Decentralized Machine Learning Data Sharing and modeling

by ***Gopala KR ,*** ***7th March-2018***

Dell EMC |Accenture BDC7A

Agenda

PART 1 : AI brief Overview

- In the extremes
- General intelligence
- Artificial intelligence-examples
- History
- 1 lever deeper
- Realtime use cases
- Enterprise use cases
- Potential of AI- what it can do in future
- Research-consumer tech-VC research-datascientists

PART 2 : The Problem/business opportunities

- Current issues
- Current state of the problems

PART 3 : The Solution/business model

- Smart contratcs – blockchain tech
- Architecture

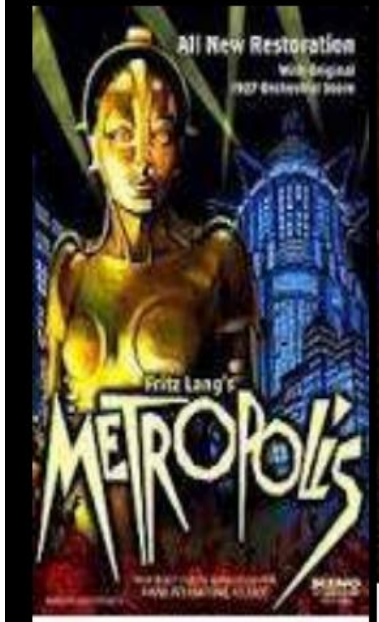
PART 4 : Competitive Analysis

- Why this is importatnt to accenture?

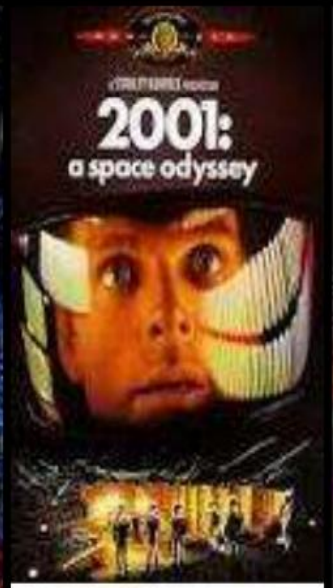
Artificial Intelligence Overview

What is AI?

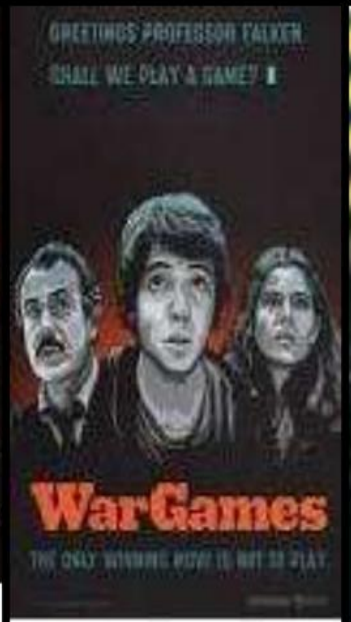
In the extreme…



Metropolis
1927

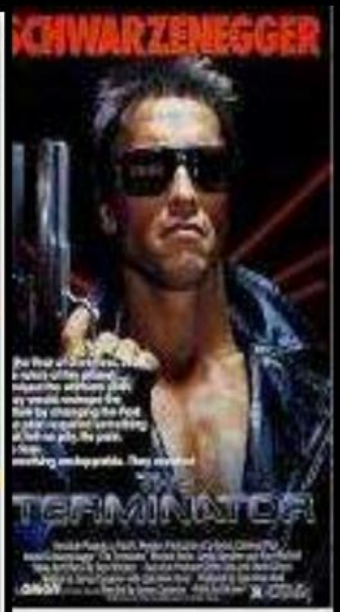


2001: A Space Odys...
1968



WarGames
1983

Blade Runner
1982



The Terminator
1984

The Matrix
1999

Ex Machina
2015

This guy is awesome

What is Human Intelligence?

It's a composition of abilities like:

Clip slide

What is Human Intelligence?

It's a composition of abilities like:

LEARNING

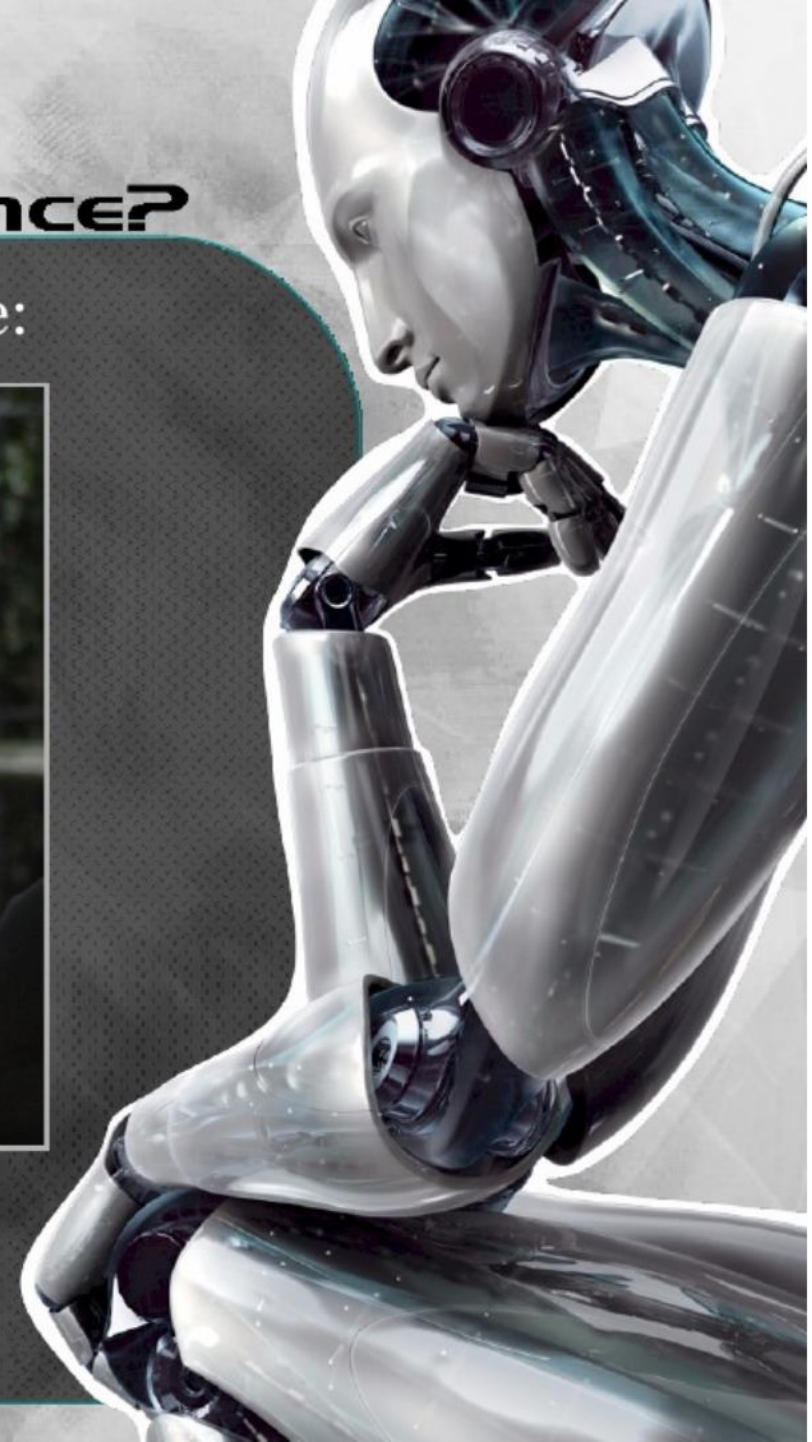
Clip slide

Apple iPhone
Select to choose what happens with this device.

What is Human Intelligence?

It's a composition of abilities like:

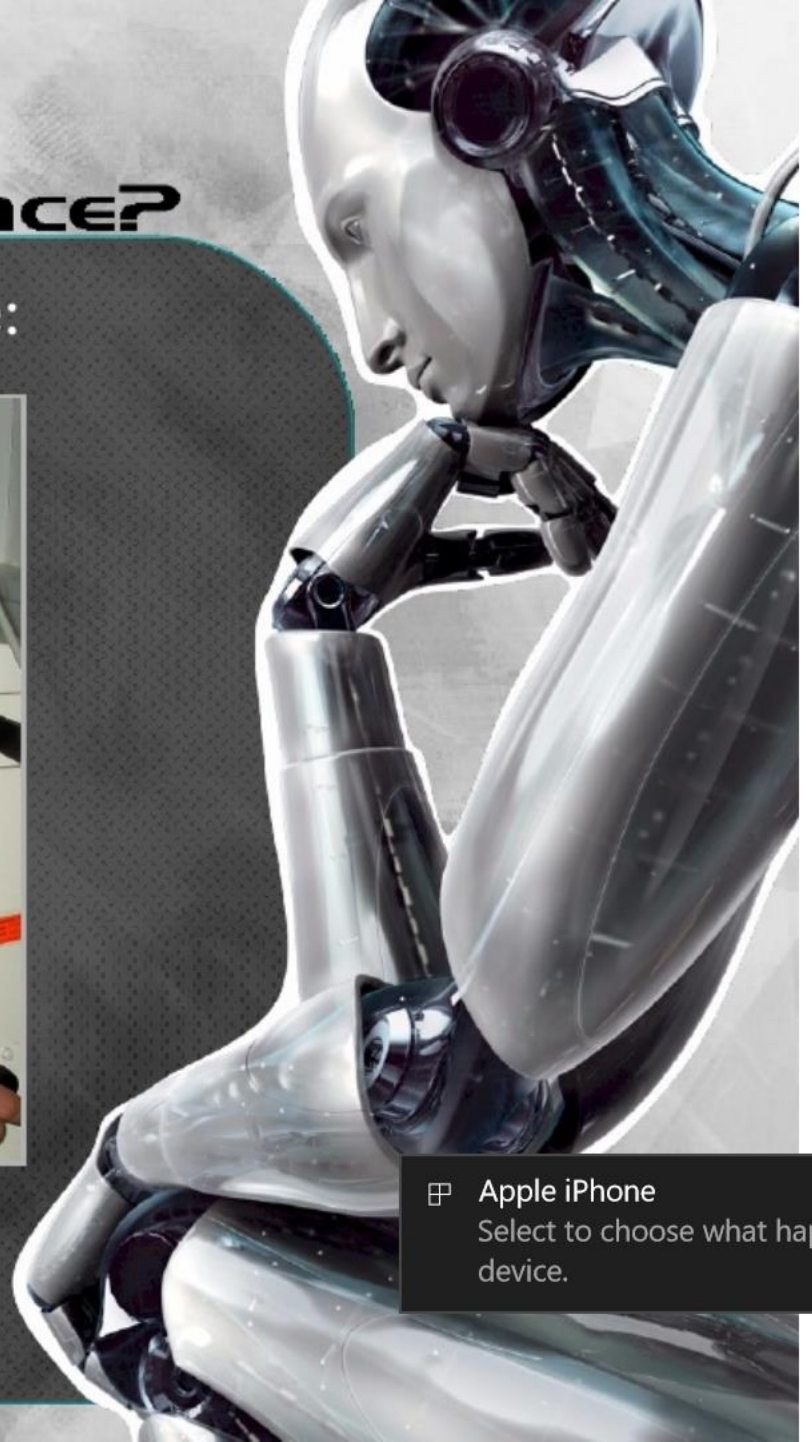
REASONING



What is Human Intelligence?

It's a composition of abilities like:

PERCEIVING



Clip slide

Apple iPhone
Select to choose what happens with this device.

What is Human Intelligence?

It's a composition of abilities like:

UNDERSTANDING OF LANGAUGE

What is Human Intelligence?

It's a composition of abilities like:

FEELING

Clip slide

Apple iPhone
Select to choose what happens with this device.

What is Intelligence Composed of?

The intelligence is intangible. It is composed of −

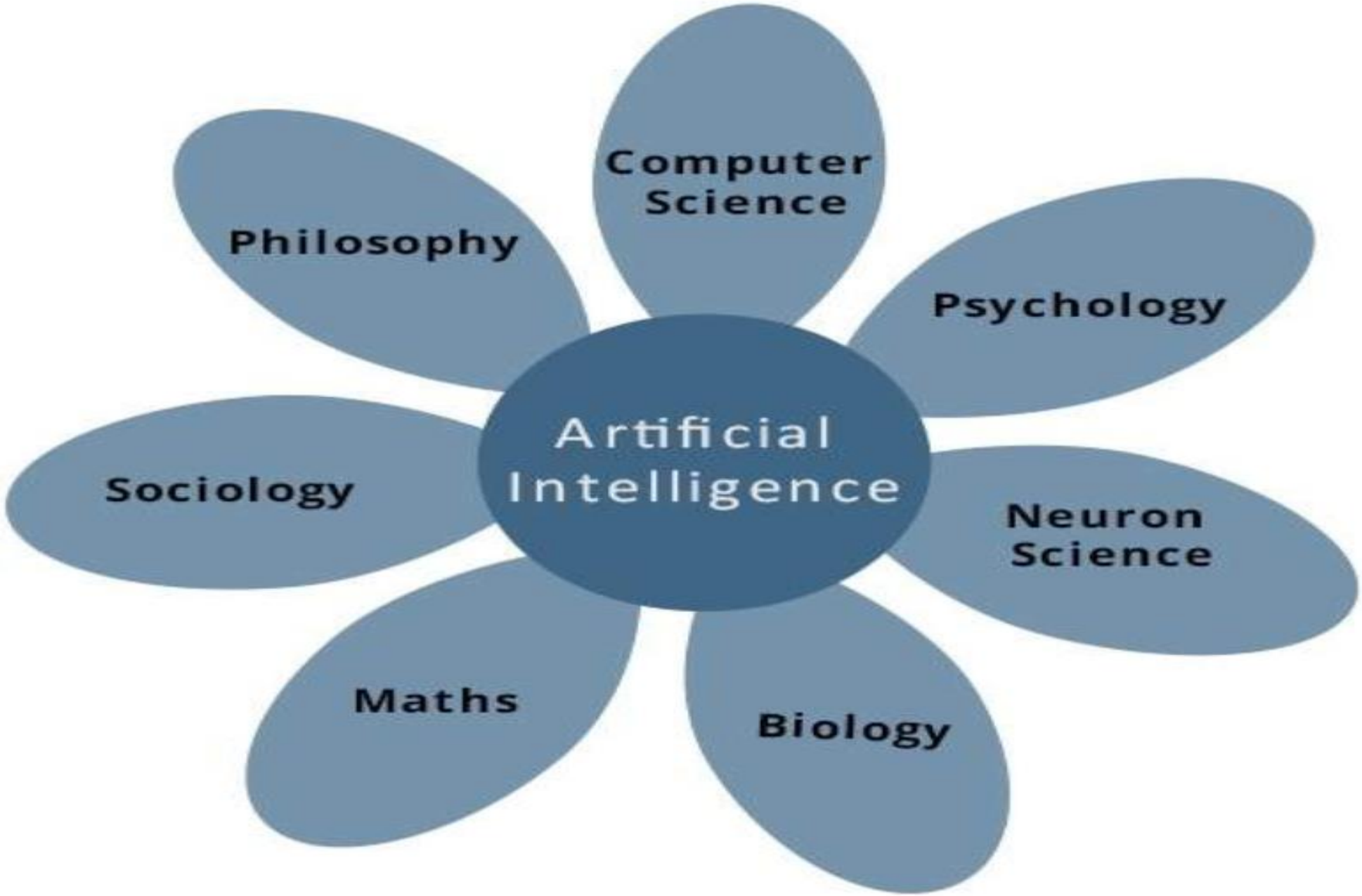
- Reasoning
- Learning
- Problem Solving
- Perception
- Linguistic Intelligence

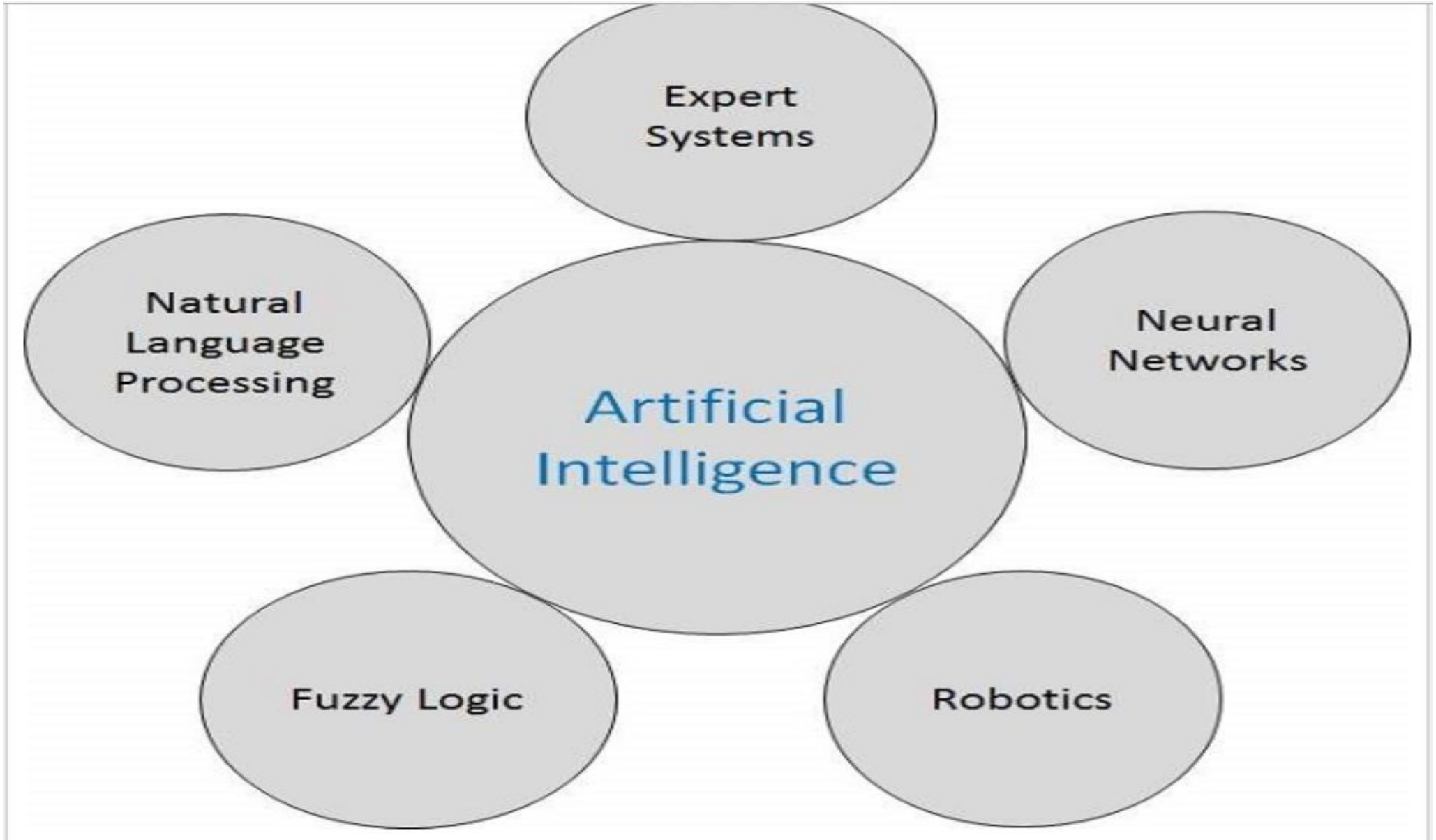
Types of Intelligence

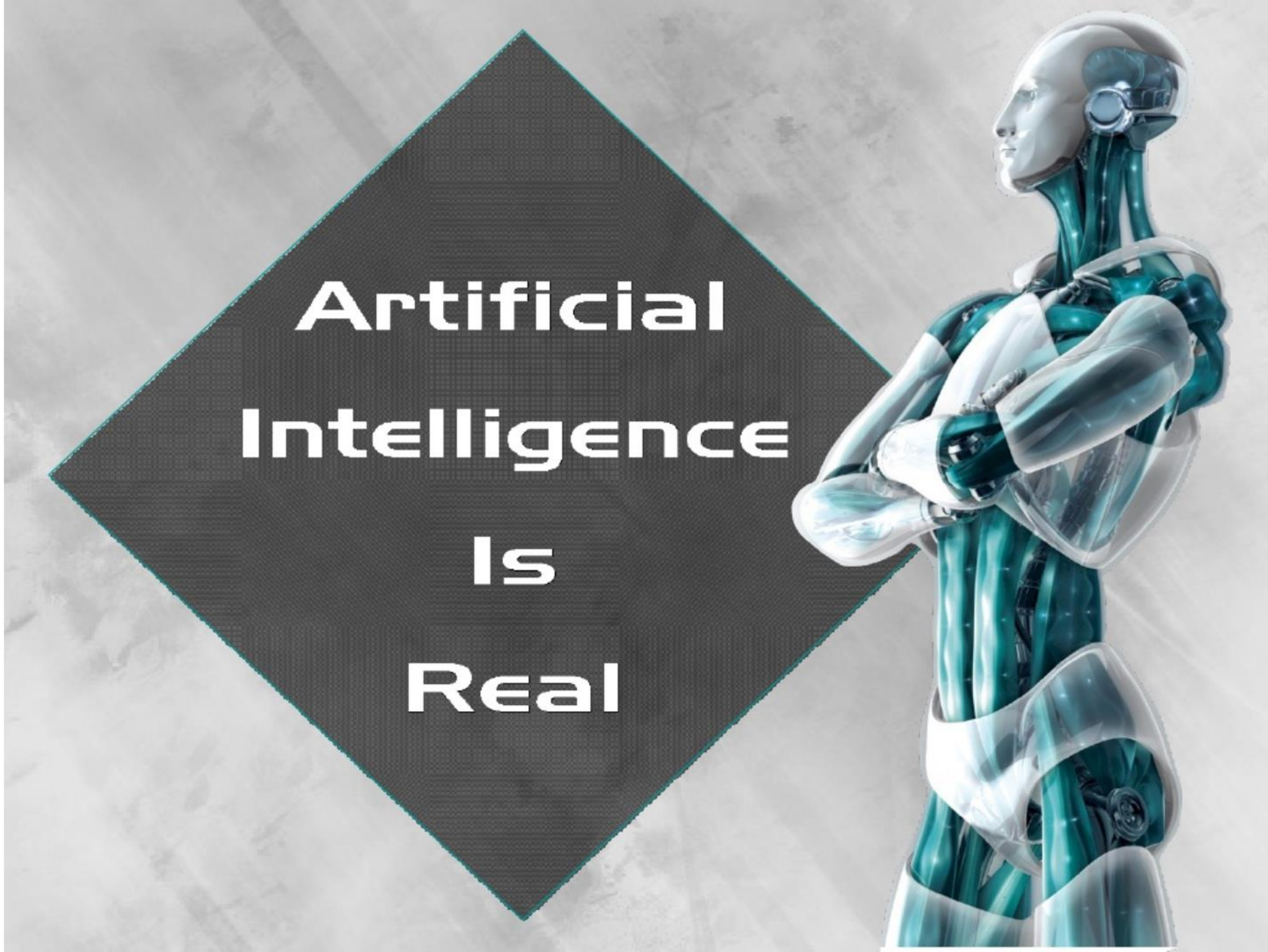
As described by Howard Gardner, an American developmental psychologist, the Intelligence comes in multifold –

