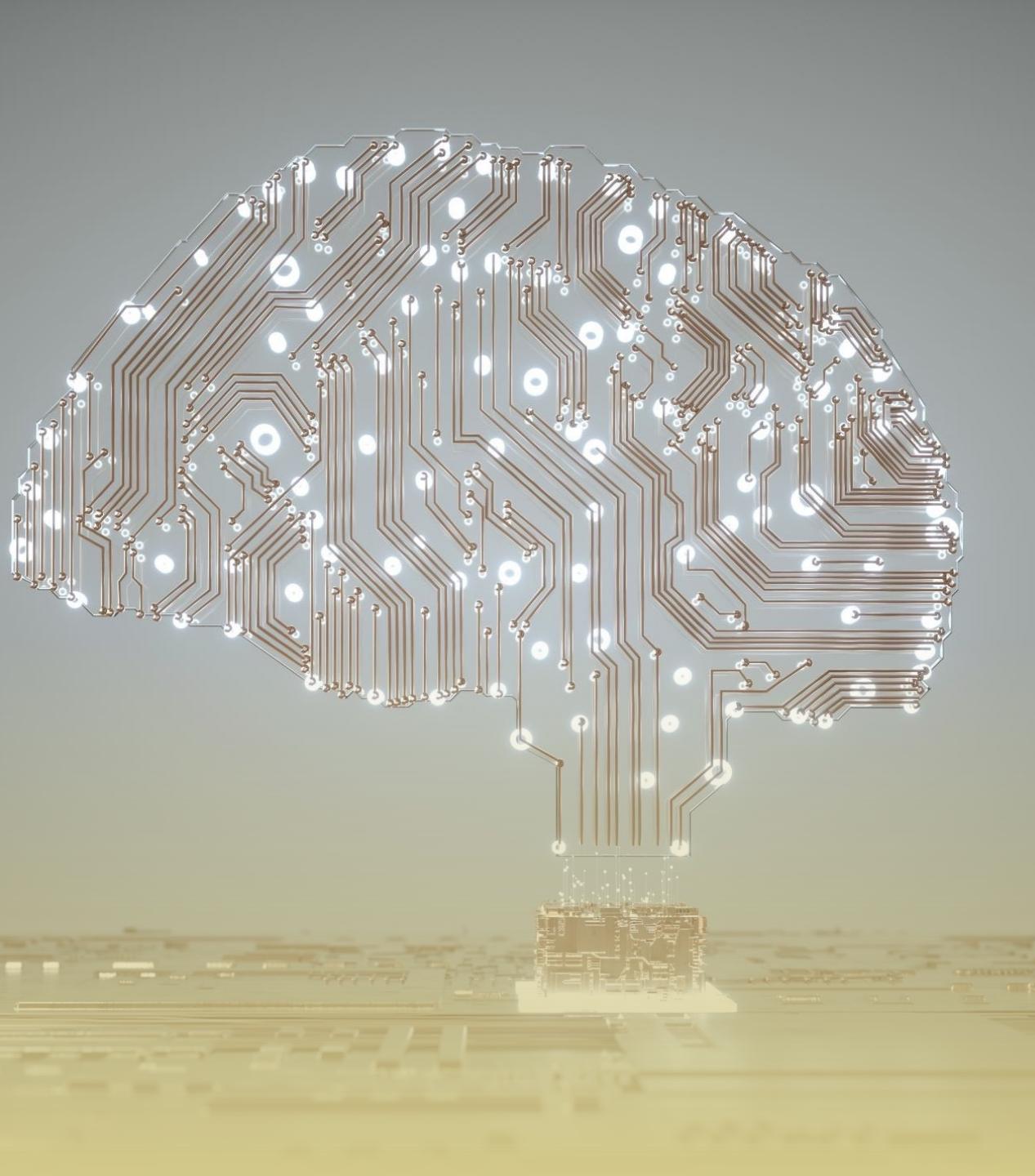


PENGANTAR DATA MINING

Nisa Hanum

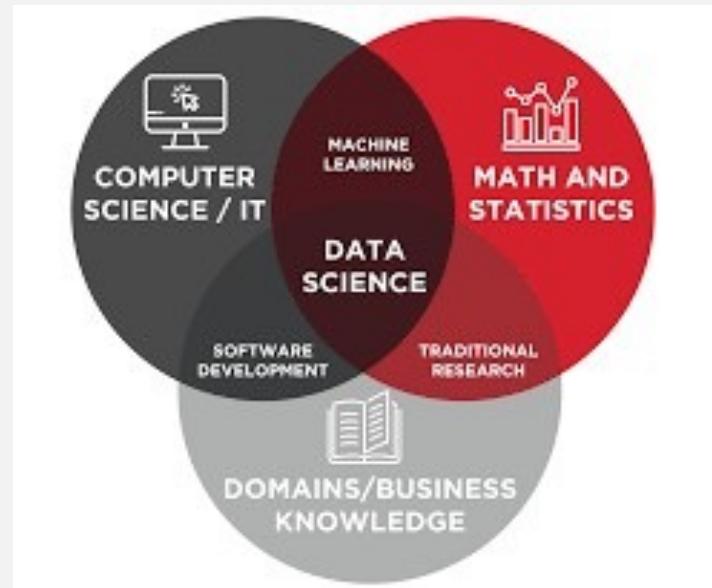


DAFTAR ISI

- Data Science (DS) vs Artificial Intelligence (AI)
- Hubungan antara AI dan DS
- Konsep, Teori, Teknik , dan Tools DS
- Konsep, Teori, Teknik , dan Tools AI.
- Aplikasi DS dan AI
- Penutup

