

# 好问题，数据治理到底解决了什么啊

原创 唐晨 唐晨说数 2025年03月30日 20:42 浙江

Hi，你好呀，见字如面，我是唐晨，本号主要更新大数据领域知识和个人观点，关注 Data + AI 领域的小伙伴，欢迎关注公众号交流互动。



唐晨说数

大数据领域从业者，分享关于数据治理、数据中台、数据应用等领域相关的，个人实践及...  
171篇原创内容

公众号

目前，我正在体系化更新数据资源管理/数据治理/数据中台系列文章。

看看都是哪些角色在订阅唐晨说数，点击下方选项查看：

请在手机微信登录投票

看看谁在看唐晨说数 单选

企业高管/CIO

数据负责人/总监/经理/主管

数据开发/数据工程师

数据/业务分析师/算法工程师

前后端/运维工程师

市场/销售/产品/运维/自媒体

其他

“唐老师，你说我们数据治理搞来搞去，到底在解决什么问题，最近工作比较迷茫。”

“好问题，属于又基础，又复杂的一个问题。”

大数据&AI 领域，从来不缺“新概念”、“新技术”，很多概念一闪而过，像时光一样。

穿过我们的耳膜，也穿过我们的脑袋，扬长而去~和我们相关，又好像无关。

以至于，很多概念，我们听过，甚至研究过一段时间，但是，我们不知道它为何存在，为何被业内共识，以及，到底为了解决什么问题。

就比如数据治理。

这绝对，不是一个新的词汇，甚至很多读者的工作就和数据治理息息相关。

今天的文章，我们回到原点，聊一聊：[数据治理，到底解决了什么问题，为什么要搞数据治理？](#)

也许你读到这里，心里就嘀咕了：数据质量不好呗，否则也没必要治理了。

如果你有类似心理活动，说明你的日常工作肯定是和数据打交道的。

但是，一句数据质量差并不能诠释，为什么企业需要搞数据治理。

如此，不妨花点时间，一起阅读一下这篇文章。

## 01

### 原点：企业数据之痛

---

每家企业的数字化进程并不同步，因此，遇到的数据问题也不尽相同。

不同之中，是否有共同点呢？

当然，几乎每家企业，在进入数字化深水区的时候，都会遇到各种各样的数据使用卡点问题。

比如数据质量、数据不一致、数据用不了等等。

导致这一切的诱因千差万别，不过都有一个共同的特点：前期不管不问，后期修修补补。

比如某银行在多个系统（如信用卡系统、理财系统、网银系统）中，一个客户存在多个ID，无法统一识别是否是同一个人。

客户 A 在信用卡系统是“李雷”，在理财系统是“李磊”，在CRM中是“LILEI123”，这就导致营销系统把 A 识别为三个客户，重复推送营销短信，且内部三个营销人员给客户打电话推销相同的活动，导致客户体验极差。

再比如某证券公司需要向证监会定期报送数十类监管数据，数据来自多个业务系统，报送口径不一致，经常出现字段缺失、逻辑错误、日期不符等问题，被多次点名，甚至收到罚单。

类似的案例，在各行各业都存在，只是换成了不同的业务场景而已。

## 02

### 那么，具体有什么问题

---

随着企业业务规模的扩展，数据量急速增长，企业内部的数据问题终将爆发。

最常见的问题为：数据使用率过低（ROI太低）、数据呈现的结论反常识、存在数据安全泄露等问题。

企业面临的数据问题，可以简单归类为：[不可信、找不到、看不懂、管不住](#)。

图片由 ChatGPT 生成

## 2.1 不可信

数据质量问题就是企业中最常见的“不可信”表现。

像数据重复、缺失、异常值或过期数据等问题频繁发生，让企业的分析成果大打折扣，也使得业务人员对数据逐渐失去信心。

比如小王负责某次运营活动的数据分析，他满怀信心地根据营销平台提供的数据，得出活动转化率高达60%的结论。

谨慎起见，他对数据进一步做了核对，发现平台的数据竟然出现重复计算，一位用户竟然被计算了多次。

原本欢天喜地准备汇报业绩的小王，瞬间陷入尴尬，还好他提前发现，如果被大老板提出异议，那么问题就严重了。

## 2.2 找不到

找不到所需的数据，不知道数据存储在哪里。

“我们天天喊着数据驱动，可是数据在哪里我都不知道，还不是靠业务经验”，这是一个潜在客户给我们吐槽自己无数可用的窘境。

对于很多传统企业来说，企业的数据经常分散在各个业务线、部门，甚至个人的电脑里。

普遍缺乏统一的索引和目录，导致员工找数据时只能靠“猜”，工作效率自然不高，更别说数据驱动了。

### 2.3 看不懂

除了找不到数据、数据反直觉外，还有一个普遍存在的问题：数据看不懂。

很多企业的资产数据，存在大量的元数据缺失的问题，尤其是管理元数据和业务元数据。

这就导致业务人员看不懂数据，有数据诉求只能逐个去找数据人员确认，这同样带来很大的工作降效。

更有甚至，业务人员直接不问了，直接去投诉说数据缺失，一口大锅狠狠地盖在了数据部门。

不过，数据部门也不无辜。

许多企业在数据创建时，没有明确的数据字典和规范，造成业务含义不明晰、命名不规范，让新手和跨部门同事在理解数据时头疼不已。

### 2.4 管不住

数据分层嵌套，成本居高不下，造成大量成本浪费。

IT负责人张总发现，公司数据分层越建越多，数据源互相嵌套重复存储，成本一年竟然上涨了40%。

更可怕的是，新增的数据越来越难维护，光是数据维护和清理，就得再招几个人，预算一算，吓出一身冷汗。

数据治理不到位，导致数据无限制复制、存储冗余严重，企业看似丰富的数据资产，实则隐藏着大量成本浪费的问题，治理难度逐年加大，严重影响企业的数字化效益。

## 03

### 未完

正常到此，都要写一个小结的。

然后，今天不同，今天分明是未完。

因为，从上面可以比较清晰地感知到“企业为什么需要做数据治理”、“数据治理主要解决什么问题”。

但是，这些问题应该如何处理，数据治理应该如何做，从哪些方面入手，这些偏向于进一步落地的问题，本文并未讲解。

让我们后续继续接着写。

欢迎关注，哦，顺便给公众号标星哦！



**唐晨说数**

大数据领域从业者，分享关于数据治理、数据中台、数据应用等领域相关的，个人实践及...  
171篇原创内容

---

公众号

(正文完)

---

今天的分享就到这里，希望对你能有多帮助和启发。

我是唐晨，日更大数据+AI 领域知识和个人观点，关注 Data + AI 领域的小伙伴，可以点击关注公众号。



**唐晨说数**

大数据领域从业者，分享关于数据治理、数据中台、数据应用等领域相关的，个人实践及...  
171篇原创内容

---

公众号

---

很多人问我有没有交流群

马上安排上

添加下方微信拉你入群

转载、投稿、商务、咨询合作均可添加微信沟通

现在文字真的越来越少用户有耐心阅读完了，后期是需要经营视频号了，欢迎关注，到 500 了启动直播，线上聊一聊。

唐晨说数

9



唐晨

“ 欢喜，值得付费 ”

喜欢作者

