

CAPÍTULO 1

Gestão de Dados

1. Introdução

Muitas organizações reconhecem que seus dados são um ativo empresarial vital. Dados e informações podem dar a elas insights sobre seus clientes, produtos e serviços. Podem ajudá-las a inovar e atingir objetivos estratégicos. Apesar desse reconhecimento, poucas organizações gerenciam ativamente os dados como um ativo do qual podem derivar valor contínuo (Evans e Price, 2012). Derivar valor dos dados não acontece no vácuo ou por acidente. Requer intenção, planejamento, coordenação e comprometimento. Requer gerenciamento e liderança.

Gerenciamento de dados é o desenvolvimento, execução e supervisão de planos, políticas, programas e práticas que fornecem, controlam, protegem e aumentam o valor dos dados e ativos de informação ao longo de seus ciclos de vida.

Um *profissional de gerenciamento de dados* é qualquer pessoa que trabalha em qualquer faceta do gerenciamento de dados (do gerenciamento técnico de dados ao longo de seu ciclo de vida até garantir que os dados sejam adequadamente utilizados e alavancados) para atender às metas organizacionais estratégicas. Os profissionais de gerenciamento de dados preenchem inúmeras funções, desde as altamente técnicas (por exemplo, administradores de banco de dados, administradores de rede, programadores) até as estratégicas de negócios (por exemplo, administradores de dados, estrategistas de dados, diretores de dados).

As atividades de gerenciamento de dados são abrangentes. Elas incluem tudo, desde a capacidade de tomar decisões consistentes sobre como obter valor estratégico dos dados até a implantação técnica e o desempenho dos bancos de dados. Portanto, o gerenciamento de dados requer habilidades técnicas e não técnicas (ou seja, 'negócios'). A responsabilidade pelo gerenciamento de dados deve ser compartilhada entre as funções de negócios e tecnologia da informação, e as pessoas em ambas as áreas devem ser capazes de colaborar

para garantir que uma organização tenha dados de alta qualidade que atendam às suas necessidades estratégicas.

Dados e informações não são apenas ativos no sentido de que as organizações investem neles para derivar valor futuro. Dados e informações também são vitais para as operações do dia a dia da maioria das organizações. Eles foram chamados de "moeda", "sangue vital" e até mesmo "novo petróleo" da economia da informação.

¹ Independentemente

de uma organização obter ou não valor de suas análises, ela não consegue nem fazer negócios sem dados.

Para dar suporte aos profissionais de gerenciamento de dados que realizam o trabalho, a DAMA International (The Data Management Association) produziu este livro, a segunda edição do *The DAMA Guide to the Data Management Body of Knowledge (DMBOK2)*. Esta edição se baseia na primeira, publicada em 2009, que forneceu conhecimento fundamental sobre o qual construir conforme a profissão avançava e amadurecia.

Este capítulo descreve um conjunto de princípios para gerenciamento de dados. Ele discute os desafios relacionados a seguir esses princípios e sugere abordagens para enfrentar esses desafios. O capítulo também descreve o DAMA Data Management Framework, que fornece o contexto para o trabalho realizado por profissionais de gerenciamento de dados em várias Data Management Knowledge Areas.

1.1 Drivers de negócios

Informação e conhecimento são a chave para a vantagem competitiva.

Organizações que têm dados confiáveis e de alta qualidade sobre seus clientes, produtos, serviços e operações podem tomar melhores decisões do que aquelas sem dados ou com dados não confiáveis. A falha em gerenciar dados é semelhante à falha em gerenciar capital. Isso resulta em desperdício e oportunidade perdida. O principal motivador para o gerenciamento de dados é permitir que as organizações obtenham valor de seus ativos de dados, assim como o gerenciamento eficaz de ativos financeiros e físicos permite que as organizações obtenham valor desses ativos.

1.2 Objetivos

Dentro de uma organização, os objetivos de gerenciamento de dados incluem:

- Compreender e dar suporte às necessidades de informação da empresa e de suas partes interessadas, incluindo clientes, funcionários e parceiros de negócios
- Capturar, armazenar, proteger e garantir a integridade dos ativos de dados
- Garantir a qualidade dos dados e informações
- Garantir a privacidade e a confidencialidade dos dados das partes interessadas
- Impedir o acesso, a manipulação ou o uso não autorizado ou inapropriado de dados e informações
- Garantir que os dados possam ser usados de forma eficaz para agregar valor à empresa

2. Conceitos Essenciais

2.1 Dados

Definições de *dados* de longa data enfatizam seu papel na representação de fatos sobre o mundo.² Em relação à tecnologia da informação, *dados* também são entendidos como informações que foram armazenadas em formato digital (embora dados não se limitem a informações que foram digitalizadas e princípios de gerenciamento de dados se apliquem a dados capturados em papel, bem como em bancos de dados). Ainda assim, como hoje podemos capturar tantas informações eletronicamente, chamamos muitas coisas de "dados" que não teriam sido chamadas de "dados" em tempos anteriores - coisas como nomes, endereços, datas de nascimento, o que alguém comeu no jantar no sábado, o livro mais recente que alguém comprou.

Esses fatos sobre pessoas individuais podem ser agregados, analisados e usados para gerar lucro, melhorar a saúde ou influenciar políticas públicas.

Além disso, a nossa capacidade tecnológica para medir uma vasta gama de eventos e atividades (desde as repercussões do Big Bang até à nossa

(próprios batimentos cardíacos) e coletar, armazenar e analisar versões eletrônicas de coisas que antes não eram consideradas dados (vídeos, fotos, gravações de som, documentos) está perto de superar nossa capacidade de sintetizar esses dados em informações utilizáveis.³ Para aproveitar a variedade de dados sem ser sobrecarregado por seu volume e velocidade, são necessárias práticas confiáveis e extensíveis de gerenciamento de dados.

A maioria das pessoas assume que, como os dados representam fatos, eles são uma forma de verdade sobre o mundo e que os fatos se encaixarão. Mas os "fatos" nem sempre são simples ou diretos. Dados são um meio de representação. Eles representam outras coisas além de si mesmos (Chisholm, 2010).

Dados são tanto uma interpretação dos objetos que representam quanto um objeto que deve ser interpretado (Sebastian-Coleman, 2013). Esta é outra maneira de dizer que precisamos de contexto para que os dados sejam significativos.

O contexto pode ser pensado como um sistema representacional de dados; tal sistema inclui um vocabulário comum e um conjunto de relacionamentos entre componentes. Se conhecermos as convenções de tal sistema, então podemos interpretar os dados dentro dele.⁴ Essas convenções são frequentemente documentadas em um tipo específico de dados chamado Metadados.

No entanto, como as pessoas frequentemente fazem escolhas diferentes sobre como representar conceitos, elas criam maneiras diferentes de representar os mesmos conceitos. A partir dessas escolhas, os dados assumem formas diferentes. Pense na gama de maneiras que temos para representar datas do calendário, um conceito sobre o qual há uma definição acordada. Agora considere conceitos mais complexos (como cliente ou produto), onde a granularidade e o nível de detalhes do que precisa ser representado nem sempre são autoevidentes, e o processo de representação se torna mais complexo, assim como o processo de gerenciamento dessas informações ao longo do tempo.

(Ver [Capítulo 10](#)).

Mesmo dentro de uma única organização, muitas vezes há várias maneiras de representar a mesma ideia. Daí a necessidade de Arquitetura de Dados, modelagem, governança e administração, e Metadados e gerenciamento de Qualidade de Dados, todos os quais ajudam as pessoas a entender e usar dados. Entre organizações, o problema da multiplicidade se multiplica.

Daí a necessidade de padrões de dados em nível industrial que possam trazer mais consistência aos dados.

As organizações sempre precisaram gerenciar seus dados, mas as mudanças na tecnologia expandiram o escopo dessa necessidade de gerenciamento, pois mudaram a compreensão das pessoas sobre o que são dados. Essas mudanças permitiram que as organizações usassem dados de novas maneiras para criar produtos, compartilhar informações, criar conhecimento e melhorar o sucesso organizacional. Mas o rápido crescimento da tecnologia e, com ela, a capacidade humana de produzir, capturar e minerar dados em busca de significado intensificou a necessidade de gerenciar dados de forma eficaz.

2.2 Dados e Informações

Muita tinta foi derramada sobre o relacionamento entre dados e informações. Dados foram chamados de “matéria-prima da informação” e informações foram chamadas de “dados em contexto”.⁵ Frequentemente, uma pirâmide em camadas é usada para descrever o relacionamento entre dados (na base), informação, conhecimento e sabedoria (no topo). Embora a pirâmide possa ser útil para descrever por que os dados precisam ser bem gerenciados, essa representação apresenta vários desafios para o gerenciamento de dados.

- Baseia-se na suposição de que os dados simplesmente existem. Mas os dados não existem simplesmente. Os dados precisam ser criados.
- Ao descrever uma sequência linear de dados até a sabedoria, ele falha em reconhecer que é preciso conhecimento para criar dados em primeiro lugar.
- Isso implica que dados e informações são coisas separadas, quando, na realidade, os dois conceitos estão interligados e são dependentes um do outro. Dados são uma forma de informação e informação é uma forma de dados.

Dentro de uma organização, pode ser útil traçar uma linha entre informações e dados para fins de comunicação clara sobre os requisitos e expectativas de diferentes usos por diferentes partes interessadas. (“Aqui está um relatório de vendas do último trimestre [informações]. Ele é baseado em dados do nosso data warehouse [dados].

No próximo trimestre, esses resultados [dados] serão usados para gerar nossas medidas de desempenho trimestrais [informações]). Reconhecer que dados e informações precisam ser preparados para diferentes propósitos reforça um princípio central do gerenciamento de dados: dados e informações precisam ser gerenciados. Ambos serão de maior qualidade se forem gerenciados em conjunto, tendo em mente os usos e os requisitos do cliente.

Em todo o DMBOK, os termos serão usados de forma intercambiável.

2.3 Dados como um Ativo Organizacional Um *ativo* é

um recurso econômico, que pode ser possuído ou controlado, e que detém ou produz valor. Ativos podem ser convertidos em dinheiro.

Os dados são amplamente reconhecidos como um ativo empresarial, embora a compreensão do que significa gerenciar dados como um ativo ainda esteja evoluindo. No início da década de 1990, algumas organizações acharam questionável se o valor do goodwill deveria receber um valor monetário. Agora, o "valor do goodwill" comumente aparece como um item na Demonstração de Lucros e Perdas (P&L). Da mesma forma, embora não seja universalmente adotada, a monetização de dados está se tornando cada vez mais comum. Não demorará muito para que vejamos isso como uma característica dos P&Ls. (Consulte [o Capítulo 3.](#))

As organizações de hoje dependem de seus ativos de dados para tomar decisões mais eficazes e operar com mais eficiência. As empresas usam dados para entender seus clientes, criar novos produtos e serviços e melhorar a eficiência operacional cortando custos e controlando riscos.

Agências governamentais, instituições educacionais e organizações sem fins lucrativos também precisam de dados de alta qualidade para orientar suas atividades operacionais, táticas e estratégicas. À medida que as organizações dependem cada vez mais de dados, o valor dos ativos de dados pode ser estabelecido mais claramente.

Muitas organizações se identificam como 'orientadas por dados'. Empresas que pretendem permanecer competitivas devem parar de tomar decisões com base em sentimentos viscerais ou instintos e, em vez disso, usar gatilhos de eventos e aplicar análises para obter insights acionáveis. Ser orientado por dados inclui o reconhecimento de que os dados devem ser gerenciados de forma eficiente e com disciplina profissional, por meio de uma parceria de liderança empresarial e conhecimento técnico.

Além disso, o ritmo dos negócios hoje significa que a mudança não é mais opcional; a disrupção digital é a norma. Para reagir a isso, as empresas devem cocriar soluções de informação com profissionais de dados técnicos trabalhando ao lado de contrapartes da linha de negócios. Eles devem planejar como obter e gerenciar dados que sabem que precisam para dar suporte à estratégia de negócios. Eles também devem se posicionar para aproveitar as oportunidades de alavancar dados de novas maneiras.

2.4 Princípios de gerenciamento de dados O

gerenciamento de dados compartilha características com outras formas de gerenciamento de ativos, como visto na Figura 1. Ele envolve saber quais dados uma organização tem e o que pode ser feito com eles, e então determinar a melhor forma de usar os ativos de dados para atingir as metas organizacionais.

