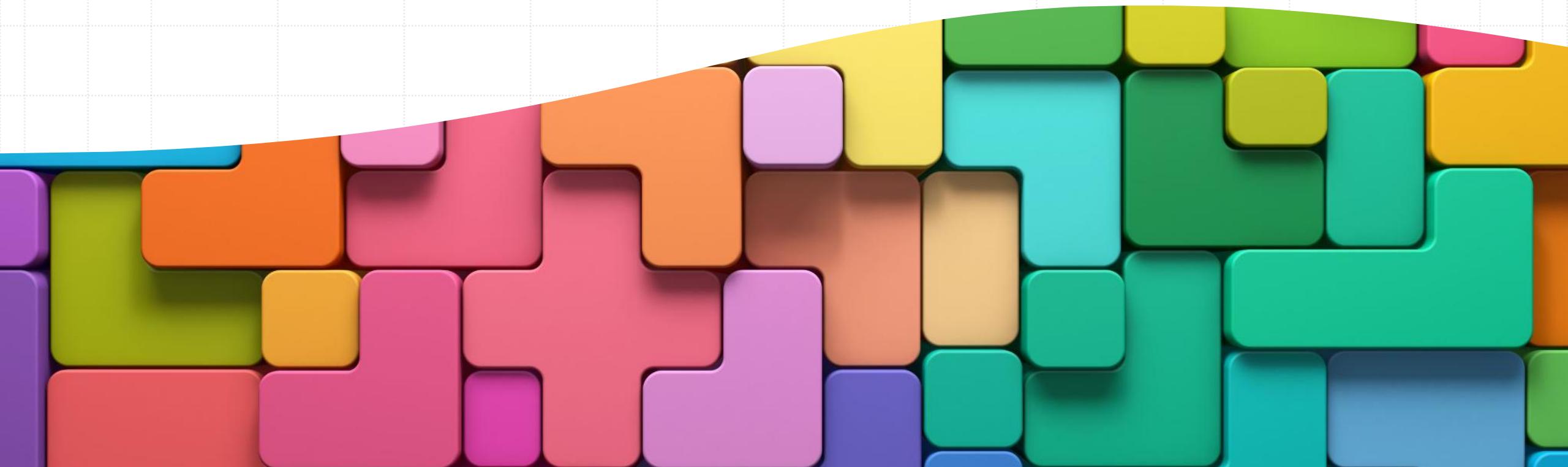


# 企业数字化转型的数据治理概述



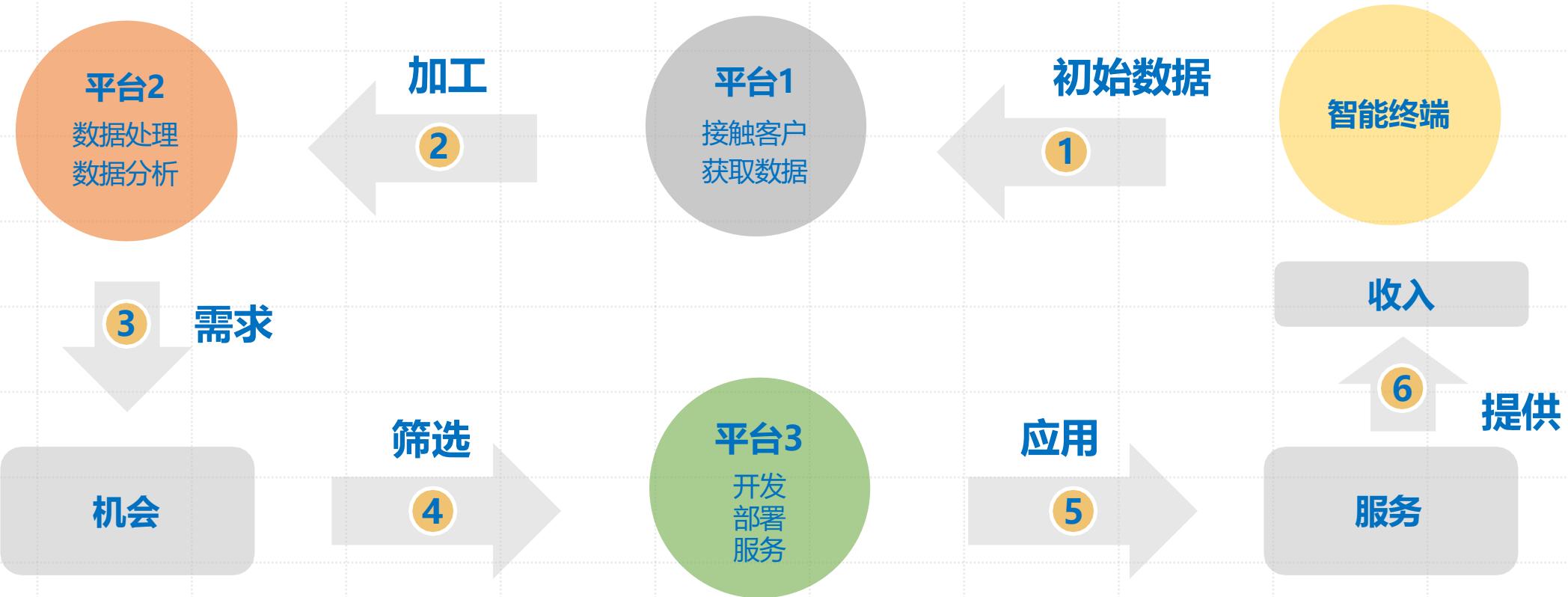
# 挑战1：企业现有数据融合 “价值有限”



- 现有数据融合：只有约10%的管理数据被打通和融合，进行数据挖掘、创新
- 物理世界数字化：约90%的OT数据并未被广泛的感知、采集、融合和价值挖掘

# 在今天，数字化生产已逐步成为普遍商业模式

数字化生产模式：以数据为处理对象，以ICT平台为生产工具，以软件为载体，以服务为目的的生产过程



**商业循环：数据变机会、机会变服务、服务变收入**

ICT : Information and Communication Technologies 信息与通讯科技

# 人工智能销售预测模型的建立

数据要准确

考量数据特性

预测目标

计算资源

业务需求

模型的评估和调优；交叉验证、超参数调整

# 人工智能销售预测模型的建立

## 1. 传统时间序列分析 - 纪要历史数据来预测未来趋势

自回归 (AR)、移动平均 (MA)、自回归移动平均 (ARMA)、季节性自回归移动平均

## 2. 机器学习模型

线性回归

决策树、随机森林（处理非线性关系）

## 3. 深度学习模型

循环神经网络 (RNN)、卷积神经网络 (CNN)

