



DTCC

2016中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2016

数据定义未来

SequeMedia
盛拓传媒

IT168.com

ChinaUnix

ITPUB



农银人寿保险股份有限公司
ABC LIFE INSURANCE CO., LTD.

新核心业务系统 数据架构规划与数据治理

演讲人: **赵华** (IT部-副总经理、新核心项目经理)

zhaohua@abchinalife.cn

种磊 (IT部-资深专员、新核心项目-数据组长)

chonglei@abchinalife.cn

总公司信息技术部

新核心项目组

2016年5月



ABC Life Insurance Co., Ltd
An Insurance Company of Agricultural Bank of China



农银人寿保险股份有限公司（简称“农银人寿”）是中国农业银行的控股子公司，依托农业银行雄厚的资金实力、庞大的经营网络、完善的金融服务、卓越的社会信誉，为客户提供高品质的保险保障和财富规划服务。

健康

重疾

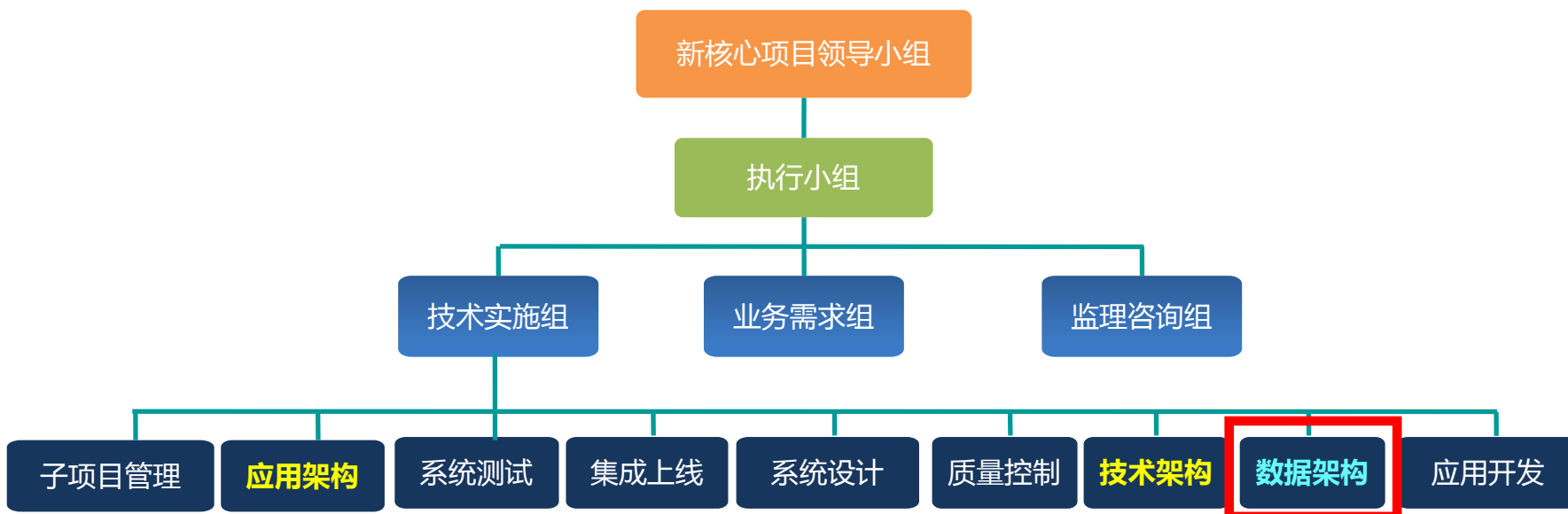
意外

年金

养老

- 国内机构布局最广的银行系寿险公司，拥有20多家分公司和300多家分支机构；
- 继中国农业银行控股农银汇理、农银租赁、农银国际后的又一股权投资力作；
- 国有五大商业银行全部拥有自己保险公司的收官之作；

组织架构 - 数据架构组 (人员、工作职责)



组内人力：11人，约占项目总人力 10%

子项目数：8*个

工作职责

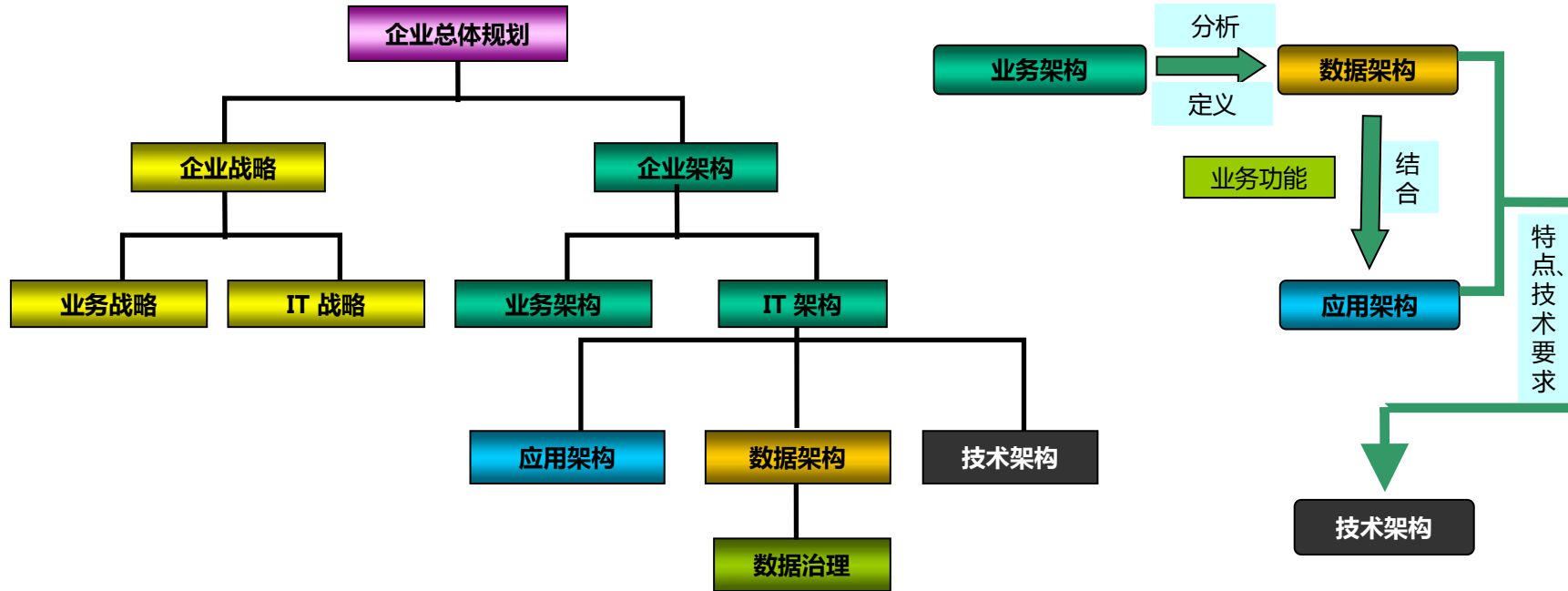
- 1、数据架构规划与设计
- 2、数据治理与标准化方案的制定及实施
- 3、老核心数据模型整理、数据治理与标准化
- 4、新核心模型设计与标准化应用
- 5、数据标准规范、数据管理流程的制定与督导
- 6、核心业务系统数据迁移方案的制定与实施

序号	角色	人数
1	组长 (负责人)	1人
2	业务模型设计	2人
3	数据模型设计、数据迁移	7人
4	老核心模型整理	1人



- 1 •数据架构 - 定位、设计目标、设计原则、设计思路
- 2 •数据架构 - 高效数据操作
- 3 •数据架构 - 规划设计
- 4 •数据架构 - 数据源、数据的准备、存储、加工、交换
- 5 •数据治理 - 概述、数据现状、分阶段实施与当前进展
- 6 •数据治理 - 元数据管理
- 7 •数据治理 - 主数据管理
- 8 •数据治理与标准化应用 - 业务建模、数据建模、数据样例
- 9 •数据治理与标准化 - 成果展示

数据架构 - 定位



■ IT总体架构需考虑**数据架构对当前业务支持**，理想规划顺序为**数据驱动**！

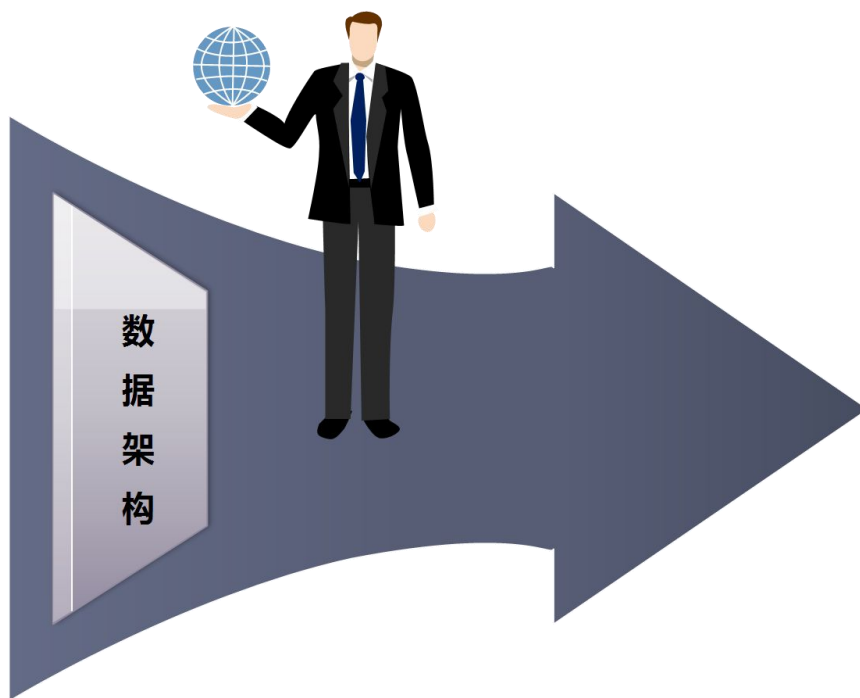
以及时的方式，获取、处理和使用数据来创造效益，不断地迭代开发新产品，借助数据探索引领企业发展（Navigating）。