Como outros processos de gestão, ele deve equilibrar as necessidades estratégicas e operacionais. Esse equilíbrio pode ser melhor atingido seguindo um conjunto de princípios que reconhecem características salientes da gestão de dados e orientam a prática de gestão de dados.

- **Dados são um ativo com propriedades únicas:** Dados são um ativo, mas diferem de outros ativos em maneiras importantes que influenciam como são gerenciados. A mais óbvia dessas propriedades é que os dados não são consumidos quando são usados, assim como os ativos financeiros e físicos.
- **O valor dos dados pode e deve ser expresso em termos econômicos:** Chamar dados de ativos implica que eles têm valor. Embora existam técnicas para medir o valor qualitativo e quantitativo dos dados, ainda não há padrões para isso. Organizações que querem tomar melhores decisões sobre seus dados devem desenvolver maneiras consistentes de quantificar esse valor. Elas também devem medir os custos de dados de baixa qualidade e os benefícios de dados de alta qualidade.
- **Gerir dados significa gerir a qualidade dos dados:** garantir que os dados são adequados à finalidade é um objetivo principal

gerenciamento de dados. Para gerenciar a qualidade, as organizações devem garantir que entendam os requisitos de qualidade dos stakeholders e medir os dados em relação a esses requisitos.

- **É preciso Metadados para gerenciar dados:** Gerenciar qualquer ativo requer ter dados sobre esse ativo (número de funcionários, códigos contábeis, etc.). Os dados usados para gerenciar e usar dados são chamados de *Metadados*. Como os dados não podem ser mantidos ou tocados, entender o que são e como usá-los requer definição e conhecimento na forma de Metadados. Os metadados se originam de uma variedade de processos relacionados à criação, processamento e uso de dados, incluindo arquitetura, modelagem, administração, governança, gerenciamento de qualidade de dados, desenvolvimento de sistemas, TI e operações comerciais e análises.

Figura 1 Gerenciamento de Dados Princípios

DATA MANAGEMENT PRINCIPLES

*Effective data
management requires
leadership
commitment*

Data is valuable

- Data is an asset with unique properties
- The value of data can and should be expressed in economic terms

Data Management Requirements are Business Requirements

- Managing data means managing the quality of data
- It takes Metadata to manage data
- It takes planning to manage data
- Data management requirements must drive Information Technology decisions

Data Management depends on diverse skills

- Data management is cross-functional
- Data management requires an enterprise perspective
- Data management must account for a range of perspectives

Data Management is lifecycle management

- Different types of data have different lifecycle characteristics
- Managing data includes managing the risks

Managing data includes managing the risks
associated with data

Figure 1 Data Management Principles

- **É preciso planejamento para gerenciar dados:** Mesmo pequenas organizações podem ter paisagens complexas de processos técnicos e de negócios. Dados são criados em muitos lugares e são movidos entre lugares para uso. Coordenar o trabalho e manter os resultados finais alinhados requer planejamento de uma perspectiva arquitetônica e de processo.
- **O gerenciamento de dados é multifuncional; ele requer uma gama de habilidades e conhecimentos:** Uma única equipe não pode gerenciar todos os dados de uma organização. O gerenciamento de dados requer habilidades técnicas e não técnicas e a capacidade de colaboração.
- **O gerenciamento de dados requer uma perspectiva empresarial:** o gerenciamento de dados tem aplicações locais, mas deve ser aplicado em toda a empresa para ser o mais eficaz possível.
Esta é uma das razões pelas quais o gerenciamento de dados e a governança de dados estão interligados.
- **O gerenciamento de dados deve levar em conta uma série de perspectivas:** Dados são fluidos. O gerenciamento de dados deve evoluir constantemente para acompanhar as maneiras como os dados são criados e usados e os consumidores de dados que os usam.
- **Gerenciamento de dados é gerenciamento de ciclo de vida:** Dados têm um ciclo de vida e gerenciar dados requer gerenciar seu ciclo de vida. Como dados geram mais dados, o ciclo de vida dos dados em si pode ser muito complexo. Práticas de gerenciamento de dados precisam levar em conta o ciclo de vida dos dados.
- **Diferentes tipos de dados têm características de ciclo de vida diferentes:** E por esse motivo, eles têm diferentes requisitos de gerenciamento. Práticas de gerenciamento de dados

é preciso reconhecer essas diferenças e ser flexível o suficiente para atender a diferentes tipos de requisitos do ciclo de vida dos dados.

- **Gerenciar dados inclui gerenciar os riscos associados aos dados:** além de ser um ativo, os dados também representam risco para uma organização. Os dados podem ser perdidos, roubados ou mal utilizados. As organizações devem considerar as implicações éticas de seus usos de dados. Os riscos relacionados aos dados devem ser gerenciados como parte do ciclo de vida dos dados.
- **Os requisitos de gerenciamento de dados devem orientar as decisões de Tecnologia da Informação:** Dados e gerenciamento de dados estão profundamente interligados com a tecnologia da informação e o gerenciamento de tecnologia da informação. Gerenciar dados requer uma abordagem que garanta que a tecnologia atenda, em vez de impulsionar, as necessidades estratégicas de dados de uma organização.
- **O gerenciamento de dados eficaz requer comprometimento da liderança:** O gerenciamento de dados envolve um conjunto complexo de processos que, para serem eficazes, exigem coordenação, colaboração e comprometimento. Chegar lá requer não apenas habilidades de gerenciamento, mas também a visão e o propósito que vêm da liderança comprometida.

2.5 Desafios de Gerenciamento de Dados Como o

gerenciamento de dados tem características distintas derivadas das propriedades dos dados em si, ele também apresenta desafios em seguir esses princípios. Detalhes desses desafios são discutidos nas Seções 2.5.1 a 2.5.13. Muitos desses desafios se referem a mais de um princípio.

*2.5.1 Dados diferem de outros ativos*⁶ Os ativos –

físicos podem ser apontados, tocados e movidos. Eles podem estar em apenas um lugar por vez. Os ativos financeiros devem ser contabilizados em um balanço patrimonial. No entanto, os dados são diferentes. Os dados não são tangíveis. No entanto, são duráveis; não se desgastam, embora o valor de

os dados frequentemente mudam conforme envelhecem. Os dados são fáceis de copiar e transportar. Mas não são fáceis de reproduzir se forem perdidos ou destruídos. Como não são consumidos quando usados, podem até ser roubados sem desaparecer. Os dados são dinâmicos e podem ser usados para vários propósitos. Os mesmos dados podem até ser usados por várias pessoas ao mesmo tempo – algo que é impossível com ativos físicos ou financeiros. Muitos usos de dados geram mais dados. A maioria das organizações deve gerenciar volumes crescentes de dados e a relação entre conjuntos de dados.

Essas diferenças tornam desafiador atribuir um valor monetário aos dados. Sem esse valor monetário, é difícil mensurar como os dados contribuem para o sucesso organizacional. Essas diferenças também levantam outras questões que afetam o gerenciamento de dados, como definir a propriedade dos dados, inventariar a quantidade de dados que uma organização tem, proteger contra o uso indevido de dados, gerenciar o risco associado à redundância de dados e definir e impor padrões para a Qualidade de Dados.

Apesar dos desafios com a medição do valor dos dados, a maioria das pessoas reconhece que os dados, de fato, têm valor. Os dados de uma organização são únicos para si mesmos. Se dados organizacionalmente únicos (como listas de clientes, inventários de produtos ou histórico de reivindicações) fossem perdidos ou destruídos, substituí-los seria impossível ou extremamente custoso. Os dados também são o meio pelo qual uma organização se conhece – é um meta-ativo que descreve outros ativos. Como tal, eles fornecem a base para o insight organizacional.

Dentro e entre organizações, dados e informações são essenciais para conduzir negócios. A maioria das transações comerciais operacionais envolve a troca de informações. A maioria das informações é trocada eletronicamente, criando uma trilha de dados. Essa trilha de dados pode servir a propósitos além de marcar as trocas que ocorreram. Ela pode fornecer informações sobre como uma organização funciona.

Devido ao papel importante que os dados desempenham em qualquer organização, eles precisam ser gerenciados com cuidado.

2.5.2 Valoração de Dados

Valor é a diferença entre o custo de uma coisa e o benefício

derivado daquela coisa. Para alguns ativos, como ações, calcular o valor é fácil. É a diferença entre o que as ações custaram quando foram compradas e o que foram vendidas. Mas para dados, esses cálculos são mais complicados, porque nem os custos nem os benefícios dos dados são padronizados.

Como os dados de cada organização são únicos, uma abordagem para avaliação de dados precisa começar articulando categorias gerais de custos e benefícios que podem ser aplicadas consistentemente dentro de uma organização.

As categorias de amostra incluem7 :-

- Custo de obtenção e armazenamento de dados
- Custo de substituição de dados em caso de perda
- Impacto para a organização se os dados estivessem faltando
- Custo de mitigação de risco e custo potencial de riscos associados aos dados
- Custo de melhoria de dados
- Benefícios de dados de maior qualidade
- Quanto os concorrentes pagariam pelos dados
- Por quanto os dados poderiam ser vendidos
- Receita esperada de usos inovadores de dados

Um desafio primário para a avaliação de ativos de dados é que o valor dos dados é contextual (o que é valioso para uma organização pode não ser valioso para outra) e frequentemente temporal (o que era valioso ontem pode não ser valioso hoje). Dito isso, dentro de uma organização, certos tipos de dados provavelmente serão consistentemente valiosos ao longo do tempo. Tomemos como exemplo informações confiáveis do cliente. As informações do cliente podem até se tornar mais valiosas ao longo do tempo, à medida que mais dados se acumulam relacionados à atividade do cliente.

Em relação à gestão de dados, estabelecer formas de associar valor financeiro aos dados é fundamental, uma vez que as organizações precisam entender os ativos em termos financeiros para fazerem análises consistentes.

decisões. Colocar valor nos dados se torna a base para colocar valor nas atividades de gerenciamento de dados.⁸ O processo de avaliação de dados também pode ser usado como um meio de gerenciamento de mudanças. Pedir aos profissionais de gerenciamento de dados e às partes interessadas que eles apoiam para entender o significado financeiro de seu trabalho pode ajudar uma organização a transformar sua compreensão de seus próprios dados e, por meio disso, sua abordagem ao gerenciamento de dados.

2.5.3 Qualidade dos dados

Garantir que os dados sejam de alta qualidade é essencial para o gerenciamento de dados.

As organizações gerenciam seus dados porque querem usá-los. Se não puderem confiar neles para atender às necessidades comerciais, o esforço para coletar, armazenar, proteger e habilitar o acesso a eles será desperdiçado. Para garantir que os dados atendam às necessidades comerciais, elas devem trabalhar com os consumidores de dados para definir essas necessidades, incluindo características que tornam os dados de alta qualidade.

Em grande parte porque os dados têm sido associados tão intimamente à tecnologia da informação, o gerenciamento da Qualidade de Dados tem sido historicamente tratado como uma reflexão tardia. As equipes de TI geralmente desdenham os dados que os sistemas que elas criam devem armazenar. Provavelmente foi um programador que primeiro observou "lixo entra, lixo sai" - e que sem dúvida queria deixar para lá. Mas as pessoas que querem usar os dados não podem se dar ao luxo de desprezar a qualidade. Elas geralmente assumem que os dados são confiáveis e confiáveis, até que tenham um motivo para duvidar dessas coisas. Uma vez que perdem a confiança, é difícil recuperá-la.

A maioria dos usos de dados envolve aprender com eles para aplicar esse aprendizado e criar valor. Exemplos incluem entender os hábitos do cliente para melhorar um produto ou serviço e avaliar o desempenho organizacional ou tendências de mercado para desenvolver uma melhor estratégia de negócios, etc. Dados de baixa qualidade terão um impacto negativo nessas decisões.

