

The background is a light gray gradient. It features several realistic water droplets of various sizes, some with highlights and shadows, scattered around the edges. In the center, there is a faint, circular, textured pattern that resembles a lens flare or a subtle watermark.

Big Data

A nova ciência dos dados

Big Data?

Big data se trata de um conceito onde o foco é grande armazenamento de dados e maior velocidades.

Os 5 V's

1. Velocidade
2. Volume
3. Variedade
4. Veracidade
5. Valor

Velocidade

Em muitos casos, as respostas precisam ser rápidas, praticamente em tempo real para o tratamento de um volume massivo de dados;

Para um sistema anti-fraude, 2 minutos pode ser muito tempo.

Volume

Empresas podem facilmente produzir GB e mais GB de informações por dia.

O Twitter pode facilmente produzir diariamente alguns TB de informações pertinentes para análise de sentimento;

Variedade

São informações de diversas fontes, como imagens, tuites e outros resultados de redes sociais, textos, etc. Os dados não seguem uma modelagem padrão;

Veracidade

A maioria dos gestores não confiam 100% nas informações que recebem para tomada de Decisões;

Ao organizar as informações e usar filtros apropriados, é possível aumentar a confiabilidade dos dados;

Valor

Depois que os dados são tratados e filtrados, podem representar um valor inestimável para empresa, seja para tomada de decisões ou definições de novos produtos;

Como funciona?

Big data é capaz de capturar, armazenar e processar um massivo volume de informações que podem ser espalhados em diversos hardware de commodity;

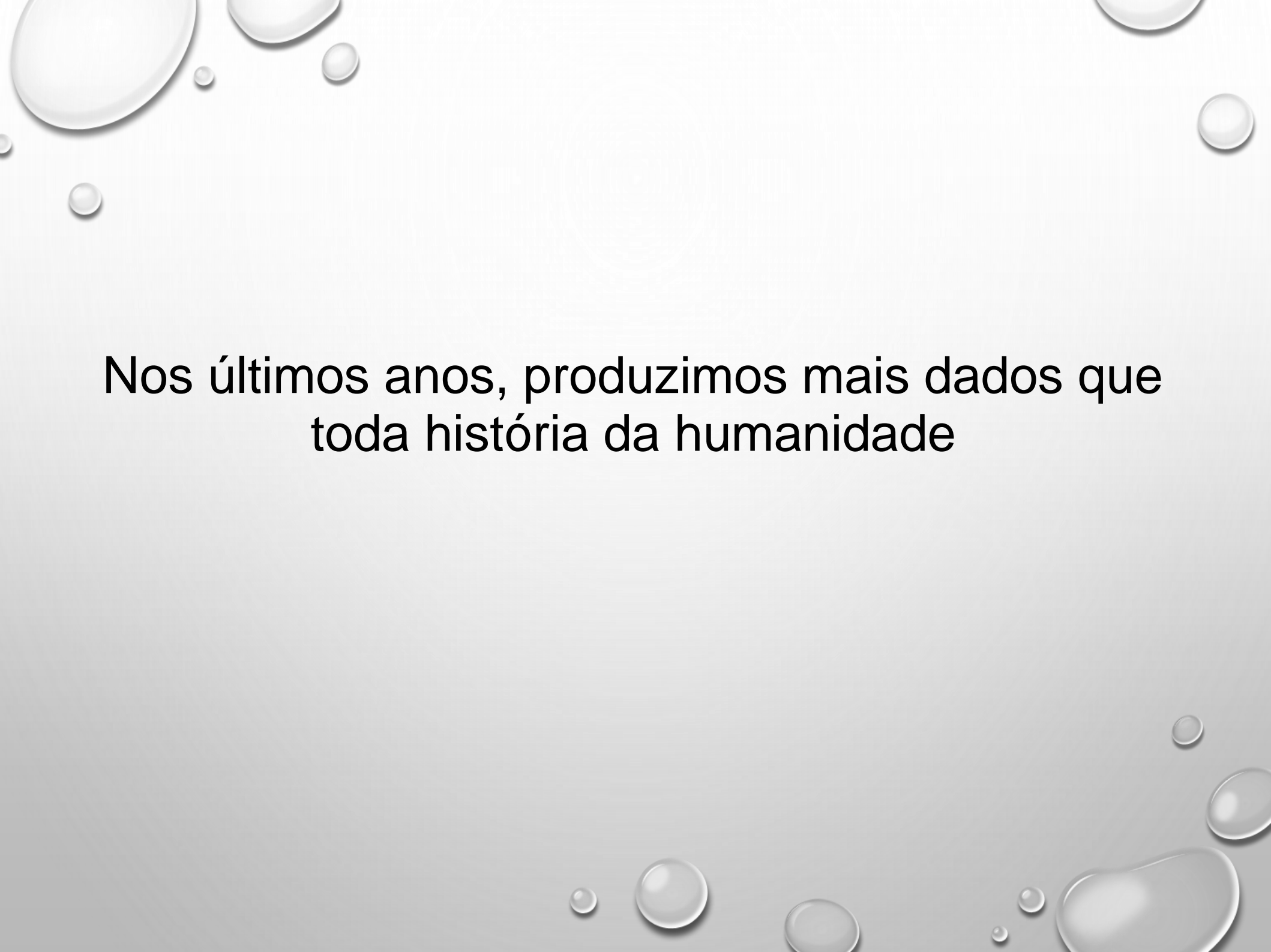
Uma vez capturado, os dados podem ser processados através de dezenas de ferramentas;

