

6th Malaysia Statistics Conference

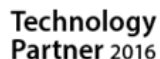
19 November 2018

Sasana Kijang, Bank Negara Malaysia

2018

Embracing Data Science and Analytics to Strengthen
Evidence-Based Decision Making

Humanizing Data



6th Malaysia Statistics Conference

The background of the image is a landscape with mountains and a cloudy sky, all tinted in shades of blue. The mountains are in the distance, and the sky is filled with soft, white clouds. The overall mood is serene and contemplative.

IT IS NOT THE STRONGEST OF THE SPECIES
THAT SURVIVES, NOR THE MOST INTELLIGENT
THAT SURVIVES. IT IS THE ONE THAT IS THE
MOST ADAPTABLE TO CHANGE.

CHARLES DARWIN

BIG DATA & AI LANDSCAPE 2018

INFRASTRUCTURE

The collage is organized into three main sections, each with a title and a collection of logos:

- HADOOP ON-PREMISE**
 - cloudera
 - Hortonworks
 - MAPR
 - Pivotal
 - IBM InfoSphere
 - bluedata
 - jethro
- HADOOP IN THE CLOUD**
 - aws
 - Microsoft Azure
 - IBM InfoSphere BigInsights
 - Google Cloud
 - treasure data
 - Dubate
 - altiscale
 - CAZENA
 - CenturyLink
- STREAMING / IN-MEMORY**
 - aws
 - databricks
 - stream
 - confluent
 - GridGain
 - DATA TORRENT
 - dataArtisans
 - ORACLE
 - hazelcast
 - TERRACOTTA

The collage displays logos for several database categories:

- NoSQL DATABASES:** Google Cloud, AWS, Oracle, Microsoft Azure, MongoDB, MarkLogic, KESPERIO, QARTS3, ArangoDB, Couchbase, Redis Labs, and SCYLLA.
- NewSQL DATABASES:** SAP, Clustrix, Pivotal, Cockroach Labs, MEMSQL, Influxdata, Aerospike, VOLTO, eXtremeDB, CUBRID, Splice Machine, and paradigm.
- GRAPH DBs:** Neo4j, Amazon Neptune, IBM, Oracle, and Gremlin.
- MPP DBs:** Teradata, Vertica, IBM Data Warehouse Systems, C-Action, Kognitio, Exasol, and Informatica.
- CLOUD EDW:** AWS, Google Cloud, Microsoft Azure, Pivotal, Snowflake, and Alteryx.

DATA TRANSFORMATION

- talend
- pentaho
- alteryx
- TRIFARCA
- tamr
- Paxata
- StreamSets
- UNIFI

DATA INTEGRATION

- Informatica
- Microsoft
- snapLogic
- TEALUM
- enigma
- Segment
- data
- redmud data
- ZALONI
- Alacorn
- Alation
- import.io
- Stitch

DATA GOVERNANCE

- Informatica
- SailPoint
- IBM
- McAfee Skyhigh Security Cloud
- colibra
- Waterintel
- OKERA

MGMT / MONITORING

- AWS
- New Relic
- actifio
- APPROXYMANS
- rubrik
- WAVEFRONT
- dynatrace
- in-vivo
- splunk
- SignalFX
- datadog
- Moogsoft
- pagerduty
- univention
- Numeracy
- bluebird

STORAGE  Google Cloud  PURESTORAGE     COHESITY	CLUSTERS/VCS    MEZOSPHERE  	APP DEV      	CROWD - SOURCING     	HARDWARE     MYTHIC    	GPU DBs      PG-Stream
--	--	--	--	--	---

CROSS-INFRASTRUCTURE/ANALYTICS



ANALYTICS

[illegible]

The collage displays logos for various data science and machine learning platforms, organized into three main categories:

- DATA PLATFORMS:** Includes logos for Microsoft, AWS, Domo, Veeva Analytics, Looker, ATSCALE, Arc4Data, Databricks, Google Data, and Birst.
- VISUALIZATION:** Includes logos for Tableau, SAP Analytics Cloud, Google Cloud, Celonis, Qlik, Palenq Data, ZEPH, Xignite, Charlio, and Plotly.
- MACHINE LEARNING:** Includes logos for AWS, Google Cloud, H2O, DataRobot, Gamalan, Element, Vizeen, Versive, and Bonsai.

