



——大数据与安全技术沙龙-长沙站——

湖南移动数据标准化实践

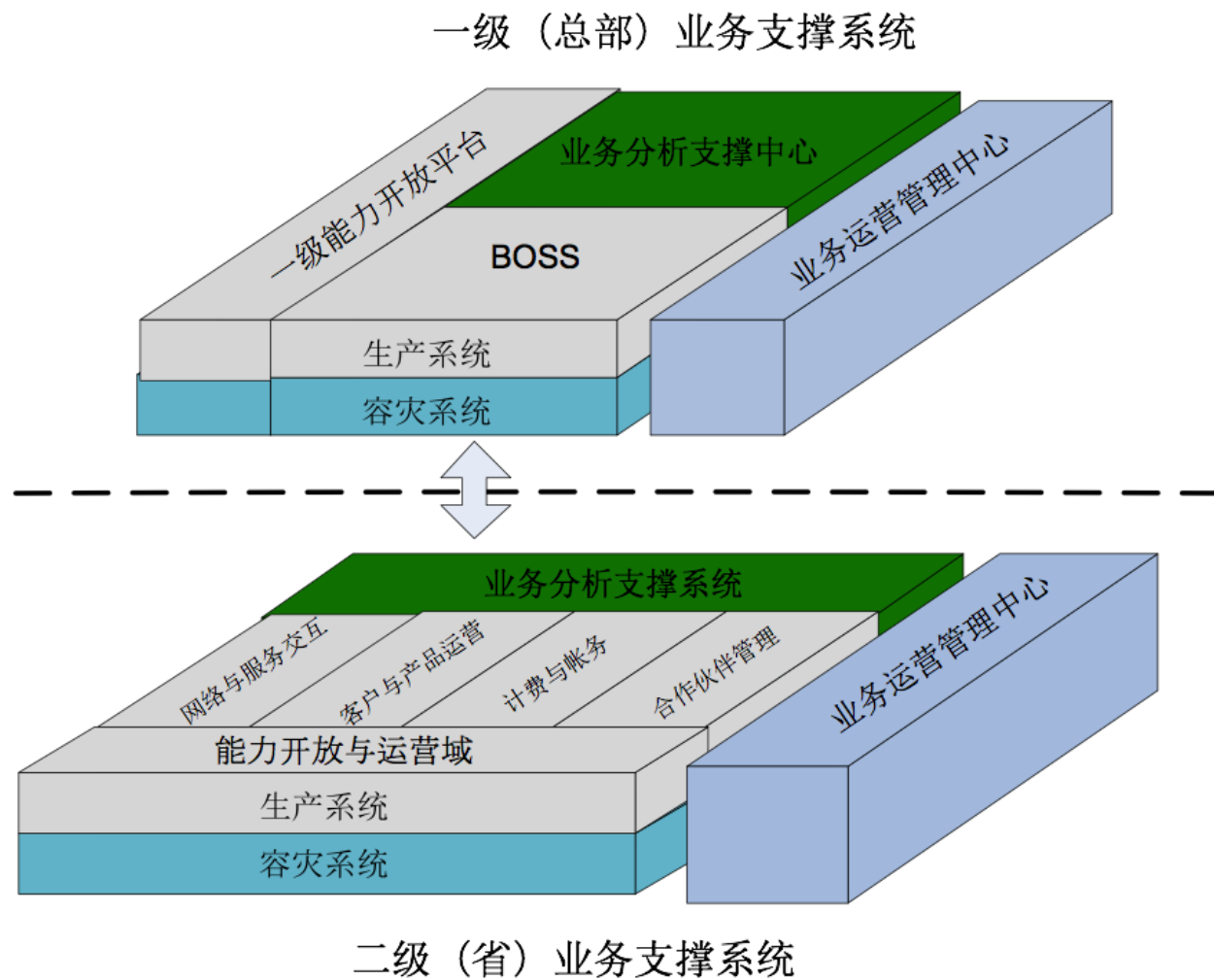
郭远胜

PART 1

中国移动业务支撑系统介绍

中国移动的业务支撑网体系架构采用两级架构，即一级业务支撑系统、省级业务支撑系统两级，一级业务支撑系统归属于有限公司，省级业务支撑系统归属于省公司。

在业务支撑网的两级架构中，一级业务支撑系统具有管理、实体和枢纽功能，为有限公司进行全网业务管理和业务运营提供支撑和保障，实现全网信息的交换和管理。省级业务支撑系统具有管理和实体功能，为省公司进行省内业务管理和业务运营提供支撑和保障。省级业务支撑系统与一级业务支撑系统通过接口相联，实现数据交换，二者相互协作与配合，共同支撑移动业务的运营与管理。



