

A Survey on Smart Analytics: Method, Tools, and Open Research Issues

Nova Nurviana

School of Electrical Engineering and Informaticss

Bandung Institute of Technology

Bandung, Indonesia

nova.nurviana@s.itb.ac.id

Abstract—Smart analytic nowadays was very popular because today big data is booming. Back several years ago, people were fascinating with statistics which surveys or census become a gun for describing or forecasting some issues as of decision making or planning. By migrating to big data, smart analytic is very needed to filter which data is useful and producing some worthy information.

Keywords— *analytic tool, smart analytic, big data, survey*

I. INTRODUCTION

Smart analytic diartikan sebagai proses dari mengumpulkan dan menganalisis data guna untuk membantu bisnis dalam proses pengambilan keputusan dalam hal produk atau service yang akan disediakan [1]. Oleh sebab itu, dalam proses smart analitik juga disertakan pre-processing data. Sebagaimana kita ketahui bahwa data sangat banyak dikarenakan semua serba digital [2] sehingga banyak sensor, banyak report yang dikumpulkan, dalam artian data terus bertumbuh, sehingga disebut sebagai big data [3]. Menurut penelitian, tujuh puluh persen waktu yang digunakan untuk memproses sebuah data adalah pada tahap pre-processing [4]. Pertama, filtering untuk data yang tidak akan digunakan. Kemudian mengubah struktur data dari yang semula unstructured menjadi structured [5].

Perbedaan smart analitik dengan analitik biasa adalah, pada smart analitik terdapat artificial intelligent guna menganalisis secara smart [6]. Dalam analisis biasa, digunakan metode statistik deskriptif dan inferensia [7] [8]. Statistik deskriptif akan menggambarkan data dalam bentuk [9] :

- Tabular: data disajikan dalam bentuk tabel sehingga dapat diketahui sebuah informasi berdasarkan atribut, disajikan dalam bentuk kolom dan baris.
- Chart: dalam bentuk diagram batang, pie charts, line charts, tujuannya untuk menampilkan proporsi/jumlah antar variabel atau antar waktu
- Histogram: untuk menunjukkan frekuensi dari suatu variabel dengan interval tertentu, disajikan dengan tampilan batang
- Scatter plot: menunjukkan keterkaitan antar dua variabel disertai dengan konfidensi interval tertentu
- Cartogram: untuk menampilkan informasi geografis dan geospasial

Selain yang disebutkan di atas, sekarang ini sangat familiar media untuk membuat data “berbicara” melalui animasi-animasi menarik, seperti infografis dan videografis [10]. Tujuannya adalah agar masyarakat lebih peka terhadap data bahwa data tidak hanya sederet angka yang membosankan, melainkan mengandung informasi yang tanpa disadari sebenarnya sangat dekat dengan kehidupan kita.

Data adalah story, bagaimana kita membawakannya (story telling) akan berdampak pada bagaimana orang lain menerimanya. Data tersebar dimana-mana, seperti pada data-data statistik, pada bisnis proses, data riset, peta lokasi, maupun data keuangan. Sederet huruf-huruf atau sederet angka-angka mayoritas membuat orang malas dan tidak tertarik dikarenakan kita membutuhkan waktu untuk membaca keseluruhan dan memprosesnya pada otak kita, tepatnya pada otak kiri karena otak kiri bertugas untuk pemrosesan bahasa, sedangkan otak kanan. Sedangkan otak kanan berfungsi untuk mengenali objek dan spasial [28]. Oleh karena perbedaan itu, data dibuat sedemikian menarik sehingga bagaimana data yang tadinya terlihat membosankan karena diproses oleh otak kiri,

dapat diproses oleh otak kanan dengan visualisasi yang menarik agar otak kita mudah menangkap dan lebih cepat masuk memori otak serta akan bertahan lama dalam otak (mudah diingat).

Berbeda dengan analitik yang bersifat konvensional, smart analitik dilengkapi dengan mesin intelijen untuk memunculkan informasi dari big data yang secara diam-diam menyembunyi informasi tersebut secara otomatis, sehingga user tidak perlu query yang kompleks, cukup menggunakan drag and drop. Hasilnya, waktu pengerjaan analisis menjadi lebih singkat.

II. SURVEY

A. Metode Smart Analitik

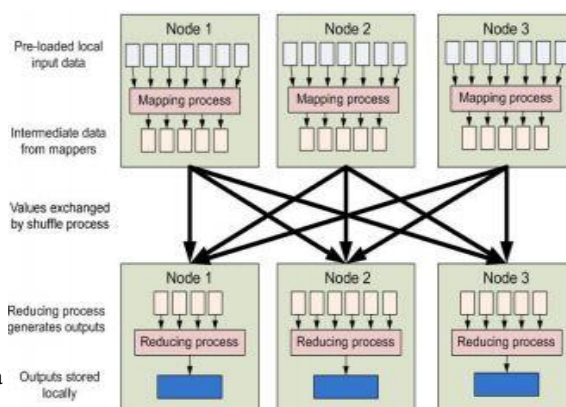
Ada banyak metode untuk menganalisis data, diantaranya:

1. Clustering
Metode jenis ini sangat fundamental. Beberapa peneliti telah fokus dalam pengembangan jenis analitik clustering [11][12][13]. Metode dasarnya adalah mengidentifikasi dataset yang memiliki persamaan yang signifikan sehingga dari data tersebut muncul kelompok-kelompok. Implementasi dari metode ini adalah mencari segmentasi konsumen sehingga pihak perusahaan dapat menawarkan produk atau servis yang berbeda antar individu tergantung dari tipe kelompoknya. Cara ini sangat ampuh untuk menghasilkan *benefit* yang lebih besar.
2. Decision Trees
Metode ini sangat bermanfaat untuk pengambilan keputusan berdasarkan kombinasi atribut yang terlibat. Ada satu yang menggunakan metode ini (Nova, 2012) dimana untuk mengklasifikasikan sebuah berita ke dalam indikator melalui metode tree, pertama difilter apakah aktivitas itu mengandung kekerasan, kemudian difilter kembali aktor yang melakukan apakah pemerintah atau masyarakat, terakhir aksi tersebut apakah mengandung unsur keagamaan, politik, atau pelayanan umum [14].
3. Column Dependency
Untuk mengetahui seberapa kuat *dependency* antar satu variabel terhadap aspek tertentu. Misalnya perusahaan Gojek ingin mengetahui apakah kesopanan para pengemudi berpengaruh penting terhadap kepercayaan konsumen kepada Gojek. Peneliti lain yang meneliti hal serupa adalah [15].
4. Mesin Rekomendasi
Idenya adalah mempelajari *historical data*, kemudian dimodelkan, kemudian untuk memprediksi selanjutnya aktivitas konsumen akan bagaimana, atau produk apa yang selanjutnya akan dicari oleh konsumen. Hasilnya akan menjadi masukan para *provider* produk maupun servis untuk mengatur strategi bisnis sesuai dengan rekomendasi yang disarankan agar mendapatkan keuntungan sebesar-besarnya [16].