COMPUTER VISION

- Microsoft Azure
- Amazon Rekognition
- Clarifai
- Google Cloud Vision
- EVER AI
- deepomatic

HORIZONTAL AI

- IBM Watson
- Cortana
- Facebook
- Sentient
- Voyager
- Affective
- Petuum
- Numenta
- Naraling
- Curious AI
- OSARO

SPEECH & NLP

- Google Cloud
- Twilio
- Amazon Lex
- Semantic Machines
- SoundHound
- IBM Watson
- Nuance
- others

SEARCH	LOG ANALYTICS	SOCIAL ANALYTICS	WEB / MOBILE / COMMERCE ANALYTICS
 elasticsearch  EXALENGE  Lucidworks  swiftype  alphasense  omni-us	 ORACLE  splunk  sumologic  LOGGY  SIBERS  kibana  logz.io	 Hootsuite  sprinklr  NETBASE  synthesio  simplesearch  billy  SimilarWeb	 Google Analytics  mixpanel  sumal  RESCI  granity  amplitude  Airtable  SIGOPT  custora

OPEN SOURCE

The diagram illustrates a comprehensive ecosystem of data science and engineering tools, organized into 12 functional categories:

- FRAMEWORK**: Includes Hadoop, MapReduce, YARN, Flink, Spark, Mesos, and CDAP.
- QUERY / DATA FLOW**: Includes Spark SQL, Presto, SLAM DATA, and Google Cloud Dataflow.
- DATA ACCESS**: Includes Hadoop, Nifi, mongoDB, Cassandra, CouchDB, HBase, and others.
- COORDINATION**: Includes Talend, Apache Zookeeper, and Apache Ambari.
- STREAMING**: Includes Spark, Flink, Beam, Kafka, and Storm.
- STAT TOOLS**: Includes Python, ScalaLab, and SciPy.
- AI / MACHINE LEARNING / DEEP LEARNING**: Includes TensorFlow, Theano, Caffe, Microsoft Cognitive Toolkit, OpenAI, FeatureFu, DMSM, VELES, DIMSUM, Chainer, MAHOUT, and Aerosolve.
- SEARCH**: Includes Elasticsearch and Solr.
- LOGGING & MONITORING**: Includes Elasticsearch, Kibana, Sentry, Logstash, and Prometheus.
- VISUALIZATION**: Includes BeakerX, Rodeo, and others.
- COLLABORATION**: Includes Jupyter, Koyala, and Anaconda.
- SECURITY**: Includes Apache Ranger and Knox.

DATA SOURCES & APIs

APPLICATIONS – ENTERPRISE

APPLICATIONS – INDUSTRY

Advertising	Location	Government	FINANCE - LENDING	FINANCE - INVESTING	REAL ESTATE	INSURANCE
AppNexus	Location	OPENGOV	oneck	REDFIN	INSURANCE	
critico	Knowtton	mark43	oneck	OpenDoor	Intimate	
ORACLE	Clever	mark43	Kreditech	Datamint	Intimate	
MGAT	Declarra	mark43	AVANT	Quantiplex	Intimate	
theTradeDesk	kidapip	mark43	UpStart	VTS	Intimate	
distillery	PANDORA	mark43	INSIKT	CREDIBLE	Intimate	
YAFAR	GOVOR	mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Intimate	
		mark43	Wecount	ADAPAR	Intimate	
		mark43	UpStart	ADAPAR	Intimate	
		mark43	INSIKT	ADAPAR	Int	

The diagram illustrates the convergence of various industries into a central 'HEALTHCARE' hub. The industries are:

- HEALTHCARE**: flitron, Clover, AYUUS, HealthHero, MYABIOta, Gingerlio, Glove, baby2you, 3DMed, zebro, Patha, OviG
- LIFE SCIENCES**: Scientific Color, BioGenetix, NeoGenetix, verily, WuxiNextBio, ZEPHYRUS, Genentech, UCB, Eisai, Clear Labs, Novartis, DANAHEA, CITRINE, Alkermes, GSK, COWIN
- TRANSPORTATION**: UBER, TESLA, CLEARPATH, drive.ai, naute, PILOTAI, NIO, PTIMUS, moovit, nexcar, comma, a.i., netradyn
- AGRICULTURE**: FARMERS, instacart, RetailMeat, BLUE RIVER, Farmer's Edge, FarmLogs, Mavrx, ByteDance, Agribotix, neta, UN, remesh, ASAPP
- COMMERCE**: Instacart, RetailMeat, BLUE RIVER, Farmer's Edge, FarmLogs, Mavrx, ByteDance, Agribotix, neta, UN, remesh, ASAPP
- INDUSTRIAL**: Cyjot, PREDIX, UPTAKE, TACHYUS, Allurion, SCOOTER

40 ZETTABYTES

(43 TRILLION GIGABYTES)
of data will be created by
2020, an increase of 300
times from 2005