APA ITU DATA SCIENCE DAN ARTIFICIAL

- **INTELLIGENCE ?**
- Data science adalah sebuah bidang interdisipliner yang menggunakan metode scientific, process, algoritma dan system untuk mengekstrak pengetahuan dan wawasan dari data yang terstruktur dan tidak terstruktur
- Data Science berkaitan juga dengan , data mining,machine learning dan big data

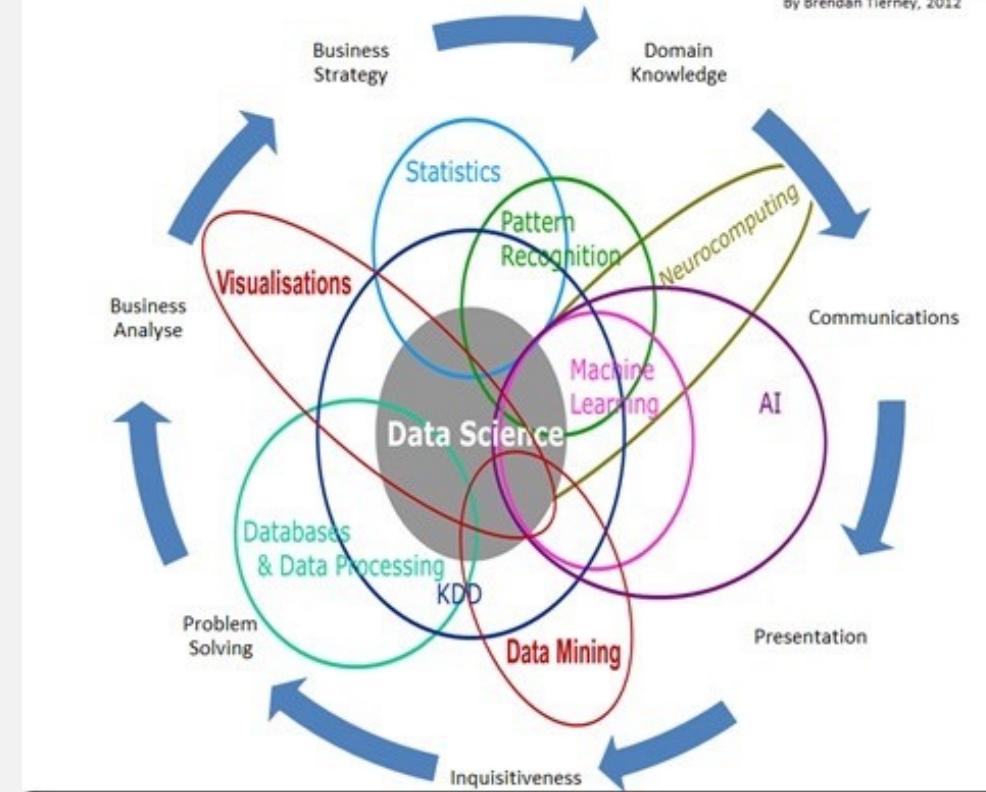


DATA SCIENCE ADALAH MULTIDIPLIN

- Bekerja baik dalam data terstruktur dan tidak terstruktur
- Bidang yang mencangkup segala sesuatu yang terkait dengan membersihkan data (*cleasing*), persiapan (*preparation*), dan akhirnya analisis data.
- Menggabungkan kemampuan pemrograman, penalaran logis matematika dan statistika

Data Science Is Multidisciplinary

By Brendan Tierney, 2012



DATA SCIENCE BISA DITERAPKAN PADA HAMPIR SEMUA DISIPLIN

Astronomi

Bisnis

Kedokteran

Ekonomi

Olahraga

Cuaca

Financial

...

