Introduction to Big Data

Dr. Junya Michanan

CRMA

Why Big?

Every 60 seconds

- **98,000+** tweets
- **695,000** status updates
- **11million** instant messages
- **698,445** Google searches
- **168 million+** emails sent
- **1,820TB** of data created
- **217** new mobile web users

Reference: www.practicalanalytics.wordpress.com

Why it Big?

BIG Data Component

Complexity of Big Data

***Source**: Contents of above graphic created in partnership with Teradata, Inc.*

4 Vs of Big Data

Big Data คืออะไร

- ข้อมูลมีขนาดใหญ่ (Volume) มีจำนวนหลาย TeraByte หรืออาจเป็น PetaByte ซึ่งอาจเป็นข้อมูลหลายสิบล้านหรือเป็นพันล้านชุด
- ข้อมูลเข้ามาอย่างต่อเนื่อง (Velocity) มีข้อมูลใหม่ที่เข้ามาตลอดเวลาอาจเป็นทุกวินาทีจำนวนมาก
- ข้อมูลมีความหลากหลาย (Variety) ข้อมูลมีทั้งแบบโครงสร้าง หรือไม่มีโครงสร้าง อาทิเช่นรูปภาพ ข้อความ หรืออาจเป็นกึ่มีโครงสร้างเช่น ไฟล์ CSV
- ข้อมูลมีความไม่แน่นอน (Veracity) ข้อมูลอาจไม่ครบถ้วน มีความไม่แน่นอน และอาจซ้ำซ้อน

ตัวอย่างอุตสาหกรรมที่มีข้อมูลขนาดใหญ่

- โทรคมนาคม ผู้ให้บริการโทรศัพท์เคลื่อนที่ก็อาจมีข้อมูลลูกค้า หรือข้อมูลการใช้งานโทรศัพท์ของลูกค้า (CDR: Call Detail Record) ที่แต่ละวันอาจเป็นหลาย Terabyte หรือหลายหมื่นล้านเรคอร์ด
- การเงินการธนาคาร มีข้อมูลการทำธุรกรรมของลูกค้าที่มีจำนวนมาก หรือข้อมูลการซื้อขายหลักทรัพย์ในตลาดทุน
- ค้าปลีก ห้างสะดวกที่มีรายการซื้อสินค้า ข้อมูลสต็อกสินค้า ข้อมูลลูกค้า ที่มีการเคลื่อนไหวในแต่ละวันจำนวนมาก
- การแพทย์ ข้อมูลการรักษาพยาบาลของผู้ป่วย ข้อมูลด้านยา ข้อมูลโรคระบาดต่างๆ

เหตุผลที่มีการกล่าวถึง Big Data ในตอนนี้

Source: William EL KAIM, Enterprise Architecture and Technology Innovation

ความก้าวหน้าทางเทคโนโลยี

- การใช้ชีวิตประจำวัน และการทำธุรกิจ ก็มีการสร้างข้อมูลขึ้นมาตลอดเวลา
- ข้อจำกัดเทคโนโลยีในแบบเดิมไม่สามารถเก็บข้อมูลในรูปดิจิทัลได้
- ความก้าวหน้าของเทคโนโลยีอย่าง Smart Phone, IoT, Social Media ทำให้มีการส่งและเก็บข้อมูลในรูปดิจิทัลมากขึ้น
- ราคา Storage และ CPU ก็ถูกลงทำให้สามารถเก็บข้อมูลได้มากขึ้นและประมวลผลได้

ข้อมูล Big Data เน้นที่ Transactional Data

- การวิเคราะห์ข้อมูลขนาดใหญ่ ต้องใช้ Transactional Data โดยจะเก็บข้อมูลที่มีรายละเอียดมากที่สุด เช่นข้อมูลการทำธุรกรรมการเงิน การซื้อสินค้า
- ข้อมูลสรุป (Summary Data) เป็นข้อมูลระดับสูงที่ให้ผู้บริหารใช้ในการตัดสินใจ การส่งข้อมูลประเภทนี้มาเพื่อใช้ในการวิเคราะห์ Big Data จะไม่สามารถทำการคาดการณ์ข้อมูลใดๆ
- ตัวอย่างใน Slide ถัดไป แสดงให้เห็น ข้อมูลสรุป และข้อมูลทำธุรกรรม จะเห็นได้ว่าข้อมูลสรุปจะไม่สามารถคาดการณ์ได้ว่า ลูกค้ารายนี้มาทานกาแฟบ่อยแค่ไหน การใช้จ่ายแต่ละครั้งเป็นอย่างไร มาเวลาใด ซึ่งต่างกับข้อมูลการทำธุรกรรม

ข้อมูลสรุป (Summary Data)

My Starbucks Rewards

The Stars will expire : **25/11/2018**

Card number : [redacted]

Status : Registered

Reload your card

Card number : 6[redacted]

Status : Registered

Reload your card

Card number : 6087728[redacted]

Status : Lost Card

Reload your card

MY STARBUCKS REWARDS

Card Information (609[redacted]

Balance = 1,871.25
Status = Registered

ข้อมูลธุรกรรม (Transactional Data)

TRANSACTION DATE	TRANSACTION TYPE	AMOUNT	BALANCE
12/7/2017 8:57:00 AM	Redeem Card	120.00	**1,871.25**
12/6/2017 7:48:00 AM	Redeem Card	120.00	**1,991.25**
12/4/2017 7:30:00 PM	Redeem Card	76.50	**2,111.25**
12/4/2017 9:12:00 AM	Redeem Card	153.00	**2,187.75**
12/3/2017 9:52:00 AM	Redeem Card	378.00	**2,340.75**
12/1/2017 2:37:00 PM	Redeem Card	76.50	**2,718.75**
12/1/2017 9:06:00 AM	Redeem Card	144.00	**2,795.25**
11/30/2017 12:14:00 PM	Redeem Card	135.00	**2,939.25**
11/29/2017 7:14:00 AM	Redeem Card	94.50	**3,074.25**
11/28/2017 7:25:00 AM	Redeem Card	105.00	**3,168.75**
11/27/2017 7:10:00 AM	Redeem Card	105.00	**3,273.75**
11/25/2017 1:56:00 PM	Redeem Card	150.00	**3,378.75**
11/24/2017 8:48:00 AM	Redeem Card	200.00	**3,528.75**
11/22/2017 10:27:00 AM	Redeem Card	170.00	**3,728.75**

