



Enap

Governança de Dados

Módulo

4

Gerenciamento de Metadados e da qualidade 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 2021.



Enap, 2021

Enap Escola Nacional de Administração Pública

Diretoria de Educação Continuada

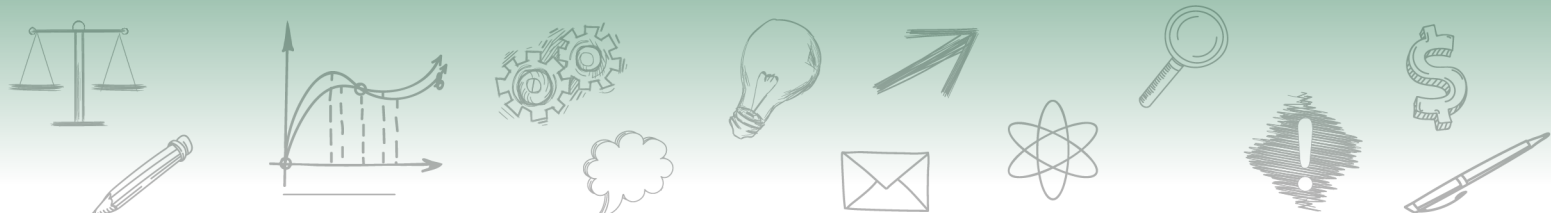
SAIS - Área 2-A - 70610-900 — Brasília, DF



Sumário

1. Gerenciamento de metadados	5
1.1 Metadados e seus Benefícios	5
1.2 Tipos de metadados	6
1.3 Metadados e Gerenciamento de Dados	7
1.4 Metadados e Interoperabilidade	9
1.5 Estratégia de Metadados	9
1.6 Requisitos de Metadados	10
1.7 Arquitetura de Metadados	11
1.8 Qualidade de Metadados	12
1.9 - Governança de Metadados	13
2. Gerenciamento da qualidade dos dados	13
2.1 Qualidade dos Dados	14
2.2 Dimensões da Qualidade dos Dados	14
2.3 Gerenciamento da Qualidade dos Dados	16
2.4 Qualidade dos Dados e outras Funções de Gerenciamento de Dados	18
2.5 Qualidade e regulação de dados	19
2.6 Ciclo de melhoria da qualidade dos dados	20
2.7 Qualidade dos Dados e Comprometimento da Liderança	22
2.8 Organização e mudança cultural	24
3. Maturidade em gerenciamento de dados	25
3.1 Avaliar o estado atual	25
3.2 Uso dos resultados para planejar melhorias	29
3.3 Gerenciamento de mudanças organizacionais	32
3.4 Governança da Gestão da Maturidade	34
3.5 Elementos do modelo de maturidade	35
Referências	38





Módulo

4 Gerenciamento de Metadados e da qualidade de Dados

1. Gerenciamento de metadados

Os metadados são essenciais para o gerenciamento de dados. O gerenciamento de metadados é uma atividade fundamental a ser realizada durante todo o ciclo de vida dos dados.

Os metadados no gerenciamento de dados equivale a um catálogo de uma biblioteca. Os leitores irão começar a procurar uma publicação específica a partir desse catálogo, que fornece as informações necessárias (código, sua localização e disponibilidade). Além disso, o próprio código possui outras informações como área temática, autor e ano de edição. Sem o catálogo, o processo de localização e identificação de uma publicação seria muito complicado, se não impossível. Com base nesta analogia, o gerenciamento de dados sem metadados equivale a uma biblioteca sem um catálogo.

À medida que a capacidade das organizações de coletar e armazenar dados se expande, a importância do uso dos metadados se amplia.

1.1 Metadados e seus Benefícios

Os metadados possuem informações sobre processos de negócio, regras de restrição de acesso e estruturas de dados lógicas e físicas. Os metadados auxiliam a organização a compreender seus dados e sistemas; avaliar a qualidade dos dados e quais aplicações estão vinculados.

Os metadados confiáveis auxiliam a:

- Aumentar a confiança nos dados, fornecendo contexto, permitindo uma representação consistente dos mesmos conceitos e permitindo a mensuração da qualidade dos dados.
- Identificar o valor das informações estratégicas (por exemplo, dados mestre).
- Ampliar a eficiência operacional e identificando dados e processos redundantes.
- Identificar a utilização de dados desatualizados ou incorretos.
- Fomentar a proteção de informações confidenciais.
- Melhorar a comunicação entre consumidores de dados e profissionais de TI.
- Criar uma análise de impacto precisa, reduzindo assim o risco de falha de projeto.



- Reduzir o tempo do ciclo de vida do desenvolvimento de sistema.
- Disponibilizar uma documentação completa do contexto, histórico e origem dos dados.
- Dar suporte à conformidade regulamentar.

Os metadados são críticos para a governança de dados, a qual fomenta o processo de qualidade de dados. Se os metadados forem gerenciados de forma inadequada, podem resultar em:

- Dicionários, repositórios e outros armazenamentos de metadados replicados e redundantes.
- Definições inconsistentes de elementos de dados e riscos associados ao uso indevido.
- Origens e versões concorrentes e conflitantes dos metadados que reduzem a confiança dos consumidores.

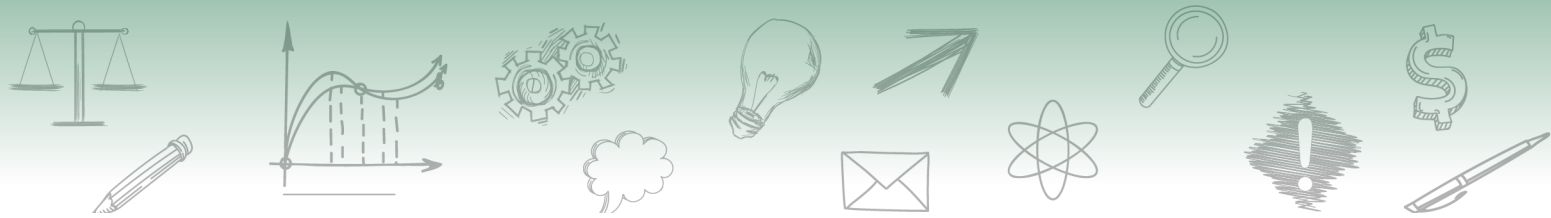
O gerenciamento eficiente de metadados permite uma compreensão consistente dos potenciais recursos da utilização dos dados.

1.2 Tipos de metadados

Os metadados são geralmente classificados em três tipos: estratégico, técnico ou operacional.

Os metadados estratégicos se concentram no conteúdo e na condição dos dados que incluem detalhes relacionados à governança de dados. Os metadados incluem nomes, definições, áreas temáticas, entidades e atributos; tipos de atributos e regras de negócios. Exemplos de metadados estratégicos:

- Modelos de dados, definições e descrições de conjuntos de dados, tabelas e colunas.
- Regras de negócios, regras de qualidade de dados e regras de transformação.
- Origem e linhagem.
- Padrões e restrições.
- Nível de segurança/privacidade.
- Problemas identificados.
- Notas de utilização.



Os metadados técnicos fornecem informações sobre os detalhes técnicos, os sistemas que o armazenam e os processos vinculados. Exemplos de metadados técnicos:

- Nomes e propriedades da tabela e da coluna do banco de dados físico.
- Direitos de acesso a dados, grupos, funções.
- Regras de CRUD de dados (criar, substituir, atualizar e excluir).
- Detalhes de processos de ETL.
- Documentação da linhagem, incluindo informações de impacto de alterações.
- Dependências do ciclo de atualização de conteúdo.

Metadados operacionais descrevem detalhes do processamento e acesso:

- Logs de execução de tarefas para programas em lote.
- Resultados de auditoria, balanceamento, indicadores de controle e registros de erros.
- Relatórios e padrões de acesso à consulta, frequência e tempo de execução.
- Patches e plano e execução de manutenção de versão, nível de patch atual.
- Provisões de backup, retenção, data de criação, recuperação de desastres.

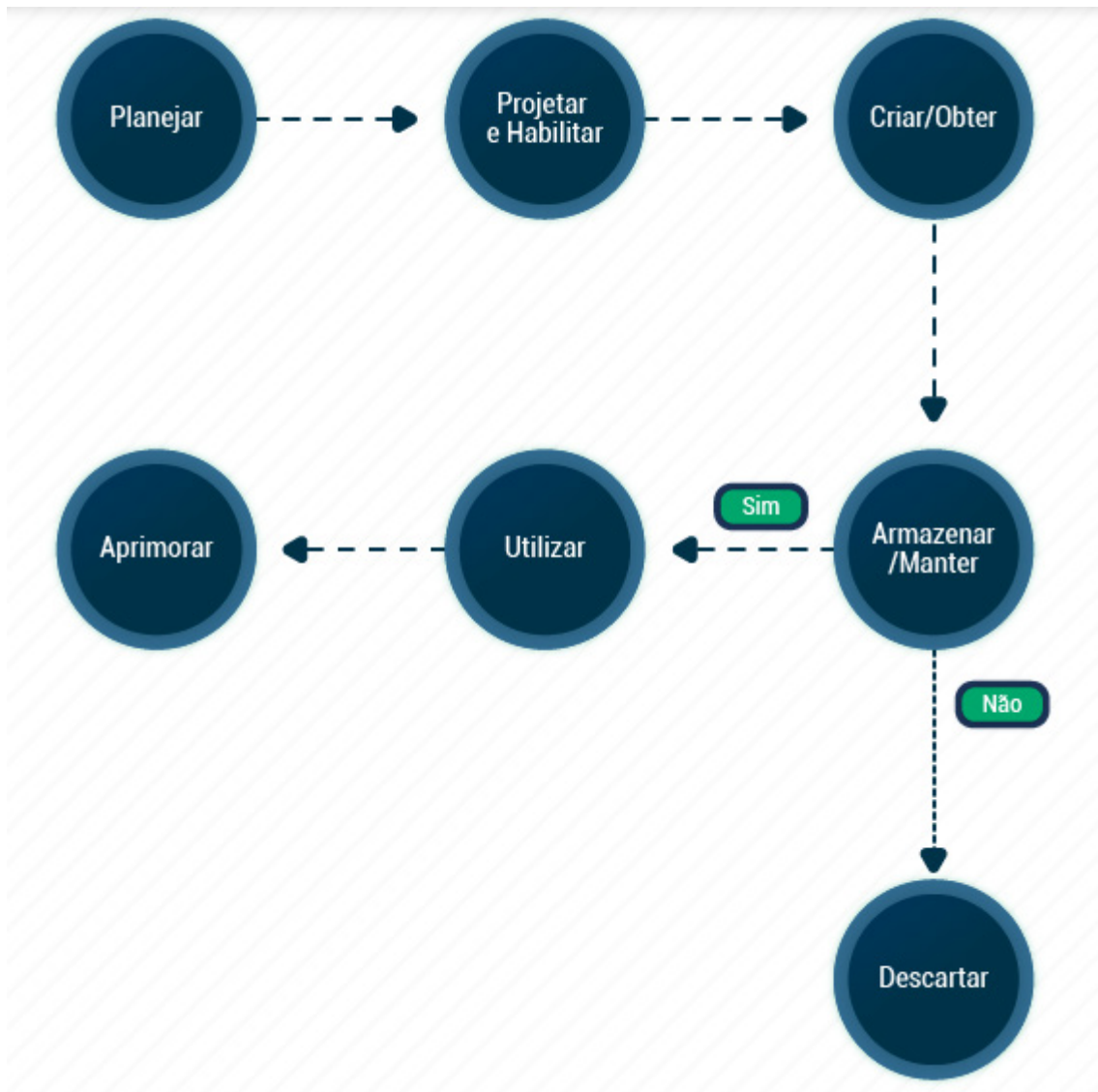
Essas categorias auxiliam na compreensão do conjunto de informações, bem como as funções que produzem metadados. Em relação ao uso, as distinções entre os tipos de metadados não são rígidas. Pode-se utilizar metadados técnicos e/ou operacionais, conforme o tipo de demanda a ser atendida.

