

Big Data Analytics

Varude Boonprayong

Ph.D. (candidate) ICT for Education, KMUTNB

Boromarajonani College of Nursing Chakriraj

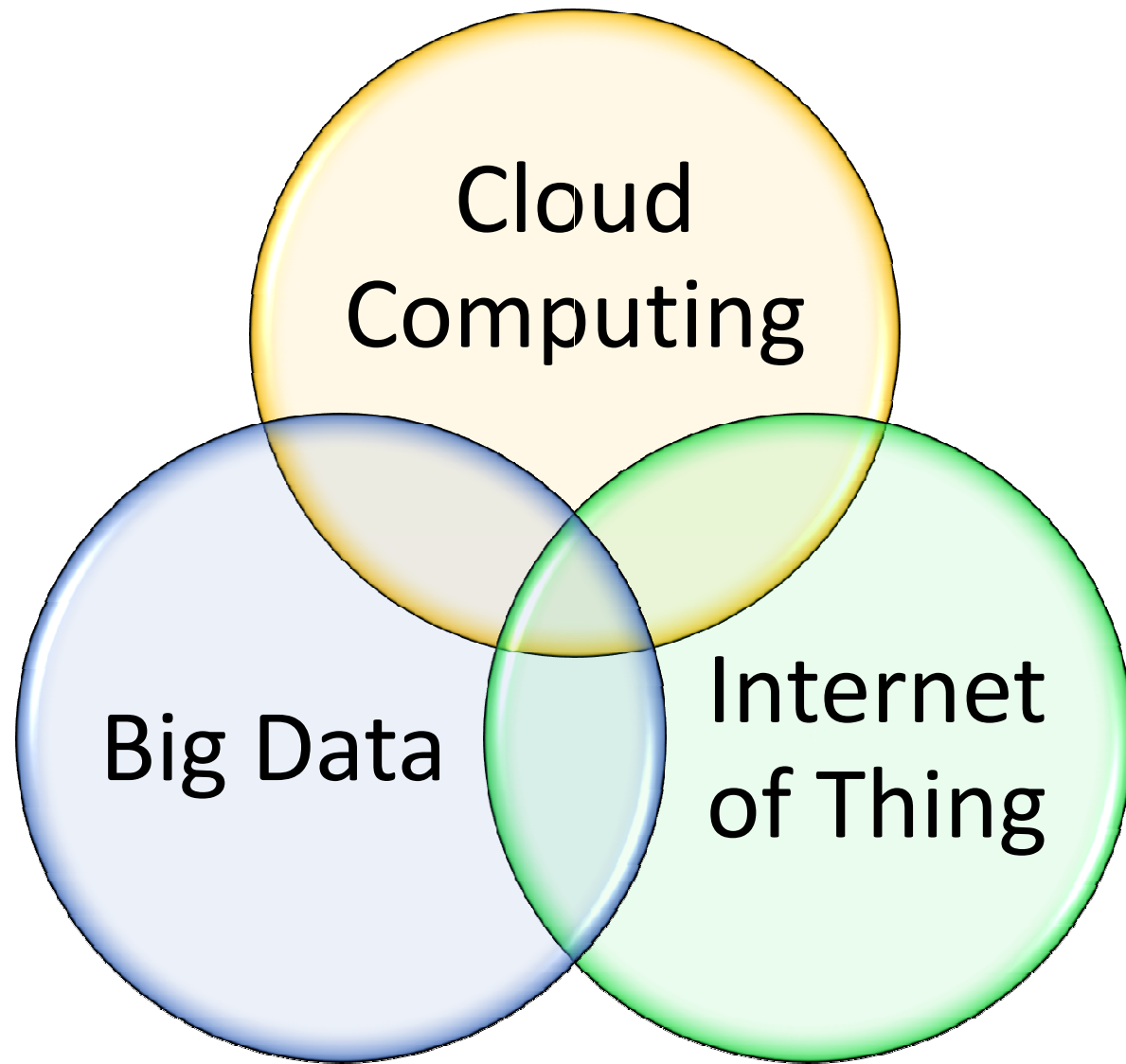
081-1655646

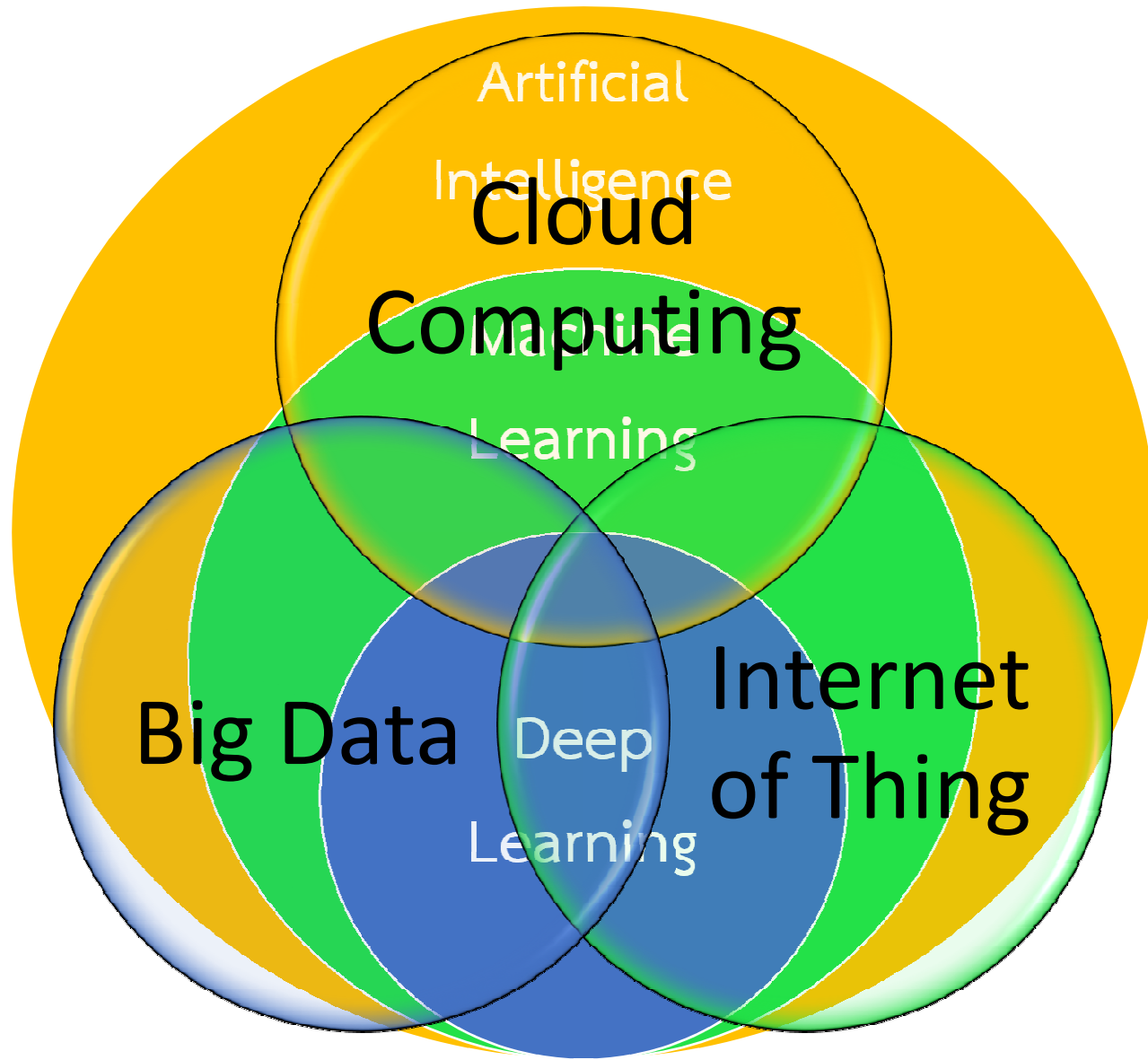
The Kahoot! logo background is a square divided into four quadrants: top-left is red, top-right is blue, bottom-left is yellow, and bottom-right is green. The word "Kahoot!" is written in white, bold, sans-serif font across the center.