Tão importante quanto isso, dados de baixa qualidade são simplesmente caros para qualquer organização. As estimativas diferem, mas os especialistas acreditam que as organizações gastam entre 10 e 30% da receita lidando com problemas de qualidade de dados. A IBM estimou que o custo de dados de baixa qualidade nos EUA em 2016 foi de US\$ 3,1 trilhões.⁹ Muitos dos custos de dados de baixa qualidade são ocultos, indiretos e, portanto, difíceis de

medida. Outros, como multas, são diretos e fáceis de calcular. Os custos vêm de:

- Sucata e retrabalho
- Soluções alternativas e processos de correção ocultos
- Ineficiências organizacionais ou baixa produtividade
- Conflito organizacional
- Baixa satisfação no trabalho
- Insatisfação do cliente
- Custos de oportunidade, incluindo incapacidade de inovar
- Custos de conformidade ou multas
- Custos de reputação

Os benefícios correspondentes de dados de alta qualidade incluem:

- Melhor experiência do cliente
- Maior produtividade
- Risco reduzido
- Capacidade de aproveitar oportunidades
- Aumento da receita
- Vantagem competitiva obtida a partir de insights sobre clientes, produtos, processos e oportunidades

Como esses custos e benefícios implicam, gerenciar a Qualidade de Dados não é um trabalho único. Produzir dados de alta qualidade requer planejamento, comprometimento e uma mentalidade que incorpore qualidade em processos e sistemas. Todas as funções de gerenciamento de dados podem influenciar a Qualidade de Dados, para o bem ou para o mal, então todas elas devem levar isso em conta ao executar seu trabalho. (Consulte [o Capítulo 13](#)).

2.5.4 Planejamento para Melhores Dados

Conforme declarado na introdução do capítulo, derivar valor de dados não acontece por acidente. Requer planejamento em muitas formas. Começa com o reconhecimento de que as organizações podem controlar como obtêm e criam dados. Se elas veem os dados como um produto que elas criam, elas tomarão melhores decisões sobre eles ao longo de seu ciclo de vida. Essas decisões requerem pensamento sistêmico porque envolvem:

- As formas como os dados conectam processos de negócios que, de outra forma, poderiam ser vistos como separados
- A relação entre os processos de negócios e a tecnologia que os suporta
- O design e a arquitetura dos sistemas e os dados que eles produzem e armazenam
- As formas como os dados podem ser usados para promover a estratégia organizacional

O planejamento para melhores dados requer uma abordagem estratégica para arquitetura, modelagem e outras funções de design. Também depende da colaboração estratégica entre a liderança de negócios e de TI. E, claro, depende da capacidade de executar efetivamente em projetos individuais.

O desafio é que geralmente há pressões organizacionais, assim como as pressões perenes de tempo e dinheiro, que atrapalham um melhor planejamento. As organizações devem equilibrar objetivos de longo e curto prazo à medida que executam suas estratégias. Ter clareza sobre as compensações leva a melhores decisões.

2.5.5 Metadados e gerenciamento de dados

As organizações precisam de metadados confiáveis para gerenciar dados como um ativo.

Metadados neste sentido devem ser compreendidos de forma abrangente. Eles incluem não apenas os Metadados comerciais, técnicos e operacionais descritos no [Capítulo 12](#), mas também os Metadados incorporados na Arquitetura de Dados, modelos de dados, requisitos de segurança de dados,

padrões de integração e processos operacionais de dados. (Ver Capítulos 4 – 11.)

Metadados descrevem quais dados uma organização tem, o que eles representam, como são classificados, de onde vieram, como se movem dentro da organização, como evoluem por meio do uso, quem pode e não pode usá-los e se são de alta qualidade. Dados são abstratos. Definições e outras descrições de contexto permitem que sejam compreendidos. Eles tornam os dados, o ciclo de vida dos dados e os sistemas complexos que contêm dados compreensíveis.

O desafio é que Metadados são uma forma de dados e precisam ser gerenciados como tal. Organizações que não gerenciam bem seus dados geralmente não gerenciam seus Metadados de forma alguma. O gerenciamento de metadados geralmente fornece um ponto de partida para melhorias no gerenciamento de dados em geral.

2.5.6 O gerenciamento de dados é multifuncional

O gerenciamento de dados é um processo complexo. Os dados são gerenciados em diferentes lugares dentro de uma organização por equipes que têm responsabilidade por diferentes fases do ciclo de vida dos dados. O gerenciamento de dados requer habilidades de design para planejar sistemas, habilidades altamente técnicas para administrar hardware e construir software, habilidades de análise de dados para entender questões e problemas, habilidades analíticas para interpretar dados, habilidades de linguagem para trazer consenso para definições e modelos, bem como pensamento estratégico para ver oportunidades de atender clientes e atingir metas.

O desafio é fazer com que pessoas com essa gama de habilidades e perspectivas reconheçam como as peças se encaixam para que colaborem bem enquanto trabalham em direção a objetivos comuns.

2.5.7 Estabelecendo uma Perspectiva Empresarial

Gerenciar dados requer entender o escopo e o alcance dos dados dentro de uma organização. Dados são um dos "horizontais" de uma organização. Eles se movem por verticais, como vendas, marketing e operações... Ou pelo menos deveriam. Dados não são exclusivos apenas de uma organização; às vezes, são exclusivos de um departamento ou outra subparte de uma organização. Como dados são frequentemente vistos simplesmente como um subproduto,

produto de processos operacionais (por exemplo, registros de transações de vendas são o subproduto do processo de venda), nem sempre é planejado além da necessidade imediata.

Mesmo dentro de uma organização, os dados podem ser díspares. Os dados se originam em vários lugares dentro de uma organização. Departamentos diferentes podem ter maneiras diferentes de representar o mesmo conceito (por exemplo, cliente, produto, fornecedor). Como qualquer pessoa envolvida em um projeto de integração de dados ou Master Data Management pode testemunhar, diferenças sutis (ou flagrantes) em escolhas representacionais apresentam desafios no gerenciamento de dados em uma organização. Ao mesmo tempo, as partes interessadas presumem que os dados de uma organização devem ser coerentes, e um objetivo do gerenciamento de dados é fazê-los se encaixar de maneiras de senso comum para que sejam utilizáveis por uma ampla gama de consumidores de dados.

Um dos motivos pelos quais a governança de dados se tornou cada vez mais importante é ajudar as organizações a tomar decisões sobre dados em todos os setores. (Consulte [o Capítulo 3.](#))

2.5.8 Contabilização de outras perspectivas As

organizações de hoje usam dados que criam internamente, bem como dados que adquirem de fontes externas. Elas precisam contabilizar diferentes requisitos legais e de conformidade em linhas nacionais e setoriais. Pessoas que criam dados frequentemente esquecem que outra pessoa usará esses dados mais tarde. O conhecimento dos usos potenciais dos dados permite um melhor planejamento para o ciclo de vida dos dados e, com isso, para dados de melhor qualidade. Os dados também podem ser mal utilizados. A contabilização desse risco reduz a probabilidade de uso indevido.

2.5.9 O ciclo de vida dos dados

Como outros ativos, os dados têm um ciclo de vida. Para gerenciar efetivamente os ativos de dados, as organizações precisam entender e planejar o ciclo de vida dos dados. Dados bem gerenciados são gerenciados estrategicamente, com uma visão de como a organização usará seus dados. Uma organização estratégica definirá não apenas seus requisitos de conteúdo de dados, mas também seus requisitos de gerenciamento de dados. Isso inclui políticas e expectativas de uso, qualidade, controles e segurança; uma abordagem empresarial para

arquitetura e design; e uma abordagem sustentável tanto para infraestrutura quanto para desenvolvimento de software.

O ciclo de vida dos dados é baseado no ciclo de vida do produto. Não deve ser confundido com o ciclo de vida do desenvolvimento de sistemas. Conceitualmente, o ciclo de vida dos dados é fácil de descrever (veja [a Figura 2](#)). Ele inclui processos que criam ou obtêm dados, aqueles que os movem, transformam e armazenam e permitem que sejam mantidos e compartilhados, e aqueles que os usam ou aplicam, bem como aqueles que os descartam.¹⁰ Ao longo de seu ciclo de vida, os dados podem ser limpos, transformados, mesclados, aprimorados ou agregados. À medida que os dados são usados ou aprimorados, novos dados são frequentemente criados, então o ciclo de vida tem iterações internas que não são mostradas no diagrama. Os dados raramente são estáticos. O gerenciamento de dados envolve um conjunto de processos interconectados alinhados com o ciclo de vida dos dados.

As especificidades do ciclo de vida dos dados dentro de uma determinada organização podem ser bastante complicadas, porque os dados não só têm um ciclo de vida, como também têm uma linhagem (ou seja, um caminho ao longo do qual se movem do seu ponto de origem para o seu ponto de uso, às vezes chamado de *cadeia de dados*). Entender a linhagem dos dados requer documentar a origem dos conjuntos de dados, bem como seu movimento e transformação por meio de sistemas onde são acessados e usados. O ciclo de vida e a linhagem se cruzam e podem ser entendidos em relação um ao outro. Quanto melhor uma organização entender o ciclo de vida e a linhagem de seus dados, mais capaz ela será de gerenciar seus dados.

O foco do gerenciamento de dados no ciclo de vida dos dados tem várias implicações importantes:

- **Criação e uso são os pontos mais críticos no ciclo de vida dos dados:** o gerenciamento de dados deve ser executado com um entendimento de como os dados são produzidos, ou obtidos, bem como como os dados são usados. Custa dinheiro para produzir dados. Os dados são valiosos somente quando são consumidos ou aplicados. ([Consulte os Capítulos 5, 6, 8, 11 e 14.](#))

Figura 2 Chave do ciclo de vida dos dados Atividades

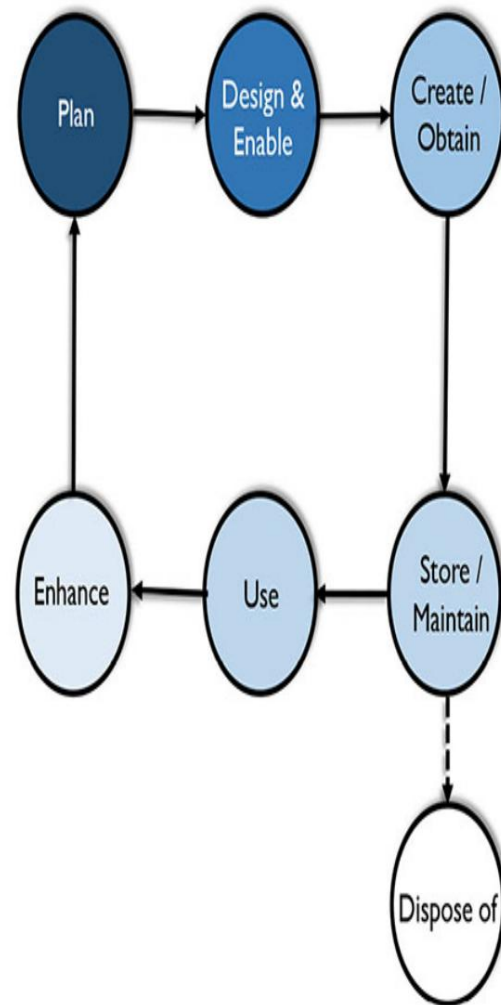


Figure 2 Data Lifecycle Key Activities

- **A Qualidade dos Dados deve ser gerenciada durante todo o ciclo de vida dos dados:** O Gerenciamento da Qualidade dos Dados é essencial para o gerenciamento de dados. Dados de baixa qualidade representam custo e risco, em vez de valor. As organizações geralmente acham desafiador gerenciar a qualidade dos dados porque, conforme descrito anteriormente, os dados geralmente são criados como um subproduto ou

processos operacionais e organizações frequentemente não definem padrões explícitos para qualidade. Como a qualidade da qualidade pode ser impactada por uma série de eventos do ciclo de vida, a qualidade deve ser planejada como parte do ciclo de vida dos dados (veja [Capítulo 13](#)).

- **A qualidade dos metadados deve ser gerenciada durante todo o ciclo de vida dos dados:** como os metadados são uma forma de dados e as organizações dependem deles para gerenciar outros dados, a qualidade dos metadados deve ser gerenciada da mesma forma que a qualidade de outros dados (consulte [o Capítulo 12](#)).
- **A Segurança de Dados deve ser gerenciada durante todo o ciclo de vida dos dados:** O gerenciamento de dados também inclui garantir que os dados estejam seguros e que os riscos associados aos dados sejam mitigados. Os dados que exigem proteção devem ser protegidos durante todo o seu ciclo de vida, da criação ao descarte (consulte [o Capítulo 7 Segurança de Dados](#)).
- **Os esforços de gerenciamento de dados devem se concentrar nos dados mais críticos:** As organizações produzem muitos dados, uma grande parte dos quais nunca é realmente usada. Tentar gerenciar cada pedaço de dados não é possível. O gerenciamento do ciclo de vida requer foco nos dados mais críticos de uma organização e minimização do ROT (dados redundantes, obsoletos, triviais) dos dados (Aiken, 2014).

