Data Science dan *Artificial Intelligence*: Konsep, Teori, Teknik, Tools, dan Aplikasi

Oleh. Prof. Zainal A. Hasibuan, PhD
Ketua APTIKOM
Universitas Dian Nuswantoro
Dewan TIK Nasional (WANTIKNAS)

Materi ini disampaikan pada acara *Training for Trainer Microcredential Certification* Bidang *Data Science* dan *Artificial Intelligence*, Direktorat Jenderal Pendidikan Tinggi, 22-26 Februari, 2021

Daftar Isi

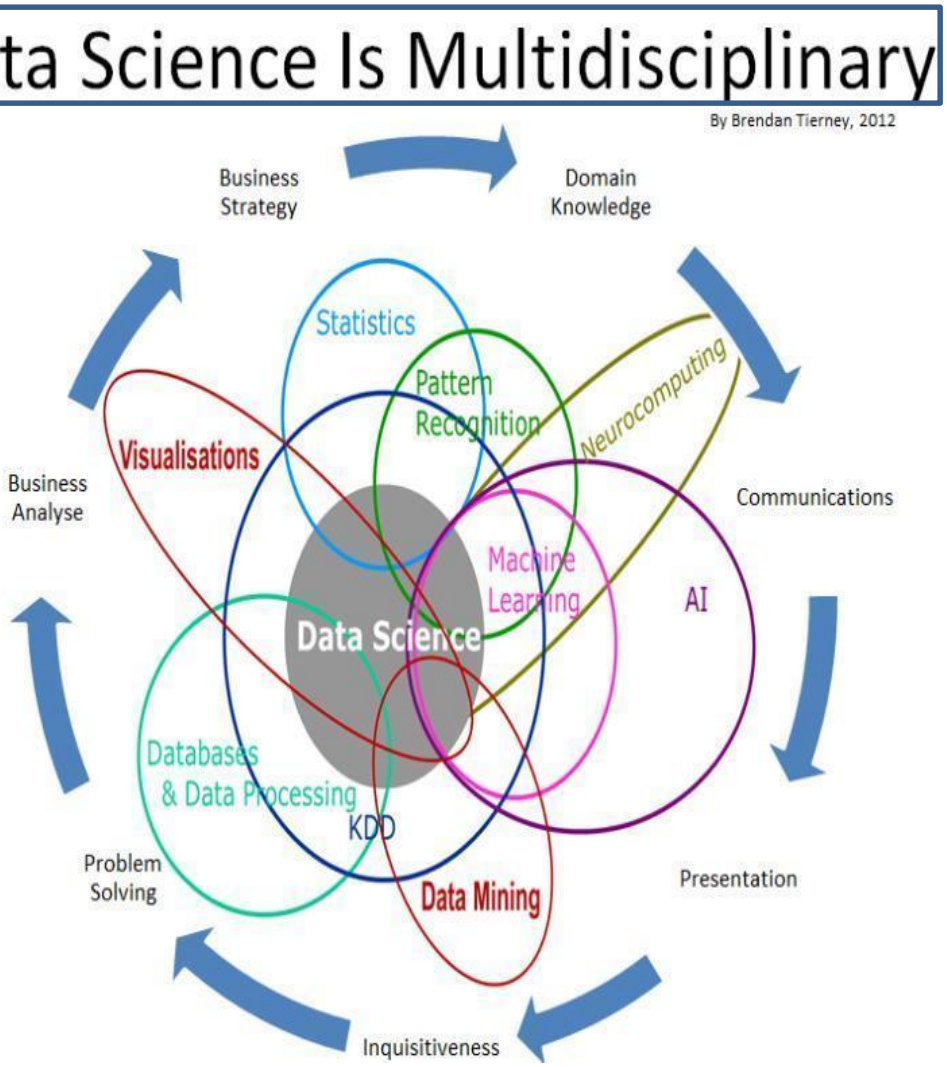
- Apa itu *Data Science* (DS) dan *Artificial Intelligence* (AI)?
- Bagaimana Hubungan DS dan AI?
- Konsep, Teori, Teknik, dan Tools DS
- Konsep, Teori, Teknik, dan Tools AI.
- Aplikasi DS dan AI.
- Peluang Pengembangan IPTEK-BIS Berbasis DS dan AI.
- Penutup

#1. Apa itu *Data Science* (DS) dan *Artificial Intelligence* (AI)?

Data science *is an inter-disciplinary field that uses* ***scientific*** *methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured* ***data****.*

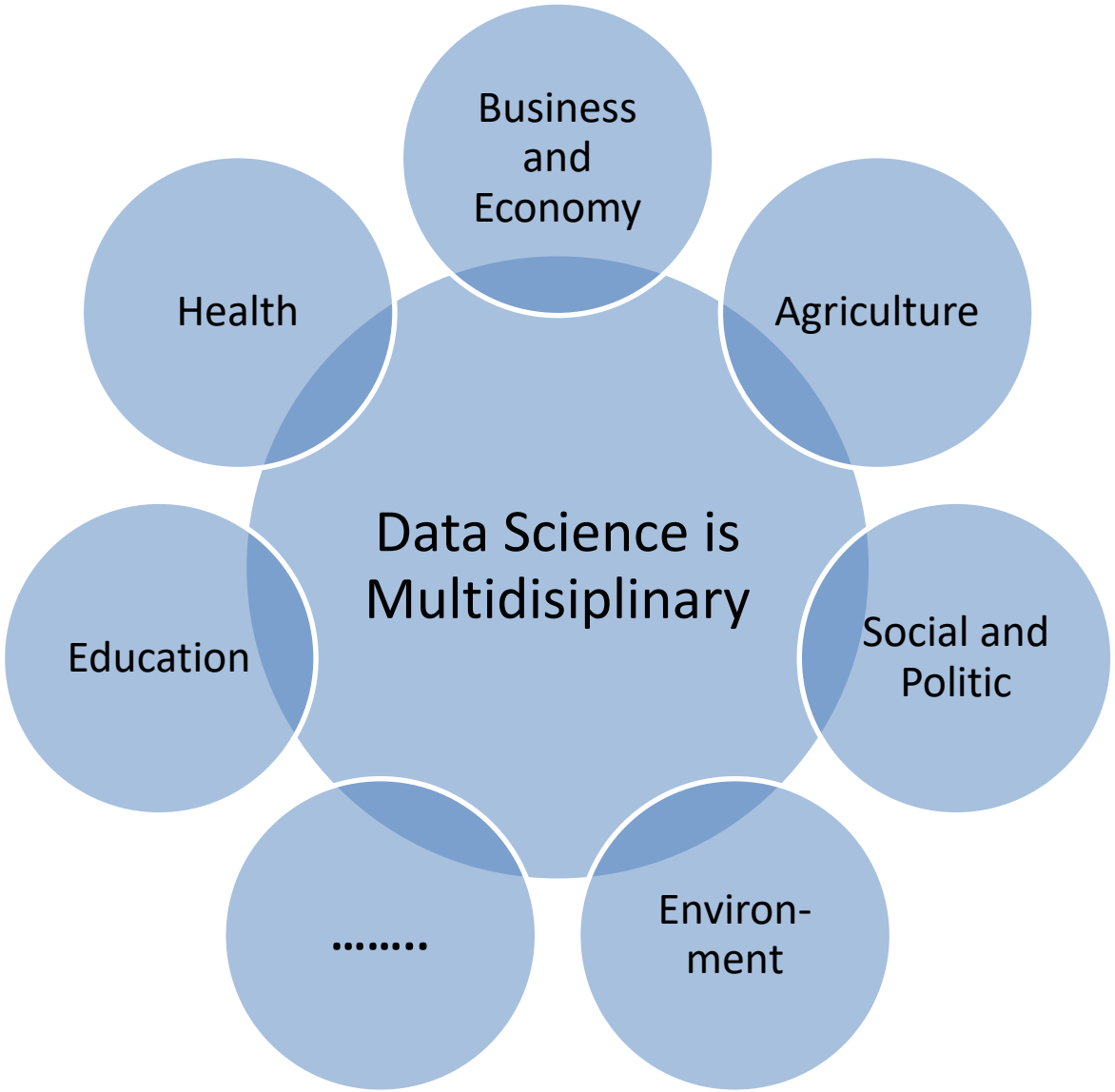
Data science *is related to* ***data*** *mining, machine learning and big* ***data****.*

- Deals with both structured and unstructured data
- A field that includes everything that is associated with the cleansing, preparation and final analysis of data
- Combines the programming, logical reasoning, mathematics and statistics
- Cleanses, prepares and aligns the data
- An umbrella of several techniques that are used for extracting the information and the insights of data



Source: Leonard Heiler, 2017. https://www.datasciencecentral.com/profiles/blogs/difference-of-data-science-machine-learning-and-data-mining

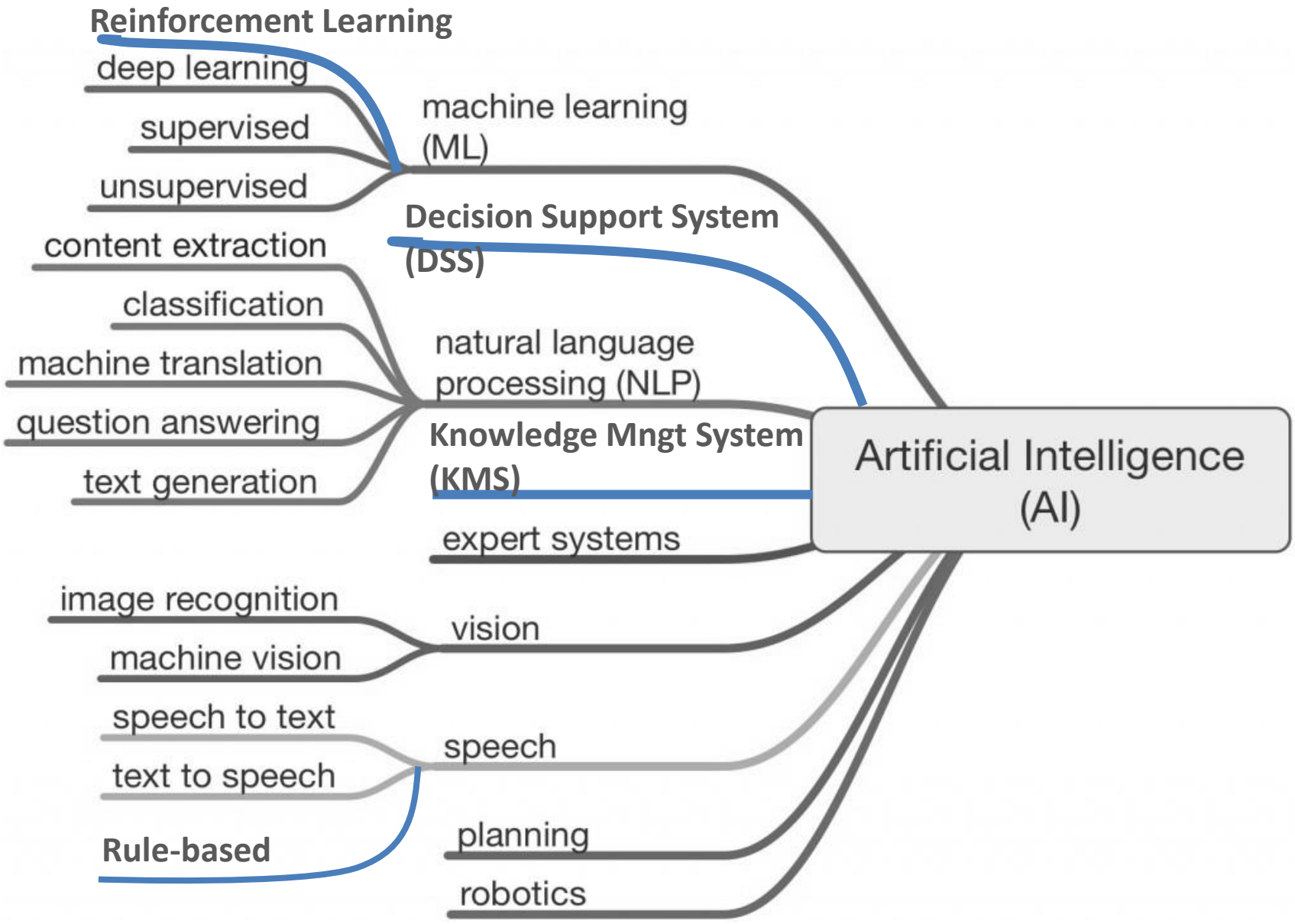
DS bisa Diaplikasikan pada Hampir Disemua Disiplin Ilmu



What is Artificial Intelligence?