ข้อมูลในโลกปัจจุบันคือ Big Data ประเด็นสำคัญคือเราสามารถที่จะเก็บข้อมูลขนาดใหญ่และสามารถทำการวิเคราะห์ได้หรือไม่

Descriptive Analytics
การวิเคราะห์เชิงบรรยาย
Predictive Analytics
การวิเคราะห์เชิงพยากรณ์
Data Analytics
Diagnostic Analytics
การวิเคราะห์เชิงค้นหาเหตุ
Prescriptive Analytics
การวิเคราะห์เชิงบรรเทาเหตุ

การวิเคราะห์ Big Data

- การทำ Business Intelligence (BI) คือการทำ Descriptive Analytics ที่อยากได้ข้อมูลสรุปว่าเกิดอะไรขึ้น เลยต้องมีการทำ Data Warehouse และ Data Cleansing
- ประโยชน์ที่สำคัญของ Big Data คือการนำข้อมูลดิบ (Raw Data) มาทำ Predictive Analytics ที่เป็นการคาดการณ์ในเรื่องต่างๆจากข้อมูลขนาดใหฯ่ ซึ่งต้องมีการทำ Data Lake และใช้หลักการของ Data Science

How can we make it happen?
Prescriptive Analytics
What will happen?
Predictive Analytics
Why did it happen?
Diagnostic Analytics
What happened?
Descriptive Analytics
Optimization
Foresight
Insight
Information
Hindsight
Value
Difficulty
Gartner

กลยุทธ์ด้านข้อมูล

- การใช้ข้อมูลเพื่อใช้ในการตัดสินใจเชิงธุรกิจให้ดีขึ้น
- การใช้ข้อมูลเพื่อให้กระบวนการทางธุรกิจเป็นไปอย่างมีประสิทธิภาพมากขึ้น
- การเปลี่ยนแปลงธุรกิจโดยทำให้ข้อมูลเป็นสินทรัพย์ขององค์กร

1) การใช้ข้อมูลเพื่อใช้ในการตัดสินใจเชิงธุรกิจให้ดีขึ้น

- เริ่มต้นด้วยการจัดความสำคัญของธุรกิจ และตั้งคำถามที่เราต้องการคำตอบ
- แบ่งออกมาเป็นสี่ด้าน:
    - ด้านลูกค้า การตลาด และคู่แข่ง
    - ด้านการเงิน
    - ด้านกระบวนการทำงานภายใน
    - ด้านบุคลากร

2) การใช้ข้อมูลเพื่อให้กระบวนการทางธุรกิจเป็นไปอย่างมีประสิทธิภาพมากขึ้น

- การทำให้งานประจำรายวันมีประสิทธิภาพมากขึ้น เช่น
  - การนำเสนอสินค้าให้กับลูกค้าที่เข้ามาในร้าน
  - การเพิ่มประสิทธิภาพการผลิตในโรงงาน
  - การจัดส่งสินค้าให้มีประสิทธิภาพมากขึ้น
  - การทำการตลาด แนะนำสินค้าที่ดีขึ้น
- กลยุทธ์ด้านนี้จะยากกว่าแบบแรก และต้องมีข้อมูลรายละเอียดที่มากขึ้น เช่น อาจต้องมีระบบ IoT

3) การเปลี่ยนแปลงธุรกิจโดยทำให้ข้อมูลเป็นสินทรัพย์ขององค์กร

- องค์กรจะใช้ข้อมูลเป็นหลักในการบริหาร ตัดสินใจ และเพิ่มคุณค่าให้กับองค์กร
- อาจเป็นการนำข้อมูลไปร่วมกันใช้กับคู่ค้า หรือเป็นทรัพย์สินให้กับองค์กร
- ตัวอย่างธุรกิจที่ใช้ข้อมูลในการทำงาน
    - FitBit
    - Uber
    - Facebook

Data Sources

Technology

Analytics

เทคโนโลยีด้าน Big Data

Data Collection/ Ingestion	Data Storage	Data Analysis/ Processing	Data visualisation

Big Data
Data Warehouse
BI
Data Lake
Data Science

Business Intelligence

- เน้น Descriptive Analytics และ Diagnostic Analytics
- เน้นความรู้สถิติ
- ลักษณะข้อมูล
    - มีจำกัด
    - Cleaned data
    - Structure
    - Simple model
- เทคโนโลยี
    - Data warehouse

Big Data Analytics

- เน้น Predictive Analytics และ Prescriptive Analytics
- เน้น Data Science
- ลักษณะข้อมูล
    - ขนาดใหญ่
    - ข้อมูลดิบ
    - หลากหลายชนิด
    - Complex model
- เทคโนโลยี
    - Data Lake

การทำ Data Science

- บุคลากรที่จะทำด้านนี้ต้องมีทักษะสามด้านคือ คณิตศาสตร์, การเขียนโปรแกรม และความรู้เชิงธุรกิจในด้านนั้น
- การทำ Data Science ต้องเขียนโปรแกรมอย่างภาษา Python หรือ R
- ภาษา SQL ไม่สามารถที่จะมาเขียนโปรแกรท Machine Learning ได้ เน้นในการสอบถาม (Query) ข้อมูล

Source: Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.

Data Solution Life Cycle

Manager / Team Leader

- Problem Analysis
- Data Prep
- Evaluation / Monitoring

Data Engineer

- Data Prep
- Modeling

Developer

- Data Solutions
- Deployment
- Evaluation / Monitoring

Data Analyst

- Data Prep
- Modeling
- Data Analytics
- Data Solutions

Big Data Technology

HADOOP

Reference: http://blog.agroknow.com/wp-content/uploads/2015/07/Hadoop_Ecosystem.jpg

BIG DATA LANDSCAPE 2017

INFRASTRUCTURE

HADOOP ON-PREMISE
cloudera, Hortonworks, MAPR, Pivotal, IBM InfoSphere, bluedata, jethro

HADOOP IN THE CLOUD
amazon web services, Microsoft Azure, Google Cloud Platform, IBM InfoSphere BigInsights, TREASURE DATA, Qubole, altiscale, CAZENA, CenturyLink

