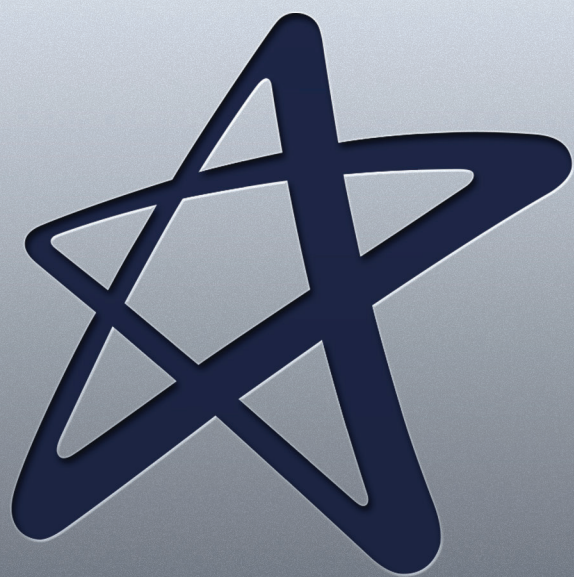


Pós-Graduação

Qualidade de Dados





Conceituação de Qualidade de Dados

Responsável pelo Conteúdo:

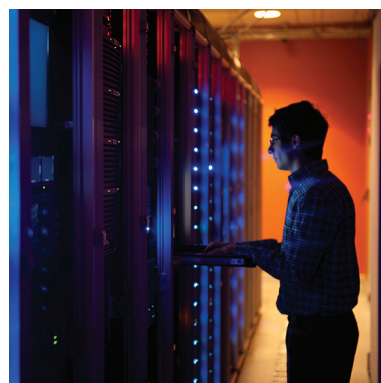
Prof. Dr. Artur Ubaldo Marques júnior

Revisão Textual:

Profa. Esp. Kelciane da Rocha Campos

Nesta unidade, trabalharemos os seguintes tópicos:

- Introdução ao Tema
- Orientações para Leitura Obrigatória
- Material Complementar



Fonte: iStock/Getty Images



Objetivos

- Conhecer o que é qualidade de dados e a importância em tê-la.

Caro Aluno(a)!!

Normalmente, com a correria do dia a dia, não nos organizamos e deixamos para o último momento o acesso ao estudo, o que implicará o não aprofundamento no material trabalhado ou, ainda, a perda dos prazos para o lançamento das atividades solicitadas.

Assim, organize seus estudos de maneira que entrem na sua rotina. Por exemplo, você poderá escolher um dia ao longo da semana ou um determinado horário todos ou alguns dias e determinar como o seu “momento do estudo”.

No material de cada Unidade, há videoaulas e leituras indicadas, assim como sugestões de materiais complementares, elementos didáticos que ampliarão sua interpretação e auxiliarão o pleno entendimento dos temas abordados.

Após o contato com o conteúdo proposto, participe dos debates mediados em fóruns de discussão, pois estes ajudarão a verificar o quanto você absorveu do conteúdo, além de propiciar o contato com seus colegas e tutores, o que se apresenta como rico espaço de troca de ideias e aprendizagem.

Bons Estudos!

Introdução ao Tema

Qualidade de dados tem ganhado importância crescente desde que me envolvi com bancos de dados em geral e com empresas de *contact Center* em particular, isso foi por volta de 2008.

Grandes nomes na época, como Richard Wang, Madnick, Peter Aiken, Larry English e John Talburt, consolidaram e colocaram as pedras fundamentais no que hoje conhecemos como Qualidade de Dados e deram uma boa ajuda no que se convencionou chamar Governança de Dados, com capítulos em todos os lugares do mundo tipo PMI, com livros de boas práticas chamados DMBOK, no melhor estilo de escrita PMBOK. Mas como sabemos, não somente a história se repete, mas os modelos também.

Conheci e, mais próximo ou distante, convivi com todos os autores acima, principalmente quando trouxemos o capítulo da DAMA – Data Management Association para o Brasil, onde tive a satisfação de ser o vice-presidente 2010-2012. A finalidade dessa organização é fomentar a Governança de dados, todavia é preciso ter dados de qualidade.

Assim, começa aqui a nossa jornada!

Você já se fez esta pergunta: Por que a qualidade dos dados é importante?

Se não fez, vou ajudá-lo(a) a entender por que e para que precisamos:

- possuir informações precisas e oportunas para gerenciar serviços e responsabilidade;
- boa informação para gerenciar a eficácia do serviço;
- priorizar e garantir a melhor utilização dos recursos;
- relatórios de auditoria e inspeção para fazer julgamentos sobre o nosso desempenho e governança.