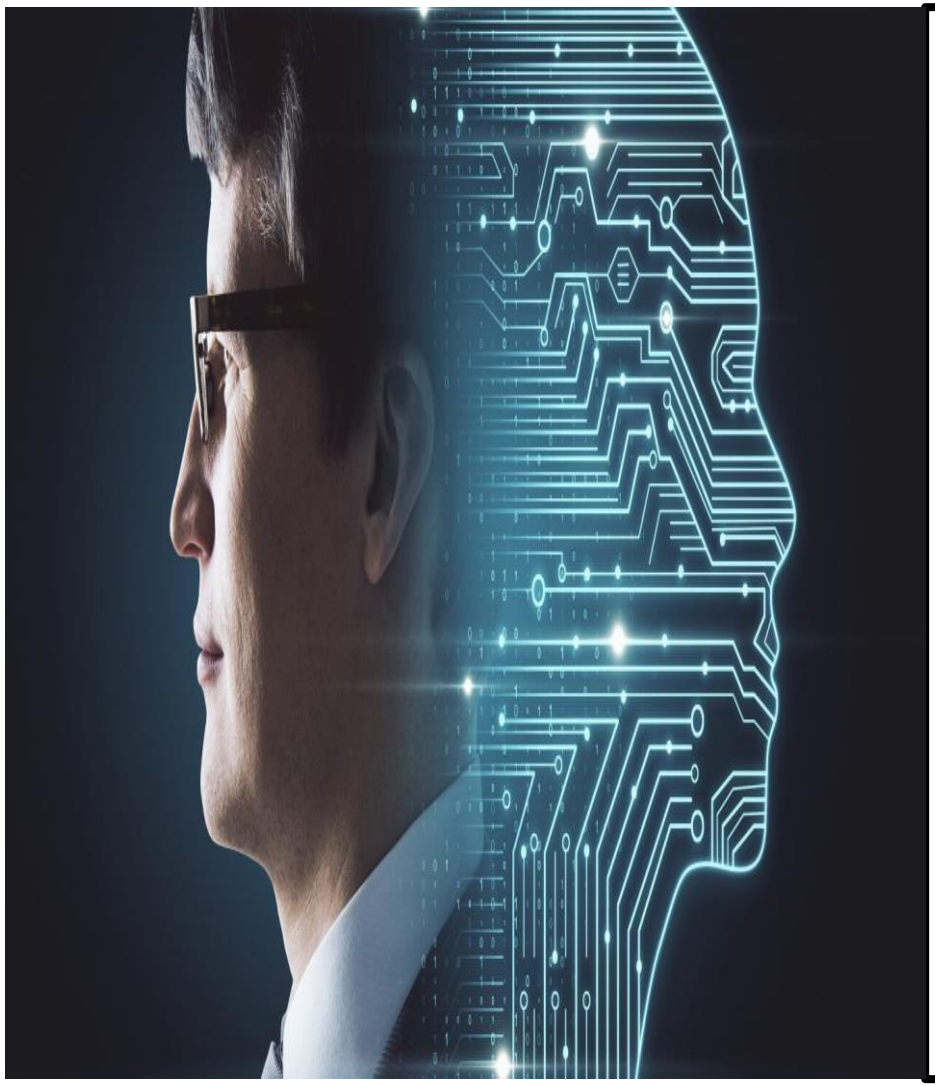
- It is the study of how to train the computers so that computers can do things which at present human can do better.
- The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision making, language translation, etc.
- It is the ability of computer program or a machine to think and learn.
- It is the branch of computer science that emphasizes the development of intelligence machines, thinking and working like humans.

Scope of Artificial Intelligence (AI)



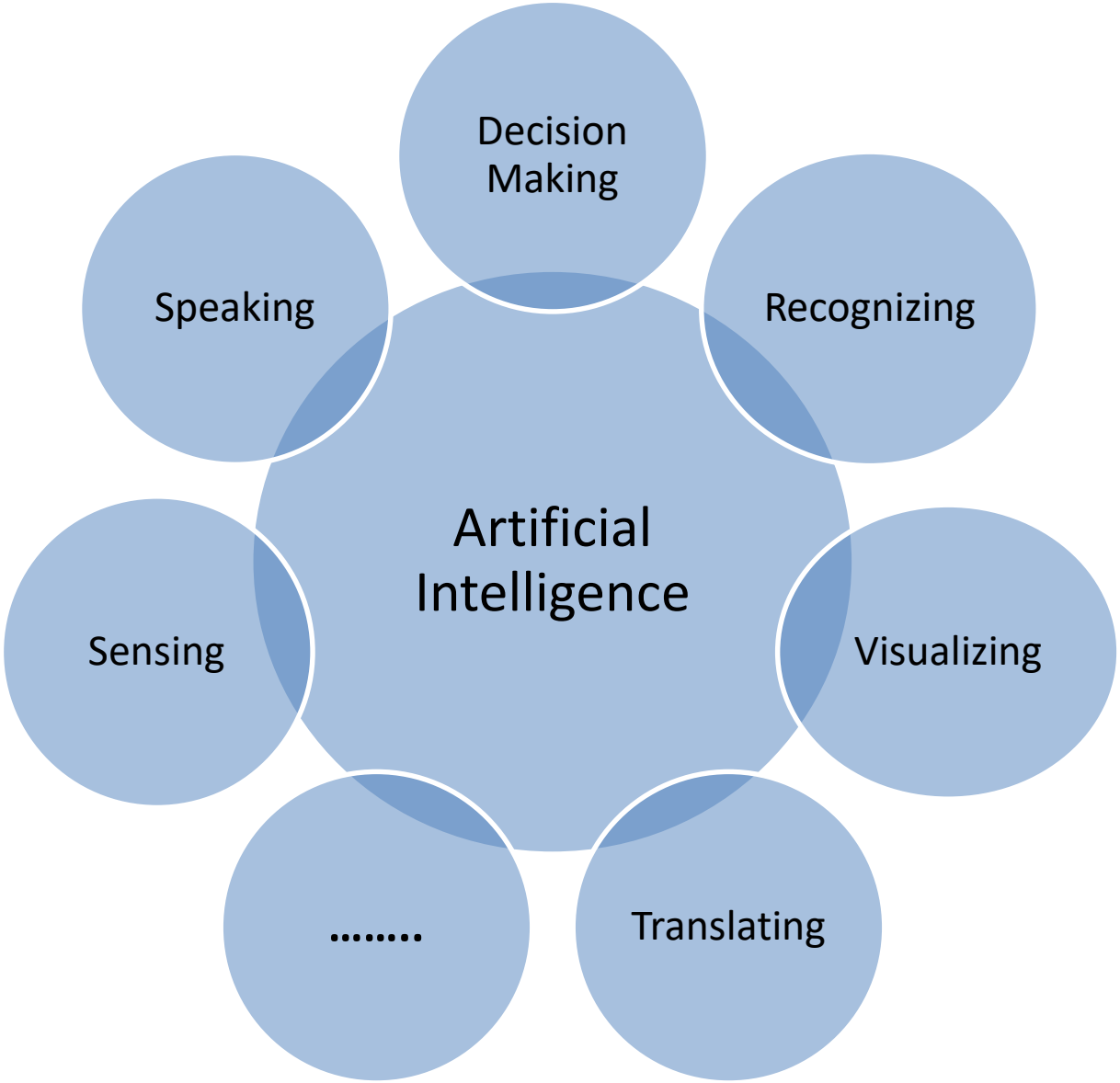
Source: ZAH, Modified from Neota Logic, 2016

Artificial Intelligence



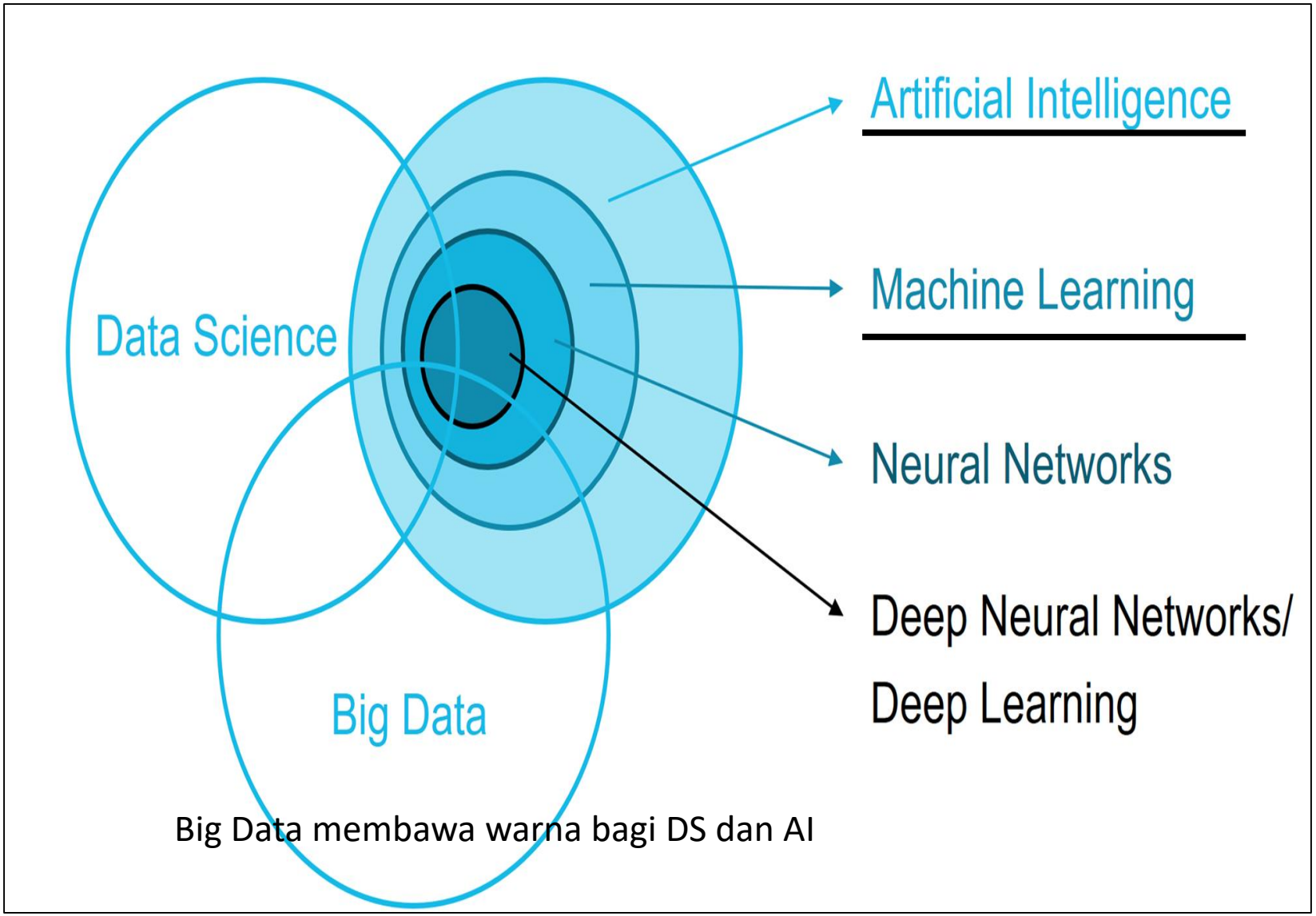
Artificial Intelligence is the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages

AI bisa Berpikir Seperti Layaknya Manusia Berpikir



#2. Bagaimana Hubungan DS dan AI?

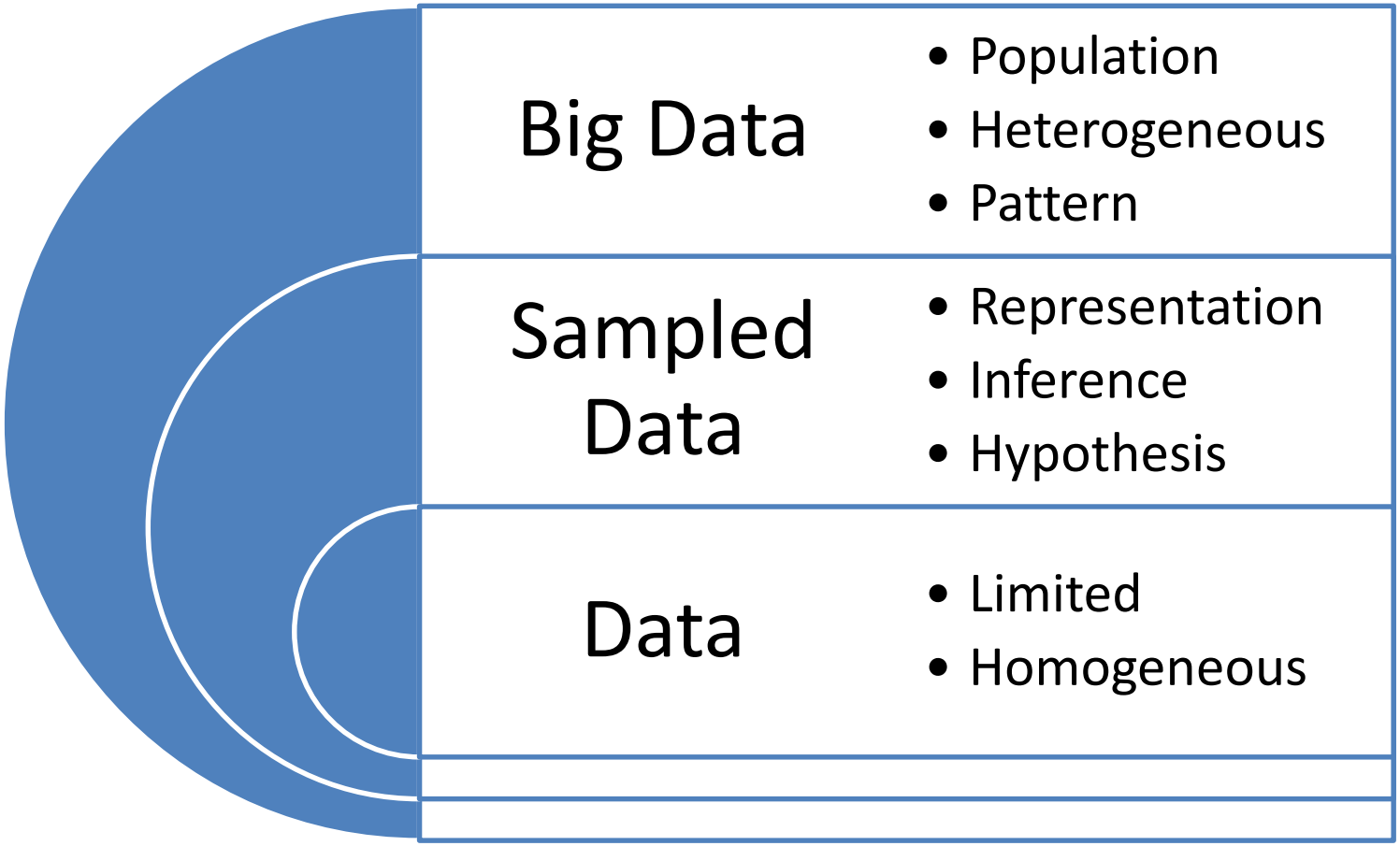
Relationship DS, Big Data, dan AI



Big Data membawa warna bagi DS dan AI

- DS memerlukan data terstruktur maupun tidak terstruktur untuk mendapatkan “knowledge” dan “insight”.
- AI memerlukan “knowledge” dan “insight” tersebut untuk berpikir dan belajar sebagaimana layaknya manusia berpikir dan belajar.
- Oleh karena itu, AI memerlukan hasil komputasi dari DS untuk dikomputasi lebih lanjut sehingga bisa digunakan untuk “meniru” bagaimana manusia berpikir dan belajar.
- Cara berpikir AI ini dapat terus ditingkatkan dengan cara “mengajarinya” dengan berbagai data masa lalu.
- Semakin besar jumlah data yang diperoleh, baik terstruktur, semiterstruktur, maupun tidak terstruktur, semakin baiklah kinerja dari AI. Data yang jumlahnya besar tersebut disebut Big Data.
- Big Data dengan 5-V nya bersifat heterogen, multiple sources, multi types, multi-temporal, multi-spatial.