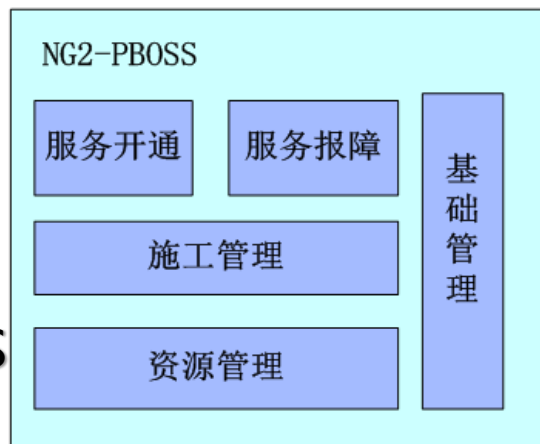
中国移动业务支撑网体系结构

CRM



以统一的CRM承载融合的全业务、全产品客户运营，实现全渠道一体化的营销、服务、业务受理。

宽带 P-BOSS



专业化、高起点的宽带P-BOSS弥补中国移动在有线业务长流程上的短板。



全业务运营

NG2-BOSS



BOSS

以强大的融合计费引擎和灵活的帐务配置支撑融合的全业务产品设计和实现。

NG2-经营分析系统

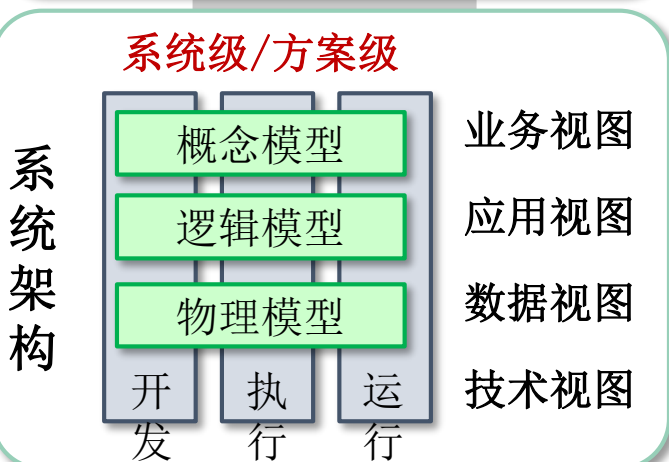
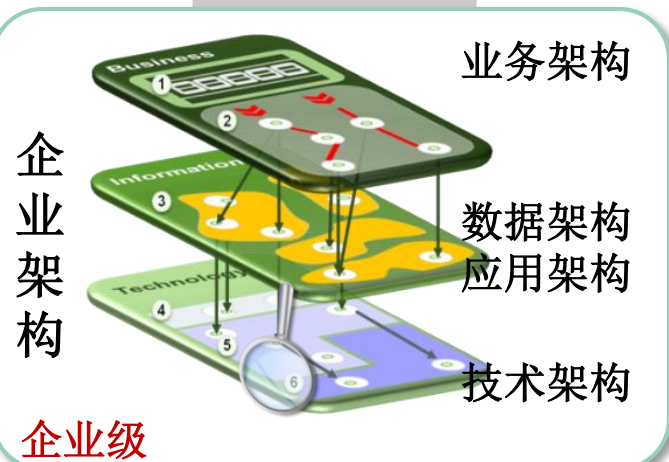


经分系统

通过数据分析挖掘，了解客户和市场，实现精确营销、精细管理。

架构的理念应用于企业信息化建设，演化为企业架构和系统架构，伴随业务的快速发展和各种新技术的引入，企业IT系统复杂程度高，具有数据量大、数据种类多、数据处理的复杂性高等特点。以往IT系统的建设主要关注业务架构、应用架构和技术架构，而现在还必须重视数据架构的管理。

细化落地



- ◆数据架构是**业务需求**、**业务规则**的重要体现，是新的系统开发建设的重要指引。
- ◆数据架构中的概念视图可以为项目建设快速的提供**范围定义**。
- ◆数据架构中的模型标准、业务术语标准为项目的规划、设计提供了**标准准则**。
- ◆数据架构为新的系统建设中的**数据集成**提供规范和要求。
- ◆数据架构为新的系统建设中的数据规划、设计、管理提供相关的**管理方法、模板和工具**。

中国移动数据架构问题现状

数据结构、数据标准、管理体系、数据共享性等多种因素的不完善，导致系统整体性能和数据质量越来越差；数据模型问题导致的后果：1、全网业务开展困难，系统改造工作量大；2、影响大数据分析的数据质量。