1.3 Metadados e Gerenciamento de Dados

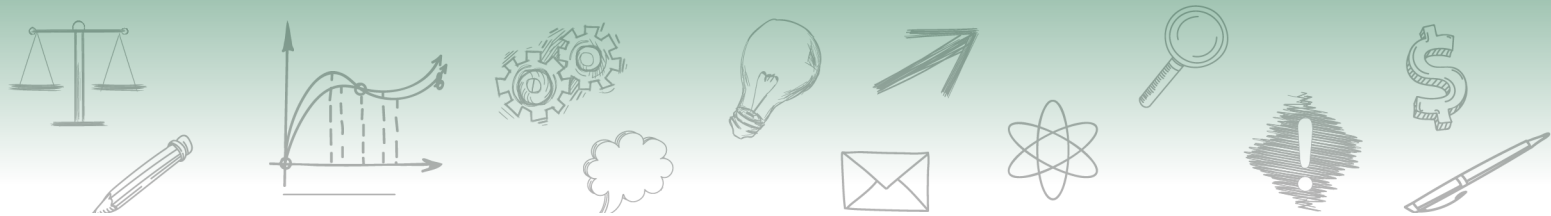
Os metadados fornecem o principal meio de coletar e gerenciar o conhecimento organizacional sobre dados.

O gerenciamento de metadados faz parte do gerenciamento de riscos. Os metadados são necessários para garantir que uma organização possa identificar quais dados são sensíveis ou confidenciais e, com isso, gerenciar o ciclo de vida, visando atender aos requisitos de conformidade e minimizar a exposição a riscos.

Sem metadados confiáveis, uma organização não consegue identificar quais dados são do seu domínio, o que eles representam, sua origem, como são interoperáveis entre sistemas, quem tem acesso e quais são qualificados.



- **PLANEJAR:** Definir requisitos de metadados.
- **PROJETAR e HABILITAR:** Identificar oportunidades para criar e gerenciar metadados como parte das atividades de gerenciamento de dados em andamento.
- **CRIAR / OBTER:** Verifique se os metadados são criados e atendem aos requisitos de qualidade.
- **ARMAZENAR / MANTER:** Verificar se os metadados permanecem atuais e continuam a atender aos requisitos.
- **DESCARTAR:** Limpar ou arquivar metadados obsoletos.
- **UTILIZAR:** Utilizar metadados para obter valor dos dados. Habilitar um *loop* de *feedback* para melhorar a qualidade dos metadados.



- **APRIMORAR:** Evolua os metadados existentes com novos conhecimentos. Implemente novos requisitos de metadados.

1.4 Metadados e Interoperabilidade

Os metadados tornaram-se essenciais para a maneira de como são interoperáveis e integrados. O Padrão de Registro de Metadados da ISO, ISO/IEC 11179, visa permitir a troca de dados orientada por metadados em um ambiente heterogêneo, com base em definições exatas de dados. Os metadados presentes no XML e em outros formatos permitem o uso dos dados. Outros tipos de marcação de metadados permitem que os dados sejam interoperáveis, mantendo os registros de propriedade, requisitos de segurança, etc.

SAIBA MAIS

Padrão de Registro de Metadados ISO/IEC 11179

O Padrão de Registro de Metadados da ISO, ISO/IEC 11179, fornece uma estrutura para definir um registro de Metadados. Este padrão foi projetado para permitir a troca de dados orientada por metadados, com base em definições exatas e com elementos de dados. O padrão está estruturado em seis partes:

Parte 1: Estrutura para geração e padronização de elementos de dados.

Parte 2: Fornece um modelo conceitual para gerenciar esquemas de classificação.

Parte 3: Atributos básicos dos elementos de dados.

Parte 4: Regras e diretrizes para a formulação de definições de dados.

Parte 5: Princípios de nomeação e identificação para elementos de dados.

Parte 6: Registro de elementos de dados.

Mais informações: <http://metadata-standards.org/11179/>.

1.5 Estratégia de Metadados

Uma estratégia de metadados descreve como uma organização pretende gerenciá-los e como ela passará do estado atual para as práticas avançadas e inovadoras. Uma estratégia de metadados deve fornecer uma estrutura para os profissionais responsáveis aprimorem o gerenciamento de metadados. O desenvolvimento de requisitos de metadados auxilia o nível estratégico a identificar possíveis obstáculos à sua implementação.

A estratégia inclui definir o conteúdo e a arquitetura de metadados da situação atual e a que pode ser alcançada pela organização, bem como as fases de implementação, que podem incluir:



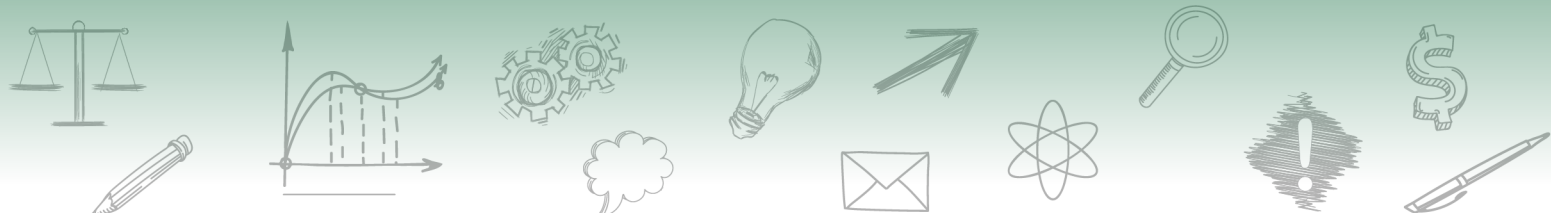
- **Iniciar o planejamento da estratégia de metadados:** A área estratégica deve definir as metas de curto e longo prazo. O planejamento inclui a elaboração do escopo, dos objetivos alinhados ao Programa de Governança de Dados e o estabelecimento de um plano de comunicação. As partes interessadas devem estar envolvidas no planejamento.
- **Realizar entrevistas com as partes interessadas:** as entrevistas com as partes interessadas das áreas técnicas e estratégica fornecem uma base de conhecimento para a estratégia de Metadados.
- **Avaliar fontes de metadados existentes e a arquitetura da informação:** A avaliação determina o grau de dificuldade na solução dos problemas relacionado à metadados.
- **Desenvolva um plano de implementação em fases:** Valide, integre e priorize os resultados das entrevistas e das análises. Documente a estratégia de metadados e defina uma abordagem de implementação em fases.

A estratégia evoluirá com o tempo, à medida que os requisitos de metadados, a arquitetura e o ciclo de vida dos metadados forem implementados e internalizados na organização.

1.6 Requisitos de Metadados

Os requisitos de metadados são amplos e serão obtidos por meio de coleta de informações dos consumidores de dados estratégicos a técnicos. As informações a serem coletadas poderão contemplar o seguinte roteiro:

- Com que frequência os atributos e conjuntos de metadados são atualizados.
- Período das atualizações em relação aos dados de origem.
- Se o histórico das atualizações dos metadados são mantidos.
- Quem pode acessar os metadados.
- Como os usuários acessam (funcionalidade específica da interface de acesso).
- Como os metadados são modelados e armazenados.
- O grau de integração de metadados de diferentes fontes; regras para integração.
- Processos e regras para atualização de metadados (registros).
- Funções e responsabilidades para gerenciar metadados.



- Requisitos de qualidade de metadados.
- Políticas de Segurança para gestão de metadados.

1.7 Arquitetura de Metadados

Conceitualmente todas as soluções de gerenciamento de metadados incluem camadas de arquiteturas que correspondem a pontos no ciclo de vida dos metadados:

- Geração de metadados.
- Armazenamento de metadados em um ou mais repositórios.
- Integração de metadados.
- Entrega de metadados.
- Acesso e uso de metadados.
- Controle e gerenciamento de metadados.

Um sistema de gerenciamento de metadados deve ser capaz de reunir metadados de várias fontes diferentes. Os sistemas diferem dependendo do grau de integração e do papel do sistema de integração na manutenção dos metadados.

Um ambiente de metadados gerenciado deve isolar o usuário final das várias e diferentes fontes de metadados. A arquitetura deve fornecer um único ponto de acesso, mediante adoção de um dos três modelos para implementação de um repositório comum de metadados:

- **Centralizada**
Uma arquitetura centralizada consiste em um único repositório que contém cópias dos metadados das várias fontes. As organizações que buscam um alto grau de consistência no repositório comum de metadados podem se beneficiar de uma arquitetura centralizada.
- **Distribuído**
Uma arquitetura distribuída mantém um único ponto de acesso. O mecanismo de recuperação de metadados responde às solicitações do usuário, recuperando dados dos sistemas de origem em tempo real; não há repositório persistente. Nessa arquitetura, o ambiente de gerenciamento de metadados mantém os catálogos do sistema de origem e as informações de pesquisa necessárias para processar efetivamente as consultas e pesquisas do usuário. Um broker de solicitação de objeto comum ou protocolo de middleware semelhante acessa esses sistemas de origem.



- **Híbrido**

Uma arquitetura híbrida combina características de arquiteturas centralizadas e distribuídas. Os metadados ainda transitam diretamente dos sistemas de origem para um repositório centralizado. No entanto, o design do repositório considera apenas os metadados adicionados pelo usuário, que podem ser itens padronizados e/ou críticos.

O conteúdo do repositório deve ser genérico no design. Não deve meramente refletir os projetos de banco de dados do sistema de origem. O planejamento deve considerar a integração de metadados para que os consumidores de dados tenham acesso a diferentes fontes de dados. Essa capacidade será um dos principais recursos estratégicos do repositório.

1.8 Qualidade de Metadados

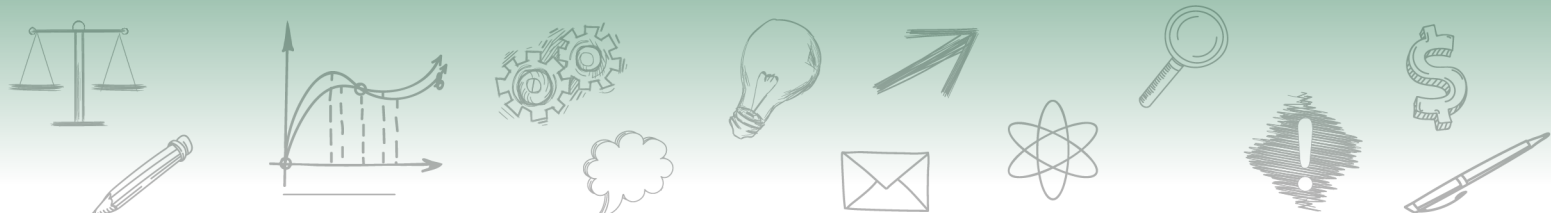
Obter metadados de alta qualidade decorre da mudança conceitual de que os metadados devem ser avaliados como um requisito do processo de geração de dados. As fontes de metadados, como o modelo de dados, a documentação de mapeamento de origem para o destino, os logs de ETL e similares devem ser tratados como fontes de dados. Deve-se implementar processos e controles para garantir que produzam dados confiáveis e utilizáveis.

Os processos, sistemas e dados precisam de algum nível de meta-informação; isto é, uma descrição de suas partes componentes e como elas funcionam. Além disso, à medida que o processo, sistema ou dados são usados, as meta-informações crescem e mudam. Com isso, o sistema precisa ser mantido e aprimorado. O uso de metadados geralmente resulta no reconhecimento de requisitos para metadados adicionais. Por exemplo, os consumidores que utilizam dados de usuários de dois sistemas diferentes podem precisar saber de onde os dados foram originados para entender melhor o perfil dos usuários.

Princípios gerais do gerenciamento de metadados descrevem os meios para gerenciar sua qualidade:

- **Responsabilidade:** Reconhecer que os metadados geralmente são produzidos por meio de processos existentes (modelagem de dados, SDLC, definição de processos de negócios) e responsabilizar os proprietários dos processos pela qualidade dos metadados (tanto na geração quanto na manutenção).
- **Padrões:** definir, impor e auditar os padrões dos metadados para simplificar a integração e permitir sua utilização.
- **Melhorar:** Criar mecanismos de feedback para que os consumidores possam informar a equipe de Gerenciamento de Metadados de que estão incorretos ou desatualizados.

