

考试资料职业发展
技术读书笔记分享

B站/闲鱼：大西洋活跃的锅巴
公众号：不太甜

DAMA-DMBOK

数据管理知识体系指南CDGA/CDGP认证

第一章 数据管理（完整课程视频请扫描二维码）



前言/PREFACE



国际数据管理协会（DAMA国际）是一个全球性的专业组织，由数据管理和相关的专业人士组成，非营利性机构。协会自1980年成立以来，一直致力于数据管理和数字化的研究、实践及相关知识体系的建设，先后发行了《DAMA 数据管理字典》和《DAMA数据管理知识体系》等。该知识体系目前已被广泛使用，并已成为业界的标杆和权威。

2020年10月起，国际数据管理协会中国分会（DAMA中国）对DAMA国际数据管理专业人员认证（CDMP）的考试语言、考试形式、考试内容、证书类型等进行考试认证，认证考试分为两个级别，完成第一级别CDGA认证后方可进行CDGP认证。

数据治理工程师—CDGA(Certified Data Governance Associate)

数据治理专家—CDGP(Certified Data Governance Professional)

前言/PREFACE



本课程为《DAMA-DMBOK》数据管理知识体系指南配套的讲义课程，旨在全面讲解了数据管理知识体系的12个数据管理原则和11个知识领域的专业基础理论。是培养数据团队体系化思考能力，提升专业数据素养的必修基础内容。

内容适用于从事数据管理的专业人员：数据治理岗，数据质量岗，数据架构岗、元数据管理岗、主数据管理岗、数据分析岗、数据建模岗、数据安全岗，文件档案管理岗，各类咨询顾问、各类IT相关岗位、知识工作者、教育工作者、法律工作者、各级关注数据的管理人员。

第一章 数据管理/Contents



01

引言

02

基本概念

03

数据管理框架

04

附录



01

第一章 数据管理 引言

数据管理的定义和目标



数据管理的定义：

数据管理是为了交付、控制、保护并提升数据和信息资产的价值，在其整个生命周期中制定计划、制度、规程和实践活动，并执行和监督的过程。

数据管理专业人员的定义：

是指从事数据管理各方面的工作（从数据全生命周期的技术管理工作，到确保数据的合理利用及发挥作用），并通过其工作发挥数据管理价值。

业务驱动因素：

数据管理的主要驱动力：使组织能够从其数据资产中获取价值。



管理 目标

- 1) 理解并支撑企业机器利益相关方的信息需求得到满足。
- 2) 获取、存储、保护数据和确保数据资产的完整性
- 3) 确保数据和信息的质量
- 4) 确保利益相关方的数据隐私和保密性
- 5) 防止数据和信息未经授权或被不当访问、操作及使用
- 6) 确保数据能有效地服务于企业增值的目标



02

基本概念

数据、数据和信息、管理原则和挑战

基本概念

什么是数据？

数据既是对其所代表的对象的解释，也是必须解释的对象。
数据代表事实，是这个世界中与某个事实结合在一起的一种真实表达。

什么是语境？

语境可被视为数据的表示系统，该系统包括一个公共词汇表和一系列组件之间的关系，如果知道这样一个系统的约定，就可解释其中的数据。这些数据通常记录在一种特殊类型的数据——元数据中。

数据是一种资产

资产是一种经济资源，能被拥有或控制、持有或产生价值。资产可以转化为货币。数据是一种组织资产。

1、关于数据的几个异议

- 1) 基于数据是简单存在的假设。但数据并不是简单存在，而是要被创造出来的。
- 2) 将数据到智慧描述为一个自下而上的逐级序列，未认识到创建数据需要知识。
- 3) 金字塔模型意味着数据和信息是分开的，但数据和信息是相互交织并相互依赖的，数据是信息的一种形式，反之也是。