集团视角
省公司视角

电商模式等新业务的发展
导致信息模型复杂



各省数据标准不统一，信息模
型存在差异



全网数据资产无法集中管理和展
现，共享性差



欠缺数据模型审核机制



- 管理体系不完善
- 已有标准落实差
- 管理角色分工不清晰

数据模型变为黑匣子



- 缺乏数据标准：表和字段命名定义混乱
- 历史数据庞大复杂：表结构相似，表间关系不清晰
- 扩展性差：增加新业务时就需要增加相应的表，数据处理变得越来越复杂

沟通成本高理解难度大



- 数据模型复杂，业务人员难以理解
- 表结构变更、系统改造时，对应用造成的影响难于甚至无法评估

PART 2

数据标准化刻不容缓

数据标准化是对分散在各系统中的数据提供一套统一的数据命名、数据定义、数据类型、赋值规则等的定义基准，并且提供运维这套数据标准成果的科学流程。

数据标准化管理包括标准单词、标准域、标准命名规则、标准用语、标准编码的新建、修改、删除、停用、申请审批流程等。



标准单词

- 具有业务含义的最小单位
- 构成实体、属性、字段名的最小单位



标准域

- 对于表示一般特征的数据进行分类，限制属性值的数据类型和数据长度、数据取值
- 明确用语含义和属性值范围



标准用语

- 将业务中使用的用语转换为信息系统中使用的技术用语，使其保持一致性，定义标准
- 由一个以上的标准单词和一个域组成

IT系统中数据要素的命名混乱、定义不清，造成一系列沟通障碍和关联错误，并最终导致IT系统开发和运维的成本升高。只有制定数据标准，规范数据要素的命名、定义、取值类型，实现数据标准化，才能彻底解决上述问题。

确保表达精确、沟通顺畅

- 两个表示相同意义的同义词，只选用一个作为标准单词
- 定义标准名称，预防数据的重复定义，确保沟通顺畅
- 明确单词之间的定义区分，确保理解一致，提高沟通效

提高业务处理效率

- 通过统一的用语定义，提高同一单词的重复使用率
- 使用标准用语，提高数据查询等业务效率
- 在命名新用语时，缩短人工选择最合适单词的时间

提升数据质量

- 通过标准化，容易检验数据结构、数据流等的准确性，提高数据质量
- 使用标准用语，保持数据模型中对象用语名称的一致性
- 明确数据的特征，提高数据的可信度

- 实现企业级大数据平台省级系统数据治理过程中数据的标准、质量、价值、管控的有效性、高效性

有效性

- 建设架构统一、标准统一、元数据统一、质量流程统一、资产价值统一的一体化管控平台体系

统一性

- 体现平台化、开放性运维思想，实现人人参与数据治理、人人参与数据运维的理念

开放性

- 体现安全的重要性、必要性，保障数据治理过程中数据的安全可控

安全性



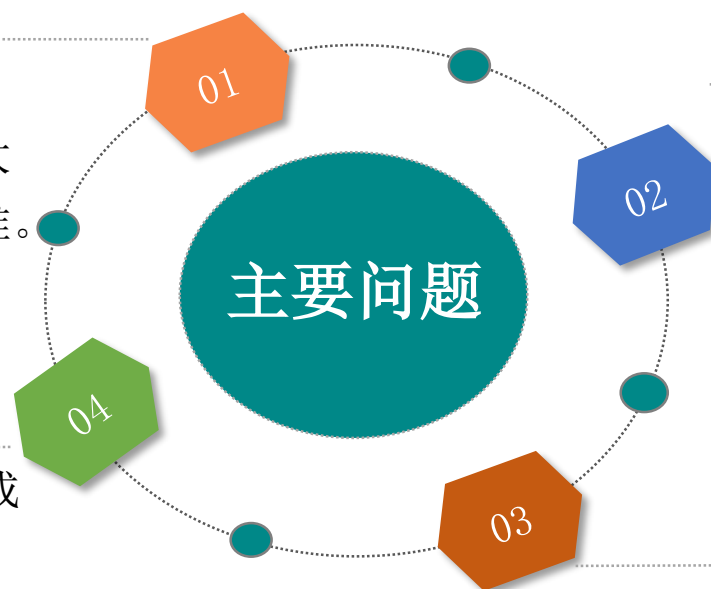
系统越来越多，越来越复杂，数据模型流向失控，如何生成及维护一套数据命名、数据定义、数据类型的数据标准，并有效地进行推行，维持数据源模型的统一性，提升数据正确性和数据质量，成为一个急需解决的难题。

