



UTILIZATION OF BIG DATA IN IMPROVING MARKETING STRATEGIES

PEMANFAATAN BIG DATA DALAM MENINGKATKAN STRATEGI PEMASARAN

Riri Fajriah^{1*}

^{1*} Universitas Mercu Buana, Jakarta Barat

riri.fajriah@mercubuana.ac.id

INFORMASI ARTIKEL

Submitted:
13-12-2023

Accepted:
22-02-2024

Published:
06-05-2024

Keywords:
Big Data; Analysis; Marketing

Kata Kunci:
Big Data; Analisis; Pemasaran

ABSTRACT

Big Data is data characterized by very large, very varied, fast-growing of data and may not need to be treated specially structured with innovative technology to obtain in-depth information and can help in making a better decision. The important things of Big Data is not just about technical capability to process data, but the benefits of its that can be realized by the company by using Big Data Analytics Terminology and it is believed come from web search company which process data the aggregation of distributed very large and unstructured. The challenge of Big Data includes the acquisition, accuracy of data, storage, search (search), the distribution, transfer, data analysis, and data visualization. The trend of Big Data is increasingly enlarged set of data caused by the increase of information from large sets of inter-related, compared with other small associations with the total number of the same data. One of the utilization of Big Data is a boost in sales with improved marketing strategy in matters such as: product campaigns, consumer segmentation and markets analysis, location-based marketing, sentiment analysis, personalized marketing and lure customers.

ABSTRAK

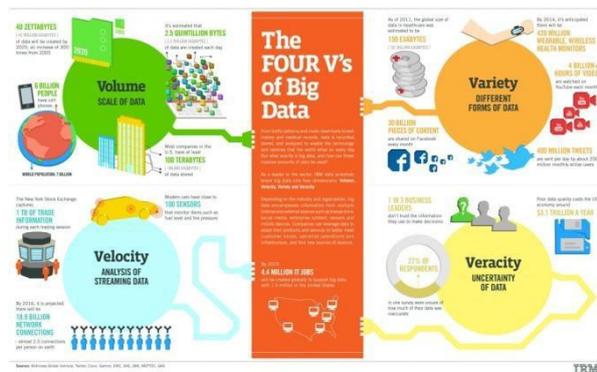
Big Data adalah data dengan ciri berukuran sangat besar, sangat variatif, sangat cepat pertumbuhannya dan mungkin tidak terstruktur yang perlu diolah khusus dengan teknologi inovatif sehingga mendapatkan informasi yang mendalam dan dapat membantu pengambilan keputusan yang lebih baik. Hal terpenting dari Big Data bukanlah sekedar kemampuan teknis untuk mengolah data melainkan manfaat yang dapat disadari oleh perusahaan dengan menggunakan Big Data Analytics Terminologi Big Data diyakini berasal dari perusahaan pencarian web yang mengolah data dengan agregasi yang terdistribusi sangat besar dan tidak terstruktur. Tantangannya meliputi pemerolehan, kurasi, penyimpanan, penelusuran (search), pembagian, pemindahan, analisis, dan visualisasi data. Tren kian membesarnya himpunan data terjadi akibat bertambahnya informasi dari himpunan-himpunan besar yang saling terkait, dibandingkan dengan himpunan-himpunan kecil lain dengan jumlah total data yang sama. Salah satu pemanfaatan Big Data adalah dalam meningkatkan penjualan dengan perbaikan strategi pemasaran dalam hal-hal semisalnya kampanye produk, segmentasi konsumen dan pasar, pemasaran berbasis lokasi, analisis sentiment, personal marketing dan menarik simpati pelanggan.

INTRODUKSI

Awalnya *Big Data* adalah sebuah sistem teknologi yang diperkenalkan untuk menanggulangi „ledakan informasi“ seiring dengan semakin bertumbuhnya ekosistem pengguna perangkat mobile dan data internet. Pertumbuhan perangkat mobile dan data internet ternyata sangat mempengaruhi perkembangan volume dan jenis data yang terus meningkat secara signifikan di dunia maya (Bahar, 2016).

Berbagai jenis data, mulai data yang berupa teks, gambar atau foto, video hingga bentuk data-data lainnya membanjiri sistem komputasi. Tentunya hal ini perlu jalan keluar dan *Big Data* adalah solusi yang kerap digaungkan beberapa waktu belakangan ini, dan kemunculannya memang dianggap solusi dari fakta yang menunjukkan bahwa pertumbuhan data dari waktu ke waktu telah melampaui batas kemampuan media penyimpanan maupun sistem database yang ada saat ini.

Sementara itu IBM di situs resminya mendefinisikan *Big Data* ke dalam empat karakteristik yang dikenal dengan 4V's of *Big Data*. Keempat karakteristik tersebut adalah : berukuran sangat besar (*high-volume*), atau sangat bervariasi (*high-variety*), atau kecepatan pertumbuhan tinggi (*high-velocity*), dan sangat tidak jelas (*high-veracity*).



Gambar 1. Empat Karakteristik *Big Data*

Sumber :

<http://www.ibmbigdatahub.com/infographic/four-vs-big-data>

Dengan begitu, *Big Data* dapat diasumsikan sebagai sebuah media penyimpanan data yang menawarkan ruang tak terbatas, serta kemampuan untuk mengakomodasi dan memproses berbagai jenis data dengan sangat cepat.

KONSEP *BIG DATA*

Dalam mengimplementasikan teknologi *Big Data* di suatu organisasi ada 4 elemen penting yang menjadi tantangan, yaitu data, teknologi, proses dan SDM (Aryasa, 2015).

▪ Data

Deskripsi dasar dari data menunjuk pada benda, event, aktivitas, dan transaksi yang terdokumentasi, terklarifikasi, dan tersimpan tetapi tidak terorganisasi untuk dapat memberikan suatu arti yang spesifik. Data yang telah terorganisir sehingga dapat memberikan arti dan nilai kepada penerima disebut informasi. (Rainer, cegielski, 2009 p10). Ketersediaan data menjadi kunci awal bagi teknologi *Big Data*. Ada beberapa organisasi yang memiliki banyak data dari proses bisnisnya yang dilakukan, baik data terstruktur maupun tidak terstruktur, seperti industri telekomunikasi maupun perbankan. Namun, ada pula organisasi yang perlu membeli atau bekerjasama dengan pihak lain untuk mendapatkan data. Selain itu *Big Data* didukung oleh sumber data yang bersumber dari pemakaian data mobile, email, sosial media serta sensor pertukarang data melalui jaringan internet. Hal ini dapat terlihat pada ilustrasi di Gambar 2. berikut ini :



Gambar 2. *Big Data* Platform

Sumber :

<http://www.ibmbigdatahub.com/infographic/big-data-making-world-go-round>

▪ Teknologi

Hal ini terkait dengan infrastruktur dan *tools* dalam pengoperasian *Big Data*, seperti teknik komputasi dan analitik, serta media penyimpanan (*storage*). Biasanya, organisasi tidak akan mengalami

kendala yang berarti dalam hal teknologi karena teknologi bisa didapatkan dengan membeli atau kerja sama dengan pihak ketiga.

