

Big Data

Anjali Saputra¹, Muhammad Isa Firdaus², Riko Wahyudi³, Lorina Mohdo⁴,
Muhammad Erlangga Gunawan⁵, Muhammad Encep, Mulil Khaira

Program Studi Ilmu Komputer, Universitas Djuanda Bogor;

¹Annnxjs@gmail.com; ²muhamadisafirdaus123@gmail.com; ³rikowahyudi36@gmail.com;
⁴lorin.mohdo@gmail.com; ⁵gunawanerlanggamuhammad@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 Definisi, Konsep dan Manfaat Big Data. Selanjutnya akan dijelaskan Sejarah Big Data, Penerapan Big Data, serta tahapan untuk mengelolanya. Dibagian akhir terdapat sebuah kesimpulan mengenai Big Data.

Kata kunci: big data, kecepatan, volume, variasi

Abstract

Big Data is a large data collection with a high diversity of data sources, so it needs to be managed with appropriate methods and tools. This article will begin with the Definition, Concept, and Benefits of Big Data. Next, we will explain the history of big data, the application of big data, and the steps for managing it. In the end, there is a conclusion regarding Big Data.

Keywords: big data, velocity, volume, variety

I. 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. (Palatino Linotype, 12, spasi 1, rata tengah)

II. DEFINISI BIG DATA

Big Data adalah kombinasi teknologi yang bisa memanajemen data yang mempunyai variasi dalam jumlah besar, kecepatan yang tepat, dan memiliki ketepatan saat melakukan analisis dan reaksi.

Bisa dilihat pada bidang teknologi, banyak manfaat yang diperoleh dalam memproses big data seperti teknik akses dan penyimpanan data menggunakan *key value store* (KVS) dan teknik komputasi paralel menggunakan *MapReduce* (sujana, 2013).

Konsep ini mulai muncul sebagai tren dalam pengolahan informasi dalam kurun 5 tahun terakhir mengingat begitu besarnya pertumbuhan data diinternet, khususnya melalui media sosial.

Secara umum Big Data dapat dimanfaatkan sebagai sebuah kumpulan data yang berukuran sangat besar (volume), sangat cepat berubah (velocity), hadir dalam beberapa bentuk atau format (veriety), serta memiliki nilai tertentu (value), dengan catatan jika berasal dari sumber yang akurat (veracity).

Big Data merupakan istilah umum untuk sekumpulan data yang jumlahnya sangat besar dan kompleks sehingga tidak mudah untuk ditangani atau proses hanya dengan metode pemrosesan data biasa. Terdapat tiga masalah utama yang diselesaikan oleh big data, antara lain (Pawitra, 2016):

1) *Volume*

Ukuran data yang disimpan atau diproses.

2) *Velocity*

Kecepatan membuat data. Kecepatan data dibuat umumnya berbanding lurus dengan volume data.

3) *Variety*

Keberagaman data yang diolah. Dari segi format maupun struktur data.

III. KONSEP BIG DATA

Big Data memiliki konsep untuk mengumpulkan seluruh data yang dihasilkan dan kemudian mengolahnya dengan tepat agar menghasilkan sebuah nilai diharapkan.

Konsep big data terbagi menjadi tiga macam, yaitu:.

1) *Integrasi Data*

Integrasi data merupakan suatu proses untuk mengumpulkan data yang telah dibuat yang akan menjadi **BIG DATA**.

2) *Pengolahan Data*

Seluruh data yang dihasilkan harus bisa dikelola dengan benar, baik saat proses penyimpanan maupun saat data akan diakses kembali.

3) *Analisis Data*

Terakhir adalah analisa data, yaitu semua data yang telah disimpan dan dikelompokkan sesuai jenisnya dapat kamu gunakan untuk kebutuhan selanjutnya.

IV. MANFAAT BIG DATA

Jika kamu sudah mengetahui jenis, dan konsep dari sebuah *big data*. Kamu juga perlu memahami apa saja manfaat yang ditimbulkan dari penggunaan *big data* bagi penggunaannya. Empat manfaat tersebut antara lain:

1) *Meningkatkan Inovasi*

Dengan menggunakan big data, kamu dapat mengenali minat konsumen, bisa dilihat melalui jenis produk atau dengan banyaknya transaksi pada suatu produk.

2) *Merancang UX untuk yang User-friendly*

Interaksi yang banyak dilakukan oleh konsumen melalui aplikasi atau *website* bisa menentukan tingkat penjualan produknya. Dengan adanya big data, akan membantu kamu menyimpan riwayat belanja dari konsumen yang datang.

