

DBAplus

沙龙标题

浅析企业数据运营之元数据管理

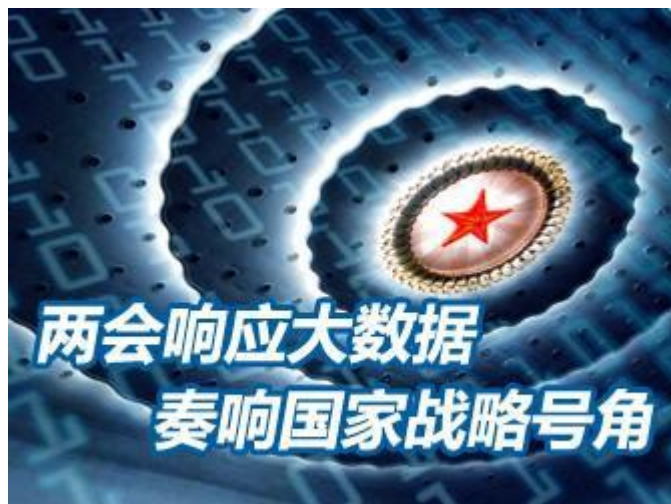
胡日平

01 企业数据治理现状与问题

02 数据运营之元数据的重要性

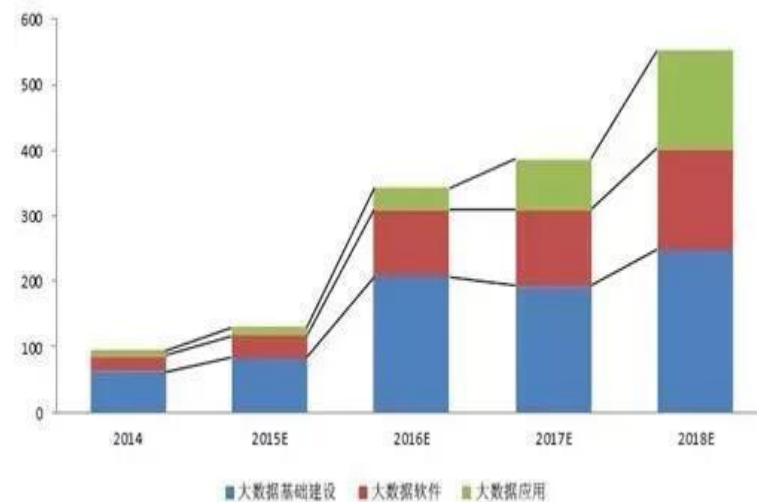
03 元数据管理范围及方法

04 元数据给企业带来的收益

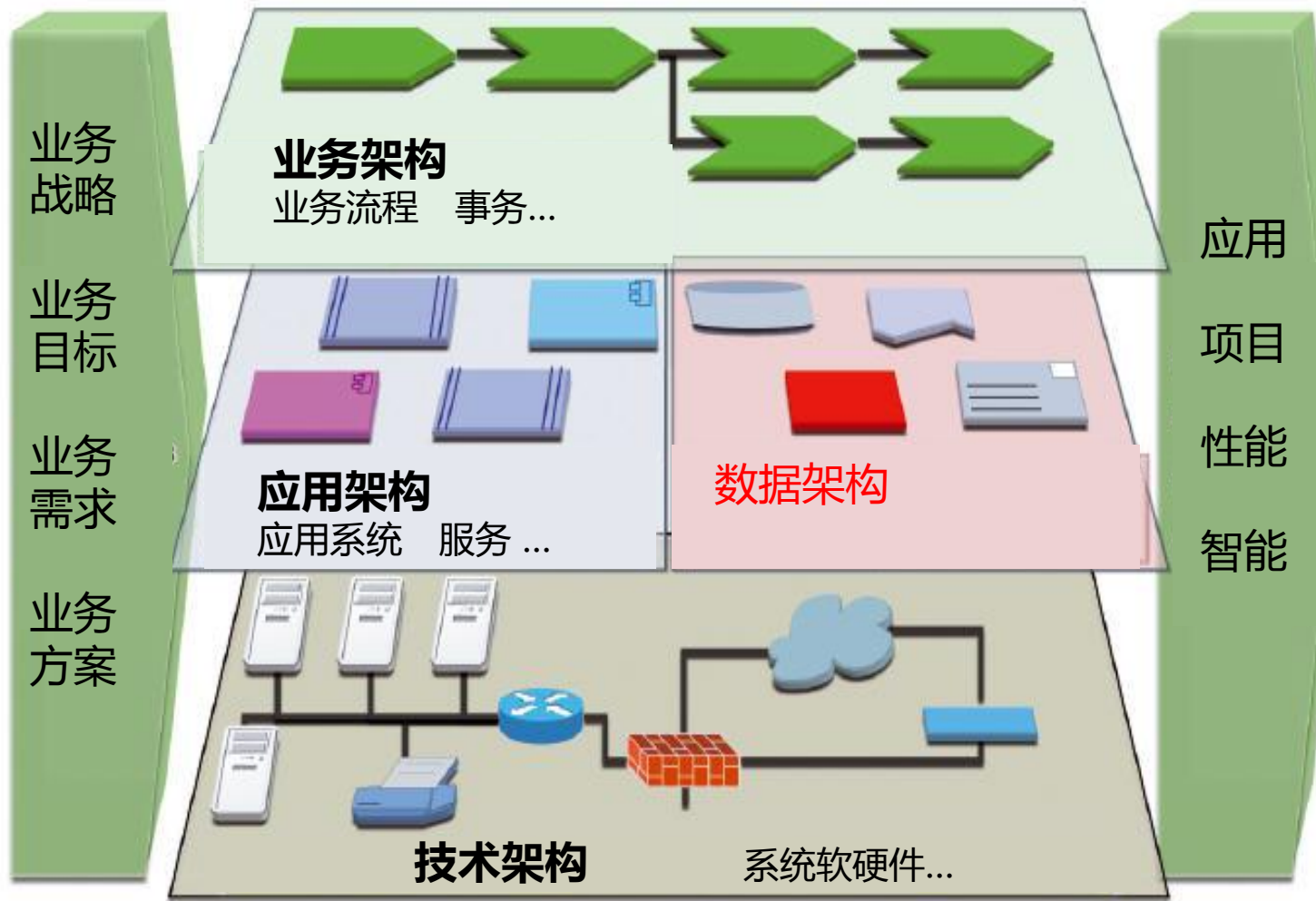


- 数据已经成为新的资源、宝贵的资源。中国要通过发展新经济为主要方向，从依赖自然资源到依赖人力资源，实现可持续发展。
- 政府拥有80%的数据资源，应该大力向外开放。推动政府信息共享，数据可以比作钻石矿，把互不相连的信息孤岛连接起来。
- 共享经济不仅在做加法，而且是乘法，充分利用闲置资源和过剩产能，提供劳动效率，让人人都受益。

- ◆ 2016.5，美国发布了《联邦大数据研究与开发战略计划》，围绕大数据研发的七大关键领域，包括促进人类对科学、医学和安全所有分支的认识，确保美国在研发领域继续发挥领导作用，通过研发来提高美国和世界解决紧迫社会和环境问题的能力。



TOGAF企业架构框架定义了数据架构位于业务架构与基础技术架构之间，通过数据架构的治理实现的数据和应用的更有效管理。



◆ 数据架构基线

- 数据架构梳理
- 数据架构优化
- 数据实体关系图
- 概念数据模型(图)
- 逻辑数据模型(图)
- 物理数据模型(图)

