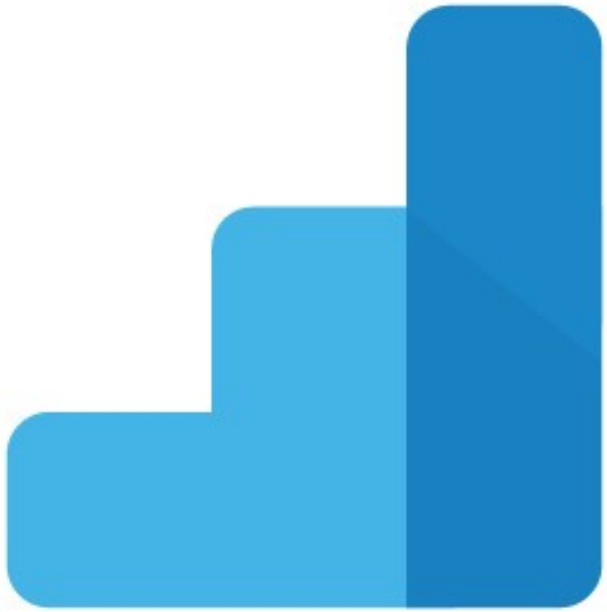
2017 IIA INDONESIA NATIONAL CONFERENCE
M.O.V.E. FORWARD
Mastery, Open-Minded, Value, and Ethics
Big Data becomes Smart Data
Komang B. Aryasa
Head of Big Data, IoT & Research
PT. Telekomunikasi Indonesia, Tbk
komang@telkom.co.id
The Institute of Internal Auditors Indonesia
Telkom Indonesia
IIA INDONESIA
2017 NATIONAL CONFERENCE
INTERCONTINENTAL DAGO PAKAR BANDUNG ● 10-11 OCTOBER 2017

Content

Re-calling Big Data

Opportunities & Challenges

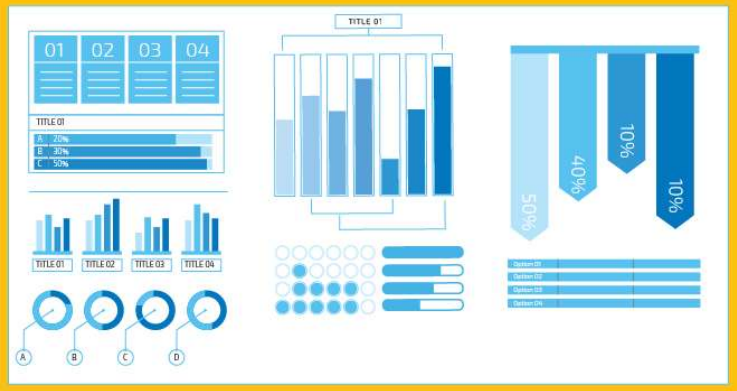
Big Data Implementation in Telkom



Re-Calling Big Data

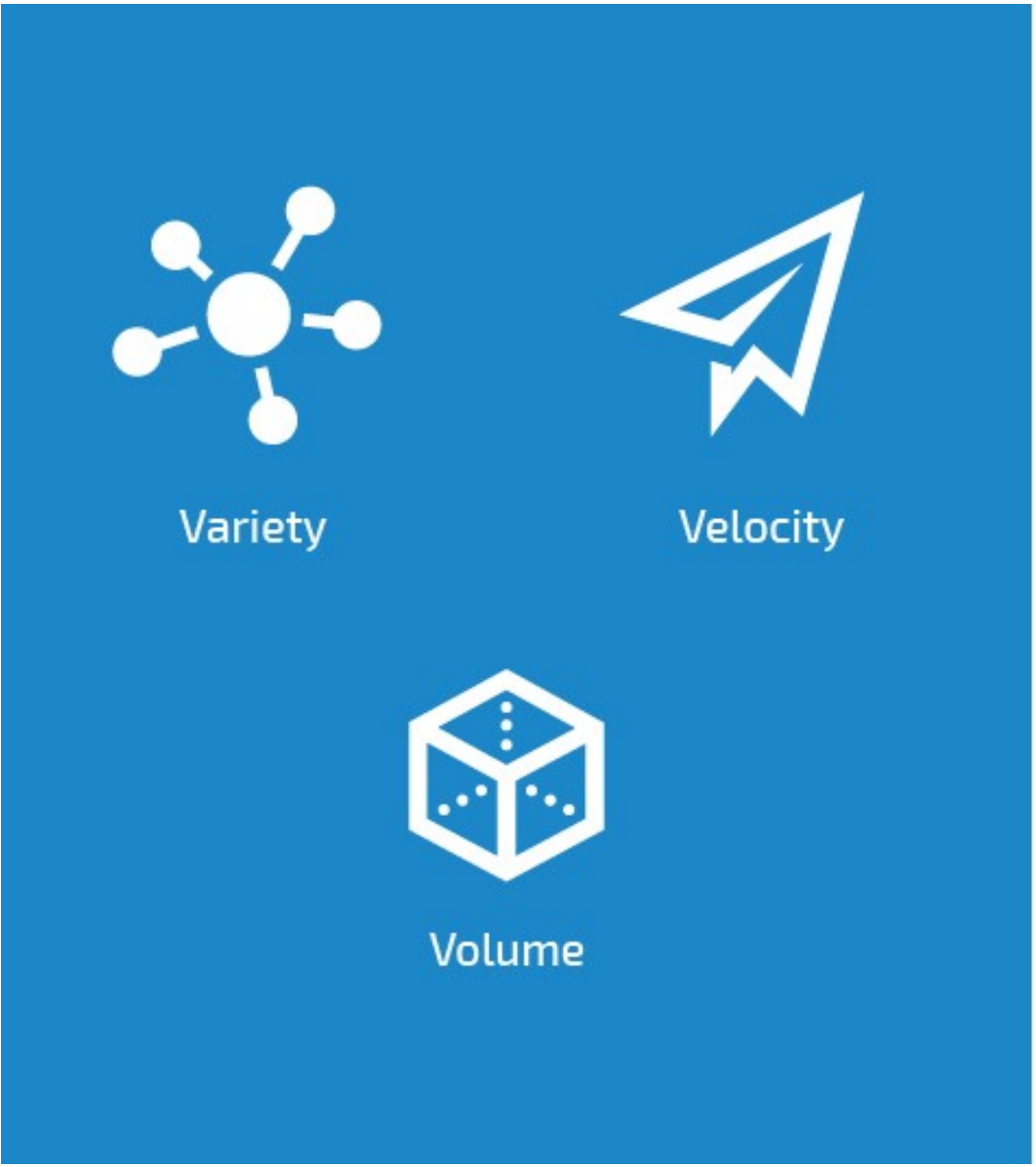
SMALL DATA

ID	CUST_NAME	EMAIL	PHONE	ADDRESS
00010990	Andi	andi@email.com	08219392813	Jakarta
08915688	Dini	dini@email.com	0811144455	Jakarta
36758091	Ujang	ujang@email.com	0855685949	Bogor
76800014	Ucok	ucok@email.com	0954732403	Bandung
87398663	Tina	tina@email.com	08219392813	Jakarta
91386554	Siti	Siti@email.com	0811144455	Jakarta
93098443	Rndy	rendy@email.com	0855685949	Bogor
98866422	Dinda	dinda@email.com	0954732403	Bandung



