Specialist Programme on Artificial Intelligence for IT & ITES Industry

AI Applications & Best Practices

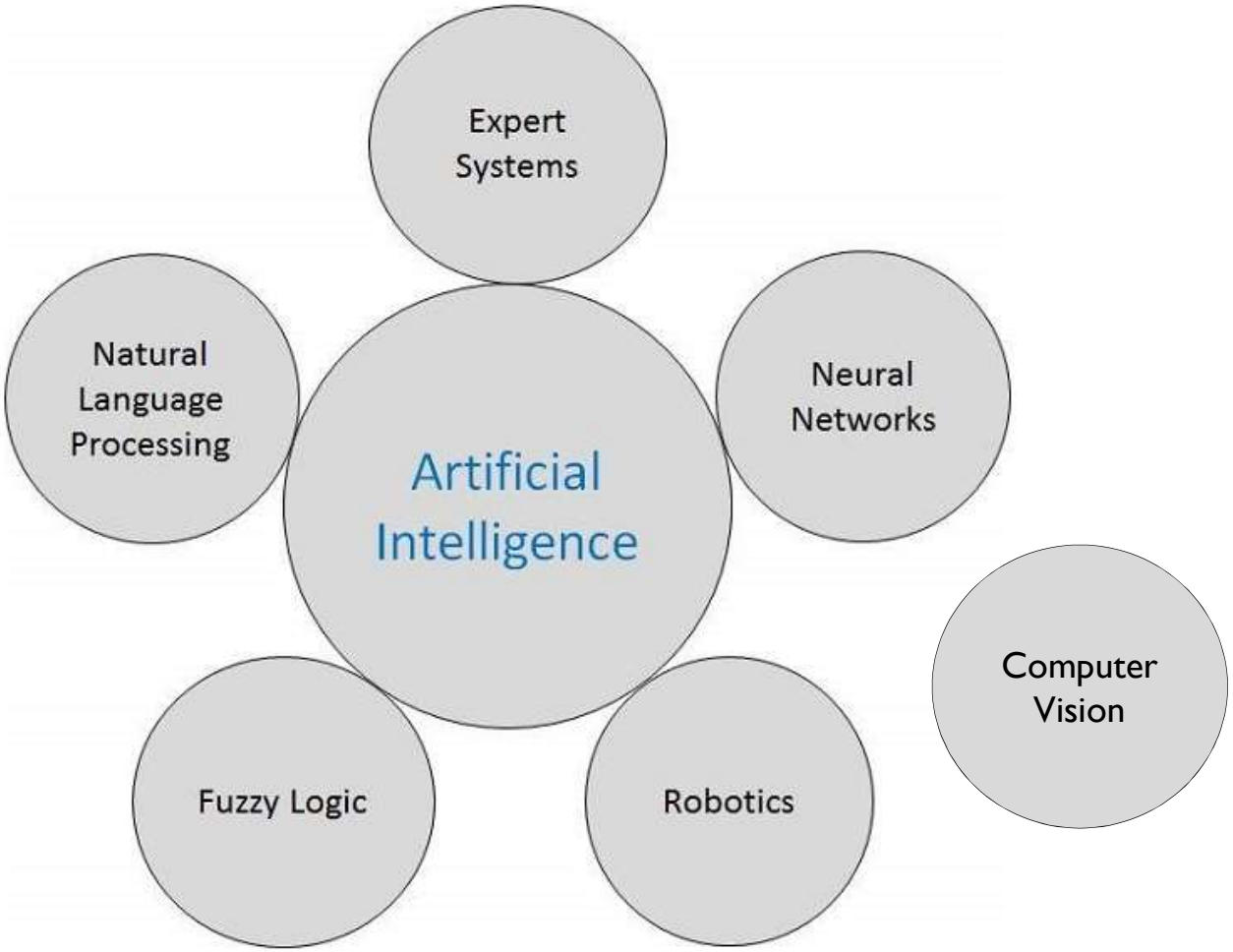
Dr Barry Shepherd
barryshepherd@nus.edu.sg

Singapore e-Government Leadership Centre
National University of Singapore

Agenda

- AI Applications
- Process and best practices
- Challenges & Issues

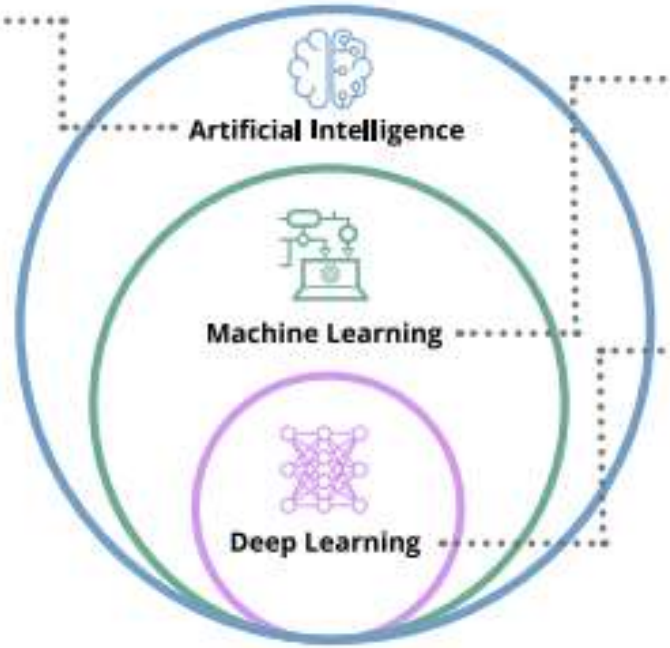
AI – Traditional View



AI – Current View

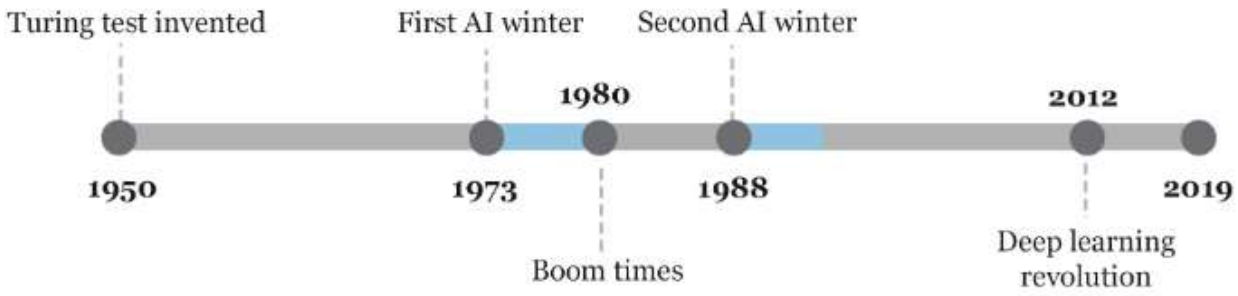
Artificial Intelligence (AI)
A computer process that has learned to solve tasks in a way that mimics human decisions

AI solutions today are mostly used for very specific tasks, versus general applications

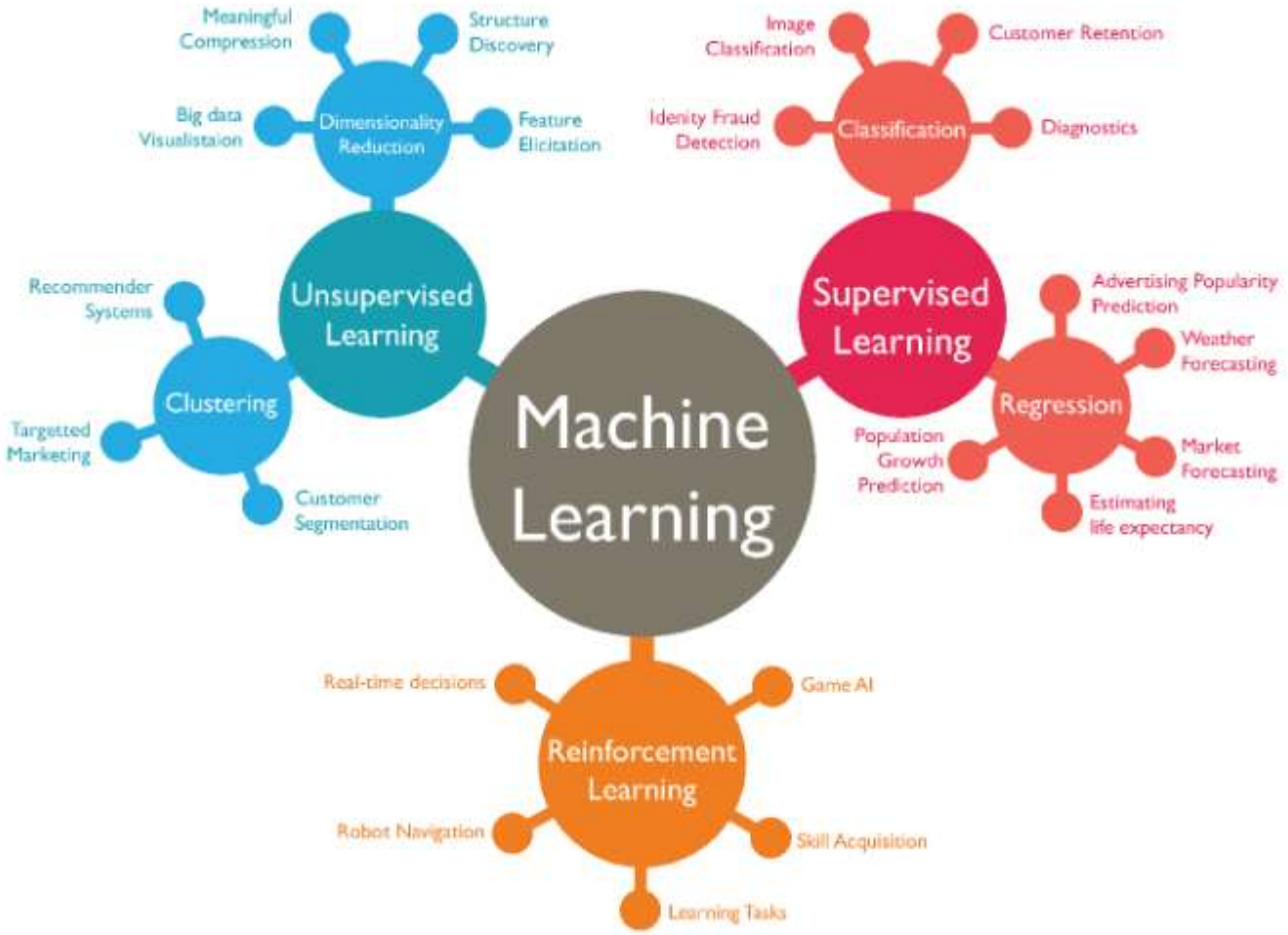


Machine Learning (ML)
Uses algorithms to help computers learn by task-specific examples and progressive improvements, without explicit programming

Deep Learning (DL)
Uses a cascade of processing layers modeled on neural networks to learn data representations such as features or classifications



The World through an ML lens



https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications

AI is not just Machine Learning

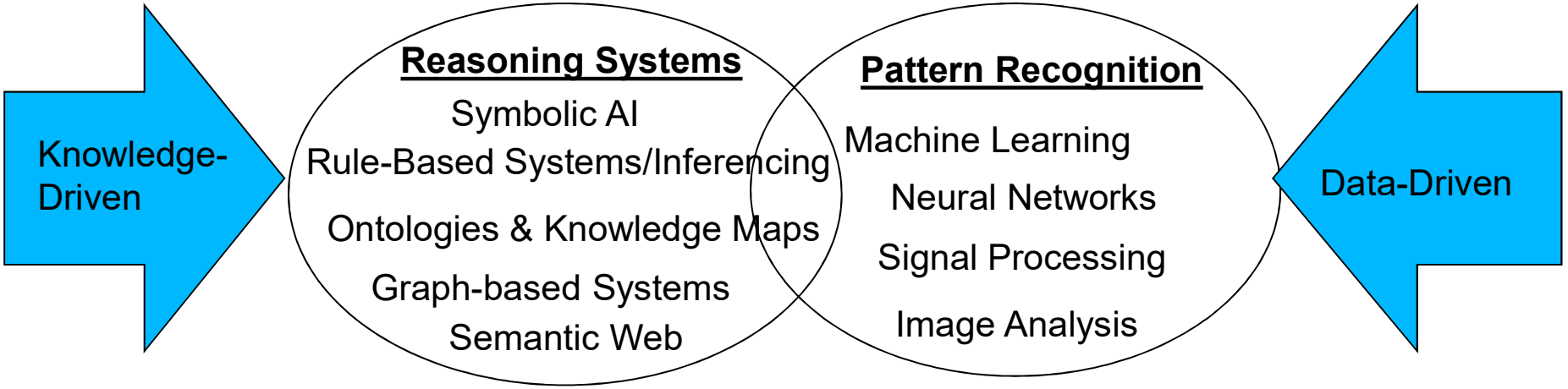
GOVTECH SINGAPORE

https://www.tech.gov.sg/media/technews/practical-applications-of-artificial-intelligence

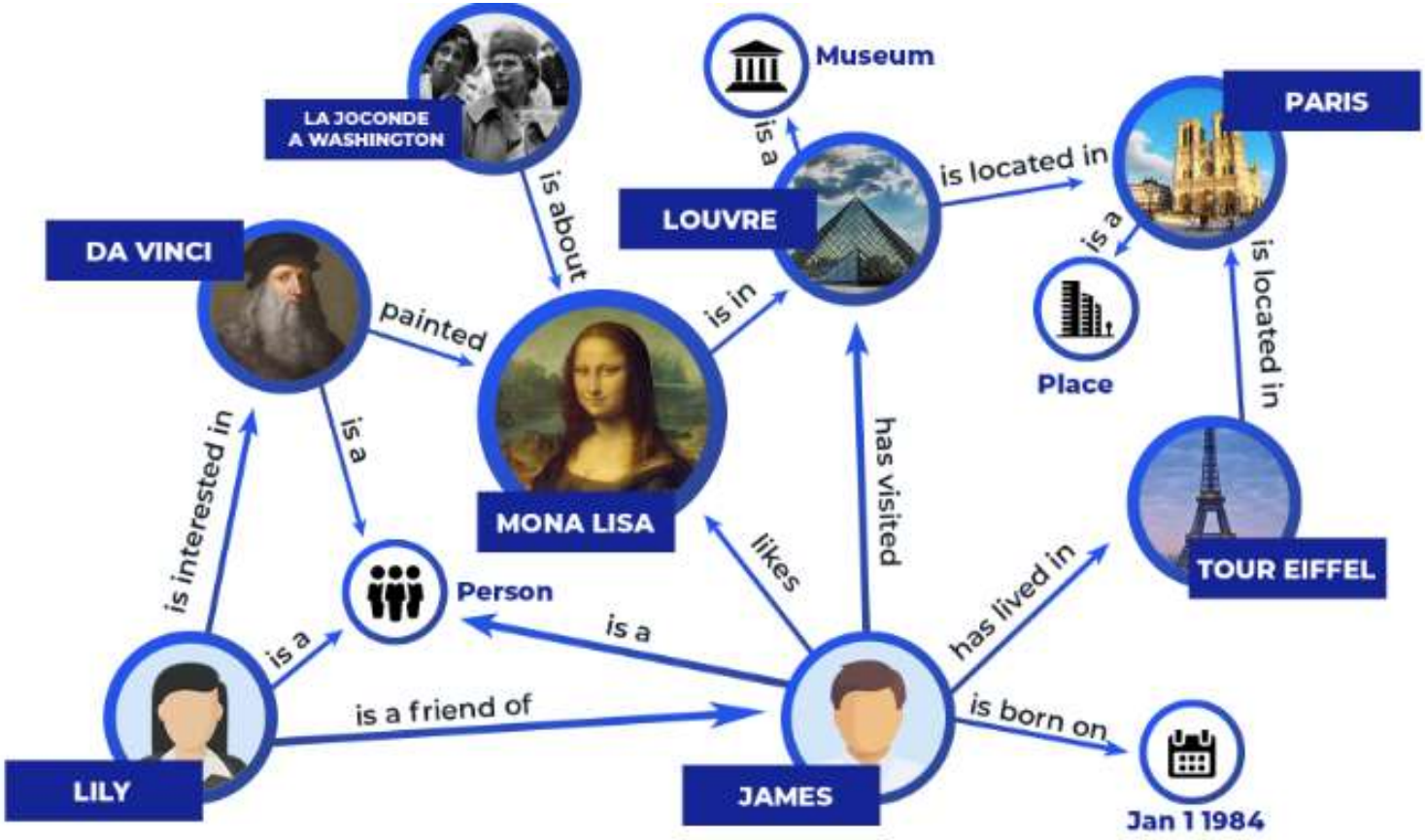
Practical applications of artificial intelligence

01 JUL 2019

"Contrary to popular sentiment, AI is not a homogenous technology, nor is it a magic bullet for all problems. Rather, ***AI is an umbrella term*** *encompassing innovations in the fields of natural language processing, machine translation, image and graph analytics, visualisation and much more, Dr Enzell noted. Hence,* ***matching the appropriate technological capability with the right use case is critical****."*

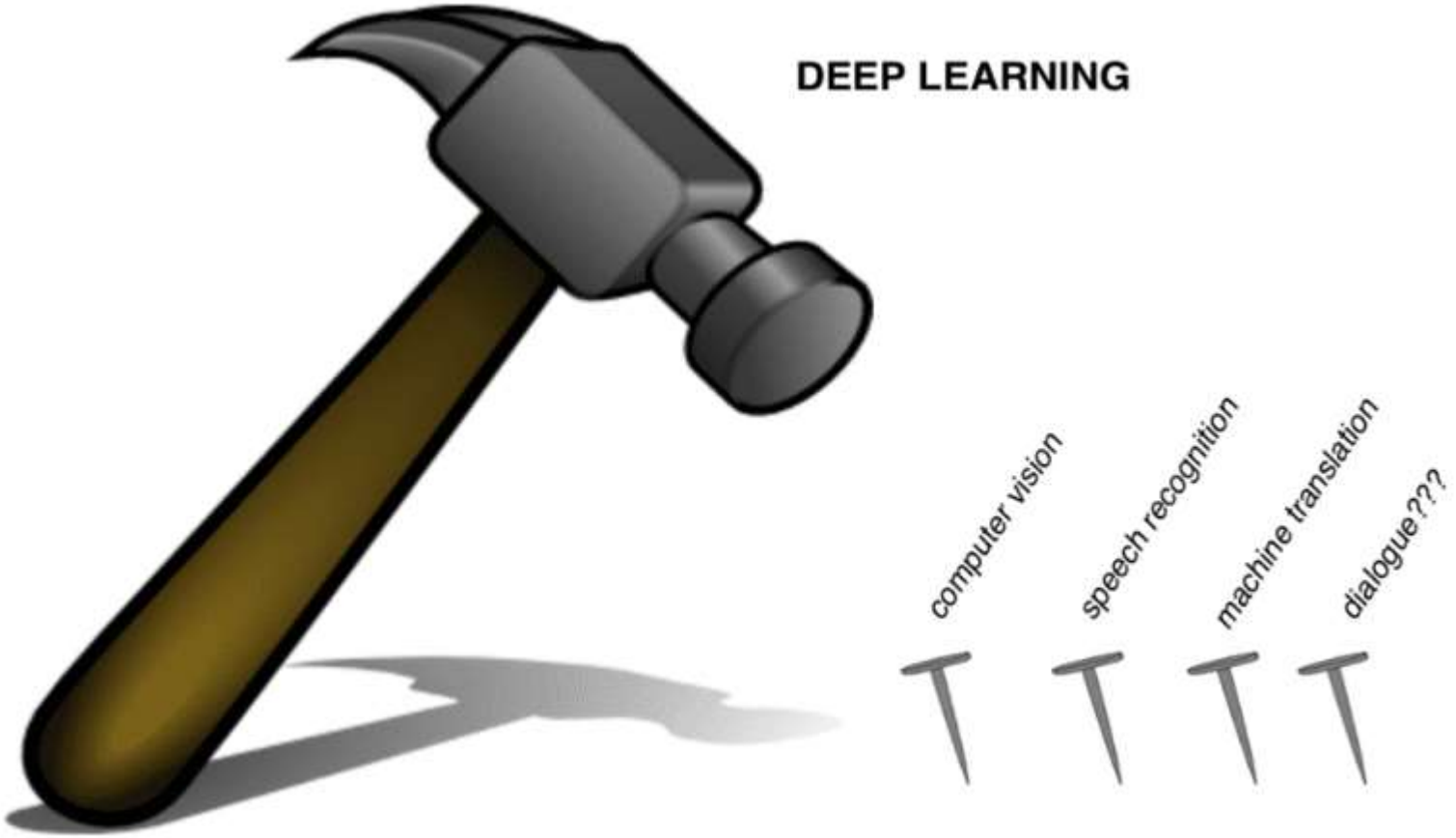


Many Overlaps ....



The Knowledge Graph

Not Every Problem is a Nail



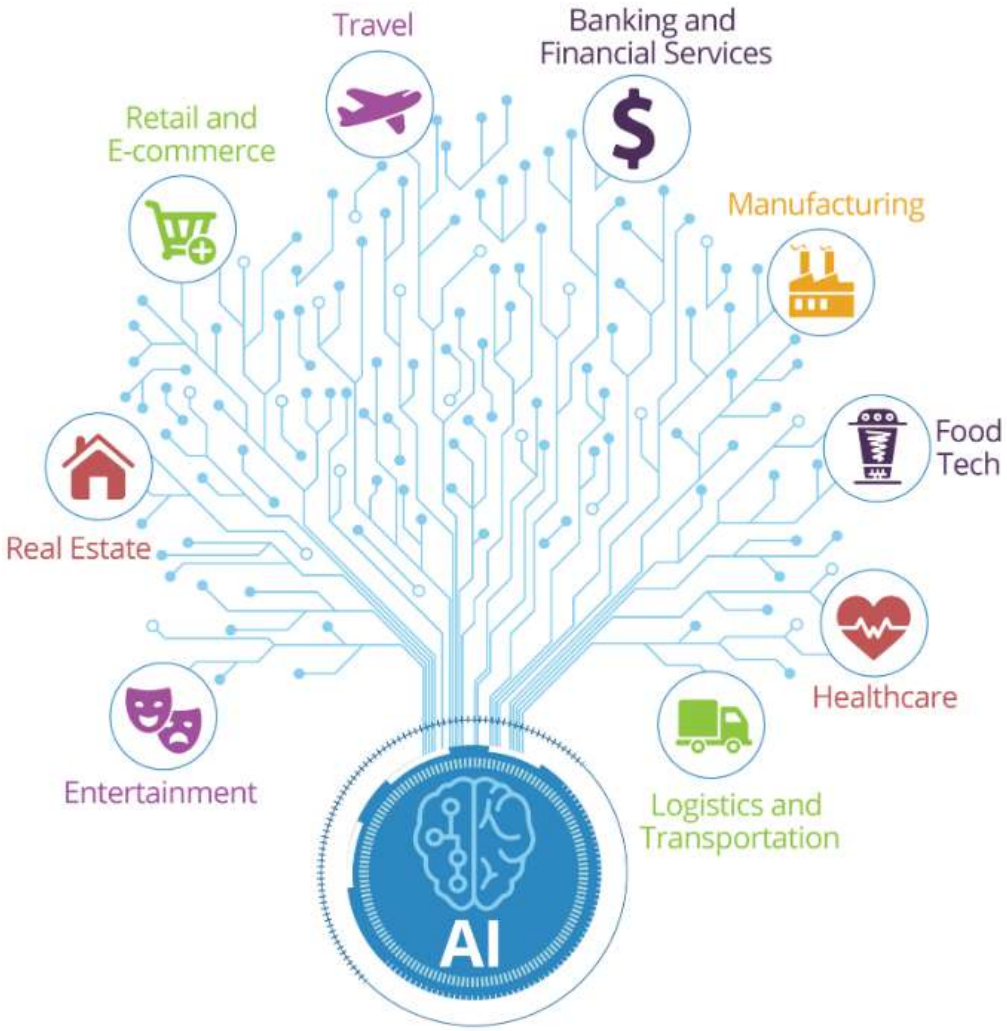
Data (lots) is crucial for AI…..

"the foundation for developing practical AI applications lies in the ***ability to collect and integrate data****"*

And the ability to explore and understand it…..

Visualisation of data*,* ***once it has been analysed by AI, is just as important****, said Mr Lu. He showed how the command centre of Jing'an district in Shanghai, China, has an integrated interface which displays event assessment and analysis reports as well as real-time information about the goings-on in the city.*

AI Applications Overview



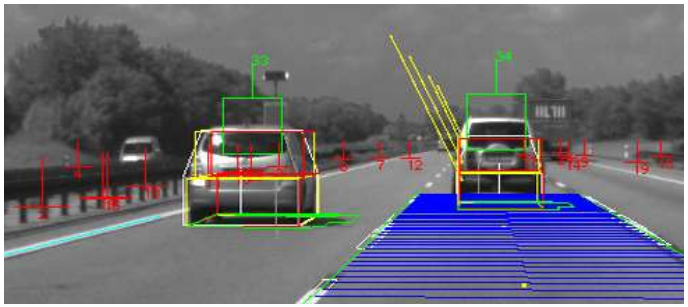
Many taxonomies are possible, lets consider a simple one…..