2、要为不同目的准备数据和信息，使得数据管理形成一个核心原则：

数据和信息都需要被管理；

如果再将两者的使用和客户的需求结合在一起进行管理，则两者应具有更高的质量。

(1) 数据是有独特属性的资产

数据是一种资产，但相比其他类型资产，其在管理方式的某些方面有很大差异，对比金融和实物资产，其中最明显的一个特点是数据资产在使用过程中不会产生消耗

(2) 数据的价值可以用经济术语来表示

将数据视为资产意味着它有价值，虽然有技术手段可以测量数据的数量和质量，但还未形成这样做的标准来衡量其价值。

(3) 管理数据意味着对数据的质量管理

首要目标是确保数据符合应用的要求，为了管理质量，组织必须确保他们了解利益相关者对质量的要求，并根据这些要求度量数据

(4) 管理数据需要元数据

用于管理和如何使用的数据都称为元数据，元数据源于与数据创建、处理和使用相关的一系列流程，包括架构、建模、管理、治理、质量管理、系统开发、IT和业务运营以及分析

(5) 数据管理需要规划

数据在多个地方被创建，但因为使用需要在很多存储位置间移动，需要一些协调工作来保持最终结果的一致，需要从架构和流程的角度进行规划

(6) 数据管理须驱动信息技术决策

数据和数据管理与信息技术和信息技术管理紧密结合，管理数据需要一种方法，确保技术服务于而不是驱动组织的战略数据需求。

（7）数据管理是跨职能的工作

需要一系列的技能和专业知识、因此单个团队无法管理组织的所有数据。数据管理需要技术能力、非技术能力以及写作能力

（8）数据管理需要企业级视角

虽然数据管理存在很多本地应用程序，但它必须能够有效地应用于整个企业，这就是为什么数据管理和数据治理是交织在一起的原因之一。

（9）数据管理需要多角度思考

数据是流动的，数据管理必须不断发展演进，以跟上数据创建的方式、应用的方式和消费者的变化。

（10）数据管理需要全生命周期的管理，不同类型数据有不同的生命周期特征

不同类型数据有不同的生命周期特征，因此它们有不同的管理需求、数据管理实践需要基于这些差异，保持足够的灵活性来满足不同类型数据的生命需求。

（11）数据管理需要纳入与数据相关的风险

数据除了是一种资产外，还代表着组织的风险。数据可能丢失、被盗或误用。组织必须考虑其使用数据的伦理影响，数据相关风险必须作为数据生命周期的一部分进行管理。

（12）有效的数据管理需要领导层承担责任

数据管理涉及一组复杂的过程，需要协调、协作和承诺。为了达到目标，不仅需要管理技巧，还需要来自领导层的愿景和使命。

1、数据与其他资产的区别：无形的，不会损耗，动态的，可以同时被多人使用。引出以下问题：

- 定义数据所有权
- 列出组织拥有的数据量
- 防止数据滥用
- 管理与数据冗余相关的风险
- 定义和实施数据质量标准

2、数据价值

是一件事物的成本和从中获得利益的差额；评估数据的价值需要在组织内部持续付出的一般性成本和各类收益：

- 1) 获取和存储数据的成本
- 2) 如果丢失, 更换数据需要的成本
- 3) 数据丢失对组织的影响
- 4) 风险缓解成本和与数据相关的潜在风险成本
- 5) 改进数据的成本
- 6) 高质量数据的优势
- 7) 竞争对手为数据付出的费用
- 8) 数据潜在的销售价格
- 9) 创新性应用数据的预期收入

主要挑战是：数据的价值是上下文相关的。在一个组织中，某些类型的数据可能会随着时间的推移而具有一致的价值。

3、数据质量

确保高质量数据是数据管理的核心。

低质量数据的成本主要来源于：

- 1) 报废和返工
- 2) 解决方法和隐藏的纠正过程
- 3) 组织效率低下或生产力低下
- 4) 组织冲突
- 5) 工作满意度低
- 6) 客户不满意
- 7) 机会成本，包括无法创新
- 8) 合规成本或罚款
- 9) 声誉成本