2.5.10 Diferentes Tipos de Dados O

gerenciamento de dados se torna mais complicado pelo fato de que há diferentes tipos de dados que têm diferentes requisitos de gerenciamento de ciclo de vida. Qualquer sistema de gerenciamento precisa classificar os objetos que são gerenciados. Os dados podem ser classificados por tipo de dados (por exemplo, dados transacionais, Dados de Referência, Dados Mestres, Metadados; alternativamente, dados de categoria, dados de recursos, dados de eventos, dados de transações detalhadas) ou por conteúdo (por exemplo, domínios de dados, áreas de assunto) ou por formato ou pelo nível de proteção que os dados exigem. Os dados também podem ser classificados por como e onde são armazenados ou acessados. (Consulte

Capítulos [5](#) e [10](#).)

Como diferentes tipos de dados têm diferentes requisitos, são associados a diferentes riscos e desempenham diferentes papéis dentro de uma organização, muitas das ferramentas de gerenciamento de dados são focadas em aspectos de classificação e controle (Bryce, 2005). Por exemplo, Master Data tem diferentes usos e, consequentemente, diferentes requisitos de gerenciamento do que dados transacionais. (Consulte os Capítulos 9, [10](#), [12](#) e [14](#).)

2.5.11 Dados e Risco

Dados não representam apenas valor, eles também representam risco. Dados de baixa qualidade (imprecisos, incompletos ou desatualizados) obviamente representam risco porque suas informações não estão corretas. Mas dados também são arriscados porque podem ser mal compreendidos e mal utilizados.

As organizações obtêm o máximo valor dos dados da mais alta qualidade – disponíveis, relevantes, completos, precisos, consistentes, oportunos, utilizáveis, significativos e compreendidos. No entanto, para muitas decisões importantes, temos lacunas de informação – a diferença entre o que sabemos e o que precisamos saber para tomar uma decisão eficaz. As lacunas de informação representam passivos empresariais com impactos potencialmente profundos na eficácia operacional e na lucratividade. As organizações que reconhecem o valor de dados de alta qualidade podem tomar medidas concretas e proativas para melhorar a qualidade e a usabilidade de dados e informações dentro de estruturas culturais regulatórias e éticas.

O papel crescente da informação como um ativo organizacional em todos os setores levou a um foco maior por parte de reguladores e legisladores sobre os usos e abusos potenciais da informação. De Sarbanes-Oxley (com foco em controles sobre precisão e validade de dados de transações financeiras da transação ao balanço) a Solvência II (com foco na linhagem de dados e qualidade dos dados que sustentam modelos de risco e adequação de capital no setor de seguros), ao rápido crescimento na última década de regulamentações de privacidade de dados (abrangendo o processamento de dados sobre pessoas em uma ampla gama de indústrias e jurisdições), está claro que, enquanto ainda estamos esperando que a Contabilidade coloque a Informação no balanço como um ativo, a

o ambiente regulatório espera cada vez mais vê-lo no registro de riscos, com mitigações e controles apropriados sendo aplicados.

Da mesma forma, à medida que os consumidores se tornam mais conscientes de como seus dados são usados, eles esperam não apenas uma operação mais suave e eficiente dos processos, mas também proteção de suas informações e respeito por sua privacidade. Isso significa que o escopo de quem são nossos stakeholders estratégicos como profissionais de gerenciamento de dados pode frequentemente ser mais amplo do que tradicionalmente pode ter sido o caso. (Consulte [os Capítulos 2 Ética no Manuseio de Dados](#) e [7 Segurança de Dados](#).)

Cada vez mais, o impacto do gerenciamento de informações no balanço patrimonial, infelizmente, surge com muita frequência quando esses riscos não são gerenciados e os acionistas votam com suas carteiras de ações, os reguladores impõem multas ou restrições às operações e os clientes votam com suas carteiras.

2.5.12 Gerenciamento de Dados e Tecnologia

Conforme observado na introdução do capítulo e em outros lugares, as atividades de gerenciamento de dados são abrangentes e exigem habilidades técnicas e comerciais. Como quase todos os dados de hoje são armazenados eletronicamente, as táticas de gerenciamento de dados são fortemente influenciadas pela tecnologia. Desde o início, o conceito de gerenciamento de dados tem sido profundamente interligado ao gerenciamento de tecnologia. Esse legado continua.

Em muitas organizações, há uma tensão constante entre o desejo de construir novas tecnologias e o desejo de ter dados mais confiáveis — como se os dois fossem opostos em vez de necessários um ao outro.

O gerenciamento de dados bem-sucedido requer decisões sólidas sobre tecnologia, mas gerenciar tecnologia não é o mesmo que gerenciar dados. As organizações precisam entender o impacto da tecnologia sobre os dados, a fim de evitar que a tentação tecnológica conduza suas decisões sobre dados. Em vez disso, os requisitos de dados alinhados com a estratégia de negócios devem conduzir as decisões sobre tecnologia.

2.5.13 A gestão eficaz de dados requer liderança e comprometimento

O Manifesto de Dados do Líder (2017) reconheceu que “as melhores oportunidades de crescimento orgânico de uma organização estão nos dados”. Embora a maioria das organizações reconheça seus dados como um ativo, elas estão longe de serem orientadas por dados. Muitas não sabem quais dados têm ou quais dados são mais críticos para seus negócios. Elas confundem dados e tecnologia da informação e administram mal ambos. Elas não abordam os dados estrategicamente. E subestimam o trabalho envolvido com o gerenciamento de dados. Essas condições aumentam os desafios do gerenciamento de dados e apontam para um fator crítico para o potencial de sucesso de uma organização: liderança comprometida e o envolvimento de todos em todos os níveis da organização.¹¹ Os desafios descritos aqui devem

ênfaticamente este ponto: o gerenciamento de dados não é fácil nem simples. Mas como poucas organizações o fazem bem, é uma fonte de oportunidade amplamente inexplorada.

Para se tornar melhor nisso, é preciso visão, planejamento e disposição para mudar. (Veja os Capítulos [15-17.](#))

A defesa do papel de Chief Data Officer (CDO) decorre do reconhecimento de que o gerenciamento de dados apresenta desafios únicos e que o gerenciamento de dados bem-sucedido deve ser orientado para os negócios, em vez de para a TI. Um CDO pode liderar iniciativas de gerenciamento de dados e permitir que uma organização alavanque seus ativos de dados e obtenha vantagem competitiva deles. No entanto, um CDO não apenas lidera iniciativas. Ele ou ela também deve liderar uma mudança cultural que permita que uma organização tenha uma abordagem mais estratégica para seus dados.

2.6 Estratégia de Gerenciamento de Dados

Uma estratégia é um conjunto de escolhas e decisões que, juntas, traçam um curso de ação de alto nível para atingir objetivos de alto nível. No jogo de xadrez, uma estratégia é um conjunto sequenciado de movimentos para vencer por xeque-mate ou sobreviver por impasse. Um *plano* estratégico é um curso de ação de alto nível para atingir objetivos de alto nível.

Uma estratégia de dados deve incluir planos de negócios para usar informações para vantagem competitiva e dar suporte a objetivos empresariais. A estratégia de dados deve vir de uma compreensão das necessidades de dados inerentes à estratégia de negócios: quais dados a organização precisa, como ela obterá

os dados, como serão gerenciados e garantirá sua confiabilidade ao longo do tempo, e como serão utilizados.

Normalmente, uma estratégia de dados requer uma estratégia de programa de Gerenciamento de Dados de suporte – um plano para manter e melhorar a qualidade dos dados, integridade dos dados, acesso e segurança, ao mesmo tempo em que mitiga riscos conhecidos e implícitos. A estratégia também deve abordar desafios conhecidos relacionados ao gerenciamento de dados.

Em muitas organizações, a estratégia de gerenciamento de dados é de propriedade e mantida pelo CDO e promulgada por meio de uma equipe de governança de dados, apoiada por um Data Governance Council. Frequentemente, o CDO elaborará uma estratégia de dados inicial e uma estratégia de gerenciamento de dados antes mesmo de um Data Governance Council ser formado, a fim de obter o comprometimento da alta gerência em estabelecer a administração e governança de dados.

Os componentes de uma estratégia de gerenciamento de dados devem incluir:

- Uma visão convincente para a gestão de dados
- Um caso de negócios resumido para gerenciamento de dados, com exemplos selecionados
- Princípios orientadores, valores e perspectivas de gestão
- A missão e os objetivos direcionais de longo prazo da gestão de dados
- Medidas propostas para o sucesso da gestão de dados
- Objetivos do programa de gerenciamento de dados de curto prazo (12 a 24 meses) que sejam SMART (específicos, mensuráveis, acionáveis, realistas e com limite de tempo)
- Descrições de funções e organizações de gerenciamento de dados, juntamente com um resumo de suas responsabilidades e direitos de decisão
- Descrições dos componentes e iniciativas do programa de gerenciamento de dados
- Um programa de trabalho priorizado com limites de escopo

- Um rascunho do roteiro de implementação com projetos e itens de ação

Os resultados do planejamento estratégico para gerenciamento de dados incluem:

- **Uma Carta de Gerenciamento de Dados:** Visão geral, caso de negócios, metas, princípios orientadores, medidas de sucesso, fatores críticos de sucesso, riscos reconhecidos, modelo operacional, etc.
- **Uma Declaração de Escopo de Gerenciamento de Dados:** Metas e objetivos para um horizonte de planejamento (geralmente 3 anos) e as funções, organizações e líderes individuais responsáveis por atingir esses objetivos.
- **Um roteiro de implementação de gerenciamento de dados:** identificação de programas, projetos, atribuições de tarefas e marcos de entrega específicos (consulte o [Capítulo 15](#)).

A estratégia de gerenciamento de dados deve abordar todas as áreas de conhecimento do DAMA Data Management Framework relevantes para a organização. (Consulte a [Figura 5 O DAMA-DMBOK2 Data Management Framework \(The DAMA Wheel e as Seções 3.3 e 4.\)](#))

3. Estruturas de gerenciamento de dados

O gerenciamento de dados envolve um conjunto de funções interdependentes, cada uma com seus próprios objetivos, atividades e responsabilidades. Os profissionais de gerenciamento de dados precisam levar em conta os desafios inerentes à tentativa de derivar valor de um ativo empresarial abstrato, ao mesmo tempo em que equilibram objetivos estratégicos e operacionais, requisitos comerciais e técnicos específicos, demandas de risco e conformidade e entendimentos conflitantes sobre o que os dados representam e se são de alta qualidade.

Há muito o que acompanhar, por isso é útil ter uma estrutura para entender o gerenciamento de dados de forma abrangente e

veja relacionamentos entre suas peças componentes. Como as funções dependem umas das outras e precisam estar alinhadas, em qualquer organização, as pessoas responsáveis pelos diferentes aspectos do gerenciamento de dados precisam colaborar se a organização quiser derivar valor de seus dados.

Frameworks desenvolvidos em diferentes níveis de abstração fornecem uma gama de perspectivas sobre como abordar o gerenciamento de dados. Essas perspectivas fornecem insights que podem ser usados para esclarecer estratégias, desenvolver roteiros, organizar equipes e alinhar funções.

As ideias e conceitos apresentados no DMBOK2 serão aplicados de forma diferente em todas as organizações. A abordagem de uma organização para o gerenciamento de dados depende de fatores-chave, como seu setor, a gama de dados que ela usa, sua cultura, nível de maturidade, estratégia, visão e os desafios específicos que ela está abordando. As estruturas descritas nesta seção fornecem algumas lentes através das quais é possível ver o gerenciamento de dados e aplicar os conceitos apresentados no DMBOK.