Os metadados podem ser perfilados e inspecionados quanto à sua qualidade.



1.9 - Governança de Metadados

O gerenciamento de metadados deve ser prioridade nas organizações, no sentido de definir e coletar requisitos específicos para o gerenciamento do ciclo de vida de metadados críticos e estabelecer processos de governança. Recomenda-se atribuir funções e responsabilidades formais a recursos humanos capacitados e dedicados, especialmente em áreas críticas.

2. Gerenciamento da qualidade dos dados

O gerenciamento eficaz de dados envolve um conjunto de processos inter-relacionados que permitem a organização utilizar seus dados para atingir objetivos estratégicos. O gerenciamento de dados inclui a capacidade de projetar dados para desenvolver aplicações, armazená-los e acessá-los com segurança, compartilhá-los adequadamente e aprender com eles para atender aos objetivos estratégicos e operacionais. As organizações que almejam transformar seus dados em ativos devem priorizar a implementação de processos de melhoria na qualidade dos dados. Seguem alguns fatores que impedem ou prejudicam à adoção de processos com essa finalidade:

- Ausência de informações sobre os impactos decorrentes da utilização de dados de baixa qualidade.
- Planejamento ineficiente ou insuficiente.
- Projeto isolado de processos e sistemas (silos).
- Processos de desenvolvimento técnico inconsistentes.
- Documentação incompleta e metadados.
- Ausência de padrões e de práticas de governança.

Uma parcela das organizações não prioriza em definir o que torna os dados adequados ao objetivo estratégico e, conseqüentemente, possuem baixa maturidade em práticas de gestão da qualidade dos dados.

As organizações enfrentam desafios relacionados a implantação da gestão da qualidade de dados. A ausência de gestão pode gerar prejuízos de cunho financeiro e de confiabilidade.

A qualidade dos dados tornou-se uma prática estratégica. Demonstrando que a organização adota processos consolidados de qualidade e de segurança da informação, em conformidade com as normas e regulamentos em vigor. Além disso, amplia sua confiabilidade na gestão de dados junto as partes interessadas como a alta administração e os consumidores de dados.



2.1 Qualidade dos Dados

O termo qualidade dos dados é usado para se referir à implementação e gestão de boas práticas relacionadas aos processos utilizados para mensurar e melhorar essa qualidade.

Os dados são de alta qualidade na proporção em que atendam às expectativas e necessidades dos consumidores de dados. Quando esse propósito não for atendido ou parcialmente atendido, podemos avaliar como dados de baixa qualidade. A qualidade dos dados depende, portanto, do contexto e das necessidades dos consumidores de dados.

Um dos desafios no gerenciamento da qualidade dos dados é que as expectativas relacionadas à qualidade nem sempre são conhecidas ou mensuradas. Os profissionais que gerenciam dados nem sempre fazem levantamento sobre esses requisitos e, tampouco, inexistente uma avaliação consolidada de informações dos consumidores finais sobre a confiabilidade dos dados. O processo de avaliação da qualidade e das expectativas das partes interessadas precisam ser contínuos, visto que os requisitos sofrem mudanças com o tempo, à medida que as necessidades dos negócios e a influência de forças externas evoluem.

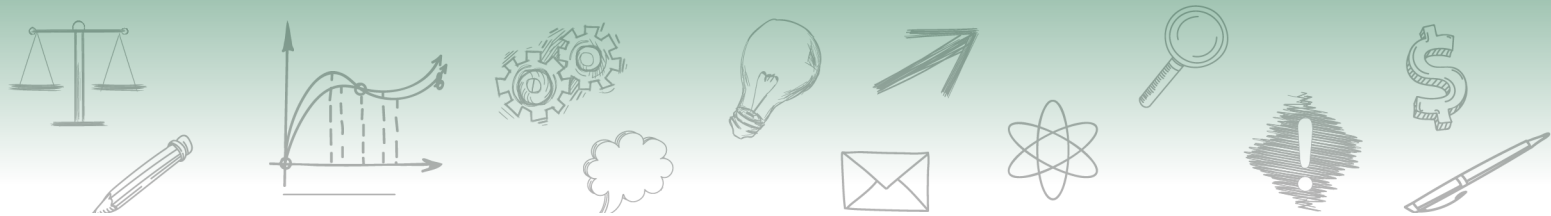
2.2 Dimensões da Qualidade dos Dados

Uma dimensão da qualidade de dados é um recurso ou característica mensurável dos dados. O termo dimensão é usado para fazer a conexão com as dimensões na mensuração de objetos físicos (por exemplo, comprimento, largura, altura). As dimensões da qualidade dos dados fornecem um vocabulário para definir os requisitos de qualidade dos dados. A partir daí, eles podem ser utilizados para definir os resultados da avaliação inicial da qualidade, bem como das medições em andamento. Para se medir a qualidade dos dados, uma organização deve estabelecer práticas mensuráveis e factíveis vinculadas aos processos de negócios.

As dimensões fornecem uma base para regras mensuráveis, as quais devem estar diretamente vinculadas aos riscos potenciais em processos críticos. Por exemplo:

- **Risco:** Se os dados “endereço de e-mail” do usuário estiverem incompletos, haverá prejuízo no processo de comunicação.
- **Meios de mitigação do risco:** Mensurar o percentual de usuários para os quais temos endereços de e-mail utilizáveis e aplicar as práticas para melhorar a qualidade, de forma que o resultado desse processo seja um endereço de e-mail confiável, utilizável para pelo menos 98% de usuários.

Muitos autores escreveram ou escrevem sobre as dimensões da qualidade dos dados. Embora não haja um único conjunto de dimensões de qualidade de dados consolidado, boa parte dos conjuntos contém ideias comuns. As dimensões incluem algumas características que podem ser medidas objetivamente (integridade, validade, conformidade) e outras que dependem muito do contexto ou da interpretação subjetiva (usabilidade, confiabilidade, relevância). Quaisquer



que sejam os nomes utilizados, as dimensões se concentram na existência de dados suficientes (integridade), se estão corretos (precisão, validade), são interoperáveis (consistência, integridade, exclusividade), são atualizados (tempestividade), acessível, utilizável e seguro.

Em 2013, o framework *Data Management Body of Knowledge* (DAMA DMBOK®) do Reino Unido produziu um documento técnico propondo seis dimensões principais da qualidade dos dados. O conjunto inclui:

- **Completeude:** a proporção de dados armazenados contra o potencial de 100%.
- **Exclusividade:** nenhuma instância de entidade será registrada mais de uma vez após identificada.
- **Oportunidade:** o grau em que os dados representam a realidade a partir do momento requerido.
- **Validade:** os dados são válidos se estiverem em conformidade com a sintaxe (formato, tipo, intervalo) relacionado à sua definição.
- **Precisão:** o grau em que os dados descrevem corretamente o objeto ou evento do 'mundo real' sendo descrito.
- **Consistência:** A ausência de diferença ao comparar duas ou mais representações de uma instancia com uma definição.

O documento oficial da DAMA UK (<https://www.dama-uk.org/>) também descreve outras características que têm impacto na qualidade.

- **Usabilidade:** os dados são compreensíveis, relevantes, acessíveis, mantidos no nível certo de precisão?
- **Problemas de tempo (além da própria tempestividade):** é estável, mas responde a solicitações de mudanças legítimas?
- **Flexibilidade:** os dados são comparáveis e compatíveis com outros dados? Possui agrupamentos e classificações úteis? Pode ser reaproveitado? É fácil de manipular?
- **Confiança:** a governança, a proteção e a segurança dos dados estão vigentes? Qual é a credibilidade dos dados? São verificáveis?
- **Valor:** existe um levantamento de custo/benefício para os dados? Serão utilizados da melhor maneira possível? Há riscos a segurança ou a privacidade das pessoas ou as responsabilidades legais da organização? São aderentes à imagem corporativa ou a visão estratégica corporativa?



A organização que almeja melhorar a qualidade de seus dados deve adotar ou desenvolver um conjunto de dimensões por meio das quais possa mensurar a qualidade e alcançar um consenso sobre as dimensões que podem fornecer um ponto de partida para a construção de um vocabulário comum sobre o tema.

2.3 Gerenciamento da Qualidade dos Dados

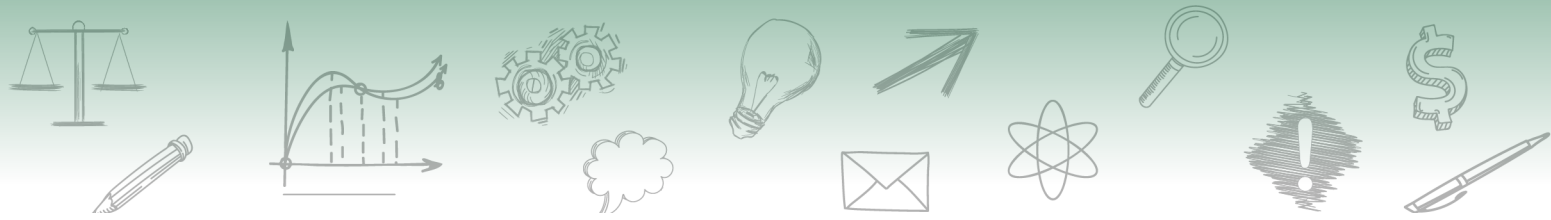
Como observado, o termo qualidade dos dados é usado para se referir aos processos usados para medir ou melhorar a qualidade dos dados, o qual representa o próprio gerenciamento da qualidade dos dados. Embora todas as funções de gerenciamento de dados tenham potencial para afetar a qualidade. O gerenciamento se concentra em auxiliar a organização a:

- **Definir dados de alta qualidade:** por meio de padrões, regras e requisitos de qualidade de dados.
- **Avaliar os dados:** em relação a esses padrões e comunicar resultados às partes interessadas.
- **Monitorar e relatar:** a qualidade dos dados em aplicações e armazenamentos de dados.
- **Identificar problemas:** e avaliar e defender oportunidades de melhoria.