3) *Pengembangan produk lebih mudah*

BIG DATA dapat membantu perusahaan untuk mengembangkan produk yang memiliki nilai dan kualitas lebih baik serta sesuai dengan kebutuhan konsumen, sehingga proses penjualan dapat ditingkatkan lagi.

4) *Menghindari Risiko Manipulasi Data*

Tidak hanya ketiga manfaat yang sudah disebutkan di atas, ternyata **BIG DATA** juga memiliki peran terhadap masalah keamanan. Dengan kondisi data yang besar, kamu bisa mengidentifikasi pola data yang tidak sesuai dengan data yang seharusnya dan dinilai bisa merugikan bisnis yang kamu kelola.

V. SEJARAH BIG DATA

1) *Awal mula sejarah big data dimulai pada tahun 1663*

Sejarah big data di mulai pada tahun 1663 Jhon Graunt mencatat dan memeriksa semua informasi tentang penyebab kematian di London. Jhon ingin mendapatkan pemahaman dan membangun sistem peringatan untuk wabah pes yang sedang berlangsung saat itu. Dalam catatan analisis data statistik pertamanya yang tercatat ia mengumpulkan temuannya dalam buku *Natural and Political Observation Made on the Bills of Mortality*, yang memberikan wawasan besar tentang penyebab kematian pada abad 17, karna karyanya tersebut Jhon Graunt dapat dianggap sebagai bapak statistik.

2) *Era Tahun 1887*

Herman Hollerith menemukan mesin komputasi untuk dapat membaca lubang yang di buat pada kertas untuk mengatur data sensus.

3) *Tahun 1937*

Proyek mengenai data yang cukup besar pertama dibuat pada tahun 1937 milik Administrasi Frangklin D. Roosevelt di Amerika serikat. Proyek ini muncul setelah undang-undang jamnian sosial menjadi hukum yang berlaku pada tahun 1937, akibat proyek ini pemerintah harus melacak kontribusi 26 juta orang amerika dan lebih dari 3 juta pemberi kerja. Dar itu IBM dipercaya untuk menyelesaikan proyak besar dengan mebuat mesin hole punch card.

4) *Tahun 1943*

Mesin pengelola data pertama kali muncul pada tahun 1943 yang dikembangkan oleh inggris untuk memecahkan kode tentara nazi selama perang dunia ke dua, perangkat ini di beri nama Colossus yang bertugas untuk mencari pola dalam pesan yang disadap oleh inggris. Perangkat ini dapat membaca 5000 karakter perdetik yang mampu mengurangi waktu pengerjaan yang tadinya membutuhkan berminggu-minggu hanya menjadi dlam hitungan jam.

5) *Tahun 1965*

Pada tahun 1965 pemerintah amerika serikat memutuskan membangun pusat data pertama untuk menyimpan lebih dari 742 juta pengembalian pajak dan 175 juta set sidik jari dengan menteransfer semua catatan tersebut ke pita komputer magnetik yang harus disimpan di satu sisi lokasi.

6) *Perkembangan big data yang cukup pesat pada tahun 1989*

Sejarah big data berkembang pada tahun 1989 karna di tahun ini ilmuan komputer inggris Tim Berners-lee menciptakan word wide web, bertujuan untk memfsilitasi proses berbagi informasi menggunakan sistem 'hypertext'.

7) *Tahun 1995*

Tahun 1995 data yang ada di dunia sangat banyak karna semakiin banyak perangkat yang terhubung ke internet baik itu Iot devices ataupun berbentuk pc. Pada tahun in super komputer pertama dibangun yang mampu melakukan banyak pekerjaan dalam satu detik daipada kalkulator yang dioperasikan oleh satu orang dalam 30.000 tahun.

8) *Tahun 2005*

Pada tahun ini roger Mougals dari O'reilly median menciptakan istilah Big Data untuk pertama kali, bigdata ini mengacu pada kumpulan data besat yang hampir tidak mungkin untuk dikelola dan diperoses menggunakan alat intelegn bisnis tradisional.

Pada tahun 2005 juga merupakan tahun dimana Hadoop diciptakan oleh Yahoo di bangun diatas google Mapreduce tujuannya untuk mengindek seluruh world wide web dan saat ini hadoop open-source digunakan oleh banyak organisasi untuk mengelola data dalam jumlah besar.