- Physical Domain
  - sensor-driven systems, e.g. robots, drones
- Business Domain
  - non-sensor data, e.g. transactions, customer data
  - …but also sensor-data

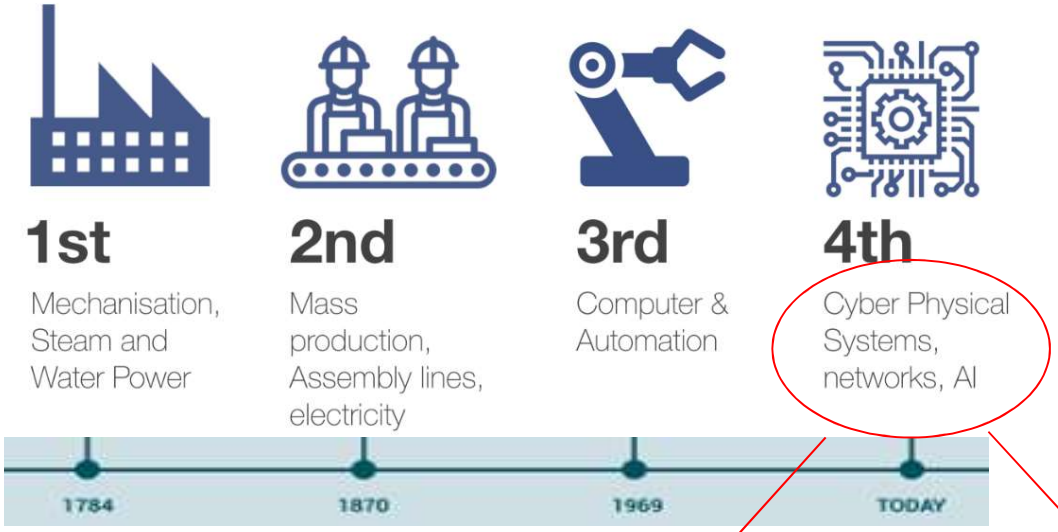
AI Applications: Sensor-Driven Systems

- Applications usually involving ***physical interaction*** in the real-world
- Often involve processing of sensory data

- ***Smart Manufacturing*** - Industry 4.0
- ***Smart City*** – smart buildings, transport networks, IOT..
- ***Smart Healthcare*** – medical devices, robot nurses, robot surgeons, tele-medicine (VR and AR tools),..
- ***Service Robots*** – shopping assistants, restaurant workers, care for the elderly, domestic help,..
- ***Autonomous Vehicles*** – self-driving cars, trucks, buses,..
- ***Drones*** – inspection, surveillance, deliveries, …
- ***Vision Systems*** - *surveillance, inspection, navigation,.*



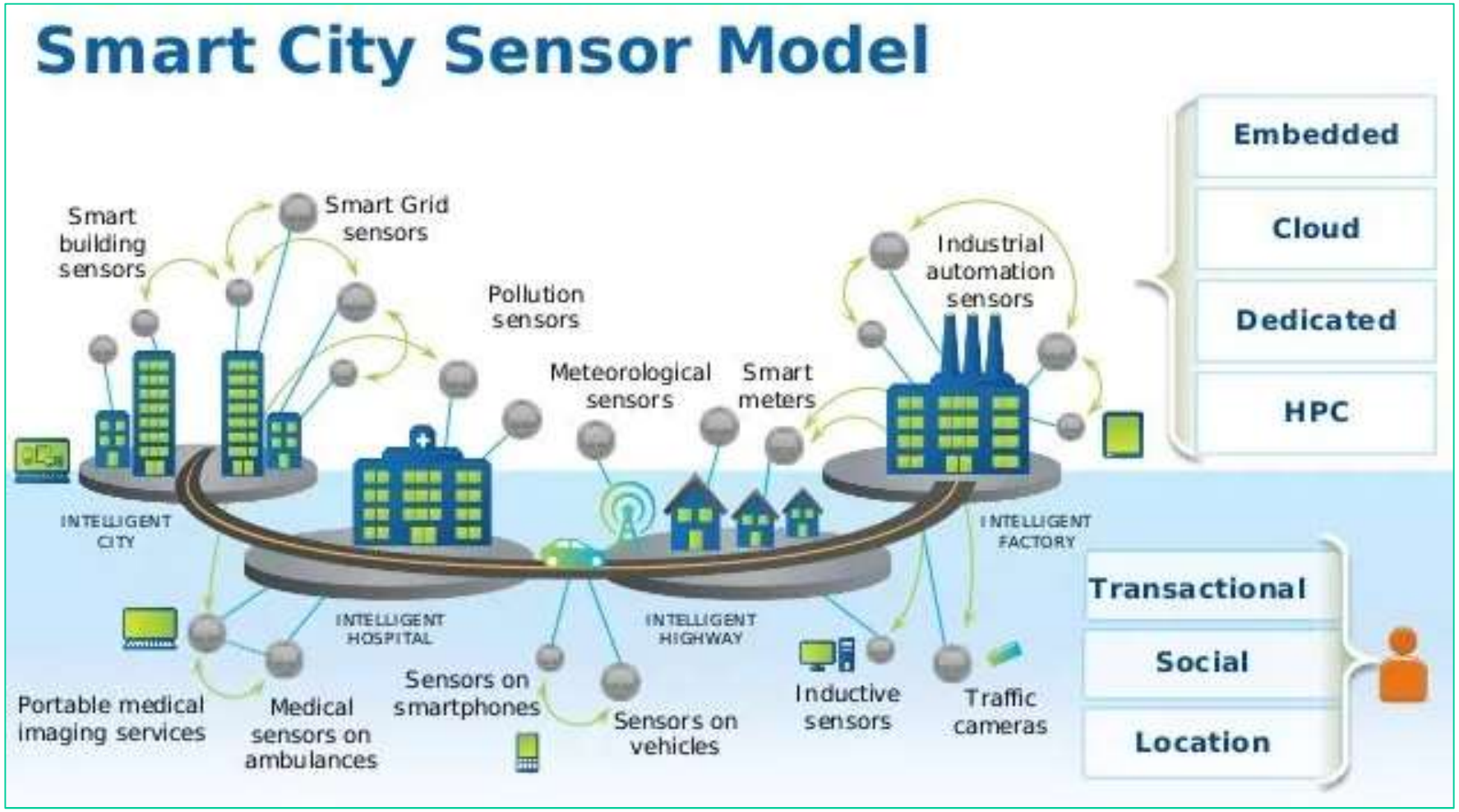
Smart Industry – Industry 4.0



The Fourth Industrial Revolution is marked by emerging technology breakthroughs in a number of fields, including:
Robotics, *Artificial Intelligence*, *Nanotechnology*, *Quantum Computing*, *Biotechnology*, *the Internet of Things*, *3D printing and Autonomous Vehicles*
(Wikipedia)

Cyber Physical Systems ~ typically a ***network*** of interacting elements ……, closely tied to concepts of ***robotics*** and ***sensor networks*** with ***intelligence mechanisms*** of ***computational intelligence*** leading the pathway

Human-Robot Cooperation - Cobots



Smart cities use Internet of Things (**IoT**) devices such as connected sensors, lights, and meters to collect and analyze data. The **cities** then use this data to improve infrastructure, public utilities and services, and more

Smart City Data Sources

Examples…

THE STRAITS TIMES

Building a Smart Nation

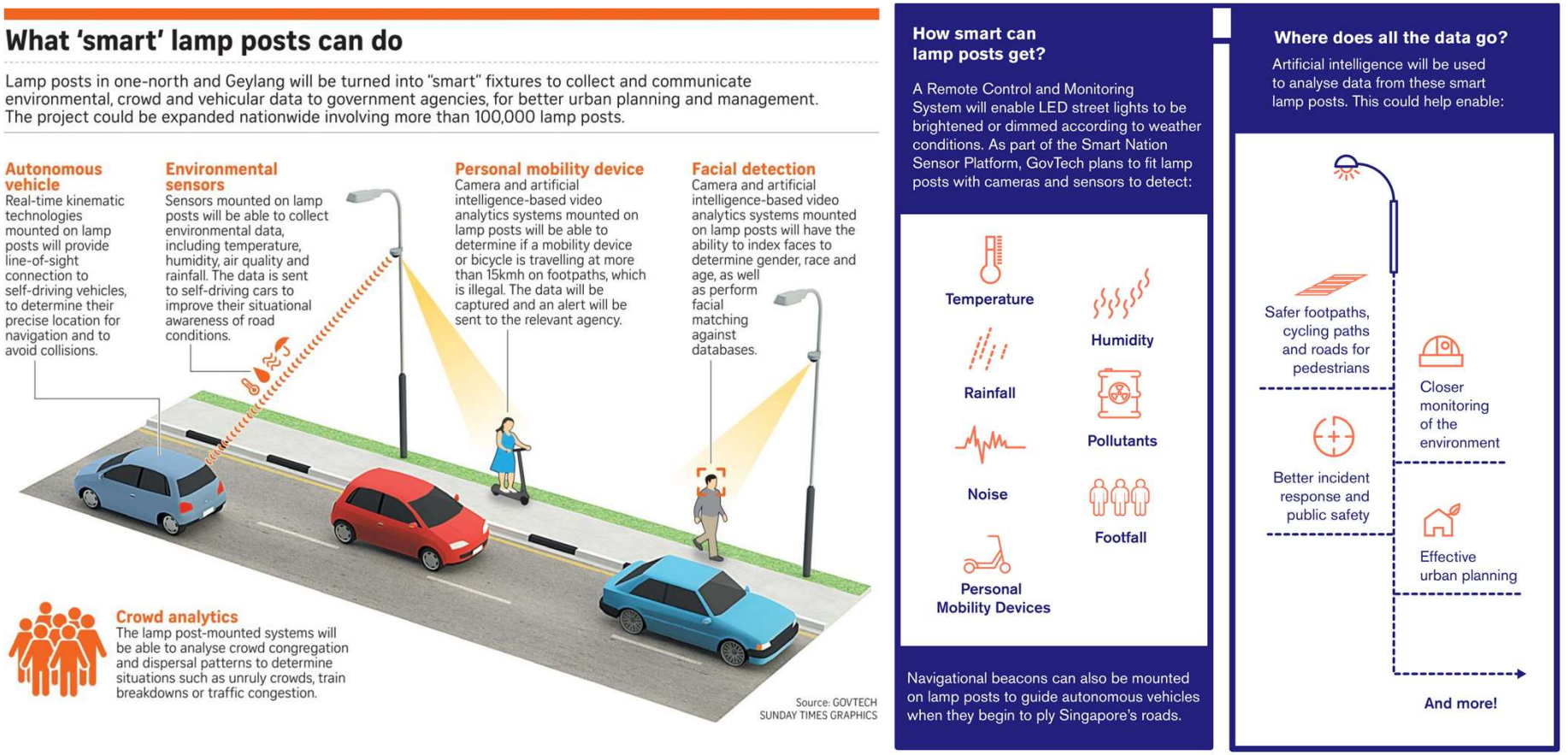
'Smart' lamp posts to transmit data

PUBLISHED AUG 21, 2017, 5:00 AM SGT

Data collected from a network of interconnected lamp posts can be used to detect anomalies and predict situations such as potentially unruly crowds and traffic congestion. PHOTO: ST FILE

"This network of interconnected lamp posts could form the spine of the Smart Nation Sensor Platform (SNSP), which aims to use artificial intelligence (AI) technologies to analyse, for instance, video footage collected by various government agencies"

Smart Lamp Posts



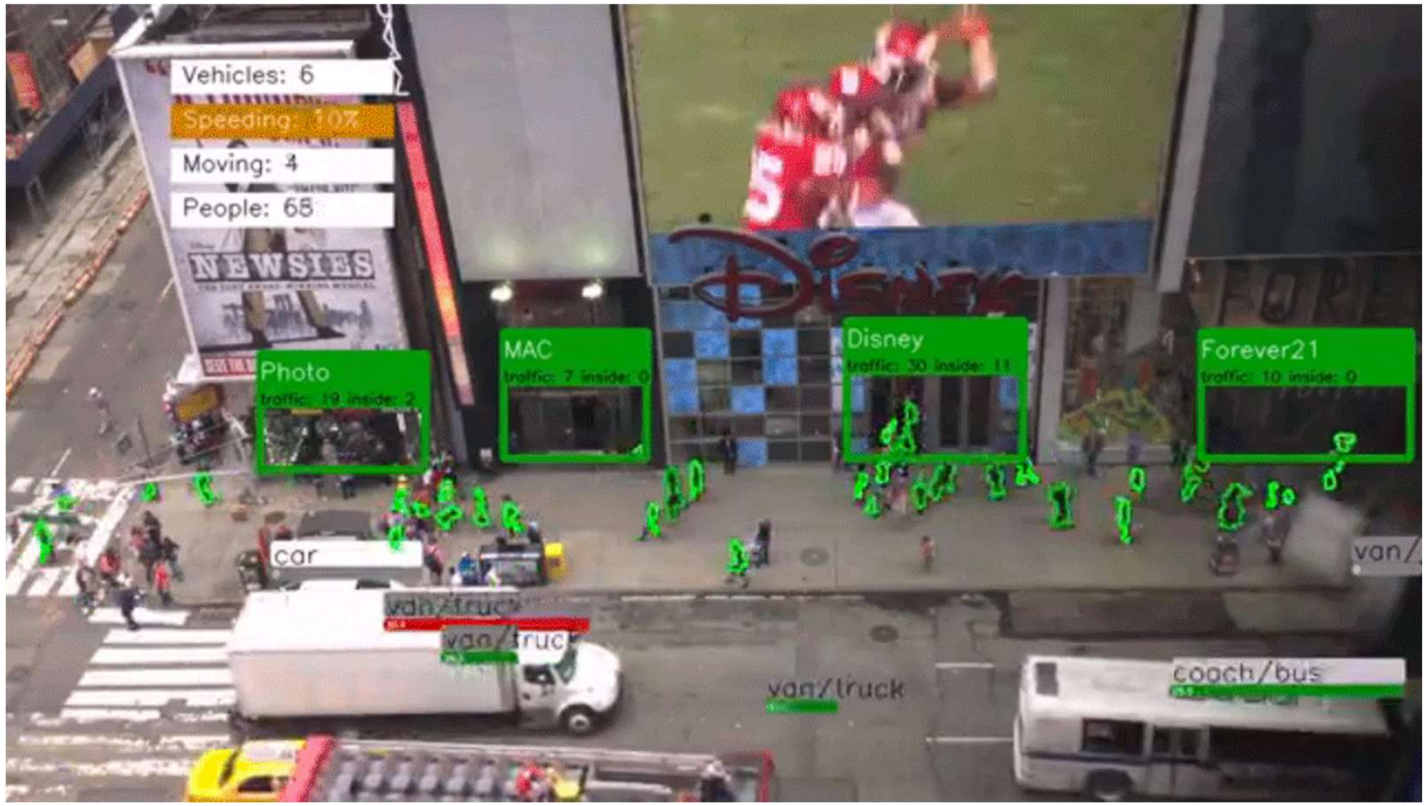
"Government agencies can increase their situational awareness, detect potential problems and respond quickly to incidents, such as unruly crowds, train breakdowns or traffic congestion… The trial aims to assess how we can make use of existing public infrastructure such as lamp posts to fit, connect and power various kinds of sensors that can help us improve the living environment"

From: https://www.tech.gov.sg/media/technews/infographics-just-how-smart-can-lamp-posts-get
See also: https://www.straitstimes.com/singapore/smart-street-lamps-with-high-tech-sensors-set-for-trial

Human Surveillance (via camera)

- Issues?
  - Needs a human!
  - Tired
  - Mistakes
  - MIA
  - Habituated (selective attention)

AI Surveillance



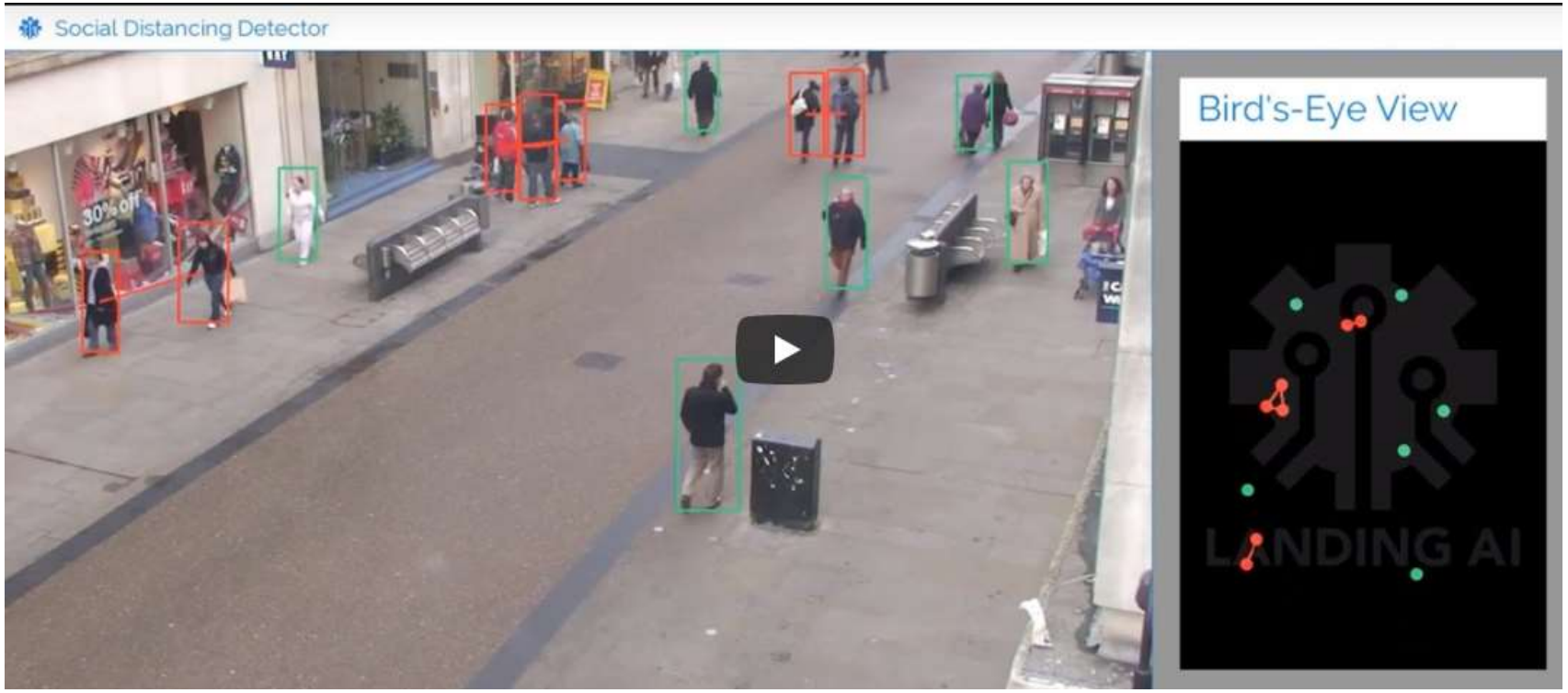
CNBC MARKETS BUSINESS INVESTING TECH POLITICS CNBC TV

WORLD NEWS

One billion surveillance cameras will be watching around the world in 2021, a new study says

PUBLISHED FRI, DEC 6 2019•1:38 PM EST

Social Distancing Detector

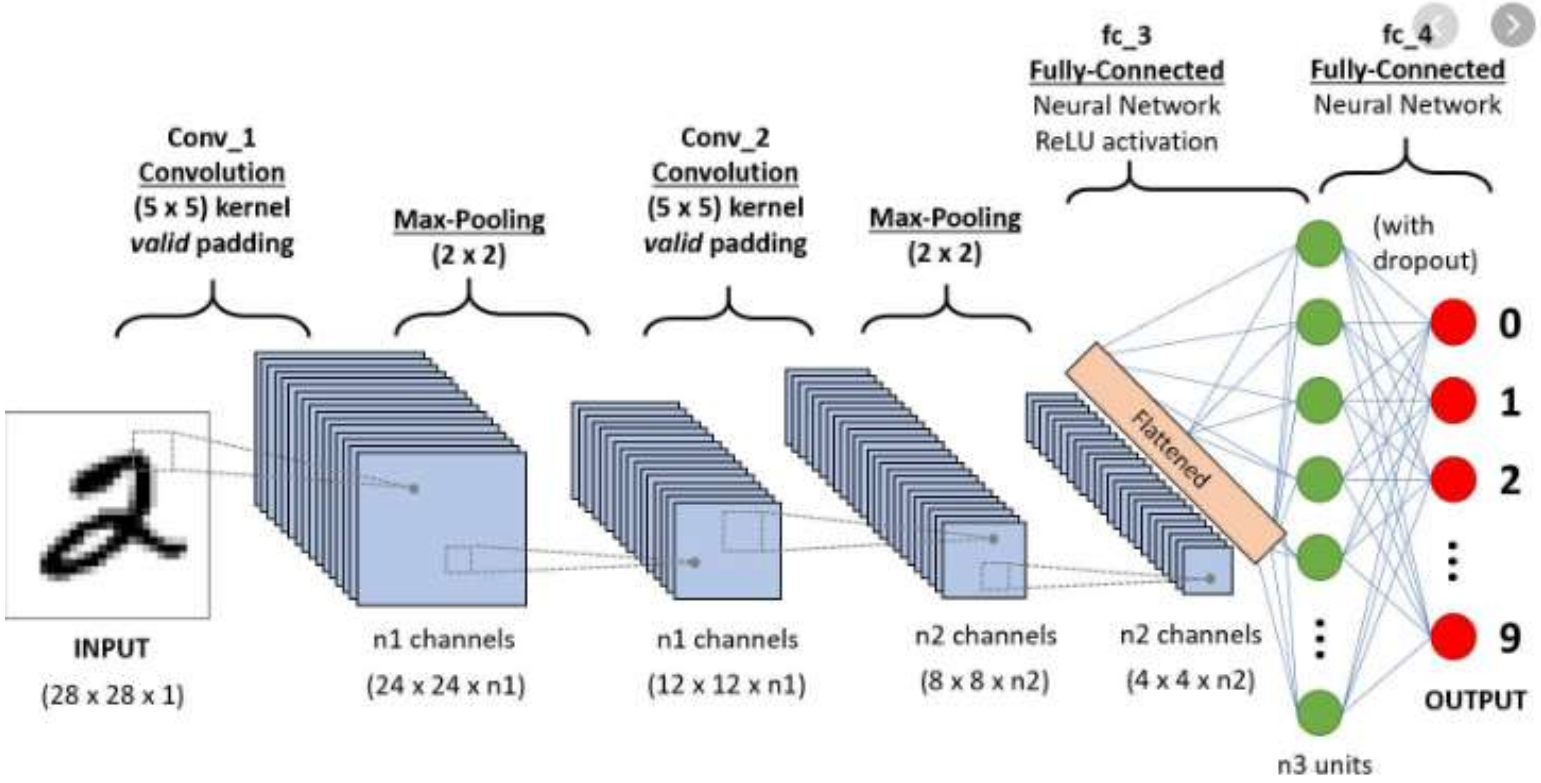


https://youtu.be/15ilV1Lff-M

https://landing.ai/landing-ai-creates-an-ai-tool-to-help-customers-monitor-social-distancing-in-the-workplace/

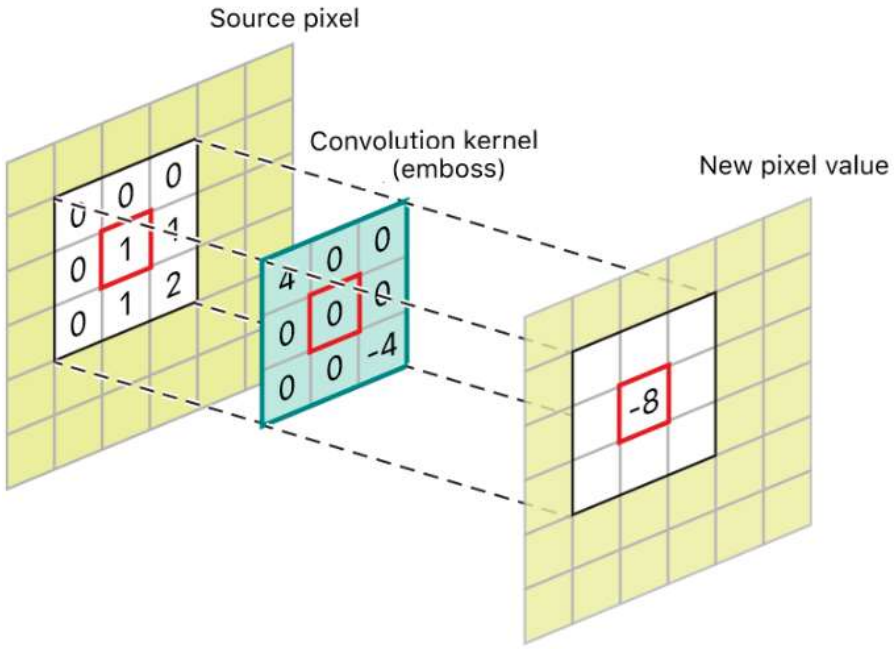
AI Vision – The CNN

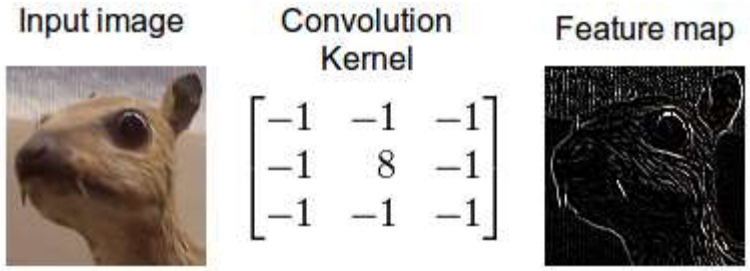
- Convolutional Neural Network



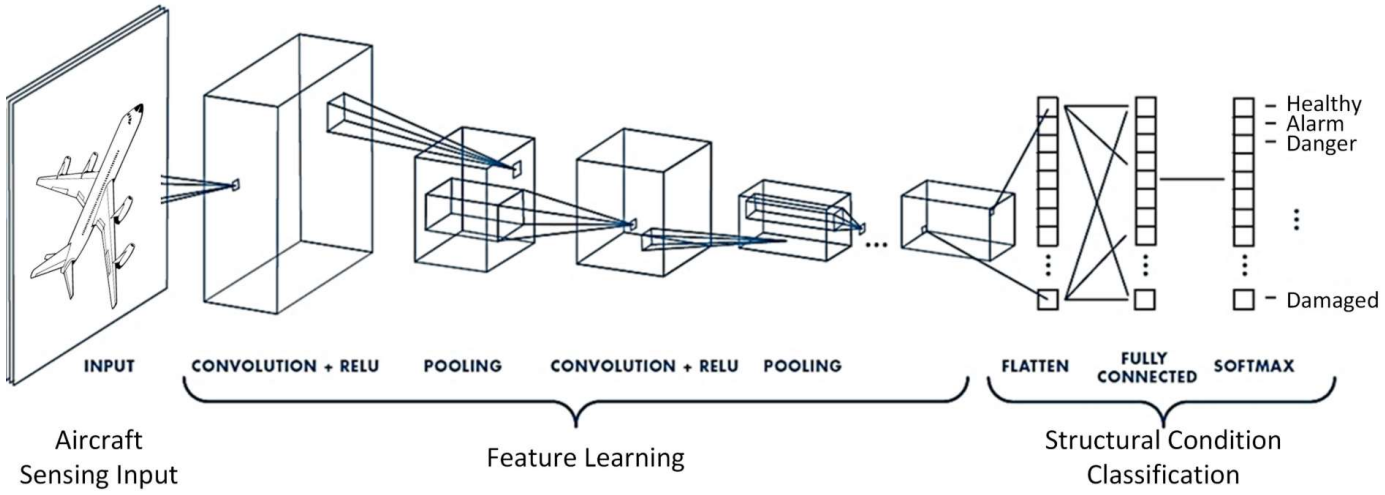
When was the CNN invented? A **convolutional**-like **network** trained by backprop **was** proposed in 1989, the Time-delay **neural network** (TDNN) by Waibel et al. It can be considered a **convolutional network** without pooling

Image Convolutions





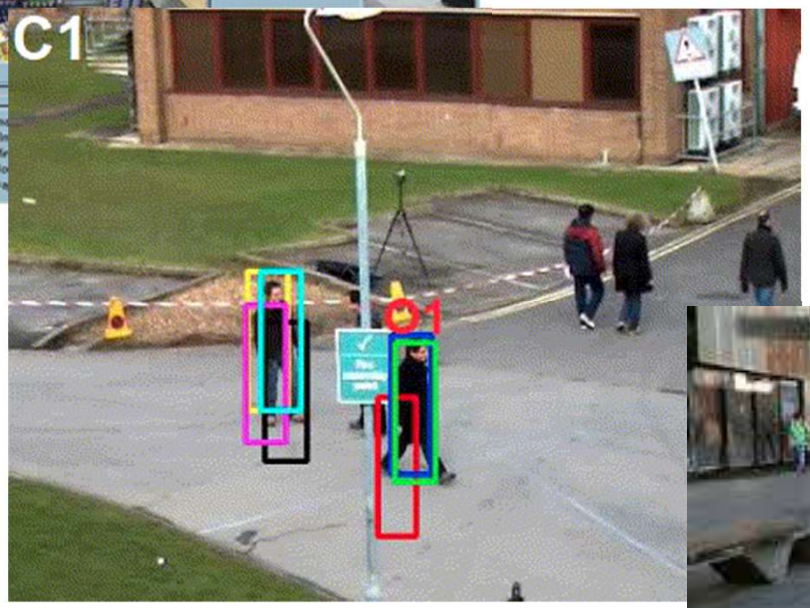
Convolving an image with an edge detector kernel.



