Ismail Fahmi, S.T., M.A., Ph.D.

Ismail.fahmi@gmail.com

|	
1992 – 1997	Undergraduate, Electrical Engineering, ITB, Indonesia
2003 – 2004	Master, Information Science, University of Groningen, NL
2004 – 2009	Doctor, Information Science, University of Groningen, NL
2009 – Now	Engineer at **Weborama** (Paris/Amsterdam)
2014 – Now	Founder **PT. Media Kernels Indonesia**, a Drone Emprit Company
2015 – Now	Konsultan di **Perpustakaan Nasional**, Inisiator Indonesia OneSearch
2017 – Now	Dosen di IT Magister Program of the **Universitas Islam Indonesia**
2021 – Now	Wakil Ketua Komisi Infokom, **Majelis Ulama Indonesia Pusat**
2022 – Now	Wakil Ketua Majelis Pustaka dan Informasi, **PP Muhammadiyah**

AGENDA

Media Kernels Indonesia

- Tentang Kami
- Democracy 4.0 berbasis Big Data dan AI
- Big Data Analytics dan AI dalam Pemerintahan
- Big Data untuk Official Statistics
- Tantangan dan Solusi
- Platform Big Data Analytics dan AI untuk Pemerintahan
- Kesimpulan

TENTANG KAMI

Media Kernels Indonesia

SISTEM 'AIS' KOMINFO
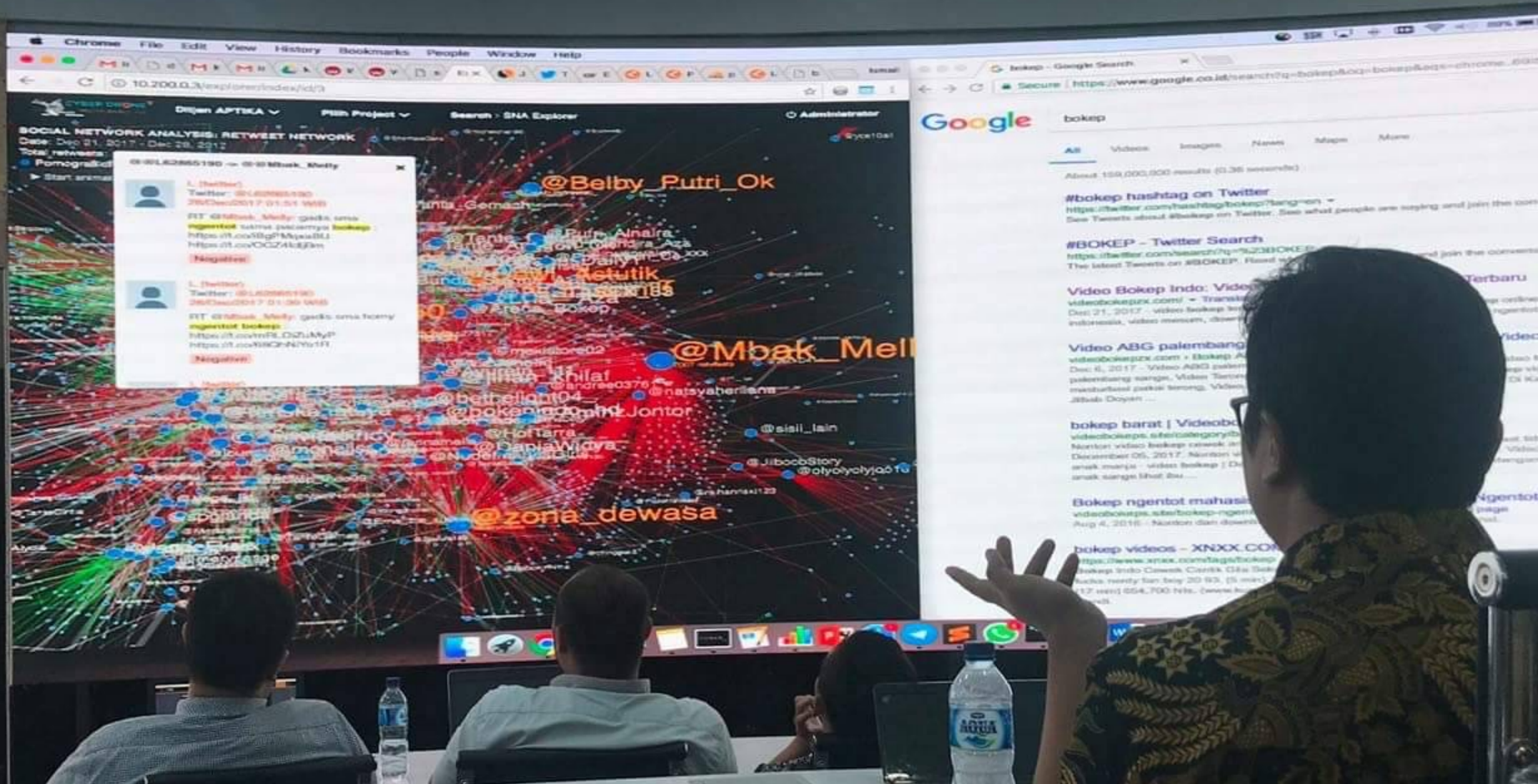
PIKOBAR JAWA BARAT

Drone Emprit Customized

Drone Emprit

DEMOCRACY 4.0
BERBASIS BIG DATA ANALYTICS DAN AI

Media Kernels
Indonesia

QUESTION:
IN INDUSTRY 4.0, WHERE IS OUR DEMOCRACY?

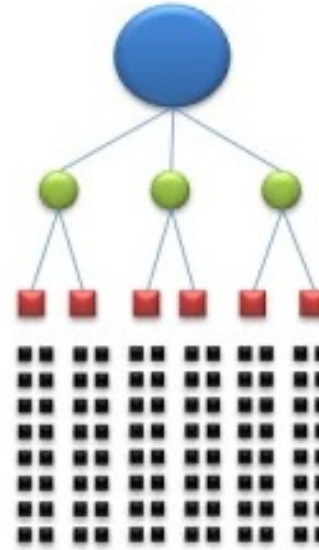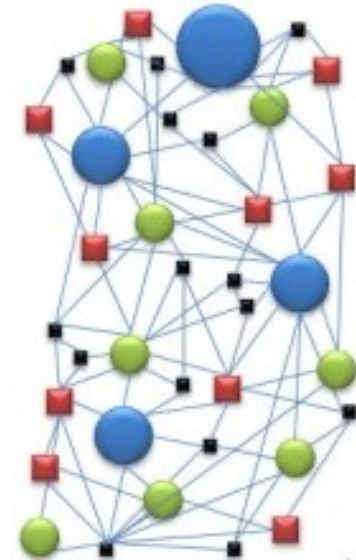| 1st	2nd	3rd	4th
Mechanization, water power, steam power	Mass production, assembly line, electricity	Computer and automation	Cyber Physical Systems
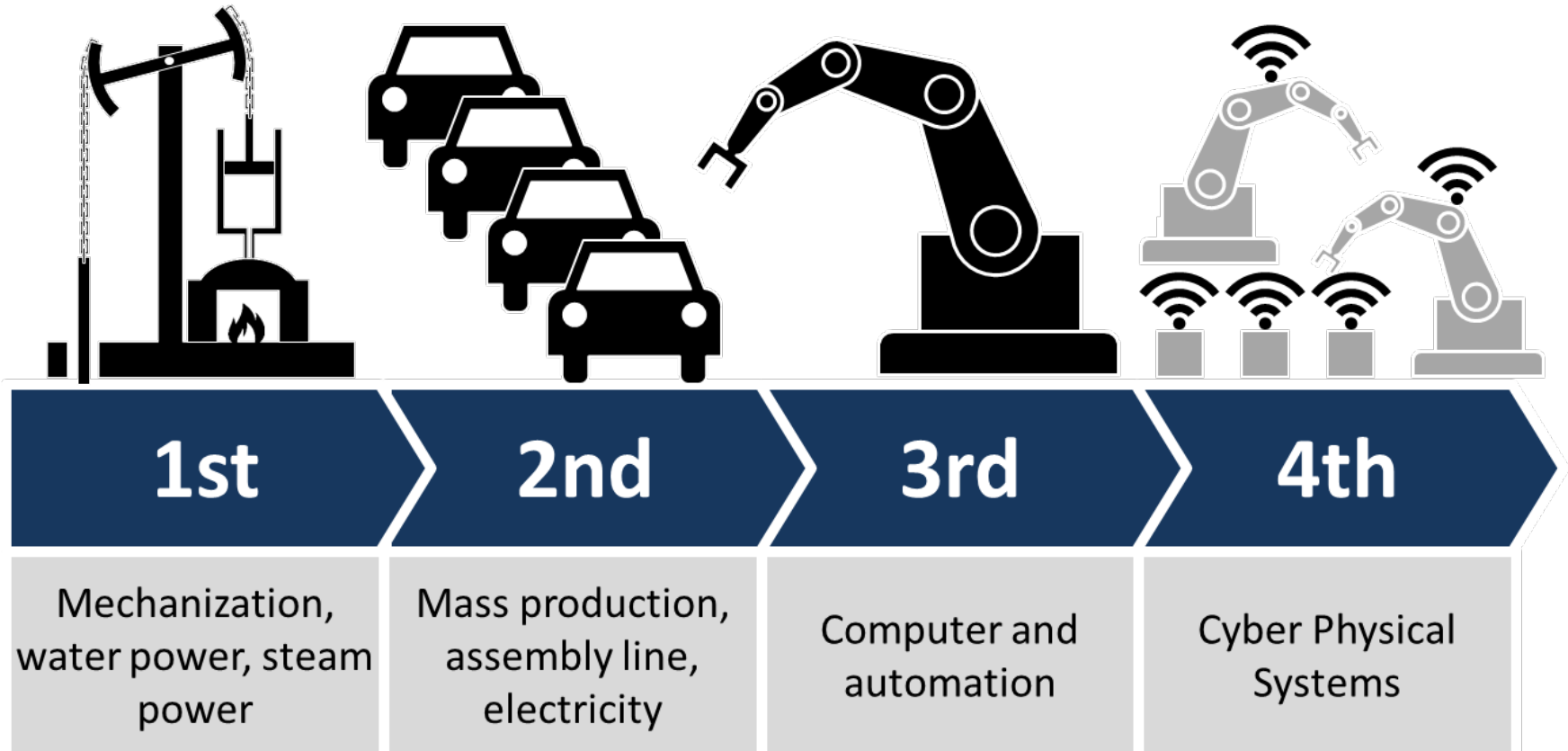
WHICH SOCIETY IS OURS?
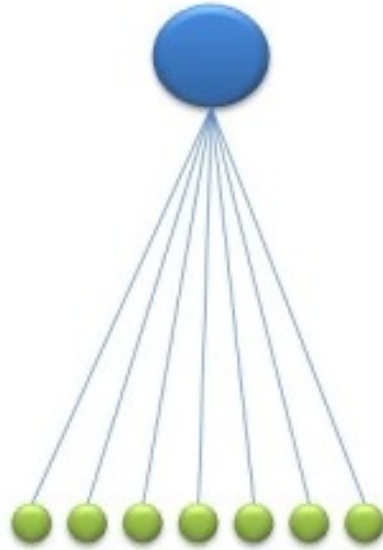
Tribal Society
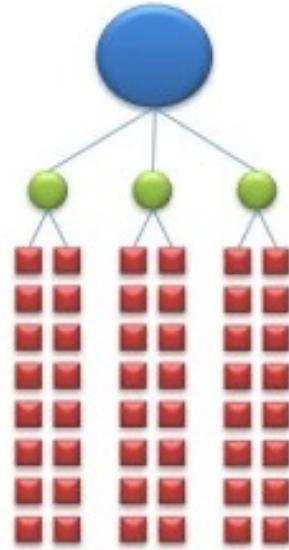
Ancient Civilization

