



# 数据治理实践

---

任增涛 百信银行 数据治理架构师



# 目录 CONTENT

**01** 数据治理概览

**03** 数据治理组织架构

**02** 数据治理痛点

**04** 数据治理落地

# 01 数据治理概览



## 国家及监管动态分享

数字化时代下金融科技迅速发展，新金融环境与大量复杂金融交易产生的海量数据，对传统金融机构监管数据治理带来前所未有的挑战，监管机构今年纷纷发布数据治理相关要求，对金融机构监管愈发严格，处罚层级与力度不断增强

2018年5月

银保监会

## 《银行业金融机构数据治理指引》

首次将数据治理与监管评级挂钩，  
标志着数据治理在我国银行业全  
面实践时代的到来

2019年8月

人民银行

**《金融科技发展规划(2019-2021年) 》**

规划指出要加强金融科技审慎，建立健全监管基本规则体系，增强金融监管的专业性、统一性和穿透性

2020年1月

## 某农村商业银行因数据治理问题被罚

监管首次针对数据治理开出罚单，中小银行因数据治理问题被监管开出罚单，意味着金融机构数据治理越来越受到监管的重视

2020年5月

银保监会

## 《关于开展监管数据质量专项治理工作的通知》

要求各金融机构开展为期一年监管数据质量专项治理工作，开展自评自查，并向监管机构提交自查报告

2020年6月

银保监会

## 《关于印发保险业监管数据标准化规范（人身保险公司版）的通知》

规范要求各人身保险公司于2021年1月15日前按要求完成2017-2020年全量标准化监管数据报送,并对报送数据一致性、真实性、

2017年4月

保监会

## 《关于加强中国保险业保单 登记管理信息平台数据报送 和质量管理工作的通知》

要求各保险机构落实数据报送职责，建立数据报送和质量管理责任机制，加强数据源头管理

## 2019年5月

银保监会

## 《中国保险业保单登记管理信息平台数据治理情况的通报》

指出部分保险公司在数据报送中  
存在数据报送不及时、不完整、  
不准确，数据整改不积极、效率  
低等问题，要求各保险机构及时整  
改，完善数据报送监控和数据质  
量检查机

2019年12月

银保监会

## 《中国银保监会现场检查办法(试行)》

要求各金融机构应当加强数据治理，完成检查分析系统所需数据整理、报送等工作，保证相关数据的全面、真实、准确、规范和及时。

2020年5月

## 银保监会首次就EAST报送开具罚单

8家银行合计遭罚1770万元，本次处罚可以视为《指引》中要求的“不满足要求的金融机构可依法采取监管措施及实施行政处罚”的实际执行

2020年5月

银保监会

## 关于《配合做好监管标准化数据（EAST）质量稽核调查相关工作的通知》

要求21家全国性银行明确牵头部门，完成管理及运用情况调研，并做好数据准备

# 监管要求概览

2018年5月21日，银保监会印发《银行业金融机构数据治理指引的通知》，《银行监管统计数据质量管理良好标准(试行)》正式废止。指引将数据治理纳入公司治理范畴，将数据治理情况与公司治理评价和监管评级挂钩



## 治理架构:

银行业金融机构应当建立组织架构健全、职责边界清晰的数据治理架构

## 治理机制:

银行业金融机构应当制定全面科学有效的数据管理制度、确立数据质量管理目标并建立质量控制机制、建立数据质量监控体系和数据质量考核评价体系、建立数据治理问责机制、建立数据治理自我评估机制、建立数据质量整改制度