Surveillance System

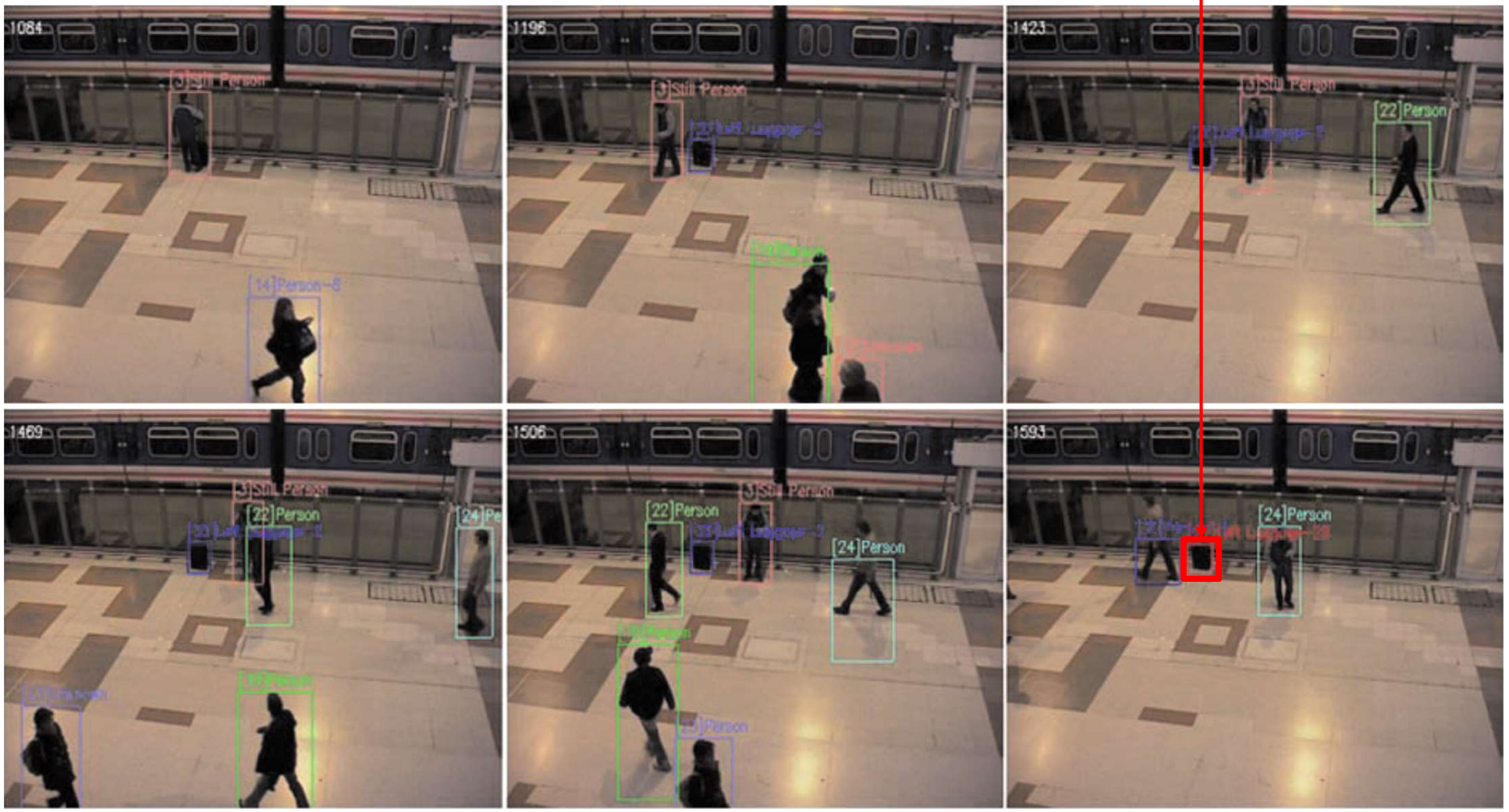
Sensing

Sense Making



Sensing → Sense Making → Decision Making

"Unattended Object"



ARTIFICIAL INTELLIGENCE (AI) SYSTEM TO REDUCE WORKSITE ACCIDENTS

Detection and analysis of potentially hazardous situations from real-time CCTV footage

- Able to automatically identify safety lapses and detect unsafe behaviours and conditions at HDB construction sites in real time.
- When lapses are detected, alerts will be immediately sent to the site safety supervisor's mobile phone via the Telegram app (right).

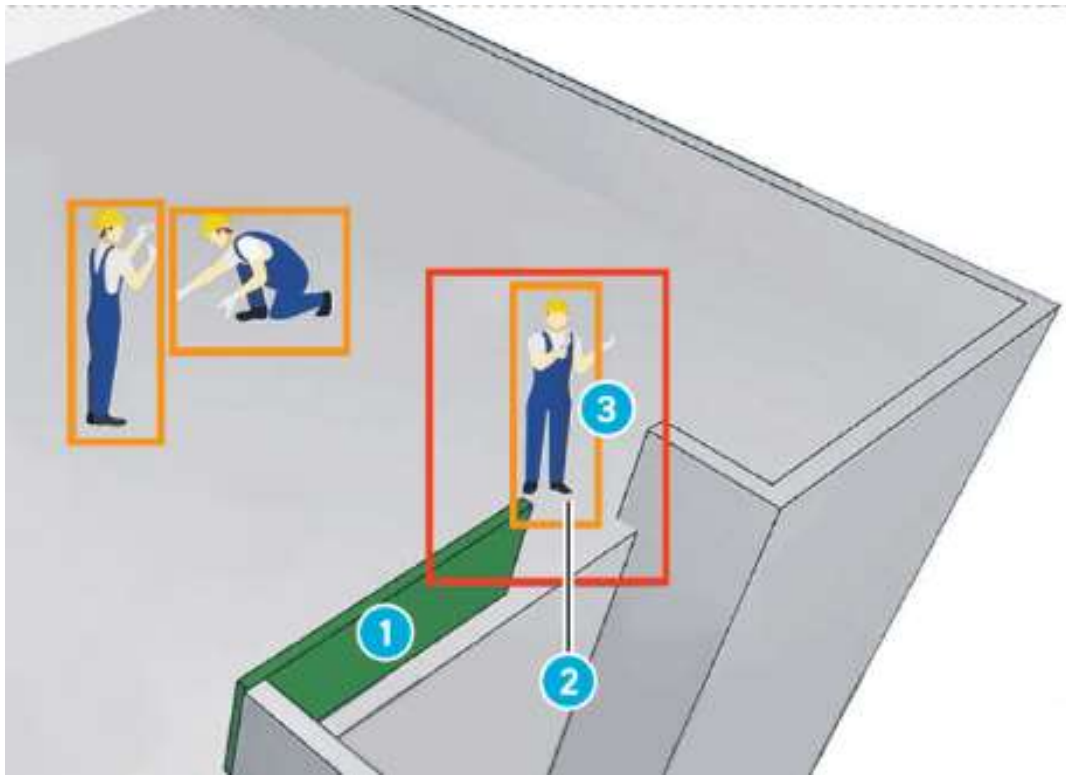
High-risk situation: Struck by falling object

1. **DETECTED LIFTED LOAD**
   Load lifted by tower cranes are automatically detected.

2. **DANGER ZONE**
   The area directly under the lifted load (i.e. fall path) will be highlighted as a high-risk zone.

3. **WORKER DETECTED UNDER LOAD**
   The AI system will immediately send an alert to the supervisor and he will tell the worker to leave the area.

https://www.straitstimes.com/singapore/housing/hdb-to-try-out-automated-alert-system-for-worksite-safety



High-risk situation: **Fall from height**

1. **BARRICADE AT BUILDING EDGES**
Open sides of a building will be barricaded to prevent workers from falling over.

2. **BARRICADE REMOVAL DETECTED**
Barricades at open edges are removed if work needs to be carried out in that area. Removal of barricades is automatically detected by the AI system, which identifies the area as a high-risk zone.

3. **WORKER AT OPEN EDGES**
Any worker who comes within 1m of a non-barricaded building edge will be highlighted so that safety supervisors can ensure that he has observed the proper safety procedures before commencing work close to the edge.

DRONES FOR BUILDING FACADE INSPECTION

- Future drone inspection system taps a cloud software platform to conduct a visual scan of building facades captured during inspection.
- Leveraging AI, it processes photos and identifies building defects in the cloud. It can detect and categorise the types of defects, and tag them to a visual of the building, including the exact location of the defect.
- The platform would then deliver a report to highlight the severity of the defects detected and recommend possible remedies.

Building Inspection

Straits Times: July 9th 2018

Inspector Drone can check out building facades

Tech system developed by JTC and H3 Zoom.AI could be making the rounds at industrial parks next year

Derek Wong

Inspecting a building's facade for possible defects can be a tedious and long-drawn process that takes up to several weeks, requires equipment such as ropes and gondolas, and may pose a safety risk to inspectors. No longer.

Welcome to the world of drones, which can be deployed to check building facades instead.

The technology was developed by statutory board JTC Corporation and H3 Zoom.AI, a digital services platform under parent company H3 Dynamics, following JTC's 2016 open innovation call for sustainable building solutions.

The H3 Zoom.AI Facade Inspector, which was unveiled to the media in a demonstration at one-north last Friday, promises to conduct checks in a faster, safer and

here in recent years, with facade components falling off residential blocks and questions raised over the combustibility of cladding.

These concerns have led the Building and Construction Authority (BCA) to announce regulations for mandatory facade inspections of buildings, which might kick in next year.

Currently, most inspections are carried out by people, and could be costly, risky and tedious. The process of marking out defects, taking photographs and churning out a report could take up to four to six weeks, depending on the building's dimensions.

The H3 Zoom.AI Facade Inspector, which is operated by a two-man team, aims to slash the time taken to a few days using artificial intelligence to mark facade defects, said H3 Zoom.AI chief technical officer and co-founder Shaun Koo,

JTC director of building management Jason Foo (left) and H3 Zoom.AI chief technical officer and co-founder Shaun Koo with the H3 Zoom.AI Facade Inspector. Mr Foo said the ease of using the system will encourage building owners to check facades more frequently. ST PHOTO: JONATHAN CHOO

coming BCA regulations will call for such inspections to be done every seven years in line with the "lifespan" of the facade material.

Still, the use of drones will not completely do away with the need for human input. For instance, a professional engineer has to endorse the findings and the defects have to be rectified by humans.

In addition, the current system can pick up only defects such as wall cracks, peeling paint and water marks – and only on concrete surfaces. It is not programmed to work on other surfaces such as glass and aluminium or detect rust and corrosion.

Most inspections are carried out by people, and could be costly, risky and tedious. The process of marking out defects, taking photographs and churning out a report could take up to four to six weeks, depending on the building's dimensions.

Mr Koo said the H3 Zoom.AI team is working hard to develop more features by the end of this year. With ***machine learning***, the more buildings the drones scan, the greater the data base of defects it can detect.

https://www.youtube.com/watch?v=gknz2zdiNCg&feature=youtu.be

Coronavirus pandemic

Singapore's first drone delivery service takes flight

A drone delivery trial flight by local start-up F-drones last December. F-drones is the first company to receive authorisation from the Civil Aviation Authority of Singapore for such deliveries. The service's first delivery here, of a parcel containing 2kg of vitamins onto a ship, took place on April 19. PHOTO: F-DRONES

PUBLISHED APR 29, 2020, 5:00 AM SGT

FRIDAY, JUNE 12, 2020 | THE STRAITS TIMES |

An Airbus Skyways drone at the Singapore Maritime Drone Estate. In a joint statement yesterday, the Infocomm Media Development Authority, the Maritime and Port Authority of Singapore and M1 announced the partnership with aviation giant Airbus for the one-year flight trial, which will begin in August. PHOTO: AIRBUS

5G trial to test drones for improved port surveillance

Cleaning Robots

Shape-shifting cleaner doesn't cut corners

Cleaning floors ought to be less of a chore now, with the many automated, self-roaming vacuum cleaning robots in the market.

The problem is that they are not that smart, blindly going over surfaces that have already been cov-ered, wasting time and electricity.

Enter hTetro, a shape-shifting, floor-cleaning robot with built-in artificial intelligence and machine-learning capabilities that let it map out, learn and remember its cleaning environment.

Making its debut at the CleanEnviro Summit Singapore, the robot is the brainchild of Dr Mohan Rajesh Elara and his research team at the Singapore University of Technology and Design (SUTD).

In experiments with a prototype conducted in four-room Housing Board flats, hTetro covered 30 per cent more floor area than current commercial models, Dr Mohan told The Straits Times.

The prototype, built from SUTD's 3D printing facilities with different plastics, took the team 10 months to complete, said the 36-year-old assistant professor in engineering product development.

Its design was inspired by popular tile-matching video game Tetris.

The robotic floor cleaner is made up of individual tiles connected to one another at the corners.

This design allows the robot to transform itself into different shapes, with the individual tiles moving around just as Tetris blocks do in the game.

This shape-shifting capability lets it clean room corners, as well as narrow and complex furniture spaces that are almost impossible for existing cleaning robots to reach.

sTetro, the staircase-cleaning counterpart of hTetro, can even climb steps.

ST

SCAN TO WATCH
See the hTetro in action.
http://str.sg/htetro

Dr Mohan Rajesh Elara with the hTetro, whose built-in artificial intelligence and machine-learning capabilities let it map out, learn and remember its cleaning environment. Its design was inspired by video game Tetris. ST PHOTO: KHALID BABA

While navigating tight corners, hTetro turns into the shape of the letter I, while its default mode is the shape of an O, Dr Mohan said.

Its wheels are also fitted with wheel encoders, which can record the number of rotations the wheel has made.

This helps the robot know how much distance it has covered.

A laser sensor on top of the robot helps it to recognise and learn its cleaning environment by emitting laser light and capturing the reflection, in a method called Light Detection and Ranging (Lidar).

Current commercial robots are programmed to clean everywhere, but hTetro is programmed to avoid solids and liquids, said Dr Mohan.

Current cleaners will run over spilt ice cream in an attempt to wipe it off, only to spread it across the floor, for instance.

Apart from Lidar, the robot has an inertia measurement unit, which lets it assess its speed, direction and tilt.

Being able to assess tilt means the robot can be reconfigured to evaluate, for example, how even floors are.

The technology can be extended to the architecture and engineering fields.

The team is already in talks with some companies to license the component technologies of hTetro, such as object identification and floor-mapping capabilities.

The team is also looking into spinning off the project into a start-up, and hopes to commercialise the product within two years.

The team members are working with cleaning equipment manufacturers and service providers to improve the prototype before launching it in the market, Dr Mohan said.

A residential-scale hTetro is likely to sell for about $600.

hTetro may also be built with capabilities to clean ice cream one day, said Dr Mohan.

"However, it is a trade-off between cost and complexity.

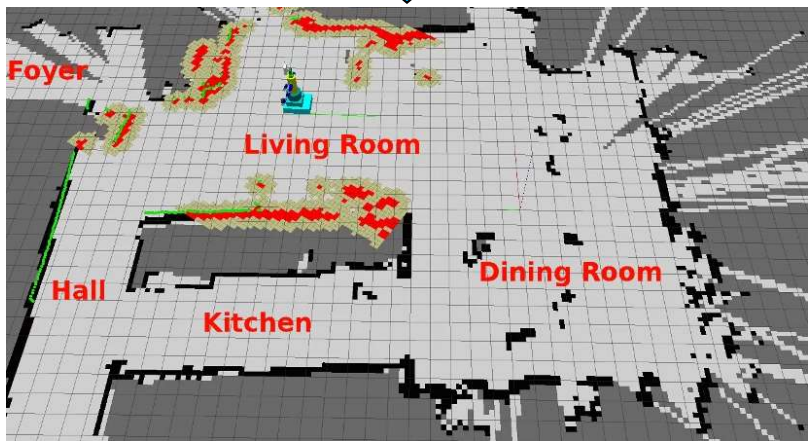
"The more complex the robot, the more expensive it will be," he added.

Straits Times: July 12th 2018

World Cities Summit, Singapore International Water Week, CleanEnviro Summit

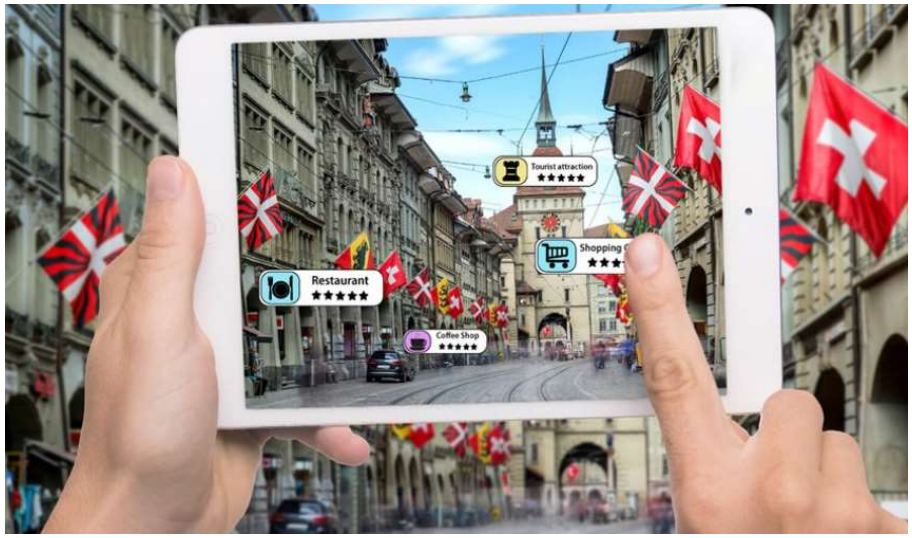
SLAM – Simultaneous Localisation & Mapping

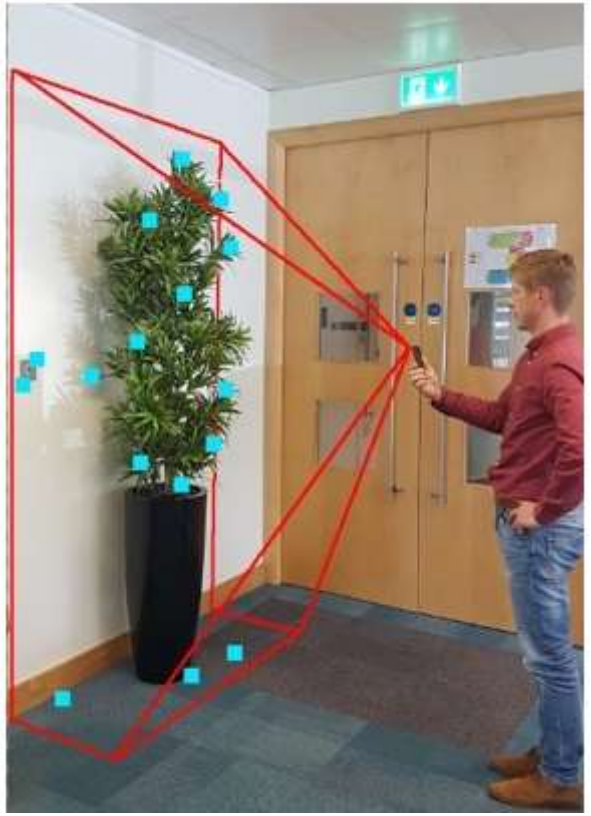
Given a set of sensory observations (O) which happen at discrete time steps (T). Create a map of the environment (M) & find the location (M) of the robot within M



AI underlies Augmented Reality

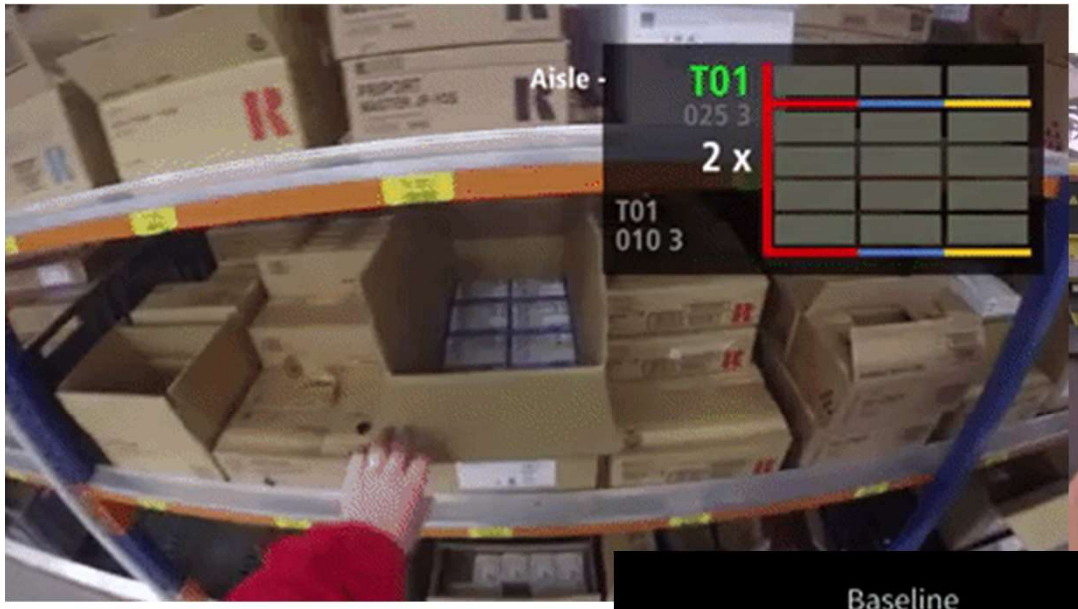
SLAM is not just for robots….





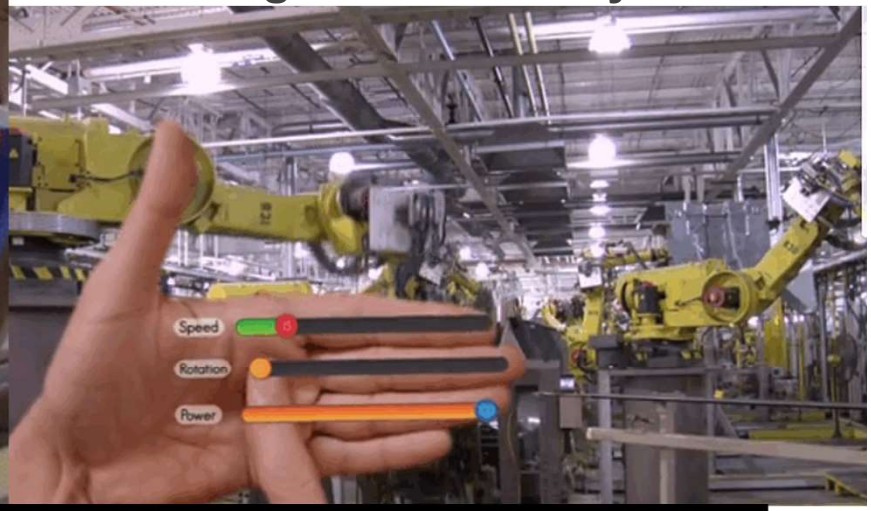
https://community.arm.com/developer/tools-software/graphics/b/blog/posts/introducing-slam-technology