B. Tools Smart Analytic

Terdapat beberapa tools yang dapat digunakan untuk mengolah big data dan tools untuk analitik, diantaranya:

- Apache Hadoop dan MapReduce
Apache hadoop merupakan pengembangan dari opensource MapReduce [17] yang fungsinya untuk mengindeks web crawl. MapReduce adalah model pemrograman untuk memproses dataset yang besar berdasarkan metode Map step dan Reduce step [18][19]. Map step bekerja dengan cara node master mengambil input, kemudian memecahnya menjadi bagian-bagian kecil sehingga dapat didistribusikan kepada node worker. Node worker kemudian mengerjakan bagiannya dan menyerahkan hasilnya pada node master. Sedangkan reduce step akan mengumpulkan hasil kemudian membentuknya menjadi output.



Gambar 1 Proses Map dan Reduce [19]

- Apache Mahout

Berawal dari machine learning berupa program komputer untuk mengoptimalkan performa komputer dengan menggunakan data histori atau sampel data [20]. Machine learning sendiri terdiri dari tiga macam, yaitu supervise, unsupervise, dan semisupervise.

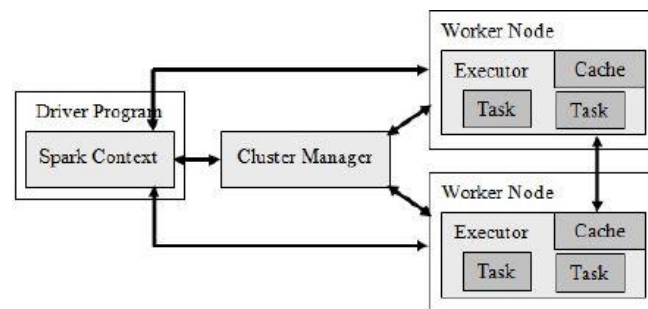
Supervise sendiri merupakan fungsi untuk memprediksi dengan cara memberi label kepada data-data history kemudian dimasukkan ke dalam machine learning. Tujuannya agar ketika ada data baru dimasukkan, machine learning sudah dapat mengenali dan dapat mengetahui apa yang harus dikerjakan terhadap data tersebut. Unsupervised merupakan fungsi tanpa melabeli data, tapi menggunakan fungsi statistika. Sedangkan semi supervised akan berangkat dari unsupervised, kemudian dari hasil yang missed atau tidak sesuai tujuan, data tersebut kemudian dilabeli, sehingga hasilnya menjadi lebih bagus.

Kemudian, Dari semua domain machine learning tersebut, Hadoop sudah menyelesaikannya dengan baik, namun gap diantaranya adalah Mahout lebih scalable. Selain itu dibandingkan machine learning yang lain, Mahout banyak komunitasnya, dokumentasi dan contohnya, skalabilitasnya, memiliki lisensi Apache, dan orientasi pada riset [21].

- Apache Spark

Apache Spark merupakan program open source yang banyak digunakan oleh programmer di Indonesia. Alasannya adalah karena open source tentunya, alasan lain adalah dapat dituliskan dalam bahasa Java, Scala, R, atau Python [18]. Dalam penelitian [22] disebut bahwa Apache Spark sangat cocok bagi modeling yang memiliki iteratif yang banyak dikarenakan ia memiliki arsitektur dengan library yang terdistribusi sehingga memiliki kecepatan pemrosesan yang bisa diandalkan. Selain itu, ia memiliki sifat fault-tolerant sehingga apabila komponen lain ada yang rusak, sistem dapat terus berjalan.

Apabila kita memiliki penelitian yang banyak mengandung statistika, Apache Spark merupakan pilihan yang tepat dikarenakan Apache Spark memiliki fitur untuk menampilkan grafik [22], sangat cocok untuk menampilkan regresi maupun deskripsi data berupa visualisasi data, maupun untuk menampilkan hubungan antar variabel dimana variabel disimbolkan dengan node, sedangkan hubungannya digambarkan dengan edge.



Gambar 2 Arsitektur Apache Spark [18]

Data disimpan dalam anis fault tolerance, kemudian cluster manager akan membagi-bagikan task kepada beberapa executor. Tak heran jika Apache Spark memang terkenal untuk distribusi programming secara real time.

Apache Spark memiliki ekosistem yang menggabungkan beberapa komponen sehingga semua kebutuhan analitik diharapkan telah mampu dilakukan oleh Spark. Berikut merupakan komponen ekosistemnya [23].

- ✓ **Spark Core**

Spark core merupakan engine dari pemrograman secara paralel dan terdistribusi. Dalam Spark Core ini dipasang untuk streaming SQL serta machine learning. Tugas dari Spark Core adalah untuk mengatur penyimpanan, fault recovery, penjadwalan, pendistribusian tugas, serta monitoring proses.

- ✓ **Spark Streaming**

Seperti dari morfologi katanya, streaming yang berarti real time, spark streaming bertanggung jawab dalam pemrosesan data secara real time.

✓ **Spark SQL**

Spark SQL adalah modul yang mengintegrasikan relational processing dan functional programming. Untuk query nya dapat menggunakan SQL atau Hive Query Language.

