

# 元数据管理

---

## 元数据管理

### 概述

#### 元数据分类

业务元数据

技术元数据

操作元数据

#### 元数据核心能力

元数据采集能力

元数据存储能力

元数据查找能力

血缘分析/关系分析能力

基于角色的访问权限控制和分层能力

业务元数据管理能力

元数据变更控制能力

元数据对比分析能力

数据生命周期管理能力

与其他系统的集成能力

### 建设目标

### 建设过程

元模型管理

元数据集成和变更

元数据应用

### 标准条款

级别 1：初始级

级别 2：受管理级

级别 3：稳健级

级别 4：量化管理级

级别 5：优化级

# 元数据管理

## 概述

元数据是“关于数据的数据”，即对数据的描述信息。元数据管理是关于元数据的创建、存储、整合与控制等一整套流程的集合。

元数据的信息范围很广，它不仅包括**技术和业务流程**、**数据规则和约束**，还包括**逻辑数据结构与物理数据结构**等。它描述了数据本身（如数据库、数据元素、数据模型）、数据表示的概念（如业务流程、应用系统、软件代码、技术基础设施）、数据与概念之间的联系（关系）。

## 元数据分类

对数据不同方面的描述形成不同类型的元数据，一般而言，元数据一般分为**业务元数据**（如业务术语、业务规则等）、**技术元数据**（如数据库表结构、文件结构等）、**操作元数据**（如数据的更新时间、更新频率等）和**管理元数据**（如数据的负责部门、负责人等）。

业务元数据主要是数据中业务语义的描述；技术元数据主要关注技术的实现；操作元数据主要关注数据处理的过程与关系。

## 业务元数据

业务元数据包括**业务名称**、**业务定义**、**业务描述**等。业务元数据使用**业务名称**、**定义**、**描述**等信息表示企业环境中的各种属性和概念。业务元数据能让用户更好地理解和使用企业环境中的数据，业务元数据广泛地存在于企业环境中。

### 1. ERP系统

企业的ERP系统中存储着大量的业务元数据，比如财务计算公式、过程逻辑、业务规则等。

### 2. 报表

报表的表头也是一种业务元数据，特别是那些包含合计、平均数等带有总结性质的列，以及报表中的一些计算公式等。

### 3. 表格

Excel的表头和公式也是很重要的业务元数据。大多数表格中会有单独一列“描述”，有些表格中还会有一列代码和代码描述，这些都是很有用的业务元数据。

### 4. 文件

文件中到处都是业务元数据，比如标题、作者、修改时间等，文件内容中的业务元数据的获取相对比较困难，涉及机器学习等技术。

## 5. BI工具

操作是“钻取”，在向上和向下钻取的过程中通常定义了企业的各种分类结构，这些都是很重要的业务元数据。

## 6. 数据仓库

数据仓库中也有业务元数据，比如，在构建数据仓库之前通常需要做大量的调研来研究如何集成多个数据源，这些与数据仓库构建过程相关的文件中存在着大量的业务元数据。

## 技术元数据

技术元数据描述有关数据的技术细节、存储数据的系统，以及在系统内和系统之间数据流过程的信息。

- (1) 物理数据库表名和字段名；
- (2) 字段属性；
- (3) 数据库对象属性；
- (4) 访问权限；
- (5) 物理数据模型，包括数据表名、键和索引。

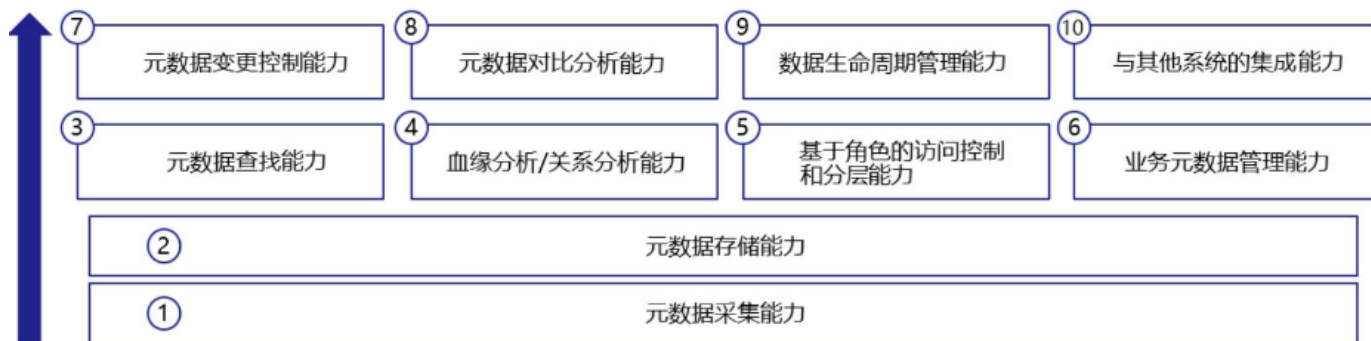
## 操作元数据

操作元数据描述了处理和访问数据的细节，例如以下内容：

- (1) 批处理程序的作业执行日志；
- (2) 抽取历史和结果；
- (3) 调度异常处理；
- (4) 审计、平衡、控制测量的结果；
- (5) 错误日志。