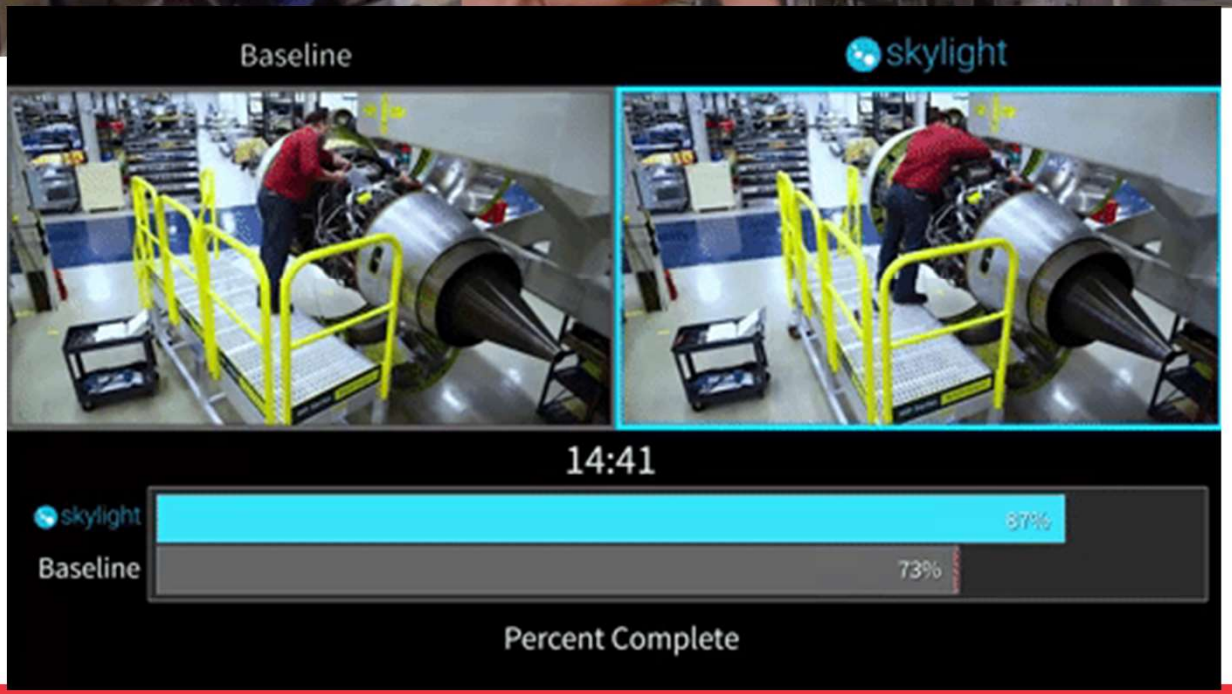
AI underlies Augmented Reality



Warehouse
Stockpicking/
Cargo management

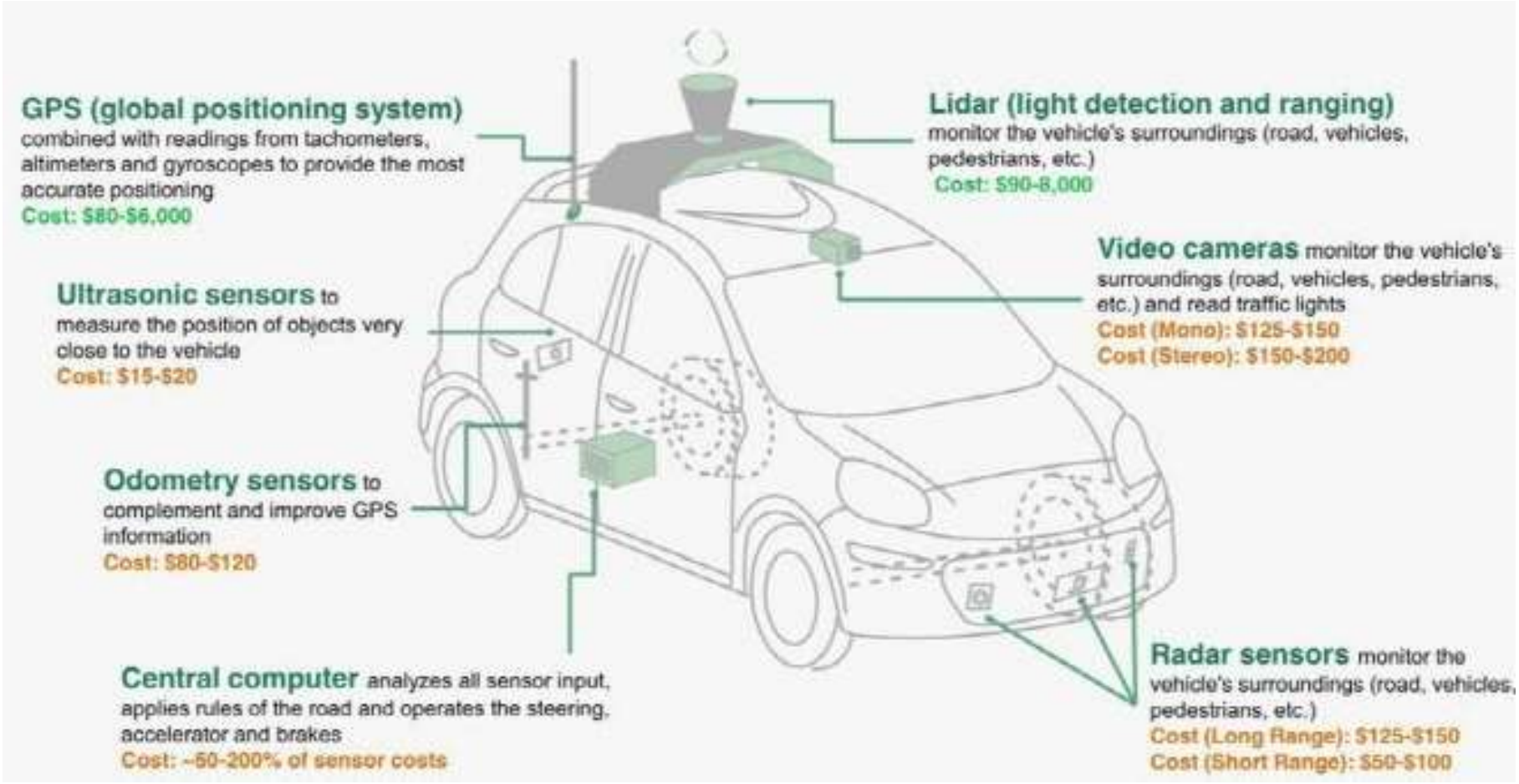
Augmented Reality Controls



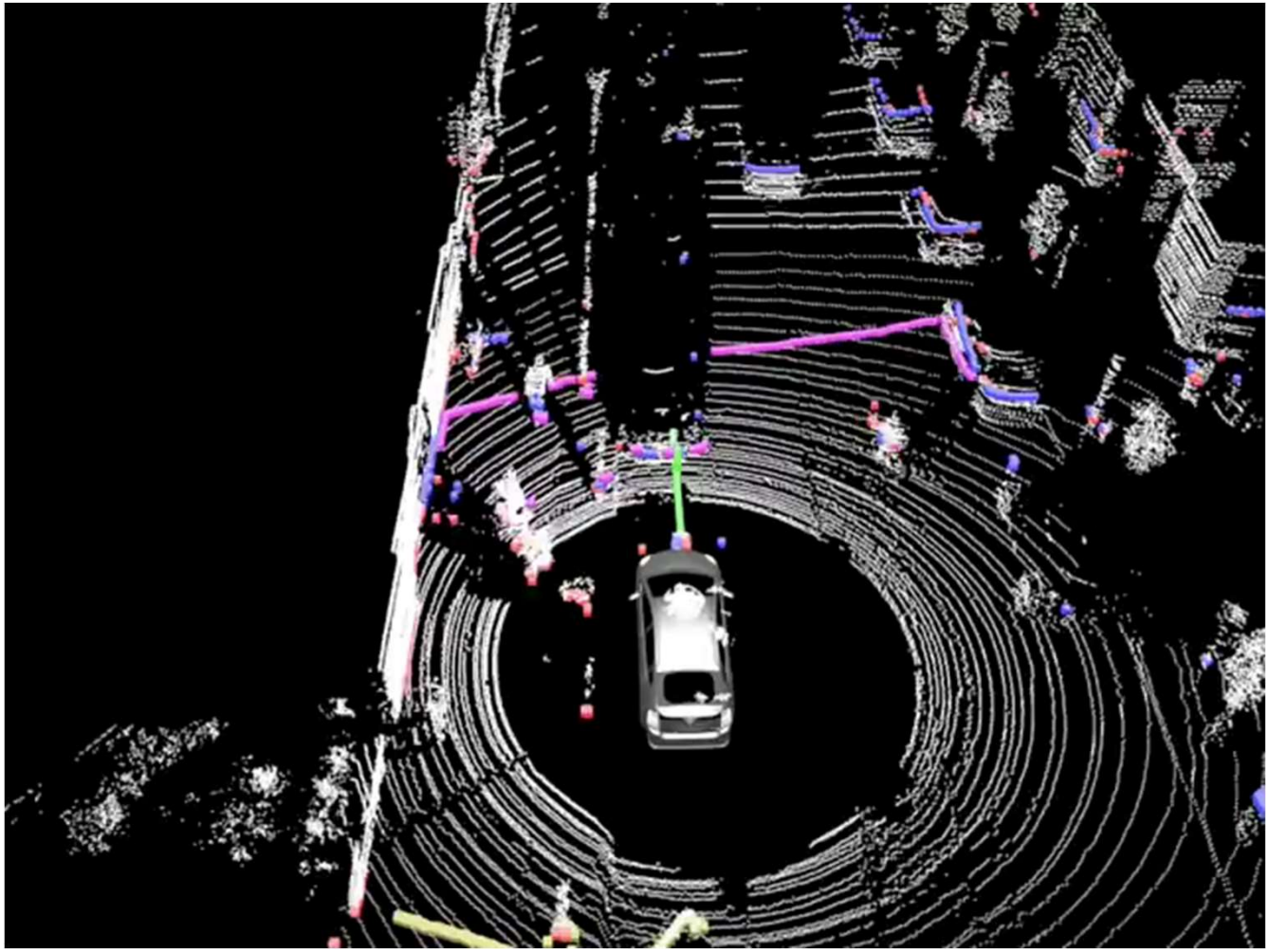


Engine repairs
completed
16% faster with
AR insructions

Sense -> Sense Making -> Decide -> Action



LIDAR – powerful sensing



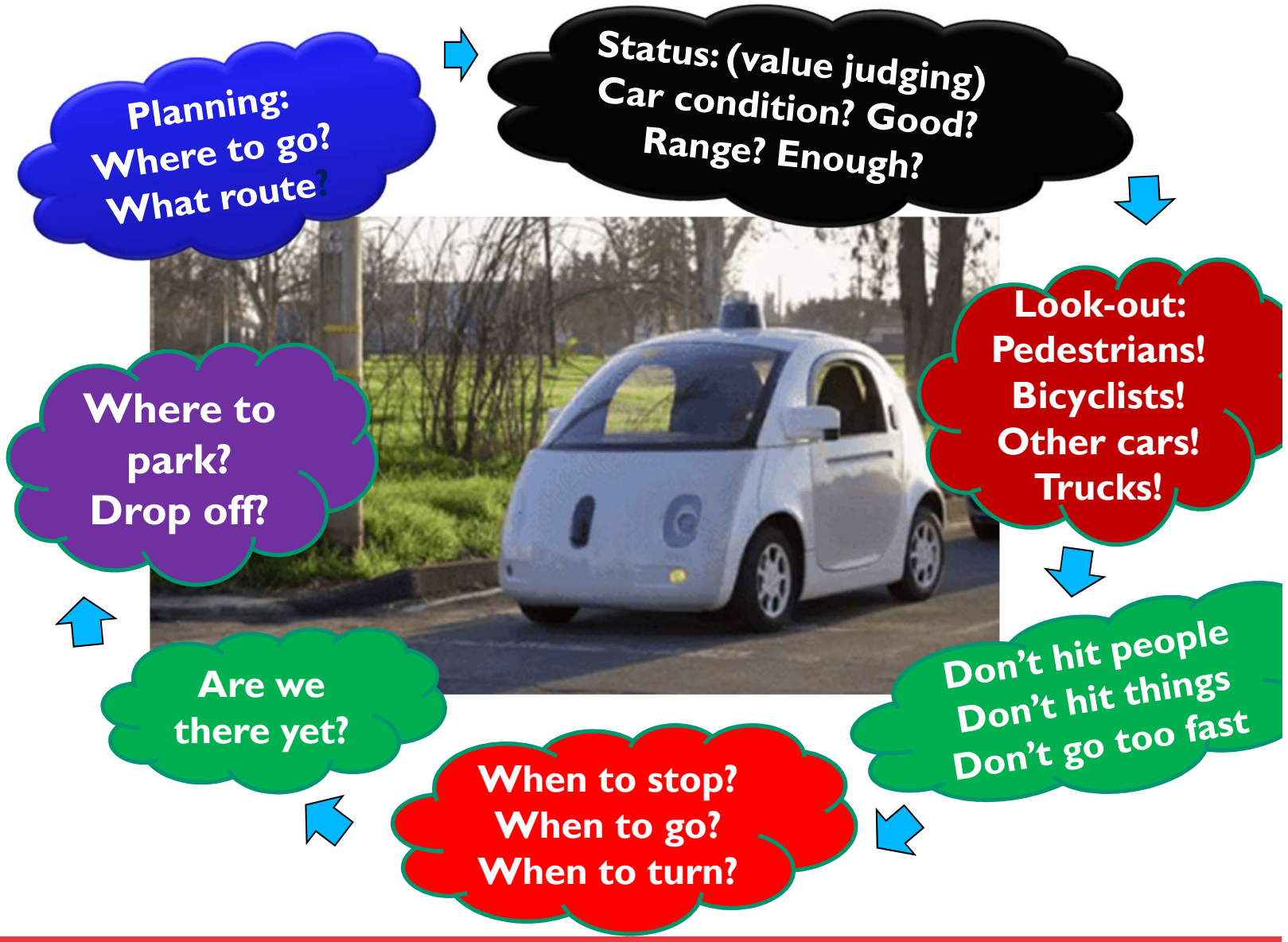
Light Detection and Ranging

Decision Making!

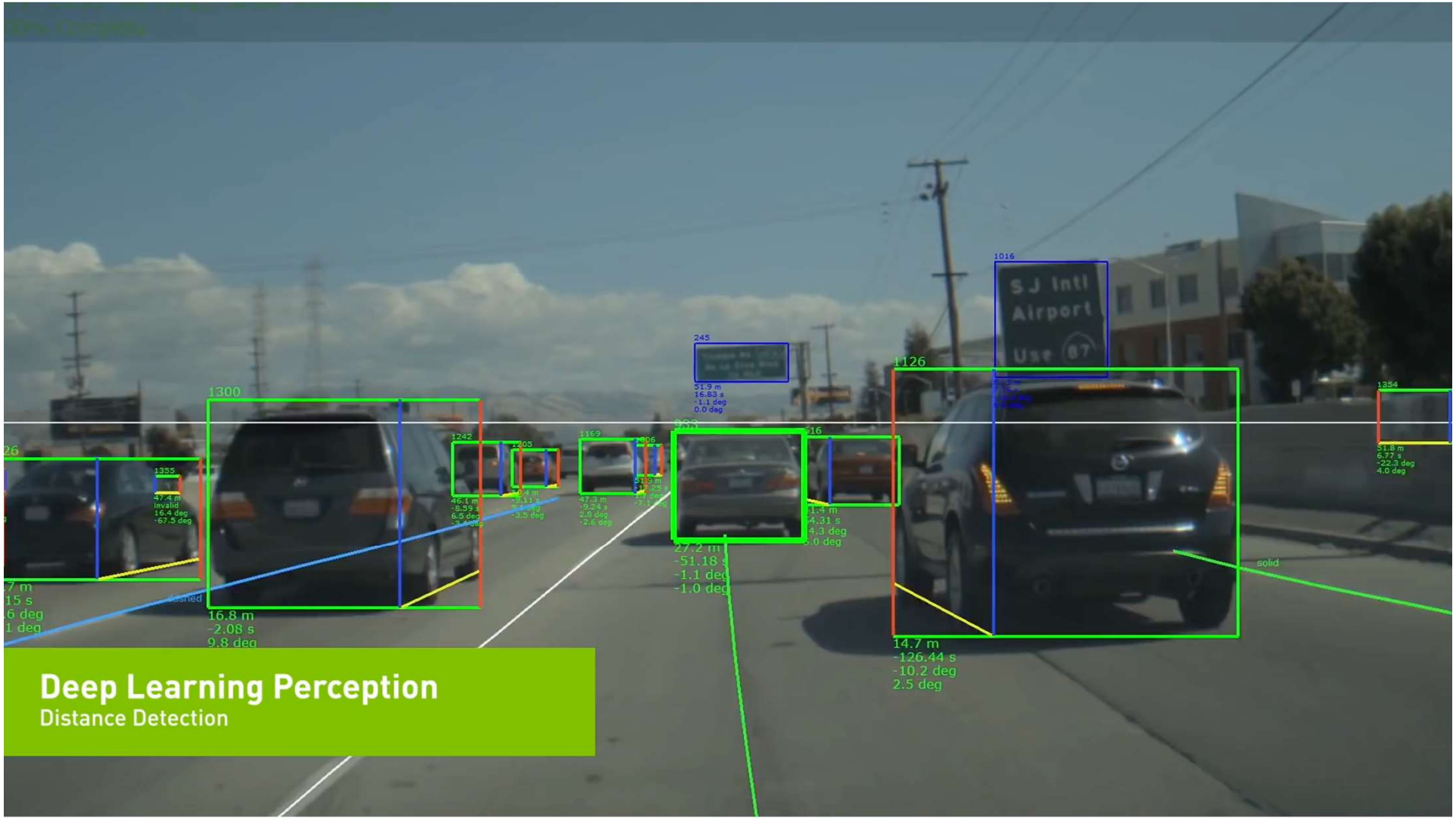
Strategic Planning: e.g. navigation

Tactical Planning: e.g. lane changes

Reactive Planning: e.g. obstacle avoidance

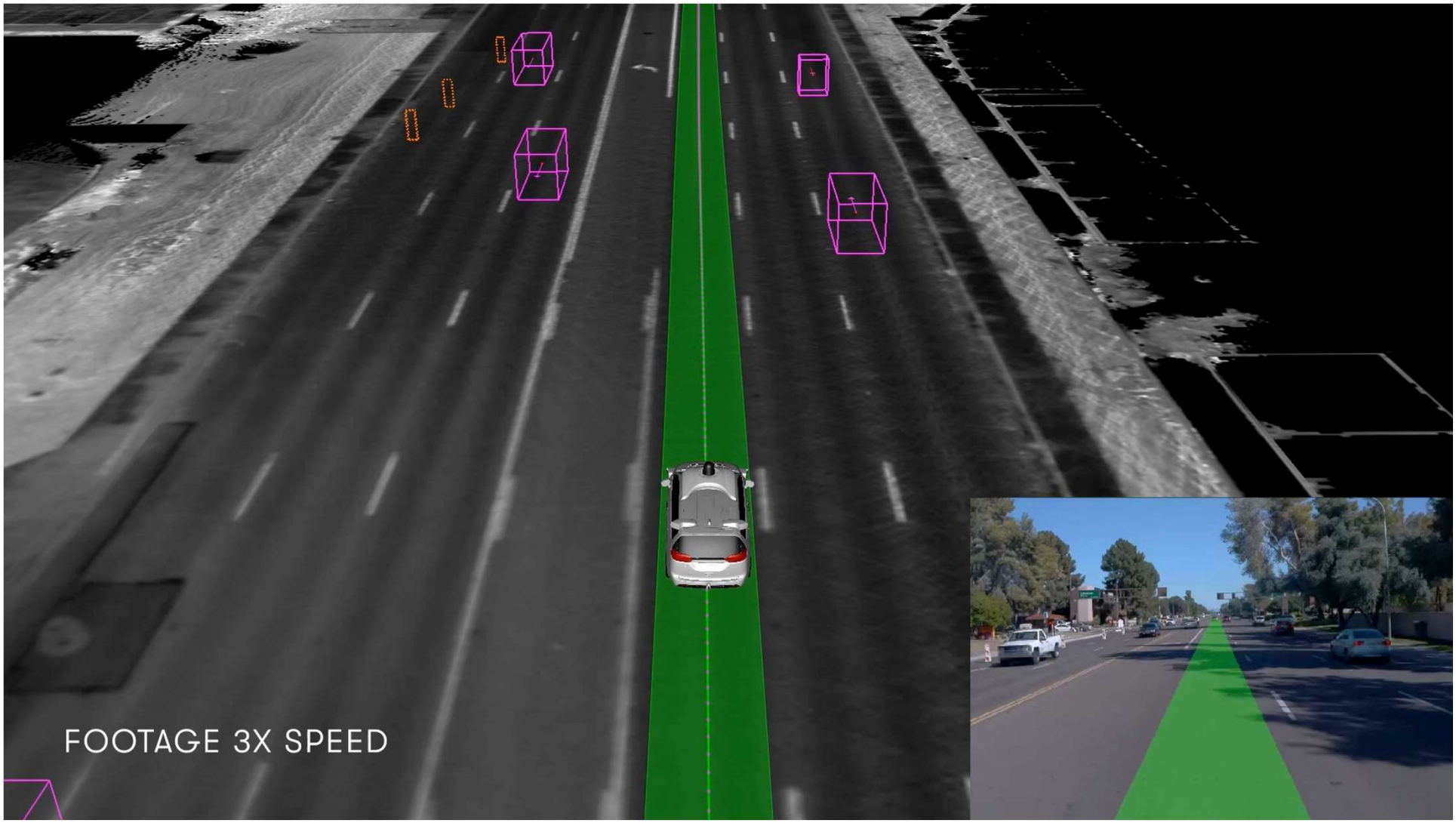


Still learning (driver in car)

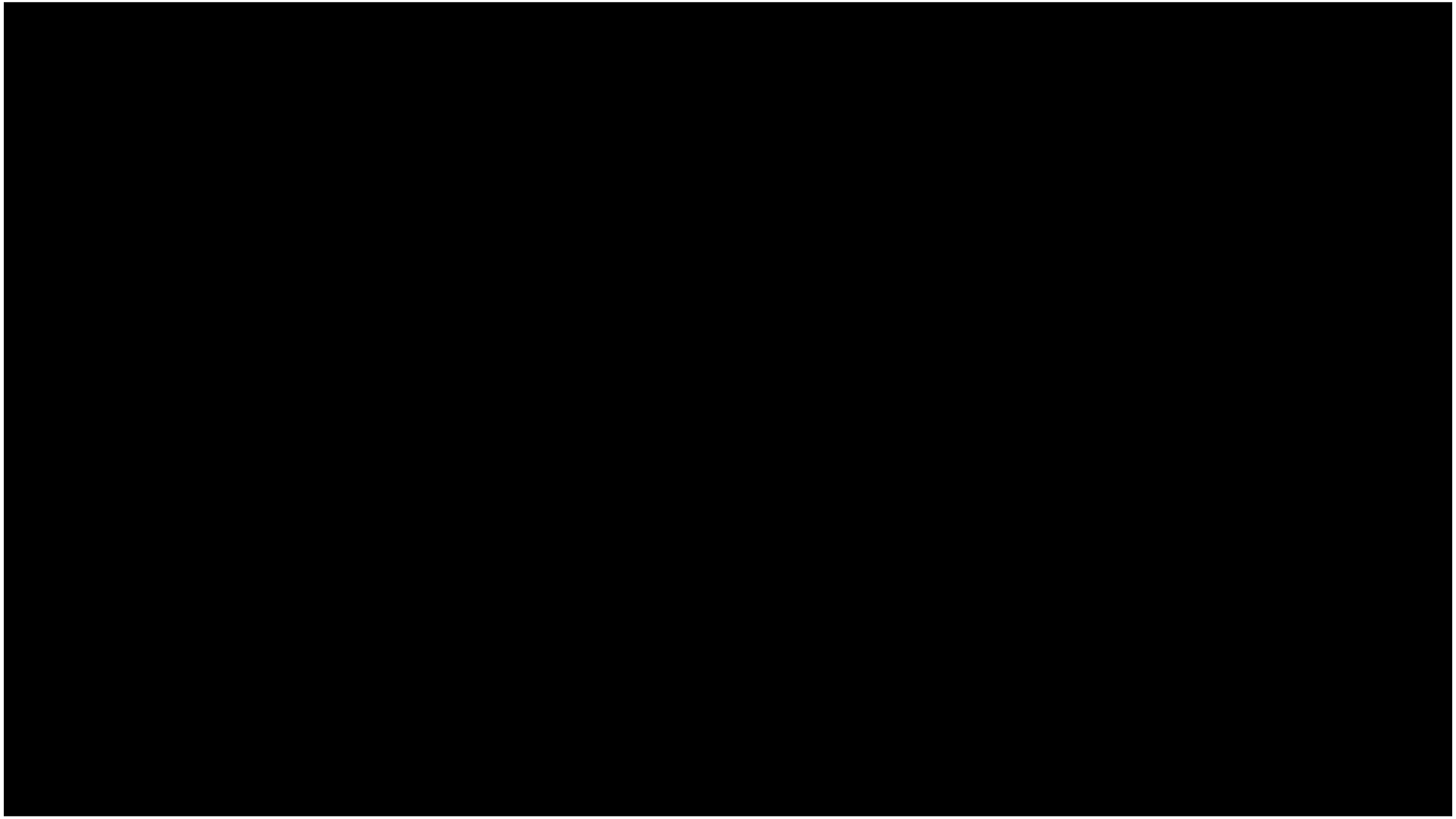


Fully autonomous (Driver in back seat!)

At an intersection…



Letting bicycle overtake…



MTech Student Project Examples…

Bernard the Robot

- Facial recognition and personalized response
- Machine learning for Human-like training
- Gesture mimicking

Intelligent Drone

- Image Recognition / Obstacle avoidance / Path planning

Smart Socks for Parkinson Patients

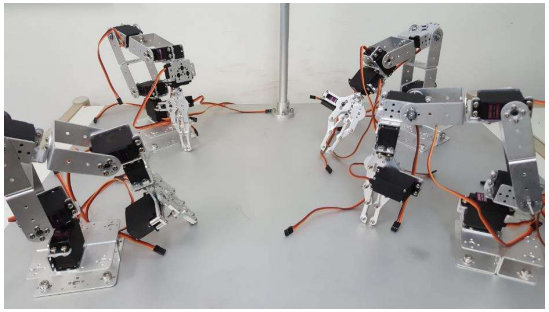
- Intelligent sensing and learning
- Biofeedback through stimuli

EEG Signal Analysis and Classification for Wheelchair Control

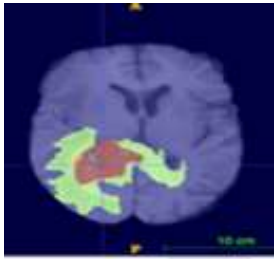
EEG Mind-typing interface for LIS Patients

Multi-Arm Surgical Cobot

- One Surgeon to operate instead of multiple
- Collaboration w/ SingHealth



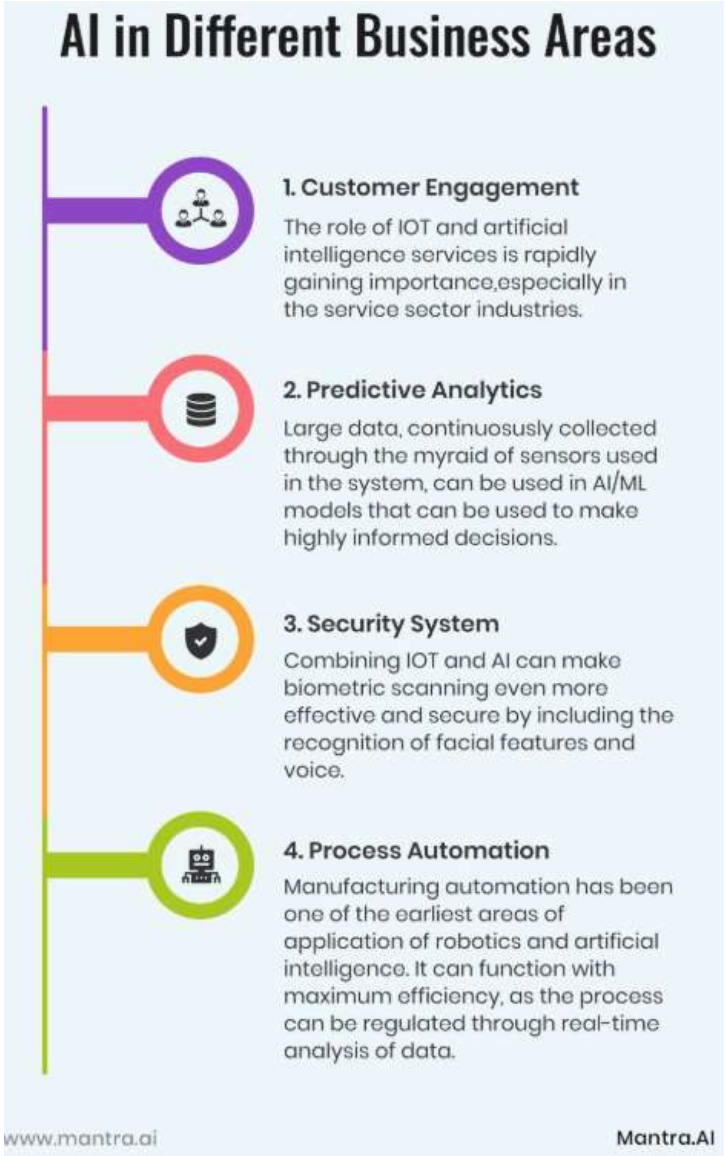
Brain Lesion Classification using Deep Learning