缺乏统一标准

- 开发商各自为政，自建标准，缺乏适用于移动大数据建设的统一数据标准。

风险不可控

- 模型变更时，对应用造成的影响无法有效评估。
- 对应用开发造成难度，且易引发故障。



无法有效管控

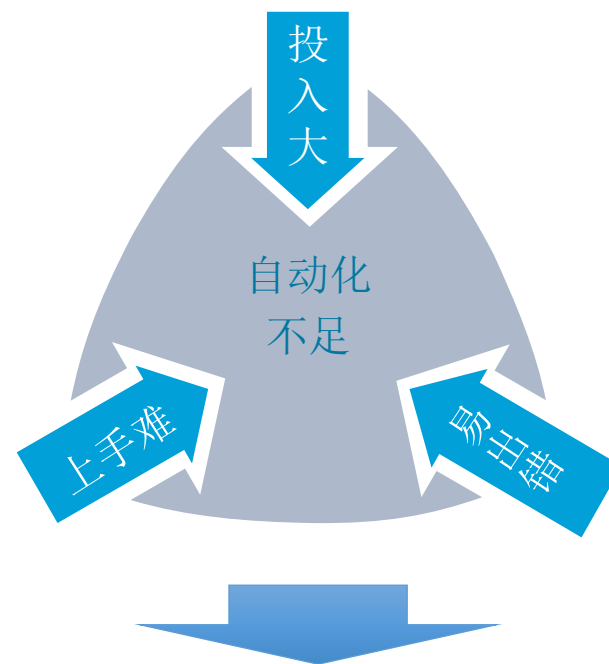
- 数据模型管理流程、指标、人员、工具缺失。
- 数据标准在数据模型中无法自动落地，无法检测及监督。