Conseguimos boa qualidade de dados através do seguinte:

- governança e liderança - funções e responsabilidades definidas para garantir a prestação de contas pela qualidade dos dados com as políticas e os procedimentos em vigor para apoiar o processo;
- sistemas e processos - no lugar em que garantam a qualidade dos dados;
- pessoas e habilidades - treinar pessoal para que eles tenham o conhecimento adequado, as competências e a capacidade de exercer seus papéis;
- uso de dados - o objetivo de coletar dados robustos e de boa qualidade e informar a administração de dados, fazer melhorias na prestação de serviços referente a fornecedores de fontes de dados fidedignas, e promover a responsabilização de clientes, partes interessadas, residentes locais e Governo;
- Segurança dos dados - os dados coletados devem ser seguros e só devem ser utilizados para fins autorizados.

Percebemos que qualidade de dados é importante porque sem ela mandamos produtos e correspondências para os endereços errados, ou para o endereço certo só que a pessoa certa não está mais lá, ou mandamos a coisa errada para a pessoa errada (nossa, isso é pior ainda!), ou no caso dos *contact centers*, entramos em contato com um número de telefone discado por uma máquina e uma vez atendida a linha, passa-se para um operador que cobra você de uma dívida ou quer falar com fulano de tal para cobrar uma dívida e nem uma coisa nem a outra é com você... Desgaste, palavras rudes, mal falação, etc., porque os dados para fazerem o processo automatizado fluir estavam desatualizados, errados na origem, gravados no lugar errado, misturados, por fim, sem qualidade.

Nessa narrativa meio insana, você deve ter visto que isso acontece recorrentemente e envolve além do desgaste das pessoas um rio de dinheiro jogado no lixo pelas empresas em geral.

Normalmente, essa má qualidade foi adquirida pela organização ao longo de anos e anos de descuido, e de puxadinhos de sistemas e de bancos de dados migrados ou incorporados sem dicionários de dados, sem estudo de modelagem, sem arquitetura, sem limpeza e de duplicação de dados, entre outros, feitos por DBAs iniciantes ou por programadores que não consolidaram as bases e chumbaram no código as regras que deveriam estar nos bancos e que uma vez fora deles e disseminadas por uma infinidade de programas, haverá sua atualização em uns e não em outros componentes do sistema, o que provocará em muitos momentos a entrada de dados errada e conseqüentemente a inutilidade desses dados.

Entendeu a importância?

Quer um exemplo além desses sobre o que acontece quando os dados são ruins?

- Talvez o pior de tudo, decisões são feitas com dados ruins. Isso pode levar a sérios problemas de negócios na linha.
- Os processos automatizados param ou não podem ser implementados. A automação de processos exige absolutamente dados limpos e consistentes. Se os dados se tornarem inconsistentes, podem quebrar os processos.
- Os fluxos de trabalho do processo são quebrados ou são direcionados incorretamente porque os dados subjacentes estão incorretos.

A qualidade dos dados começa com os processos de Gerenciamento de Dados Mestres numa empresa, pois os Dados Mestres formam a base de sua análise.

Dados mestres de qualidade são completos, precisos e consistentes, disponíveis, marcados no horário e baseados em padrões industriais.

Melhorar os dados resultará em custos reduzidos, eficiência aprimorada, melhores informações, possibilitando a colaboração entre verticais.

Dados mestres fazem referência aos dados cadastrais, eles correspondem a um set de atributos para uma ou outra característica específica.

Difícil? Bom, vamos melhorar a definição: são dados críticos da sua atividade ou negócio, estão armazenados em diferentes sistemas e são utilizados por diferentes setores organizacionais.

Voltando ao nosso tema, os dados completos significam que todos os dados relevantes para um cliente estão vinculados e inseridos no banco de dados.

Muitas vezes, obter registros de dados completos é um desafio para as organizações, já que as pessoas de vendas muitas vezes esquecem de solicitar certas informações ou os clientes não veem o benefício de fornecer todos os detalhes necessários.

Agora você já entendeu por que existe um monte de empresas querendo usar BI, *Big Data* e *DataWarehouse* e ir além com algoritmos ultrassofisticados de *machine learning* e análise inteligente de comportamento de consumidor, etc. Mas não funciona, ou se funciona, é em uma pequena parte. A raiz dessa resposta errada, dessa má solução, está no tijolo fundamental do conhecimento, o DADO.

Qualidade de entrada de dados significa saída de dados com qualidade. Em última análise, geração de riqueza.