9) *Munculnya data base biometrik terbesar didunia pada sejarah big data tahun 2009*

Pada tahun 2009 pemerintah india memutuskan untuk melakukan pemindaian iris mata, sidik jari, dan foto dari 1,2 miliar penduduk, semua data ini disimpan dalam data base biometri terbesar didunia

10) *Perkembangan era tahun 2010*

Eric schmidt berbicara pada konfrensi Teknologi di lake tahoe di california , ia menyatakan bahwa “ada 5 exbyte informasi yang diciptakan oleh seluruh dunia antara awal peradaban”

11) *Perkembangan tahun 2011*

Pada tahun 2011 laporan McKinsey tentang big data “The next frontier for innovation, competition and productivity”, menyatakan bahwa pada tahun 2018 Amerika serikat akan menghadapi kekurangan 140.00-190.00 ilmuan data scientist dan juga 1,5 juta data manager.

Pada tahun yang sama facebook meluncurkan open compute project untuk membagikan spesifikasi untuk data canter yang menggunakan energi secara efisien.

VI. PEMICU PERKEMBANGAN BIG DATA

Menurut para ahli, terdapat 3 hal utama yang memicu perkembangan teknologi Big Data diantaranya:

1) *Pesatnya Perkembangan Kemampuan Penyimpanan Data*

Kemampuan penyimpanan data telah bertumbuh sangat signifikan. Kemampuan teknologi penyimpanan data telah berkembang pesat setiap hari. Perkembangan teknologi penyimpanan data pada awalnya berbentuk analog dan sekarang telah berubah menjadi bentuk digital akibat perkembangan teknologi.

2) *Pesatnya Perkembangan Kemampuan Mesin Pemrosesan Data*

Seiring dengan pesatnya perkembangan teknologi pada *hardware*, maka kapasitas komputasi pada mesin atau perangkat komputer juga telah meningkat dengan sangat pesat. Hal tersebut dapat dilihat dari perkembangan teknologi baru yang banyak bermunculan.

3) *Ketersediaan Data yang Melimpah*

Perusahaan-perusahaan yang ada diberbagai sektor pada saat ini memiliki data yang begitu banyak, hal tersebut untuk bersaing pada pasar. 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.

VII. KLASIFIKASI BIG DATA

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

1) *Data Terstruktur*

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

2) *Data Tidak Terstruktur*

Merupakan kelompok data tekstual dengan format yang tidak menentu atau tidak memiliki struktur yang 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 social, e-mail, blog, audio dan video.

VIII. PENERAPAN BIG DATA

Saat ini kita berada pada era digital dimana sebuah data menjadi hal yang sangat penting baik itu personal maupun komersial, perusahaan tentunya memiliki data dan informasi dimana semakin besar sebuah perusahaan maka semakin besar pula data dan informasi yang dimiliki oleh perusahaan tersebut sehingga banyak perusahaan yang rela membayar untuk berinvestasi menyimpan seluruh datanya. Dan kemudian itulah yang dimaksud dengan big data, lalu bagaimana contoh penerapan atau implementasi big data tersebut? berikut beberapa contoh penerapan dan implementasi big data antara lain:

1) *Big Data dalam bidang Pendidikan*

Memanfaatkan dan menggunakan teknologi Big Data bukanlah hal yang baru dan mustahil untuk dilakukan karena faktanya big data ini sudah diimplementasikan di beberapa intitusi. Implementasinya yaitu melalui sebuah software kita dapat mengetahui perkembangan pembelajaran siswa dengan cara menganalisis penekanan pada mereka. Software ini akan mengumpulkan sebuah informasi dari perangkat yang digunakan oleh siswa, mengumpulkan nilai dan keterampilan belajar. Dengan mengimplementasikannya, Big Data mampu menyesuaikan dengan kebutuhan Pendidikan dan proses belajar siswa, selain itu big data juga mampu melihat potensi

seorang siswa sehingga mempengaruhi pengajaran yang dilakukan oleh tenaga pengajar sehingga bisa memiliki banyak waktu untuk melihat kebutuhan seorang siswa dan mendukungnya secara individu.

Sistem Pendidikan yang ada saat ini bukan hal yang mustahil berubah secara massif yaitu dengan meningkatnya inovasi teknologi dan peningkatan mobilitas siswa di dunia pendidikan. Dimana Massive Open Online Course (MOOC) menjadi sangat populer dan tercatat lebih dari 70 intitusi di Amerika Serikat menawarkan kursus MOOC sekaligus menjadi salah satu flatform pendidikan online terbesar didunia. Dan hingga saat ini Eropa telah menerapkan pengajaran online dan menggunakan implementasi teknologi Big Data untuk meningkatkan daya saingnya.