Industrial Society

Network Society

© Isabella Mader

GENERAL VOTING

- extremely open to abuse
- misused as a legitimation tool
- badly informed citizens vote
- some issues have been populistically manipulated

E.g. by hoax, propaganda, and influence operation.

RESIDU PILPRES:
VICIOUS CYCLE OF NOISES

THE NEW SOCIAL CONTRACT: FROM REPRESENTATIVE TO PARTICIPATIVE

- Participation is not as in a pure online voting.

- Participative:
  - **Personal engagement** in civil society projects
  - Join the **discussion** and development of public policy and programs.

The New Social Contract
Industrial Society vs. Network Society

Representative Democracy	Participative Democracy
Competition	Cooperation
Growing, social inequality	Tax relief for small incomes
Austerity	Public investment creates return in GDP
Overregulation	Lean regulations
Employment	Freelancers & Entrepreneurs
Shrinking education budgets	Education standard: new culture techniques

© Isabella Mader

DEMOCRACY 4.0: PARTICIPATORY DECISION MAKING

Participatory Decision Making in 21st Century Politics & Technology
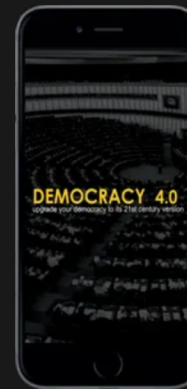
Reinventing political engagement as like never before…

- Interactive
- Cooperative
- Creative
- Bold
- Personal
- Fun and entertaining
- As easy as talking to your peers
- Fact-based
- Constructive
- Informative both for the politicians and citizens
- Advance technology with an ethical approach on data Re-imagine a political environment in the digital era of 21st Century.Democracy 4.0

14

BIG DATA ANALYTICS DAN AI UNTUK DEMOCRACY 4.0

BIG DATA ANALYTICS DAN AI DALAM PEMERINTAHAN

Media Kernels Indonesia
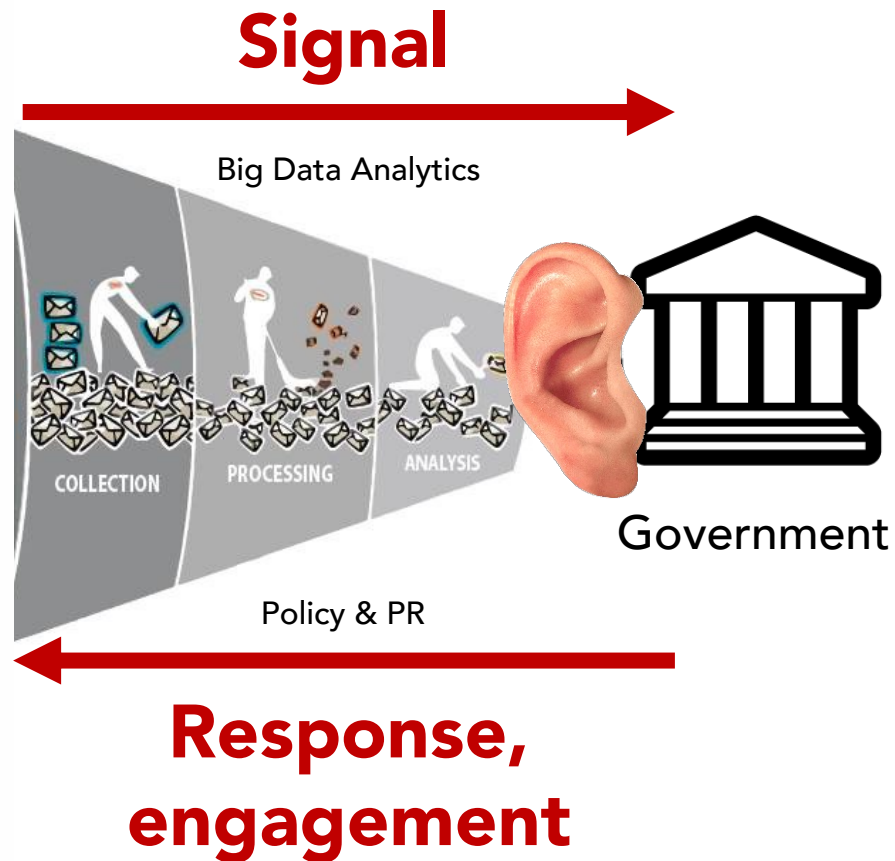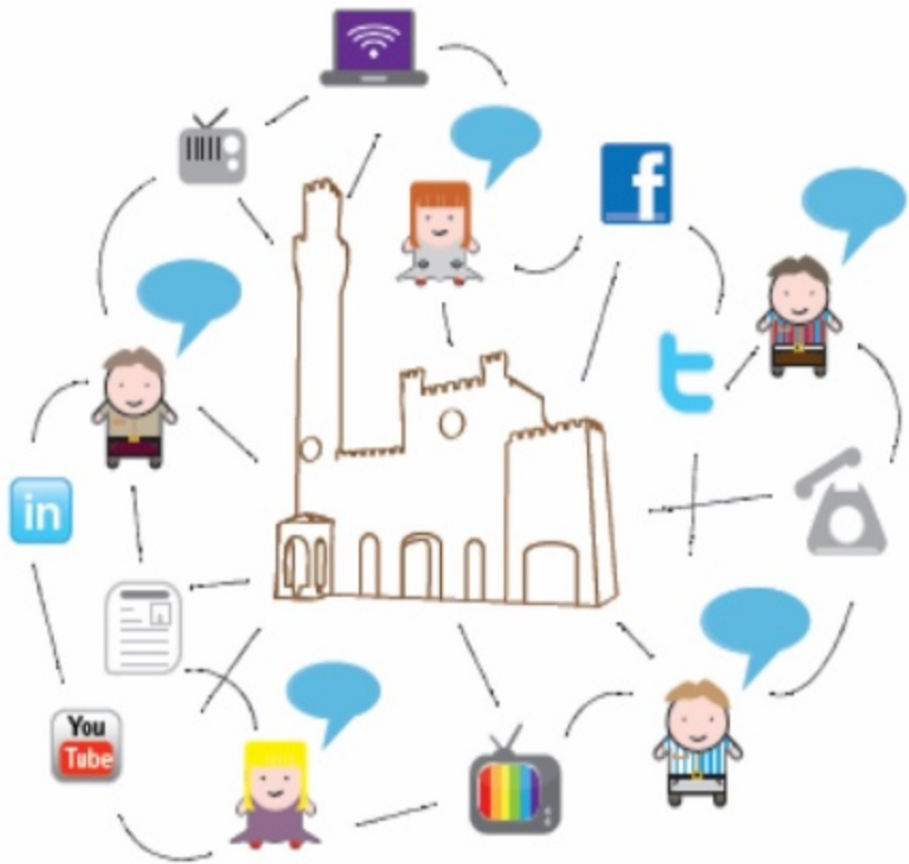
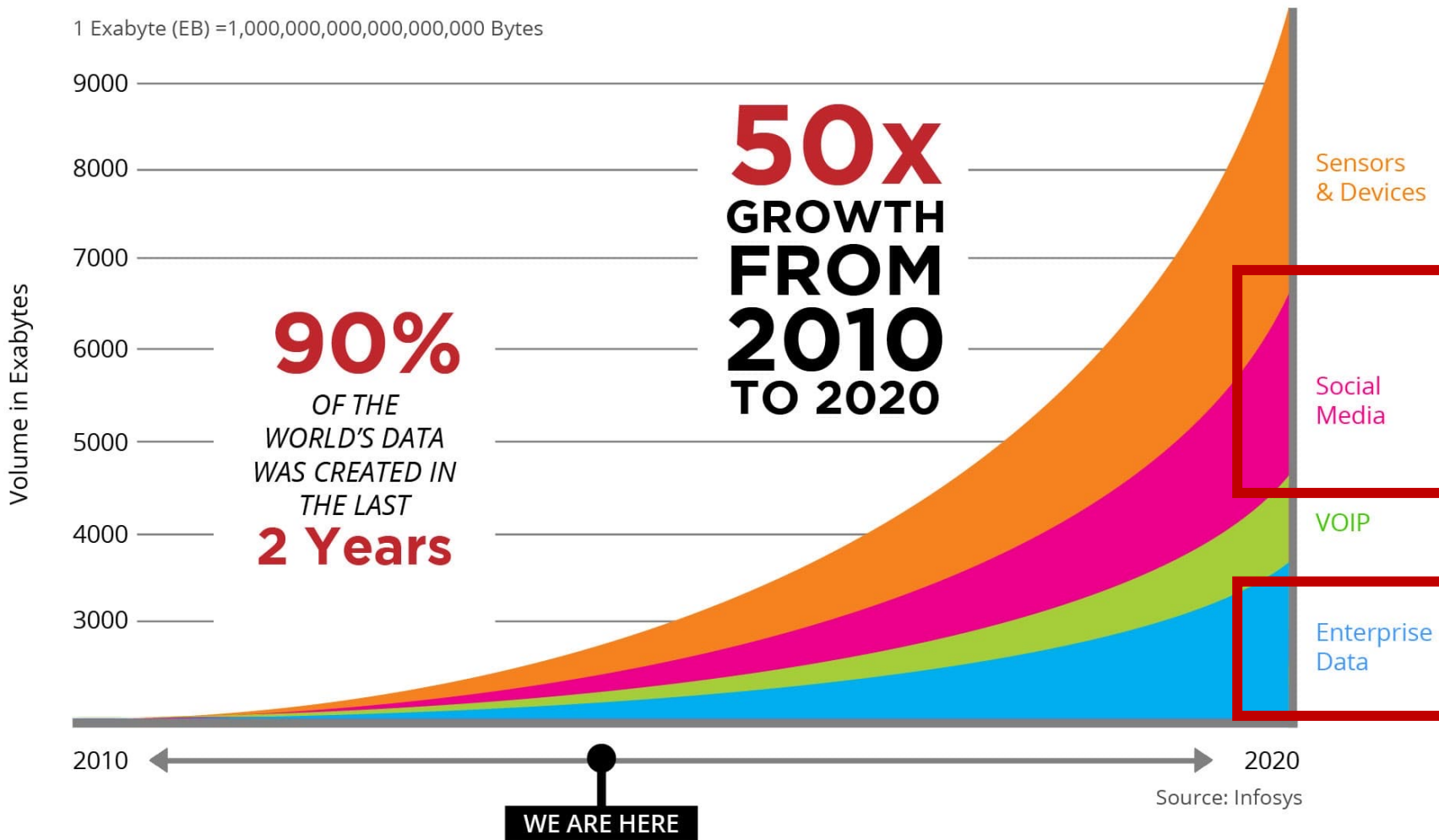# LATAR BELAKANG