判断是否为数据驱动的标准：

- (1) 产生的数据量
- (2) 使用数据的程度
- (3) 内化数据的过程及深度

数据架构 - 设计目标

数据架构是IT架构的重要组成部分，用来解决“**如何管理数据**”和“**如何使用数据**”的问题。



1

实现数据标准化

2

减少数据冗余，提升数据质量，提升系统性能

3

消除信息孤岛，实现数据在系统间的广泛共享

4

发挥数据资产价值，为企业带来高附加值回报

总体原则

在基于问题现状、改进分析的基础上，结合业务需求要点、主流技术及行业实践，进行目标的数据架构规划。

1

灵活性

- 对数据的组织划分要灵活，以满足不同业务的要求
- 快速应对业务变更

2

高效性

数据校验、数据加载、数据加工处理、数据移动的高效率，要满足特定的时间窗口。

3

可扩展性

- 充分考虑未来系统的扩展能力
- 减少新技术、新业务需求对现有架构的冲击。

4

共享性

数据项在企业范围内全局唯一，有确切的业务含义；强调对共有需求的加工，在系统内、系统间实现数据的重用和共享。

5

可用性

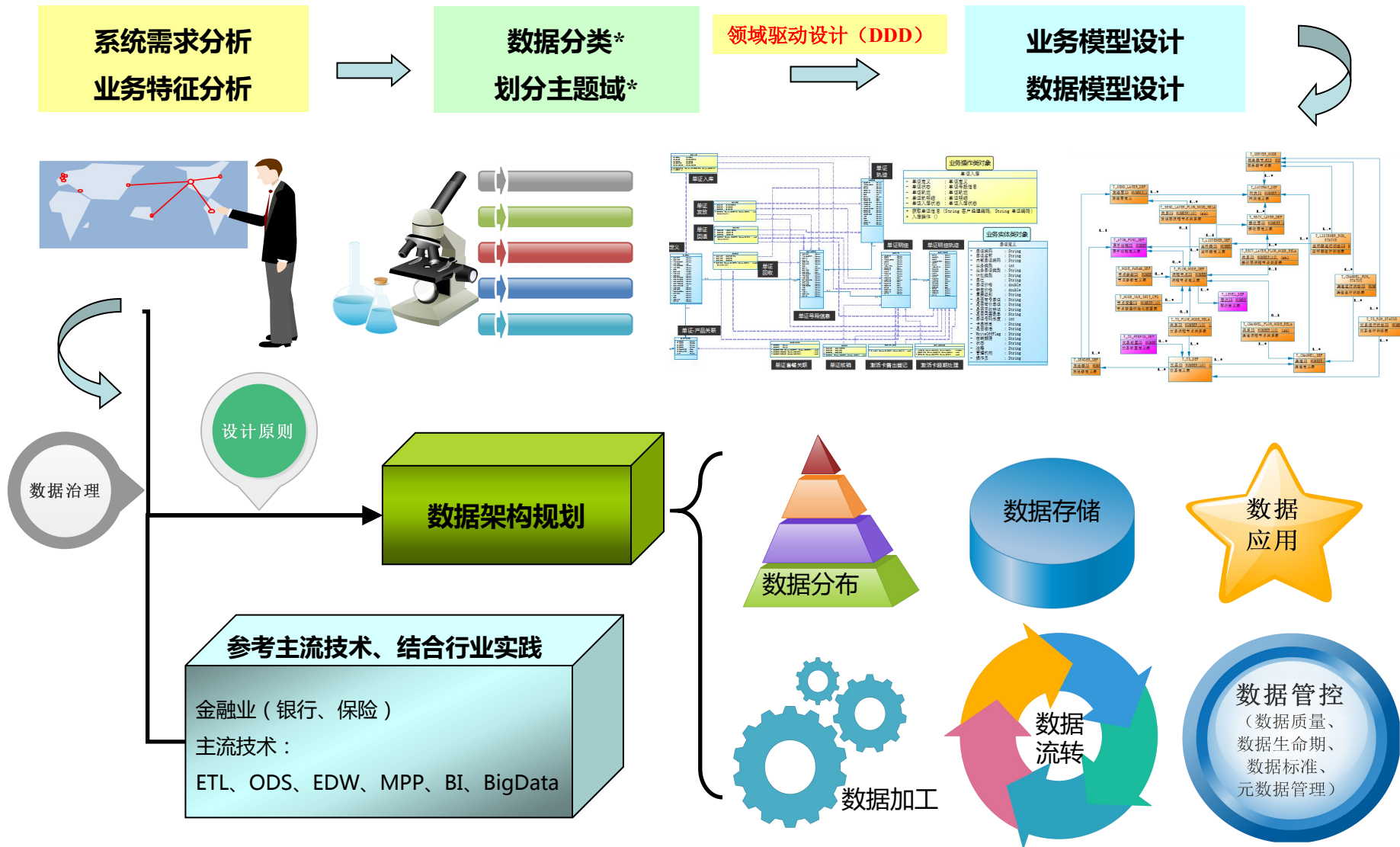
数据具有使用价值；数据的存储、分布、流转、加工要满足业务系统消费数据的要求。

6

安全性

- 区分数据安全级别、管理等级
- 区分敏感数据、非敏感数据
- 要求特定权限进行管理和访问

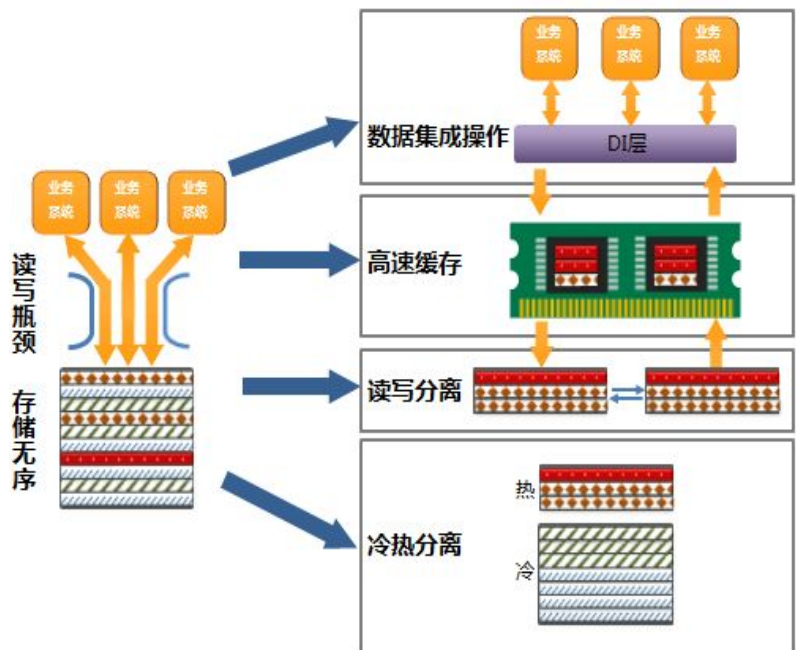
数据架构 - 设计思路





- 1 •数据架构 - 定位、设计目标、设计原则、设计思路
- 2 •数据架构 - 高效数据操作
- 3 •数据架构 - 规划设计
- 4 •数据架构 - 数据源、数据的准备、存储、加工、交换
- 5 •数据治理 - 概述、数据现状、分阶段实施与当前进展
- 6 •数据治理 - 元数据管理
- 7 •数据治理 - 主数据管理
- 8 •数据治理与标准化应用 - 业务建模、数据建模、数据样例
- 9 •数据治理与标准化 - 成果展示

数据架构 - 高效数据操作

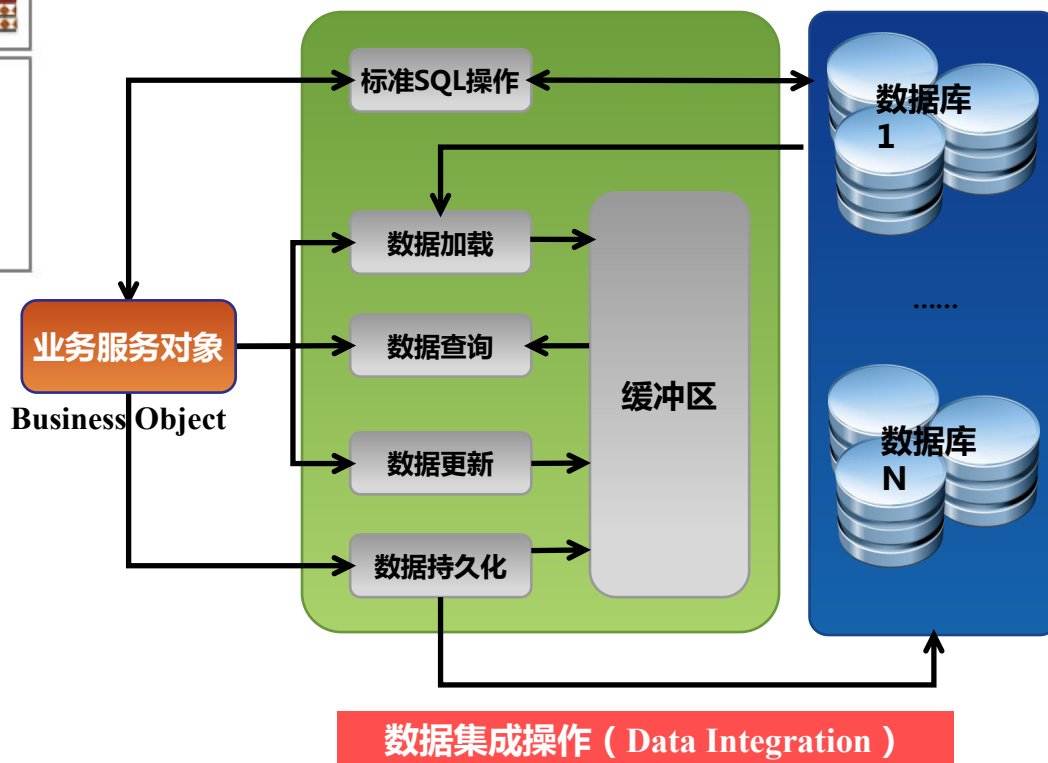
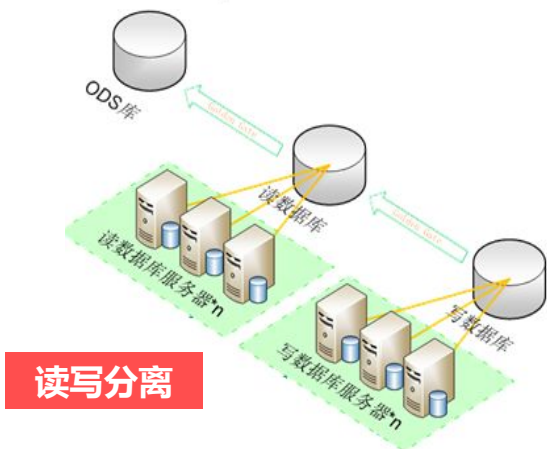