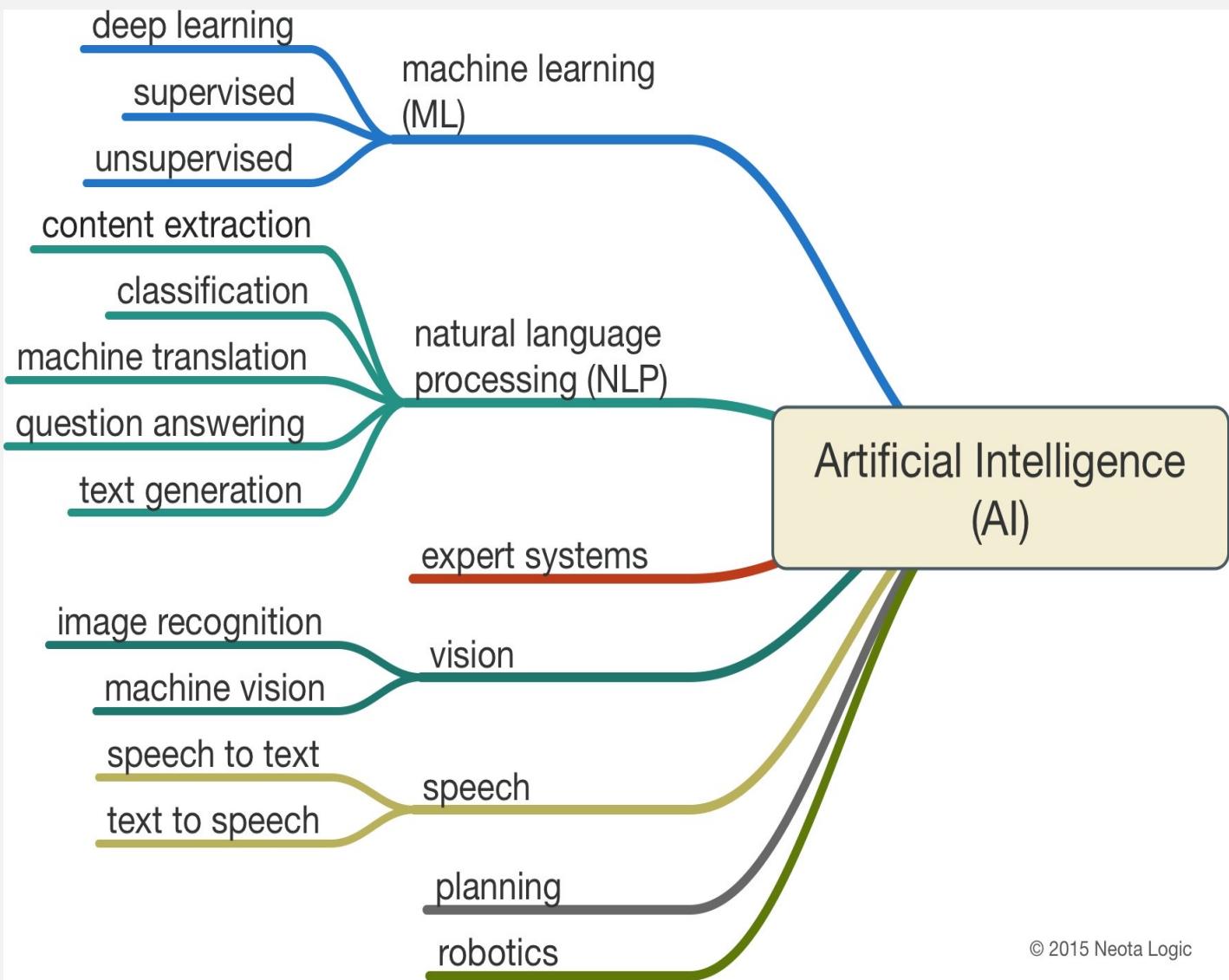


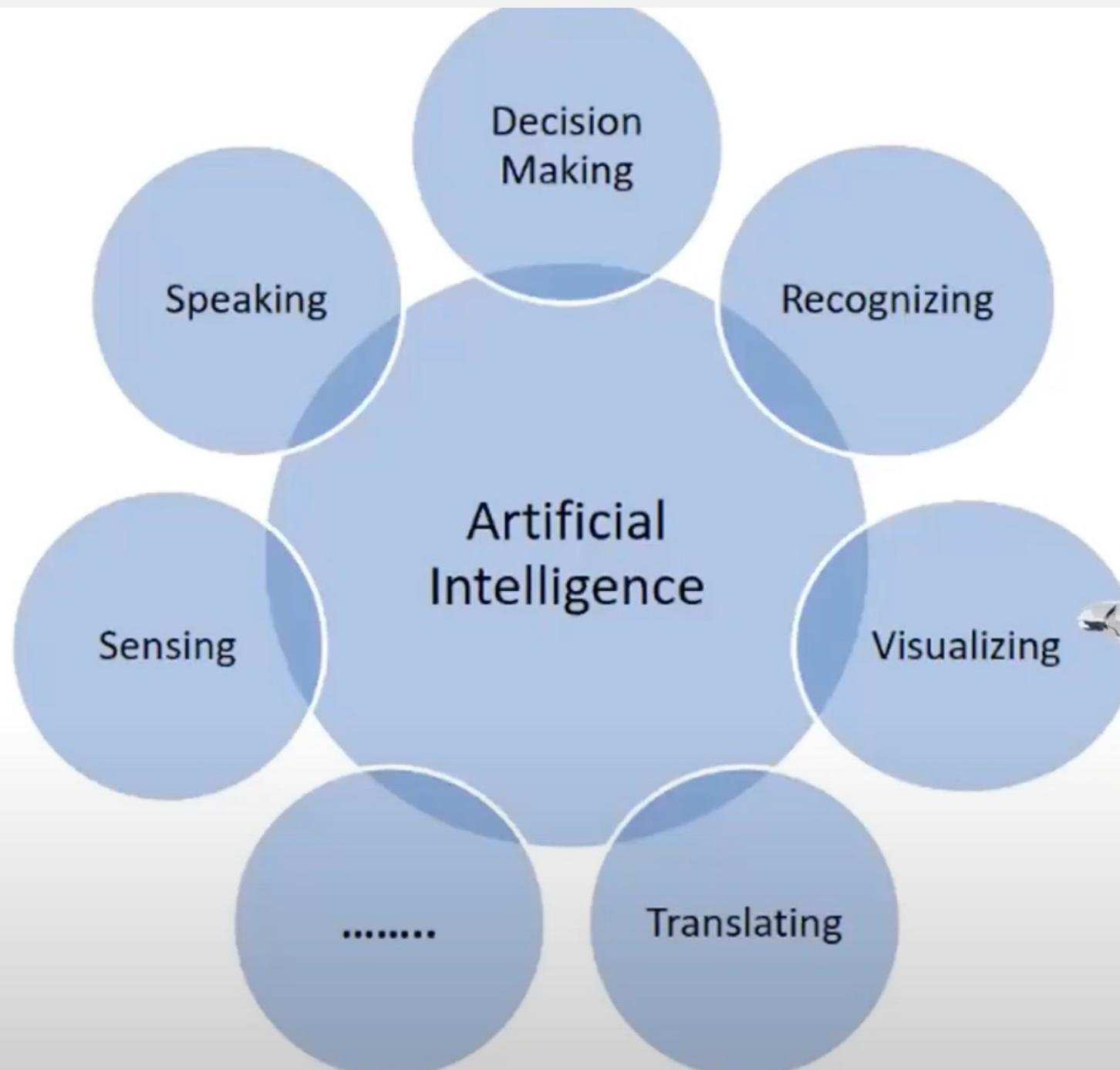


APA ITU KECERDASAN BUATAN (AI)

- AI adalah sebuah bidang studi yang mencoba untuk melatih computer sehingga computer dapat melakukan hal-hal yang saat ini dapat dilakukan manusia dengan lebih baik
- Komputer dilatih untuk belajar sehingga dapat melakukan pengenalan citra secara otomatis(visual perception), speech recognition, pengambilan keputusan, dan menterjemahkan bahasa

SCOPE OF ARTIFICIAL INTELLIGENCE (AI)