Tabel 1. Perbandingan Antara Metode Konvensional Versus Data Analitik Dalam Pengumpulan dan Analisis Data

<i>LEGACY</i>	<i>DATA ANALYTICS</i>
<i>Confirmative</i>	<i>Explorative (predictive)</i>
<i>Small data set</i>	<i>Large data set</i>
<i>Small number of variable</i>	<i>Large number of variable</i>
<i>Deductive (no predictions)</i>	<i>Inductive</i>
<i>Numeric data</i>	<i>Numeric and non-numeric data</i>
<i>Clean data</i>	<i>Data cleaning</i>

Sumber : *Data Mining and Statistics : What are the Connections ? (Jerome Friedman, 1997)*

▪ **Proses**

Dalam proses mengadopsi teknologi *Big Data* dibutuhkan perubahan budaya organisasi. Misalnya, sebelum adanya *Big Data*, seorang pimpinan dalam menjalankan organisasi melakukan pengambilan keputusan hanya berdasarkan ‘intuisi’ berdasarkan nilai, keyakinan atau asumsinya. Namun setelah adanya teknologi *Big Data*, pimpinan mampu bertindak “*data-driven decision making*” artinya mengambil keputusan berdasarkan data yang akurat dan informasi yang relevan.

▪ **SDM**

Dalam mengaplikasikan teknologi *Big Data* dibutuhkan SDM dengan keahlian analitik dan kreativitas yaitu kemampuan / ketrampilan untuk menentukan metode baru yang dapat dilakukan untuk mengumpulkan, menginterpretasi dan menganalisis data, keahlian pemrograman komputer, dan ketrampilan bisnis yaitu pemahaman tentang tujuan bisnis. Sumber data dalam teknologi *Big Data* dapat berupa data terstruktur dan tidak terstruktur.

- Data terstruktur : memiliki tipe data, format, dan struktur yang telah terdefinisi. Dapat berupa data transaksional, OLAP data, tradisional RDBMS, file CSV, *spread-sheets* sederhana.
- Data tidak terstruktur : 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, audio serta data semantik.

Sejak dikenalnya *Big Data* dalam pengumpulan dan analisis data, teknik ini mulai diperbandingkan dengan metode konvensional yang sebelumnya dilakukan, seperti survey. Perbandingan antara kedua metode tersebut terlihat sebagai berikut :

METODE PENELITIAN

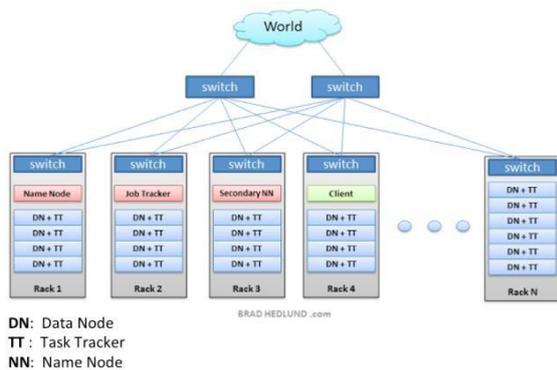
TOOLS BIG DATA

Beberapa tools yang tersedia untuk penerapan *Big Data* dapat dikategorikan menjadi 2, berbayar (*personalized software*) dan tidak berbayar (*open source software*). Tahapan aktivitas dan dukungan teknologi pada *Big Data* yaitu :

- *Acquired* berhubungan sumber dan cara mendapatkan data.
- *Accessed* berhubungan dengan daya akses data; data yang sudah dikumpulkan memerlukan tata kelola, integrasi, *storage* dan *computing* agar dapat dikelola untuk tahap berikutnya. Beberapa *processing/computing tools* : Hadoop, Nvidia CUDA, Twitter storm, GraphLab. Beberapa tools terkait *storage* : neo4J, Titan, HDFS.
- *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), *prescriptive analytic* (merekomendasikan pilihan dan implikasi dari setiap opsi). Beberapa *analytic tools* : MLPACK, Mahout, R.
- *Application* terkait visualisasi dan reporting hasil dari analitik. Contoh *application tool* : RStudio.

Salah satu aplikasi berbasis *open-source* yang populer untuk *Big Data* yaitu Hadoop. Hadoop

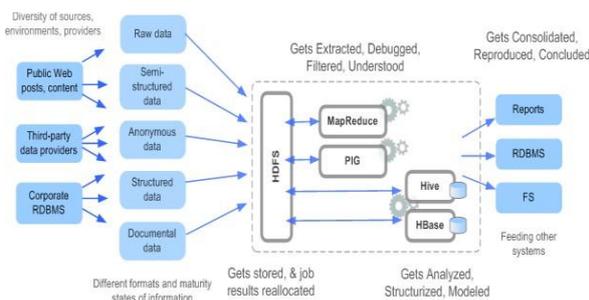
merupakan framework untuk menjalankan aplikasi pada komoditas perangkat keras yang memiliki cluster besar. Hadoop menerapkan prinsip komputasi yang disebut *Map/Reduce*, yaitu model pemrograman untuk memproses data bervolume besar secara parallel dan independen dengan cara membagi tugas (*task*) kedalam beberapa fragmen kerja. Selain itu, Hadoop menyediakan sistem file terdistribusi (HDFS) yang akan menyimpan data dalam titik-titik (*nodes*) komputasi. Dengan kedua teknik tersebut, kegagalan di satu titik diatasi secara otomatis oleh sistem.



Gambar 3. Hadoop Cluster

Sumber : <http://datacenter.opentray.com/2013/08/hadoop-tutorial-with-hortonworks-sandbox-vm-part-2-the-boss-edition/>

Komponen-komponen dalam Hadoop yaitu : Nama node, Data node, Secondary- name node, Job Tracker, Task Tracker.