Paradigm Shift in Viewing Data

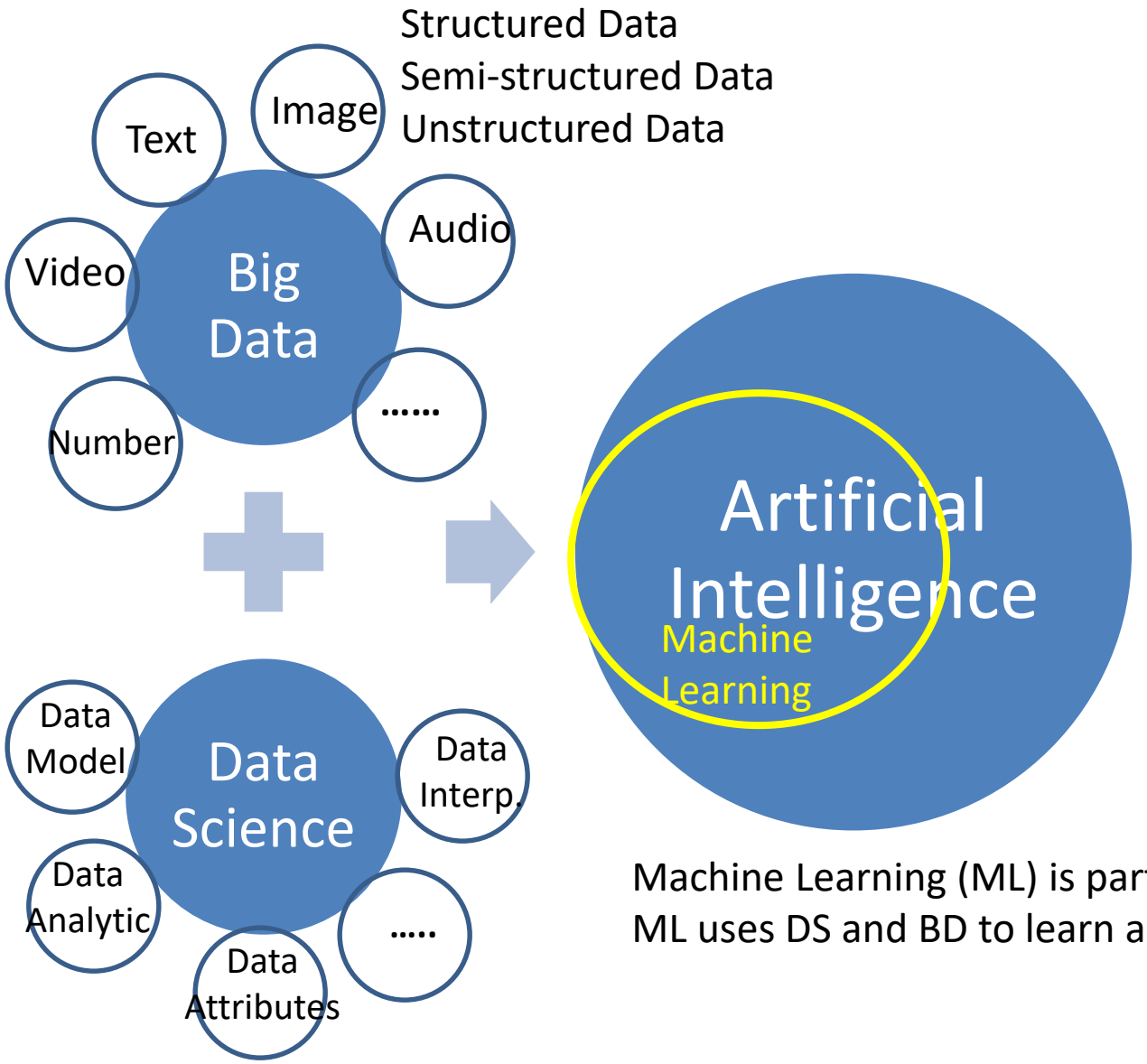


Big Data: More, Messy, Good Enough

- In this new world we can analyze far MORE data.
- Big data gives us an especially clear view of the granular: subcategories and submarkets that samples cannot assess.
- As scale increases, the number of inaccuracies increases as well (Messy).
- A move away from the search for causality to discover patterns and correlations.
- Big data is about WHAT, not WHY.
- Big data changes the nature of business, markets, and society.
- Values is shifted from physical infrastructure to intangibles such as brands and intellectual property.
- Big data is the oil of the digital economy.
- As individual shifts from privacy to probability: likelihood one get a heart attack, default on a mortgage, commit crime, climate change, eradicating diseases, fostering good governing and economic development.

BD Strengthening Relationship of DS and AI

- An organization that has big amounts of data gain competitive advantages in its playing field.
- The more data an organization has, the more accurate its descriptions, predictions, and prescriptions can be.
- Data Science, Big Data, and Artificial Intelligence play significant roles to present the solutions.
- This means making use of mathematical models to create algorithms to identify, classify, cluster, predict, learn, and to process data.



Machine Learning (ML) is part of AI
ML uses DS and BD to learn and re-learn

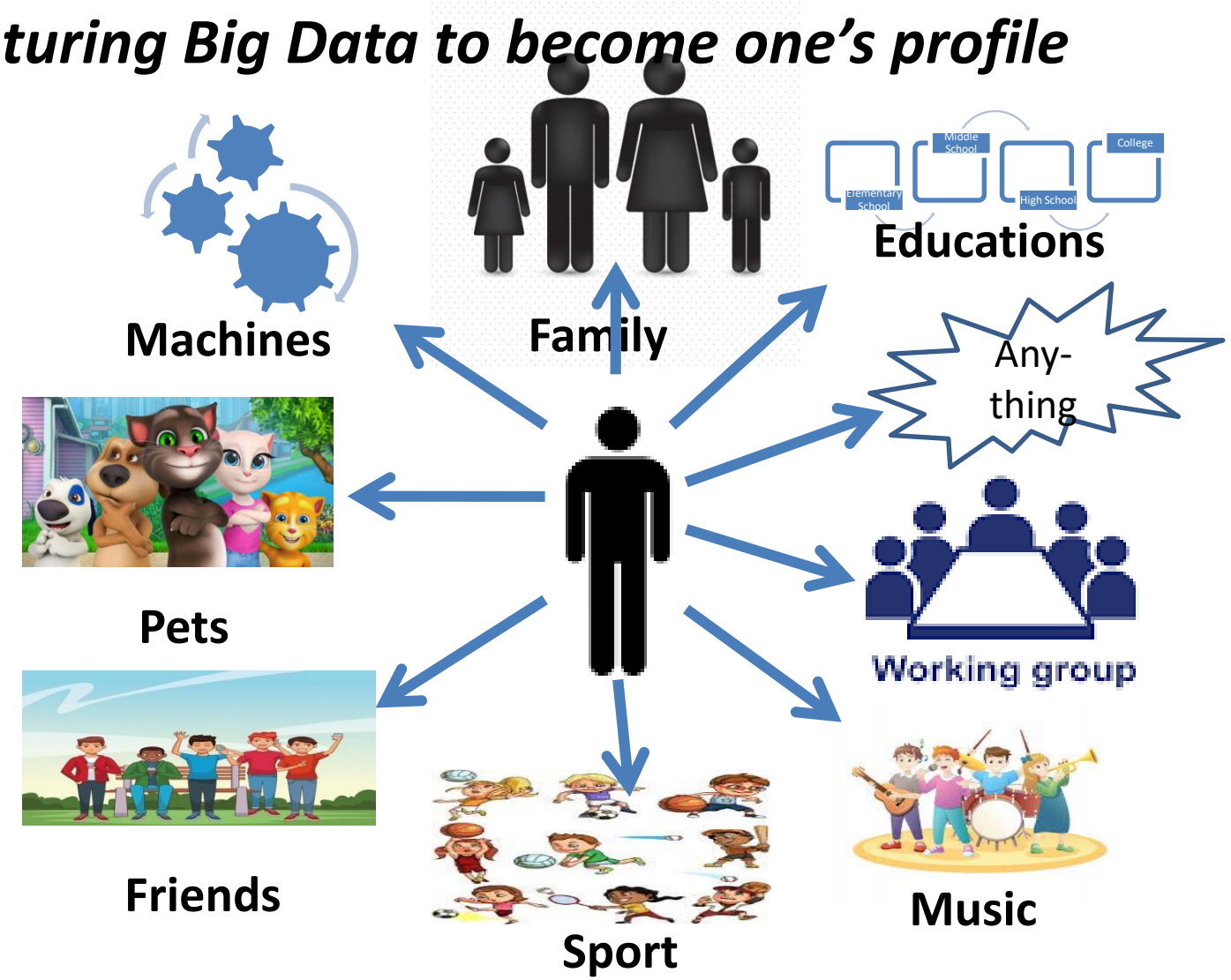
Zainal A. Hasibuan, 2021

#3. Konsep, Teori, Teknik, dan Tools untuk DS

Konsep *Data Science*

- Pada dasarnya, DS menggali dan mengekstrak “knowledge” dan “insight” dari data.
- Di era digital ini, hampir semua kegiatan manusia, interaksi manusia dengan lingkungannya, dan lingkungannya itu sendiri, bisa terekam dalam bentuk digital, atau bisa di digitalisasikan, sehingga menjadi data.
- Sehingga dapat dikatakan “everything is data”.
- Secara *inherence*, ternyata berbagai data tersebut ada strukturnya (*inherence structure,* atau *hidden structure*).
- Menjadi tantangan tersendiri bagi data scientist untuk mengungkapkan *inherence structure,* atau *hidden structure* tersebut, menjadi “knowledge” and “insight”.

Capturing Big Data to become one's profile



We are Networked Society, Most of Our Interactions Recorded Digitally to Becomes Big Data & Needs Data Science to Compute

THE FACTS OF DATA TODAY | BIG DATA

UNSTOPPABLE FLOWS OF DATA



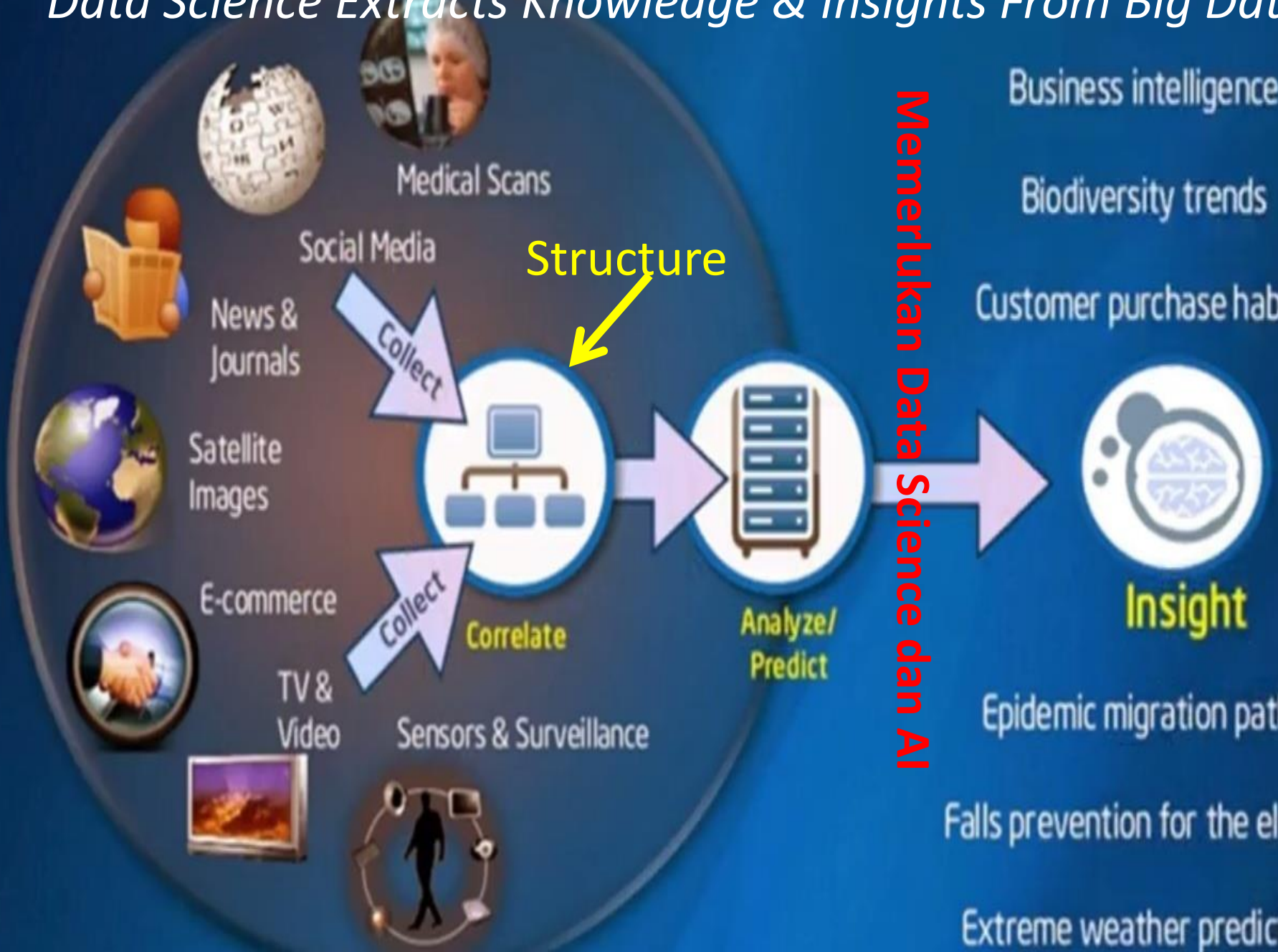
VALUABLE MEANING (INFORMATION)

STRUCTURED VS UNSTRUCTURED
DISCRETE VS CONTINUOUS
FORMAL VS INFORMAL
HUMAN VS MACHINE
SINGLE VS MULTI-FORMAT
VALID VS INVALID
RELIABLE VS UNRELIABLE
PAST VS PRESENT VS FUTURE
SLOW VS FAST-MOVE
SMALL VS BIG
STATIC VS DYNAMIC
DIGITAL VS ANALOG
OFFLINE VS ONLINE
NATURAL VS DIGITILIZED

BAGAIMANA CARA MENGOLAHNYA ?