Exemplos de uso

- Decodificar o **genoma humano** demorava cerca de 10 anos; atualmente, pode ser processado em uma fração desse tempo com big data;
- **Walmart** processa milhões de transações de clientes por dia, com big data, consegue identificar padrões de fraudes quase que imediatamente;
- **FICO** (inclui sistema de detecção de fraudes em cartão de crédito) consegue processar mais de 2 bilhões de transações por dia;

The background of the slide is a light gray gradient. In the top-left and bottom-right corners, there are several realistic-looking water droplets of various sizes, rendered with soft shadows and highlights to give them a three-dimensional appearance. The title text is centered in the middle of the slide.

A nova ciência dos dados

The background is a light gray gradient. In the top-left and bottom-right corners, there are several realistic-looking water droplets of various sizes, some overlapping. The text is centered in the middle of the slide.

Nos últimos anos, produzimos mais dados que
toda história da humanidade




- 
- Já imaginou que a maioria dos dispositivos eletrônicos já possuem um endereço de IP?

Quando você opera esse dispositivo, alguma informação é gerada.

Isso vai ficar gravado em algum lugar.

O mundo possui aproximadamente 7 bilhões de pessoas.



Estamos gerando informação o tempo todo

- Utilizando avião
- Pagando contas online
- Comprando pela internet
- Acessando redes sociais
- Assistindo TV (sim, SmartTV)
- Consultando um médico
- Realizando ligações telefônicas
- Até quando morremos

