

# AULA 17 – GOVERNANÇA DE DADOS

PROFA. DRA. LEILA BERGAMASCO

CC5232 – Banco de Dados

# GOVERNANÇA DOS DADOS

ESTADÃO

Economia & Negócios



## Governo comete série de erros na divulgação de dados e arranha credibilidade de informações oficiais

03 de fevereiro de 2020 / 13h55



JORNAL NACIONAL

### 5 milhões com renda inferior a R\$ 56 não receberam auxílio-emergencial em maio

Benefício do Governo Federal chegou a 38,7% dos domicílios do país  
há 3 dias

ECONOMIA | f t ...

## Fraudes em saques de pagamentos do auxílio dão prejuízo de mais de R\$ 60 milhões à Caixa

BRASÍLIA - Uma série de fraudes em saques e pagamentos com recursos do auxílio emergencial já leva a Caixa a amargar um prejuízo de mais de R\$ 60 milhões, segundo apurou o Estadão/Broadcast. Falhas na poupança digital e no aplicativo Caixa Tem, entre outras brechas, têm permitido

Camila Turtelli e Idiana Tomazelli

26 de junho de 2020 | 18h36

## Após corrida, CPFs superam população em 12,5 milhões

Disparada veio depois de regularização pela Receita devido ao auxílio emergencial

veja

## ONU: Brasil tem um dos 20 governos mais digitais do mundo

Embora digitalização seja importante para a cidadania, país precisa atentar à segurança dos dados

Por Sabrina Brito - Atualizado em 11 jul 2020, 19h05 - Publicado em 11 jul 2020, 19h05

FOLHA DE S.PAULO



Em balanço de 300 dias,  
governo infla medidas, usa  
dados incorretos e tem  
abordagem ideológica

## IMPACTOS FISCAIS

- **R\$ 774 bilhões** foram gastos em 2019 em benefícios previdenciários, assistenciais e trabalhistas;
- 3-10% de fraudes e erros nestes pagamentos;
- **Perda de R\$ 23-77 bilhões anuais**, apenas em programas federais.

## Principais causas – Erros e Fraudes



## O QUE É

- “Exercício de autoridade, controle, planejamento, monitoramento, disponibilidade, segurança e execução dos ativos de dados e seu respectivo consumo”
- Governança de dados é uma estrutura que coordena, orienta e define regras para criação, reuso e consumo dos dados.

# DAMA DMBOK

- Data Management Body of Knowledge (DAMA DMBOK®) é um framework de boas práticas de gestão de dados, que tem como intuito transmitir a importância da gestão de dados.





# GERENCIAMENTO DE ARQUITETURA DE DADOS

- Busca entender quais os requisitos do seu projeto de dados, ou seja, entender quais os dados que são necessários, de onde eles vêm e por onde terão de passar até chegar ao seu dashboard. É definir o caminho deste dado.

# DESENVOLVIMENTO DOS DADOS

- Consiste na análise, projeto, implementação, implantação, e manutenção de soluções de dados para maximizar o valor de recursos de dados para a organização.

# GERENCIAMENTO DE OPERAÇÕES DE DADOS

- Trata-se do planejamento, controle, manutenção e suporte ao ativo dado, durante todo o seu ciclo de vida

# GERENCIAMENTO DE SEGURANÇA DE DADOS

- Aqui é muito claro. É necessário planejar e projetar uma estrutura capaz de garantir a privacidade, confidencialidade e acesso apropriado ao dado.

# GERENCIAMENTO DE DADOS MESTRE E REFERÊNCIA

- Dados mestre: trata-se de projetar uma estrutura que promova uma “Central Única da Verdade”, garantindo uma versão consistente e confiável do dado, onde possa ser distribuído (ou compartilhado) para outros contextos.
- Já a gestão de dados de referência inclui o controle de termos padronizados, códigos e outros identificadores únicos. Inclui definições de negócio para cada valor de código, relacionamento de negócios dentre as listas de domínios, uso compartilhado dos dados. Garante a consistência, precisão e temporalidade de atualização necessária para classificar e categorizar os dados.
- Dados de referência e dados mestre incluem glossários, dicionários, classificadores, índices, identificadores e codificadores. Além de dicionários padrão, diretórios e classificadores.