VS

BIG DATA

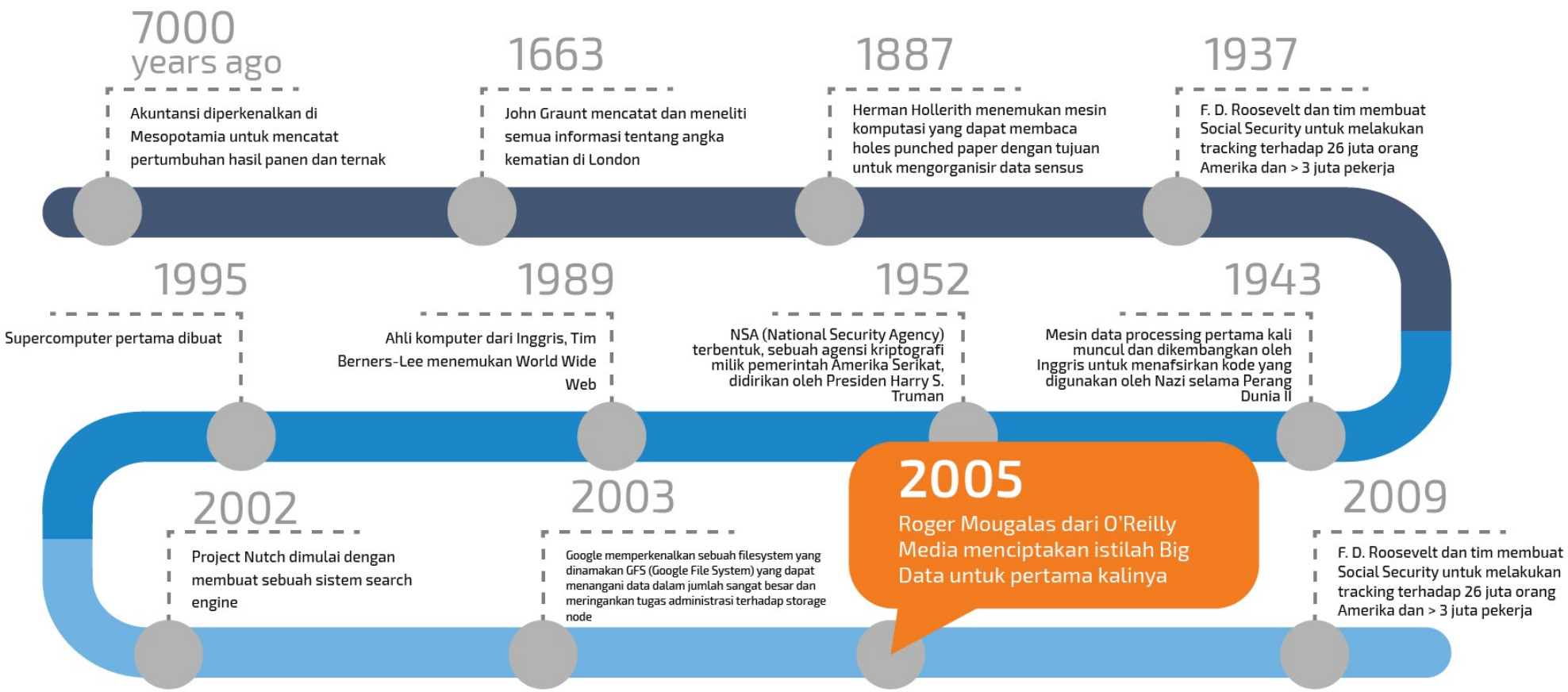


Big Data Era

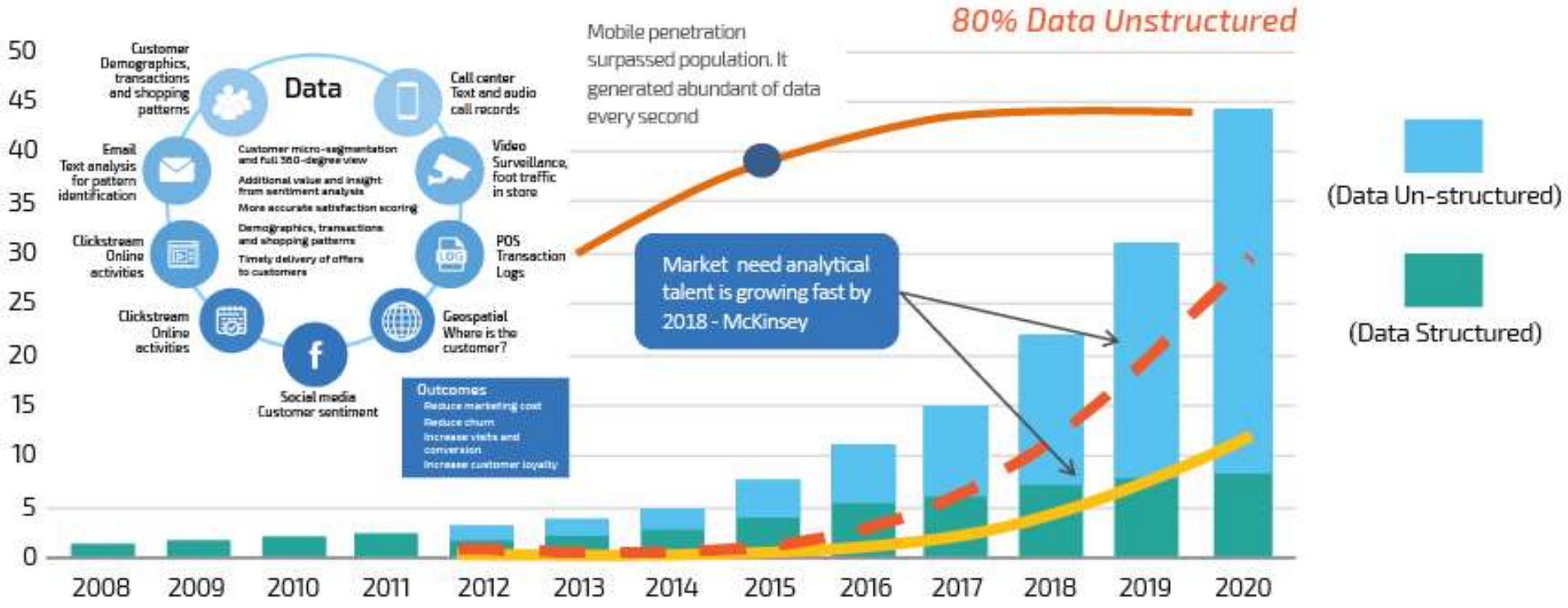
Big Data is extremely large **volume** of data, wide **variety** of data and high **velocity** of data, thus it needs specific technical architecture and analytical method to gain insight that gives new value to the business.

Big data is often associated with 3Vs, whereby the definition of Big Data is not only in terms of large volume of data, but also other indicators. Others 4Vs : Veracity, Visualization, Variability, and Value.

Big Data Journey

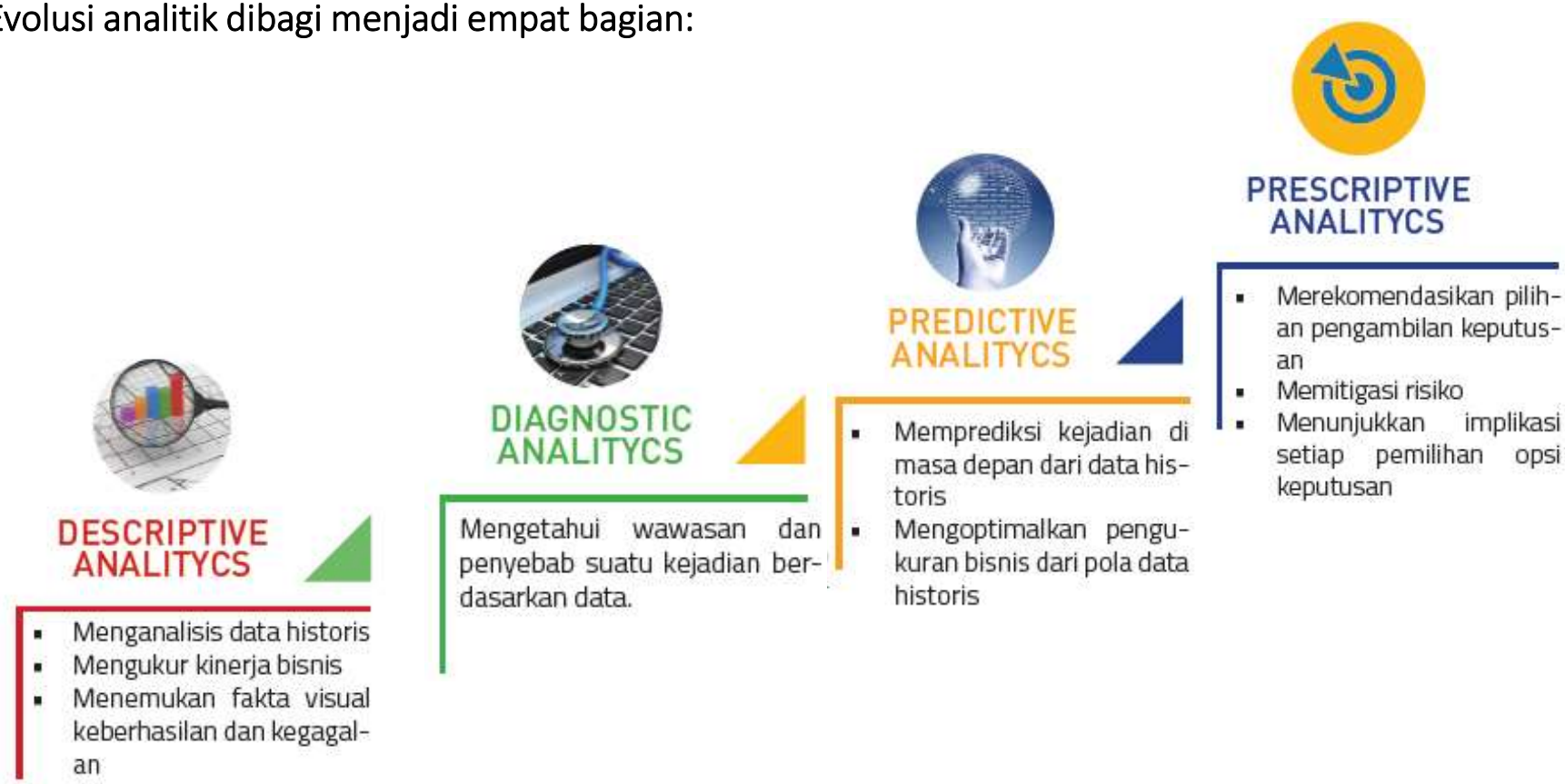


Data Volume : Exponential Growth



Evolusi Big Data Analytics

Evolusi analitik dibagi menjadi empat bagian:



Sumber : Gartner

Manfaat Big Data

Keunggulan Perusahaan Pemakai Big Data Analytics

2x lebih baik dalam kinerja finansial

2x

5x lebih cepat dalam pengambilan keputusan

5x

3x

3x lebih berhasil dalam eksekusi keputusan sesuai keinginan

Sumber : Bain & Company

Apa Kata para Eksekutif ?

45%

45% eksekutif mengatakan: analisis dan pengaplikasian data menjadi tantangan terbesar.

83%

Sebanyak 83% eksekutif mempertimbangkan penggunaan data *real time*.

68%

68% eksekutif pemasaran berencana meningkatkan belanja untuk data.

56%

56% eksekutif akan merekrut tenaga kerja baru untuk posisi-posisi terkait data.

Sumber : Infogroup/Yesmail interactive, 2013

Opportunities & Challenges

Data Driven Companies Have Become the Most Valuable

COMPANY: MARKET CAPITALIZATION

RANK	APRIL 2017	Q4 2011	Q4 2006
1	Apple: 741	Exxon Mobil: 406	Exxon Mobil: 447
2	Alphabet: 585	Apple: 376	General Electric: 384
3	Microsoft: 505	PetroChina: 277	Microsoft: 294
4	Amazon: 432	Royal Dutch Shell: 237	Citigroup: 274
5	Facebook: 408	ICBC: 228	Gazprom: 271
6	Berkshire Hathaway: 404	Microsoft: 218	ICBC: 255

Apple

Source : Click behavior on IOS target
Value : Improve customer experience and product versioning

Alphabet

Source : Android device, search, YouTube, Chrome
Value : Improve customer experience

Microsoft

Source : Search engine, corporate product
Value : Sales Improvement

Recommendation engine

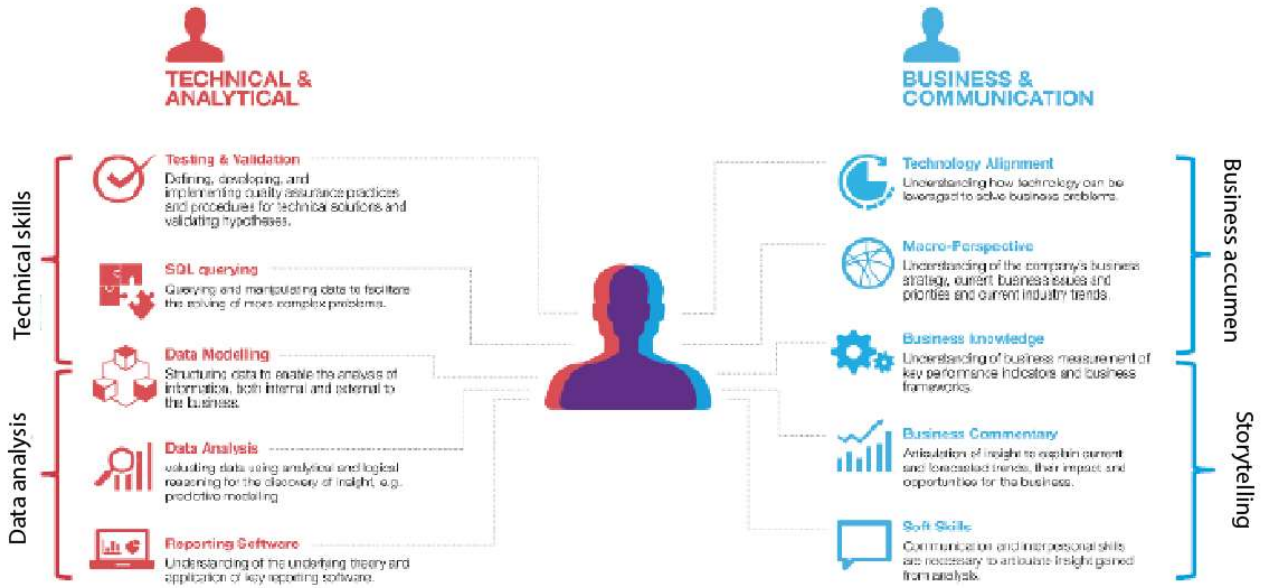
IOS user behavior
Targeted advertising

Shopping behavior, chat behavior,
payment behavior
Targeted sales and recommendation

Challenge #1 : People

In 2012, **Harvard Business Review** named **DATA SCIENTIST** is the "**sexiest job of the 21st century.**" More recently, Glassdoor named it the "**best job of the year**" for 2016. "**It isn't a big surprise**," Dr. Andrew Chamberlain, Glassdoor's chief economist, told Business Insider.