Gambar 4. Bagan Mekanisme Map Reduce Hadoop

Sumber : <http://hadooper.blogspot.co.id/>

Saat ini sudah tersedia tools yang mampu menghasilkan koding *map reduce* dari Hadoop. Tools tersebut memiliki user interface yang memungkinkan pengguna untuk meng-klik dan drag untuk membuat aliran pengolahan data dan secara otomatis menghasilkan dan mengeksekusi kode pada prosesor

yang didistribusikan secara parallel. Dengan tools tersebut pengolahan data dapat dilakukan tanpa/sedikit koding.

Selain Hadoop, *tools* yang populer digunakan serta memiliki fungsi dalam analitik dan penyajian data yang beragam adalah R (Aryasa, 2015). R merupakan bahasa pemrograman dan *software framework* untuk analisis statistical dan grafik. Keunggulan dari software R terletak pada pendekatan analisis data dengan menggunakan visualisasi untuk menemukan karakteristik yang penting dari suatu dataset.

Dalam hal sebagai perangkat lunak berbayar (*Proprietary Software*) *Big Data Analytics* menjadi model bisnis baru bagi beberapa vendor/perusahaan. Seperti yang dilakukan oleh PT. Telkom Tbk. sejak tahun 2014, dengan menawarkan Big Data untuk menunjang pertumbuhan di industri lain. Telkom telah menjadi penyedia *data analytics* bagi instansi/lembaga pemerintah pusat dan daerah, industri airlines, dan health care. Telkom juga memfasilitasi pertukaran data miliknya dengan sumber data lainnya, sebab dalam dunia operator telekomunikasi data yang dikelola sangat banyak seperti data transaksi, data lokasi serta data kunjungan ke website. Sebagai contoh, dengan jumlah pelanggan 156 juta orang, diketahui dalam satu menit terjadi kicauan di Twitter sebanyak 98 ribu, update status Facebook 695 ribu, serta pembicaraan di instant messaging 11 juta, data ini tentu dapat bercerita banyak jika diolah.

Selain itu, ada pula vendor/perusahaan yang secara spesifik bergerak di bidang IT solution yang menyediakan solusi berbasis teknologi *Big Data* sesuai dengan kebutuhan klien (*personalized*). Ada vendor yang membuat tool analytic secara parsial, misalnya tool untuk integrasi data saja, penyimpanan data, analisis data, dan presentasi data hasil analisis. Tetapi ada pula vendor yang menyediakan tool secara lengkap, mulai dari integrasi data sampai analisis dan presentasi data, seperti Oracle, SAP, IBM, Microsoft, Teradata.

DIMENSI-DIMENSI BIG DATA

■ Volume

Perusahaan tertimbulk dengan data yang terus tumbuh dari semua jenis sektor, dengan mudah mengumpulkan terabyte bahkan petabyte informasi.

- a) Mengubah 12 terabyte Tweet dibuat setiap hari ke dalam peningkatan sentimen analisis produk.

- b) Mengkonvert 350 miliar pembacaan tahunan untuk lebih baik dalam memprediksi kemampuan beli pasar.

Mungkin karakteristik ini yang paling mudah dimengerti karena besarnya data. Volume juga mengacu pada jumlah massa data, bahwa organisasi berusaha untuk memanfaatkan data untuk meningkatkan pengambilan keputusan yang banyak perusahaan di banyak negara. Volume data juga terus meningkat dan belum pernah terjadi sampai setinggi ini sehingga tidak dapat diprediksi jumlah pasti dan juga ukuran dari data sekitar lebih kecil dari petabyte sampai zetabyte. *Dataset Big Data* sekitar 1 terabyte sampai 1 petabyte perperusahaan jadi jika *Big Data* digabungkan dalam sebuah organisasi / group perusahaan ukurannya mungkin bisa sampai zetabyte dan jika hari ini jumlah data sampai 1000 zetabyte, besok pasti akan lebih tinggi dari 1000 zetabyte.

▪ Variety

Volume data yang banyak tersebut bertambah dengan kecepatan yang begitu cepat sehingga sulit bagi kita untuk mengelola hal tersebut. Kadang-kadang 2 menit sudah menjadi terlambat. Untuk proses dalam waktu sensitif seperti penangkapan penipuan, data yang besar harus digunakan sebagai aliran ke dalam perusahaan Anda untuk memaksimalkan nilainya.

- a) Meneliti 5 juta transaksi yang dibuat setiap hari untuk mengidentifikasi potensi penipuan
- b) Menganalisis 500 juta detail catatan panggilan setiap hari secara real-time untuk memprediksi gejala pelanggan lebih cepat. Berbagai jenis data dan sumber data. Variasi adalah tentang mengolah kompleksitas beberapa jenis data, termasuk *structured data*, *unstructured data* dan *semi-structured data*. Organisasi perlu mengintegrasikan dan menganalisis data dari array yang kompleks dari kedua sumber informasi Tradisional dan non tradisional informasi, dari dalam dan luar perusahaan. Dengan begitu banyaknya sensor, perangkat pintar (*smart device*) dan teknologi kolaborasi sosial, data yang dihasilkan dalam bentuk yang tak terhitung jumlahnya, termasuk *text*, *web data*, *tweet*, *sensor data*, *audio*, *video*, *click stream*, *log file* dan banyak lagi.

▪ Velocity

Big Data adalah setiap jenis data - data baik yang terstruktur maupun tidak terstruktur seperti *teks*, *data sensor*, *audio*, *video*, *klik stream*, *file log* dan banyak lagi. Wawasan baru ditemukan ketika menganalisis kedua jenis data ini bersama-sama.

- a) Memantau 100 video masukan langsung dari kamera pengintai untuk menargetkan tempat tujuan.
- b) Mengeksploitasi 80% perkembangan data dalam gambar, video, dan dokumen untuk meningkatkan kepuasan pelanggan. Data dalam gerak. Kecepatan di mana data dibuat, diolah dan dianalisis terus menerus. Berkontribusi untuk kecepatan yang lebih tinggi adalah sifat penciptaan data secara real-time, serta kebutuhan untuk memasukkan streaming data ke dalam proses bisnis dan dalam pengambilan keputusan. Dampak *Velocity Latency*, jeda waktu antara saat data dibuat atau data yang ditangkap, dan ketika itu juga dapat diakses. Hari ini, data terus-menerus dihasilkan pada kecepatan yang mustahil untuk sistem tradisional untuk menangkap, menyimpan dan menganalisis. Jenis tertentu dari data harus dianalisis secara real time untuk menjadi nilai bagi bisnis.