## GERENCIAMENTO DE DW E BI

- Trata-se de planejar e projetar modelos de dados que permitam a geração de informações para tomada de decisão sob várias perspectivas (dimensões).
- Fornecer o armazenamento atuais e históricos integrado organizado por áreas temáticas; garantir credibilidade, qualidade, estabilidade, alto desempenho e ambiente para aquisição, gestão e acesso a dados, e para todos os recursos de acesso adequados; proporcionar um ambiente de acesso fácil de usar, flexível e abrangente, entre outros.

# GERENCIAMENTO DA DOCUMENTAÇÃO E CONTEÚDO

- Trata-se de planejar e projetar a implantação e a gestão a dados que estão fora de um banco de dados. É importante que seja definido um plano para armazenamento, proteção e acesso a estes dados.

# GERENCIAMENTO DE METADADOS

- Metadado é o dado a respeito de outro dado, ou seja, são informações que complementam um dado. Por exemplo: 2 maçãs (dado), Foto destas mesmas 2 maçãs (Metadado). Portanto, o metadado tem a mesma importância do dado e o gerenciamento de metadados segue as mesmas atividades do gerenciamento de dados.
- Os metadados descrevem a estrutura e significados a respeito de dados e, assim contribuem para que seu uso seja eficiente ou ineficiente, oferecendo contexto aos dados relacionados, ou seja informações que gerem conhecimento