Source R. Eko Indrajit, 2021

Data Science Extracts Knowledge & Insights From Big Data



Teori *Data Science*

- Teori yang mendasari DS boleh dikatakan tidak ada yang berubah, antara lain seperti:
  - Teori peluang dan teori statistik
  - Teori komputasi matriks
  - Aljabar linear
  - Kalkulus: limit, integral, dan differential
  - Fungsi, discriminant analysis
  - Principle Component Analysis, Factor Analysis
  - Data kategori dan distribusi, dll.
- Dengan tersedianya computing power yang besar, volume data yang besar dan heterogen dan bisa terkumpul dengan cepat, serta bahasa pemrogaman yang canggih, membuat DS semakin powerful.
- Data Science menjadi semakin menarik dan penting dengan hadirnya Big Data, Computing Power, IoT, Blockchain, AI, dll.

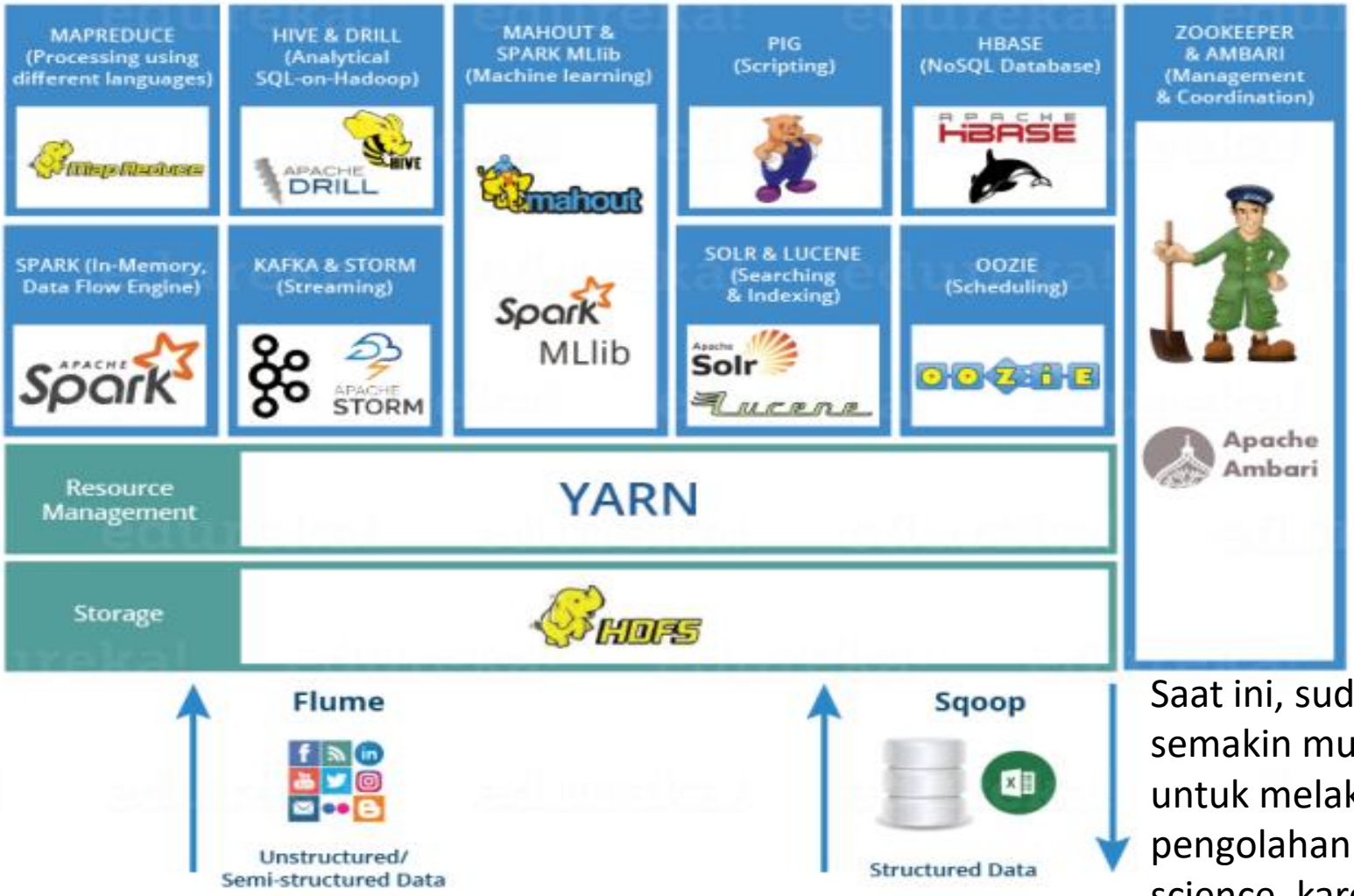
Teknik-2 Yang Digunakan untuk *Data Science*

- Probability and Statistics
- Distribution
- Regression analysis
- Descriptive statistics
- Inferential statistics
- Non-Parametric statistics
- Hypothesis testing
- Linear Regression
- Logistic Regression
- Neural Networks
- K-Means clustering
- Decision Trees, dll.

Dasar-2 teorinya sama persis dengan dasar-2 teori yang selama ini dipelajari dalam pelajaran statistic-matematik

Bedanya, belakangan ini ditambah dengan pengetahuan dan keterampilan memprogam, terutama dengan Bahasa Statistical programming.

Ecosystem Tools of Big Data Analytics



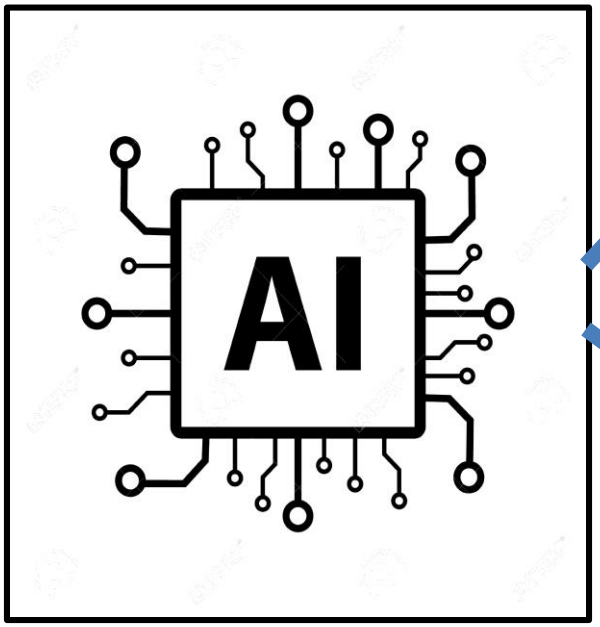
Saat ini, sudah semakin mudah untuk melakukan pengolahan Data science, karena sudah banyak tools nya.

#4. Konsep, Teori, Teknik, dan Tools AI.

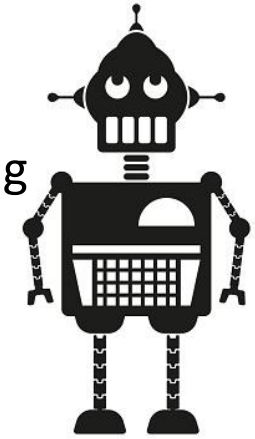
Konsep Artificial Intelligence (AI)

- Pada dasarnya, AI (Kecerdasan Artificial) adalah kecerdasan yang dimiliki oleh komputer, meniru sebagaimana kecerdasan yang dimiliki oleh manusia.
- Komputer bisa memiliki kecerdasan artificial tersebut karena “diajarkan” terlebih dahulu oleh penciptanya.
- Semua kemungkinan yang akan terjadi (teori himpunan) terlebih dahulu di “tanamkan” AI-nya ke computer tersebut.
- Selanjutnya, ada AI-nya yang bisa belajar sendiri setelah diajarkan (dilatih), itulah yang disebut *Machine Learning*.

Inherent Structure of Big Data and
Business Insight Embed in Big Data
Become Input to Train Machine Learning



Artificial Intelligence



AI-Robot



AI-Automation System:
Q/A System, Machine Learning…

Information Systems Development

Penerapan AI

- Pada dasarnya, penerapan AI dibagi dua:
  - Pada Robotic
  - Pada Automation System
- Umumnya, Robot menggantikan pekerjaan manusia yang sifatnya berulang-ulang, atau yang membahayakan nyawa manusia (pemadam kebakaran), dll.
- Sedangkan Automation System, seperti Expert System, Machine Learning System, Deep Learning System, dll., digunakan untuk mengidentifikasi, mendiagnosa, mengklasifikasi, memprediksi, pengambilan keputusan, dan preskripsi, dll., dari berbagai permasalahan (problem solving) yang dihadapi.
- Pembahasan aplikasi AI selanjutnya, difokuskan ke automation system ini, yaitu Machine Learning.

Three Types of Artificial Intelligence

- Artificial Narrow Intelligence (ANI), focuses on one single narrow task. For example: machine translation, Q/A machine, decision making, etc.
- Artificial General Intelligence (AGI), would be able to think more or less on the same level of human. For example, machine learning, deep learning. reinforcement learning.
- Artificial Super Intelligence (ASI), would have more capability than a human, they would have emotion and relationship (Human to Machine and Machine to Human Touch).

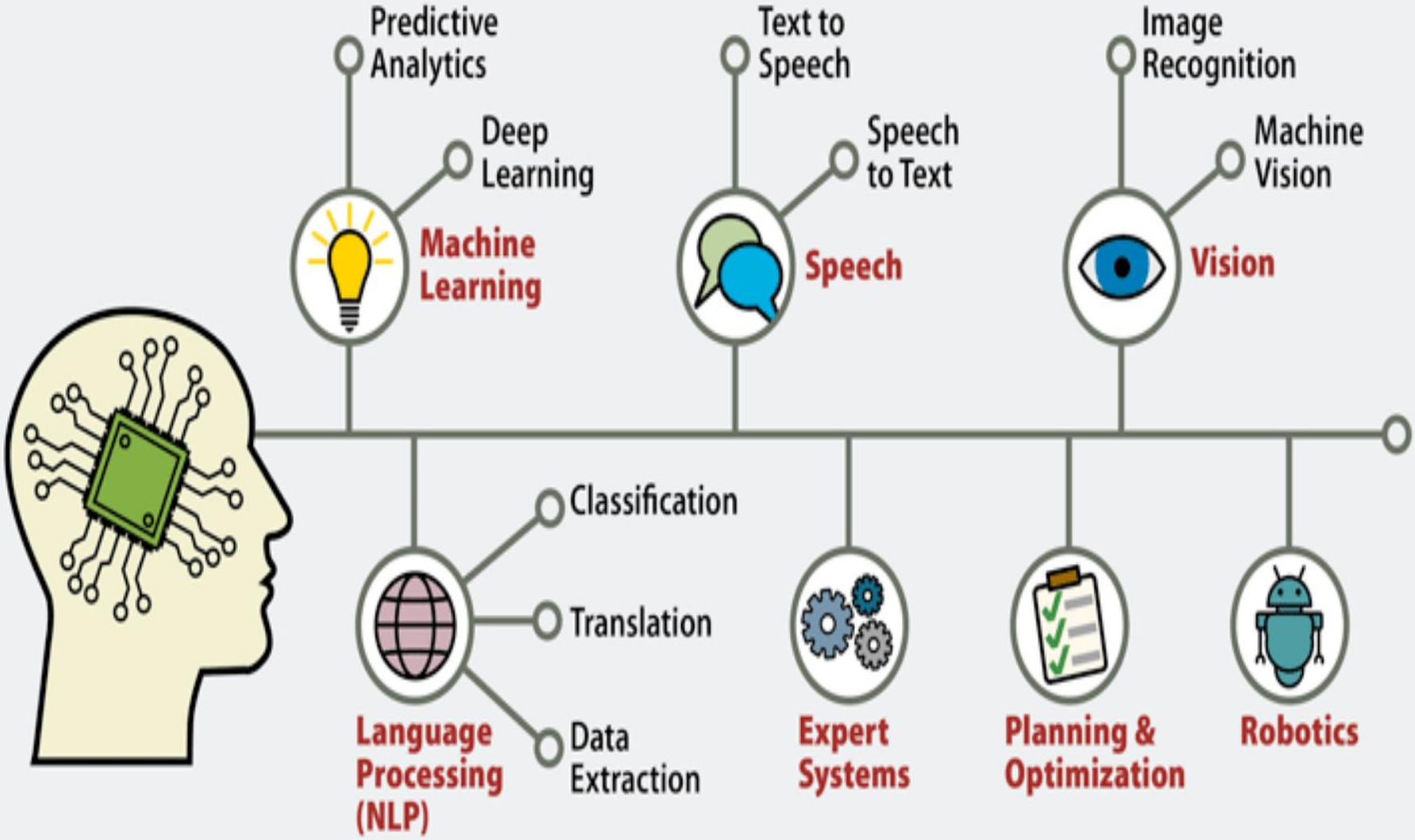
Teori Dasar untuk AI

- *Intelligent agents*
- *Problem-solving through Search*
- *Knowledge Representation and Reasoning*
- *Representing and Reasoning with Uncertain Knowledge*
- *Machine Learning and Knowledge Acquisition*
- Dll.