## 4. 混合模型

# 数字化转型的核心：“业务优化”和“业务转型”

## Digital Business Strategy

### Digital Business Optimization

Improved Productivity and Existing Revenue

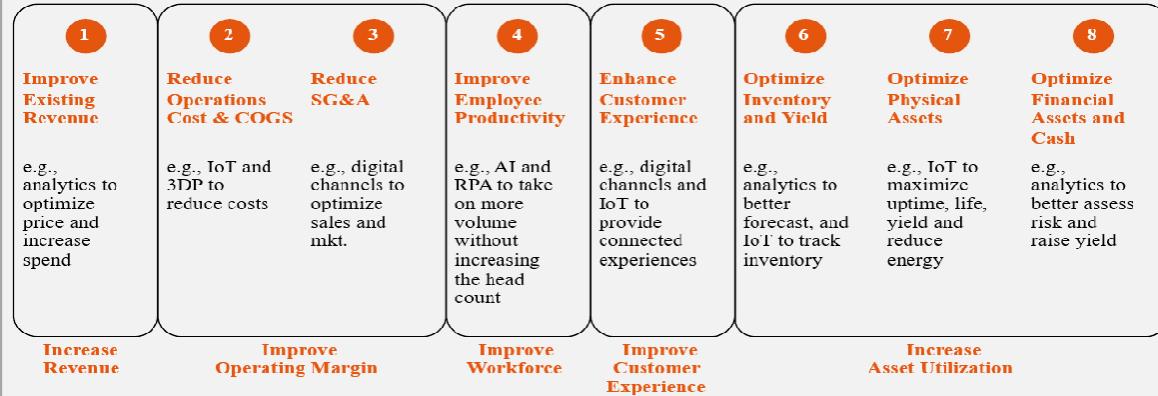
Better Customer Experience

### Digital Business Transformation

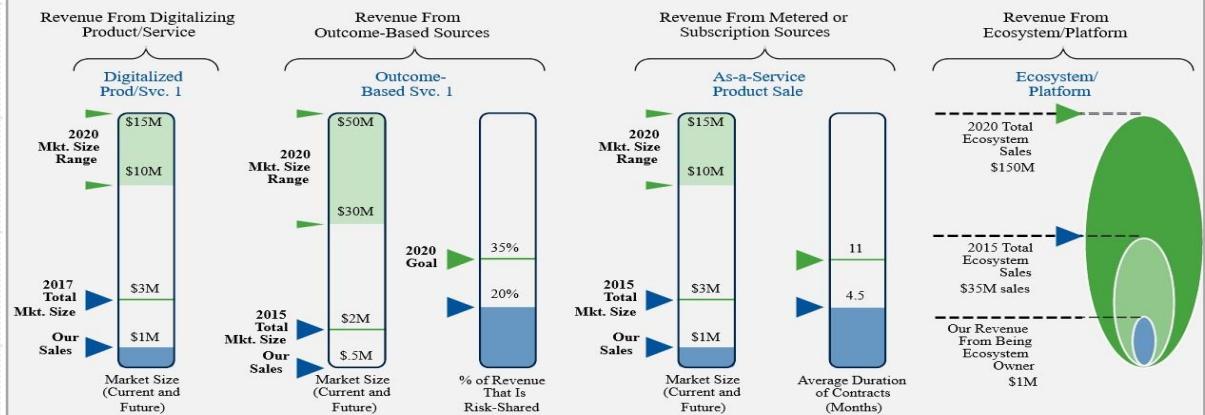
Net-New Revenue Product and Services

New Business Models

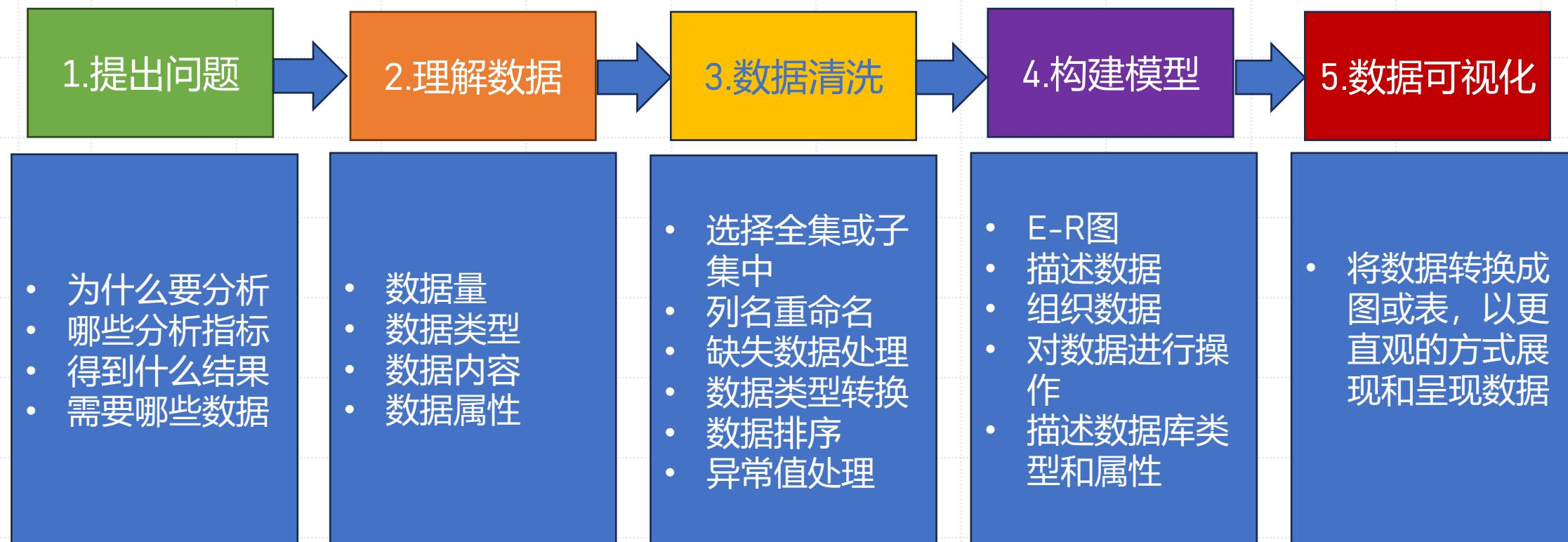
#### Digital Business Optimization Adds Significant Value Without Changing Business Model



#### KPIs for New Revenue and Models

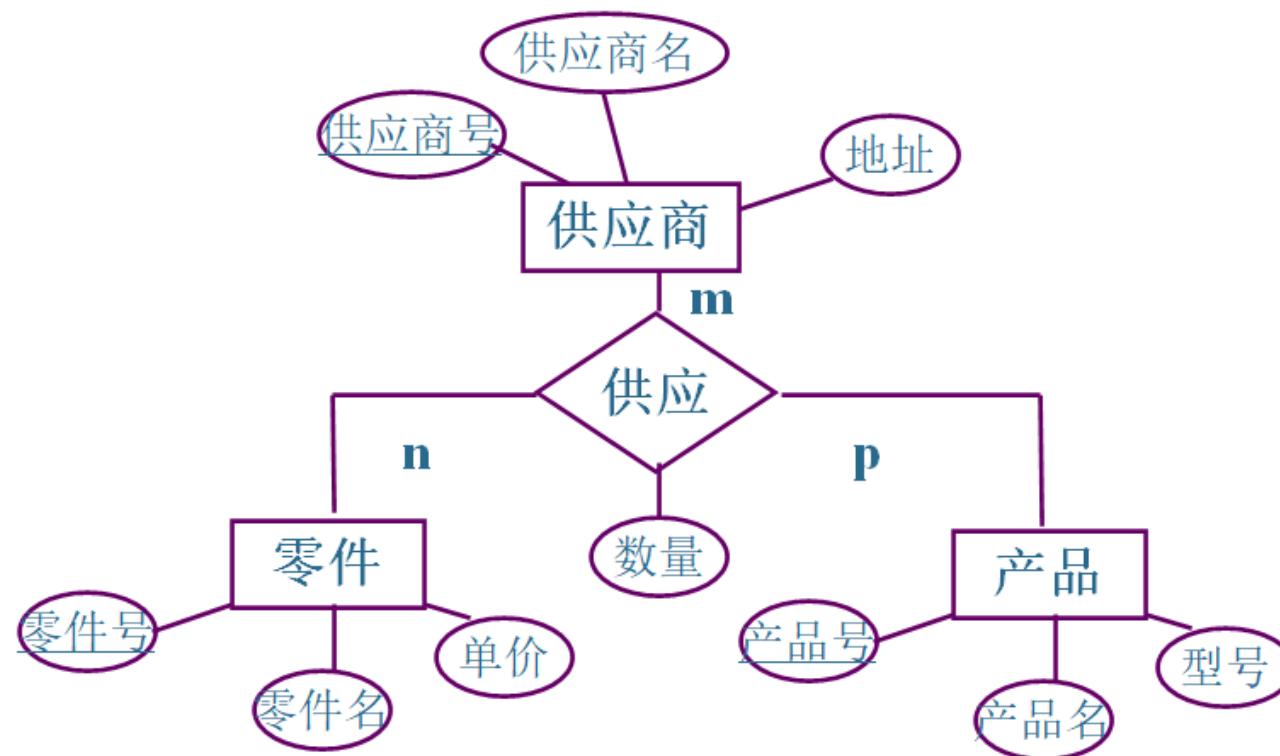


# 数据分析



# E-R图 (实体关系图)

- 产品 (产品号, 产品名, 型号)
- 供应商 (供应商号, 供应商, 地址)
- 零件 (零件号, 零件名, 单价)
- 供应 (供应商号, 零件号, 产品号, 数量)



ER图是一种描述现实世界的概念模型的方法，用来表示实体，属性和联系

# 数据可视化

表格：由行和列组成，用于比较变量。表格以结构化方式展示大量信息。

饼图和堆积条形图：这些图形分成多个部分来表示一个整体的不同部分。它们提供了一种简单的方法来组织数据并比较组件的大小。

线形图和面积图：这些视觉显示通过绘制一段时间内的一系列数据点，来展示一个或多个数量的变化。线形图利用线条展示这些变化，而面积图用线段将数据点连接起来，然后将变量堆叠起来，并用颜色区分不同的变量。