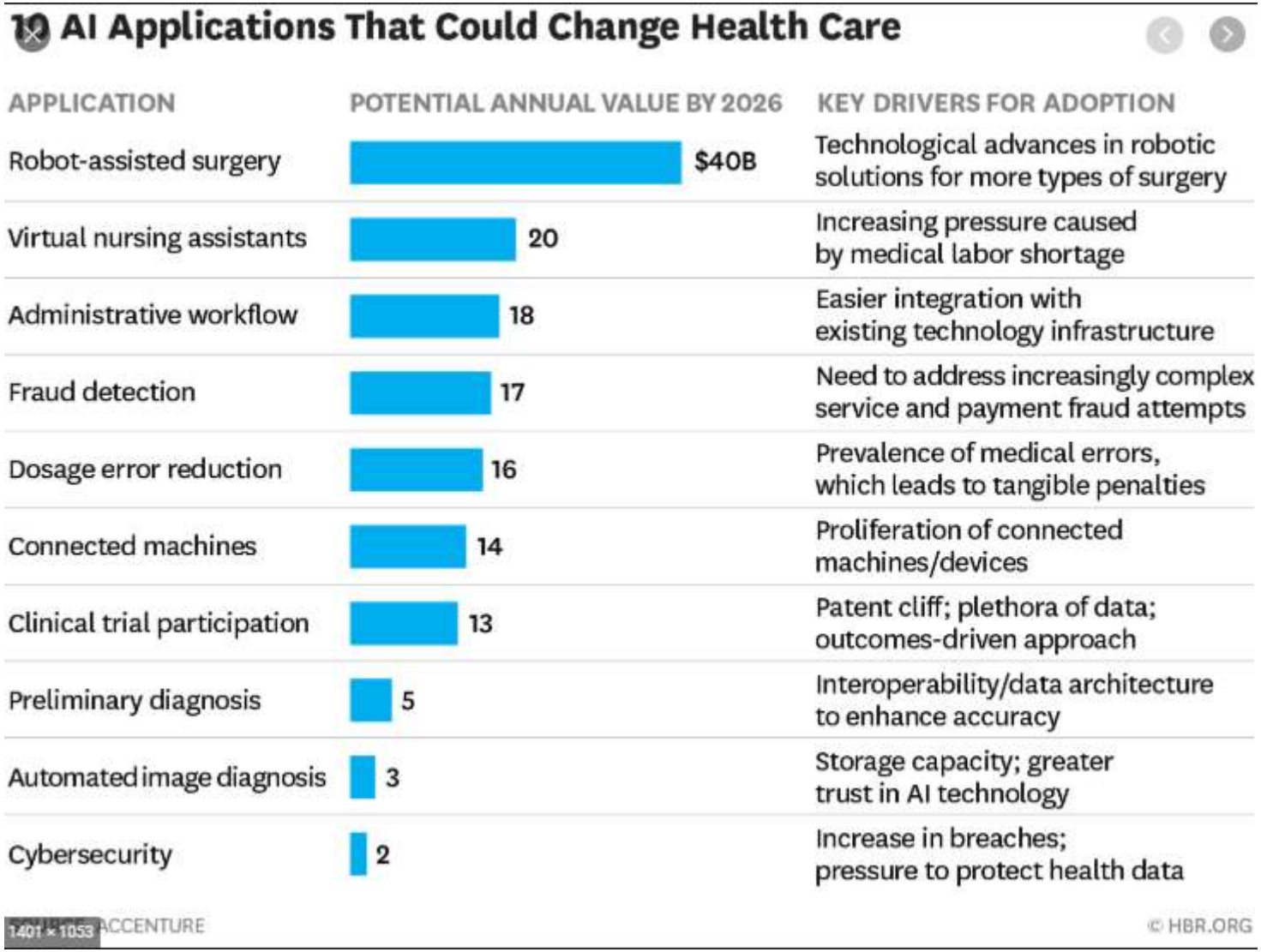


Applications: Business Domain

Driven by Business Data (but also sensor data too)

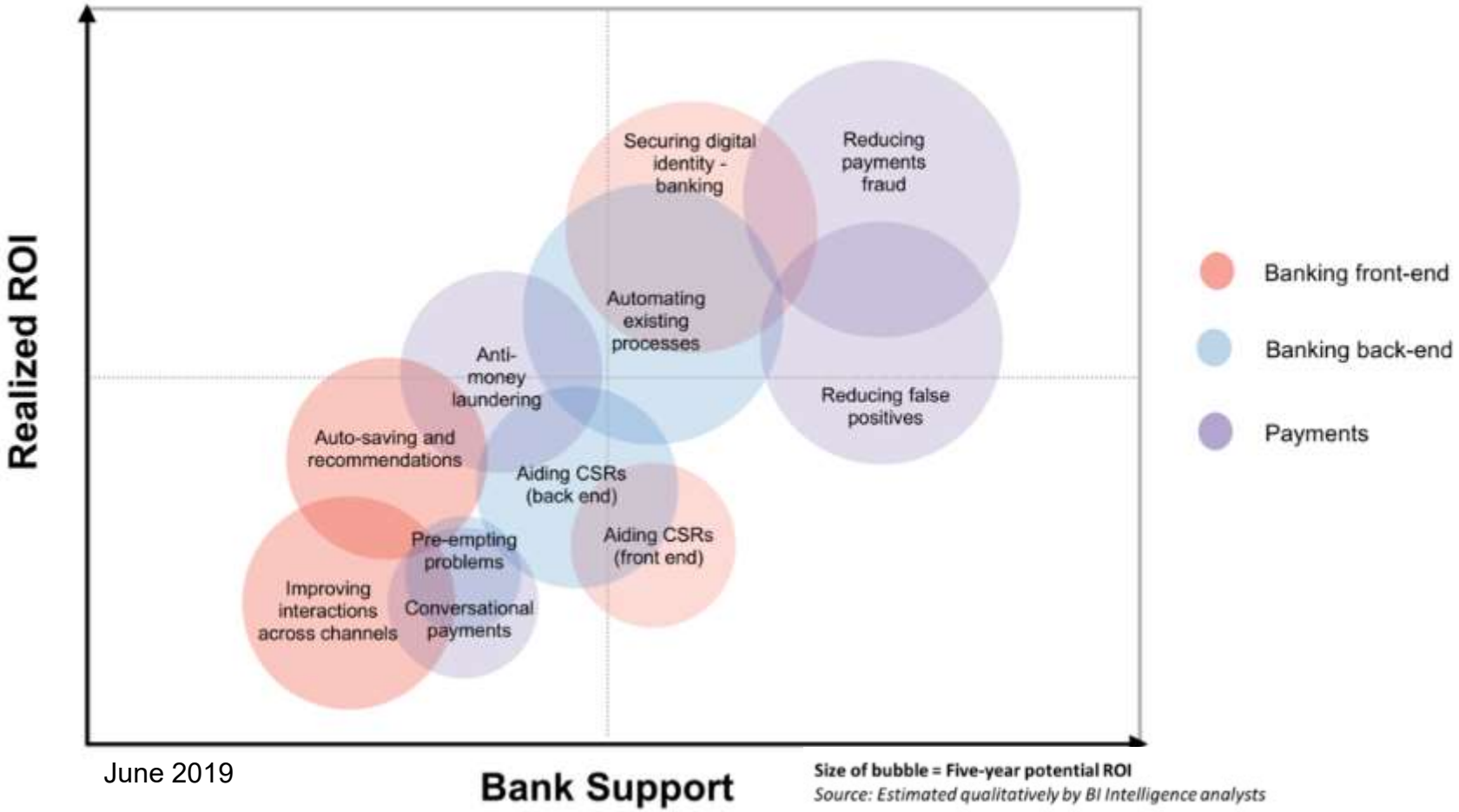
- Customer Analytics
  - Acquire & keep customers, loyalty programs
  - Personalisation - customise products, sell-more, make recommendations
  - Campaign Optimisation - campaign targeting & response models
- Banking & Finance
  - New products/services
  - Better operations - forecasting, algorithmic trading, market analysis, risk assessment, …
  - Fraud detection
- Healthcare
- Human Resource Management
- Transportation / Logistics
- Business Process Automation
- ……





https://hbr.org/2018/05/10-promising-ai-applications-in-health-care

Maturity Of Uses Of Artificial Intelligence In Banking And Payments

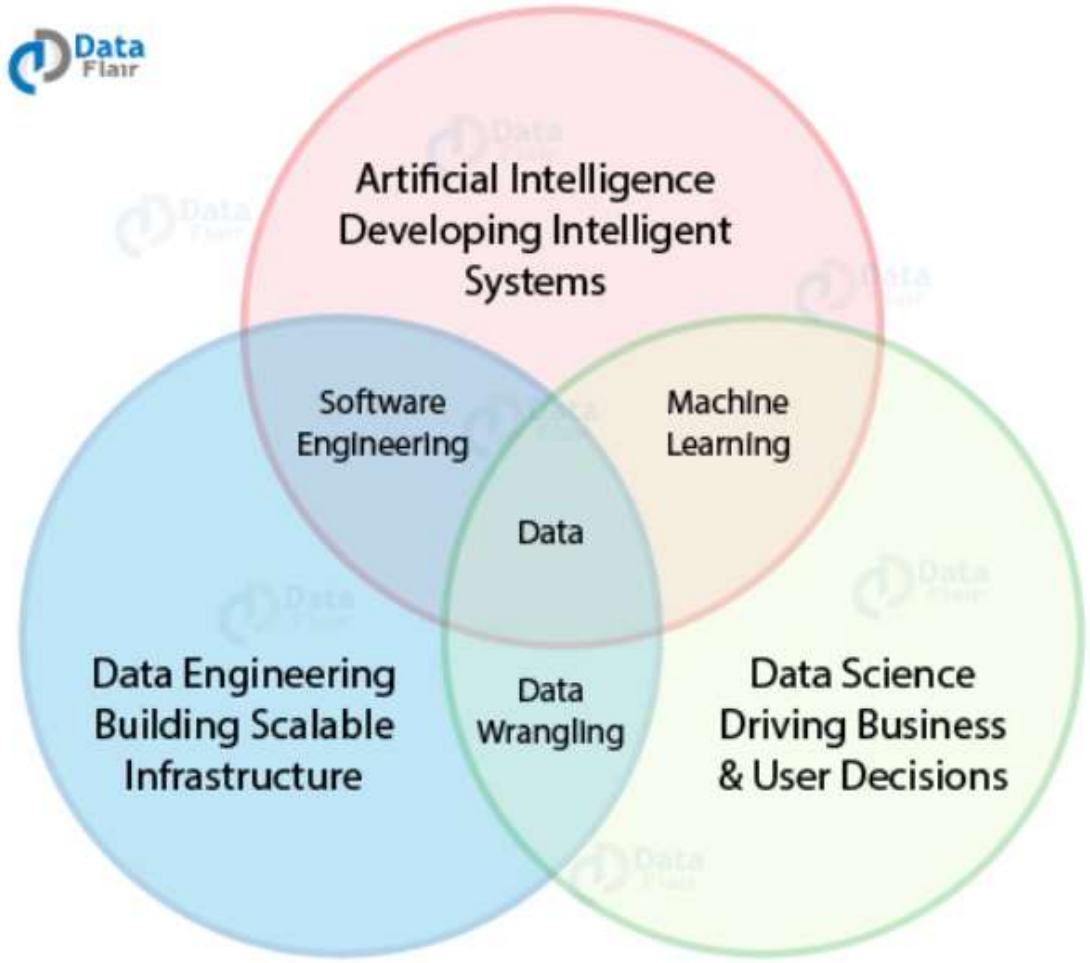


June 2019

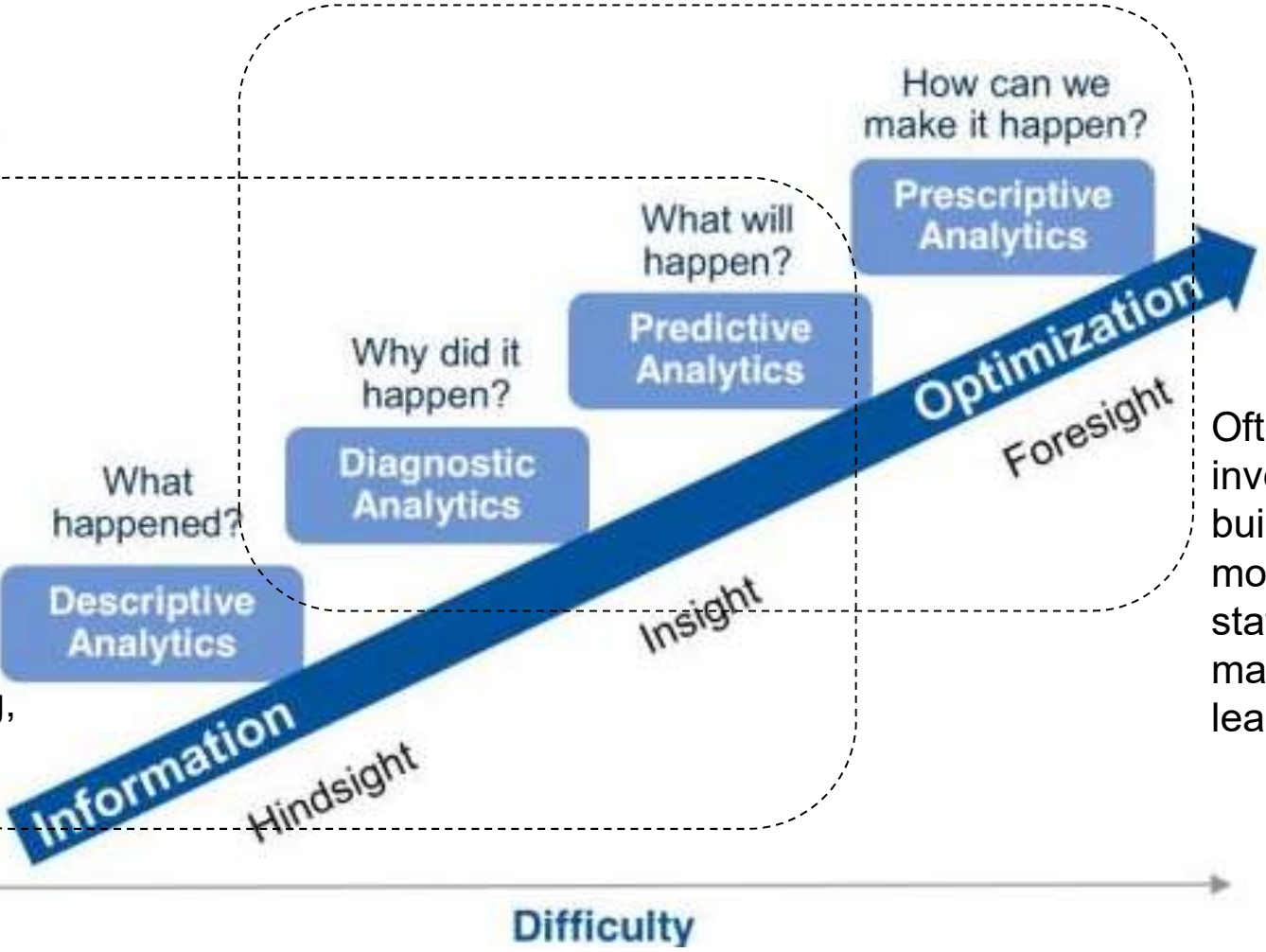
The aggregate potential cost savings for banks from AI applications is estimated at $447 billion by 2023, with the front and middle office accounting for $416 billion of that total,

https://www.businessinsider.com/artificial-intelligence

AI versus Data Science/Analytics



Business Analytics



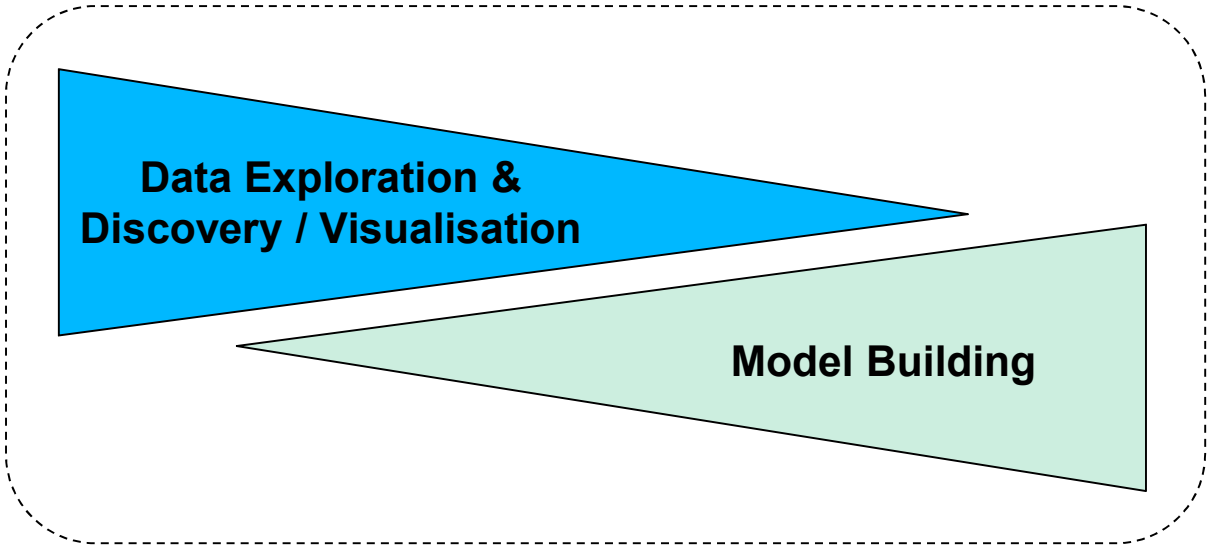
Data Understanding,
Data Visualisation,
KPI montoring,
Dashboarding

Often involves building models using statistics and machine learning

AI versus Data Analytics Applications

Tools, techniques, infrastructure are often similar

Graphics Statistics Data Engineering AI



Data Analytics Applications

(focus on business decisions)

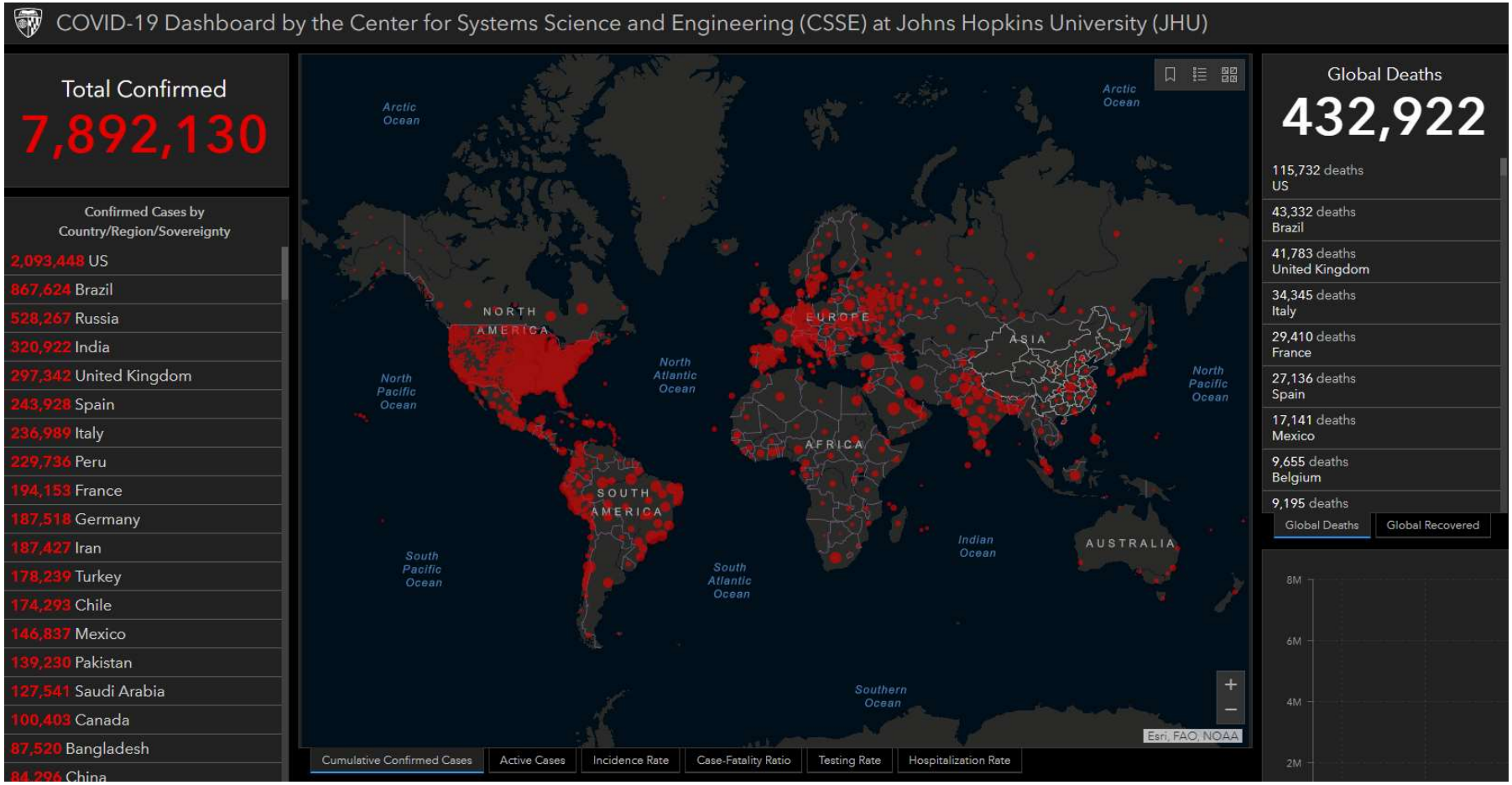
AI Applications

(focus on automation)

Best practices and process are similar

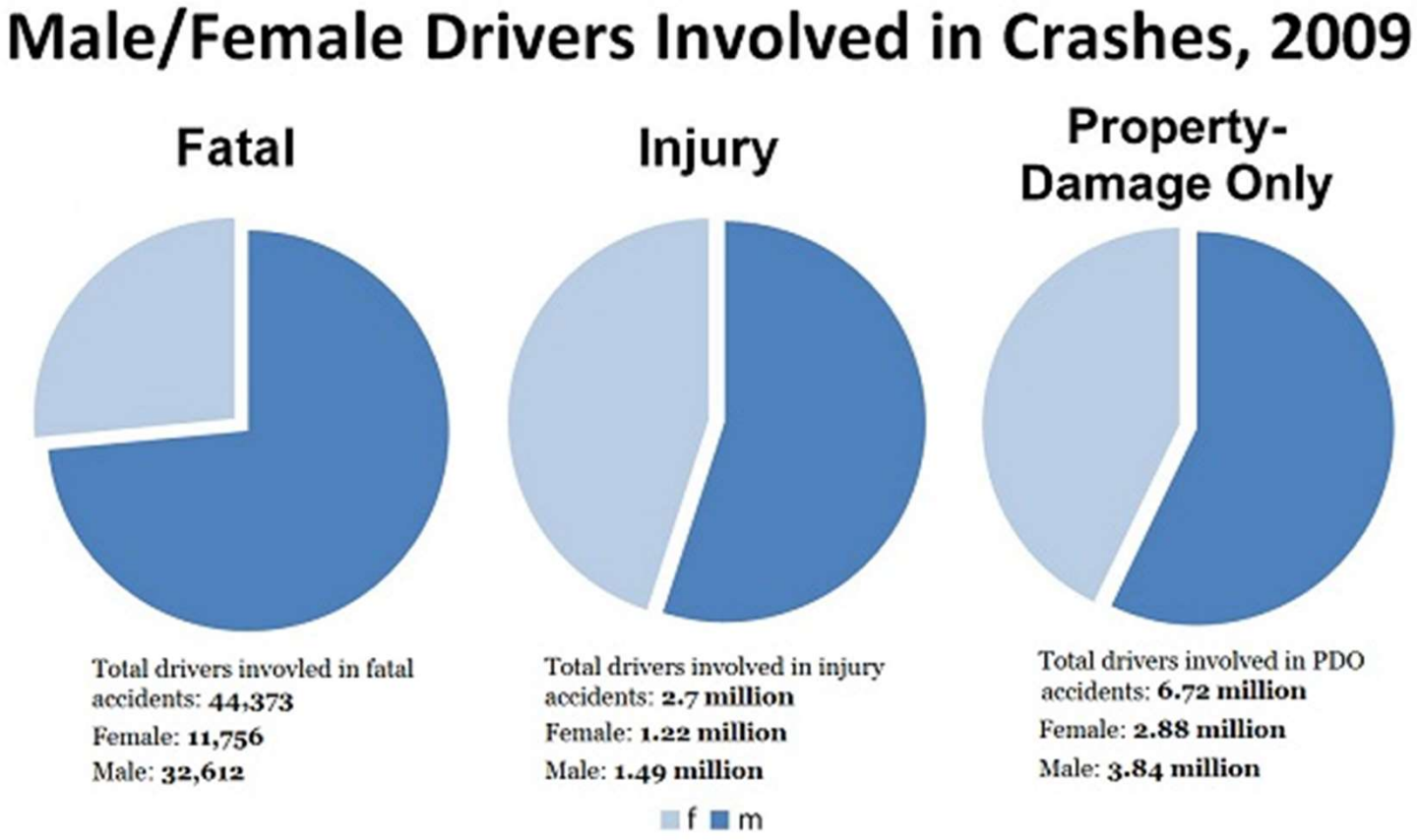
Data Dashboards

- Common tools ~ Tableau, QlikView …



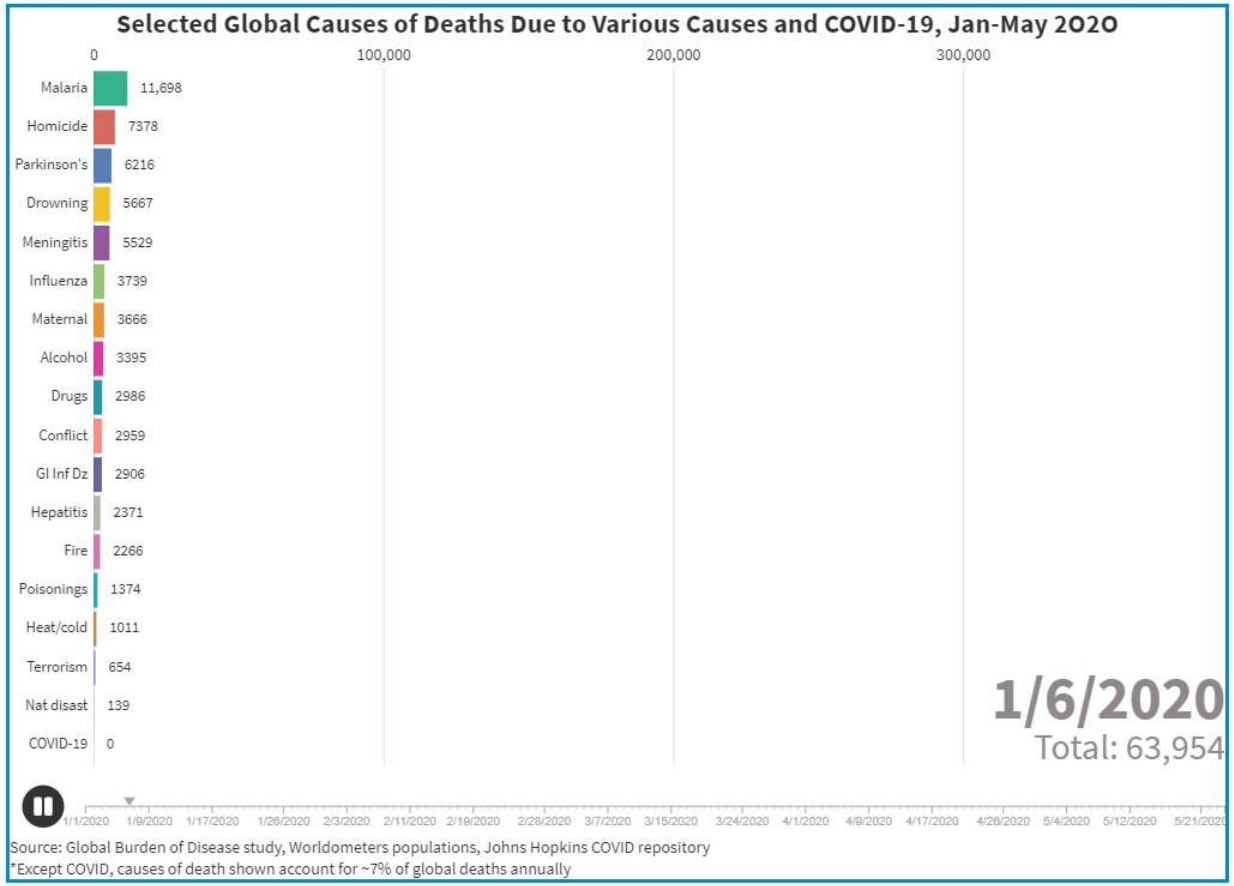
https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

Discovery through Visualisation



SG: Life Expectancy x Income by Time

Global Deaths Due to Various Causes and COVID-19 in 2020



https://public.flourish.studio/visualisation/2637725/

The Mysterious Case of the Braking Train

PROBLEM SIMILAR TO SEPT INCIDENT; CAUSE STILL UNKNOWN, SAY LTA, SMRT

Signal interference issue plagues Circle Line again

Mobile services may be suspended along parts of the line for short periods if similar faults recur

problem that disrupted train services for five days in September returned yesterday during the morning rush hour, leaving six Circle Line (CCL) stations, including major interchanges like Bishan and Serangoon, without train services.