Kahoot!



Mentimeter





Cloud



SAAS
Software
as a Service



CONSUME



PAAS
Platform
as a Service



BUILD ON IT



IAAS
Infrastructure
as a Service



MIGRATE TO IT

Internet of Things Connecting the World



Big Data



จะเริ่มค้นหา Big Data Project ต้องหาอย่างไร

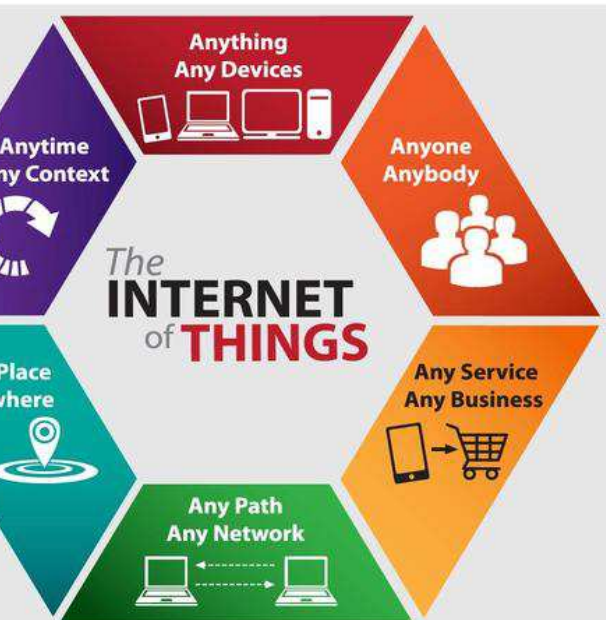
สำรวจแต่ละส่วนขององค์กร ว่ามีส่วนใดบ้างที่ยัง Manual หรือเกิดการตัดสินใจโดย “มนุษย์” อยู่บ้าง

สำรวจดูว่าเรามีข้อมูลอยู่ตรงไหนอย่างไรบ้าง หรือจะมีแนวทางในการเก็บข้อมูลเพิ่มเติมจากส่วนใดได้บ้าง

หา Technology ใหม่ ๆ

เปลี่ยน Mindset หรือแนวคิดในการทำงาน

Where Does Big Data Come From?



Technical Selection

Locate

- On-Premise Cluster
- Cloud-Based Cluster

Node

- Single Node
- Multiple Node

Data

- Structured
- Unstructured
- Semi structured

A Good Big Data

- Transactional data : no summary data
- Not Data Center
- Must predict and analyst
- real-time Big Data Analytics + IoT = value Creation

การจัดการ Big Data



การ Clean Data สำคัญไฉน?

ขั้นตอนการ Clean Data



Duplicate Elimination (ลบข้อมูลซ้ำซ้อน)



Standardizing (การทำให้อยู่รูปแบบเดียวกัน)



Correct Data(การแก้ไข)



Parsing(การแจกแจง)