✓ **GraphX**

Komponen ini berguna untuk memproses grafik. Pada selanjutnya, GraphX berkembang menjadi Graph RDD dimana graph tidak hanya menunjukkan node yang saling berhubungan yang ditunjukkan dengan edge, tapi Graph RDD dapat menunjukkan jenis relationship apa yang terjadi di edge tersebut beserta besarnya.

✓ **MLlib**

MLlib merupakan kepanjangan dari Machine Learning library. MLlib digunakan untuk menjalankan machine learning dalam Apache Spark

✓ **SparkR**

Dalam bahasa pemrograman R terdapat paket yang menyediakan fungsi-fungsi seperti filtering, aggregation, selection, sama seperti pada SQL namun untuk dataset yang besar.

• Dryad

Model programming lain yang digunakan untuk memproses paralel computing serta program terdistribusi menggunakan grafik dataflow [24]. Dalam artikel lain [25] juga disebutkan bahwa dryad adalah mesin untuk mengeksekusi secara terdistribusi atau paralel data kasar (coarse grain data) . Dryad akan mengkombinasikan antara komputasi verteks dengan channel komunikasi untuk membuat grafik dataflow. Pertama-tama dryad akan mengeksekusi verteks dari graph pada beberapa set komputer, cara dryad mengetahui komputer yang available adalah melalui komunikasi file, saluran TCP, serta shared-memory FIFOS dimana data dipublikasikan lalu yang pertama kali datang berarti pertama yang dieksekusi [26]. Beberapa komputer itu disusun dari single komputer secara clustering, kemudian di dalam single komputer itu terdapat beberapa core di dalamnya sehingga pekerjaan menjadi lebih cepat.

• Storm

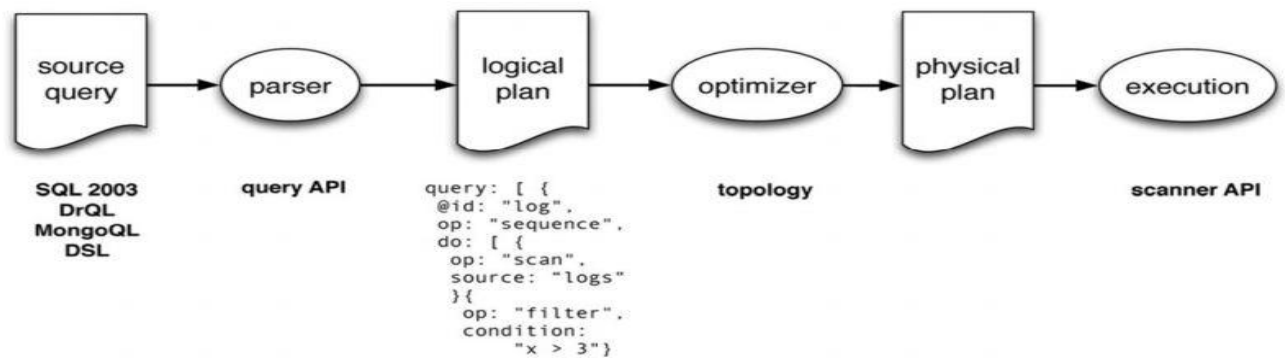
Sistem komputasi lain untuk memproses *large data* secara real time dan fault tolerant adalah storm. Storm sangat berbeda dengan Hadoop dimana Hadoop bekerja dengan cara mengumpulkan data terlebih dahulu kemudian diproses, sedangkan storm sistem komputasinya bersifat real time, sehingga setiap ada input data baru yang masuk, nilai komputasinya akan langsung berubah sesuai dengan fungsi model yang diterapkan. Selain dari itu, perbedaan dalam mengerjakan task dapat dilihat dari arsitektur sistem dimana Hadoop menggunakan sistem Map dan Reduce (dapat dilihat pada gambar 1), sedangkan Storm arsitekturnya bersifat dinamis [18], beda task, beda pula arsitekturnya.

• Apache Drill

Kelebihan dari Apache Drill adalah Apache Drill memiliki support yang fleksibel dan dinamis untuk urusan bahasa query, format data, serta sumber data. Hal ini justru merupakan kelebihan yang dilirik banyak programmer karena programmer seperti yang diketahui, memiliki kemampuan coding yang berbeda-beda, sehingga programmer tidak perlu mempelajari bahasa query tertentu untuk dapat menggunakan fasilitas dari Apache Drill.

Gambar 3 menjelaskan bahwa interface Apache Drill menyediakan query language, bisa berupa Java Database Connectivity (JDBC) atau Open Database Connectivity (ODBC). Mereka adalah Application Programming Interface, yang JDBC menggunakan bahasa, sedangkan ODBC khusus untuk mengakses database relational menggunakan sistem operasi windows.

Setelah melalui API, data ditransformasi dan digeneralisasi ke logical plan dimana proses



Gambar 3 Query Execution di Apache Drill [27]

berlangsung dalam memori namun sebagaimana dalam Java, logical plan dianggap sebagai objek. Setelah itu, ditransformasi ke physical plan yang merepresentasikan struktur data sebenarnya. Optimizer berguna untuk membantu pembentukan physical plan dari logical plan dengan cara restrukturisasi data. Selain itu, optimizer akan mempercepat proses dengan cara mengeksploitasi data kolom.

- Jaspersoft

Jasprsoft merupakan platform untuk analisis data dengan skala besar dan menggunakan struktur database berbasis kolom. Kelebihan dari Jaspersoft adalah kita tidak perlu melakukan ekstraksi, transformasi, serta loading (ETL) untuk mengolah big data. Selain itu, Jaspertext dapat secara langsung membuat laporan dalam bentuk Hypertext Markup Language (HTML) serta memiliki fitur dashboard secara interaktif langsung dari big data tanpa melakukan ETL.

- Splunk

Splunk adalah smart platform untuk menganalisis big data. Splunk sangat *powerful* dalam memproses data dengan volume besar sehingga disebut sebagai mesin data. Untuk ilustrasi bahwa Splunk dapat diandalkan adalah, dapat menyelamatkan data center dikala web mengalami down, namun kita tidak sedang berada di data center, melainkan kita berada di tempat lain. Splunk dapat menyelamatkan kita karena kita dapat kontrol web servis, database, firewall, router, maupun load balancer dari jarak jauh. Sama layaknya kita membuat konfigurasi device tertentu dari device lain, atau membuat konfigurasi pada sistem operasi tertentu, maupun aplikasi lain dari jarak jauh.

