



Enap

Governança de Dados

Módulo

2

Princípios, importância e desafios do Gerenciamento de Dados



Fundação Escola Nacional de Administração Pública

Presidente

Diogo Godinho Ramos Costa

Diretor de Educação Continuada

Paulo Marques

Coordenador-Geral de Educação a Distância

Carlos Eduardo dos Santos

Conteudista/s

Wellington Luiz Barbosa (conteudista, 2019)

Roberto Shayer Lyra (conteudista, 2019)

Curso produzido em Brasília 2019.



Enap, 2019

Enap Escola Nacional de Administração Pública

Diretoria de Educação Continuada

SAIS - Área 2-A - 70610-900 — Brasília, DF



Sumário

1. A Importância do Gerenciamento de Dados	5
1.1. Os Dados Estão em Todo Lugar	5
1.2. Dados Como um Ativo	6
1.3. Gerenciamento de Dados x Gerenciamento de Tecnologia.....	6
1.4. Atividades de Gerenciamento de Dados	7
1.5. Áreas de Conhecimento em Gerenciamento de Dados.....	8
2. Desafios no Gerenciamento de Dados	13
2.1 Os dados diferem de outros ativos.....	14
2.2 Dados representam risco.....	15
2.3 Dados de baixa qualidade custam dinheiro.....	18
2.4 A avaliação de dados não é padronizada.....	21
2.5 Gerenciamento de dados significa gerenciar o ciclo de vida dos dados.....	22
2.6 Diferentes tipos de dados têm diferentes requisitos de ciclo de vida	24
2.7 Os metadados devem ser gerenciados como parte do ciclo de vida dos dados...26	
2.8 O gerenciamento de dados geralmente é confundido com o gerenciamento de tecnologia da informação.....	26
2.9 O que o gerenciamento de dados requer.....	27
3. Princípios de gerenciamento de dados do DAMA (DATA MANAGEMENT ASSOCIATION)	30
3.1. Os dados são valiosos	31
3.2. Requisitos de gerenciamento de dados são requisitos de negócios	32
3.3. Gerenciamento de dados é gerenciamento do ciclo de vida	32
3.4. O gerenciamento de dados depende de diversas habilidades	33
3.5. Princípios e maturidade em gerenciamento de dados.....	33
3.6. Reavaliar a maturidade.....	38
3.7. Ferramentas de avaliação de maturidade	39
3.8. Técnicas de avaliação de maturidade	39
3.9. Uso da estrutura DAMA-DMBOK.....	41
4. Ética na Gestão de Dados	41
4.1. Ética e gerenciamento de dados	42
4.2. Princípios éticos subjacentes à regulamentação da privacidade.....	43



4.3. Ética e vantagem competitiva.....47

4.4. Estabelecer uma cultura de manipulação ética de dados48



Módulo

2

Princípios, importância e desafios do Gerenciamento de Dados

1. A Importância do Gerenciamento de Dados

Mesmo antes do surgimento da Tecnologia da Informação, a informação e o conhecimento foram fundamentais para a vantagem competitiva.

As organizações que possuem informações confiáveis e de alta qualidade sobre seus usuários, produtos, serviços e operações podem tomar melhores decisões do que aquelas sem dados (ou com dados não confiáveis). Porém, produzir dados de alta qualidade e gerenciá-los de maneiras que permitam os seus usos com eficiência não é um processo simples.

Para aprimoramento dos recursos de gerenciamento de dados de qualquer organização, faz-se necessário considerar os seguintes aspectos:

DESTAQUE

- Os dados estão em todo lugar.
- Dados como um ativo - valor.
- Gerenciamento de dados versus gerenciamento de tecnologia.
- Atividades de gerenciamento de dados.
- Áreas de conhecimento em gerenciamento de dados.

1.1. Os Dados Estão em Todo Lugar

As organizações sempre precisaram gerenciar seus dados, mas os avanços na tecnologia, considerando a onipresença dos dados, expandiram o escopo dessa necessidade de gerenciamento.

Os dados são difundidos nas organizações. Quase todos os processos de negócios - desde a criação de clientes, transações de compras, contato com clientes para feedback e serviços - usam dados como entrada e produzem dados como saída. A maioria desses dados está em formato eletrônico, o que significa maleabilidade: podem ser armazenados em grandes quantidades, manipulados, integrados e agregados para diferentes usos, incluindo inteligência comercial e análise preditiva. Eles também fornecem evidências de conformidade de uma organização (ou falta de conformidade) com leis e regulamentos.

As mudanças técnicas permitiram que as organizações usassem os 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 extrair dados para obter significado intensificaram a necessidade de gerenciar dados de maneira eficaz.

1.2. Dados Como um Ativo

Um ativo é um recurso econômico, que pode ser de propriedade ou controlado e que detém ou produz valor. Os ativos são frequentemente vistos como propriedade, mas com a forte implicação de que podem ser convertidos em dinheiro. Os dados são amplamente reconhecidos como um ativo da organização, embora muitas organizações ainda lutem para gerenciar os dados como tal. Um exemplo disso é o fato de que os dados ainda não são contabilizados no balanço da maioria das organizações.

Se questionados, muitos usuários diriam que os dados de sua organização são um ativo valioso.

Os dados não são apenas necessários para as operações de negócio, mas também podem fornecer informações sobre usuários, produtos e serviços. No entanto, pesquisas mostram que poucas organizações tratam seus dados como um ativo e para muitas eles podem até ser um passivo. A falha no gerenciamento de dados é semelhante à falha no gerenciamento de capital, isso resulta em desperdício e em oportunidades perdidas. Além do mais, dados mal gerenciados apresentam riscos éticos e de segurança.

Mesmo usuários que reconhecem os dados como um ativo podem não ser capazes de descrever exatamente o que isso significa, pois os dados diferem de outros ativos de maneiras importantes. No entanto, o principal fator para o gerenciamento de dados é permitir que as organizações obtenham valor com seus dados, assim como o gerenciamento eficaz de ativos financeiros e físicos permite que as organizações obtenham valor com esses ativos. A obtenção de valor dos dados não ocorre no vácuo ou por acidente, pois além de gerenciamento, ela requer compromisso organizacional e liderança.

1.3. Gerenciamento de Dados x Gerenciamento de Tecnologia

É necessário diferenciar Gerenciamento de Dados e Gerenciamento de Tecnologia!

Gerenciamento de Dados é o desenvolvimento, execução e supervisão de planos, políticas, programas e práticas que entregam, controlam, protegem e aprimoram o valor dos ativos de dados e informações durante todo o seu ciclo de vida.

Você pode se questionar: “Isso não é o que nosso departamento de tecnologia da informação tem feito?” A resposta é: infelizmente, não.

A TI geralmente não se concentra nos dados. O foco é a tecnologia, os processos, as pessoas que desenvolvem aplicações e as ferramentas usadas para fazer isso. Historicamente, a TI não se concentra nos dados criados ou armazenados nas aplicações que gerencia. Não há direcionamento estratégico para isso.



Embora o gerenciamento de dados seja altamente dependente da tecnologia e cruze com o gerenciamento de TI, as atividades são separadas, independentes de processos e ferramentas utilizadas.

Com isso, o que o Gerenciamento de Dados realmente envolve? O que significa gerenciar dados de maneira eficaz?

O Gerenciamento de Dados envolve o planejamento e a coordenação de recursos e atividades para atender aos objetivos organizacionais. As atividades em si variam desde as altamente técnicas, como garantir que grandes bancos de dados sejam acessíveis com desempenho e segurança, até as altamente estratégicas, como propor a interoperabilidade de dados entre entes de governo de maneira eficiente e segura. Essas atividades de gerenciamento devem se esforçar para disponibilizar dados confiáveis e de alta qualidade para a organização, garantindo que sejam acessíveis aos usuários autorizados e protegidos contra uso indevido.

1.4. Atividades de Gerenciamento de Dados

Existe uma gama de atividades de Gerenciamento de Dados. Elas podem ser entendidas em três grupos:

- As que focam na governança para garantir que a organização tome decisões consistentes e alinhadas (atividades de governança).
- As que se concentram no gerenciamento do ciclo de vida dos dados, desde a obtenção dos dados até a sua eliminação (atividades de ciclo de vida).
- As fundamentais, que focam em permitir o gerenciamento, manutenção e uso de dados ao longo do tempo.

As atividades de governança ajudam a controlar o desenvolvimento de dados e reduzem os riscos associados ao seu uso, ao mesmo tempo em que permitem que uma organização aproveite-os estrategicamente.

Essas atividades estabelecem um sistema de direitos e responsabilidades de decisão por dados para que uma organização possa tomar decisões consistentes nas verticais dos negócios.

Atenção, as atividades de governança incluem as seguintes ações:

- Definir a estratégia de dados.
- Definir a política de governança.
- Definir o valor dos dados para a organização.
- Preparar a organização para obter mais valor de seus dados.



- Avaliar práticas de gerenciamento de dados.
- Evoluir a mentalidade da organização em torno dos dados.

As atividades do ciclo de vida se concentram no planejamento e design dos dados, permitindo seu uso, garantindo que eles sejam efetivamente mantidos e realmente utilizados. Geralmente, o uso de dados resulta em aprimoramentos e inovações que têm seus próprios requisitos de ciclo de vida.

Esteja atento às atividades do ciclo de vida que consideram:

- Arquitetura de dados.
- Modelagem de dados.
- Construção e gerenciamento de data *warehouses e marts*.
- Integração de dados para uso por analistas de inteligência de negócios e cientistas de dados.
- Gerenciamento do ciclo de vida de dados compartilhados altamente críticos, como dados de referência e dados mestres.

Atividades básicas são necessárias para o gerenciamento consistente dos dados ao longo do tempo.

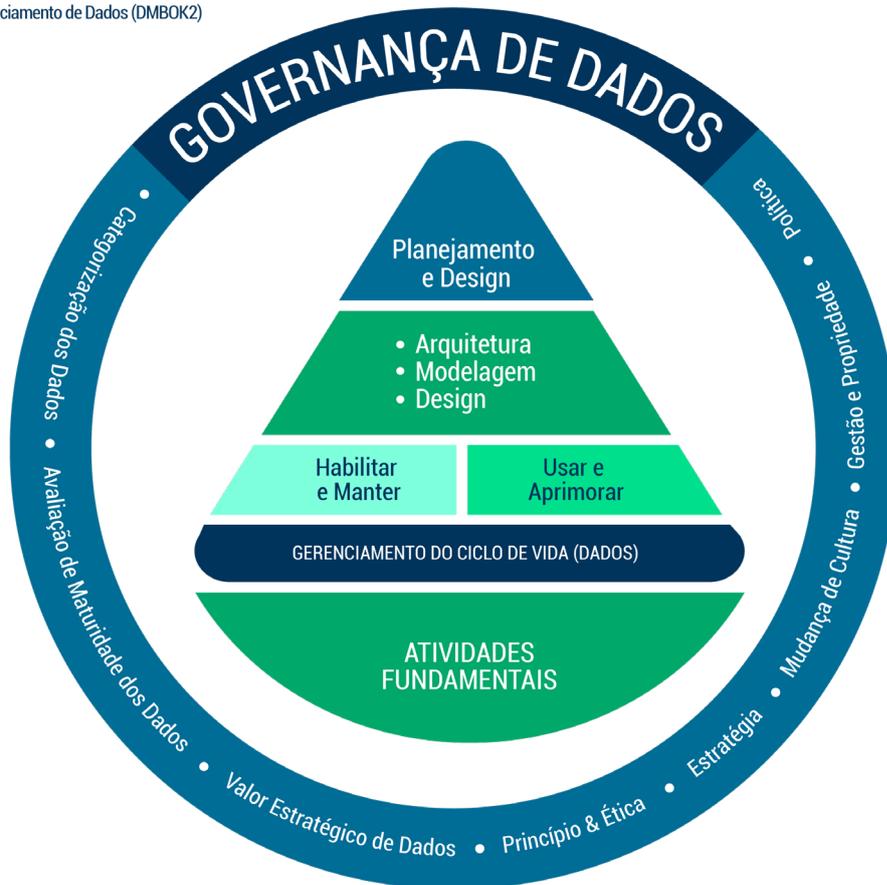
Abaixo, ressaltamos as ações que são vitais a todo o ciclo de vida dos dados:

- Garantir a proteção dos dados.
- Gerenciar metadados.
- Gerenciar a qualidade dos dados.

As atividades fundamentais devem ser consideradas como parte do planejamento e design e devem ser realizadas operacionalmente. Essas atividades também são essenciais para o sucesso das estruturas de governança.

1.5. Áreas de Conhecimento em Gerenciamento de Dados

Atividades de Gerenciamento de Dados são realizadas por profissionais que atuam em diferentes áreas de conhecimento, logo necessitam das habilidades e dos conhecimentos apresentados na imagem a seguir.



Habilitar e Manter

- Armazenamento de Big Data;
- Data Warehousing;
- Gestão de Dados Mestres;
- Armazenamento de Dados & Operações;
- Gerenciamento de Dados de Referência;
- Integração de Dados & Interoperabilidade.

Usar e Aprimorar

- Ciência de Dados;
- Visualização de Dados;
- Valor Estratégico do Dado;



- Análise Preditiva;
- Uso de Dados Mestres;
- Business Intelligence;
- Gerenciamento de Documentos e Conteúdo.

Atividades Fundamentais

- Privacidade e Proteção de Dados;
- Segurança da Informação;
- Gestão de Risco;
- Gestão de Metadata;
- Gestão da Qualidade de Dados.

O guia criado pela DAMA Internacional, organização que visa promover a compreensão, o desenvolvimento e a prática da gestão de dados e informações como ativos chave de uma organização, propõe o conjunto de dez funções para a Governança de Dados. Acesse, para conhecer cada uma dessas funções, as abas a seguir:

- **Governança de Dados**
A Governança de Dados fornece orientação e supervisão para atividades e funções de gerenciamento de dados, estabelecendo um sistema de direitos e responsabilidades de decisão por dados. Esses direitos e responsabilidades devem responder às necessidades da empresa como um todo.
- **Arquitetura de Dados**
A Arquitetura de Dados define o modelo para o gerenciamento de ativos de dados, alinhando-se à estratégia organizacional e estabelecendo projetos para atender aos requisitos de dados estratégicos.
- **Modelagem e Design de Dados**
A Modelagem e Design de Dados é o processo de descoberta, análise, representação e comunicação de requisitos de dados de uma forma precisa chamada Modelo de Dados.
- **Armazenamento e Operações de Dados**
O Armazenamento e Operações de Dados inclui o design, implementação e suporte de dados armazenados para maximizar seu valor. As operações fornecem suporte durante todo o ciclo de vida dos dados, do planejamento à disposição dos dados.