高质量数据的作用包括：

- 1) 改善客户体验
- 2) 提高生产力
- 3) 降低风险
- 4) 快速响应商机
- 5) 增加收入
- 6) 洞察客户、产品、流程和商机，获得竞争优势

4、数据优化计划

决策需要系统思考，因为涉及：

- 1) 数据也许被视为独立于业务流程存在
- 2) 业务流程与支持它们的技术之间的关系
- 3) 系统的设计和架构及其所生成和存储的数据
- 4) 使用数据的方式可能被用于推动组织战略

5、元数据和数据管理

元数据描述了一个组织拥有什么数据，它代表什么、如何被分类、它来自哪里、在组织之内如何移动、如何在使用中演进、谁可以使用它以及是否为高质量数据。

元数据是全面改进数据管理的起点。

6、数据管理是跨职能的工作

数据管理需要系统规划的设计技能、管理硬件和构建软件的高技术技能、利用数据分析理解问题和揭示数据的技能、通过定义和模型达成共识的语言技能、发现客户服务商机和实现目标的战略思维。

挑战在于让具备这一系列技能的人认识到各部分是如何结合在一起的。

7、建立企业的视角

管理数据需要理解一个组织中的机会和数据范围。

重要的原因之一是帮助组织跨垂直领域作出决策。

8、数据管理需要多角度思考

自有和外购数据，以及外国法律法规，以及误用风险

9、数据生命周期

数据生命周期基于产品生命周期，不应该与系统开发生命周期混淆。它包括创建或获取、移动、转换和存储数据并使其得以维护和共享的过程，使用数据的过程，以及处理数据的过程。在整个生命周期中，可以清理、转换、合并、增强或聚合数据。随着数据的使用或增强，通常会生成新的数据，因此其生命周期具有内部迭代，而这些迭代没有显示在图表上。

数据管理对生命周期的关注有几个重要影响：

- 1) 创建和使用是数据生命周期中的关键点
- 2) 数据质量管理必须贯穿整个数据生命周期
- 3) 元数据质量管理必须贯穿整个数据生命周期
- 4) 数据管理还包括确保数据安全，并降低与数据相关的风险
- 5) 数据管理工作应聚焦于关键数据，将数据ROT降至最低

10、不同种类的数据

按数据类型分类，如交易数据、参考数据、主数据、元数据，或者类别数据、源头数据、事件数据、详细交易数据；

也可以按数据内容、数据所需格式或保护级别、存储或访问的方式和位置进行分类

11、数据和风险

数据的风险在于，它可能被误解和误用；

最高质量的数据带给组织最大的价值——可获得、相互关联、完整、准确、一致、及时、适用、有意义和易于理解。

12、数据管理和技术

管理技术和管理数据不同。组织需要了解技术对数据的影响，以防止技术诱惑推动他们对数据的决策。相反，与业务战略一致的数据应该推动有关技术的决策。

13、高效的数据管理需要领导力和承诺

《领导者的数据宣言》：组织有机增长的最佳机会在于数据；

倡导首席数据官（CDO）的作用源于认识到管理数据会带来独特的挑战，成功的数据管理必须由业务驱动，而不是由IT驱动。

战略是一组选择和决策，它们共同构成了实现高水平目标的高水平行动过程。

战略计划是为实现高水平目标而采取的高水平行动。

数据战略应该包括使用信息以获得竞争优势和支持企业目标的业务计划。

数据战略需要一个支持性的数据管理战略——

一个维护和改进数据质量、数据完整性、访问和安全性的规划，同时降低已知和隐含的风险。

该战略还必须解决与数据管理相关的已知挑战。

数据管理战略的组成应包括：

- 1) 令人信服的数据管理愿景
- 2) 数据管理的商业案例总结
- 3) 指导原则、价值观和管理观点
- 4) 数据管理的使命和长期目标
- 5) 数据管理成功的建议措施
- 6) 符合SMART原则（具体、可衡量、可操作性、现实、有时间限制）的短期数据管理计划目标
- 7) 对数据管理角色和组织的描述，以及对其职责和决策权的总结
- 8) 数据管理程序组件和初始化任务
- 9) 具体明确范围的优先工作计划
- 10) 一份包含项目和行动任务的实施路线图草案