STREAMING / IN-MEMORY
amazon web services, databricks, confluent, Striim, GridGain, METAMARKETS, DATATORRENT, dataArtisans, ORACLE Coherence, hazelcast, TERRACOTTA

NOSQL DATABASES
Google Cloud Platform, ORACLE, amazon DynamoDB, Microsoft Azure, MarkLogic, mongoDB, DATASTAX, AEROSPIKE, Couchbase, redislabs, influxdata

NEWSQL DATABASES
SAP HANA, Clustrix, Pivotal, NuoDB, Cockroach LABS, memsql, splice machine, MariaDB, VOLTDB, citusdata, Trafodion, deepdb, paradigm4

GRAPH DBS
neo4j, IBM, ORACLE, DIRAGRAPH, OrientDB, InfiniteGraph, Objectivity

MPP DBS
TERADATA, VERTICA, NETEZZA, action, kognitio, EXASOL, dremio

CLOUD EDW
amazon web services, Google Cloud Platform, Microsoft Azure, Pivotal, snowflake, Infoworks

DATA TRANSFORMATION
talend, pentaho, alteryx, TRIFACTA, tamr, Paxata, StreamSets, UNIFI

DATA INTEGRATION
informatica, MuleSoft, snaplogic, TEALIUM, Segment, enigma, podium data, alooma, xplenty, ZALONI, import.io, Stitch

DATA GOVERNANCE
informatica, IBM, skyhigh, collibra, Alation, Waterline Data

MGMT / MONITORING
amazon web services, New Relic, APPDYNAMICS, actifio, DATADOG, WAVEFRONT, unravel, splunk, Anodot, Prometheus, pagerduty, Numerify

STORAGE
amazon web services, Google Cloud Platform, Microsoft Azure, ALLUXIO, nimblestorage, Qumulo, COHO, panasas

CLUSTER SERVICES
amazon web services, kubernetes, docker, MESOSPHERE, CoreOS, pepperdata

APP DEV
Lightbend, Keen IO, rainforest, CASK

CROWDSOURCING
amazon mechanical turk, Upwork, WorkFusion, CrowdFlower

HARDWARE
Google TPU, ARM, nervana, Graphcore, MYTHIC, cerebras, NVIDIA, Movidius, SCORTEX

ANALYTICS

DATA ANALYST PLATFORMS
Microsoft, pentaho, alteryx, Digital Reasoning, GUAVUS, AYASDI, MATTIVO, Datameer, Quid, ClearStory, OrigamiLogic, inter|ana, Bottlenose, ARIMO, ENDOR, MODE

DATA SCIENCE PLATFORMS
IBM, KNIME, data iku, DOMINO, yhat, rapidminer, CONTINUUM ANALYTICS, Alpine, ALGORITHMIA, Angoss

BI PLATFORMS
Microsoft, DOMO, amazon web services, looker, Wave Analytics, ARCADIA DATA, atscale, GoodData, SISENSE

VISUALIZATION
tableau, SAP Lumira, Google Cloud Platform, Qlik, MAPD, celonis, Periscope, ZEPL, CHARTIO, plotly, GOOMDATA

VERTICAL ANALYTICS
PREDIX, IoT, CAPE, UPTAKE, Orbital Insight, TACHYUS, Alluvium, datorama

STATISTICAL COMPUTING
sas, SPSS, MATLAB

DATA SERVICES
Palantir, Mu Sigma, DATA SCIENCE, OPERA, DATA SCIENCE, Fractal, kaggle, EXL, DataKind, FF

MACHINE LEARNING
Azure Machine Learning, amazon web services, Google Cloud Platform, H2O.ai, DataRobot, context relevant, VISENZE, bonsai, DATARPM, deepsense.io, nutonian

HORIZONTAL AI
IBM Watson, Cortana, Face++, 旷视, sentient, Voyager Labs, vicarious, clarifai, Affectiva, SENSETIME, Cognitive Scale, Numenta, cronocom, PETUUM, nara, OSARO, THE CURIOUS COMPANY AI, BLUE VISION LABORATORIES

SPEECH & NLP
Google Cloud Platform, twilio, amazon alexa, NarrativeScience, Wolfram Alpha, semantic machines, MindMeld, NUANCE, ARRIA, iDIBON, TalkIQ, cortical.io, maluuba, snips, yseop, Gridspace, SoundHound Inc.

SEARCH
Autonomy, elastic, EXALEAD, ORACLE ENDECA, ThoughtSpot, Algolia, Lucidworks, swiftype, MAANA, alphasense, Searchink, SINEQUA

LOG ANALYTICS
splunk, sumologic, loggly, kibana, logz.io

SOCIAL ANALYTICS
Hootsuite, sprinklr, NETBASE, DATASIFT, tracx, synthesio, simplereach, bitly, predata

WEB / MOBILE / COMMERCE ANALYTICS
Google Analytics, mixpanel, AMPLITUDE, sumall, Airtable, retention, SIGOPT, granify, custora

APPLICATIONS – ENTERPRISE

SALES
einstein, CHORUS, INSIDESALES.COM, conversica, clari, AVISO, TACT, fuse|machines, TROOPS

MARKETING - B2B
RADIUS, App Annie, EVERSTRING, Lattice, infer, MINTIGO, 6sense, tubular, Reflektion, DataFox, NGAGIO

MARKETING - B2C
Zeta, bloomreach, blueyonder, PERSADO, ACTIONIQ, kahuna, BLUECORE, SAILTHRU, QUANTIFIND, mparticle, Amplero

CUSTOMER SERVICE
MEDALLIA, zendesk, CLARABRIDGE, Gainsight, ClickFox, NGDATA, DigitalGenius, appuri, automat, frame.ai, msg.ai, INTERCOM

HUMAN CAPITAL
HireVue, entelo, hiQ, GIGSTER, textio, RESTLESS BANDIT, Wade&Wendy, Clustree, Stella, pymetrics

LEGAL
RAVEL, Seal, Everlaw, JUDICATA, Brevia, ROSS, casetext

FINANCE
anaplan, uora, tidemark, S/4HANA, TRADESHIFT

ENTERPRISE PRODUCTIVITY
slack, facebook, X., ORACLE, lumiata, diffbot, Clara, talla, butter.ai, KASISTO