- **Segurança de Dados**
A Segurança de Dados garante que a privacidade e a confidencialidade dos dados sejam mantidas, que os dados não sejam violados e que os dados sejam acessados adequadamente.
- **Integração e Interoperabilidade de Dados**
A Integração e Interoperabilidade de Dados inclui processos relacionados à movimentação e consolidação de dados dentro e entre repositórios de dados, aplicativos e organizações.
- **Gerenciamento de Documentos e Conteúdo**
O Gerenciamento de Documentos e Conteúdo inclui atividades de planejamento, implementação e controle para gerenciar o ciclo de vida de dados e informações encontrados em uma variedade de mídias não estruturadas, especialmente os documentos necessários para suportar os requisitos de conformidade legal e regulamentar.
- **Gerenciamento de Referência e Dados Mestre**
O Gerenciamento de Referência e Dados Mestre inclui reconciliação e manutenção contínuas dos principais dados compartilhados críticos para permitir o uso consistente entre os sistemas da versão mais precisa, oportuna e relevante da verdade sobre entidades essenciais de negócios.
- **Data Warehousing e Business Intelligence**
O *Data Warehousing* e o *Business Intelligence* incluem os processos de planejamento, implementação e controle para gerenciar os dados de suporte à decisão e permitir que os profissionais do conhecimento obtenham valor dos dados por meio de análise e relatório.
- **Gerenciamento de Metadados**
O Gerenciamento de Metadados inclui atividades de planejamento, implementação e controle para permitir o acesso a metadados integrados de alta qualidade, incluindo definições, modelos, fluxos de dados e outras informações críticas para entender os dados e os sistemas através dos quais são criados, mantidos e acessados.
- **Gerenciamento da Qualidade dos Dados**
O Gerenciamento da Qualidade dos Dados inclui o planejamento e a implementação de técnicas de gerenciamento da qualidade para medir, avaliar e melhorar a adequação dos dados para uso em uma organização.

Essas áreas de conhecimento representam atividades no núcleo do Gerenciamento de Dados. Qualquer organização que busque obter valor com seus dados deverá se envolver neles.

É bom lembrar-se que os dados também estão evoluindo.



Alterações em nossa capacidade de criar e usar dados significam que outros conceitos também podem ser considerados "áreas de conhecimento" de Gerenciamento de Dados, tais como Ética de Dados, Ciência de Dados, Gerenciamento de *Big Data* e Tecnologias Emergentes.

IMPORTANTE

A estrutura DAMA define as áreas de conhecimento de Gerenciamento de Dados. Ela centraliza a disciplina Governança de Dados das atividades de Gerenciamento de Dados, já que a governança é necessária para a consistência e o equilíbrio entre as funções. As outras áreas de conhecimento (Arquitetura de Dados, Modelagem de Dados, etc.) são balanceadas em torno do centro. Todas são partes necessárias de uma função madura de Gerenciamento de Dados, mas podem ser implementadas em momentos diferentes, dependendo dos requisitos da organização.

Os profissionais de Gerenciamento de Dados que trabalham nessas áreas de conhecimento devem ter como objetivo:

DESTAQUE

- Compreender e dar suporte às necessidades de informações da organização e de suas partes interessadas, incluindo clientes, funcionários e parceiros de negócios.
- Capturar, armazenar e garantir a integridade e a qualidade dos dados para permitir seu uso pela organização.
- Garantir a segurança, a privacidade e a confidencialidade dos dados, impedindo acesso, manipulação ou uso inadequado.

SAIBA MAIS

DataOps é uma metodologia automatizada, orientada a processos, usada por equipes analíticas e de dados, para melhorar a qualidade e reduzir o tempo de ciclo da análise de dados. Embora o DataOps tenha começado como um conjunto de práticas recomendadas, ele agora amadureceu e se tornou uma abordagem nova e independente da análise de dados. O DataOps se aplica a todo o ciclo de vida dos dados, da preparação dos dados aos relatórios, e reconhece a natureza interconectada da equipe de análise de dados e das operações de tecnologia da informação.

O DataOps incorpora a metodologia Agile para reduzir o tempo de ciclo do desenvolvimento de análises, alinhado às metas de negócios.



O DevOps se concentra na entrega contínua, aproveitando os recursos de TI sob demanda e automatizando o teste e a implantação de análises. Essa fusão do desenvolvimento de software e das operações de TI melhorou a velocidade, a qualidade, a previsibilidade e a escala da engenharia e implantação de software. Empréstando métodos do DevOps, o DataOps procura trazer esses mesmos aprimoramentos para a análise de dados.

O DataOps utiliza o controle estatístico do processo (CEP) para monitorar e controlar o pipeline de análise de dados. Com o SPC instalado, os dados que fluem através de um sistema operacional são constantemente monitorados e verificados como funcionando. Se ocorrer uma anomalia, a equipe de análise de dados poderá ser notificada por meio de um alerta automatizado.

O DataOps não está vinculado a uma tecnologia, arquitetura, ferramenta, linguagem ou estrutura específica. As ferramentas que suportam o DataOps promovem colaboração, orquestração, qualidade, segurança, acesso e facilidade de uso.

2. Desafios no Gerenciamento de Dados

O gerenciamento de dados é uma disciplina focada em aprimorar o valor dos dados e informações. O gerenciamento eficaz de dados beneficia a organização e sociedade de várias maneiras. No caso da Administração Pública, é disciplina fundamental para alavancar o planejamento eficiente de Políticas Públicas e de disponibilização de serviços digitais ao cidadão por meio da interoperabilidade de dados. Existem dez funções principais de gerenciamento de dados que envolvem uma variedade de indivíduos com habilidades diferentes. O gerenciamento de dados é mais uma função que, quando implementado de forma eficiente, promove inúmeros resultados positivos.

Em termos gerais, podemos definir Gerenciamento de dados como o desenvolvimento, execução e supervisão de planos, políticas, programas e práticas que entregam, controlam, protegem e aprimoram o valor dos ativos de dados e informações ao longo de seus ciclos de vida.

Os dados são uma necessidade estratégica e operacional. Para gerenciar qualquer ativo, é necessário trabalhar para obter valor. Neste contexto, abordaremos os seguintes tópicos:

- Os dados diferem de outros ativos.
- Dados representam risco.
- Dados de baixa qualidade custam tempo e dinheiro.
- A avaliação de dados não é padronizada.



- Gerenciamento de dados significa gerenciar o ciclo de vida dos dados.
- Diferentes tipos de dados têm diferentes requisitos de ciclo de vida.
- Os metadados devem ser gerenciados como parte do ciclo de vida dos dados.
- O gerenciamento de dados geralmente é confundido com o gerenciamento de tecnologia da informação.
- O que o gerenciamento de dados requer.

2.1 Os dados diferem de outros ativos

Os dados têm características únicas que os diferenciam de outros ativos: os ativos físicos, por exemplo, podem ser apontados, tocados e movimentados. Os ativos financeiros, por sua vez, são contabilizados no balanço patrimonial. Mas os dados são diferentes, porque não são tangíveis. No entanto, são duráveis e não se desgastam; são fáceis de copiar e transportar, mas não de serem reproduzidos se perdidos ou destruídos. Por não serem consumidos quando usados, podem até ser roubados sem desaparecer. Os dados são dinâmicos e podem ser usados para várias finalidades. 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.

Essas diferenças tornam desafiador simplesmente manter o controle dos dados e muito mais ainda atribuir um valor monetário a eles. Sem esse valor monetário, pode ser difícil medir como os dados contribuem para o sucesso organizacional.

Essas diferenças também levantam outros problemas que afetam o gerenciamento de dados, como:

- Inventariar a quantidade de dados que uma organização possui.
- Definir propriedade e responsabilidade dos dados.
- Proteger dados contra o uso indevido.
- Gerenciar riscos associados aos dados.
- Definir e aplicar padrões de qualidade para dados.



2.2 Dados representam risco

Os dados não representam apenas valor e oportunidade, também retratam riscos. Dados imprecisos, incompletos ou desatualizados indicam riscos, porque suas informações não estão corretas. Além disso, os dados apresentam outras ameaças, incluindo:

- **Uso indevido:** se os consumidores de dados não tiverem informações suficientes e corretas (metadados) sobre os dados que eles usam, existe o risco de serem mal utilizados e/ou compreendidos.
- **Confiabilidade (*Unreliability*):** Se a qualidade e a confiabilidade dos dados não forem estabelecidas por meio de normas e métricas, então existe risco de que decisões sejam baseadas em dados não confiáveis.
- **Uso inadequado:** se os dados não estiverem protegidos, existe o risco de que eles sejam utilizados por pessoas não autorizadas para fins não autorizados.

O fato de os dados poderem ser facilmente copiados e replicados significa que eles podem ser violados sem serem "retirados" de seus legítimos proprietários. Além disso, como os dados representam pessoas, produtos e dinheiro, os legisladores e reguladores reconheceram os possíveis usos e abusos das informações e criaram leis destinadas a mitigar riscos óbvios. Por exemplo:

- A Sarbanes-Oxley nos EUA concentra-se nos controles sobre a precisão e a validade dos dados das transações financeiras, da transação ao balanço.
- O Solvência II na União Europeia concentra-se na linhagem de dados e na qualidade dos dados subjacentes aos modelos de risco e adequação de capital no setor de seguros.
- Em todo o mundo, os regulamentos de privacidade de dados descrevem obrigações em relação ao tratamento de dados de identificação pessoal (tais como: nome, endereço, afiliação religiosa ou orientação sexual) e privacidade (acesso ou restrição a essas informações):
 - ✓ Lei de Portabilidade e Responsabilidade do Seguro de Saúde (HIPPA), nos EUA.
 - ✓ Lei de Proteção de Informações Pessoais e Documentos Eletrônicos (PIPEDA), no Canadá.
 - ✓ O Regulamento Geral de Proteção de Dados (RGPD), na União Europeia.
 - ✓ Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709/2018, é a legislação brasileira que regula as atividades de tratamento de dados pessoais e que também altera os artigos 7º e 16 do Marco Civil da Internet, Lei nº 12.965/2014.



Destacamos a relevância da publicação do Decreto nº 10.046, de 09 de outubro de 2019 que estabeleceu as regras e as diretrizes de compartilhamento de dados no âmbito da Administração Pública Federal - APF. Esse decreto, instituiu também o Cadastro Base do Cidadão (CBC) e o Comitê Central de Governança de Dados (CCGD). Sendo este último, o fórum legítimo para propor políticas e normas sobre Governança de Dados, mediar e dirimir conflitos sobre compartilhamento de dados entre órgãos e entidades da administração pública federal e demais Poderes da União.

Acesse as abas a seguir e conheça sobre cada um desses dispositivos do decreto:

- **Regras e diretrizes do compartilhamento de dados na APF**

As regras contidas no decreto não são aplicáveis ao compartilhamento de dados com os conselhos de fiscalização de profissões regulamentadas e com o setor privado. Além disso, as regras do texto excluem os dados protegidos por sigilo fiscal que estão sob gestão da Secretaria Especial da Receita Federal do Brasil do Ministério da Economia.

A política de compartilhamento de dados terá como finalidades:

- ✓ simplificar a oferta de serviços públicos;
- ✓ orientar e otimizar a formulação, a implementação, a avaliação e o monitoramento de políticas públicas;
- ✓ possibilitar a análise das condições de acesso e manutenção de benefícios sociais e fiscais;
- ✓ promover a melhoria da qualidade e da fidedignidade dos dados custodiados pela administração pública federal; e,
- ✓ aumentar a qualidade e a eficiência das operações internas da administração pública federal.

O Decreto traz alguns conceitos de atributos dos dados, como o de atributos biográficos, que são os dados de pessoa natural relativos aos fatos da sua vida, tais como nome civil ou social, data de nascimento, filiação, naturalidade, nacionalidade, sexo, estado civil, grupo familiar, endereço e vínculos empregatícios; o de atributos biométricos que são as características biológicas e comportamentais mensuráveis da pessoa natural que podem ser coletadas para reconhecimento automatizado, tais como a palma da mão, as digitais dos dedos, a retina ou a íris dos olhos, o formato da face, a voz e a maneira de andar; e o de atributos genéticos, que são as características hereditárias da pessoa natural, obtidas pela análise de ácidos nucleicos (DNA) ou por outras análises científicas. Além disso, traz o conceito de dados cadastrais, que são informações identificadoras perante os cadastros de órgãos públicos, tais como CPF, CNPJ, PIS, NIS, número de título de eleitor etc.

O compartilhamento de dados pelos entes da administração pública federal deverá seguir algumas diretrizes.



No caso de informação de Estado, este compartilhamento deverá ser o mais amplo possível, observadas as restrições legais, os requisitos de segurança da informação e comunicações e o disposto na Lei nº 13.709, de 2018 (Lei Geral de Proteção de Dados Pessoais – LGPD).

E no caso dados sigilosos, será necessária a assunção (aceitação) pelo recebedor de dados, dos deveres de sigilo e auditabilidade impostos ao custodiante dos dados. O Decreto assegura que nas hipóteses de tratamento de dados pessoais serão observados o direito à preservação da intimidade e da privacidade da pessoa natural, a proteção dos dados e as normas e os procedimentos previstos na legislação e que a coleta, o tratamento e o compartilhamento de dados por cada órgão do Governo Federal serão realizados nos termos do disposto na LGPD.

Está estabelecido também que o compartilhamento de dados entre os órgãos e as entidades do Governo Federal é categorizado em três níveis, de acordo com sua confidencialidade, que serão definidas pelo gestor dos dados:

- ✓ compartilhamento amplo – quando se tratar de dados públicos que não estão sujeitos a nenhuma restrição de acesso, cuja divulgação deve ser pública e garantida a qualquer interessado, na forma da legislação;
 - ✓ compartilhamento restrito – quando se tratar de dados protegidos por sigilo, nos termos da legislação, com concessão de acesso a todos os órgãos e entidades do Governo Federal para a execução de políticas públicas, cujo mecanismo de compartilhamento e regras sejam simplificados e estabelecidos pelo Comitê Central de Governança de Dados;
 - ✓ compartilhamento específico – quando se tratar de dados protegidos por sigilo, nos termos da legislação, com concessão de acesso a órgãos e entidades específicos, nas hipóteses e para os fins previstos em lei, cujo compartilhamento e regras sejam definidos pelo gestor de dados.
- **Cadastro Base do Cidadão**

O Cadastro Base do Cidadão constitui um meio unificado de identificação do cidadão para a prestação de serviços públicos. Ele funciona a partir do cruzamento de informações das bases dos dados cadastrais partindo do número do CPF do cidadão.

A proposta envolve disponibilizar uma ferramenta unificada de atualização cadastral, suportada tecnologicamente por todas as entidades e órgãos públicos participantes do cadastro.

Os dados que inicialmente serão disponibilizados serão os dados biográficos que constam na base do CPF. São eles: número de inscrição no CPF; situação cadastral no CPF; nome completo; nome social; data de nascimento; sexo; filiação; nacionalidade; naturalidade; indicador de óbito; data de óbito, quando cabível; e data da inscrição ou da última alteração no CPF.