数据集成层 (DI) Powered by DEVELA

- ◆ 高效、透明、可独立部署
- ◆ 降低开发难度，提升开发效率
- ◆ 解决重复查询导致的性能问题

高速缓存 Powered by CACHE

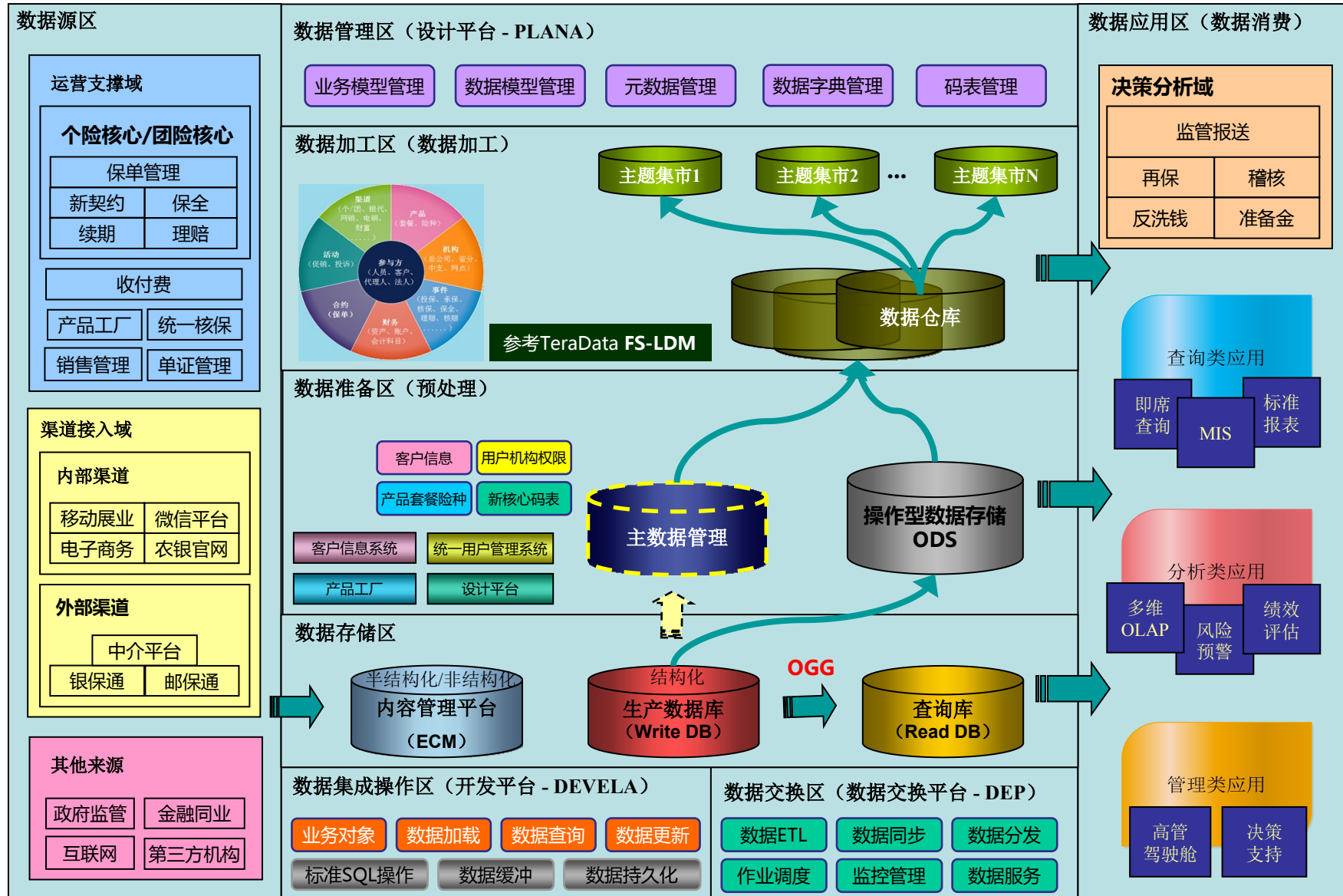
- ◆ 存放访问频率很高、但数据量较少的数据。如：码表、用户机构权限、费率相关。





- 1 •数据架构 - 定位、设计目标、设计原则、设计思路
- 2 •数据架构 - 高效数据操作
- 3 •数据架构 - 规划设计
- 4 •数据架构 - 数据源、数据的准备、存储、加工、交换
- 5 •数据治理 - 概述、数据现状、分阶段实施与当前进展
- 6 •数据治理 - 元数据管理
- 7 •数据治理 - 主数据管理
- 8 •数据治理与标准化应用 - 业务建模、数据建模、数据样例
- 9 •数据治理与标准化 - 成果展示

数据架构 - 规划设计



•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换

•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展

•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

•数据治理与标准化应用 - 业务建模、数据建模、数据样例

•9

•数据治理与标准化 - 成果展示



数据架构 - 数据源

农银人寿核心业务系统（运营支撑）

内部数据



个险核心/团险核心

（保单、契约、保全、理赔、续期）

收付费、销管、单证、统一核保
产品工厂.....

农银人寿核心业务系统（渠道接入）



内部渠道

移动展业、微信平台、农银官网、电子商务

外部渠道

中介平台、银保通、邮保通

金融机构（银行、保险同业）

外部数据



政府部门、监管机构



第三方机构（电商、互联网公司）



互联网

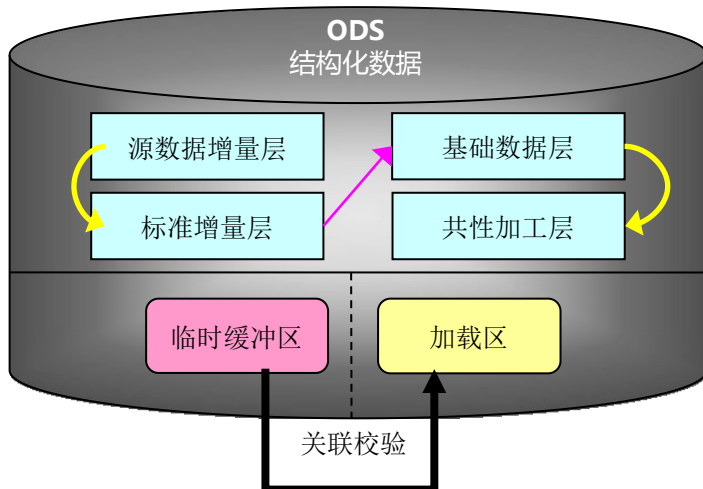


数据源 - 非功能属性（对数据存储设计起参考作用！）

- 数据源
- 获取方式（接口 or 文件交换；全量 or 增量）
- 变动频率（极少、偶尔、固定周期）、变动形式（新增、更新）
- 数据格式（结构化、半结构化、非结构化）
- 数据量及变动量
- 共享程度：根据业务模块对数据的需求强度

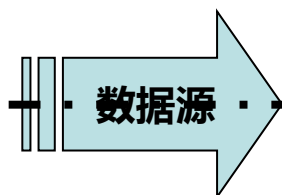
数据架构 - ODS、内容管理平台 (ECM)

数据准备区 (预处理)

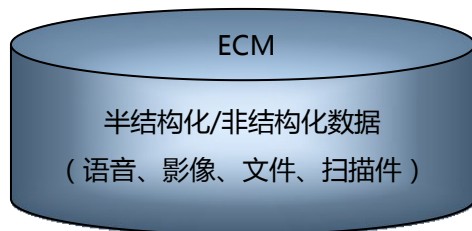


操作型数据存储 (Operational Data Storage)

- 面向主题的、集成的、反映当前细节变化的数据集
- 贴源设计
- 整合业务数据全貌，提供跨系统的、细节查询
- 数据清洗及标准化，提升数据质量
- EDW主要数据源



数据存储区



数据架构 - 数据仓库与集市 (规划)

EDW : 面向主题的、集成的、稳定的、分析型的、不可更新的、非实时的、反应一定历史时期内的数据集合。

EDW - 设计要点 (Points)

- 数据源 (数据采集)
- 数据模型 (存储结构)
- ETL过程 (设计关键)
- 元数据管理 (BI中尤为重要)
- 技术手段 (数据挖掘、知识发现、多维分析)
- 数据应用 (复杂查询及报表、绩效评估、高管驾驶舱)

EDW - 数据源 (Data Source)

- 从OLTP DB中捕捉业务变化数据
- 清洗、标准化、统一编码 (确保数据一致性、消除冗余)
- 数据整合 (计算、汇总、集成、形成指标体系)
- 重点关注: 数据类型、更新方式、更新频率、更新数据量、数据质量、抽取范围、每次抽取量、消费系统的实时性要求。

EDW - Transform in ETL

- 数据不一致的转换
- 数据粒度的转换
- 计算指标的转换

转换的具体实现

- 业务数据按照EDW粒度进行聚合
- 多个SQL语句表示复杂转换规则
- 数据集市字段与业务系统字段关联映射 (转换函数及参数)



参考 TeraData FS-LDM 分主题构建数据模型

数据仓库



EDW建设阶段

一、规划分析阶段

- 从用户角度进行需求分析（决策分析的内容和功能）
- 构建概念模型（参考模型，划分主题，确定实体粒度和层级）
- 从技术角度分析（主题需要哪些业务数据源支持、ETL工具选取）

二、设计实现

数据源选取、模型设计、元数据设计、ETL设计、数据清洗、统一编码、填充测试数据、数据应用

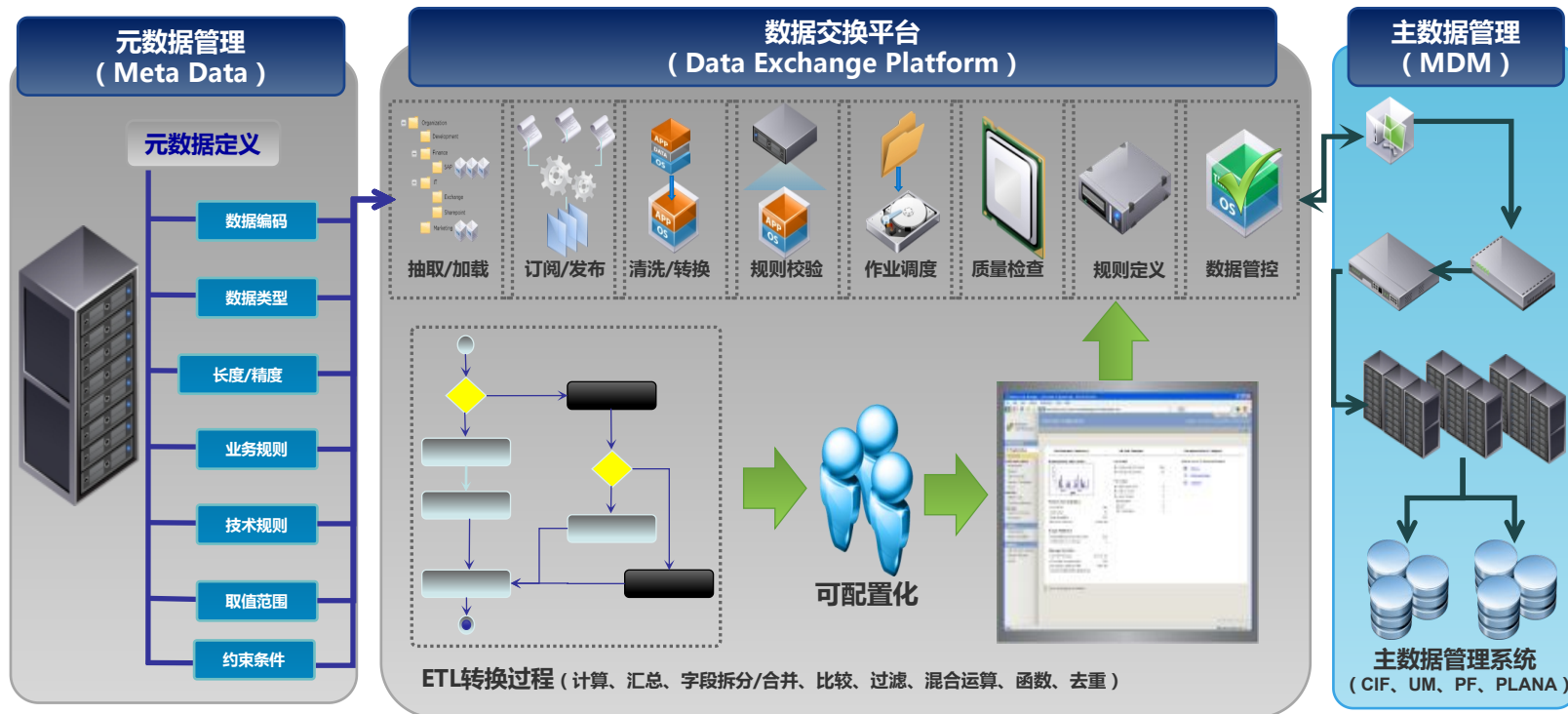
三、运维阶段

- 定期备份与恢复
- 归档（离线存储）
- 性能监控（数据备份与恢复时间、ETL时间、报表生成时间）

EDW建设过程

- 项目规划
- 需求分析
- 概念模型设计
- 逻辑模型设计
- 数据架构设计
- 元数据设计
- 物理模型设计
- ETL设计
- 填充测试数据
- 开发及实施
- 性能监控与优化
- 运维交付

