

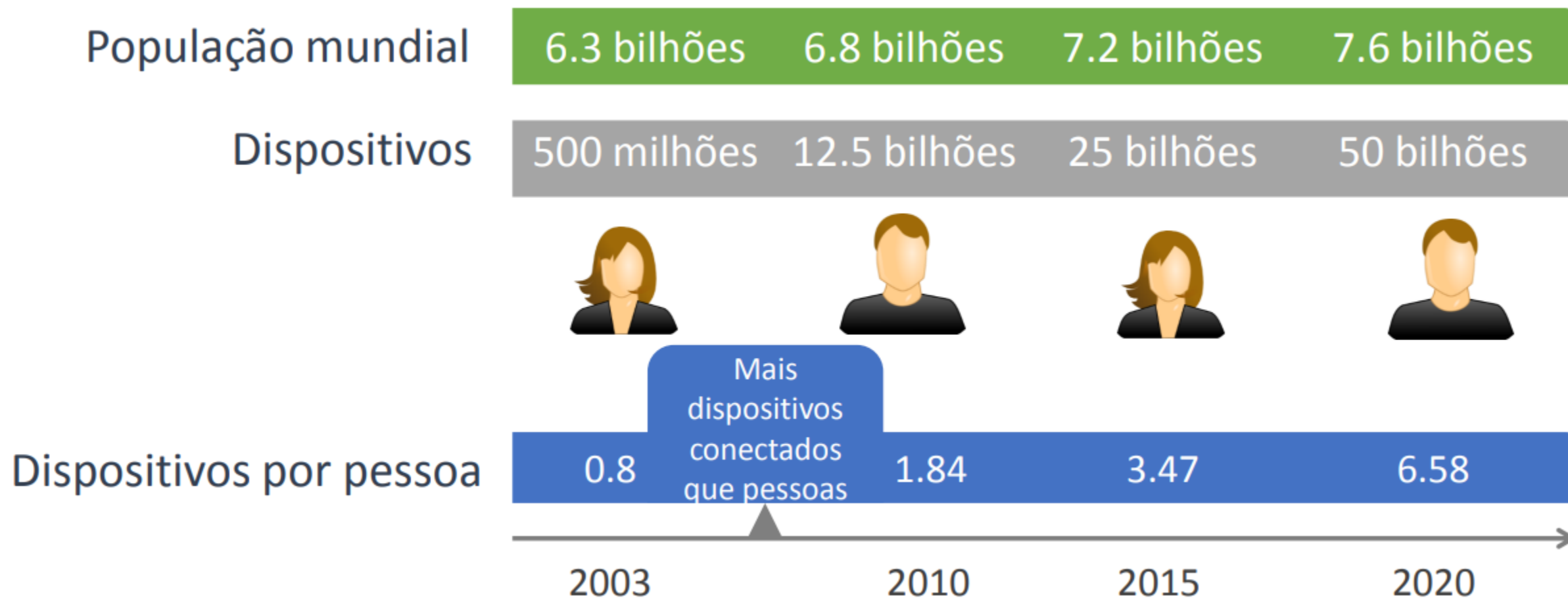
Big Data

Professor Vitor Kessler

Agenda

- Big Data.
 - Fundamentos.
 - Tipos de dados: estruturados, semiestruturados e não estruturados.
 - Vs do Big Data.
 - Pipeline de dados.
 - Conceito de data lake.
 - ETL x ELT.
 - NoSQL.
 - Hadoop.

Fundamentos – Big Data



Fundamentos – Big Data



Fundamentos – Big Data



Fundamentos – Big Data

30
segundos
de
internet



Fundamentos – Big Data



Fundamentos – Big Data

- Muito volume de dados.
- Alta velocidade.
- Diversas fontes de dados.
- O termo é recente e o mercado está em formação.
- Representa um conjunto de dados que não pode ser facilmente gerenciado.
- As ferramentas, métodos e arquiteturas de dados atuais não são suficientes.
- Demanda formas inovadoras e rentáveis de processamento de dados.

Fundamentos – Big Data

- Exemplos de fontes:
 - E-mails.
 - Web log.
 - Dados de sensores.
 - Logs de call centers.
 - Imagens / vídeos.
 - Dados RFID.
 - Dados de localização.

Fundamentos – Big Data

- Aplicações possíveis:
 - Machine Learning.
 - Análise de sentimentos.
 - Processamento de texto, de imagens e de vídeos.
 - Análise de log.
 - Análise de e-mails.

Fundamentos – Big Data

- Prova: CESPE - 2015 - TRE-GO - Técnico Judiciário - Área Administrativa
- A Big Data pode ser utilizada na EAD para se entender as preferências e necessidades de aprendizagem dos alunos e, assim, contribuir para soluções mais eficientes de educação mediada por tecnologia.

Fundamentos – Big Data

- Prova: CESPE - 2015 - TRE-GO - Técnico Judiciário - Área Administrativa
- A Big Data pode ser utilizada na EAD para se entender as preferências e necessidades de aprendizagem dos alunos e, assim, contribuir para soluções mais eficientes de educação mediada por tecnologia.

Fundamentos – Big Data

- Prova: CESPE - 2018 - EBSEH - Analista de Tecnologia da Informação
- As soluções de big data focalizam dados que já existem, descartam dados não estruturados e disponibilizam os dados estruturados.

Fundamentos – Big Data

- Prova: CESPE - 2018 - EBSEH - Analista de Tecnologia da Informação
- As soluções de big data focalizam dados que já existem, descartam dados não estruturados e disponibilizam os dados estruturados.