The very ability to derive meaningful insights from data is about acquiring, developing, and retaining a capability that is hard to find



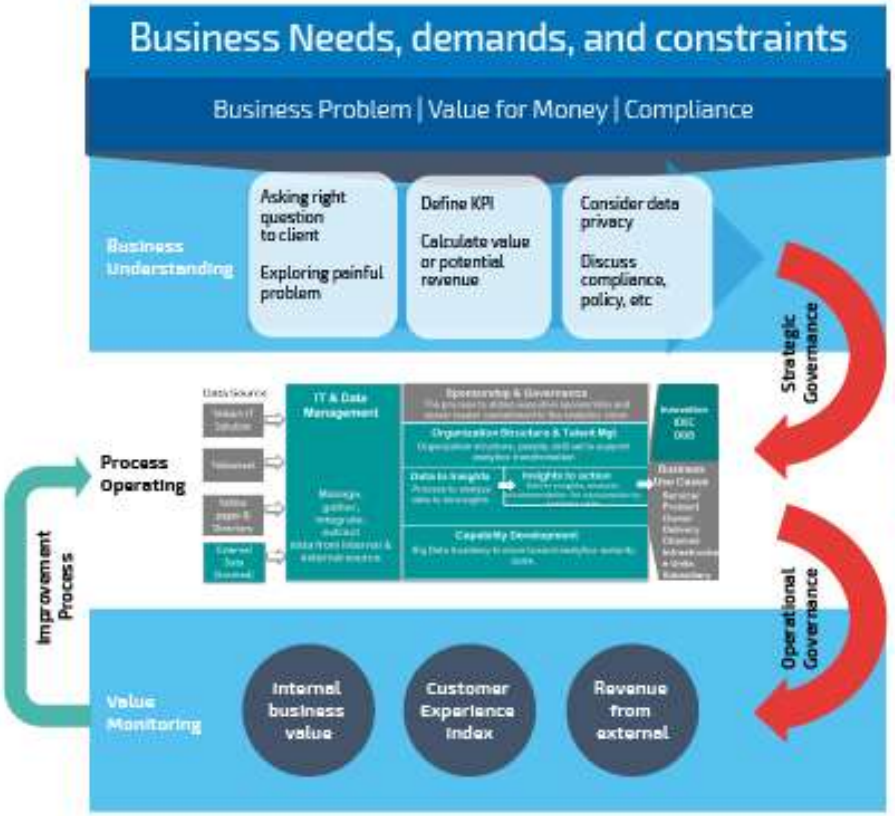
According to Gartner's 2012 CIO Agenda Report, analytics and business intelligence (BI) are the highest tech priority across the globe for CIOs. This has created a talent crunch for a new breed of data scientists who can handle sophisticated data analysis (red skills), but who also have fluent communication skills, business acumen and political nous (blue skills).

Global Situation :
1. Limited number of data scientists
2. Expensive man hour
3. Campus just started to set up Data Science Faculty

Indonesia Situation :
1. Several (very few) number of data scientists
2. Data scientist is treated as programming perspective
3. No university (institution) address data scientist subject

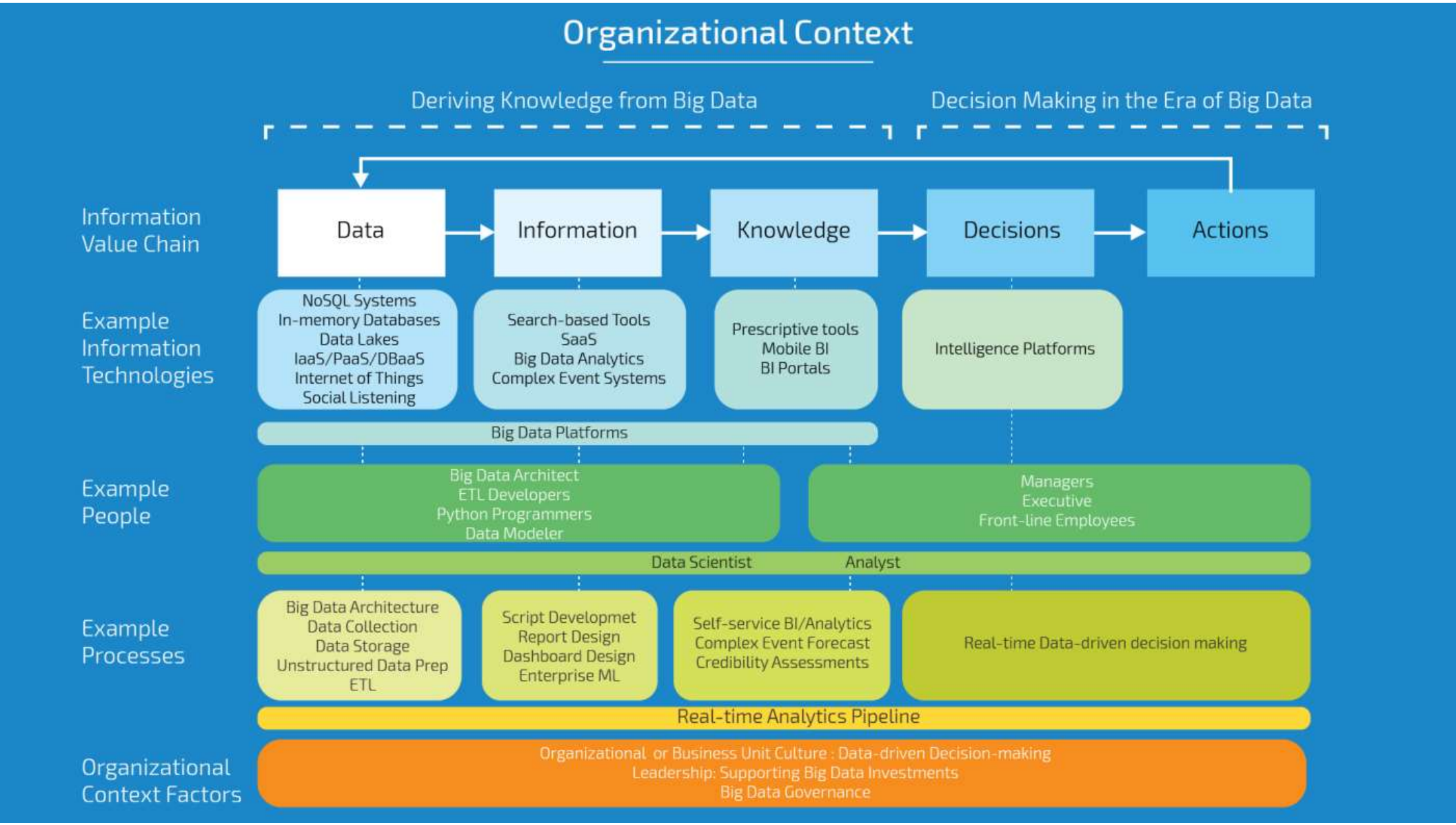
Best Approach :
Data scientist as a team (technical, business, analytics

Challenge #2 : Process



Always Start from the business goal

...End with value evaluation

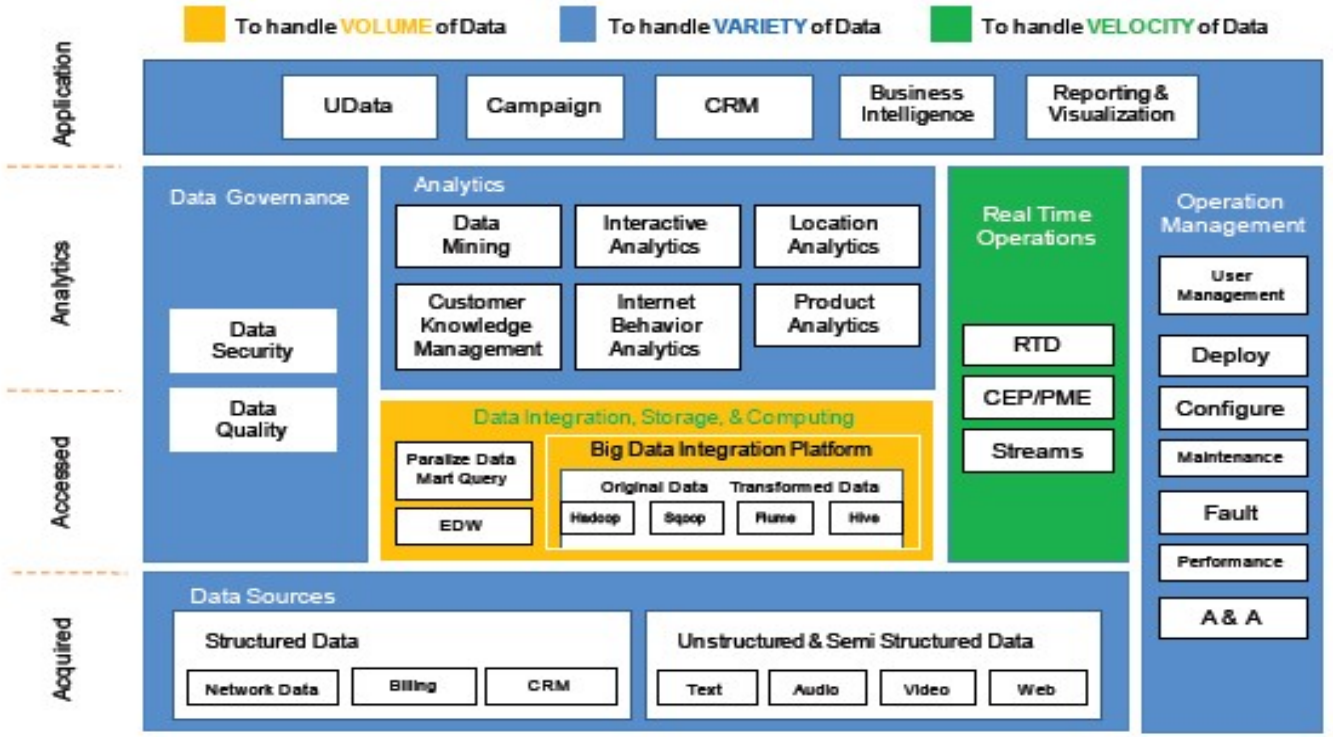

Organizational Context
Deriving Knowledge from Big Data
Decision Making in the Era of Big Data
Information Value Chain
Data
Information
Knowledge
Decisions
Actions
Example Information Technologies
NoSQL Systems
In-memory Databases
Data Lakes
IaaS/PaaS/DBaaS
Internet of Things
Social Listening
Search-based Tools
SaaS
Big Data Analytics
Complex Event Systems
Prescriptive tools
Mobile BI
BI Portals
Intelligence Platforms
Big Data Platforms
Example People
Big Data Architect
ETL Developers
Python Programmers
Data Modeler
Managers
Executive
Front-line Employees
Data Scientist
Analyst
Example Processes
Big Data Architecture
Data Collection
Data Storage
Unstructured Data Prep
ETL
Script Developmet
Report Design
Dashboard Design
Enterprise ML
Self-service BI/Analytics
Complex Event Forecast
Credibility Assessments
Real-time Data-driven decision making
Real-time Analytics Pipeline
Organizational Context Factors
Organizational or Business Unit Culture : Data-driven Decision-making
Leadership: Supporting Big Data Investments
Big Data Governance

Challenge #3 : Technology

Big Data needs support of reliable technology, especially to handle large volume of data, wide variety of data and high velocity of data.

Big Data membutuhkan dukungan teknologi yang handal, terutama dalam penanganan *volume* data yang besar, *variety* yang kompleks dan *velocity* yang cepat.

- **Acquired** berhubungan dengan mendapatkan data, baik data terstruktur maupun tidak terstruktur.
- **Accessed** berhubungan dengan daya akses data, data yang sudah dikumpulkan diperlukan tata kelola, integrasi, storage dan computing agar dapat dikelola untuk tahap berikutnya.
- **Analytic** behubungan dengan insight yang akan didapatkan, hasil pengelolaan data yang telah diproses.
- **Application** merupakan tahapan akhir, dimana hasil dari analytic dilakukan visualisasi dan reporting guna mendukung pengambilan keputusan..