数据架构 - 数据交换平台 (DEP)



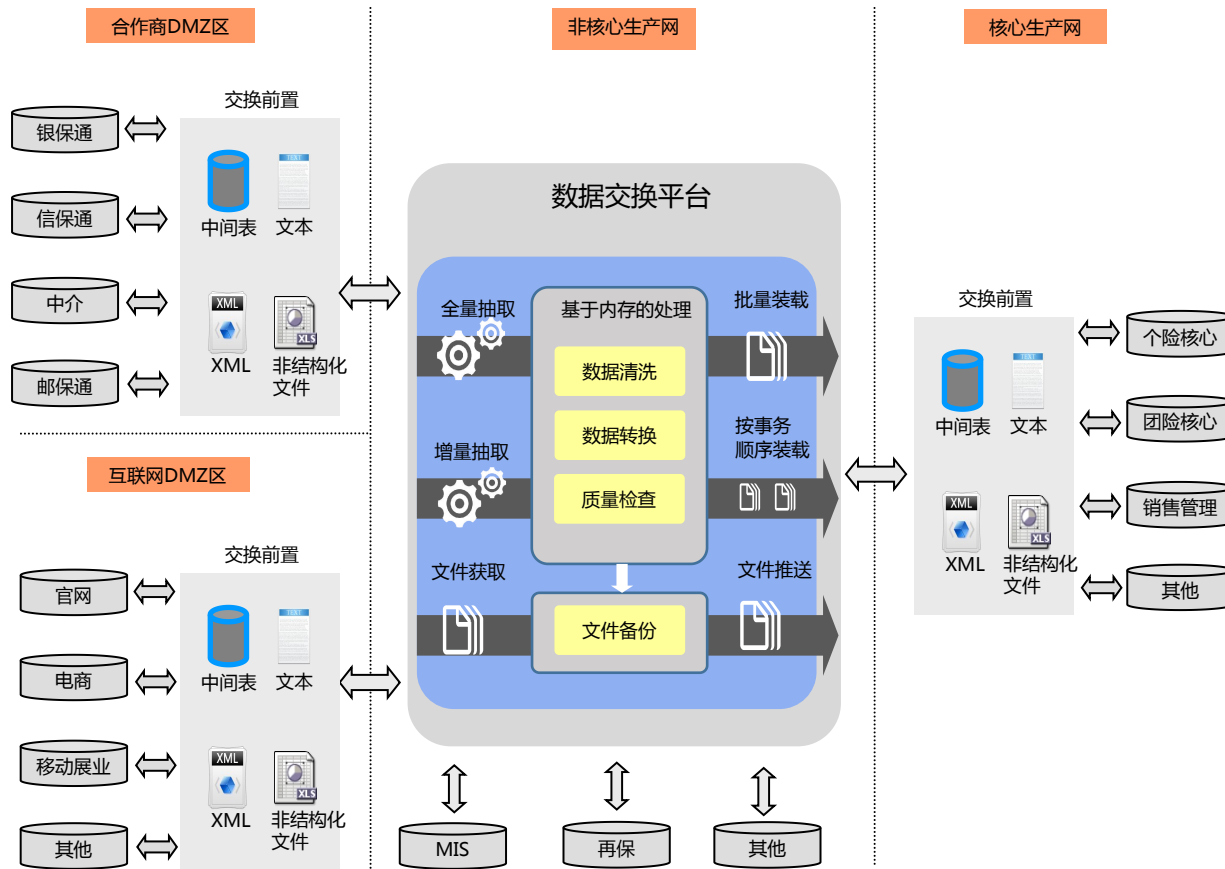
问题现状

- 数据集成、数据同步、数据共享困难
(核心内各子系统之间、核心与外围系统之间)
- 系统耦合度较高 (A系统数据结构发生变化, 需要B系统协同改造)
- dblink使用泛滥 (性能开销大、存在数据安全隐患)

建设目标

- ◆ 提供统一的**数据集成规范**、**数据获取与分发**、**数据交换与共享**、**数据监控**。
- ◆ 提高**数据加工**、**数据流转**效率, 加快数据在系统内、系统间的快速移动。
- ◆ 改变传统的“多对多交换模式”, 实现“**一源多目标**”的数据更新。

数据架构 - 数据交换平台 (DEP)



DEP功能定位 - 数据交换枢纽

通过一定的“数据获取及分发”策略、作业调度策略，结合内置于各系统内的数据交换区，实现：

- (1) 核心业务系统内各子系统
 - (2) 核心业务系统与外围各系统
- 之间的数据交换、数据同步、数据共享，为OLTP、OLAP系统提供数据接口服务或数据文件交换服务。

设计要点

- **增量数据捕获：**包括新增、更新；判断依据（时间戳、日志分析）
- **细化调度作业：**分清任务依赖关系，增大并行度，提高交换效率

各系统配合DEP

内置数据或文件交换区，提供中间接口表（源数据供应表、消费数据获取表）、格式化文件等。

•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换



•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展

•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

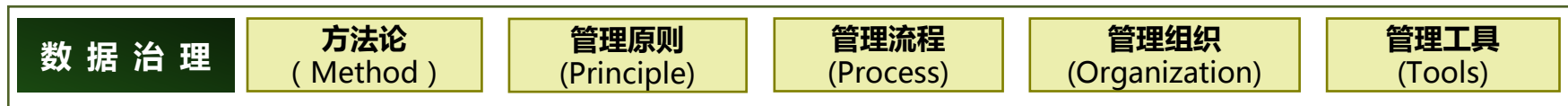
•数据治理与标准化应用 - 业务建模、数据建模、数据样例

•9

•数据治理与标准化 - 成果展示

数据治理 - 概述

数据治理框架 - Data Governance Framework



> 关于如何制定、实施针对整个企业内部数据的商业应用和技术管理的一系列政策和程序。
 > 一套持续改善的管理机制，通常包括数据架构组织、数据模型、政策及体系制定、技术工具、数据标准、数据质量、影响度分析、作业流程、监督及考核等。

数据治理

> 描述数据的数据
 > 比一般意义上的数据范畴更加广泛
 > 在分析型项目中，帮助数据仓库设计和开发人员快速查找数据
 > 在OLAP系统中，用来描述分析型应用的内部数据结构、建立方法和流程

元数据管理

Data Driven World

EN·CORE

MetaData (Data about Data)

数据名称、数据类型、数据归属、取值范围、校验规则、数据源

- ❖ 建立数据应用标准
- ❖ 消除不一致性，实现数据广泛共享
- ❖ 提升组织的数据质量
- ❖ 将数据资产应用到业务、管理和战略决策
- ❖ 发挥数据资产的商业价值

数据治理作用

作为一种语义工具，为数据治理建立一套资料库，存储：数据定义、负责人、来源、转换关系、目标、质量等级、依赖关系、安全权限等。对于实现**商业整合、数据质量提升、可审计性**等数据治理目标至关重要。

元数据管理作用

*MetaM: 根据业务规则变化，识别系统需要改变的部分，评估改变部分对其他系统或系统其他部分产生的影响，确保数据一致性，实现数据可追溯

数据现状

信息技术 vs. 数据质量

- ❖ 信息技术可以帮助企业决策，快速适应业务变化和市场需要；
- ❖ 但数据质量低下，将导致决策延迟或决策错误，使企业错过商机或造成损失；

数据现状

Situation

- 不符合业务或技术规则
- 键值重复、属性重复
- 数据格式错误
- 无效数据
- 多套重复编码*，码表取值范围不统一
- 命名规则不统一*
- 数据类型不统一*（类型、长度/精度/小数位）
- 数据列冗余，难以找寻基准数据，一致性难以维护
- 大量废弃不用的表*、空值字段