Intelligence	Description	Example
Linguistic intelligence	The ability to speak, recognize, and use mechanisms of phonology (speech sounds), syntax (grammar), and semantics (meaning).	Narrators, Orators
Musical intelligence	The ability to create, communicate with, and understand meanings made of sound, understanding of pitch, rhythm.	Musicians, Singers, Composers
Logical-mathematical intelligence	The ability of use and understand relationships in the absence of action or objects. Understanding complex and abstract ideas.	Mathematicians, Scientists

Spatial intelligence	The ability to perceive visual or spatial information, change it, and re-create visual images without reference to the objects, construct 3D images, and to move and rotate them.	Map readers, Astronauts, Physicists
Bodily-Kinesthetic intelligence	The ability to use complete or part of the body to solve problems or fashion products, control over fine and coarse motor skills, and manipulate the objects.	Players, Dancers
Intra-personal intelligence	The ability to distinguish among one's own feelings, intentions, and motivations.	Gautam Buddhha
Interpersonal intelligence	The ability to recognize and make distinctions among other people's feelings, beliefs, and intentions.	Mass Communicators, Interviewers


Computer Science
Philosophy
Psychology
Artificial Intelligence
Sociology
Neuron Science
Maths
Biology


Expert Systems
Natural Language Processing
Artificial Intelligence
Neural Networks
Fuzzy Logic
Robotics


Artificial
Intelligence
Is
Real

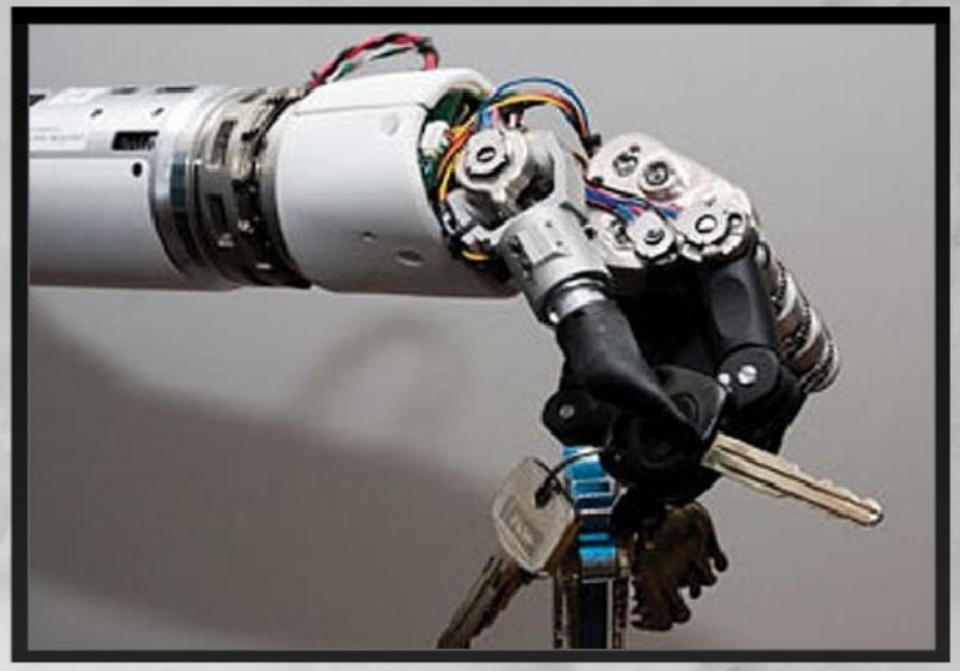
AI IS REAL!

Space Program

The Mars Lander:
Being able to Navigate on the Red Planet, the robot arm has been digging in the Martian Soil & Ice for 2.5 hours a day.

AI IS REAL!

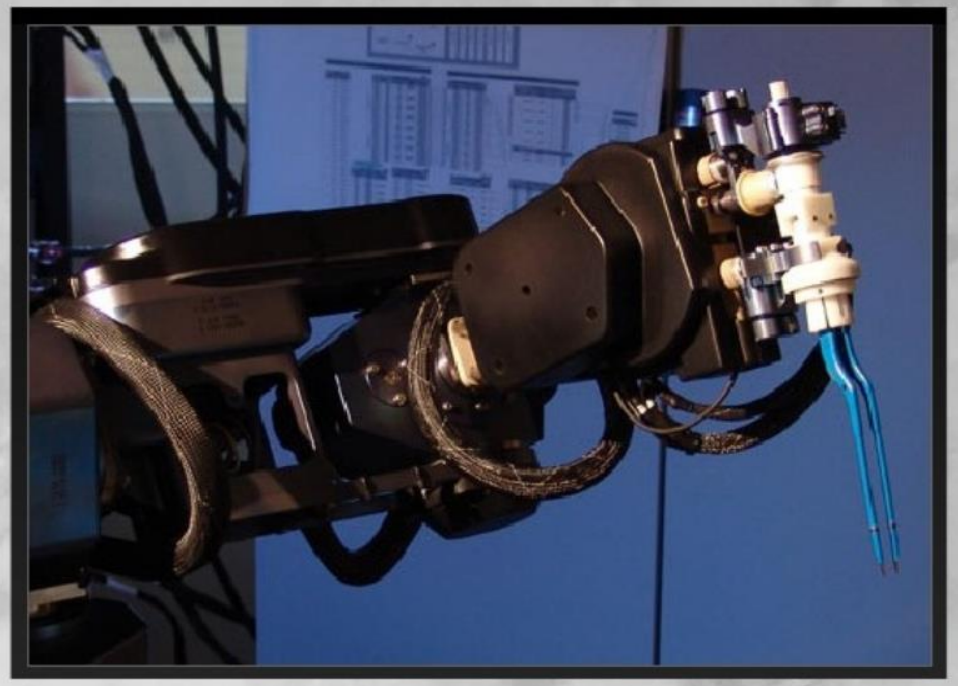
Prosthetic Limbs



Mind-controlled prosthetic arm :
The user controls the arm through existing nerves and it is sensitive enough to pick up even a piece of paper!

AI IS REAL!

Surgical System



Neuroarm :
It is a surgical robotic system, that can do delicate brain surgery, also enabling physicians to manipulate tools at microscopic levels.

AI IS REAL!

All-in-one Floor Cleaner

iRobot :
It performs the duties of a vacuum cleaner, floor washer and gutter cleaner.

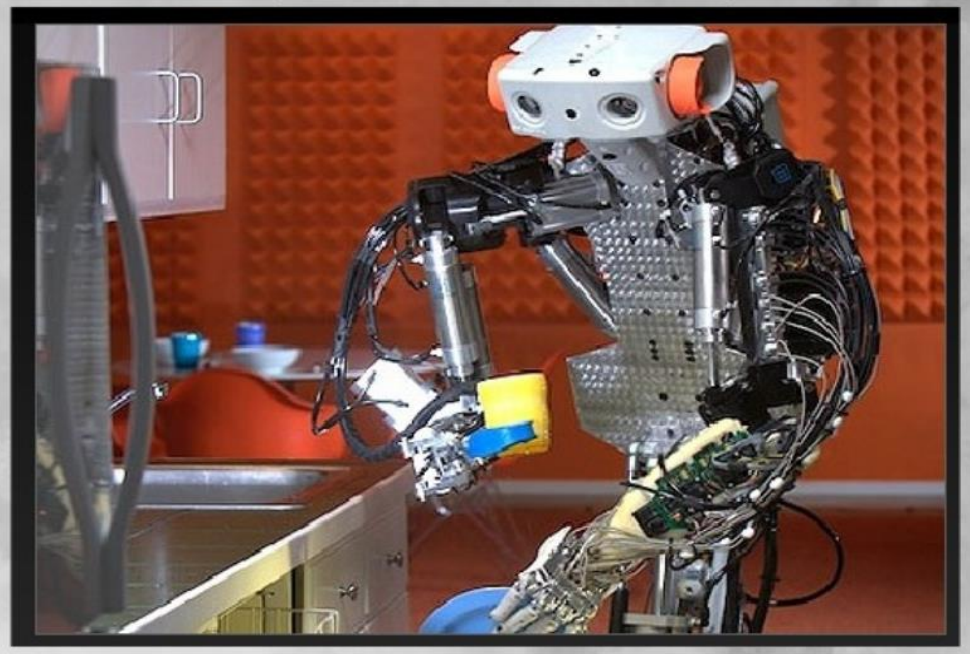
AI IS REAL!

Your Digital Secretary

Ultra Hal Assistant :
A digital secretary that uses AI to understand spoken English commands and to learn over time..

AI IS REAL!

Walk Like a Man



Dexter :
Having flexible joints, driven by air cylinders, it can walk, run and even jump.

Military Technology

Aiming, Defending, and Moving are done simultaneously by us.

But with AI technology, the tank will be able to auto-detect enemy units and take action.

Space Exploration:

Humans need provisions for: Water, Recycling, Food, Space Radiation and Psychological Issues of confinement.

None of these issues concern a robot.

Medical Research

AI Technology can perform delicate operations more precisely and efficiently.

A brief history of AI

"The term artificial intelligence was first coined by John McCarthy in 1956 when he held the first academic conference on the subject."

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made

A brief History of AI

What Is AI?

It has grown up out of many different disciplines and has many definitions. Here is ours.

Artificial intelligence, as we use the term in this report, is an evolving constellation of technologies that enable computers to simulate elements of human thinking — learning and reasoning among them. Regular improvements in Google's search algorithm, for example, come from machine learning, a type of AI that programs systems to learn from data, find patterns in it, and make predictions about it. The same technology has been pivotal to voice and image recognition as well as advances in self-driving cars. Integral to many recent improvements in the field is a form of machine learning called deep learning. Loosely modeled on the way neurons and synapses in the brain change as they are exposed to new input, it has been used independently or in combination with other AI approaches to help machines tackle tricky tasks and exhibit something resembling intuition, in some cases performing tasks better than humans.

A Time Line of AI

After a century of ups and downs, artificial intelligence is getting smarter.

1914

In what would come to be described as the world's first computer game, Spanish inventor Leonardo Torres y Quevedo debuts El Ajedrecista, a machine that can automatically play chess thanks to a simple algorithm built into its mechanical design.

1943

Neuroscientist Warren McCulloch and logician Walter Pitts present a calculus based on neuron-like "logic units" that can be connected together in networks to model the action of a real brain.

1950

In a paper that helps establish a practical goal for artificial-intelligence research, Alan Turing proposes a game to answer the question "Can machines think?" He predicts that by 2000 computers will be able to pass as human more than 30 percent of the time.

1956

John McCarthy, Marvin Minsky, and Claude Shannon organize a summertime research meeting at Dartmouth that brings together the leading thinkers on information theory, artificial neural networks, and symbolic logic, christening the field "artificial intelligence."

1958

Oliver Selfridge presents a paper in England describing Pandemonium, a new model of a neural network based on lower-level "data demons" working in parallel with higher-level "cognitive demons" in order to perform pattern recognition and other tasks.

1960

Frank Rosenblatt demonstrates the Mark I Perceptron, an attempt to create an artificial neural network for image recognition that the New York Times calls the first step toward a computer "able to walk, talk, see, write, reproduce ... and be conscious of its existence."

1961

Marvin Minsky publishes his foundational paper, "Steps Toward Artificial Intelligence."

A brief History of AI

2000

Cynthia Breazeal designs a sociable humanoid robot named Kismet that is able to express emotion and recognize cues from interaction with humans.

1987

Ernst Dickmanns and collaborators equip a Mercedes van with video cameras, microprocessors, and other electronics to demonstrate autonomous driving at almost 60 miles per hour. After much other AI research falls short, DARPA cuts the project's budget.

1979

A backgammon program developed by Hans Berliner defeats the reigning world champion in a match, the first time a computer has defeated a champion-level competitor in an intellectual game.

1966

Joseph Weizenbaum demonstrates ELIZA, the world's first chat program, which is able to converse using a series of preprogrammed phrases, sometimes to comic effect.

2004

DARPA sponsors its first "Grand Challenge," which pits research teams against each other to design driverless vehicles capable of independently traversing the Mojave Desert.

1997

IBM's Deep Blue chess computer avenges its prior defeat to world champion Garry Kasparov in a tense match commemorated in a documentary film, The Man vs. the Machine.

1984

Douglas Lenat begins the Cyc project, an ambitious attempt to create a common-sense knowledge base that can eventually become self-educating. Little progress is seen for decades.

1972

AI takes a hit when philosopher Hubert Dreyfus publishes "What Computers Can't Do," a manifesto challenging the predictions of AI researchers, and scientist James Lighthill pens a pessimistic review of progress in AI research in the U.K., leading to funding cuts.

2011

IBM's Watson defeats Jeopardy! champions Ken Jennings and Brad Rutter in a televised two-game, three-night face-off that ends with the computer amassing more than three times the winnings of its human competitors.

2012

A team from Geoff Hinton's lab wins the ImageNet Large Scale Visual Recognition Challenge with deep-learning software that could within five guesses identify a thousand types of objects about 85 percent of the time, a huge improvement in accuracy.

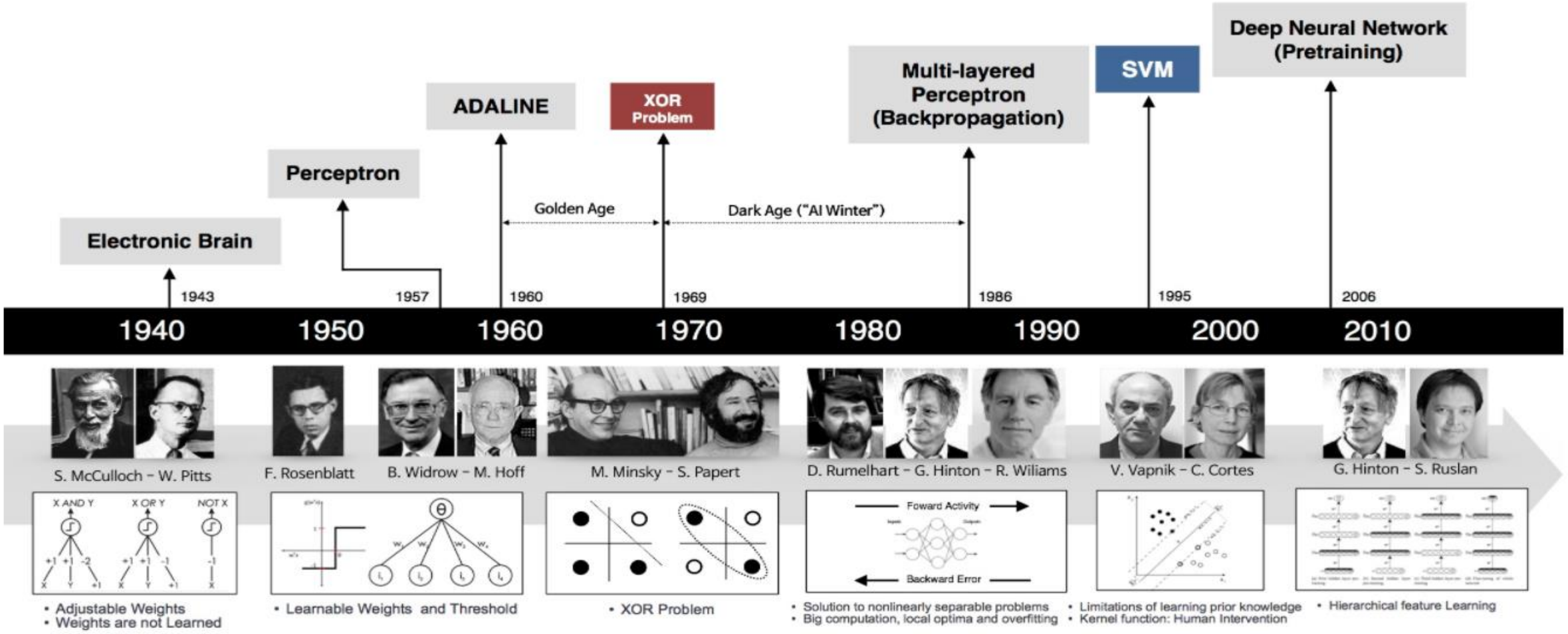
2014

Google acquires DeepMind Technologies, a small London-based startup focused on deep learning, a relatively new field of artificial intelligence that aims to achieve tasks like recognizing faces in video or words in human speech.

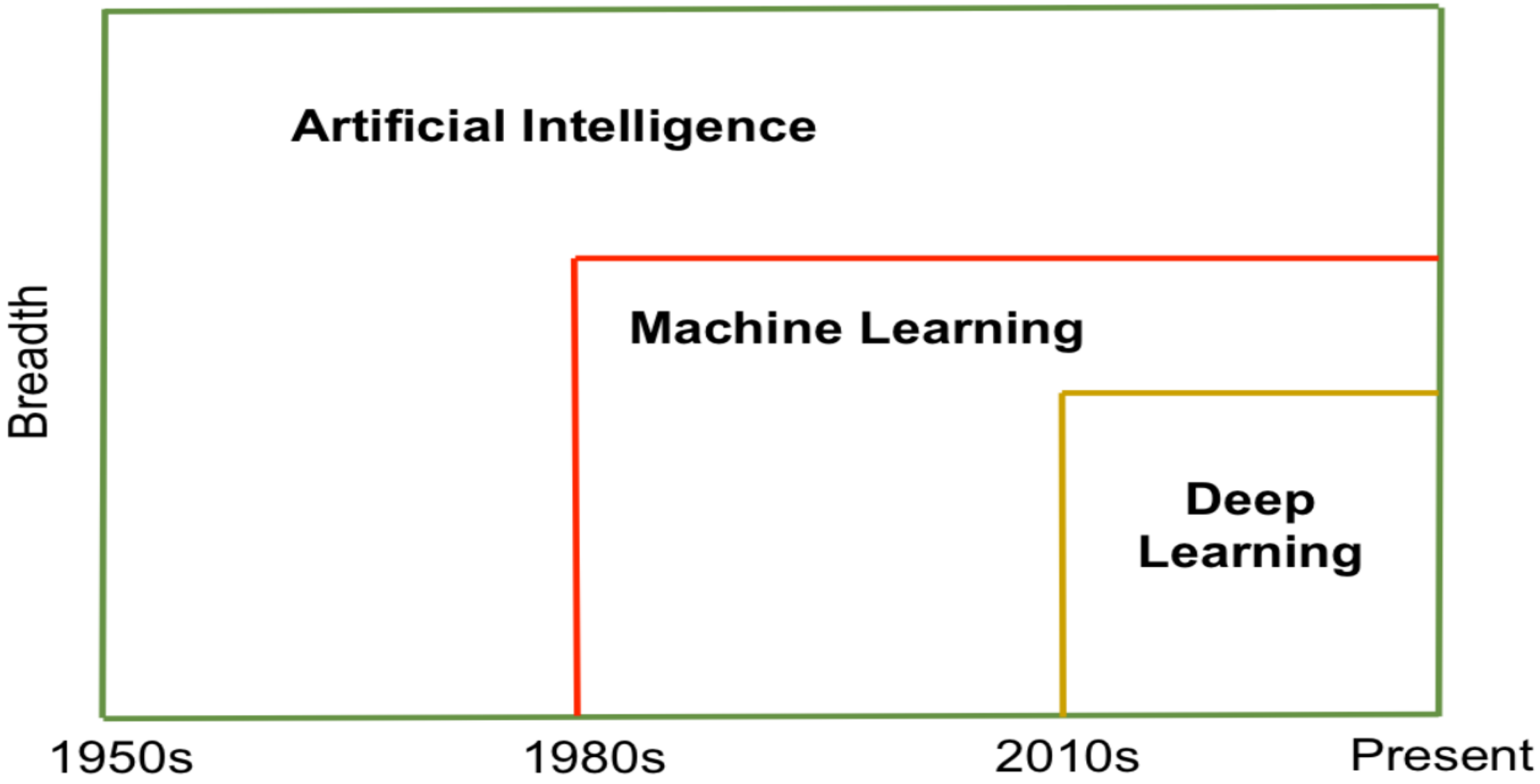
2016

Google's AlphaGo decisively beats the world champion of the complex board game Go.