- **Comitê Central de Governança de Dados**

O Comitê Central de Governança de Dados (CCGD) é responsável, dentre outras tarefas, de elaborar as políticas e as diretrizes para a categorização de compartilhamento específico dos dados. A entidade também será responsável pela forma e o meio de publicação dessa categorização, tendo em mente a legislação pertinente, referente à proteção de dados pessoais e as regras e os parâmetros para o compartilhamento restrito, incluídos os padrões relativos à preservação do sigilo e da segurança.

O CCGD será composto sete membros: dois do Ministério da Economia, sendo um da Secretaria Especial de Desburocratização, Gestão e Governo Digital, que será o presidente do Comitê, e um da Secretaria Especial da Receita Federal do Brasil; um da Casa Civil da Presidência da República; um da Secretaria de Transparência e Prevenção da Corrupção da Controladoria-Geral da União; um da Secretaria Especial de Modernização do Estado da Secretaria-Geral da Presidência da República; um da Advocacia-Geral da União; e um do Instituto Nacional do Seguro Social (INSS).

O Comitê terá reuniões ordinárias a cada dois meses e extraordinárias sempre que for convocado pelo seu presidente ou um de seus membros. O quórum de reunião será dois terços de seus membros e o quórum de aprovação é por consenso. Suas decisões serão resoluções que serão publicadas pela Secretaria Especial de Desburocratização, Gestão e Governo Digital do Ministério da Economia.

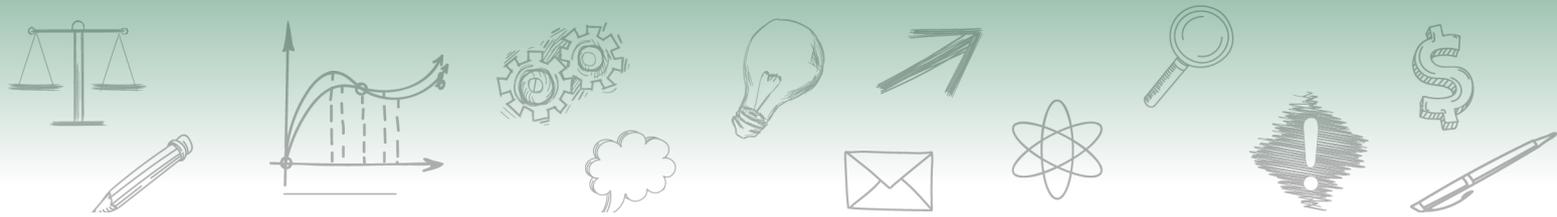
2.3 Dados de baixa qualidade custam dinheiro

Garantir que os dados sejam de alta qualidade é essencial para seu gerenciamento. Se os dados não atenderem às necessidades de seus consumidores - se não forem "adequados ao objetivo" -, então o esforço para coletar, armazenar, proteger e permitir o acesso a eles é desperdiçado. Para garantir que os dados atendam às necessidades dos negócios, as equipes de gerenciamento de dados devem trabalhar com os consumidores para definir as características que tornam os dados de alta qualidade.

A maioria dos usos dos dados envolve aprender com eles, a fim de aplicar esse aprendizado e criar valor. Por exemplo, entender os hábitos do cliente para melhorar um produto ou serviço; avaliar o desempenho organizacional ou as tendências do mercado, a fim de desenvolver uma melhor estratégia de negócios, etc.

Dados de baixa qualidade são simplesmente caros para qualquer organização. As estimativas diferem, mas os especialistas acreditam que as organizações gastam entre 10 a 30% da receita lidando com problemas de qualidade de dados. A IBM estimou o custo de dados de baixa qualidade nos Estados Unidos da América, em 2016, em US \$ 3,1 trilhões.

Muitos dos custos de dados de baixa qualidade são ocultos e indiretos e, portanto, difíceis de mensurar. No entanto, outros custos podem ser provenientes de:



- Retrabalho.
- Soluções alternativas e processos de correção ocultos.
- Ineficiências organizacionais ou baixa produtividade.
- Conflito organizacional.
- Insatisfação do usuário.
- Custos de oportunidade, incluindo a incapacidade de inovar.
- Custos ou multas de conformidade.

Os benefícios correspondentes de dados de alta qualidade incluem:

- Experiência aprimorada do usuário.
- Maior produtividade.
- Risco reduzido.
- Capacidade de atuar em oportunidades.
- Aumento de oportunidades.
- Vantagem competitiva obtida com insights sobre políticas, serviços, processos e oportunidades.
- Vantagem competitiva obtida com segurança e qualidade demonstráveis dos dados.

Não é algo único como esses custos e benefícios implicam gerenciar a qualidade dos dados. A produção de dados de alta qualidade requer planejamento, comprometimento e uma mentalidade que agrega qualidade aos processos e sistemas, em conformidade com o padrão ISO.

SAIBA MAIS

Padrão ISO de Qualidade de Dados

A International Organization for Standardization - ISO 8000, o padrão internacional para qualidade de dados, está sendo desenvolvido para permitir o intercâmbio de dados complexos de forma neutra em aplicações. Na introdução do padrão, a ISO afirma: “A capacidade de criar, coletar, armazenar, manter, transferir, processar e apresentar dados para dar suporte aos processos de negócios de



maneira oportuna e econômica exige uma compreensão das características dos dados que determinar sua qualidade e a capacidade de medir, gerenciar e relatar a qualidade dos dados".

A ISO 8000 define características que podem ser testadas por qualquer organização na cadeia de fornecimento de dados para determinar objetivamente a conformidade dos dados.

A primeira parte publicada da ISO 8000 (*parte 110, publicada em 2008*) enfocou a sintaxe, a codificação semântica e a conformidade com a especificação de dados dos Dados Mestres. Outras partes projetadas para o padrão incluem a parte 100 - Introdução, parte 120 - Proveniência, parte 130 - Precisão e parte 140 - Completude.

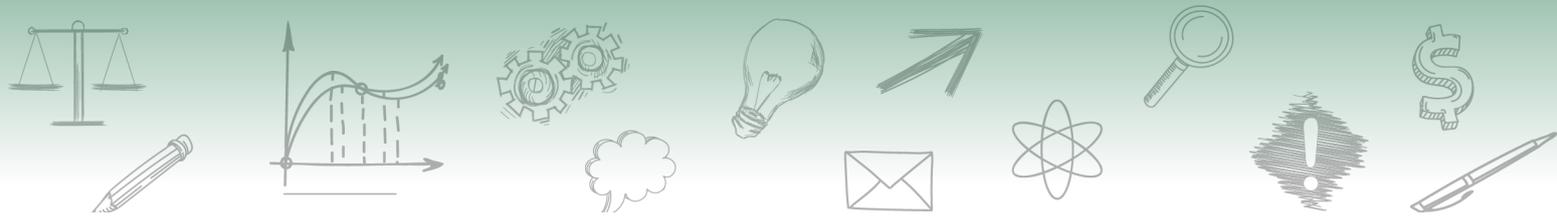
A ISO define dados de qualidade como “dados portáteis que atendem aos requisitos estabelecidos”. O padrão de qualidade dos dados está relacionado ao trabalho geral da ISO em portabilidade e preservação de dados. Os dados são considerados “portáteis” se puderem ser separados de uma aplicação de software. Os dados que só podem ser usados ou lidos usando um aplicativo de software licenciado específico estão sujeitos aos termos da licença do software. Uma organização pode não conseguir usar os dados criados, a menos que esses dados possam ser desanexados do software utilizado para criá-los.

Para atender aos requisitos estabelecidos é necessário que esses requisitos sejam definidos de maneira clara e inequívoca. O ISO 8000 é suportado pelo ISO 22745, um padrão para definir e trocar Dados Mestre. A ISO 22745 define como as declarações de requisitos de dados devem ser construídas, fornece exemplos em XML e define um formato para a troca de dados codificados. A ISO 22745 cria dados portáteis rotulando os dados usando um Dicionário Técnico Aberto compatível com ISO 22745, como o ECCMA *Open Technical Dictionary (eOTD)*.

A intenção da ISO 8000 é ajudar as organizações a definir o que é e o que não é dados de qualidade, permitir que eles solicitem dados de qualidade usando convenções padrão e verificar se eles receberam dados de qualidade usando os mesmos padrões. Quando os padrões são seguidos, os requisitos podem ser confirmados por meio de um programa de computador.

ISO 8000 - Parte 61 O modelo de referência do processo de gerenciamento da qualidade das informações e dos dados está em desenvolvimento. Esta norma descreverá a estrutura e organização do gerenciamento da qualidade dos dados, incluindo:

- Planejamento da qualidade dos dados;
- Controle de Qualidade de Dados;



- **Garantia de Qualidade de Dados;**
- **Melhoria da qualidade dos dados.**

eOTD é o acrônimo do ECCMA Open Technical Dictionary: dicionário é um banco de dados independente de idioma, com conceitos, termos e definições associados, usados para descrever de maneira inequívoca indivíduos, organizações, locais, bens, serviços, processos, regras e regulamentos. O EOTD é mantido pelo Comércio Eletrônico Código Management Association (ECCMA).

2.4 A avaliação de dados não é padronizada

Você sabe quanto custa coletar e gerenciar dados? E quanto custaria reconstruir o histórico, caso esses dados fossem perdidos?

Ainda assim, colocar atribuir valor aos dados é útil porque as decisões estratégicas se tornam a base para a compreensão do valor das atividades de gerenciamento de dados. Uma abordagem para a avaliação de dados é definir categorias gerais de custo e benefício que podem ser aplicadas de maneira consistente em uma organização.

As categorias de exemplo incluem:

- Custo de obtenção e armazenamento de dados.
- Custo da substituição de dados se eles forem perdidos.
- Impacto para a organização se houver perda dados.
- Custos potenciais de riscos associados aos dados.
- Custo de mitigação de risco.
- Custo de melhoria de dados.
- Benefícios de dados de maior qualidade.
- Vantagens de compartilhar esses dados.
- Retorno esperado pelo uso inovador desses dados.

A avaliação de ativos de dados também deve reconhecer que o valor dos dados é contextual (ou seja, o que é valioso para uma organização pode não o ser para outra) e geralmente temporal (ou seja, o uso foi vantajoso ontem e pode não o ser hoje). Apesar disso, dentro de uma organização, é provável que certos tipos de dados, como dados cadastrais, sejam consistentemente valiosos



ao longo do tempo; portanto, a maioria das organizações se concentra primeiro em garantir a qualidade desses dados altamente críticos.

2.5 Gerenciamento de dados significa gerenciar o ciclo de vida dos dados

Um dos motivos pelos quais as pessoas confundem gerenciamento de dados com gerenciamento de tecnologia é que geralmente avaliam dados apenas de um ponto de vista: a aplicação do qual os dados são acessados.

Pessoas não reconhecem que os dados podem ser separados das aplicações em que são criados ou armazenados e que os dados têm um ciclo de vida. O ciclo de vida é baseado no ciclo de vida do resultado. Concentra-se em garantir que os dados sejam criados, movidos e mantidos de maneira a torná-los utilizáveis pelas pessoas e processos que os exigem. Embora dados e tecnologia estejam interligados, o ciclo de vida dos dados não deve ser confundido com o ciclo de vida de desenvolvimento de sistemas (systems development lifecycle - SDLC), que se concentra na conclusão de projetos dentro do prazo e do orçamento.

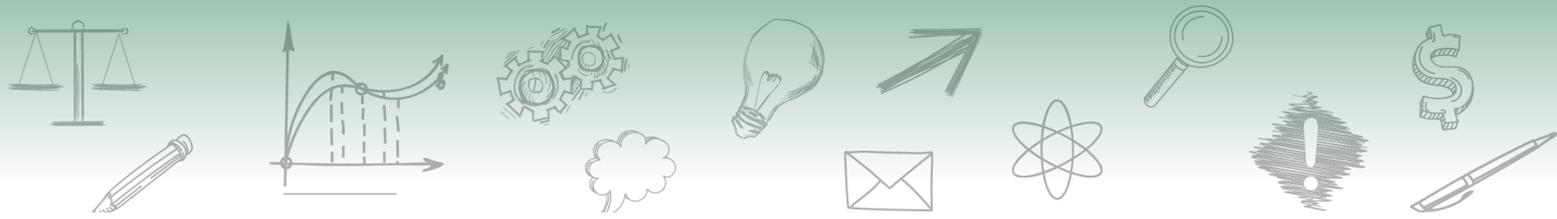
Conceitualmente, o ciclo de vida dos dados é fácil de descrever: inclui processos que criam ou obtêm dados, aqueles que os movem, transformam e armazenam e permitem que sejam mantidos e compartilhados, e aqueles que os usam ou os aplicam, bem como aqueles que os descartam.

Os dados raramente são estáticos: durante todo o seu ciclo de vida, os dados podem ser limpos, transformados, mesclados, aprimorados ou agregados. Os dados geralmente se movem horizontalmente dentro da organização: à medida que os dados são usados ou aprimorados, novos dados são criados, de modo que seu ciclo de vida possui iterações internas e os 'mesmos' dados podem ter requisitos de ciclo de vida diferentes em diferentes partes da organização.

A complexidade é adicionada ao conceito do ciclo de vida pelo fato de que diferentes tipos de dados têm requisitos de ciclo de vida diferentes. Por exemplo, os dados transacionais podem ser controlados em grande parte por meio da aplicação de regras básicas, enquanto os Dados Mestre requerem curadoria.

