

BIG DATA DAN PEMANFAATANNYA DALAM BERBAGAI SEKTOR

Budi Maryanto

Sekolah Tinggi Manajemen Informatika dan Komputer LIKMI
Jl. Ir. H. Juanda 96 Bandung 40132

E-mail : budimailyanto@gmail.com

ABSTRAK

Big Data merupakan kumpulan data yang *volume* datanya super besar, memiliki keragaman sumber data yang tinggi, sehingga perlu dikelola dengan metode dan perangkat bantu yang kinerjanya sesuai.

Artikel ini akan diawali dengan pengertian dan karakteristik Big Data. Selanjutnya akan dipaparkan faktor-faktor yang memicu perkembangan pesatnya, klasifikasi data yang terdapat dalam Big Data, serta tahapan untuk mengelolanya. Bagian akhir artikel memuat berbagai contoh pemanfaatan teknologi Big Data dalam sektor bisnis dan sektor layanan publik.

Kata-kata kunci : big data, volume-velocity-variety.

1 PENDAHULUAN

Big Data adalah istilah yang diberikan pada kumpulan data yang beukuran sangat besar dan kompleks, sehingga tidak memungkinkan untuk diproses menggunakan perangkat pengelola database konvensional ataupun aplikasi pemroses data lainnya.

Dalam Gartner IT Glossary, Big Data didefinisikan sebagai berikut [1] : *Big Data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation.*

Dengan mengacu pada definisi tersebut, maka dapat disimpulkan bahwa karakteristik utama Big Data meliputi tiga hal – biasa disingkat dengan 3V – yaitu *volume*, *velocity*, dan *variety*. *Volume* terkait dengan besaran data yang harus dikelola berukuran super besar. *Velocity* berkenaan dengan kecepatan pemrosesan data yang harus mengimbangi pesatnya pertumbuhan jumlah data. Sedangkan *variety* merujuk pada

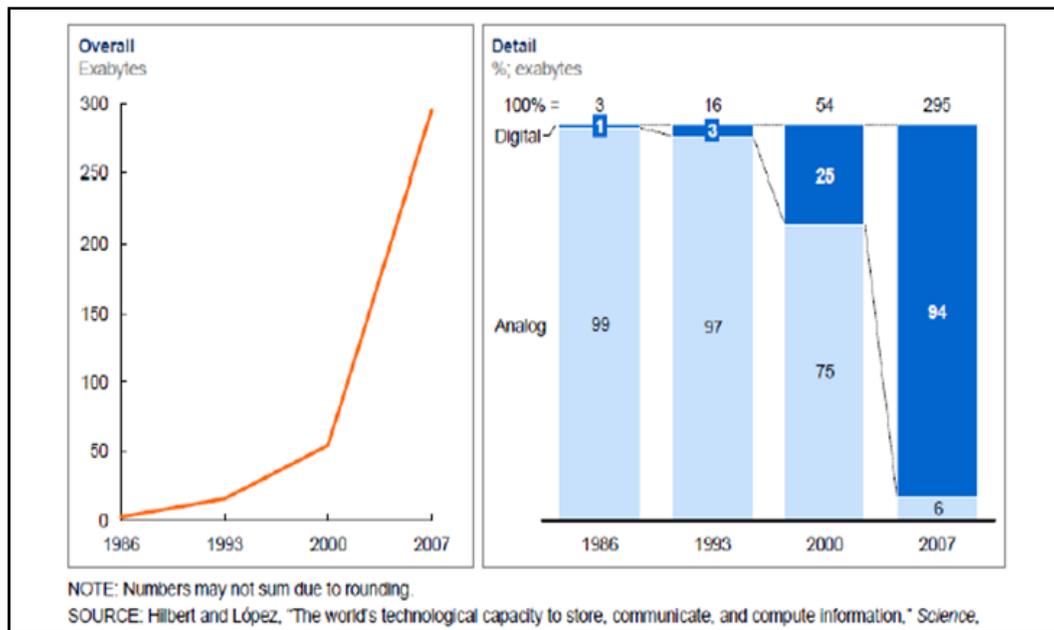
karakteristik sumber data yang sangat beragam, baik itu yang berasal dari basis data yang terstruktur maupun juga dari data-data yang tidak terstruktur.