Arquitetura

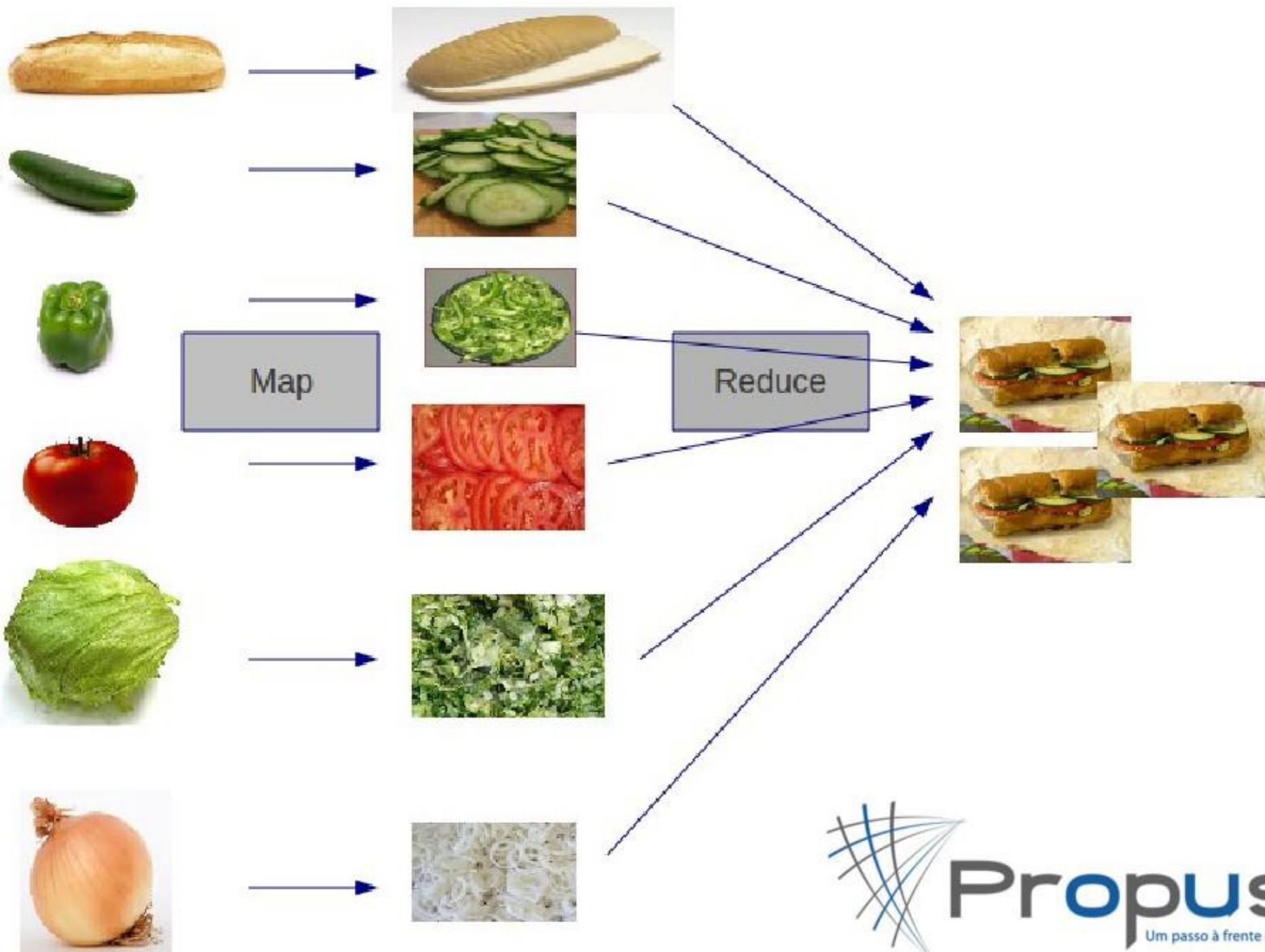
- Inúmeros sistemas legados;
- Falta de padronização dos dados;
- Cada produto tem um objetivo e guarda os dados no seu formato;
- Dificuldade de um produto “conversar” com outro;

Arquitetura

- Quando existe a necessidade de buscar informações de diversas fontes, geralmente o trabalho é manual;
- Alguém centraliza as informações de diferentes ferramentas fazendo um “copy & paste”;
- Trabalho é lento, pode ter falha humana;

Arquitetura

- Soluções de Big Data podem automatizar esse trabalho...
- ... Não apenas automatizar, mas colocar inteligência...
- ... Identificar novos padrões...
- ... Aumentar a precisão das informações ...
- ... Fazer o gestor enxergar novas possibilidades.



Propus
Um passo à frente em TI

Algunas tecnologías de Big Data

- Hadoop
- Pig
- Hive
- Hbase
- Storm
- Solr
- Ambari
- Zookeeper, etc...

Big Data Landscape (Version 2.0)