Big Data Implemetation in Telkom

Use Case Big Data Progress

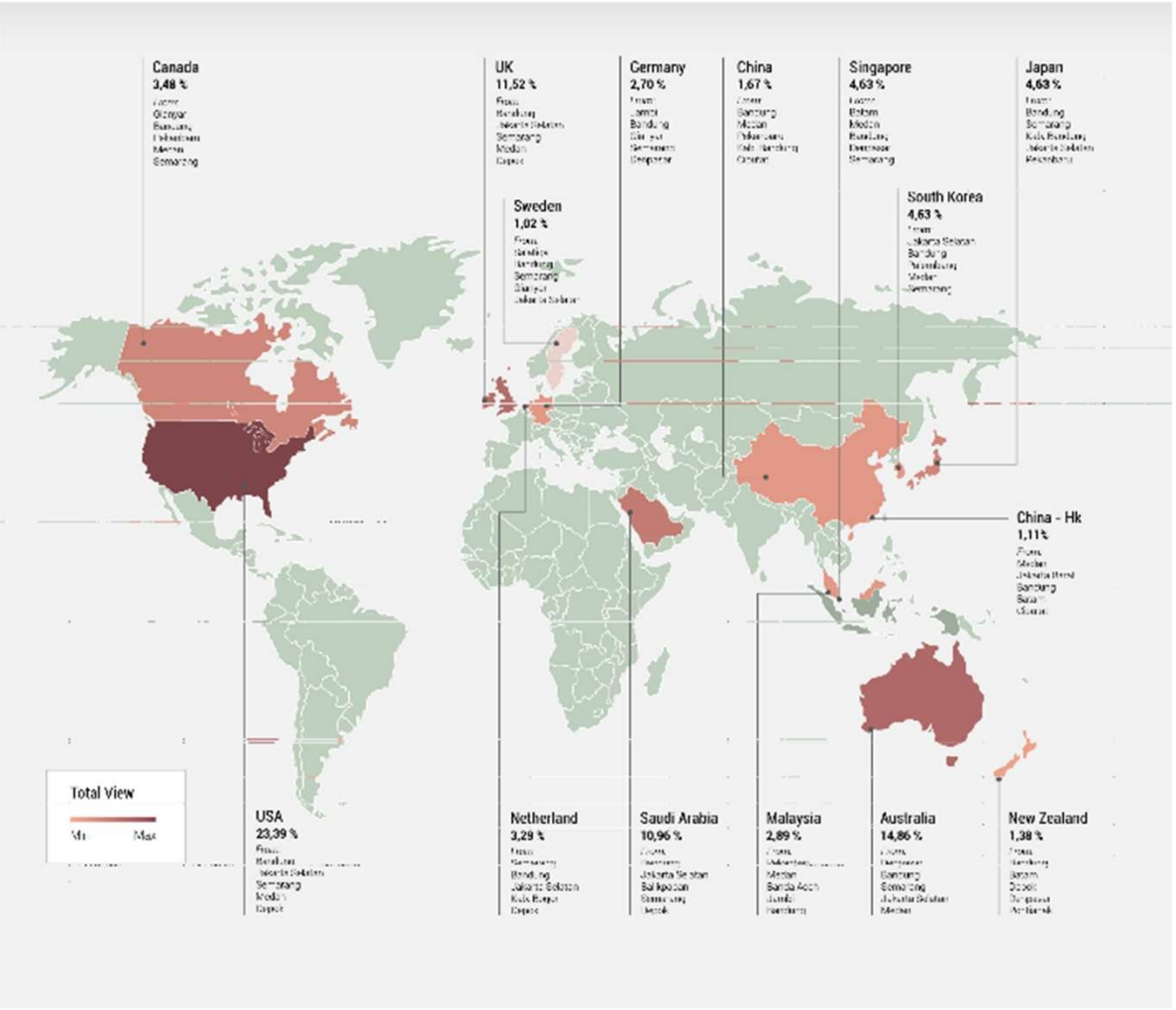
Insight Category	Key Use Cases
Sentiment	General Mood Analysis (Developing) · Review/top trends extraction (Developing) · Top Of Mind (TOM)Analysis (Developing)
Customer Profiling	360 Profile insights (In-Operation) · Micro Segmentation (Data Mgt. Platform) (Not Yet) · Real-time Segmentation (Developing) · Location-based Segmentation (Developing)
Targeting & Advertising	Dynamic offer creation (post call mkt.) (In-Operation) · Pricing analysis (Not Yet) · Basket analysis (Real-time upselling) (Not Yet) · Campaign Optimization (UseeTV Analytics) (In-Operation) · Channel Management (Developing)
Risk & Security	Fraud Detection (Developing) · Default Risk Management (Risk Scoring) (In-Operation) · System threat Management (Developing)
Operation & Performance	Churn Prediction (In-Operation) · Network Performance Monitoring (Developing) · Infrastructure Deployment (Developing) · BI & Reporting (In-Operation) · Data Storage Retrieval (In-Operation)
Industry Specific & Others	Insurance Analytics (Not Yet) · Traffic Vehicle mgmt. (GIS) (In-Operation) · People Pedestrian Mgmt. (Not Yet) · Mediation Platform (Developing) · Weather Solution (Not Yet) · Smart City Mgmt. (Developing) · Photogrammetry (Developing)

In-Operation · Developing · Not Yet

Diagnostic Analytics

Customer Profiling, Customer Segmentation,
Know Your Customer Better

STUDY ABROAD DESTINATIONS FOR INDONESIAN STUDENTS 2017



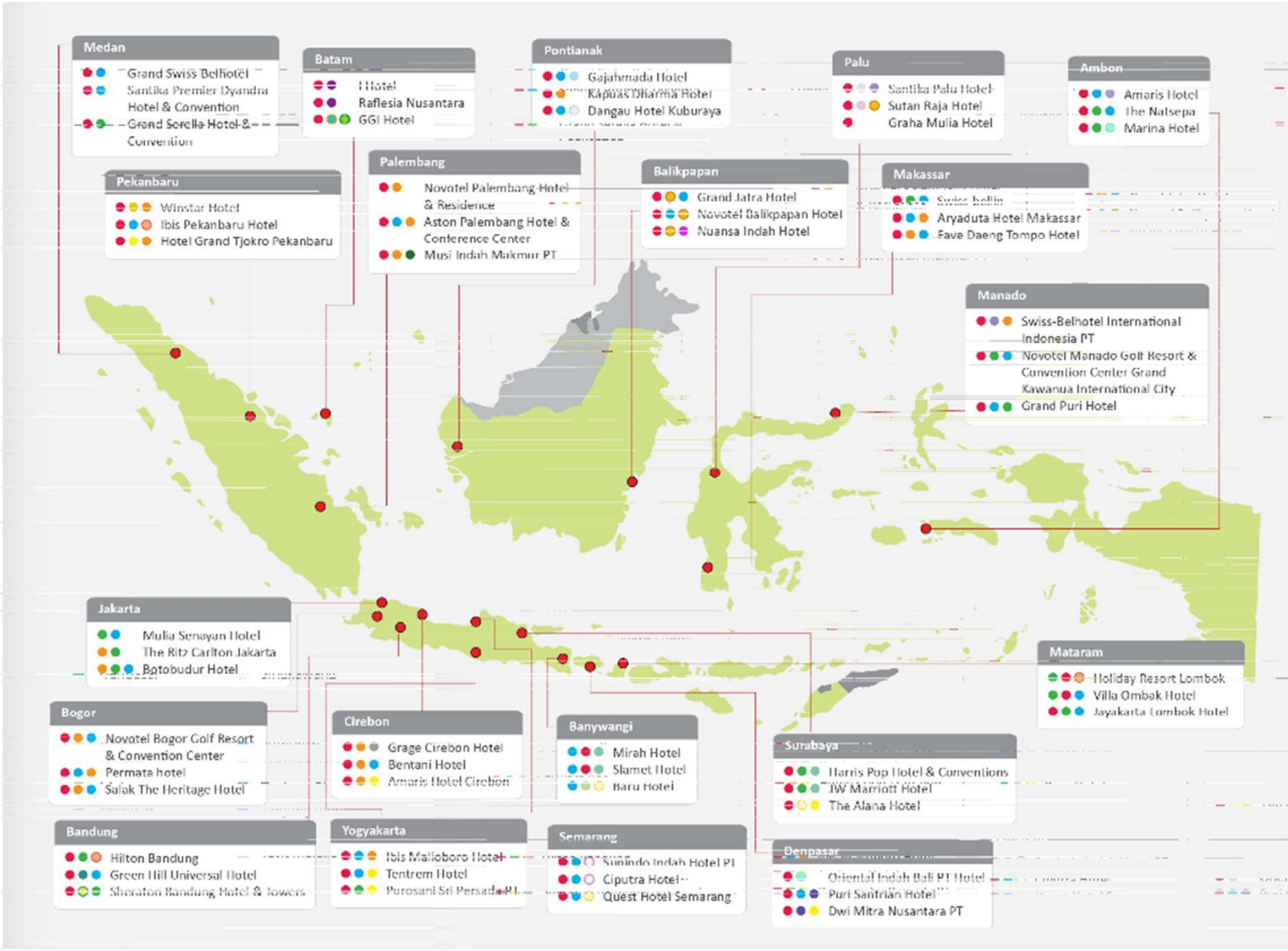
Sumber Data:

1. Searching Internet
2. URL Click terkait situs pendidikan
3. Lokasi user yang melakukan click

Row Labels	Sum of TOTAL_VIEW	% of TOTAL_VIEW
KOTA BANDUNG	23641	19.13%
JAKARTA SELATAN	16246	13.14%
KOTA SEMARANG	14870	12.03%
MEDAN SUMUT	14046	11.36%
DENPASAR	10478	8.48%
PEKANBARU RIAU	10056	8.14%
BATAM RIAU	10033	8.12%
DEPOK	9348	7.56%
KAB BANDUNG	7440	6.02%
GIANYAR	7438	6.02%

Row Labels	% of TOTAL_VIEW	Sum of TOTAL_VIEW
USA	23.39%	80031
Australia	14.86%	50839
UK	11.52%	39409
Saudi Arabia	10.96%	37485
Singapore	4.63%	15828
Japan	4.20%	14368
Canada	3.48%	11904
Netherlands	3.29%	11264
Malaysia	2.89%	9879
Germany	2.70%	9232