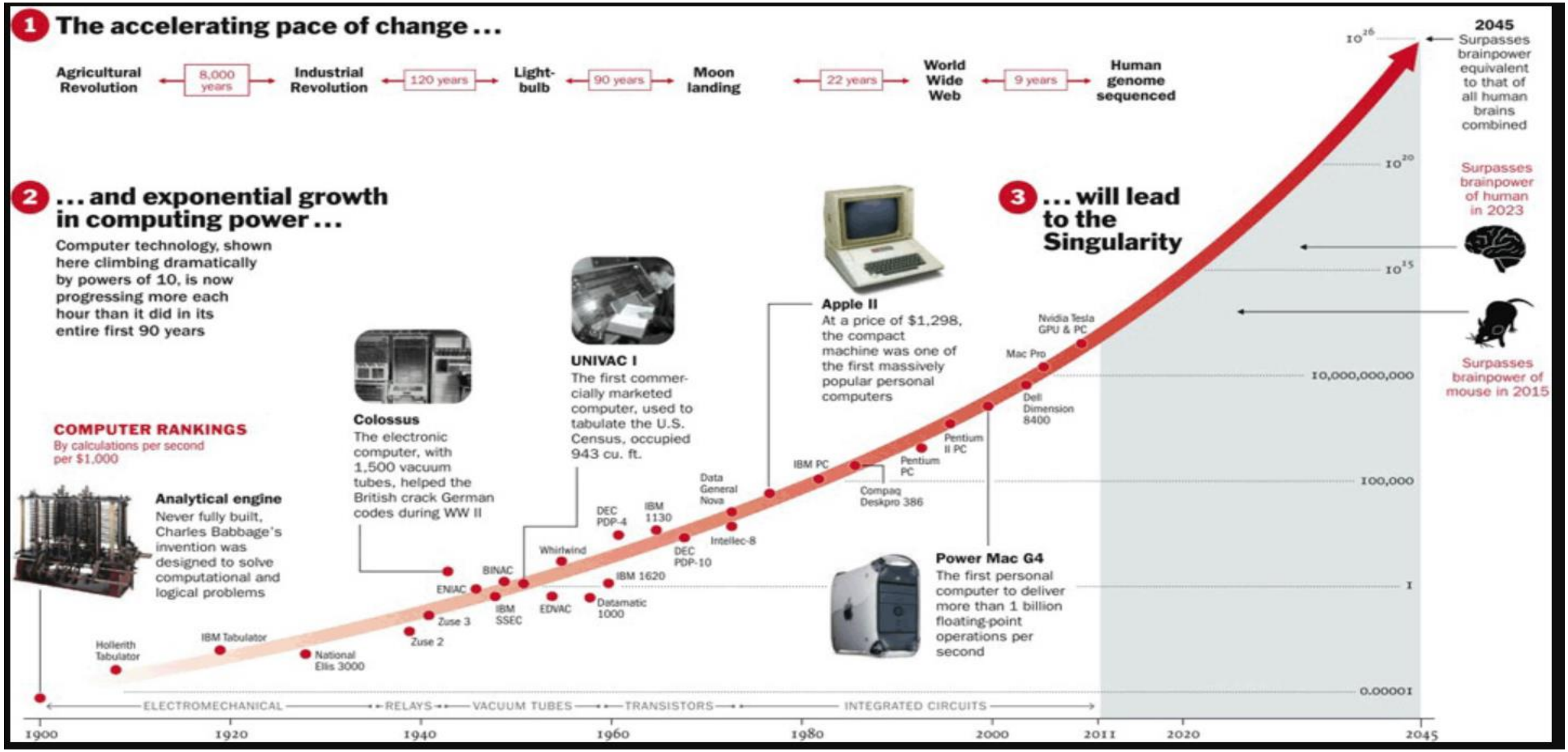
A brief History of AI



A brief History of AI



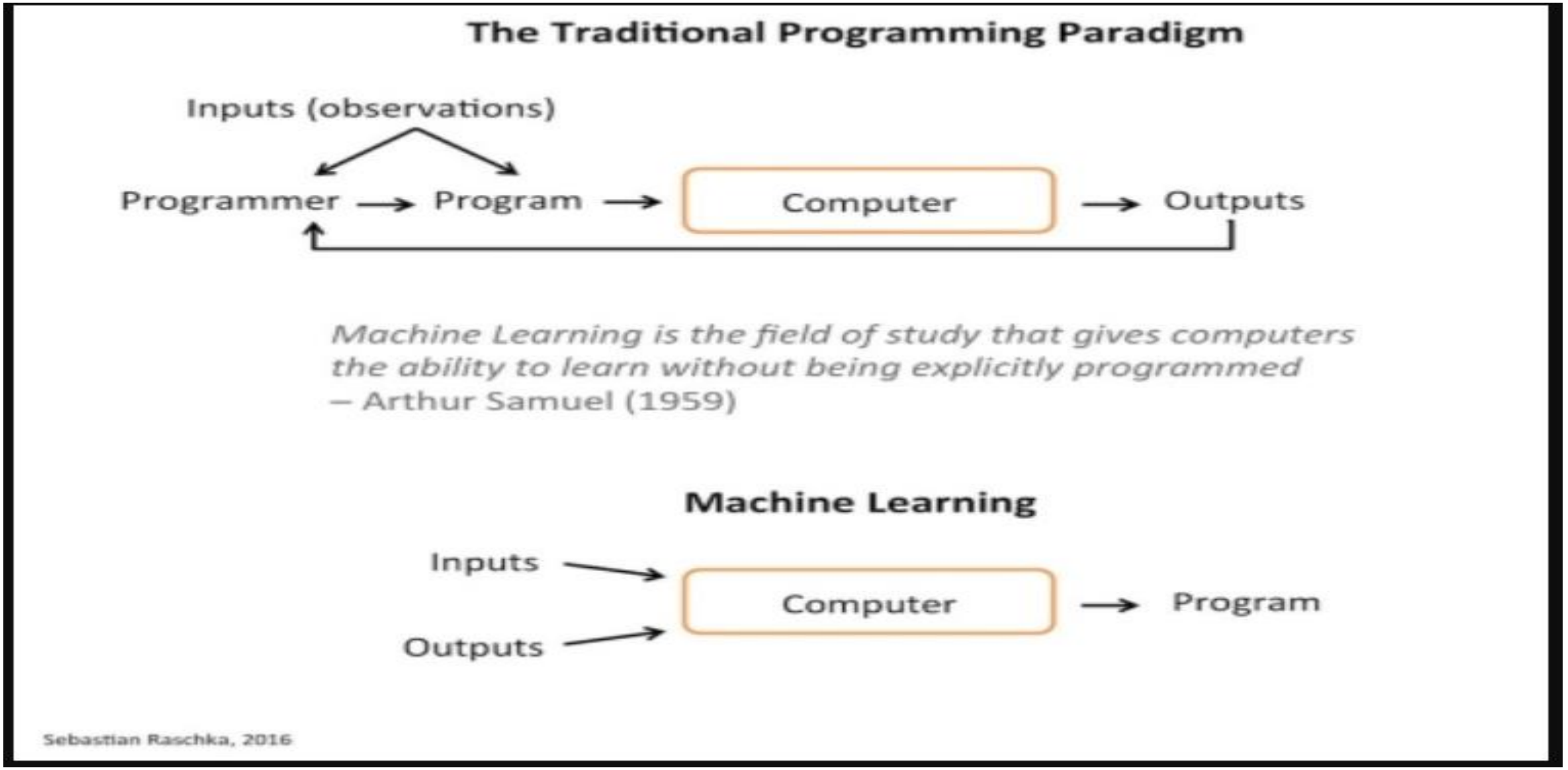
A brief history – computing power



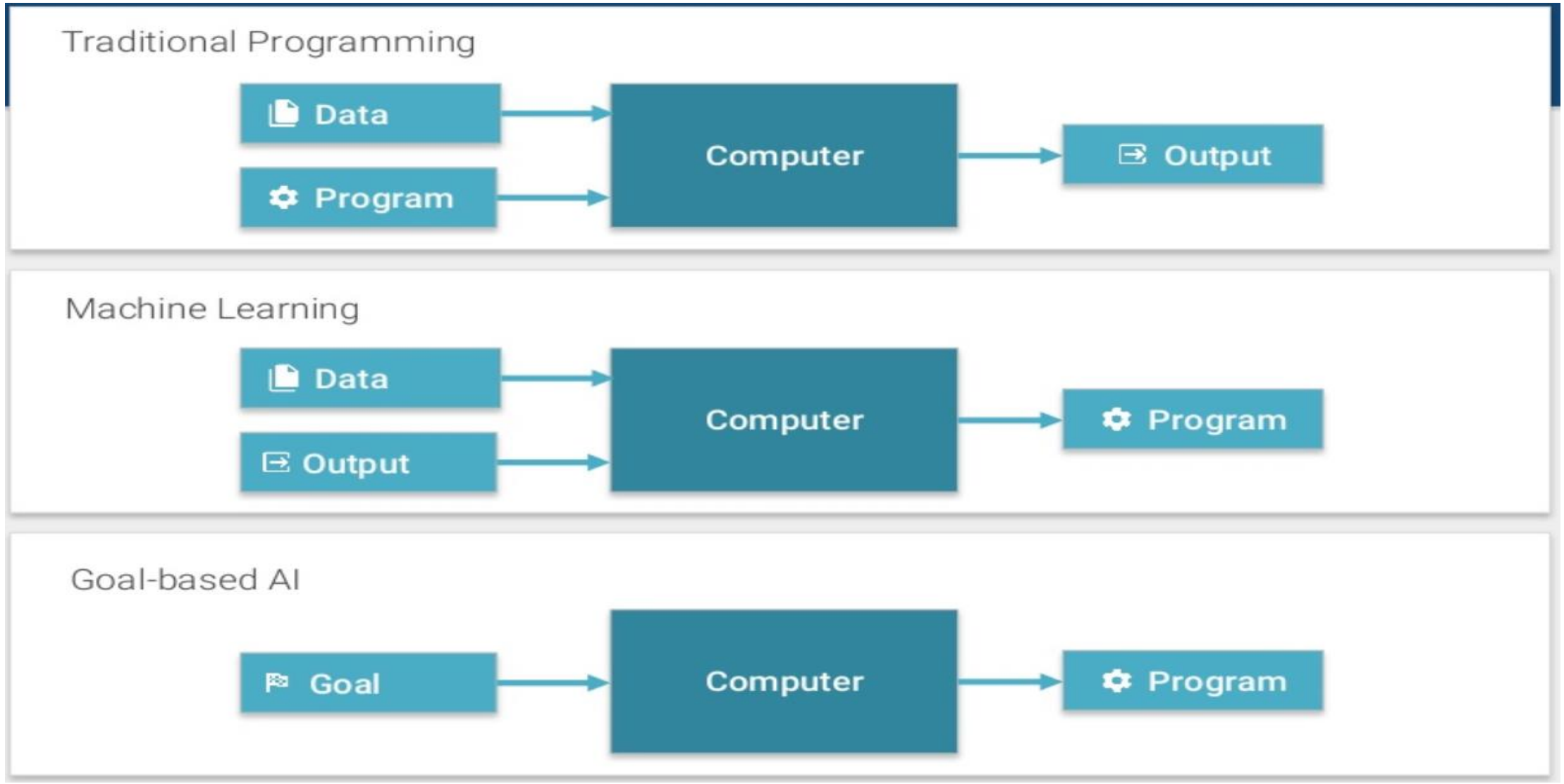
How it works?


Input:
Data
Sensors
Images
Artificial
Intelligence
Output:
Action
Movement
Text

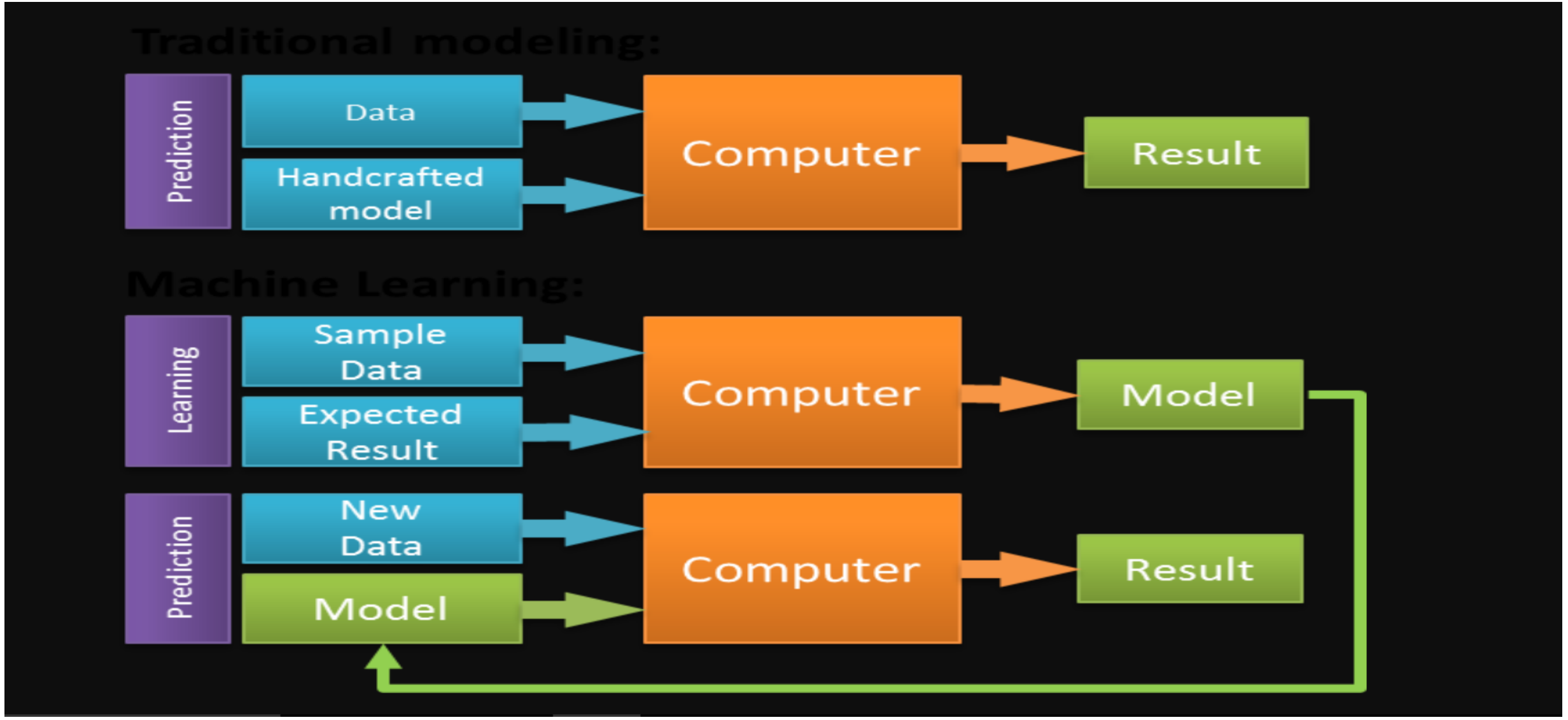
How it works?


The Traditional Programming Paradigm
Inputs (observations)
Programmer
Program
Computer
Outputs
Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed
– Arthur Samuel (1959)
Machine Learning
Inputs
Outputs
Computer
Program
Sebastian Raschka, 2016

How it works?

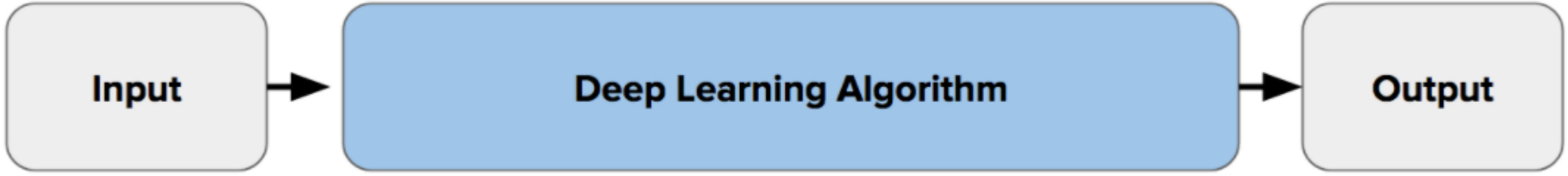

Traditional Programming
Data
Program
Computer
Output
Machine Learning
Data
Output
Computer
Program
Goal-based AI
Goal
Computer
Program

How it works?


Traditional modeling:
Prediction
Data
Handcrafted model
Computer
Result
Machine Learning:
Learning
Sample Data
Expected Result
Computer
Model
Prediction
New Data
Model
Computer
Result