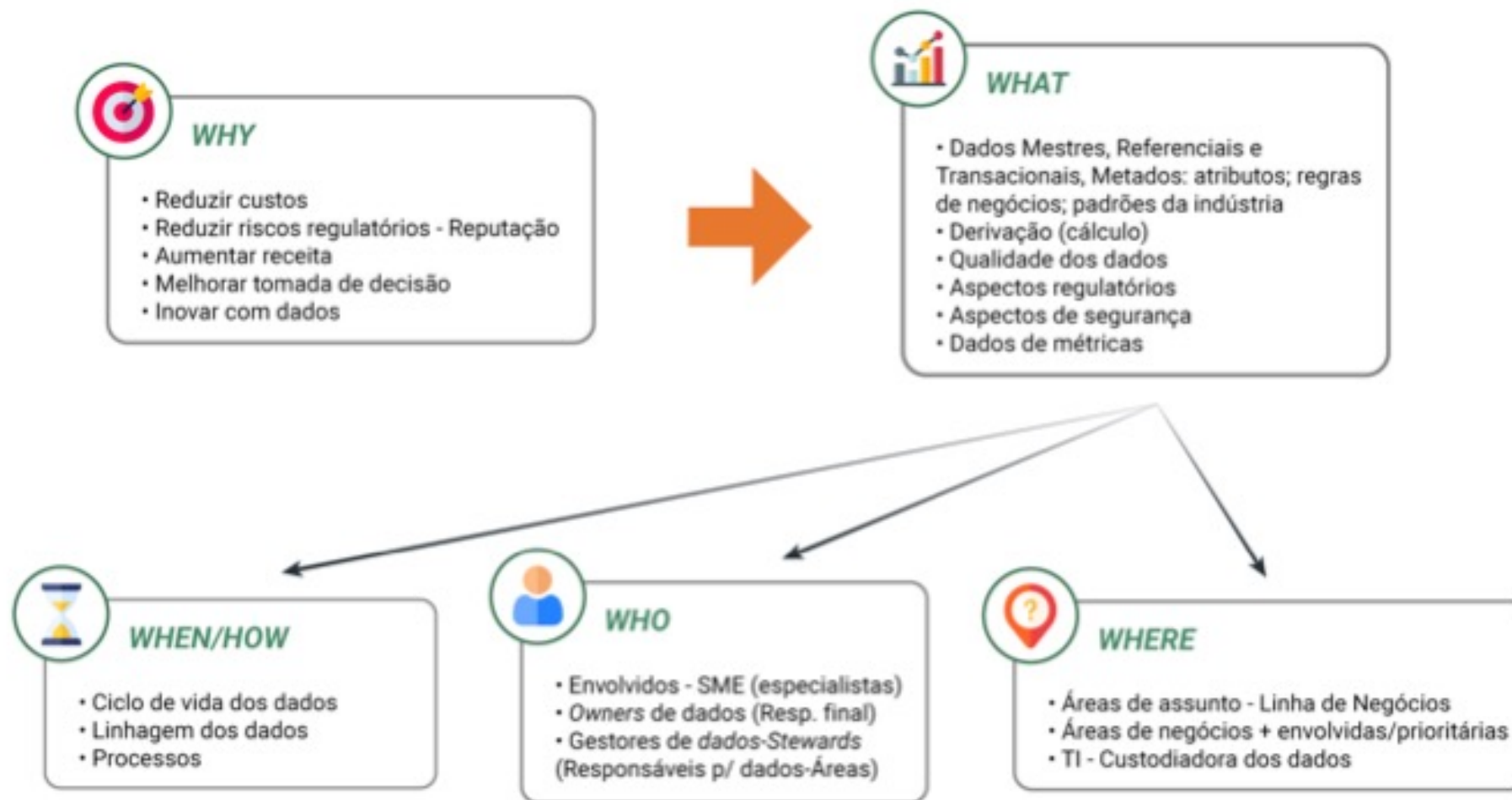


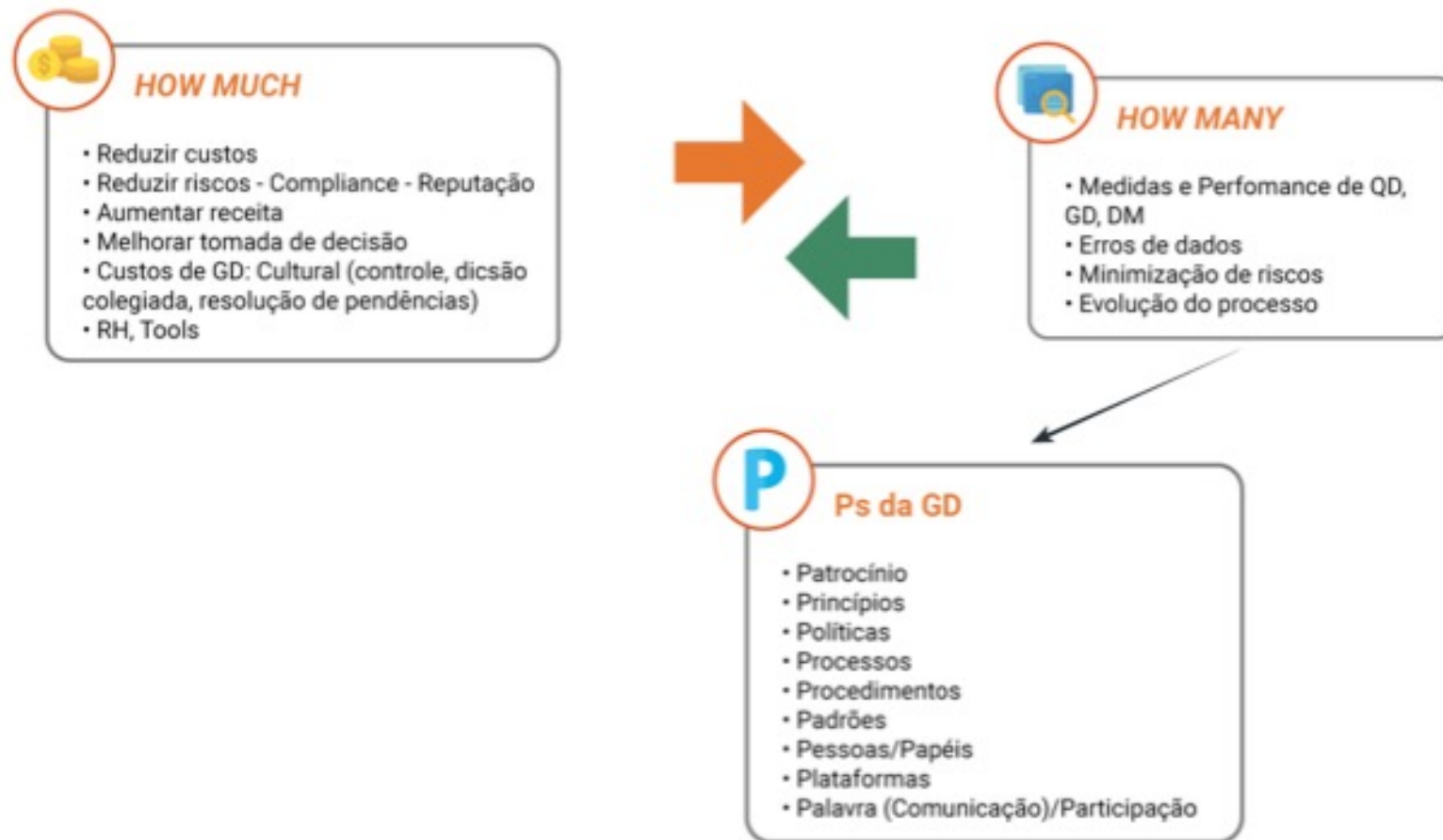
# GERENCIAMENTO DE QUALIDADE DE DADOS

- Trata-se de planejar e projetar o saneamento do dado, provendo qualidade ao mesmo, para que este dado possa gerar informações confiáveis para suportar a tomada de decisão. O objetivo é planejar, implementar e controlar atividades que apliquem técnicas de gerência de qualidade de dados para medir, avaliar, melhorar e garantir a adequação dos dados ao seu uso pretendido.