Dalam E – learning tentu akan menghasilkan sebuah data yang harus disimpan dan dikelola, solusi dari hal tersebut adalah dengan memanfaatkan sistem Big Data dimana data yang disimpan akurat, terstruktur dan data yang dihasilkan kompleks sehingga membutuhkan sebuah teknologi komputer yang canggih.

Dalam penerapan Big Data lembaga pendidikan seperti prestasi individu pelajar, proses belajar dengan baik dan mudah apabila wali yang ingin melihat data tentang siswa secara akurat kapanpun serta tanpa dibatasi oleh waktu dan jarak.

2) *Big Data dalam bidang komunikasi krisis*

Untuk pelaksanaan membuat sebuah kebijakan pemerintah dapat mempercepat kegiatan tersebut dengan memanfaatkan teknologi dari Big Data. Pemerintah dapat memberdayakan warga untuk meningkatkan tranparansi dan partisipasi semua pemilik kepentingan melalui sebuah data yang tersedia dalam sistem Big Data. Dalam mengambil sebuah keputusan pemerintah perlu mengambil sebuah langkah yang efisien dan efektif seperti mengelola data – data yang tersedia menjadi sebuah informasi yang berguna dimana nantinya akan berguna untuk pertimbangan dalam membuat sebuah keputusan. Selain itu cara yang paling efektif dan efisien yang sangat dibutuhkan oleh pemerintahan saat ini yaitu dengan cara memanfaatkan media (Big Data).

3) *Big Data Analytic*

Aktivitas mahasiswa seperti registrasi perkuliahan, pembayaran, partisipasi di kelas, belajar online dan penilaian dapat dianalisis secara real time berdasarkan pengalaman mahasiswa menggunakan Big data analytic.

Learning analytic, academic analytic dan proses mining merupakan contoh penggunaan big data didunia pendidikan. Learning analytic yaitu untuk memprediksi mahasiswa sukses dan mahasiswa beresiko akademik dengan cara melakukan analisis data pembelajaran secara real time. Dengan sistem ini perguruan tinggi dapat melakukan sebuah tindakan preventif atau dukungan tambahan untuk meningkatkan

keberhasilan dan memberikan keyakinan kepada mereka mahasiswa yang beresiko dropping out sebelum benar-benar increase results.

Learning analytic bertujuan untuk memberdayakan pendidikan serta membuka pintu pengalaman belajar yang bermakna sehingga dapat melibatkan, menginspirasi dan menghasilkan siswa yang akan sukses untuk masa kini dan masa, yaitu dengan cara menyediakan tools, teknologi dan platform yang tersedia pada learning analytic. Dengan adanya Learning analytic ini diharapkan dapat meningkatkan persentase keberhasilan dari mahasiswa dan institusi.

4) *Big Data Dalam Internet*

Kita tentu tidak bisa lepas dari internet, dimana internet sudah mendarah daging dan setiap hari kita terhubung dengan internet. Dalam penggunaannya tentu kita akan memakai sebuah search engine dan yang paling banyak digunakan saat ini yaitu Google, pada mesin pencarian ini kita bisa melakukan pencarian tentang berbagai macam informasi dan data hasil pencarian yang kita lakukan akan disimpan dalam sistem big data milik Google.

5) *Big Data dalam Media Sosial*

Siapa yang tidak kenal dengan media sosial, hampir seluruh manusia di bumi ini menggunakan layanan satu ini dimana kita bisa mengunggah segala macam hal baik itu berupa teks, images, audio, video. dan lainnya dimana data – data tersebut akan disimpan dalam big data milik perusahaan media sosial tersebut.

IX. TAHAPAN PENGOLAHAN BIG DATA

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

1) *Acquired*

Berhubungan dengan sumber dan cara mendapatkan data.

2) *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.

3) *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. 18 Budi Maryanto Big Data dan Pemanfaatannya dalam Berbagai Sektor.

4) *Application*

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

X. METODOLOGI (PALATINO LINOTYPE, 12, BOLD)

Penelitian ini adalah penelitian kepustakaan (library research) dengan menggunakan metode analisis deskriptif, yaitu dengan jalan mengumpulkan data, menyusun atau mengklarifikasi, menyusun dan menginterpretasinya. Analisis deskriptif ini dilakukan dengan mula-mula mengumpulkan data, mengklasifikasi data, lalu merumuskan kaedah-kaedah terhadap keteraturan yang terdapat pada keteraturan data itu.