直方图：该图形使用条形图绘制数据分布（条形之间没有间隔），表示属于特定范围的数据数量。这种视觉表示使终端用户容易识别给定数据集内的异常值。

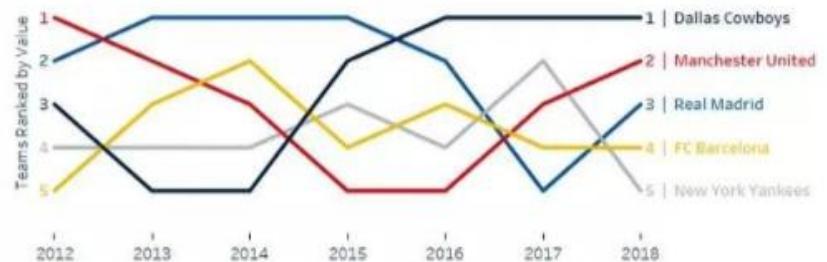
散点图：这些视觉表示对于揭示两个变量的关系十分有用，通常用于回归数据分析。但是，有时可能会与气泡图混淆，气泡图用于通过 x 轴、y 轴和气泡大小来呈现三个变量。

热图：这些图形显示有助于按位置呈现行为数据。位置可能是地图上的地点，甚至是网页。

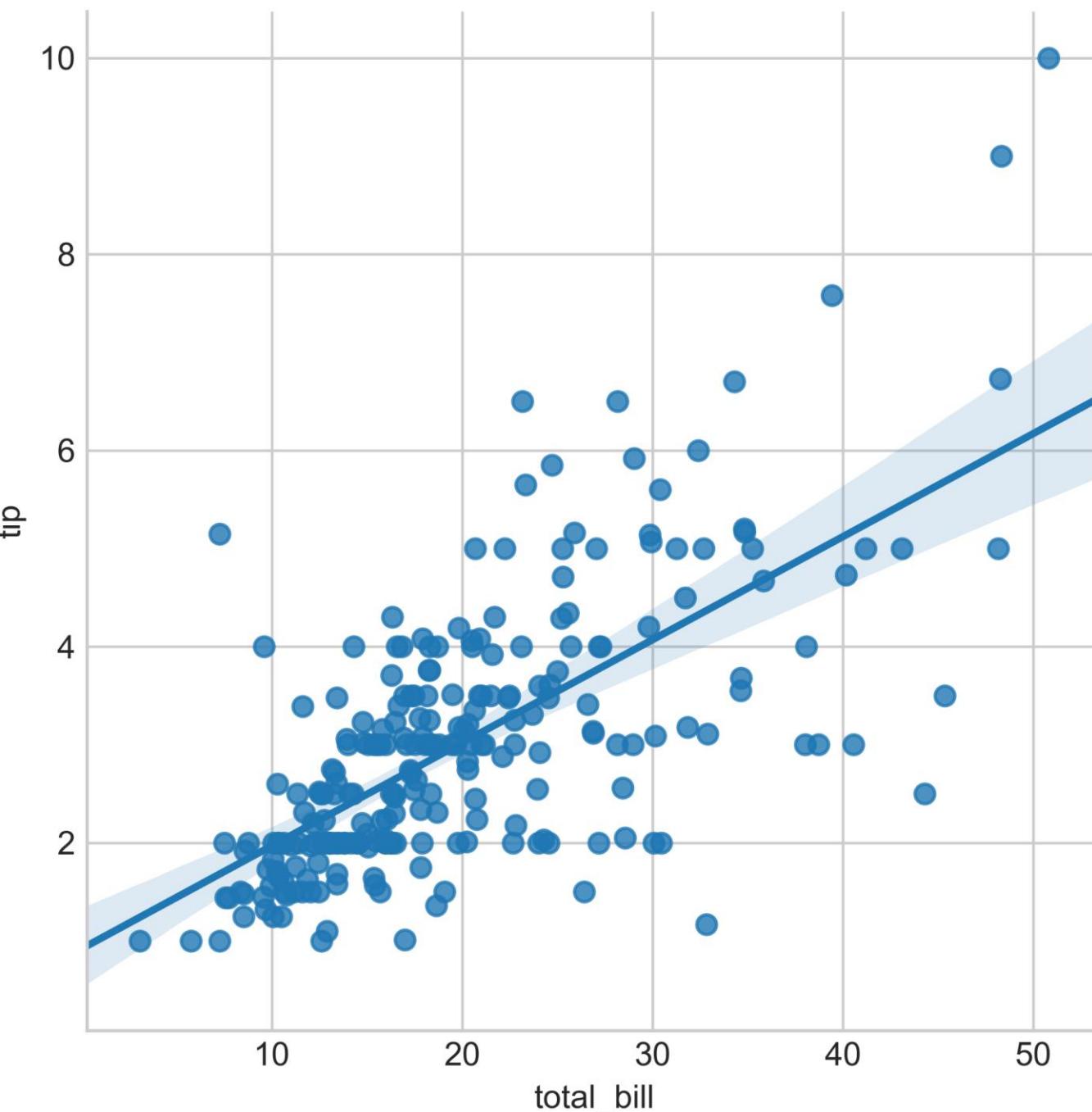
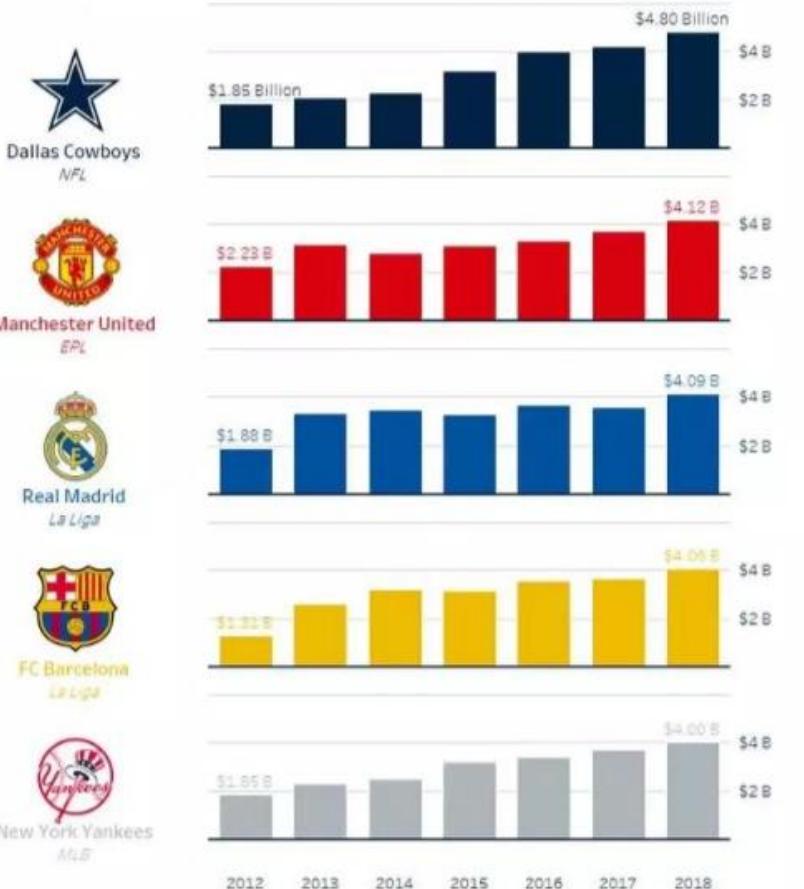
树形图：以一组嵌套形状（通常是矩形）展示层级数据。树形图非常适合根据面积大小比较类别比例。