▪ Veracity

Ketidakpastian data. *Veracity* mengacu pada tingkat keandalan yang terkait dengan jenis tertentu dari data. Berjuang untuk kualitas data yang tinggi merupakan syarat *Big Data* penting dan tantangan, tapi bahkan metode pembersihan data yang terbaik tidak dapat menghapus ketidakpastian yang melekat pada beberapa data, seperti cuaca, ekonomi, atau aktual keputusan membeli pelanggan masa depan. Kebutuhan untuk mengakui dan merencanakan ketidakpastian adalah dimensi data besar yang telah diperkenalkan sebagai eksekutif berusaha untuk lebih memahami dunia di sekitar mereka. Beberapa data tidak pasti, misalnya : sentimen dan kebenaran pada manusia, sensor GPS memantul antara pencakar langit Manhattan, cuaca kondisi- kondisi, faktor ekonomi, dan masa depan. Ketika berhadapan dengan jenis data, tidak ada metode pembersihan

data dapat memperbaiki untuk semua itu. Namun, meski ketidakpastian, data masih mengandung informasi yang berharga.

Kebutuhan untuk mengakui dan menerima ketidakpastian ini merupakan ciri dari data. *Uncertainty* besar memmanifestasikan dirinya dalam data besar dalam banyak cara. Sekarang dalam skeptisisme yang mengelilingi data yang dibuat dalam lingkungan manusia seperti jaringan sosial, dalam ketidaktahuan bagaimana masa depan akan terungkap dan bagaimana orang-orang, alam atau kekuatan pasar yang tak terlihat akan bereaksi terhadap variabilitas dari dunia di sekitar mereka. Untuk mengelola ketidakpastian, analisis perlu menciptakan konteks sekitar data. Salah satu cara untuk mencapai ini adalah melalui data fusion, di mana menggabungkan beberapa sumber yang kurang dapat diandalkan menciptakan lebih akurat dan berguna point data, seperti komentar sosial ditambahkan ke geospasial informasi lokasi. Cara lain untuk mengelola ketidakpastian adalah melalui matematika canggih yang mencakup hal itu, seperti teknik optimasi yang kuat dan pendekatan *fuzzy logic*. Manusia secara alami, tidak menyukai ketidakpastian, tetapi hanya mengabaikannya dapat menciptakan lebih banyak masalah daripada ketidakpastian itu sendiri. Dalam era *Big Data*, eksekutif akan perlu pendekatan dimensi ketidakpastian berbeda. Mereka akan perlu untuk mengakuinya, menerimanya dan menentukan.

ARSITEKTUR BIG DATA

Untuk memahami level aspek arsitektur yang tinggi dari *Big Data*, sebelumnya harus memahami arsitektur informasi logis untuk data yang terstruktur. Pada gambar di bawah ini menunjukkan dua sumber data yang menggunakan teknik integrasi (ETL / Change Data Capture) untuk mentransfer data ke dalam DBMS *data warehouse* atau *operational data store*, lalu menyediakan bermacam-macam variasi dari kemampuan analisis untuk menampilkan data.

Beberapa kemampuan analisis ini termasuk ; *dashboards*, laporan, EPM/BI *Applications*, ringkasan dan *query statistic*, *interpretasi semantic* untuk data tekstual, dan alat visualisasi untuk data yang padat. Informasi utama dalam prinsip arsitektur ini termasuk cara memperlakukan data sebagai asset melalui nilai, biaya, resiko, waktu, kualitas dan akurasi data.



Gambar 5. Bagan *Traditional Information Architecture Capabilities*

Sumber : Sun & Heller, 2012, p. 11

Mendefinisikan kemampuan memproses untuk *Big Data Architecture*, diperlukan beberapa hal yang perlu dilengkapi; volume, percepatan, variasi, dan nilai yang menjadi tuntutan. Ada strategi teknologi yang berbeda untuk *real-time* dan keperluan batch processing. Untuk *real-time*, menyimpan data nilai kunci, seperti NoSQL, memungkinkan untuk performa tinggi, dan pengambilan data berdasarkan indeks. Untuk batch processing, digunakan teknik yang dikenal sebagai Map Reduce, memfilter data berdasarkan pada data yang spesifik pada strategi penemuan. Setelah data yang difilter ditemukan, maka akan dianalisis secara langsung, dimasukkan ke dalam *unstructured database* yang lain, dikirimkan ke dalam perangkat mobile atau digabungkan ke dalam lingkungan *data warehouse* tradisional dan berkolerasi pada data terstruktur.



Gambar 6. Bagan *Big Data Information Architecture Capabilities*

Sumber : Sun & Heller, 2012, p. 11

Sebagai tambahan untuk *unstructured data* yang baru, ada dua kunci perbedaan untuk *Big Data*. Pertama, karena ukuran dari *data set*, *raw data* tidak dapat secara langsung dipindahkan ke dalam suatu *data warehouse*. Namun, setelah proses *Map Reduce* ada kemungkinan akan terjadi reduksi hasil dalam lingkungan *data warehouse* sehingga dapat memanfaatkan pelaporan *Business Intelligence* (BI), statistik, semantik, dan kemampuan korelasi yang biasa. Akan sangat ideal untuk memiliki kemampuan analitik yang mengkombinasikan perangkat BI bersamaan dengan visualisasi big data dan kemampuan query. Kedua, untuk memfasilitasi analisis dalam lingkungan *Hadoop*, lingkungan *sandbox* dapat dibuat.

Untuk beberapa kasus, *Big Data* perlu mendapatkan data yang terus berubah dan tidak dapat diperkirakan, untuk menganalisis data tersebut, dibutuhkan arsitektur yang baru. Dalam perusahaan retail, contoh yang bagus adalah dengan menangkap

jalur lalu lintas secara real-time dengan maksud untuk memasang iklan atau promosi toko di tempat strategis yang dilewati banyak orang, mengecek peletakan barang dan promosi, mengamati secara langsung pergerakan dan tingkah laku pelanggan.

Dalam kasus lain, suatu analisis tidak dapat diselesaikan sampai dihubungkan dengan data perusahaan dan data terstruktur lainnya. Sebagai contohnya, analisis perasaan pelanggan, mendapatkan respon positif atau negatif dari social media akan memiliki suatu nilai, tetapi dengan mengasosiasikannya dengan segala macam pelanggan (paling menguntungkan atau bahkan yang paling tidak menguntungkan) akan memberikan nilai yang lebih berharga. Jadi, untuk memenuhi kebutuhan yang diperlukan oleh *Big Data* BI adalah konteks dan pemahaman. Menggunakan kekuatan peralatan statistik dan semantik akan sangat memungkinkan untuk dapat memprediksikan kemungkinan – kemungkinan di masa depan.