影响数据质量

- 数据质量差，严重影响大数据应用。

- ❑ 投入大量DBA人力
- ❑ 脚本开发、调试、实施
- ❑ 周期化固定的资源投入
- ❑ 重用率低，工作低效重复

- ❑ 要求用户熟悉物理模型
- ❑ 使用门槛较高



- ❑ 人工编辑脚本，脚本正确性难保证
- ❑ 操作人工执行，易出现误操作
- ❑ 执行结果验证难，执行过程易遗漏



PART3

数据标准化实践

大数据平台及数据
交换中心正常运作
的规则库



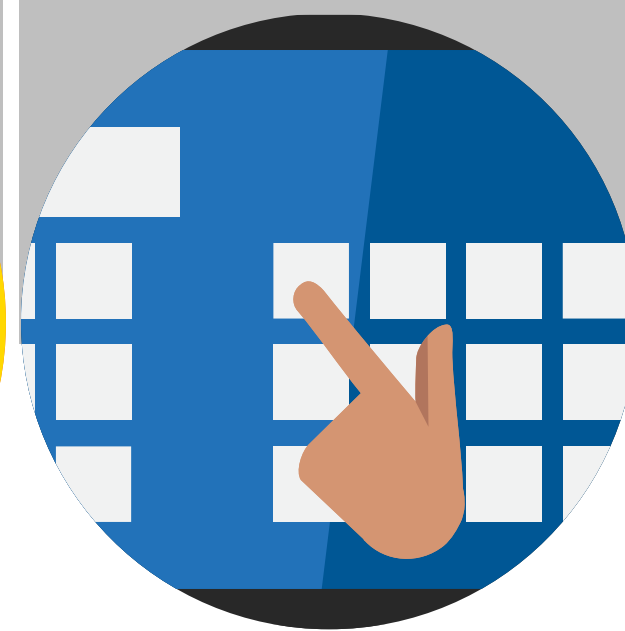
系统间数据比对及
差异化分析的前提
条件



各层级业务、技术
人员顺畅沟通的术
语规范，跨行业数
据交易、交换的沟
通桥梁



数据治理的基石，
信息化系统建设长
期有序发展的档案
馆

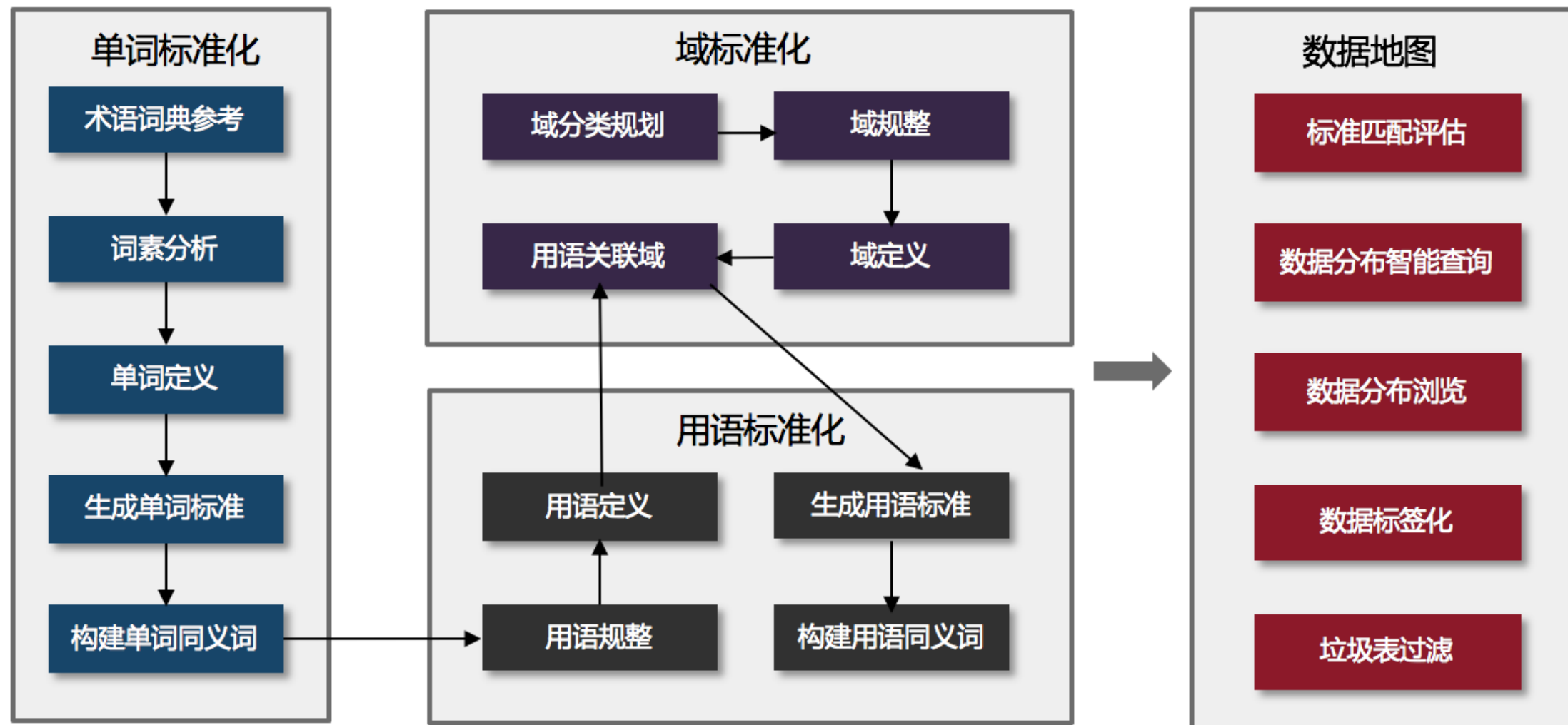


总体目标

- 建立一个面向业务人员、简单易用的数据标准管理平台，并对元数据进行图形化展现和多维度分析，满足灵活多变的业务分析需求。

系列目标

- 获取数据模型变更信息，帮助运维人员及时了解数据模型的变更内容及其影响范围、发现变更冲突，对相关部分做出调整，避免数据模型变更引起的数据质量问题，并为数据质量问题的分析定位提供帮助，提升运维人员解决问题的效率。
- 建立标准化管理机构，定义数据标准化步骤
- 对业务用语、单词、域、编码进行标准化





数据标准总览



数据标准地图



数据标准档案库



数据标准变更审批管理



数据标准应用



数据标准总览 ×

湖南移动NGBOSS ▾



数据标准档案容量

2017-12-04 15:35:13

✓ 域数:280

✓ 单词数:407

✓ 用语数:824

数据标准档案治理进程



■ 已标准化 28.8%

■ 未标准化 69.5%

■ 同义词标准 1.7%



本月标准审核通过率

0%

本月新增图列



已纳入数据标准化管理系统



电渠系统



CRM系统



BOSS系统



计费系统

🕒 数据标准总览

📊 数据标准地图

📁 数据标准档案库

📄 数据标准变更审批管理

📄 数据标准应用

数据标准地图 ×

🏠 / 数据标准管理 / 数据标准地图 / 数据地图详情

🔔 CRM系统

📊 关系地图



89235

非标准用语



47865

标准用语



2537

同义词标准用语

包含的主题域：

客户服务域

营销域

渠道域

零售库存域

客户交互域

订单域

客户域

销售域

产商品域

公共域

- 数据标准总览
- 数据标准地图
- 数据标准档案库
- 数据标准变更审批管理
- 数据标准应用

数据标准档案库

数据标准管理 / 数据标准档案库

用语 单词 域

查询 筛选 导出 批量导入

用语物理名: --

用语逻辑名: --

点击添加

用语类型: 变长字符串

用语物理名: EMAIL

用语逻辑名: 电子邮件

编辑 删除

用语类型: 变长字符串

用语物理名: CHECK_REMARK

用语逻辑名: 检验备注

编辑 删除

用语类型: 变长字符串

用语物理名: SHORT_CODE

用语逻辑名: 短码

编辑 删除

用语类型: 小数

用语物理名: SERVICE_ID

用语逻辑名: 服务标识

编辑 删除

用语类型: 变长字符串

用语物理名: AREA_TYPE

用语逻辑名: 区域类型

编辑 删除

用语类型: 变长字符串

用语物理名: CITY

用语逻辑名: 城市

编辑 删除

用语类型: 变长字符串

用语物理名: DEVELOP_CITY_CODE

用语逻辑名: 发展城市编码

编辑 删除

[数据标准总览](#)[数据标准地图](#)[数据标准档案库](#)[数据标准变更审批管理](#)[数据标准应用](#)[数据标准变更审批管理](#)[数据标准管理](#) / [数据标准变更审批管理](#)