引发问题

Trouble

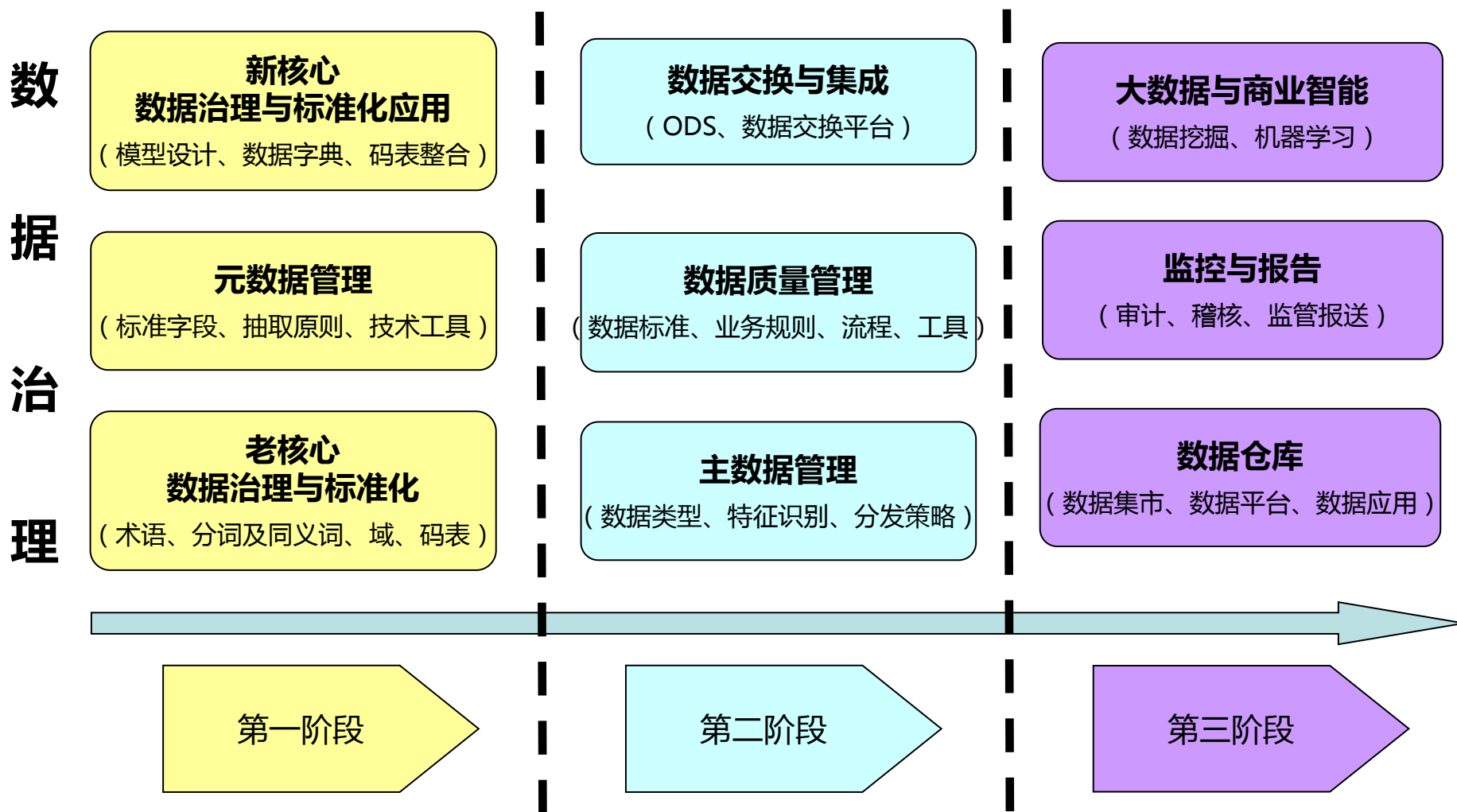
- 影响数据一致性
- 影响数据完整性
- 影响数据准确性
- 系统间数据交换共享困难
- 重复统计造成误差
- 数据迁移困难
- 开发运维效率低
- 数据库性能降低

数据治理目标

Goal

- 数据管理原则
- 数据规范标准
- 数据管理流程
- 数据管理组织
- 老核心数据治理与标准化
- 新核心应用数据治理与标准化成果，进行数据质量管控，实现数据广泛共享，发挥数据资产价值。

数据治理 - 分阶段实施、当前进展



•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换

•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展



•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

•数据治理与标准化应用 - 业务建模、数据建模、数据样例

•9

•数据治理与标准化 - 成果展示

标准字段 (Standard Field)

属“元数据管理”范畴，是仅包含最本质业务含义的信息项。它规定了新核心业务系统中，数据项的中英文名称、业务含义、所属应用模块、数据类型（长度/精度/小数位）、数据格式、校验规则（技术规则及基本业务规则）、默认值、取值范围（码表项）等属性特征。

识别方法

遵照《新核心业务系统 - 标准字段抽取规范》执行。

管理工具

由新核心设计平台 (PLANA) 对标准字段进行统一维护。

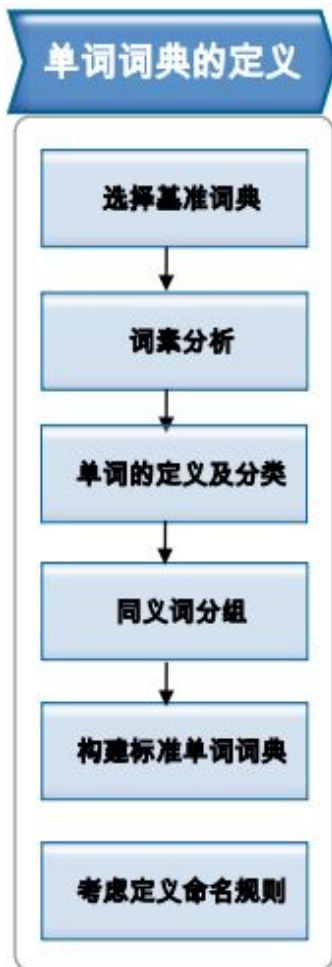
使用方式

数据字段（表字段）继承自“标准字段”，并根据被引用数据表的业务场景进行扩展。

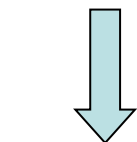
举例

“投保人号”、“被保人号”，本质为“客户号”。在设计表字段时，它们都将从标准字段（客户号）继承，具有与标准字段相同的数据类型（长度/精度）、数据格式、校验规则、默认值。





源系统信息采集



源系统
数据字段名称

保险合同号

保单号

保单号码

保单合同号

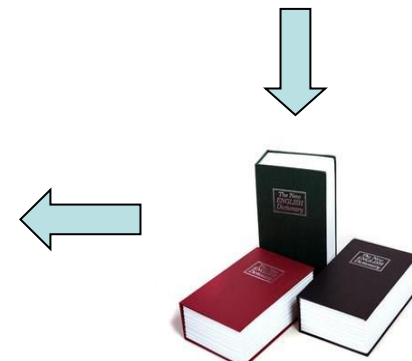


命名规则

区分	用法
实体	<ul style="list-style-type: none"> (修饰语 + 主题语 + 修饰语) + 实体分类语 由中文、英文、数字来组成
属性	<ul style="list-style-type: none"> (修饰语 + 主题语 + 修饰语) + 域 (域=修饰语+分类语) 有时属性名和域名会相同，因此一般以修饰语+域的形式来定义

标准化 基础词汇 (中文)	标准化 基础词汇 (英文)	标准化 基础词汇 (英文缩写)	标准化 同义词
保单	POLICY	POL	保险合同、保单合同
号码	NUMBER	NO	号