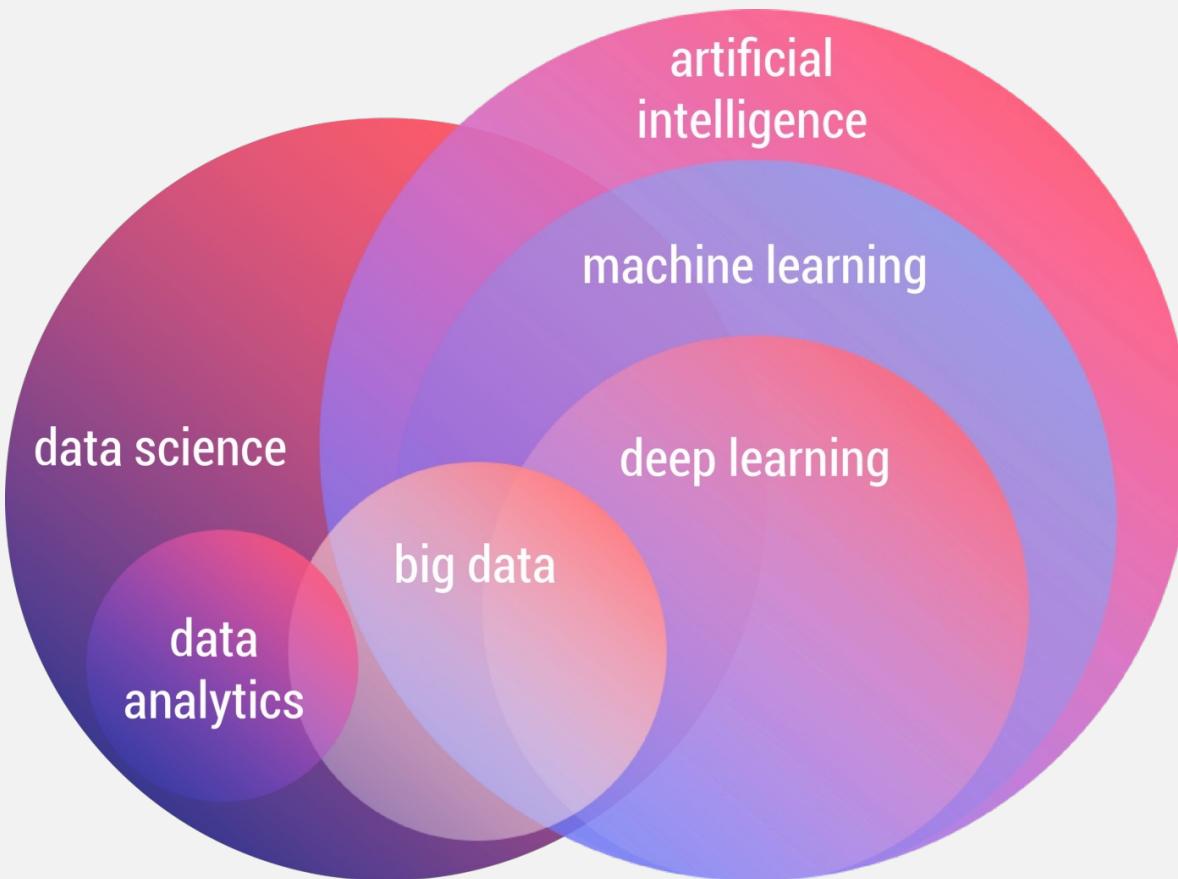




AI BISA BERFIKIR
SEPERTI LAYAKN YA
MANUSIA BERFIKIR

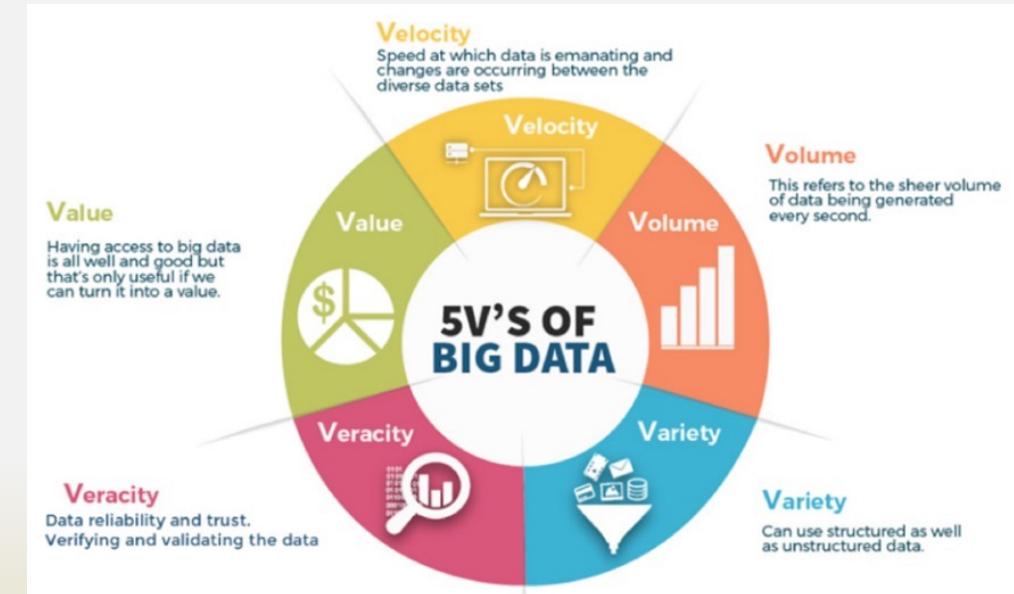
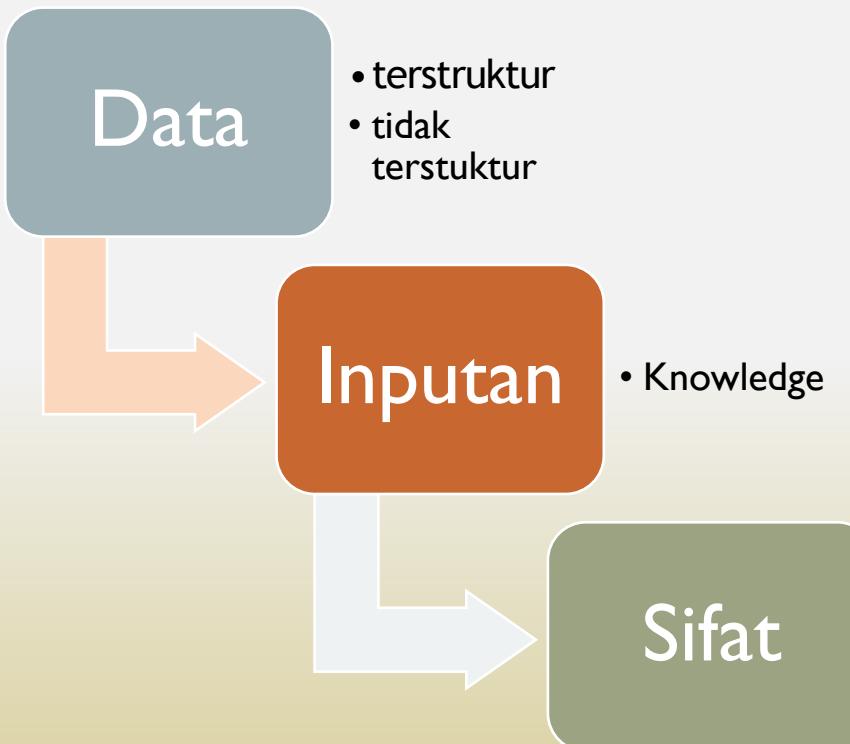