The interference caused a loss of signal between trains and stations, forcing the trains to come to a halt. The Land Transport Authority (LTA) and SMRT yesterday said that preliminary investigations indicate that the nature of the fault in the latest incident could be similar.

With some trains down between stations, crowds began to build at the stations. SMRT staff had to make their way to the stalled trains to drive them to station platforms, while drivers had to be deployed on the usually-unmanned CCL trains during the evening rush hour as a precautionary measure.

Commuters at Bishan bus interchange during the morning rush hour yesterday. The signal fault was first announced at 7.37am and full service resumed only about three hours later. PHOTO: KOH MUI FONG

The LTA and SMRT, which are still unable to pinpoint the cause of the problem, are looking to suspend telecommunication signals along stretches of the CCL, should signal interference occur again.

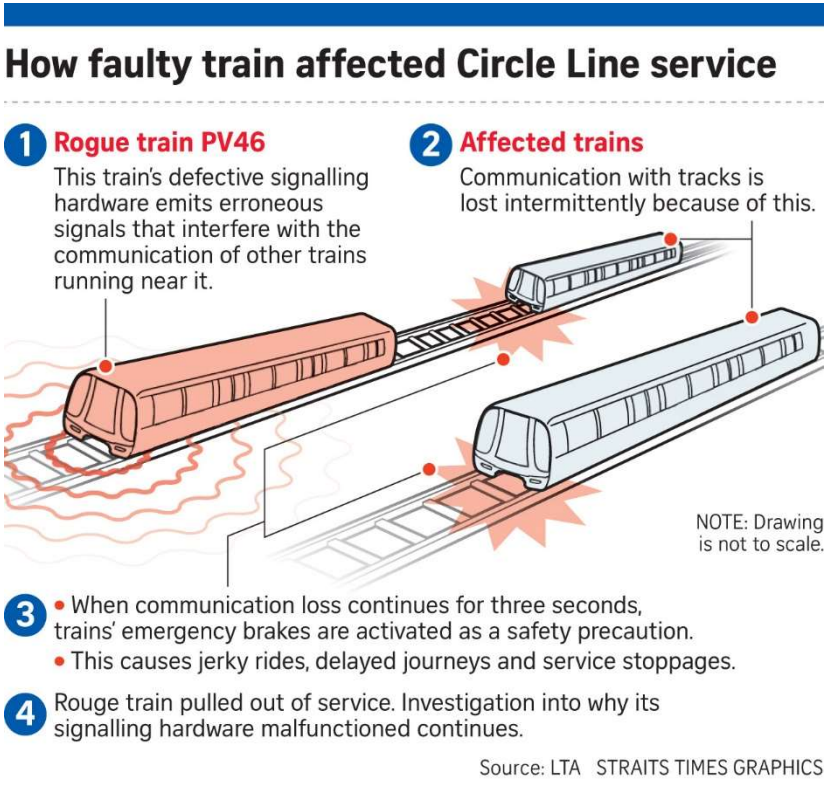
SMRT would not say how many commuters were affected by the disruption. The daily ridership of the CCL is 398,000 as of last May.

The signal fault was first announced at 7.37am, with SMRT saying over Twitter that there was a 10-minute delay between Pasir Panjang and one-north stations on the CCL. This escalated into delays on the whole line of up to 30 minutes.

At about 9.50am, SMRT announced on Twitter that service had been halted at six

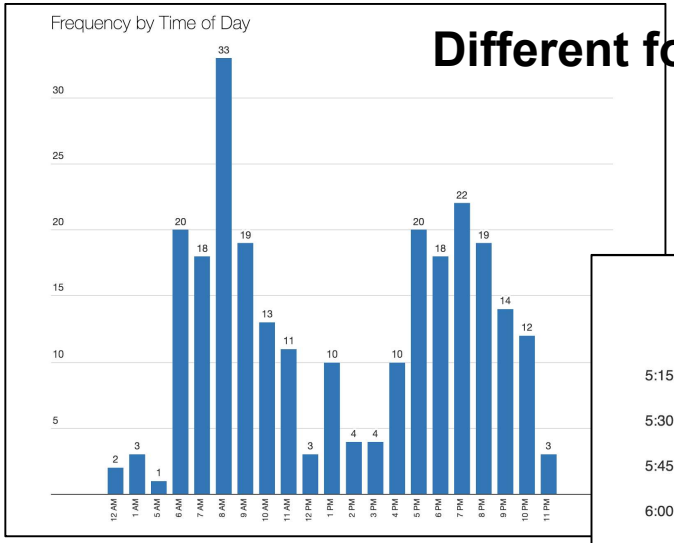
CONTINUED ON PAGE 3

from: https://www.allsingaporestuff.com/article/track-fault-train-fault-now-even-our-hp-signal-fault



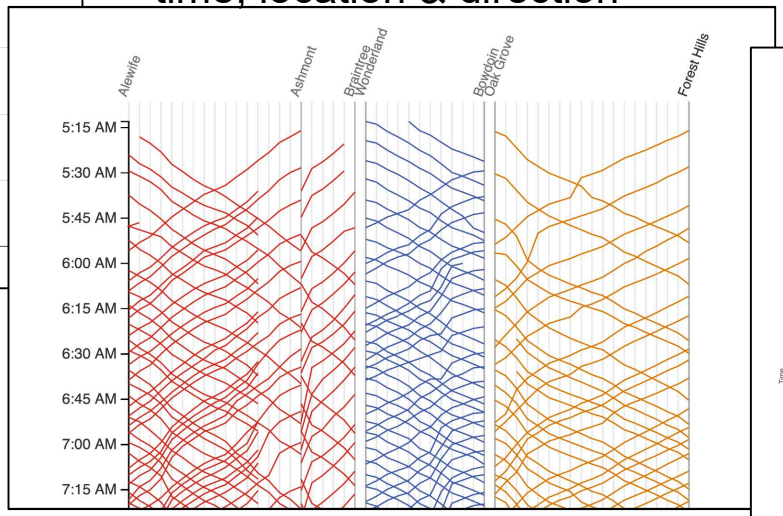
Using time, location, direction data

Different forms of Visualization (description)



Frequency of faults

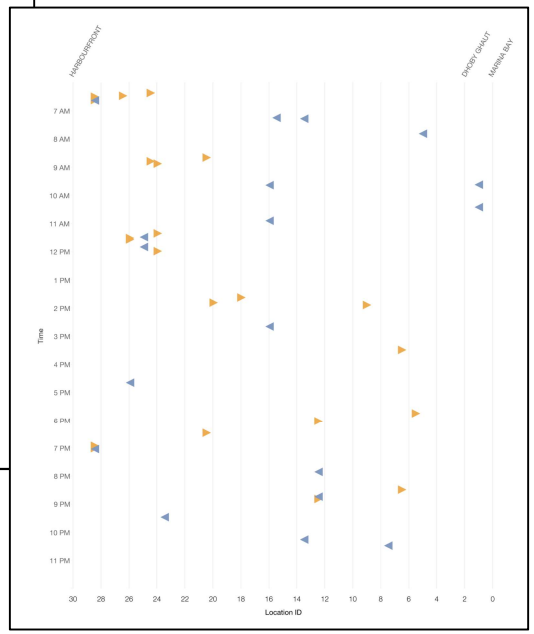
Marey Chart:Visualising time, location & direction



Mapping the faults:

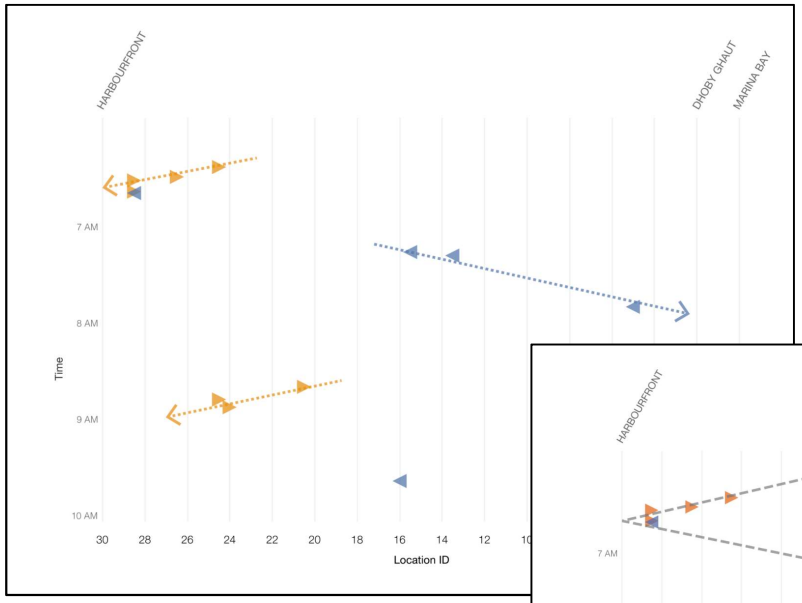
- horizontal axis are stations
- vertical axis is time
- diagonal paths are train movement

Scatterplot with train direction

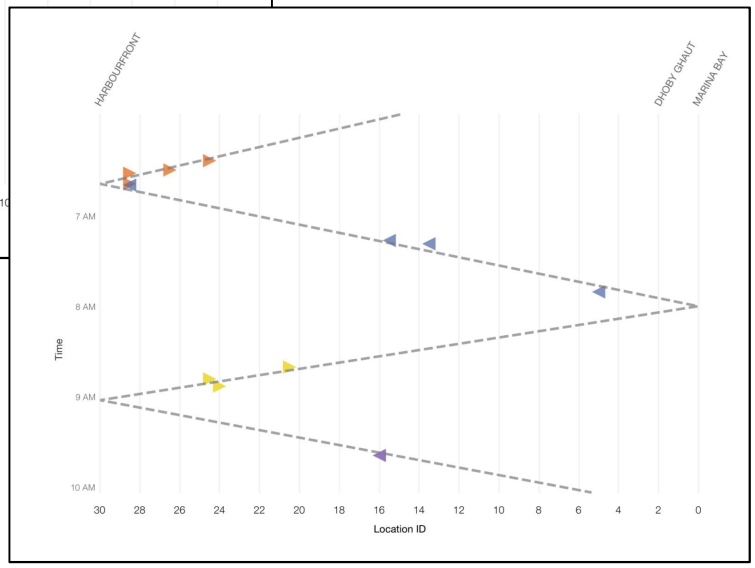


From: https://blog.data.gov.sg/how-we-caught-the-circle-line-rogue-train-with-data-79405c86ab6a

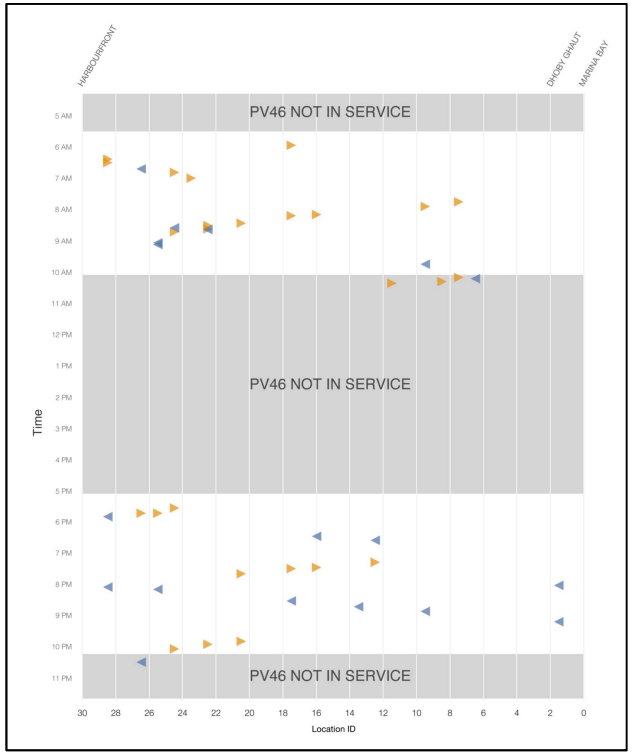
Catching the culprit…PV46!



A pattern emerges…



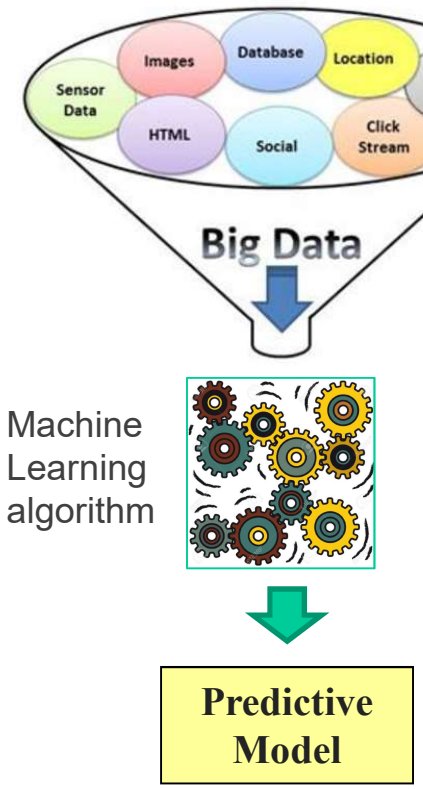
Seems to be a single train…



Testing the hypothesis…

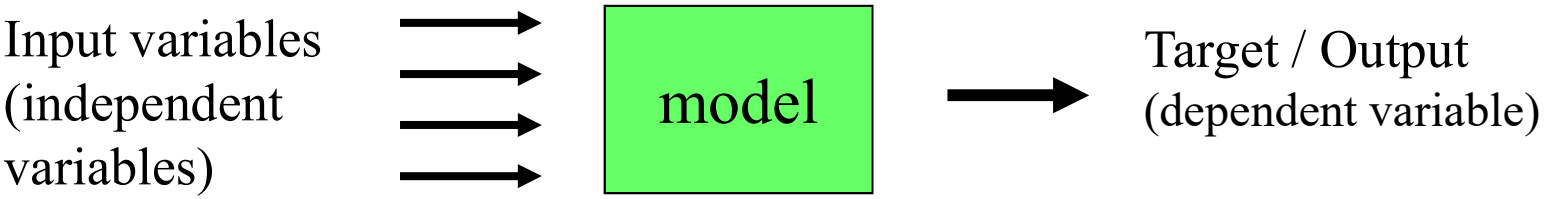
From: https://blog.data.gov.sg/how-we-caught-the-circle-line-rogue-train-with-data-79405c86ab6a

Predictive Modelling



- Who will buy my product?
- How many widgets should I buy?
- Who will close their account (churn)?
- Which loan requests are high risk?
- Which insurance claim is a likely fraud?
- Which grant application should I approve?
- Which companies will become bankrupt?
- Which patients will respond best to drug Y?
- How many beds are required next week?
- When will my device fail?
- Which part will become defective?
- …….

Predictive Modelling



- Event Prediction - will something happen?
  - E.g. will customer respond to a campaign? (Response Modelling)
  - E.g. will customer buy a product?
- Value Prediction - predicting the value of something
  - E.g. how much money will a person spend?

The accuracy required is entirely business dependent

Forecasting ATM cash demands

DBS Bank has 80 percent fewer cash-outs, improves process efficiency by 33 percent

In Singapore, customers of banking giant DBS conduct 25 million transactions a month at more than 1,100 ATMs. They rely on DBS – the island nation's largest bank – for convenient, ready access to funds day or night. To make sure that its ATMs – and its customers – don't come up empty-handed (i.e., a cash-out), DBS uses SAS to forecast withdrawal activity and to optimize the reloading process.

As a result, more than 30,000 hours of customer wait time have been eliminated annually as customers are spared the inconvenience of waiting while empty ATMs are reloaded.

> Using this innovative solution, a first in the banking world, DBS is now able to convert valuable ATM usage and customer behavior data into a daily execution plan, allowing optimal reloading at non-peak times.

David Gledhill
Managing Director and Head of Group Technology and Operations

"We serve over 4 million customers in Singapore, and it is important for us to place customers at the heart of the banking experience across all our touch points," says David Gledhill, Managing Director and Head of Group Technology and Operations at DBS Bank.

"DBS' ATMs have one of the highest utilization rates in the world. Any downtime in a single ATM would mean inconvenience for our customers. Hence, we have to continually improve the efficiency of our ATM network and operational process."

DBS analyzes withdrawal data from each ATM to forecast upcoming activity. The forecasts allow the bank to make smarter decisions about reloading its network of machines.

Predicting Lift Failure

Predictive system may ease lift failure woes

SATURDAY, NOVEMBER 26, 2016 | THE STRAITS TIMES

It was tested on lifts in several town councils and successfully predicted two breakdowns

Janice Heng

Imagine a system which can predict which lift will fail, and when. This would help town councils perform preventive maintenance before a problem breaks out, said a spokesman for Surbana Jurong (SJ).

Several town councils are hoping to achieve this by tying up with the infrastructure consultancy, which monitors more than 24,000 lifts in HDB estates, and which believes it may just have this predictive system.

The consultancy said it worked with multiple town councils to place sensors in 720 lifts, of varying models and ages, in HDB blocks. The sensors detected and recorded lift speed, acceleration and deceleration, and jerky movements. SJ declined to name the town councils.

After collecting the data, SJ put it through an algorithm, and predicted the specific days for two lift breakdowns, two weeks before they actually happened. SJ also predicted three other breakdowns but those lifts underwent scheduled maintenance beforehand.

The consultancy said it can do even better. As more data is collected over time, the algorithm will improve, allowing breakdowns to be predicted even earlier, said a spokesman.

SJ declined to say how much its predictive system will cost.

Lift maintenance came under the spotlight recently after a series of incidents, including a case where a woman's hand was severed by lift doors in October last

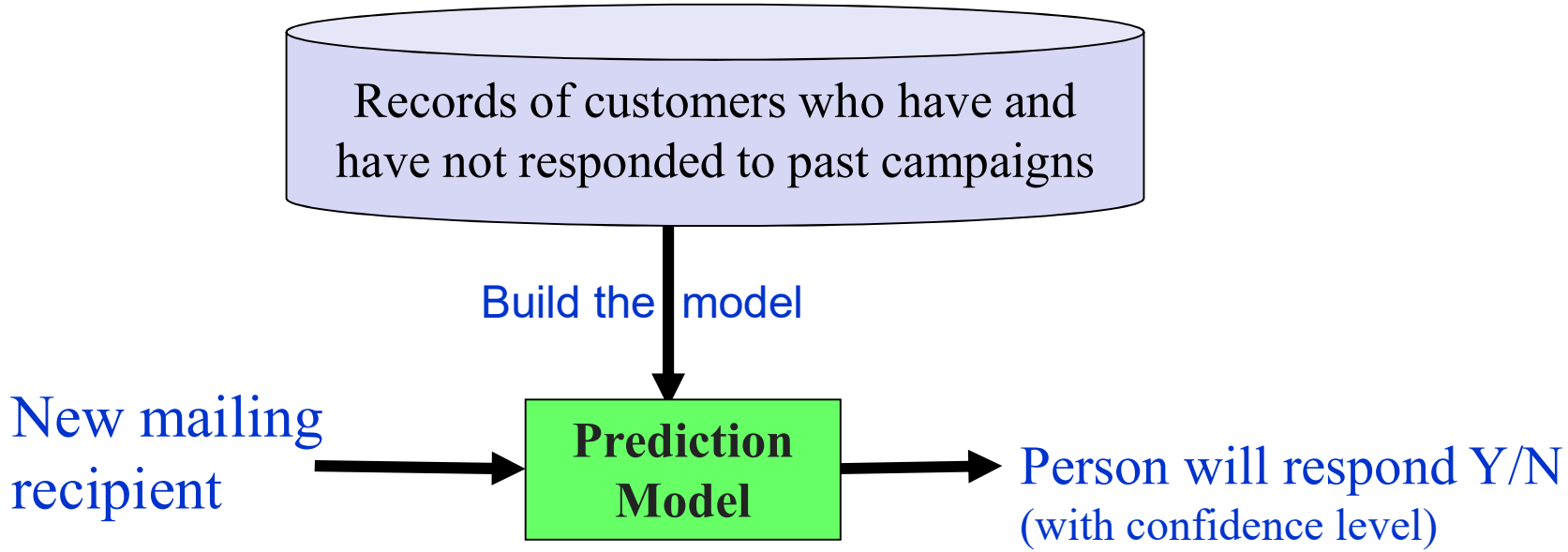
MICROSOFT TIE-UP continued on B2

"Sensors were placed in 720 lifts of varying models and ages. They recorded lift speed, acceleration, deceleration and jerky movements. Data was put through an algorithm and predicted specific days for two lift breakdowns two weeks before they happened"

(system dev. by Surbana Jurong)

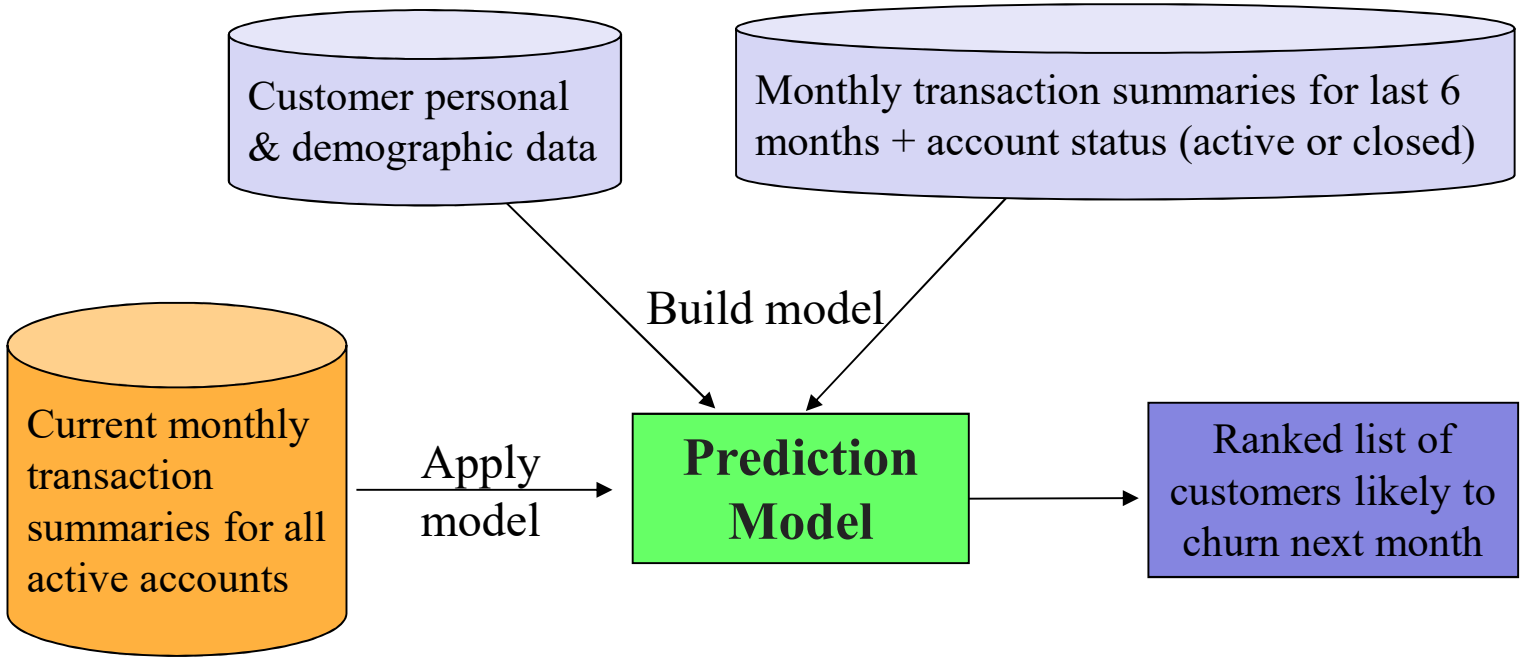
Common Prediction Apps: Response Modelling

- Predicting a T/F event, e.g. a persons response to a mailing



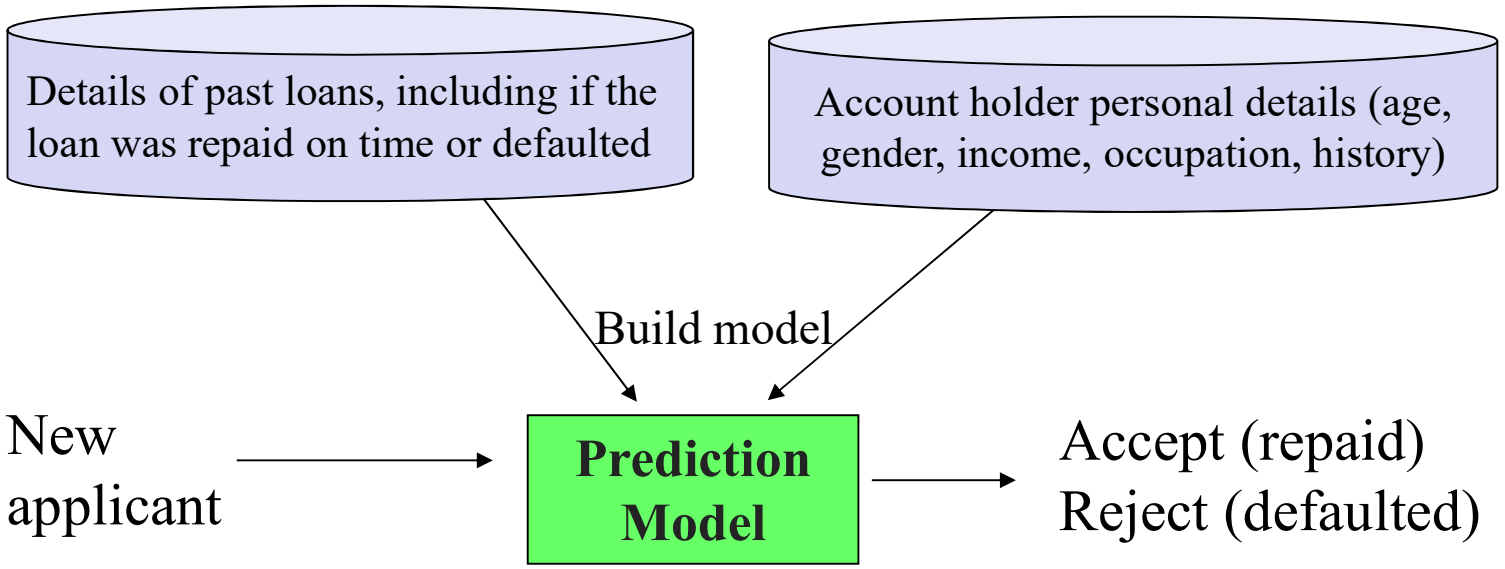
Common Prediction Apps: Churn Modelling

- Prediction if a customer will close their account in the near future
  - If yes then make them a special offer (or similar) to encourage them to stay
- The Telco industry has been a big user (pioneer) of churn modeling