Cara kerja Splunk adalah,

- Splunk akan mengumpulkan data dan data tersebut kemudian diberi indeks.
- Dengan indeks, Splunk dapat mencari permasalahan dari log server. Karena Splunk memiliki skalabilitas yang besar dan tugas dapat dikerjakan secara cepat, Splunk dapat menentukan secara cepat problem apa yang terjadi.
- Dari history problem yang ada, kemudian penyebab masalah, dapat dibuat peraturan jika terjadi lagi, maka kirim alert ke device kita agar kita dapat maintenance problem dari jarak jauh.

III. ANALYSIS

A. Done Work

Teknologi yang diciptakan dan dikumpulkan kini telah memberikan *benefit* pada bisnis. Sebagaimana bisnis akan menyesuaikan terhadap gaya hidup kebanyakan orang, maka perlu diketahui apa saja gaya hidup orang berkaitan dengan teknologi. Apakah teknologi mempengaruhi gaya hidup, atau gaya hidup yang mempengaruhi perkembangan teknologi.

Penelitian [29] menyatakan bahwa teknologi dapat menangkap gaya hidup perorangan, kemudian dari hasil tangkapan dapat dianalisis, misalnya membuat segmentasi customer berdasarkan gaya hidup.

Penelitian lainnya [30] menghasilkan sebuah model yang menggambarkan dampak dari informasi dan teknologi terhadap gaya hidup dan energi yang dikonsumsi. Secara sederhana dapat dianalogikan sebagai berikut. Teknologi yang ada memungkinkan untuk terjadinya e-commerce sehingga customer tidak perlu lagi mengunjungi pusat perbelanjaan untuk memenuhi kebutuhannya, cukup dengan mengunjungi situs belanja online, kemudian transfer, barang akan diantar sampai pada tempat yang kita kehendaki.

Berbeda dengan penelitian kebanyakan, bahwa analisis dari gaya hidup tidak cocok bila disajikan dalam bentuk grafik, charts, atau peta. Peneliti [31] membuat suatu model yang diambil dari analisis gaya hidup, kemudian disajikan berupa teks sehingga lebih menggambarkan pengalaman karena berbentuk cerita, seperti di blog.

Secara spesifik, Miwako Doi, 2012, menjelaskan dampak dari teknologi terhadap masyarakat [32]. Ia juga memprediksi teknologi masa depan beserta dampaknya. Analitik dari teknologi yang dapat dikembangkan adalah, pertama, dulu untuk mendengarkan musik harus berada di suatu tempat yang terdapat alat pemutar musik, misal tape atau radio. Seiring dengan kebutuhan masyarakat bahwa orang sedang bepergian juga sangat senang mendengarkan musik, maka sekarang ini pemutar musik dapat dibawa kemana-mana, misalnya ipod. Berkembang lagi penelitian bahwa untuk lagu-lagu yang disimpan di memory internal membuat boros ruang penyimpanan, sehingga sekarang teknologinya adalah menggunakan cloud memory seperti spotify, i-music, Google play music, Apple music, Youtube music, Amazon music, Vevo, Soundcloud, Joox, dll [32].

Independensi. Contoh lainnya adalah mesin cuci, dimana dahulu mesin cuci masih membutuhkan campur tangan manusia secara manual. Ketika kita tekan tombol on, mesin cuci akan berputar sesuai dengan durasi waktu yang kita kehendaki, kemudian kita akan membuang airnya secara manual dengan cara menekan tombol untuk membuang air bekas cucian. Dilanjutkan dengan pengisian air untuk pembilasan. Selanjutnya, pemindahan pakaian dari tempat bilas ke tempat pengering dimana mesin akan berputar sekitar 700 rpm, artinya kira-kira dalam satu menit mesin cuci dapat berputar sebanyak 700 kali. Penemuan selanjutnya adalah mesin cuci matic dimana manusia tidak lagi terlibat dalam mengisi air dan memindahkan pakaian dari bak pembilasan pada bak pengering, pengguna cukup menekan tombol mulai di awal, dan mesin cuci akan berhenti sendiri apabila tugasnya telah selesai dilakukan, sampai tahap pengeringan baju. Lalu pengembangan dilakukan lagi dengan cara mesin cuci diberi kemampuan analitik sehingga dapat bekerja secara independen [34]. Sekarang ini mesin cucidapat bekerja secara mandiri (otomatis menyalakan dirinya sendiri) ketika tuannya sedang berangkat kerja. Sepulang kerja, baju telah bersih dan kering. Ada pula mesin cuci yang dilengkapi fungsi tambahan yaitu dapat menyetrika baju dan melipat baju.

Interaktif. Kebutuhan manusia tidak hanya sandang, pangan, serta papan. Namun, berinteraksi dan berkomunikasi juga merupakan kebutuhan manusia yang harus dipenuhi karena sesungguhnya manusia itu adalah makhluk sosial dimana manusia tidak dapat hidup sendiri. Muncullah media sosial seperti Twitter, Facebook, dan Instagram dimana kita dapat memberikan atau menyebarkan informasi melalui sosial media, dan orang lain dapat mengomentarnya, bahkan menyebarluaskan kembali. Seperti pada gambar 4 bahwa seorang bocah dari Nairobi, Kenya yang bernama Lawrence sangat mengidolakan Mesut Ozil [33] sehingga ia mengenakan kaos yang tertulis nama Ozil beserta nomor punggungnya dengan tinta pulpen. Kemudian Eric Njiru pada 3 Desember 2018 mengunggah foto Lawrence dengan deskripsi foto sebagai berikut.

“Aku lihat bocah ini menggembala sapi di jalanan Nairobi, kota yang tak punya rumput. Dia ngomong suka sekali dengan @MesutOzil1088 (Kau bisa lihat bajunya). Mungkin suatu hari nanti Ozil dapat memberi seragam @Arsenal betulan”

Sekitar 14.000 pengguna menyukai postingan Eric ini, serta 9.800 orang membagikan kembali postingan ini. Hal ini membuat Ozil terenyuh dan memberikan Lawrence dan keluarganya sebuah paket berisi seragam Arsenal bertanda tangan, sepatu, topi, pernah pernik lainnya pada 11 Maret 2019.