Big Data Analytics

Data Sources

Database 

Data Warehouse 

ERP 

CRM 

Audio, Video 

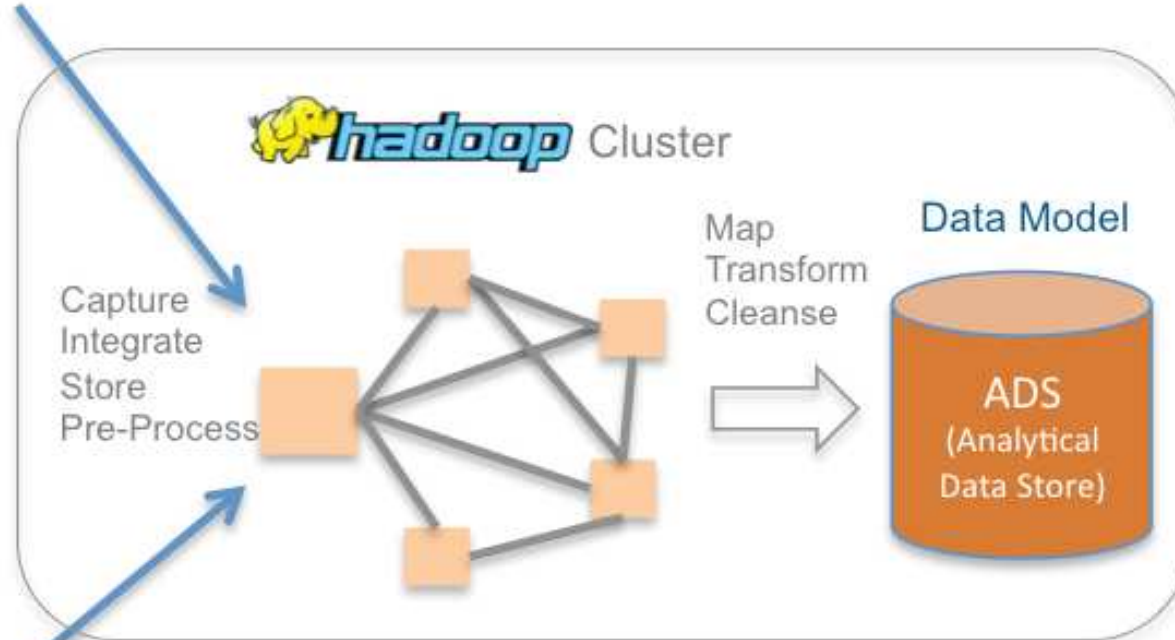
Clickstream 

Weblogs, Documents, Text 

Excel, CSV 


Social Media 

Data Platform



Analytics

Reports, Scorecards 

Dashboards 

Continuous
Real-Time
Historical

Ad-Hoc Query 

Structured

Unstructured

Advantage

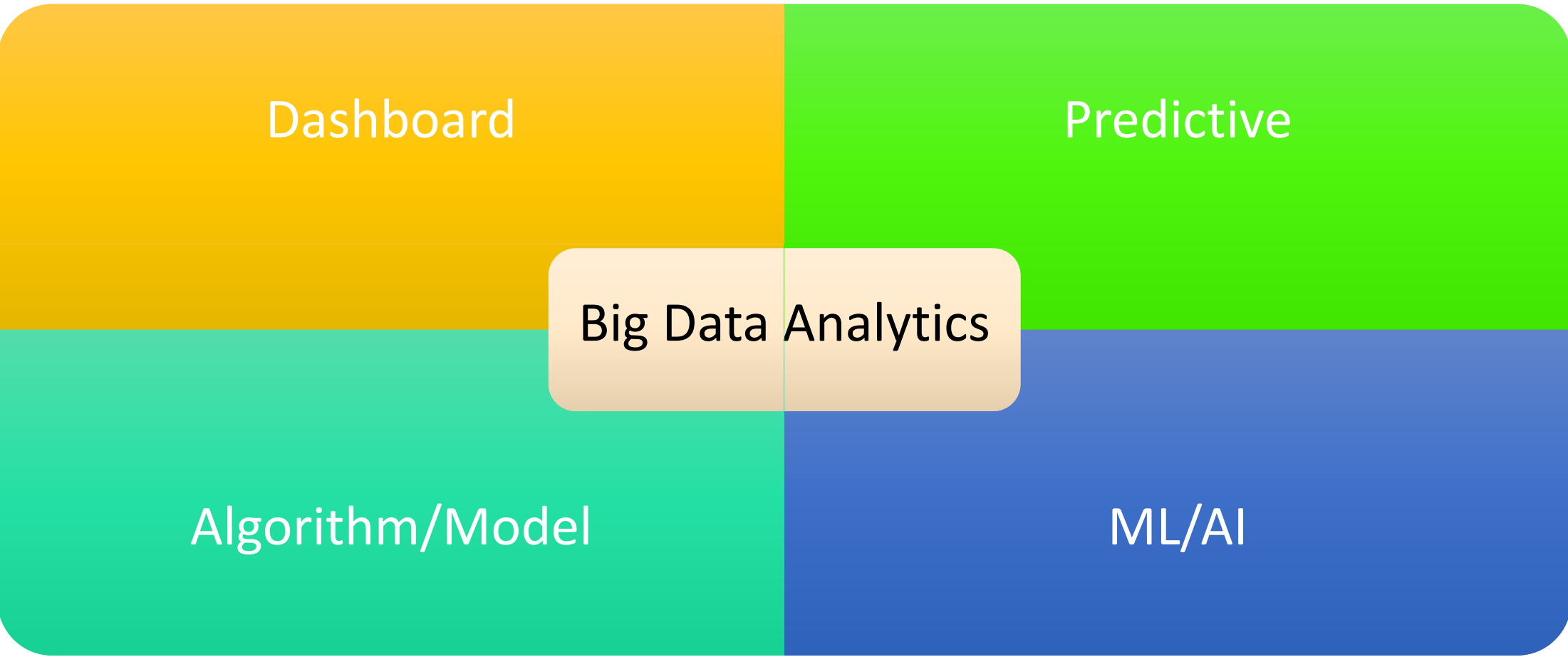
Dashboard

Predictive

Big Data Analytics

Algorithm/Model

ML/AI



Setup for Big Data

- OS : Ubuntu, CentOS, other Linux
- JAVA
- Configure SSH
- Hadoop
- Start the Cluster (name node, HDFS Services, YARN Services, Hadoop services)
- Visualization : data studio, Power BI, Qlick View, Tableau, etc.