◆ 数据架构关系

- 业务流程与数据流向
- 业务功能/数据关系矩阵
- 应用/数据关系矩阵
- 数据架构展现与变更管理平台

◆ 数据管理

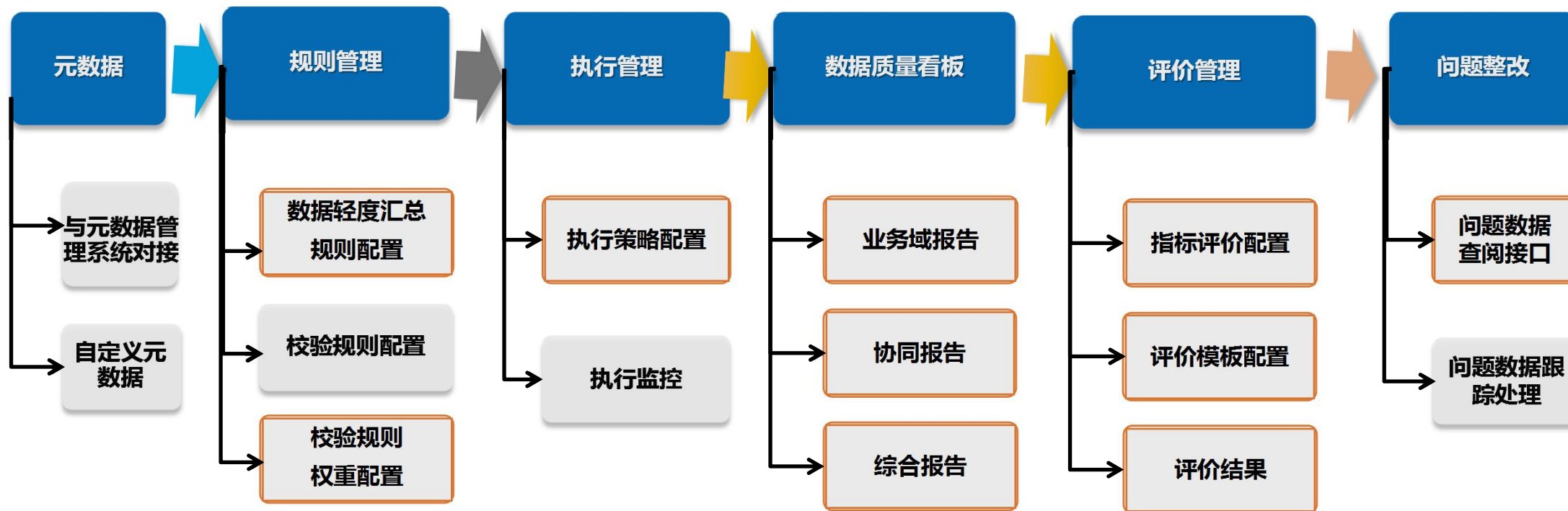
- 数据分布管理
- 数据质量管理
- 数据容量管理
- 数据迁移.....



原因一：企业的IT系统经历了数据量高速膨胀的时期，但大多数数据企业最初没有对数据架构重视，导致企业海量的、分散在不同角落，企业无法从统一的业务视角去概览整个企业内部的数据信息。暴露出来的只是一个个独立的系统，系统与系统之间的关系、标准数据从哪里获取都无从知晓

原因二：传统企业的数据资产问题归根到底是由于企业中对企业数据缺少统一而有力的组织、制度、流程的管控，引起的“**数据孤岛**”问题。

元数据管理是数据治理工作是重中之重，为什么企业内部的数据质量总是不高？其实只要有数据存在就有数据质量问题存在。但是也可以通过一个有效的管理流程来提高企业数据质量，其中最核心的核心就是元数据的管理；



01 企业数据治理现状与问题

02 数据运营之元数据的重要性

03 元数据管理范围及方法

04 元数据给企业带来的收益

元数据是企业中用来描述数据的数据。它可理解为比一般意义的数据库范畴更加广泛的数据，不再仅仅表示数据的类型、名称、值等信息，它可以进一步提供数据的上下文描述信息，比如数据的所属域、取值范围、数据间的关系、业务规则，甚至是数据的来源。

在数据分析中，元数据可以帮助DW管理员和DW开发人员非常方便地找到他们所关心的数据

- ✓ 有用的数据在哪里
- ✓ 提供一份数据结构定义和元数据的详细示意图
- ✓ 数据来龙去脉、关系
- ✓ 数据质量审计



**元数据提供
企业数据DNA**

- ✓ 减少数据冗余性,增加数据共享,
- ✓ 使应用开发过程更有效
- ✓ 参照性、引用性、血缘分析、影响分析、变化分析...
- ✓ 利用数据解决企业问题的能力