- Os dois primeiros, o Modelo de Alinhamento Estratégico e o Modelo de Informação de Amsterdã, mostram relacionamentos de alto nível que influenciam a maneira como uma organização gerencia dados.
- O DAMA DMBOK Framework (The DAMA Wheel, Hexagon, and Context Diagram) descreve as Áreas de Conhecimento de Gerenciamento de Dados, conforme definidas pelo DAMA, e explica como sua representação visual dentro do DMBOK.
- Os dois últimos tomam a Roda DAMA como ponto de partida e reorganizam as peças para entender e descrever melhor as relações entre elas.

3.1 Modelo de Alinhamento Estratégico O

Modelo de Alinhamento Estratégico (Henderson e Venkatraman, 1999) abstrai os impulsionadores fundamentais para qualquer abordagem de dados

gerenciamento. Em seu centro está o relacionamento entre dados e informações. As informações são mais frequentemente associadas à estratégia de negócios e ao uso operacional de dados. Os dados são associados à tecnologia da informação e aos processos que dão suporte ao gerenciamento físico de sistemas que tornam os dados acessíveis para uso.

Em torno desse conceito estão os quatro domínios fundamentais da escolha estratégica: estratégia de negócios, estratégia de tecnologia da informação, infraestrutura e processos organizacionais e infraestrutura e processos de tecnologia da informação.

O Modelo de Alinhamento Estratégico totalmente articulado é mais complexo do que o ilustrado na Figura 3. Cada um dos hexágonos de canto tem suas próprias dimensões subjacentes. Por exemplo, dentro da estratégia de negócios e de TI, há uma necessidade de levar em conta o escopo, as competências e a governança. As operações devem levar em conta a infraestrutura, os processos e as habilidades. Os relacionamentos entre as peças ajudam uma organização a entender tanto o ajuste estratégico dos diferentes componentes quanto a integração funcional das peças. Até mesmo a representação de alto nível do modelo é útil para entender os fatores organizacionais que influenciam as decisões sobre dados e gerenciamento de dados.

Figura 3 Modelo de Alinhamento Estratégico¹²

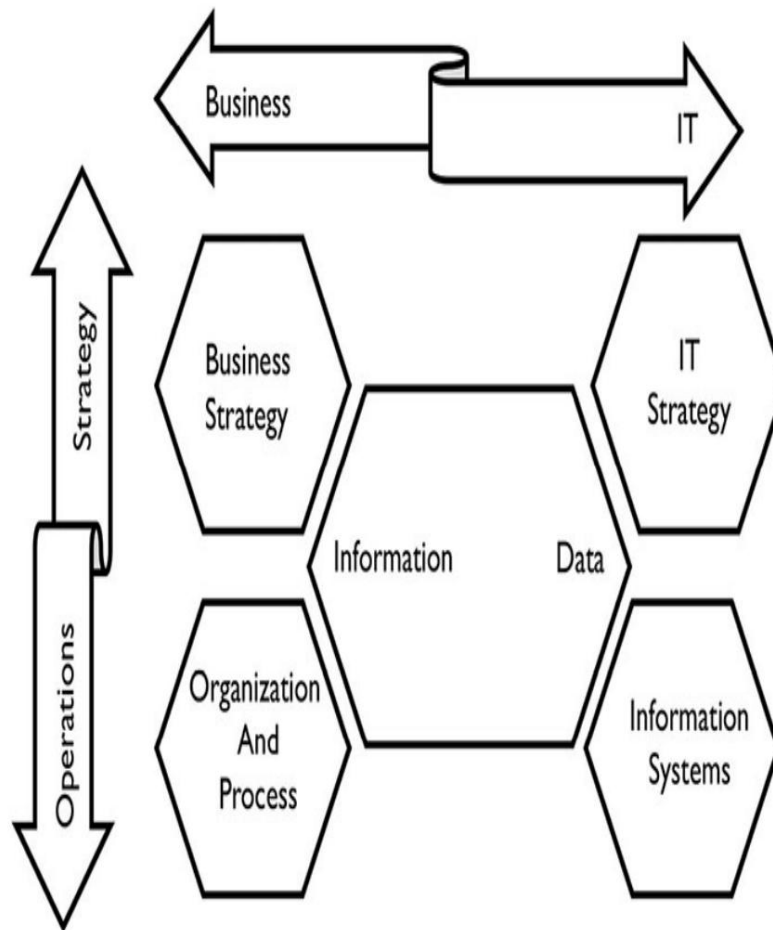


Figure 3 Strategic Alignment Model¹²

3.2 O Modelo de Informação de Amsterdã

O Amsterdam Information Model, assim como o Strategic Alignment Model, adota uma perspectiva estratégica sobre o alinhamento de negócios e TI (Abcouwer, Maes e Truijens, 1997),¹³ conhecido como 9-cell, ele reconhece uma camada intermediária que se concentra em estrutura e táticas, incluindo planejamento e arquitetura. Além disso, ele reconhece a necessidade de comunicação de informações (expressa como o pilar de governança de informações e qualidade de dados na Figura 4).

Os criadores das estruturas SAM e AIM descrevem em detalhes a relação entre os componentes, tanto de uma perspectiva horizontal (estratégia de negócios/TI) quanto vertical (estratégia de negócios/operações de negócios).

Figura 4 Modelo de Informação de Amsterdã¹⁴

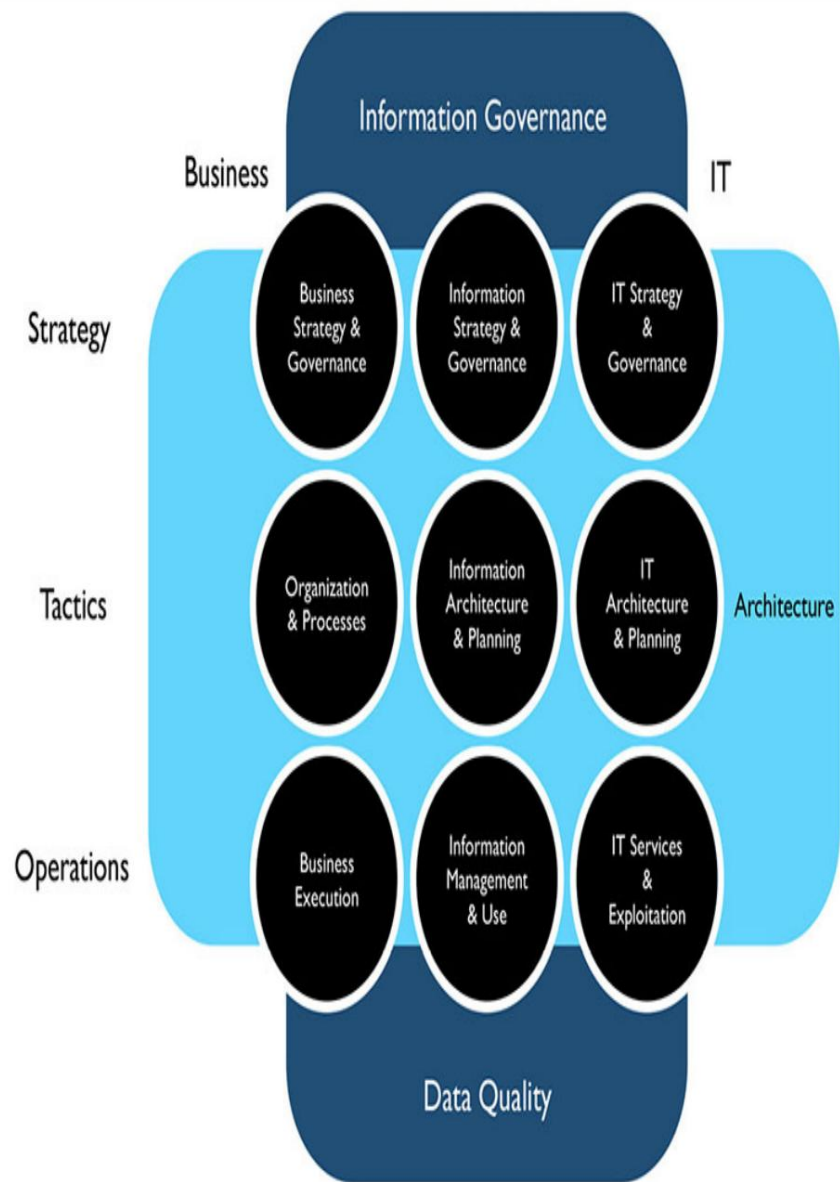


Figure 4 Amsterdam Information Model¹⁴

3.3 A estrutura DAMA-DMBOK

O DAMA-DMBOK Framework se aprofunda mais nas Knowledge Areas que compõem o escopo geral do gerenciamento de dados. Três visuais descrevem o Data Management Framework do DAMA:

- A Roda DAMA ([Figura 5](#))
- O hexágono dos Fatores Ambientais ([Figura 6](#))
- Diagrama de Contexto da Área de Conhecimento ([Figura 7](#))

A Roda DAMA define as Áreas de Conhecimento de Gerenciamento de Dados. Ela coloca a governança de dados no centro das atividades de gerenciamento de dados, já que a governança é necessária para consistência dentro e equilíbrio entre as funções. As outras Áreas de Conhecimento (Arquitetura de Dados, Modelagem de Dados, etc.) são equilibradas em torno da Roda.

Todas elas são partes necessárias de uma função de gerenciamento de dados madura, mas podem ser implementadas em momentos diferentes, dependendo dos requisitos da organização. Essas Áreas de Conhecimento são o foco dos Capítulos 3 a 13 do DMBOK2. (Veja [a Figura 5](#).)

O hexágono Fatores Ambientais mostra a relação entre pessoas, processos e tecnologia e fornece uma chave para a leitura dos diagramas de contexto do DMBOK. Ele coloca objetivos e princípios no centro, uma vez que estes fornecem orientação sobre como as pessoas devem executar atividades e usar efetivamente as ferramentas necessárias para o gerenciamento de dados bem-sucedido. (Veja [a Figura 6](#).)

Figura 5 O DAMA-DMBOK2
Estrutura de gerenciamento de dados
(A Roda DAMA)

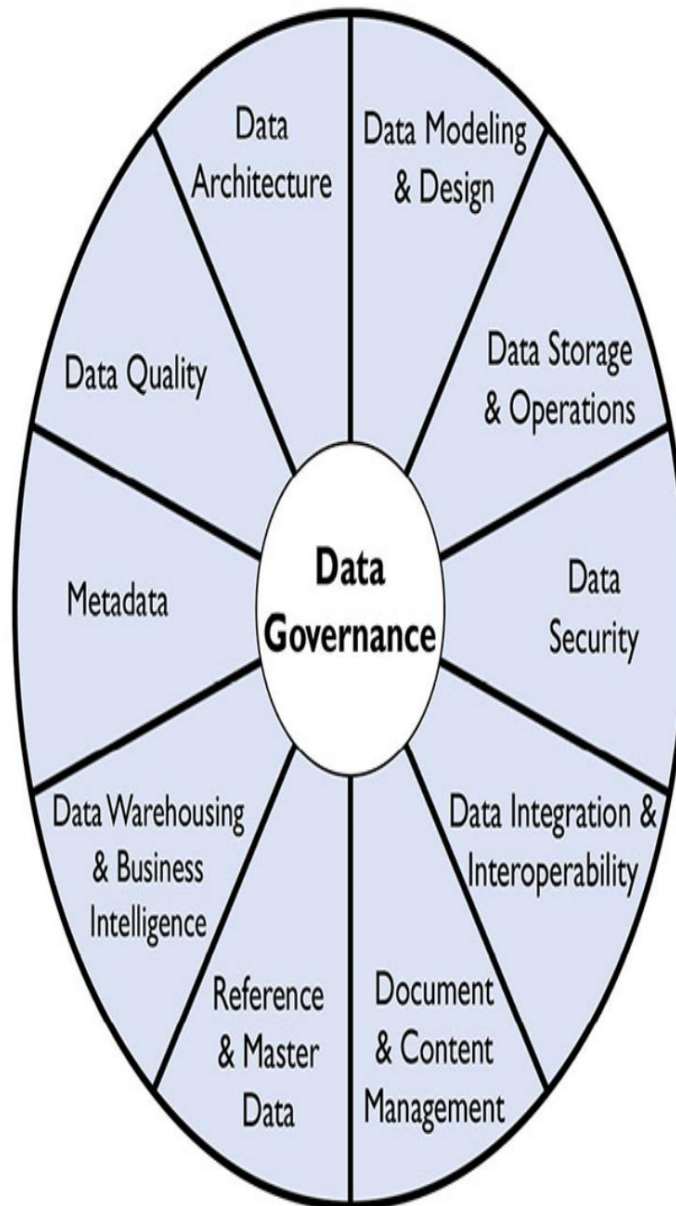


Figure 5 The DAMA-DMBOK2 Data Management Framework (The DAMA Wheel)