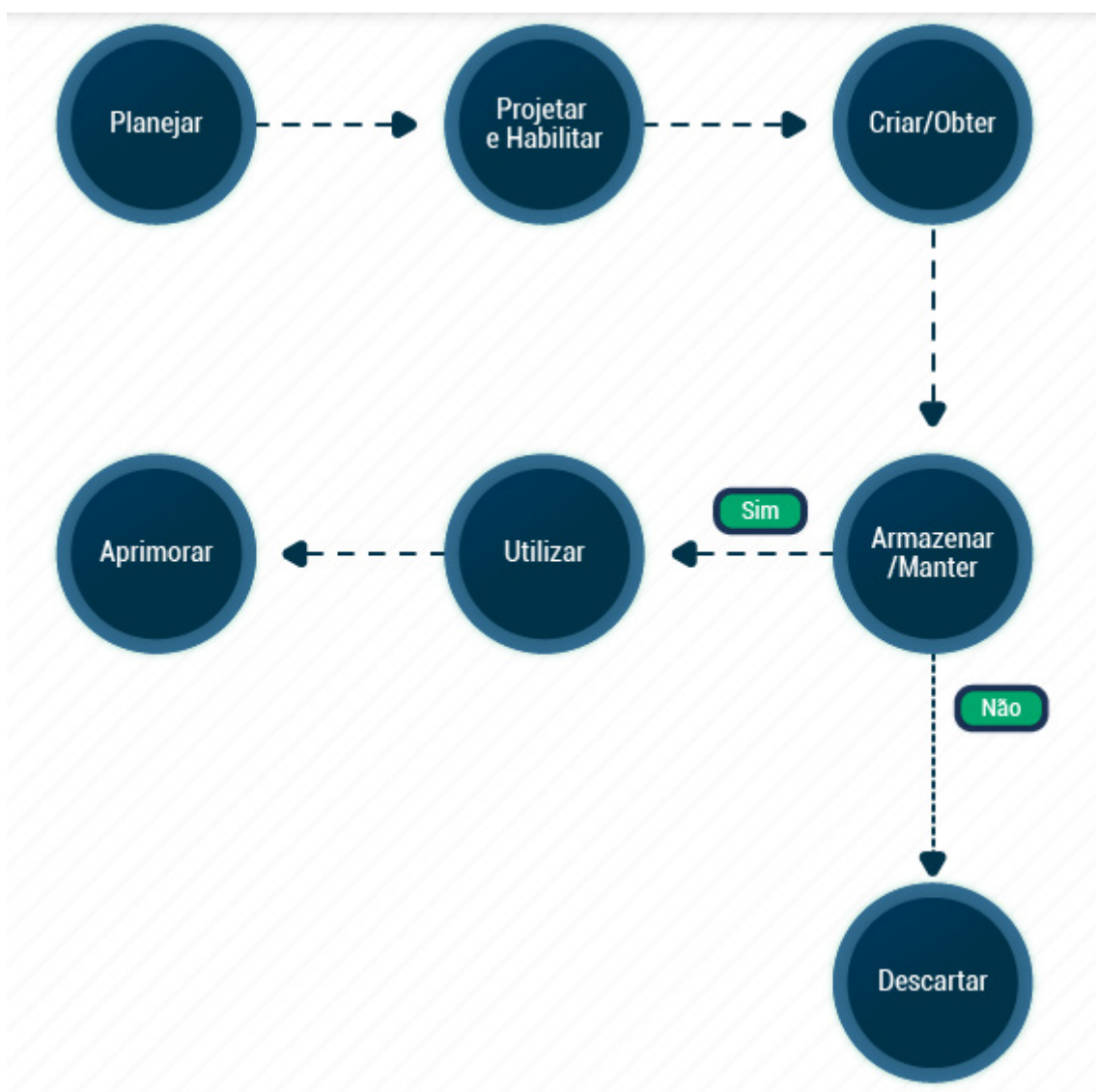
O gerenciamento da qualidade dos dados é semelhante ao gerenciamento contínuo da qualidade de outros serviços. Isso inclui o gerenciamento durante o seu ciclo de vida, definindo padrões, incorporando qualidade aos processos gerados, transformam e armazenam, mensura em relação aos padrões. Gerenciar dados para esse nível geralmente requer uma equipe de programa de qualidade de dados. Essa equipe é responsável por envolver os profissionais de gerenciamento de dados e representantes das áreas de negócio, bem como conduzir os trabalhos de aplicação das boas práticas de gerenciamento de qualidade aos dados, a fim de garantir que os dados sejam adequados ao consumo para uma variedade de propósitos.

A mesma equipe poderá estar envolvida em uma série de projetos através dos quais poderão estabelecer processos e práticas recomendadas, enquanto abordam problemas de dados de alta prioridade. Como o gerenciamento da qualidade dos dados envolve o gerenciamento do ciclo de vida dos dados, um programa de qualidade também terá responsabilidades operacionais relacionadas ao uso dos dados. Por exemplo, relatórios sobre os níveis de qualidade dos dados e participação na análise, quantificação e priorização de riscos relacionados a dados.

A equipe também é responsável por auxiliar aqueles que utilizam dados em suas atividades, para garantir que os dados atendam às suas necessidades e por auxiliar aqueles que geram, atualizam ou excluem dados. Essa responsabilidade alcança atividades de verificação dos procedimentos de manipulação dos dados.



O gerenciamento da qualidade dos dados é um programa, não um projeto. Deve incluir um plano de comunicação e de capacitação contínua. Mais importante, o sucesso a longo prazo do programa de melhoria da qualidade dos dados depende de uma organização internalizar uma mentalidade de qualidade de dados. Conforme declarado no *Manifesto de dados do líder*: mudanças fundamentais e duradouras exigem liderança e envolvimento comprometidos de pessoas em todos os níveis de uma organização. As pessoas que utilizam dados para realizar seu trabalho – o que significa um percentual muito grande – precisam promover mudanças. E uma das mudanças mais críticas a serem focadas é como as organizações gerenciam e melhoram a qualidade de seus dados.



- **PLANEJAR:** Definir características de dados de alta qualidade.
- **PROJETAR e HABILITAR:** Definir controle de sistema e de processo para evitar erros que prejudiquem o suporte a qualidade contínua dos dados.



- **CRIAR / OBTER:** Medir ou inspecionar os dados para garantir que atendam aos requisitos de qualidade.
- **ARMAZENAR / MANTER:** Monitorar a qualidade dos dados nos sistemas e dos processos para garantir que continuem a atender às expectativas.
- **DESCARTAR:** Identificar e limpar corretamente os dados com base nos requisitos.
- **UTILIZAR:** Habilitar um *loop* de *feedback* que identifique maneiras de melhorar a qualidade dos dados.
- **APRIMORAR:** Atuar nas oportunidades de melhoria da qualidade dos dados.

Os princípios de gerenciamento de dados descritos no *Data Management Body of Knowledge* (DAMA DMBOK®) afirmam que o gerenciamento de dados é o gerenciamento do ciclo de vida dos dados e que gerenciar dados significa gerenciar a qualidade dos dados. Durante o ciclo de vida dos dados, as atividades de gerenciamento da qualidade dos dados auxiliam uma organização a definir e medir as expectativas relacionadas aos seus dados. Essas expectativas podem mudar ao longo do tempo, conforme a utilização dos dados evoluem.

2.4 Qualidade dos Dados e outras Funções de Gerenciamento de Dados

As áreas do gerenciamento de dados têm o potencial de afetar a qualidade dos dados. Governança e gestão, modelagem de dados e gerenciamento de metadados têm efeitos diretos na geração de dados de alta qualidade. A execução ineficiente desses processos, torna-se inviável gerar dados confiáveis. Visto que para isso é necessário estabelecer padrões, definições e regras. A qualidade de dados visa atender às expectativas de forma coletiva descrevendo um conjunto de expectativas comuns.

A qualidade dos dados é baseada na premissa de que atende aos requisitos dos consumidores de dados. Ter um processo robusto pelo qual os dados são definidos suporta a capacidade de uma organização formalizar e documentar os padrões e requisitos pelos quais a qualidade dos dados pode ser medida.

Os metadados definem o que os dados representam. A gestão de dados e os processos de modelagem de dados são fontes de metadados críticos. Metadados bem gerenciados também podem apoiar o esforço para melhorar a qualidade dos dados. Um repositório de metadados pode abrigar resultados de medições de qualidade que podem ser compartilhados por toda a organização, de forma que a equipe de qualidade de dados possa ter insumos para produzir consenso sobre as prioridades e fatores de melhoria do Programa de Qualidade de Dados.

Um programa de qualidade de dados é mais eficaz quando está incluído em um programa de governança de dados. Visto que esse alinhamento visa identificar problemas de qualidade de dados e propor soluções a serem discutidas e avaliadas a luz da política de governança de dados.



A incorporação dos esforços de qualidade dos dados no escopo da governança permite que a equipe do programa de qualidade dos dados trabalhe com uma variedade de partes interessadas e facilitadores, que podem ser:

- Equipe de Gestão de Riscos e Segurança que pode apoiar na identificação de vulnerabilidades organizacionais relacionadas a dados.
- Equipe de engenharia e treinamento de processos de negócios que pode auxiliar as equipes a implementar melhorias de processos que aumentam a eficiência e resultam em dados mais adequados para utilização posterior.
- Gestores de dados podem identificar dados críticos, definir padrões e expectativas de qualidade e priorizar a correção de problemas.

Uma organização de governança pode acelerar o trabalho de um programa de qualidade de dados:

- Definindo as prioridades.
- Desenvolvendo e agindo na manutenção de padrões e políticas para a qualidade dos dados.
- Estabelecendo mecanismos de comunicação e compartilhamento de conhecimento.
- Monitoramento e gerando relatórios sobre desempenho e indicadores de qualidade de dados.
- Compartilhar resultados da inspeção da qualidade dos dados para conscientizar e identificar oportunidades de melhoria.

O programa de governança define responsabilidades no Gerenciamento de Dados Mestre e Gerenciamento de Dados de Referência. Vale ressaltar que o gerenciamento de dados mestre e o gerenciamento de dados de referência são exemplos de processos focados na curadoria de tipos específicos de dados com o objetivo de garantir sua qualidade. Definir um conjunto de dados de Dados Mestre implica em gerar expectativas sobre seu conteúdo e confiabilidade.

2.5 Qualidade e regulação de dados

Usuários e gestores de negócios necessitam de dados íntegros, precisos e legítimos. A qualidade dos dados pode ser um requisito regulatório. Os regulamentos diretamente relacionados às práticas de qualidade dos dados incluem os seguintes exemplos:

- *Sarbanes-Oxley* (EUA), que se concentra na precisão e validade das transações financeiras.



- *Solvência II* (UE), que se concentra na linhagem de dados e na qualidade dos dados subjacentes aos modelos de risco.
- O Regulamento Geral de Proteção de Dados (GDPR, UE) afirma que os dados pessoais devem ser precisos e, quando necessário, mantidos atualizados. Devem ser tomadas medidas razoáveis para apagar ou corrigir dados pessoais imprecisos.
- A Lei de Proteção de Informações Pessoais e Documentos Eletrônicos (PIPEDA, Canadá) afirma que os dados pessoais devem ser tão precisos, completos e atualizados para seus fins.
- A Lei Geral de Proteção de Dados Pessoais (LGPD ou LGPDP), Lei nº 13.709/2018, 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, que regula o uso da Internet no Brasil por meio da previsão de princípios, garantias, direitos e deveres para quem usa a rede, bem como da determinação de diretrizes para a atuação do Estado.

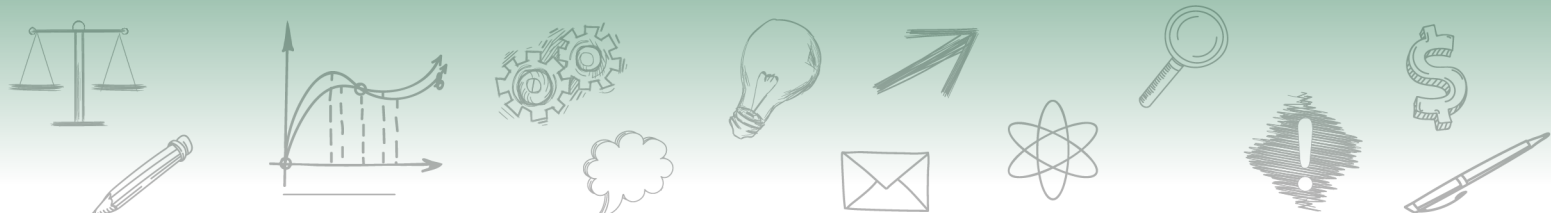
Vale ressaltar que, mesmo quando os requisitos de qualidade dos dados não são especificamente mencionados, a capacidade de proteger dados pessoais depende, em parte, da gestão da qualidade.

2.6 Ciclo de melhoria da qualidade dos dados

A maioria das abordagens para melhorar a qualidade dos dados se baseia nas técnicas de melhoria da qualidade na fabricação de produtos físicos. Nesse paradigma, os dados são entendidos como o produto de um conjunto de processos. Na sua forma mais simples, um processo é definido como uma série de etapas que transforma entradas em saídas. Um processo que gera dados pode consistir em uma etapa (coleta de dados) ou várias etapas: coleta de dados, integração em um *data warehouse*, agregação em um *data mart*, etc. Em qualquer etapa, os dados podem ser afetados negativamente. Pode ser coletado incorretamente, eliminado ou duplicado entre sistemas, alinhado ou agregado incorretamente, etc.

Melhorar a qualidade dos dados requer a capacidade de avaliar o relacionamento entre entradas e saídas para garantir que as entradas atendam aos requisitos do processo e as saídas estejam em conformidade com as expectativas. Como as saídas de um processo se tornam entradas para outros processos, os requisitos devem ser definidos ao longo de toda a cadeia de dados.