- Big data Analytic telah menjadi kebutuhan seluruh Instansi Pusat Pemerintah Daerah (IPPD).

- Saat ini beberapa IPPD tengah mengusulkan pembiayaan bagi pembangunan Big data Analytic untuk kepentingan sektor Pemda masing-masing.

- Pembangunan sistem Big data Analytic secara masing-masing oleh IPPD memiliki konsekuensi Inefisiensi anggaran negara dan Inefektivitas analisis data.

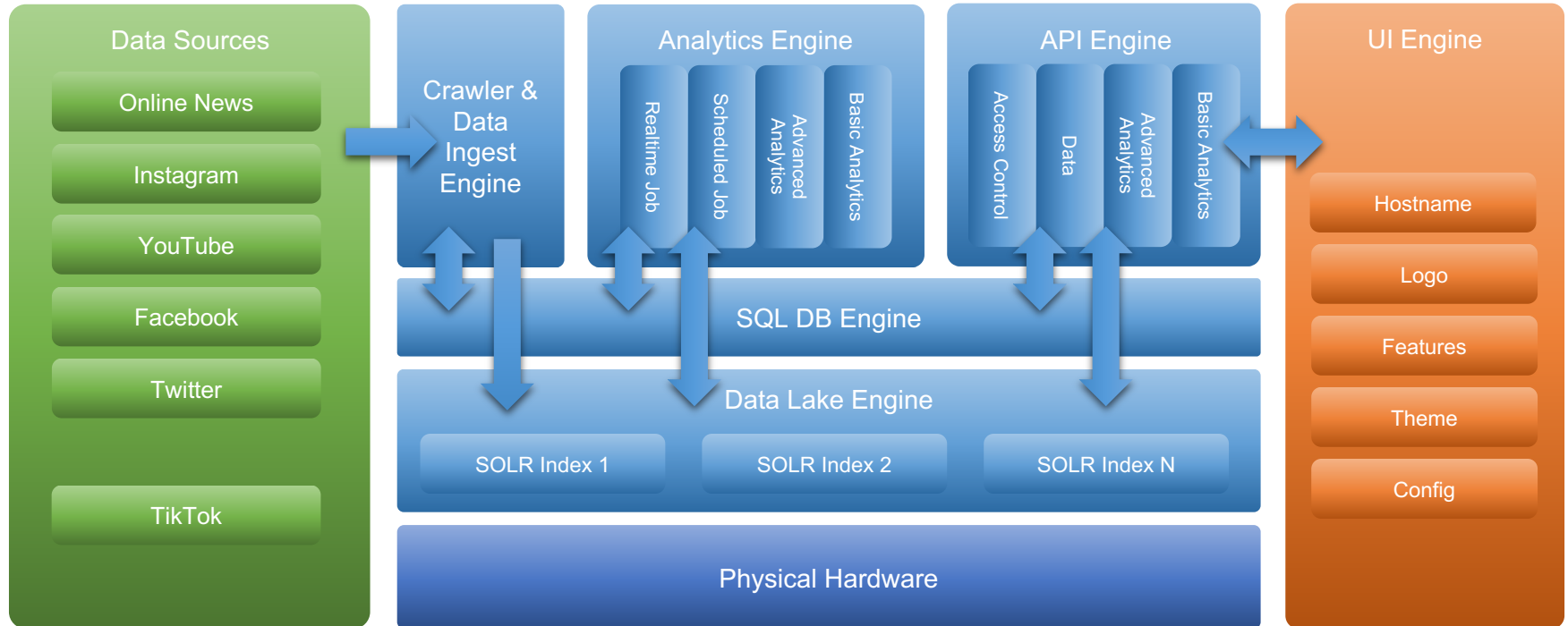# CONTOH OUTCOME BIG DATA ANALYTICS DAN AI DI PEMERINTAHAN

1. **Peningkatan Efisiensi dan Efektivitas**: Melalui analisis data yang mendalam, pemerintah dapat mengidentifikasi inefisiensi dalam operasional dan alokasi sumber daya, sehingga dapat mengambil tindakan untuk mengoptimalkan penggunaan sumber daya.

2. **Pelayanan Publik yang Lebih Baik**: Dengan memahami kebutuhan dan preferensi masyarakat melalui analisis data, pemerintah dapat menyediakan layanan yang lebih sesuai dan responsif terhadap kebutuhan masyarakat.

3. **Pengambilan Keputusan Berbasis Data**: Keputusan yang didasarkan pada analisis data cenderung lebih akurat dan relevan, sehingga dapat meningkatkan kualitas kebijakan publik.

4. **Deteksi dan Pencegahan Penipuan**: Sistem AI dapat membantu mendeteksi pola-pola penipuan dalam pengeluaran atau klaim pemerintah, sehingga mengurangi kerugian finansial.

5. **Peningkatan Transparansi**: Dengan menyediakan akses ke data pemerintah melalui platform big data, masyarakat dapat memiliki wawasan yang lebih baik tentang bagaimana pemerintah bekerja, yang pada gilirannya dapat meningkatkan kepercayaan publik.

6. **Prediksi dan Antisipasi Perubahan**: Dengan teknologi AI, pemerintah dapat membuat prediksi tentang tren masa depan, seperti pertumbuhan penduduk, kebutuhan infrastruktur, atau risiko bencana alam, dan merencanakan strategi proaktif.

7. **Otomatisasi Tugas-tugas Rutin**: Banyak tugas administratif dan rutin dapat diotomatisasi dengan AI, membebaskan sumber daya manusia untuk fokus pada tugas-tugas yang lebih kompleks.