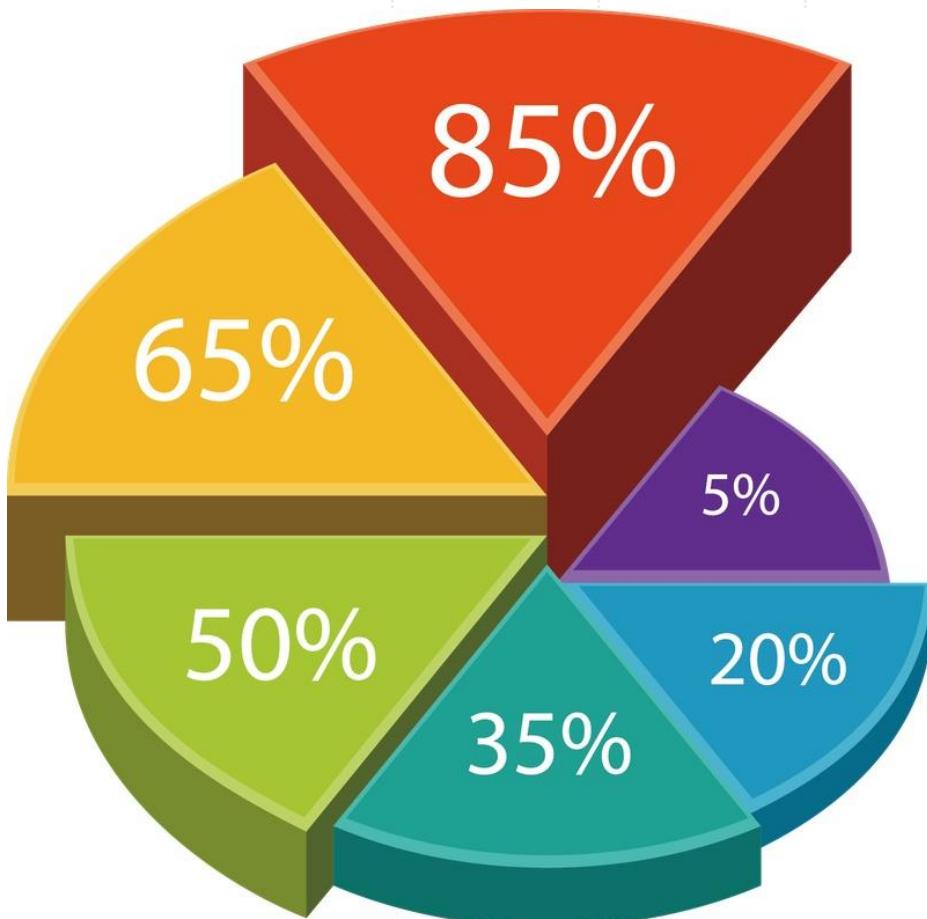
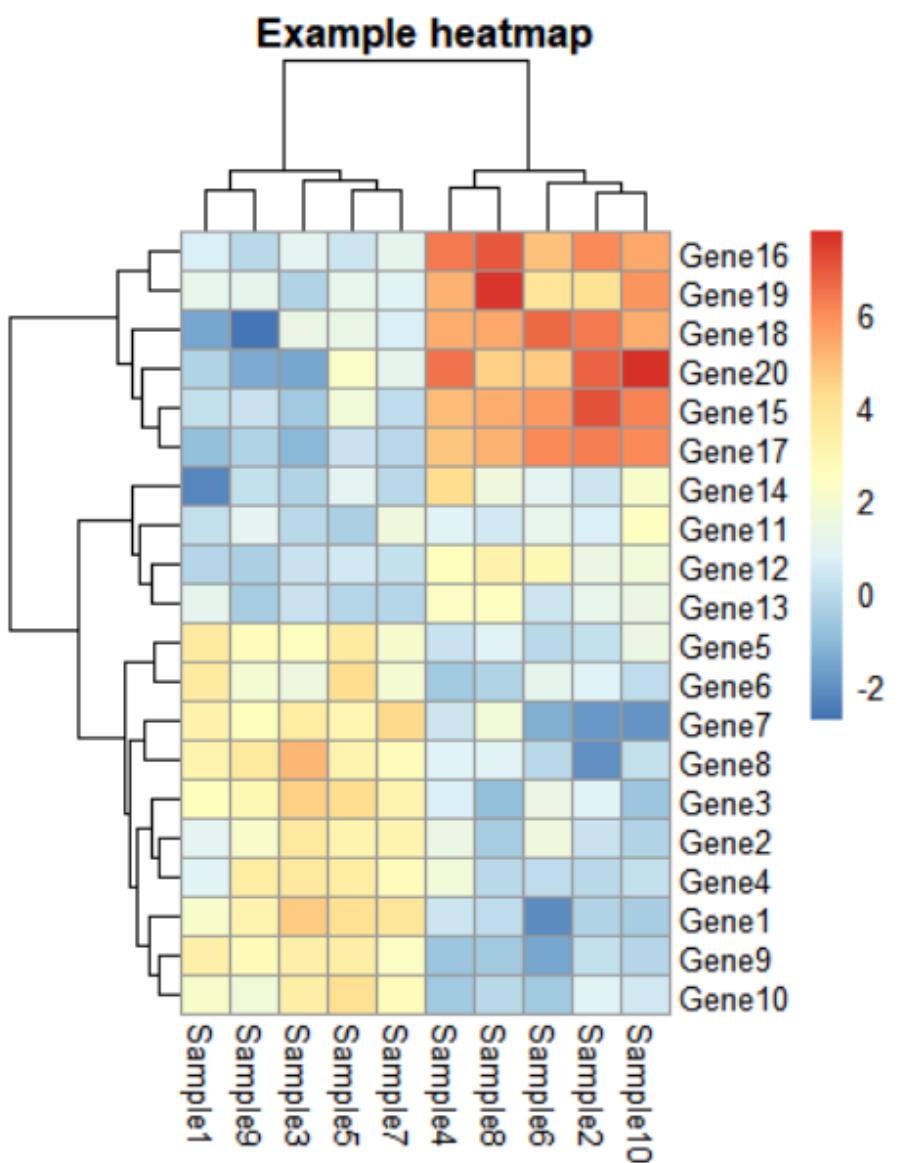
# The Most Valuable Teams