Figura 6 DAMA Ambiental Fatores Hexágono

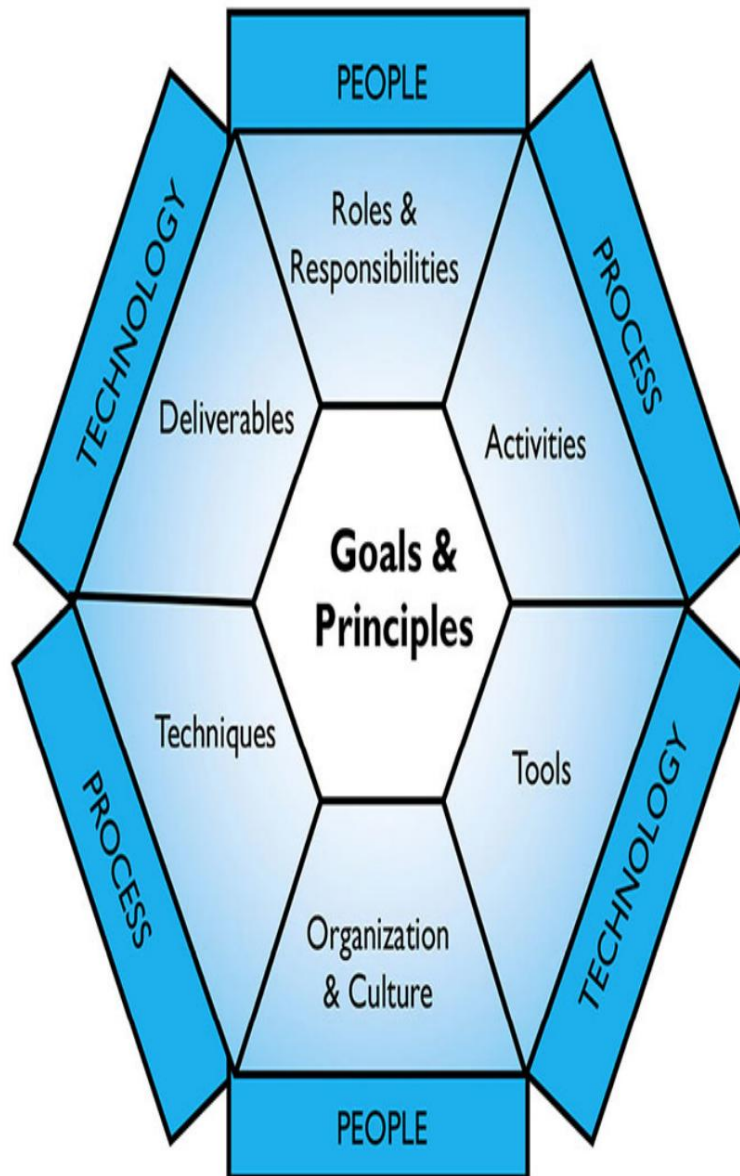


Figure 6 DAMA Environmental Factors Hexagon

Os Diagramas de Contexto da Área de Conhecimento (Veja [a Figura 7](#)) descrevem os detalhes das Áreas de Conhecimento, incluindo detalhes relacionados a pessoas, processos e tecnologia. Eles são baseados no conceito de um diagrama SIPOC usado para gerenciamento de produtos (Fornecedores, Entradas, Processos, Saídas e Consumidores). Os Diagramas de Contexto colocam as atividades no centro, uma vez que produzem os entregáveis que atendem aos requisitos das partes interessadas.

Cada diagrama de contexto começa com a definição da Área de Conhecimento

e objetivos. As atividades que impulsionam os objetivos (centro) são classificadas em quatro fases: Planejar (P), Desenvolver (D), Operar (O) e Controlar (C). No lado esquerdo (fluindo para as atividades) estão as Entradas e Fornecedores. No lado direito (fluindo das atividades) estão os Entregáveis e os Consumidores. Os Participantes são listados abaixo das Atividades. Na parte inferior estão as Ferramentas, Técnicas e Métricas que influenciam os aspectos da Área de Conhecimento.

As listas no diagrama de contexto são ilustrativas, não exaustivas. Os itens serão aplicados de forma diferente a diferentes organizações. As listas de funções de alto nível incluem apenas as funções mais importantes. Cada organização pode adaptar esse padrão para atender às suas próprias necessidades.

Figura 7 Área de conhecimento

Diagrama de Contexto

GENERIC CONTEXT DIAGRAM

Definition: High-level description of the knowledge area

Goals: Purposes of the Knowledge Area

1. Goal 1
2. Goal 2

Business
Drivers

Inputs:

- Input 1
- Input 2
- Input 3

Inputs are generally
outputs from other
Knowledge Areas

Activities:

1. **Planning Activity / Activity Group (P)**
 1. Sub activity
 2. Sub activity
2. **Control Activity / Activity Group (C)**
3. **Development Activity / Activity Group (D)**
4. **Operational Activity / Activity Group (O)**

Deliverables:

- Deliverable 1
- Deliverable 2
- Deliverable 3

Deliverables are
generally inputs to other
Knowledge Areas

Suppliers:

- Supplier 1
- Supplier 2

Participants:

- Role 1
- Role 2

Consumers:

- Role 1
- Role 2

Technical
Drivers

Techniques:

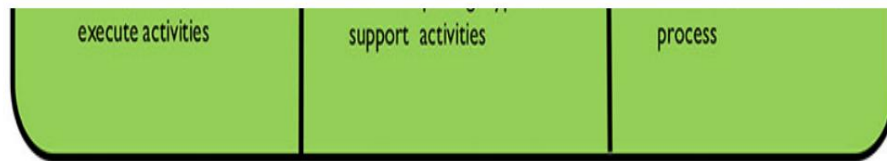
- Methods and procedures to

Tools:

- Software package types to

Metrics:

- Measurable results of the



(P) Planning, (C) Control, (D) Development, (O) Operations

Figure 7 Knowledge Area Context Diagram

Os componentes do diagrama de contexto incluem:

1. **Definição:** Esta seção define concisamente o Conhecimento Área.
2. **Os objetivos** descrevem o propósito da Área de Conhecimento e os princípios fundamentais que orientam o desempenho das atividades dentro de cada Área de Conhecimento.
3. **Atividades** são as ações e tarefas necessárias para atingir os objetivos da Área de Conhecimento. Algumas atividades são descritas em termos de subatividades, tarefas e etapas.
As atividades são classificadas em quatro categorias: Planejar, Desenvolver, Operar e Controlar.
 - **(P) As atividades de planejamento** definem o curso estratégico e tático para atingir as metas de gerenciamento de dados. As atividades de planejamento ocorrem de forma recorrente.
 - **(D) As atividades de desenvolvimento** são organizadas em torno do ciclo de vida de desenvolvimento do sistema (SDLC) (análise, design, construção, teste, preparação e implantação).
 - **(C) As atividades de controle** garantem a qualidade contínua dos dados e a integridade, confiabilidade e segurança dos sistemas por meio dos quais os dados são acessados e usados.
 - **(O) As Atividades Operacionais** dão suporte ao uso, manutenção e aprimoramento de sistemas e

processos pelos quais os dados são acessados e usados.

4. **As entradas** são as coisas tangíveis que cada Área de Conhecimento requer para iniciar suas atividades. Muitas atividades requerem as mesmas entradas. Por exemplo, muitas requerem conhecimento da Estratégia de Negócios como entrada.
5. **Os produtos finais** são os resultados das atividades dentro do Área de Conhecimento, as coisas tangíveis que cada função é responsável por produzir. Os entregáveis podem ser fins em si mesmos ou entradas em outras atividades. Vários entregáveis primários são criados por múltiplas funções.
6. **Funções e responsabilidades** descrevem como indivíduos e equipes contribuem para atividades dentro da área de conhecimento. As funções são descritas conceitualmente, com foco em grupos de funções necessárias na maioria das organizações. As funções para indivíduos são definidas em termos de habilidades e requisitos de qualificação. O Skills Framework for the Information Age (SFIA) foi usado para ajudar a alinhar os títulos das funções. Muitas funções serão multifuncionais.¹⁵ (Consulte o [Capítulo 16](#)).
7. **Fornecedores** são as pessoas responsáveis por fornecer ou viabilizar o acesso aos insumos para as atividades.
8. **Consumidores:** aqueles que se beneficiam diretamente dos principais resultados criados pelas atividades de gerenciamento de dados.
9. **Os participantes** são as pessoas que executam, gerenciam o execução ou aprovação das atividades na Área de Conhecimento.
10. **Ferramentas** são os aplicativos e outras tecnologias que permitem os objetivos da Área de Conhecimento.¹⁶
11. **Técnicas** são os métodos e procedimentos usados para executar atividades e produzir entregas dentro de uma Área de Conhecimento. Técnicas incluem convenções comuns, recomendações de melhores práticas, padrões

e protocolos e, quando aplicável, abordagens alternativas emergentes.

12. **Métricas** são padrões para medição ou avaliação de desempenho, progresso, qualidade, eficiência ou outro efeito. As seções de métricas identificam facetas mensuráveis do trabalho que é feito dentro de cada Área de Conhecimento. Métricas também podem medir características mais abstratas, como melhoria ou valor.

Enquanto a Roda DAMA apresenta o conjunto de Áreas de Conhecimento em alto nível, o Hexágono reconhece componentes da estrutura das Áreas de Conhecimento, e os Diagramas de Contexto apresentam os detalhes dentro de cada Área de Conhecimento. Nenhuma das peças da estrutura de Gerenciamento de Dados DAMA existente descreve o relacionamento entre as diferentes Áreas de Conhecimento. Esforços para abordar essa questão resultaram em reformulações da Estrutura DAMA, que são descritas nas próximas duas seções.

3.4 Pirâmide DMBOK (Aiken)

Se perguntadas, muitas organizações diriam que querem obter o máximo de seus dados – elas estão se esforçando para alcançar aquela pirâmide dourada de práticas avançadas (mineração de dados, análise, etc.). Mas essa pirâmide é apenas o topo de uma estrutura maior, um pináculo em uma fundação. A maioria das organizações não tem o luxo de definir uma estratégia de gerenciamento de dados antes de começar a ter que gerenciar dados. Em vez disso, elas constroem em direção a essa capacidade, na maioria das vezes em condições abaixo do ideal.

A estrutura de Peter Aiken usa as áreas funcionais do DMBOK para descrever a situação em que muitas organizações se encontram.

Uma organização pode usá-lo para definir um caminho a seguir para um estado em que eles tenham dados e processos confiáveis para dar suporte a objetivos estratégicos de negócios. Ao tentar atingir esse objetivo, muitas organizações passam por uma progressão lógica semelhante de etapas ([consulte a Figura 8](#)):

- **Fase 1:** A organização adquire um aplicativo que

inclui capacidades de banco de dados. Isso significa que a organização tem um ponto de partida para modelagem/design de dados, armazenamento de dados e segurança de dados (por exemplo, deixar algumas pessoas entrarem e manter outras fora). Para fazer o sistema funcionar dentro do ambiente deles e com seus dados, é necessário trabalhar na integração e interoperabilidade.

- **Fase 2:** Depois que começarem a usar o aplicativo, eles encontrarão desafios com a qualidade dos dados. Mas chegar a dados de maior qualidade depende de Metadados confiáveis e Arquitetura de Dados consistente. Eles fornecem clareza sobre como os dados de diferentes sistemas funcionam juntos.
- **Fase 3:** Práticas disciplinadas para gerenciar qualidade de dados, metadados e arquitetura exigem governança de dados que forneça suporte estrutural para atividades de gerenciamento de dados. A Governança de Dados também permite a execução de iniciativas estratégicas, como Gerenciamento de Documentos e Conteúdo, Gerenciamento de Dados de Referência, Gerenciamento de Dados Mestres, Armazenamento de Dados e Business Intelligence, que viabilizam totalmente as práticas avançadas dentro da pirâmide dourada.
- **Fase 4:** A organização aproveita os benefícios de dados bem gerenciados e aprimora seus recursos analíticos.