BIG DATA – BIG GROWTH

Media Kernels Indonesia

1 Exabyte (EB) =1,000,000,000,000,000,000 Bytes

Volume in Exabytes

9000
8000
7000
6000
5000
4000
3000

90% OF THE WORLD'S DATA WAS CREATED IN THE LAST **2 Years**

50x GROWTH FROM **2010** TO **2020**

Sensors & Devices

Social Media

VOIP

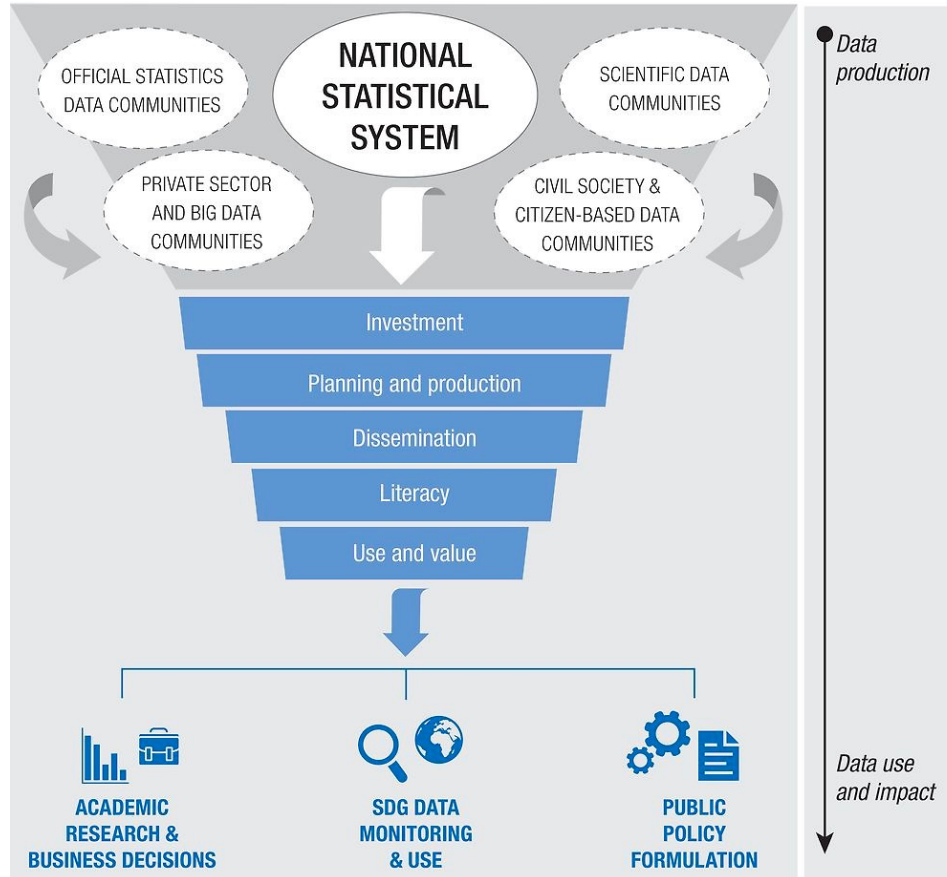Enterprise Data

2010

WE ARE HERE

2020

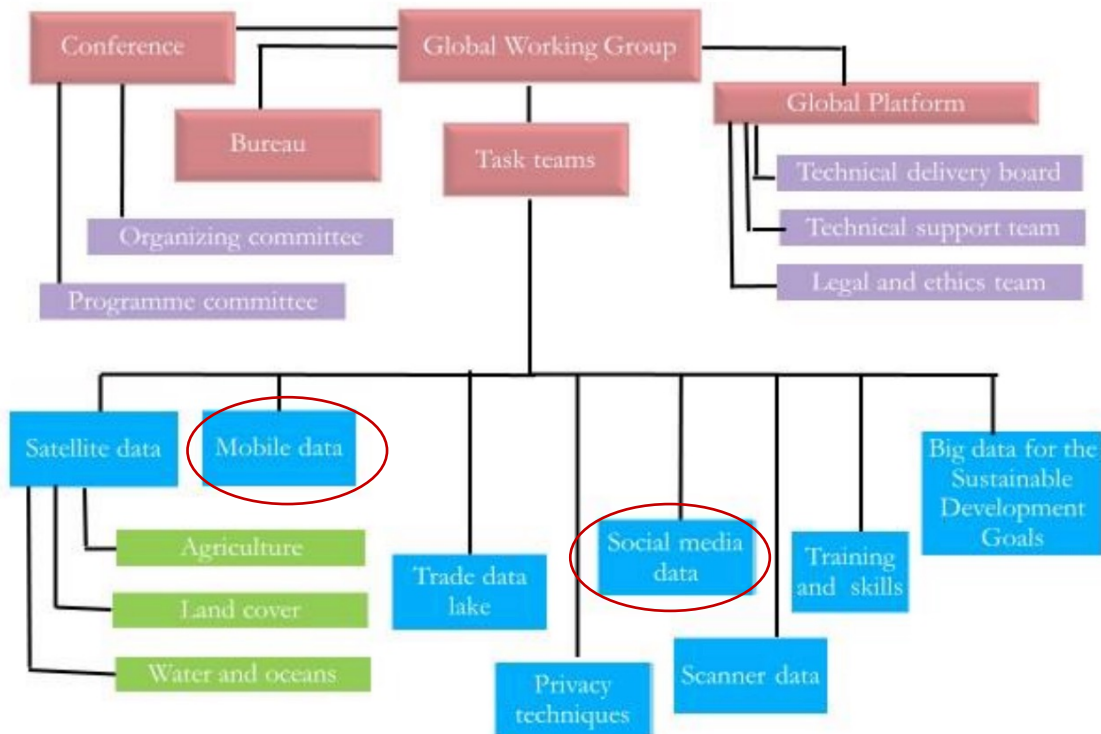Source: Infosys

CONTOH ARSITEKTUR SISTEM BIG DATA DAN AI

BISAKAH BIG DATA UNTUK STATISTIK RESMI?

LAYANAN PUBLIK BUTUH STATISTIK RESMI

UN GLOBAL WORKING GROUP ON BIG DATA FOR OFFICIAL STATISTICS

https://unstats.un.org/bigdata/bureau/documents/reports/GWG%20report%20-%202019-27-BigData-E.pdf

BIG DATA SOURCE: SOCIAL NETWORK

Social Networks **(Human-sourced information)**

Mostly digitized data, loosely structured and often ungoverned
- **Social Networks: Facebook, Twitter, Tumblr etc.**
- **Blogs and comments**
- **Personal documents**
- **Pictures: Instagram, Flickr, Picasa etc.**
- **Videos: You-tube etc.**
- **Internet searches**
- **Mobile data content: text messages**
- **User-generated maps**
- **E-Mail**

Internet of Things (Machine-generated data)

Sensors and machines used to measure and record events and situations in the physical world

- o Fixed sensors
  - ✓ **Home automation; Weather/ pollution sensors; Traffic sensors / webcam; Scientific sensors; Security / Surveillance cameras**
- o Mobile sensors
  - ✓ **Mobile Phone Location; Cars; Satellite images**
- o Data from computer systems
  - ✓ **Logs / Web logs**

CONTOH BIG DATA UNTUK STATISTIK RESMI

CONTOH PEMANFAATAN BIG DATA UNTUK STATISTIK RESMI

Media Kernels Indonesia

- ✓ **Canada –** Smart meter data

- ✓ **Netherlands –** Road sensor and Mobile Phone Data

- ✓ **USA –** Web-scraping and partnerships

- ✓ **Ireland –** Mobile Phone data

- ✓ **Mexico –** Twitter

CONTOH 1:
TWITTER FOR TOURISM MONITORING IN MEXICO

In 2014, a working group on Big Data at INEGI conducted a pilot study to track domestic tourism from Twitter data, in order to contribute to the empirical modelling of individual tourist behavior. The objective of this pilot program was to identify the characteristics of an average Tweeting tourist in order to identify how many people travelled to Puebla and Guanajuato during the holiday weekend of February 1-3, 2014. The team of researchers from INEGI, in collaboration with the Mexican Ministry of Tourism, analysed 60 million Tweets from January to July 2014, from the continuous 1% georeferenced sample that Twitter makes available for free.[65] From this data, INEGI collected Tweets from the 7,955 Twitter users who Tweeted in Guanajuato (48%) and Puebla (52%) during the holiday. They then gathered all the Tweets sent by those users in the remainder of the target period (amounting to 827,424 total Tweets), and identified which users Tweeted from another state (presumably their homestate) after being in Guanajuato or Puebla, in order to map the origin of domestic tourism to those two areas during the holiday.[66] The resulting estimates of domestic tourism to Guanajuato and Puebla were compared to estimates made by the respective offices of tourism of those two states.[67]

- GOAL: Menghitung Jumlah Turis
- Menggunakan 1% sampel twit (georeferenced)
- 60 juta twit (Januari-Juli)
- Ambil sampel 7.955 user
- Lihat lokasi twit mereka
- Lihat pergerakan turis domestik

CONTOH 2:
USING GOOGLE TRENDS TO NOWCAST ECONOMIC DATA IN COLOMBIA

The economic indicators used by Colombia's Administrative Department for National Statistics (DANE) to analyze economic activity at the sectorial level have an average time lag of 10 weeks. In order to obtain more real-time estimates of economic activity, the Colombian Ministry of Finance is looking into ways to real-time forecast (i.e. nowcast) activity based on data from Google web searches.

Researchers at the Ministry of Finance analyzed the relative frequency of particular search terms, using *Google Trends*, a tool maintained by Google that keeps track of prevalent search terms over time and provides an index of how common the queries are in each part of the world over a given time period. Based on a methodology for short-term forecasting of economic series developed by Choi and Varian (2011), the researchers used *Google Trends* data to infer economic activity at the sectorial level by choosing certain keywords to act as proxies for consumer behavior,[81] thus providing alternative indicators to traditional statistics in a much timelier manner. These indicators were produced for certain sectors of the economy, such as agriculture, industry, commerce, construction, and transports; other economic sectors such as mining or financial services cannot be assessed with this method.[82]

The resulting sectorial indicators, known by their Spanish acronym ISAAC, were validated against DANE's official indicators of economic activity, and both sets of indicators were made publically available. The ISAAC data, which pertains to the sectorial level, is aggregated to produce a single leading indicator for economic activity, known as ISAAC+. The project team, led by Luis Fernando Mejía, continues to publish the ISAAC and ISAAC+ on a monthly basis.

A major limitation of such web-based measures is that they risk not being representative in countries where internet penetration is low, as is the case in Colombia (~60%). However, as internet penetration continues to grow, the caveat of non-representativeness becomes less of an issue.

Thus, Colombia's exploration of timelier economic indicators shows promise and has attracted the attention of other countries interested in implementing their own Big Data-based forecasts of economic indicators.[83]

- Statistik resmi butuh 10 minggu
- Google Trend digunakan untuk now cast (realtime)
- Certain keywords to act as proxies for consumer behavior

CONTOH 3:
ANALYSIS OF SOCIAL MEDIA MESSAGES BY STATISTICS NETHERLANDS

The Statistics Netherlands studied publically available social media messages created on various social media platforms, such as Twitter and Facebook, as well as the public messages posted on news sites, web forums and blogs. The messages were obtained from a commercial company that routinely harvested all publically available messages written in Dutch on the Dutch-language part of the web.