Conheça além desses, outros princípios que se aplicam ao ciclo de vida de qualquer dado:

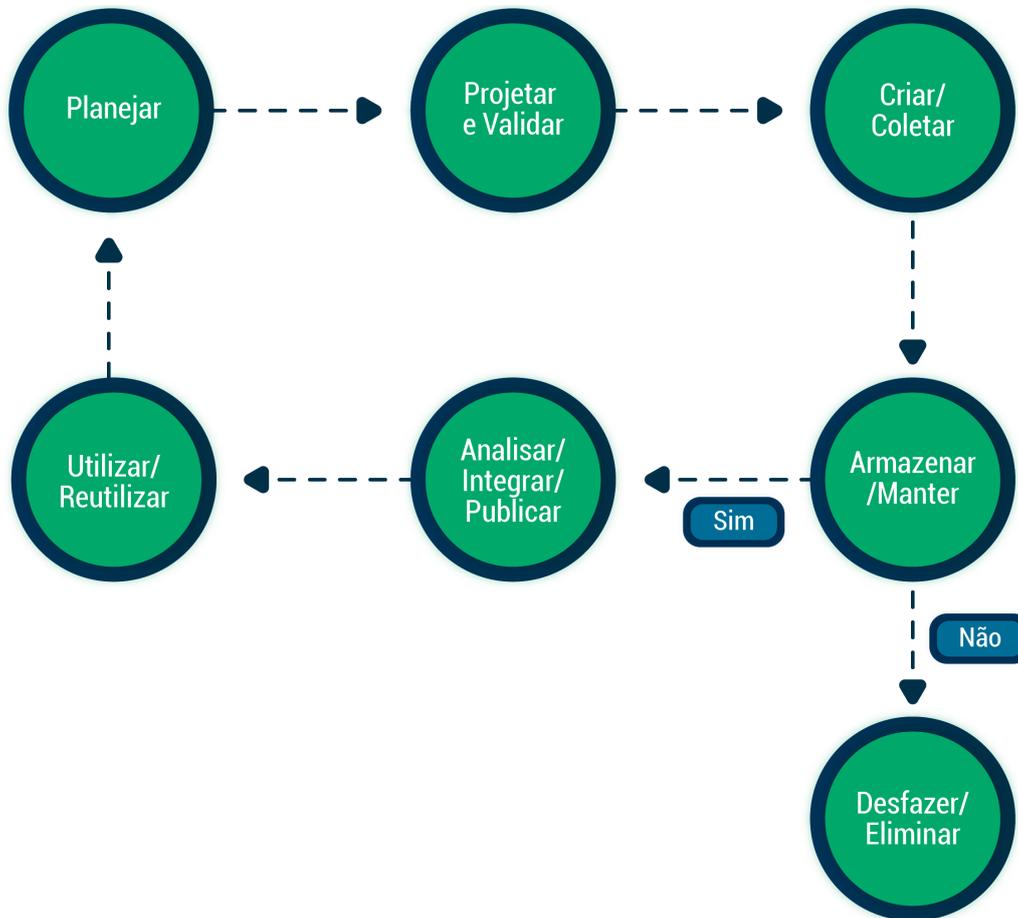
- **Criação e uso são os pontos mais críticos no ciclo de vida dos dados**
O gerenciamento de dados deve ser executado com um entendimento de como os dados são produzidos ou obtidos, assim como os dados são usados. Custa dinheiro para produzir dados. Os dados são valiosos somente quando são consumidos ou aplicados.



- **A qualidade dos dados deve ser gerenciada durante todo o ciclo de vida dos dados**
Como a qualidade dos dados pode ser afetada por vários eventos do ciclo de vida, a qualidade deve ser planejada como parte do ciclo de vida dos dados. Não é um complemento ou algo a ser 'feito mais tarde'.
- **A qualidade dos metadados deve ser gerenciada por meio do ciclo de vida dos dados**
Os metadados são tipos de dados utilizados para descrever outros dados. Como tal, é essencial para todas as funções de gerenciamento de dados. Os metadados geralmente são criados como parte do ciclo de vida de outros dados e devem ser vistos como um produto (e não como um subproduto) desse ciclo de vida. A qualidade dos metadados deve ser gerenciada da mesma maneira que a qualidade de outros dados.
- **A segurança dos dados deve ser gerenciada durante todo o ciclo de vida dos dados**
O gerenciamento de dados inclui a garantia de segurança dos dados e redução dos riscos associados. Os dados que requerem proteção devem ser protegidos por todo o ciclo de vida, desde a criação até o descarte.
- **Os esforços de gerenciamento de dados devem se concentrar nos dados mais críticos**
As organizações produzem muitos dados, muitos dos quais nunca são realmente usados. Tentar gerenciar todos os dados não é possível nem desejável. O gerenciamento do ciclo de vida requer o foco nos dados mais críticos de uma organização e a minimização do ROT de dados (ou seja, dados redundantes, obsoletos ou triviais).



Ciclo de Vida dos Dados



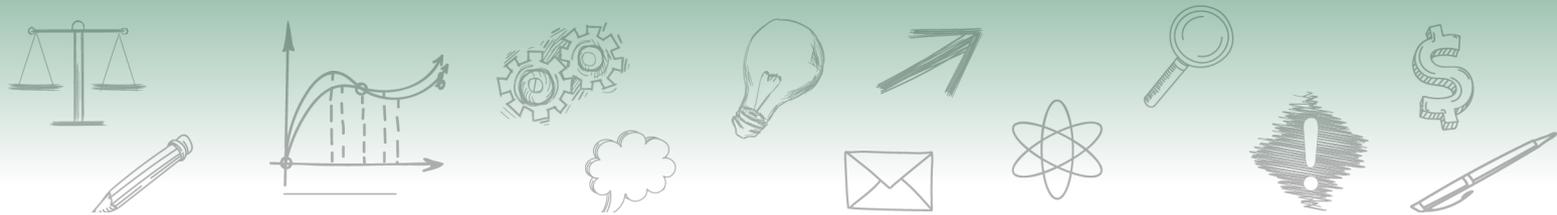
The Data Lifecycle (DMBOK2, p.29, com adaptações).

2.6 Diferentes tipos de dados têm diferentes requisitos de ciclo de vida

O gerenciamento de dados se torna mais complicado pelo fato de diferentes tipos de dados apresentarem requisitos diferentes de gerenciamento do ciclo de vida.

Os dados podem ser classificados conforme:

- função que servem (por exemplo, dados transacionais, dados de referência, dados mestre, metadados; alternativamente, dados de categoria, dados de recursos, dados de eventos, dados detalhados de transações);
- conteúdo (domínios de dados, áreas de conhecimento);
- formato ou nível de proteção exigido pelos dados; e
- forma e onde são armazenados ou acessados.



Como diferentes tipos de dados têm requisitos diferentes, eles estão associados a riscos diferentes e desempenham papéis diferentes dentro de uma organização. A maioria das ferramentas de gerenciamento de dados estão focadas nos aspectos de classificação e controle. Por exemplo, os Dados Mestre têm usos diferentes e, conseqüentemente, requisitos de gerenciamento diferentes dos dados transacionais.

Como variados tipos de dados têm requisitos diferentes, eles estão associados a riscos distintos e desempenham papéis diferentes dentro de uma organização. A maioria das ferramentas de gerenciamento de dados estão focadas nos aspectos de classificação e controle.

Por exemplo, os Dados Mestre têm usos discrepantes e, conseqüentemente, requisitos de gerenciamento desiguais dos dados transacionais.

O quadro a seguir, apresenta a análise sobre as fases do ciclo de vida dos dados, considerando a Lei nº 13.709/2018, Lei Geral de Proteção de Dados (LGPD):

CICLO DE VIDA DOS DADOS		
Fase do Ciclo	Antes da LGPD	Com a LGPD
Coleta	Os dados pessoais são coletados indiscriminadamente.	Os dados pessoais coletados devem obedecer ao princípio da necessidade e da finalidade.
Processamento	Os dados podem ser processados sem um tratamento específico.	O processamento de dados só poderá ser realizado se o tratamento estiver enquadrado no Art. 7º da LGPD.
Análise	A análise de dados é feita para entender o mercado, conhecer o perfil das pessoas e definir estratégias para oferecer bens e serviços para o público-alvo.	A análise dos dados deve levar em consideração a finalidade da coleta. Devem ser obedecidos os princípios de tratamento com propósito legítimo específico e explícito.
Compartilhamento	Os dados pessoais são compartilhados sem a necessidade do consentimento de seus titulares.	O compartilhamento de dados deve ser consentido pelos seus titulares* *Ver Inciso II do Art. 3º do Decreto nº 10.046/2019.
Armazenamento	Os dados pessoais são reutilizados sem a necessidade de consentimento de seus titulares.	Os dados pessoais devem ser armazenados e mantidos por prazos definidos, ou seja, até que a finalidade seja alcançada ou deixem de ser necessários ou pertinentes ao alcance da finalidade.



Reutilização	Os dados pessoais são reutilizados sem a necessidade de consentimento de seus titulares.	Um novo consentimento deve ser solicitado sempre que houver mudança de finalidade.
Eliminação	Os dados pessoais são mantidos sem a obrigatoriedade de serem eliminados.	Os dados pessoais devem ser eliminados após o término de seu tratamento.

A depender do objetivo da coleta do dado, segue-se um ciclo de vida. Após o finalizado o objetivo de coleta de dados, os trabalhos devem continuar sobre os dados produzidos, a fim de que eles sejam reutilizados após análises, citações e atualizações. Dessa forma, existe a possibilidade de reutilização dos dados para outros fins, que seja até para alimentar sistemas que compõem um ecossistema relacionado a uma política pública ou de pesquisas estatísticas para fins de planejamento de outras políticas para atendimento de demandas de interesse da sociedade.

2.7 Os metadados devem ser gerenciados como parte do ciclo de vida dos dados.

Existe uma atenção especial dos profissionais de gerenciamento de dados em relação aos metadados porque sabem da sua importância no processo de gestão de dados. No entanto, é um fato que estes profissionais se limitam em utilizar o termo Metadados ao se direcionar à Alta Administração, visto que essa gestão deve ser transparente aos usuários/consumidores e bastante eficiente no contexto técnico.

Os metadados incluem uma variedade de informações que permitem às pessoas entender dados e os sistemas que os contêm. Os metadados descrevem quais dados uma organização possui, o que representa, como é classificado, sua origem, sua interoperabilidade dentro da organização, como evolui por meio do uso, quem pode ou não acessá-los e se são de alta qualidade.

O desafio não é evidenciar a importância dos metadados para gerenciá-los, mas a forma de ser gerenciados como tal. As organizações que não gerenciam bem seus dados geralmente não gerenciam seus metadados. A resposta para esse desafio é que o gerenciamento de metadados geralmente fornece um ponto de partida para melhorias no gerenciamento de dados em geral.

2.8 O gerenciamento de dados geralmente é confundido com o gerenciamento de tecnologia da informação

Como quase todos os dados atuais são armazenados eletronicamente, o gerenciamento de dados está intimamente ligado ao gerenciamento de tecnologia. Eles precisam ser vistos em relação um ao outro, porque as decisões sobre a tecnologia afetam muitas facetas de como os dados são gerenciados.



Observe que:

- O gerenciamento de dados se concentra em garantir que os dados em si sejam utilizáveis e confiáveis; e
- O gerenciamento de tecnologia se concentra na criação e manutenção de infraestrutura, sistemas e aplicações.

Os dois estão fundamentalmente conectados pelo fato de que sistemas e aplicações automatizam processos de negócios que coletam ou criam dados e diferentes opções tecnológicas impõem restrições diferentes aos próprios dados. Os requisitos de gerenciamento de dados e gerenciamento de tecnologia devem estar enraizados nos processos de negócios que criam ou usam dados e nas necessidades das pessoas e processos que consomem dados.

Em muitas organizações, existe uma tensão contínua entre o desejo de construir novas tecnologias e o desejo de ter dados mais confiáveis - como se os dois se opusessem um ao outro, em vez de precisarem um do outro.

O gerenciamento bem-sucedido de dados requer decisões sólidas sobre tecnologia, mas gerenciar tecnologia não é o mesmo que gerenciar dados. As organizações precisam entender o impacto da tecnologia nos dados, para impedir que a tentação tecnológica direcione suas decisões sobre os dados. Em vez disso, os requisitos de dados alinhados à estratégia de negócios devem orientar as decisões sobre tecnologia.

2.9 O que o gerenciamento de dados requer

O gerenciamento de dados é multifuncional e requer habilidades e uma perspectiva corporativa. Vamos conhecê-las.

Habilidades

O gerenciamento de dados envolve um conjunto de processos interconectados alinhados ao ciclo de vida dos dados. Embora muitas organizações vejam o gerenciamento de dados como uma função de tecnologia da informação, na verdade, requer uma ampla gama de pessoas com um conjunto diversificado de habilidades trabalhando em diferentes partes de uma organização.

O gerenciamento de dados é um processo complexo porque é executado em toda a organização. Os dados são gerenciados em diferentes locais da organização por equipes que são responsáveis por diferentes fases do ciclo de vida dos dados.

O gerenciamento de dados requer:

- Habilidades de processos de negócios para entender e planejar a criação de dados confiáveis.



- Habilidades de design para planejar sistemas nos quais os dados serão armazenados ou usados.
- Habilidades altamente técnicas para administrar hardware e criar software onde os dados são mantidos.
- Habilidades de análise de dados para entender questões e problemas descobertos nos dados.
- Habilidades analíticas para interpretar dados e aplicá-los a novos problemas.
- Habilidades linguísticas para trazer consenso às definições e modelos para que as pessoas possam entender dados.
- Pensamento estratégico para visualizar oportunidades de uso de dados para atender usuários/consumidores e atingir metas.

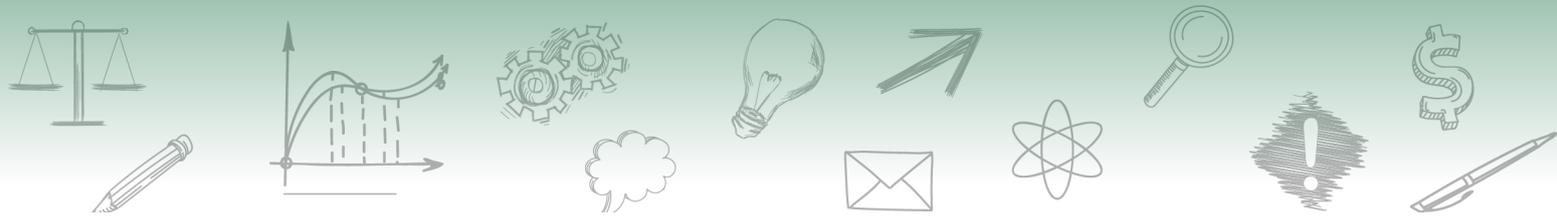
O desafio é fazer com que as pessoas com esse rol de habilidades e perspectivas reconheçam como as peças se encaixam e como o trabalho delas são transversais com o trabalho de outras partes da organização, para que elas colaborem com sucesso e alcancem objetivos comuns.

Uma perspectiva corporativa

A maioria das organizações divide o trabalho por unidades ou funções de negócios, cada uma das quais pode desenvolver suas próprias aplicações para executar seu trabalho. Como os dados geralmente são vistos simplesmente como subprodutos de processos operacionais, nem sempre são planejados para além da necessidade imediata. Pode até não ser reconhecido como algo que outras pessoas e processos utilizam.

A menos que os padrões de dados corporativos sejam estabelecidos e aplicados, haverá diferenças na forma como os dados são definidos e criados em diferentes áreas. Por exemplo, considere algo aparentemente simples como um Número de Segurança Social (Social Security Number - SSN), um identificador americano para indivíduos. Se uma aplicação captura o SSN como um valor numérico e outro a captura em um campo de texto, os dados do SSN serão formatados de maneira diferente. Isso pode resultar em problemas como a eliminação de zeros à esquerda nos SSNs.