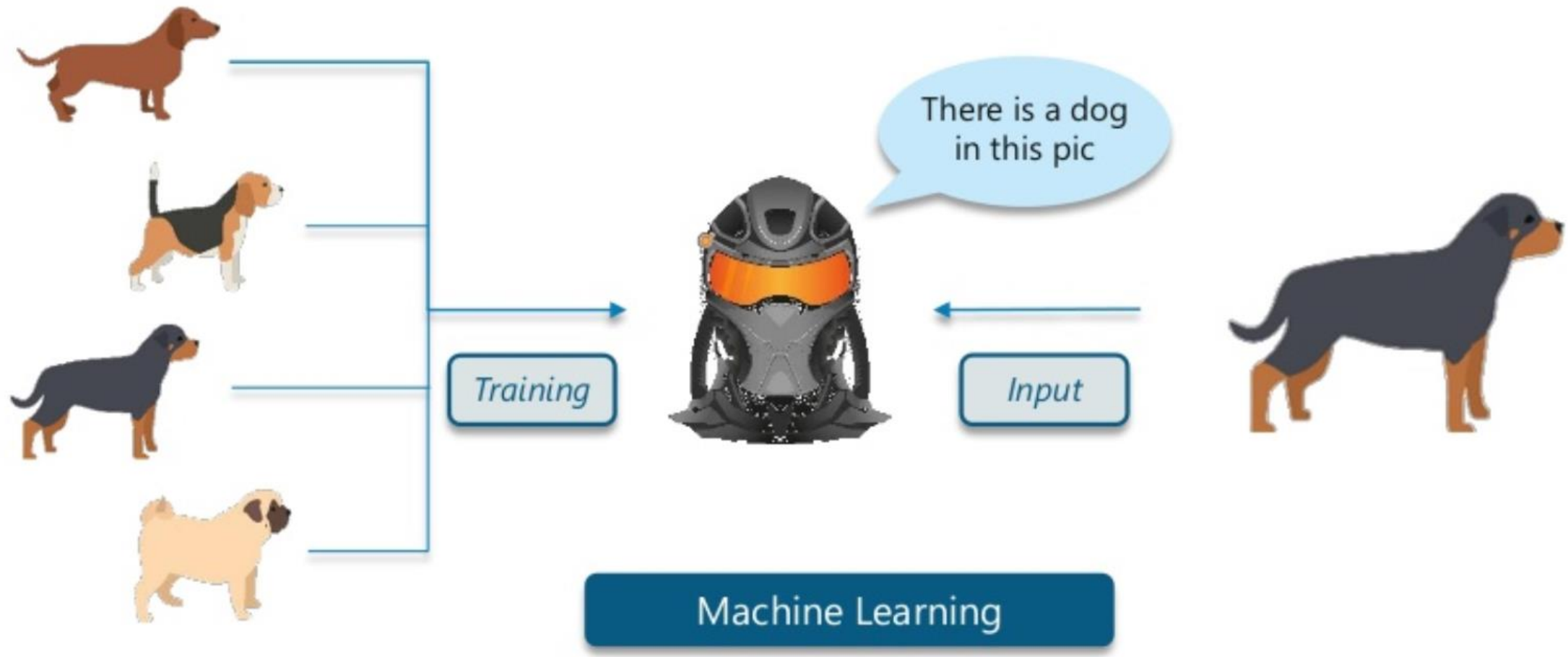
Machine learning vs Deep learning

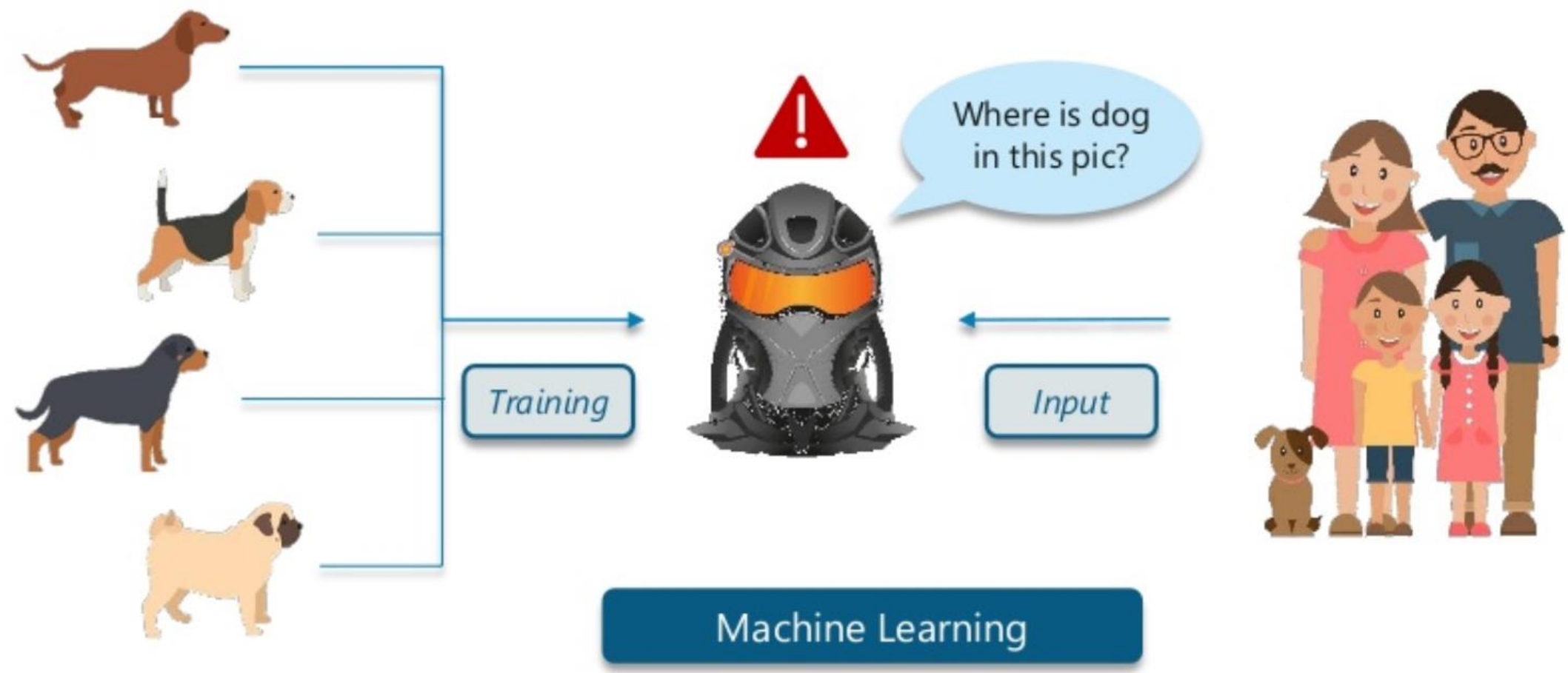
Traditional Machine Learning Flow

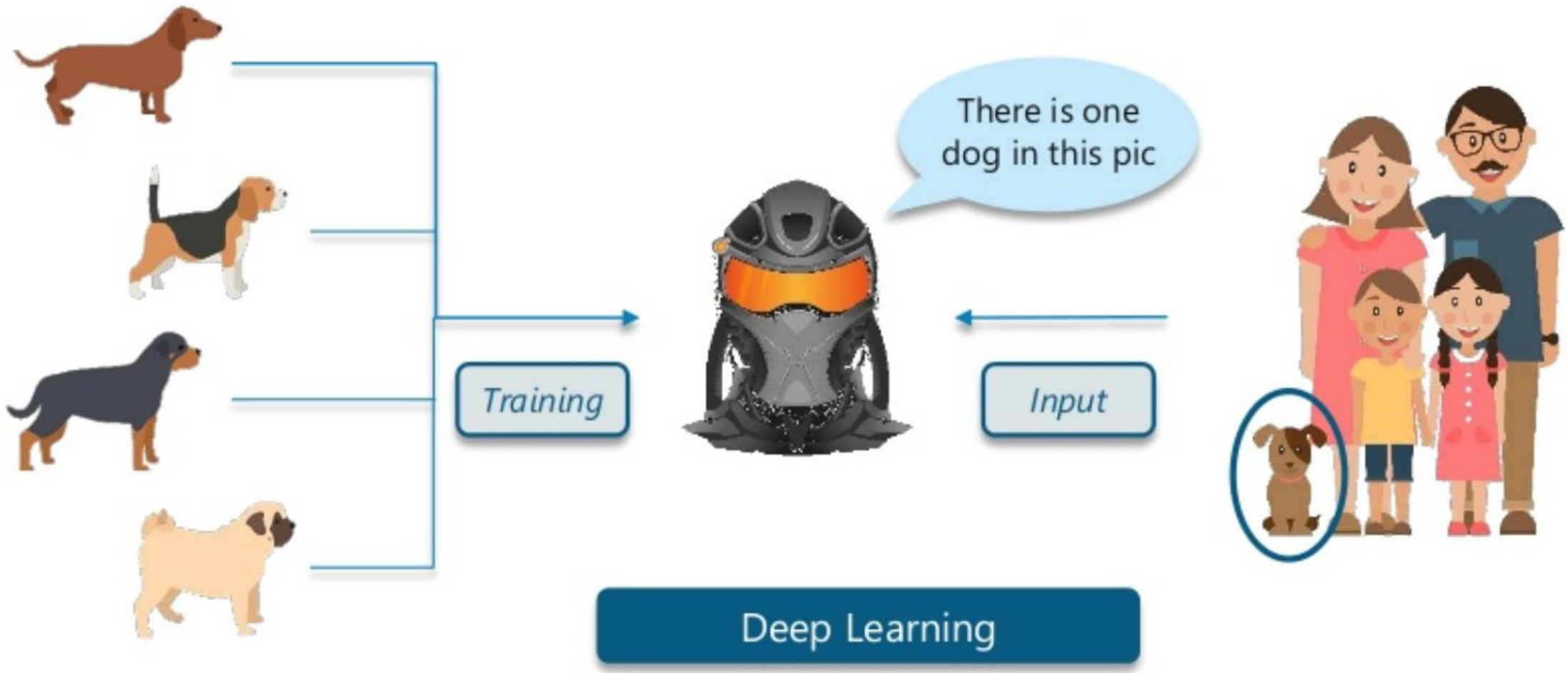


Deep Learning Flow

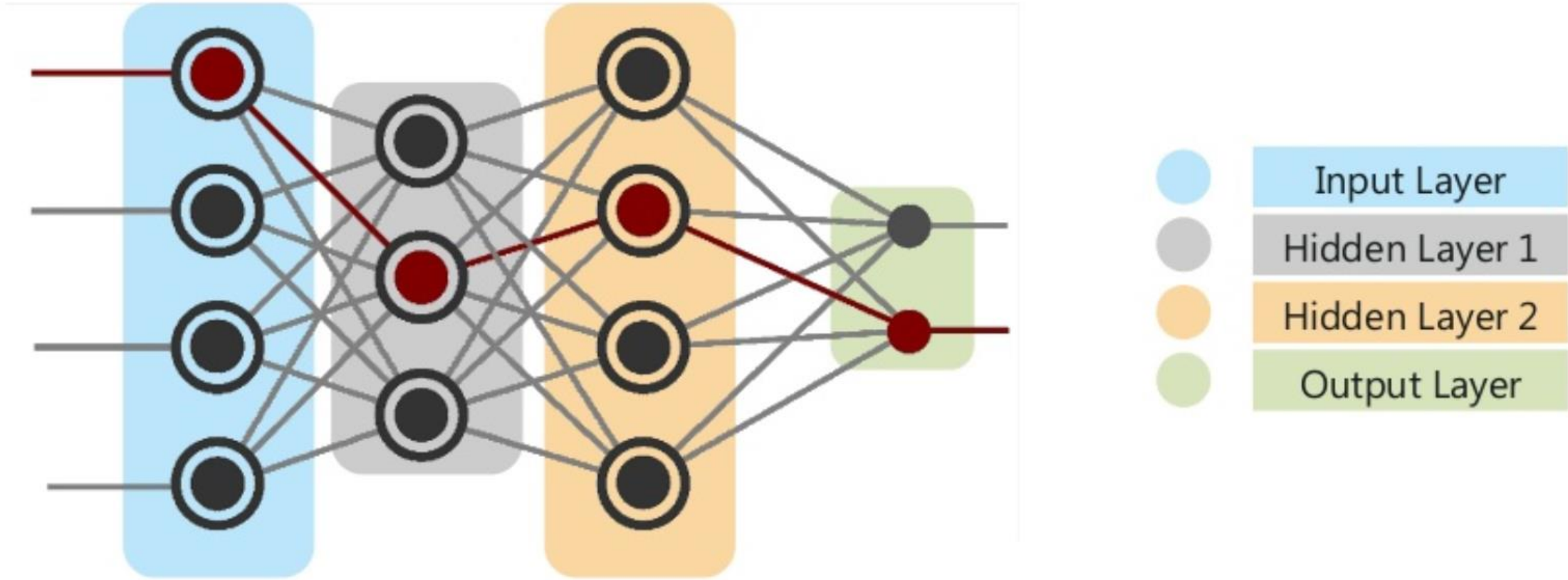
An artificial neuron contains a nonlinear activation function and has several incoming and outgoing weighted connections.


There is a dog
in this pic
Training
Input
Machine Learning


Where is dog in this pic?
Training
Input
Machine Learning

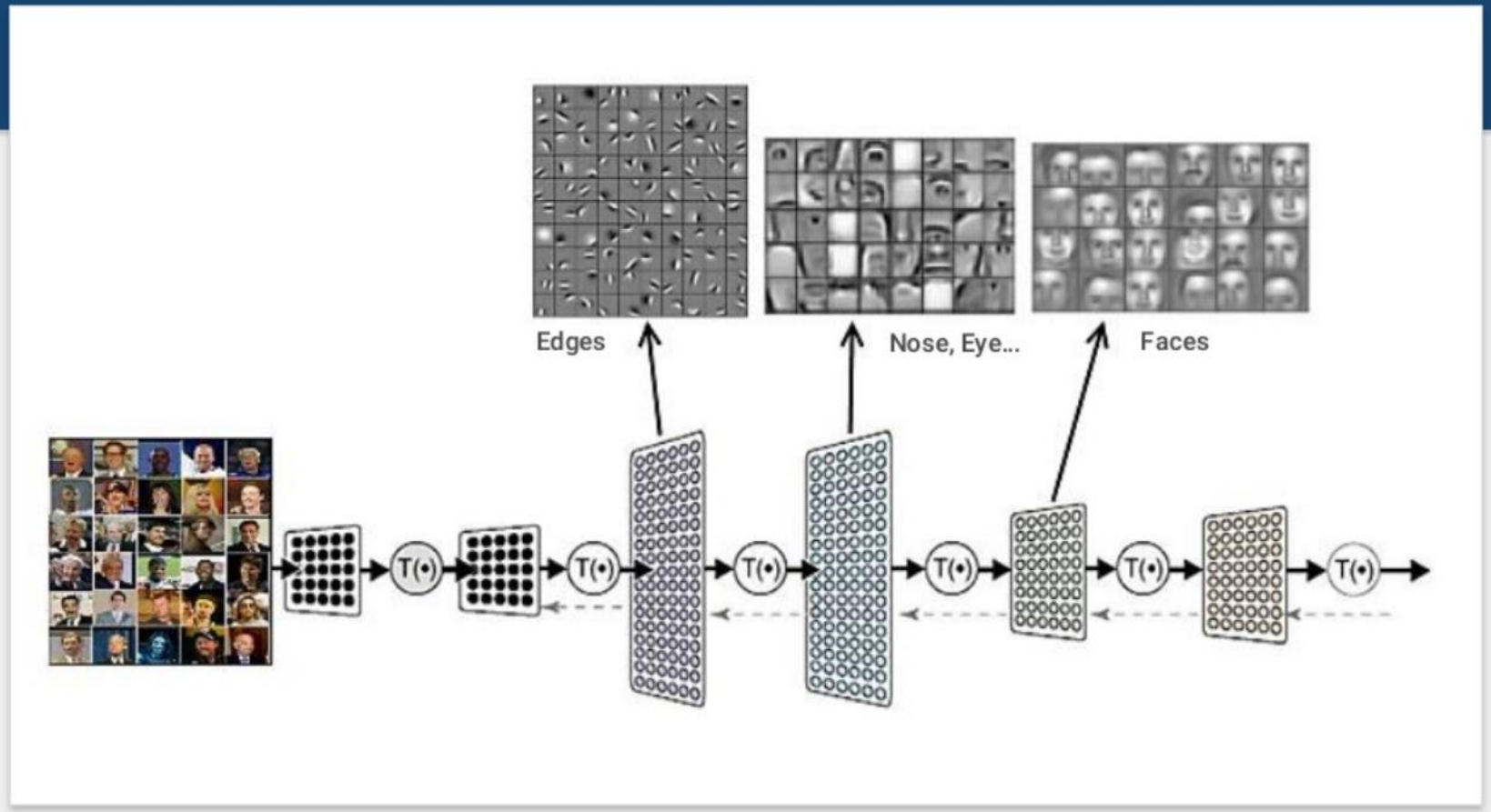

There is one dog in this pic
Training
Input
Deep Learning

A collection of statistical machine learning techniques used to learn feature hierarchies often based on artificial neural networks



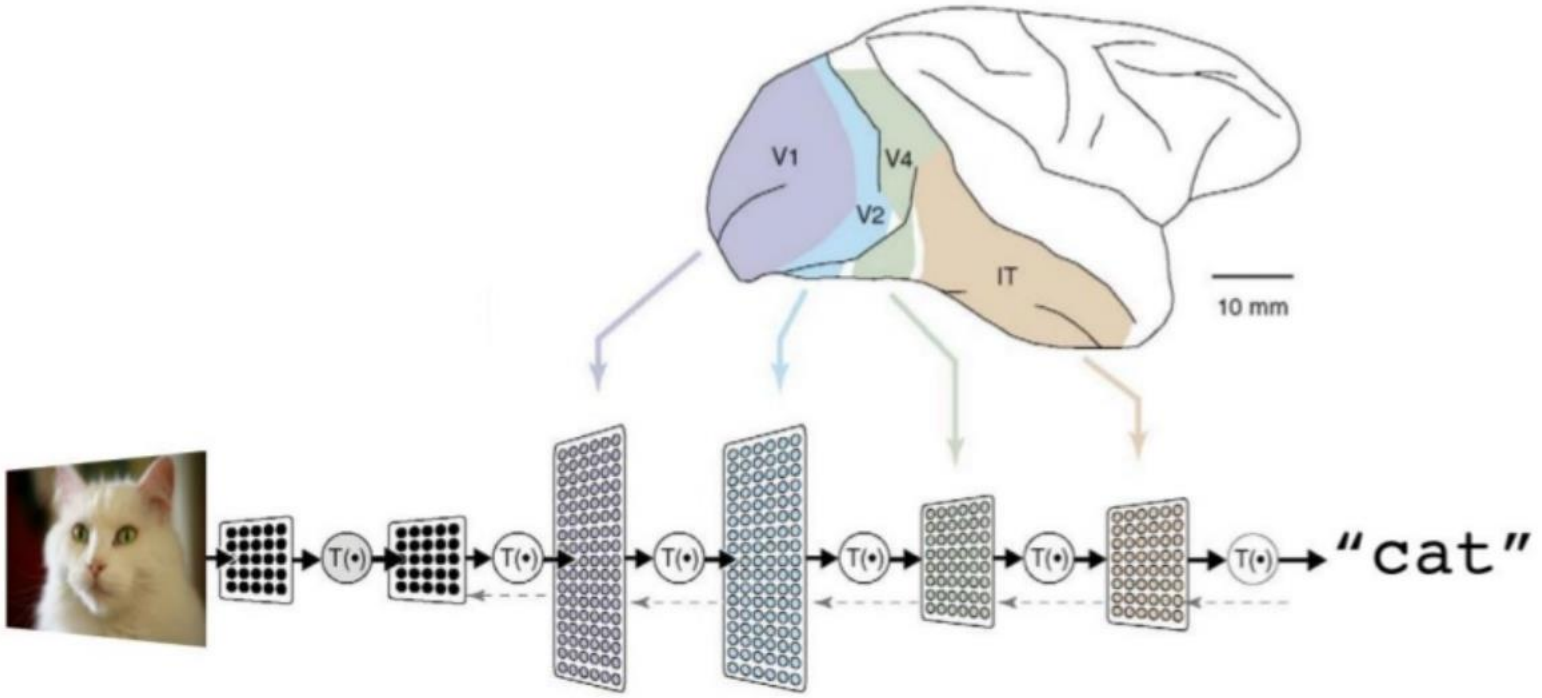
Deep Learning - Basics

What did it learn?



Deep Learning - Basics

Architecture



A deep neural network consists of a hierarchy of layers, whereby each layer transforms the input data into more abstract representations (e.g. edge -> nose -> face). The output layer combines those features to make predictions.

What Machine Learning Can Do

A simple way to think about supervised learning.

INPUT A	RESPONSE B	APPLICATION
Picture	Are there human faces? (0 or 1)	Photo tagging
Loan application	Will they repay the loan? (0 or 1)	Loan approvals
Ad plus user information	Will user click on ad? (0 or 1)	Targeted online ads
Audio clip	Transcript of audio clip	Speech recognition
English sentence	French sentence	Language translation
Sensors from hard disk, plane engine, etc.	Is it about to fail?	Preventive maintenance
Car camera and other sensors	Position of other cars	Self-driving cars

SOURCE ANDREW NG

© HBR.ORG

List of Algorithms in practice/industrial use

Applications of AI

AI For Good

AI to address socially relevant problems such as homelessness. At Stanford, researchers are using AI to analyze satellite images to identify which areas have the highest poverty levels

Aviation

Gate allocation for plane while landing.

Ticket price determination.

Education

There are a number of companies that create robots to teach subjects to children ranging from biology to computer science, though such tools have not become widespread yet

Applications of AI

Healthcare

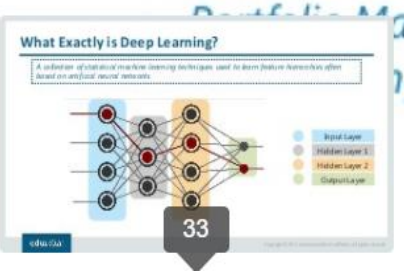
Companion robots for the care of the elderly
Mining medical records to provide more useful information
Design treatment plans
Assist in repetitive jobs including medication management
Provide consultations
Using avatars in place of patients for clinical training

Heavy Industry

Robots have become common in many industries and are often given jobs that are considered dangerous to humans. Robots have proven effective in jobs that are very repetitive which may lead to mistakes or accidents due to a lapse in concentration and other jobs which humans may find degrading.

Finance

- *Algorithmic Trading*
- *Market Analysis and Data Mining*
- *Personal Finance*



Use case – Self driving Car

GPS is used to get a rough location of the car

The car can easily recognize road marking, signs and traffic lights.

The car even detect road works, cyclists and safely navigate around them

Smartphone is used for pick up at the user's location with the destination set.

Use case – Self flying helicopter

Self-flying helicopters. The K-MAX may be used for firefighting and It can carry up to 6,000 pounds of cargo, this helicopter can drop nearly 2,800 gallons water on the fire.

Use case - vigilante

There is an interesting example of artificial intelligence technology created for stopping hijackers. This machine combines functions of a drone, helicopter, and trained snipers. This helicopter is called Vigilante.

Use case – artistic images + creative music

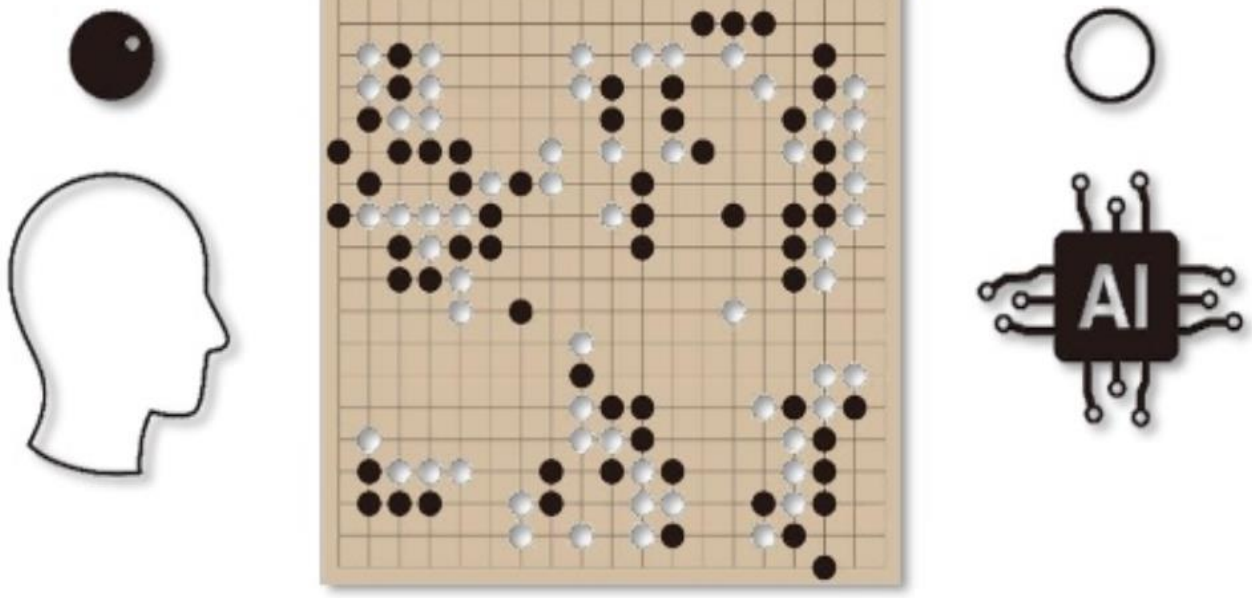
AARON is a computer program written by artist Harold Cohen that creates original artistic images

Another awesome example of using artificial intelligence creatively is composing music

Use case – IBM watson

IBM Watson, a natural language question answering computer, competes on Jeopardy! and defeats two former champions.

Use case - AlphaGo



Google DeepMind's **AlphaGo** defeats Go champion **Lee Sedol** which AI researcher thought was not possible to achieve in next **20 years**

Tactics used by **AlphaGo** are taught to GO players during their training

Use case - Google Translator

Google's AI translation tool seems to have invented its own secret internal language

Google Neural Machine Translation

Training

ENGLISH

JAPANESE

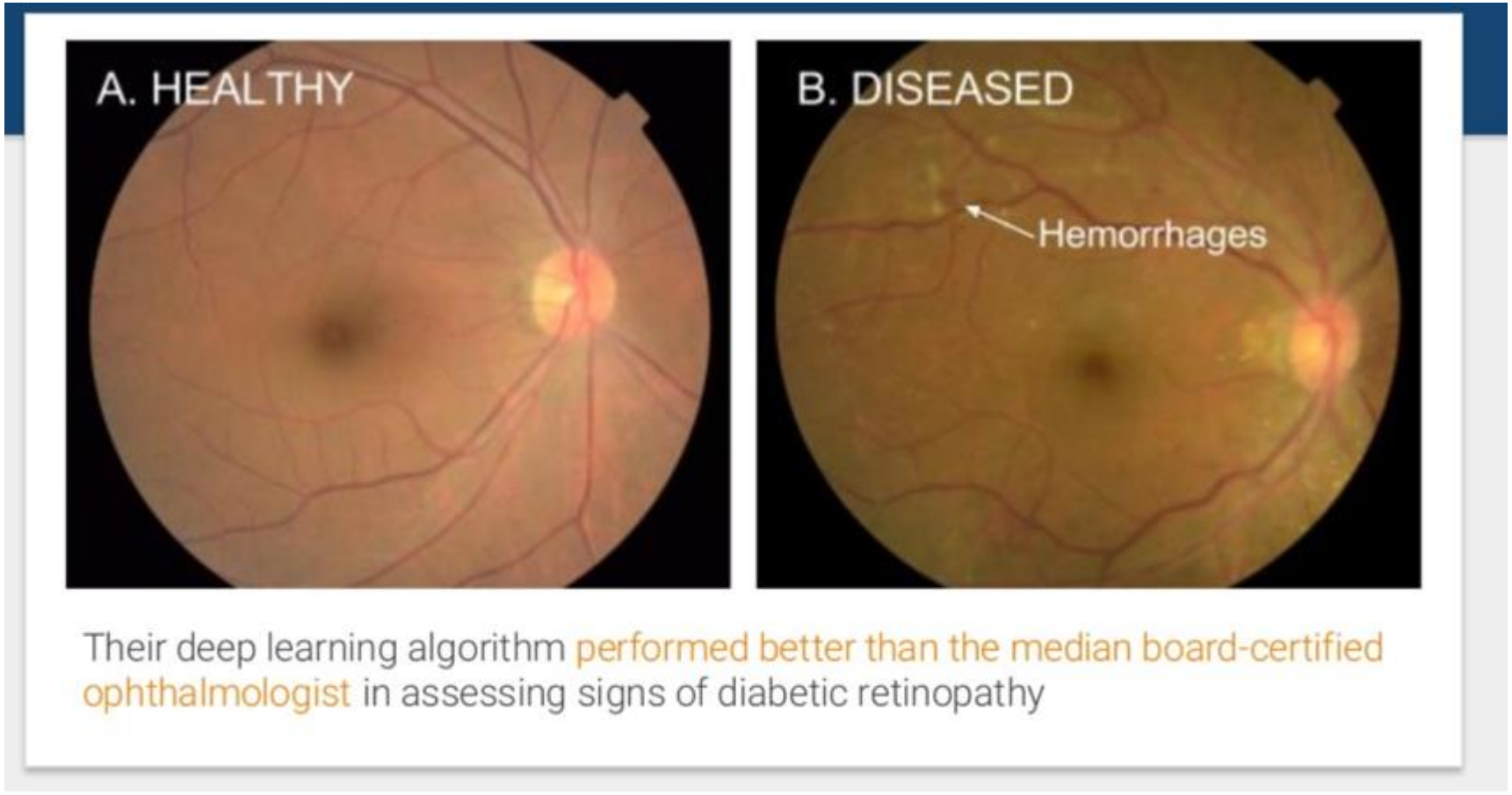
KOREAN

ENGLISH

JAPANESE

KOREAN

Use case – Detecting diseases



Use case – Siri/ Cartana

Google Now, Siri and Cortana are all intelligent digital personal assistants on various platforms (iOS, Android, and Windows Mobile). In short, they help find useful information when we ask for it using our voice.

Use case – Recommender systems

Enterprise Use cases

Automate Knowledge Work

- Transformational **HR** services

- Lights out **finance** operations

- Self-driving customer **service**
- Conversational **sales** bots
- Customer **retention** insights

Do the Impossible

- Image-based Ariba **commerce**
- Contextual Concur **travel** concierge

- Video-aware **marketing**
- Visual **store execution**

- Drone and satellite-based **asset management**

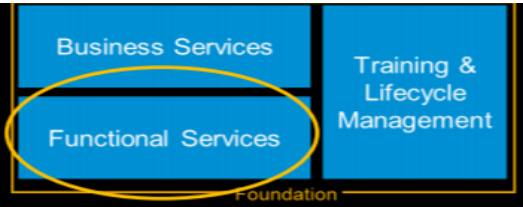
- Vision-enabled **manufacturing**
- Contextual **logistics**

Enterprise data Sources

	Text	Image/Video	Speech & Audio	Tabular
General availability 1710	▪ Topic detection ▪ Text classification ▪ Text feature extraction	▪ Image feature extraction ▪ Image classification ▪ Customizable image classification		▪ Time series change point detection ▪ Similarity scoring
Alpha 1710	▪ Machine translation ▪ Product text classification ▪ Language detection ▪ Document Clustering ▪ Customizable text classification	▪ Product image classification ▪ Image segmentation ▪ Face detection		▪ Time series forecasting ▪ Multi-dimensional time series forecasting ▪ Multi-dimensional data clustering ▪ Generic classification (text and tabular)
Roadmap In process	▪ Sentiment analysis ▪ Named entity recognition ▪ Hate speech detection ▪ File-to-text conversion	▪ Image text extraction ▪ Image NER/extraction ▪ Apparel detection ▪ Video object segmentation ▪ Video classification ▪ Video human action recognition	▪ Voice recognition (speaker identification)	▪ Time-to-failure forecasting ▪ Association rule learning ▪ Customizable recommender
Partner In process		▪ Optical character recognition	▪ Speech-to-text	

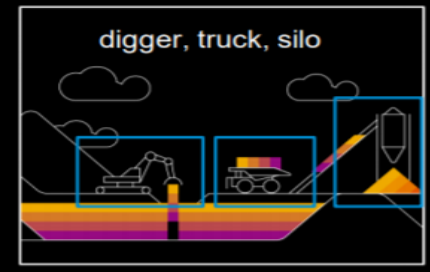
Enterprise use cases

Functional service: Image classification



Image

SAP Leonardo
Machine Learning



Classified objects in an image

Detects objects in an image and classifies them into a fixed set of categories.