Both the content and the sentiment of the messages were studied. Studies of the content of messages in Dutch on Twitter, the social media platform on which most publically available Dutch-language messages are created, revealed that nearly 50 per cent of those messages were composed of "pointless babble". The remainder predominantly discussed spare-time activities (10 per cent), work (7 per cent), media (television and radio) (5 per cent) and politics (3 per cent). Use of these more serious messages was hampered by the less serious "babble" messages.
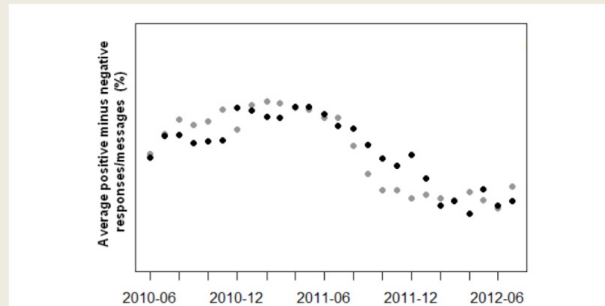

Figure 4. Dutch consumer confidence (grey) and the overall sentiment in Dutch social media messages on a monthly basis (black). Dutch articles are used as search terms. The social media sentiment in December is considerably more positive compared to the sentiment in the months before and after.

Determination of the sentiment in all messages created on all available platforms revealed a highly interesting potential use of these data for statistics. With a query language and a web interface, messages were selected from the database. Messages were classified in positive, negative and neutral. The sentiment in these messages was found to be highly correlated with Netherlands consumer confidence, in particular with sentiments regarding the economic situation. A consumer confidence index is produced every month by Statistics Netherlands using survey data from a random sample from the population register. While Social Media messages are generated by 70 percent of the Dutch population.[181] The latter relation was stable on a monthly and on a weekly basis. Daily figures, however, displayed highly volatile behaviour suggesting that it is possible to produce monthly and weekly sentiment indicators comparable with consumer confidence. The latter indicators can be produced on the first working day following the week studied, demonstrating the ability to deliver results quickly. Only in December the numbers did not relate, where a much more positive sentiment occurred in social medial, removing all messages that including words for Christmas and New Year's eve reduced these peaks.[182]

- Goal: consumer confidence index
- Lewat survey hasilnya bulanan
- Analisis Twitter, difilter hanya yang relevan
- Mengukur sentimen (positif – negatif)
- Hasilnya bisa mingguan (atau realtime)
- Tantangan: keywording

30

CONTOH MOBILE DATA UNTUK TRACKING PERGERAKAN MASYARAKAT

GOOGLE
COVID-19 COMMUNITY MOBILITY REPORTS

https://www.google.com/covid19/mobility/

APPLE
COVID-19 MOBILITY TRENDS REPORTS

Mobility Trends

Change in routing requests since January 13, 2020

Search (for example Italy, California, or New York City)
Jakarta, Indonesia

- Walking +10%
- Driving +4%

https://www.apple.com/covid19/mobility

33
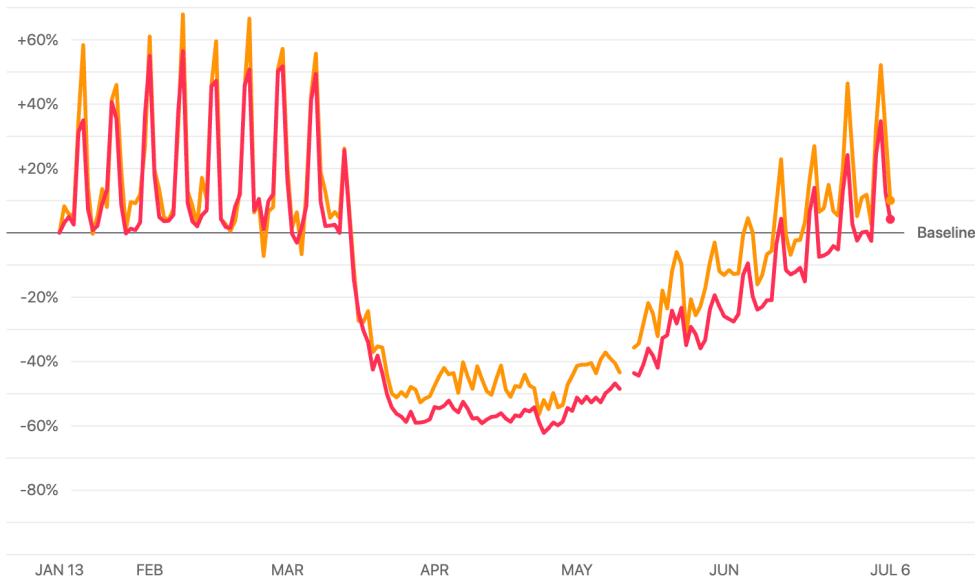
DKI JAKARTA

Dashboard Persiapan Pergerakan Mudik
DKI Jakarta
(Transportasi)

jakarta smart city

Angka Mobilitas DKI Jakarta (15 Februari 2020 - 11 April 2020)   Sumber : Google LLC "Google COVID-19 Community Mobility Reports."

Transportasi Publik   Tempat Kerja
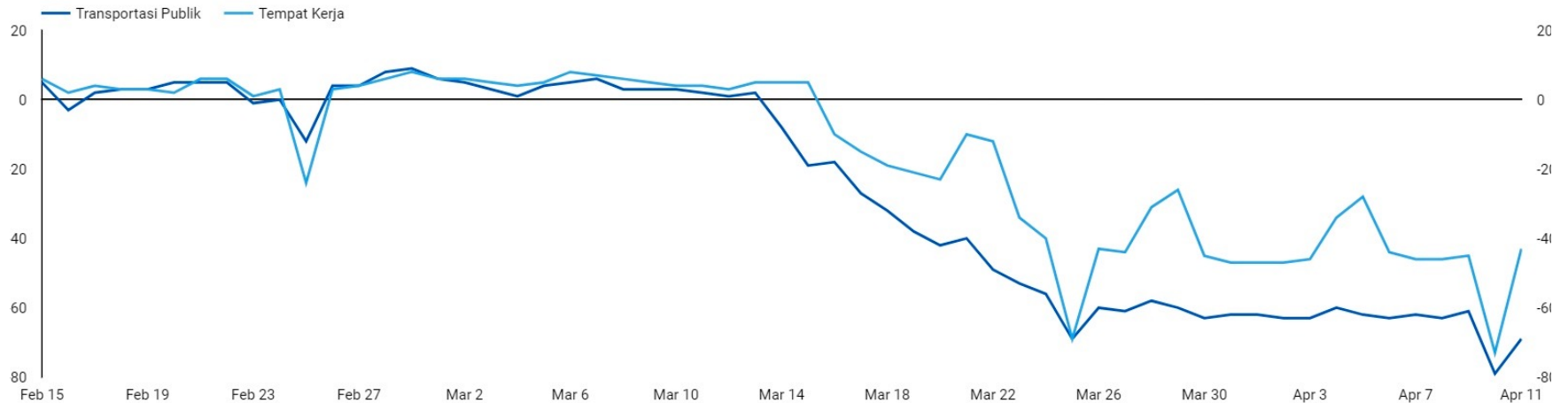
CONTOH SOCIAL MEDIA ANALYTICS UNTUK LAYANAN PUBLIK

PIKOBAR
PUSAT INFORMASI DAN KOORDINASI COVID-19 JAWA BARAT

Media Kernels Indonesia

👍 **Liked** ▾   🔊 **Following** ▾   ➤ **Share**   •••

Ridwan Kamil ✔
March 5 · 🌐

Pusat Informasi dan Koordinasi COVID-19 Jawa Barat (PIKOBAR) resmi beroperasi, menyediakan semua informasi faktual dan aktual terkait penanganan dan pencegahan Covid-19 berada di Command Center Gedung B, Gedung Sate, Kota Bandung.

Nomor Hotline COVID-19 Dinas Kesehatan Provinsi Jabar: 0811-2093-306, dan Emergency Kesehatan: 119.
Atau kunjungi website: www.pikobar.jabarprov.go.id

Tetap tenang namun waspada. Terapkan pola hidup sehat. Virus Corona / Covid-19 tidak ditularkan melalui udara, tapi melalui kontak langsung dan melalui cairan tubuh. Cuci tangan setiap selesai melakukan kegiatan, berdasar himbauan WHO, cara ini menjadi yang paling efektif untuk mencegah penularan.

Bersama jajaran kepolisian, pemerintah juga mengawasi penjualan dan penimbunan masker oleh yang tidak bertanggung jawab. Sebab untuk kondisi saat ini, masker dibutuhkan bagi para pekerja medis dan mereka yang sakit untuk melindungi diri.