1) 数据管理章程

包括总体愿景、业务案例、目标、指导原则、成功衡量标准、关键成功因素、可识别的风险、运营模式等

2) 数据管理范围声明

包括规划目的和目标，以及负责实现这些目标的角色、组织和领导

3) 数据管理实施路线图

确定特定计划、项目、任务分配和交付里程碑。

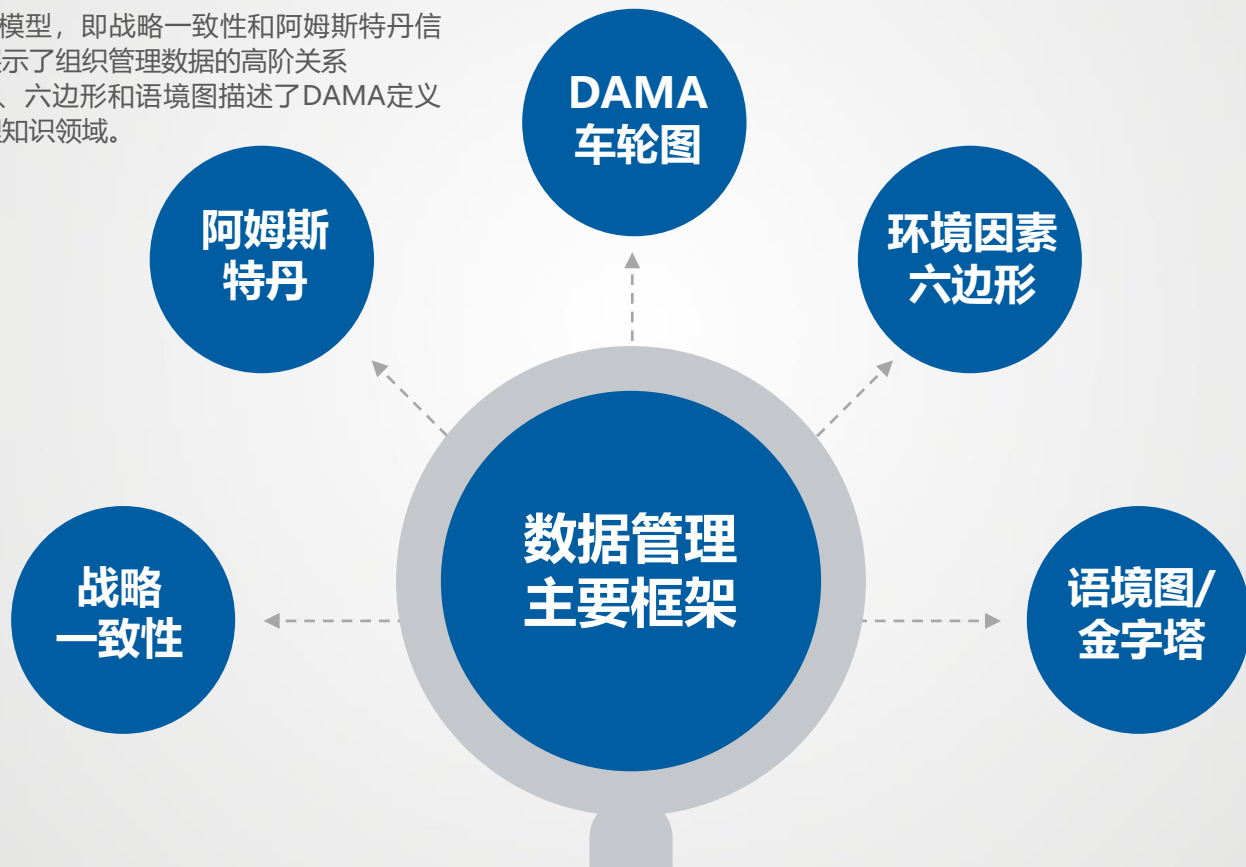


03

数据管理框架

战略一致性模型、阿姆斯特丹模型、DMBOK框架

- 1) 前两个模型，即战略一致性和阿姆斯特丹信息模型，展示了组织管理数据的高阶关系
- 2) 车轮图、六边形和语境图描述了DAMA定义的数据管理知识领域。

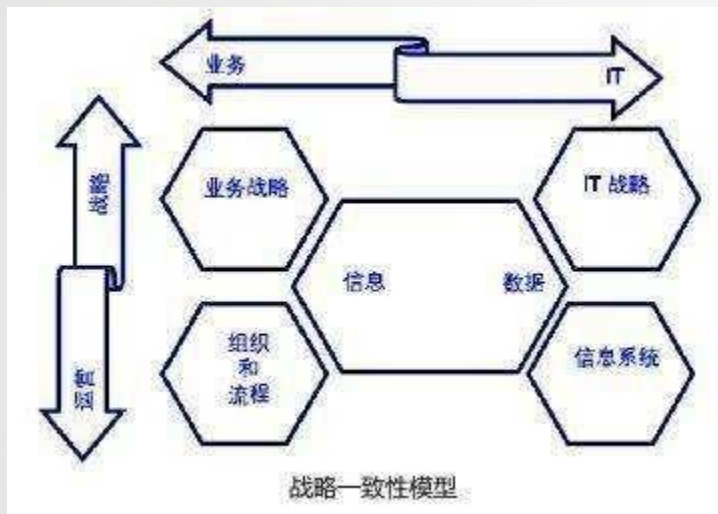


战略一致性模型 (考点。包括哪几部分)