Infrastructure



Analytics



Applications

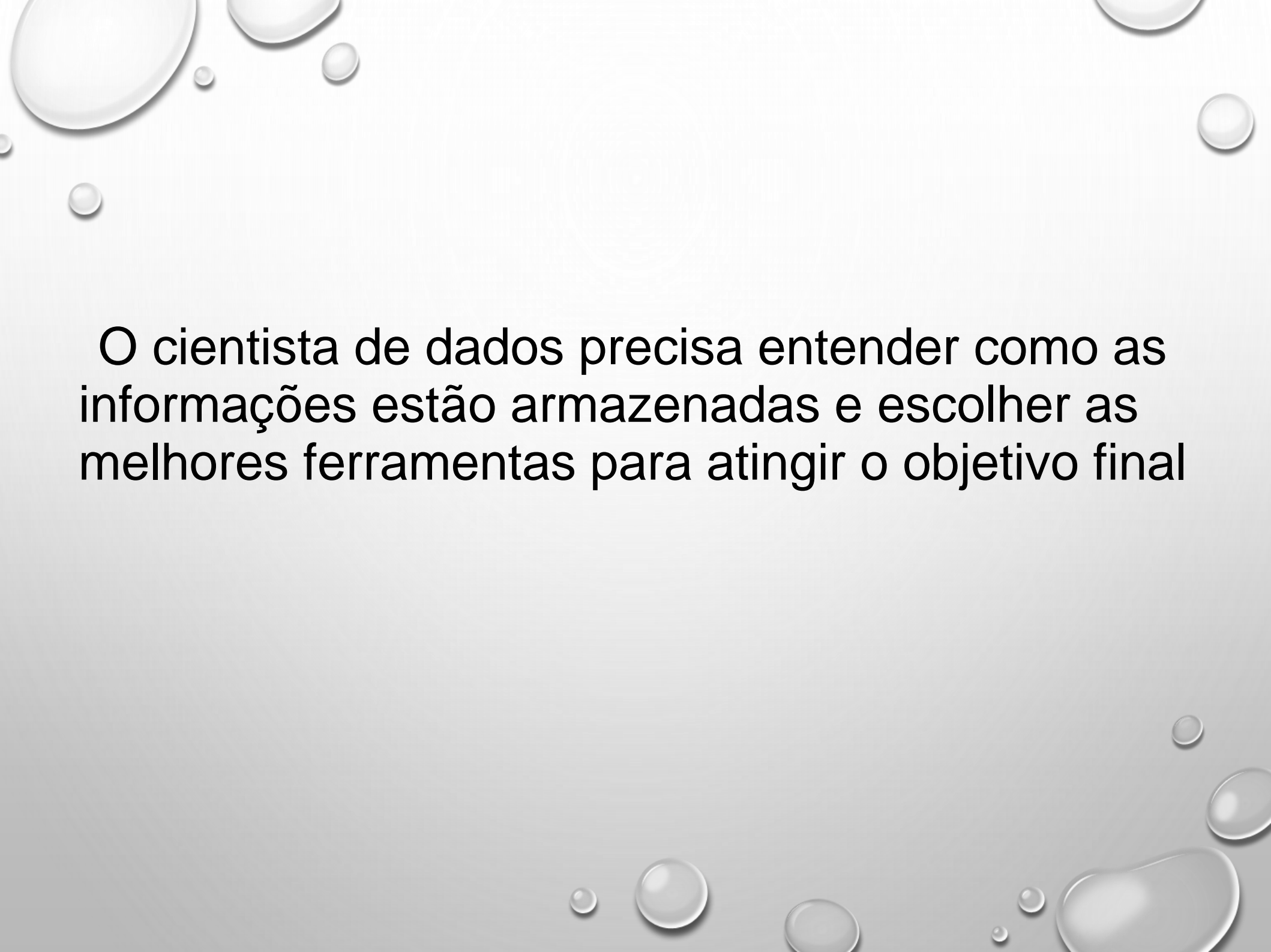


Cross Infrastructure / Analytics



Open Source Projects





O cientista de dados precisa entender como as informações estão armazenadas e escolher as melhores ferramentas para atingir o objetivo final

Hadoop

- Um framework que fornece uma maneira simples e escalável para processos distribuídos em diversos servidores;
- Hadoop é open source e inclui diversas ferramentas, como sistema de arquivos distribuído e Map/Reduce.

Big data Analytics

Informações em diversos formatos: Como as informações não precisam seguir uma modelagem, nenhuma informação é perdida;

Escalável: Hadoop já provou ser muito escalável, Facebook e Yahoo! são dois cases de sucesso;

Novas descobertas: Pela flexibilidade em cruzar informações, é simples conseguir novos insights, algo complicado quando está preso a uma modelagem;



Escalabilidade

A maioria das aplicações apresentam dificuldade para escalar ou exigem configurações complexas. O Hadoop escala facilmente, usando hardware comum, sem nenhuma necessidade de configurações complexas.

Alta Disponibilidade

O fato de ter sistema de arquivos distribuído, (HDFS) onde os dados não estão restritos em apenas um servidor, já faz o sistema com alta disponibilidade.

Empresas que investiram em Hadoop

- 1. Yahoo! - grande investidora de Hadoop;
- 2. Facebook - Utiliza para cruzar informações da rede social
- 3. Twitter - Utiliza para cruzar informações da rede social;
- 4. Adobe - Utiliza em seus sistemas de publicação online;
- 5. e-Bay - Utiliza para identificar tendências de compras;

Quem adota Big Data?

- Empresas de saúde
- Portais de serviço
- Redes Sociais
- Empresas financeiras
- Mineradoras
- Empresas de Segurança Pública

The background of the slide is a light gray gradient. It is decorated with several realistic water droplets of various sizes, primarily located in the top-left, top-right, and bottom-right corners. The droplets have highlights and shadows, giving them a three-dimensional appearance.