Diferenças de formatação, diferenças na granularidade dos dados e diferenças sobre quais atributos são obrigatórios para capturar - todas essas diferenças apresentam obstáculos à integração de dados de diversas aplicações. Os obstáculos à integração limitam o valor que uma organização pode obter com seus dados.



As organizações que veem os dados como um produto que eles criam ou compram tomarão melhores decisões sobre como gerenciá-los ao longo de seu ciclo de vida. Essas decisões exigem o reconhecimento de:

- As formas como os dados conectam processos de negócios que, de outra forma, poderiam ser vistos como separados.
- A relação entre 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 maneiras pelas quais os dados podem ser usados para avançar a estratégia organizacional.

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

O desafio é que geralmente existem pressões organizacionais, bem como as constantes pressões de tempo e dinheiro, que atrapalham o planejamento. As organizações devem equilibrar as metas de longo e curto prazo à medida que executam sua estratégia. Ter clareza sobre as compensações leva a melhores decisões.

O que você precisa saber:

DESTAQUE

- Os dados são um ativo valioso, mas também representam risco. Uma organização pode começar a entender o valor de seus dados reconhecendo os custos de dados de baixa qualidade e os benefícios de dados de alta qualidade.
- Os dados têm características únicas que dificultam o gerenciamento.
- A melhor abordagem para enfrentar esses desafios é gerenciar dados em todo o seu ciclo de vida e adotar uma perspectiva corporativa.
- A falha no gerenciamento do ciclo de vida dos dados é cara, embora muitos custos estejam ocultos.
- O gerenciamento de dados em todo o ciclo de vida requer planejamento, habilidade e trabalho em equipe.



3. Princípios de gerenciamento de dados do DAMA (DATA MANAGEMENT ASSOCIATION)

O gerenciamento de dados apresenta desafios únicos conectados à natureza dos dados em si. Mesmo com suas características exclusivas, o gerenciamento de dados ainda compartilha características com outras formas de gerenciamento. Envolve conhecer quais dados uma organização possui e a melhor forma de usar esses ativos para atingir as metas organizacionais. Como outros processos de gerenciamento, deve-se equilibrar as necessidades estratégicas e operacionais.

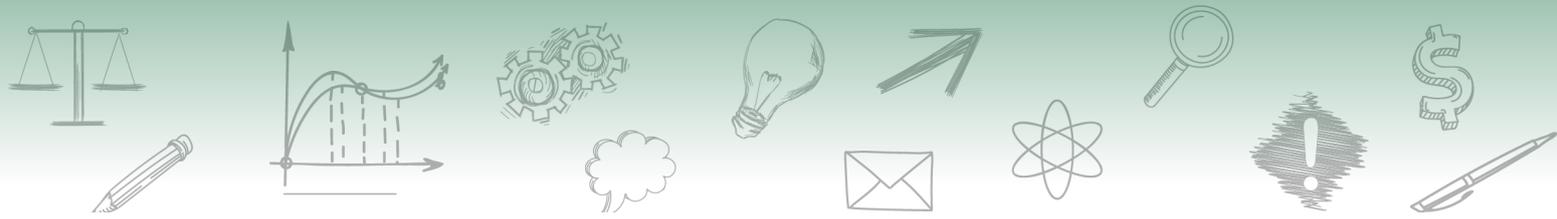
Para ajudar as organizações a alcançar esse equilíbrio, o DAMA desenvolveu um conjunto de princípios que reconhecem os desafios do gerenciamento de dados e auxiliam a orientar as práticas de gerenciamento de dados.

Em nível mais estratégico, esses princípios se resumem a quatro asserções que examinaremos neste módulo:

- Os dados possuem valor estratégico.
- Requisitos de gerenciamento de dados são requisitos de negócios.
- Gerenciamento de dados refere-se ao gerenciamento do ciclo de vida.
- O gerenciamento de dados depende de diversas habilidades.

Os princípios de gerenciamento de dados da DAMA fornecem uma lente para entender como sua organização gerencia seus dados. Depois de revisar suas implicações, as examinaremos no contexto da maturidade do gerenciamento de dados. Um modelo de maturidade define uma progressão do controle crescente sobre um conjunto de processos. Quando uma organização obtém uma compreensão das características do processo, pode estabelecer um plano para melhorar suas capacidades. Ele também pode medir a melhoria e comparar-se aos concorrentes ou parceiros, guiados pelos níveis do modelo.

Os modelos de maturidade do gerenciamento de dados descrevem detalhes dos processos de gerenciamento de dados que podem ser usados para esse tipo de avaliação.



Princípios de Gerenciamento de Dados	
Dados são estratégicos	
Dados são um ativo de propriedade exclusiva.	
O valor dos dados pode e deve ser expresso em termos econômicos/estratégicos.	
Gerenciamento eficaz de dados requer comprometimento da liderança.	
Requisitos de gerenciamento de dados são requisitos de negócios	
Gerenciar dados significa gerenciar a qualidade dos dados.	
São necessários metadados para gerenciar dados.	
É preciso planejar o gerenciamento de dados	
Requisitos de gerenciamento de dados devem orientar as decisões de Tecnologia da Informação.	
Gerenciamento de dados é gerenciamento do ciclo de vida	Gerenciamento de dados depende de diversas habilidades
Tipos diferentes de dados têm características diferentes do ciclo de vida.	Gerenciamento de dados é multifuncional.
Gerenciar dados inclui gerenciar o risco associado aos dados.	O gerenciamento de dados requer uma perspectiva corporativa.
	O gerenciamento de dados deve levar em conta uma variedade de perspectivas.

Fonte: Data Management Principles (Adapted from DMBOK2)

3.1. Os dados são valiosos

- **Os dados são um ativo com propriedades exclusivas:** os dados são um ativo, mas diferem de outros ativos de maneiras importantes que afetam a maneira como são gerenciados. A mais óbvia dessas propriedades é que os dados não são consumidos quando usados, assim como os ativos financeiros e físicos.
- **O valor dos dados pode e deve ser expresso em termos econômicos:** Chamar dados de um ativo implica ele ter valor. Embora existam técnicas para medir o valor qualitativo e quantitativo dos dados, ainda não existem padrões para isso. As organizações que desejam tomar melhores decisões sobre seus dados devem desenvolver maneiras consistentes de quantificar esse valor. Eles também devem medir os custos de dados de baixa qualidade e os benefícios de dados de alta qualidade.
- **O gerenciamento eficaz de dados requer comprometimento da liderança:** o gerenciamento de dados envolve um conjunto complexo de processos que, para serem eficazes, requerem coordenação, colaboração e comprometimento. Chegar lá exige não apenas habilidades de gerenciamento, mas também a visão e o objetivo advindos da liderança comprometida.

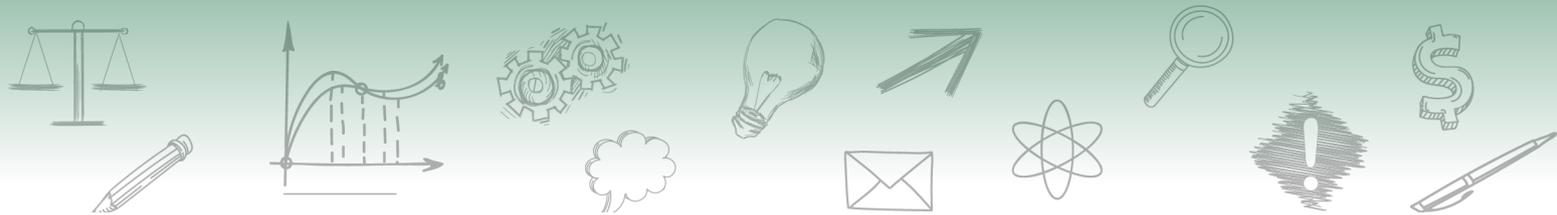


3.2. Requisitos de gerenciamento de dados são requisitos de negócios

- **Gerenciar dados significa gerir a sua qualidade:** garantir que os dados sejam adequados ao objetivo é um objetivo principal do gerenciamento de dados. Para gerenciar a qualidade, as organizações devem garantir que compreendem os requisitos de qualidade das partes interessadas e medem os dados em relação a esses requisitos.
- **São necessários 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, para entender o que são e como usá-los, requer definição e conhecimento na forma de Metadados. Os metadados são originários de uma variedade de processos relacionados à criação, processamento e uso de dados, incluindo arquitetura, modelagem, administração, governança, gerenciamento da qualidade dos dados, desenvolvimento de sistemas, operações de negócios e TI e análise.
- **É preciso planejar o gerenciamento de dados:** Mesmo pequenas organizações podem ter cenários técnicos e de negócios complexos. Os dados são criados em muitos locais e transitam entre vários locais para seu uso. Coordenar o trabalho e manter os resultados finais alinhados requer planejamento da perspectiva de arquitetura e do processo.
- **Os requisitos de gerenciamento de dados devem orientar as decisões em tecnologia da informação:** o gerenciamento de dados está extremamente interligado à tecnologia da informação e ao gerenciamento da tecnologia da informação. O gerenciamento de dados requer uma abordagem que garanta que a tecnologia atenda e direcione as necessidades estratégicas de dados de uma organização.

3.3. Gerenciamento de dados é gerenciamento do ciclo de vida

- **Gerenciamento de dados é gerenciamento do ciclo de vida:** os dados têm um ciclo de vida e o gerenciamento de dados requer a administração de seu ciclo de vida. Como os dados geram mais dados, o próprio ciclo de vida dos dados pode ser muito complexo. As práticas de gerenciamento de dados precisam levar em consideração o ciclo de vida em evolução dos dados.
- **Diferentes tipos de dados têm características diferentes de ciclo de vida:** e por esse motivo, eles têm requisitos de gerenciamento diferentes. As práticas de gerenciamento de dados precisam reconhecer essas diferenças e ser flexíveis o suficiente para atender a diferentes tipos de requisitos do ciclo de vida dos dados.



- **O gerenciamento de dados inclui o controle dos riscos associado aos dados:** além de ativos, os dados também representam riscos para a organização. Os dados podem ser perdidos, furtados ou mal utilizados. As organizações devem considerar as implicações éticas de seu uso. Os riscos relacionados aos dados devem ser gerenciados como parte do ciclo de vida dos dados.

3.4. O gerenciamento de dados depende de diversas habilidades

- **O gerenciamento de dados é multifuncional:** uma única equipe não pode gerenciar todos os dados de uma organização. Fazer isso requer uma série de habilidades e conhecimentos. O gerenciamento de dados requer habilidades técnicas e não técnicas além da capacidade de colaborar.
- **O gerenciamento de dados requer uma perspectiva corporativa:** o gerenciamento de dados possui aplicativos locais, mas deve ser aplicado em toda a instituição para ser o mais eficaz possível. Essa é uma das razões pelas quais o gerenciamento e o controle de dados estão interligados.
- **O gerenciamento de dados deve dar conta de várias perspectivas:** os dados são fluidos e estão em mudança. O gerenciamento de dados deve evoluir constantemente para acompanhar não somente a maneira como os dados são criados e usados, mas também os consumidores que os usam.

3.5. Princípios e maturidade em gerenciamento de dados

Reconhecer a importância, identificar os desafios e conhecer os princípios do gerenciamento de dados são relevantes para uma gestão eficiente. Sua organização está, sem dúvida, aplicando alguns desses princípios, pois segue, certamente, algumas das práticas descritas aqui. Mas, a menos que a organização aumente sua consciência por meio de um nível de autoavaliação, é improvável que ela seja capaz de melhorar suas práticas.

A Avaliação da Maturidade da Capacidade (CMA) é uma abordagem para a melhoria de processos com base em uma estrutura - um Modelo de Maturidade da Capacidade (CMM) - que descreve como as características de um processo evoluem de *ad hoc* para o ideal.

O conceito de CMA surgiu dos esforços do Departamento de Defesa dos Estados Unidos para estabelecer critérios através dos quais avaliar fornecedores de software. Em meados da década de 1980, o *Capability Maturity Model for Software* foi publicado pelo Instituto de Engenharia de Software da Universidade Carnegie-Mellon. Embora aplicados pela primeira vez ao desenvolvimento de software, os CMMs foram desenvolvidos para uma variedade de outros campos, incluindo gerenciamento de dados.

Modelos de maturidade são definidos em termos de uma progressão através de níveis que descrevem as características do processo. Quando uma organização obtém uma compreensão das características do processo, pode avaliar seu nível de maturidade e implementar um plano



para melhorar suas capacidades. Há possibilidade de medir a melhoria e comparar-se aos concorrentes ou parceiros, guiados pelos níveis do modelo.

A avaliação da maturidade da capacidade é um instrumento bastante eficiente para isso. Trata-se de uma abordagem para a melhoria do processo com base em uma estrutura - um Modelo de Maturidade da Capacidade - que descreve como as características de um processo evoluem de *ad hoc* para ideal.