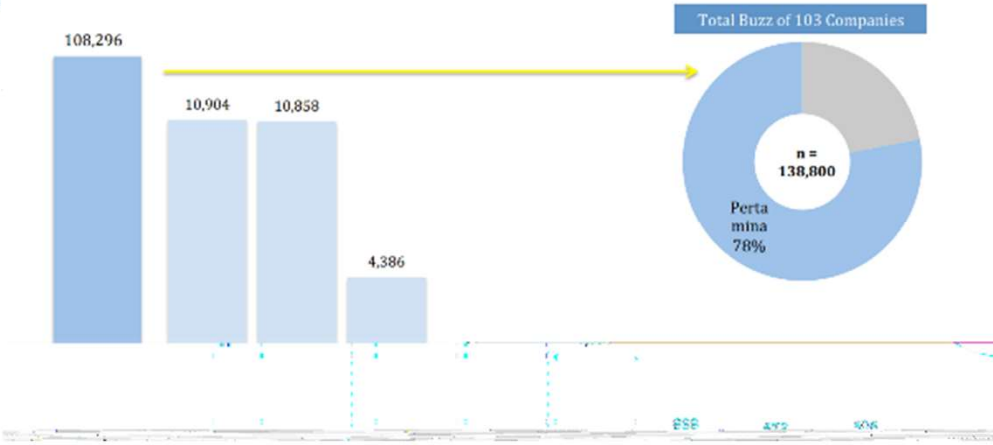
TOP HOTELS IN INDONESIA 2017

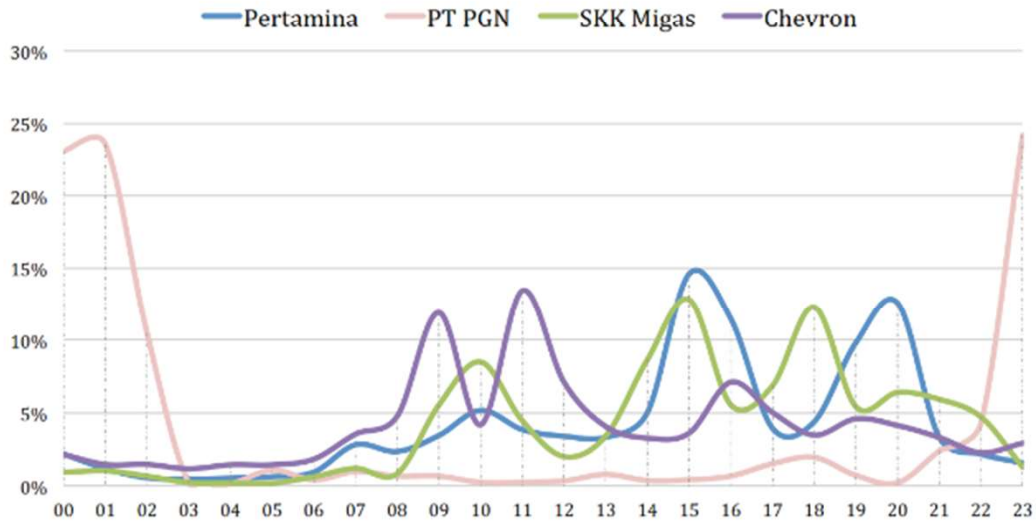


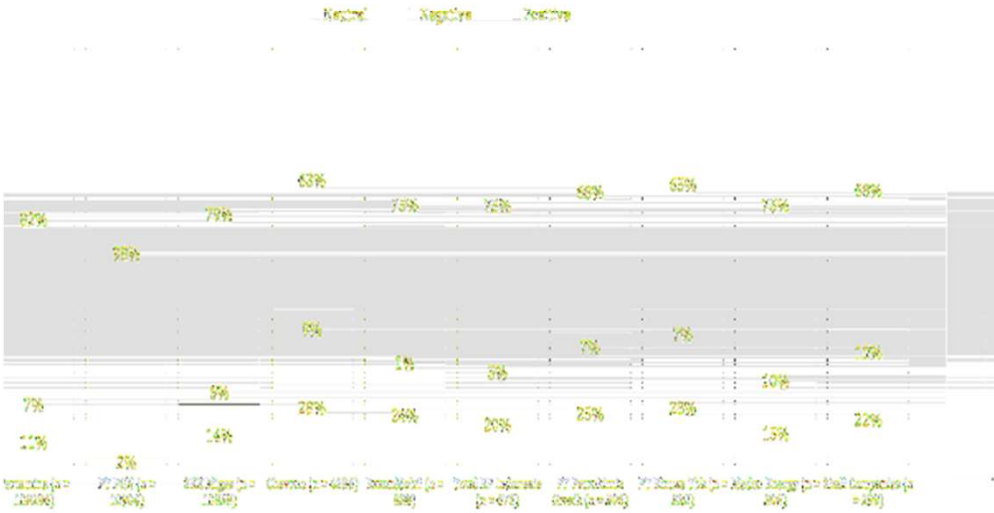
Visitor Legend

- Bandung
- Banjarmasin
- Bogor
- Denpasar
- Indramayu
- Jakarta
- Jember
- Luwuk
- Malang
- Makasar
- Medan
- Pekalongan
- Pematang Siantar
- Prabumulih
- Salatiga
- Samarinda
- Semarang
- Singkawang
- Solo
- Subang
- Surabaya
- Tanjung Balai Karmun
- Tanjung Pinang
- Tanjung Redeb
- Tapaktua
- Tasikmalaya
- Yogyakarta

Company Image Tracking (Oil & Gas)





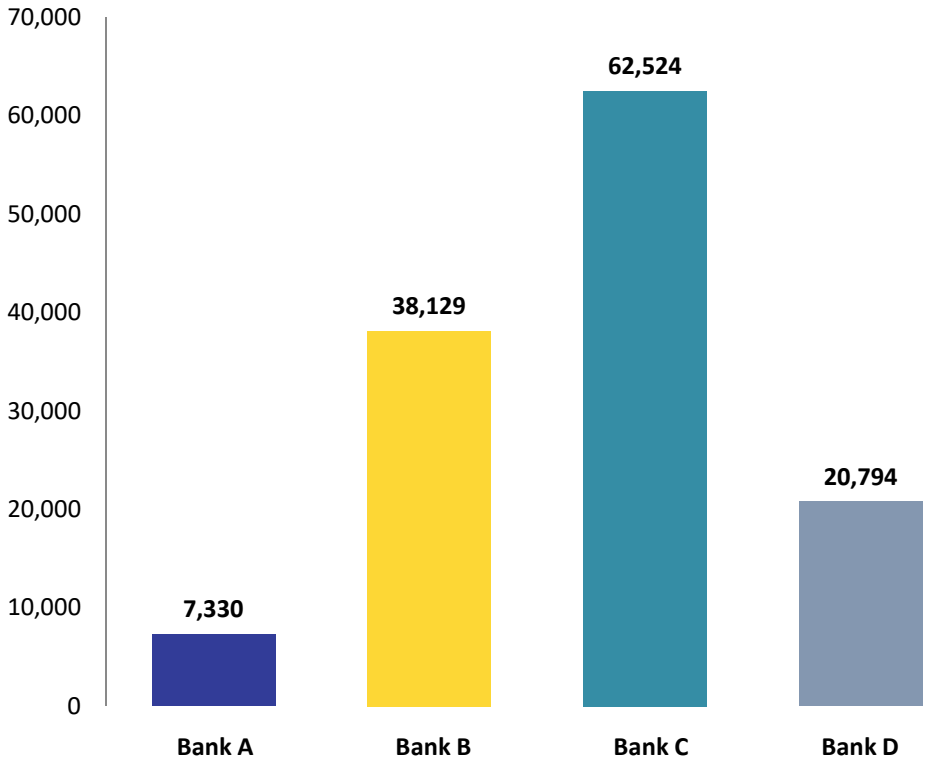


Source: Provetic & Telkom, 2015

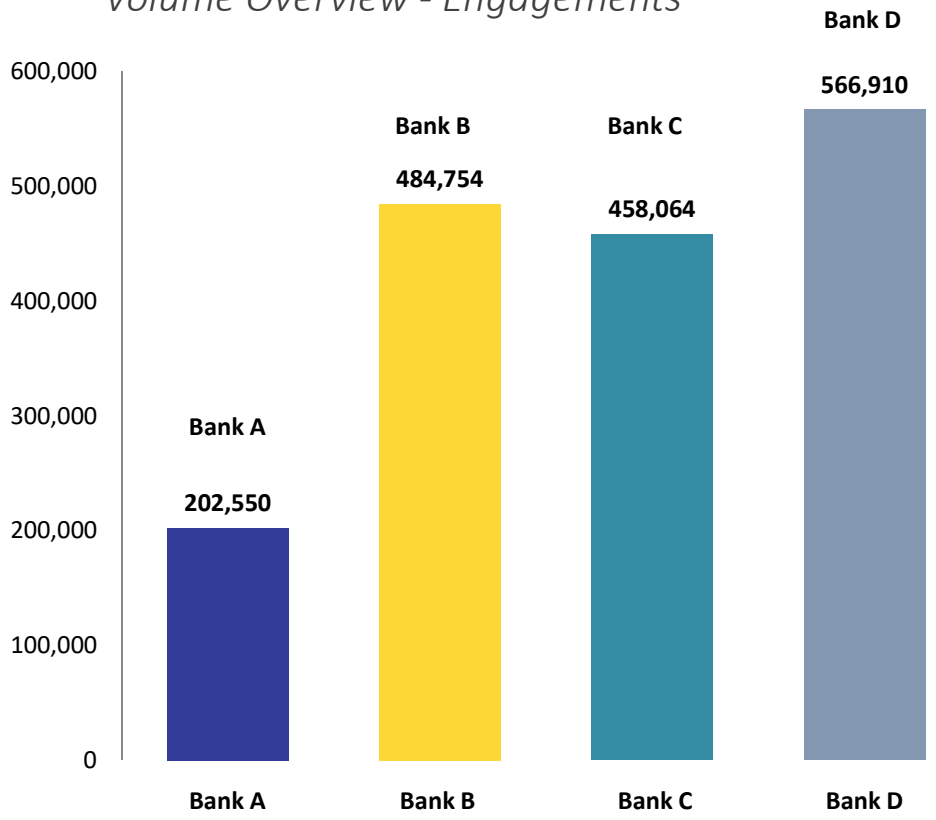
Social Media Analytics 4 Bank in Indonesia

Period tracking: Jan-Mar 2017)

Volume Overview - Post

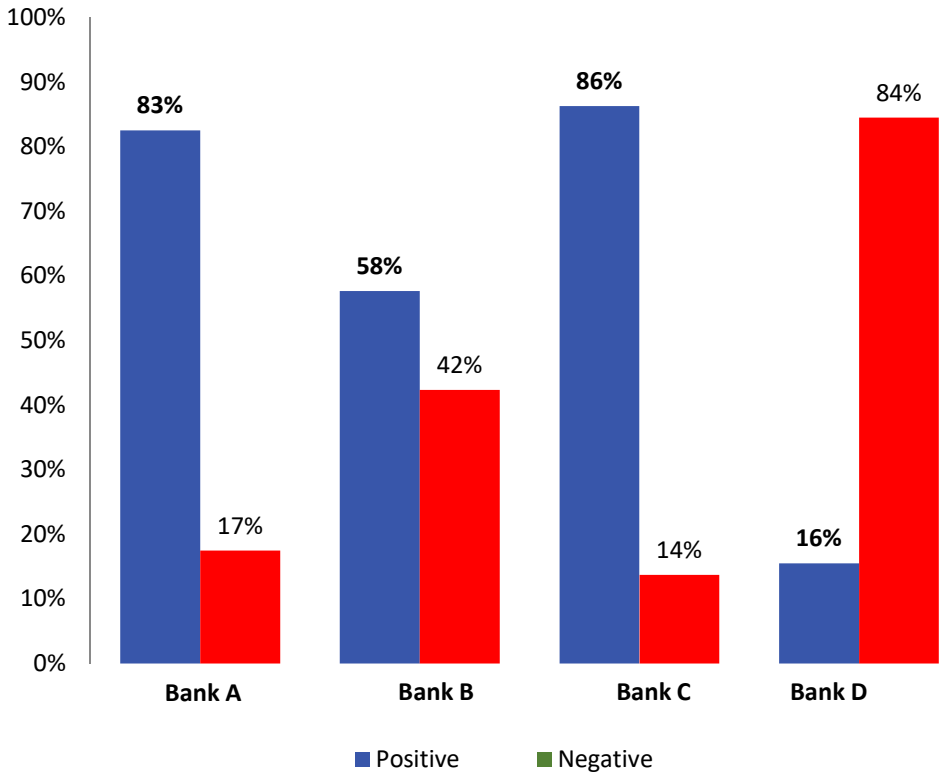


Volume Overview - Engagements



OVERALL DIGITAL BUZZ

Sentiment



Bank A

(+) POSITIF

- Promo layanan E-Toll Bank A untuk proses transaksi yang lebih cepat.
- **Jumlah : 2,107 percakapan**
- Pengangkatan direktur baru Bank A Bapak ……

(-) NEGATIF

- Komplain nasabah bahwa proses transfer uang ke dan dari Bank A sangat lama. **Jumlah : 17 percakapan**.
- Kasus korupsi program kredit kepada Lembaga Keuangan (KKLK). **Jumlah : 7 percakapan**.