The teams ranked in the top 5 haven't changed for seven years



By 2018, all of the top 5 are valued at over \$4 Billion





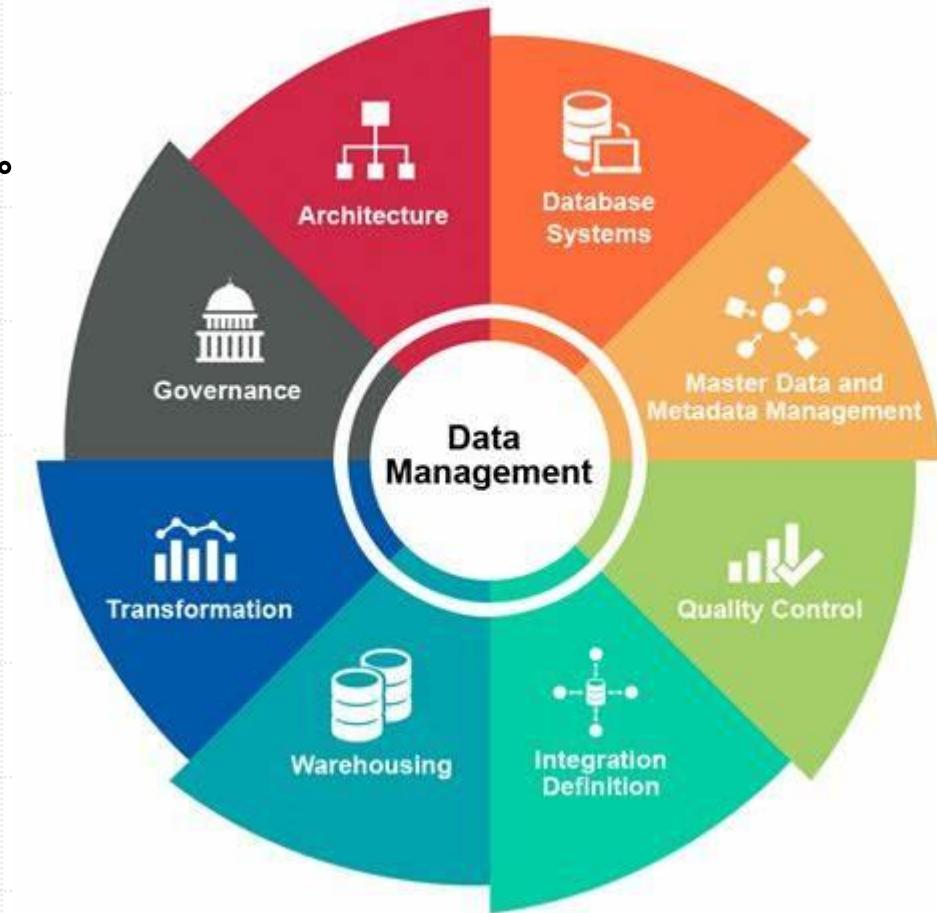
# 什么是数据治理

数据治理是一项战略计划，用于优化企业处理数据的方式。

它旨在对数据资产进行规划、监控、执行、管理；

组织和改进公司用于定义、收集、存储、保护、管理和货币化业务数据的政策和程序，并有目的性的实践活动；

良好的数据治理不仅旨在保护数据，而且旨在寻找为企业创造数据价值的新方法。



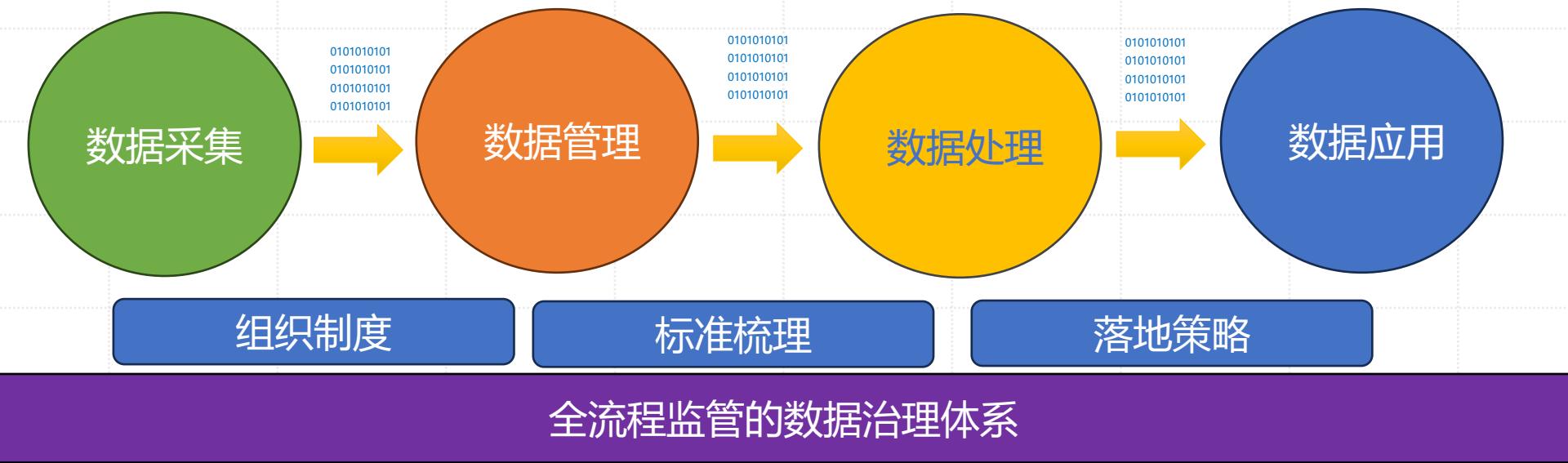
# 什么是数据治理

战略规划  
组织和角色  
理念植入  
政策和标准  
项目和服务  
技术支持  
成本和资源

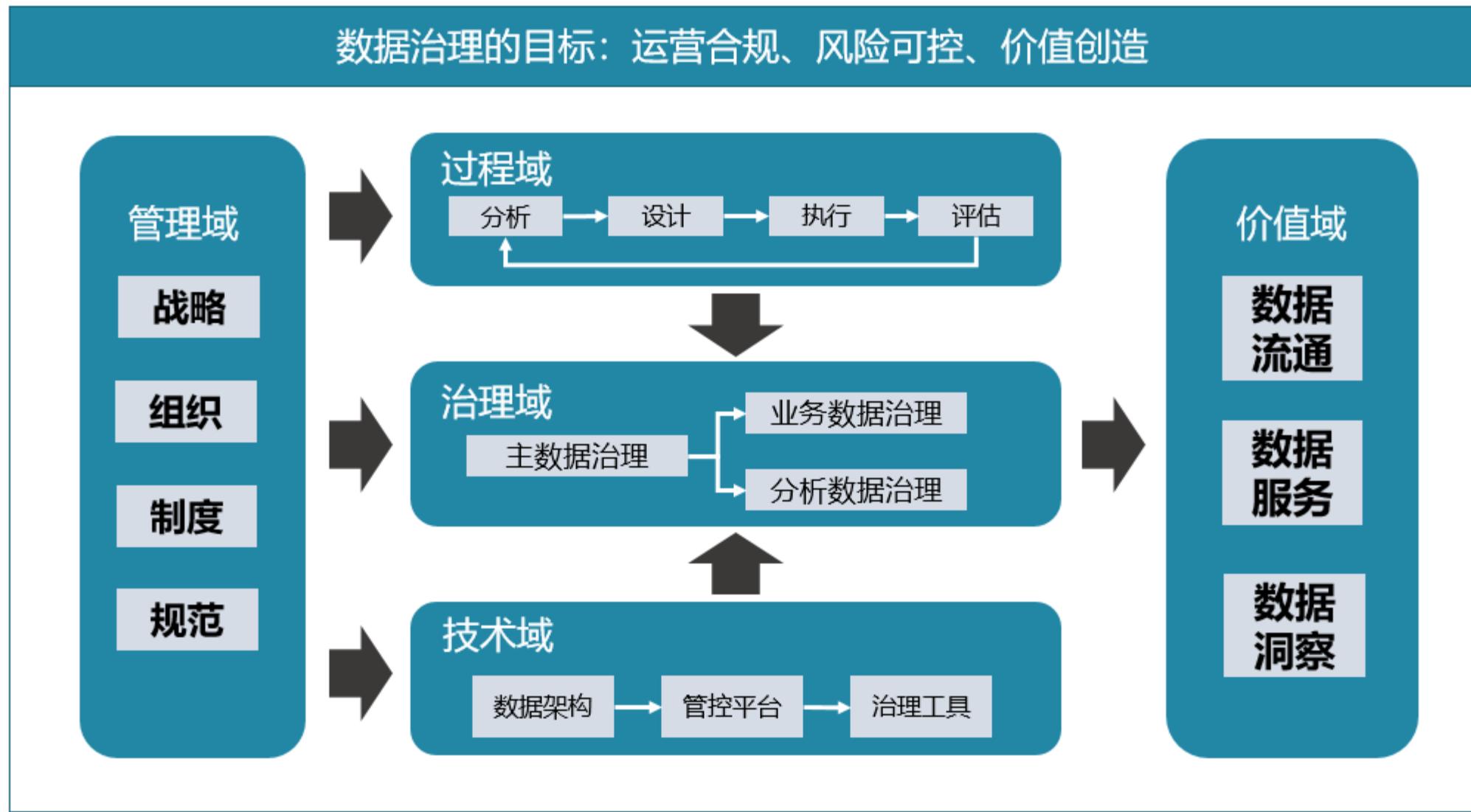


# 全流程监管的数据治理体系

- 关键数据资产有清晰的业务管理责任
- IT建设有稳定的原则和依据
- 作业人员有规范的流程和指导
- 当面临争议时，有裁决机构和升级处理机制；
- 治理过程所需的人才、组织、预算有充足的保障