2005

2020

Volume SCALE OF DATA

6 BILLION PEOPLE
have cell phones



WORLD POPULATION: 7 BILLION

It's estimated that
2.5 QUINTILLION BYTES

(2.3 TRILLION GIGABYTES)
of data are created each day



Most companies in the
U.S. have at least
100 TERABYTES

(100,000 GIGABYTES)
of data stored

1 IN 3 BUSINESS LEADERS

don't trust the information
they use to make decisions



27% OF RESPONDENTS

in one survey were unsure of
how much of their data was
inaccurate

Veracity UNCERTAINTY OF DATA

Poor data quality costs the US
economy around
\$3.1 TRILLION A YEAR



The New York Stock Exchange
captures
1 TB OF TRADE INFORMATION
during each trading session



By 2016, it is projected
there will be

18.9 BILLION NETWORK CONNECTIONS

— almost 2.5 connections
per person on earth



Velocity ANALYSIS OF STREAMING DATA



Modern cars have close to
100 SENSORS
that monitor items such as
fuel level and tire pressure

As of 2011, the global size of
data in healthcare was
estimated to be

150 EXABYTES
(151 BILLION GIGABYTES)



30 BILLION PIECES OF CONTENT
are shared on Facebook
every month



Variety DIFFERENT FORMS OF DATA

By 2014, it's anticipated
there will be

420 MILLION WEARABLE, WIRELESS HEALTH MONITORS



4 BILLION+ HOURS OF VIDEO
are watched on
YouTube each month



400 MILLION TWEETS
are sent per day by about 200
million monthly active users





Cloud (Optional)

Multiple Platform

ORACLE®

Oracle

SIEBEL®

Siebel



Social Media

SAP

SAP



CRM



SharePoint



Google Analytics



Excel



FTP



Big Data



Many others

Data Extractor & Connector

ORACLE®

Oracle Connector

SIEBEL®

Siebel Connector



Social Media Connector etc.