考试资料职业发展
技术读书笔记分享

B站/闲鱼: 大西洋活跃的锅巴
公众号: 不太甜

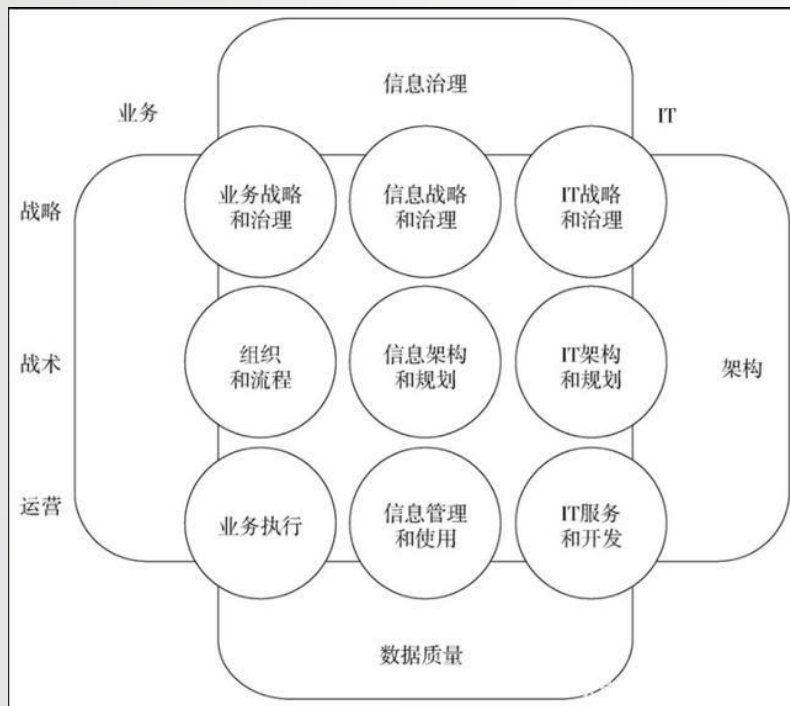
22



战略一致性模型 (Strategic Alignment Model, SAM) 抽象了各种数据管理方法的基本驱动因素，模型的中心是数据和信息之间的关系。

信息通常与业务战略和数据的操作相关，数据与信息技术和流程相关。

围绕这一概念是战略选择的4个基本领域：**业务战略、IT战略、组织和流程以及信息系统。**



与战略一致性模型一样，从战略角度看待业务和IT的一致性。共有9个单元，它抽象出一个关注结构（包括规划和架构）和策略的中间层。此外，还要认识到信息通信的必要性。

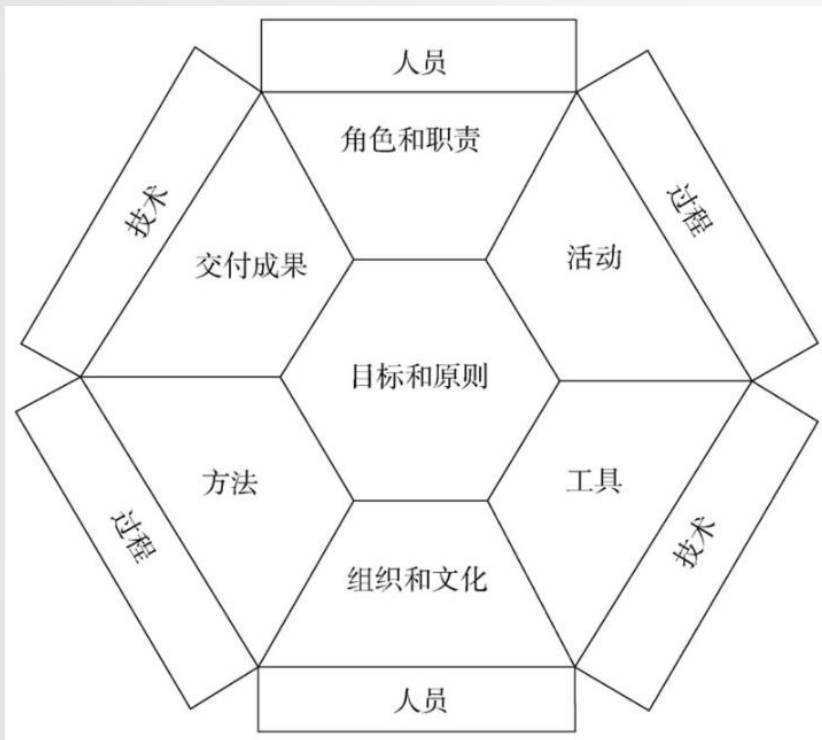
SAM（战略一致性模型）和AIM（阿姆斯特丹信息模型）框架动横轴和纵轴两个维度详细描述组件之间的关系



1) DAMA车轮图

车轮图定义了数据管理知识领域，将数据治理放在中心，其他知识领域围绕车轮平衡。

数据管理和数据治理的关系？数据治理是管理的一部分



2) 环境因素六边形图

显示了人、过程和技术之间的关系，是理解DMBOK语境关系图的关键。它将目标和原则放在中心，因为这些目标和原则为人们如何执行活动及有效地使用工具成功进行数据管理提供了指导。

考题，哪几部分：角色和职责、活动、工具、组织文化、方法、交付成果

3) 知识领域语境关系图

描述了知识领域的细节，包括与人员、流程和技术相关的细节。他们基于产品管理的SIPOC图的概念。每个语境关系图都以知识领域的定义和目标开始。目标驱动的活动分为四个阶段：计划（P）、控制（C）、开发（D）、运营（O）