标准化 数据字段 (中文名称)	标准化 数据字段 (英文名称)
保单号	POLNO



•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换

•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展

•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

•数据治理与标准化应用 - 业务建模、数据建模、数据样例

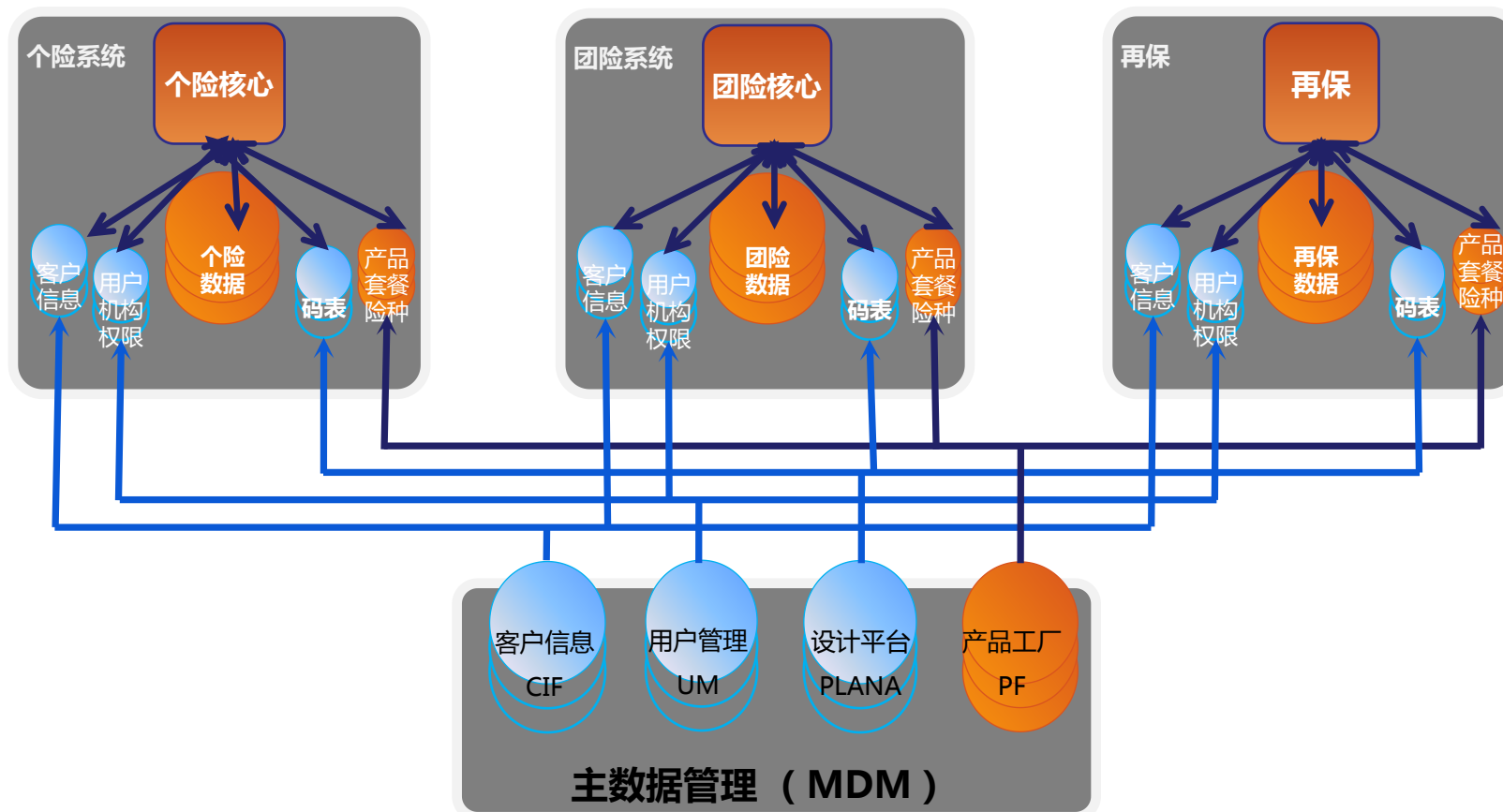
•9

•数据治理与标准化 - 成果展示

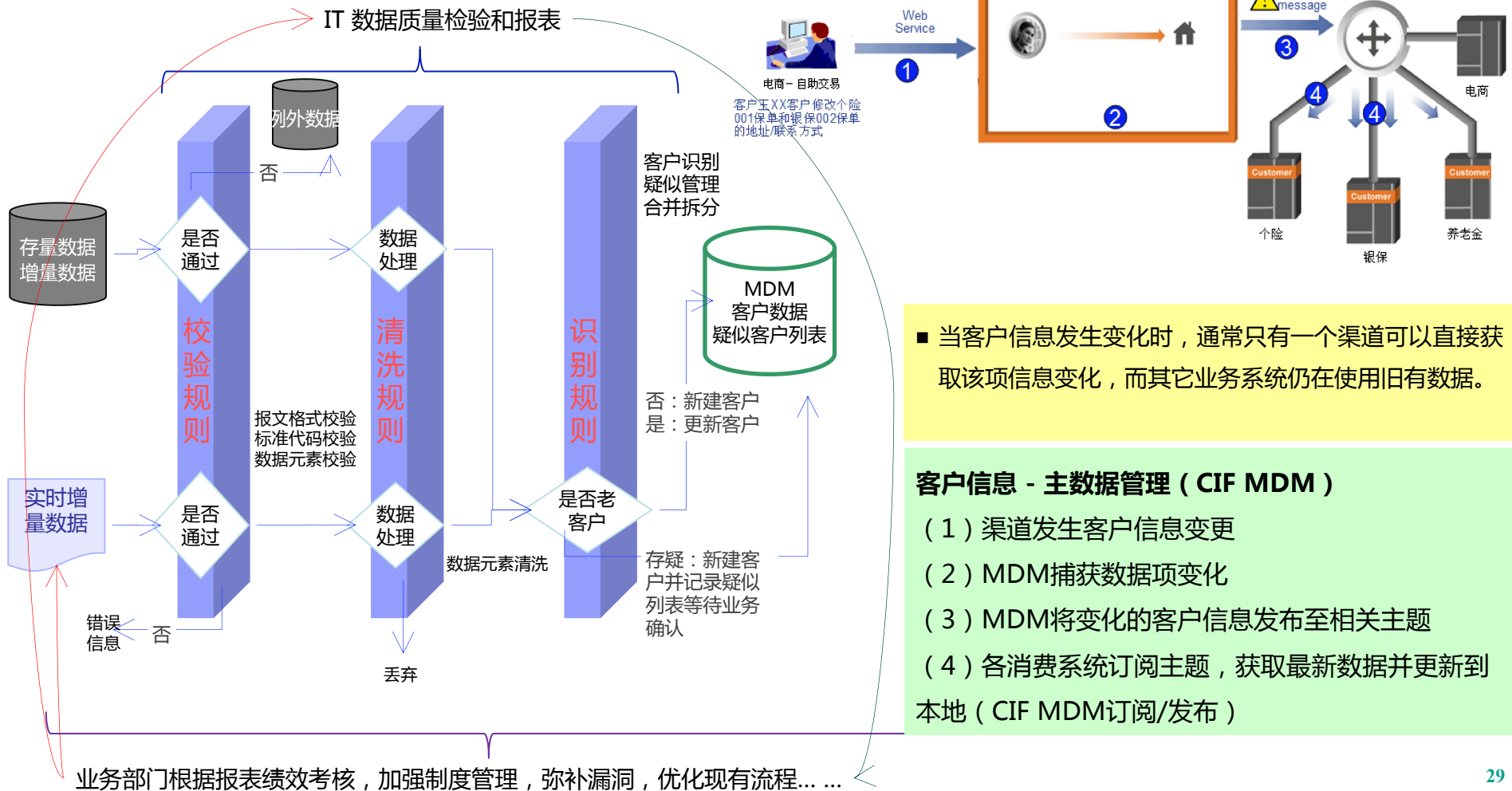


重点关注

- 1、识别主数据 - 源生产系统（获取最准确主数据）
- 2、识别主数据 - 目标消费系统
- 3、主数据 - 清洗策略（数据校验规则、清洗规则）
- 4、主数据管理原则、分发同步策略（MDM系统订阅/发布功能结合DEP实现）



客户信息数据的获取、校验、清洗、分发



■ 当客户信息发生变化时，通常只有一个渠道可以直接获取该项信息变化，而其它业务系统仍在使用旧有数据。

客户信息 - 主数据管理 (CIF MDM)

- (1) 渠道发生客户信息变更
- (2) MDM捕获数据项变化
- (3) MDM将变化的客户信息发布至相关主题
- (4) 各消费系统订阅主题，获取最新数据并更新到本地 (CIF MDM订阅/发布)

数据治理 - 主数据管理 (规则校验、质量管控)

添加标识

标识类型: 身份证

描述:

颁发地点:

标识编号:

状态:

到期日期:

有效日期:

结束日期:

按参与方标识分配:

分配条件:

标识摘要

全部 活动 不活动

标识类型	标识编号
身份证	110101197107134041
身份证	110101102606066063

1 < 第 1 页 (共 1 页) > > |

标识详细信息

标识类型: 身份证

描述:

颁发地点:

标识编号:

状态:

到期日期:

有效日期: 2015年4月15日

结束日期:

按参与方标识分配:

分配条件:

客户类型	信息项	清洗规则
个人客户	证件类型	去除两边和中间的空格
		若证件类型为驾驶证, 则转换为身份证
	证件号码	若证件类型为身份证, 则 15 位转 18 位
		全角转半角
		身份证号出现的字母均转为大写
		去除两边和中间的空格
	个人姓名	注: 按上述规则清洗后, 若证件号码为全一样的数字, 或只有 1 位, 则不合规
		英文姓名, 去除前后空格
		中文姓名, 去除两边和中间的空格
		注: 若姓名中存在汉字, 则认为是中文姓名
		去除姓名中包含的特殊字符
		“*”、“?”、“?”、“_”、“_”、“~”、“.”、“!”、“@”、“/”、“+”、“\”、“.”、“.”
		全角转半角
	去除第一个字符为“.”, 去除最后一个字符为“.”	
	去除第一个字符为“.”, 去除最后一个字符为“.”	
注: 按上述规则清洗后, 姓名全是数字, 则不合规		
手机	去除两边和中间的空格	
	全角转半角	
	只保留数字	
固定电话	去掉“86”开头	
	去除两边和中间的空格	
电子邮件	全角转半角	
	只保留数字	
性别	去除两边和中间的空格	
	全角转半角	
出生日期	若性别为空, 从合法身份证号中获取性别	
	将历史数据中“空”的性别清洗为“未知”	
地址	若出生日期为空, 从合法身份证号获取出生日期	
	清洗“空”的地址	

清洗效果有限

- 局限于数据格式的清洗, 对文本的具体内容无能为力。
- 校验规则只能作为事后清洗和弥补的手段。
- 不如从源头上获取高质量数据

数据质量提升

- 不仅仅是技术层面的事情
- 更需要业务部门加强监管, 实行“数据质量与外包录入”的绩效挂钩!

•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换

•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展

•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

•数据治理与标准化应用 - 业务建模、数据建模、数据样例

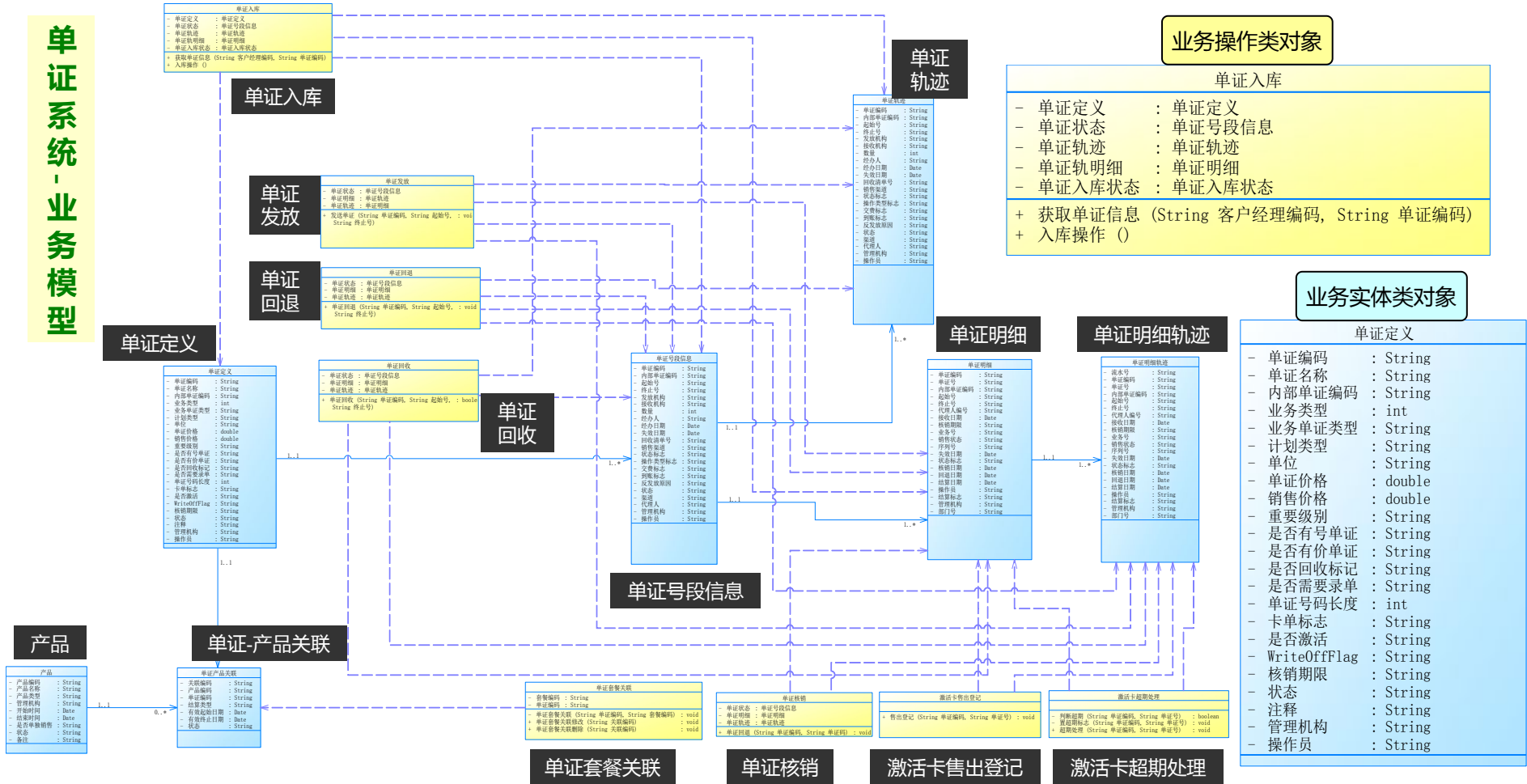
•9

•数据治理与标准化 - 成果展示



数据治理与标准化应用 - 新核心模型设计 - 业务建模

单证系统 - 业务模型



Transform 业务模型 into 数据模型

- ER模型分析方法 (面向结构)
- 对象类模型分析方法 (面向对象)

Tips : 在端到端的流程分析和业务用例分析过程中会衍生出数据对象。

数据治理与标准化应用 - 新核心模型设计 - 业务数据样例



渠道接入平台 - 业务数据样例 (Business Data Sample Case)