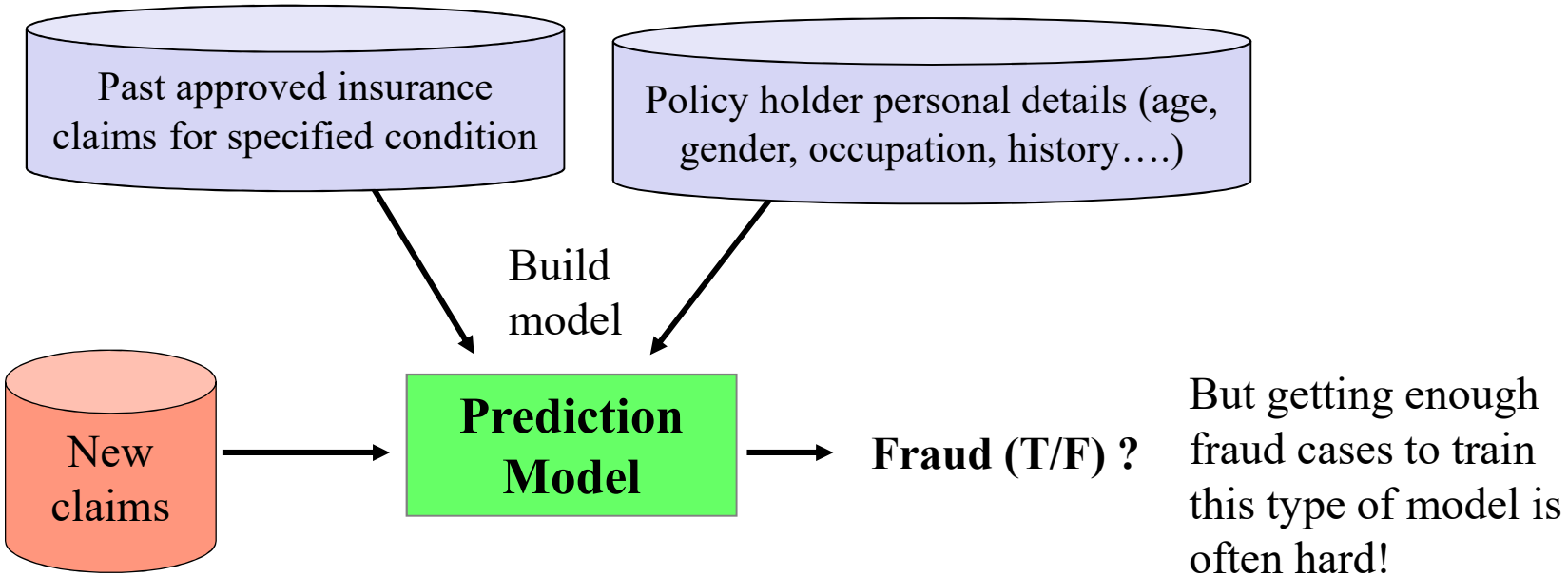
Common Prediction Apps: Credit Risk Modelling

- A person applies for a loan, should it be approved?
  - People who have the best credit don't need the loans, and people with worst credit are not likely to repay. Bank's best customers are in the middle
  - Traditional credit scoring is often a poor predictor



Common Prediction Apps: Insurance Fraud

- A large insurance company has seen a big increase in the number and size of claims against medical policies over the last year – it needs a better mechanism for detecting inflated claims

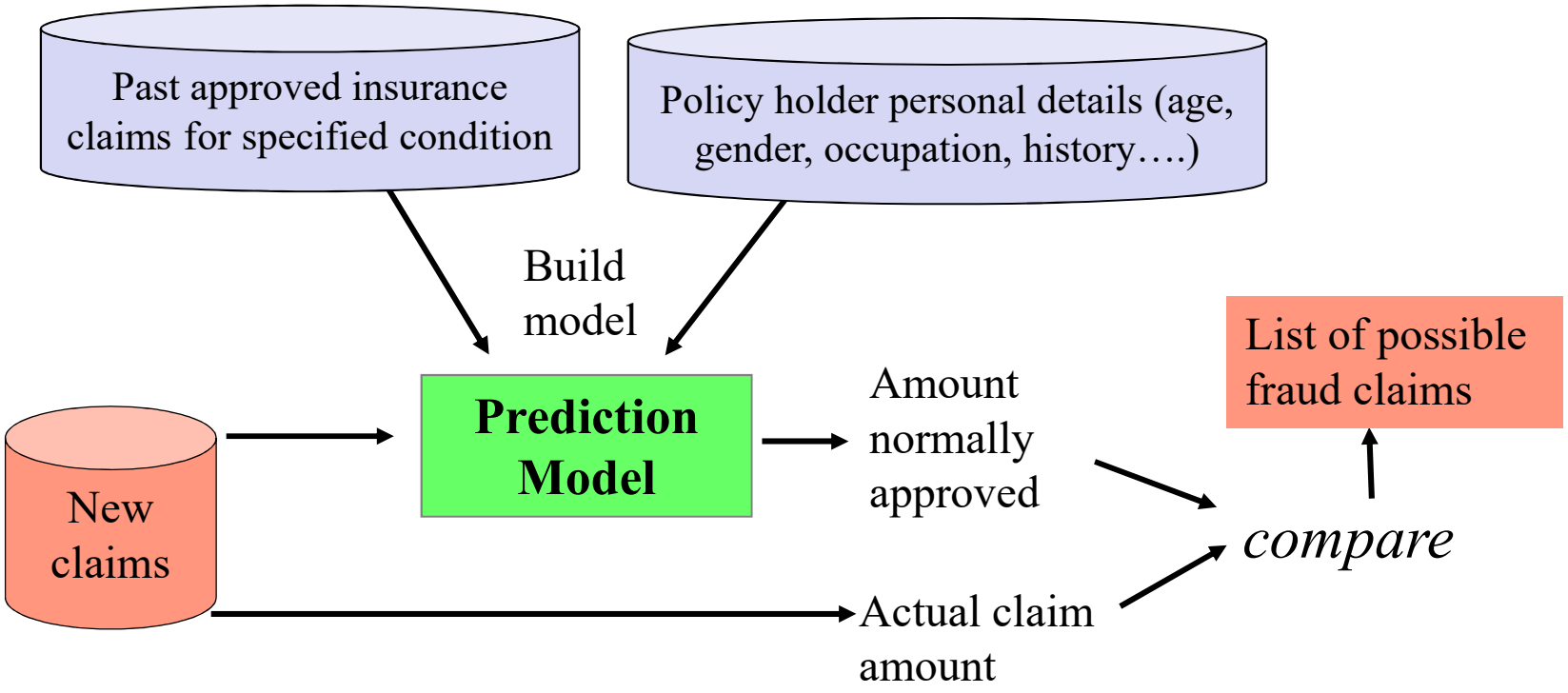


The ratio between fraud and non-fraud cases can sometimes be as much as 1 to 10,000

https://emerj.com/ai-sector-overviews/artificial-intelligence-fraud-detection-insurance/

Common Prediction Apps: Insurance Fraud

- A large insurance company has seen a big increase in the number and size of claims against medical policies over the last year – it needs a better mechanism for detecting inflated claims



Predictive Modeling Methods

- Linear regression
- Logistic regression
- Advanced Regression (GLM etc.)
- Decision tree (C5, C4.5, CART, CHAID)
- Rule sets (C5, C4.5, Association rules)
- Memory-based reasoning (K-NN)
- Naïve Bayes
- (Shallow) Neural Network (MLFF, Radial Basis,…)
- Support Vector Machine
- Ensembles (Random Forests, Bagging, Boosting, e.g. XGBoost)
- ….

Adding more AI!

Add more data…

- **Large amounts of new & disparate data is available for investigating fraud, credit risk, churn ….**

https://i-sight.com/resources/101-osint-resources-for-investigators/

E.g. Social Media data for Fraud detection

- An employee claims to have injured her shoulder on the job. While on compensation pay, a friend posts on Facebook “Can’t wait to play tennis tonight!” and she responds “Me neither!”.
- A person claims that his car was stolen. However, days later he posts a photo of himself on Instagram with the car (and license plate showing)

Social Network methods can automate the process of drawing connections from disparate data sources and visually representing them as a network.

In the UK, a large P&C insurer made £7 million savings per annum by uncovering groups of collaborating fraudsters using **network analytics**.

Text Data brings new challenges

Predict click-through rates on display ads

Display advertising is a billion dollar effort and one of the central uses of machine learning on the Internet. However, its data and methods are usually kept under lock and key. In this research competition, CriteoLabs is sharing a week's worth of data for you to develop models predicting ad click-through rate (CTR). Given a user and the page he is visiting, what is the probability that he will click on a given ad?

Kaggle Competition

Display-Ad Click Prediction

- Use ***past user behaviors*** + ***current context*** + ***ad properties*** to predict clicks on specific ads or ad categories. Can execute in real-time to help Ad selection when a user visits a website in the ad network
- Each past Ad impression is a potential training example:

• Date & Time of impression • User-Id or Cookie-Id • Advertiser characteristics • Ad Properties ~ size, shape, type (text, banner, image, video)… • Placement properties ~ location on the page, duration on screen, … • Past User Behaviors ~ pageviews, searches, ads clicked,.. • User Demographics (if available)	Possible model inputs
• Did click occur (T/F) ---------->	Model target

- Logistic Regression ~~is~~ was commonly used due to its scalability

http://people.csail.mit.edu/romer/papers/TISTRespPredAds.pdf

Simple and scalable response prediction for display advertising

OLIVIER CHAPELLE, Criteo†
EREN MANAVOGLU, Microsoft
ROMER ROSALES, LinkedIn

Display-Ad Click Prediction

- How to turn user behaviors into suitable inputs for a prediction model?
- E.g. Past events for one user:

- ❑ Searched for *"cheap flights to New York"*, 11.15am, Jan2nd
- ❑ Clicked on *"Travelocity"* link in search results, 11.16am, Jan2nd
- ❑ Viewed http://www.straitstimes.com/breaking-news/money/story/ps4-and-xbox-fuel-sales-20140117, 10pm, Jan3rd
- ❑ Received Ad2317 (SingTel) 10.00pm, Jan 3d
- ❑ Searched for *"tourist hotels in New York"*, 8.10pm, Jan4th
- ❑ Clicked on *"Holiday Inn"* link in search results, 8.12pm, Jan4th
- ❑ Clicked on Ad2523 (discount offer from Hilton), 8.15pm, Jan 4th
- ❑ Received email from *Target.com*, 9am, Jan 5th
- ❑ Viewed http://www.amazon.com/gp/goldbox/cell-phoines , 2.32pm, Jan5th
- ❑ Viewed http://www.amazon.com/categories/scifi-books/id3456, 2.32pm, Jan5th

- How many days or weeks to look-back? When does behavior go stale?

Encoding Behaviors: One Method…

- Text Categorisation - Categorise the user events (page-views, searches…) into a topic taxonomy

```
Clothing & Shoes
Clothing & Shoes/Baby Clothing & Accessories
Clothing & Shoes/Boys Clothing & Accessories
Clothing & Shoes/Costumes
Clothing & Shoes/Girls Clothing & Accessories
Clothing & Shoes/Girls Clothing & Accessories/Formal Wear
Clothing & Shoes/Gloves & Mittens
Clothing & Shoes/Hosiery & Socks
```

- Taxonomies can have many thousands of entries (nodes)
- Typically there is an accuracy/granularity versus computation trade-off

- Count the user events in each category in (say) the last 14 days
  - e.g. User2341, Clothing/Babywear=1, Clothing/Maternity=2,… etc.
- Combine with all other features, one record per ad impression

Possible model inputs

Impression	User	Demographics			Recent Behaviors			Advertiser and Ad Details			**Target**
Date-Time	ID	Sex	Age	…	Node1	Node2	…	Size	Type	…	Clicked
1/3/15:07:15	1	M	31	…	0	1	…	200*600	image	…	T
5/8/15:16:15	2	F	18	…	2	3	…	400*300	banner	…	F

But this can result in records with huge dimensionality!

…or, using a Bag-of-Words Approach

- Bag of words approach eliminates the need for the event categorisation

Impression	User	Demographics			Recent Behaviors			Advertiser and Ad Details			**Target**
Date-Time	ID	Sex	Age	…	Term1	Term2	…	Size	Type	…	Clicked
1/3/15:07:15	1	M	31	…	0	1	…	200*600	image	…	T
5/8/15:16:15	2	F	18	…	2	3	…	400*300	banner	…	F

E.g. the terms extracted from all of the users recent page-views and search queries

budget	cheap	chicago	hotel	motel	new york	….
1	0	0	1	0	1	
0	2	1	0	1	0	
1	1	1	0	1	1	

But this can result in records with even larger dimensionality!

Basic NLP – Bag of Words (BOW)

Bag of Words Example

Document 1

The quick brown fox jumped over the lazy dog's back.

Document 2

Now is the time for all good men to come to the aid of their party.

Term	Document 1	Document 2
aid	0	1
all	0	1
back	1	0
brown	1	0
come	0	1
dog	1	0
fox	1	0
good	0	1
jump	1	0
lazy	1	0
men	0	1
now	0	1
over	1	0
party	0	1
quick	1	0
their	0	1
time	0	1

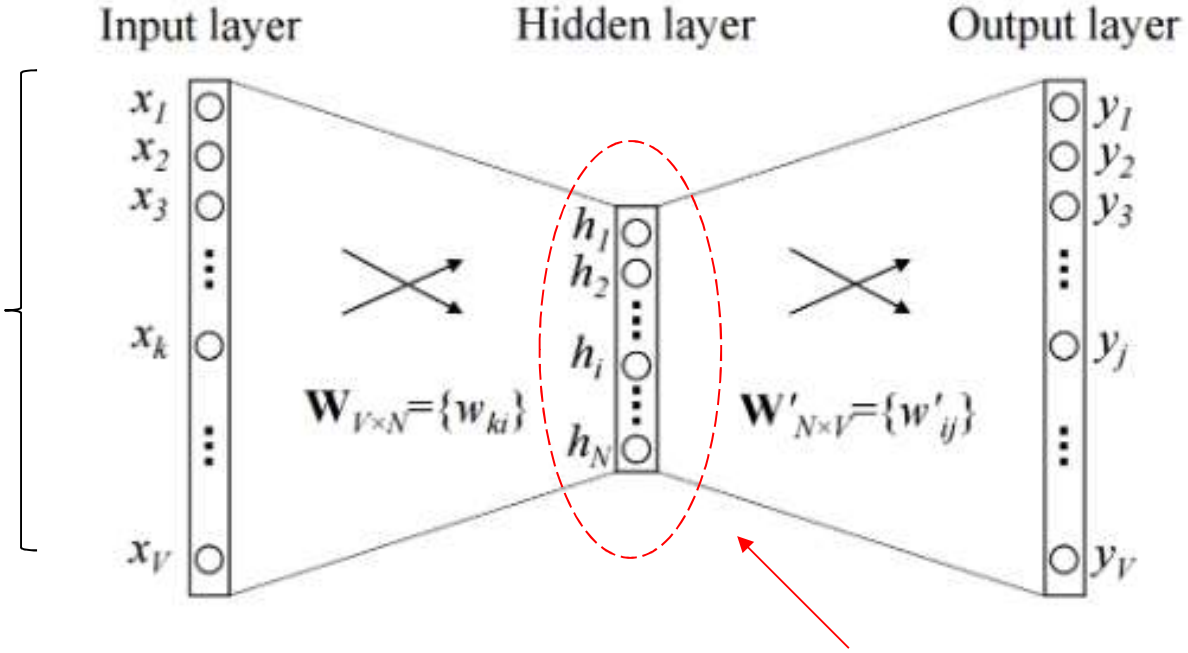
Stopword List

for
is
of
the
to

Data Reduction using Deep Learning

- One common DL architecture is the Auto-encoder
- The inputs and outputs are the same, the hidden layer(s) learn a compressed version of the input data which can then be used to recreate the inputs
- E.g. use a BOW as inputs , the smaller hidden layer then represents an encoding (embedding) of the inputs

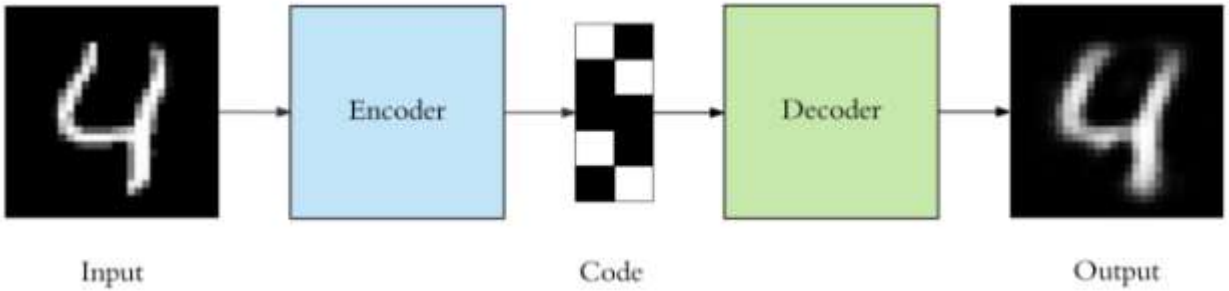
BOW

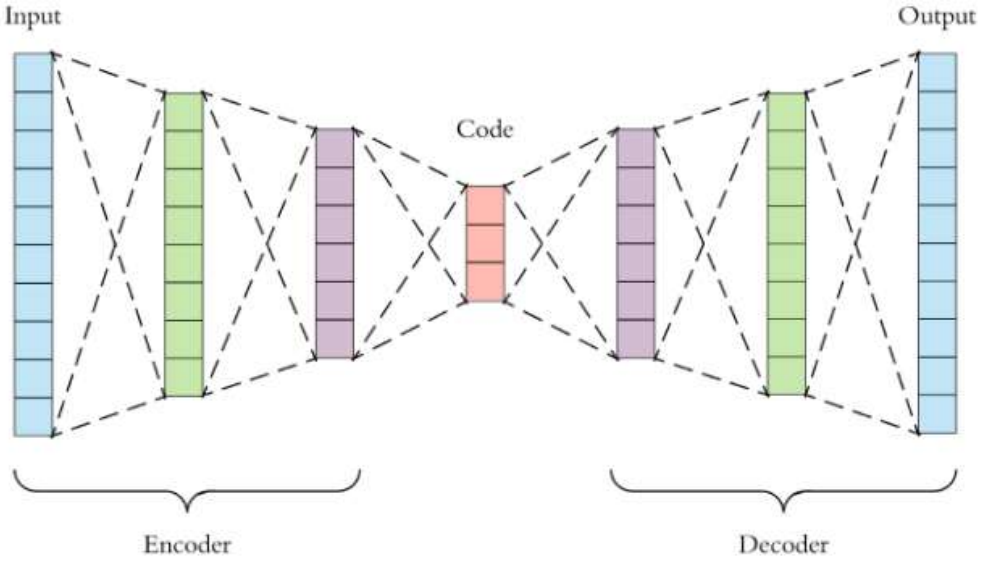


e.g. the encoding / embedding

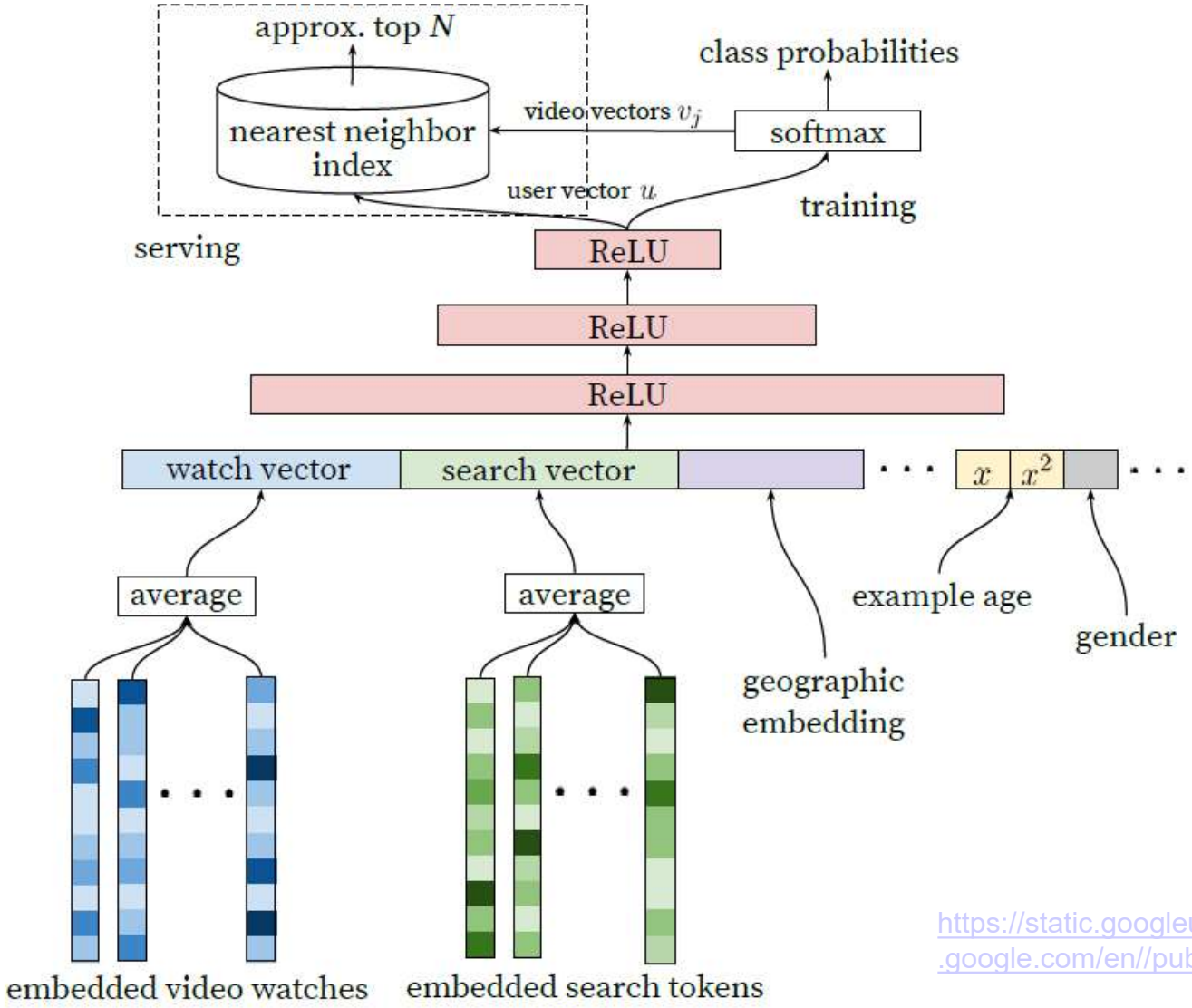
Other Applications of Auto-Encoders

- Learning to compress images





DL for YouTube Recommendations



Deep candidate generation model architecture showing embedded sparse features concatenated with dense features.

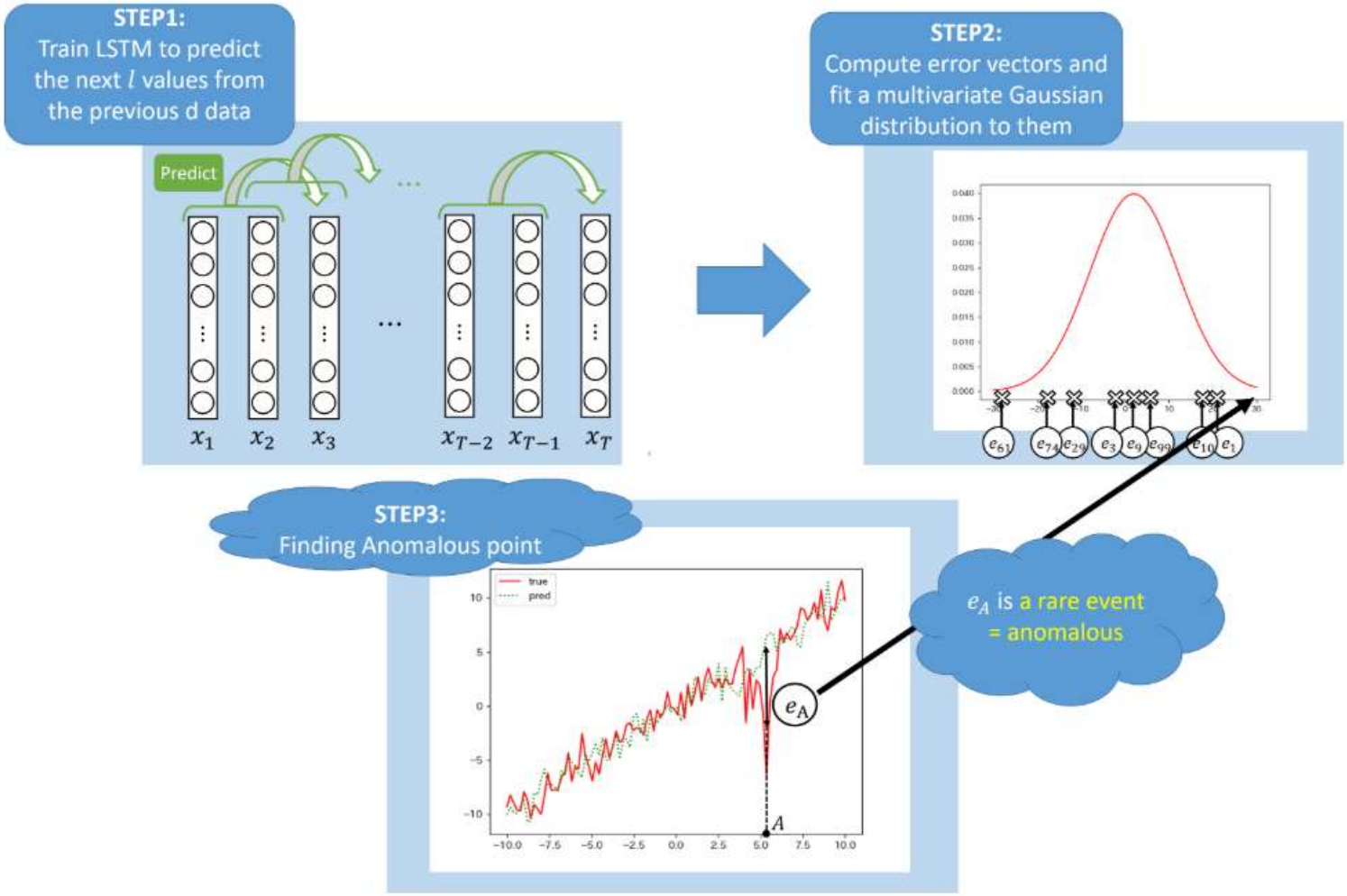
Embeddings are averaged before concatenation to transform variable sized bags of sparse IDs into xed-width vectors suitable for input to the hidden layers.