BACK OFFICE AUTOMATION
HyperScience, captricity, AppZen

SECURITY
TANIUM, CYLANCE, StackPath, DARKTRACE, illumio, CODE42, ThreatMetrix, DataGravity, VECTRA, CipherCloud, Guardian Analytics, cyberreason, ANOMALI, siftscience, SICNIFYD, SentinelOne, SecurityScorecard, BlueTalon, Recorded Future, feedzai, sqrrl, SOCURE, AREA 1 SECURITY, FORTSCALE, Keybase, sparkcognition

APPLICATIONS – INDUSTRY

ADVERTISING
AppNexus, MediaMath, criteo, xAd, rocketfuel, Integral Ad Science, OpenX, MOAT, theTradeDesk, Algorithms, dstillery, LiveIntent, drawbridge, gumgum, TAPAD, DataXu, databerries, Appier, DYNAMIC YIELD, Yieldmo

EDUCATION
KNEWTON, Clever, declara, kidaptive, PANORAMA, knowre, gradescope

GOVERNMENT
Socrata, OPENGOV, mark43, FiscalNote, OpenDataSoft

FINANCE - LENDING
OnDeck, Affirm, Kreditech, AVANT, zest finance, INSIKT, TALA, MoneyLion, TrueAccord, trooly, cignifi, aire, Active.AI

FINANCE - INVESTING
Dataminr, KENSHO, Quantopian, NUMERAI, iSENTIUM, claritymoney, ALGORIZ, AIDYA, RavenPack

REAL ESTATE
Opendoor, VTS, CREDIFI, reonomy, COMPSTAK

INSURANCE
Metromile, Lemonade, CYENCE, Shift Technology, Tractable

HEALTHCARE
FLATIRON, KYRUUS, HealthTap, METABIOTA, Ginger.io, Glow, BABYLON, COTA, zebra, ovia, AiCure, enlitic, Qventus, imagia, ARTERYS, BAYLABS, Kang Health, IMAGEN, freenome, DIAGNOS'S PROJECT

LIFE SCIENCES
23andMe, color, iCarbonX, zymergen, Benevolent AI, PATHWAY GENOMICS, ZEPHYR HEALTH, 3Scan, Clear Labs, Phosphorus, Citrine, twoXAR, deep genomics, Atomwise

TRANSPORTATION
UBER, ZOOX, TESLA, nuTonomy, CLEARPATH, nauto, drive.ai, pilot.ai, AUTOLIV, OTTO, OPTIMUS RIDE, nexar, Civil Maps, comma.ai, netradyne, NIO

AGRICULTURE
FARMERS BUSINESS NETWORK, FarmersEdge, FarmLogs, BLUE RIVER, mavrx, terrAvion, prospera

COMMERCE
instacart, STITCH FIX, RetailNext, HowGood

OTHER
eHarmony, stem, rethink robotics, AMPER MUSIC, SIGHT MACHINE, hopper, BOXEVER, celect, VERDIGRIS, duetto, Unbabel, Jukedeck, Second Spectrum

CROSS-INFRASTRUCTURE/ANALYTICS
amazon web services, Google Cloud Platform, Microsoft, IBM, SAP, Hewlett Packard Enterprise, sas, 1010data, vmware, TIBCO, TERADATA, ORACLE, NetApp

OPEN SOURCE

FRAMEWORK
hadoop HDFS, hadoop MapReduce, Apache Kylin, Flink, YARN, TEZ, Spark, MESOS, CDAP

QUERY / DATA FLOW
Spark SQL, presto, HIVE, SLAMDATA, APACHE DRILL, Google Cloud Dataflow

DATA ACCESS
cassandra, nifi, mongoDB, CouchDB, OPENTSDB, riak, HBASE, Spanner, accumulo

COORDINATION
talend, Apache Zookeeper, OOZIE, Apache Ambari

STREAMING
Spark, APEX, Flink, beam, kafka, druid, STORM

STAT TOOLS
python, R, ScalaLab, NumPy, SciPy

AI / MACHINE LEARNING / DEEP LEARNING
theano, torch, MAGnet, learn, TensorFlow, Caffe, CNTK, DMTK, Keras, PaddlePaddle, Apache SINGA, OpenAI, mxnet, VELES, WEKA, neon, FeatureFu, Chainer, DIMSUM, DSSTNE, mllib, DL4J, Intel nervana, Aerosolve

SEARCH
elasticsearch, Solr, lucene

LOG ANALYSIS
elasticsearch, kibana, logstash

VISUALIZATION
BEAKER, Rodeo

COLLABORATION
jupyter, Zeppelin, ANACONDA

SECURITY
Apache Ranger, KNOX, Sentry

DATA SOURCES & APIS

HEALTH
JAWBONE, VALIDIC, practice fusion, fitbit, GARMIN, Human API, kinsa

IOT
GE Digital, UPTAKE, ThingWorx, helium, samsara, AUGURY, estimote

FINANCIAL & ECONOMIC DATA
Bloomberg, THOMSON REUTERS, DOW JONES, S&P CAPITAL IQ, CB INSIGHTS, xignite, quandl, YODLEE, PREMISE, estimize, SECOND MEASURE, Eagle Alpha, StockTwits, PLAID, mattermark, Thinknum

AIR / SPACE / SEA
PLANET LABS, Airware, spire, SKYCATCH, AIROBOTICS, Descartes Labs, UNDERSTORY, TELLUSLABS, WINDWARD, DroneDeploy, MarineTraffic

PEOPLE / ENTITIES
acxiom, Experian, EPSILON, InsideView, Crimson Hexagon, BASIS TECHNOLOGY, quantcast, SAFEGRAPH

LOCATION INTELLIGENCE
FOURSQUARE, MapAnything, Sense, PlaceIQ, esri, factual, CARTO, Mapillary, STREETLINE

OTHER
qualtrics, DATA.GOV, data.world, panjiva, enigma

DATA RESOURCES

INCUBATORS & SCHOOLS
PLURALSIGHT, GA, galvanize, DataCamp, DataElite, INSIGHT, The Data Incubator, METIS

RESEARCH
facebook research, OpenAI, MIRI, CSAIL, DFKI, AI2 ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Big Data Architecture