**第一章 总则(7条)**

**第二章 数据治理架构(9条)**

**第三章 数据管理(12条)**

**第四章 数据质量控制(9条)**

**第五章 数据价值实现(12条)**

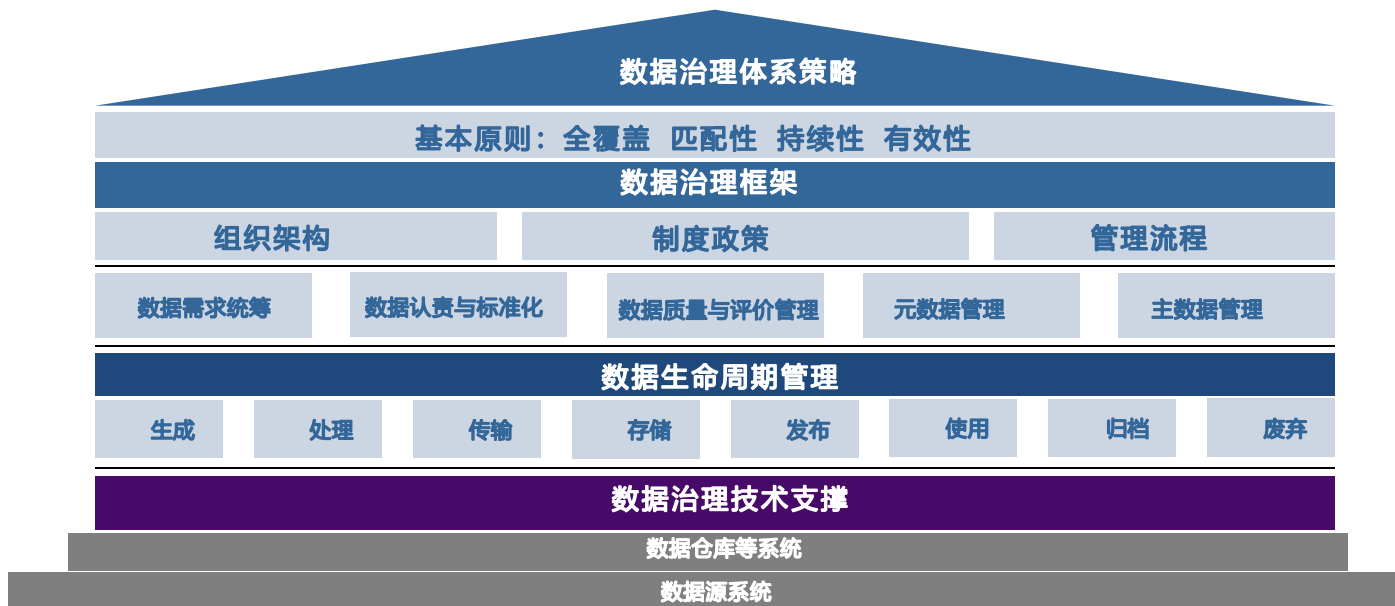
**第六章 监督管理(3条)**

**第七章 附则(3条)**

**指引全文共计7章55条**

# 数据治理体系

数据治理是指银行业金融机构通过建立组织架构，明确董事会、监事会、高级管理层及内设部门等职责要求，制定和实施系统化的制度、流程和方法，确保数据统一管理、高效运行，并在经营管理中充分发挥价值的动态过程。