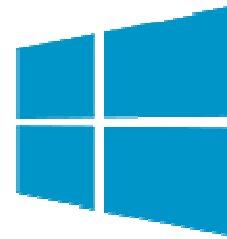


OPEN SOURCE
BIG DATA TOOLS



ORACLE®

IBM



Microsoft
Azure



Google Cloud Platform

Hosting +
Compute

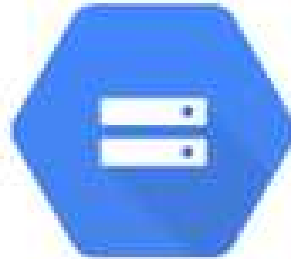


App Engine



Compute Engine

Storage



Cloud Storage



Cloud Datastore



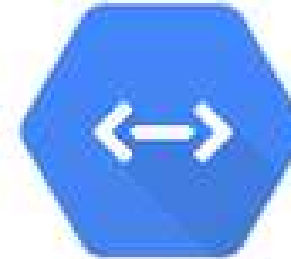
Cloud SQL

Big Data



BigQuery

Services



Cloud Endpoints



Translate API



Prediction API



Mar 14, 2018 - Apr 12, 2018

Account Overview

Earnings

0.98

-70.2%

Page views

8,886

10.0%

Page views RPM

0.11

-72.9%

Ad impressions

1,690

5.1%

Ad clicks

6

-53.8%

CPC

0.16

-35.5%

Page views CTR (%)