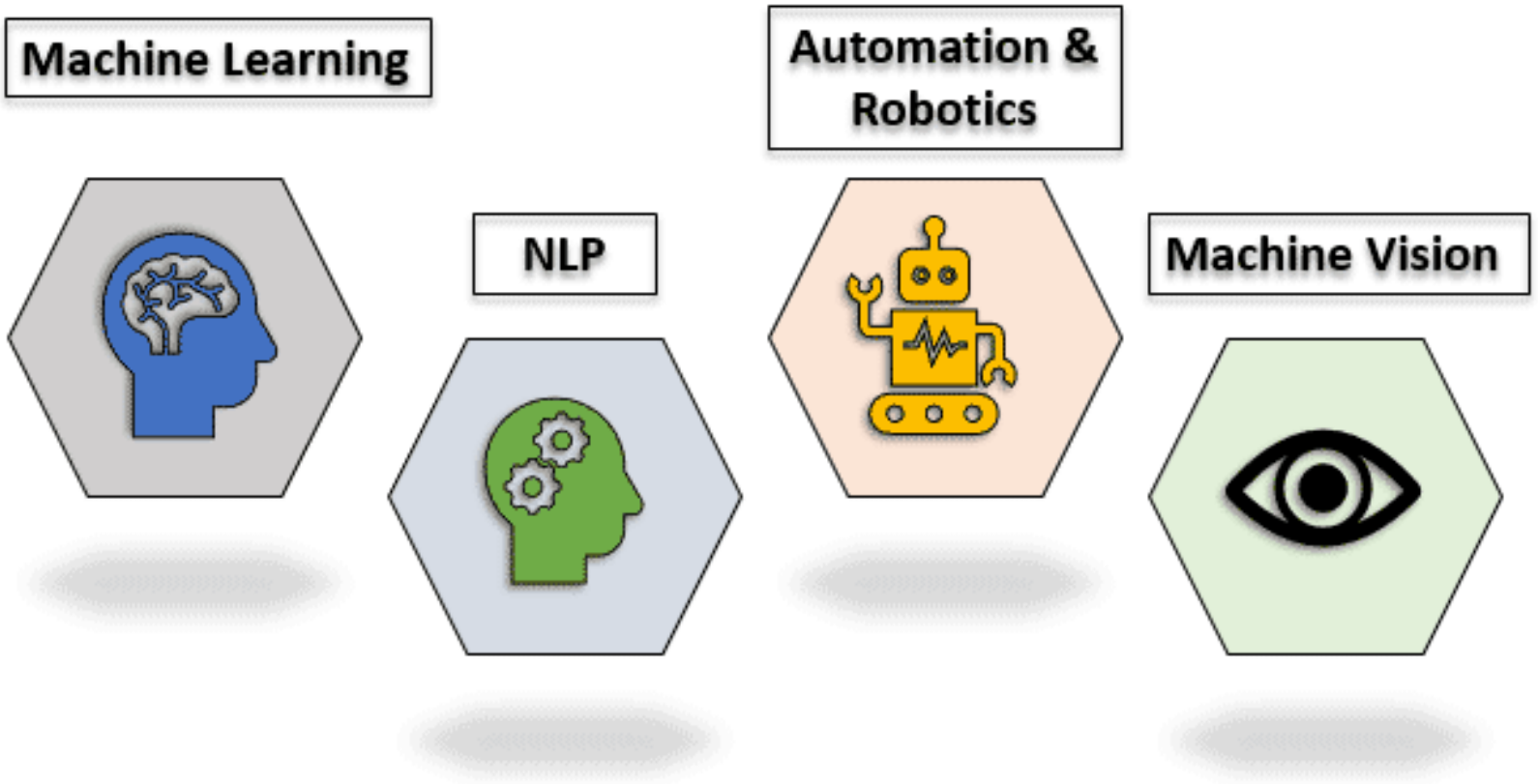
Some Fundamental AI techniques

- **Heuristics,**
- **Support Vector Machines,**
- **Neural Networks,**
- **Markov Decision Process,**
- **Natural Language Processing**
- **Dll.**

Artificial Intelligence



Top 4 Techniques of Artificial Intelligence



www.educba.com


Springboard
Machine Learning libraries
torch
scikit learn
PyTorch
Deep Learning libraries
Keras
LUDWIG
Caffe MODELS
AI for Mobile
TensorFlow Lite
ML
OpenCV
AI Platforms
TensorFlow
ML
H2O.ai
NLP, audio, video, text
SimpleCV
TESSERACT
Detectron2
Top 15 open-source artificial intelligence tools to learn in 2020

#5. Aplikasi Data Science dan Artificial Intelligence

Prinsip Aplikasi *Data Science*

- Data driven business
- Data driven decision making
- To reveal the inherent structure
- To gain new business insights

Data Science Applications



01 E-Commerce

02 Healthcare

03 Gaming

04 Energy

05 Travel

06 Finance

TechVidva



Data Science in E-commerce

Mempengaruhi trend pasar dengan cara mengetahui perilaku pelanggan terhadap berbagai produk komersial dengan menerapkan analisa regressi

Mempengaruhi customer dengan cara melakukan sentiment analysis terhadap suatu produk komersial dengan menggunakan teknik K-means untuk mengelompokkan customer



Data Science in Healthcare

Memprediksi dampak dari obat terhadap suatu penyakit dengan menerapkan uji hopetesis statistik

Melakukan gnenomic analysis sebagai bagian dari upaya preventive penyakit menggunakan neural networks

Prinsip Aplikasi AI

- Sama dengan DS, hampir semua bidang kehidupan manusia bisa juga mengaplikasikan AI.
- Berbeda dengan DS, dimana DS digunakan untuk membangun model (misalnya, regressi, decision tree, dll.) dalam rangka menganalisis data, sedangkan AI menggunakan hasil analisis data sebelum2nya tersebut dan dimasukkan kedalam computer untuk membantu (menggantikan) pekerjaan manusia, seperti melakukan identifikasi, klasifikasi, prediksi, preskripsi.
- Dasar teori dan teknik yang digunakan oleh DS dan AI hampir sama, bedanya, teori dan teknik untuk DS digunakan untuk mendapatkan *new knowledge dan new insight*, sedangkan teori dan teknik untuk AI digunakan untuk membantu pekerjaan manusia dengan cara sistemnya dilatih terlebih dahulu (learning).

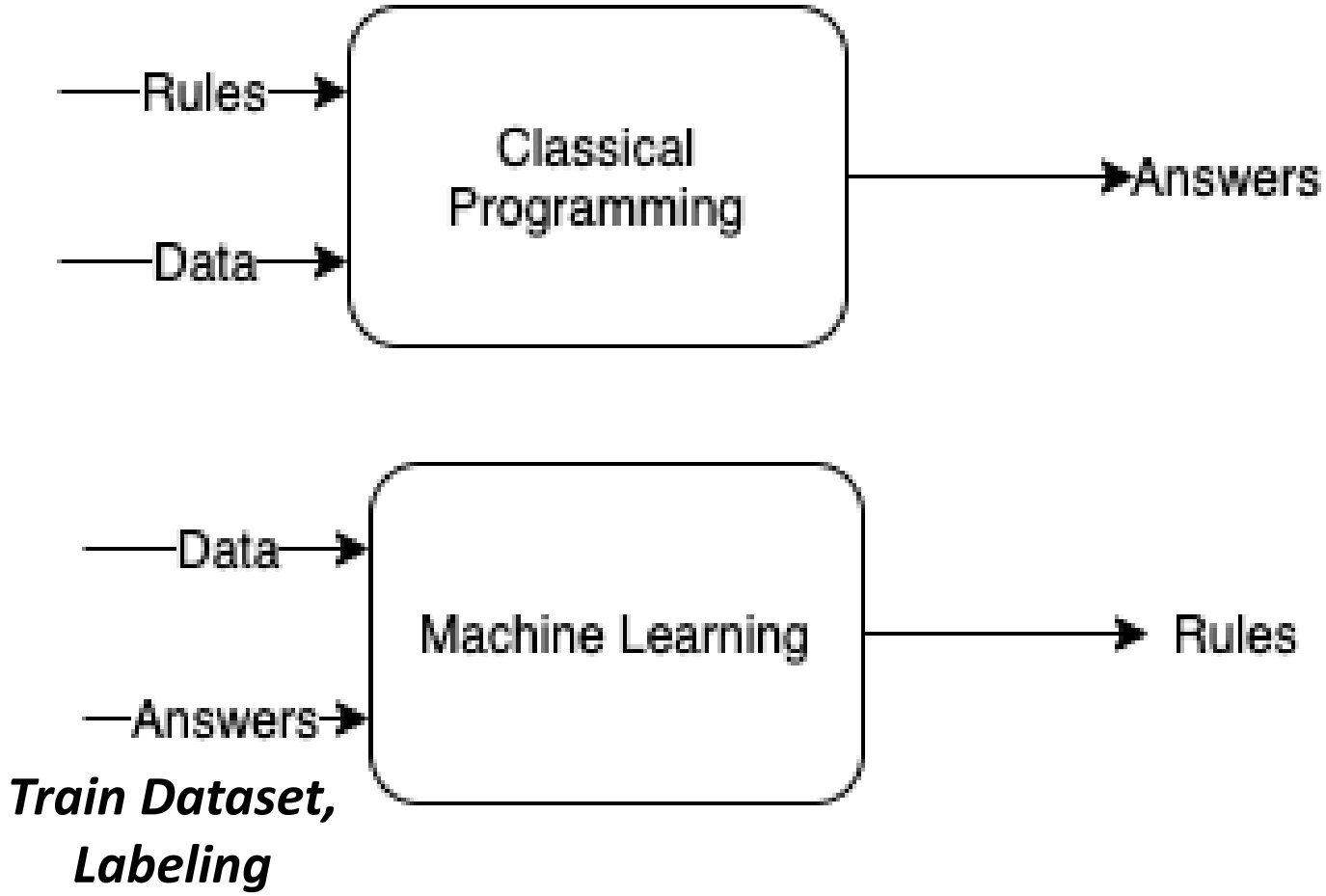
Applications of AI

- Healthcare
- Automobile
- Finance
- Surveillance
- Social Media
- Entertainment