Salah satu tantangan yang diteliti dalam pemakaian Hadoop dalam perusahaan adalah kurangnya integrasi dengan ekosistem BI yang ada. Saat ini BI tradisional dan ekosistem *Big Data* terpisah dan menyebabkan analisis data terintegrasi mengalami kebingungan. Sebagai hasilnya, hal ini tidaklah siap untuk digunakan oleh pengguna bisnis dan eksekutif biasa. Pengguna *Big Data* yang pertama kali mencoba menggunakan, seringkali menulis kode khusus untuk memindahkan hasil big data yang telah diproses kembali ke dalam database untuk dibuat laporan dan dianalisa. Pilihan – pilihan ini mungkin tidak layak dan ekonomis untuk perusahaan IT. Pertama, karena menyebabkan penyebaran salah satu data dan standar yang berbeda, sehingga arsitekturnya mempengaruhi ekonomi IT. *Big Data* dilakukan secara independen untuk menjalankan resiko investasi yang redundan, sebagai tambahannya, banyak bisnis yang sama sekali tidak memiliki staff dan ketrampilan yang dibutuhkan untuk pengembangan pekerjaan yang khusus.

Pilihan yang paling tepat adalah menggabungkan hasil *Big Data* ke dalam *data warehouse*. Kekuatan informasi ada dalam kemampuan untuk asosiasi dan korelasi. Maka yang dibutuhkan adalah kemampuan untuk membawa sumber data yang berbeda-beda, memproses kebutuhan bersama – sama secara tepat waktu dan analisis yang berharga.

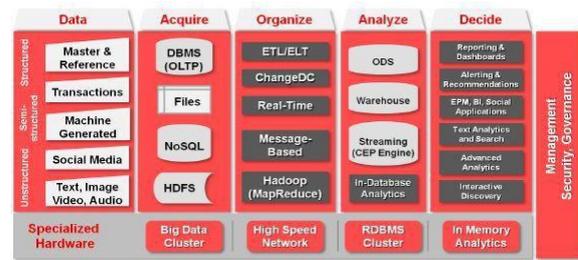


Figure 3. Oracle Integrated Information Architecture Capabilities

Gambar 7. Bagan Oracle Integrated Information Architecture Capabilities

Sumber : Sun & Heller, 2012, p. 13

Ketika bermacam – macam data telah didapatkan, data tersebut dapat disimpan dan diproses ke dalam DBMS tradisional, simple files, atau sistem *cluster* terdistribusi seperti NoSQL dan *Hadoop Distributed File System* (HDFS). Secara arsitektur, komponen kritikal yang memecah bagian tersebut adalah layer integrasi yang ada di tengah. Layer integrasi ini perlu untuk diperluas ke seluruh tipe data dan domain, dan menjadi jembatan antara data penerimaan yang baru dan tradisional, dan pengolahan kerangka. Kapabilitas integrasi data perlu untuk menutupi keseluruhan spektrum dari kecepatan dan frekuensi. Hal tersebut diperlukan untuk menangani kebutuhan ekstrim dan volume yang terus bertambah banyak. Oleh karena itu diperlukan teknologi yang memungkinkan untuk mengintegrasikan *Hadoop / Map Reduce* dengan *data warehouse* dan data transaksi.

Layer berikutnya digunakan untuk *Load* hasil reduksi dari *Big Data* ke dalam *data warehouse* untuk analisis lebih lanjut. Diperlukan juga kemampuan untuk mengakses data terstruktur seperti informasi profil pelanggan ketika memproses dalam *Big Data* untuk mendapatkan pola seperti mendeteksi aktivitas yang mencurigakan. Hasil pemrosesan data akan dimasukkan ke dalam ODS tradisional, *data warehouse*, dan *data marts* untuk analisis lebih lanjut seperti data transaksi. Komponen tambahan dalam layer ini adalah *Complex Event Processing* untuk menganalisa arus data secara real-time. Layer *business intelligence* akan dilengkapi dengan analisis lanjutan, dalam analisis *database* statistik, dan visualisasi lanjutan, diterapkan dalam komponen tradisional seperti laporan, *dashboards*, dan *query*. Pemerintahan, keamanan, dan pengelolaan operasional juga mencakup seluruh spektrum data dan lanskap informasi pada tingkat enterprise. Dengan arsitektur ini, pengguna bisnis tidak melihat suatu pemisah, bahkan tidak sadar akan perbedaan antara data transaksi tradisional dan big data. Data dan arus analisis akan terasa mulus tanpa halangan ketika dihadapkan pada bermacam – macam data dan set

informasi, hipotesis, pola analisis, dan membuat keputusan.

HASIL DAN PEMBAHASAN

2. MEMBANGUN *BIG DATA PLATFORM*

Seperti data pergudangan, toko web atau platform TI, infrastruktur untuk data yang besar memiliki kebutuhan yang unik. Dalam mempertimbangkan semua komponen platform data yang besar, penting untuk diingat bahwa tujuan akhir adalah untuk dengan mudah mengintegrasikan data yang besar dengan data perusahaan untuk memungkinkan dalam melakukan analisis mendalam pada set data gabungan.

▪ *Infrastructure Requirements*

Requirement dalam *Big Data* infrastruktur adalah sebagai berikut :

a) *Data Acquisition*

Tahap akuisisi adalah salah satu perubahan besar dalam infrastruktur pada hari-hari sebelum *Big Data*. Karena *Big Data* mengacu pada aliran data dengan kecepatan yang lebih tinggi dan ragam yang bervariasi, infrastruktur yang diperlukan untuk mendukung akuisisi data yang besar harus disampaikan secara perlahan, dapat diprediksi baik di dalam menangkap data dan dalam memprosesnya secara cepat dan sederhana, dapat menangani volume transaksi yang sangat tinggi, sering dalam lingkungan terdistribusi, dan dukungan yang fleksibel, struktur data dinamis.