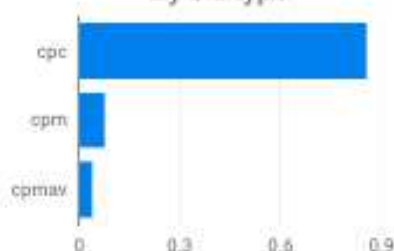
0.07%

-58.1%

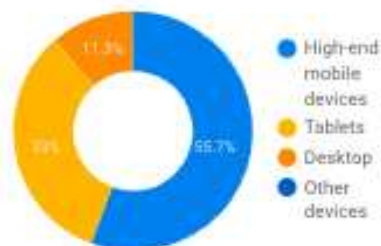


Earnings splits

By bid type



By platform



By country



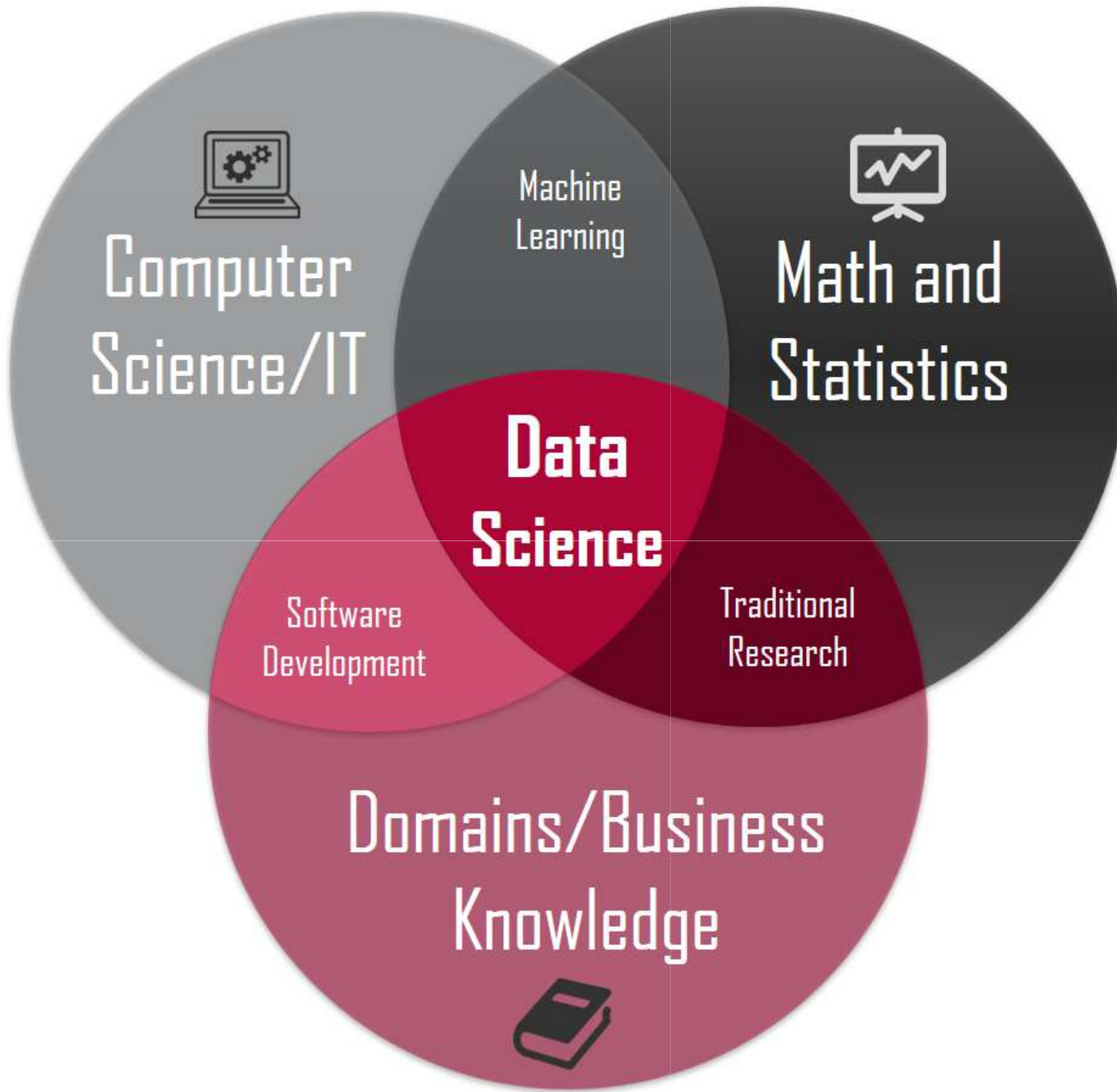
| Unit name | Ad clicks | Earnings | CPC | Ad impressions | Matched ad request | Page views |
|----------------------------|-----------|----------|------|----------------|--------------------|------------|
| Custom Square | 2 | 0.09 | 0.05 | 1,151 | 1,188 | 735 |
| Right-Responsive | 1 | 0.37 | 0.37 | 287 | 284 | 116 |
| Damos-Creative-1 | 0 | 0 | null | 9 | 5 | 12 |
| Leaderboard Preview Footer | 0 | 0 | null | 1 | 1 | 29 |