SAP

SAP Connector



CRM Connector



SharePoint Connector



Google Analytics Connector



Excel Connector



FTP Connector



Big Data



Many others

Data Processor



Data Integrity



Metadata



Validation



Audit Trail



Exception Handling



Many others

Cloud/Hybrid

Visualisation Renderer



Motion Bubbles



Gauges



Cylinders



Charts



Maps



Trees



Radars



Many others

Multi Channel



IE



Firefox



Safari



Chrome



Android



Tablet



Smartphone



Blackberry



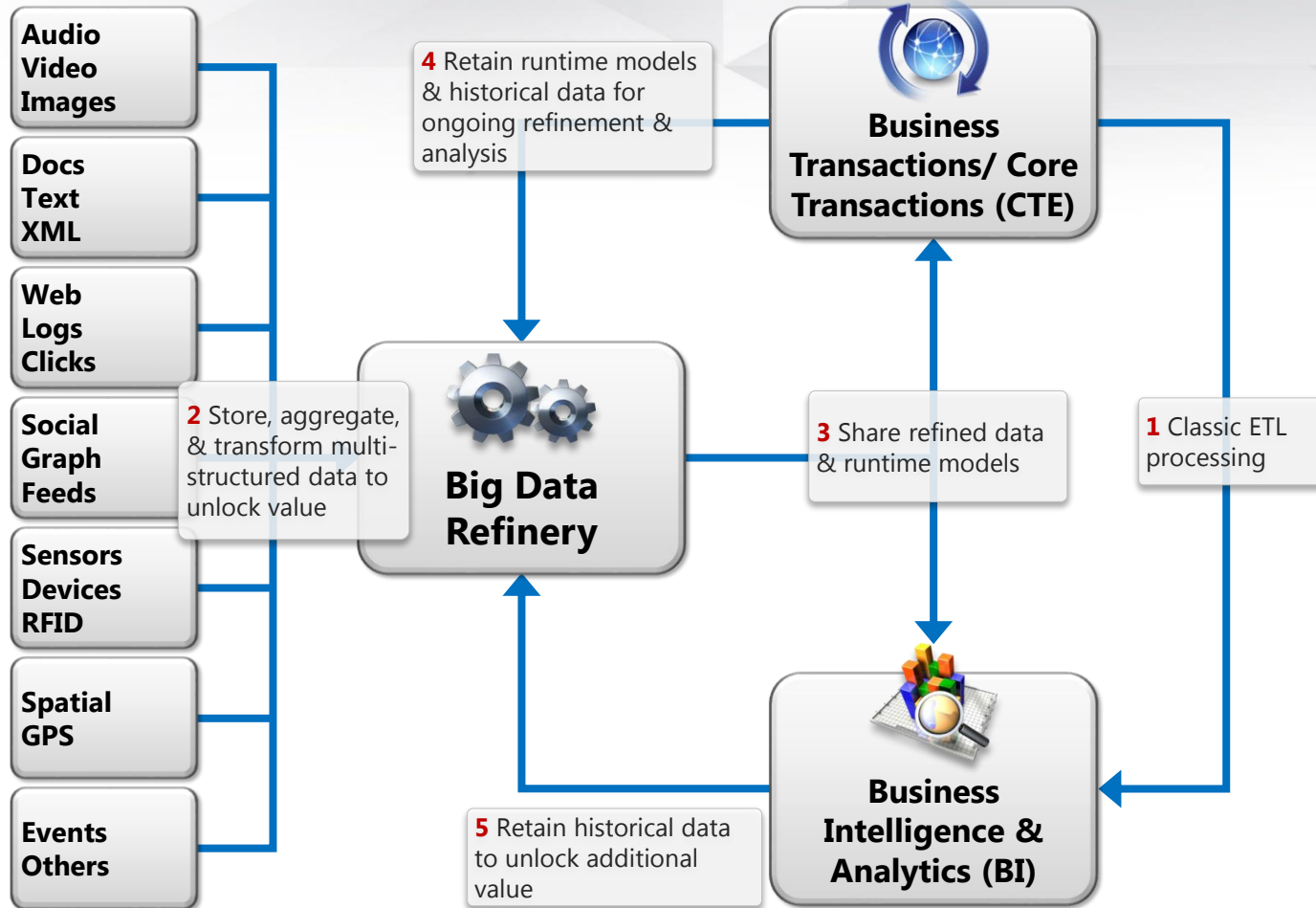
Microsoft Surface



Windows 8



Many others



A world map with a large, stylized question mark overlaid in the center. The question mark is dark blue with a thick white outline. The map shows various countries in different colors: Canada (purple), USA (orange), Brazil (orange), Argentina (yellow), Russia (pink), India (purple), Australia (purple), and others. The oceans are labeled: North Atlantic, South Atlantic, Indian Ocean, Arctic Ocean, and Pacific Ocean. The text "How's our world looks like today.." is written in white, bold, sans-serif font across the middle of the map, partially overlapping the question mark.

How's our world looks like today..

