

Big Data

Big Data adalah istilah yang mengacu pada volume data yang sangat besar, baik terstruktur maupun tidak terstruktur, yang melebihi kemampuan perangkat lunak database tradisional untuk menangkap, mengelola, dan memproses data dalam waktu yang wajar. Karakteristik Big Data sering diringkas dengan "5V", yaitu:

1. **Volume:** Merujuk pada jumlah data yang sangat besar yang dihasilkan setiap hari oleh perusahaan, media sosial, perangkat sensor, dll. Ini adalah ciri paling dasar dari Big Data, yang menekankan pada skala data yang harus dikelola.
2. **Velocity** menunjukkan laju di mana data baru dihasilkan dan seberapa cepat data tersebut perlu diproses. Dalam banyak kasus, data harus dianalisis secara real-time atau hampir real-time, seperti data dari media sosial, sensor, atau transaksi online.
3. **Variety** menunjukkan berbagai jenis dan sumber data yang tersedia, termasuk data terstruktur (seperti database), tidak terstruktur (seperti teks, gambar, dan video), dan semi-terstruktur (seperti XML, JSON). Keanekaragaman ini membutuhkan teknik pengolahan yang berbeda dan lebih kompleks.
4. **Veracity** berkaitan dengan kualitas dan keakuratan data. Dalam kumpulan data besar, seringkali ada data yang tidak lengkap, tidak akurat, atau tidak relevan, yang dapat mengganggu proses analisis jika tidak ditangani dengan benar.
5. **Value** adalah nilai atau informasi yang bisa diperoleh dari pengolahan Big Data. Memiliki volume data yang besar dengan kecepatan tinggi, variasi, dan veracity tidak berguna tanpa kemampuan untuk mengubahnya menjadi nilai bisnis atau pengetahuan yang bermanfaat.