#JabarTanggapCovid19

admin

Ridwan Kamil ✔
@mochamadridwankamil

Home
About
Videos
Posts
Events
Notes
Photos

👍 5.5K   356 Comments 720 Shares

SOCIAL MEDIA ANALYTICS DI PIKOBAR

Media Kernels Indonesia

Drone Emprit Customized

Drone Emprit

SOCIAL MEDIA ANALYTICS DI PIKOBAR

Media Kernels
Indonesia

Drone Emprit Customized

SOCIAL MEDIA ANALYTICS DI PIKOBAR

Drone Emprit Customized

Drone Emprit

DRONE EMPRIT CUSTOMIZED

Media Kernels Indonesia

Sentiment

Sentiment	Mentions
Negative	8,725
Positive	28,833
Neutral	12,279

Locations
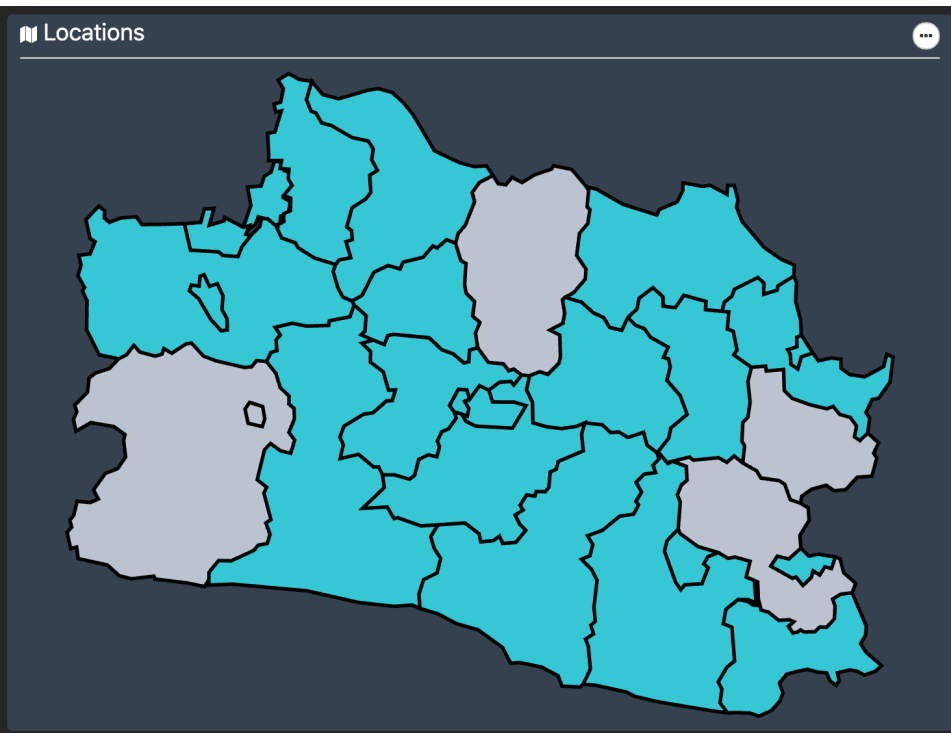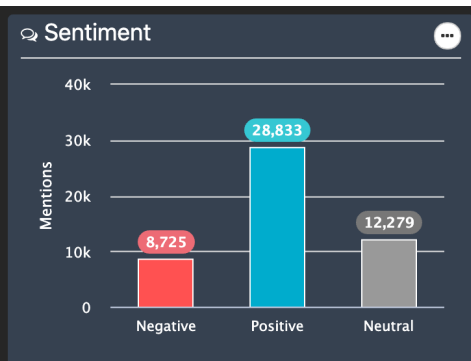
Top Problem By Area

	Area	
1	Kota Bandung	Penyakit 512
2	Kota Bogor	Penyakit 246
3	Kota Depok	Penyakit 128
4	Kota Bekasi	Penyakit 121
5	Kota Sukabumi	Penyakit 86
6	Kabupaten Bekasi	Penyakit 54
7	Kota Cirebon	Penyakit 54

Top Account

	Account	
1	Antara_Jabar	622
2	fokusjabar	284
3	120493494629294	153
4	polsekcibatuPWK	108
5	hansedbert97	76
6	Ksamanhudi	68
7	Blood4LifeID	66

Top Problems

	Problem	
1	Penyakit	2,663
2	Macet	202
3	Gempa Bumi	94
4	UMP	82
5	Biaya Sekolah	71
6	Sampah	59
7	Banjir	51

Top Active

	Account	
1	disperindag_jbr	32
2	humasjabar	23
3	humas_jabar	18
4	disperindag_jbr	13
5	8427554635	12
6	Humas Jabar	10

Top Issue (Most Retweeted)

	Issue	
1	Kemensos menyalurkan 1.000 paket sembako kepada Lansia yg perekonomian maupun kesehatannya terdampak Covid-19, di Kota Cimahi, Jawa Barat. Terima kasih @juliaribatubara @KemensosRI Hadir..!!! https://t.co/hqHm5tYtIZ ⏱ 2020-07-02 18:01 @__RismaWidiono_	1,959
2	Kemensos menyalurkan 1.000 paket sembako kepada Lansia yg perekonomian maupun kesehatannya terdampak Covid-19, di Kota Cimahi, Jawa Barat. Terima kasih @juliaribatubara @KemensosRI Hadir..!!! https://t.co/hqHm5tYtIZ ⏱ 2020-07-02 18:01 @__RismaWidiono_	1,105
3	Kemensos menyalurkan 1.000 paket sembako kepada Lansia yg perekonomian maupun	777

Top Programs

	Program	
1	JABAR MASAGI	301
2	GURILAPS	182
3	MAGRIB MENGAJI	153
4	SADESHA	92
5	OPOP	72
6	DESA DIGITAL	67

CONTOH PEMANFAATAN BIG DATA
SOCIAL MEDIA ANALYTICS UNTUK KOMUNIKASI PUBLIK JABAR

Sistem big data social media digunakan untuk melakukan analisis, reporting, penyusunan strategi komunikasi, dan evaluasi, dalam bentuk:

1. Communication Brief: Goal, point penting, dan draft konten untuk dimobilisasikan oleh Perangkat Daerah, Dinas, dan Relawan.

2. KPI: Evaluasi hasil mobilisasi.
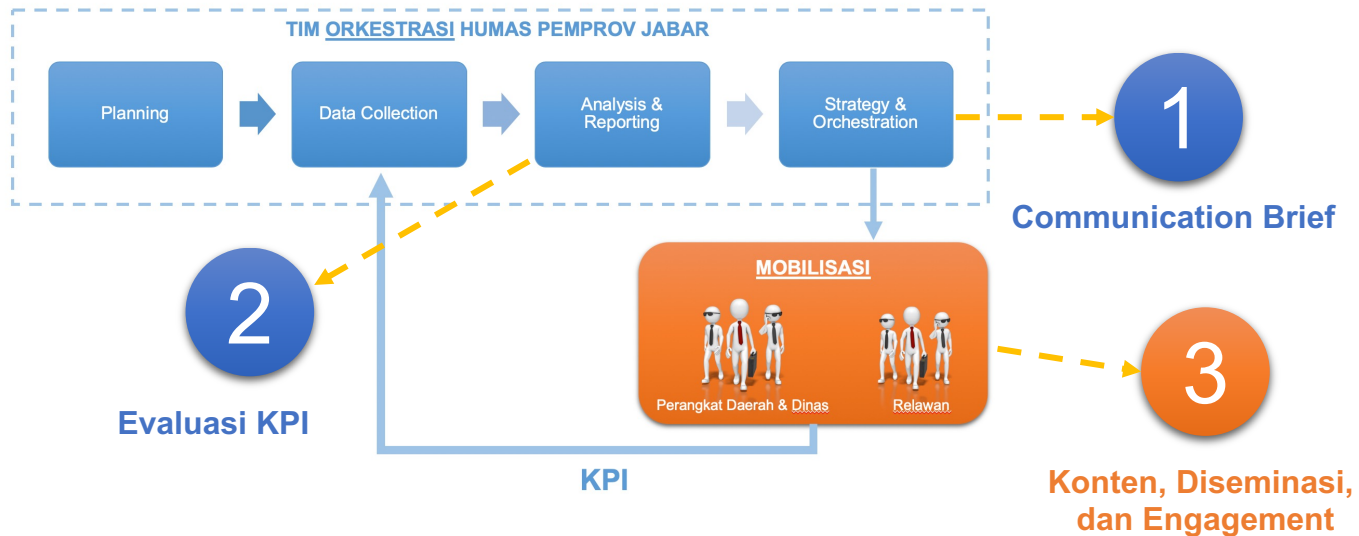
TANTANGAN DAN SOLUSI

OBTAINING EXTERNAL BIG DATA SOURCE IS A STRUGGLE

- Most companies struggle with **data from outside** the enterprise: from suppliers and from the field as products are in use.

- Most companies know they must improve product quality, customer satisfaction, and new product introduction time.

- However; for more than half of the respondents, **data on genealogy**, root causes, and supplier engineering changes takes weeks to obtain or is not available at all.
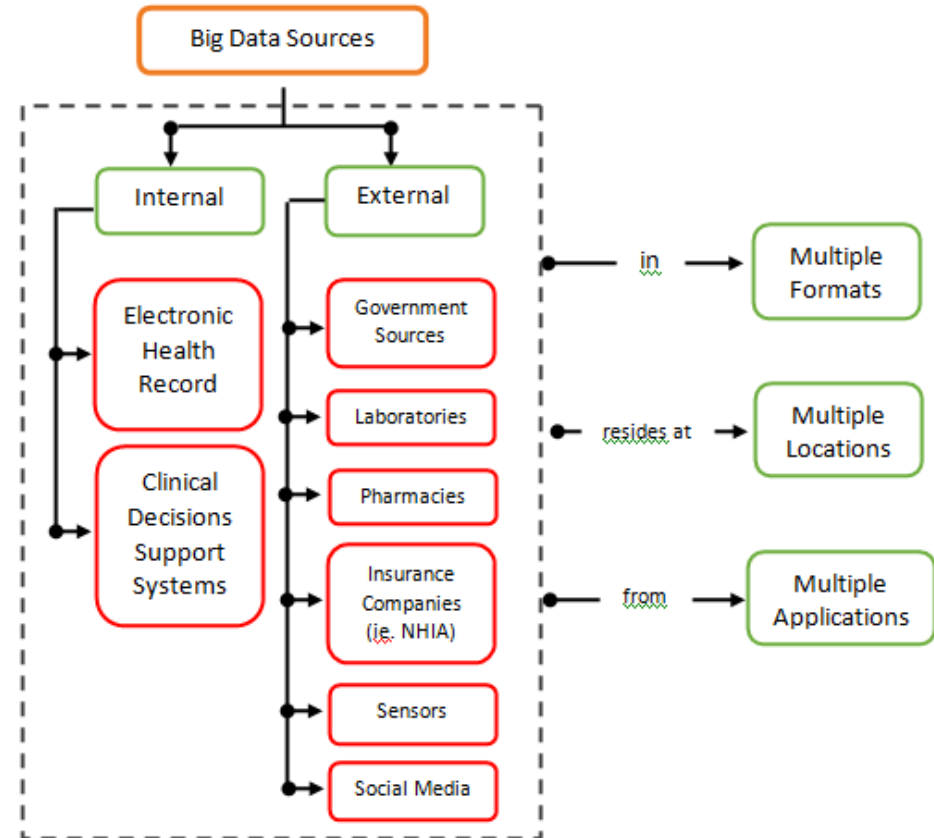
INSTITUTIONS ARE STRUGGLING TO ACCESS USEFUL DATA

- Once the data is obtained, the further step of cleansing that data can prove even more difficult. Fifty-seven percent of manufacturers don't have access to reliable product genealogy information, and 72% believe their product data is less than 80% accurate.