JAN
2018

DIGITAL AROUND THE WORLD IN 2018

KEY STATISTICAL INDICATORS FOR THE WORLD'S INTERNET, MOBILE, AND SOCIAL MEDIA USERS

TOTAL
POPULATION



we
are
social

7.593
BILLION

URBANISATION:
55%

INTERNET
USERS



4.021
BILLION

PENETRATION:
53%

ACTIVE SOCIAL
MEDIA USERS



we
are
social

3.196
BILLION

PENETRATION:
42%

UNIQUE
MOBILE USERS



5.135
BILLION

PENETRATION:
68%

ACTIVE MOBILE
SOCIAL USERS



2.958
BILLION

PENETRATION:
39%



APR
2018

GLOBAL ANNUAL DIGITAL GROWTH

YEAR-ON-YEAR CHANGE IN KEY STATISTICAL INDICATORS

INTERNET
USERS



we
are
social

+7%

SINCE APR 2017

+276 MILLION

ACTIVE SOCIAL
MEDIA USERS



+13%

SINCE APR 2017

+390 MILLION

UNIQUE
MOBILE USERS



we
are
social

+2%

SINCE APR 2017

+100 MILLION

ACTIVE MOBILE
SOCIAL USERS



+14%

SINCE APR 2017

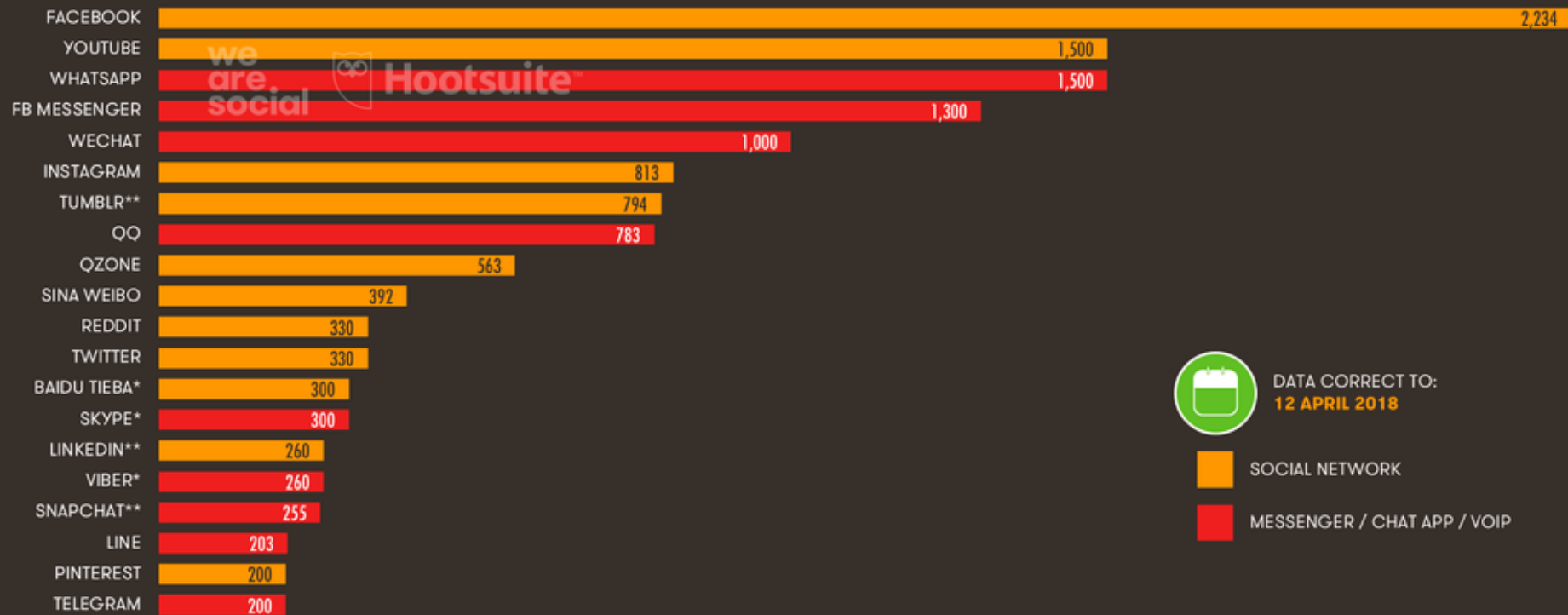
+389 MILLION



APR
2018

ACTIVE USERS OF KEY GLOBAL SOCIAL PLATFORMS

BASED ON THE MOST RECENTLY PUBLISHED MONTHLY ACTIVE USER ACCOUNTS FOR EACH PLATFORM, IN MILLIONS



DATA CORRECT TO:
12 APRIL 2018



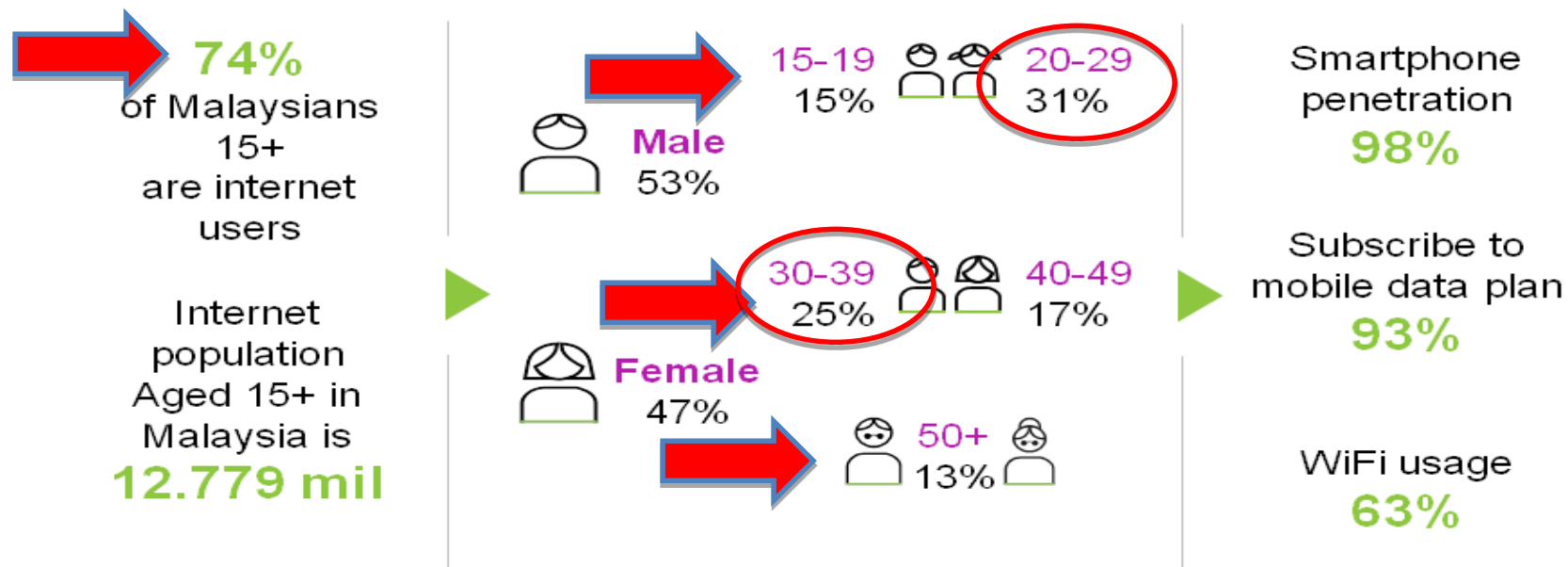