Via Rest API this functionality can be integrated in any software stack.

- Recognize patterns and unlock new insights from images
- Exploit knowledge from external sources e.g. images on social media
- Improve business processes such as visual quality inspection

Enterprise use cases

Functional service: Topic detection

Document

Topics/Keywords of document

Detects topics in a document and provides scores for each topic
Extracts keywords describing each topic

Via Rest API this functionality can be integrated in any software stack.

Unlock insights from natural language text

Understand what customers are talking about you on the web

Improve search and organization of text contents

Enterprise use case

SAP Leonardo Machine Learning Foundation: Use Case Sample

Trend Color Prediction powered by functional services

Social Media Data
Influencers Fashion Data

Apparel type detection
Color detection

Color Trend Prediction
In-Store Manufacturing
Micro Season Forecast

SAP Leonardo Machine Learning analyzes images to better understand fashion trends for increasing the quality of apparel production forecast

Enterprise use cases

SAP Leonardo
Machine Learning

Predict
Recommend
Classify
Identify
Tag
… and more

Unlock insights from Big Data and
infuse your apps with intelligent services

Scalable and managed services

Process Big Data without worries about the infrastructure

Flexibility and agility

Make immediate use of intelligent services, or build your own ML models

Integration

Smooth integration and access to data sources from SAP as well as non-SAP software

End-to-end coverage

Create, run and maintain intelligent applications – the entire lifecycle is covered

Enterprise use cases

SAP Leonardo Machine Learning Foundation: Use Case Sample

Broken product similarity search

01011
11010
10
01101

Service Ticket, e-mail incl. image of broken product

SAP Leonardo Machine Learning

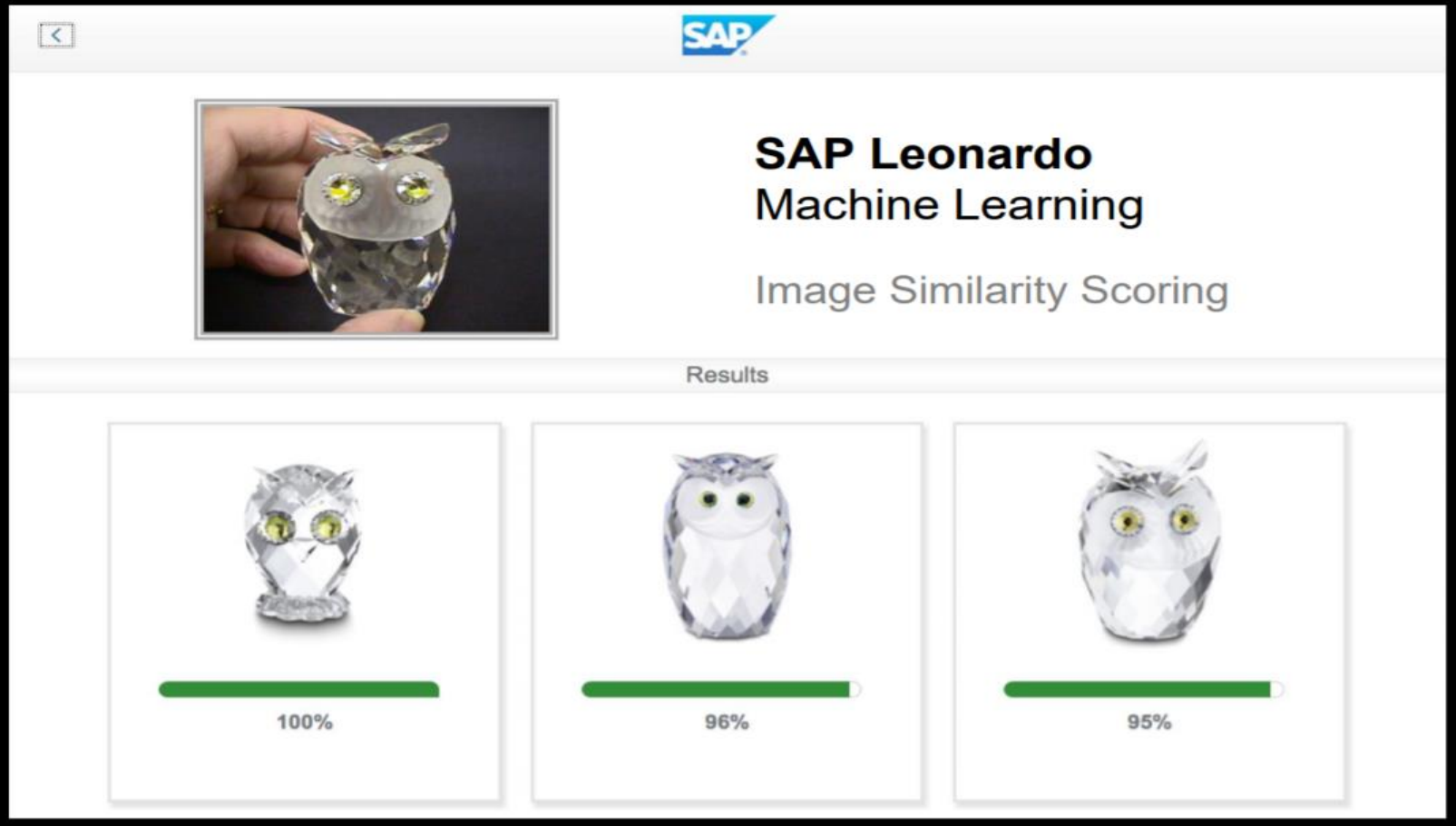
Image Feature Extraction
Similarity Scoring

Product Identification and automatic classification

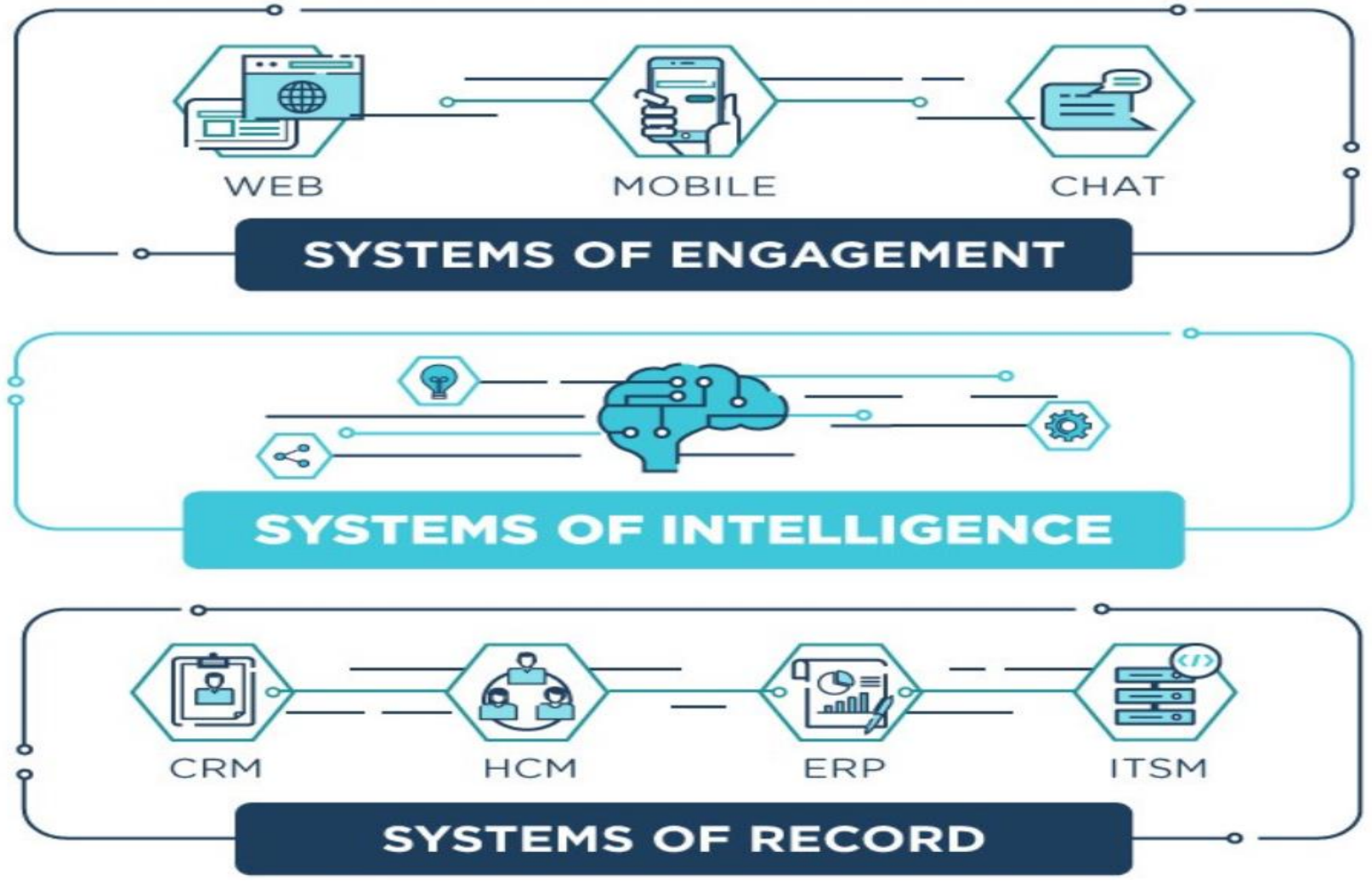
SAP's Machine Learning automatically classifies product images and enables faster customer interaction with precise information on potential product repair cost or item substitution.

Enterprise use cases

Functional Service Demo: Image Similarity Scoring



eEnterprise use cases



Kaggle Examples

290 Competitions

Passenger Screening Algorithm Challenge — $1,500,000, 518 teams
Improve the accuracy of the Department of Homeland Security's threat recognition algorithms
Featured · 7 months ago · terrorism, image data, object detection

Zillow Prize: Zillow's Home Value Prediction (Zestimate) — $1,200,000, 3,779 teams
Can you improve the algorithm that changed the world of real estate?
Featured · 6 months ago · housing, real estate

Data Science Bowl 2017 — $1,000,000, 1,972 teams
Can you improve lung cancer detection?
Featured · a year ago · healthcare, image data, binary classification

Heritage Health Prize — $500,000, 1,353 teams
Identify patients who will be admitted to a hospital within the next year using historical claims data. (Ent...
Featured · 5 years ago

Second Annual Data Science Bowl — $200,000, 773 teams
Transforming How We Diagnose Heart Disease
Featured · 2 years ago · healthcare, image data

National Data Science Bowl — $175,000, 1,049 teams
Predict ocean health, one plankton at a time
Featured · 3 years ago · oceanography, image data, multiclass classification

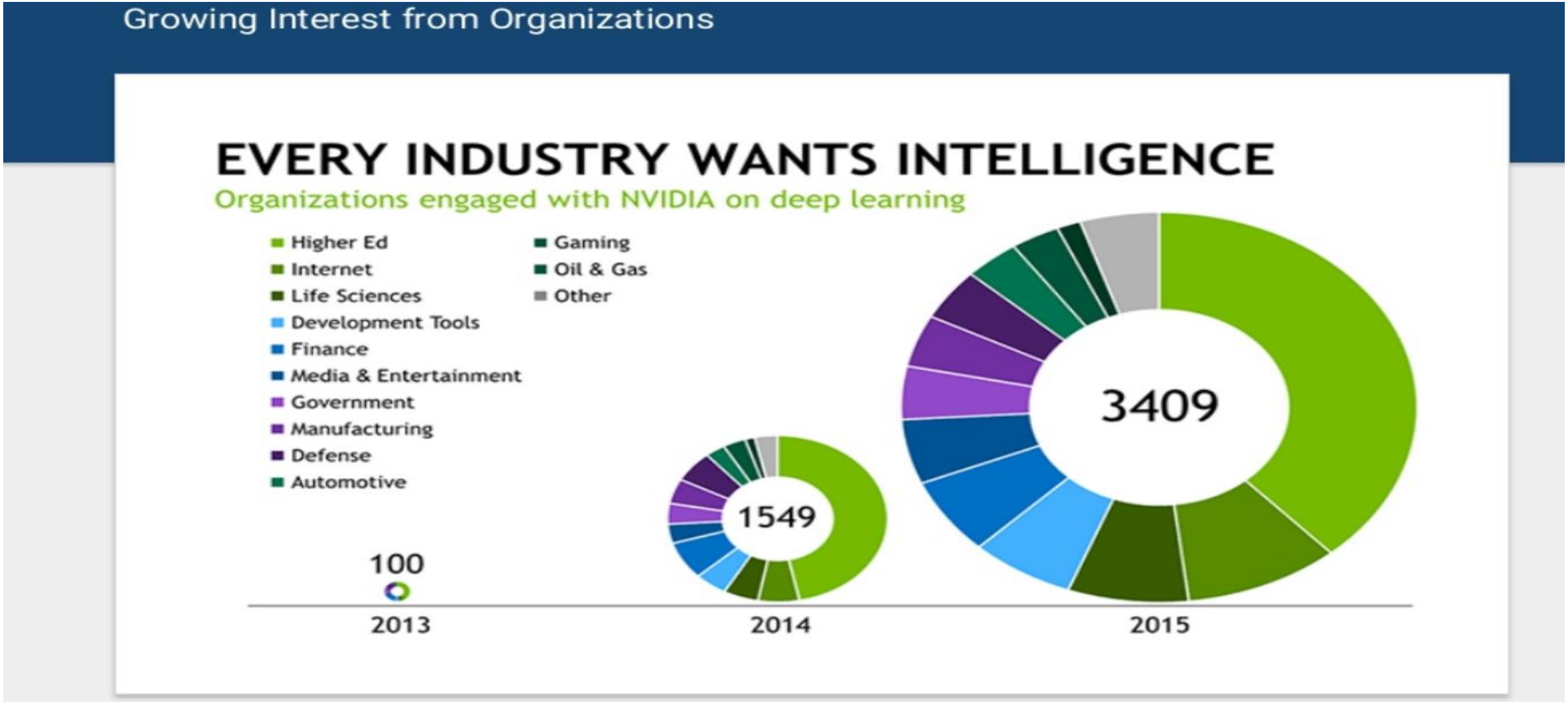
The Nature Conservancy Fisheries Monitoring — $150,000, 2,293 teams
Can you detect and classify species of fish?
Featured · a year ago · fishing, image data, multiclass classification, object detection

Industrial

Use Cases	Industries
Static image recognition, classification, and tagging	Advertising, Consumer
Algorithmic trading strategy performance improvement	Investment
Efficient, scalable processing of patient data	Finance, Healthcare
Predictive maintenance	Defense, Manufacturing, Telecommunications, Aerospace, Automotive
Object identification, detection, classification, tracking from geospatial images	Defense, Aerospace
Text query of images	Advertising,
Automated geophysical feature detection	Oil, Gas and Mining
Content distribution on social media	Media & Entertainment
Object detection and classification - avoidance, navigation	Defense, Manufacturing, Aerospace, Agriculture, Automotive
Prevention against cybersecurity threats	Business
Contract analysis	Legal
Text-based automated bots	Consumer
Sensor data analysis (IoT)	Agriculture, Automotive
Sensor data fusion	Defense, Manufacturing, Aerospace, Agriculture, Transportation
Human emotion analysis	Business, Media & Business, Media & Entertainment

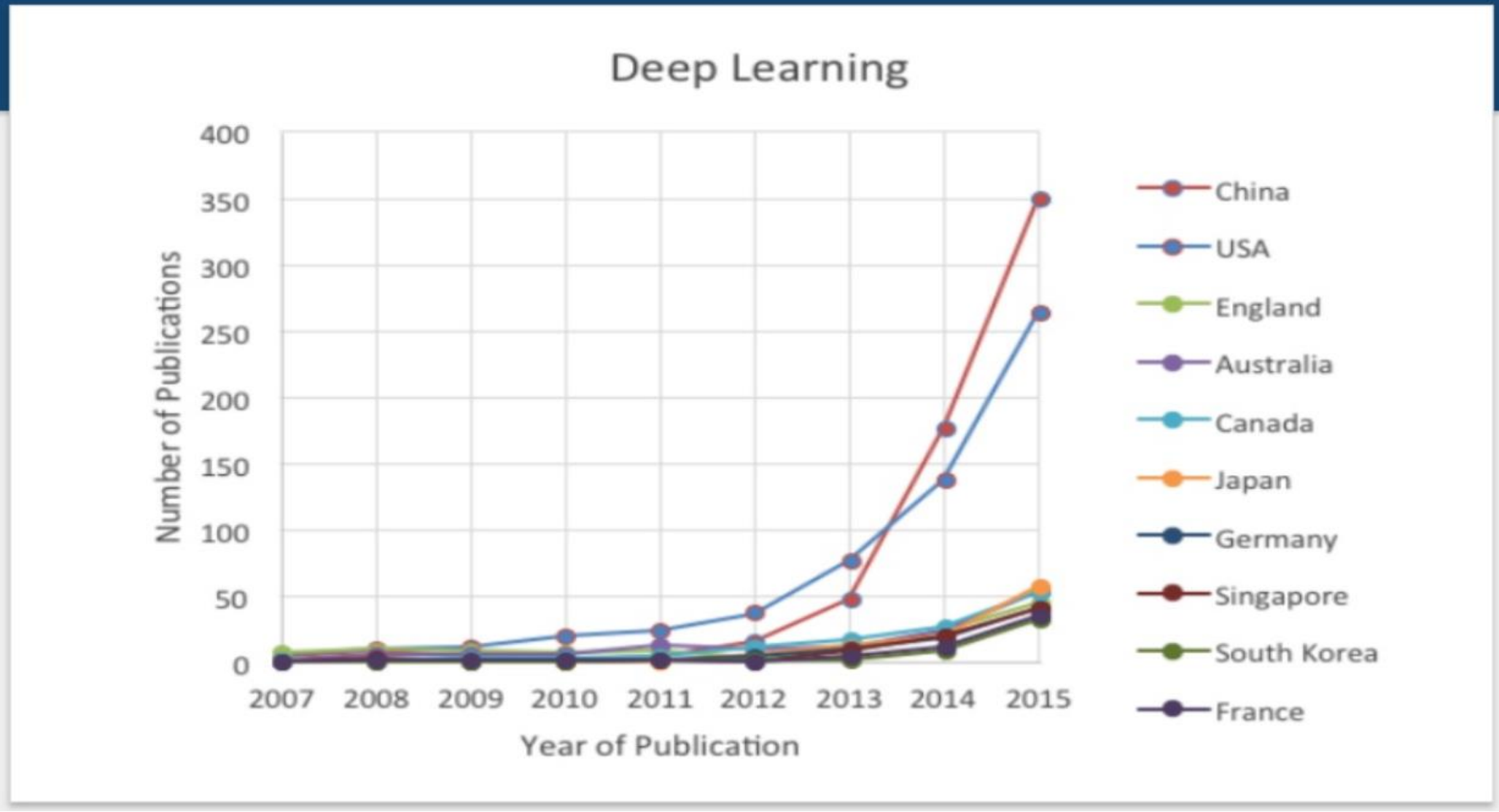
(Source: Tractica)

Potential of AI – why it is so important?