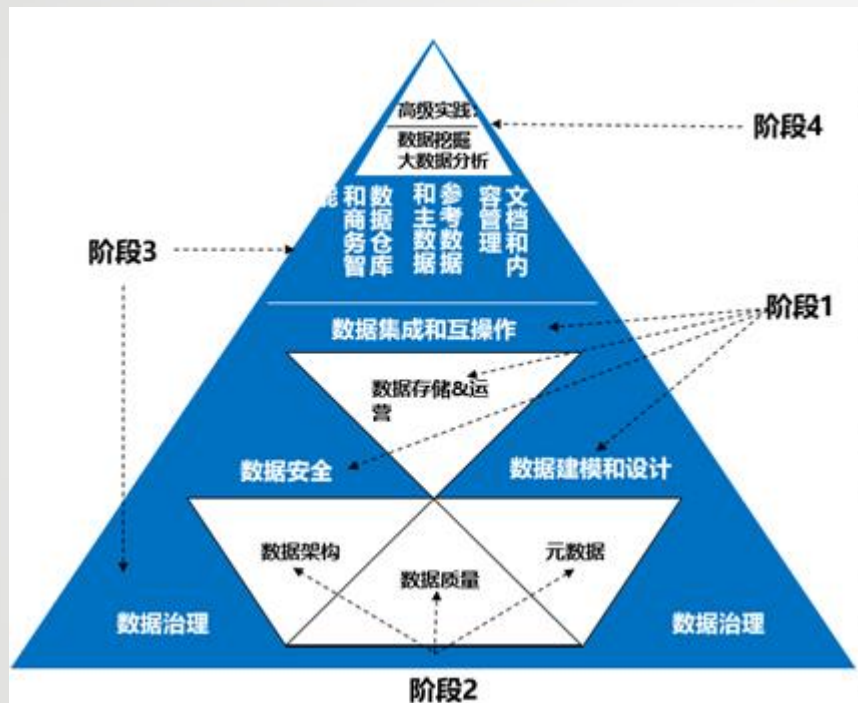
组成部分包括12个部分，都要记住：

- 1) 定义
- 2) 目标
- 3) 活动：①计划活动P ②控制活动C ③开发活动D ④运营活动O
- 4) 输入
- 5) 交付成果
- 6) 角色和职责
- 7) 供给者 负责提供或允许访问活动输入的人员
- 8) 消费者 直接受益于数据管理互动产生主要交付成果的消费方
- 9) 参与者 执行、管理或批准知识领域活动的人员。
- 10) 工具 它是实现知识领域目标的应用程序和其他技术
- 11) 方法 它是用于在知识领域内执行活动和产生可交付成果的方法和程序。它还包括共同约定、最佳实践建议、标准和协议以及新出现的一些合适的替代方法。
- 12) 度量指标 它是衡量或评估绩效、进度、质量、效率或其他影响的标准。这些指标用于定义每个知识领域内完成工作的可量化事实。



3) 知识领域语境关系图





第一阶段

组织购买包含数据库功能的

应用程序

第二阶段

一旦组织开始使用应用程序，他们将面临数据质量方面的挑战，但获得更高质量的数据还取决于可靠的元数据和一致的数据结构，它们说明了来自不同系统的数据是如何协同工作的。

第三阶段

管理数据质量、元数据和架构需要严格地实践数据治理，为数据管理活动提供体系性支持。

第四阶段

该组织充分利用了良好管理数据的好处，并提高了其分析能力

该框架从数据管理的指导目标开始：使组织能够像从其他资产中获取价值那样，从其数据资产中获取价值。

与数据生命周期相关的数据管理功能在图的中心进行了描述，包括：为可靠、高质量的数据进行规划和设计；建立过程和功能来使用和维护数据；在各种类型的分析活动以及这些过程中使用数据，以提高价值。

DAMA数据管理框架也被描述为另一种形式的DAMA车轮图，数据治理范围内的应用活动围绕着数据管理生命周期内的各项核心活动进行。

核心活动位于框架中心，包括元数据管理、数据质量管理 and 数据结构定义

生命周期管理活动可以从多个方面定义，如计划的角度（风险管理、建模、数据涉及、参考数据管理），实现的角度（数据仓库、主数据管理、数据存储和操作、数据集成和互操作、数据开发技术）