Figura 8 Capacidade de banco de dados adquirido ou construído¹⁷

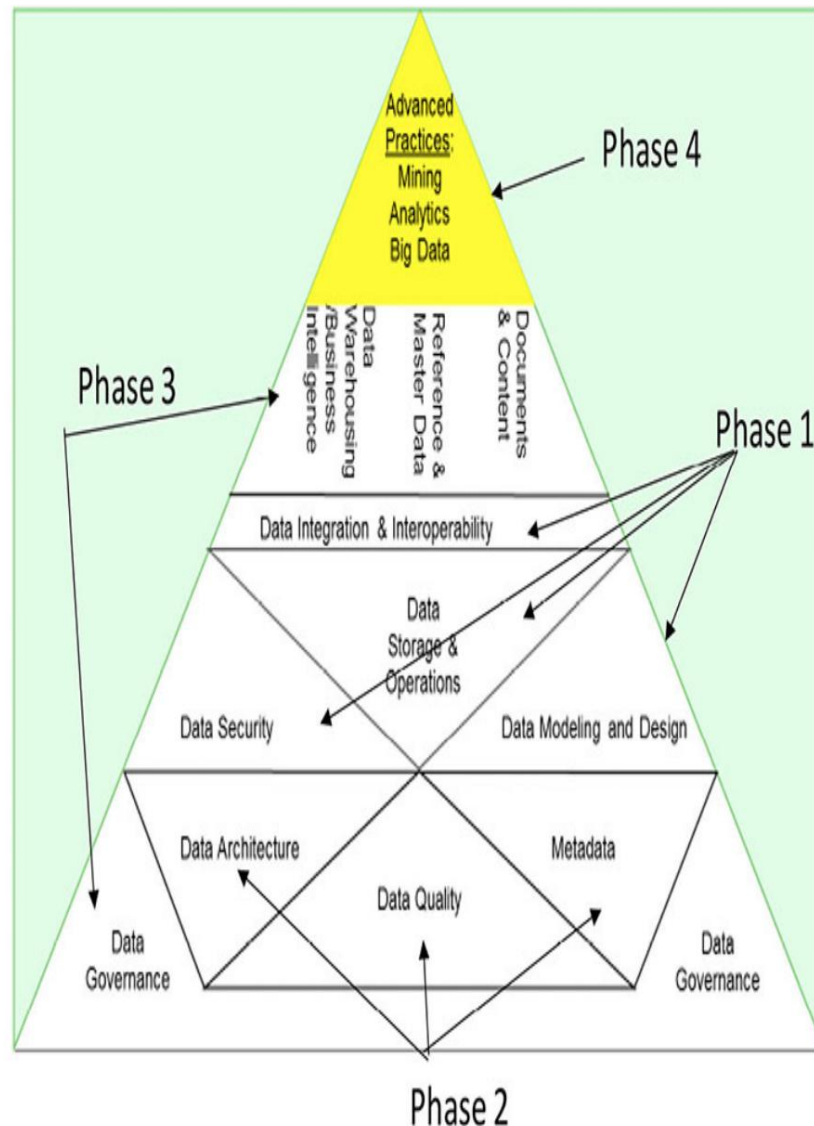


Figure 8 Purchased or Built Database Capability¹⁷

A pirâmide de Aiken extrai da Roda DAMA, mas também a informa ao mostrar a relação entre as Áreas de Conhecimento. Elas não são todas intercambiáveis; elas têm vários tipos de interdependências.

A estrutura da Pirâmide tem dois drivers. Primeiro, a ideia de construir sobre uma fundação, usando componentes que precisam estar nos lugares certos para apoiar uns aos outros. Segundo, a ideia um tanto contraditória de que estes podem ser colocados em ordem arbitrária.

3.5 Estrutura de gerenciamento de dados DAMA evoluída

A pirâmide de Aiken descreve como as organizações evoluem em direção a melhores práticas de gerenciamento de dados. Outra maneira de olhar para as DAMA Knowledge Areas é explorar as dependências entre elas. Desenvolvido por Sue Geuens, o framework na Figura 9 reconhece que as funções de Business Intelligence e Analytic têm dependências de todas as outras funções de gerenciamento de dados. Elas dependem diretamente de soluções de Master Data e data warehouse. Mas essas, por sua vez, são dependentes de sistemas de alimentação e aplicativos. Práticas de Reliable Data Quality, design de dados e interoperabilidade de dados são a base de sistemas e aplicativos confiáveis. Além disso, a governança de dados, que dentro deste modelo inclui Metadata Management, segurança de dados, Data Architecture e Reference Data Management, fornece uma base da qual todas as outras funções são dependentes.

Figura 9 DAMA Funcional

Dependências de área

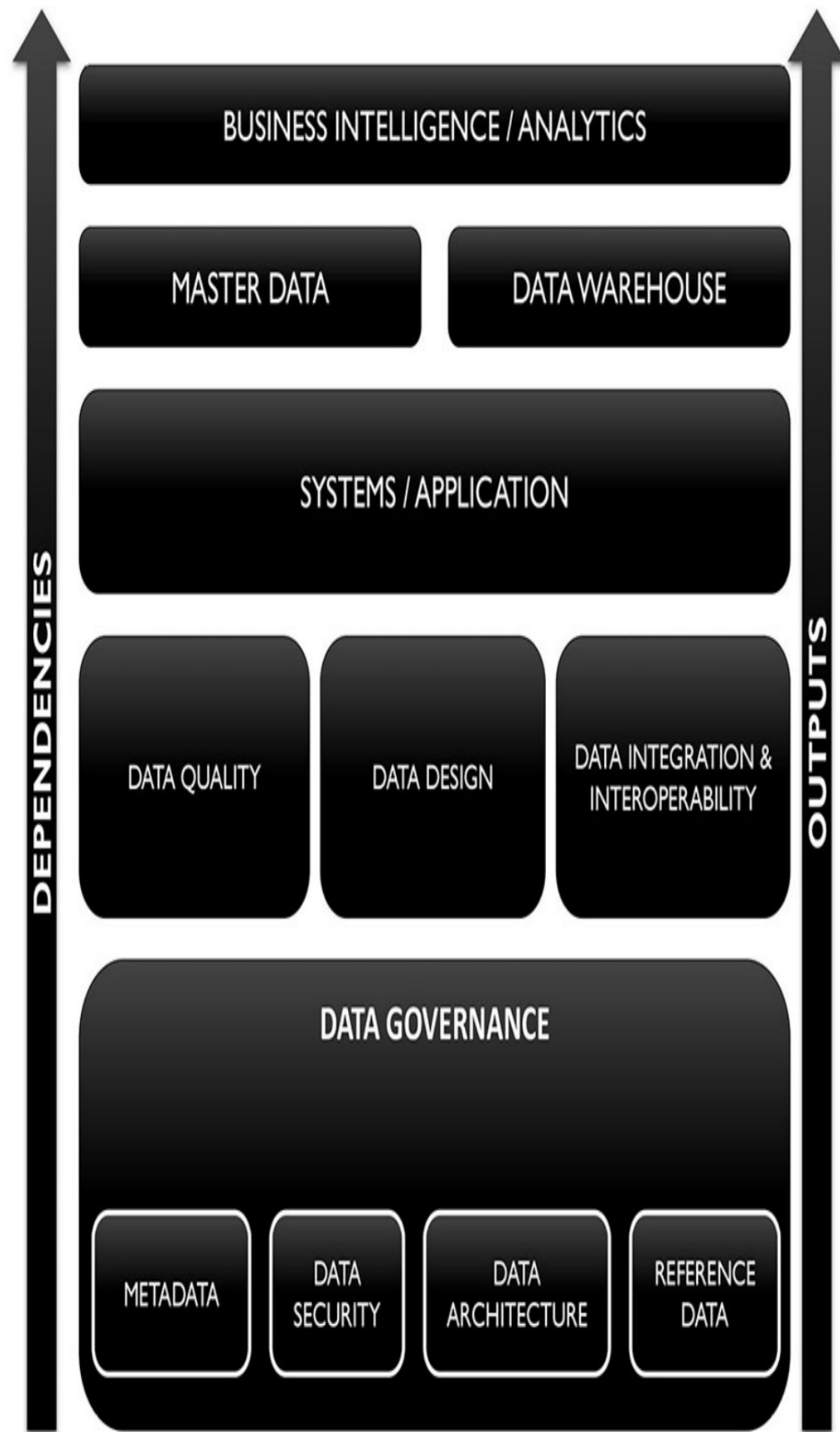


Figure 9 DAMA Functional Area Dependencies

Uma terceira alternativa para DAMA Wheel é retratada na Figura 10. Ela também se baseia em conceitos arquitetônicos para propor um conjunto de relacionamentos entre as DAMA Knowledge Areas. Ela fornece detalhes adicionais sobre o conteúdo de algumas Knowledge Areas para esclarecer esses relacionamentos.

A estrutura começa com o propósito orientador do gerenciamento de dados: Permitir que as organizações obtenham valor de seus ativos de dados como fazem com outros ativos. Derivar valor requer gerenciamento de ciclo de vida, então funções de gerenciamento de dados relacionadas ao ciclo de vida de dados são descritas no centro do diagrama. Isso inclui planejamento e design para dados confiáveis e de alta qualidade; estabelecimento de processos e funções por meio dos quais os dados podem ser habilitados para uso e também mantidos; e, finalmente, uso dos dados em vários tipos de análise e por meio desses processos, aumentando seu valor.

A seção de gerenciamento do ciclo de vida descreve o design de gerenciamento de dados e as funções operacionais (modelagem, arquitetura, armazenamento e operações, etc.) que são necessárias para dar suporte aos usos tradicionais de dados (Business Intelligence, gerenciamento de documentos e conteúdo). Ela também reconhece funções emergentes de gerenciamento de dados (armazenamento de Big Data) que dão suporte aos usos emergentes de dados (Data Science, análise preditiva, etc.). Em casos em que os dados são realmente gerenciados como um ativo, as organizações podem obter valor direto de seus dados vendendo-os para outras organizações (monetização de dados).

Figura 10 Dados DIREITOS
Função de Gestão
Estrutura

DATA MANAGEMENT FUNCTIONS

OVERSIGHT : Data Governance

Strategy

Data Valuation

Principles & Ethics

Policy

Stewardship

Culture Change

LIFECYCLE MANAGEMENT

PLAN & DESIGN

Architecture

Data
Modeling &
Design

USE & ENHANCE

Data Storage
& Operations

Data
Warehousing

Data
Integration &
Interoperability

Big Data
Storage

Master Data
Management

Reference
Data
Management

ENABLE & MAINTAIN

Business
Intelligence

Data Science

Master Data
Usage

Data
Monetization

Document
& Content
Management

Predictive
Analytics



Figure 10 DAMA Data Management Function Framework

Organizações que se concentram apenas em funções diretas do ciclo de vida não obterão tanto valor de seus dados quanto aquelas que dão suporte ao ciclo de vida dos dados por meio de atividades fundamentais e de supervisão. Atividades fundamentais, como gerenciamento de risco de dados, Metadados e gerenciamento de Qualidade de Dados, abrangem o ciclo de vida dos dados. Elas permitem melhores decisões de design e tornam os dados mais fáceis de usar. Se forem bem executadas, os dados são menos dispendiosos para manter, os consumidores de dados têm mais confiança neles e as oportunidades de usá-los se expandem.

Para dar suporte com sucesso à produção e uso de dados e garantir que as atividades fundamentais sejam executadas com disciplina, muitas organizações estabelecem supervisão na forma de governança de dados. Um programa de governança de dados permite que uma organização seja orientada por dados, colocando em prática a estratégia e apoiando princípios, políticas e práticas de administração que garantem que a organização reconheça e atue em oportunidades para obter valor de seus dados. Um programa de governança de dados também deve se envolver em atividades de gerenciamento de mudanças organizacionais para educar a organização e encorajar comportamentos que permitam usos estratégicos de dados. Assim, a necessidade de mudança de cultura abrange a amplitude das responsabilidades de governança de dados, especialmente à medida que uma organização amadurece suas práticas de gerenciamento de dados.