## 02 数据治理痛点





# 推动难

数据治理范围大、治理领域广、涉及部门多，造成在实际治理过程中推动难



## 范围大

数据治理涉及数据范围广，主要包括业务主题基础数据、指标数据、特征数据、标签数据

## 领域广

涉及含数据标准、数据质量、数据安全、主数据、数据生命周期、数据资产等领域

## 部门多

需全行各部门协同配合，各管理部门、各业务部门及各科技部门



# 落地难

数据治理各领落地过程中都遇到各种各样的挑战，造成数据治理落地难



## 数据标准落地难

由于业务系统的频繁变动和新系统的上线，如果要在流程中进行人工审批，需要大量的人力物力，往往处于疲于奔命的状态

## 数据质量整改难

推动数据质量问题整改经常被诟病是否有价值，尤其是在解决业务主题基础数据质量问题过程中尤为突出

## 数据安全管控难

数据安全经常受到强势业务部门的挑战，通常以影响业绩为由对管理部门施压，造成安全原则难坚守

# 数据治理 “两张皮”

数据治理往往是组织、流程、制度很完善，但是落地效果不佳或完全未落地



## 组织、流程、制度

组织，流程制度、管理、绩效考核  
方案非常健全



## 推动、落地、工具

组织不能推动、流程制度不能落地、  
工具不健全

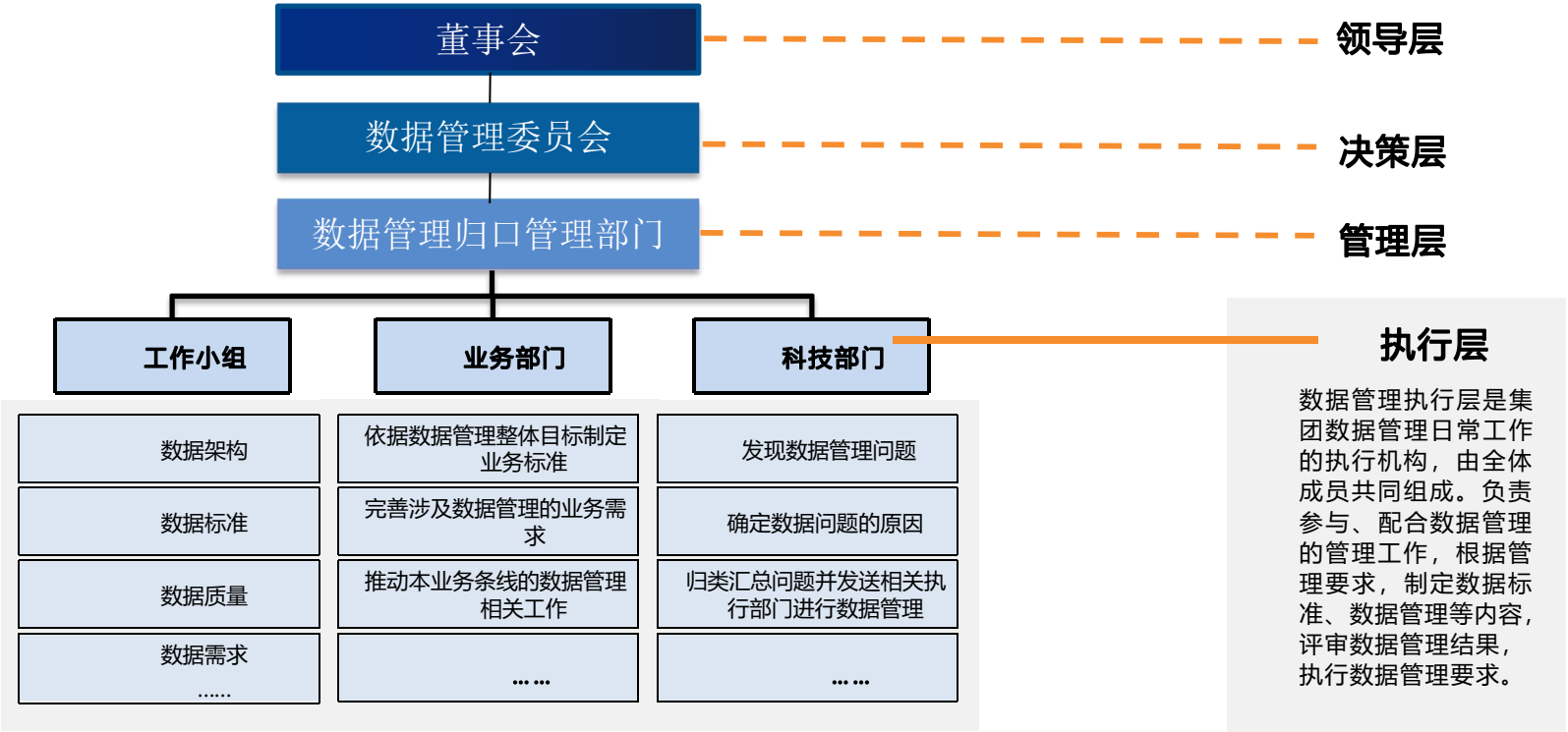


## 03 数据治理组织架构



# 数据治理组织架构建设

数据治理组织总体上分为领导层、决策层、管理层、执行层从上而下推动数据治理的管理和落地



# 组织推动

在数据治理各领域工作开展过程中，想要做到数据治理真正落地，必须依赖于组织的力量，获取高层支持，才能做到数据治理能推动、可落地。

## 01

### 数据安全

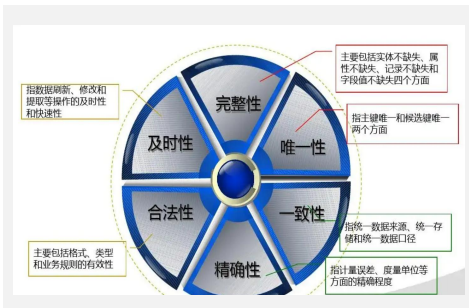