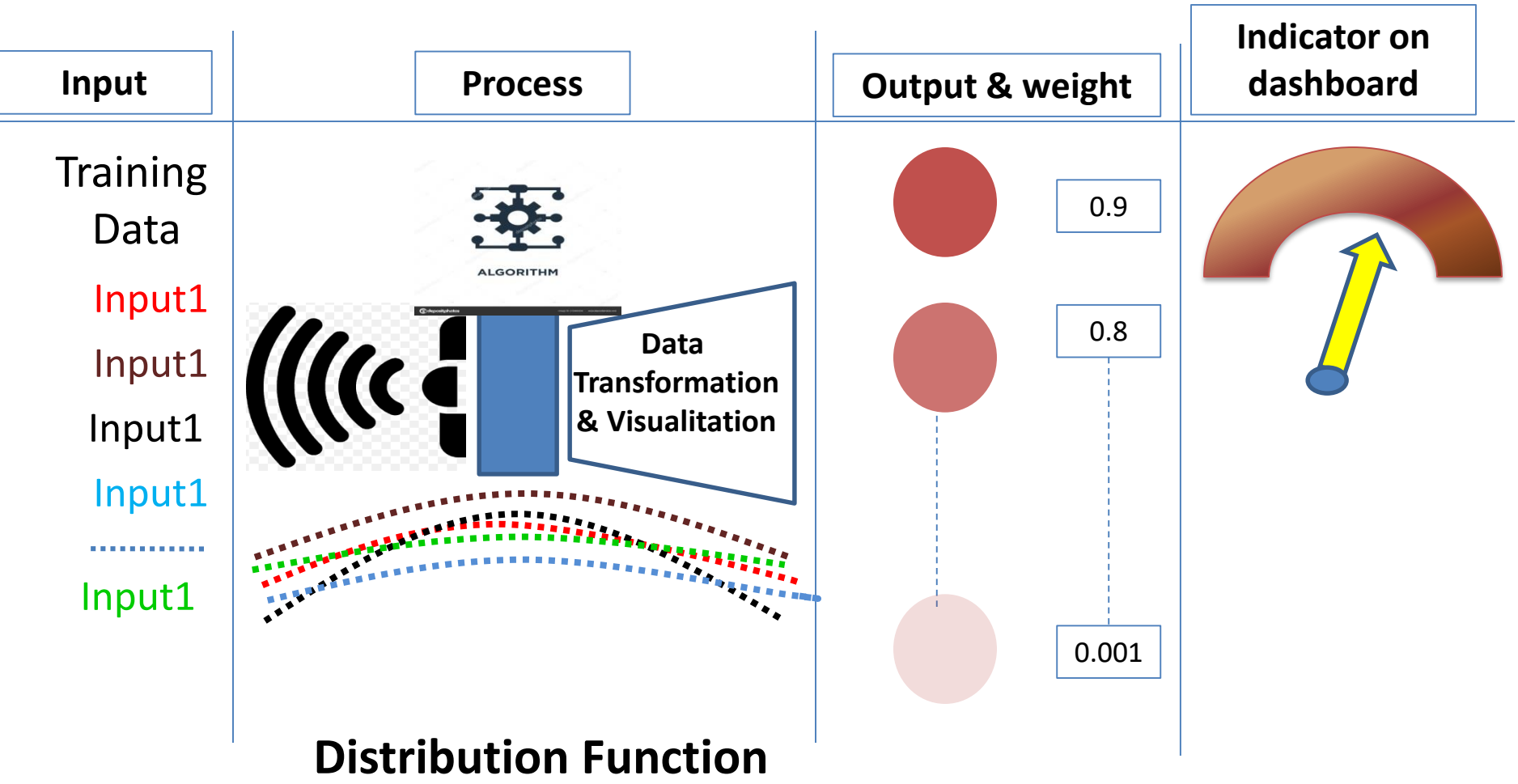


- Education
- Space Exploration
- Gaming
- Robotics
- Agriculture
- E-commerce

How Machine Learning Works

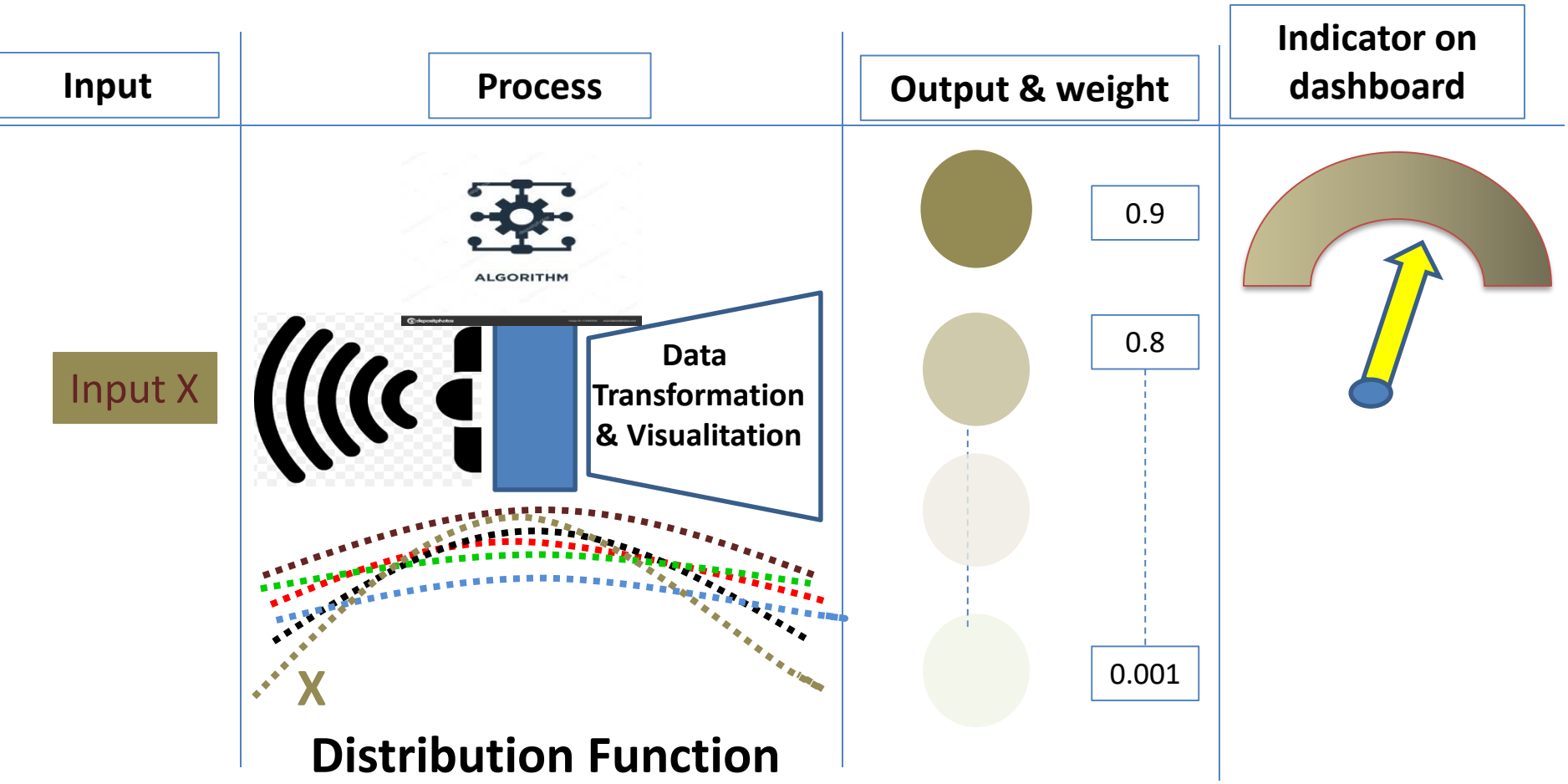


How Machine Learning Learn?



Zainal A. Hasibuan, 2020

Illustration of How Machine Learning Classify



Zainal A. Hasibuan, 2020

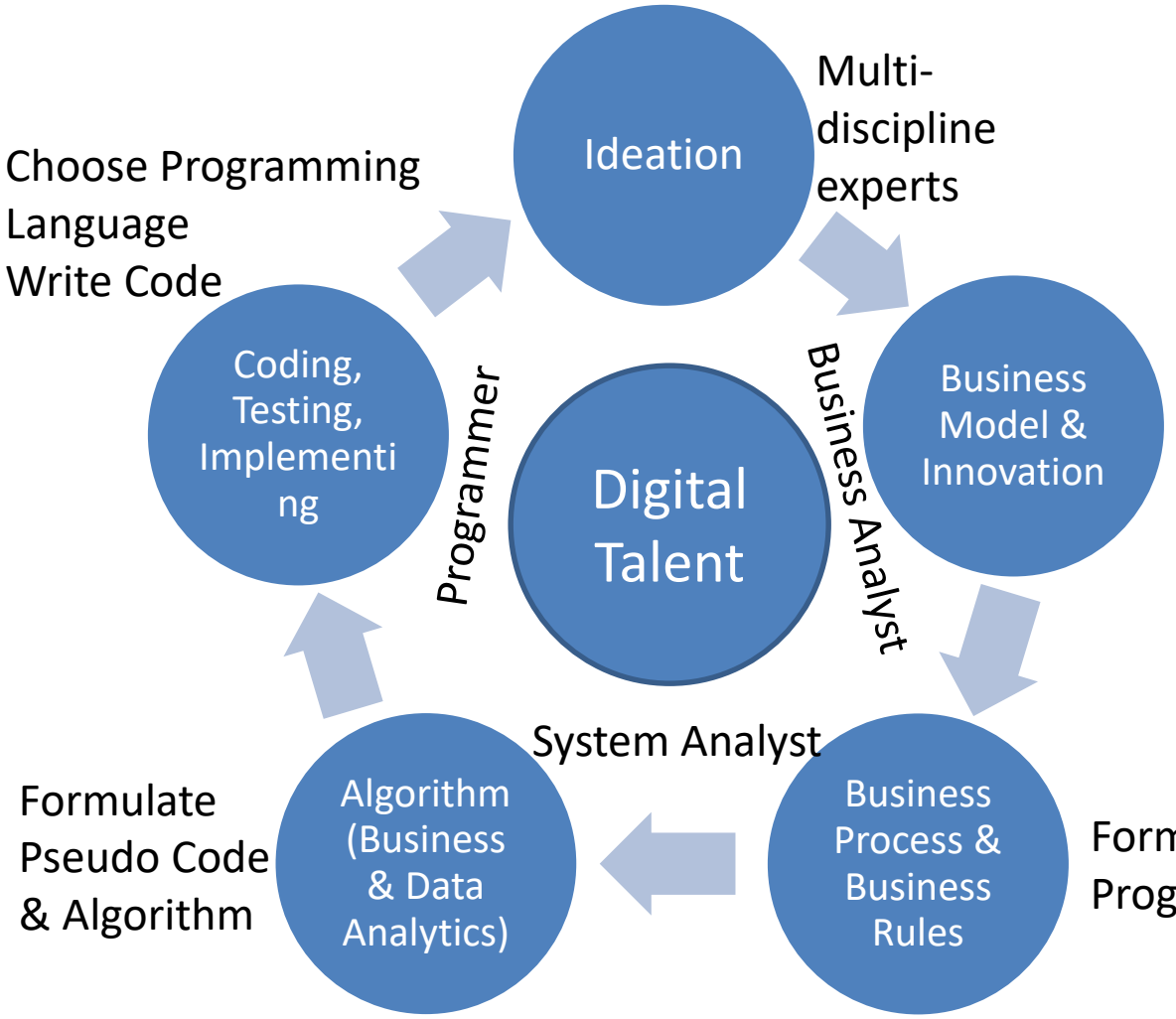
The Top 10 Algorithms for Machine Learning Newbies

1. Linear Regression
2. Logistic Regression
3. Linear Discriminant Analysis
4. Classification and Regression Trees
5. Naive Bayes
6. K-Nearest Neighbors
7. Learning Vector Quantization
8. Support Vector Machines
9. Bagging and Random Forest
10. Boosting and AdaBoost

(https://towardsdatascience.com/

#6. Peluang Pengembangan Ilmu pengetahuan, Teknologi dan Bisnis (IPTEK-BIS) Berbasis DS dan AI

Develop Automation System : Multi-discipline Approach”



Choose Programming Language
Write Code

Formulate Pseudo Code & Algorithm

Formulate Logic Programming

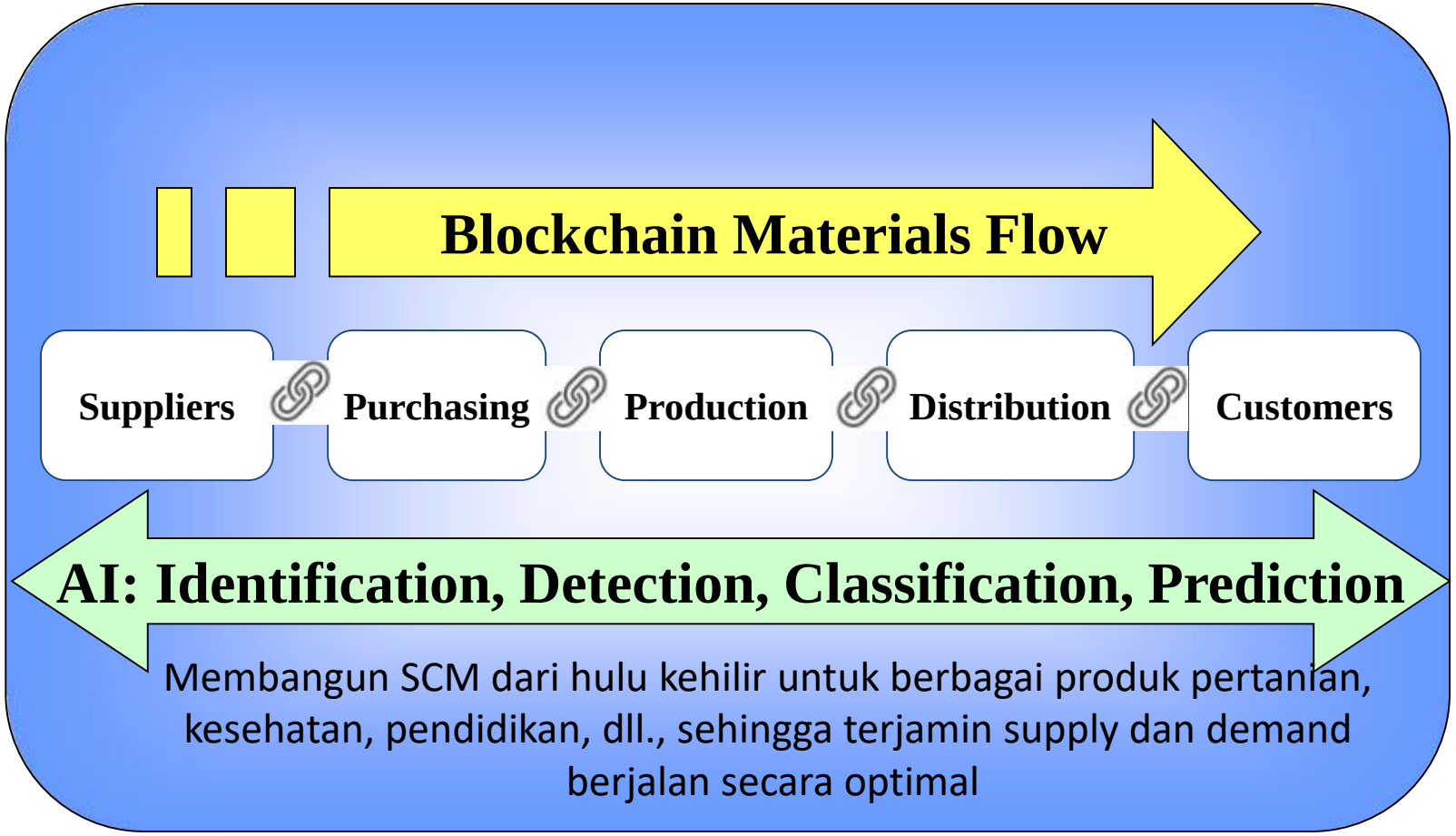
List of Digital Talent

Complex problem solving
Critical thinking
Creativity
People management
Coordinating with others
Etc….