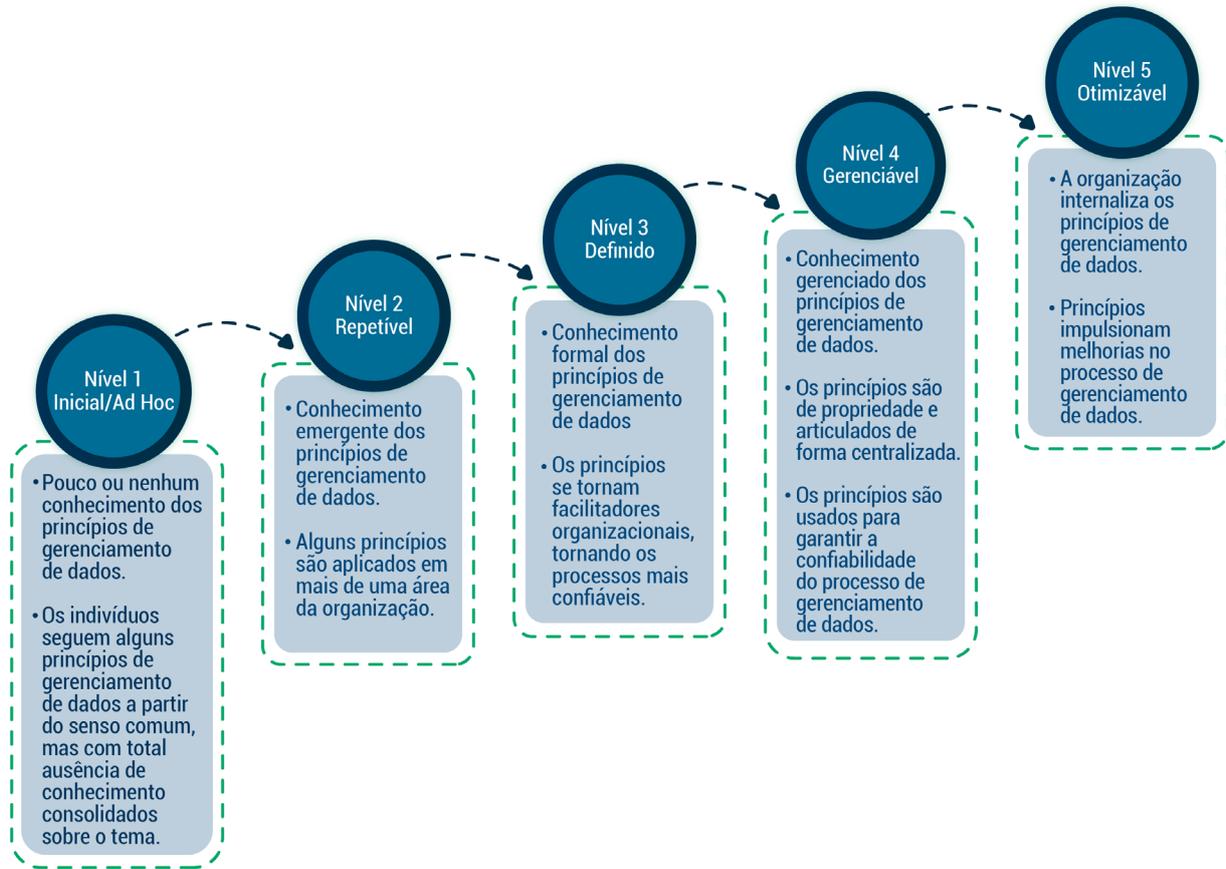
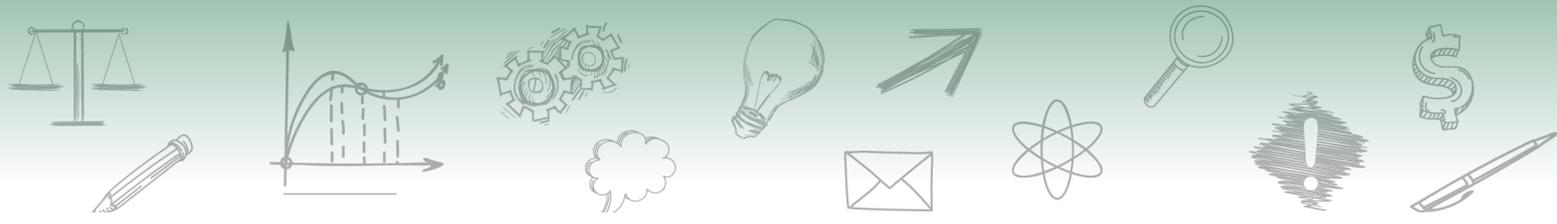
A cada novo nível, a execução do processo se torna mais consistente, previsível e confiável. Os processos melhoram à medida que assumem as características desses níveis. A progressão acontece em uma ordem definida que deve ser cumprida a cada fase conforme segue:

DESTAQUE

- **Nível 0 Ausência de capacidade**
- **Nível 1 Inicial ou ad hoc: o sucesso depende da competência dos indivíduos**
- **Nível 2 Repetível: a disciplina mínima do processo está em vigor**
- **Nível 3 Definido: os padrões são definidos e usados**
- **Nível 4 Gerenciável: os processos são quantificados e controlados**
- **Nível 5 Otimizável: as metas de melhoria de processo são quantificadas**

Dentro de cada nível, os critérios são descritos nos recursos do processo. Por exemplo, um modelo de maturidade pode incluir critérios relacionados à forma como os processos são executados, incluindo o nível de automação desses processos. Pode se concentrar em políticas e controles, além de detalhes do processo. Essa avaliação ajuda a identificar o que está funcionando bem, o que não está funcionando bem e onde uma organização tem lacunas.

A maturidade do uso dos princípios de gerenciamento de dados pode progredir conforme ilustrado na figura abaixo, onde uma organização passa do conhecimento limitado dos princípios de gerenciamento de dados para um estado em que os princípios se tornam motores da melhoria organizacional .



Fonte: Maturity Model for the Application of Data Management Principles (Adapted from DMBOK2)

Um DMMA (*Data Management Maturity Assessment*) pode ser usado para avaliar o gerenciamento de dados em geral, ou para se concentrar em uma única área funcional ou de conhecimento, ou mesmo em um único processo ou ideia (como o grau em que uma organização segue os dados).

Qualquer que seja o foco, um DMMA pode ajudar a preencher a lacuna entre as perspectivas de negócios e de TI sobre a saúde e a eficácia das práticas de gerenciamento de dados. Um DMMA fornece uma linguagem comum para descrever como é o progresso nas funções de gerenciamento de dados e oferece um caminho para melhorias baseado em etapas, que pode ser adaptado às prioridades estratégicas de uma organização. Assim, pode ser usado para definir e medir objetivos organizacionais, bem como comparar a organização com outras organizações ou referências do setor.

Conheça um pouco mais sobre o DMMA



SAIBA MAIS

Frameworks DMMA existentes

Uma estrutura de avaliação de maturidade do gerenciamento de dados é segmentada em tópicos de gerenciamento de dados discretos. O foco e o conteúdo da estrutura variam dependendo de terem um foco geral ou específico do setor. No entanto, a maioria aborda assuntos que podem ser mapeados para as áreas de conhecimento do DAMA-DMBOK. Os exemplos abaixo destinam-se a ilustrar a variedade de Modelos de Maturidade de Capacidade que foram desenvolvidos no espaço de gerenciamento de dados. Muitos fornecedores desenvolveram seus próprios modelos. As organizações devem avaliar vários modelos antes de escolher um fornecedor ou antes de desenvolver sua própria estrutura.

01. Modelo de maturidade (DMM) do *CMMI Data Management*

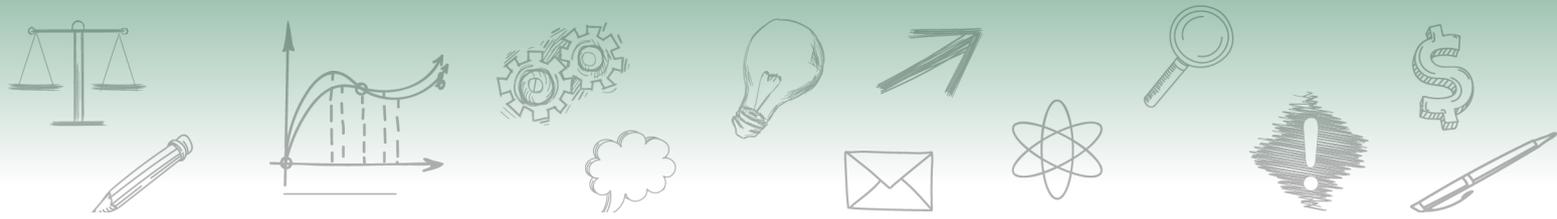
O CMMI (*Capability Maturity Model Institute*) desenvolveu o CMMI-DMM (Data Management Maturity Model), que fornece critérios de avaliação para as seguintes áreas de gerenciamento de dados:

- Estratégia de Gerenciamento de Dados;
- Gestão de dados;
- Qualidade dos dados;
- Plataforma e Arquitetura;
- Operações de dados;
- Processos de Suporte.

Dentro de cada um desses processos, o modelo identifica subprocessos para avaliação. Por exemplo, a seção Qualidade dos dados é responsável por Estratégia de qualidade dos dados e Avaliação, perfil e limpeza da qualidade dos dados. O modelo também explica a relação entre as áreas de gerenciamento de dados. Por exemplo, a necessidade de alinhamento das partes interessadas e a relação entre processos de negócios e Gerenciamento de qualidade de dados.

02. Conselho EDM DCAM

O *Enterprise Data Management Council*, uma organização de defesa do



setor para serviços financeiros com sede nos Estados Unidos, desenvolveu o DCAM (Data Management Capability Assessment Model). Resultado de um esforço dirigido por membros para obter consenso sobre as melhores práticas de gerenciamento de dados, o DCAM descreve 37 recursos e 115 sub-recursos associados ao desenvolvimento de um programa sustentável de Gerenciamento de Dados. A pontuação se concentra no nível de engajamento das partes interessadas, formalidade do processo e existência de artefatos que demonstram a conquista de recursos.

03. Modelo de Maturidade do IBM Data Governance Council

O Modelo de Maturidade do IBM Data Governance Council foi baseado nas contribuições de um conselho de 55 (cinquenta e cinco) organizações. Os membros do conselho colaboraram para definir um conjunto comum de comportamentos observáveis e desejados que as organizações podem usar para avaliar e projetar seus próprios programas de governança de dados. O objetivo do modelo é ajudar as organizações a criar consistência e controle de qualidade na governança por meio de tecnologias comerciais comprovadas, métodos colaborativos e melhores práticas. O modelo está organizado em torno de quatro categorias principais:

Resultados: Gerenciamento e conformidade de riscos de dados, criação de valor.

Facilitadores: Estrutura organizacional e conscientização, política, administração.

Disciplinas principais: gerenciamento da qualidade dos dados, gerenciamento do ciclo de vida das informações, segurança e privacidade das informações.

Disciplinas de Suporte: Arquitetura de Dados, classificação e Metadados, informações de auditoria, registro e geração de relatórios.

O modelo IBM é apresentado como um *Maturity Framework* e como um conjunto de perguntas de avaliação com respostas construídas para indicar níveis de maturidade.

04. Modelo de maturidade em governança de dados da Stanford

O *Stanford Data Governance Maturity Model* foi desenvolvido para uso da Universidade; não era para ser um padrão da indústria. Mesmo assim, serve como um exemplo sólido de um modelo que fornece orientação e um padrão de medição. O modelo se concentra na governança de dados, não no gerenciamento de dados, mas, no entanto, fornece uma base para avaliar o gerenciamento de dados em geral. O modelo diferencia os componentes fundamentais (conscientização, formalização, metadados) e do projeto



(gerenciamento de dados, qualidade dos dados, dados mestre). Dentro de cada um, articula drivers para pessoas, políticas e capacidades. Em seguida, articula características de cada nível de maturidade. Ele também fornece medições qualitativas e quantitativas para cada nível.

05. Modelo de maturidade do gerenciamento de informações corporativas do Gartner

O *Gartner* publicou um modelo de maturidade EIM, que estabelece critérios para avaliar visão, estratégia, métricas, governança, funções e responsabilidades, ciclo de vida e infraestrutura.

Um Framework, ou arcabouço conceitual, é um conjunto de conceitos usado para resolver um problema de um domínio específico. Framework conceitual não se trata de um software executável, mas sim de um modelo de dados para um domínio.

O framework atua onde há funcionalidades em comum a várias aplicações, porém para isso as aplicações devem ter algo razoavelmente grande em comum para que o mesmo possa ser utilizado em várias aplicações.

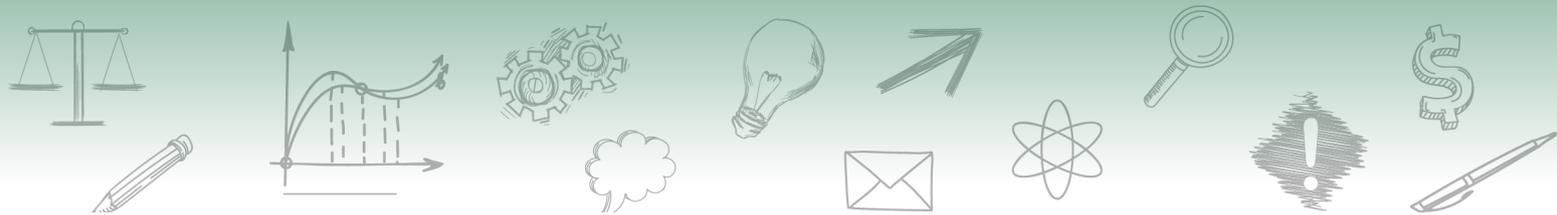
3.6. Reavaliar a maturidade

As reavaliações devem ser realizadas em intervalos regulares, convergindo para um ciclo de melhoria contínua:

DESTAQUE

- Estabelecer uma classificação de linha de base através da primeira avaliação.
- Definir parâmetros de reavaliação, incluindo o escopo organizacional.
- Repetir a avaliação do DMM conforme necessário em um planejamento o publicado.
- Acompanhar tendências em relação à linha de base inicial.
- Desenvolver recomendações com base nos resultados da reavaliação.

A reavaliação também pode revigorar ou reorientar o esforço. O progresso mensurável ajuda a manter o compromisso e o entusiasmo em toda a organização. Alterações em estruturas regulatórias, políticas internas ou externas ou inovações que possam mudar a abordagem de governança e estratégias são razões adicionais para reavaliar periodicamente.



3.7. Ferramentas de avaliação de maturidade

- **Estrutura de maturidade do gerenciamento de dados:** A principal ferramenta usada em uma avaliação de maturidade é a própria estrutura do DMM.
- **Plano de comunicação:** Um plano de comunicação inclui um modelo de envolvimento das partes interessadas, o tipo de informação a ser compartilhada e o cronograma para o compartilhamento de informações.
- **Ferramentas de colaboração:** As ferramentas de colaboração permitem que os resultados da avaliação sejam compartilhados. Além disso, evidências de práticas de gerenciamento de dados podem ser encontradas nos e-mails, nos modelos concluídos e nos documentos de revisão criados por meio de processos padrão para design colaborativo, operações, rastreamento de incidentes, revisões e aprovações.
- **Repositórios de gerenciamento e de conhecimento em metadados:** padrões de dados, políticas, métodos, agendas, atas de reuniões ou decisões e artefatos de negócio e técnicos que servem como prova de prática podem ser gerenciados nesses repositórios. Em alguns CMMs, a falta desses repositórios é um indicador de menor maturidade na organização. Os repositórios de metadados podem existir em várias construções, o que pode não ser óbvio para os participantes. Por exemplo, alguns aplicativos de Business Intelligence dependem completamente dos metadados para compilar suas visualizações e relatórios, sem se referir a ele como um repositório distinto separado.

3.8. Técnicas de avaliação de maturidade

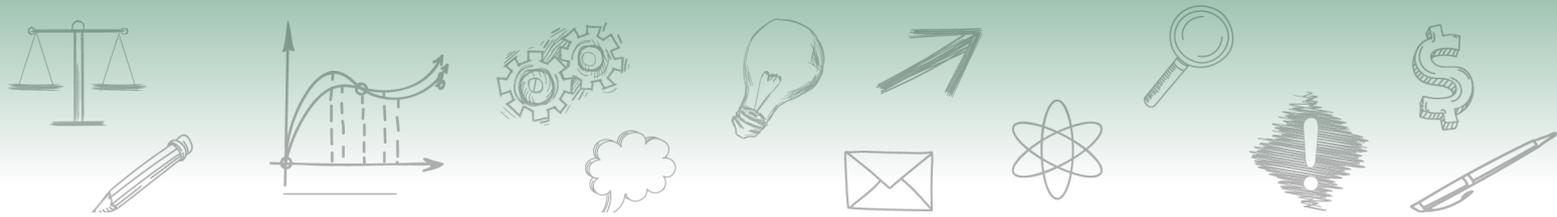
Diversas técnicas relacionadas à execução de um DMMA são definidas pela metodologia da estrutura de DMM escolhida. Técnicas mais gerais são descritas aqui.