Gambar 4 Bocah Penggembala Sapi Dapat Hadiah dari Mesut Ozil (Foto twitter @erick_njiru)

Interaktif lainnya yang sangat dibutuhkan pada saat terjadi bencana. Bencana dapat datang kapan saja dan dapat menimpa pada negara mana saja, namun dari kejadian yang telah terjadi di negara lain, dapat kita antisipasi untuk memitigasi bencana atau merencanakan apa yang harus dilakukan oleh masyarakat jika terjadi bencana. Contohnya di Jepang, dikarenakan seringnya terjadi bencana gempa bumi, tidak hanya bangunan saja yang didesain untuk tahan gempa bumi, tapi dipikirkan pula bagaimana agar berita dapat tetap bersifat interaktif antara di daerah yang mengalamibencana dengan daerah luar. Apabila terjadi bencana, seluruh daya akan mati sehingga tidak ada kabar dari tempat itu ke tempat lain, atau sebaliknya. Hal ini menyulitkan masyarakat dan pemerintah dalam hal pemberian bantuan kepada masyarakat. Dengan demikian, untuk barang elektronik yang sifatnya alat komunikasi seperti TV maupun radio dipasang baterai sehingga masyarakat masih dapat mengetahui apa yang sedang terjadi dan dapat mengetahui apa yang harus dilakukan. Misalnya, gempa bumi ini disebabkan oleh gearan yang berada di bawah laut, dan akan menimbulkan tsunami. Maka masyarakat dapat mengetahui apa yang harus dilakukan, yaitu menyelamatkan diri dengan cara segera mungkin pergi ke tempat yang lebih tinggi. Berbeda dengan gempa bumi yang diakibatkan gunung meletus, maka masyarakat tidak akan mendekati tempat yang tinggi. Pentingnya komunikasi ini menyebabkan peneliti bekerja sama dengan pemerintah untuk memitigasi bencana. Bayangkan apabila tidak ada power cadangan, tempat lain tidak bisa berkomunikasi, semua jaringan putus. Pemerintah pun tidak bisa mengetahui ada bencana. Sedangkan apabila ada power cadangan, selain listrik yang dikelola oleh negara, pemerintah akan mengetahui segera mungkin karena sistemnya seperti alarm. Begitu ada bencana, seluruh acara stasiun tv akan berhenti secara otomatis dan berganti menjadi sirine tanda kedaruratan yang dilengkapi dengan peta dimana bencana itu datang sehingga seluruh masyarakat yang memiliki sanak saudara di tempat yang terjadi bencana, dapat mengetahui kabar bahwa terjadi bencana di daerah itu dan pemerintah serta relawan – relawan dapat segera berdatangan untuk membantu pemulihan daerah bencana.

Gabungan fisik dan informasi. Dimulai sejak adanya e-commerce, sistem pembayaran pun berkembang menjadi non fisik. Orang didesain tidak lagi membawa dompet tebal, namun cukup dengan aplikasi, lalu login dengan menginput username dan password, maka transaksi pembelian maupun transfer sudah dapat dilakukan. Bahkan semenjak ada fintech, keberadaan bank konvensional semakin tergeser.

Fintech merupakan inovasi yang menggabungkan bisnis dengan teknologi, khususnya di bidang finansial [35]. Keberadaan fintech merupakan jawaban dari adanya gaps antara bisnis dan finansial yang semula terpisah, menjadi lebih terintegrasi. Misalnya untuk berbelanja di toko online, kita harus transfer atau mendatangi merchant tertentu yang sudah bekerja sama dengan toko online tersebut. Banyak kerugian yang dialami oleh customer. Pertama, adanya biaya admin. Transfer melalui bank seringkali menyebabkan adanya penambahan dana yang harus dikirimkan untuk melakukan suatu pembayaran. Kedua, effort tenaga dan waktu yang dihabiskan untuk mencapai atm terdekat atau merchant tertentu. Dengan adanya fintech, customer dimudahkan karena uang sudah ada di aplikasi. Tidak hanya berbelanja online, berbelanja secara fisik pun sudah dapat dilakukan dengan fintech ini. Dari mulai transportasi, belanja kebutuhan makanan, pakaian, perabotan rumah tangga, semua hanya semudah menggerakkan jari.

Tidak hanya menyediakan sarana untuk menyimpan uang kita, fintech ini juga dapat menyediakan fasilitas pinjaman uang. Kini untuk mengajukan pinjaman tidak perlu antre di bank kemudian memberikan persyaratan atau survei ke tempat tinggal atau tempat usaha. Kini hanya dengan aplikasi, pengajuan pinjaman dapat secara mudah dilakukan.

B. Gaps

Uncertainty. Analisis big data menggunakan operasi statistika yang berarti menggunakan metode peluang. Sebagai contoh, dari suatu pemodelan diperoleh atribut-atribut apa saja yang berpengaruh. Kita akan memisalkan atribut tersebut adalah sebuah node dimana hanya node yang berada di pemodelan saja yang saling terhubung, sehingga node baru dianggap sebagai node yang saling lepas (tidak terhubung). Sementara, hal ini belum bisa diyakini sebagai kebenarannya [36]. Hal ini dikarenakan pemodelan dilakukan secara makro.

Berbeda halnya jika kita menggunakan konsep micro-process. Micro process adalah proses yang berskala kecil [37]. Dengan demikian, node-node yang kecil akan digabungkan menjadi sebuah network sehingga lebih bisa menggambarkan keadaan atau behaviour yang sebenarnya. Konsep ini sama dengan konsep Service Oriented Architecture yang dimunculkan oleh tokoh SOA yang bernama Thomas Erl [38]. Arsitektur SOA yang diusung terdiri dari bottom up dan top down. Yang dimaksud dengan penelitian yang diusulkan oleh [37] adalah arsitektur Thomas Erl yang jenisnya bottom up.

