 Database 及其对象	元数据服务、Kerberos 或平台权限是否同时满足
PostgreSQL、Greenplum	连接目标库、使用目标 Schema、读取系统目录	采样时再增加目标表 SELECT；新增表的默认权限是否需要配置
openGauss	读取 pg_catalog 中逆向所需目录	目录视图范围是否与当前插件版本匹配
MongoDB	对目标 Database 读取集合及结构信息	是否允许样本读取；认证库是否正确
Vertica、ClickHouse、Doris 等	读取目标 Schema 或 Database 的结构信息	按内部清单和当前插件说明确认具体权限

上表说明授权目标，不替代数据库厂商或项目的正式授权脚本。对于生产环境，DBA 应先审查授权语句，再在限定范围内执行。

授权语句参考

以下命令用于说明授权范围和对象，不包含任何真实账号或密码。命令中的 user_name、database_name、schema_name 和 password 均为占位符。

MySQL

全库结构和数据读取：

❏sql

```
GRANT SELECT ON *.* TO 'user_name'@'%';
```

限定一个数据库：

❏sql

```
GRANT SELECT ON database_name.* TO 'user_name'@'%';
GRANT REFERENCES ON database_name.* TO 'user_name'@'%';
```

如果只采集结构，先由 DBA 按插件实际读取的系统表授予目录权限；需要采样时再授予目标表的 SELECT，不要直接扩大到所有业务库。

Oracle

Oracle 逆向至少需要建立会话并读取数据字典：

□sql

```
CREATE USER user_name IDENTIFIED BY "password";
GRANT CREATE SESSION TO user_name;
GRANT SELECT ANY DICTIONARY TO user_name;
```

如果只采集指定 Schema，优先使用项目批准的对象级授权方案，并确认跨 Schema 视图、存储过程和依赖关系是否需要额外读取权限。

达梦

常用的元数据和对象读取权限参考：

□sql

```
CREATE USER user_name IDENTIFIED BY "password";
GRANT SELECT ANY DICTIONARY,
      SELECT ANY MATERIALIZED VIEW,
      SELECT ANY SEQUENCE,
      SELECT ANY VIEW,
      SELECT ANY TABLE TO user_name;
```

生产环境应结合目标模式收窄范围；指定数据库的授权脚本由 DBA 按当前达梦版本确认。

SQL Server

采集账号需要能够读取目标数据库的对象定义和依赖信息。内部适配示例使用数据库级读取角色：

□sql

```
USE database_name;
ALTER ROLE [db_datareader] ADD MEMBER [user_name];
GRANT VIEW DEFINITION TO [user_name];
```

部分旧环境使用 db_owner 作为快速验证方式，但该角色权限过大，不应直接作为生产默认配置。验证完成后应由 DBA 按目标库、Schema 和对象范围改为最小权限。

Hive 和 Impala

限定数据库的读取示例：

□sql

```
GRANT SELECT ON DATABASE database_name TO USER user_name;
```

还需确认 Hive Metastore、认证方式、Kerberos 或平台权限已满足插件连接要求。不同发行版的授权入口可能不同。

PostgreSQL 和 Greenplum

创建账号并准备系统目录和目标 Schema 的读取权限：

□sql

```
CREATE USER user_name;
ALTER USER user_name WITH PASSWORD 'password';
GRANT CONNECT ON DATABASE database_name TO user_name;
GRANT USAGE ON SCHEMA schema_name TO user_name;
GRANT SELECT ON ALL TABLES IN SCHEMA schema_name TO user_name;
ALTER DEFAULT PRIVILEGES IN SCHEMA schema_name
    GRANT SELECT ON TABLES TO user_name;
```

逆向插件还需要读取 pg_catalog 中与当前版本匹配的对象目录。采样权限只授予批准的目标 Schema 或表。Greenplum 按 PostgreSQL 方案准备后，再结合集群版本验证。

openGauss

逆向所需的目录读取示例：

□sql

```
GRANT SELECT ON pg_catalog.pg_namespace TO user_name;
GRANT SELECT ON pg_catalog.pg_class TO user_name;
GRANT SELECT ON pg_catalog.pg_attribute TO user_name;
GRANT SELECT ON pg_catalog.pg_proc TO user_name;
GRANT SELECT ON pg_catalog.pg_constraint TO user_name;
GRANT SELECT ON pg_catalog.pg_inherits TO user_name;
GRANT SELECT ON pg_catalog.pg_index TO user_name;
GRANT SELECT ON pg_catalog.pg_type TO user_name;
GRANT SELECT ON pg_catalog.pg_roles TO user_name;
GRANT SELECT ON pg_catalog.pg_user TO user_name;
GRANT SELECT ON pg_catalog.pg_extension TO user_name;
```