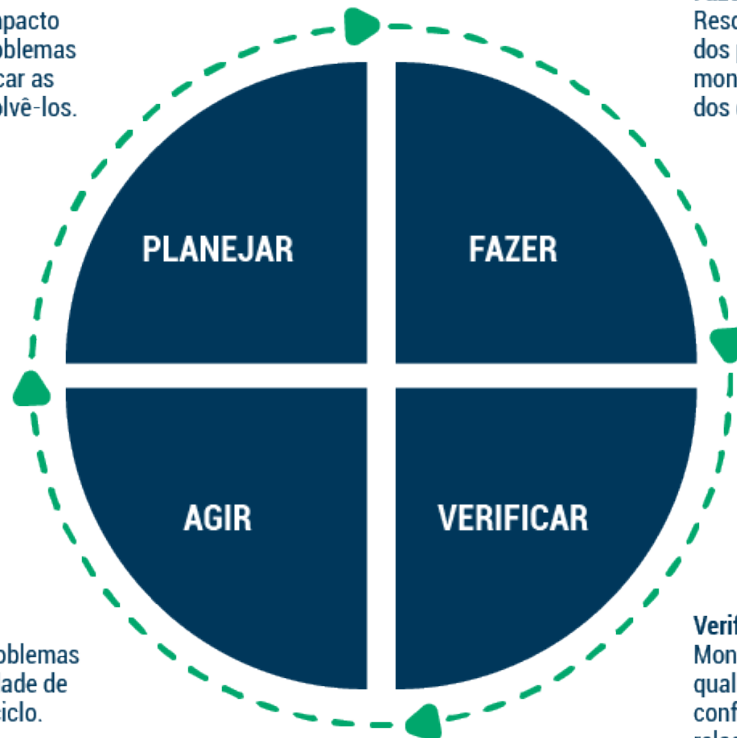
Uma abordagem geral para a melhoria da qualidade dos dados é uma versão do ciclo *Shewhart/Deming*. Com base no método científico, o ciclo de *Shewhart/Deming* é um modelo de solução de problemas conhecido como 'planejar-fazer-verificar-agir'. A melhoria ocorre através de um conjunto definido de etapas. A condição dos dados deve ser medida em relação aos padrões e, se não atender aos padrões, a(s) causa(s) raiz(es) da discrepância em relação aos padrões deve(m) ser identificada(s) e remediada(s). As causas-raízes podem ser encontradas em qualquer uma das etapas do processo, técnicas ou não técnicas. Uma vez remediados, os dados devem ser monitorados para garantir que continuem a atender aos requisitos.

**Planejar:**

Definir o escopo, o impacto e a prioridade dos problemas conhecidos e identificar as alternativas para resolvê-los.

Fazer:

Resolver as causas principais dos problemas e planejar o monitoramento contínuo dos dados.

**Agir:**

Abordar e resolver problemas emergentes de qualidade de dados e continuar o ciclo.

Verificar:

Monitorar ativamente a qualidade dos dados, conforme mensurado em relação aos requisitos.

Um Ciclo de Gerenciamento de Qualidade de Dados baseado no ciclo de Shewhart (Adptado do DMBOK2)

Para um determinado conjunto de dados, um ciclo de melhoria da qualidade começa identificando os dados que não atendem aos requisitos dos consumidores, que são problemas e geram obstáculos à consecução dos objetivos de negócios. Os dados precisam ser avaliados em relação às principais dimensões da qualidade e aos requisitos de negócio conhecidos. As causas dos problemas deverão ser identificadas para que as partes interessadas possam compreender os custos da correção e os riscos de não os remediar. Esse trabalho geralmente é realizado em conjunto com o Data Stewards e outras partes interessadas.

- **Planejar**

No estágio Planejar (Plan), a equipe de qualidade dos dados (QD) avalia o escopo, o impacto e a prioridade dos problemas conhecidos e avalia alternativas para resolvê-los. Esse plano deve ser baseado em uma base sólida de análise das causas dos problemas. A partir do conhecimento das causas e dos impactos, é possível avaliar o custo / benefício, determinar a prioridade e formular um plano para resolvê-los.

- **Fazer**

No estágio Fazer (Do), a equipe de qualidade de dados lidera os esforços para abordar as causas principais dos problemas e planejar o monitoramento contínuo dos dados. Para causas raízes baseadas em processos não técnicos, a equipe de Qualidade de Dados (QD) pode trabalhar com os responsáveis pelos processos para implementar mudanças. Para causas raiz se exigem alterações técnicas, a equipe de QD deve



trabalhar com equipes técnicas e garantir que os requisitos sejam implementados corretamente e que nenhum erro não intencional seja introduzido por alterações técnicas.

- **Verificar**

O estágio de Verificar (Check) envolve o monitoramento ativo da qualidade dos dados, conforme medido em relação aos requisitos. Desde que os dados atendam aos limites definidos para a qualidade, ações adicionais não são necessárias. Os processos serão considerados sob controle e atendendo aos requisitos de negócios. No entanto, se os dados estiverem abaixo dos limites de qualidade aceitáveis, ações adicionais deverão ser tomadas para elevá-lo a níveis aceitáveis.

- **Agir**

O estágio Agir (Act) são as atividades de abordagem e resolução de problemas emergentes de qualidade de dados. O ciclo é reiniciado, à medida que as causas dos problemas são avaliadas e as soluções propostas. A melhoria contínua é alcançada iniciando um novo ciclo. Novos ciclos começam como:

- ✓ As medidas existentes são insuficientes.
- ✓ Novos conjuntos de dados estão sob investigação.
- ✓ Novos requisitos de qualidade de dados emergem para conjuntos de dados existentes.
- ✓ Regras, padrões ou expectativas de negócios mudam.

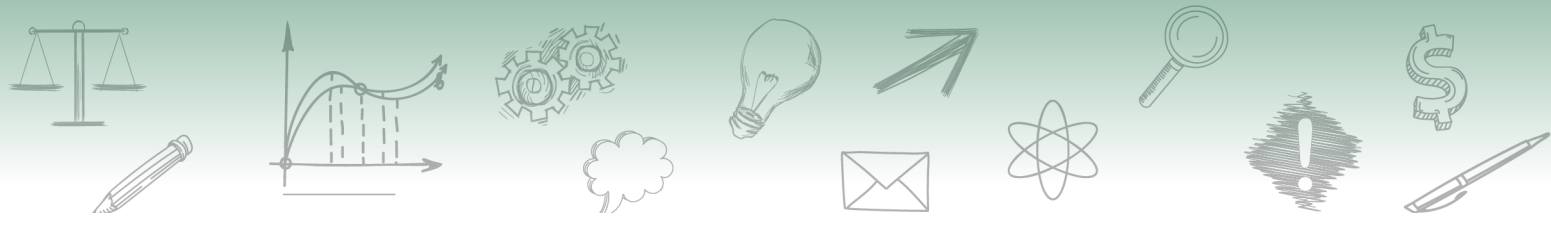
Estabelecer critérios para a qualidade dos dados no início de um processo ou construção do sistema é um sinal de uma organização com maturidade no gerenciamento de dados. Fazer isso exige governança e disciplina, além de colaboração multifuncional.

Integrar a qualidade nos processos de gerenciamento de dados desde o início custa menos do que modernizá-la. Manter dados de alta qualidade em todo o ciclo de vida dos dados é menos arriscado do que tentar melhorar a qualidade em um processo existente. Isso também cria um impacto muito menor na organização.

É melhor fazer as coisas corretamente da primeira vez. Mesmo se o fizerem, gerenciar a qualidade é um processo contínuo. As demandas de mudança e o crescimento orgânico ao longo do tempo podem causar problemas de qualidade dos dados gerando prejuízos.

2.7 Qualidade dos Dados e Comprometimento da Liderança

Os problemas de qualidade dos dados podem surgir em qualquer fase do ciclo de vida dos dados, desde a geração até o reuso ou descarte. Ao investigar as causas, os analistas devem identificar possíveis pontos de atenção, como problemas com entrada de dados, processamento, design de sistema e intervenção manual em processos automatizados. Os problemas poderão ter múltiplas causas e fatores contribuintes (como problemas mascarados e não resolvidos). Essas causas implicam em rever os processos e identificar práticas que possam ser aplicadas para evitar futuros problemas:



- Melhoria no design de interface.
- Teste de regras de qualidade de dados como parte do processamento.
- Foco na qualidade dos dados no design do sistema.
- Controles rígidos de intervenção manual em processos automatizados.

Obviamente, táticas preventivas devem ser usadas. No entanto, o senso comum diz e pesquisas indicam que muitos problemas de qualidade dos dados são causados pela falta de comprometimento organizacional, que por sua vez resultam da falta de liderança da governança e do gerenciamento.

Toda organização possui ativos de informação e dados que são estratégicos para suas operações. De fato, as operações dependem da capacidade de compartilhar informações. Apesar disso, poucas organizações gerenciam esses ativos com rigor.

Muitos programas de controle e ativos de informações são orientados apenas pela conformidade e não pelo valor potencial dos dados como um ativo. A falta de reconhecimento por parte da liderança significa uma falta de compromisso dentro de uma organização para gerenciar dados como um ativo, incluindo o gerenciamento de sua qualidade. As barreiras para o gerenciamento eficaz da qualidade dos dados incluem:

- Falta de conscientização por parte da liderança e da equipe.
- Falta de governança de negócios.
- Falta de liderança e gestão.
- Dificuldade em justificar melhorias.
- Instrumentos inadequados ou ineficazes para mensurar o valor.

Essas barreiras têm efeitos negativos na experiência do usuário, produtividade, eficácia organizacional e vantagem competitiva. Aumentam os custos e produzem riscos.

O reconhecimento dessas barreiras - as causas dos dados de baixa qualidade - fornece à organização uma visão de como melhorar sua qualidade. Motivos que reforcem a necessidade de a organização implementar uma governança atuante e consolidada.

O reconhecimento de um problema é o primeiro passo para resolvê-lo. Para isso, é necessário um processo de gerenciamento de mudança organizacional.

Barreiras ao gerenciamento de dados como um ativo estratégico. (DMBOK2, p. 467)

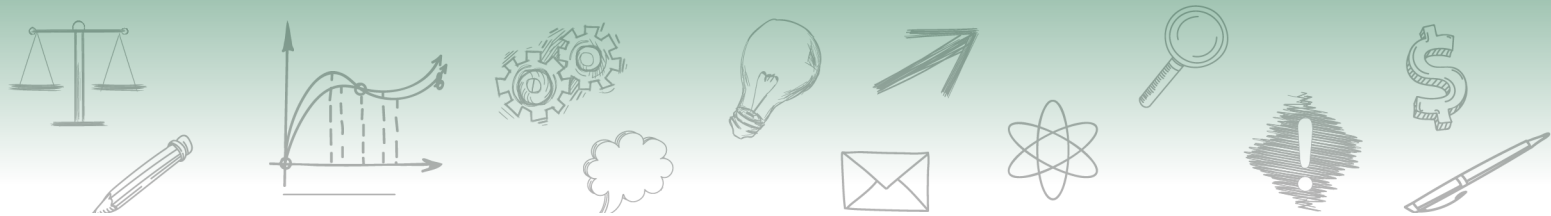


Causas mais comuns observadas - Barreiras que retardam / impedem as organizações de gerenciar suas informações como um ativo estratégico (DMBOK2)

2.8 Organização e mudança cultural

A qualidade dos dados não será aprimorada por meio de uma coleção de ferramentas e conceitos, mas por uma mentalidade que auxilia os consumidores e as partes interessadas a priorizarem a qualidade dos dados necessários para atender os objetivos estratégicos. A organização deve compreender que a qualidade dos dados requer uma mudança cultural significativa. Essa mudança requer visão e liderança.

O primeiro passo é promover a conscientização sobre o papel e a importância dos dados para a organização e definir as características dos dados de alta qualidade. Os consumidores de dados devem agir com responsabilidade e levantar problemas de qualidade. Esse processo de mudança deve ser conduzido por uma equipe de qualidade de dados em parceria com a área de governança com o apoio da área de TI.