## 元数据核心能力

元数据管理工具基础能力如图：



## 元数据采集能力

元数据采集能力是指从错综复杂的企业环境中自动实时解析和采集各种元数据的能力。

为应对各种数据环境，需要使用各种技术和语法来支持大数据平台、关系型数据库、第三方工具、存储过程、脚本、文本文件、表格文件中的自动化数据采集。

## 元数据存储能力

元数据存储能力是将采集到的元数据进行统一存储的能力。

为支持各种元数据，以及元数据之间关系的存储，元数据存储需要灵活、可扩展的架构支撑，能够实时更新存储。

## 元数据查找能力

元数据查找能力是一种提供统一的端口对元数据进行查找的能力。完善的元数据管理工具应该能支持按照企业的各种分类方法来对元数据进行查找。

## 血缘分析/关系分析能力

血缘分析/关系分析能力是指分析数据的来源和数据的流向，揭示数据的上下游关系的能力。

在元数据管理工具中可以分析、描述并可视化其中的细节，方便用户对关键信息进行跟踪。

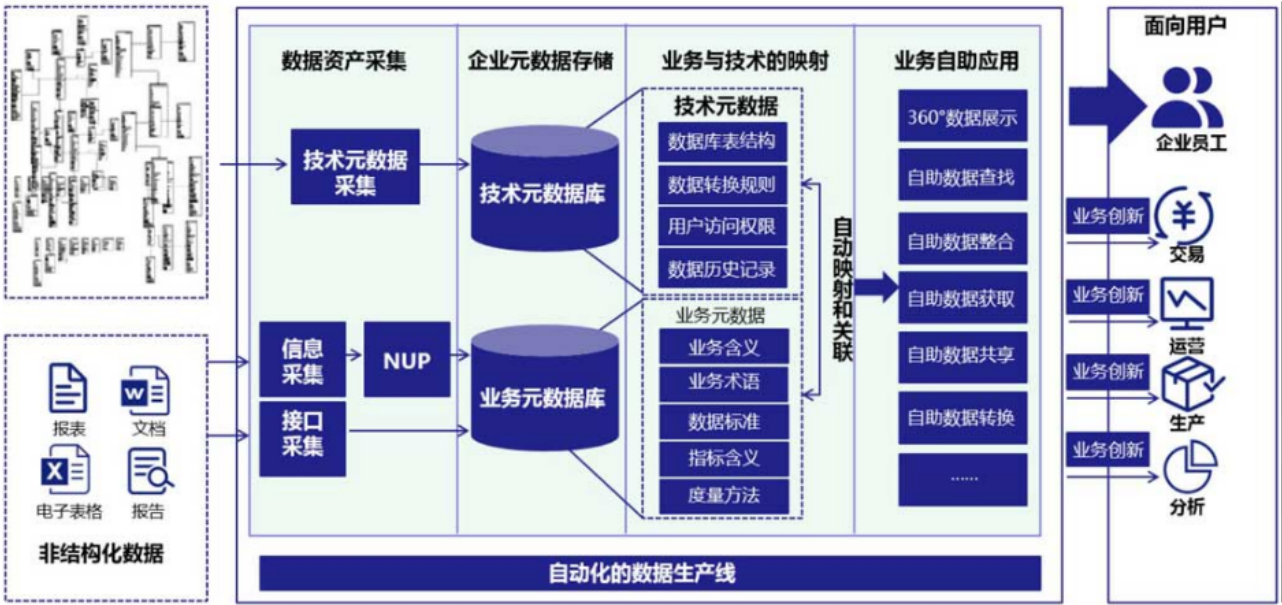
完善的血缘分析需要是**横向（当前）**和**纵向（历史）**皆可用的，以方便对同一时期的不同对象进行分析和不同时期的同一对象的变化进行分析。

