



Interações entre Big data e cloud computing



O que é *Big Data*?

Deivid Sardinha



➤ *Big Data*

- *Big Data* significa um grande volume de informações de alta variedade e velocidade que exige formas inovadoras e econômicas de análise para melhor tomada de decisões e automação de processos (GARTNER, 2019).

Figura 1 – Representação de dados e pessoas



Fonte: [katemangostar/Freepik.com](https://www.freepik.com/free-vector/illustration.htm#fromquery=katemangostar&from_view=detail&from_from=detail&from_id=11857&from_format=png&from_size=full)



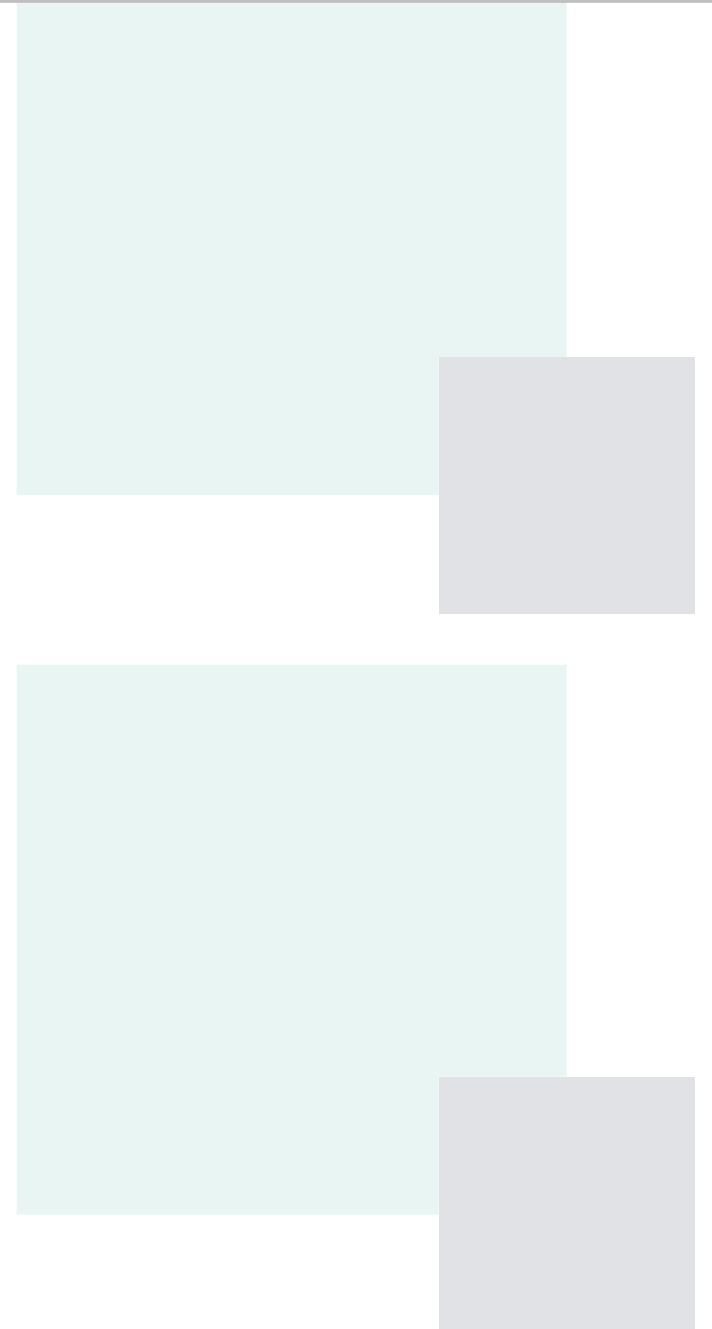
Uma breve história

- **1663** – Pesquisa sobre peste bubônica.
- **1980** – Primeiros equipamentos para processar dados.
- **1945** – Máquina digital de processamento de dados.
- **1965** – Criação dos primeiros centros de dados.
- **1989** – Criação da *Word Wide Web*.



Evolução do big data

- A virada no século 21.
- Sistemas de armazenamento.
- Difusão do IoT (*Internet of Things*).
- Evolução do hardware.



➤ *Cloud computing*

- A computação em nuvem (*cloud computing*) utiliza toda essa engrenagem que possibilita o compartilhamento de recursos pela internet.
- Com ela, após o ano de 1990, temos início da oferta de tecnologia como serviço, são elas:
 - IaaS.
 - SaaS.
 - PaaS.