- The report notes that IT infrastructure makes it difficult to turn current data into usable information.

The Process of Data Cleansing
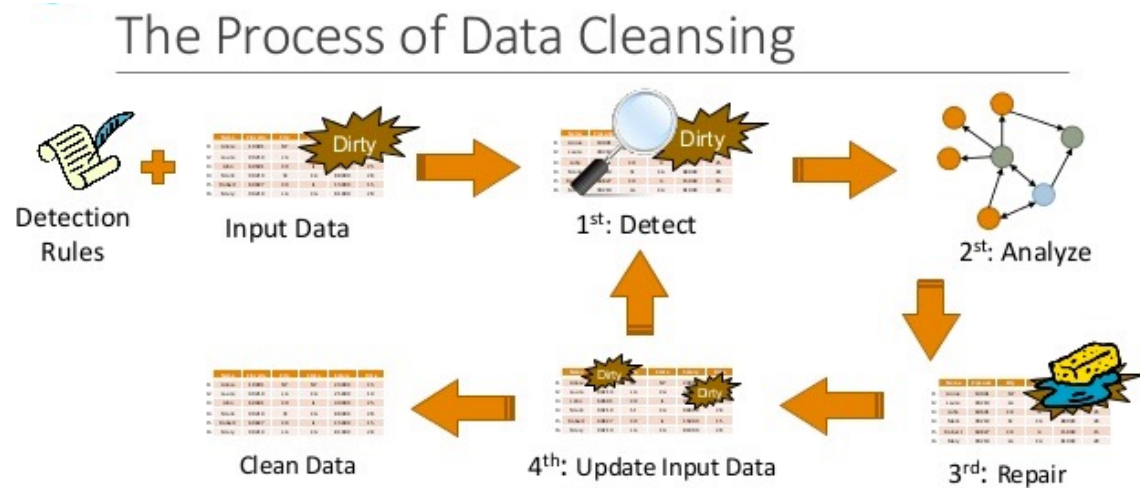
DIFFICULT TO GET DATA FROM DIFFERENT DEPARTMENTS WITHIN THE ORGANIZATION

What many companies struggle with is aggregating and correlating data from different operational areas or applications within the organization, what is often called the silo effect.

Data Silos – All Storage No Action

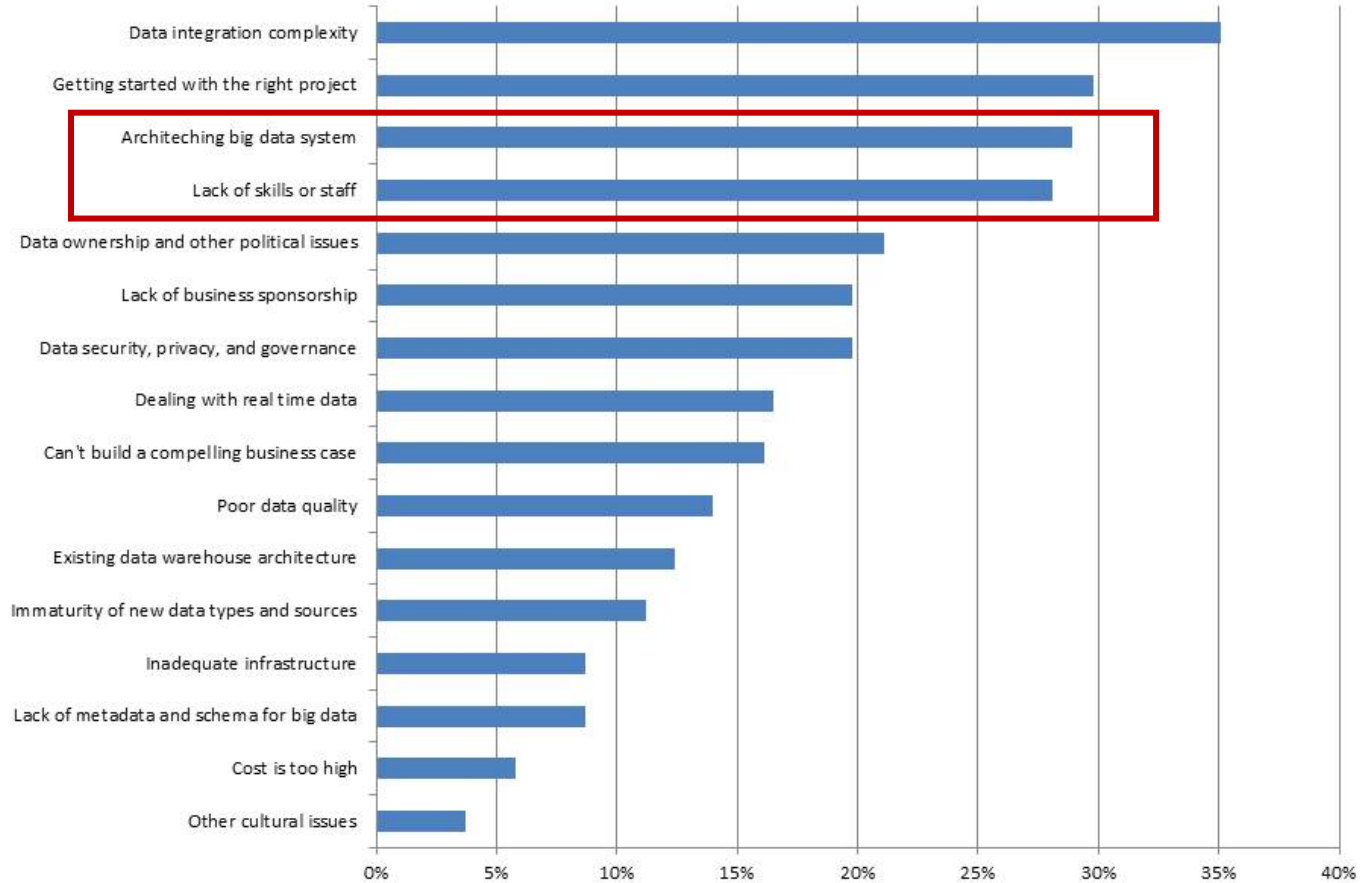➡ Large Enterprises that have inherited a legacy of data silos through mergers and acquisitions have no single source of truth at any one time on their customers.

➡ 20+ CRM systems is not uncommon in Large Enterprise Environments

Sources: Forrester, IDG, Squirro

BIG DATA CHALLENGES

(source: TDWI Predictive Analytics Study, 2013)

BIG DATA SCIENCE SKILL SET

MATH & STATISTICS

- Machine Learning
- Statistical Modeling
- Exploratory Analysis
- Clustering
- Regression Analysis

PROGRAMMING & DATABASE

- Computer Science Fundamentals
- Database Management System
- Data Visualization
- Python
- Big Data

DOMAIN KNOWLEDGE & SOFT SKILLS

- Inclination towards business operations
- Keen on working with data
- Problem solver
- Strategic, proactive, and cooperative
- Interested in hacking

COMMUNICATION & VISUALIZATION

- Storytelling skills
- Convert data-based insights into decisions
- Collaborative with Sr. Management
- Knowledge of tools like Tableau
- Visual art design

BIG DATA SKILL SHORTAGES

Big data skills shortages – and how to work around them

Businesses are finding it hard to recruit enough people with big data and analytics skills, but for CIOs there are alternative strategies to consider

By **Shamus Rae,** KPMG   Published: **06 Jun 2018**

For the fourth year running, the *Harvey Nash/KPMG CIO Survey* has found that big data and analytics are top of the skills shortage critical list. This is having a significant impact on all organisations, with two-thirds of IT leaders saying it is preventing them from keeping up with the pace of change.

Given the newness of the discipline, it is not surprising that there aren't enough skilled data scientists. Very few universities offer pure data science degrees (as opposed to computer science). Many schools still don't even offer computer science at A-level or GCSE. It's going to take some years before there are enough skilled data scientists in the workforce.
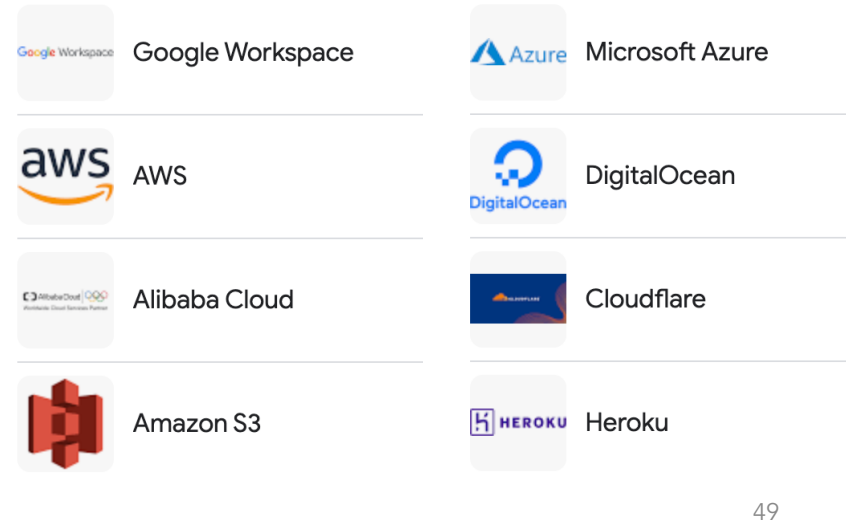
48

SOLUTION: RENT

One of the key questions CIOs often ask themselves is: do we buy or build? In fact, with data analytics, you can rent. Because data analytics, big data and AI are new areas, there is a natural tendency to think that the organisation has to prove itself and build its own system.