生命周期管理活动源于数据的使用：主数据使用、文件和内容管理、商务智能、数据科学、预测分析、数据可视化。

督

(1) 数据治理

通过建立一个能够满足企业需求的数据决策体系，为数据管理提供指导和监

(2) 数据架构

定义了组织战略协调的管理数据资产蓝图，以建立战略性数据需求及满足需求的总体设计

(3) 数据建模和设计

以数据模型的精确形式，进行发现、分析、展示和沟通数据需求

(4) 数据存储和操作

以数据价值最大化为目标，包括存储数据的设计、实现和支持活动以及在整个数据生命周期中，从计划到销毁的各种操作活动。

(5) 数据安全

确保数据隐私和机密性得到维护，数据不被破坏，数据被适当访问

(6) 数据集成和互操作

包括与数据存储、应用程序和组织之间的数据移动和整合相关的过程

(7) 文件和内容管理

用于管理非结构化媒体数据和信息的生命周期过程，包括计划、实施和控制活动，尤其是指支持法律法规遵从性要求所需的文档

(8) 参考数据和主数据

包括核心共享数据的持续协调和维护，使关键业务实体的真实信息以准确、及时和相关联的方式在各系统间得到一致使用。

(9) 数据仓库和商务智能

包括计划、实施和控制流程来管理决策支持数据，并使知识工作者通过分析报告从数据中获得价值。

(10) 元数据

包含规划、实施和控制活动，以便能够访问高质量的集成元数据，包括定义、模型、数据流和其他至关重要的信息

(11) 数据质量

包括规划和实施质量管理技术，以测量评估和提高数据在组织内的适用性。

数据处理伦理

描述数据伦理规范在促进信息透明、社会责任决策中的核心作用。数据采集、分析和使用过程中的伦理意识对所有数据管理专业人员有指导作用。

大数据和数据科学

描述了针对大型的、多样化数据集收集和分析能力的提高而出现的技术和业务流程。

数据管理成熟度评估

概述了评估和改进组织数据管理能力的方法

数据管理组织和角色期望

为组建数据管理团队、实现成功的数据管理活动提供了实践指导和参考。

数据管理和组织变革管理

描述了如何计划和成功地推动企业文化变革。



04

附录

- (1) 数据治理：通过建立一个能够满足企业需求的数据决策体系，为数据管理提供指导和监督。
- (2) 数据架构：定义了与组织战略协调的管理数据资产蓝图，以建立战略性数据需求及满足需求的总体设计。
- (3) 数据建模设计：以数据模型的精确形式，进行发现、分析、展示和沟通数据需求。
- (4) 数据存储和操作：以数据价值最大化为目标，包括存储数据的设计、实现和支持活动以及在整个数据生命周期中，从计划到销毁的各种操作活动。
- (5) 数据安全：确保数据隐私和机密性得到维护，数据不被破坏，数据被适当访问。
- (6) 数据集成和互操作：包括与数据存储、应用程序和组织之间的数据移动和整合相关的过程。
- (7) 文件和内容管理：用于非结构化媒体数据和信息的生命周期过程，包括计划、实施和控制活动，尤其是指支持法律法规遵从性要求所需的文档。
- (8) 参考数据和主数据：包括核心共享数据的持续协调和维护，使关键业务实体的真实信息以准确、及时和相关联的方式在各系统间得到一致使用。
- (9) 数据仓库和商务智能：包括计划、实施和控制流程来管理决策支持数据，并使知识工作者通过分析报告从数据中获得价值。
- (10) 元数据：包含规划、实施和控制活动，以便能够访问高质量的集成元数据，包括定义、模型、数据流和其他至关重要的信息（对理解数据及其创建、维护和访问系统有帮助）
- (11) 数据质量：包括规划和实施质量管理技术，以测量、评估和提高数据在组织内的适用性。

- （1）数据处理伦理：描述了关于数据及其应用过程中，数据伦理规范在促进信息透明、社会责任决策中的核心作用。数据采集、分析和使用过程中的伦理意识对所有数据管理专业人员有指导作用。
- （2）大数据和数据科学：描述了针对大型的、多样化数据集收集和分析能力的提高而出现的技术和业务流程。
- （3）数据管理成熟度评估：概述了评估和改进组织数据管理能力的方法。
- （4）数据管理组织和角色期望：为组建数据管理团队、实现成功的数据管理活动提供了实践指导和参考。
- （5）数据管理和组织变革管理：描述了如何计划和成功地推动企业文化变革。文化的变革是将数据管理实践有效地嵌入组织中的必然结果。

本章完结 感谢观看

完整课程视频请扫描二维码咨询