HDFS (*Hadoop Distributed File System*)

Deivid Sardinha





Estrutura do *Big Data*

Figura 2 – Significado dos 5 Vs



Fonte: elaborada pelo autor.

➤ Interface e possibilidade

1. Entendendo e direcionando o cliente.
2. Entendendo e otimizando os processos de negócios.
3. Qualificação pessoal e otimização de desempenho.
4. Melhorar a saúde individual e a saúde pública.
5. Melhora no desempenho esportivo.

➤ Interface e possibilidade

6. Melhorando a ciência e a pesquisa.
7. Otimizando máquinas e desempenho de dispositivos.
8. Melhora da segurança e aplicação da lei.
9. Melhorando e otimizando as cidades.
10. Negociação financeira.

Teoria em Prática

Bloco 3

Deivid Sardinha



➤ Reflita sobre a seguinte situação

Existem muitos multiplicadores de dados, incluindo humanos, máquinas e processos de negócios, e o volume de dados cresce exponencialmente. Espera-se que os dados de saúde, seguros e os dados de fabricação cresçam enormemente a cada ano, sendo que mais de 80% desses dados são desestruturados e incapazes de serem processados por soluções existentes. As informações valiosas estão escondidas em documentos, e-mails, bate-papos, transcrições de centrais de atendimento, conteúdo de mídia social, comentários de clientes e relatórios de setor.

Enquanto a análise estruturada fornece o que, onde e quando de um desafio de negócios, análise de conteúdo não estruturada fornece o porquê e como. Isso ajuda as empresas a antecipar e identificar defeitos de produtos, melhorar o design de produtos, o gerenciamento de recursos e serviços, reduzir a rotatividade, identificar concorrentes e otimizar os gastos com marketing (REDDY, 2018). Forneça um exemplo e uma solução de uma situação em que dados podem ser analisados para resolver um problema.

Forneça um exemplo e uma solução de uma situação em que dados podem ser analisados para resolver um problema.



► Base de dados (*datasets*)

- Kaggle.
- Dados.gov.br.
- Portal da Transparência.
- CIFAR-10.



Dicas do(a) Professor(a)

Bloco 4

Deivid Sardinha



Leitura Fundamental

Prezado aluno, as indicações a seguir podem estar disponíveis em algum dos parceiros da nossa Biblioteca Virtual (faça o login através do seu AVA). Algumas indicações também podem estar disponíveis em sites acadêmicos como o Scielo, repositórios de instituições públicas, órgãos públicos, anais de eventos científicos ou periódicos científicos, acessíveis pela internet.

Isso não significa que o protagonismo da sua jornada de autodesenvolvimento deva mudar de foco. Reconhecemos que você é a autoridade máxima da sua própria vida e deve, portanto, assumir uma postura autônoma nos estudos e na construção da sua carreira profissional.

Por isso, te convidamos a explorar todas as possibilidades da nossa Biblioteca Virtual e além! Sucesso!





Indicação de leitura 1

O artigo sugerido apresenta os conceitos do *Big Data* e questiona quais problemas ele pode resolver. Embora haja muitos dados circulando no mundo, desde dados históricos, mídia social, sensor de IoT entre outros, o autor afirma que *Big Data* não é sobre dados, e sim sobre o seu significado. O conceito envolve um ecossistema de tecnologia que inclui coletar dados brutos de várias fontes, armazenar para uso posterior e gerar significado.

Referência:

PERRY, J., S. What is Big Data? More than volume, velocity and variety.... **IBM**, Armonk, 22 de maio de 2017.



Indicação de leitura 2

O artigo indicado traz uma revisão de conceitos que envolvem o *Big Data*, faz uma revisão das tecnologias envolvidas, tais como virtualização, processamento paralelo, Hadoop, *cloud computing*, IoT etc. Apresenta a cadeia de valor com início na geração de dados, seguida de armazenamento e análise. Discute os possíveis desafios envolvendo o *Big Data* e aborda as aplicações que tratam as redes sociais, inteligência coletiva e outras. É sugerida especialmente a leitura do item 5 *Big Data Analysis* em que alguns métodos de análise são apresentados.