RELASI DS, BIG DATA , DAN AI



BIG DATA ERA

Karakteristik AI



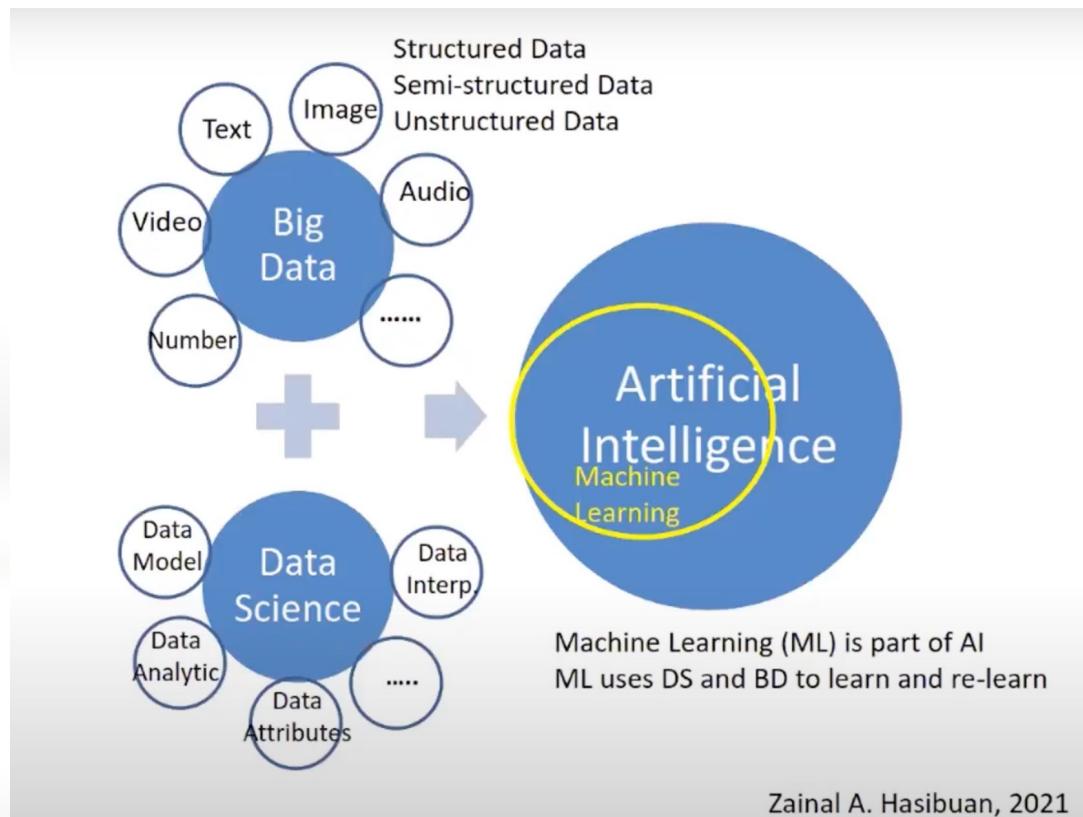
Zainal S, Hasibuan 2021

BIG DATA

- Di era teknologi yang semakin berkembang ini kita menganalisis LEBIH BANYAK data
- Big data “**Is The oil of the digital economy**”
- Organisasi yang memiliki data dalam jumlah besar akan mendapatkan keunggulan kompetitif dibidangnya.
- DS , Big Data dan AI memainkan peran penting untuk menyajikan solusi.
- Semua memanfaatkan model matematika untuk mengidentifikasi, mengklasifikasi, mengelompokan, memprediksi dan mempelajari dan mengolah data