待审批申请

开始时间

至

结束时间

请选择申请人

请选择变更标准体系

请选择审核状态

查询

申请时间分布图



审核申请数统计

**mf42637028: 申请编号**

申请时间: 2017-11-28 18:11:06.0

申请人: admin

被变更的标准体系: 湖南移动NGBOSS

查看详情

mf85385142: 申请编号

申请时间: 2017-11-28 18:10:47.0

申请人: admin

被变更的标准体系: 湖南移动NGBOSS

查看详情

PART 4

数据标准化价值和收益

一

■提高了数据品质、减少了管理成本、提高了数据治理效率

二

■定义了数据标准化的步骤，形成了完善的数据标准管理机制

三

■建立了数据标准体系，从结构上构建标准数据字典

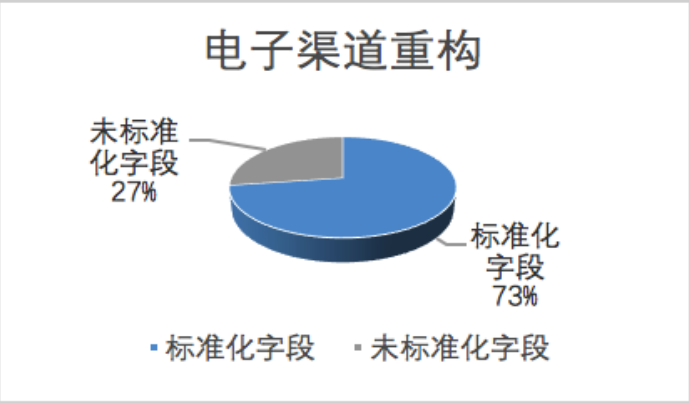
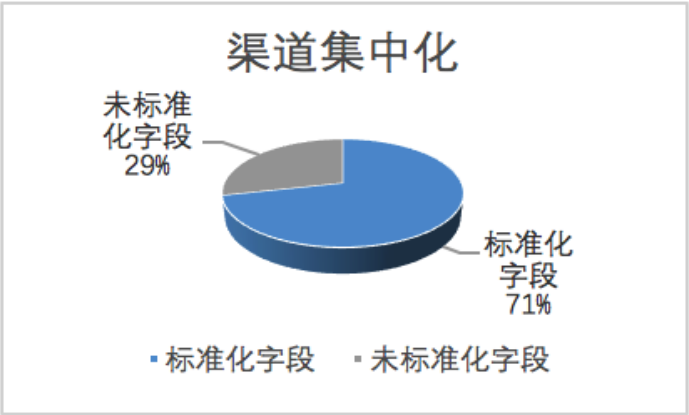
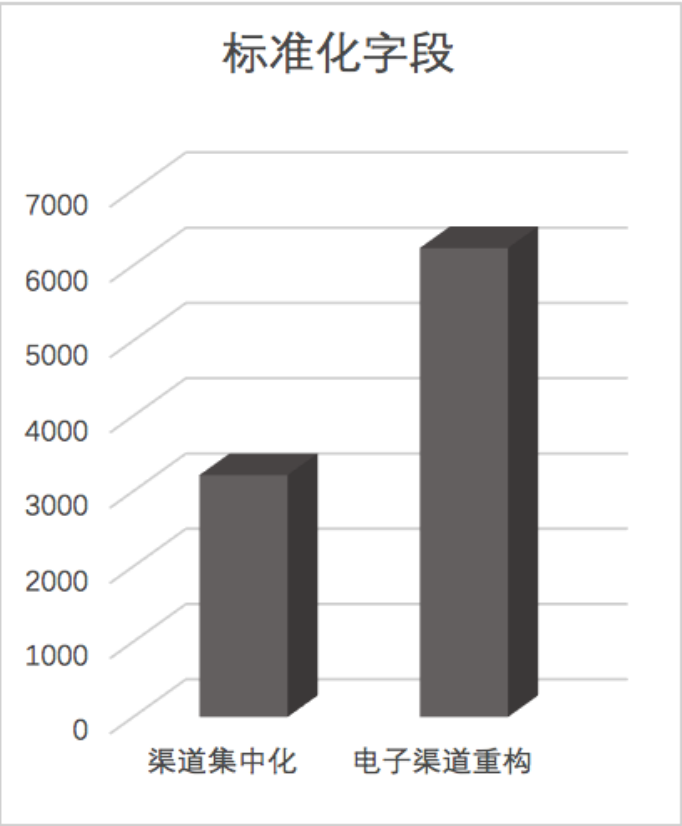
四