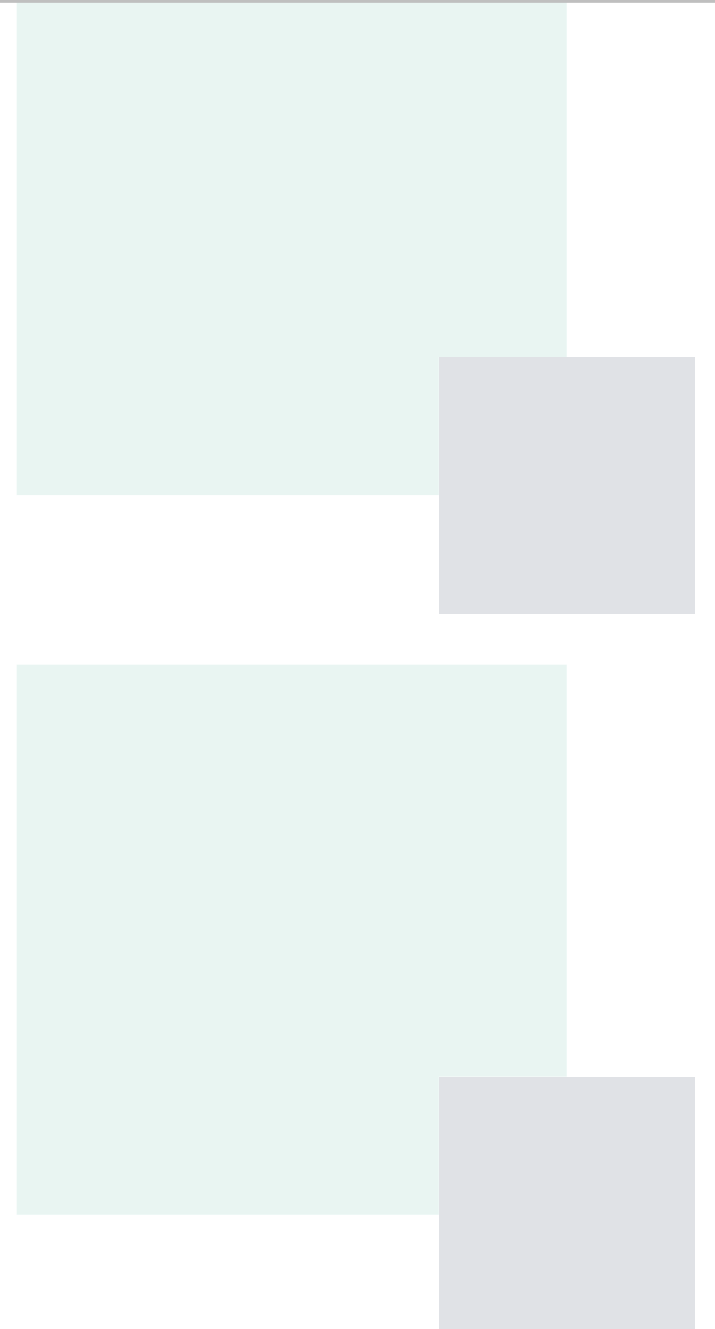
Referência:

CHEN, M.; MAO, S.; LIU, Y. Big Data: A Survey. **Mobile Netw Appl**, v. 19, p. 171-209, 2014.



Dica do(a) Professor(a)

O *Big Data* tem o princípio de que quanto mais se sabe de algo, mais confiança se tem para obter novos insights e fazer previsões sobre o futuro. Quando se comparam dados, relacionamentos ocultos começam a surgir, e eles permitem aprender a tomar decisões inteligentes. O processo envolve criação de modelos com base nos dados, execução de simulações, aprimoramento do valor e monitoramento dos resultados (MARR, 2019b).





Referências

MARR, B. Big Data: The 5Vs Everyone Must Know. **LinkedIn**, [s.l.], 6 de março de 2014. Disponível em: <https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know>. Acesso em: 24 jun. 2022.

PERRY, J., S. What is Big Data? More than volume, velocity and variety.... **IBM**, Armonk, 22 de maio de 2017. Disponível em: <https://developer.ibm.com/blogs/what-is-big-data-more-than-volume-velocity-and-variety/>. Acesso em: 24 jun. 2022.

PROVOST, F.; FAWCETT, T. Data Science for Business: What you need to know about Data Mining and Data-Analytic think. PDF Drive, [s.l.], [s.d.]. Disponível em: <https://www.pdfdrive.com/data-science-for-business-what-you-need-to-know-about-data-mining-anddata-analytic-thinking-d170193185.html>. Acesso em: 24 jun. 2022.

Bons estudos!