Database NoSQL sering digunakan untuk mengambil dan menyimpan *Big Data*. Mereka cocok untuk struktur data dinamis dan sangat terukur. Data yang disimpan dalam database NoSQL biasanya dari berbagai variasi/ragam karena sistem dimaksudkan untuk hanya menangkap semua data tanpa mengelompokkan dan *parsing* data. Sebagai contoh, database NoSQL sering digunakan untuk mengumpulkan dan menyimpan data media sosial. Ketika aplikasi yang digunakan pelanggan sering berubah, struktur penyimpanan dibuat tetap sederhana. Alih-alih merancang skema dengan hubungan antar entitas, struktur sederhana sering hanya berisi kunci utama untuk mengidentifikasi titik data, dan kemudian wadah konten memegang data yang relevan. Struktur sederhana dan dinamis ini

memungkinkan perubahan berlangsung tanpa reorganisasi pada lapisan penyimpanan.

b) *Data Organization*

Dalam istilah Data pergudangan klasik, pengorganisasian data disebut integrasi data. Karena ada volume/jumlah data yang sangat besar, ada kecenderungan untuk mengatur data pada lokasi penyimpanan aslinya, sehingga menghemat waktu dan uang dengan tidak memindah-midahkan data dengan volume yang besar. Infrastruktur yang diperlukan untuk mengatur data yang besar harus mampu mengolah dan memanipulasi data di lokasi penyimpanan asli. Biasanya diproses didalam *batch* untuk memproses data yang besar, beragam format, dari tidak terstruktur menjadi terstruktur.

Apache Hadoop adalah sebuah teknologi baru yang memungkinkan volume data yang besar untuk diatur dan diproses sambil menjaga data pada cluster penyimpanan data asli. *Hadoop Distributed File System* (HDFS) adalah sistem penyimpanan jangka panjang untuk log web misalnya. Log web ini berubah menjadi perilaku browsing dengan menjalankan program *MapReduce* di cluster dan menghasilkan hasil yang dikumpulkan di dalam cluster yang sama. Hasil ini dikumpulkan kemudian dimuat ke dalam sistem DBMS relasional.

c) *Data Analysis*

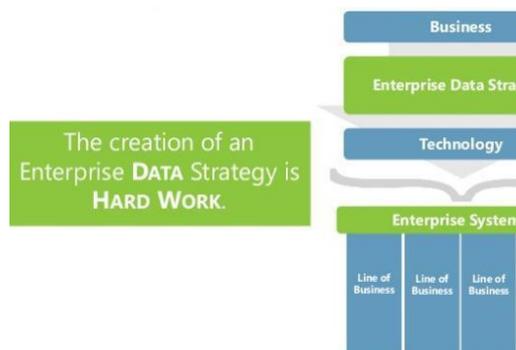
Karena data tidak selalu bergerak selama fase organisasi, analisis ini juga dapat dilakukan dalam lingkungan terdistribusi, di mana beberapa data akan tinggal di mana data itu awalnya disimpan dan diakses secara transparan dari sebuah *data warehouse*. Infrastruktur yang diperlukan untuk menganalisis data yang besar harus mampu mendukung analisis yang lebih dalam seperti analisis statistik dan *data mining*, pada data dengan jenis yang beragam dan disimpan dalam sistem yang terpisah, memberikan waktu respon lebih cepat didorong oleh perubahan perilaku; dan mengotomatisasi keputusan berdasarkan model analitis. Yang paling penting, infrastruktur harus mampu mengintegrasikan analisis pada kombinasi data yang besar dan data perusahaan tradisional.

Wawasan baru datang bukan hanya dari analisis data baru, tapi dari menganalisisnya dalam

konteks yang lama untuk memberikan perspektif baru tentang masalah lama. Misalnya, menganalisis data persediaan dari mesin penjual otomatis cerdas dalam kombinasi dengan acara kalender untuk tempat di mana mesin penjual otomatis berada, akan menentukan kombinasi produk yang optimal dan jadwal pengisian untuk mesin penjual otomatis.

PENERAPAN *BIG DATA* DALAM BIDANG PEMASARAN

Big Data menawarkan kesempatan bisnis yang tidak tertandingi dalam mengelola wawasan perilaku pelanggan (konsumen) dan mengubahnya menjadi hasil bisnis. Sebagai pebisnis, pelaku usaha (perusahaan) perlu memahami keberadaan data bukanlah suatu permasalahan. Kebanyakan bisnis memiliki lebih banyak dari cukup data yang membangun, namun perusahaan tidak tahu bagaimana menggunakannya. Kenyataan adalah bahwa kebanyakan bisnis sudah kaya data, tapi miskin wawasan. Mungkin benar bahwa perusahaan-perusahaan seperti Amazon, Google dan Facebook menikmati keunggulan kompetitif karena memiliki data yang cukup untuk diakses dan memiliki anggaran besar serta tim *Data scientist* yang tugasnya hanya untuk menganalisis data tersebut. Bagi sebagian besar bisnis, hal tersebut tidak mungkin, realistis atau diperlukan. Ada mungkin cukup data dalam bisnis sekarang bagi perusahaan untuk memanfaatkan kekuatan dari *Big Data* tanpa teknologi tinggi atau anggaran yang besar. Dan bahkan jika perusahaan dalam menjalankan bisnis tidak menyimpan catatan yang sangat baik atau tidak memegang sejumlah besar data yang ada pasti ada sumber eksternal lain yang cukup untuk memanfaatkan kekuatan *Big Data* dalam bisnis suatu perusahaan. Terkait hal tersebut pada penjelasan di atas maka dapat dilihat pada gambar dibawah ini perihal mekanisme dari suatu strategi dalam mengadopsi data yang berpengaruh dalam strategi bisnis yaitu:



Gambar 8. Bagan Mekanisme Data Strategy Adoption in Business

Sumber :

<http://www.slideshare.net/cmilroy/data-strategy-in-a-big-data-world>

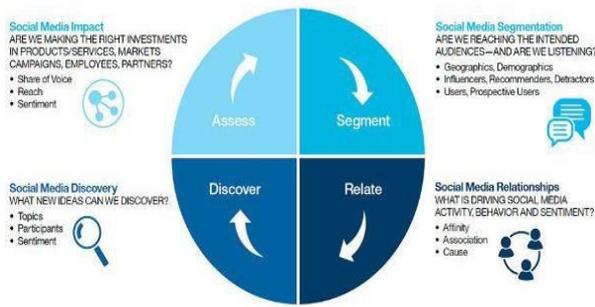
Dapat dilihat dari bagan diatas bahwa pengembangan bisnis yang optimal adalah dengan mengelola dengan baik *Enterprise Data Strategy* yang didukung dengan pemanfaatan teknologi untuk menciptakan *output enterprise system* yang mendukung keputusan strategis bagi perusahaan. Hingga saat ini banyak start up yang memiliki latar belakang teknologi yang menyediakan layanan berbasis *Big Data*. Hal ini menjadi alternatif dimana perusahaan-perusahaan menengah dapat merasakan serta menggunakan manfaat dari *Big Data* yang awalnya sebagian besar hanya dapat digunakan oleh perusahaan-perusahaan besar.

