





5.数据安全确保数据隐私和机密性得到保护，数据不被破坏，数据被适当访问。

6.数据集成和互操作性包括与数据存储、应用程序和组织内部之间的数据移动和整合相关的过程。

7.文档和内容包括计划、实施和控制活动，用于管理在一系列非结构化媒体中发现的数据和信息的生命周期，特别是支持法律和法规遵从性要求所需的文档。

8.参考数据和主数据包括核心关键共享数据的持续核对和维护，使系统能够一致地使用关于基本业务实体的最准确、最及时和最相关版本的真相。

9.数据仓库和商务智能包括计划、实施和控制过程，以管理决策支持数据，并使知识工作者能够通过分析和报告从数据中获取价值。

10.元数据包括规划、实施和控制活动，以便能够访问高质量的综合元数据，包括定义、模型、数据流和其他对理解数据及其创建、维护系统至关重要的信息并访问。

11.数据质量包括质量管理技术的计划和实施，以测量、评估和提高数据在组织内的适用性。

根据DAMA的说法，数据质量包括质量管理技术的计划和实施，以测量、评估和提高数据在组织内的适用性。

实际上，数据质量在大多数其他数据管理领域中都发挥着作用。要保证数据安全，就不可能没有好的数据质量，元数据与数据质量的关系也是双向的，数据架构也会对数据的质量起到一定作用，反之亦然。处于中心地位的数据治理确保所有这些数据管理领域结合在一起。

## 2 无论治不治理，数据质量都在那里

到现在为止，你可能会说：“好吧，我有点明白了，但我们没有做数据治理，数据质量也挺好。”如果是这样的话，我认为有两种可能的现实：

•真的没有数据治理

•有隐形的数据治理

让我们更详细地看一下这两个案例。

### 2.1 真的没有数据治理

如果数据治理不存在，那么我们可能会遇到以下一种、部分或全部情况：

•**数据质量不是企业级全域的**：你也可能认为是覆盖企业范围的，比如ERP，这并不意味着数据质量是企业范围的。尤其是在像我们这样的大型组织中，因为可能存在数据质量规则，这些规则仅根据该企业系统的利益相关者的输入创建，但这可能会影响那些不是系统利益相关者或用户的人，比如有集成相关的系统用户，也可能是数仓或者下游数据用户。

•**缺乏通用标准**：可能会创建数据质量标准，但前提是它们与特定部门、业务单位或系统的需求有关。一旦其他部门或系统加入到数据质量计划的工作中，这些标准就会发生冲突，需要解决方法或完全更改。

•**没有明确的角色和职责**：可能会分配负责清洗和维护数据的资源，但资源类型因团队而异。当标准、定义和优先级发生冲突时，也可能不清楚谁拥有所有权，谁负责解决。

•**数据质量管理大多是被动的**：这主要是数据质量管理处于早期阶段，也意味着缺乏数据治理。数据质量问题是被动的方式识别和处理的，并不总是解决问题的根本原因。

### 2.2 有隐形的数据治理

就目前我们公司的现状开看，作为产研、数仓、数据集市、数据分析等团队成员，你实际上可能会说：“好吧，我们并没有像你上面描述的那样糟糕”。您实际上可能有：

•数据质量要求

•数据质量校验规则

•数据质量测试，甚至是监控

•数据质量问题处理

那么，你可能就有了我喜欢称之为“卧底数据治理”的东西。你可能有很多数据治理的工作内容，但没有正式、正规和专业化的数据治理。

## 3 数据质量和数据治理关系

我想我已经从上一节中明确指出，数据治理和数据质量是共生关系。它们是同一枚硬币的两面。**没有数据治理，就不可能有良好的数据质量，而数据治理实施必须非常有效，才能从根本上解决数据质量问题。**

实际上，数据质量和数据治理之间存在相当多的重叠，如下图所示：





数据治理描述了谁需要做什么、在什么条件下对什么数据执行操作，以及使用哪些流程、程序、工具和总体最佳实践。因此，很多因素都会对数据质量产生有益的影响，但不仅如此。业务规则、数据标准、指标、角色和职责等都有利于数据质量，因此是重叠的，但不仅仅是数据质量。主数据管理、数据可访问性、数据集成、元数据管理、BI，甚至数据安全等等都有直接的好处。

当然，也有一些领域只与数据质量有关，例如：数据概要分析、数据匹配、根因分析和数据清洗。

## 4 总结

很多时候，数据质量是数据治理的驱动因素之一，这是数据治理计划的最初重点，因此两者之间可能会混淆。但它们并不相同，它们是同一枚硬币的两面，而不是两枚硬币。

作者： 京东科技 李然辉

来源： 京东云开发者社区 转载请注明来源

标签： 数据分析

### 评论 2

- 登录 / 注册

即可发布评论!
- 最热 | 最新
- 小Tomkk 大数据讲师

京东科技就在上海北外滩来福士

11月前 | 点赞 评论

...
- 小Tomkk 大数据讲师

真不错

12月前 | 点赞 评论

...

### 为你推荐

- 如何理解数据治理的一点见解

Ros1n | 3年前 | 1.9k 6 2

大数据 后端
- 数据治理：说起来容易，做起来难？这个方法论值得收藏

IT技术管理的那些事儿 | 3年前 | 2.2k 3 评论

数据分析
- 数据治理流程中，最重要的3点都在这

马哥说数据 | 3年前 | 1.4k 点赞 评论

大数据
- 数据治理怎么做？这篇万字长文终于讲清楚了！

云祁 | 3年前 | 2.0k 6 评论

后端
- 火山引擎DataLeap一站式数据治理解决方案及平台架构

字节跳动数据平台 | 1年前 | 707 3 1

大数据 数据库 数据分析
- 数据分析前端向的业务理解

VioletJack | 10月前 | 153 点赞 评论

数据分析 数据可视化
- 数据类型错误带来严重后果！一文读懂怎么正确选择数据类型

Squids数据库云服务... | 1年前 | 517 点赞 评论

大数据 数据库 MySQL
- 京东大数据治理探索与实践 | 京东零售技术实践

京东零售技术 | 2月前 | 96 2 评论

大数据
- 计算性能提升百倍 火山引擎数智平台这款产品助力企业员工更好看数用数

字节跳动数据平台 | 2年前 | 11k 60 评论

数据库 数据挖掘 运营
- 基于主动元数据 构建智能数据治理体系

京东零售技术 | 3月前 | 73 点赞 评论

算法
- ELT已死，EtLT才是现代数据处理架构的终点!

白鲸开源 | 1年前 | 1.7k 4 评论

大数据



### 数据治理怎么做？这篇万字长文终于讲清楚了！

云祁 | 3年前 | 201 | 1 | 评论

后端

### 大数据之数据治理体系全面指南

五分钟学大数据 | 8月前 | 1.2k | 2 | 评论

大数据

### 数据治理设计理论

LightGao | 10月前 | 819 | 4 | 1

数据库

### 治理要从数据建模抓起 | 京东云技术团队

京东云开发者 | 1年前 | 3.2k | 2 | 评论

数据库数据挖掘