■实现了数据地图可视化，能直观展现各系统元数据信息

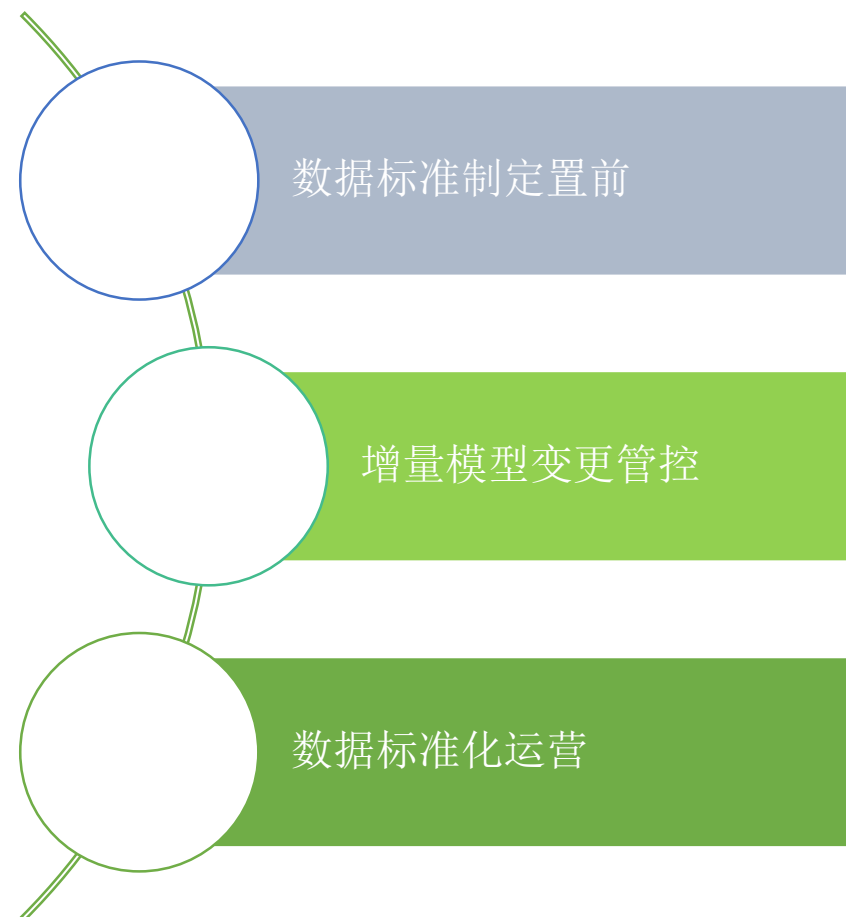
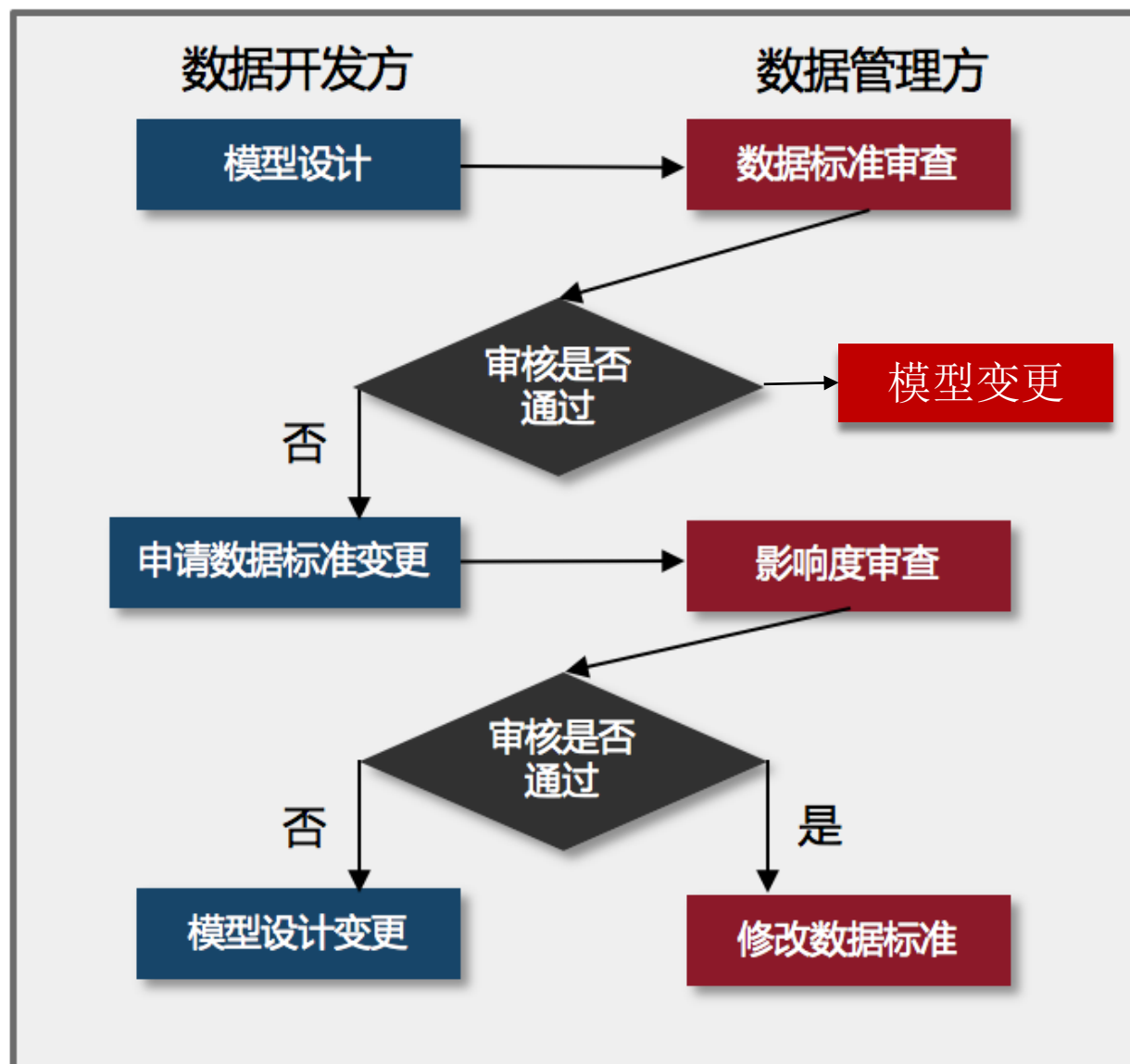
五