3. Maturidade em gerenciamento de dados

Dados confiáveis não são produzidos por acidente. Dados gerenciados com eficiência dependem do planejamento, governança e comprometimento com a qualidade e segurança, bem como da execução disciplinada dos processos contínuos de gerenciamento de dados.

3.1 Avaliar o estado atual

Antes de definir uma nova gestão de dados ou tentar aprimorar a existente, é necessário avaliar o estado atual da gestão dos dados, Embora as especificidades e a condições para mudança sejam diferentes de organização para organização. A avaliação do estado atual foca na melhoria do gerenciamento de dados deve considerar:

- **O papel dos dados na organização:** quais processos são orientados por dados? Como os requisitos de dados são definidos e compreendidos? De que forma os dados impactam na estratégia organizacional? De que maneira a organização tem ciência dos custos decorrentes da gestão de dados de baixa qualidade?
- **Normas sobre dados:** Existem obstáculos normativos para se implementar ou melhorar as estruturas de gerenciamento e governança de dados? Os processos de negócios upstream estão cientes do uso downstream de seus dados?
- **Práticas de gerenciamento e governança de dados:** quem e como é executado as atividades relacionadas à gestão de dados? De que forma as decisões sobre o gerenciamento de dados são tomadas?
- **Como o trabalho é organizado e executado:** A execução das atividades é baseada em projetos? Quais estruturas de comitê existem para apoiar o esforço de gerenciamento de dados? Qual é o modelo operacional para interações de TI/ Negócios? Como os projetos são patrocinados?
- **Estrutura e Capital Humano:** a organização é centralizada ou descentralizada, hierárquica ou matricial? As equipes são colaborativas?
- **Níveis de habilidade:** qual é o nível de conhecimento de dados e de gerenciamento de dados das partes interessadas? Existem especialistas no assunto?

SAIBA MAIS

Upstream é a designação que se dá ao fluxo de dados do cliente para o servidor, enquanto **Downstream** é o fluxo de dados do servidor até o cliente.



A avaliação do estado atual também deve incluir o nível de satisfação. Isso fornecerá informações sobre as necessidades e prioridades de gerenciamento de dados da organização. Por exemplo:

- **Tomada de decisão:** a organização possui as informações necessárias para tomadas de decisões estratégicas, oportunas e sólidas?
- **Relatórios:** A organização confia em seus relatórios?
- **Principais indicadores de desempenho (KPIs):** Com que eficácia a organização rastreia seus KPIs?
- **Conformidade:** a organização está em conformidade com todas as normas legais relacionadas à governança de dados?

O meio mais eficaz para conduzir essa avaliação é usar um modelo confiável de maturidade em gerenciamento de dados que forneça informações sobre como a organização se compara a outras organizações; quais são as orientações para avançar do estado atual para outro nível.

Os modelos de maturidade definem cinco ou seis níveis de maturidade, cada um com suas próprias características, que variam de “maturidade inexistentes (nível 0)”, “ad hoc (nível 1)” a “otimizados ou de alto desempenho (nível 5)”.

Os níveis macro da escala de maturidade do gerenciamento de dados estão resumidos no destaque abaixo. Uma avaliação detalhada incluiria critérios para categorias amplas, como pessoas, processos e tecnologia; e para subcategorias como estratégia, política, padrões, definição de função, tecnologia/automação etc.

- **Nível 0**
Ausência de Capacidade: Não existem práticas organizadas ou processos corporativos formais para gerenciamento de dados. Há uma parcela considerável de organizações no nível 0. Esse nível é reconhecido apenas para fins de conceito.
- **Nível 1 Inicial/Ad Hoc**
Gerenciamento de dados utilizando um conjunto limitado de ferramentas, com pouca ou nenhuma governança. O tratamento dos dados é altamente dependente de especialistas. Papéis e responsabilidades são definidos em silos. Cada gestor de dados recebe, gera e envia dados de forma autônoma. Os controles, se existirem, são aplicados de forma inconsistente. As soluções para gerenciamento de dados são limitadas. Os problemas de qualidade dos dados são difundidos, mas não tratados e solucionados. O suporte de infraestrutura está no nível da unidade de negócio. Os critérios de avaliação podem incluir a presença de qualquer controle de processo, como registro de problemas de qualidade de dados.
- **Nível 2 Repetível**
Surgimento de ferramentas consistentes e definição de função para apoiar a execução do processo. No nível 2, a organização começa a usar ferramentas centralizadas e a



fornecer mais supervisão para o gerenciamento de dados. As funções são definidas e os processos não dependem apenas de um especialista. Existe uma conscientização organizacional sobre questões e conceitos de qualidade de dados. Os conceitos de gerenciamento de dados mestre e de referência começam a ser reconhecidos. Os critérios de avaliação podem incluir a definição formal de função em artefatos, como descrições de tarefas, a existência de documentação do processo e a capacidade de alavancar conjuntos de ferramentas.

- **Nível 3 Definido**

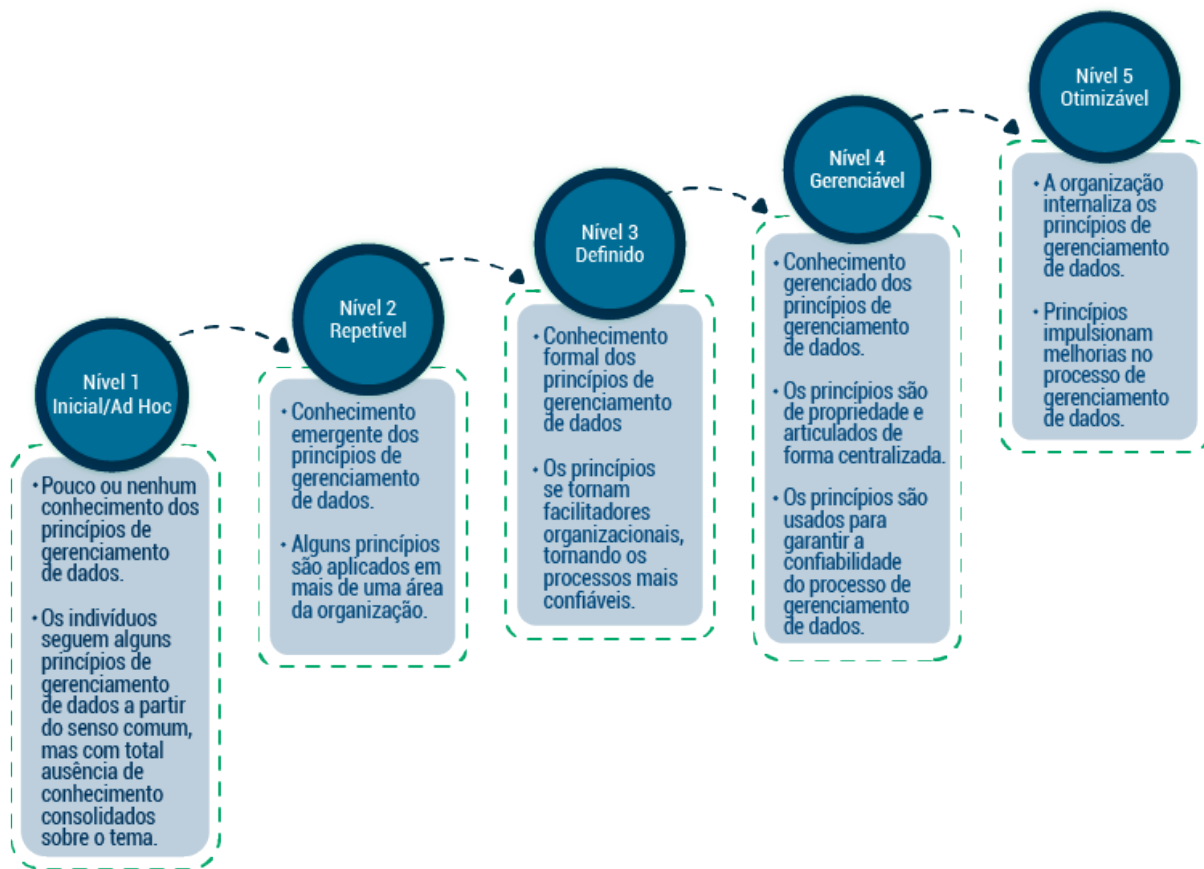
Capacidade emergente de gerenciamento de dados. O Nível 3 vê a introdução e a institucionalização de processos escaláveis e uma visão do gerenciamento de dados como um facilitador organizacional. As características incluem a replicação de dados em uma organização com alguns controles em vigor e um aumento na qualidade dos dados, juntamente com definição e gerenciamento coordenados de políticas. Uma definição de processo mais formal conduz a uma redução significativa na intervenção manual. Isso, junto com um processo de design centralizado, significa que os resultados do processo são mais previsíveis. Os critérios de avaliação podem incluir a existência de políticas de gerenciamento de dados, o uso de processos escalonáveis e a consistência de modelos de dados e controles do sistema.

- **Nível 4 Gerenciado**

O conhecimento institucional adquirido com o crescimento nos níveis 1 a 3 permite que a organização preveja resultados ao abordar novos projetos e tarefas e comece a gerenciar os riscos relacionados aos dados. O gerenciamento de dados inclui métricas de desempenho. As características do nível 4 incluem ferramentas padronizadas para gerenciamento de dados, das estações de trabalho à infraestrutura, juntamente com uma função centralizada de planejamento e governança. As expressões desse nível são um aumento mensurável na qualidade dos dados e nos recursos de toda a organização, como auditorias de dados de ponta a ponta. Os critérios de avaliação podem incluir métricas relacionadas ao sucesso do projeto, métricas operacionais para sistemas e métricas de qualidade dos dados.

- **Nível 5 Otimizado**

Quando as práticas de gerenciamento de dados são otimizadas, elas são altamente previsíveis, devido à automação de processos e gerenciamento de mudanças tecnológicas. As organizações nesse nível de maturidade se concentram na melhoria contínua. No nível 5, as ferramentas permitem uma visualização dos dados entre os processos. A proliferação de dados é controlada para evitar duplicação desnecessária. Métricas bem entendidas são usadas para gerenciar e medir a qualidade e os processos dos dados. Os critérios de avaliação podem incluir artefatos e métricas de gerenciamento de mudanças na melhoria do processo.



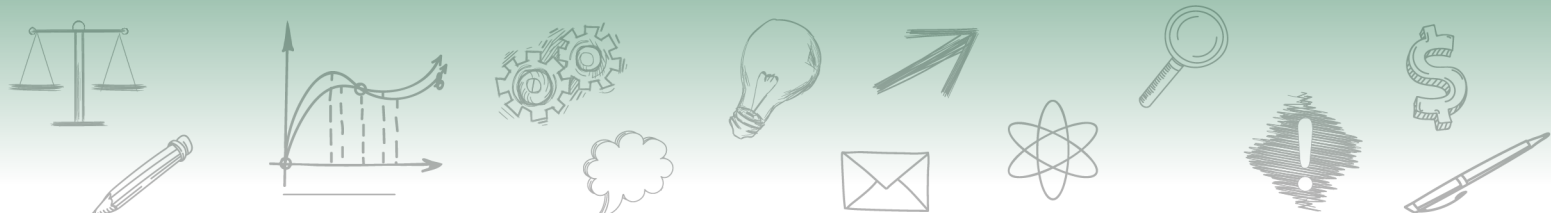
Fonte: Maturity Model for the Application of Data Management Principles (Adapted from DMBOK2)

O gráfico espiral ilustra uma maneira de apresentar um resumo das descobertas de uma Avaliação de Maturidade de Gerenciamento de Dados (*Data Management Maturity Assessment - DMMA*). Para cada um dos recursos (Governança, Arquitetura, etc.), o anel externo da teia mostra o nível de capacidade que a organização determinou que precisa evoluir. O anel interno exibe o nível de capacidade conforme determinado pela avaliação. As áreas em que a distância entre os dois anéis é maior representam os maiores riscos para a organização. Esse relatório pode ajudar a definir prioridades, como também, ser utilizado para medir o progresso ao longo do tempo.