Solving Problem
Creating Opportunity

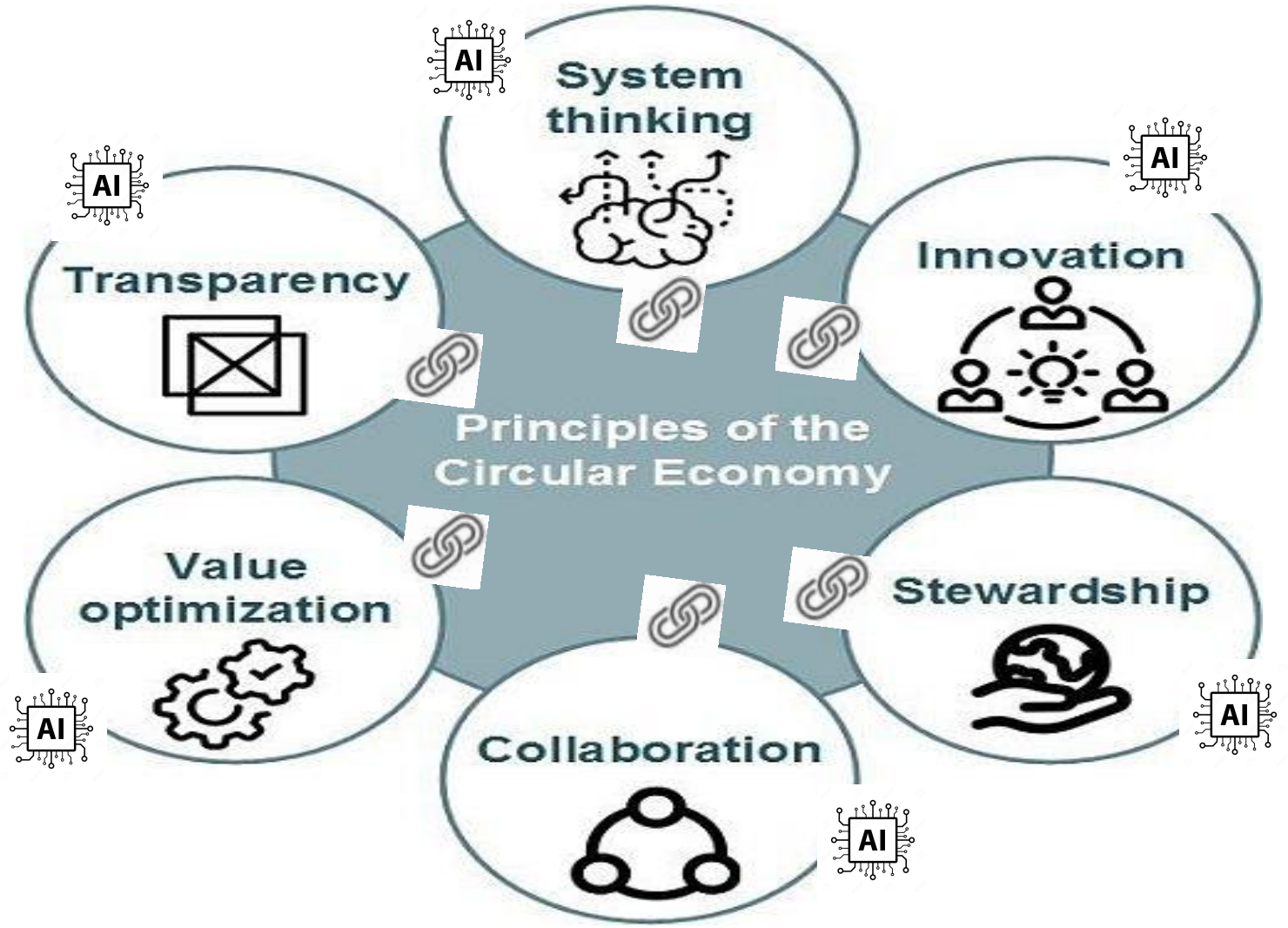
Zainal A. Hasibuan, 2019

Integrating Supply Chain Management Using Blockchain, Data Science and Artificial Intelligence

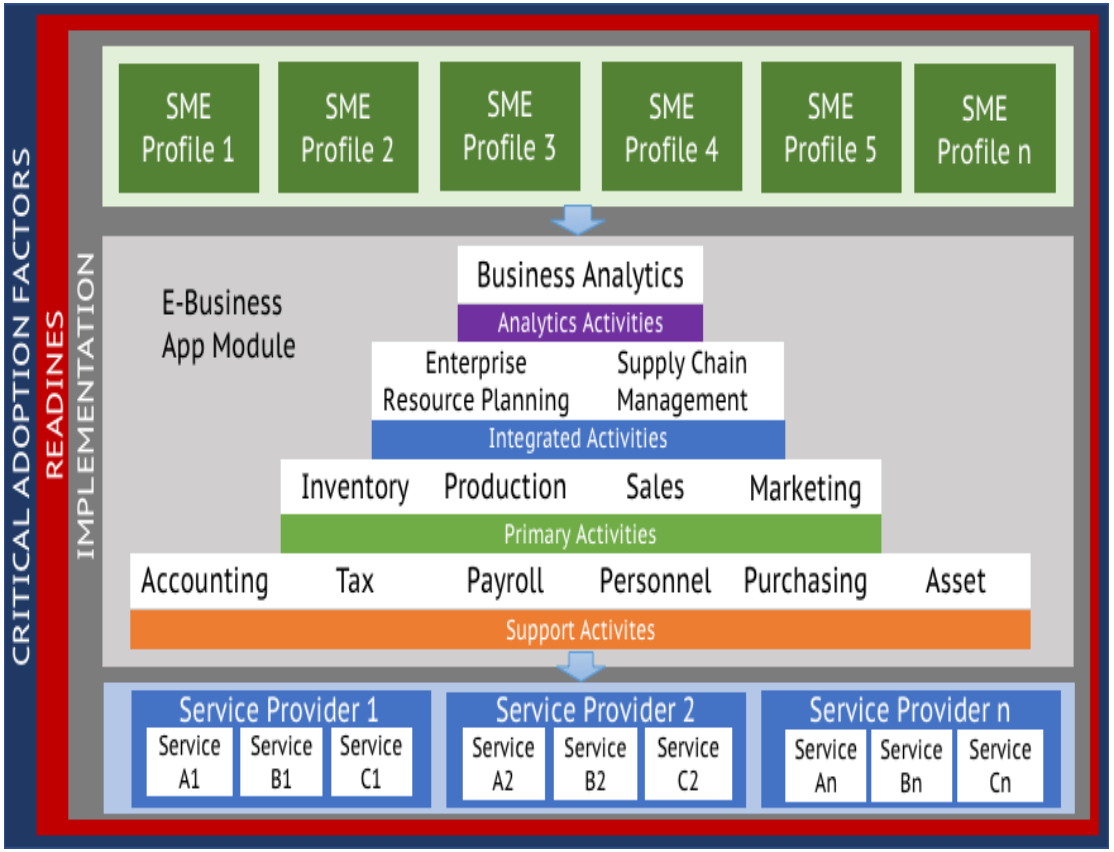


Modified from Zainal A. Hasibuan, 2020

Application of Blockchain, DS and AI in Systems Thinking of Circular Economy



Design dari Aplikasi Untuk UMKM



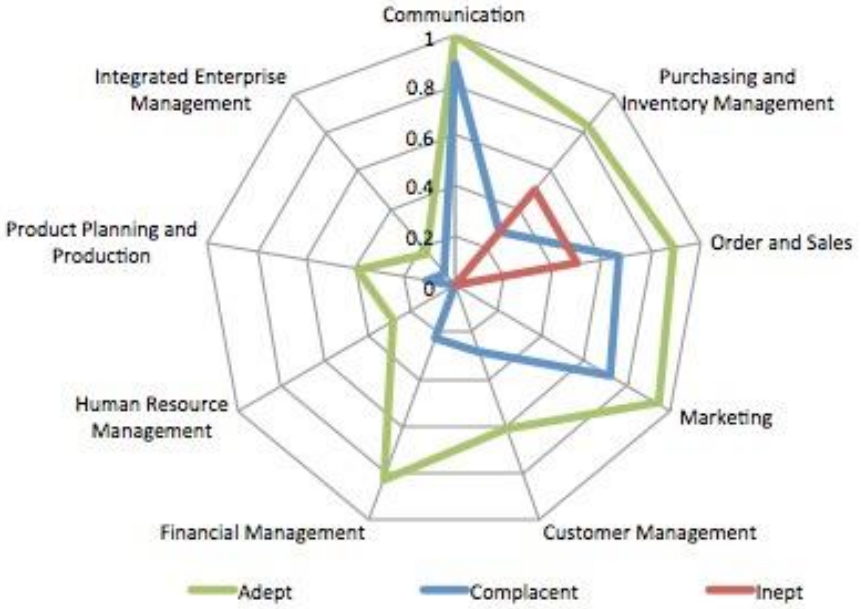
Pada tahap ini dirancang sebuah *conceptual framework* untuk mengidentifikasi dan mendefinisikan variable-variable penting yang diangkat dalam penelitian ini, dan hubungan antar variable satu sama lainnya.

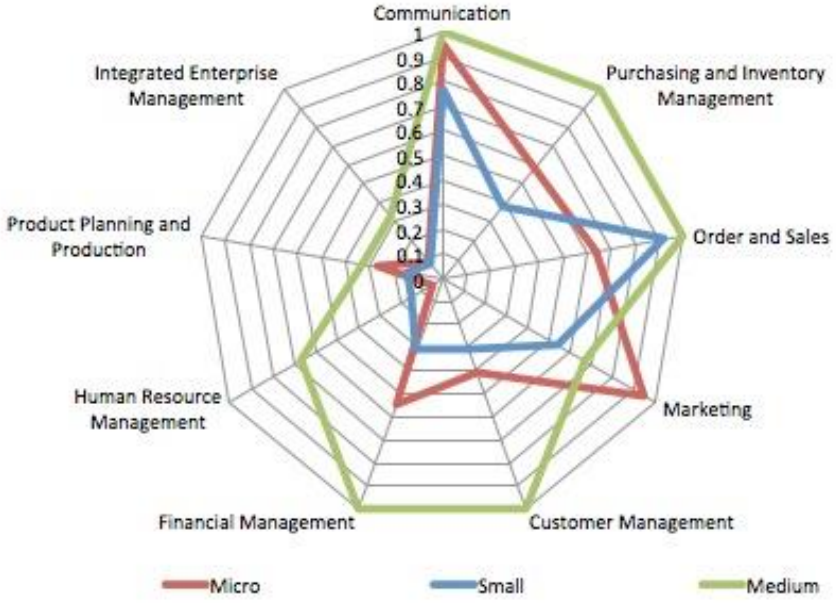
Adapun variable yang dibahas dalam penelitian ini terdiri dari :

1. Keberagaman UMKM
2. Plug and Play Services
3. Service Push

Ni Made Satvika Iswari, Eko Budiarjo, Zainal A. Hasibuan, 2017

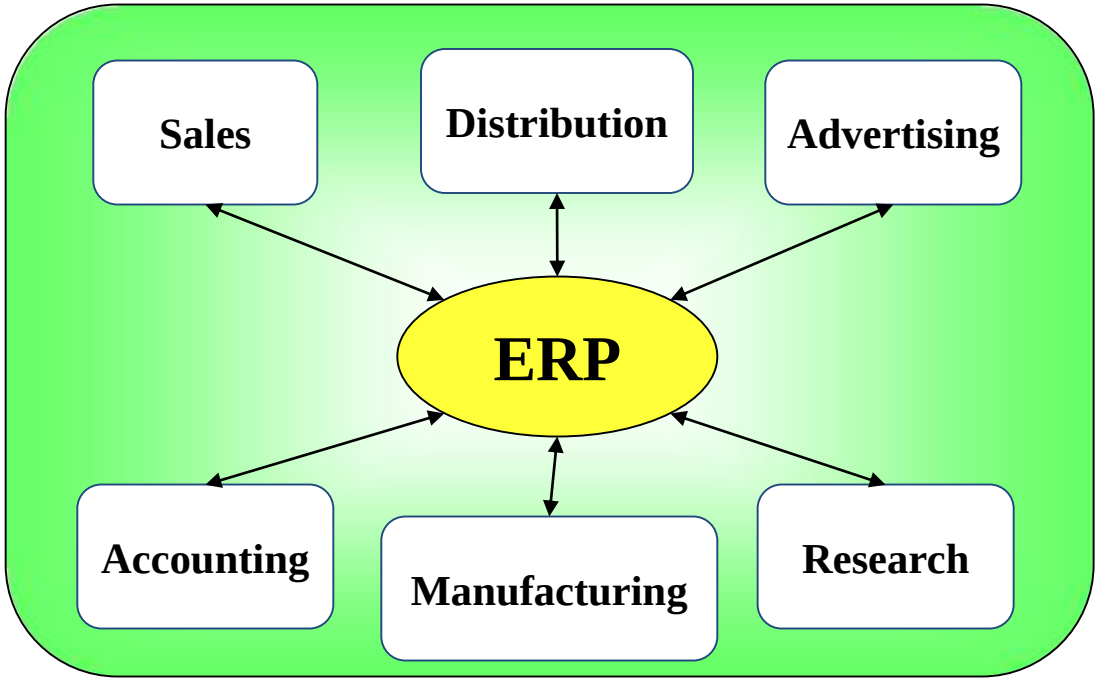
SME Classification vs. Implementation Patterns