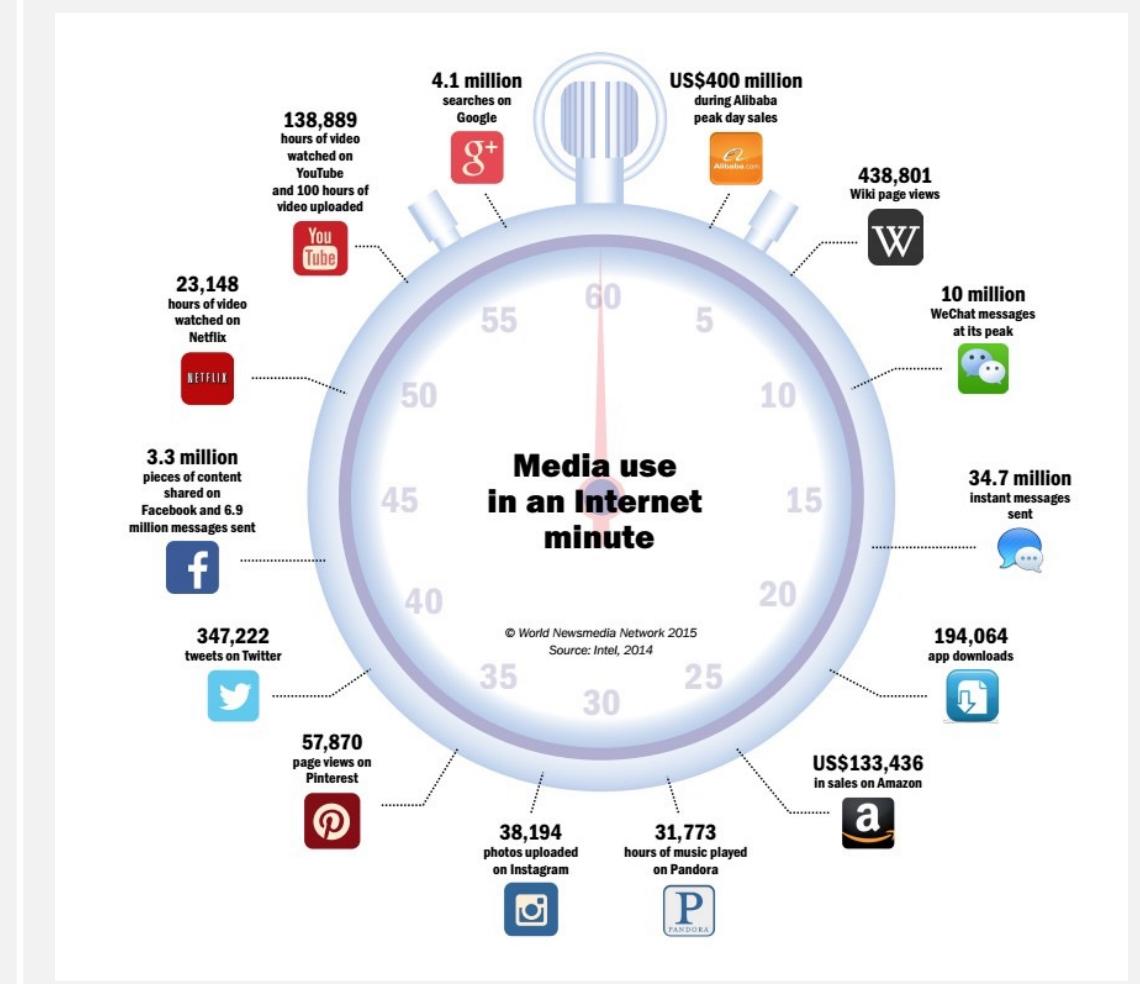


HUBUNGAN AI + BIG DATA+ DS

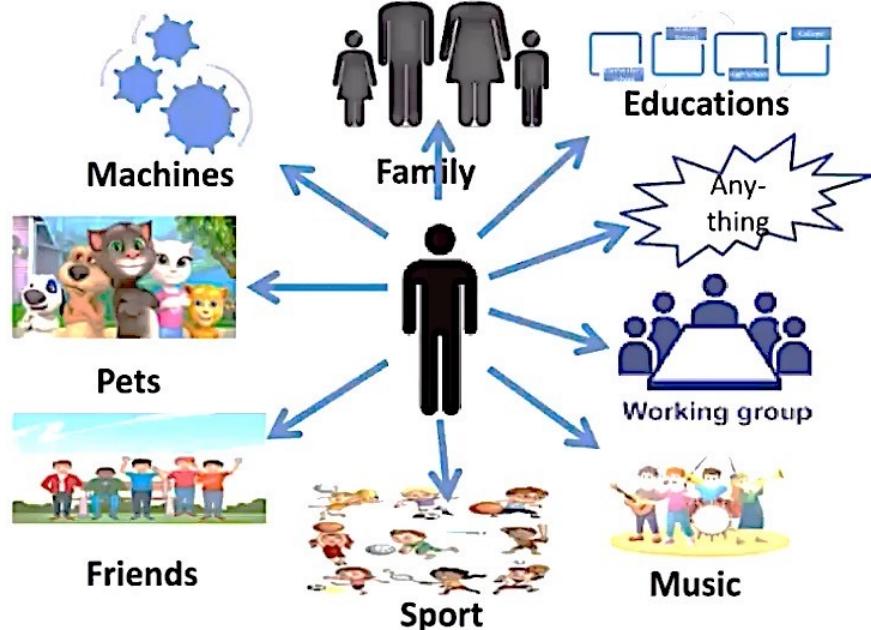


KONSEP DATA SCIENCE

- Di era digital ini,hampir semua kegiatan manusia, interaksi manusia dan lingkungannya dan lingkungan itu sendiri bisa terekam dalam bentuk digital,atau bisa di digitalisasikan sehingga terbentuk dalam data
- Dapat dikatakan “everything is data”



Capturing Big Data to become one's profile



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

How Does Google Track You?