2 PEMICU PERKEMBANGAN BIG DATA

Menurut Hilbert dan Lopez, ada tiga hal utama yang memicu perkembangan teknologi Big Data [2] :

a. Pesatnya pertumbuhan kemampuan penyimpanan data.

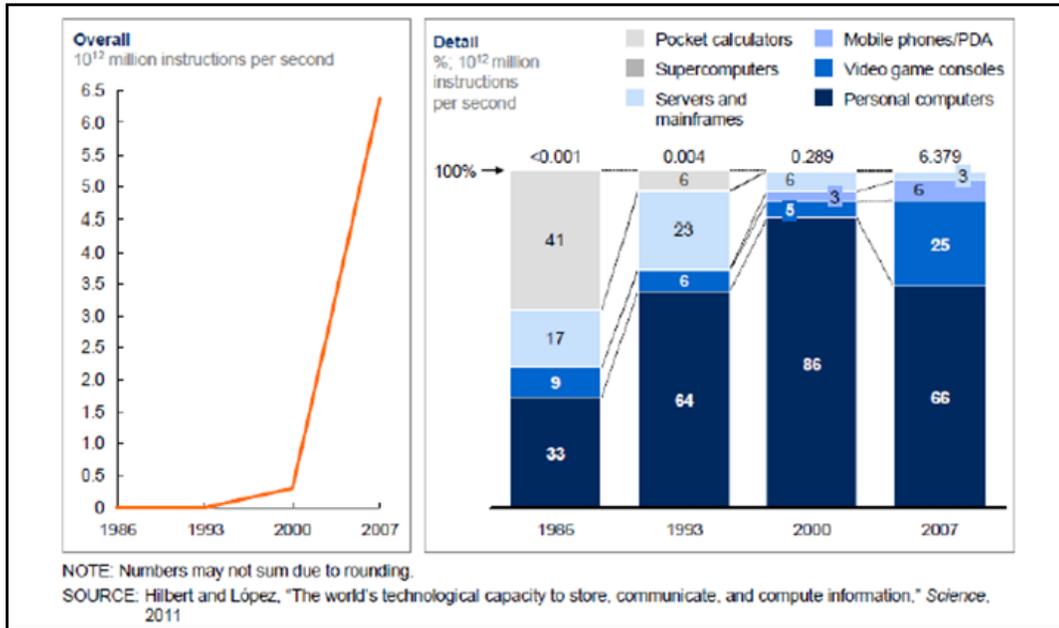
Kemampuan penyimpanan data telah bertumbuh sangat signifikan. Grafik sebelah kanan pada Gambar 1 memperlihatkan pergeseran penyimpanan data dari bentuk analog menjadi digital.



Gambar 1
Grafik Pertumbuhan Data Storage [2]

b. Pesatnya pertumbuhan kemampuan mesin pemrosesan data.