# VÁRIOS FRAMEWORKS

- 5W2H
- IBM
- COBIT
- CMMI





## RESULTADOS

Riscos relativos aos dados e aspectos de aderência e criação de valor para os ativos de informação.

Riscos relativos aos dados e aspectos de aderência

Criação de valor para os ativos de informação

## VIABILIZADORES

Conscientização e Estrutura Organizacional

Política

Custódia de Dados

Envolve conscientização e estrutura organizacional com relação à política e custódia de dados. Aqui, são destacadas as políticas que definem direcionamentos estratégicos que apontam o comportamento desejado pela empresa a respeito da forma de atuação dos envolvidos no programa. Há também a *data stewardship* (gestão de dados), disciplina de controle de qualidade para garantir um cuidado focado nos dados, visando à sua melhoria e a gerência de riscos envolvidos, e a *data custodian* ("custodiadores" de dados).

## DISCIPLINAS CENTRAIS

Surge a gerência de qualidade dos dados, gerência do ciclo de vida da informação, e segurança e privacidade da informação.

Gerência da Qualidade dos Dados

Gerência do Ciclo de Vida da Informação

Segurança e Privacidade da Informação

## DISCIPLINAS DE APOIO

Disciplinas de suporte, como a arquitetura dos dados com documentação e definição de modelos de dados, classificação e definição de metadados, e procedimentos de controle com informações de auditoria, *logging* e relatórios.

Arquitetura dos Dados

Classificação e Metadados

Informações de auditoria, *Logging* e Relatórios





# DICIONÁRIO DE DADOS (DD)

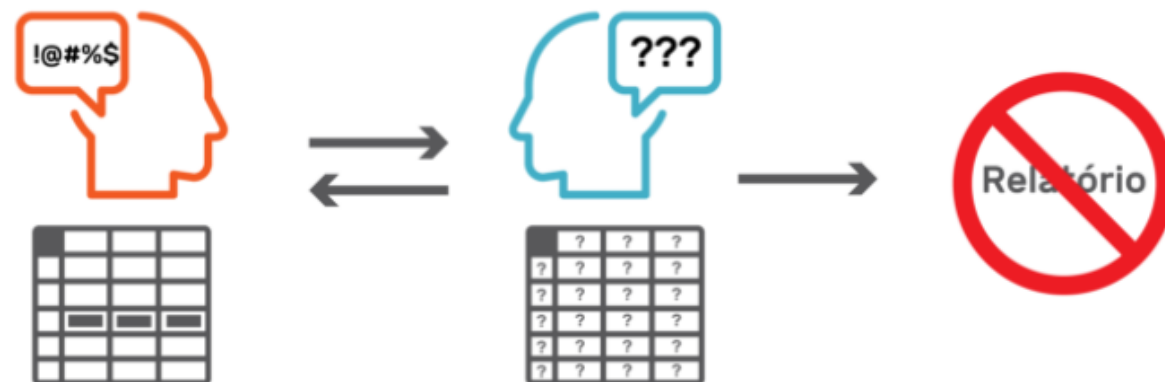
- Um dicionário de dados é uma coleção de metadados que contém definições e representações de elementos de dado.
- Documento de texto ou planilha que centraliza informações sobre o conjunto de dados
- O dicionário de dados é um repositório que descreve, de forma estruturada, o significado, origem, relacionamento e uso dos dados



# MOTIVAÇÃO

- “Consegui uma planilha (conjunto de dados, banco de dados ou dataset) do departamento X mas não entendo o que significa cada uma das colunas e linhas... Tenho que ligar lá toda vez para saber.”

Assim como um dicionário idiomático que explica a origem das palavras, seus significados e sinônimos, um dicionário de dados organiza o conhecimento necessário a respeito dos dados, onde são coletados, suas características, funções e etc. Compartilha significado!