Reference: www.openbi.com/data-warehouse-architecture

BIG DATA STORYMAP
EMC²
CURRENT STATE
BI/DW CHALLENGES
• Batch oriented, inflexible, high latency
• Brittle & labor intensive
• Performance & scalability challenged
• Aggregated, silo'd, structured data
BUSINESS CHALLENGES
• Rigid architecture impedes immediately available business opportunities
• Retrospective reporting doesn't guide business decisions
• Social, mobile, or machine-generated data insights not available
• Aggregation destroys valuable insights
FUTURE STATE
• Big Data analytics continuously drive business innovations
• Real-time insights optimize operations and boost profits
• Unprecedented customer insights improve products and user experience
TRADITIONAL TOWN
DATA WAREHOUSE
ETL
BUSINESS INTELLIGENCE
BUSINESS-MODEL-MATURATION
BUSINESS MONITORING
BUSINESS INSIGHTS
BUSINESS OPTIMIZATION
DATA MONETIZATION
BUSINESS METAMORPHOSIS
Value Creation City
BIG DATA VALUE CHAIN "MILE"
PRODUCT DEVELOPMENT
PROCUREMENT
FINANCE
MANUFACTURING
SALES & MARKETING
DISTRIBUTION
CUSTOMER EXPERIENCE
Continuously Refine Analytic Models
Publish Insights in Apps
Operationalize Big Data
Acquire & Prepare Data
Capture new Business Requirements
Measure Decision Effectiveness
Implement Big Data Infrastructure
Incorporate Insights into apps
Define data strategy
Build analytic models
Identify targeted business initiatives
Determine required insights
VALUE JOURNEY
CONDUCT DATA SCIENTIST TRAINING & CERTIFICATION
VOLUME
VARIETY
VELOCITY
Records Terabytes
Transactions Petabytes
Tables, Files
Structured
Unstructured
Semi-Structured
Real-Time
Near Real-Time
Batch
DEEP DATA
DIVERSE DATA
FAST DATA
BIG DATA RIVER
Risk of Not Preparing for BIG DATA...
• Competitors Innovate Faster
• Profits & Margins Degenerate
• Missed Business Opportunities
• Market Share Declines
Big Data concepts by William Schmarzo
Visualization concept by Mark A. Lawson and Glenn Steinhandler
© EMC Corp. Created by The Grove Consultants International

การลงทุนด้านเทคโนโลยี

- การทำ Big Data Analytics คือการลงทุนด้าน Data Lake
- การลงทุนด้าน Data Warehouse จะเก็บได้เฉพาะข้อมูล Structure และเน้นการทำ BI
- งบประมาณด้านเทคโนโลยีจะหมดไปเก็บ Data storage และ Data Processing
- Hadoop คือตัวอย่างของเทคโนโลยีที่เหมาะสมในการทำ Data Lake
- Big data as a service บน Public cloud ช่วยลดค่าใช้จ่ายได้มาก

Digital Age

Digitalize of Everything
needle & thread
Casper
airbnb
HOUSING
POSTMATES
DELIVEROO
grubHub
Blue Apron
amazon fresh
nest
Spotify
Google play
amazon instant video
HBO NOW
NETFLIX
SnapScan
instamojo
NUMBER26
bitcoin
mint
WATCH
fitbit
Tencent
slack
box
Dropbox
Palantir
ZENEFITS
Square
BOSCH
FANUC
Rolls-Royce
carma
UBER
waze
lyft
快
OLA

Intel CPU Trends
(sources: Intel, Wikipedia, K. Olukotun)
Dual-Core Itanium 2
Pentium 4
Pentium
386
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
0
1970
1975
1980
1985
1990
1995
2000
2005
2010
Transistors (000)
Clock Speed (MHz)
Power (W)
Perf/Clock (ILP)

MOORE'S LAW

50 years of electronic devices

Exponential growth of Technology products

Digital Banking มา “สาขา” ธนาคารในไทยจะหายไป

นับจากนี้ “ธนาคารไทย” กำลังจะมีการเปลี่ยนแปลงครั้งใหญ่เกิดขึ้นเพราะ “ดิจิทัล แบงก์กิ้ง” ได้เข้ามามีบทบาทสำคัญ ในการผลิกโฉมรูปแบบบริการของ “ธนาคาร” ให้ต่างไปจากเดิม ผู้บริโภคไม่จำเป็นต้องเดินไปที่ “สาขา” ก็สามารถทำธุรกรรมทางการเงินได้จากทุกที่ ความสะดวกสบายนี้เอง ส่งผลให้จำนวนผู้ใช้บริการ “ดิจิทัล แบงก์กิ้ง” ไม่ว่าจะเป็น “อินเทอร์เน็ต แบงก์กิ้ง” และ “โมบาย แบงก์กิ้ง” เติบโตแบบฉุดไม่อยู่สะท้อนจากข้อมูลของธนาคารแห่งประเทศไทย ที่ระบุว่า ปัจจุบัน “อินเทอร์เน็ต แบงก์กิ้ง” มีผู้ใช้งานเกือบ 13 ล้านคน เติบโตเฉลี่ย 20% ต่อปี ในขณะที่ “โมบาย แบงก์กิ้ง” มีผู้ใช้งานเกือบ 12 ล้านคน เติบโตถึง 50% ซึ่งเป็นผลจากการเข้ามาของ 4G และการพัฒนาของสมาร์ทโฟนรุ่นต่างๆ ได้เข้ามากระตุ้นการใช้งานของผู้บริโภคให้เพิ่มขึ้น โดยเฉพาะกลุ่มคนรุ่นใหม่ มี “คนไทย” ใช้บริการ “Digital Banking” เท่าไหร่ ?

SCB เตรียมลดสาขาเหลือ 400 แห่ง ในปี 63 พนักงานหายวูบกว่า 10,000 คน

22 มกราคม พ.ศ. 2561 เวลา 16:00 น.