业务部门违反《数据安全法》或监管制度，管理部门受到强势业务部门的挑战，通常需要有更高层组织进行审议



## 02

### 数据质量

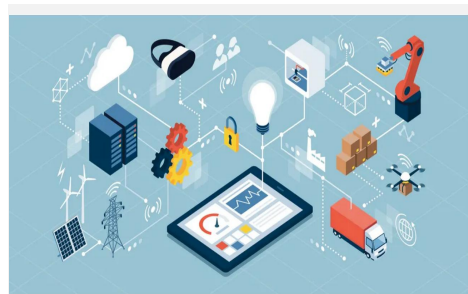
推动业务系统数据质量问题整改，难以体现价值，业务和科技部门配合度低，需借助强需求向上争取管理力度



## 03

### 数据标准

数据标准前置管控以牺牲开发效率为代价，需要向更高层组织讲请价值，获取管理力度。



## 总结

- 数据治理组织的核心是如何向上获取管理力度，否则数据治理很容易陷入“两张皮”的窘境
  - 如果管理力度较弱，建议以价值驱动进行数据治理落地，如指标数据治理
- 



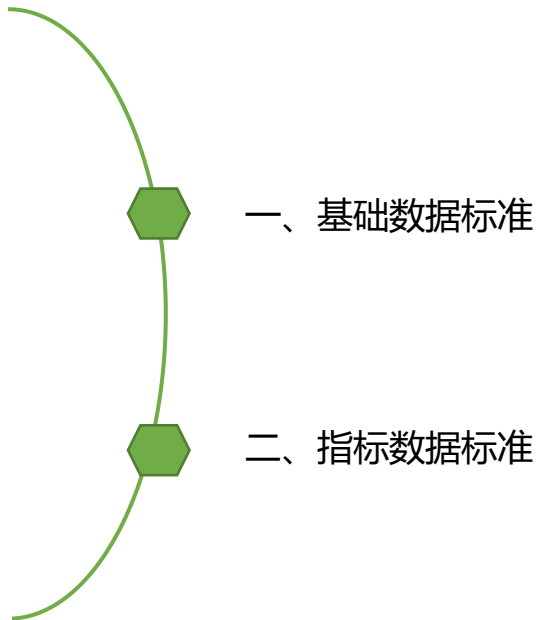


## 04 数据治理落地





# 数据标准落地-分类和痛点



1. 如何制定适合企业发展的数据标准?
2. 数据标准检查工作量大, 如果确保落地?
3. 动辄几十万字段, 有没有必要全部按照标准执行?

1. 指标是否有必要全部管理?
2. 是否有必要对数据仓库进行管理?
3. 如何确保指标质量?

## 数据标准落地-业务主题基础数据-标准制定

- 主要依据行业规范进行数据标准的制定，如综合east、反洗钱、金融数据安全分级指南、个人信息技术保护规范等相关规范
- 基于自然语言处理技术，遍历全行元数据（字段编码和备注），生成数据项（粗版数据标准）和字段的映射管关系，用于辅助数据标准的制定，以自动化工具为辅、人工制定为主，大幅提升数据资产盘点能力。



# 数据标准落地-业务主题基础数据-资产盘点

建立字段和数据标准映射关系，可根据字段编码、字段备注、字段内容中任意一项识别字段归属的数据标准，为后续精细化管理打造核心能力。以自动化工具为主、人工识别为辅，大幅提升数据资产盘点能力。