Big Data yang di implementasikan oleh perusahaan besar umumnya digunakan secara multi fungsi, dengan implementasi ini analisa baru digunakan untuk suatu kepentingan tanpa mengubah infrastruktur yang sudah mereka bentuk secara signifikan. Namun, jika perusahaan menengah atau SME (*Small Medium Enterprise*) mengimplementasikan *Big Data* untuk menganalisis suatu keperluan tertentu secara spesifik atau terperinci.

STRATEGI PEMASARAN MENGGUNAKAN *BIG DATA*

▪ Analisis Data Social (*Social Data Analysis*)

Pada bagian ini menjadi solusi yang sangat cemerlang bagi start up karena kemudahan dalam mengakses dari media sosial. Tidak hanya itu, dengan menggunakan Analisis Data Sosial perusahaan dapat mengembangkan kegiatan analisisnya mulai dari *Customer Segmentation*, mengukur efektifitas kegiatan marketing, dan hal-hal lainnya.



Gambar 9. Bagan IBM *Social Media Analytic Data Discovery*

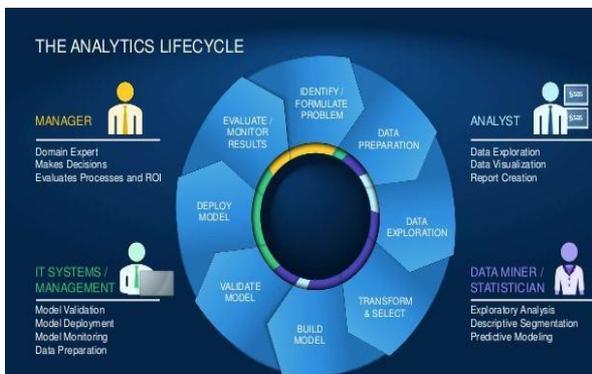
Sumber : [http://www-](http://www-01.ibm.com/software/analytics/solutions/customer-analytics/images/social-media-analytics-solution600x308.jpg)

[01.ibm.com/software/analytics/solutions/customer-analytics/images/social-media-analytics-solution600x308.jpg](http://www-01.ibm.com/software/analytics/solutions/customer-analytics/images/social-media-analytics-solution600x308.jpg)

Pada bagan diatas terlihat bahwa analisa dalam pengolahan data melalui media social media terdiri dari 4 aspek *Social Media Impact* (Evaluasi dari layanan pelanggan, kampanye produk, kerja sama dengan pegawai dalam pemanfaatan marketing menggunakan *social media*), *Social Media Segmentation* (Pembagian kategori pelanggan berdasarkan indikator tertentu), *Social Media Relationships* (Analisa kebiasaan/trend yang menjadi pembicaraan di *social media* untuk kepentingan penempatan strategi pemasaran), *Social Media Discovery* (Menemukan ide-ide baru dalam pemasaran produk melalui *social media*).

▪ **Analisis Data Riwayat (*Historical Data Analysis*)**

Fungsi dari tools ini untuk menganalisis data-data lampau dari sebuah perusahaan yang tujuannya untuk mencari suatu ide atau gagasan dari data lampau sehingga kita dapat mempelajari serta mengetahui apa yang telah terjadi terhadap perusahaan dimasa lalu. Pada gambar di bawah ini ditampilkan proses siklus dari suatu analisis data dalam suatu operasional bisnis di perusahaan.



Gambar 10. Bagan *Data Analytics Lifecycle*

Sumber:

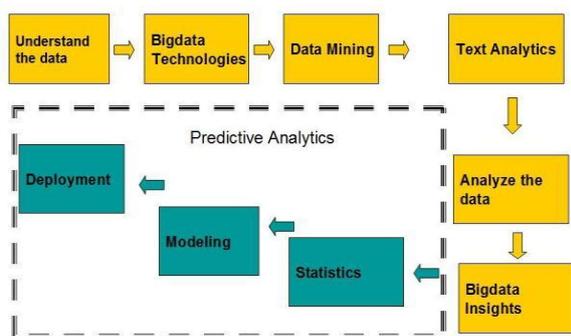
<https://image.slidesharecdn.com/itsecretarymeetinghpava-130827072155-phpapp02/95/big-data-analytics-ingovernment-4-638.jpg?cb=1377588176>

▪ **Analisis Prediksi (*Predictive Analysis*)**

Pada dasarnya Tools ini termasuk dalam bagian Historical Data Analysis, namun perbedaannya dengan menggunakan Predictive Analysis kita dapat menggunakan data-data perusahaan dimasa lalu untuk dijadikan sebuah ide atau gagasan serta memunculkannya kembali atau dijadikan sebuah trend pada masa sekarang atau untuk kedepannya. Dalam melakukan proses analisa data dengan jumlah yang besar, terdapat peran seorang data analisis yang berfungsi dalam membantu mendesain proses keseluruhan analisis serta domain expert yang tidak bisa dihilangkan. Domain expert berperan sebagai ahli yang mengerti bidang usaha sehingga dapat menerjemahkan hasil dari analisis data yang bisa dimengerti oleh management, sehingga mereka dapat membuat keputusan maupun tindakan strategis yang diperlukan. Langkah-langkah yang perlu dilakukan sebelum mengimplementasikan big data diantaranya :

- a) Memastikan adanya masalah yang perlu segera diselesaikan atau dicari solusinya dengan menggunakan *Big Data*.
- b) Terdapat data yang akan dianalisis, sehingga melahirkan solusi dari masalah yang sedang dihadapi.
- c) Adanya peran Data Analisis serta *Domain Expert* yang akan membantu dalam proses implementasi serta pemanfaatan hasil dari *Big Data*.
- d) Sesuaikan proses implementasi *Big Data* dengan kondisi financial perusahaan, dalam hal ini apakah perusahaan ingin membangun infrastruktur sendiri atau memanfaatkan *Third Party* yang menyediakan solusi *Big Data* untuk masalah yang sedang dihadapi.