Fundamentos – Big Data

- Provas: CESPE / CEBRASPE - 2021 - Polícia Federal - Escrivão de Polícia Federal
- Acerca dos conceitos de mineração de dados, aprendizado de máquina e bigdata, julgue o próximo item.
- As aplicações de bigdata caracterizam-se exclusivamente pelo grande volume de dados armazenados em tabelas relacionais.

Fundamentos – Big Data

- Provas: CESPE / CEBRASPE - 2021 - Polícia Federal - Escrivão de Polícia Federal
- Acerca dos conceitos de mineração de dados, aprendizado de máquina e bigdata, julgue o próximo item.
- As aplicações de bigdata caracterizam-se exclusivamente pelo grande volume de dados armazenados em tabelas relacionais.

Fundamentos – Big Data

- Prova: CESPE - 2018 - Polícia Federal - Escrivão de Polícia Federal
- Em um big data, alimentado com os dados de um sítio de comércio eletrônico, são armazenadas informações diversificadas, que consideram a navegação dos usuários, os produtos comprados e outras preferências que o usuário demonstre nos seus acessos.
- Tendo como referência as informações apresentadas, julgue o item seguinte.
- O big data consiste de um grande depósito de dados estruturados, ao passo que os dados não estruturados são considerados data files.

Fundamentos – Big Data

- Prova: CESPE - 2018 - Polícia Federal - Escrivão de Polícia Federal
- Em um big data, alimentado com os dados de um sítio de comércio eletrônico, são armazenadas informações diversificadas, que consideram a navegação dos usuários, os produtos comprados e outras preferências que o usuário demonstre nos seus acessos.
- Tendo como referência as informações apresentadas, julgue o item seguinte.
- O big data consiste de um grande depósito de dados estruturados, ao passo que os dados não estruturados são considerados data files.

Fundamentos – Big Data

- Prova: CESPE / CEBRASPE - 2021 - SEED-PR - Professor - Informação e Comunicação
- Assinale a opção que apresenta a denominação de um conjunto de dados que contém tipos de dados não estruturados e semiestruturados, como texto, áudio e vídeo, e que exige um pré-processamento adicional para obter significado e dar suporte a metadados, com o objetivo de que estes sejam usados para resolver problemas de negócios.
- A banco de dados NoSQL
- B machine learning (aprendizado de máquina)
- C text mining (mineração de texto)
- D data warehouse
- E Big Data

Fundamentos – Big Data

- Prova: CESPE / CEBRASPE - 2021 - SEED-PR - Professor - Informação e Comunicação
- Assinale a opção que apresenta a denominação de um conjunto de dados que contém tipos de dados não estruturados e semiestruturados, como texto, áudio e vídeo, e que exige um pré-processamento adicional para obter significado e dar suporte a metadados, com o objetivo de que estes sejam usados para resolver problemas de negócios.
- A banco de dados NoSQL
- B machine learning (aprendizado de máquina)
- C text mining (mineração de texto)
- D data warehouse
- E Big Data

Fundamentos – Big Data

- Prova: IADES - 2018 - APEX Brasil - Analista - Prospeção de Projetos
- Assinale a alternativa que apresenta o conceito de Big Data.
- A Conjuntos de dados de grande volume que se utilizam de ferramentas especiais de processamento, pesquisa e análise, e que podem ser aproveitados no tempo necessário, com precisão e grande velocidade.
- B São bancos de dados de fácil acesso e rápida velocidade, operados como computadores pessoais.
- C Manuseio de informações necessárias às empresas e aos negócios do mundo moderno, que podem ser armazenadas em computadores pessoais, utilizando-se a técnica de nuvem de dados.
- D São apenas grandes volumes de dados que precisam ainda ser mais bem aproveitados pelo mundo corporativo.
- E Refere-se a um grande número de computadores pessoais (PC) interligados entre si em uma grande rede de informação.

Fundamentos – Big Data

- Prova: IADES - 2018 - APEX Brasil - Analista - Prossecção de Projetos
- Assinale a alternativa que apresenta o conceito de Big Data.
- A Conjuntos de dados de grande volume que se utilizam de ferramentas especiais de processamento, pesquisa e análise, e que podem ser aproveitados no tempo necessário, com precisão e grande velocidade.
- B São bancos de dados de fácil acesso e rápida velocidade, operados como computadores pessoais.
- C Manuseio de informações necessárias às empresas e aos negócios do mundo moderno, que podem ser armazenadas em computadores pessoais, utilizando-se a técnica de nuvem de dados.
- D São apenas grandes volumes de dados que precisam ainda ser mais bem aproveitados pelo mundo corporativo.
- E Refere-se a um grande número de computadores pessoais (PC) interligados entre si em uma grande rede de informação.

Fundamentos – Big Data

- Prova: CESPE / CEBRASPE - 2021 - TCE-RJ - Analista de Controle Externo - Especialidade: Tecnologia da Informação
- Com relação a Big Data, julgue o item seguinte.
- Os fatores críticos de sucesso da análise de Big Data incluem uma sólida infraestrutura de dados, além de ferramentas analíticas e pessoal habilitado para lidar com elas.

Fundamentos – Big Data

- Prova: CESPE / CEBRASPE - 2021 - TCE-RJ - Analista de Controle Externo - Especialidade: Tecnologia da Informação
- Com relação a Big Data, julgue o item seguinte.
- Os fatores críticos de sucesso da análise de Big Data incluem uma sólida infraestrutura de dados, além de ferramentas analíticas e pessoal habilitado para lidar com elas.