Processamento de dados em Big Data

Parte 3 - Governança, privacidade e segurança Luiz Henrique Zambom Santana, D.Sc.

INE | CTC



Agenda

- Apresentações dos grupos
- Governança de dados
 - Leis de proteção
 - Apache Atlas

Como podemos usar essas ideias em domínios reais?

Domínios

- o ESG
- eHealth
- Apostas em esporte
- Reconhecimento facial para segurança pública
- Gestão de força de trabalho em empresas terceirizadas

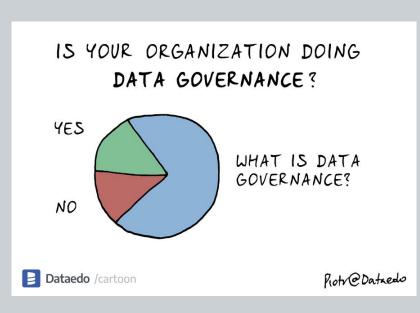
Perguntas

- Esse domínio precisa de Big Data?
- Existe alguma oportunidade para uso de Streaming?
- O domínio se beneficiaria de Machine Learning? Qual tipo?
- Como os dados poderiam ser armazenados?



Governança de dados

- Falta de controle de ativos
- Falhas na segurança
- Custos elevados
- Falta de integração entre os setores e sistemas
- Problemas judiciais



governança

Ação, resultado ou efeito de governar ou de se governar (orientar); governo.

Tendência ou capacidade de ter o poder sobre alguma coisa.

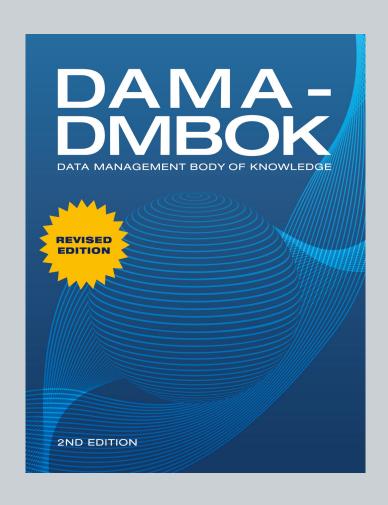


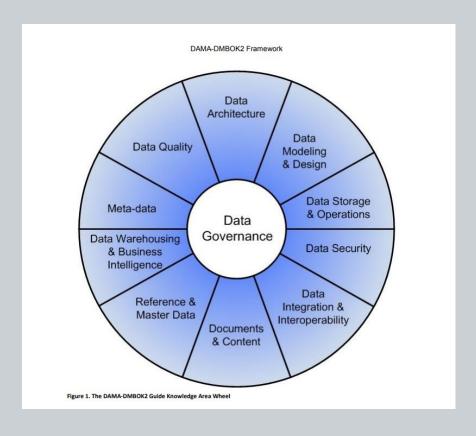


Governança de dados

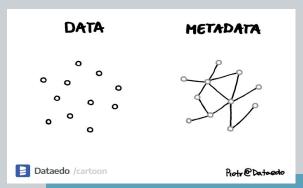


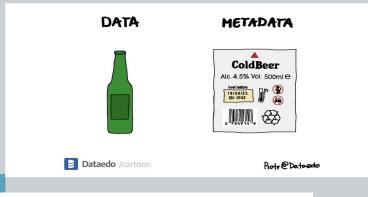
Governança de dados

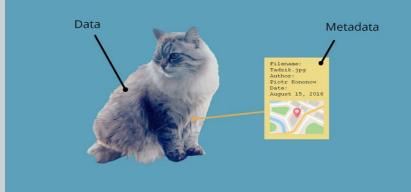


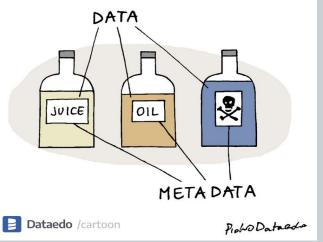


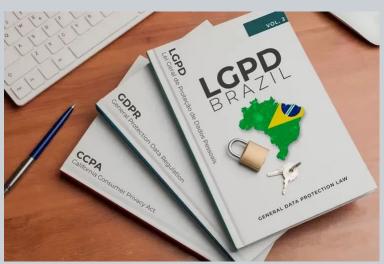
Quais dados?













- O mínimo que precisamos saber:
 - Classificação de dados
 - o Príncipios
 - Hipóteses para tratamento



- O mínimo que precisamos saber:
 - Príncipios

01	Finalidade especificada e informada explicitamente ao titular
02	Adequação à finalidade previamente acordada e divulgada
03	Necessidade do tratamento, limitado ao uso de dados essenciais para alcançar a finalidade inicial
04	Acesso livre, fácil e gratuito das pessoas à forma como seus dados são tratados
05	Qualidade dos dados, deixando-os exatos e atualizados, segundo a real necessidade no tratamento
06	Transparência, ao titular, com informações claras e acessíveis sobre o tratamento e seus responsáveis
07	Segurança para coibir situações acidentais ou ilícitas como invasão, destruição, perda, difusão
08	Prevenção contra danos ao titular e a demais envolvidos
09	Não discriminação, ou seja, não permitir atos ilícitos ou abusivos
10	Responsabilização do agente, obrigado a demonstrar a eficácia das medidas adotadas

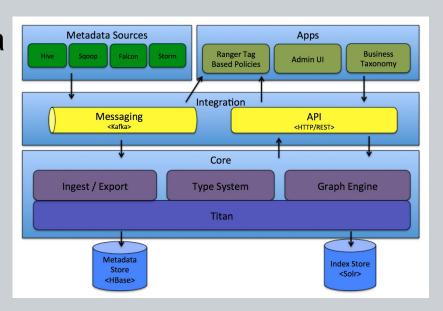


- O mínimo que precisamos saber:
 - Hipóteses para tratamento



Apache Atlas

- Governança para Big Data
- Oferece:
 - Gestão centralizada de metadados
 - Classificação de dados
 - Gestão do ciclo de vida de dados
 - Linhagem de dados
 - Segurança centralizada



Luiz Henrique Zambom Santana

Ihzsantana@gmail.com

https://www.linkedin.com/in/luizsantana/