O objetivo de uma avaliação do estado atual é compreender o ponto de partida da organização para planejar a melhoria. Uma avaliação precisa é mais importante do que uma pontuação alta. Uma avaliação formal da maturidade do gerenciamento de dados coloca a organização na escala de maturidade, demonstrando os pontos fortes e fracos de atividades críticas. Auxilia a organização a identificar, priorizar e implementar oportunidades de melhoria.

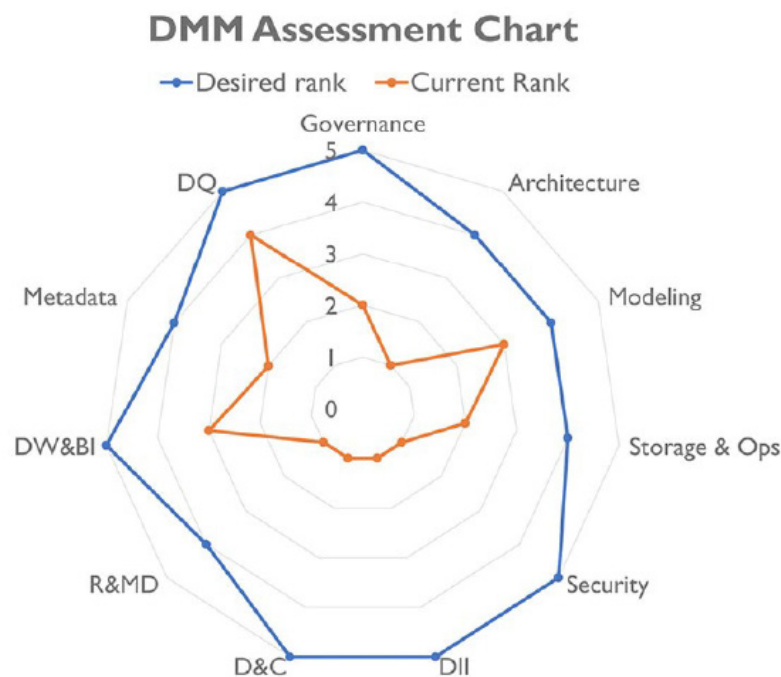
Ao atingir seu objetivo, um DMMA pode ter um impacto positivo na cultura organizacional. Isso auxilia a:

- Conscientizar as partes interessadas sobre os conceitos, princípios e práticas de gerenciamento de dados.



- Esclarecer as funções e responsabilidades das partes interessadas em relação aos dados organizacionais.
- Destacar a necessidade de gerenciar dados como um ativo crítico.
- Ampliar o reconhecimento das atividades de gerenciamento de dados em toda a organização.
- Contribuir para melhorar a efetividade da governança de dados.

Com base nos resultados da avaliação, uma organização pode aprimorar seu programa de gerenciamento de dados, de forma a apoiar as atividades estratégicas da organização. Normalmente, os programas de gerenciamento de dados são desenvolvidos em silos organizacionais. Raramente começam com uma visão corporativa dos dados. Um estudo de DMMA pode prover à organização meios para desenvolver uma visão coesa e estratégica; além de permitir uma avaliação clara e criteriosa de suas prioridades, redefinição de objetivos e desenvolvimento de um plano integrado de melhoria.



Exemplo de demonstração de resultado da aplicação da metodologia DMMA.

3.2 Uso dos resultados para planejar melhorias

Uma avaliação do estado atual auxilia e identificar o que está funcionando e o que não está, onde uma organização tem lacunas de eficiência. As conclusões fornecem a base para os objetivos do



programa de mapeamento de trilhas que auxiliam a determinar por onde começar e com qual velocidade avançar. Os objetivos devem se concentrar em:

- Oportunidades de melhoria de relevância relacionadas a processos, métodos, recursos e automação.
- Recursos alinhados à estratégia de negócios.
- Processos de governança para avaliação periódica do progresso organizacional com base nas características do modelo.

As especificidades dos planos de ação dependerão dos resultados da avaliação.

A tabela a seguir apresenta um modelo muito simplificado que considera apenas a adoção de uma metodologia padrão e o grau de automação do processo.

Por exemplo: Uma organização que reconheça a necessidade de melhorar a qualidade de seus dados. No entanto, sua avaliação atual do estado mostra que está no Nível 1. Ainda não estabeleceu práticas repetíveis em torno da mensuração da qualidade dos dados, mas consumidores testaram a qualidade dos dados e identificaram inconsistências. Com base neste evento, deve-se avaliar as soluções possíveis, alterar a estratégia de avaliação da maturidade e estabelecer metas para avançar do Nível 1 para o Nível 3 dentro de um prazo definido.

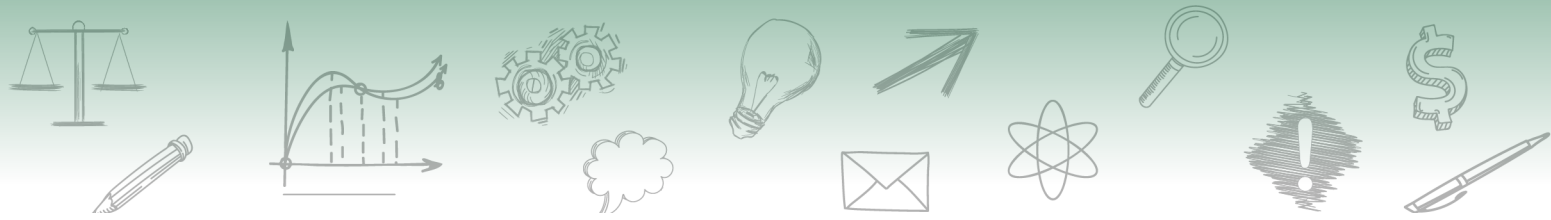
Atingir esse objetivo requer um plano de ação que detalhe cada fluxos de trabalho:

- Pesquisar métodos para se mensurar a qualidade dos dados e adoção de soluções para os pontos problemáticos da organização.
- Capacitar equipes sobre a metodologia implementada.
- Identificar e adotar ferramentas para apoiar a execução da metodologia.

A execução do plano de ação tem o objetivo de atingir as metas de correção dos problemas identificados, bem como definir as ações de evolução para o próximo nível.

As avaliações de maturidade podem ter diferentes áreas de foco. Se avaliar de forma abrangente as práticas de gerenciamento de dados, o resultado identificará diversas oportunidades de melhoria. Entretanto, será necessário priorizar as que terão impacto na estratégia do negócio.

Um modelo de maturidade de gerenciamento de dados inclui orientações internas, descrevendo como é o progresso dentro e entre as áreas de gerenciamento de dados. O processo de melhoria é realizado em etapas que podem ser adaptadas às necessidades e prioridades da organização.



Nível de Maturidade	Qualidade de Dados (QD) – Características de Medição	Grau de Automação
<p>Nível 5 – Otimizado: As metas de melhoria e de processo são quantificadas</p>	<p>Os relatórios de QD são amplamente compartilhados entre os acionistas da gerência. Os resultados da medição da qualidade dos dados são usados para identificar oportunidades de melhorias no sistema e nos processos de negócios e o impacto dessas melhorias é relatado.</p>	<p>Relatórios e Painéis otimizados e automatizados, incluindo alertas.</p>
<p>Nível 4 – Gerenciado: Processos são quantificados e controlados</p>	<p>Os gestores de sistemas e de negócios são obrigados a medir a qualidade de seus dados e a reportar resultados, para que os consumidores tenham um conhecimento consistente da qualidade dos dados.</p>	<p>O processo de medição é totalmente automatizado.</p>
<p>Nível 3 – Definido: Os padrões são definidos e usados</p>	<p>Os padrões são definidos para estabelecimento das métricas de qualidade dos dados e se estão sendo aplicados nas equipes.</p>	<p>Uma abordagem padrão foi adotada para iniciar o processo de automação.</p>
<p>Nível 2 – Repetível: Existe uma prática mínima repetível no processo</p>	<p>As partes envolvidas aprenderam práticas para medir a qualidade dos dados e estão desenvolvendo habilidades para replicar essas práticas.</p>	<p>Os processos ainda são amplamente manuais, mas algumas equipes testaram a automação.</p>
<p>Nível 1 - Inicial (Ad Hoc): O sucesso depende da competência dos indivíduos</p>	<p>Há um esforço dos indivíduos em medir a qualidade dos dados, mas isso não faz parte diretamente do rotina de atividades e não existe uma metodologia definida.</p>	<p>Nenhum. Os processos de acompanhamento dos indicadores são realizados de forma manual.</p>
<p>Nível 0 - Ausência de capacidade</p>	<p>Não existe definição de métricas e acompanhamento da qualidade dos dados.</p>	<p>Não aplicável.</p>



3.3 Gerenciamento de mudanças organizacionais

Uma parte das organizações que visam melhorar suas práticas de gerenciamento e governança de dados estão situadas na média dos indicadores da escala de maturidade, ou seja, entre 1 e 3. O que significa que quase todas necessitam melhorar suas práticas.

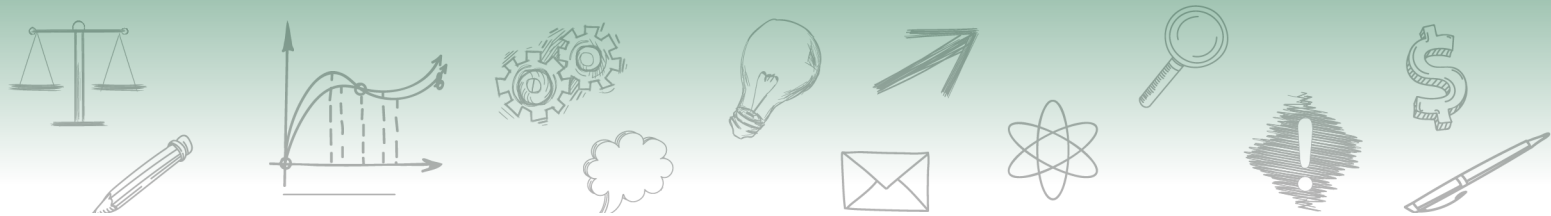
Melhorar as práticas de gerenciamento de dados requer alterar a maneira como os colaboradores desempenham suas funções de forma coletiva e como compreendem a importância da gestão dos dados; como utilizam os dados e implementam a tecnologia para dar suporte aos processos organizacionais. As práticas de gerenciamento de dados exigem, entre outros fatores:

- Aprender a gerenciar de forma horizontal, alinhando responsabilidades ao longo da cadeia de valor das informações.
- Mudar o foco da responsabilidade vertical (silo) para a gestão compartilhada de informações.
- Evoluir a qualidade da informação de uma expectativa do negócio de nicho para um valor estratégico de interesse da organização.
- Mudar o pensamento sobre a qualidade das informações de 'limpeza de dados e *scorecards*' para uma capacidade organizacional mais fundamentada em transformar a qualidade em processos.
- Implementar processos para mensurar o custo do gerenciamento de dados informal e o valor do gerenciamento de dados formalizado.

Esse nível de mudança não é alcançado por meio da tecnologia (mesmo que o uso apropriado de ferramentas de software possa dar suporte a execução e ao resultado). Em vez disso, é realizado através de uma abordagem cuidadosa e estruturada para o gerenciamento de mudanças na organização. A mudança será necessária em todos os níveis. É fundamental que seja gerenciado e coordenado a fim de evitar iniciativas sem saída, perda de confiança e danos à credibilidade da função de gerenciamento de informações e de sua liderança.

São fatores críticos para que as organizações possam iniciar um processo de mudança organizacional de forma eficiente, independentemente de sua estrutura:

- **Patrocínio da alta administração**
A alta administração deve internalizar e acreditar na iniciativa. Deve envolver-se efetivamente e mobilizar outros líderes no apoio às mudanças.
- **Visão clara**
Os líderes devem garantir que todas as partes interessadas afetadas pelo gerenciamento de dados - tanto internas quanto externas - compreendam e internalizem o que é gerenciamento de dados e a sua importância.