Proses analisa prediksi dalam mekanisme big data dapat dilihat pada skema bagan gambar di bawah ini :



Gambar 11. Bagan Predictive Analytics Processing in Big Data System

Sumber : <http://predictive-5a88.kxcdn.com/wp-content/uploads/2013/11/Big-data-and-predictive-analytics-processing.jpg>

TANTANGAN DALAM PEMANFAATAN

BIG DATA

Dalam usaha pemanfaatan *Big Data* dapat terdapat banyak hambatan dan tantangan, beberapa hal diantaranya berhubungan dengan data dimana melibatkan acquisition, sharing dan privasi data, serta dalam analisis dan pengolahan data.

▪ Privasi

Privasi merupakan isu yang paling sensitif, dengan konseptual, hukum, dan teknologi, Privasi dapat dipahami dalam arti luas sebagai usaha perusahaan untuk melindungi daya saing dan konsumen mereka. Data-data yang digunakan atau disimpan sebagai *Big Data*.

▪ Access dan Sharing

Akses terhadap data, baik data lama maupun data baru dapat menjadi hambatan dalam mendapatkan data untuk *Big Data*, terlebih pada data lama dimana data-cdata tersimpan dalam bentuk – bentuk yang berbeda- beda dan beragam ataupun dalam bentuk fisik, akses terhadap data baru juga membutuhkan usaha yang lebih kerana diperlukannya izin dan lisensi untuk mengakses data-data non-public secara legal.

▪ Analisis

Bekerja dengan sumber data baru membawa sejumlah tantangan analitis, relevansi dan tingkat keparahan tantangan akan bervariasi tergantung pada jenis analisis sedang dilakukan, dan pada jenis keputusan yang akhirnya akan bisa diinformasikan

oleh data. Tergantung dari jenis data terdapat 3 kategori dalam analisis data yaitu sebagai berikut :

a) Penentuan Gambaran Yang Benar

Masalah ini biasanya ditemukan dalam penanganan *unstructured user-generated text-based data* dimana data yang didapatkan belum tentu benar karena data atau sumber yang salah.

b) Interpreting Data

Kesalahan-kesalahan seperti Sampling selection bisa merupakan hal yang sering ditemukan dimana data yang ada tidak dapat digunakan untuk mepresentasikan semua populasi yang ada, dan apophenia, melihat adanya pola walaupun tidak benar-benar ada dikarenakan jumlah data yang besar, dan kesalahan dalam menginterpretasikan hubungan dalam data.

c) Definiting and Detecting Anomalies

Tantangan sensitivitas terhadap spesifisitas pemantauan sistem. Sensitivitas mengacu pada kemampuan sistem pemantauan untuk mendeteksi semua kasus sudah diatur untuk mendeteksi sementara spesifisitas mengacu pada kemampuannya untuk mendeteksi hanya kasus-kasus yang relevan.

Kegagalan untuk mencapai hasil yang terakhir "Tipe I kesalahan keputusan", juga dikenal sebagai "positif palsu"; kegagalan untuk mencapai mantan "Type II error", atau "negatif palsu."

Kedua kesalahan yang tidak diinginkan ketika mencoba untuk mendeteksi malfungsi atau anomali, bagaimanapun didefinisikan, untuk berbagai alasan. Positif palsu merusak kredibilitas sistem sementara negatif palsu dilemparkan ragu pada relevansinya. Tapi apakah negatif palsu lebih atau kurang bermasalah daripada positif palsu tergantung pada apa yang sedang dipantau, dan mengapa itu sedang dipantau.

KESIMPULAN

Dapat disimpulkan bahwa *Big Data* sudah mulai dimanfaatkan dan akan sangat berguna untuk dipahami lebih dalam untuk mengimbangi perkembangan jaman ke arah teknologi dan analisis yang lebih praktis. Teknologi *Big Data* bertujuan agar setiap bisnis, perusahaan ataupun individu yang mampu mengolah data tersebut bisa mendapatkan informasi lebih mendalam (*insights*) yang akan

memicu pengambilan keputusan (*decision making*) dan tindakan (*action*) bisnis yang mengandalkan insights tersebut, bukan berdasarkan insting semata.



Gambar 12. Bagan Decision Making Process

Sumber : <http://www.apaitubigdata.com/p/apa-itu-big-data.html>

Selain itu, Proyek *Big Data* merupakan bagian dari terpenting dalam kehidupan perusahaan di masa mendatang. Oleh karena itu, perusahaan perlu menentukan pendekatan dasar yang tepat sebelum memulai mengerjakan berbagai proyek mahal.

Big Data dapat merevolusi bisnis dan mendukung strategi pemasaran yang efektif. Dalam rangka untuk melakukan itu perusahaan perlu kerangka praktis yang dapat membantu untuk menerapkan *Big Data* sehingga dapat memanfaatkannya untuk mendapatkan wawasan baru yang dapat membangun bisnis ke depan menjadi lebih baik dan berkembang dengan pesat.

DAFTAR PUSTAKA

- 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.
- Aryasa, Komang. “Bisnis membutuhkan Strategi Big Data”. 2015 [online], available : <http://komangaryasa.com/2015/10/bisnis-membutuhkan-strategi-big-data/> dilihat tanggal, 1 Februari 2017.
- Aryasa, Komang. “Apa yang ditawarkan Big Data?”. 2015 [online], available : <http://komangaryasa.com/2015/06/apa-yang-ditawarkan-big-data/> dilihat tanggal, 1 Februari 2017.
- Dumbill, E. (2012). *Big Data Now Current*. Perspective. O’Reilly Media.
- Eaton, C., Dirk, D., Tom, D., George, L., & Paul, Z. (n.d). *Understanding Big Data*. Mc. Graw Hill.
- Global Pulse. (2012). *Big Data for Development : Challenges & Opportunities*. Global Pulse.
- H. Immon, W. (2005). *Building the Data Warehouse*, 4th Edition. Indianapolis, Indiana : Wiley Publishing, Inc.
- IBM. (n.d.). *Analytics : The real-world use of big data*. Retrieved from How innovative enterprises extract value from uncertain data : <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-big-data-at-work.html>.
- Kementrian Komunikasi dan Informatika (Kominfo). (2015). *Buku Saku Big Data*. Kementrian Komunikasi dan Informatika.
- Nyssa, Chrisvania. “Strategi Dalam Implementasi Big Data”. 2017 [online], available : <https://socialmediaweek.org/jakarta/2017/01/06/strategi-dalam-implementasi-big-data/> dilihat tanggal, 2 Februari 2017.
- R. Kelly Rainer, C. (2011). *Introduction to Information Systems*. John Wiley & Sons (Asia) Pte Ltd