关键业务数据在不同系统的业务定义？
它们的数据结构是怎样？

已存在系统，有哪些关键业务数据？
它们在哪里？

报表A指标值怎么算出来的？原始数据从哪得到的？

一旦对象定义发生变化，其他开发团队会不会有影响？

数据实体所有者是谁？谁作过修改或变更？

同一实体在不同系统中的命名规则及其他属性一致？

同一实体与本系统或其他系统对象是否存在依赖关系？
具体的依赖关系描述是怎么样的？



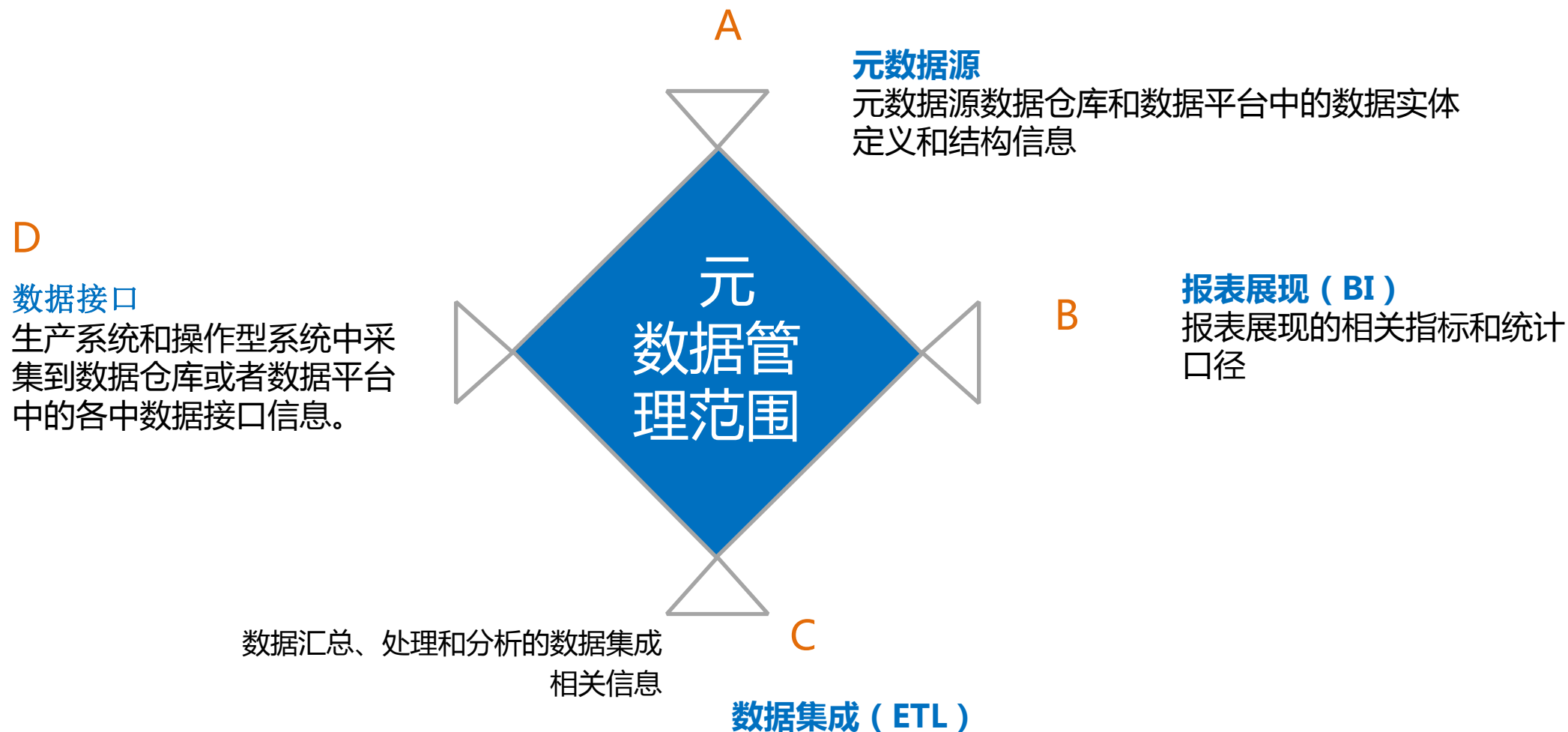
01 企业数据治理现状与问题

02 数据运营之元数据的重要性

03 元数据管理范围及方法

04 元数据给企业带来的收益

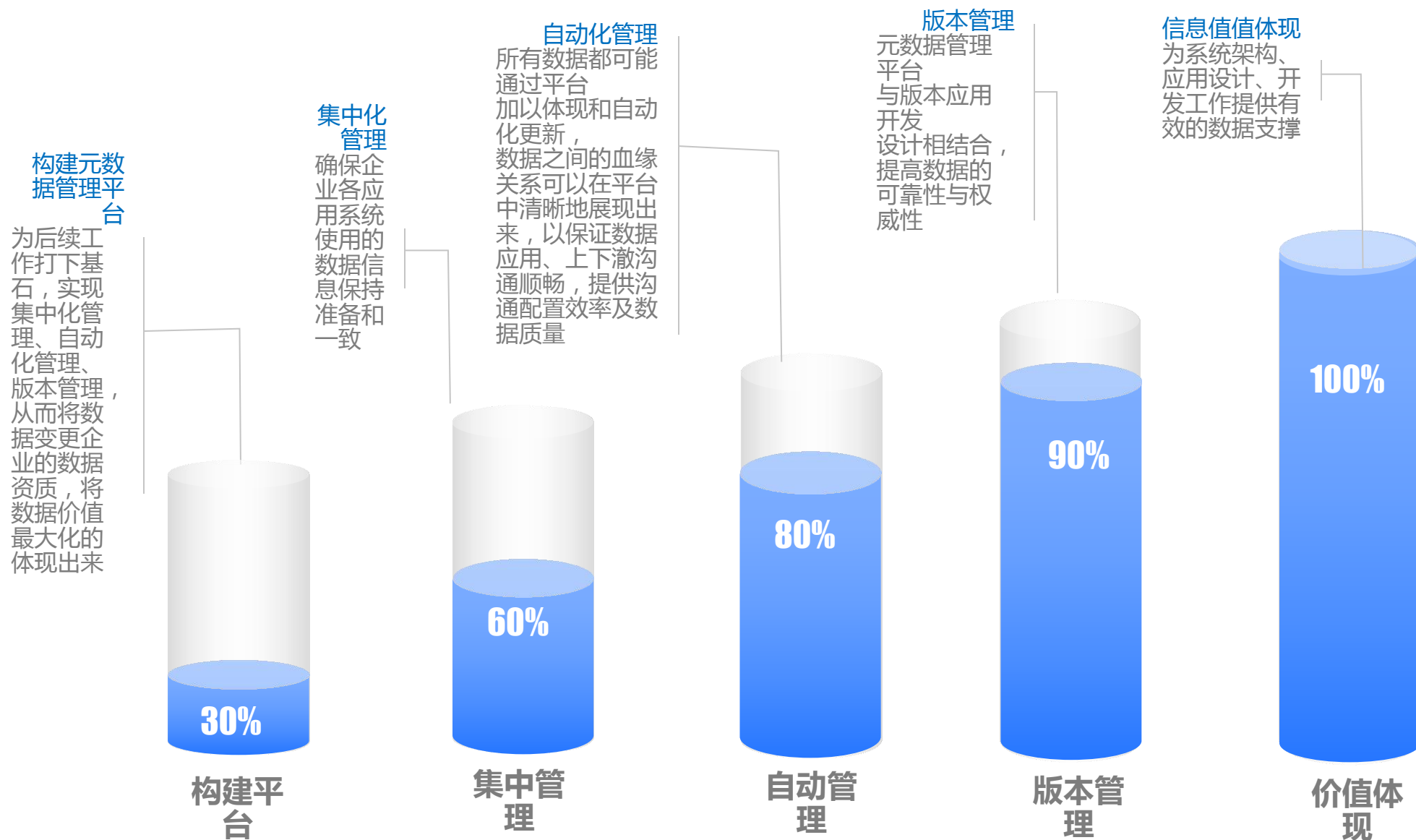
一般而言，就数据仓库或者大数据平台中的元数据可以按不同的维度分为技术元数据、业务元数据、操作元数据等，元数据管理的范围大致如下四个方面



DBAplus 实施元数据管理阶段性目标

DBAplus
数据连接未来

IT大咖说
知识分享平台



01 企业数据治理现状与问题

02 数据运营之元数据的重要性

03 元数据管理范围及方法

04 元数据给企业带来的收益