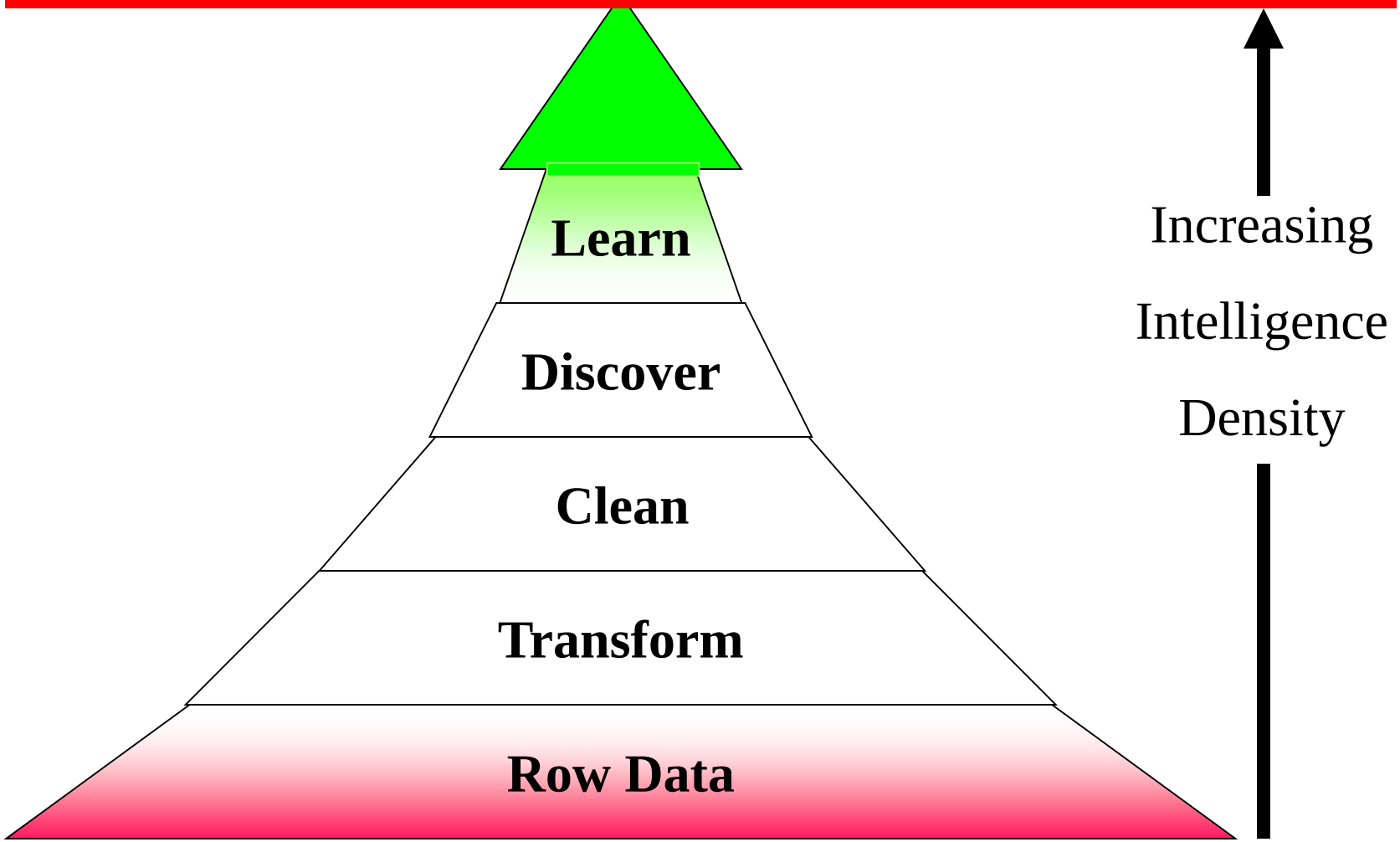
Through the uncovered implementation patterns, we can see clearer e-business implementation tendencies of the three clusters of SMEs than those using traditional SME classification, implying that SMEs should not be treated as homogenous groups.

Enterprise Resource Planing (ERP)



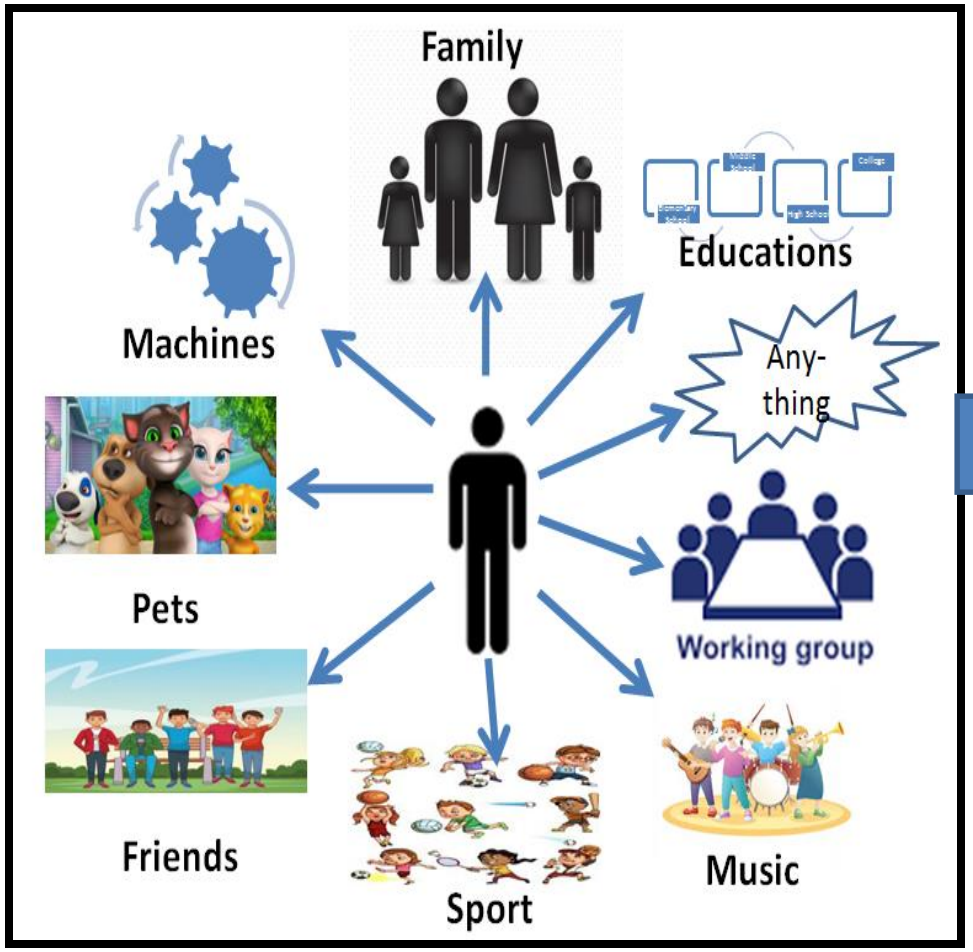
Membangun Intelligent System on top dari ERP untuk mendapatkan Kinerja yang Optimal kerjasama dan integrasi antar bagian dalam suatu Organisasi dengan menerapkan DS, BD, dan AI

Customer Relationship Management (CRM)

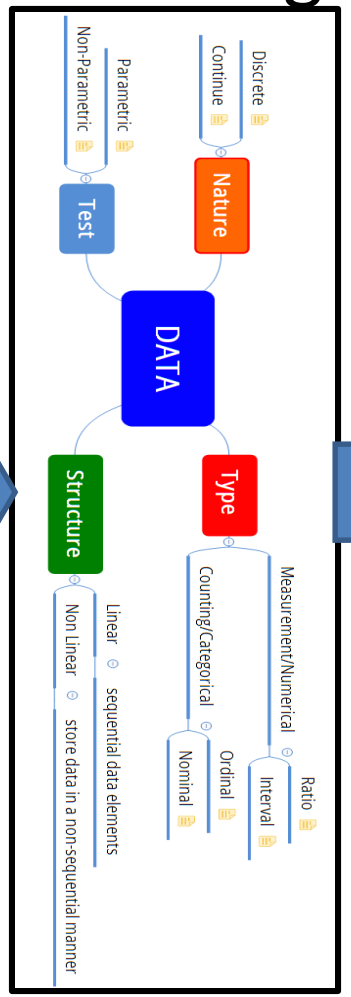


Membangun CRM dengan mengenai perilaku setiap individu customer dengan menggunakan face recognition yang diintegrasi dengan perilaku belanja seseorang, sehingga bisa dilakukan pendekatan personalisasi

Everything is Data (Multi-Stakeholders-ABGC): Universal Big Data and Artificial Intelligence (AI)



Raw Data Collection of Our Daily Life

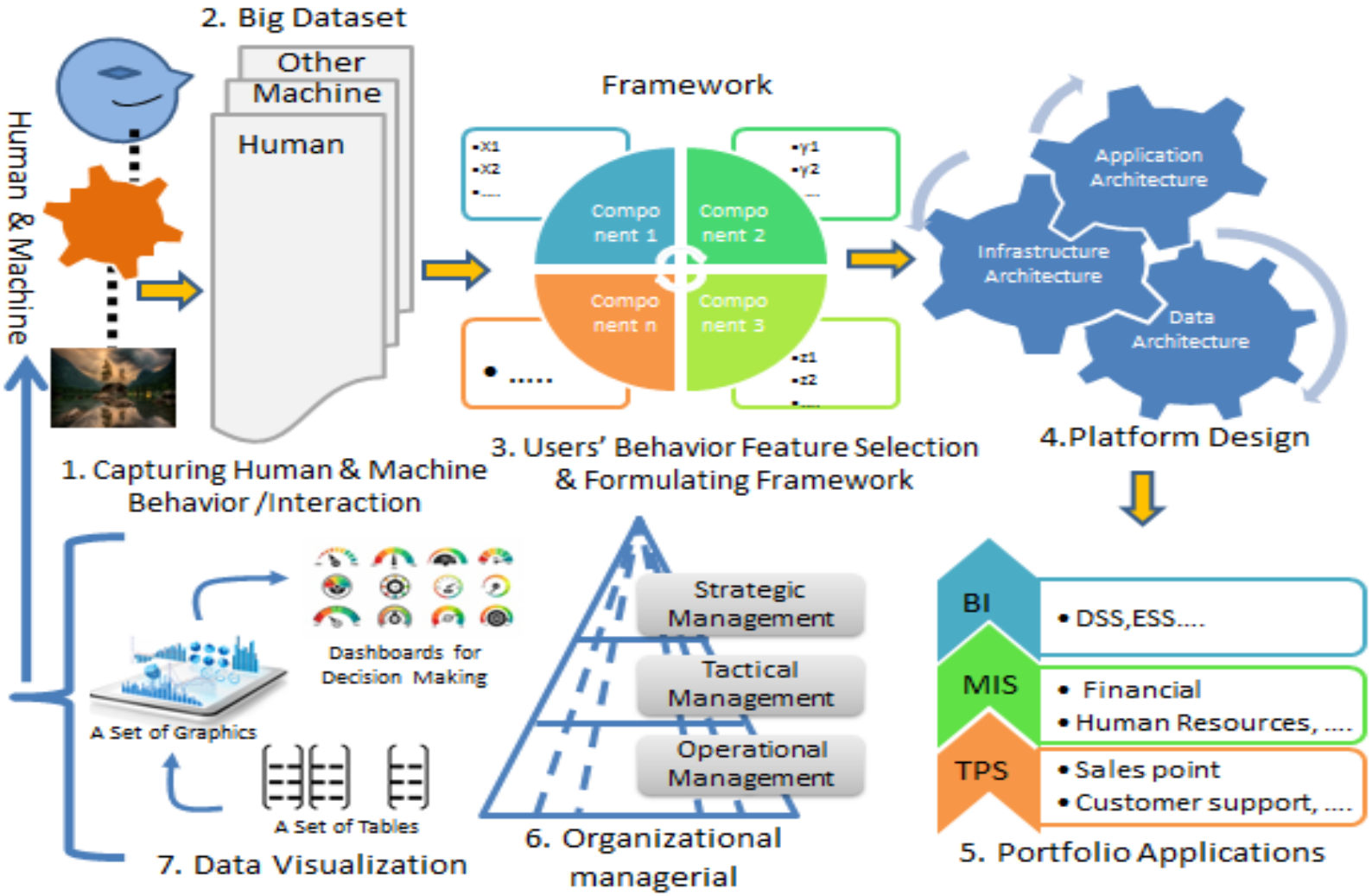


Data Organization

Data Computation to get pattern & insights

AI Applications: Machine Learning and Deep Learning etc.

Automation System Methodology: End-to-End Cycle of Pattern & Meaningful Insights



Hasibuan, IWBIS 2020

#7. Penutup

- Penggunaan Data Science, Big Data, dan Artificial Intelligence mendapatkan momentum yang baik oleh karena hadirnya teknologi 4.0 (IoT, Blockchain, Cloud Computing, dll.)
- Hasil Analysis Data Science dan Big Data menjadi input bagi AI untuk dijadikan bahan untuk melatih sistem agar sistem bisa "belajar" dan membantu pekerjaan manusia.
- Big Data menghubungkan DS dan AI.
- Teori, Teknik, dan Tools yang digunakan DS dan AI hampir sama, tetapi tujuannya berbeda. DS untuk data analytic, sedangkan AI untuk berpikir dan belajar seperti manusia (machine learning).
- Banyak peluang IPTEK-Bis bisa dikembangkan di Indonesia, mengingat secara etnis Indonesia ini sangat beragam, secara geografis sangat luas, dengan syarat data yang besar sudah harus dikumpulkan secara teratur.
- Big data adalah "bahan bakar" di era ekonomi digital dan DS serta AI adalah "engine" dari automation system yang menggunakan big data tersebut