再按目标库和 Schema 配置连接、使用 Schema 及目标表读取权限，并用小范围对象验证当前插件版本是否还需要其他目录。

MongoDB

对指定数据库授予读取角色：

□ javascript

```
db.grantRolesToUser(
  "user_name",
  [{ role: "read", db: "database_name" }]
);
```

连接时确认认证数据库、目标数据库和账号所属用户域一致。若启用采样，按组织要求控制可读取的集合和样本范围。

Vertica、ClickHouse 和 Doris

Vertica 的 Schema 读取示例：

□ sql

```
GRANT SELECT ON SCHEMA schema_name TO user_name;
```

ClickHouse 的读取示例：

□ sql

```
GRANT SELECT ON *.* TO user_name;
```

Doris 的单库读取示例：

□ sql

```
CREATE USER 'user_name'@'%' IDENTIFIED BY 'password';
GRANT SELECT_PRIV ON database_name.* TO 'user_name'@'%';
```

Doris 全库授权只适合经过审批的管理场景。Vertica、ClickHouse 和 Doris 的对象目录、视图及采样需求仍需按实际插件版本验证。

MaxCompute

MaxCompute 需要配置 Endpoint、Access ID、Access Key 和 Project Name，并在项目侧授予元数据读取权限。Access Key 不得写入手册、截图或共享文档；应通过安全配置或密钥管理方式提供给数据源连接。

授权与验证步骤

1. 确认本次要采集的应用系统、数据库、Schema、对象类型，以及是否需要采样。
2. 创建专用采集账号，设置符合组织要求的密码轮换、来源地址和有效期策略。
3. 根据数据库类型和逆向插件，只授予连接、系统目录及目标对象所需权限。
4. 在 DAM 数据源中使用该账号测试连接。连接成功只证明登录可用，不代表逆向权限完整。
5. 创建只包含少量已知对象的采集任务，执行后核对表、字段、视图、存储过程、函数、主外键和索引数量。
6. 如需采样，再对批准的目标表补充数据读取权限，并使用少量记录验证。
7. 验证通过后扩大采集范围，记录账号、授权范围、插件版本和验证日期，供后续变更复核。

逆向失败或不完整时排查

现象	优先检查
连接测试失败	网络、地址、端口、认证方式、账号状态和登录权限
连接成功但没有对象	系统目录权限、目标库 / Schema 范围、逆向插件和账号默认库
只有表和字段，缺少视图或过程	对象定义、例程或系统目录权限是否完整
只能采集部分 Schema	账号可见范围、采集筛选、白名单及跨 Schema 权限
结构采集成功但采样失败	目标表 SELECT 权限、行级安全策略及查询限制
权限补充后结果仍未变化	是否重新执行采集、任务是否仍使用旧账号、缓存或插件是否需要重新加载

不要通过反复增加全局权限来试错。先从任务日志确认失败对象和读取动作，再由 DBA 补充对应的最小权限，并重新执行小范围采集。

将字段关联到数据标准

通过字段与标准的映射，把业务定义、技术要求及代码值域关联到具体元数据。关联前确认字段语义与标准一致，不能只按名称相似程度判断。

使用前提

字段已采集，可访问目标字段和适用标准。标准及标准代码应处于适合使用的状态；对有多个业务口径的字段，先与业务责任人确认。

建立映射

1. 在“数据集”找到目标表，单击字段名进入详情。
2. 在“技术属性”中单击数据标准的编辑按钮。
3. 搜索并查看候选数据标准或指标标准，核对定义、类型、长度和值域。
4. 单击“选定”建立映射。
5. 如果系统询问是否引用标准的标准代码，先确认当前字段的取值语义再选择。
6. 返回表详情，检查该字段的标准关联。

需要单独管理标准代码时，可通过技术属性中的标准代码编辑入口操作。

验证引用关系

打开对应标准的“引用信息”，确认目标字段及所属来源正确。建立映射不会自动修改源数据库字段，也不代表数据取值已经通过质量检查。技术属性是否符合标准，应结合数据标准的核标任务与数据质量检核判断。

调整或解除映射

发现业务口径错误时，先核对标准和字段，再调整关联。在标准的引用信息中可对字段解绑；解绑只处理引用关系，应同时检查后续核标、质量及治理工作是否需要调整。

相关文档：[基础标准](#)、[标准代码](#)、[落标与核标](#)。