Bank B

(+) POSITIF

- Pembagian laba 2016 Rp.6,2 Triliun. Jumlah: 1,094 percakapan.
- Titik terang kasus pembobolan Bank B oleh Kejagung

(-) NEGATIF

- Anjloknya laba Bank B pada 2016 sebesar 32.1%. Jumlah : 1,219 percakapan.
- Meningkatnya jumlah Kredit Macet(NPL)
- Komplain saldo E-Money hilang. Jumlah: 851 percakapan

Bank C

(+) POSITIF

- Campaign Bank C : #Quotes, #atGATF, #UKUKSR009, #DebitCard, #JF2017, #TwitJazzQuiz dan #SmartPayment menjadi daya tarik
- Invoasi Bank C Syariah dalam menghadirkan empat produk baru. Jumlah: 906 post.

(-) NEGATIF

- Dirut Bank C Syariah yang dipecat
- gangguan mobile banking. Jumlah: 327 percakapan

Bank D

(+) POSITIF

- Campaign #HUT dengan promo-promo untuk kartu kredit/debit
- Promo kartu kredit . Jumlah: 199 percakapan

(-) NEGATIF

- Gangguan Mobile Banking. Jumlah 1,126 post dengan 1,991 interaksi
- Kasus korupsi pajak Bank D Jumlah 164 percakapan dengan 1,290 interaksi

TOP CHANNEL DAN PROGRAM DI INDONESIA

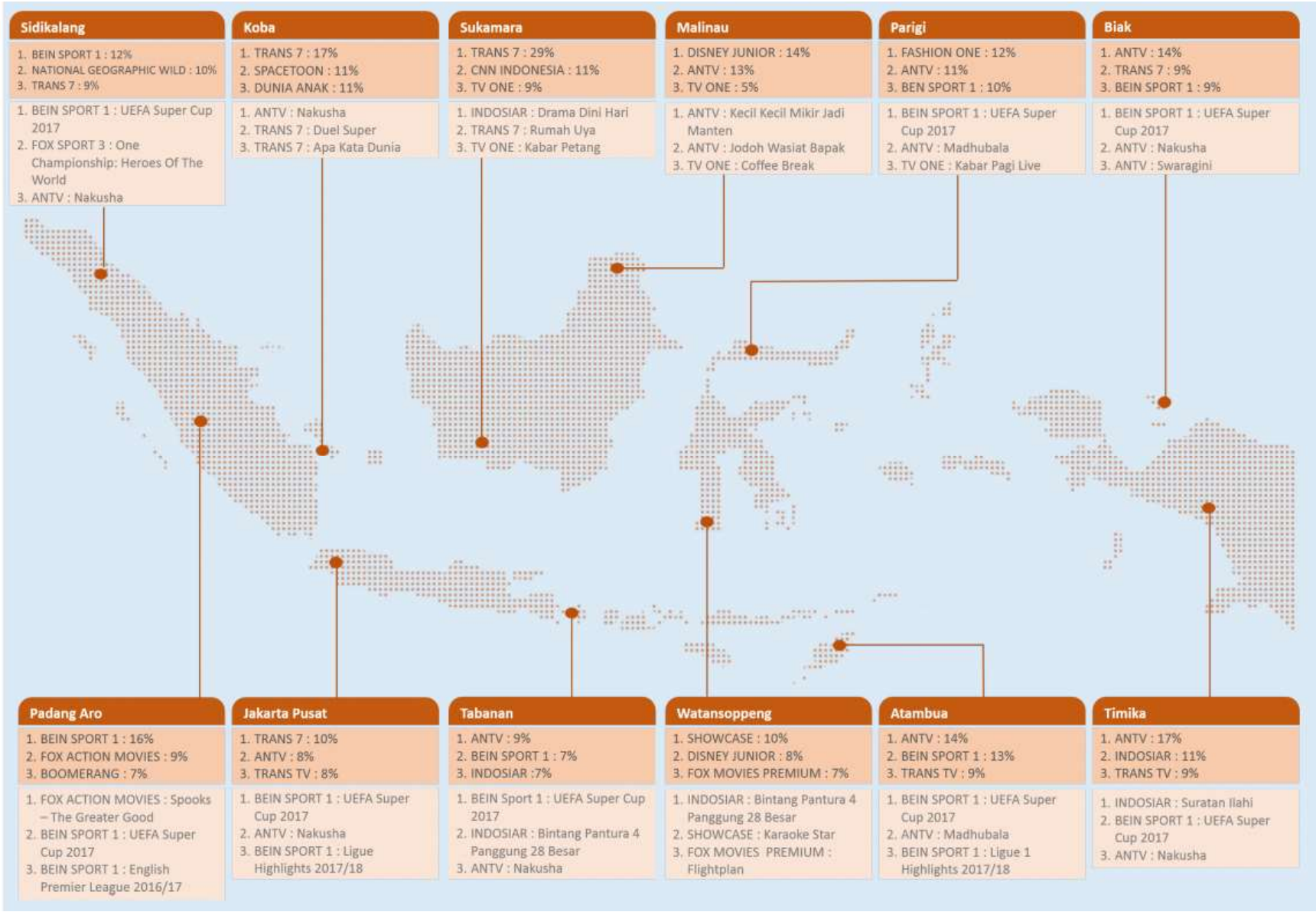
***Source**: UseeTV 9 Agustus 2017*

1. TRANS 7
2. antv
3. TRANSTV

1. beIN 1 SPORT
   UEFA Super Cup 2017
2. antv
   Nakusha
3. antv
   Madhubala

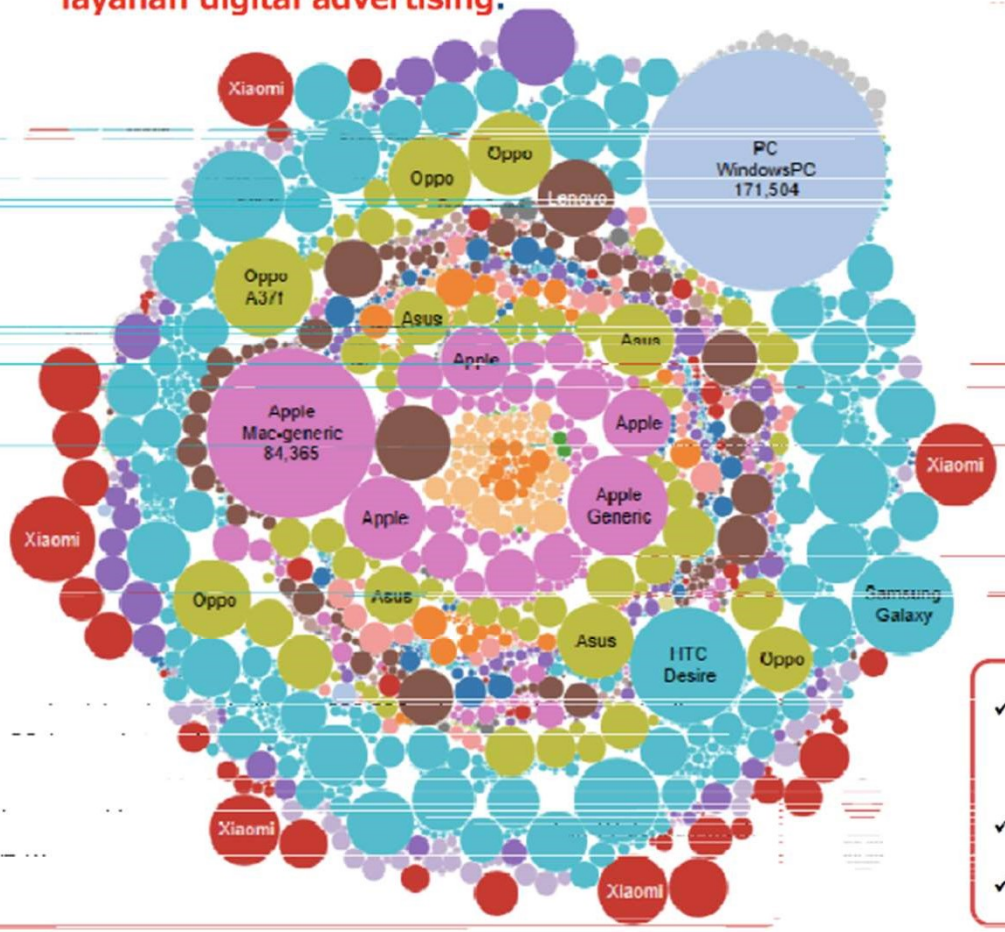
Top Channel

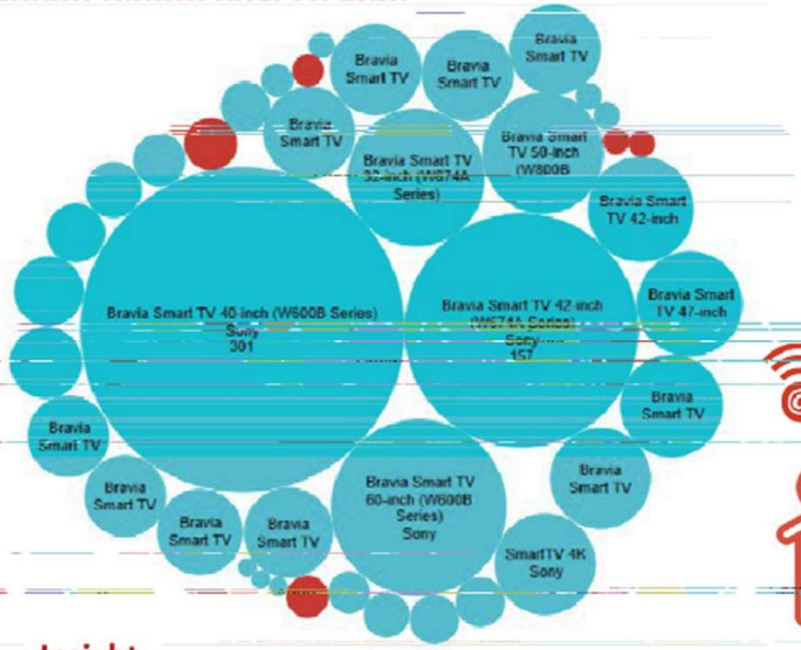
Top Program



Big Data Insight - Broadband Customer Profile

Pelanggan Indihome dan wifi.id nasional memiliki insight 53 juta unique device, insight behavior dalam menikmati video / musik streaming dan browsing Internet yang berpotensi untuk di-monetize melalui layanan digital advertising.