Os seguintes critérios devem ser considerados ao selecionar uma estrutura DMM.

- **Acessibilidade**
As práticas são declaradas em termos não técnicos que transmitem a essência funcional da atividade.
- **Abrangência**
A estrutura trata de um amplo escopo de atividades de gerenciamento de dados e inclui o envolvimento dos negócios, não apenas os processos de TI.
- **Extensível e flexível**
o modelo está estruturado para permitir o aprimoramento de disciplinas específicas do setor ou adicionais e pode ser usado no todo ou em parte, dependendo das necessidades da organização.



- **Caminho do progresso futuro incorporado**
embora as prioridades específicas sejam diferentes de organização para organização, a estrutura do DMM descreve um caminho lógico a seguir em cada uma das funções que descreve.
- **Agnóstico e específico do setor**
algumas organizações se beneficiarão de uma abordagem específica do setor, outras de uma estrutura mais genérica. Qualquer estrutura de DMM também deve aderir às melhores práticas de gerenciamento de dados que cruzam verticais.
- **Nível de abstração ou detalhe**
as práticas e os critérios de avaliação são expressos em um nível de detalhe suficiente para garantir que eles possam estar relacionados à organização e ao trabalho que ela realiza.
- **Não prescritivo**
A estrutura descreve o que precisa ser executado, não como deve ser executado.
- **Organizado por tópico**
A estrutura coloca as atividades de gerenciamento de dados em seu contexto apropriado, permitindo que cada uma seja avaliada separadamente, enquanto reconhece as dependências.
- **Repetível**
A estrutura pode ser interpretada de forma consistente, suportando resultados repetíveis para comparar uma organização com outras do setor e acompanhar o progresso ao longo do tempo.
- **Suportado por uma organização independente e neutra**
O modelo deve ser neutro em relação ao fornecedor, a fim de evitar conflitos de interesse, e amplamente disponível para garantir uma ampla representação das melhores práticas.
- **Tecnologia neutra**
O foco do modelo deve estar nas práticas, e não nas ferramentas.
- **Suporte de treinamento incluído**
O modelo é suportado por um treinamento abrangente para permitir que os profissionais dominem a estrutura e otimizem seu uso.



3.9. Uso da estrutura DAMA-DMBOK

O DAMA-DMBOK pode ser usado para preparar ou estabelecer critérios para um DMMA.

Os responsáveis pela execução identificarão ligação direta entre funções segmentadas (as Áreas de Conhecimento) e as tarefas correspondentes (atividades).

As áreas de conhecimento, atividades e produtos (produtos de trabalho) do DMBOK podem ser configuradas para uma estrutura específica do DMM com base nas áreas medidas, suas atividades de suporte, relevância e tempo disponível. Essa abordagem rápida, em conjunto com a lista de verificação, pode ser usada para determinar áreas que precisam de análise mais profunda, que representam lacunas ou que indicam pontos de acesso para correção.

O DMBOK oferece vantagem adicional como ferramenta de planejamento de avaliação: Há uma grande comunidade de profissionais do conhecimento usando o DMBOK como um guia em vários setores, criando uma comunidade de prática em torno de seu uso.

4. Ética na Gestão de Dados

Definida de maneira simples, a ética é um princípio de comportamento baseado em ideias de certo e errado.

Os princípios éticos geralmente se concentram em ideias como justiça, respeito, responsabilidade, integridade, justiça, qualidade, confiabilidade, transparência e confiança. A ética de manipulação de dados preocupa-se em como adquirir, armazenar, gerenciar, usar e descartar dados de maneira alinhada aos princípios éticos. Em outras palavras, eles se preocupam em fazer as coisas certas com os dados e impedir que coisas erradas sejam feitas com os dados, mesmo quando ninguém está olhando.

Manipular dados - não apenas gerenciá-los, mas usá-los e compartilhá-los com outras entidades - de maneira ética é necessário para o sucesso a longo prazo de qualquer organização que deseja obter valor de seus dados.

O tratamento antiético de dados pode resultar em perda de reputação e clientes, pois coloca em risco as pessoas cujos dados estão expostos. Em alguns casos, práticas antiéticas também são ilegais. Dada a conexão entre o direito à privacidade e outros direitos humanos, a ética dos dados também é uma questão de responsabilidade social.

O tratamento ético dos dados, abrangerá:



DESTAQUE

- O porquê é importante gerenciar dados eticamente.
- Os princípios subjacentes ao tratamento ético de dados.
- Os benefícios de adotar uma abordagem ética para o gerenciamento de dados.
- O como estabelecer uma abordagem ética para o gerenciamento de dados.

4.1. Ética e gerenciamento de dados

A ética relacionada ao tema “manipulação de dados” apresenta alguns conceitos básicos. São eles:

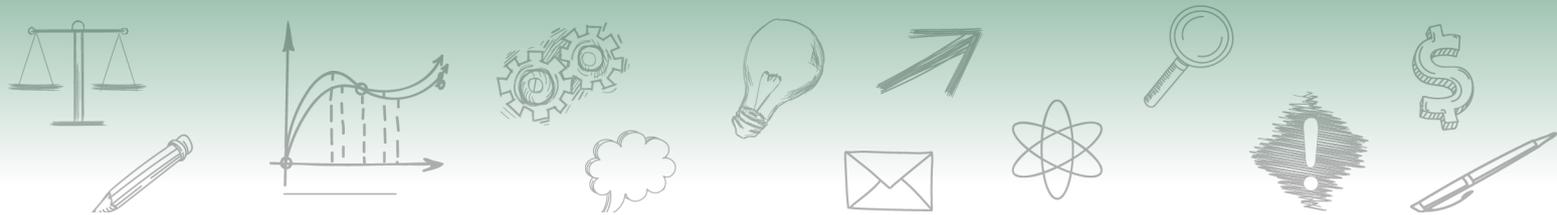
- **Impacto nas pessoas:** Os dados geralmente representam características de pessoas (Exemplos servidores, pacientes, fornecedores etc.) e são usados para tomar decisões que afetam a vida das pessoas. A ética exige que os dados sejam usados apenas de maneira a preservar a dignidade humana.
- **Potencial de uso indevido:** O uso indevido de dados pode afetar negativamente pessoas e organizações. Isso leva a um imperativo ético para impedir o uso indevido de dados, principalmente por meio de ações que prejudicam o bem maior.
- **Valor econômico dos dados:** os dados têm valor econômico. A ética da propriedade dos dados deve determinar como esse valor pode ser acessado e por quem.

As organizações protegem os dados com base principalmente em normas legais e regulamentos. No entanto, como os dados afetam as pessoas, os profissionais de gerenciamento de dados devem reconhecer que existem razões éticas (e legais) para proteger os dados e garantir que não sejam mal utilizados.

Mesmo dados que não representam diretamente indivíduos, como dados sobre acessibilidade ou distribuição de recursos, ainda podem ser usados para tomar decisões que afetam a vida das pessoas.

Há um imperativo ético não apenas para proteger os dados, mas também para gerenciar sua qualidade. As pessoas que tomam decisões, bem como as que são impactadas por elas, esperam que os dados sejam completos e precisos para que eles tenham uma base sólida para as decisões.

Do ponto de vista comercial e técnico, os profissionais de gerenciamento de dados têm uma responsabilidade ética de gerenciar os dados de uma maneira que reduz o risco de que eles possam deturpar, ser mal utilizados ou ser mal interpretados. Essa responsabilidade se estende ao longo do ciclo de vida dos dados, da criação à destruição dos dados.



Infelizmente, muitas organizações falham em reconhecer e responder às obrigações éticas inerentes ao gerenciamento de dados. Elas podem adotar uma perspectiva técnica tradicional e declarar não entender os dados. Ou elas assumem que, se seguirem a letra da lei, não terão riscos relacionados ao tratamento de dados. Esta é uma suposição perigosa.

O ambiente de dados está evoluindo rapidamente. As organizações estão usando os dados de maneiras que elas nem imaginariam há alguns anos atrás. O *Google Analytics* pode aprender coisas com dados que diversas pessoas ainda não considerariam possíveis.

Embora as leis codifiquem alguns princípios éticos, a legislação não pode acompanhar os riscos associados à evolução do ambiente de dados. As organizações devem reconhecer e responder à sua obrigação ética de proteger os dados a eles confiados, promovendo e sustentando uma cultura que valorize o manejo ético das informações.

4.2. Princípios éticos subjacentes à regulamentação da privacidade

As políticas e normas legais tentam codificar o certo e o errado com base em princípios éticos, mas não é possível codificar todas as circunstâncias. Por exemplo, as leis de privacidade na União Europeia, Canadá e Estados Unidos mostram diferentes abordagens para codificar a ética dos dados. Esses princípios também podem fornecer uma estrutura para a política organizacional.

Destaca-se os princípios subjacentes ao Regulamento Geral de Proteção de Dados (RGPD) da União Europeia incluem:

- **Justiça, legalidade, transparência:** os dados pessoais devem ser processados de maneira legal, justa e transparente em relação ao titular dos dados.
- **Limitação de finalidade:** Os dados pessoais devem ser coletados para finalidades especificadas, explícitas e legítimas, e não devem ser processados de maneira incompatível com essas finalidades.
- **Minimização de dados:** os dados pessoais devem ser adequados, relevantes e limitados ao necessário em relação aos propósitos para os quais são processados.
- **Precisão:** os dados pessoais devem ser precisos e, quando necessário, atualizados. Todas as medidas razoáveis devem ser tomadas para garantir que os dados pessoais imprecisos... sejam apagados ou retificados sem demora.
- **Limitação de armazenamento:** Os dados devem ser mantidos de uma forma que permita a identificação dos titulares de dados [indivíduos] por não mais que o necessário para os fins para os quais os dados pessoais são processados.



- **Integridade e confidencialidade:** Os dados devem ser processados de maneira a garantir a segurança adequada dos dados pessoais, incluindo proteção contra processamento não autorizado ou ilegal e contra perda, destruição ou dano acidental, usando medidas técnicas ou organizacionais apropriadas.
- **Responsabilidade:** Os controladores de dados devem ser responsáveis por e ser capazes de demonstrar conformidade com esses princípios.

Os princípios do RGPD são equilibrados e suportam certos direitos qualificados que os indivíduos têm de seus dados, incluindo os direitos de acesso, retificação de dados imprecisos, portabilidade, o direito de se opor ao processamento de dados pessoais que podem causar danos ou problemas e apagamento.

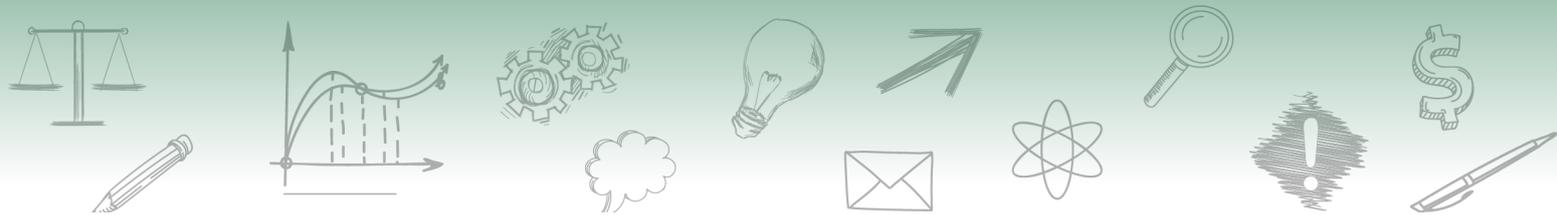
Quando os dados pessoais são processados com base no consentimento, esse consentimento deve ser uma ação afirmativa fornecida livremente, específica, informada e inequívoca. O RGPD exige governança e documentação eficazes para permitir e demonstrar conformidade e determina a Privacidade do Design.

A Lei de Proteção de Dados Pessoais e Documentos Eletrônicos (PIPEDA) é a lei federal canadense de privacidade que combina um regime abrangente de proteção da privacidade com a autorregulação do setor.

A PIPEDA (Lei de Proteção de Informações Pessoais e Documentos Eletrônicos) se aplica a toda organização que coleta, usa e divulga informações pessoais no decorrer de atividades comerciais. Ela estipula regras, com exceções, que as organizações devem seguir no uso das informações pessoais dos consumidores.

Ressalta-se ainda que as obrigações estatutárias, baseadas na PIPEDA, consideram :

- **Prestação de contas:** Uma organização é responsável pelas informações pessoais sob seu controle e deve designar um indivíduo para ser responsável pela conformidade da organização com o princípio.
- **Objetivos de identificação:** Uma organização deve identificar os propósitos para os quais as informações pessoais são coletadas antes ou no momento em que as informações são coletadas.
- **Consentimento:** Uma organização deve obter o conhecimento e o consentimento do indivíduo para a coleta, uso ou divulgação de informações pessoais, exceto quando inapropriado.
- **Limitando a coleta, o uso, a divulgação e a retenção:** A coleta de informações pessoais deve ser limitada àquela necessária para os fins identificados pela organização. As informações devem ser coletadas por meios justos e legais. As informações pessoais não devem ser usadas ou divulgadas para outros fins que não aqueles para os quais foram coletadas, exceto com o consentimento do indivíduo



ou conforme exigido por lei. As informações pessoais serão retidas apenas pelo tempo necessário para o cumprimento desses propósitos.

- **Precisão:** as informações pessoais devem ser precisas, completas e atualizadas conforme o necessário para os fins para os quais devem ser usadas.
- **Salvaguardas:** as informações pessoais devem ser protegidas por salvaguardas de segurança adequadas à sensibilidade das informações.
- **Abertura:** Uma organização deve disponibilizar prontamente às pessoas informações específicas sobre suas políticas e práticas relacionadas ao gerenciamento de suas informações pessoais.
- **Acesso Individual:** Mediante solicitação, um indivíduo deve ser informado da existência, uso e divulgação de suas informações pessoais e ter acesso a essas informações. Um indivíduo deve ser capaz de contestar a precisão e a integridade das informações e alterá-las conforme apropriado.
- **Desafios de conformidade:** Um indivíduo deve ser capaz de enfrentar um desafio relativo à conformidade com os princípios acima ao indivíduo ou indivíduos designados responsáveis pela conformidade da organização.