Analisis big data yang menggunakan prinsip bottom up biasanya merupakan penelitian terkait behaviour dan real time data. Penelitian [38] mengumpulkan log kegiatan dari customer dalam menggunakan e-commerce, dari mulai pencarian barang, maupun pemilihan barang, dan cara pembayaran barang. Dari hasil analitiknya menggunakan Hadoop disertai Pig dan Hive, dapat dibuat suatu penawaran untuk customer menurut jenis barang sesuai dengan history pencarian atau barang-barang yang telah dibeli. Selain itu, dapat digunakan untuk

analisis promo atau diskon. Penelitian [39] menyebutkan bahwa promosi di awal akan mengakrabkan customer dengan produk baru, kemudian selanjutnya akan menjad customer tetap. Secara lebih dalam, peneliti menyebutkan bahwa dalam membuat suatu promosi, perlu ada “sesuatu” yang dapat diingat oleh customer: musik, cara mengekspresikan fitur, teknik menggunakan, konten, stimulating affect, ketersediaan informasi mengenai produk tersebut, serta slogan.

Penelitian lain yang meneliti serupa dengan [38] adalah [40] yang mencoba untuk memprediksi berita apa yang akan dibaca sehingga jenis iklan akan disesuaikan dengan berita yang dibaca (anggapan: berita yang dibaca adalah menurut selera dari pembaca sehingga iklan yang ditawarkan sesuai dengan selera pembaca pula. Penelitian [40] menggunakan algoritma Vector Space untuk memprediksi iklan yang ditampilkan dengan analisis data metode pengelompokan. Semua iklan dikelompokkan menjadi beberapa kelompok. Setelah itu, berita yang dibaca, dihitung probabilitasnya masuk ke kelompok iklan mana.

Industrial Peer. Implementasi hasil penelitian merupakan bukti nyata kebaruan dan kebermanfaatannya hasil penelitian. Menurut kamus Webster [40] definisi dari implementasi adalah suatu proses pemindahan ide/gagasan dari sebuah konsep menjadi realita. Tidak dapat dipaksakan sebuah penelitian harus diakhiri dengan implementasi, tetapi seyogyanya atau secara sempurna tahapan penelitian itu diawali dengan perencanaan, diikuti oleh perancangan, dan diakhiri dengan implementasi. Ada pula yang menambahkan evaluasi sebagai tahap akhir. Ada terdapat macam-macam metode penelitian, tergantung dari objek yang diteliti. Buku [41][42] memaparkan metode yang berbeda untuk data kualitatif, kuantitatif, maupun gabungan.

- **Metode Kualitatif**
Metode kualitatif digunakan biasanya untuk meneliti data yang berbentuk teks maupun gambar. Biasanya peneliti akan melakukan observasi atau wawancara, kemudian hasil dari observasi atau wawancara akan direkam dengan struktur data yang tidak terstruktur dan semi terstruktur. Instrumennya berupa pertanyaan terbuka yang membuat partisipan dapat menjawab pertanyaan sesuai dengan ide dan kemampuan menjawabnya. Selain instrumen di atas, dapat juga dengan mempelajari dokumen—dokumen yang tersedia. Untuk data jenis gambar dapat menggunakan visual etnografi, living stories, metaphorical visual naratif, dan digital archive.
- **Metode Kuantitatif**
Untuk mendapatkan data kuantitatif biasanya dilakukan pengambilan data dengan cara sensus dan survei. Adapun penelitian yang menggunakan metode kuantitatif biasanya perihal mencari kethubungan antara dua atau lebih variabel. Selain itu, dapat juga mengenai pembuktian sebuah hipotesis atau asumsi tertentu. Tentunya apabila data diambil dengan cara survey, maka perlu diperhatikan mengenai jumlah dan karakteristik dari sampel sehingga sampel memenuhi persyaratan minimumnya dan dapat merepresentasikan populasi seluruhnya. Oleh sebab itu, perlu dilakukan analisis sampling method agar sampel yang diambil dapat menggambarkan populasi objek penelitian seluruhnya. Variabel yang diikutsertakan terdiri dari variabel dependen dan variabel independen. Misalnya pengaruh banyaknya pemakaian pulsa listrik terhadap besarnya uang yang dikeluarkan. Maka yang menjadi variabel independen adalah pulsa listrik yang digunakan, sedangkan uang yang harus dibayarkan adalah variabel dependen.
- **Metode Gabungan**
Metode gabungan dilakukan dikarenakan data yang tersedia gabungan dari open-ended answer dan closed-ended answer. Ketika objek yang dijadikan penelitian merupakan gabungan dari kualitatif dan kuantitatif, maka yang dikerjakan adalah dengan menggunakan metode gabungan. Menurut penelitian dari [43] bahwa biasanya untuk penelitian yang meneliti keadaan sosial atau behavior dari masyarakat, biasanya datanya berupa kualitatif kemudian ada juga data kuantitatif sehingga diperlukan metode gabungan kuantitatif dan kualitatif.

Implementasi adalah tahapan yang akan dilalui baik itu menggunakan metode kualitatif, kuantitatif, maupun gabungan. Implementasi juga cakupannya sangat luas, dapat merambah ke berbagai bidang keilmuan. Pentingnya implementasi sehingga berbagai penelitian menimbulkan “efek” bagi bisnis proses, efektifitas, keilmuan, metodologi, kehidupan sosial, dan sebagainya. Terkadang untuk melakukan implementasi, akan sedikit mengubah konsep perancangan karena setelah di dunia nyata, ternyata perancangan yang dibuat tidak dapat diimplementasi. Adapun kunci sukses dari implementasi adalah memiliki karakteristik sebagai berikut:

- **Acceptable**
Penelitian yang dibuat diharapkan dapat diterima oleh subject matter karena yang menggunakan adalah user sehingga diharapkan penelitian dapat membantu pekerjaan user, memudahkan segala urusan user.
- **Adoption**
Penelitian diharapkan dapat diadopsi atau direusability oleh peneliti lain sehingga peneliti lain dapat mengembangkan, meneruskan, tanpa perlu mengulangi dari awal. Oleh sebab itu, dokumentasi sangatlah penting dalam hal ini.
- **Appropriateness**

Kecocokan ini berhubungan dengan study field. Oleh sebab itu, dalam setiap penelitian biasanya disebutkan bahwa penelitian dilakukan untuk study kasus dimana. Tidak dapat dipungkiri bahwa setiap daerah memiliki karakteristik tertentu sehingga penelitian di daerah A belum tentu cocok untuk diterapkan di daerah B.