Casos de sucesso



NOVOS EXPERIMENTOS

Uso do big data na segurança pública é bem-vindo

01 de julho de 2014, 09:18h

 [Imprimir](#)

 [Enviar por email](#)

 [85](#)

 [23](#)

 [1](#)

[Por Elisa Mombelli](#)

Soluções de *big data*, quando aplicadas à prevenção de crimes, transformam em realidade aquilo que antes era ficção. O policiamento preditivo (*predictive policing*) é uma prática de segurança pública utilizada em diversos países, e que consiste, basicamente, na indicação de futuras e prováveis cenas criminosas com base em análise de dados que são obtidos de diversas fontes (até mesmo de redes sociais, suspeita-se).

Você está em: Home » Notícias » Pronunciamentos

▪ Mais Notícias

▪ Pronunciamentos

▪ Busca por Notícias

▪ Sala de Imprensa

- Press Releases
- Cadastro de Jornalistas
- Assessoria de Imprensa
- Assessorias de Imprensa dos Bancos
- Informe FB

Pronunciamentos

21/05/2013

KEYNOTES TRAIARÃO DE BIG DATA DURANTE O CIAB FEBRABAN 2013

Um dos principais itens de investimentos do setor financeiro brasileiro é a formação de Big Data e formas de manipular, cruzar e usar informações como vantagem competitiva. Este vai ser um dos temas principais do Ciab FEBRABAN 2013, maior evento para os setores financeiro e de tecnologia da América Latina, que será realizado nos dias 12, 13 e 14 de junho, em São Paulo. Entre os especialistas em Big Data estarão presentes Jason White (HP), Mike Huckaby (EMC) e Hyoung Kim (Ernst & Young).

BUSCA

BUSCAR



Sistema de Divulgação de Tarifas
de Serviços Financeiros da Febraban



Bancos usam Big Data para combater fraudes no Brasil

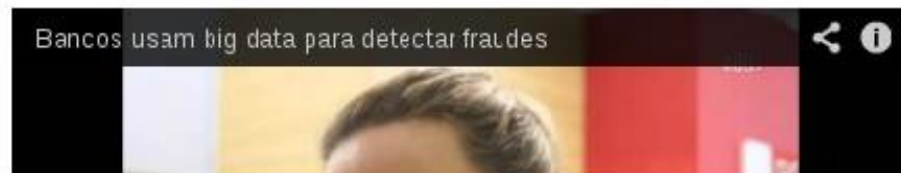
Convergência Digital - Hotsite Cloud Computing
:: Ana Paula Lobo e Roberta Prescott :: 10/06/2014



Com big data amadureceu no Brasil e as empresas de Telecomunicações e os Bancos são as que lideram o uso das soluções de análise, afirma Ana Claudia de Oliveira, responsável pelo desenvolvimento de negócios da Pivotal no Brasil. "As ofertas premium feitas pelas teles para os seus clientes já são formadas a partir de análise efetiva dos dados. Há um conhecimento da vontade do cliente", salienta a executiva.

Os bancos, por sua vez, contratam o big data como forma de detectar fraudes. "Cada vez mais os cibercriminosos estão sofisticados. E a análise de dados nos ajuda a combater a ação deles", salienta. O setor público também é uma vertice onde a análise de dados cresce, apesar da complexidade do dia a dia da gestão governamental.

Atuando no Brasil desde o final de 2012, a Pivotal, foi criada pela EMC, com investimentos conjuntos da VMware e da General Electric (GE). Assistam a entrevista concedida por Ana Cláudia Oliveira, da Pivotal Brasil, à CD TV, do portal Convergência Digital, durante o CIAB 2014.



Soluções de TI Simples

Patrocínio: IBM

Malware como serviço prolifera na Internet

No Brasil foram contabilizados mais de 30 mil ataques na modalidade

» Internet das Coisas movimentará US\$ 2 bilhões

» Cliente mal atendido custa US\$ 100 bilhões

» NFL marca um touchdown na área de TI



OPINIÃO

Como será a rede SDN na América Latina?

:: Por Hector Silva*

VÍDEOS

Mais vídeos



The background of the slide is a light gray gradient. It is decorated with several realistic water droplets of various sizes, primarily located in the top-left, top-right, and bottom-right corners. The droplets have highlights and shadows, giving them a three-dimensional appearance.

Muito obrigado!

Manoel Ribeiro
manoelra66@gmail.com