@wifi.id
Indonesia WiFi

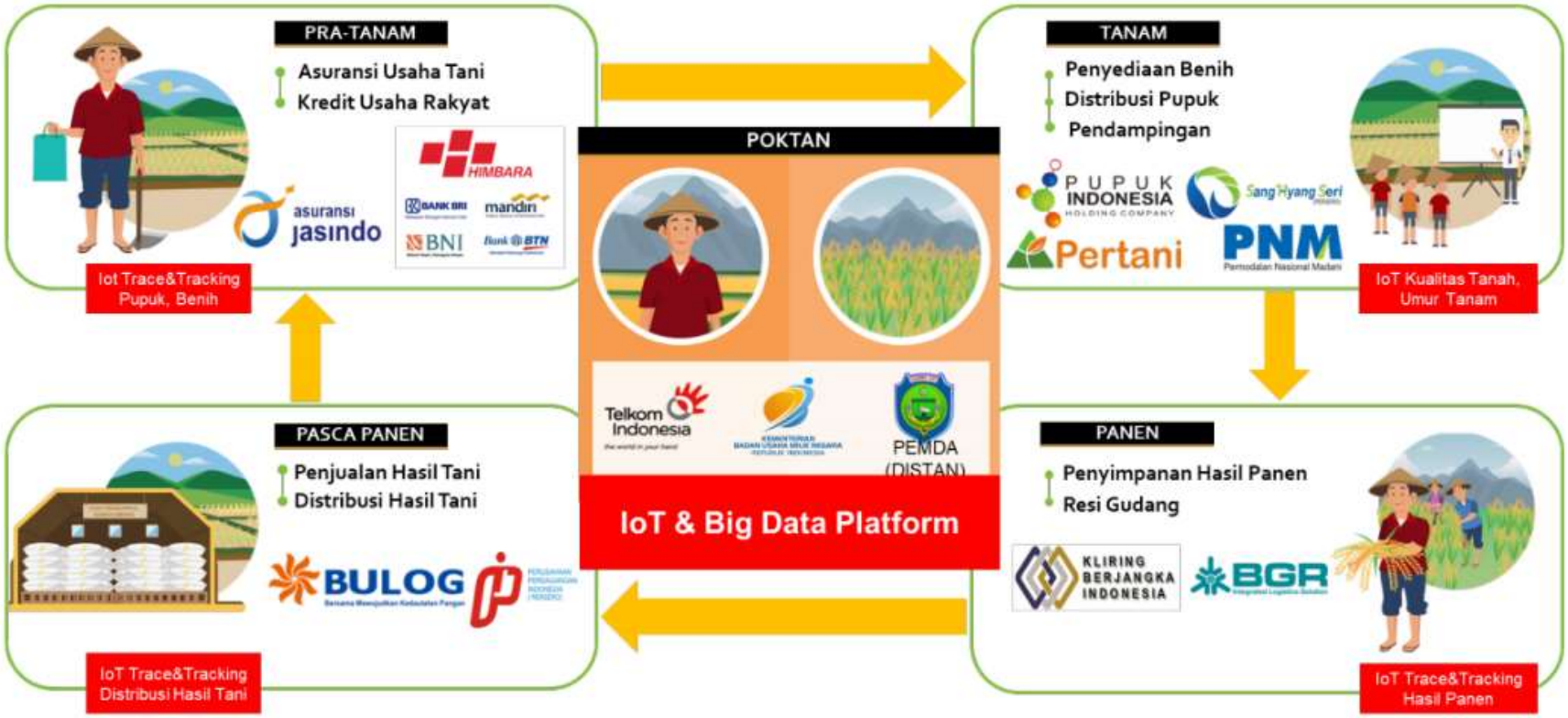
INDI HOME

Insight

- ✓ Jawa Tengah dengan jumlah pelanggan Indihome 223.526 pelanggan, dapat memberikan insight pengguna PC dengan kategori :
  - ➢ 171.504 pemakai PC Windows
  - ➢ 84.365 pemakai Apple
- ✓ Dari brand TV, pelanggan Indihome di Jawa Tengah mayoritas menggunakan TV dengan merk Sony Bravia Smart TV
- ✓ Data lain yang di-capture : Program TV, Music Streaming, Behaviour user

DATA TUNGGAL PERTANIAN (LOGISTIK TANI) – BIG DATA & IOT USECASE

Flow of Process pada setiap entitas mengoptimalkan peran IoT melalui perangkat Point of Sales (PoS) dan Data Analytic pada platform Big Data



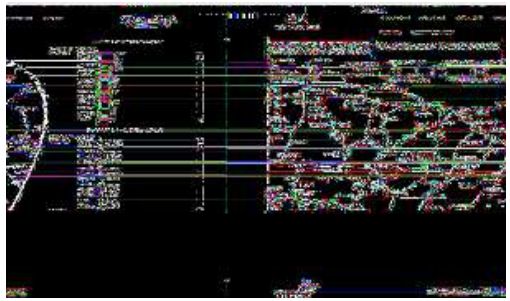
DATA ANALYTICS & VISUALISATION DATA TUNGGAL PERTANIAN (LOGISTIK TANI)

Executive Summary

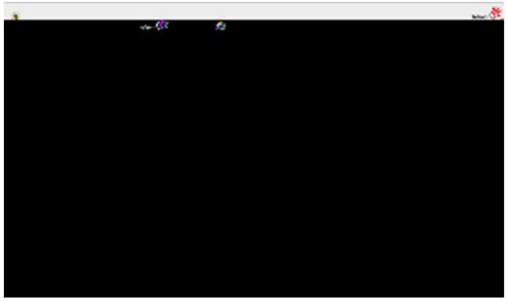
Profile Petani

Poligon Lahan

Profile Lahan



Sumber Permodalan



Penyerapan Benih

Produktivitas Petani

Trend Harga

Data Analytics

Data-data yang terkumpul dikorelasikan dengan berbagai data yang tersedia (data cuaca, data harga-harga komoditas dan dilakukan analisis untuk memperoleh insight –insight baru

Data Visualisation

Insight ditampilkan dalam bentuk dashboard untuk memudahkan pengambilan keputusan

CUSTOMER PROFILING & MICROSEGMENTATION

Framework From Data to Value

Multi Data sources → Data management layer → Analytical / visualisation layer → Examples

Internal sources
- CDRs
- CRM
- IPTV CDRs
- Deep Packet Inspection-1
- Network Alarm
- …

External sources
- Census data
- Social Media

Data management layer

Hive DB
- Staging tables
- L1 data-mart tables
- L2 data-mart

MySQL DB
- Other tables
- L2 data-mart

Analytical / visualisation layer

R
- Algorithms
- Segmentation models

Tableau
- Visualisation dashboards

Examples

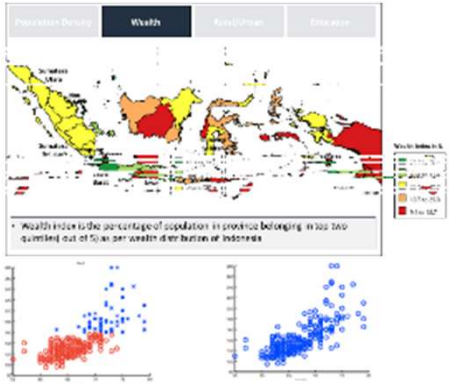
Classification & Regression :
- Naïve Bayes – Decision Trees
- Log Regression – Etc.

Clustering :
- K-means - PIC
- Gaussian - LDA

Collaborative Filtering

Optimization

Etc.



Customer Profiling:

Total of 265 variables are available for Customer household level profiling

Age band, gender, income and education level etc.

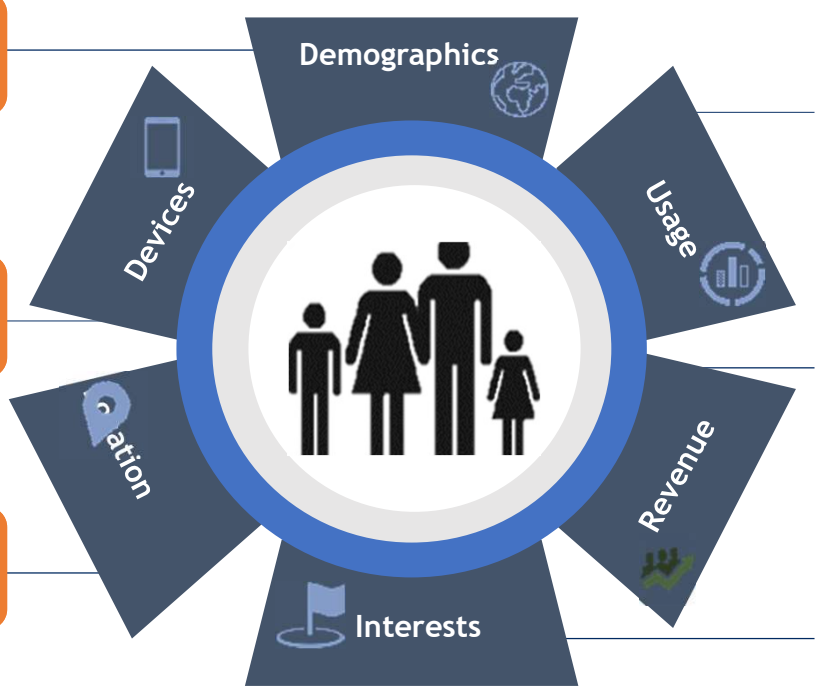
- 8 variables

Device type, brand, characteristics etc.

- 8 variables

Passive and active location – home, office, POI etc.

- 6 variables



- 82 variables

Behavioural data based on internet (DPI) and TV content

- 16 variables

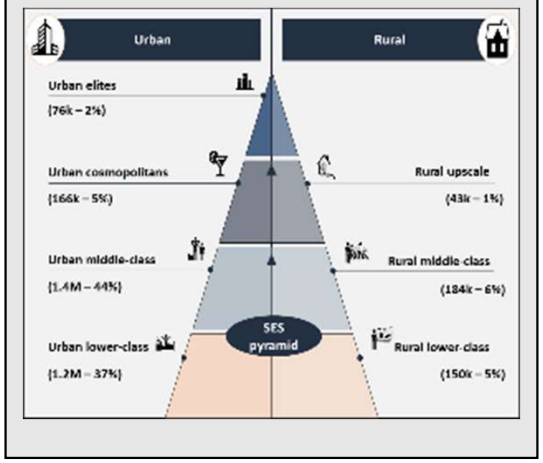
Revenue breakdown by products, usage etc.

- 145 variables

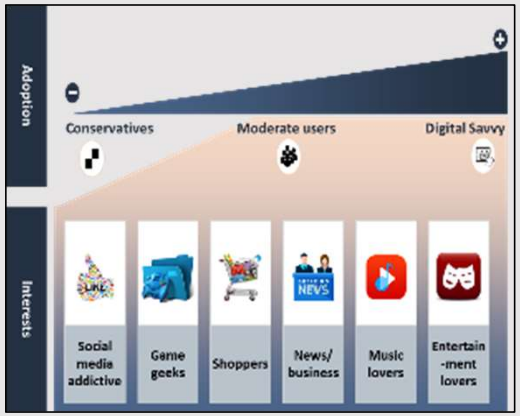
Behavioural data based on internet (DPI) and TV content