# MELHORES PRÁTICAS

- Faça o versionamento dos dicionários e mantenha-os atualizados com alguém responsável pela gestão do documento. O fato de existir pessoas trabalhando com versões diferentes do mesmo dicionário pode causar ruído.
- Tenha um repositório acessível por todos os analistas, pois o nível da cultura de data analytics só aumenta quando todos os indivíduos estão cientes dos significados dos processos baseados em dados e este significado é compartilhado.
- Tenha um dicionário para cada conjunto de dados e que este obedeça uma padronização mínima dentro do setor ou organização.
- Simplifique ao máximo o nome das colunas. Utilize nomes curtos, sem espaços, sem acentos. Espaços e acentos geram problemas técnicos nos diversos sistemas e que sejam intuitivos, para que seus colegas de trabalho possam entender os dados o mais facilmente possível. Evite jargões técnicos.
- Quando compartilhar dados, tenha o costume de enviar os dicionários junto

## MELHORES PRÁTICAS

- No modelo de dicionário de dados, os campos a serem preenchidos são os seguintes:
  - Chave: Tipo da chave (PK ou FK)
  - Campo: nome do campo, como consta no banco de dados. Normalmente estes nomes são curtos, sem acentos e nem sempre são autoexplicativos.
  - Tipo: trata da natureza da variável, como por exemplo: categórica, binária, escalar, numérica.
  - Descrição: o que significa a variável e como ela se relaciona com outras variáveis do banco de dados. Se houver regras de negócio importantes atreladas ao dado, elas devem estar descritas aqui. Formatos e padrões também.

| Field Name    | Data type | Field Length | Constraint  | Description               |
|---------------|-----------|--------------|-------------|---------------------------|
| Client_id     | Int       | 10           | Primary key | Client id, Auto generated |
| Client_name   | Varchar   | 20           | Not null    | Name of client            |
| Password      | Varchar2  | 30           | Not null    | Login Password for client |
| Contact_no    | Int       | 15           | Not null    | Landline or mobile number |
| Email_id      | Varchar2  | 30           | Not null    | Any email id              |
| Max_Users     | Int       | 10           | Not null    | Maximum number of users   |
| Current_users | Int       | 10           | Not null    | Currently present user    |

```
SELECT *
FROM information_schema.columns
WHERE table_schema =
'<nome_do_esquema>'
```

nishiqmo/nishiqmo@CC6240

Query Query History Scratch Pad x

```
1 SELECT *
2 FROM information_schema.columns
3 WHERE table_schema = 'cc5232'
```

Data output Messages Notifications

|    | table_catalog | table_schema | table_name | column_name | ordinal_position | column_default | is_nullable | data_type         | character_maximum_length | character_octet_length |
|----|---------------|--------------|------------|-------------|------------------|----------------|-------------|-------------------|--------------------------|------------------------|
| 27 | nishiqmo      | cc5232       | aula2022   | nome        | 1                | [null]         | YES         | character varying | 255                      | 1                      |
| 28 | nishiqmo      | cc5232       | aula2022   | qtde_alum   | 2                | [null]         | YES         | integer           | [null]                   | [null]                 |
| 29 | nishiqmo      | cc5232       | aula2022   | codigo_disc | 3                | [null]         | YES         | integer           | [null]                   | [null]                 |
| 30 | nishiqmo      | cc5232       | aluno      | alu_numero  | 1                | [null]         | NO          | integer           | [null]                   | [null]                 |
| 31 | nishiqmo      | cc5232       | aluno      | alu_nome    | 2                | [null]         | YES         | character varying | [null]                   | 1073741                |
| 32 | nishiqmo      | cc5232       | aluno      | alu_rua     | 3                | [null]         | YES         | character varying | [null]                   | 1073741                |
| 33 | nishiqmo      | cc5232       | aluno      | alu_cidade  | 4                | [null]         | YES         | character varying | [null]                   | 1073741                |
| 34 | nishiqmo      | cc5232       | professor  | prof_nome   | 1                | [null]         | YES         | character varying | 2000                     | 8                      |
| 35 | nishiqmo      | cc5232       | professor  | prof_rua    | 2                | [null]         | YES         | character varying | 2000                     | 8                      |
| 36 | nishiqmo      | cc5232       | professor  | prof_cidade | 3                | [null]         | YES         | character varying | 2000                     | 8                      |
| 37 | nishiqmo      | cc5232       | professor  | prof_tel    | 4                | [null]         | YES         | character varying | 2000                     | 8                      |
| 38 | nishiqmo      | cc5232       | professor  | prof_numero | 5                | [null]         | NO          | integer           | [null]                   | [null]                 |

OBRIGADO E ATÉ A PRÓXIMA AULA!