XI. HASIL DAN DISKUSI (PALATINO LINOTYPE, 12, BOLD)

Setelah mengumpulkan data/informasi mengenai big data, didapati 5 hal berikut yaitu:

- 1) *Definisi, Konsep dan Manfaat Big Data*
- 2) *Sejarah Big Data*
- 3) *Pemicu Perkembangan Big Data*
- 4) *Penerapan Big Data*
- 5) *Tahapan pengelolaan Big Data*

5 hal diatas merupakan topik tentang dengan big data yang dicantumkan dalam artikel ini setelah kami mendiskusikan bagian penting dari topik big data yang telah dicari dan dikumpulkan.

XII. KESIMPULAN (PALATINO LINOTYPE, 12, BOLD)

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.

REFERENSI

- The Gartner IT Glossary: What is Big Data? <https://www.gartner.com>, diakses 2 Juni 2007.
- Hilbert and Lopez, 2011, *The World's Technological Capacity to Store, Communicate and Compute Information*, Science.
- Tim Penyusun Kominfo, 2015, *Buku Saku Big Data*, Kementerian Komunikasi dan Informatika.
- Aryanto, Andre Febrian. 2020. "Implementasi Big Data Dalam E- Learning Implementation of Big Data in E-Learning."
- Efgivia, Mohammad Givi. 2020. "Pemanfaatan Big Data Dalam Penelitian Teknologi Pendidikan." *Educate: Jurnal Teknologi Pendidikan* 5(2): 107–19.
- Solihin, Olih. 2021. "Implementasi Big Data Pada Sosial Media Sebagai Strategi Komunikasi Krisis Pemerintah." *Jurnal Common* 5(1): 56–66.
- Vience Mutiara Rumata, "Peluang dan Tantangan Big Data dalam Penelitian Ilmu Sosial: sebuah Kajian Literatur", *Jurnal Penelitian Komunikasi dan Opini Publik*, Vol 20, No 1. Agustus 2016.
- Haines Toba, "Big Data: Menuju Evaluasi Era Informasi Selanjutnya", Conference paper, April 2015, ([www.Researchgate.net/publication/2747126686](http://www.researchgate.net/publication/2747126686))
- Aleksandra Krajnovic, "Digital Marketing and behavioral Economics", Conference paper, October 2016, <http://www.researchgate.net/publication/309490126>.
- Emyana Ruth Eritha S, "Implementasi Teknologi Big Data di Lembaga Pemerintahan Indonesia", *Jurnal Penelitian Pos dan Informatika, JPPI* Vol 6 No. 2, 2016.
- Sue YYadav "The Impact of Digital Technology on Consumer Purchase Behavior", the journal of Financial Perspectives: F InTech, Vol 3- issue 3, 2015.
- John F.Quigley, "Consumer Behavior in Digital Markets", Research paper, southern Illinois University Carbondale, OpenSIUC, 2015.
- Veikko Halttunen, "Consumer Behavior in Digital Era", Academic dissertation, the Faculty of Information Technology of the University of Jyväskylä, 2016.
- Ashari, Mochamad, 2015. *Big Data Industry and Academic Point of View*. Konferensi Big Data Indonesia.
- IBM Big Data & Analytics Hub, 2013. 4v's of Big Data. www.ibmbigdatahub.com/infographic/four-vs-big-data. Diakses tanggal 17 Februari 2016.
- Aryasa, Komang. "Big Data dan Bisnis Masa Depan". 2015 [online], available : <http://komangaryasa.com/2015/10/big-data-dan-bisnis-masa-depan/> dilihat tanggal, 1 Februari 2017.
- Dumbill, E. (2012). *Big Data Now Current Perspective*. O'Reilly Media.
- Kementerian Komunikasi dan Informatika (Kominfo). (2015). *Buku Saku Big Data*. Kementerian Komunikasi dan Informatika.
- Nyssa, Chrisvania. "Strategi Dalam Implementasi Big Data". 2017 [online], available : <https://socialmediaweek.org/jakarta/2017/01/06/strategi-dalamimplementasi-big-data/> dilihat tanggal, 2 Februari 2017.
- Eaton, C., Dirk, D., Tom, D., George, L., & Paul, Z. (n.d). *Understanding Big Data*. Mc. Graw Hill.