Potential of AI – why it is so important? - research

Academic Publications about Deep Learning

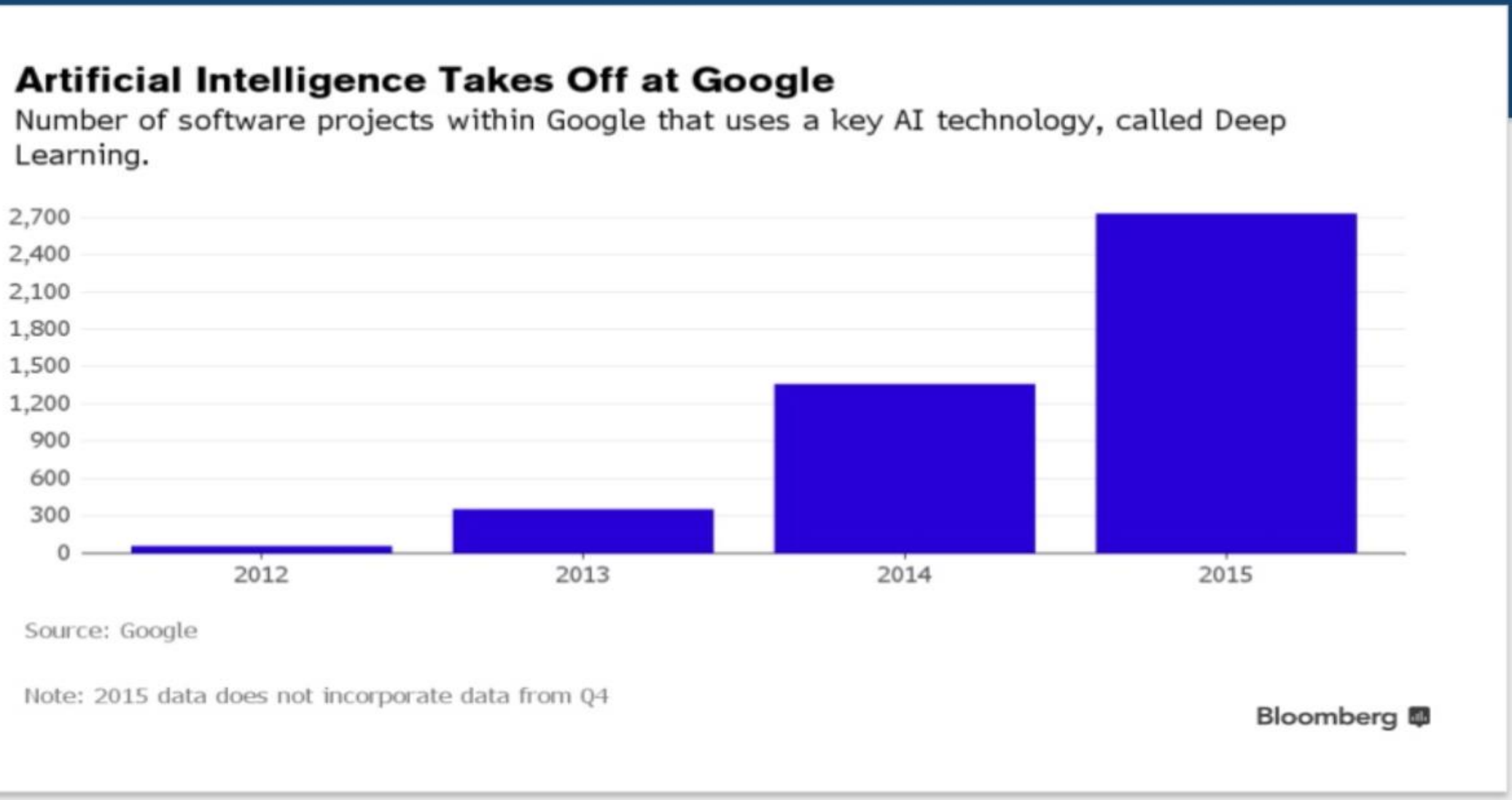


Potential of AI – why it is so important? – consumer tech

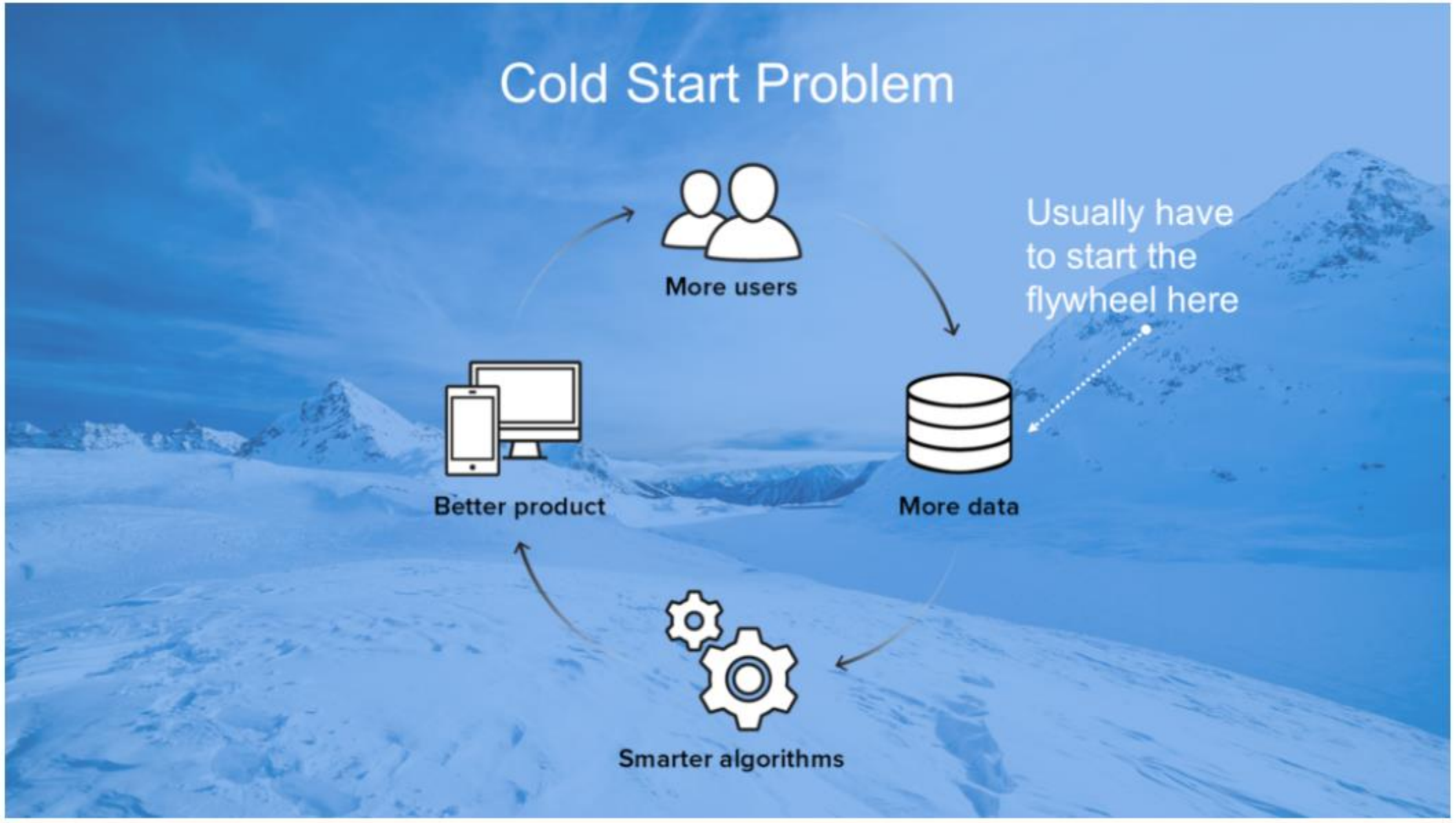
Potential of AI – why it is so important?

Potential of AI – why it is so important?

Potential of AI – why it is so important?



Potential of AI – why it is so important?



Potential of AI – why it is so important?

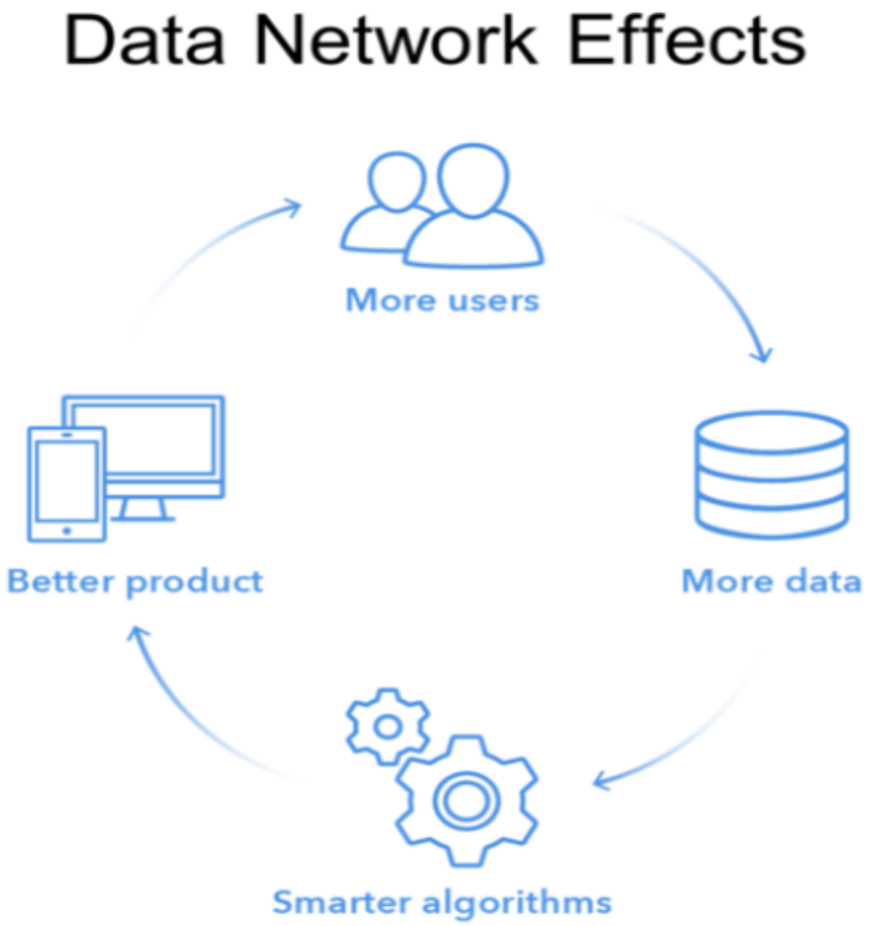
Exemplified by Industry Giants

Google

amazon

NETFLIX

UBER



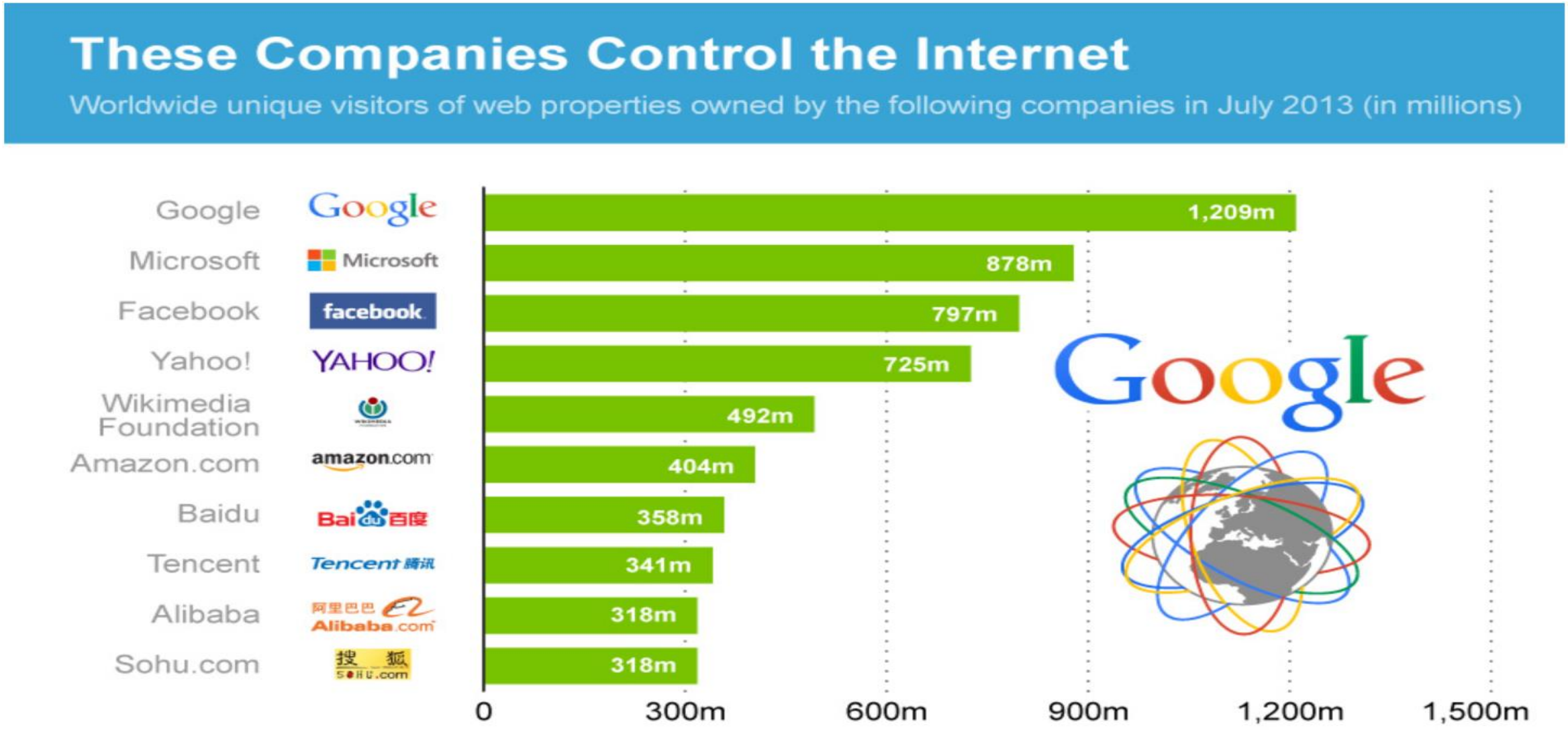
But also available to startups

waze

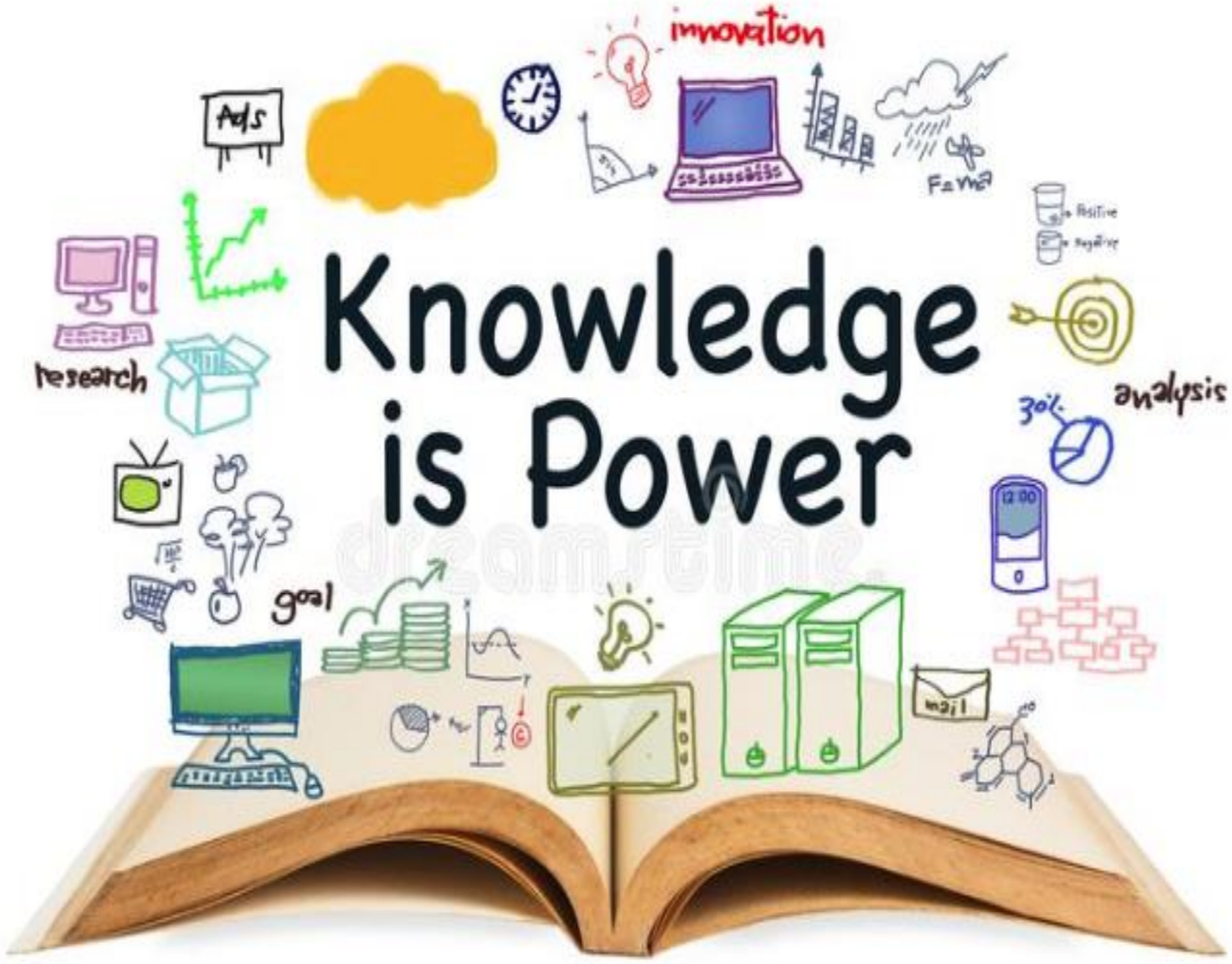
Phosphorus

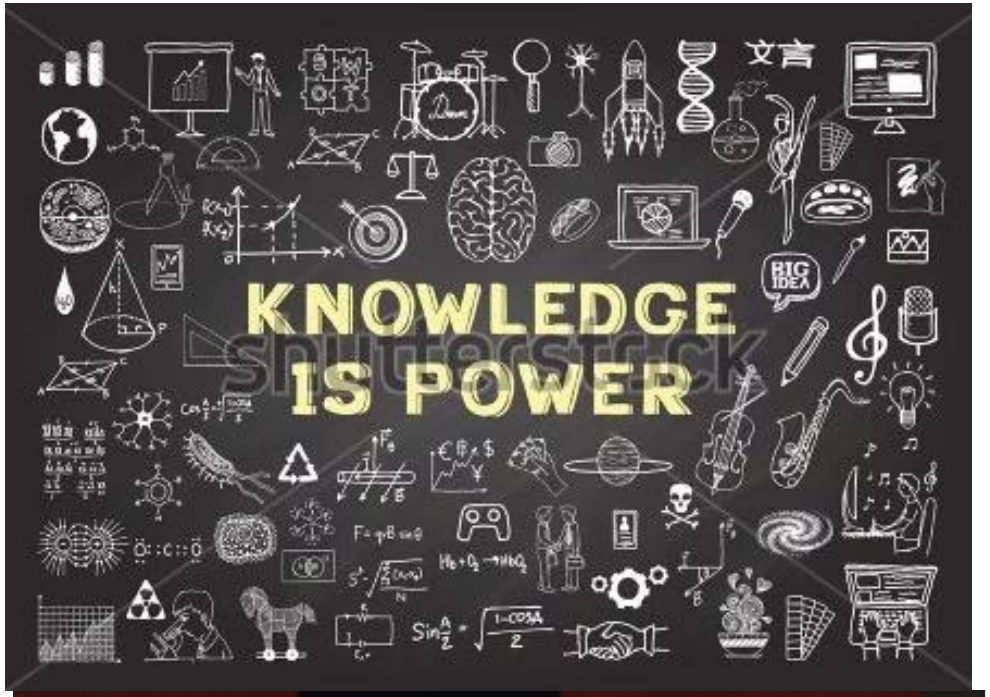
(h [s]) HyperScience

Potential of AI – why it is so important?



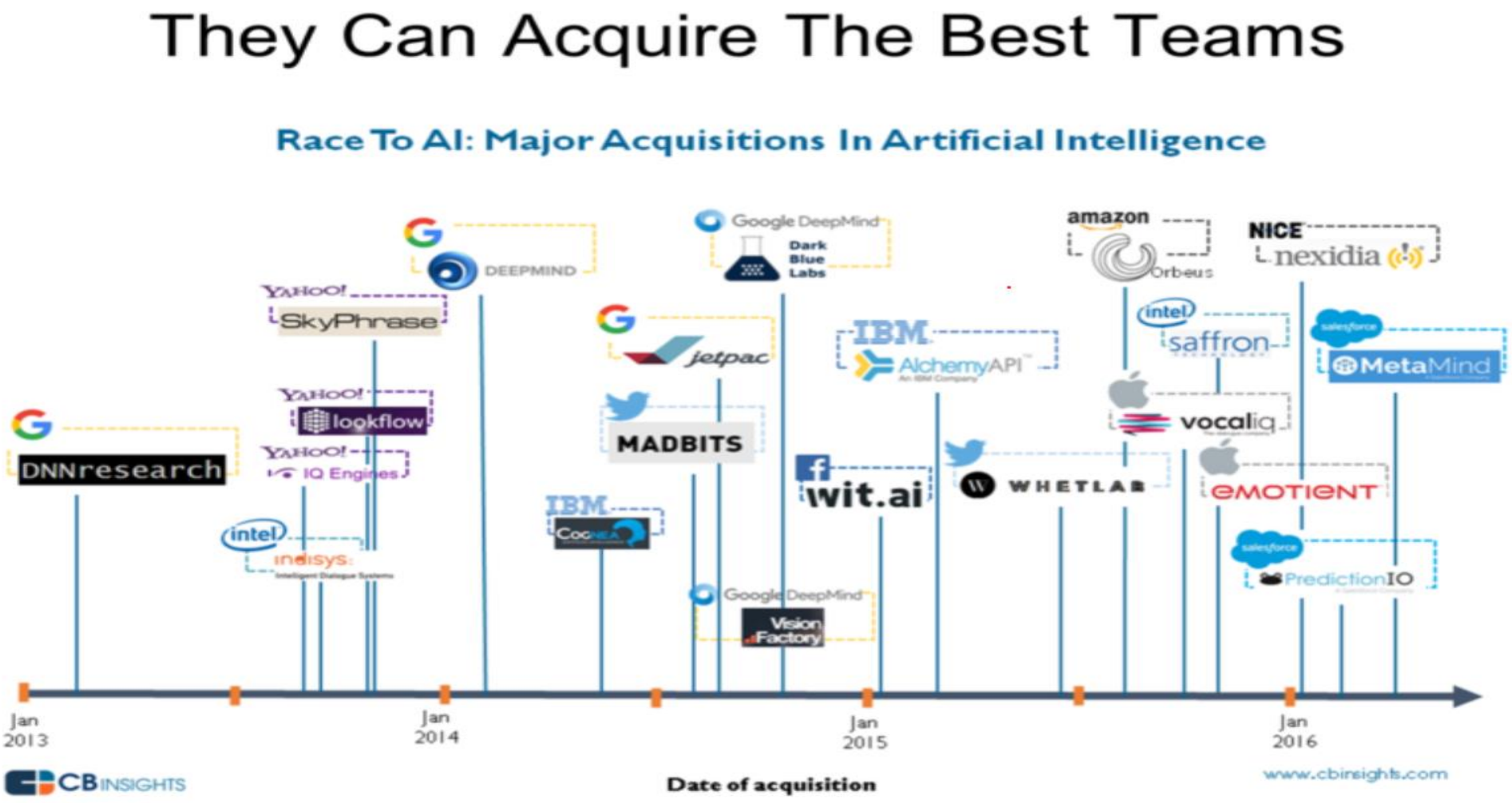
Potential of AI – why it is so important?