SOCIAL NETWORK



MESSENGER / CHAT APP / VOIP



MALAYSIAN INTERNET USERS – WHO ARE THEY?



Nielsen Consumer & Media View(CMV), Jul'16-Jun'17
Base: All 15+, accessed Internet past month (12.779 MM)

WORLD POPULATIONS

-  1.CHINA
-  2.INDIA
-  3.FACEBOOK
-  4.TENCENT
-  5.WHATSAPP
-  6.UNITED STATES
-  7.GOOGLE+
-  8.INDONESIA
-  9.LINKEDIN
-  10.TWITTER

Marrying the customer and employee experience is instrumental in firms' preparedness for digital disruption



90%

90% of respondents **agree** or **strongly agree** that digital technologies are **disrupting** the industry to a great or moderate extent.

AND



93%

90% of those with a digital strategy **agree** or **strongly agree** that the objective of their digital strategy is to **enhance** customer experience and engagement.

YET



46%

Only 46% **agree** or **strongly agree** that they are adequately preparing for digital disruption.



89%

of negative customer experiences stem from mistakes by customer service employees

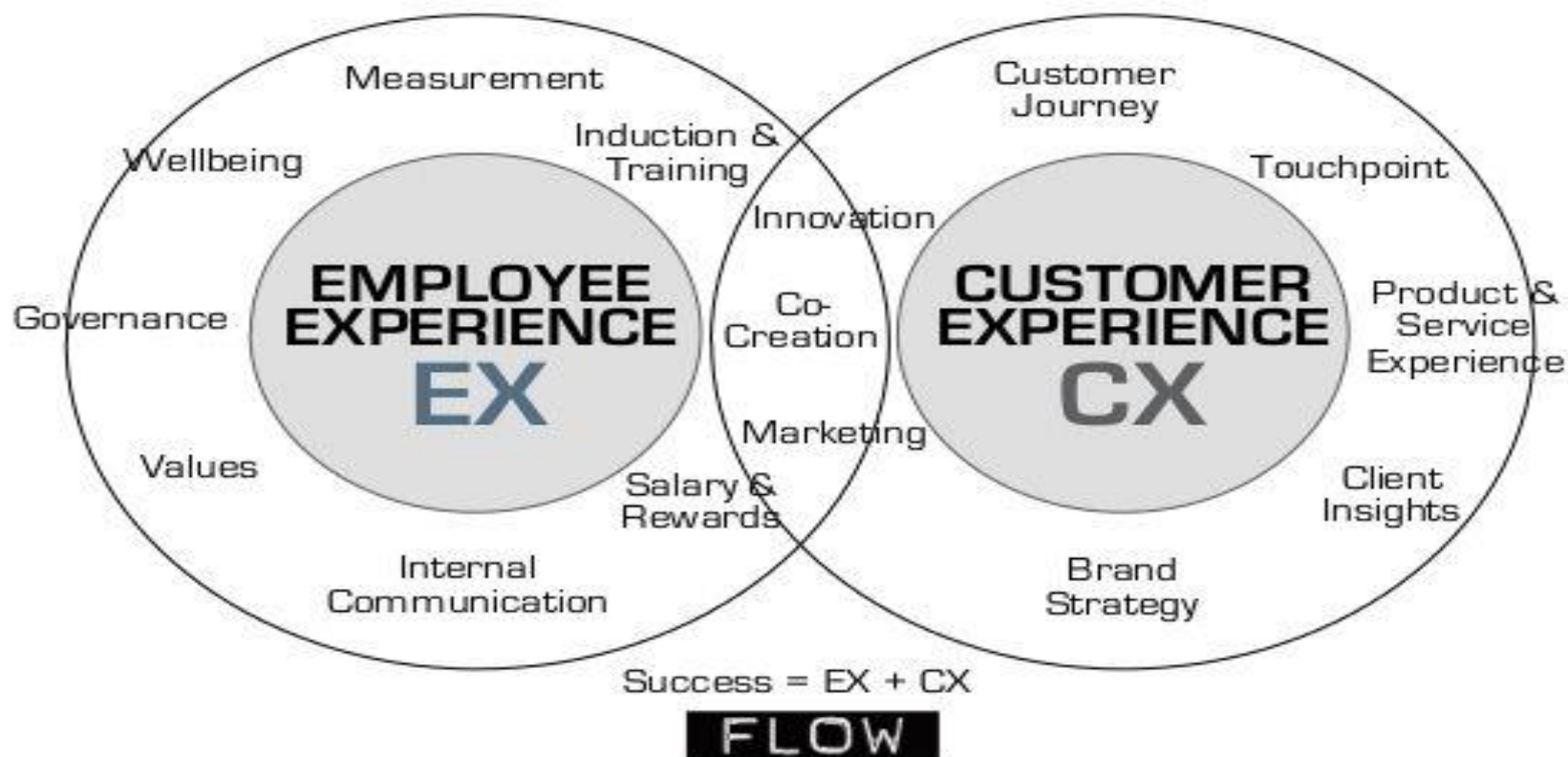
UNDERSTAND THE EMPLOYEE EXPERIENCE...