G+ แชร์ LINE it! Tweet แชร์ 396

SCB เผยตั้งเป้าลดสาขาภายในปี 63 เหลือ 400 แห่ง ลดพนักงานเหลือ 15,000 คน

22 ม.ค. 61 - นายอาทิตย์ นันทวิทยา กรรมการผู้จัดการใหญ่และประธานเจ้าหน้าที่บริหาร ธนาคารไทยพาณิชย์ จำกัด (มหาชน) (SCB) เปิดเผยว่า ธนาคารตั้งเป้าภายในปี 63 จะลดสาขาลงเหลือ 400 แห่ง จากปัจจุบันอยู่ที่ 1,153 แห่ง เนื่องจากปริมาณธุรกรรมและการใช้บริการของลูกค้าลดลง ทั้งนี้ จะมีการปรับเปลี่ยน และโยกย้ายพนักงานที่ทำงานในสาขาที่ปิดมาอยู่ศูนย์อบรม (SCB Academy) รวมถึงลดพนักงานเหลือ 15,000 คน จากปัจจุบันอยู่ 27,000 คน ซึ่งปีที่ผ่านมาพนักงานลาออกเฉลี่ยที่ 2,000-3000 คน โดยทางธนาคารไม่มีนโยบายปลดพนักงานออก

อย่างไรก็ตาม การปรับตัวของธนาคารในช่วง 2 ปีที่ผ่านมา ได้มีการลงทุนด้านไอที 40,000 ล้านบาท โดยเป็นการทยอยลงทุน 5 ปี (59-63) ได้ส่งผลกระทบต่อผลการดำเนินงานในปีที่ผ่านมาประมาณ 4,000 ล้านบาท ซึ่งในปีนี้เป็นต้นไปจนถึงปี 63 มองว่าจะกระทบ

Round UP

SCB ประกาศวิสัยทัศน์ 2020 ปรับรูปแบบธุรกิจ เดินหน้าดิจิทัลเต็มตัว คาดลดสาขาเหลือ400สาขาใน3ปี

วันที่ 22 มกราคม 2561 - 11:44 น.

4.1K SHARES Facebook 4.1K Twitter G+ Google+ LINE

นายอาทิตย์ นันทวิทยา กรรมการผู้จัดการใหญ่ และประธานเจ้าหน้าที่บริหาร ธนาคารไทยพาณิชย์ แถลงข่าว "2020 SCB VISION" ว่า นอกจากการปรับเปลี่ยนรูปแบบการทำธุรกิจของไทยพาณิชย์ให้ไปสู่ดิจิทัลมากขึ้น ทั้งการนำดิจิทัลเข้ามาให้บริการทางการเงินแทน และการใช้ดิจิทัลเข้ามาวิเคราะห์ในการทำธุรกิจในอนาคตแล้ว ภายใน 3 ปีข้างหน้า จะเห็นการเปลี่ยนแปลงของธนาคารค่อนข้างมาก คือการลดลงของสาขาอย่างมีนัยสำคัญ โดยคาดว่าจำนวนสาขาจะลดลงเหลือ 400 สาขา จากปัจจุบันที่มีสาขาอยู่ที่ 1,153 สาขา

ขณะเดียวกัน ธนาคารตั้งเป้าลดจำนวนพนักงานให้เหลือ 15,000 คน จากปัจจุบันที่มีพนักงานธนาคารทั่วประเทศ 27,000 คน

2016, Starbucks corp. hold $1.2 Billion

in Starbucks card deposit money

STARBUCKS COFFEE®

1.2
Billion US

Deposit in Starbucks Card

Credit: http://fortune.com/2016/06/10/starbucks-card-balance/
Credit:
http://www.marketwatch.com/story/starbucks-has-more-customer-money-on-cards-than-many-banks-have-in-deposits-2016-06-09
Credit: https://www.starbuckscard.in.th/cards/

venmo

PayPal™

Pay

GObank

Moven

spend, save & live smarter

Square

VISA

DISCOVER NETWORK

MasterCard

SIMPLE

Google wallet

AMERICAN EXPRESS

เกมพลิก เมื่อธนาคาร 'ฟรีค่าธรรมเนียม'

หุ้นส่วนประเทศไทย

ดร.กมล เกียรติเรืองอกมลา
รองอธิการบดี
และคณะไม่ได้รับรองผลการสำรวจนี้ เพื่อวัตถุประสงค์ทางการค้าพระใจ

ธนาคารพาณิชย์ในประเทศไทยได้มี การประกาศยกเว้น หรือฟรีค่าธรรมเนียมการชำระเงินที่ผ่านระบบอิเล็กทรอนิกส์ ซึ่งเรื่องนี้ถือเป็นเรื่องใหญ่มาก เพราะเปรียบเสมือนการปฏิวัติการทำธุรกรรม **ในรอบสิบกว่าปีที่ผ่านมา** เนื่องจากที่ผ่านมารายได้ของธนาคารจำนวนไม่น้อยหรืออาจจะมากถึงเกือบ 1 ใน 4 ในบางธนาคาร เป็นรายได้ที่มาจากค่าธรรมเนียมการให้บริการผ่านระบบอิเล็กทรอนิกส์ต่างๆ โดยเฉพาะตู้ ATM, Internet Banking หรือ Mobile Banking ดังนั้นเมื่อธนาคารประกาศยกเลิกค่าธรรมเนียมจึงย่อมส่งผลกระทบโดยตรงต่อรายได้ของธนาคารไม่มากก็น้อย ...**แต่ทำไมธนาคารถึงยอมขาดทุนด้วยการยกเลิกการเก็บค่าธรรมเนียมดังกล่าว สาเหตุคงจะมาจากกระแสของคลื่นสามระลอกใหญ่ ที่ธนาคารไม่สามารถต้านทานได้**

คลื่นลูกที่หนึ่ง... พฤติกรรมที่เปลี่ยนแปลงไปของผู้บริโภค : พฤติกรรมของคนรุ่นใหม่ ที่ส่วนใหญ่เป็นคนในยุค Gen X หรือต่ำกว่า Gen X ลงมานั้น มีความต้องการอะไรที่ง่ายสะดวก รวดเร็ว และเสพติดการใช้สื่อโซเชียลเป็นหลัก มีการใช้งานแอพพลิเคชั่นต่างๆ บนโทรศัพท์มือถือเป็นสำคัญ และเปิดใจยอมรับการใช้งานเทคโนโลยีสมัยใหม่ได้ง่ายกว่า จึงเป็นสาเหตุสำคัญที่ทำให้หลายองค์กรธุรกิจรวมถึงธนาคาร จำเป็นจะต้องมีการปรับเปลี่ยนการเข้าถึงผู้บริโภคให้มีความรวดเร็วและสะดวกเพิ่มขึ้น ผ่านการเข้าถึงในรูปแบบโทรศัพท์มือถือ