## 基于角色的访问权限控制和分层能力

元数据的增、删、改等权限的控制，元数据管理工具应该支持访问权限控制。

## 业务元数据管理能力

采集企业环境中的业务元数据，并完成业务元数据与技术元数据的映射，为元数据赋予业务属性。



技术元数据与业务元数据的映射

元数据变更控制能力

当元数据需要变更时，元数据管理工具提供了变更控制能力，明确元数据版本，保存元数据的历史状态，在发生任何问题时可以自动恢复到之前的版本。

元数据对比分析能力

元数据管理工具还可以对不同环境中的元数据进行对比分析，分析其中的异同，必要时还能根据分析结果产生相应的分析报告。

数据生命周期管理能力

元数据管理工具应该保留数据“创建→存储→过时”这个过程中被删除/备份等各种状态下的元数据，从而管理数据在整个生命周期中的流动。

与其他系统的集成能力

元数据管理工具与其他系统的集成能力，让元数据管理系统发挥业务价值。

建设目标

建设目标应包括以下内容：

1. 根据业务需求、数据管理和应用需求，对元数据进行分类，建立元模型标准，保障不同来源的元数据集成和互操作；元模型变更实现规范管理；
2. 实现不同来源的元数据有效集成，形成组织的数据资产全景图，能够从业务、技术、操作、管理不同视角管理和使用数据资产；元数据变更遵循规范化流程并与信息系统开发生命周期实现良好融合；
3. 建立元数据应用和元数据服务，提升相关方对数据资产的理解，辅助数据管理和数据应用。

## 建设过程

元数据管理主要包括以下活动：

### 元模型管理

对包含描述元数据属性定义的元模型进行分类并定义每一类元模型，元模型可采用或参考元模型国际标准。

### 元数据集成和变更

基于定义的元模型对元数据进行收集，对不同类型、不同来源的元数据进行集成，形成对数据资产描述的统一视图，并且元数据随着数据资产的变化而更新，基于规范化的流程对元数据变更过程加以管理。

### 元数据应用

基于数据管理和数据应用需求，对于组织管理的各类元数据进行分析应用，如查询、血缘（溯源）分析、影响分析、一致性分析、质量分析等。

## 标准条款

度量标准应包含以下内容：

### 级别 1：初始级

条款 1) 元模型的定义遵循应用系统项目建设需要和工具已有定义；

条款 2) 在项目层面生成和维护各类元数据,如业务术语、数据模型、接口定义、数据库结构等；

条款 3) 在项目层面收集和实现元数据应用需求,如数据字典查询、业务术语查询等。

级别 2：受管理级

- 条款 1) 在某个业务领域,对元数据分类并设计每一类元数据的元模型；
- 条款 2) 元模型设计参考国际、国内和行业元模型规范；
- 条款 3) 在某个业务领域建立了集中的元数据存储库，统一采集不同来源的元数据；
- 条款 4) 在某个业务领域制定了元数据采集和变更流程；
- 条款 5) 在某个业务领域,初步制定了元数据应用需求管理的流程,统筹收集、设计和实现，元数据应用需求；
- 条款 6) 实现了部分元数据应用,如血缘分析、影响分析等,初步实现本领域内的元数据共享。

级别 3：稳健级

- 条款 1) 制定了组织级的元数据分类及每一类元数据的范围，设计相应的元模型；
- 解读：首先制定了**元数据管理制度**，明确了组织元数据分类，一般包含业务元数据、技术元数据和管理元数据；其次针对每一类元数据设计了**元模型**，确定每一类元数据的属性。
- 条款 2) 规范和执行组织级元模型变更管理流程，基于规范流程对元模型进行变更；
- 解读：首先制定了**元模型变更管理流程**，可放在元数据管理制度中，也可形成单独的元模型变更流程，规范元模型变更需求、审批、评审等流程；其次实际**元模型变更时，要遵从元模型流程**。

**第二十条 元模型的修改和变更**

(一) xx 负责对元模型的修改和变更；

(二) 元模型执行变更之前需组织对变更影响的应用、程序、接口等进行评估确认，告知相关公司业务部门做好必要准备，避免造成业务故障；

(三) 修改和变更后的元模型应通过接口传输方式将变更后的元模型推送其他合作单位，通过数据资源共享平台进行统一管理，以确保元模型数据的准确性。

**第二十一条 元模型的废止和下线**

(一) 当各业务部门的系统库表/接口/指标等已不再满足当前业务需求，需要对其进行废止下线时，可向公司 xx 发起下线申请，由公司 xx 进行审批确认下线需求，同时根据元模型废止和下线流程进行元模型废止下线操作。