■实现了安全管控业务数据，为新上模型评审提供依据

湖南移动在渠道集中化、电子渠道重构项目中，使用数据标准管理平台对新上线系统的数据实施了标准化管控。输出标准单词442个，标准域527个，标准用语915个。



通过该项工作的开展，湖南移动公司的数据治理和数据标准化工作获得业界的肯定。在5月26日贵州举办的2017中国国际大数据产业博览会上荣获2016年度中国数据治理实践贡献奖。



建立数据标准档案馆
降低对业务、系统的理解难度，
系统迁移有迹可循

实现自动化数据标准审核，
推进数据治理进程，
为数据开放创造前提条件

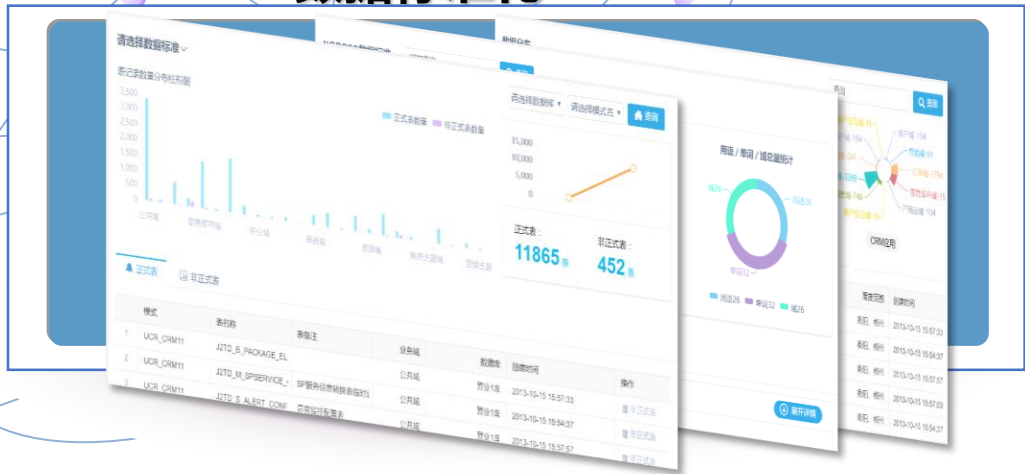
稽核数据一致性，提升
报表指标口径准确性

业务数据标准地图，统一用语语义，
降低需求提出方与开发商的沟通难
度

通过对数据标准中口径、脚
本、调度配置等信息的记录
及变更监控，预警指标差异

基于数据标准定义，为
模型整合、降维提供分
析支撑

数据标准化



The logo for DBAplus, featuring the letters 'DBA' in red, blue, and orange respectively, followed by 'plus' in green. A thin white horizontal line is positioned below the letters.

DBAplus

www.dbaplus.cn

THANK YOU!