คลื่นลูกที่สอง... Disruptive Technology ทางด้านการเงิน หรือ FinTech (Financial Technology) : ปัจจุบันนี้หลายท่านอาจจะเคยใช้บัตรสตาร์บัคส์การ์ด (Starbucks Card) เพื่อซื้อสินค้าในร้านสตาร์บัคส์ ซึ่งบัตรสตาร์บัคส์การ์ดนี้ เคยมีการลงข่าวในประเทศสหรัฐอเมริกาว่ามีเงินที่สะสมอยู่ในบัตรเพื่อใช้ซื้อสินค้าจากทุกๆ คนรวมกัน เทียบเท่ากับเงินฝากในธนาคารชั้นนำบางแห่งของสหรัฐอเมริกา และมีบางบริษัทที่มีธุรกรรมทางการเงินบางอย่างไม่แตกต่างจากธนาคารพาณิชย์เลย **แต่บริษัทเหล่านี้กลับไม่จำเป็นจะต้องดำเนินการภายใต้การควบคุมของธนาคารกลาง และไม่จำเป็นต้องมีสาขาจำนวนมากแบบธนาคารพาณิชย์ แต่ก็สามารถให้บริการทางการเงินที่มีมูลค่างวงเงินเป็นจำนวนมากได้ ซึ่งบริษัทที่ให้บริการในลักษณะนี้จะถูกเรียกว่า Non Bank**

บางภาพประกอบ กราฟฟิก

โดยในปัจจุบันนี้ จะมีบริษัท หรือ Startup ที่เป็น Non Bank ให้บริการทางการเงินอยู่เป็นจำนวนมาก อาทิ PayPal, Alipay, Apple Pay, Sumsung Pay, อื่นๆ โดย Non Bank เหล่านี้ สามารถติดต่อกับผู้ใช้บริการเพื่อทำธุรกรรมทางการเงินต่างๆ ผ่านการใช้เทคโนโลยี FinTech หรือ Mobile Application เป็นหลัก ธนาคารธุรกรรมการเงินส่วนใหญ่จะไม่ได้อยู่ที่ธนาคารสาขาแล้ว แต่จะย้ายมาเป็นการให้บริการบนโทรศัพท์มือถือ อีกทั้งพฤติกรรมผู้บริโภคมีความคุ้นเคยการใช้เทคโนโลยีอยู่แล้ว ทำให้ธนาคารพาณิชย์เร่งพัฒนาการให้บริการธนาคารผ่าน Mobile Application แทน และหันไปลดจำนวนธนาคารสาขาลง

คลื่นลูกที่สาม... PromptPay : จากนโยบายของธนาคารแห่งประเทศไทยที่ต้องการจะผลักดันให้ประเทศไทยเป็นสังคมไร้เงินสด (Cashless Society) จึงได้ผลักดันการโอนเงิน PromptPay ซึ่งกำหนดให้การโอนเงินที่มีมูลค่าไม่เกิน 5,000 บาท (ซึ่งเป็นธุรกรรม 95% ของธุรกรรมทั้งหมด) ไม่มีการคิดค่าธรรมเนียม จึงเป็นการเร่งผลักดันทางอ้อมให้ธนาคารพาณิชย์จำต้องลดหรือยกเลิกการเก็บค่าธรรมเนียมลงนั่นเอง

จากคลื่นทั้งสามลูกที่พัดต่อเนื่องกันมา ส่งผลให้ธนาคารพาณิชย์ต้องเร่งปรับตัวอย่างเร่งด่วน หากต้องการจะรอดพ้นจากการถูกทัดพาของคลื่นทั้งสามลูก ด้วยการนำเทคโนโลยีทางการเงิน หรือ FinTech มาให้บริการแก่ลูกค้า ซึ่งเทคโนโลยี FinTech ที่สำคัญสำหรับธนาคาร คือ **การทำตัวเองให้เป็นแพลตฟอร์ม** และนำธนาคารมาอยู่บนโทรศัพท์มือถือ (Mobile Application) รวมถึงการปรับปรุงกระบวนการทำงาน **ของธนาคารให้ผอม (Lean) มากที่สุด** เพื่อให้ธนาคารเกิดความคล่องตัวหรือมีความยืดหยุ่นในการรับผลกระทบที่จะเกิดขึ้นในอนาคต

ที่ผ่านมาธนาคารมีการแข่งขันกันอย่างรุนแรงเพื่อแย่งชิงลูกค้าให้มาใช้บริการ Mobile Application ของตนเอง ซึ่งธนาคารใดหากมีผู้ใช้จำนวนมาก ย่อมจะมีความได้เปรียบในการแข่งขันในอนาคต และทำให้ธนาคารมีความคุ้มค่าในการลงทุนเพื่อพัฒนาระบบ Mobile Application นั้นๆ จึงไม่เป็นที่แปลกใจหากธนาคารจะนำกลยุทธ์ที่มีประสิทธิภาพที่สุดในหลัก 4P ซึ่งก็คือ กลยุทธ์ด้านราคา (Price) มาใช้ในการดึงดูดลูกค้าให้มาใช้บริการ Mobile Application ของตนเอง ด้วยการยกเว้นค่าธรรมเนียม เพราะอย่างไรเสียธนาคารก็ไม่สามารถเรียกเก็บค่าธรรมเนียมได้อยู่แล้ว เนื่องจากลูกค้ามีทางเลือกในการโอนผ่าน PromptPay แทนได้

อย่างไรก็ตาม การฟรีค่าธรรมเนียมย่อมส่งผลให้ในระยะสั้นธนาคารจะต้องสูญเสียรายได้อย่างหลีกเลี่ยงไม่ได้ แต่ในระยะยาวแล้วนั้นหากธนาคารมีลูกค้าที่มาใช้บริการ Mobile Application ของตนเองในระดับที่มากพอแล้ว ย่อมสามารถนำมาใช้เป็นอำนาจต่อรอง เพื่อเรียกเก็บค่าธรรมเนียมจากผู้ขายสินค้า แทนที่จะเรียกเก็บจากฝั่งผู้ใช้บริการดังเช่นที่ผ่านมาได้ ซึ่งรูปแบบนี้จะเป็น Business Model ที่จะเกิดขึ้นในอนาคต และช่วยให้ธนาคารลดการขาดทุนลงได้