提供一份企业级的详细和应用。数据定义和元素的详细地图, 有利于企业的数据应用人员进行数据分析

帮助用户理解数据来龙去脉、关系及相关属性, 统一所有报表的统计口径, 减少数据冲突

数据影响分析, 帮助用户识别数据实体会影响的相关处理过程和报表, 有利于增强企业运营效率、规避错误及风险



制定企业元数据标准, 减少数据冗余性, 增加数据共享, 使应用开发过程更有效, 开发费用更低、避免异构结构定义、提高不同工具之间定义的重复利用率

数据血缘分析, 为数据分析提供直观的数据处理先后逻辑, 协助提升数据质量

通过元数据变更, 自始至终确保元数据的准确性和权威性

总结：

元数据管理能够增强数据理解，可以架起企业内业务与IT部门之间的一座桥梁，无论是企业的业务部门还是IT部门，很少能完整的拿出一套企业各项数据的业务含义、口径、技术标准、分布情况等说明，使用元数据管理可以自动化的获取整个企业的数据业务含义，帮助理解数据，增加分析的敏捷性，使用元数据产品能够方便内部管理、审计或外部监管的需求追溯业务指标、报表的数据来源和加工过程，追述数据的来源；同时还可以针对企业内部、外部的数据需求，快速建立业务与技术之间的衔接，为企业管理提供重要的保障。

DBAplus

www.dbaplus.cn