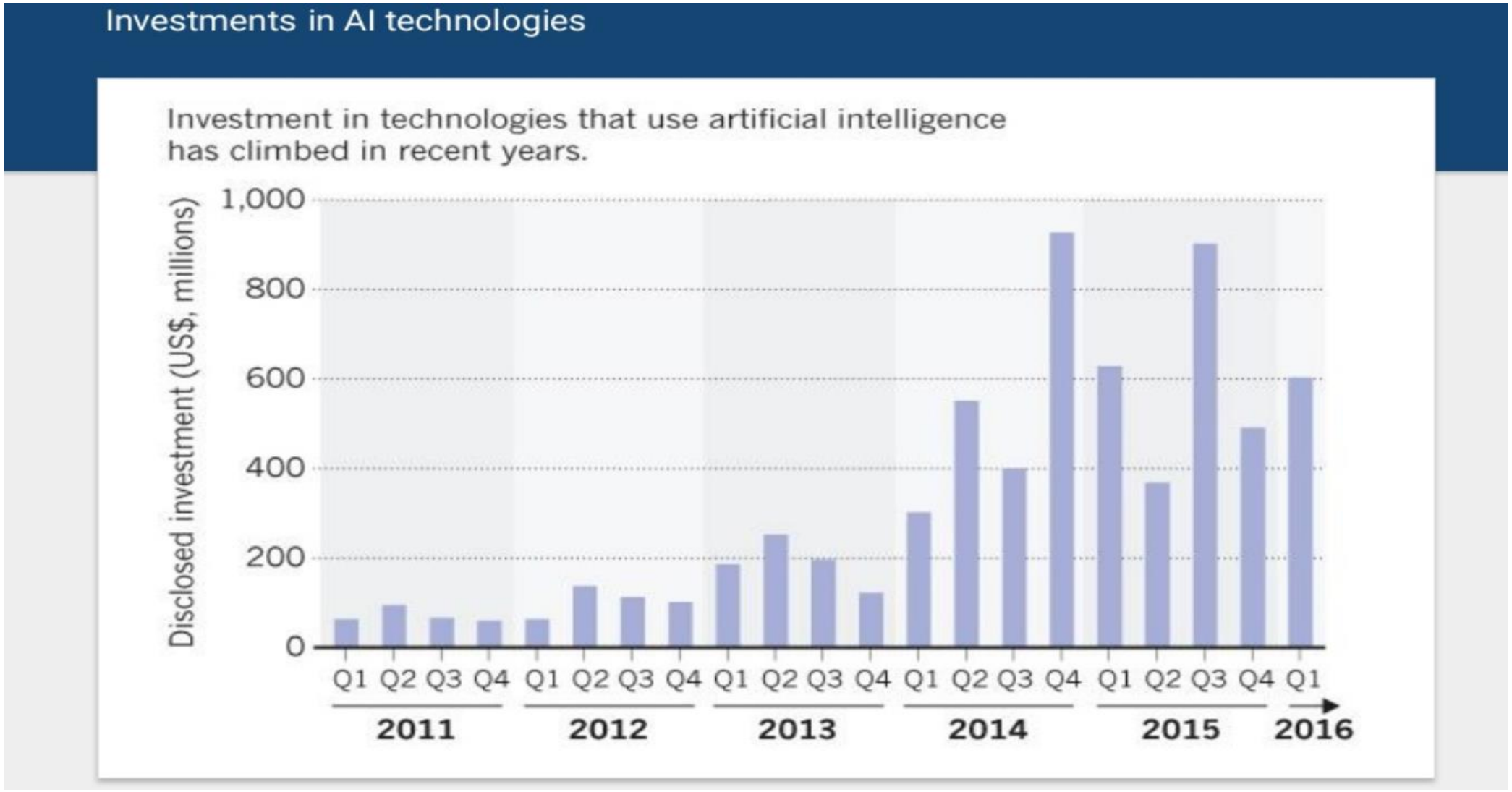


Potential of AI – why it is so important? - Opensource

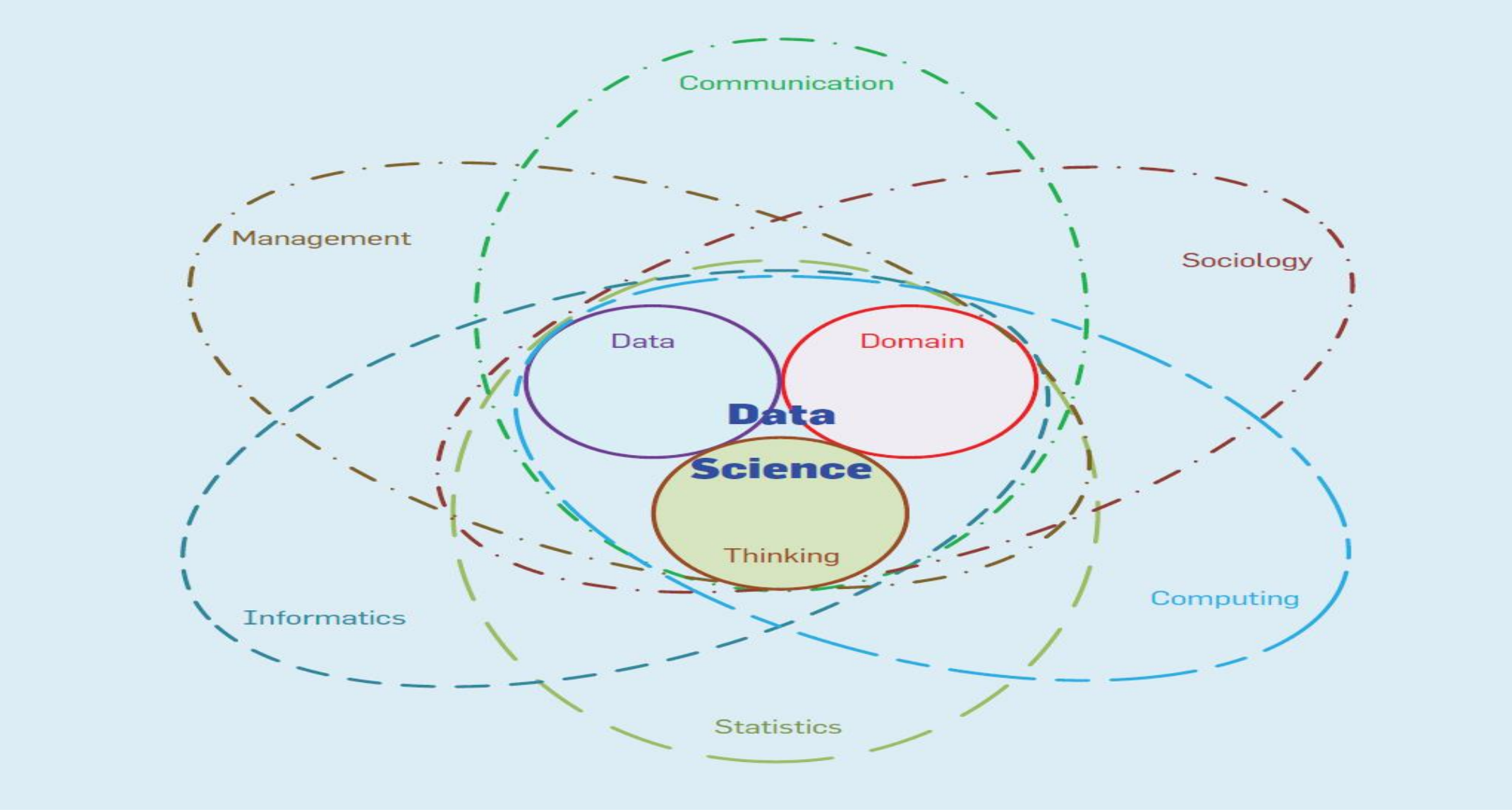
Potential of AI – why it is so important?



Potential of AI – why it is so important?



Data Science skills



Key stakeholders in AI based products

Table 1. Key stakeholders and challenges for each phase of a building a ML based product feature

	Ideation	Execution	Operation
Key Stakeholders	Product Managers Designers	Data Scientists Data Engineers Developers	SRE Support
Challenges	Data availability Privacy concerns Project risk estimation	Build vs rent Scalability issues Productionisation	Maintain accuracy Stability of data sources

- **The questions to be answered include: what is the right metric we try to optimise for?**
- **Are the data and infrastructure required for operation currently available?**
- **How confident are we that a ML system can be built with some minimum level of performance in key metrics like mean square error, or precision and recall?**
- **Currently there is not much research that helps practitioners estimate the anticipated performance of the system given the data available, in terms of samples, features and some methodology. They need to rely on rules of thumb, their experience and previous efforts in similar systems to make an estimation**

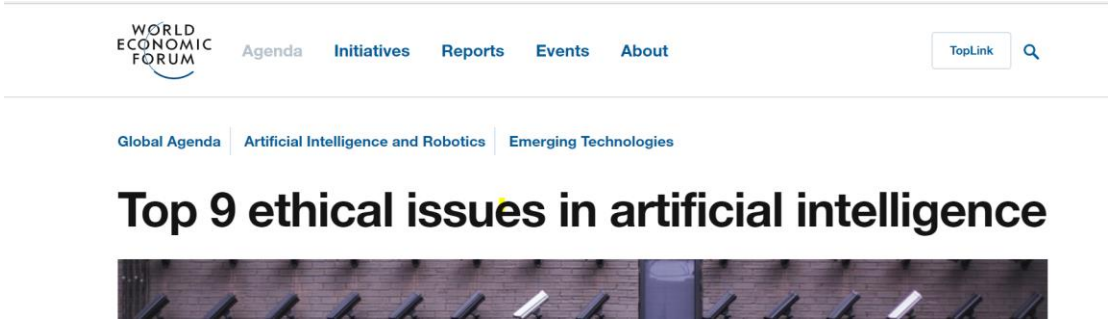
The Problem

To find ideas, find the problems.

To find the problems, talk to people.

Great idea pitches starts by defining change in the world

Problem – current issues



1. Unemployment. What happens after the end of jobs?
2. Inequality. How do we distribute the wealth created by machines?
3. Humanity. How do machines affect our behaviour and interaction?
4. Artificial stupidity. How can we guard against mistakes?
5. Racist robots. How do we eliminate AI bias?
6. Security. How do we keep AI safe from adversaries?
7. Evil genies. How do we protect against unintended consequences?
8. Singularity. How do we stay in control of a complex intelligent system?
9. Robot rights. How do we define the humane treatment of AI?

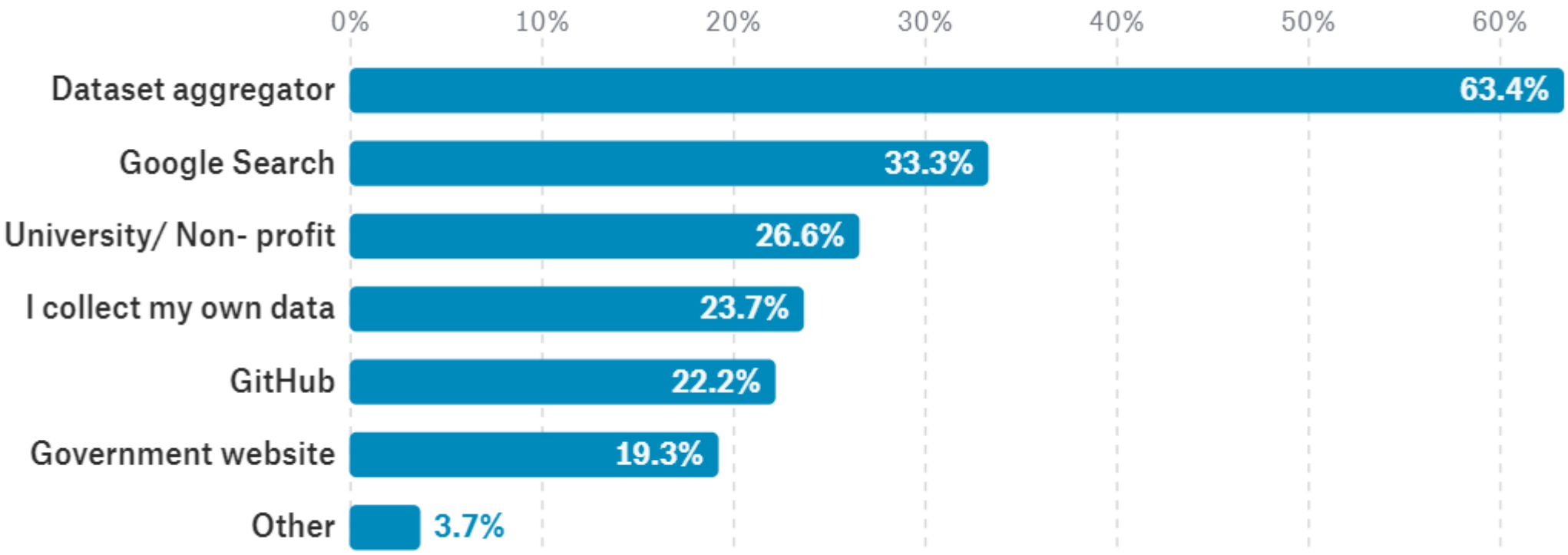
Problem – current issues

Ethical Issues

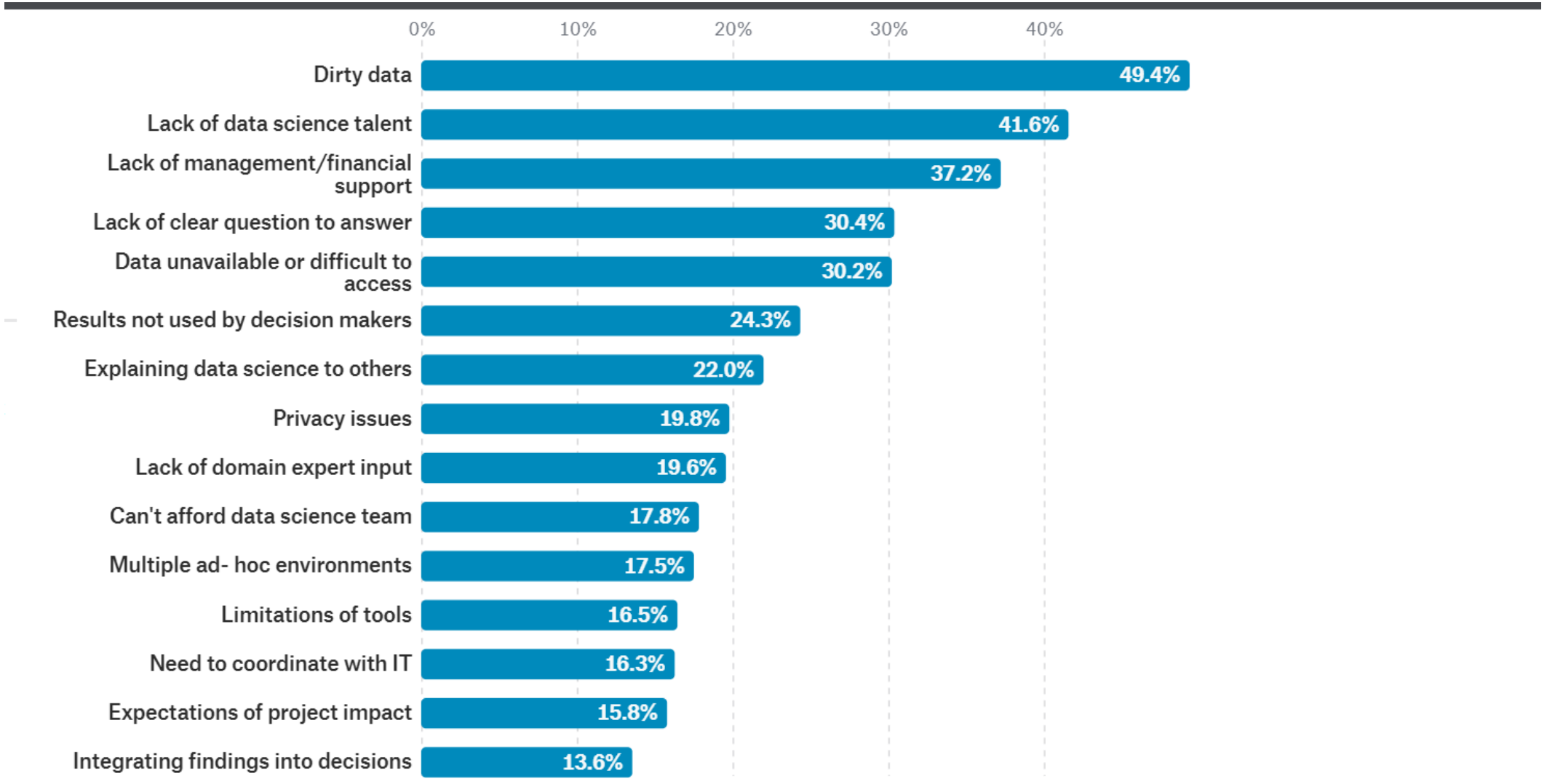
Technical issues

Issues for research advancements

Technical issues: Open Data Scraping



Survey - What barriers data scientists are faced at work?



Problem Statement – Data Monopoly

- Big tech firms like Google, Facebook, and Microsoft are today's coal mines. They have abundant data and so can afford to run inefficient machine learning systems, and improve them. Smaller startups might have good ideas, but they won't be able to follow through without data.
- Take health care, for example, where AI is being used for machine vision tasks like recognizing

  tumors in X-ray scans, but where digitized data can be sparse.

the National Health Service in the UK so significant.) The problem, says Lawrence, is not really about finding ways to distribute data, but about making our deep learning systems more efficient and able to work with less data. And just like Watt's improvements, that might take another 60 years.

Problem Statement - Privacy

A very sensitive area that affects decisions on the feasibility or the design of a ML system are concerns around the privacy of users and their data.

The legal framework within which ML systems operate is changing fast as governments try to regulate the access of companies to the private data of their citizens. Legal and risk departments of companies are getting involved and the design of ML based systems is severely affected in an effort to provide compliance with the privacy policy of the organisation and the legal framework

Users should also become more sensible about what they share with companies or post on the

In the hunt for more and more data, companies may trek into uncharted territory and cross privacy boundaries. Such was the case of a retail store that found out about a teenage girl's secret pregnancy, and the more recent case of UK National Health Service's patient data sharing program with Google's DeepMind, a move that was supposedly aimed at improving disease prediction.

making it any

Problem Statement - Algorithmic bias

As has been proven on several accounts in the past years, AI can be just as — or even more — biased than humans.

The problem is, if the information trainers feed to these algorithms is unbalanced, the system will eventually adopt the covert and overt biases that those data sets contain. And at present, the AI industry is suffering from diversity troubles that some label the "white guy problem," or largely dominated by white males.

This is the reason why an AI-judged beauty contest turned out to award mostly white candidates, a name-ranking algorithm ended up favoring white-sounding names, and advertising algorithms preferred to show high-paying job ads to male visitors.

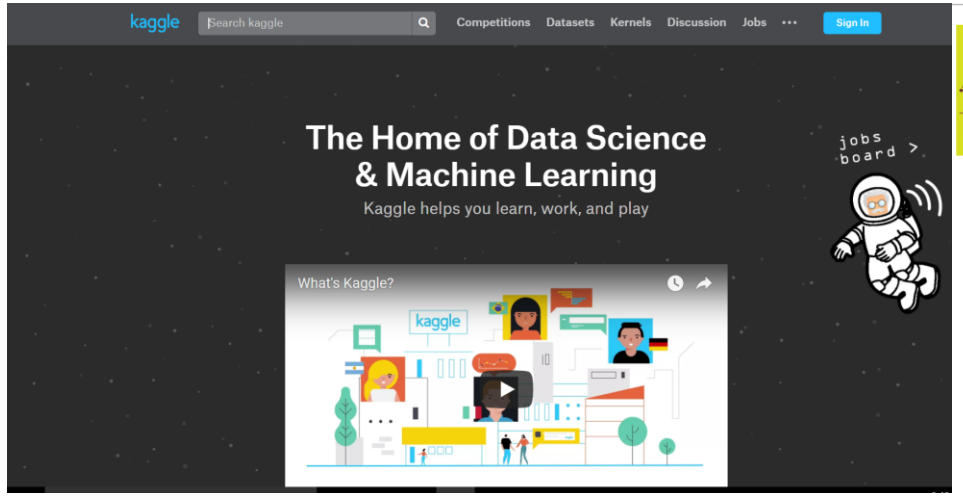
Another problem that caused much controversy in the past year was the "filter bubble" phenomenon that was seen in Facebook and other social media that tailored content to the biases and preferences of users, effectively shutting them out from other viewpoints and realities that were out there.

This can be achieved by promoting transparency and openness in algorithmic datasets. Shared data repositories that are not owned by any single entity and can be vetted and audited by independent bodies can help move toward this goal.

Problem Statement - No reward for sharing proper data

- **In the data sharing case, data producers make efforts to prepare the data for deposit, but the benefit of the data preparation largely goes to secondary users.**
- **In addition, data producers are at risk of being harmed by the misuse and misinterpretation of data by unqualified users, or by being charged with misconduct.**
- **To motivate data producers to prepare and share data, there must be some incentive mechanisms.**

Current Open data science systems



2018 Data Science Bowl
Find the nuclei in divergent images to advance medical discovery
Featured · 20 days to go · biology
$100,000
3,067 teams

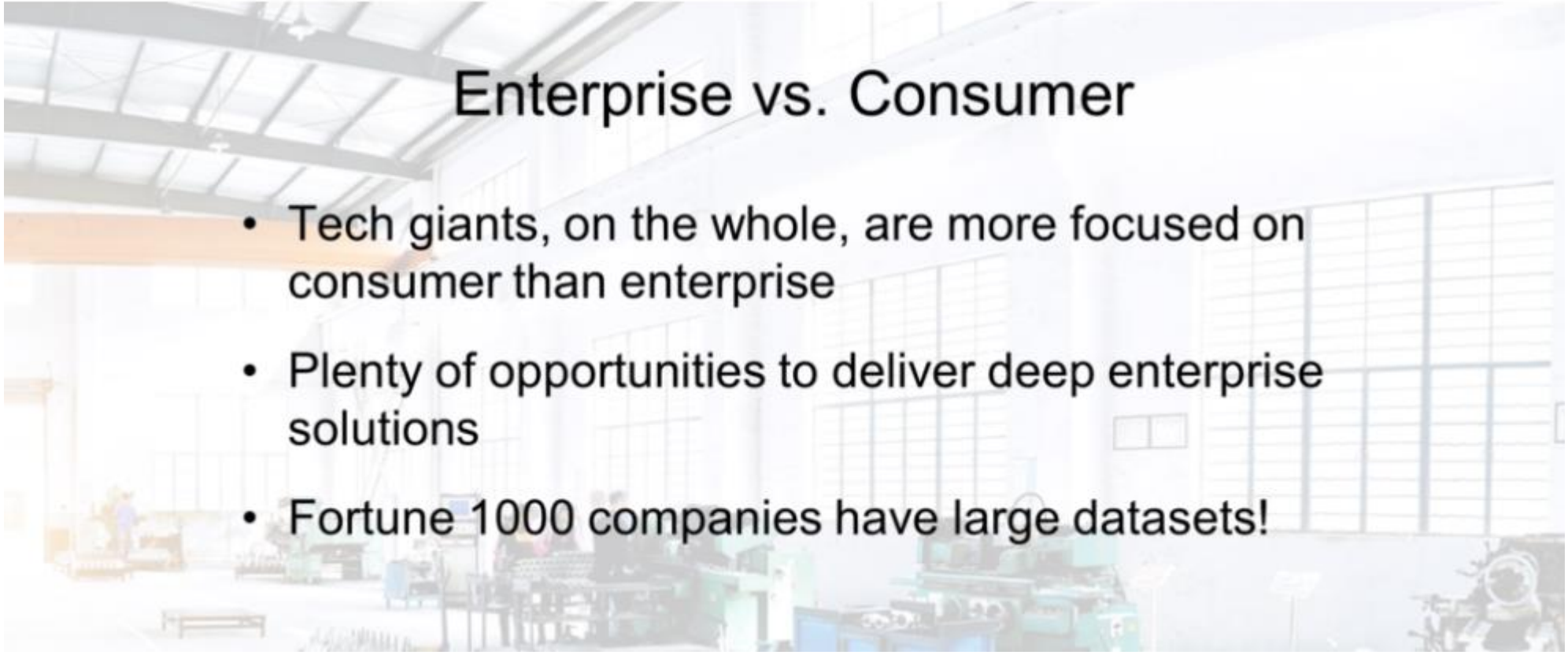
TalkingData AdTracking Fraud Detection Challenge
Can you detect fraudulent click traffic for mobile app ads?
Featured · a month to go ·
$25,000
1,698 teams

Toxic Comment Classification Challenge
Identify and classify toxic online comments
Featured · 7 days ago · arguments, text data
$35,000
4,551 teams