TO TRANSFORM THE CUSTOMER EXPERIENCE.

Desktop analytics allows companies to focus on areas for improvement in business processes, compliance, training, and application usage as well as support other initiatives such as capturing the customer journey at various touch points within an organization. Here are some ways companies are using desktop analytics to improve the employee and customer experience:

IHMINEN KESKIÖSSÄ

$$EX + CX = UX$$

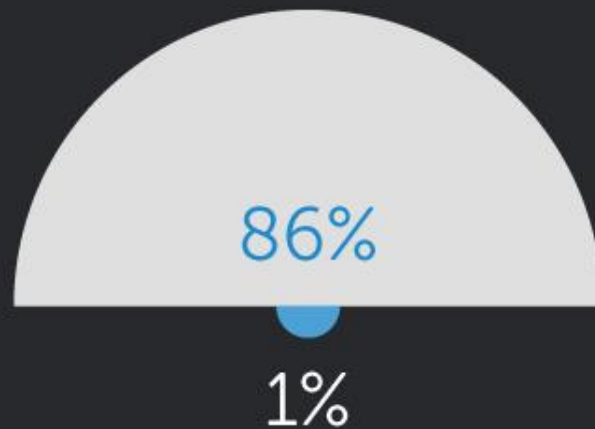


86%

of buyers will pay more
for a better customer experience,

but only **1%**
of customers feel
that vendors consistently meet
their expectations.