Template provided by:



1 - 4 / 4





มุมมองการใช้ข้อมูลของ 3 คน

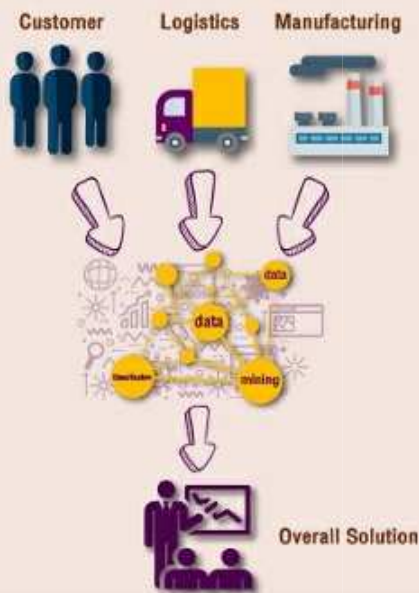
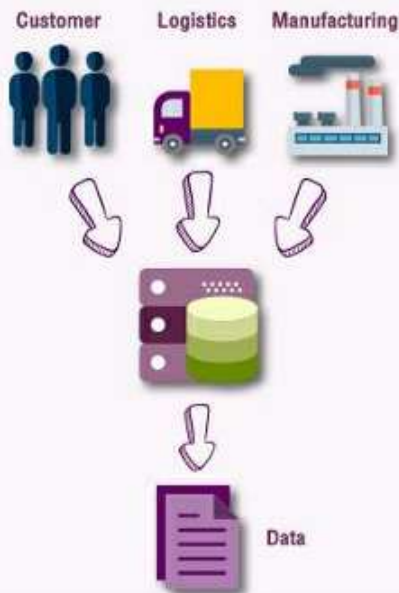
Data Engineer