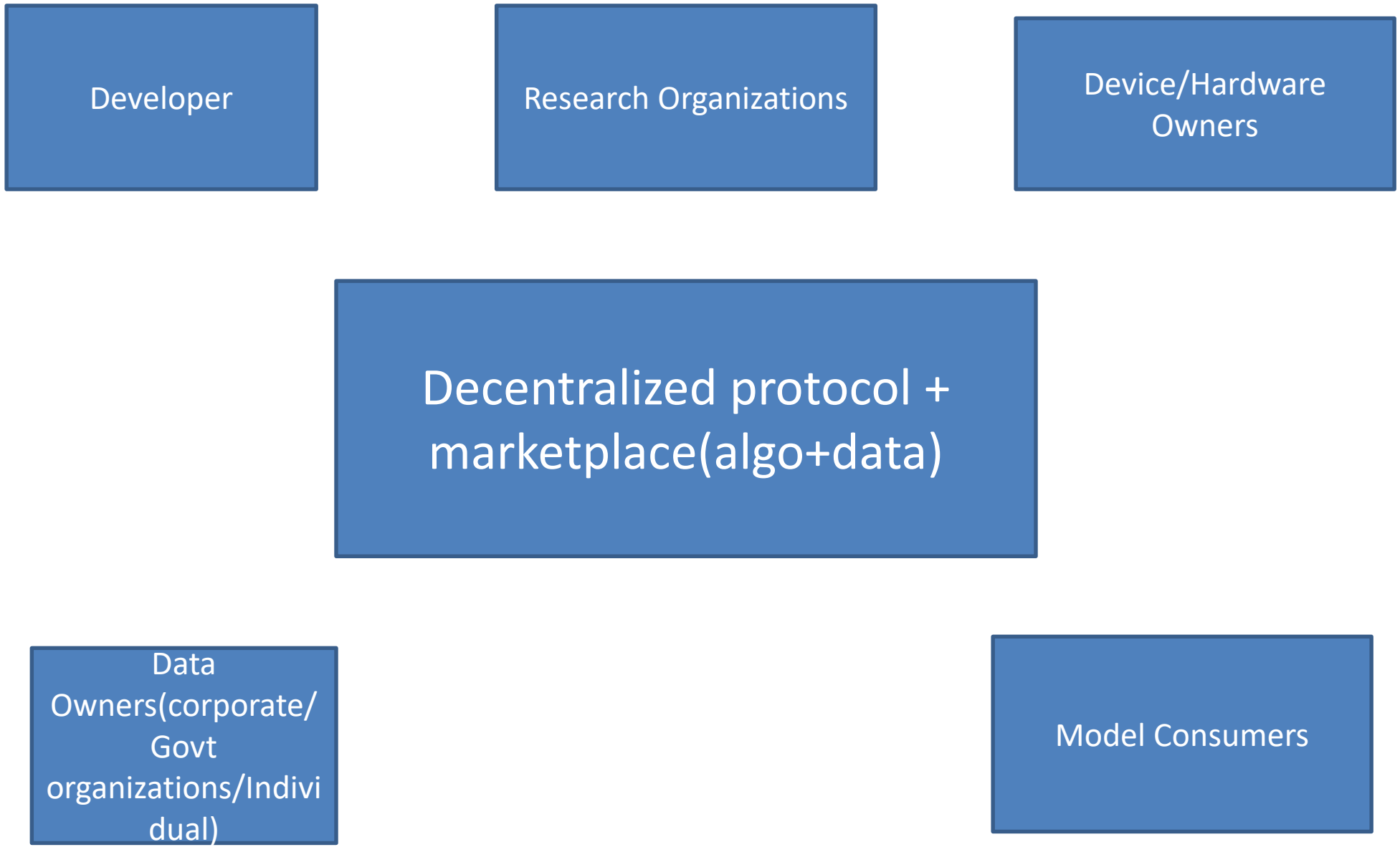
Mercari Price Suggestion Challenge
Can you automatically suggest product prices to online sellers?
Featured · a month ago ·
$100,000
2,384 teams

Zillow Prize: Zillow's Home Value Prediction (Zestimate)
Can you improve the algorithm that changed the world of real estate?
Featured · 3 months ago · housing, real estate
$1,200,000
3,779 teams

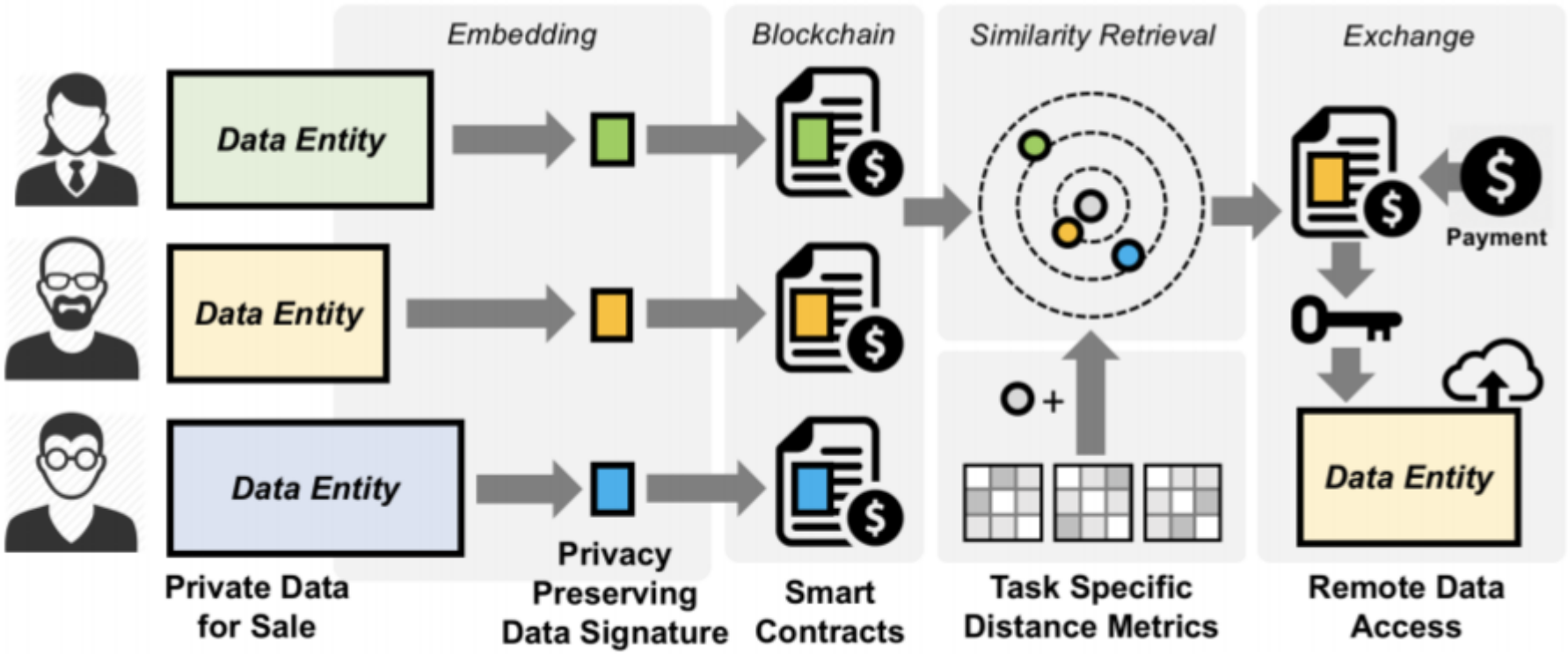
Problem statement - opportunities



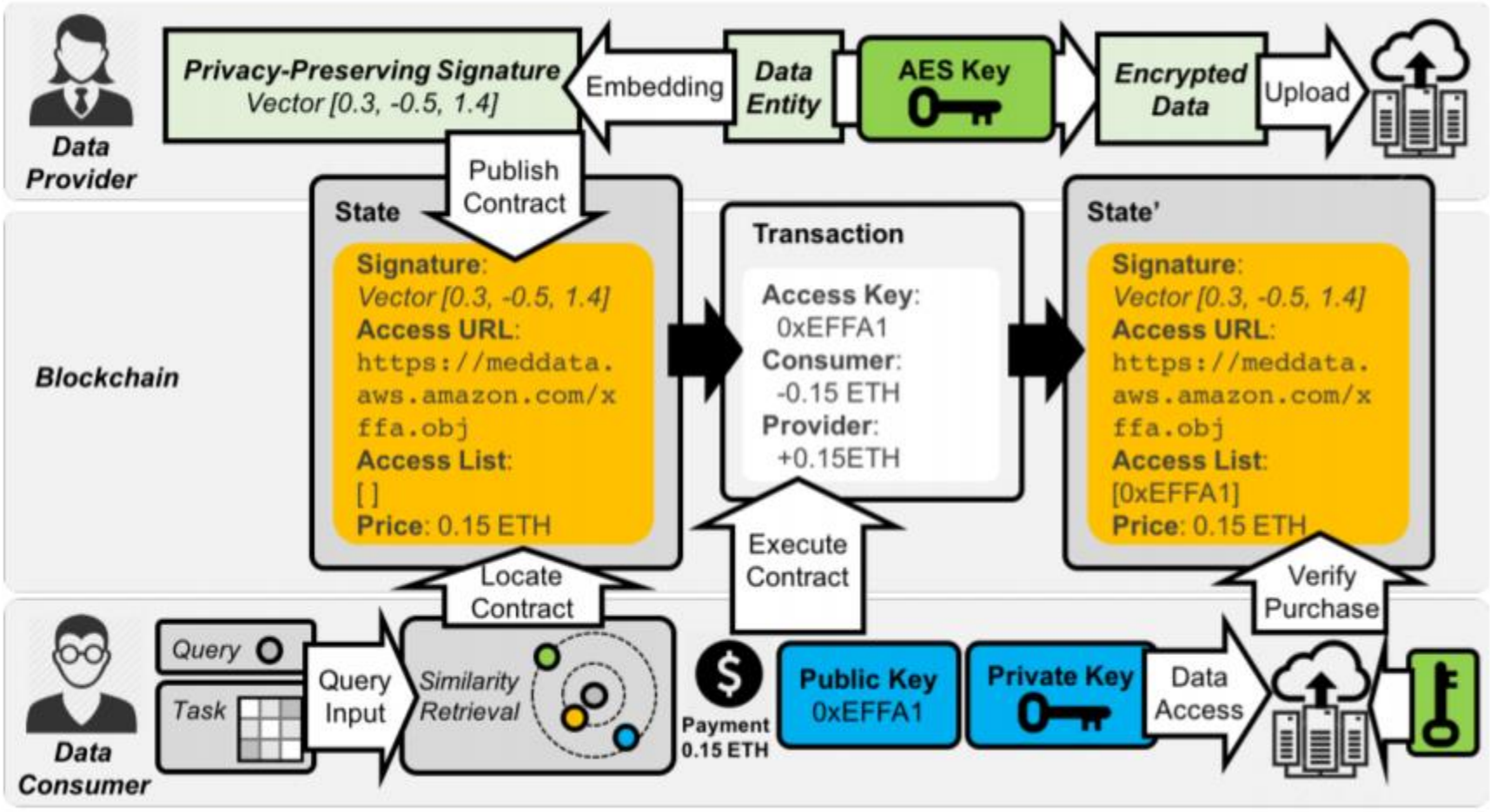
The Solution



The Solution - Data Vending



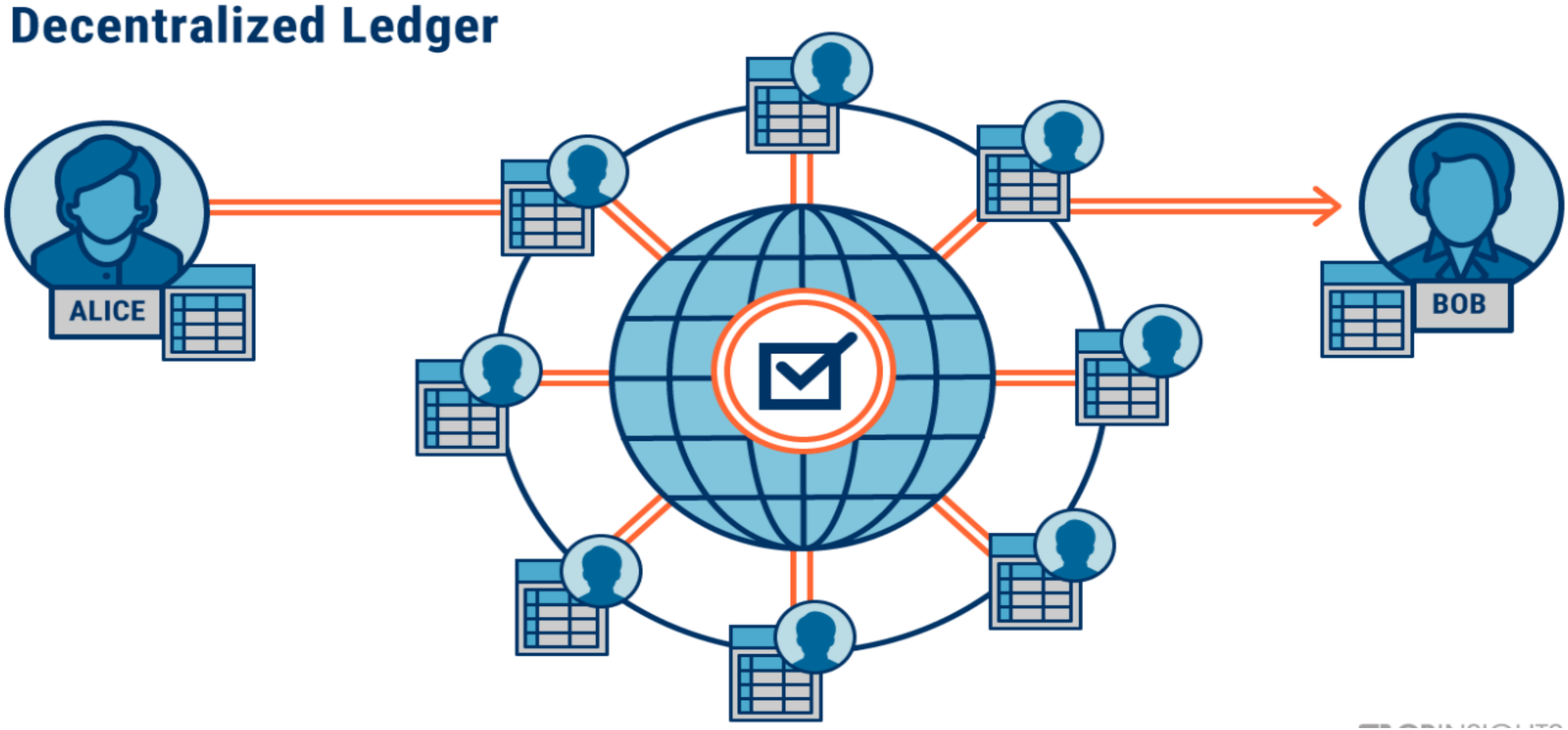
The Solution - Data Vending – deep dive



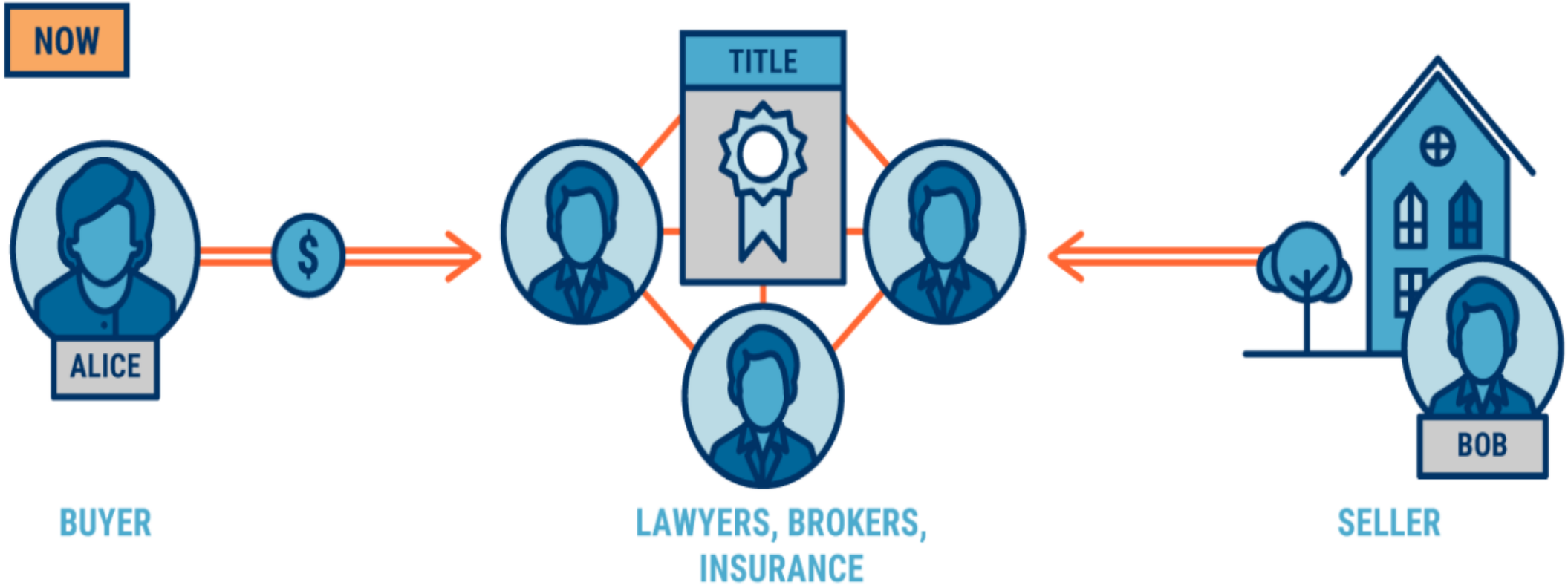
Solution Overview

- **Utilizing untapped private data for machine learning while protecting data privacy**
- **connecting and leveraging idle processing power of individual devices for machine learning**
- **encouraging involvement from the periphery by creating a developer community and algorithm marketplace that promotes innovation to build machine learning algorithms that match practical utilities**
- **improving and correcting existing machine learning algorithms and models through crowdsourced fine-tuning model trainers**
- **Secured data vending/Transfer without data/information leakage**
- **creating a new DML utility token and leveraging on blockchain smart contract technology to provide a trustless and middle-man free platform that connects potential contributors in machine learning from all aspects**

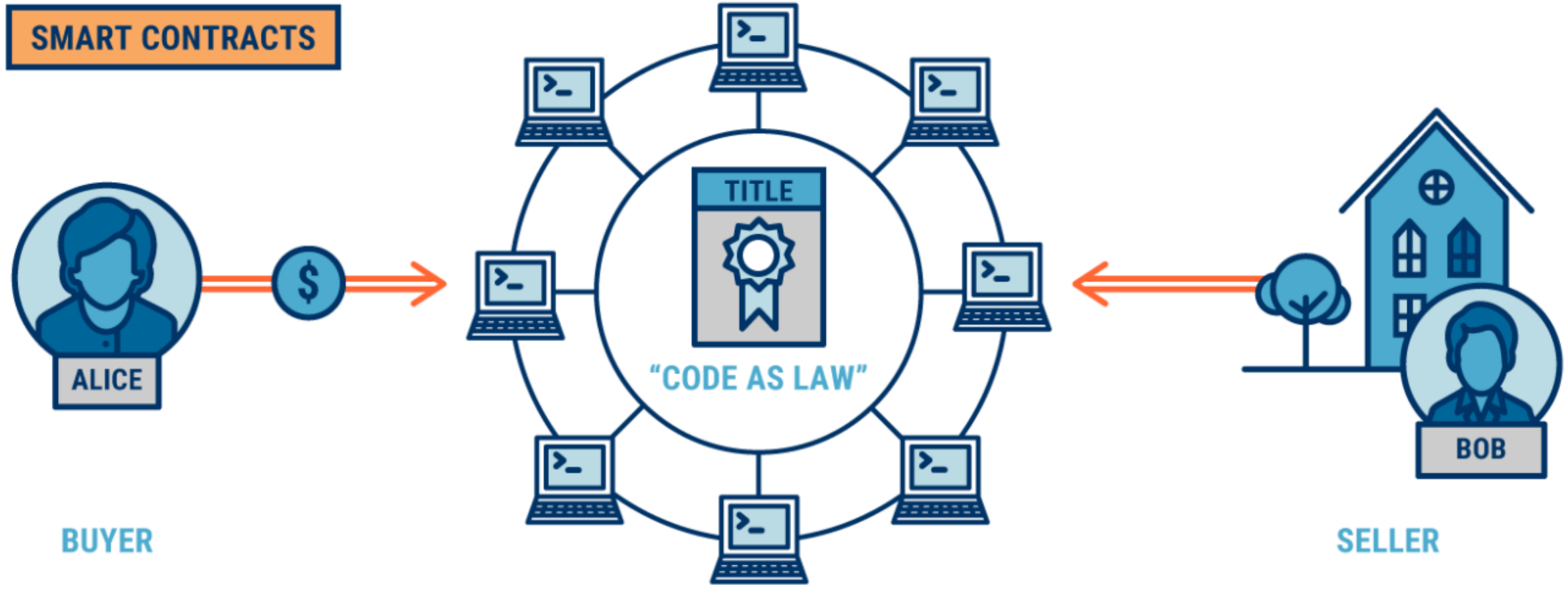
Blockchain - Explained



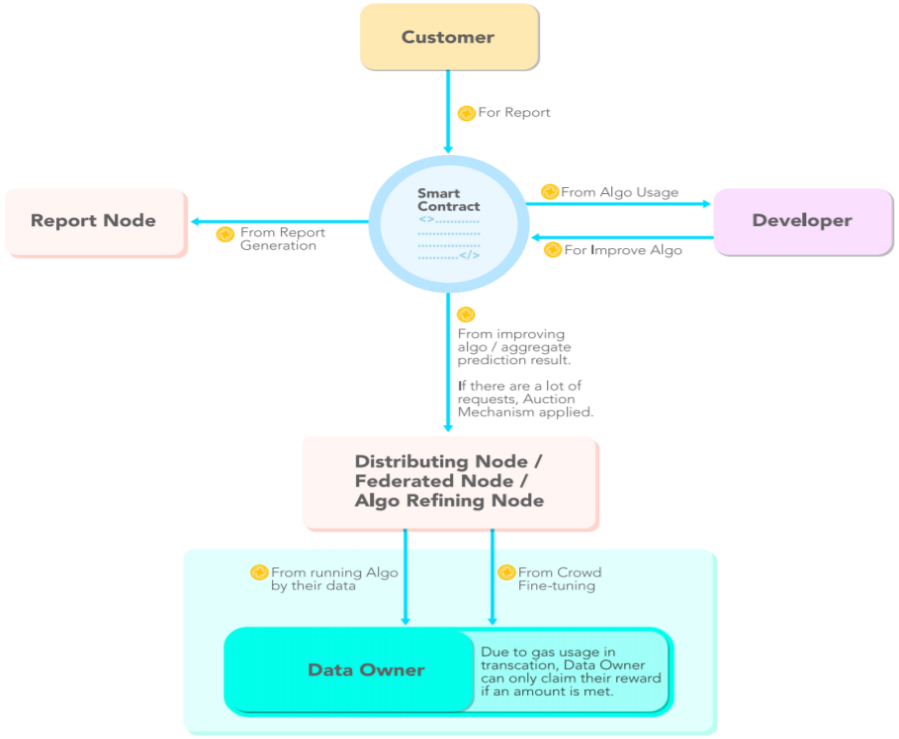
Smart Contract - Old way



Smart Contracts – New way



Working


Customer
For Report
Smart Contract
Report Node
From Report Generation
From Algo Usage
Developer
For Improve Algo
From improving algo / aggregate prediction result.
If there are a lot of requests, Auction Mechanism applied.
Distributing Node / Federated Node / Algo Refining Node
From running Algo by their data
From Crowd Fine-tuning
Data Owner
Due to gas usage in transcation, Data Owner can only claim their reward if an amount is met.

Thank you