Orientações para Leitura Obrigatória



Leia o texto intitulado **Plano revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde**, de Claudia Risso de Araujo Lima, Joyce Mendes de Andrade Schramm, Claudia Medina Coeli e Márcia Elizabeth Marinho da Silva. O texto aponta que a garantia de uma informação de qualidade é condição essencial para a análise objetiva da situação sanitária, para a tomada de decisões baseadas em evidências e para a programação de ações de saúde: <https://goo.gl/MjrSrX>



Leia o texto intitulado **Qualidade de dados para data warehouse – roteiro de implementação**, de Mozart da Silva Britto e Jorge Rady de Almeida Júnior.

Apesar do consenso geral da importância da qualidade de dados, em muitas organizações ainda não se percebe que a falta de qualidade de dados é um dos principais causadores de perda de tempo, dinheiro e oportunidades dentro dos negócios. Mesmo dentre as organizações que têm essa cultura, em muitas delas ainda não existe uma metodologia consolidada e aplicada para garantir a melhoria contínua da qualidade de dados. O insucesso de muitas investidas empresariais tem sido causado por decisões estratégicas erradas, baseadas em dados organizacionais de baixa qualidade: <https://goo.gl/4V676E>



Leia o texto intitulado **Gestão da qualidade dos dados e informações dos Sistemas de Informação em Saúde: subsídios para a construção de uma metodologia adequada ao Brasil**, de Claudia Risso de Araujo Lima.

Atualmente, não se adota no Brasil uma abordagem formal para a avaliação da qualidade dos dados e informações em saúde, nem existe uma coordenação dos estudos de avaliação e dos esforços para aprimoramento da qualidade das informações da área, resultando em iniciativas isoladas e não regulares. A adoção de uma metodologia para gestão da qualidade das informações em saúde poderá trazer diversos benefícios, como estabelecer a regularidade das avaliações, ampliar a abrangência das regiões e dos sistemas analisados, adotar a padronização dos conceitos e das dimensões de qualidade avaliadas, possibilitar a prevenção e a identificação oportuna de problemas e garantir maior transparência no processo utilizado para assegurar a qualidade dos dados e informações: <https://goo.gl/tFA2mm>

Material Complementar

Indicações para saber mais sobre os assuntos abordados nesta Unidade:

Leitura

Um processo de avaliação de dados em um Data Warehouse

Sugiro a leitura do artigo Um processo de avaliação de dados em um Data Warehouse, de Tania M. Cernach, Edit Grassiani, Renata M. de Oliveira e Carlos H. Arima, que trata de efetuar a avaliação do estado atual de uma amostra de dados, contidos em sistemas reais que dependem fundamentalmente de dados confiáveis, dentro de uma instituição financeira. Dessa forma, é proposto um processo para a avaliação da qualidade dos dados tomando como base métodos e processos pesquisados na literatura. A definição de uma amostra de dados, de dimensões, de métricas e de regras de qualidade relevantes ao contexto do negócio, bem como a medição dos dados e análise dos resultados que foram executados.

<https://goo.gl/1jZYxk>

Sintomas de Falta de Qualidade de Dados

Leia o artigo Sintomas de falta de qualidade de dados, de Oswaldo Godinho, que trata de sintomas semelhantes de má qualidade dos dados, que surgem repetidamente durante discussões com os clientes e que essencialmente relacionam-se com baixos padrões de informações de contato.

<https://goo.gl/AtCehQ>

Uma análise das Dimensões da Qualidade de Dados em Projetos de Dados Governamentais Abertos

Leia o artigo Uma análise das dimensões da qualidade de dados em projetos de dados governamentais abertos, de Edson Carlos Germano e Hiroo Takaoka, que trata quais aspectos da Qualidade de Informação (QI) são contemplados na literatura, discute os princípios dos dados governamentais abertos e destaca as dimensões da QI que devem receber uma maior atenção em projetos de Dados Governamentais Abertos (DGA).

<https://goo.gl/7vTfJE>

Referências

OLSON J. E. **Data Quality** – The accuracy dimension. 1ª ed. Morgan Kaufmann, USA, 2003.

REDMAN, T. **Data Quality** – The field guide. 1ª ed. Elsevier, Digital Press: Boston, USA., 2001.

WANG, R. Y.; ZIAD, M.; LEE, Y. W. **Data Quality**. 1ª ed. Springer USA, 2002.



Cruzeiro do Sul Virtual
Educação a Distância

www.cruzeirodosulvirtual.com.br

Campus Liberdade

Rua Galvão Bueno, 868

CEP 01506-000

São Paulo - SP - Brasil

Tel: (55 11) 3385-3000



Cruzeiro do Sul
Educatonal