THE FACTS OF DATA TODAY | BIG DATA

UNSTOPPABLE FLOWS OF DATA



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

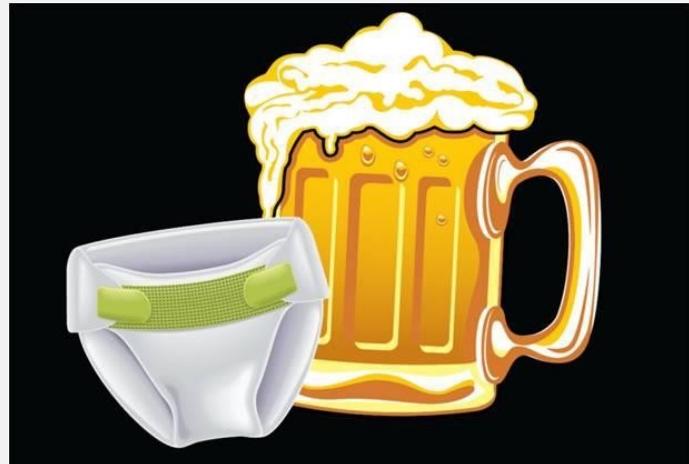


VALUABLE MEANING (INFORMATION)

BAGAIMANA CARA ?
MENGOLAHNYA

Source R. Eko Indrajit, 2021

WHY DATA IS GOLD?

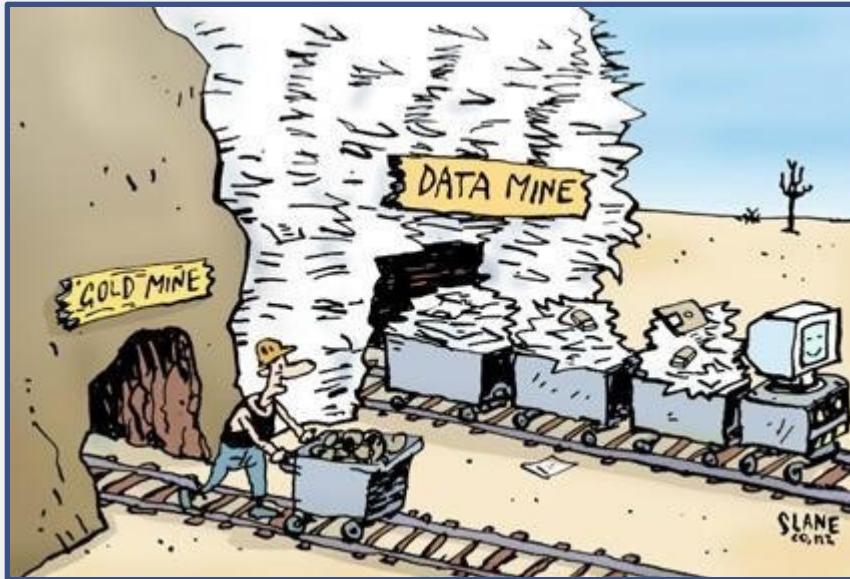


- Wal-mart discovered through data mining that the sales of **diapers and beer** were correlated on friday nights.
- It determined that the correlation was based on working men who had been asked to pick up **diapers** on their way home from work.
- AN example of Data mining!

LIKE GOLD MINING



Like digging gold from a
mine...



Finding important information from
a huge reservoir of data!!

4 HAL UNTUK MENJADI DATA SCIENTIST



```
var y = o.top;
var ax = settings.accX;
var ay = settings.accY;
var th = t.height();
var wh = w.height();
var tw = t.width();
var ww = w.width();

if (y + th + ay >= b &&
    y <= b + wh + ay &&
    x + tw + ax >= a &&
    x <= a + ww + ax) {

    //trigger the custom event
```

Programming Skills

Ada berbagai bahasa pemrograman, tetapi yang umum digunakan adalah **R dan Python**. Kedua bahasa pemrograman ini cukup mudah digunakan oleh orang-orang yang tidak memiliki background programming. Kedua bahasa pemrograman ini juga memiliki platform seperti **Rstudio untuk (R), Jupyter notebook, Pycharm, dan spider untuk (Python)**.



Math and Statistics

Merupakan acuan dan pondasi dari pemodelan data science yang dibuat. R dan Python memiliki library khusus yang akan **memudahkanmu dalam membuat pemodelan tanpa melibatkan rumus matematika secara manual.**



Business Analytics

Selain programming dan matematika, kemampuan analisis yang kuat juga dibutuhkan untuk menjadi Data Scientist. Hal itu dibutuhkan untuk **menentukan variabel serta faktor yang harus digunakan dalam pemodelan agar hasil pemodelan bisa membuat keputusan yang lebih tepat.**



Portofolio

Rekruter cenderung lebih memilih untuk melakukan hiring kepada applicants yang memiliki portofolio. Karena hal itu akan **membantu kamu dalam membuktikan keahlian yang kamu miliki.**

Platform yang dapat kamu gunakan untuk publikasi portofoliomu antara lain **github, medium, dan kaggle.**

REFERENSI

- www.algorit.ma

TUGAS PERORANGAN

- Cari 3 jurnal Minimal Q2 dan berbahasa inggris (ex; Elsevier), max 5 tahun kebelakang
 - 3. Tantangan dan peluang dari keilmuan data science
 - 2. Trend Keilmuan data science
 - 1. Study kasus dari penerapan data Science (Harus ada repository data setnya)

Buat Paper review nya.

Format IEEE

Yg dikumpulkan (paper review dan jurnal-jurnal acuannya)

Dikumpulkan paling lambar minggu depan H-1 .