Big Data - Spark, Hadoop e afins.

Leonardo Gloria

Aula o1



Leo Gloria

- * Formação:
 - * MBA em Machine Learning (2019)
 - * Bacharel em CC (2011)
- * Certificações:
 - * AWS Architect Certified 2017
 - * OCWCD (2010)
 - * SCJP (2008)

Experiência Profissional

- * Arquiteto de Soluções Credilink Informações de Crédito (2010)
- * Instrutor Caelum Alura (2017)
- * Consultor Marinha do Brasil
- * Consultor AWS
- * Professor Instituto Infnet

Compreendendo os fundamentos de BigData

Evolução da Gestão de dados.

- * Volume de dados extremamente grande.
- * Velocidade de dados extremamente alta.
- * Variedade de dados extremamente ampla.

Ponto atual do Big Data.



Ondas do Big data

Onda 1: Criando estruturas de dados administráveis

Onda 2: Administração web e de conteúdo.

Onda 3: Administrando Big data

Definindo Big Data

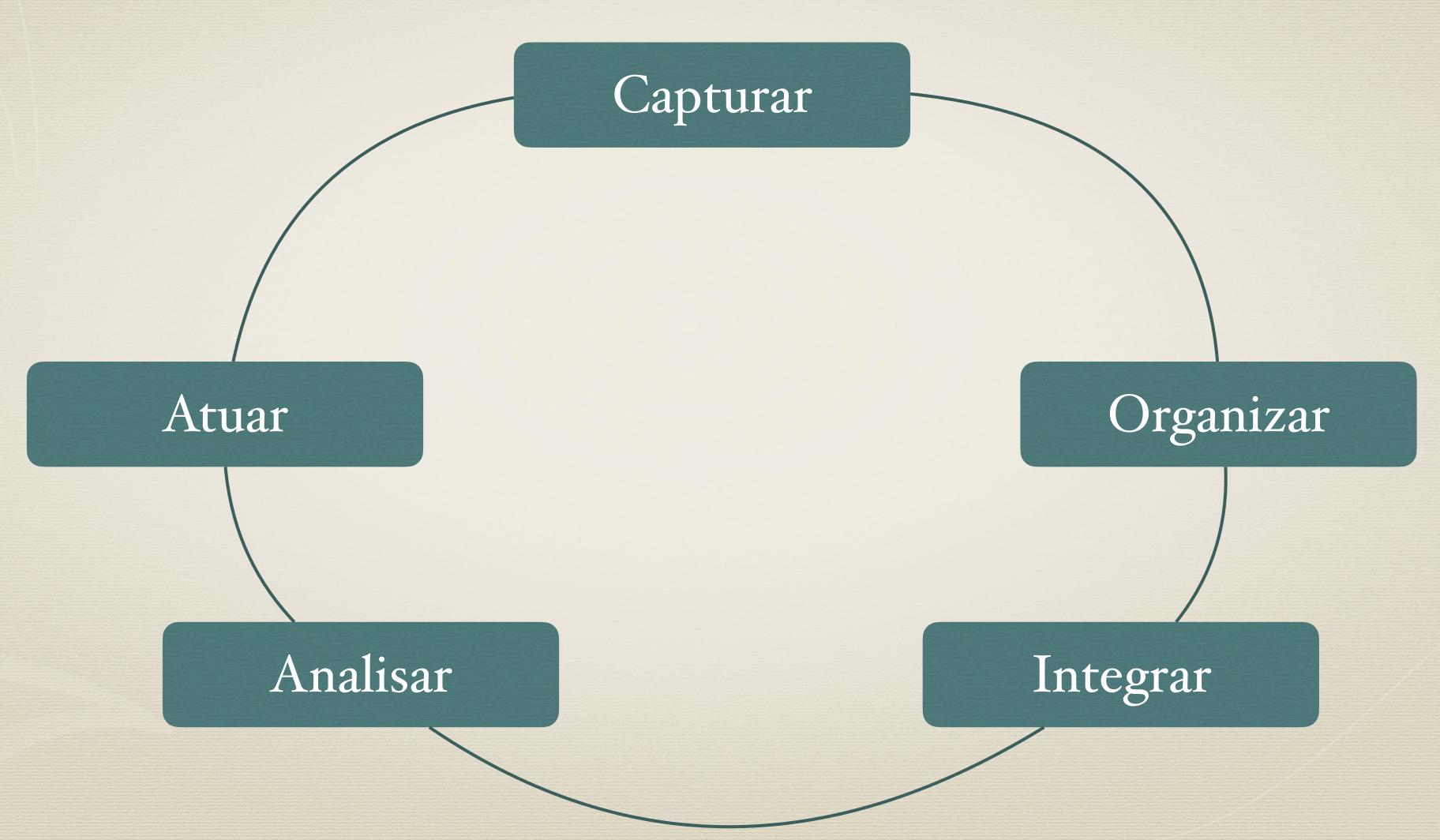
- * Volume: Quantidade de dados
- * Velocidade: A rapidez com que esses dados são processados
- * Variedade: Os vários tipos de dados.

Definindo Big Data

- * Volume: Quantidade de dados
- * Velocidade: A rapidez com que esses dados são processados
- * Variedade: Os vários tipos de dados.
- * Veracidade: O quão precisos são esses dados.

Construindo uma arquitetura de administração de Big Data

Ciclo de Gestão de Big Data



Estabelecendo a base arquitetônica

- * Quantos dados administrar agora e no futuro?
- * Qual frequência precisarei administrar dados em tempo real ou próximo disso.
- * Qual a quantidade de risco que consigo suportar? Conformidades de segurança, LGPD, etc.
- * Quão certos ou precisos devem ser os dados?

Stack de Tecnologias para Big Data

Infra física

Infra de Segurança

Fontes de Dados

Desempenho e persistência poliglota

Map Reduce

Big Table

Hadoop

Análises Tradicionais e avançadas

Machine Learning

Relatórios e visualizações

Aplicativos que utilizam BD.

Dúvidas?

"Stay hungry, stay foolish."

-Steve Jobs