- **Gerenciamento proativo de mudanças**
O gerenciamento de mudança organizacional estabelece as práticas e os desafios para as partes interessadas, aumentando sua adesão e a sustentabilidade a longo prazo.
- **Alinhamento da liderança**
O alinhamento da liderança garante que haja acordo e suporte a necessidade de um programa de gerenciamento de dados. O alinhamento da liderança inclui o alinhamento entre as metas estratégicas e os resultados do gerenciamento de dados.
- **Comunicação**
A organização deve garantir que as partes interessadas sejam comunicadas e tenham a compreensão da importância do gerenciamento de dados e que impactos as transformações e mudanças afetaram o comportamento.
- **Engajamento das partes interessadas**
A iniciativa de mudanças organizacionais promovidas pela gestão de dados afeta de forma diferente os colaboradores e partes interessadas. O modo como a organização trata a condução dessas mudanças, terá um impacto significativo no sucesso da iniciativa.
- **Orientação e treinamento**
A capacitação é essencial para que o gerenciamento de dados aconteça. Diferentes grupos de pessoas (líderes, gestores de dados, equipes técnicas) exigirão diferentes tipos e níveis de capacitação. Essa capacitação deve contemplar novas políticas, processos, técnicas, procedimentos e ferramentas.
- **Avaliação**
Definir métricas para avaliar o desempenho das diretrizes de gerenciamento de dados visando manter a continuidade da execução do programa. O aspecto da inovação do gerenciamento de dados pode se concentrar na melhoria na tomada de decisão e na análise por meio de dados aprimorados e confiáveis.
- **Adesão aos princípios orientadores**
Princípios de gerenciamento de dados do DAMA servem como pontos de referência para tomada de decisões. Estabelecê-los antes da criação de um programa de gerenciamento de dados pode representar um ganho de efetividade.
- **Evolução, não revolução**
Em todos os aspectos do gerenciamento de dados, a filosofia de 'evolução não revolução' auxilia a minimizar os riscos e incertezas da prática de mudanças de alto impacto. O estabelecimento de uma metodologia de mudança que evolua a maturidade da organização a longo prazo, de forma incremental, permitirá que novas políticas e processos sejam adotados e que as mudanças comportamentais sejam sustentáveis.



3.4 Governança da Gestão da Maturidade

O Data Management Maturity Assessment (DMMA) corresponde a um conjunto de atividades relacionadas ao processo de mensuração da maturidade no contexto de governança de dados. O ciclo de vida de um DMMA consiste no planejamento e na avaliação inicial, seguido de recomendações, de um plano de ação e reavaliação periódica.

Supervisão do Processo DMMA

A supervisão do processo DMMA é exercida pela área de Governança de Dados. Na ausência de uma área formal de Governança de Dados ou ausência de know-how, os procedimentos de supervisão serão adotados pelo Comitê Diretor ou pela camada de gerenciamento que iniciou o DMMA. O processo deve ter um patrocinador da alta Administração para garantir que as melhorias nas atividades de gerenciamento de dados sejam mapeadas conforme os objetivos de negócio.

A amplitude e profundidade da supervisão dependem do escopo da DMMA. Cada função envolvida no processo tem sua responsabilidade na execução do método. A área de gerenciamento de dados estará envolvida de forma interdependente, utilizando uma linguagem comum.

Métricas

Além de ser um componente principal de qualquer estratégia de melhoria, as métricas são uma ferramenta de comunicação essencial. As métricas do DMMA são as categorizações que representam o estado atual do gerenciamento de dados. As métricas podem ser reavaliadas periodicamente para demonstrar tendências de melhoria. Cada organização deve desenvolver métricas adaptadas a sua realidade. Métricas podem incluir:

- **Classificações DMMA**
As classificações DMMA apresentam, por meio de níveis, a capacidade de maturidade da organização. As classificações podem ser acompanhadas de uma descrição.
- **Taxas de utilização de recursos**
Utilizar métricas que auxiliam a demonstrar o custo do gerenciamento de dados com base em cálculos percentuais.
- **Exposição ao risco ou a capacidade de responder a cenários de risco**
Expressa os recursos de uma organização em relação às classificações de DMMA.
- **Gerenciamento de gastos**
Expressa como o custo do gerenciamento de dados é alocado em uma organização e identifica os impactos desse custo na sustentabilidade e na agregação de valor. Essas métricas se sobrepõem às métricas de governança de dados.
 - ✓ Sustentabilidade no gerenciamento de dados
 - ✓ Realização das metas e objetivos



- ✓ Eficácia da comunicação
 - ✓ Eficácia da capacitação de pessoal
 - ✓ Velocidade de adoção das mudanças
 - ✓ Valor agregado com práticas de gerenciamento de dados
 - ✓ Contribuições para os objetivos de negócios
 - ✓ Reduções de riscos
- **Entradas para o DMMA**

São importantes para gerenciar, pois indicam a abrangência da cobertura, nível de investigação e detalhes do escopo relevante para a interpretação dos resultados. As entradas principais podem incluir: contagem, cobertura, disponibilidade, número de sistemas, volumes de dados, equipes envolvidas etc.
 - **Taxa de mudança**

A taxa na qual uma organização está evoluindo em sua capacidade. Uma linha de base é estabelecida por meio do DMMA. A reavaliação periódica é usada para avaliar a tendência.

3.5 Elementos do modelo de maturidade

Podemos utilizar os dados para obter vários tipos de resultados, conforme a necessidade. É necessário verificar se os resultados estão de acordo com os parâmetros da metodologia do modelo adotado. Caso contrário, pode gerar extrapolações equivocadas.

Proceder uma avaliação da maturidade dos dados para auxiliar na interpretação do cenário no qual está inserido. Para isso é necessário que os dados coletados sejam os mais realísticos possíveis. A avaliação da maturidade dos dados é uma ferramenta útil não só para compreender a situação atual da organização, mas para demonstrar o progresso das iniciativas de melhoria. Visto que essas informações serão utilizadas como insumos da linha de base. Muitas mudanças falham porque não houve tempo para compreender a real posição da organização. Uma lagarta pode ser uma lagarta perfeitamente boa, mas para atingir seu verdadeiro potencial, ela precisa se transformar em uma borboleta. Embora não estamos sugerindo que é possível saber tudo com antecedência, é possível em pouco tempo, compreender a posição da sua organização sobre os fatores importantes em relação à forma como trata os dados.

A linha de base pode demonstrar o que a organização tem realizado e em qual momento apresentou uma evolução. Com isso é possível determinar se existe retorno relevante para a organização e decidir se mantém ou não esse processo. É necessário apresentar os resultados da mudança e seus benefícios. Sem isso, pode passar a imagem de que a organização não está no caminho certo ou atingindo seus objetivos. As partes interessadas e os usuários podem consumir uma grande quantidade de energia para gerenciar e também podem inviabilizar completamente as iniciativas de mudança organizacional. Uma maneira de auxiliar a compreensão é demonstrar o progresso em relação a linha de base - que corresponde ao primeiro resultado da avaliação de maturidade.

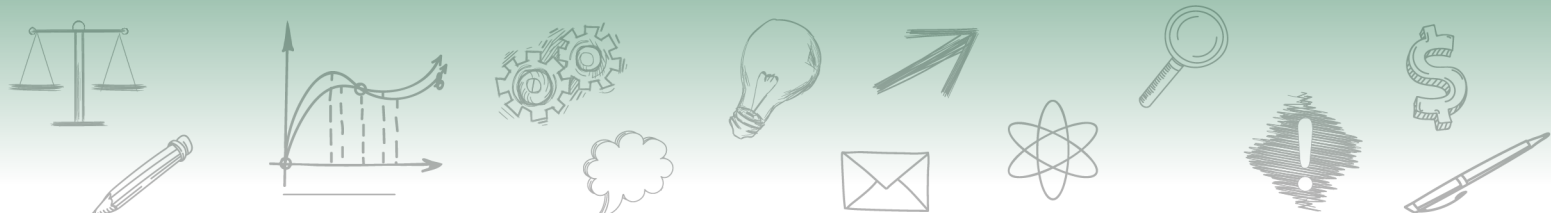


Alguns pontos relevantes ao se conduzir um modelo de maturidade:

- Toda interação é importante. O trabalho de preparação foi realizado para que haja alinhamento das motivações e expectativas.
- Nem todos serão defensores ou positivos sobre a proposta do modelo. É necessário avaliar se o esforço de convencimento é factível.
- Se forem positivos, adotar uma estratégia de mantê-los envolvidos para garantir um bom nível de aceitação.
- Aproveite as oportunidades para mudar a atitude das partes interessadas em relação aos dados.

Os principais elementos do modelo de maturidade são:

- Estratégia – Organização com níveis de maturidade altas comunicam a visão. A liderança segue uma estratégia de negócio bem delineada e fornece os princípios para elaboração de plano setoriais dos quais a estratégia de dados deve se alinhar.
- Governança corporativa – Os elementos principais da boa governança corporativa estão vigentes e estão bem implantados? Operam isoladamente ou adotam uma abordagem sincronizada? As atividades relevantes e adaptáveis são realizadas regularmente e de forma eficiente? A governança deve alinhar-se conformidade.
- Liderança e patrocínio – Existem pessoas-chaves nos níveis estratégicos da organização que compreendem a importância e o valor das iniciativas relacionadas a gestão de dados? Garantem apoio para que se negocie prazos para execução desse processo.
- Estrutura, processo e ferramentas – Sua organização possui estrutura coerente para fazer com que os demais setores se mantenham alinhados? Como as políticas são atualizadas? As ferramentas disponíveis auxiliam ou atrapalham? Quantos sistemas relacionados a dados estão em operação no momento? As pessoas utilizam esses sistemas de dados de forma correta? Compreendem o ciclo de vida das informações na sua organização?
- Políticas – Existem políticas, padrões, procedimentos para garantir que as pessoas da organização tenham um instrumento de alinhamento estratégico? Essas políticas, procedimentos e padrões são claros, consistentes e acessíveis? Existe uma estrutura para demonstrar a inter-relação entre essas políticas?
- Risco de informação – O risco das informações está bem definido e em que nível? Tendo em vista a importância do negócio. As ferramentas disponíveis para auxiliar o gerenciamento de dados são adequadas para mitigar os riscos de forma adequada?



- Arquitetura – Como as informações de toda a organização são estruturadas e como transitam entre os silos? As arquiteturas de dados e informações estão mapeadas? Quais são os responsáveis?
- Organização, funções e responsabilidade – As funções são bem delineadas e formalizadas em toda a organização? Existe uma estrutura dedicada pela gestão dos dados da organização? Foram definidas funções para abordar elementos de gerenciamento e de informações em seus diferentes domínios? Existe um conselho diretor com poderes para tomar decisões relacionadas a gestão de dados e informações?
- Habilidades – Quais habilidades são necessárias para alcançar os objetivos de gerenciamento de dados definidos pela organização? A organização dispõe de programas de treinamento para os profissionais de dados e de informações?
- Métricas – As funções de gerenciamento de dados são mensuradas do ponto de vista de desempenho e demonstrada por meio de relatórios que apontam os benefícios alcançados? O resultado do processo de mensuração é direcionado para as mudanças comportamentais pretendidos pela cultura de valorização de dados?
- Comportamento da tutela da informação – Valoriza-se os dados e o que pode ser realizado pelo seu uso? Os investimentos no gerenciamento de dados foram positivos a ponto de serem mensuráveis?
- Tecnologia – A tecnologia disponível é facilmente escalável e adequada para atender os processos de gerenciamento de dados?



Referências

Data Driven Business Transformation by Caroline Carruthers, Peter Jackson, Publisher: Wiley, Release Date: May 2019.

Navigating the Labyrinth: An Executive Guide to Data Management by Laura Sebastian-Coleman for DAMA International Publisher: Technics Publications Release Date: June 2018.

DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition) by DAMA International Publisher: Technics Publications, Release Date: July 2017.