## 数据标准项

名称	编码
类型	枚举值
是否个人金融信息等级	金融信息等级
是否加密	质量规则
是否行级	其他

## 发现涉及标准项字段



基于  
自然语言的  
治理引擎

## 打通字段和标准项的关系

表一

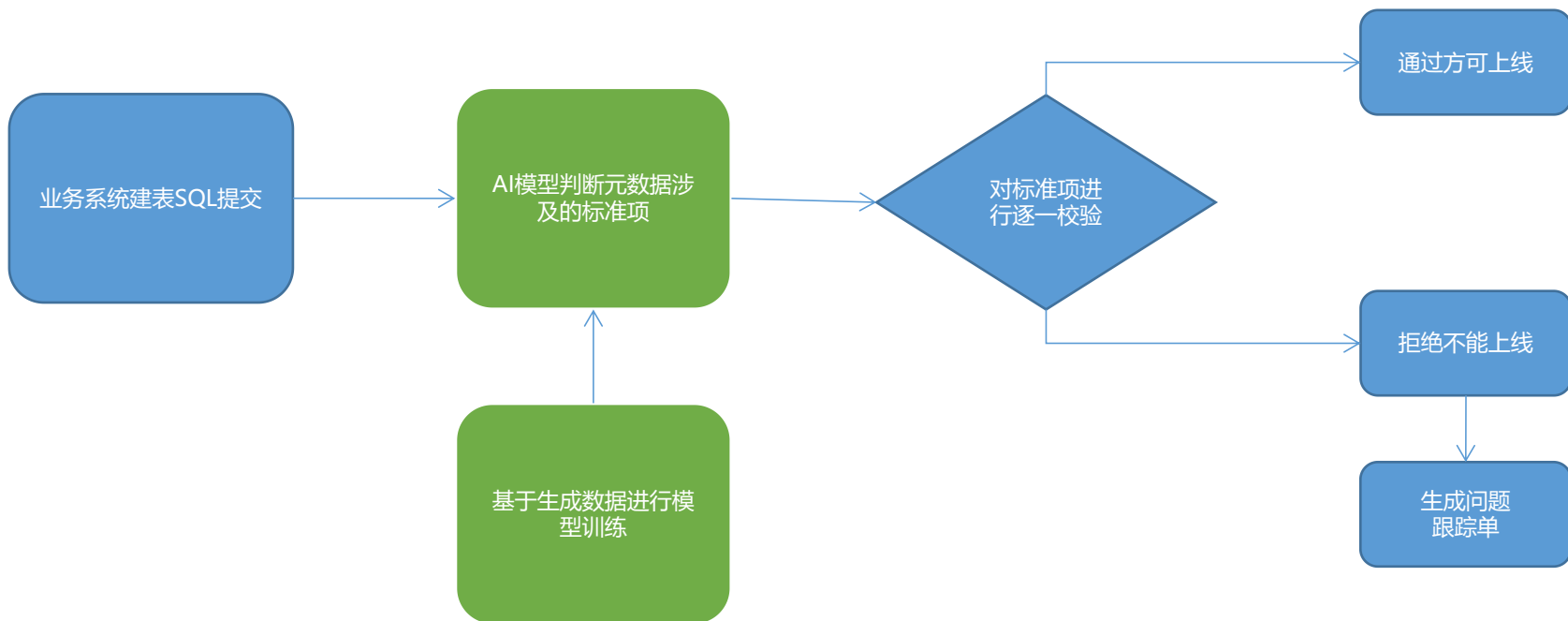
字段	数据标准项
cust_id	客户名称
idnumber	身份证号
bankcard	银行卡号

表二

字段	数据标准项
customer_id	客户名称
id_number	身份证号
card	银行卡号

# 数据标准落地-业务主题基础数据-前置管控

基于自然语言处理技术开发SQL检查工具，设立数据标准前置检查点，自动校验建表和修改表语句，发现问题并要求整改，杜绝不符合数据标准的情况发生，大幅提升数据标准检查效率。