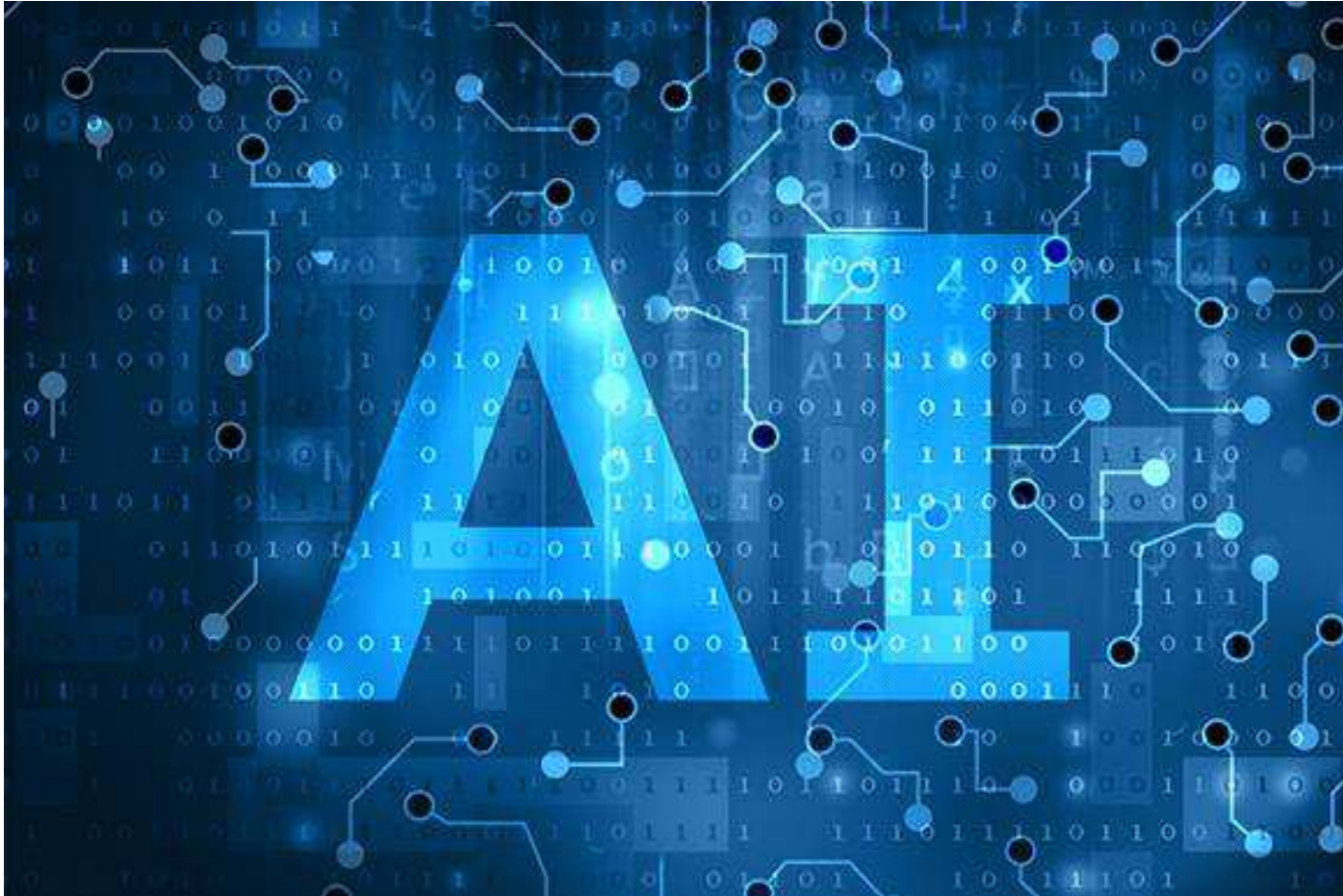
Data Scientist



Business and Data Analyst



แนวโน้มของ Big Data Project



<https://www.thirdsector.co.uk/caf-says-charities-ethical-guardians-ai->

สรุปขั้นตอนการดำเนินการ

Data

Google
cloud

BigQuery

Predictive
Dashboard

Case Study

คนที่ประชุมแห่งหนึ่งถ้านายกรัฐมนตรีอยากทราบว่า ทั้งปีมีคนที่มารับยา Paracetamol ก็เฝ้า
ระยะเวลาในการรอการตรวจแต่ละจุดก็น่าที่
แนวโน้มที่ผู้ป่วยจะเป็นคนใช้จิตเวชเป็นอย่างไร
คนที่เล่น Facebook มีความสัมพันธ์กับการป่วยอย่างไร
บุคลากรที่ไปช่วยราชการ การเลื่อนขึ้นเงินเดือน
เรื่องการลดค่าใช้จ่ายในการดำเนินงานของสถานบริการ

Heart of Human

Praboromarajchanok Institute
for Health Workforce Development