O DAMA Data Management Framework também pode ser descrito como um

evolução da Roda DAMA, com atividades principais cercadas por atividades de ciclo de vida e uso, contidas dentro das restrições de governança. (Veja [Figura 11.](#))



As principais atividades, incluindo gerenciamento de metadados, gerenciamento de qualidade de dados e definição de estrutura de dados (arquitetura), estão no centro da estrutura.

As atividades de gerenciamento do ciclo de vida podem ser definidas a partir de uma perspectiva de planejamento (gerenciamento de riscos, modelagem, design de dados, gerenciamento de dados de referência) e uma perspectiva de capacitação (gerenciamento de dados mestres, desenvolvimento de tecnologia de dados, integração e interoperabilidade de dados, armazenamento de dados e armazenamento e operações de dados).

Os usos emergem das atividades de gerenciamento do ciclo de vida: uso de dados mestres, gerenciamento de documentos e conteúdo, Business Intelligence, ciência de dados, análise preditiva, visualização de dados. Muitos deles criam mais dados ao aprimorar ou desenvolver insights sobre dados existentes. Oportunidades para monetização de dados podem ser identificadas como usos de dados.

As atividades de governança de dados fornecem supervisão e contenção, por meio de estratégia, princípios, política e administração. Elas permitem consistência por meio da classificação e avaliação de dados.

A intenção em apresentar diferentes representações visuais do DAMA Data Management Framework é fornecer perspectiva adicional e abrir discussão sobre como aplicar os conceitos apresentados no DMBOK. À medida que a importância do gerenciamento de dados cresce, tais frameworks se tornam ferramentas de comunicação úteis tanto dentro da comunidade de gerenciamento de dados quanto entre a comunidade de gerenciamento de dados e nossos stakeholders.

4. DAMA e o DMBOK

Embora o gerenciamento de dados apresente muitos desafios, poucos deles são novos. Desde pelo menos a década de 1980, as organizações reconheceram que o gerenciamento de dados é essencial para seu sucesso. À medida que nossa capacidade e desejo de criar e explorar dados aumentaram, também aumentou a necessidade de dados confiáveis

práticas de gerenciamento de dados.

Figura 11 Roda DAMA Evoluído

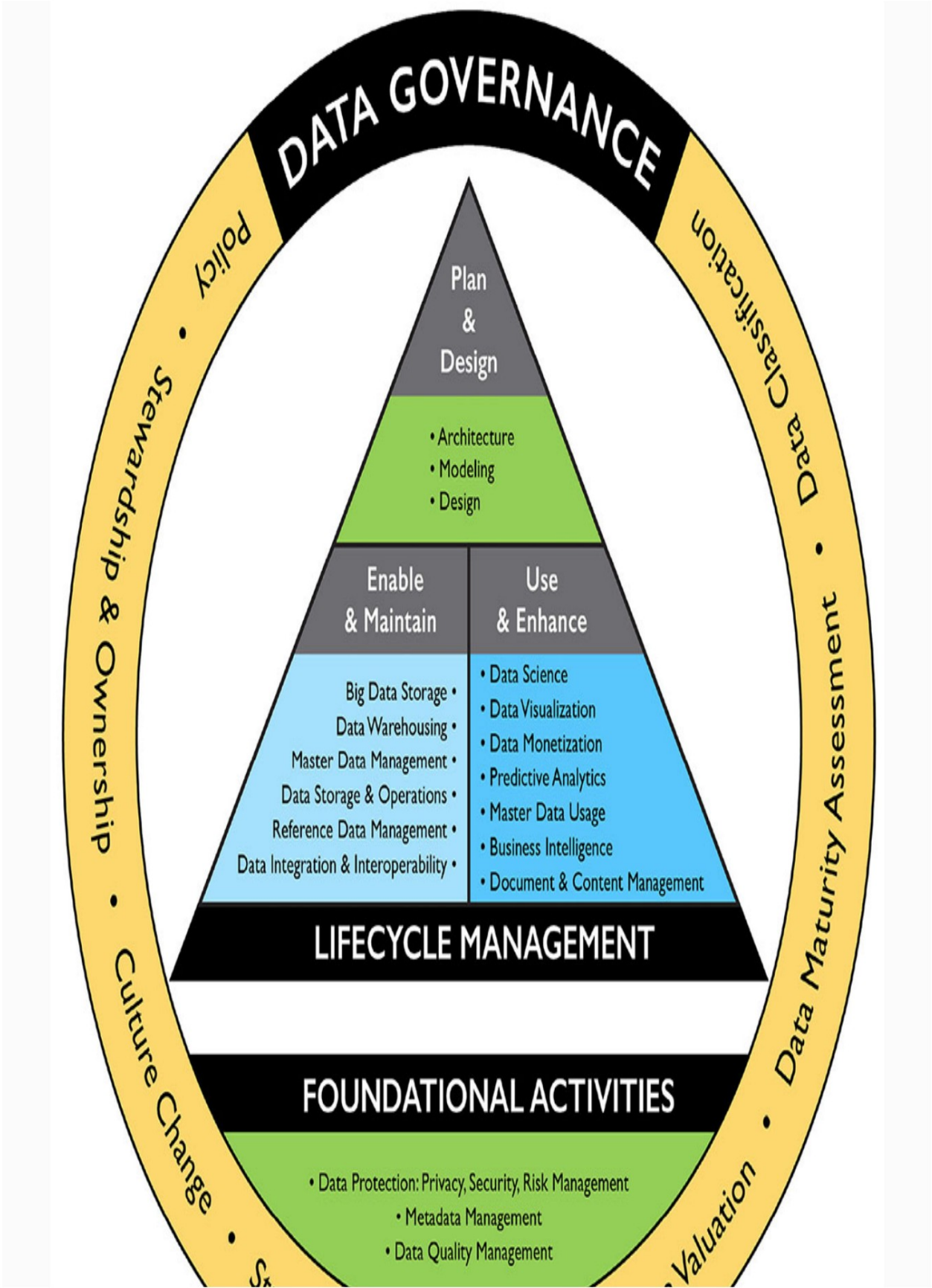




Figure 11 DAMA Wheel Evolved

A DAMA foi fundada para lidar com esses desafios. O DMBOK, um livro de referência acessível e confiável para profissionais de gerenciamento de dados, apoia a missão da DAMA ao:

- **Fornecer uma estrutura funcional** para a implementação de práticas de gerenciamento de dados corporativos; incluindo princípios orientadores, práticas amplamente adotadas, métodos e técnicas, funções, papéis, entregas e métricas.
- **Estabelecer um vocabulário comum** para conceitos de gerenciamento de dados e servir como base para melhores práticas para profissionais de gerenciamento de dados.
- **Servindo como guia de referência fundamental** para o CDMP (Certified Data Management Professional) e outros exames de certificação.

O DMBOK é estruturado em torno das onze Áreas de Conhecimento do DAMA-DMBOK Data Management Framework (também conhecido como DAMA Wheel, veja [a Figura 5](#)). Os capítulos 3 a 13 são focados nas Áreas de Conhecimento. Cada capítulo da Área de Conhecimento segue um modelo comum estrutura:

1. Introdução

- Motoristas de negócios

- Objetivos e Princípios
- Conceitos Essenciais

2. Atividades

3. Ferramentas

4. Técnicas 5.

Diretrizes de implementação

6. Relação com a Governança de Dados

7. Métricas

As Áreas de Conhecimento descrevem o escopo e o contexto de conjuntos de atividades de gerenciamento de dados. Incorporados nas Áreas de Conhecimento estão os objetivos e princípios fundamentais do gerenciamento de dados. Como os dados se movem horizontalmente dentro das organizações, as atividades das Áreas de Conhecimento se cruzam entre si e com outras funções organizacionais.

1. **A Governança de Dados** fornece direção e supervisão para o gerenciamento de dados, estabelecendo um sistema de direitos de decisão sobre dados que leva em conta as necessidades da empresa. ([Capítulo 3](#))
2. **A Arquitetura de Dados** define o modelo para gerenciar ativos de dados alinhando-se com a estratégia organizacional para estabelecer requisitos de dados estratégicos e designs para atender a esses requisitos. ([Capítulo 4](#))
3. **Modelagem e design de dados** é o processo de descobrir, analisar, representar e comunicar requisitos de dados em um formato preciso chamado *modelo de dados*. ([Capítulo 5](#))
4. **O armazenamento e as operações de dados** incluem o design, implementação e suporte de dados armazenados para maximizar seu valor. As operações fornecem suporte em todo o processo de dados

ciclo de vida desde o planejamento até o descarte dos dados. [\(Capítulo 6\)](#)

5. **A Segurança de Dados** garante a privacidade dos dados e a confidencialidade seja mantida, que os dados não sejam violados e que os dados sejam acessados adequadamente. [\(Capítulo 7\)](#)
6. **A integração e a interoperabilidade de dados** incluem processos relacionados à movimentação e consolidação de dados dentro e entre armazenamentos de dados, aplicativos e organizações. [\(Capítulo 8\)](#)
7. **O gerenciamento de documentos e conteúdo** inclui atividades de planejamento, implementação e controle usadas para gerenciar o ciclo de vida de dados e informações encontrados em uma variedade de mídias não estruturadas, especialmente documentos necessários para dar suporte a requisitos de conformidade legal e regulatória. [\(Capítulo 9\)](#)
8. **Os Dados Mestres e de Referência** incluem dados contínuos reconciliação e manutenção de dados essenciais compartilhados para permitir o uso consistente em todos os sistemas da versão mais precisa, oportuna e relevante da verdade sobre entidades comerciais essenciais. [\(Capítulo 10\)](#)
9. **Data Warehousing e Business Intelligence** incluem os processos de planejamento, implementação e controle para gerenciar dados de suporte à decisão e permitir que trabalhadores do conhecimento obtenham valor dos dados por meio de análises e relatórios. [\(Capítulo 11\)](#)
10. **Metadados** incluem atividades de planejamento, implementação e controle para permitir acesso a Metadados integrados de alta qualidade, incluindo definições, modelos, fluxos de dados e outras informações críticas para a compreensão dos dados e dos sistemas por meio dos quais eles são criados, mantidos e acessados. [\(Capítulo 12\)](#)
11. **A qualidade dos dados** inclui o planejamento e a implementação de técnicas de gestão da qualidade para medir, avaliar e melhorar a adequação dos dados para uso dentro de uma organização. [\(Capítulo 13\)](#)

Além dos capítulos sobre as Áreas de Conhecimento, o DAMA-DMBOK contém capítulos sobre os seguintes tópicos:

- **A Ética de Manipulação de Dados** descreve o papel central que a ética de dados desempenha na tomada de decisões informadas e socialmente responsáveis sobre dados e seus usos. A conscientização sobre a ética da coleta, análise e uso de dados deve orientar todos os profissionais de gerenciamento de dados. [\(Capítulo 2\)](#)
- **Big Data e Data Science** descrevem as tecnologias e processos de negócios que surgem à medida que nossa capacidade de coletar e analisar conjuntos de dados grandes e diversos aumenta. [\(Capítulo 14\)](#)
- **A Avaliação de Maturidade em Gerenciamento de Dados** descreve uma abordagem para avaliar e melhorar as capacidades de gerenciamento de dados de uma [organização](#). [\(Capítulo 15\)](#)
- **A Organização de Gerenciamento de Dados e Expectativas de Função** fornecem as melhores práticas e considerações para organizar equipes de gerenciamento de dados e permitir práticas de gerenciamento de dados [bem-sucedidas](#). [\(Capítulo 16\)](#)
- **Gerenciamento de Dados e Gerenciamento de Mudanças Organizacionais** descreve como planejar e passar com sucesso pelas mudanças culturais que são necessárias para incorporar práticas eficazes de gerenciamento de dados dentro de uma [organização](#). [\(Capítulo 17\)](#)

Como uma organização específica gerencia seus dados depende de seus objetivos, tamanho, recursos e complexidade, bem como de sua percepção de como os dados dão suporte à sua estratégia geral. A maioria das empresas não realiza todas as atividades descritas em cada Área de Conhecimento. No entanto, entender o contexto mais amplo do gerenciamento de dados permitirá que as organizações tomem melhores decisões sobre onde focar enquanto trabalham para melhorar as práticas dentro e entre essas funções relacionadas.

5. Trabalhos Citados / Recomendados