# 数据标准落地-指标数据-痛点

## 质量堪忧

在实际业务场景中，经常出现同一指标口径不一致，同一指标内容不一致、指标产出延迟、指标准确性差等一系列问题，造成管理者、运营分析人员对指标可信度质疑。

## 开发周期长

指标开发周期往往不能满足业务需要，如在报表中增加一个指标或维度，往往需要完成一个完整的需求开发流程，很难支持企业快速决策

## 指标管理痛点

## 复用率低

每张报表的每个指标保存在不同结果表中，即同一指标为满足不同场景，需要开发多次，浪费开发资源的同时，也增大了管理难度。

## 管理难度高

指标往往分布于数据仓库、数据集市、报表平台等多个系统中，很难保证同一指标在多个系统中的一致和准确

# 数据标准落地-指标数据-架构



## 指标平台

指标平台集成了【指标管理】、【维度管理】、【模型管理】、【OLAP分析】、【固化报表】等功能

对数仓中的表进行建模管理，依托模型进行【指标】、【维度】创建，统一口径和名称，将多元的表存储和设计抽象为模块化的字段管理，通过对字段级别的管理和加工，赋能上层应用

用户自由拖拽【指标】、【维度】即可完成即席查询

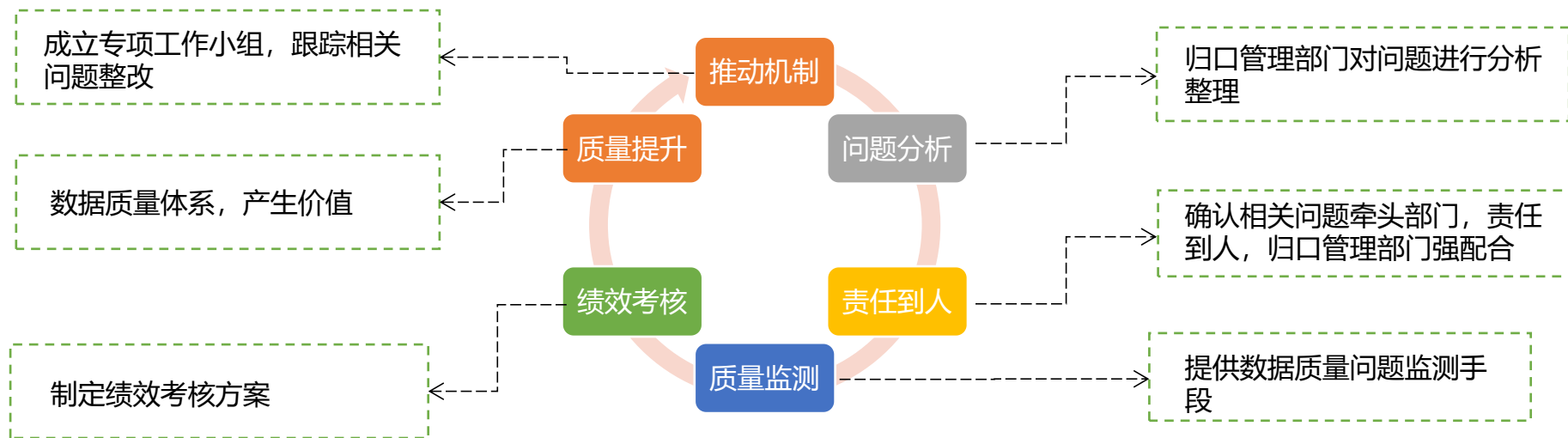
- 1.新需求：加工新指标即可
- 2.更改指标：同步更新模型中所有对应字段

无论是加工效率亦或是维护成本均大幅度提升，应用层的报表、看板、归因分析、移动端数据口径绝对统一，保障用户数据的一致性

# 数据质量落地

**监管政策** 在《银行业金融数据治理指引》中明确要求监管数据应建立工作机制和流程，确保监管数据报送工作有效组织开展，监管数据质量持续提升，法定代表人或主要负责人对监管数据质量承担最终责任

**核心思路** 依托数据治理委员会推动解决，责任到人，推动数据质量问题的整改





# 数据安全落地

依托于数据标准项和字段的映射关系，以数据标准项为核心对采取数据安全管理工作



# 非常感谢您的观看

---

 DataFun.