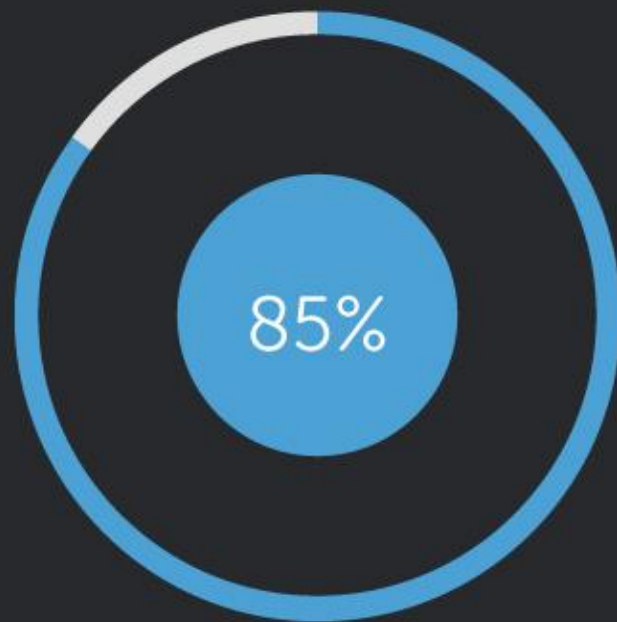
Source: CEI



By 2020,
the customer will manage

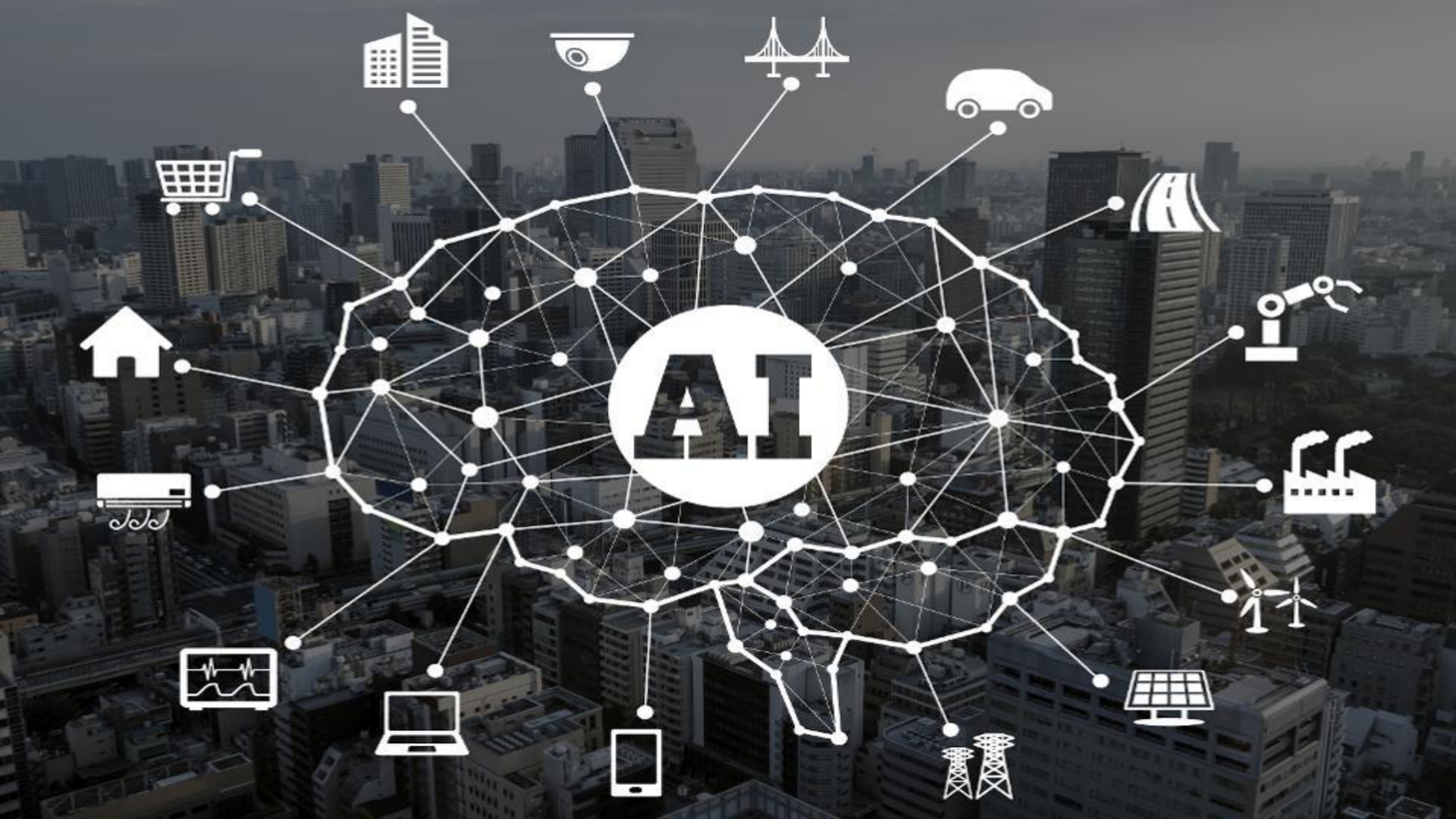
85%

of its relationship with an
enterprise without interacting
with a human.



Source: Gartner





The Future Of A.I.

Forecasted cumulative global artificial intelligence revenue 2016-2025, by use case (U.S. dollars)



TOP BIG DATA USE CASES

Customer

Financial

Marketing

Retail

Security

Pharma

Customer
Analytics
48%

45%
Experience
Analytics

Threat
Analysis
30%

Risk
Analysis
37%

28%
Regulatory
Compliance
Analysis

Campaign
Optimization
26%

23%
Location-based
Targeting

Fraud
Analysis
22%

16%
Brand
Sentiment
Analysis

Product
Placement
Optimization
16%

9%
Other

Drug
Discovery

1%

The future of big data and AI boils down to one thing

When it comes to big data, analytics and AI, the value does not come from collecting the data, or even from deriving some insight from it — value comes from just one thing: action.



splunk>

Transforming IT Ops with Predictive Analytics

Expand [+]



MORE LIKE THIS ::



AI-powered data analytics: Inside this transformative trend



知己知彼， 百战不殆

孙子兵法

TAKE HOME POINTS

1.



2.



3.



Thank You

Stan Lee



WINNER
Business
Intelligence



Technology
Partner 2016