But companies such as Amazon Web Services (AWS), Google, Microsoft, IBM and others are building applications for voice recognition, image classification, facial recognition and more that are available for rent in the cloud with no long-term commitment.
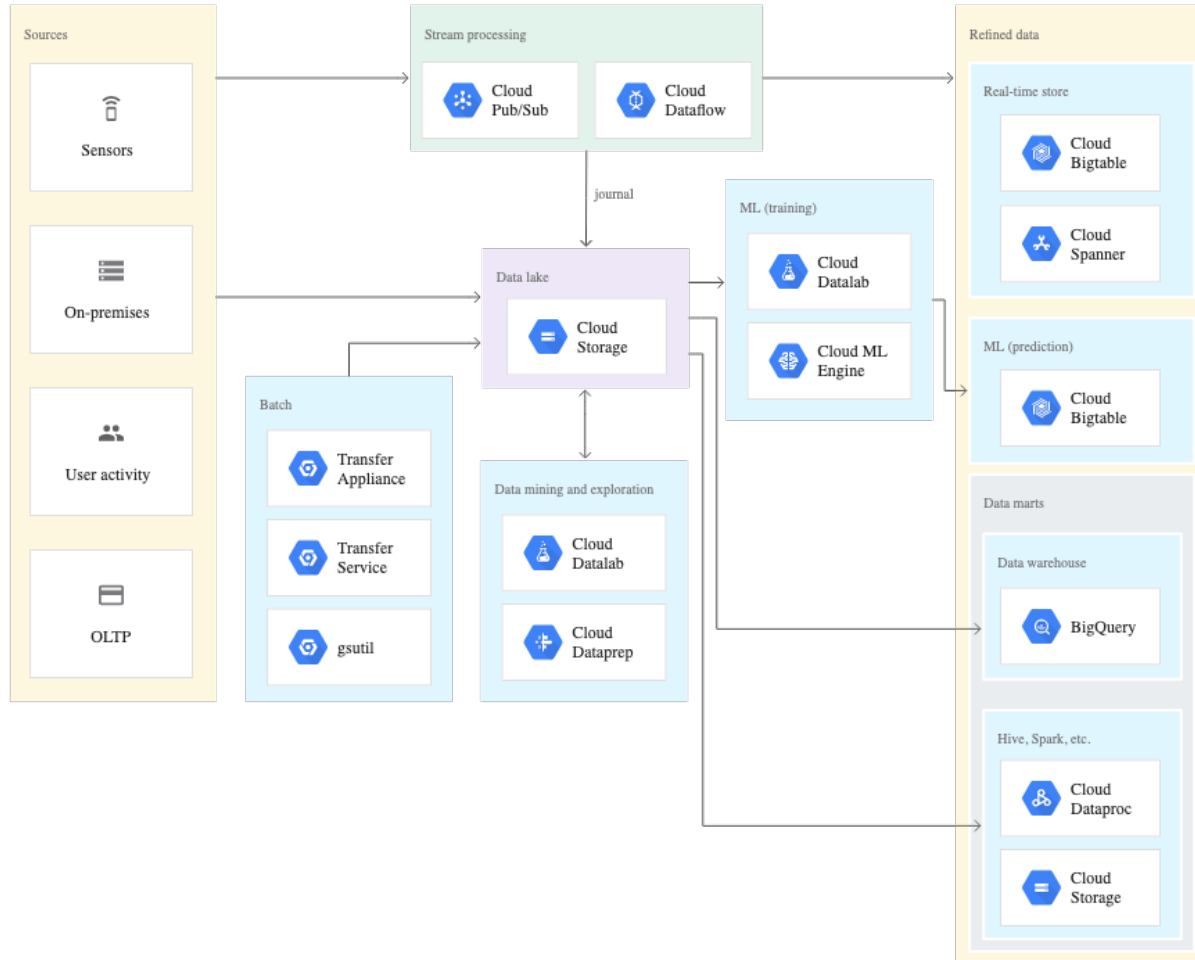
One example is Google's AutoML. Google's systems automatically create a machine learning model from content uploaded by the client. Rather than in-house teams spending a lot of time fine-tuning algorithms in search of the best solution, this – and similar services being developed by other organisations – could save time and ease the pressure on resources.

Google Workspace	Google Workspace	Azure	Microsoft Azure
aws	AWS	DigitalOcean	DigitalOcean
Alibaba Cloud	Alibaba Cloud		Cloudflare
	Amazon S3	HEROKU	Heroku

49

CONTOH PLATFORM

Cost Effective

Serverless

Highly Scalable

Cloud AI

Custom Machine Types

Big Data Analytics

Internet of Things

API Platform and Ecosystem

Google Cloud Platform

CONTOH GOOGLE CLOUD PLATFORM

PLATFORM BIG DATA ANALYTICS DAN AI UNTUK PEMERINTAHAN

SHARED PLATFORM

SHARED PLATFORM

- Cost Effective
- Highly Scalable
- Custom Machine Types
- Internet of Things
- Serverless
- Cloud AI
- Big Data Analytics
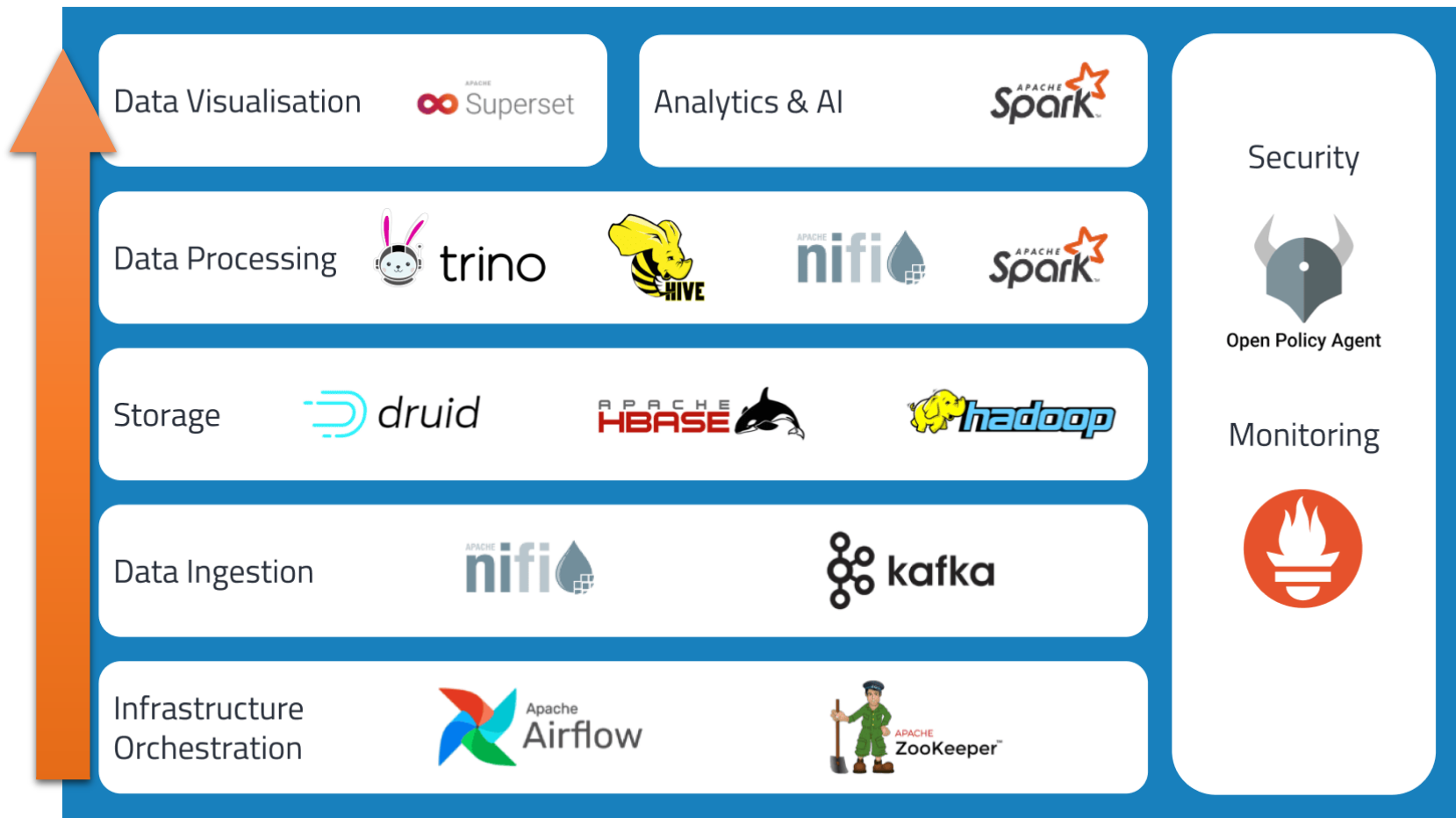- API Platform and Ecosystem

Instansi Pusat Pemerintah Daerah

Instansi Pusat Pemerintah Daerah

Instansi Pusat Pemerintah Daerah

OPEN SOURCE SOLUTIONS

KUNCI KEBERHASILAN

- Keterlibatan Pimpinan
- Pendidikan dan Pelatihan
- Infrastruktur yang Kuat
- Kerjasama Lintas-sektor
- Regulasi dan Kebijakan yang Jelas
- Kualitas Data
- Inovasi Berkelanjutan

KESIMPULAN

- Big Data Analytics adalah kebutuhan penting untuk Instansi Pusat Pemerintah Daerah (IPPD).

- Efisiensi dan efektivitas pelayanan dapat ditingkatkan dengan analisis data.

- Konsep Democracy 4.0 menyoroti bagaimana teknologi mempengaruhi demokrasi.

- Tantangan utama: mendapatkan sumber data eksternal, mengakses data yang berguna, dan mengatasi silo data antar departemen.

- Solusi: menggunakan platform berbagi dan open-source untuk big data analytics.

- Kunci keberhasilan: keterlibatan pimpinan, pendidikan, infrastruktur yang kuat, kerjasama lintas-sektor, regulasi yang jelas, kualitas data, dan inovasi berkelanjutan.

THANK YOU