- Feasibility
Yang dimaksud dengan feasible adalah layak. Diharapkan penelitian minimal memiliki predikat “good enough” sehingga dapat digunakan. Namun, terdapat beberapa kasus yang sangat darurat atau memiliki kepentingan emergency sehingga predikat dari penelitian, khususnya software, harus berpredikat “perfect”, misalnya software untuk pesawat terbang sehingga menyangkut keamanan dari semua penumpang pesawat.
- Fidelity
Maksudnya adalah jangan sampai penelitian yang dikerjakan membahayakan keamanan dari user. Misalnya software yang dibuat jangan sampai merupakan virus demi mencuri data user.
- Implementation cost
Perlu dipikirkan biaya untuk implementasi sehingga jangan sampai biaya implementasi tidak melebihi total budget seluruhnya
- Coverage
Untuk coverage ini perlu diperhatikan bahwa implementasi dari perancangan telah memenuhi kebutuhan user
- Sustainability
Diharapkan hasil dari penelitian dapat digunakan user untuk waktu yang lama sehingga mungkin pihak pengembang dapat memberikan garansi untuk rentang waktu tertentu dalam rangka maintenance program.

C. Opportunities

Prospek. Menurut penelitian dari [45] bahwa bidang analitis memberikan prospek yang sangat cerah bagi masa depan para pelaku di dalamnya. Para manager atau pimpinan dari perusahaan atau institusi membutuhkan logika analitis karena para pimpinan akan menjamin kebenaran dan kesesuaian dari proyek yang dipimpin sehingga mencapai apa yang ditargetkan. Ada tiga tujuan yang ingin dicapai dalam setiap proyek yang dikerjakan: on time, on budget, dan on desire. Artinya proyek yang dipimpin harus sesuai dengan target penyelesaian sehingga kemampuan penjadwalan dan strategi penyelesaian perlu dimiliki oleh manager atau pimpinan institusi. Selain waktu, hal yang terpenting adalah budgeting dimana biaya yang dikeluarkan tidak boleh melebihi budget yang direncanakan, harus mempertimbangkan pengeluaran-pengeluaran yang diluar perencanaan. Terakhir yaitu, proyek yang dihasilkan harus sesuai dengan yang diinginkan oleh user sehingga perlu dengan hati-hati dalam evaluasi bukan hanya menguji fungsionalitasnya, namun juga acceptance dari user. Untuk mencapai ketiga tujuan tersebut, diharapkan pimpinan memiliki skill analitis yang memadai. Ditambah lagi sekarang zamannya serba big data. Big data sangat erat hubungannya dengan smart analytic. Jika kita memiliki keahlian di bidang ini, maka dapat diperkirakan kita memiliki masa depan yang cerah.

Everyone Can Be A Data Scientist. Untuk menjadi data scientist tidak harus menjadi programmer, tidak juga harus menjadi ahli statistik, atau tidak juga harus jago di bidang tertentu. Percaya atau tidak, bidang analitik merupakan proses gabungan atau kolaborasi dari semua aktor yang terlibat. Untuk mengumpulkan dan menjalankan tools diperlukan ahli programmer, untuk mengolah data statistik diperlukan orang yang ahli di bidang statistik, sedangkan untuk dapat mengartikan dari sebuah data atau hubungan antar variabel atau efek dari sebuah perlakuan dapat dilakukan oleh yang ahli di bidang itu, misalnya ahli sosial mengamati masalah-masalah sosial sehingga ia akan memiliki asumsi atau hipotesis dari suatu variable terhadap isu tertentu. Terntunya untuk mengolah data tersebut diperlukan bantuan dari programmer dan statistisi

IV. CONCLUSION

Pembedakan smart analytic dengan analytic biasa adalah adanya fitur tambahan dengan melibatkan Artificial Intelligence dimana sistem dapat menemukan sendiri suatu informasi yang berasal dari big data. Sedangkan metode yang digunakan biasanya berupa clustering, decision tree untuk menghasilkan suatu keputusan berdasarkan atribut yang terlibat, dependency, dan sistem rekomendasi. Tools yang digunakan sangat banyak, namun hanya beberapa disampaikan disini. Kemudian dari penelitian yang ada, sekarang analitik sudah dimanfaatkan bukan hanya untuk otomatisasi, melainkan untuk kontrol dari jarak jauh, antar device saling berhubungan. Namun, masih kurang dalam implementasi sehingga perlu ada usaha untuk mengintegrasikan antara penelitian dengan dunia industri. Peluang untuk berada di zona analisis masih terbuka sangat lebar dikarenakan big data memerlukan analisis dan semua orang bisa menjadi data scientist.