(二) 通过审批后发布元模型退役。

**第二十二条 元模型的检查和稽查**

(一) xx 负责定期对元模型的准确性、完备性和及时性等稽核。

(二) 通过开发稽查工具，并固化相应的稽查规则，实现自动与真实系统进行校验、评估，确保稽查元数据与真实系统的一致性。

(三) 元模型的稽核可采用每月抽样检查的方式进行，每年需对系统内元模型进行全面的扫描稽核。

元模型修改工作联系单				
元模型修改说明	发起人	发起日期		
	联系电话			
	工作部门	应用系统	力开	
	元模型类别	预修改数据量	修改原因说明	
元模型修改审批	数据存储类	2	为了解决数据目录完整性，增加 2 字段。	
	审批意见			
元模型修改审批	审批人		初审日期	2021.6.3
	审批意见			
元模型修改审批	审批人		审批日期	2021.6.3
	其他说明			
为了解决数据目录完整性，增加 2 字段。 数据表所有者和数据表类型。				

元模型变更管理流程

条款 3) 建立了组织级集中的元数据存储库，统一管理多个业务领域及其应用系统的元数据，并制定和执行统一的元数据集成和变更流程；

解读：一是制定了**元数据管理制度**，明确了元数据采集和元数据变更的管理流程；二是在过程文件上，在实际**元数据采集和变更管理**时，要**遵从已有的管理流程**；三是在工具平台上，开发了**元数据管理平台**，统一采集和存储公司元数据。

条款 4) 元数据采集和变更流程与数据生存周期有效融合，在各阶段实现元数据采集和变更管理，元数据能及时、准确反映组织真实的数据环境现状；

解读：在**数据采集、数据存储、数据传输和数据交换**等环节实现元数据采集和变更管理，一般在数据中台-元数据管理实现。

数据治理规划与实施。

**第十六条 元数据的应用**

（一）元数据应用是对元数据进行查询和使用以满足数据管理需求的过程，包括血缘分析、影响分析、数据语义分析等。

（二）公司各业务部门在本专业领域内需要新增或修改元数据查询权限、采集权限、服务权限等，需要向 XX 提出元数据使用申请。XX 审核通过后新增或修改相应元数据使用权限。

（三）XX 负责优化提升元数据管理工作，实现对信息管理系统以及生产系统内各类元数据的采集和存储，提供元数据运维支撑能力，以业务人员视角提供体验良好的可视化界面和实时查询能力，可动态、灵活的展现元数据之间的血缘关系、影响关系，辅助各种数据问题的追溯和分析。

元数据应用需求管理流程

条款 5) 制定和执行统一的元数据应用需求管理流程，实现元数据应用需求统一管理和开发；

解读：一是制定**元数据应用需求管理流程**，包含需求收集、需求评审等内容，可与信息化项目研发流程融合，也可在元数据管理制度中体现。

条款 6) 实现了丰富的元数据应用，如基于元数据的开发管理、元数据与应用系统的一致性校验、指标库管理等；

解读：在元数据管理工具上，实现了丰富的元数据应用。**元数据与应用系统的一致性校验**，识别目前应用系统的数据定义与数据标准或者数据模型不一致；**血缘分析和影响分析**，可以查询数据来源以及数据对下游的影响。





元数据应用--血缘分析



元数据内容以服务的方式共享使用

条款 7) 各类元数据内容以服务的方式在应用系统之间共享使用。

解读：应用系统可以通过接口等服务方式获取到应用系统的元数据，比如与数据质量管理工具、数据安全运营系统、指标库集成。

## 级别 4：量化管理级

条款 1) 定义并应用量化指标，衡量元数据管理工作的有效性；

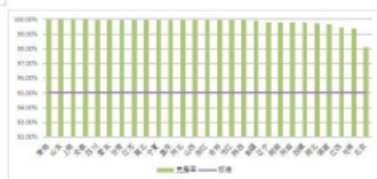
解读：在元数据管理制度或者方案，制定元数据管理的指标，如元数据覆盖率、元数据完整性、元数据规范性等指标，并定期形成元数据管理报告。

## 公司元数据管理情况分析报告

开展了数据盘点工作, 当前已取得阶段性成果。[ ]等  
目录工具已全面上线运行, 在数据应用溯源中发挥[ ]支撑作用。  
为提升数据资源目录内容维护的及时性和准确性, 促进数据资源  
的共享共建共用, 提高元数据管理能力。在各单位数据盘点  
和数据目录建设成果上, 按照“统一管理、服务应用、协同配  
合、[ ]”原则, 构建公司级数据资源目录体系,  
实现对公司数据资源的统一管

