



Analytics para área da saúde

Big Data



Big Data é um termo recente que remete ao processamento (em tempo real ou quase real) e análise de grande quantidade de dados. O termo foi primeiro usado por dois pesquisadores da NASA, Michael Cox e David Ellsworth no artigo Application-Controlled Demand Paging for Out-of-Core Visualization*

Exemplos:

- * Dinâmica de fluidos computacional.
- * Transações de cartões de crédito.

* <https://www.nas.nasa.gov/assets/pdf/techreports/1997/nas-97-010.pdf>



Os cinco Vs de Big Data: Volume

Grande quantidade de dados que são processadas a todo instante por diversos tipos de canais, empresas, aplicativos.

Exemplos:

- * Cadastros de CPF's.
- * Transações de cartões de crédito.





Os cinco Vs de Big Data: Velocidade

Velocidade com a qual informações são processadas.

Exemplos:

- * Modelos de risco.
- * Liberação de transações financeiras.
- * Sinais vitais de pacientes





Os cinco Vs de Big Data: Variedade

Capacidade de capturar diversos tipos de dados, sejam dados estruturados, tais como tabelas de dados cadastrais e sensores em plantas industriais ou dados não estruturados como imagens e áudio.

Exemplos:

- * Imagens capturadas por drones.
- * Áudio capturados por aparelhos eletrônicos.



Os cinco Vs de Big Data: Veracidade



Empresas devem se responsabilizar pelos dados que geram e adquirem, principalmente quando precisam responder para acionistas.

Exemplos:

- * Sinais vitais de pacientes internados.
- * Valores de ações da bolsa de valores.

Os cinco Vs de Big Data: Valor



Os dados gerados e armazenados devem ser relevantes e possuir significância estatística para que analistas e cientistas possam realizar análises.

Exemplos:

- * Dados da quantidade de infectados, mortos e recuperados por covid19.
- * Valores de ações da bovespa.

Computação em paralelo (MapReduce)



Para processar grandes quantidades de dados, utilizamos o conceito de MapReduce, dividimos o trabalho de processamento em diversos computadores (cluster) que realizam o processamento em paralelo.

Exemplos:

- * Hadoop File Systems (HDFS, Apache).
- * Google File Systems (GFS).

Computação em nuvem

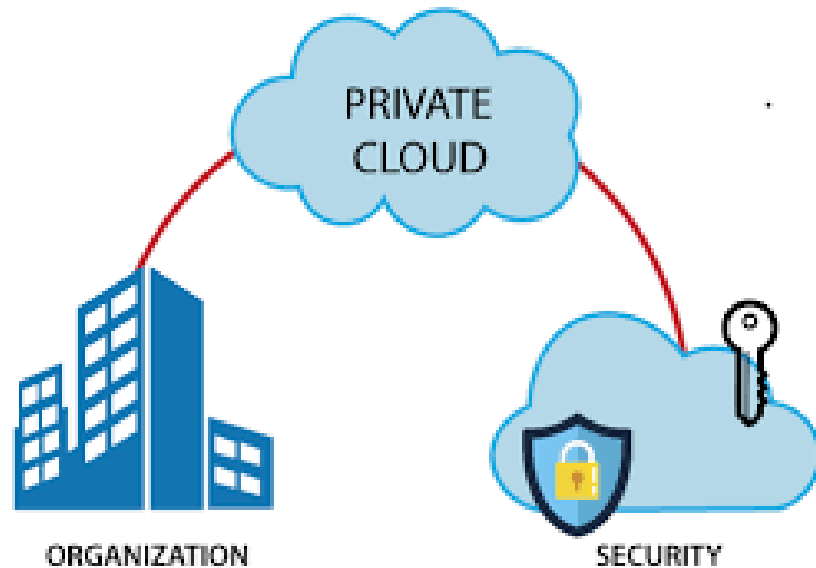


É o termo utilizado para designar a disponibilidade sob demanda de serviços conectados à internet. Os principais serviços sendo armazenamento de dados e processamento computacional em paralelo.

Exemplos:

- * Amazon Web Services.
- * Google Cloud Services.
- * IBM Cloud Computing.

Nuvem Privada

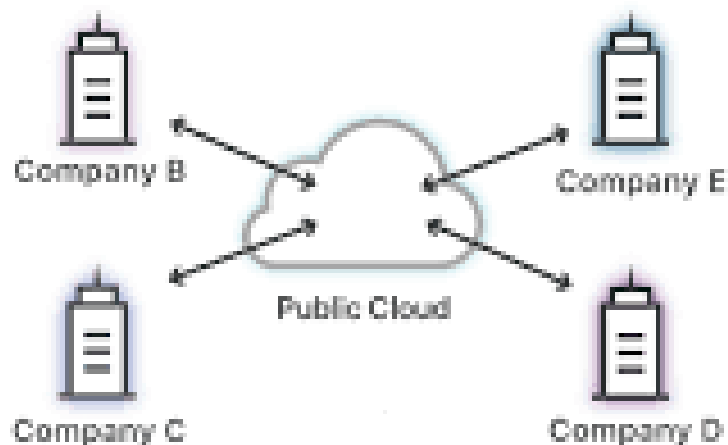


As nuvens privadas são aquelas construídas exclusivamente para um único utilizador. É uma infraestrutura em nuvem operada exclusivamente para uma única organização, e quase sempre operada pela própria organização ou por terceiros e hospedada interna ou externamente. Pode utilizar a internet ou um servidor interno.

Nuvem Pública



Public cloud shared
by multiple companies



As nuvens públicas possuem toda sua infraestrutura terceirizada operando localmente pelo provedor. Os serviços são oferecidos sob demanda para empresas através da internet. Esta é uma opção menos custosa que a nuvem privada, pois elimina a necessidade da compra, instalação e manutenção de computadores e servidores locais.



Nuvem Híbrida

Como o próprio nome diz é a combinação das nuvens privada e pública, utilizando infraestrutura local (on-premises) e terceirizada. A principal característica da nuvem híbrida é de ser eficiente em termos de custo, controle de risco e flexibilidade. Picos de demandas sazonais podem ser planejado e suprido pela nuvem pública enquanto que os assets mais sensíveis da empresa podem ser tratados localmente de forma segura e planejada.