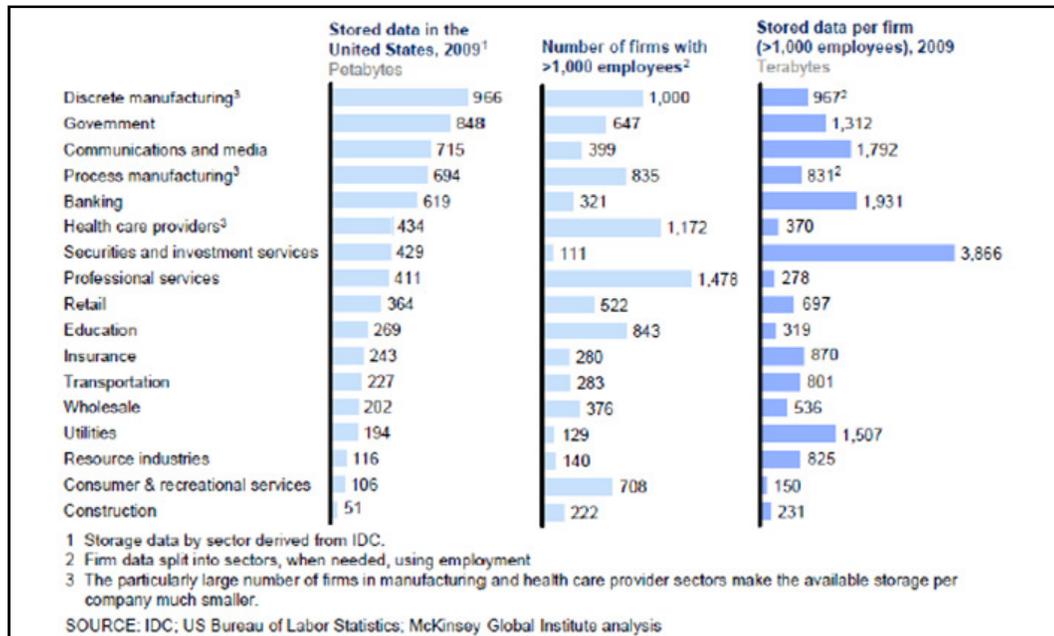
Seiring dengan pesatnya perkembangan teknologi *hardware*, maka kapasitas komputasi pada mesin/ perangkat komputer juga telah meningkat sangat tajam.



Gambar 2
Grafik Pertumbuhan Kapasitas Komputasi [2]

c. Ketersediaan data yang melimpah.

Perusahaan-perusahaan dari berbagai sektor di Amerika Serikat memiliki data setidaknya 100 terabytes. Bahkan banyak diantara perusahaan tersebut yang memiliki data lebih dari 1 petabyte.



Gambar 3
Grafik Ketersediaan Data dari Berbagai Sektor di AS [2]

3 KLASIFIKASI DATA

Teknologi Big Data memiliki kemampuan untuk menangani berbagai variasi data. Secara umum ada 2 kelompok data yang harus dikelola, yaitu :

a. **Data terstruktur**

Kelompok data yang memiliki tipe data, format, dan struktur yang telah terdefinisi. Sumber datanya dapat berupa data transaksional, OLAP data, tradisional RDBMS, file CSV, spreadsheets

b. **Data tidak terstruktur**

Kelompok data tekstual dengan format tidak menentu atau tidak memiliki struktur melekat, sehingga untuk menjadikannya data terstruktur membutuhkan usaha, *tools*, dan waktu yang lebih. Data ini dihasilkan oleh aplikasi-aplikasi internet, seperti data URL log, media sosial, e-mail, blog, video, dan audio.

4 TAHAPAN PENGELOLAAN BIG DATA

Berikut ini adalah 4 tahap pengelolaan Big Data serta perangkat bantu (*tools*) yang dapat dimanfaatkan untuk mendukung pemrosesan pada tiap tahap [3] :

a. **Acquired**

Berhubungan dengan sumber dan cara mendapatkan data.

b. **Accessed**

Berhubungan dengan daya akses data; data yang sudah dikumpulkan memerlukan tata kelola, integrasi, *storage* dan *computing* agar dapat dikelola untuk tahap berikutnya. Perangkat untuk pemrosesan (*processing tools*) menggunakan Hadoop, Nvidia CUDA, Twitter Storm, dan GraphLab. Sedangkan untuk manajemen penyimpanan data (*storage tools*) menggunakan Neo4J, Titan, dan HDFS.

c. **Analytic**

Berhubungan dengan informasi yang akan didapatkan, hasil pengelolaan data yang telah diproses. Analitik yang dilakukan dapat berupa *descriptive* (penggambaran data), *diagnostic* (mencari sebab akibat berdasar data), *predictive* (memprediksi kejadian dimasa depan) maupun *prescriptive analytics* (merekomendasikan pilihan dan implikasi dari setiap opsi).