REFERENCES

- [1] Jindal, Anish et al. 2018. Consumption Aware Data Analytical Demand Respond Scheme for Peak Load Reduction in Smart Grid. Published in IEEE Transactions on Industrial Electronics November 2018
- [2] Malvar, Hendrique and David H Staelin. 1988. Optional Pre Postfilters for Multichannel Signal Processing. Published on IEEE Transactions on Acoustic, Speech, and Signal, Febr 1988
- [3] Vatrappu, Ravi et al. 2016. Social Set Analysis: A Set Theoretical Approach to Big Data Analytics. Published in IEEE Access
- [4] Barbara Kitchenham etc, Systemic Literature Review in Software Engineering – A Tertiary Study, *Information and Software Technology*, 2010. Available from : www.elsevier.com/locate/infsof
- [5] <https://becominghuman.ai/lets-talk-about-advanced-analytics-a-brief-look-at-artificial-intelligence-bf1c7a7d3f96>
- [6] Bengt Ahlgren, 2016, Internet of Things for Smart Cities: Interoperability and Open Data. IEEE Internet Computing
- [7] J. Carroll, J Chen, C. Yuan, and B. Hanrahan. “In Search of Coproduction: Smart Services as Reciprocal Activities”. *Computer*, Vol 49 No.7, Jan 2016 pp 26-32
- [8] Michael Weyrich and Christof Ebert. 2016. Reference Architectures for the Internet of Things. IEEE Software. January/February 2016
- [9] Lankow, Jason, Josh Ritchie, and Ross Crooks. 2012. Infographics : The Power of Visual Storytelling
- [10] G. Lin and W. Tang, “Wearable sensor patches for psychological monitoring”, in *NASA Tech Briefs : Engineering Solutions for Design and Manufacturing* , 2000
- [11] R. Paradiso, G. Loriga , and Taccini, “A wearable health care system based on knitted integrated sensors”, *IEEE Trans. Inf. Technol B Vol.9 pp 337-344, 2015*
- [12] M. Pacelli, G. Loriga, N. Taccini, and R. Paradiso “Sensing fabrics for monitoring psychological and biomechanical variables: E textile solutions” in *IEEE Engineering in Medicine and Biology Society , New York, USA, 2006*
- [13] P.A. Shaltis, A.T Reisner, and H.H. Asada, “Cuffless blood pressure monitoring using hydrostatic pressure changes,” *IEEE Trans. Biomed. Eng. Vol 55 No 6 pp 1775 – 1777 June 2008*
- [14] M-Z. Poh, K.Kim, A. Goessling, N. Swenson, and P. Picard, “Cardiovascular monitoring using earphones and a mobile device”, *IEEE Pervasive Comput, Vol 11 No. 4, pp 18-26. Oct-Dec 2012*
- [15] C. Strohrmann, H. Harms, C. Kappeler-Setz, and G. Tröster, “Monitoring kinematic changes with fatigue in running using bodyworn sensors,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 5, pp. 983–990, Sep. 2012
- [16] J. Dean and S. Ghemawat. MapReduce: simplified data processing on large clusters. *Commun. ACM*, 51(1):107–113, Jan. 2008
- [17] Acharjya, Debi Prasanna and Kauser Ahmed P. 2016. A Survey n Big Data Anaytic: Challenge, Open Research Issue and Tools in *International Journal of Advanced Computer Science and Application* February 2016
- [18] Patel, Aditya B, Manashvi Birla, Ushma Nair. 2012 Addressing Big Data Problem Using Hadoop and Map Reduce. In 2012 Nirma University International Conference on Engineering, Nuicone-2012, 06-08 December, 2012
- [19] Alpaydin, Ethem. 2014. Introduction to Machine Learning 3rd Edition. Massachusetts Institute of Technology: Massachusetts
- [20] Witten, Ian H and Eibe Frank. 2005. Data Mining: Practical Machine Learning Tools and Techniques. Elsevier Inc: USA
- [21] Meng Xiangrui et al. 2016. Mlib: Machine Learning in Apache Spark. In *Journal of Machine Learning Research* 17 (2016) 1-7
- [22] <https://www.edureka.co/blog/spark-architecture/>
- [23] H, Li, G. Fox, and J, Qiu, Performance model for parallel matrix multiplication with dryad: Dataflow graph runtime, Second International Conference on Cloud and Green Computing, 2012 pp.675-683
- [24] Isard, Michael et al. 2007. Dryad: Distributed Data-Parallel Programs from Sequential Building Blocks. Published in EuroSys '07 Proceedings of the 2nd ACM SIGOPS/EuroSys European Conference on Computer Systems 2007
- [25] <https://id.wikipedia.org/wiki/FIFO>
- [26] M, Hausenblas et al. 2013. Apache Drill: Interactive Ad-Hoc Analysis at Scale. *Journal Big Data*
- [27] Krum, Randy. 2014. Cool Infographic: Effective Communication Data Visualization and Design. John Wiley & Sons, Inc: Canada
- [28] Khalil, Ashraf and Suha Glal. 2009. Step Up: A Step Couner Mobile Application to Promote Healthy Lifestyle
- [29] Nakamura, Jiro et al. 2003. Impact of Spread of Informations Technology on Lifestyle Changes. Proceeding of EcoDesign2003: Third International Symposium on Environmentally Conscious Design and Inverse Manufacturing Tokyo, Japan, December 8-11, 2003
- [30] Pavel, Dana et al. 2013. Lifestyle Stories: Correlating User Information through a Story-Inspired Paradigm.
- [31] <https://www.idntimes.com/tech/trend/jcnd/aplikasi-terbaik-untuk-streaming-musik-di-ios-android-c1c2/full>
- [32] <https://www.dream.co.id/sport/bocah-penggemballa-sapi-di-kenya-dapat-hadiah-dari-ozil-190313f.html>
- [33] Doi, Miwako et al. 2012. Personal and Home Electronics and Our Changing Lifestyles. Invited paper proceeding IEEE vol 100, May 13th, 2012
- [34] Philippon, Thomas. 2016. The Fintech Opportunity. Worpping paper series in National Bureau of Economic Research.
- [35] Cnijders, Chris. 2012. Big Data: Big Gaps of Knowledge in The Field of Internet Science. *International Journal of Internet Science* 2012, 7(1), 1-5
- [36] <https://en.oxforddictionaries.com/definition/microprocess>
- [37] Erl, Thomas. 2017. Service Oriented Architecture.
- [38] Ansari, Shahriar. 2011. The Affect of Sales Promotion on Customer Interest to Purchase in IKCO Automative Company. Published in *Journal of Knowledge Management, Economics and Information Technology*
- [39] Webster Collegiate Dictionary

- [40] Yusuf, A Muri. 2017. *Metode Peneliiian Kuantitatif, Kualitatif, dan Penelitian Gabungan*. Gramedia: Jakarta
- [41] Creswell, John W. 2014. *Research Design: Qualitative, Quantitative, and Mixed Method Approach*
- [42] Tashakkori and Teddle. 2003. *The Handbook of Mixed Methods In the Social and Behavior Sciences*.
- [43] Peters, DH et al. 2013. *Implementation Research: What it is and How to Do it*. Published in *BMJ* 2013 Nov 20;347:f6753. doi: 10.1136/bmj.f6753.