一：“服务1”中的监听器(192.168.1.1:80001)接收中介请求											
二：解密1001报文											
三：1001请求报文->1006请求报文，并进行合法性校验											
四：向核心发起1006交易请求，记录1006交易报文											
五：1006返回报文->1001返回报文，记录交易报文和交易日志											
六：返回报文加密											
七：请求成功											
1	标记交易日志状态为成功										
系统日志表T_SYS_LOG											
系统日志ID	系统流水号	网关ID	网关编码	服务器节点	渠道ID	渠道编码	业务流水号	交易编码	内部编码	交易用户	交易机构
SYS_LOG_ID	SYS_SN	GATEWAY_ID	GATEWAY_CODE	SERVER_NODE	CHANNEL_ID	CHANNEL_CODE	BIZ_SN	TX_CODE	INTER_CODE	TX_USER	TX_ORG
1	1000000001	8800001	ZJ	服务器1	1	06	20000001	1006	1006	当前用户	当前用户机构
2	1000000001	8800001	ZJ	服务器1	1	06	20000001	1006	1006	当前用户	当前用户机构
5	1	20000001	1000000001	ZJ	服务器1	06	ZJ	ZJ	1006	1006	tx_send_layer
6	2	20000001	1000000001	ZJ	服务器1	06	ZJ	ZJ	1006	1006	tx_rtn_layer
7	2	20000001	1000000001	ZJ	服务器1	06	ZJ	ZJ	1006	1006	channel_rtn_layer
5	1	20000001	1000000001	ZJ	服务器1	06	ZJ	ZJ	1006	1006	tx_send_layer
6	2	20000001	1000000001	ZJ	服务器1	06	ZJ	ZJ	1006	1006	tx_rtn_layer
原子服务日志表T_ATOM_SRV_LOG											
原子服务日志ID	系统日志ID	系统流水号	业务流水号	服务名称	应用名称	报文路径	日志时间	日志类型	执行状态	备注	
ATOM_SRV_LOG_ID	SYS_LOG_ID	SYS_SN	BIZ_SN	SRV_NAME	APP_NAME	MSG_PATH	LOG_TIME	LOG_TYPE	EXEC_STATUS	REMARK	
1	1	1000000001	20000001	?	?	c:\cip\	2015-7-7 13:24	01	01	-	
2	2	1000000001	20000001	?	?	c:\cip\	2015-7-7 13:24	02	01	-	

•1

•数据架构 - 定位、设计目标、设计原则、设计思路

•2

•数据架构 - 高效数据操作

•3

•数据架构 - 规划设计

•4

•数据架构 - 数据源、数据的准备、存储、加工、交换

•5

•数据治理 - 概述、数据现状、分阶段实施与当前进展

•6

•数据治理 - 元数据管理

•7

•数据治理 - 主数据管理

•8

•数据治理与标准化应用 - 业务建模、数据建模、数据样例

•9

•数据治理与标准化 - 成果展示



数据治理与标准化 - 成果展示 (1/7)

老核心数据治理

字段名称	表名	列名	字段中文含义	字段来源	取值范围	数据类型	数据长度
ABSDAYS	LAACCOUNTS	ACCOUNT	帐号	输入		VARCHAR2	30
ACCNAME	LAACCOUNTS	ACCOUNTNAME	帐号名称	输入		VARCHAR2	60
ACCOUNT	LAACCOUNTS	AGENTCODE	业务员编码	输入		VARCHAR2	10
ACCOUNTNAME	LAACCOUNTS	BANK	开户银行	输入		VARCHAR2	120
ACOLNAME	LAACCOUNTS	DESTORYDATE	销户日期	输入		DATE	7
ACTIVE	LAACCOUNTS	MAKEDATE	入机日期	系统		DATE	7
ACTIVECONTEXT	LAACCOUNTS	MAKETIME	入机时间	系统		VARCHAR2	8
ACTIVEFLAG	LAACCOUNTS	MODIFYDATE	上一次修改日期	系统		DATE	7
ACTIVEMAN	LAACCOUNTS	MODIFYTIME	上一次修改时间	系统		VARCHAR2	8
ACTYPE	LAACCOUNTS	OPENDATE	开户时间	输入		DATE	7
ADDBONUS	LAACCOUNTS	OPERATOR	操作人员	系统		VARCHAR2	6
ADDCOUNT	LAACCOUNTS	STATE	状态	输入		VARCHAR2	2
ADDCOUNT	LAACCOUNTS	STATE	状态	输入		VARCHAR2	2
ADDDEPSDY	LAAGENT	AFFIX	附件	输入		VARCHAR2	100
ADDFYC	LAAGENT	AFFIXNAME	附件名	输入		VARCHAR2	40
ADDRESS	LAAGENT	AGE	代理人年龄	输入		NUMBER	22
ADDVACFLAG	LAAGENT	AGENTCODE	代理人编码	输入		VARCHAR2	10
AFFIX	LAAGENT	AGENTGROUP	代理人展业机构代码	输入		VARCHAR2	12
AFFIXNAME	LAAGENT	AGENTGROUP	代理人展业机构代码	输入		VARCHAR2	12
AGE	LAAGENT	AGENTKIND	代理人类别	输入	01-银代经理02-银代	VARCHAR2	6
AGENTCODE	LAAGENT	AGENTSTATE	代理人状态	输入	代理人状态(01-增员	VARCHAR2	2
AGENTCODE	LAAGENT	APPROVEDATE	档案调入日期(复核日期)	输入		DATE	7
AGENTCODE	LAAGENT	APPROVER	复核员	输入		VARCHAR2	10
AGENTCODE	LAAGENT	ARCHIEVCODE	档案编码	输入		VARCHAR2	20
AGENTCODE	LAAGENT	ASSUMONEY	保证金	输入		NUMBER	22
AGENTCOM	LAAGENT	BANKACCNO	银行帐户	输入		VARCHAR2	50
AGENTCOM	LAAGENT	BANKADDRESS	开户银行地址	输入		VARCHAR2	100
AGENTCOM	LAAGENT	BANKCODE	银行编码	输入		VARCHAR2	20
AGENTCUSCODE	LAAGENT	BIRTHDAY	出生日期	输入		DATE	7
	LAAGENT	BLOODTYPE	血型	输入	A/B/AB/O	VARCHAR2	2
	LAAGENT	BP	传呼	输入		VARCHAR2	20
	LAAGENT	BRANCHCODE	代理人组别	输入		VARCHAR2	12
	LAAGENT	BRANCHTYPE	展业类型	输入	展业类型(1-个人营	VARCHAR2	2

表名	表中文含义	归属	是否必填	是否主键
LAAssessIndex	指标定义表	销管	配置表	是
LAAssessMain	代理人考评主表	销管	交易数据表	是

- ❖ 配合新核心建设，整理现有数据对象及资源
- ❖ 为系统设计、业务建模、数据建模提供标准化素材和参考
- ❖ 为数据迁移提供依据

数据模型整理 - 老数据字典

【数据表】
 表中文名、所属模块、表类型、活跃度、典型使用场景、表间关系 (ER图)。

【数据字段】
 字段含义、标准化名称、数据类型 (长度/精度)、码表取值、技术规则、业务规则、使用该字段的表、字段来源。

《新核心标准化字典 - 基础词汇》的应用 (一)

1	中文标准化单词	英文标准化单词	英文标准化缩写	中文单词同义词
9	保单	POLICY	POL	保险单
10	保单号	POLICY NUMBER	POLNO	
11	保额	SUMINSURED	SUMINSUR	保险金额
12	保费	PREMIUM	PREM	
13	保全	POLICY SERVICE	PS	保单服务
14	保险	INSURANCE	INSUR	
15	保险标的	OBJECT INSURED	INSUROBJ	
16	保险公司	INSURANCE COMPANY	INSURCOM	
17	保险合同	INSURANCE CONTRACT	INSURCONT	
18	保险计划	INSURANCE PLAN	PLAN	
19	保险期间	INSURANCE PERIOD	INSURPERIOD	保险期限
20	保险账户	INSURANCE ACCOUNT	INSURACC	
21	保障	PROTECTION	PROTEC	
22	保障计划	PROTECTION SCHEME	PROTEC_SCHEME	
23	保证	GUARANTEE	GUARANT	担保
24	保证金	CAUTION MONEY	CAUTIONMNY	
25	报告	REPORT	RPT	上报
26	报名	SIGN UP	SIGNUP	
27	备份	BACKUP	BAK	
28	备用	RESERVED	RESRVD	保留,预留
29	备注	REMARK		
30	倍数	MULTIPLE	MULTI	
31	被保险人	INSURED	INSURD	被保险人,出险人

基础词汇

保单管理 产品工厂 销售管理 单证管理 反洗钱 收付费 渠道接入 ... +

分词标准化实践* (1人做全程+工具支持)

- 分词 (术语拆分的断点、粒度)
- 汉译英 (英文选取、中英文同义词处理)
- 多个中 (英) 词汇对应同一英 (中) 词汇的处理
- 副词、介词、数量词的提取及表示*
- 全称 vs. 缩写、缩写程度*
- 单义词 vs. 复合词*
- 对象名称超长的截短处理
- 对象名称的标准化审核

分词结果取决于分词人的 :

- 业务知识
- 词汇量
- 工作经验
- 教育背景
- 常识
- 个人习惯

数据治理与标准化 - 成果展示 (3/7)



《新核心标准化字典 - 基础词汇》的应用 (二)

系统类别	中文全称	中文简称	英文全称	英文简称	
基础应用 (9个)	运营支撑 (7个)	渠道	岗位	岗位编码	岗位全称 (英文)
		个险	总公司个险人管岗	PI-HOHM	PersonalInsurance-HeadOfficeHumanManagement
总公司个险业管岗			PI-HOBM	PersonalInsurance-HeadOfficeBusinessManagement	
分公司个险人管岗			PI-BOHM	PersonalInsurance-BranchOfficeHumanManagement	
分公司个险业管岗			PI-BOBM	PersonalInsurance-BranchOfficeBusinessManagement	
中支综合内勤岗			PI-CBCIW	PersonalInsurance-CenterBranchComprehensiveInternalWork	
专项负责人(钻石、标准组等方案)			PI-SPR	PersonalInsurance-SpecialProjectResponsible	
营服内勤岗	PI-SDIW		PersonalInsurance-ServiceDepartmentInternalWork		
项目实施 (7个)	分析决策 (5个)	总公司个险培训岗	PI-HOT	PersonalInsurance-HeadOfficeTraining	
		分公司个险培训岗	PI-BOT	PersonalInsurance-BranchOfficeTraining	
		IT运维岗	PI-ITOM	PersonalInsurance-ITOperationMaintenance	
	代理人(预留权限, 暂不使用)	PI-AGNT	PersonalInsurance-Agent		
	个险/团险核心 (5个)	续期	总公司续期人管岗	RN-HOHM	Renew-HeadOfficeHumanManagement
总公司续期业管岗			RN-HOBM	Renew-HeadOfficeBusinessManagement	
分公司续期人管岗			RN-BOHM	Renew-BranchOfficeHumanManagement	
分公司续期业管岗			RN-BOBM	Renew-BranchOfficeBusinessManagement	
中支续期内勤岗			RN-CBIW	Renew-CenterBranchInternalWork	
渠道接入 (4个)	新核心项目整体名称	总公司续期培训岗	RN-HOT	Renew-HeadOfficeTraining	
		分公司续期培训岗	RN-BOT	Renew-BranchOfficeTraining	
		总公司团险人管岗	GI-HOHM	GroupInsurance-HeadOfficeHumanManagement	
		总公司团险业管岗	GI-HOBM	GroupInsurance-HeadOfficeBusinessManagement	
团险	分公司团险人管岗	GI-BOHM	GroupInsurance-BranchOfficeHumanManagement		
	分公司团险业管岗	GI-BOBM	GroupInsurance-BranchOfficeBusinessManagement		
	中支团险人管岗	GI-CBHM	GroupInsurance-CenterBranchHumanManagement		
	客户经理(预留权限, 暂不使用)	GI-CM	GroupInsurance-CustManager		
	分公司团险企划培训岗	GI-BOPT	GroupInsurance-BranchOfficePlanningTraining		