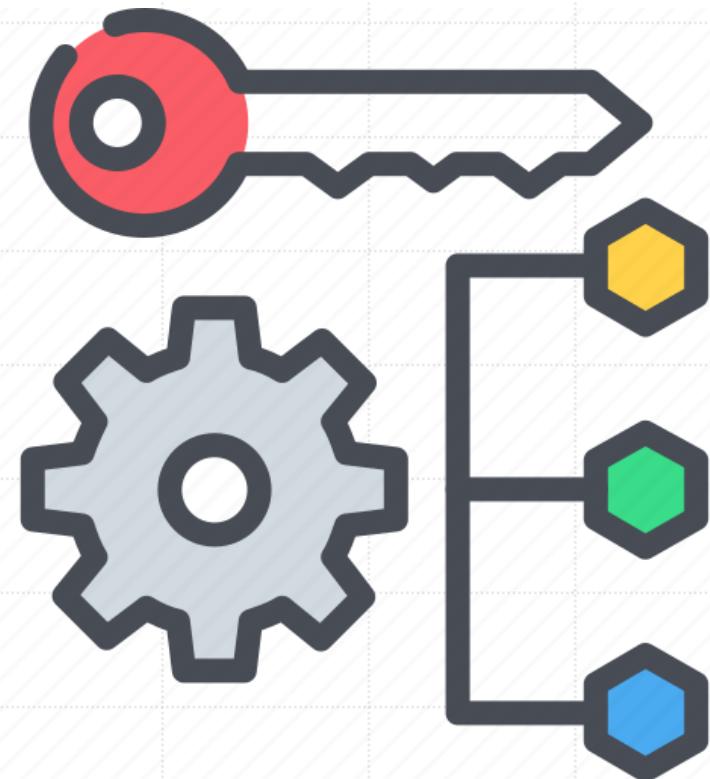


## 数据治理体系搭建



# 数据治理的核心工作

在企业的数据建设进程中，  
保障企业的数据资产得到  
**正确有效地管理。**



# 什么是数据管控

数据治理和数据管控缺一不可，治理在前、管控在后。

数据治理针对的是存量数据，是个由乱到治、建立规范的过程；

数据管控针对的增量数据，是个有章可循，行不逾矩的约束。

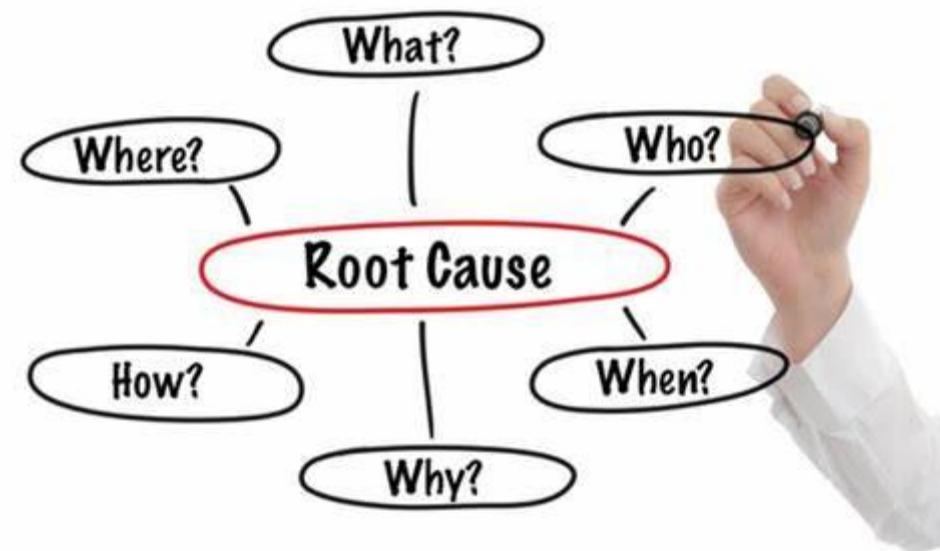
数据管控专注在纵向管控深度，上层定标准，基层去执行



# 为什么要做数据治理

## 症状

- 有垃圾数据
- 关键数据缺失
- 数据不同步
- 数据无法追溯
- 数据异常
- 业务和财务口径不一致
- 前端数据和后端数据不一致



## 根因

- 政策不清晰
- 管理力度缺失
- 体系和规模庞大，梳理困难
- 培训不到位
- 主人翁意识不够
- 标准化缺失

# 为什么要做数据治理

- 提升数据质量
- 为共享数据建立清晰的决策规则和决策流程
- 提升数据资产的价值
- 提供解决数据问题的机制
- 促进IT和非IT人员共同参与决策
- 促进部门和业务单元之间的协作和相互依赖
- 为共享数据建立共同责任制

保证数据质量

保证数据的可用性

保证数据安全



# 为什么要做数据治理

当前可能的情况：

- 问题长期积累
- 能容忍
- 重要但不紧急
- IT在管理和做补救



可能采取的模式

- 业务部门主导
- IT部门主导
- 企业数据管理部门主导

# 数据质量维度

- 完备性：是否存在所有必要的数据
- 有效性：数据指与定义的值域一致
- 准确性：数据正确表示“真实”实体的程度
- 一致性：确保数据值在数据集内和数据集之间表达的相符程度
- 完整性：即连贯性，包括与完备性、准确性、一致性相关的想法
- 及时性：数据产生到可用的时间延迟程度
- 唯一性：数据集内的任何实体不会重复出现



# 数据治理的内容

	架构管理	数据开发	操作管理	安全管理	主数据管理	数据仓库管理	内容管理	元数据管理	质量管理
数据模型 价值链分析 数据架构建立	数据分析	获取	标准	外部规范	架构	获取	架构	标准	
	数据建模	恢复	分级	内部规范	实施	存储	整理	规范	
	数据库设计	调优	管理	客户数据	培训和支持	备份和恢复	控制	分析	
	实施	保留	授权	产品数据	监控和调优	整理	交付	度量	
		清除	审计	维度管理		检索		改进	
						保留			
						清除			

# 数据类型

主数据

关于业务实体的数据，描述组织内的“物”，如供应商、客户、产品等。通过全局视角对企业主要的“人”和“物”有个全局把控，为其建立现实和数据领域的唯一标识。

交易数据

即过程数据，描述组织业务运营过程中的内外部时事件或交易记录的数据。如：销售订单，通话记录，供应商名单等数据。交易数据是BI分析的基础

参考数据

用于将其他数据进行分类或者标记整理的数据，外部对标数据

元数据

即描述数据的数据，帮助理解、获取、使用数据，分为技术元数据、业务元数据等

统计数据

即对企业业务活动进行统计分析的数值型数据，也就是指标，如年度销售额

# 主数据管理制度

明确相关的组织职责、流程规范、数据标准

- 主数据管理办法
- 各类主数据属性模板
- 主数据流程清单
- 绩效考核



# 主数据标准的梳理

编码：

建立适用全企业的编码规则，全局统一管理、辨别和使用

遵循全局性、唯一性、适度性、灵活性、扩展性等编码原则

满足编码共享、自动生成、编码扩展等使用要求，

分析现有编码问题，提出改进意见，最终确认主数据编码规则



# 主数据标准的梳理

分类：

建立统一、规范、科学的分类，提升管理效率，降低因分类不准确造成的错误

1. 调研、收集相关分类标准
2. 差异及对标分析
3. 确定信息分类
4. 确定结构及规则

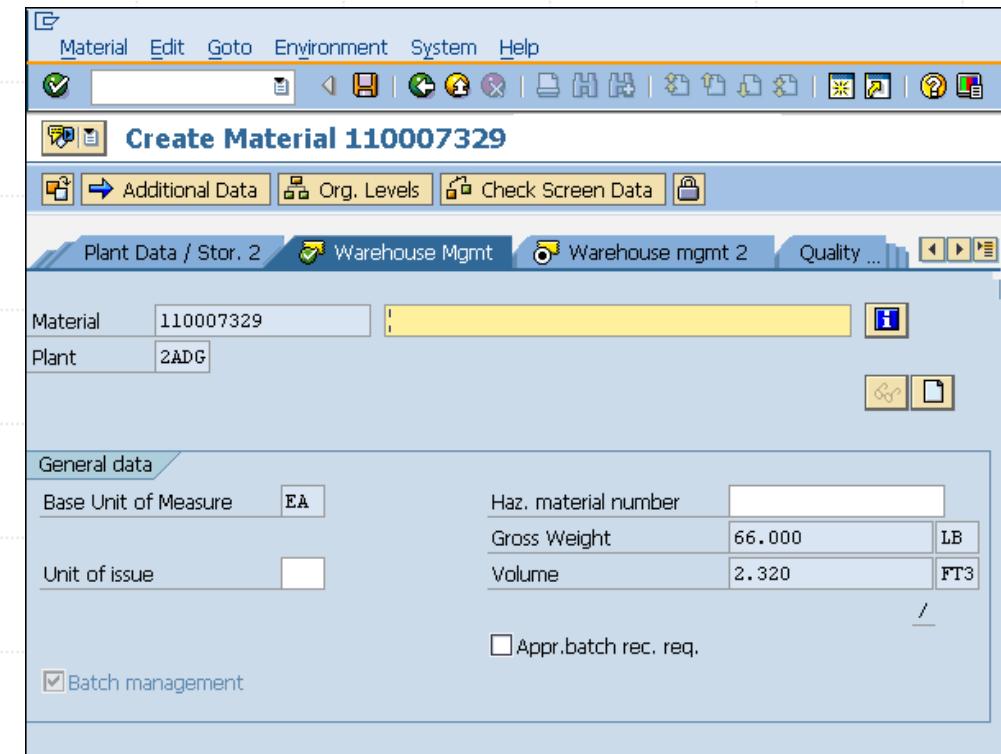
主数据类型	
物料主数据	原料, 辅料, 半成品, 服务
设备主数据	类型, 地点, 资产, 费用
财务主数据	科目, 分录
供应商	类型, 产品, 价值
客户	区域, 规模, 价值

# 主数据标准的梳理

数据三大属性：业务属性、技术属性和管理属性

属性标准的梳理是对数据的每个属性项分别定义相关标准规范，从而可以约束各系统中的属性差异。

属性标准可以参照国家、行业标准，内部的业务制度从业务标准、技术标准、管理标准等不同角度进行标准化。



# 主数据标准的梳理

数据标准核心内容

主题定义  
和分类

- 明确业务主题概念、本质和内涵（如，什么是客户）
- 明确具体的数据来源、分类体系、使用规则

标准属性

- 需要遵守的规则、标准属性、负责部门、业务定义和描述（所属主题、名称、引用的数据项名称、参考标准、数据类型、长度和其他业务定义）

标准代码

- 明确代码取值和业务含义（编号、编码规则、格式、名称、数值、描述等）

业务定义

管理信息

技术属性

谁是主？ 谁来用？ 干什么？ 什么时候干？

准确性，及时性，全面性

# 主数据标准的梳理

## 数据标准的依据

国际标准 (ISO,ITU...)