การที่ธนาคารยกเลิกการเก็บค่าธรรมเนียม ยังกระทบต่อธุรกิจรับชำระค่าบริการอื่นๆ อาทิ เคาน์เตอร์เซอร์วิส หรือตู้เติมเงินบุญเติม เป็นต้น ที่จะต้องได้รับผลกระทบโดยตรงจากการสูญเสียรายได้และสูญเสียลูกค้าให้แก่ธนาคารที่มีการยกเว้นค่าธรรมเนียม แต่หากมองให้รอบด้านการยกเว้นค่าธรรมเนียมมิเพียงแต่เกิดผลด้านลบในวงกว้างเท่านั้น แต่ยังคงมีผลด้านบวกด้วยเช่นเดียวกัน **ดังนั้นการ "ฟรีค่าธรรมเนียม" จึงถือเป็นจุดเปลี่ยนที่สำคัญของวงการธุรกิจในหลายมิติ**

สรุป**การฟรีค่าธรรมเนียม มีทั้งผลกระทบโดยตรงทั้งด้านบวกและลบต่อธนาคารและธุรกิจอื่น** โดยเฉพาะการสูญเสียรายได้เป็นจำนวนไม่น้อยของธนาคาร และการสูญเสียรายได้ของธุรกิจที่เป็นตัวกลางในการชำระสินค้า ซึ่งจะส่งผลให้ธุรกิจธนาคารและธุรกิจต่างๆ จะต้องมีการปรับลดค่าใช้จ่ายลง และค่าใช้จ่ายด้านบุคลากรที่มักจะเป็นค่าใช้จ่ายในลำดับต้นๆ ขององค์กร ย่อมต้องได้รับผลกระทบจากการปรับลด **จนอาจส่งผลให้มีการชะลอหรือลดการจ้างพนักงานลง**

อีกทั้งผลกระทบทางอ้อม จะส่งผลให้เกิดการแข่งขันอย่างรุนแรงทั้งธุรกิจ Bank และ Non Bank อีกทั้งการใช้ราคาเป็นกลยุทธ์หลัก ย่อมเป็นการกีดกันผู้แข่งขันรายใหม่ หรือทำให้คู่แข่งที่ยังไม่มีความแข็งแรงเพียงพอต้องถอนตัวออกไปจากการแข่งขันที่รุนแรงนี้ อย่างไรเสียในมุมของผู้บริโภค **การฟรีค่าธรรมเนียม และการแข่งขันที่รุนแรงย่อมเป็นประโยชน์โดยตรงต่อผู้บริโภคอย่างแท้จริง** ที่ช่วยให้ผู้บริโภคสามารถได้ใช้ของดีในราคาถูก ทั้งยังเป็นการช่วยผลักดันให้คนหันมาสนใจการใช้บริการธนาคารผ่าน Mobile App กันมากขึ้น และขับเคลื่อนให้ประเทศไทยสามารถก้าวสู่สังคมไร้เงินสดได้ในท้ายที่สุด ■

ค่าธรรมเนียมธนาคาร

ค่าธรรมเนียม	ไทยพาณิชย์ เดิม	ไทยพาณิชย์ ใหม่	กสิกรไทย เดิม	กสิกรไทย ใหม่	กรุงไทย เดิม	กรุงไทย ใหม่	กรุงเทพ เดิม	กรุงเทพ ใหม่	ทหารไทย	ออมสิน	กรุงศรี
โอนเงินต่างธนาคาร	12-50 บาท ต่อครั้ง*	Free เฉพาะช่องทางออนไลน์ ถึง 31.12.61	12-120 บาท ต่อครั้ง*	Free เฉพาะช่องทางออนไลน์ ถึง 31.12.61	25-120 บาท ต่อครั้ง*	Free เฉพาะช่องทางออนไลน์ ถึง 31.12.61	25-120 บาท ต่อครั้ง*	Free เฉพาะช่องทางออนไลน์ และกรุงไทย ถึง 1.6.61	25-120 บาท ต่อครั้ง*	60-120 บาท/ครั้ง*	60-120 บาท/ครั้ง*
โอนเงินข้ามเขต	Free-10 บาท ต่อครั้ง*		Free-10 บาท ต่อครั้ง*		Free-10 บาท ต่อครั้ง*		Free-10 บาท ต่อครั้ง*		25-50 บาท ต่อครั้ง*	10 บาท, 10,000 บาท**	Free, 30 บาท/ครั้ง*
จ่ายบิล	ขึ้นอยู่กับข้อตกลง**		ขึ้นอยู่กับข้อตกลง**		ขึ้นอยู่กับข้อตกลง**		ขึ้นอยู่กับข้อตกลง**		ขึ้นอยู่กับข้อตกลง**	ขึ้นอยู่กับข้อตกลง**	ขึ้นอยู่กับข้อตกลง**
ถอนเงิน ATM	เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท		เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท		เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท		เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท		เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท	เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท	เฉพาะเดือนละ ฟรี ค่าบริการ 15 บาท
ค่าบัตร ATM	รายปี 200-1,500 บาท***		รายปี 200-800 บาท***		รายปี 200-1,500 บาท***		รายปี 200-600 บาท***		รายปี 200 บาทฟรีถ้าใช้ 200 บาท ฟรีตลอดชีพ ถ้าใช้ 2,000 บาท***	รายปี 20-500 บาท***	รายปี 200-3,500 บาท***

*ขึ้นอยู่กับช่องทางและจำนวนเงิน **ขึ้นอยู่กับข้อตกลงระหว่างหน่วยงานหรือร้านค้า ***ขึ้นอยู่กับการทำรายการหรือค่าสมาชิก

Start-up Era

Grab
UBER
EATS
SKOOTAR
send
ranger
LALAMOVE
deliver faster
Ginja

Disruptive Technology

Thailand4.0

Thailand 1.0
Agriculture

Thailand 2.0
Light industry

Thailand 3.0
Heavy industry
Middle income Trap
Inequality & Imbalance

Thailand 4.0
Digital Economy
Middle income Trap
Inequality & Imbalance

Not change...

Not Survive

You 4.0 ???
CRMA 4.0 ???