《新核心标准化字典 - 基础词汇》

的广泛应用

- 数据库对象 (表、字段、主外键、序列、同义词、索引等)
- 新核心码表项名称
- 子系统 - 原子服务名称
- 子系统 - 重要内容编码
- 新核心项目名称

新核心业务系统 - 码表 - 公共代码集、私有代码集 (业务层面整合)

中华人民共和国金融行业标准 保险业务代码集——通用代码

- GB/T 2260 中华人民共和国行政区划代码
- GB/T 2261.1-2003 个人基本信息分类与代码 第1部分:人的性别代码
- GB/T 2261.2-2003 个人基本信息分类与代码 第2部分:婚姻状况代码
- GB/T 2261.3-2003 个人基本信息分类与代码 第3部分:健康状况代码
- GB/T 2261.4-2003 个人基本信息分类与代码 第4部分:从业状况(个人身份)代码
- GB/T 2659 世界各国和地区名称代码
- GB/T 4658-2006 学历代码
- GB/T 4754-2002 国民经济行业分类
- GB/T 8561-2001 专业技术职务代码
- GB 11714 全国组织机构代码编制规则
- GB/T 12402 经济类型分类与代码
- GB/T 12406 表示货币和资金的代码
- GB/T 16835-1997 高等学校本科、专科专业名称代码
- GA 16.1-2003 道路交通事故信息代码 第1部分:道路交通事故等级代码
- GA 16.2-2003 道路交通事故信息代码 第2部分:事故形态代码
- GA 16.6-2003 道路交通事故信息代码 第6部分:当事人伤害程度代码

公共编码

通用性较强，参考“国家标准”、“行业标准”、“保险业务代码集”，编入《新核心-通用编码集》。

(如：证件类型、性别、婚姻状况、行业)

私有编码

业务性较强，虽沿用现有码表，但需规范化整合，编入《新核心-私有编码集》。

(如：销售方式、保单类型、保单状态)

规范化整合及标准化

- 沿用准确项，标准化名称
- 剔除明显垃圾项
- 替换、扩充取值范围
- 合并同类项，消除冗余

数据治理与标准化 - 成果展示 (5/7)

新核心业务系统 - 数据码表规范 —— 准入标准、命名规范、取值编码原则、码表结构、分发策略

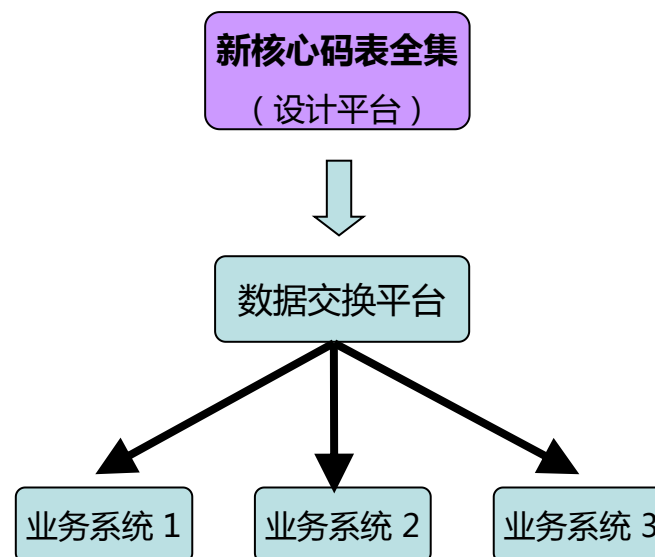
农银人寿保险股份有限公司
ABC LIFE INSURANCE CO.,LTD.



编号	
级别	<input type="checkbox"/> 限制 <input checked="" type="checkbox"/> 内部 <input type="checkbox"/> 公开
当前版本	V1.0
发布日期	
受控状态	<input checked="" type="checkbox"/> 受控 <input type="checkbox"/> 非受控

农银人寿新核心业务系统项目
数据码表规范

- 码表准入原则
- 编码项命名规范
- 码表数据结构
- 编码项归属域划分
- 编码项取值编码原则
- 码表分发策略
- 公司内外码转换原则



数据码表 - 整合与标准化

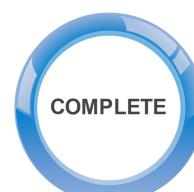
- 屏显、出单打印的信息更加准确
- 农银内部、新核心各子系统之间无需转码
- 便于开发运维人员理解

数据治理与标准化 - 成果展示 (6/7)



编号	
级别	<input type="checkbox"/> 限制 <input checked="" type="checkbox"/> 内部 <input type="checkbox"/> 公开
当前版本	V1.0
发布日期	
受控状态	<input checked="" type="checkbox"/> 受控 <input type="checkbox"/> 非受控

新核心业务系统 - 数据库设计规范



编号	
级别	<input type="checkbox"/> 限制 <input checked="" type="checkbox"/> 内部 <input type="checkbox"/> 公开
当前版本	V1.0
发布日期	
受控状态	<input checked="" type="checkbox"/> 受控 <input type="checkbox"/> 非受控

新核心业务系统 - 数据模型设计规范

■ 总体设计原则

按需设计、规范、完整、高内聚、低耦合、安全

■ 逻辑对象设计

■ 数据库对象命名

■ 物理结构设计

参数配置、数据块、数据文件、连接数、游标、优化模式

■ 数据库安全

用户管理、角色及权限、应用级设计、密码管理

■ 数据库优化

SQL优化、反规范化设计、并行查询

■ 数据生命期管理

数据分类、数据迁移、数据归档、数据分层存储

■ 数据库使用

■ 设计目标、设计输入、技术工具

■ 数据模型-设计模式

主扩展模式、主从模式、名值模式、多对多模式、继承模式、自联结模式、单表模式

■ 数据模型-设计过程

需求分析、业务模型分析、确认核心实体、确认实体属性、范式化设计、逻辑设计评审、模型转化、表设计、反范式设计、物理设计评审

■ 典型场景-模型设计

审批类、多对多、业务日志类、操作日志类、实体状态分离、历史数据存储、参数类、账户类、删除操作、日期时间、货币金额、联系信息、树形关系、接口表

数据治理与标准化 - 成果展示 (7/7)



编号	
级别	<input type="checkbox"/> 限制 <input checked="" type="checkbox"/> 内部 <input type="checkbox"/> 公开
当前版本	V1.0
发布日期	
受控状态	<input checked="" type="checkbox"/> 受控 <input type="checkbox"/> 非受控

新核心业务系统 - 主数据管理方案



编号	
级别	<input type="checkbox"/> 限制 <input checked="" type="checkbox"/> 内部 <input type="checkbox"/> 公开
当前版本	V1.0
发布日期	
受控状态	<input checked="" type="checkbox"/> 受控 <input type="checkbox"/> 非受控

新核心业务系统 - 数据迁移总体方案

■ 主数据概念

定义、特征、范围、识别分析

■ 主数据管理

定义、体系规划、方法论、现状调研、需求分析、技术工具

■ 必要性及意义

■ 关键技术

数据标准、分类编码、SOA架构、数据仓库

■ 主数据生命期管理

数据模型、业务管理、数据清洗、质量管理、安全管理

■ 主数据应用

业务系统集成、主数据管理与数据挖掘、主数据管理与大数据

■ 方案概述

基本原则、实施策略、基本步骤、技术工具、若干阶段、相关方职责

■ 准备工作

迁移范围确定（迁移范围、不迁移范围）、数据模型分析（源数据模型分析、新数据模型分析、综合分析）

■ 实施步骤

任务列表、环境准备、源数据清洗、数据检查及异常处理、数据映射转换、数据迁移、数据核对

■ 关键点控制

时间窗口、映射正确性检验、迁移演练

■ 应急处理

风险控制、系统回滚、版本管理等



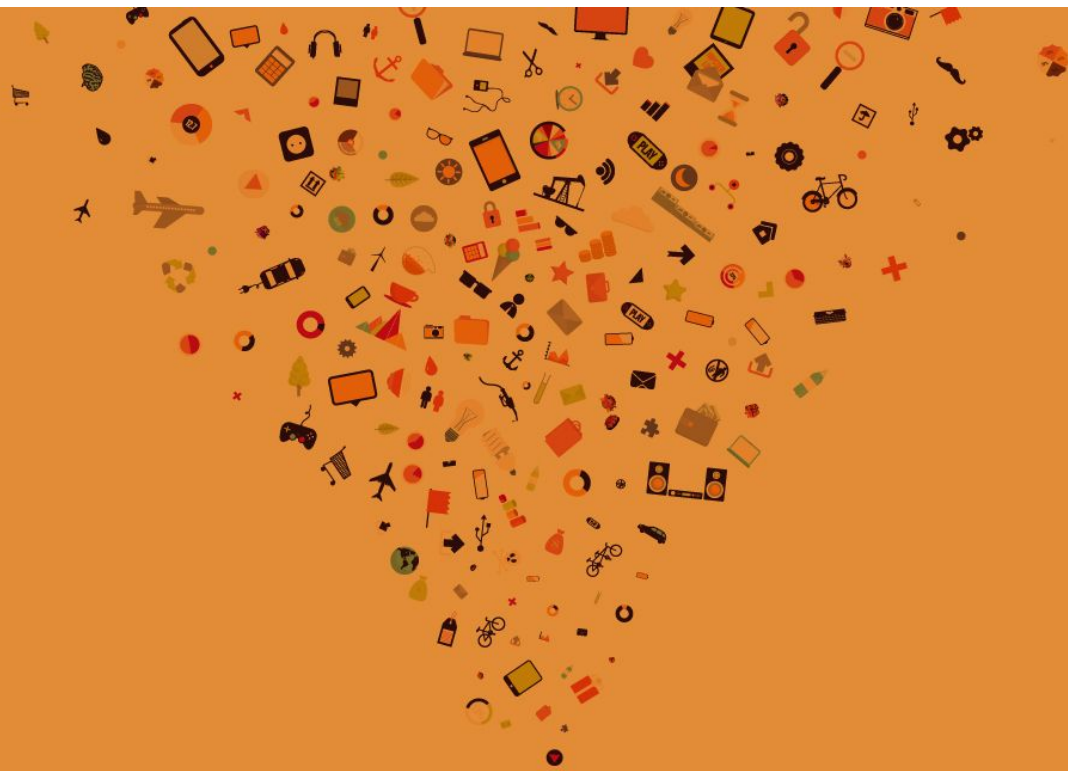
**数据治理是一项长远的工作，
我们任重而道远！**

技术层面

- 数据质量持续提升（准确性、一致性、完整性）
- 数据生命期更加清晰（数据产生及变迁的追溯）
- 系统运行更加平稳、顺畅

业务层面

- 数据资源转变为战略资产
- 减少运营成本，降低经营风险
- 支持业务运营、高层决策
- 增强可审计性、安全合规性
- 发挥数据价值，增加企业利润



THANKS

SequeMedia
盛拓传媒

IT168.com

ChinaUnix

ITPUB