Micro-Segmentation Dimensions

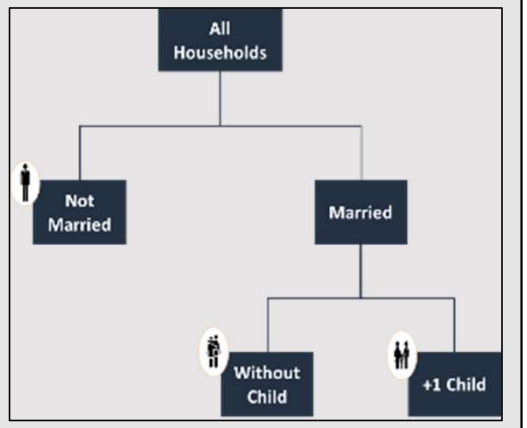
1 Socio-Economic Status

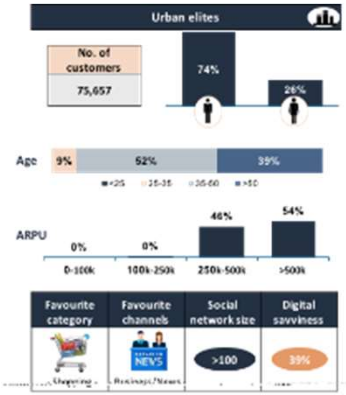


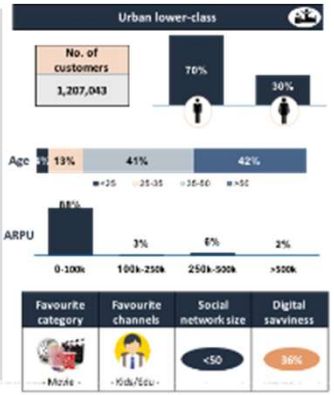
2 Digital adoption and interests

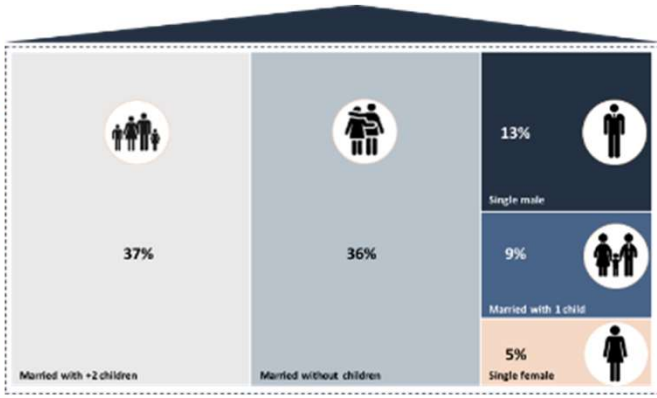


3 Household Hierarchy

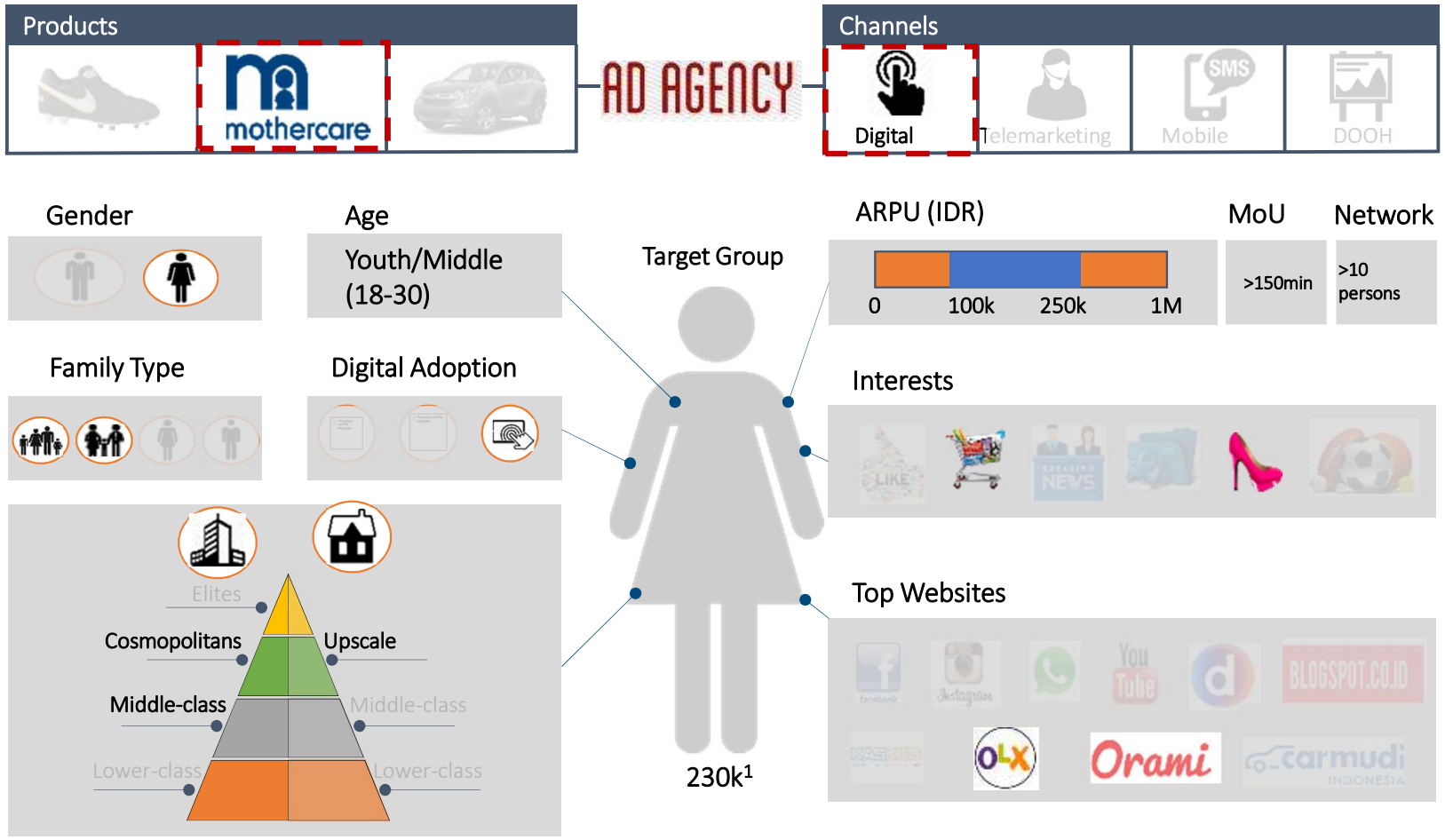




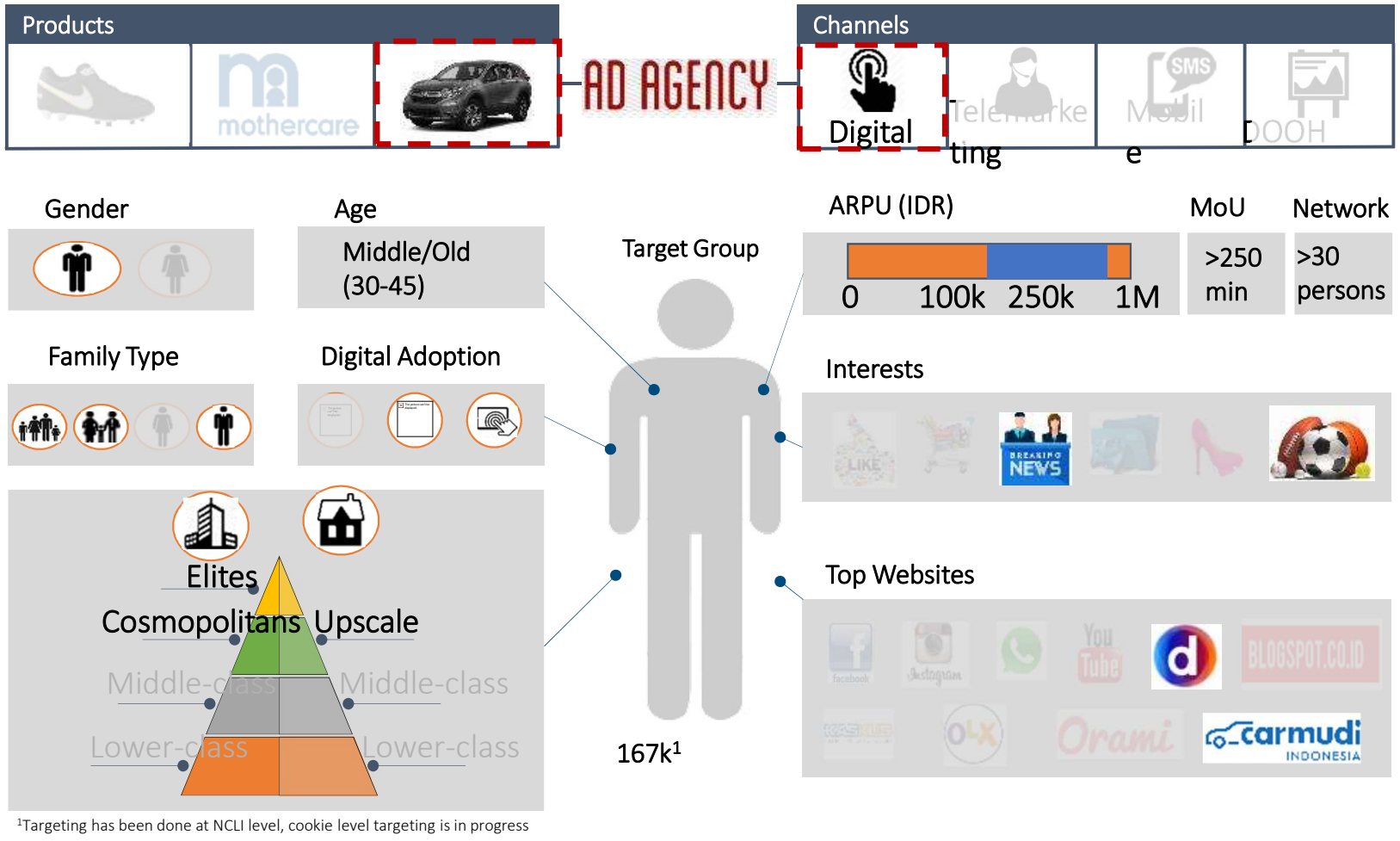




Example: Digital Ads Targeting



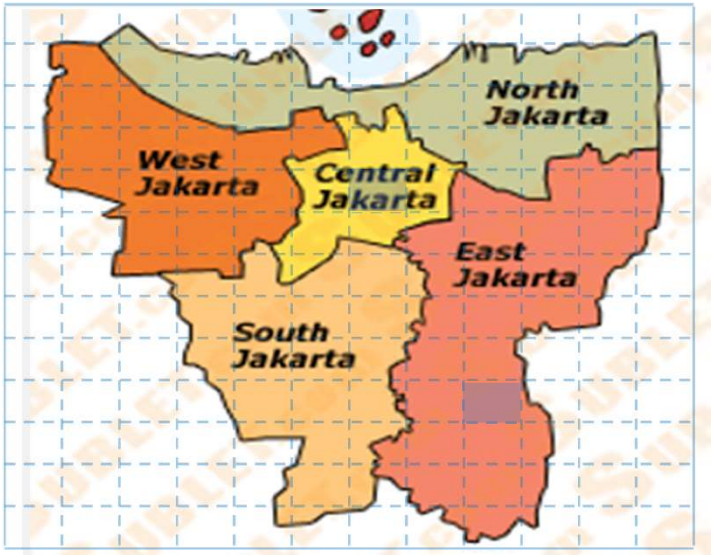
Digital Ads Targeting (2)



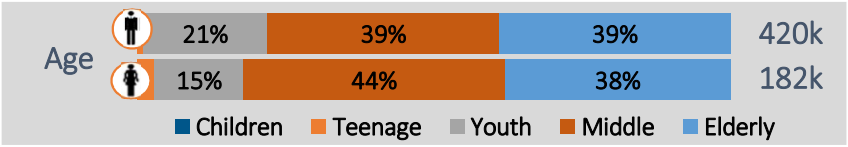
[1]Targeting has been done at NCLI level, cookie level targeting is in progress

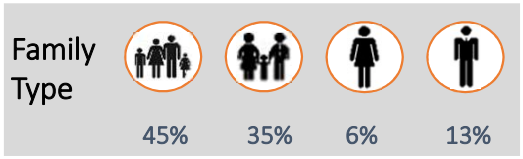
Smart Market Analysis

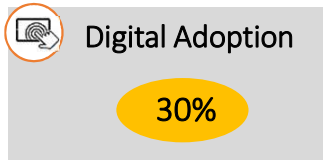
Jakarta	East Jakarta – Taman Mini

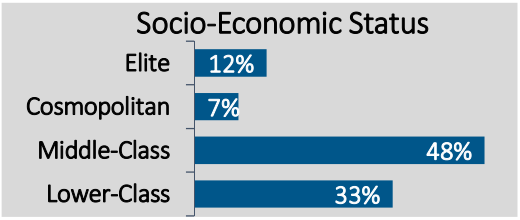


Micro-Segmentation









Area Characteristics

# Households	0.6M	Land Prices per sq. meter	IDR 10M
Population	10M	Avg. Rent for 2-Beds House	IDR 7M
Pop Density per km. meter	14k	Avg. Price for 2-Beds House	IDR 25M

Profiling Metrics

Interests	Demographics	Other
Web Categories	Education	ARPU
TV Channels	Religion	Device Type

Conclusion: Now… the roles of BIG DATA are….