国家标准

行业标准 (YD, YD/T, JT. . . )

同行标准 (GDW / Q. . . )

企业自定标准

## 数据标准的类型

编码类

代码类

标志类

文本类

金额类

比例类

数值类

日期类

时间类

日期时间类

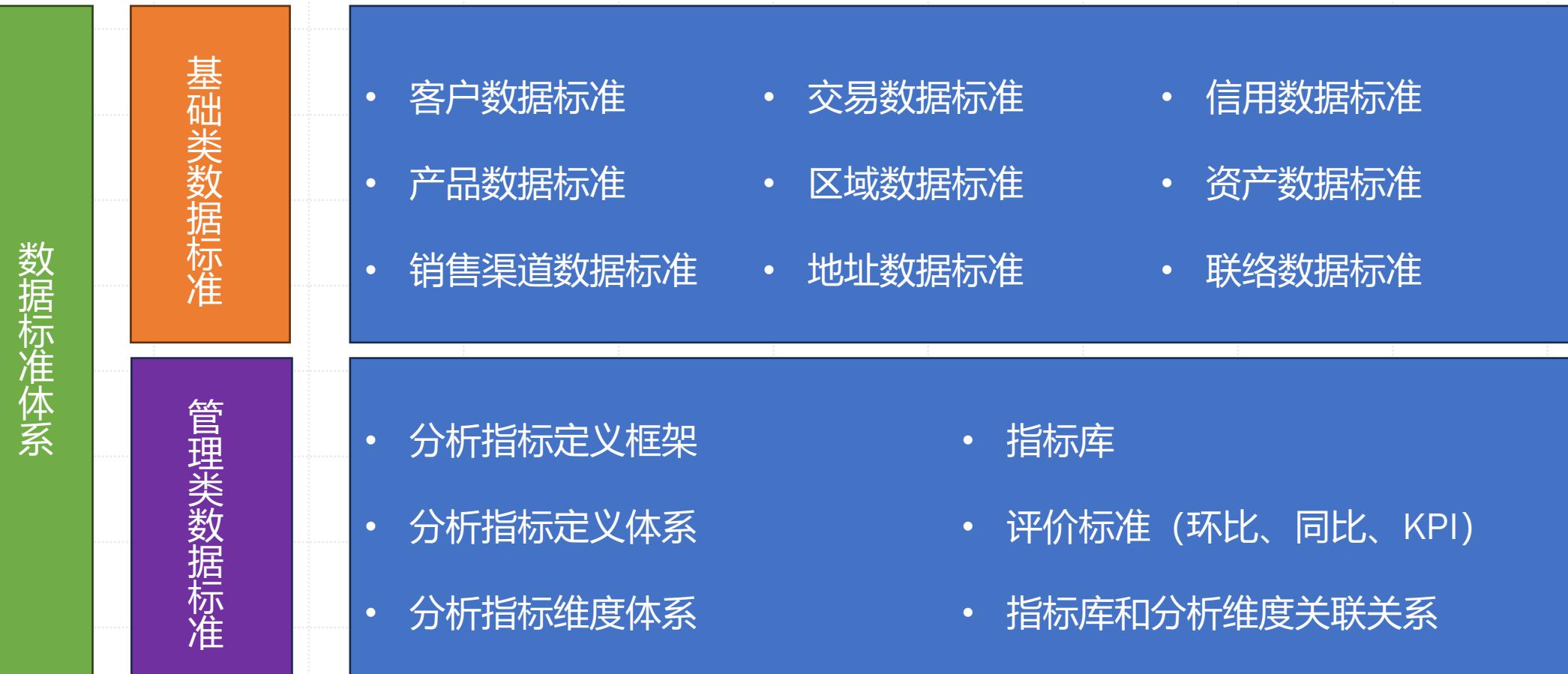
数据标准来源于业务，服务于业务。

# 一张名片的数据

公司	x x x	Company	x x x
部门	x x x	Department	x x x
职位	x x x	Title	x x x
姓名	x x x	Name	x x x
地址	x x x	Address	x x x
电话 1	x x x	PSTN	x x x
电话 2	x x x	Cell phone	x x x
邮件	x x x	Email	x x x
传真	x x x	Fax	x x x
邮编	x x x	Zip code	x x x