Saiba mais sobre proteção de dados

SAIBA MAIS

HIPAA (Lei de Portabilidade e Responsabilidade do Seguro de Saúde) nos EUA, PIPEDA (Lei de Proteção de Informações Pessoais e Documentos Eletrônicos) no Canadá, o Regulamento Geral de Proteção de Dados da UE (RGPD) e outras leis de proteção de dados / privacidade das informações descrevem obrigações em relação ao manuseio de dados de identificação pessoal (por exemplo, nome, endereço, afiliação religiosa ou orientação sexual) e privacidade (acesso ou restrição a essas informações).

Em março de 2012, a *Federal Trade Commission* (FTC) dos EUA emitiu relatório recomendando que as organizações projetassem e implementassem seus próprios programas de privacidade com base nas melhores práticas descritas no relatório (ou seja, Privacidade por Design).

Evidencia-se que esse relatório reafirma o foco da FTC nos Princípios Justos de Processamento de Informações. São eles:

- **Aviso/conscientização:** Os coletores de dados devem divulgar suas práticas de informações antes de coletar informações pessoais dos consumidores.



- **Escolha/Consentimento:** Os consumidores devem ter opções em relação a como as informações pessoais coletadas por eles podem ser usadas para fins além daqueles para os quais as informações foram fornecidas.
- **Acesso/Participação:** Os consumidores devem poder visualizar e contestar a precisão e a integridade dos dados coletados sobre eles.
- **Integridade/Segurança:** Os coletores de dados devem tomar medidas razoáveis para garantir que as informações coletadas dos consumidores sejam precisas e seguras contra uso não autorizado.
- **Execução/reparação:** O uso de um mecanismo confiável para impor sanções por não conformidade com essas práticas de informações justas.

Existe uma tendência global no sentido de aumentar a proteção legislativa da privacidade das informações dos indivíduos, seguindo os padrões estabelecidos pela legislação da União Europeia.

As leis em todo o mundo impõem diferentes tipos de restrições ao movimento de dados através das fronteiras internacionais. Mesmo dentro de uma organização multinacional, haverá limites legais para o compartilhamento de informações globalmente. Portanto, é importante que as organizações tenham políticas e diretrizes que permitam que a equipe cumpra os requisitos legais, além de usar dados dentro do apetite de risco da organização.

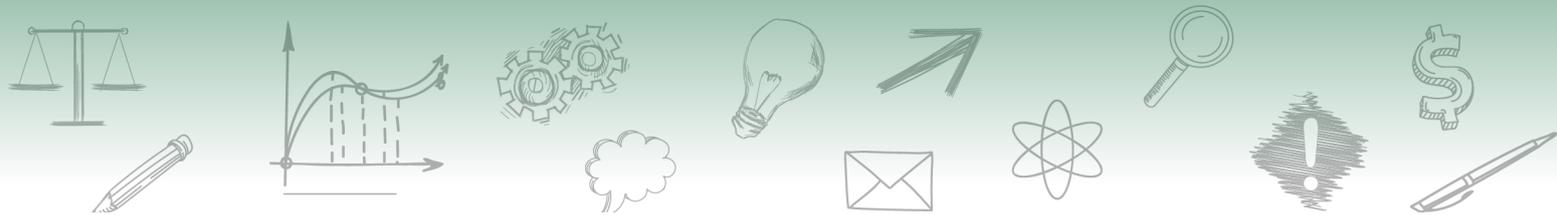
Acesse o Saiba Mais para conhecer os normativos brasileiros que regulamentam o tema em estudo.

SAIBA MAIS

Conforme estudado anteriormente, a Secretaria de Governo Digital da Secretaria Especial de Desburocratização, Gestão e Governo Digital do Ministério da Economia publicou no Diário Oficial da União (DOU) o Decreto nº 10.046, de 09 de outubro de 2019, que estabeleceu a Política de Compartilhamento de Dados no âmbito da Administração Pública Federal, instituiu o Cadastro Base do Cidadão (CBC) e o Comitê Central de Governança de Dados (CCGD).

O decreto é dirigido a todas as entidades da administração pública federal direta, autárquica e fundacional, além dos demais Poderes da União. O texto publicado no DOU traz alguns conceitos de atributos dos dados, como:

- **Atributos biográficos** - São os dados de pessoa natural relativos aos fatos da sua vida. Por exemplo: nome civil ou social, data de nascimento, filiação, naturalidade, nacionalidade, sexo, estado civil, grupo familiar, endereço e vínculos empregatícios;



- **Atributos biométricos** - São as características biológicas e comportamentais mensuráveis da pessoa natural. Podem ser coletadas para reconhecimento automatizado, tais como: a palma da mão; as digitais dos dedos; a retina ou a íris dos olhos; o formato da face; a voz e a maneira de andar;- **Atributos genéticos** - São as características hereditárias da pessoa natural, obtidas pela análise de ácidos nucleicos (DNA) ou por outras análises científicas. Além disso, traz o conceito de dados cadastrais, que são informações identificadoras perante os cadastros de órgãos públicos. São eles CPF, CNPJ, PIS, NIS, número de título de eleitor etc.

4.3. Ética e vantagem competitiva

As organizações estão cada vez mais reconhecendo que uma abordagem ética ao uso de dados é uma vantagem competitiva dos negócios. O tratamento ético dos dados pode aumentar a confiabilidade de uma organização e os resultados dos processos da organização. Isso pode criar melhores relacionamentos entre a organização e seus stakeholders. Criar uma cultura ética implica a introdução de uma governança adequada, incluindo a instituição de controles para garantir que os resultados pretendidos e resultantes do processamento de dados sejam éticos e não violem a confiança ou violem a dignidade humana.

A manipulação de dados não acontece no vácuo. Enfatizamos as quatro fortes razões comerciais para lidar com os dados de forma ética:

1. **Expectativas das partes interessadas:** os clientes e outras partes interessadas esperam comportamentos e resultados éticos das empresas e de seus processos de dados.
2. **Gerenciamento de risco:** Reduzir o risco de que os dados pelos quais a organização é responsável sejam utilizados indevidamente por usuários, clientes ou parceiros é o principal motivo para cultivar princípios éticos para o tratamento de dados.
3. **Prevenção de uso indevido:** Também existe uma responsabilidade ética de proteger dados de criminosos (ou seja, proteger contra hackers e possíveis vazamentos de dados).
4. **Respeitando a propriedade:** Diferentes modelos de propriedade de dados influenciam a ética do tratamento de dados. Por exemplo, a tecnologia melhorou a capacidade das organizações de compartilhar dados entre si. Essa capacidade significa que as organizações precisam tomar decisões éticas sobre sua responsabilidade pelo compartilhamento de dados que não lhes pertencem.

As funções emergentes de *Chief Data Officer*, *Chief Risk Officer*, *Chief Privacy Officer* e *Chief Analytics Officer* estão focadas no controle de riscos, estabelecendo práticas aceitáveis para o tratamento de dados. Mas a responsabilidade se estende além das pessoas nesses papéis. O tratamento de dados de forma ética requer o reconhecimento de toda a organização dos riscos



associados ao uso indevido de dados e o compromisso organizacional de lidar com dados com base em princípios que protegem os indivíduos e respeitam os imperativos relacionados à propriedade dos dados.

A governança de dados pode ajudar a garantir que os princípios éticos sejam seguidos para processos críticos, como decidir quem pode usar os dados e como eles podem usá-los. Os profissionais de governança de dados devem considerar os riscos éticos de certos usos de dados sobre as partes interessadas. Eles devem gerenciar esses riscos de maneira semelhante à maneira como gerenciam a qualidade dos dados.

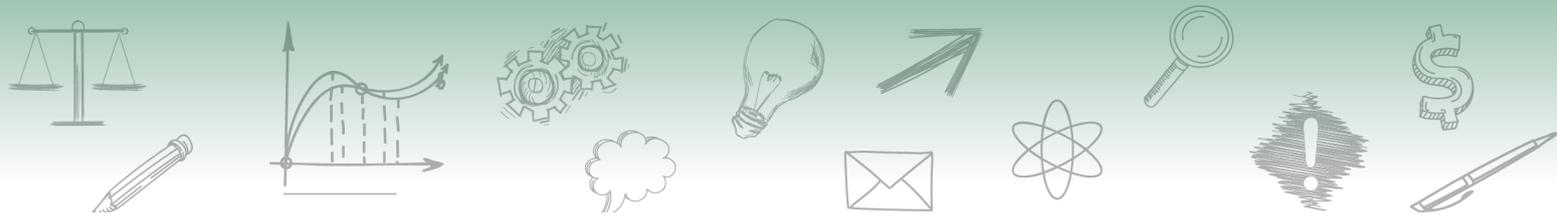
4.4. Estabelecer uma cultura de manipulação ética de dados

Estabelecer uma cultura de manipulação ética de dados requer entender as práticas existentes, definir comportamentos esperados, codificá-los nas políticas e um código de ética e fornecer treinamento e supervisão para aplicar os comportamentos esperados. Como em outras iniciativas relacionadas ao controle de dados e a mudança de cultura, esse processo exige uma liderança forte.

O tratamento ético dos dados obviamente inclui seguir a lei. Também influencia como os dados são analisados, interpretados e aproveitados interna e externamente. Uma cultura organizacional que valorize o comportamento ético não apenas terá um código de conduta, mas também garantirá a existência de controles claros de comunicação e governança para apoiar os usuários que tomam conhecimento de práticas ou riscos antiéticos. Os usuários precisam poder relatar essas circunstâncias sem medo de retaliação. Melhorar o comportamento ético de uma organização em relação aos dados geralmente requer um processo formal de Gerenciamento de Mudanças Organizacionais (*Organizational Change Management - OCM*).

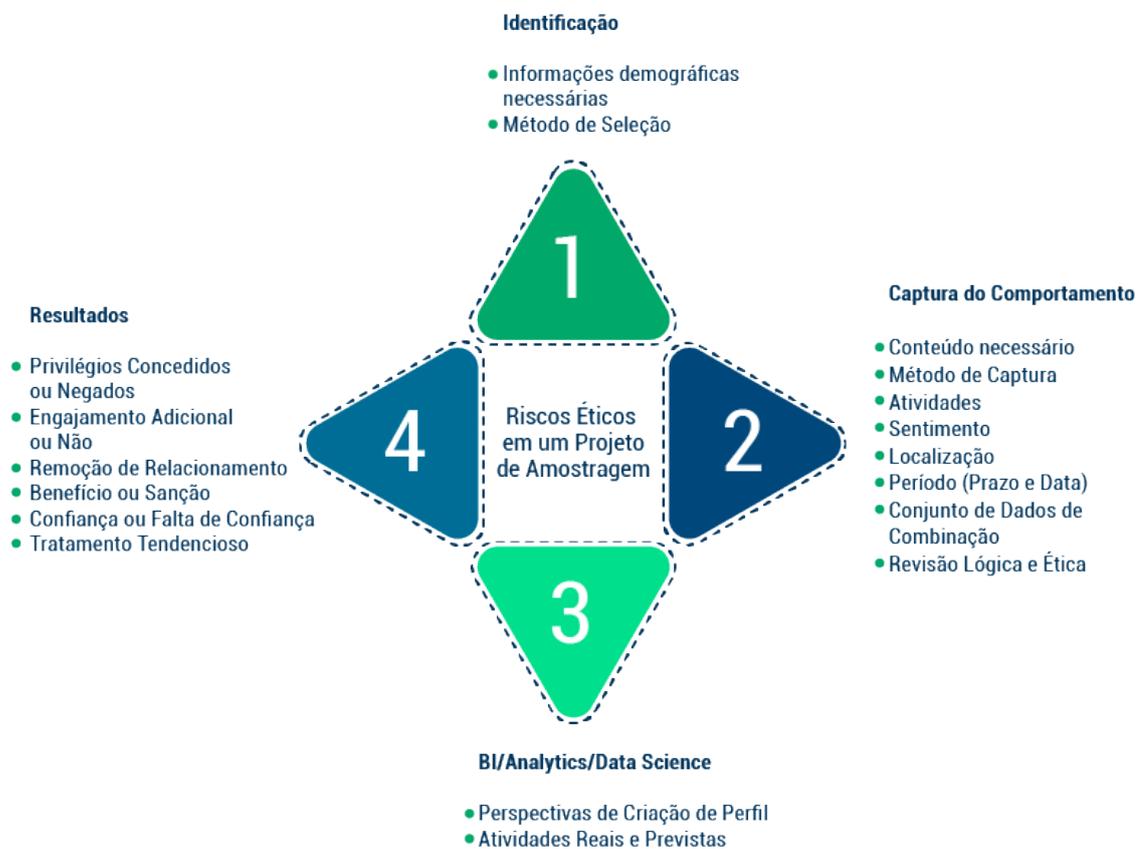
Atenção, para o estabelecimento de uma cultura de manipulação ética de dados, é necessário cumprir as seguintes etapas. Acesse cada uma das abas abaixo para conhecê-las.

- **Análise as práticas atuais de manipulação de dados do estado**
Examine até que ponto as práticas atuais estão direta e explicitamente conectadas aos fatores éticos e de conformidade; identificar até que ponto os usuários compreendem as implicações éticas das práticas existentes na construção e preservação da confiança de clientes, parceiros e outras partes interessadas.
- **Identifique princípios, práticas e fatores de risco**
Considere o risco de que os dados sejam mal utilizados e causem danos a usuários, fornecedores, outras partes interessadas ou a organização como um todo. Além dos riscos relacionados ao setor, a maioria das organizações possui riscos específicos, que podem estar relacionados ao seu know-how tecnológico, à taxa de rotatividade de usuários, aos meios pelos quais eles coletam dados de usuários ou a outros fatores. Os princípios devem estar alinhados aos riscos (prejuízos que podem acontecer se os princípios não forem respeitados) e práticas (a forma correta realizar atividades para que os riscos sejam evitados). As práticas devem ser apoiadas por controles.



- **Adote um modelo de risco ético socialmente responsável**

A execução de atividades de inteligência de negócios, análises e ciência de dados exige uma perspectiva ética que olhe além dos limites da organização e explique as implicações para a comunidade em geral. Uma perspectiva ética é necessária não apenas porque os dados podem ser facilmente mal utilizados, mas também porque as organizações têm uma responsabilidade social de não prejudicar seus dados. Um modelo de risco pode ser usado para determinar se um projeto deve ser executado, ele também influenciará como executar o projeto. Como os projetos de análise de dados são complexos, as pessoas podem não identificar os desafios éticos. As organizações precisam identificar ativamente riscos potenciais. Um modelo de risco pode ajudá-las a fazer o que a figura a seguir sugere.



Modelo de Avaliação de Risco para Gestão Ética de Dados (DMBOK2)

- **Crie uma estratégia e roteiro éticos para o tratamento de dados**

Após uma revisão do estado atual e o desenvolvimento de um conjunto de princípios, a organização pode formalizar uma estratégia para melhorar suas práticas de tratamento de dados. Essa estratégia deve refletir princípios éticos e o comportamento esperado relacionado aos dados, expresso em declarações de valores e código de comportamento ético.