### 一、各单位应做的工作

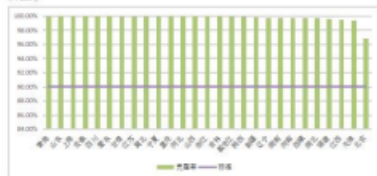
合A级单位4家,分别 东、  
B级单位23家 二、  
为 新疆、四川 、湖北、  
湖南、 京、河北、浙江、陕西、  
、指称评价C级单位0家;指称评价D级单



## 二、各指标具体情况

### (一) 元数据覆盖率

目前,元数据覆盖率为99.78%,其中,青海、山东、上海、安徽、四川、蒙东、甘肃7家单位为100%,20家单位为90%以上。



## (1) 指标元数据覆盖率

从各单位指标元数据覆盖率来看,有26家单位达到90%及以

## 元数据管理的指标

条款 2) 与外部组织合作开展元模型融合设计、开发;

解读：组织与外部机构（同行业、监管机构、跨行业等组织）合作开展的数据服务产品的技术方案（主要体现元数据设计共同参与）。

基于选定的中台或数据库产品,形成数据库结构和表设计。

### 三、设计原则

1.需求驱动,以用促建:以数据中台全链路监测、知识图谱建设、数据中台全景视图等重点应用场景为驱动,采取迭代拓展和更新的方式,开展元仓模型设计、落模应用和优化完善工作。

2. 联合设计, 持续迭代: 结合重点建设任务, 组建国网和阿里云、华为云等外部机构的联合工作团队, 共同开展元数据和元仓模型需求梳理和模型优化设计等工作, 形成统一设计成果, 实现共享共用。

3. 基线控制, 按需扩展: 基于数据中台全链路监测、知识图谱建设、数据中台全景视图等方面元数据需求, 统一组织设计元仓模型基线版本; 基线未覆盖部分, 按需各自扩充完善, 总库规范后按需统一纳入模型迭代范围。

4. 两级协同, 在线管控: 建立总部、各单位两级模型完善、应用协同和服务支撑机制, 提高问题解决和需求响应效率质量, 共同推进模型常态化和落地应用。

#### 四、元模型融合设计、开发

### (一) 元模型融合设计

元模型描述元数据的结构和语义, 描述实体的特征。

面向以数据中台为核心，构建两级中台贯通中元模型分析应用，融合省测华为云、云网两类数据中台中元数据合

库的现状 参照 CMM 模型, 与阿里、华为等外部组  
织共同设计, 元模型。将元模型分为六类, 分别是基础类、  
数据采集类、数据存储类、数据处理类、数据应用类、数据

根据业务部门的业务需求,组织开展元模型定义工作,各业务部门予以配合,对本专业领域内元模型的提出业务需求及协助定义。组织开展对业务需求元模型的识别,并根据需求完成基础类、数据采集类、数据存储类、数据处理类、数据应用类、数据保障类等二级元模型定义相关工作。

### ✓ 模型架构设计

基于元模型分类设计原则,将需求分析阶段输出的元对象进行总结完善(如:对象合并、语义统一等),将最终元对象归入到元模型二极分类中,形成元对象模型总体架构

- ✓ 模型规范设计

表命名规范: 1、表名称采用全小写 2、实体表命名规则: mdl\_{一级分类}\_{二级分类}\_{三级分类}\_{元数据类型}\_{表名}; 示例: 贴源层元数据表, mdl\_se\_cam\_ods\_adherentlayer table。

字段命名规范: 1、字段名称采用全小写, 下划线分割, 字段命名能够清晰描述其含义, 如: 项目空间名称, project\_name; 表名, table\_name; 最后更新时间, last\_modify\_t

## 元模型融合设计、开发

条款 3) 组织与少量外部机构实现元数据采集、共享、交换和应用。

解读：组织与外部机构（同行业、监管机构、跨行业等组织）合作开展的数据服务产品案例，体现元数据整个全生命周期双方共同参与（需求、接口、应用等）。

API 清单和规范					
API 清单列表:					
接口功能	接口 URL		操作方式		
规划总览数据查询	/gJf	laning-			
规划上报提交		/planingOverview/leaderview/taskId	GET		
规划审核		sps-plan/api/v1/planAudit/submit	POST		
5G 建设调度大屏新增站点数据查询	/g	anAudit/countersign/approve	POST		
区域详情数据查询	/gJgx	consti			
基站详情数据查询	/gJgx/ap	construct.	GET		

与少量外部机构实现元数据采集、共享、交换和应用

### 级别 5：优化级

- 条款 1) 参与国际、国家或行业相关元数据管理相关标准制定;
- 条款 2) 参与国际、国家、行业的元数据采集、共享、交换和应用;
- 条款 3) 在业界分享最佳实践,成为行业标杆。