# 数据标准体系的建立



# 数据标准生命周期管理



# 数据标准的制定

循序渐进  
不断完善

- 确定可通过数据实现的具体业务目标，并定义实现这些目标所需的特定数据元素
- 标准的制定和维护、标准的治理、完整性和可操作性在标准落地应用过程中逐步完善

战略规划  
实践平衡

- 清晰认定企业战略，标准的制定需要结合企业发展的战略，要有前瞻性
- 数据标准需要平衡收益和成本，需要考量投资回报

外部借鉴  
内部适配

- 数据标准需要参考国际和行业标准，也要兼顾企业内部实际需求
- 内部外部数据的标准接轨需要逐步完成，需要分析目前相关流程规范后做出相应改变

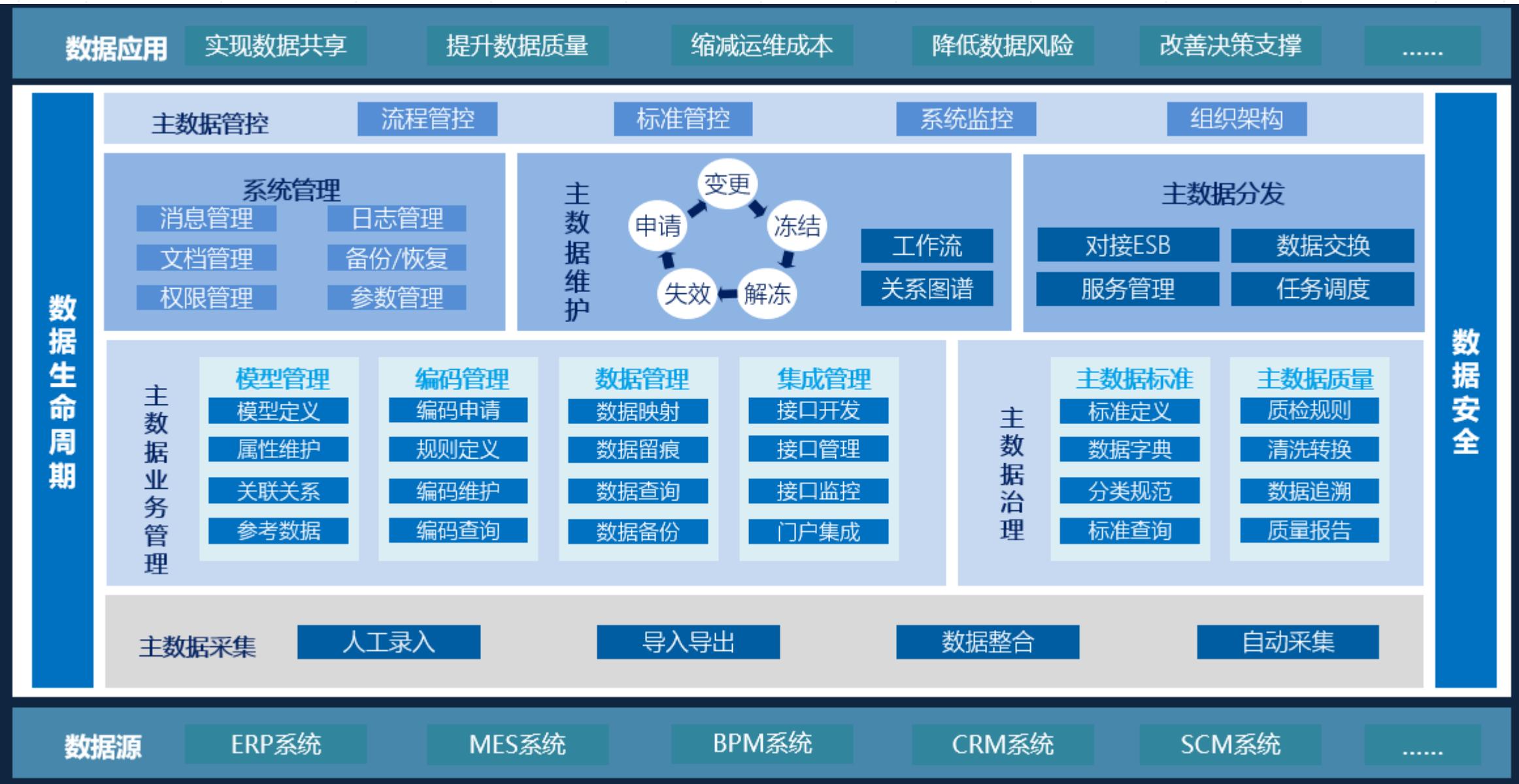
思维改变  
全员参与

- 需要对全员进行指导和提供培训，指定数据所有者和数据管理专员
- 业务需要全面参与标准的制定和实施以及逐步改善

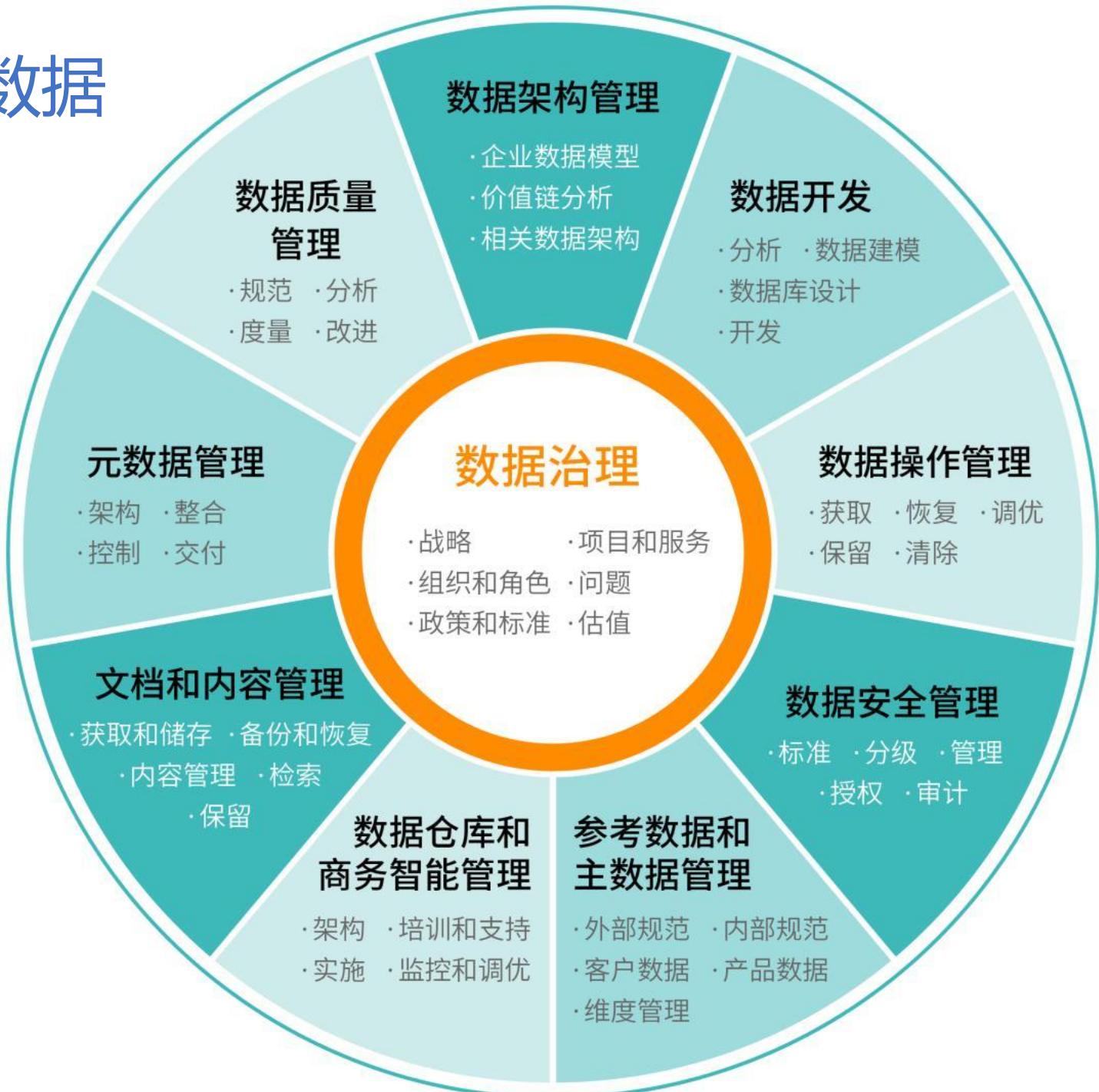
审核到位  
指标落地

- 建立关键绩效指标(KPI)来衡量数据治理计划的成功与否
- KPI与组织的特定公司战略和具体业务目标紧密联系、与时俱进

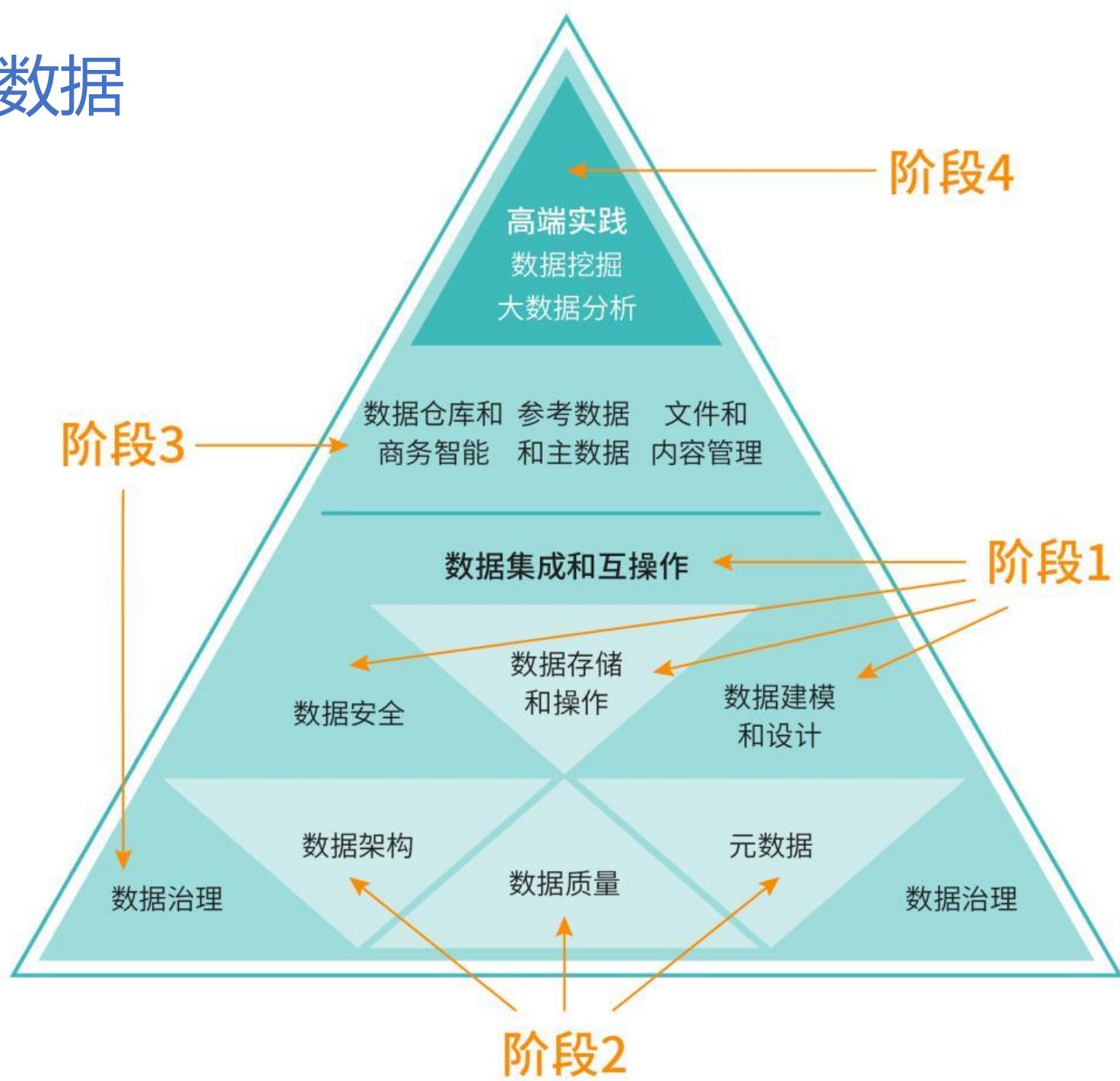
# 主数据管理平台



# DAMA(国际数据管理协会)



# DAMA(国际数据管理协会)



# 数值治理总览



# 数值治理关键点

1. 文化共识，循序渐进
2. 业务情况和能力现阶段调研
3. 纲举目张，明确主数据定义上的责任边界，在企业内部明确岗位职责、管理流程，操作规范，保证核心主数据信息的一致性和共享
4. 根据业务管理和报表需求来合理进行主数据分类
5. 权限控制，严格审计，考核到位



## 创新迭代 持续改善

高速迭代与不确定的环境下，谨慎小心投入，持续积极努力探索，发现这个时代最大的机会

保持不断持续创新的心态，全力增加内部和外部的协作和数据连接，真正创造生态链认可的价值和体验

应用新技术，风险管控，拥抱交互革命与生产力革命