Tools untuk tahap analitik menggunakan MLPACK dan Mahout.

d. **Application**

Terkait visualisasi dan *reporting* hasil dari analitik. *Tools* untuk tahap ini menggunakan RStudio.

5 PEMANFAATAN BIG DATA PADA SEKTOR BISNIS

Perusahaan-perusahaan yang bergerak dalam sektor bisnis memiliki orientasi utama pada pencapaian margin laba setinggi mungkin (*profit oriented*). Berbagai informasi penting dapat dihasilkan dari Big Data yang dapat mendukung proses pengambilan keputusan bagi pimpinan perusahaan sebagai berikut [3] :

- a. Mengetahui respons masyarakat terhadap produk-produk yang dikeluarkan melalui analisis sentimen di media sosial.
- b. Membantu perusahaan mengambil keputusan secara lebih tepat dan akurat berdasarkan data
- c. Membantu meningkatkan citra perusahaan di mata pelanggan.
- d. Perencanaan usaha, dengan mengetahui perilaku pelanggan seperti pada perusahaan telekomunikasi dan perbankan.
- e. Mengetahui trend pasar dan keinginan konsumen.

5 PEMANFAATAN BIG DATA PADA SEKTOR LAYANAN PUBLIK

Perusahaan atau institusi yang berada pada sektor layanan publik lazimnya memiliki orientasi utama pada pencapaian kepuasan klien/ pelanggan. *Resource* Big Data dapat memberikan andil dengan menyajikan berbagai informasi berharga sebagai berikut [3] :

- a. Mendapatkan *feedback* dan respon masyarakat sebagai dasar penyusunan kebijakan dan perbaikan pelayanan publik. *Feedback* tersebut dapat diperoleh dari sistem informasi layanan pemerintah maupun dari media sosial.
- b. Membuat layanan terpadu dengan segmen khusus sehingga layanan bisa lebih efektif dan efisien.
- c. Menemukan solusi atas permasalahan yang ada, berdasarkan data. Sebagai contoh : menganalisis informasi cuaca dan informasi pertanian terkait data tingkat kesuburan tanah, pemerintah dapat menetapkan atau menghimbau jenis varietas tanaman yang ditanam oleh petani pada daerah dan waktu tertentu.

6 KESIMPULAN

Big Data memiliki karakteristik 3V : *Volume* (berukuran sangat besar), *Variety* (datanya sangat beragam), *Velocity* (memiliki kecepatan akses data yang memadai). Sumber data untuk Big Data bisa berupa basis data yang terstruktur ataupun data-data yang tidak terstruktur. Ada 4 tahap yang harus dilalui untuk memproses Big Data hingga dihasilkannya output informasi bagi *user*, yaitu : *acquired*, *accessed*, *analytic*, dan *application*.

Manfaat teknologi Big Data telah dirasakan secara luas dalam berbagai sektor. Perusahaan-perusahaan yang bergerak pada sektor bisnis dapat memanfaatkan informasi-informasi berharga yang dihasilkan Big Data untuk mengoptimalkan proses pengambilan keputusan, agar target memaksimalkan raihan *profit* dapat tercapai. Sedangkan institusi yang bergerak di bidang layanan publik dapat menggunakan output informasi dari Big Data untuk memaksimalkan tingkat kepuasan layanan kepada klien/ pelanggannya.

DAFTAR PUSTAKA

- [1] *The Gartner IT Glossary : What is Big Data?*
<https://www.gartner.com>, diakses 2 Juni 2007.
- [2] Hilbert and Lopez, 2011, *The World's Technological Capacity to Store, Communicate, and Compute Information*, Science.
- [3] Tim Penyusun Kominfo, 2015, *Buku Saku Big Data*, Kementerian Komunikasi dan Informatika.