All hidden layers are fully connected. In training, a cross-entropy loss is minimized with gradient descent on the output of the sampled softmax.

At serving, an approximate nearest neighbor lookup is performed to generate hundreds of candidate video recommendations.

https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf

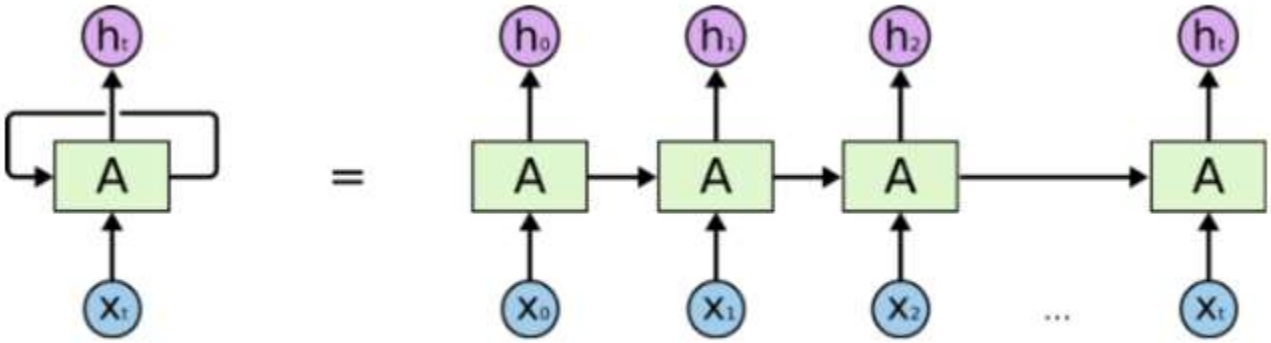
DL for Time Series Analysis



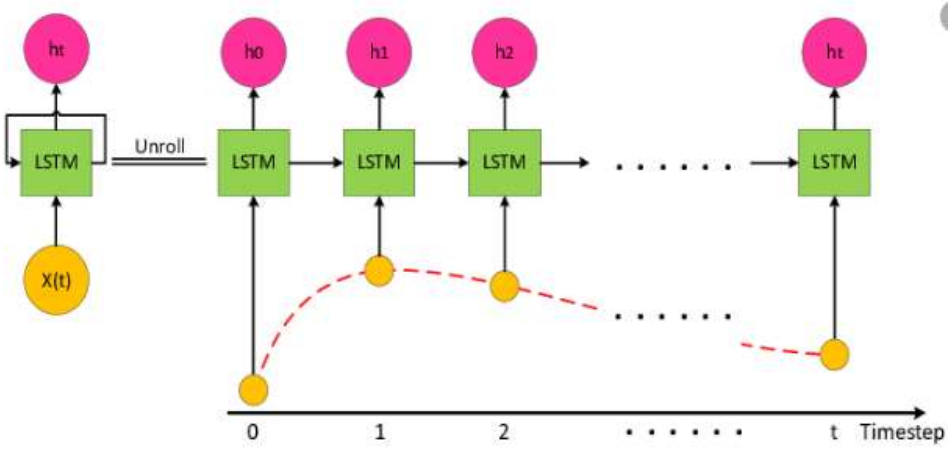
E.g. LSTM for time series prediction, or finding Anomalies in time series

Lon Short Term Memory - LSTM

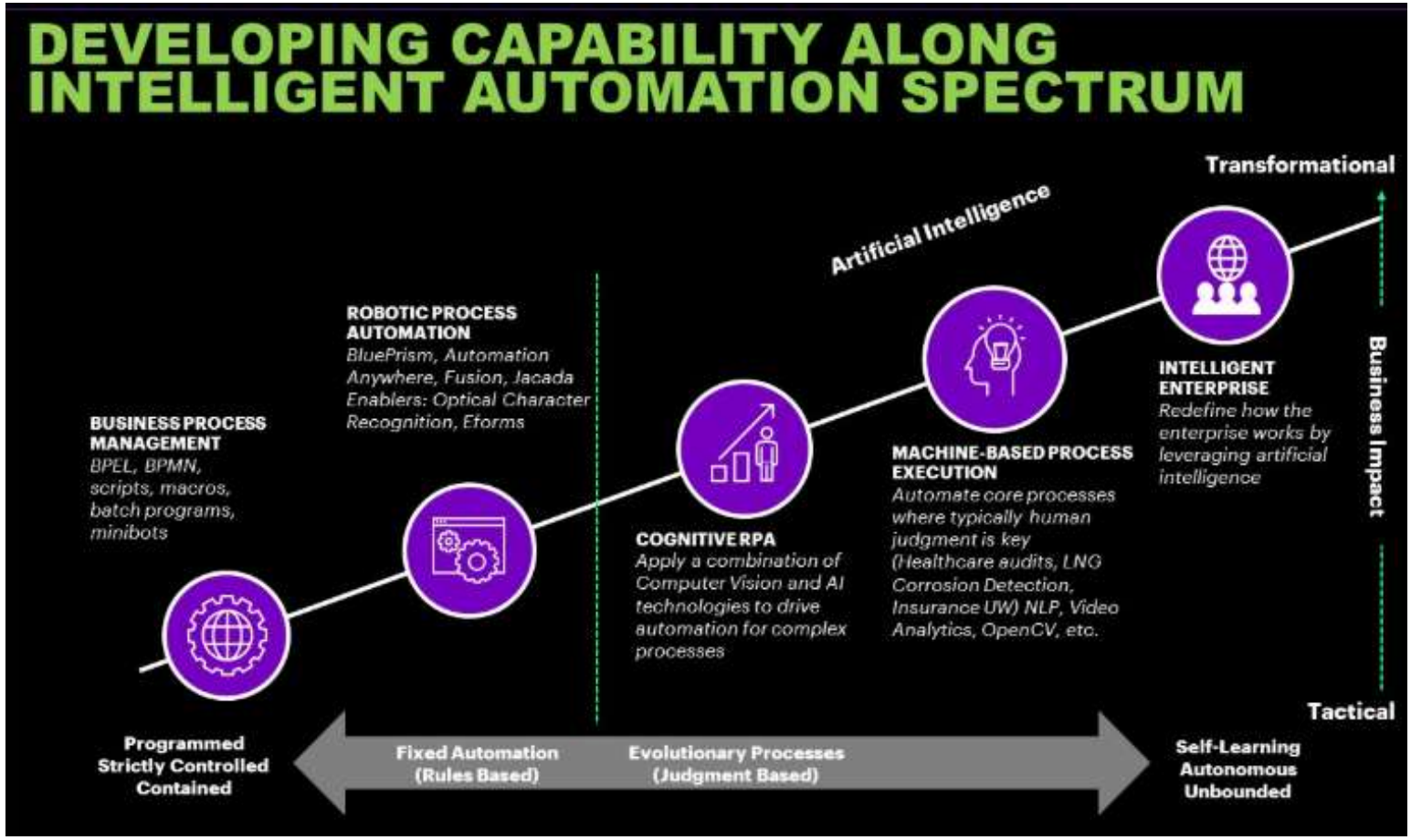
- A form of recurrent NN. Unlike standard feedforward NN's, LSTM has feedback connections. It can not only process single data points, but also entire sequences of data



when Xt comes in, the hidden state from Xt-1 will be concatenated with Xt and become the input for the Network at time t. This process will be repeated for every sample in a time-series.



AI for Business Automation

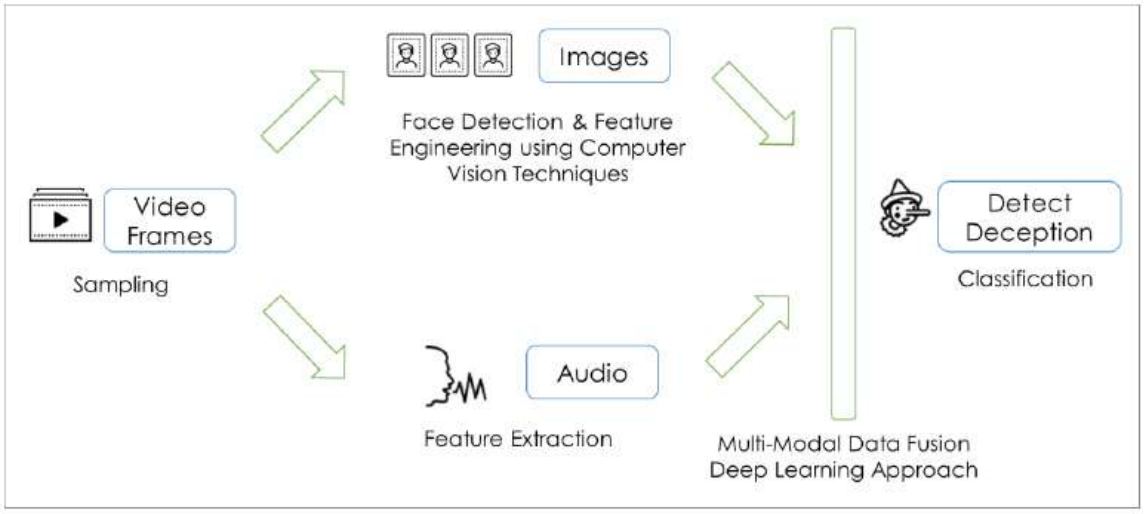


https://www.accenture.com/us-en/blogs/blogs-better-with-bots

ISS FYP: Insurance Claim Processing

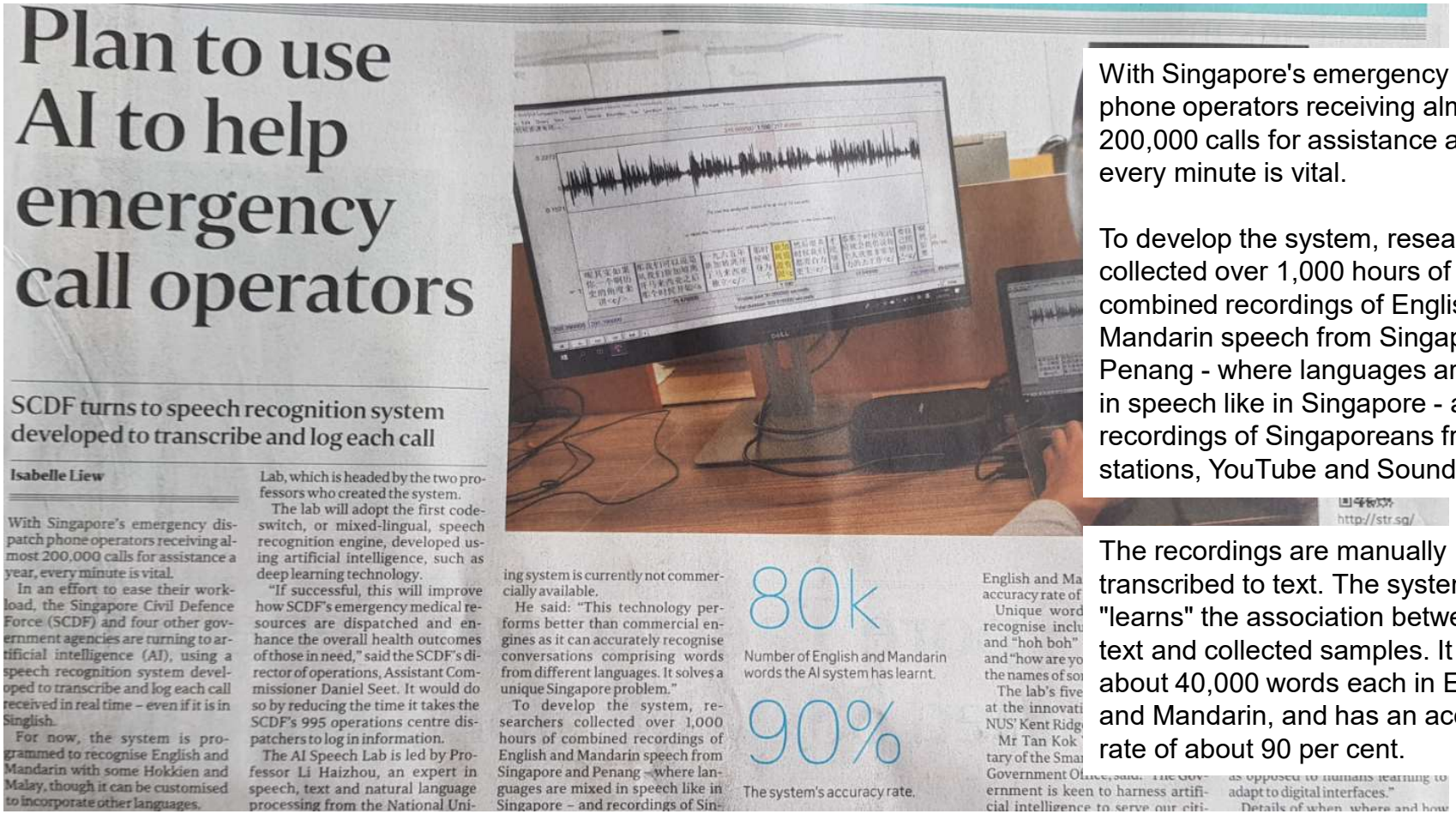
One of the traditional inefficiencies in the insurance industry is the convoluted claims assessment process which involves multiple parties for loss assessment and fraud detection. The lengthy claims process is a common pain point experienced by insurance consumers and providers alike. There is potential to improve the claims experience by making use of artificial intelligence techniques, to create a seamless claims pay-out experience and a sufficiently robust claims fraud detection and prevention system.

A possible way to re-engineer the process is to require the claimant to submit a short live video to describe the claim and the incident



AI Automation in the Call Centre

Straits Times: July 10th 2018



Plan to use AI to help emergency call operators

SCDF turns to speech recognition system developed to transcribe and log each call

Isabelle Liew

With Singapore's emergency dispatch phone operators receiving almost 200,000 calls for assistance a year, every minute is vital.

In an effort to ease their workload, the Singapore Civil Defence Force (SCDF) and four other government agencies are turning to artificial intelligence (AI), using a speech recognition system developed to transcribe and log each call received in real time – even if it is in Singlish.

For now, the system is programmed to recognise English and Mandarin with some Hokkien and Malay, though it can be customised to incorporate other languages.

Lab, which is headed by the two professors who created the system.

The lab will adopt the first code-switch, or mixed-lingual, speech recognition engine, developed using artificial intelligence, such as deep learning technology.

"If successful, this will improve how SCDF's emergency medical resources are dispatched and enhance the overall health outcomes of those in need," said the SCDF's director of operations, Assistant Commissioner Daniel Seet. It would do so by reducing the time it takes the SCDF's 995 operations centre dispatchers to log in information.

The AI Speech Lab is led by Professor Li Haizhou, an expert in speech, text and natural language processing from the National Uni-

ing system is currently not commercially available.

He said: "This technology performs better than commercial engines as it can accurately recognise conversations comprising words from different languages. It solves a unique Singapore problem."

To develop the system, researchers collected over 1,000 hours of combined recordings of English and Mandarin speech from Singapore and Penang – where languages are mixed in speech like in Singapore – and recordings of Sin-

80k

Number of English and Mandarin words the AI system has learnt.

90%

The system's accuracy rate.

English and Ma
accuracy rate of
Unique word
recognise inclu
and "hoh boh"
and "how are yo
the names of so
The lab's five
at the innovati
NUS' Kent Ridg
Mr Tan Kok
tary of the Sma
Government Of
ernment is keen to harness artifi-
cial intelligence to serve our citi-

adapt to digital interfaces."
Details of when, where and how

With Singapore's emergency dispatch phone operators receiving almost 200,000 calls for assistance a year, every minute is vital.

To develop the system, researchers collected over 1,000 hours of combined recordings of English and Mandarin speech from Singapore and Penang - where languages are mixed in speech like in Singapore - and recordings of Singaporeans from radio stations, YouTube and SoundCloud.

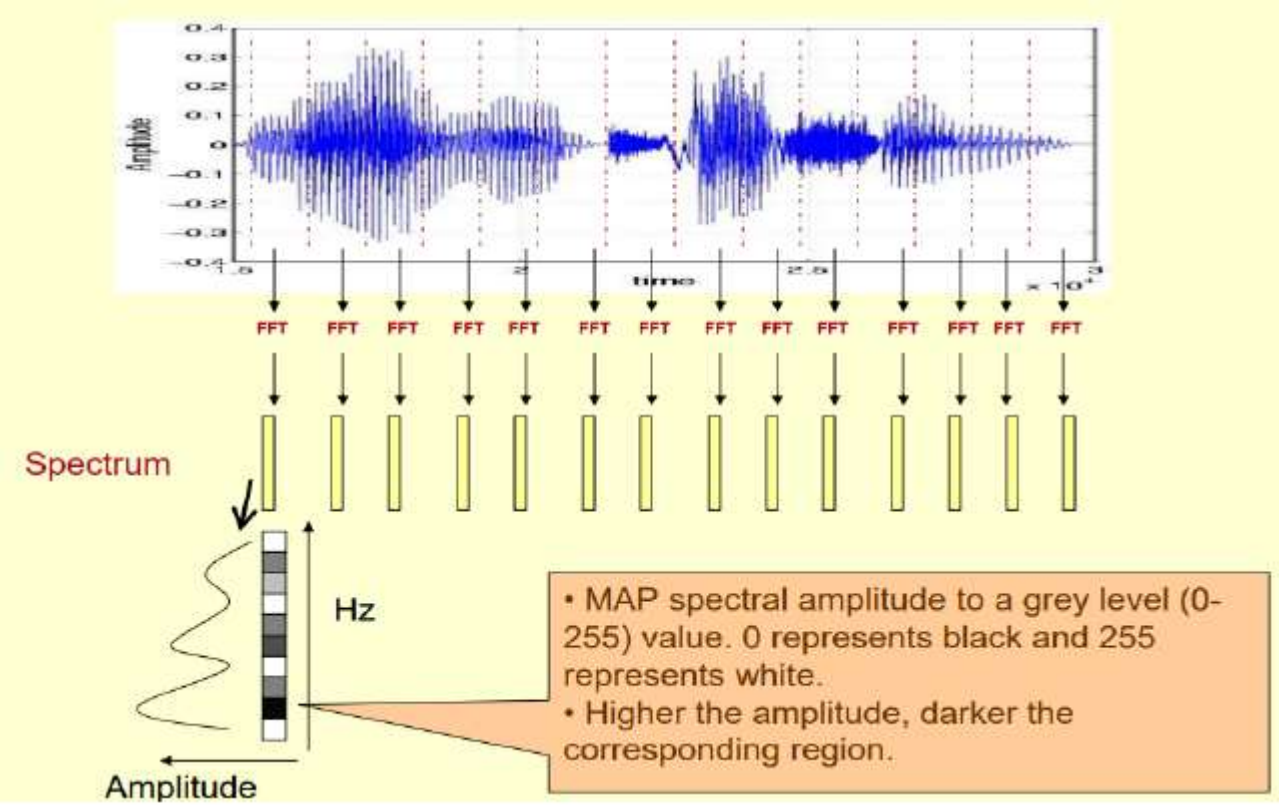
The recordings are manually transcribed to text. The system then "learns" the association between the text and collected samples. It knows about 40,000 words each in English and Mandarin, and has an accuracy rate of about 90 per cent.

AI Singapore (AISG), a programme under the National Research Foundation, is investing $1.25 million to set up the AI Speech Lab

https://youtu.be/zAdjjFCMX8Y

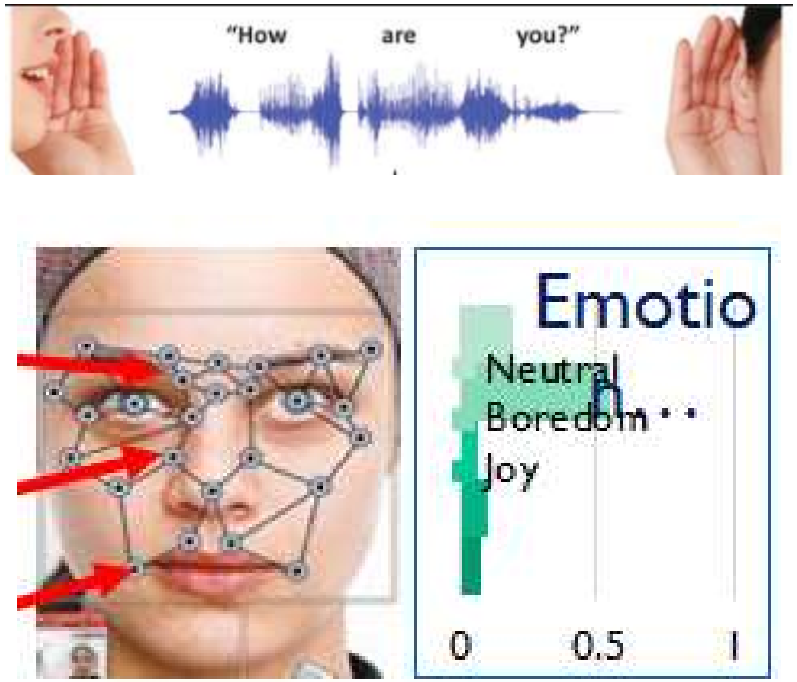
ISS FYP: Localised Real-time Neural Singaporean-Accented Speech-to-Text Classifier System

- Converting Sound to Data, then input to Gated Recurrent Unit (GRU),

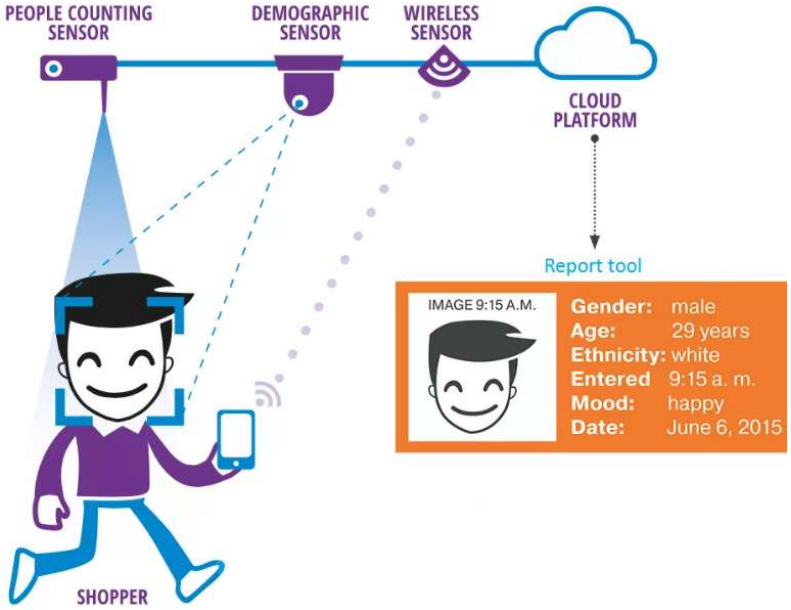


AI in the Call Centre - Emotion Detection

In the call centre

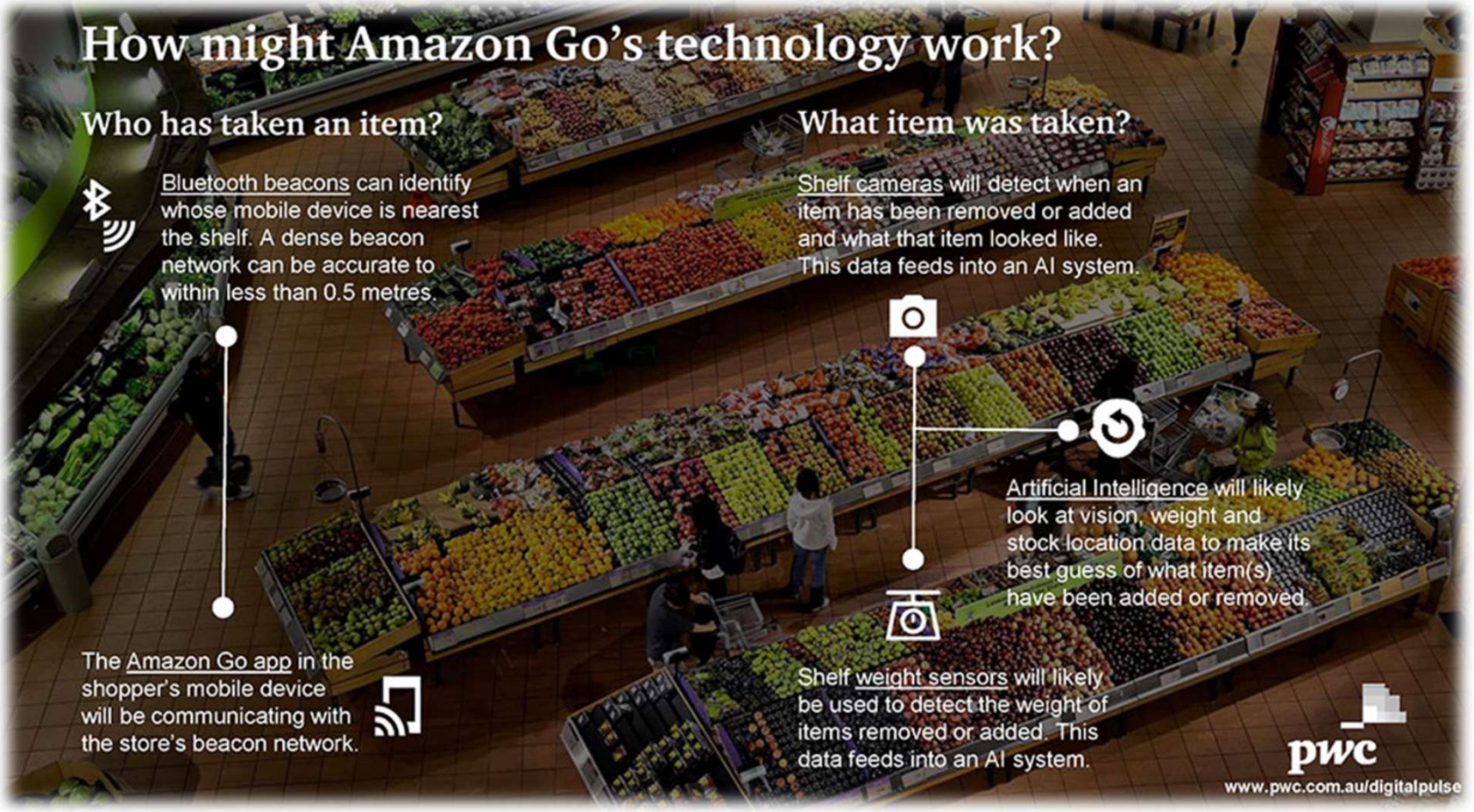


In the shopping mall



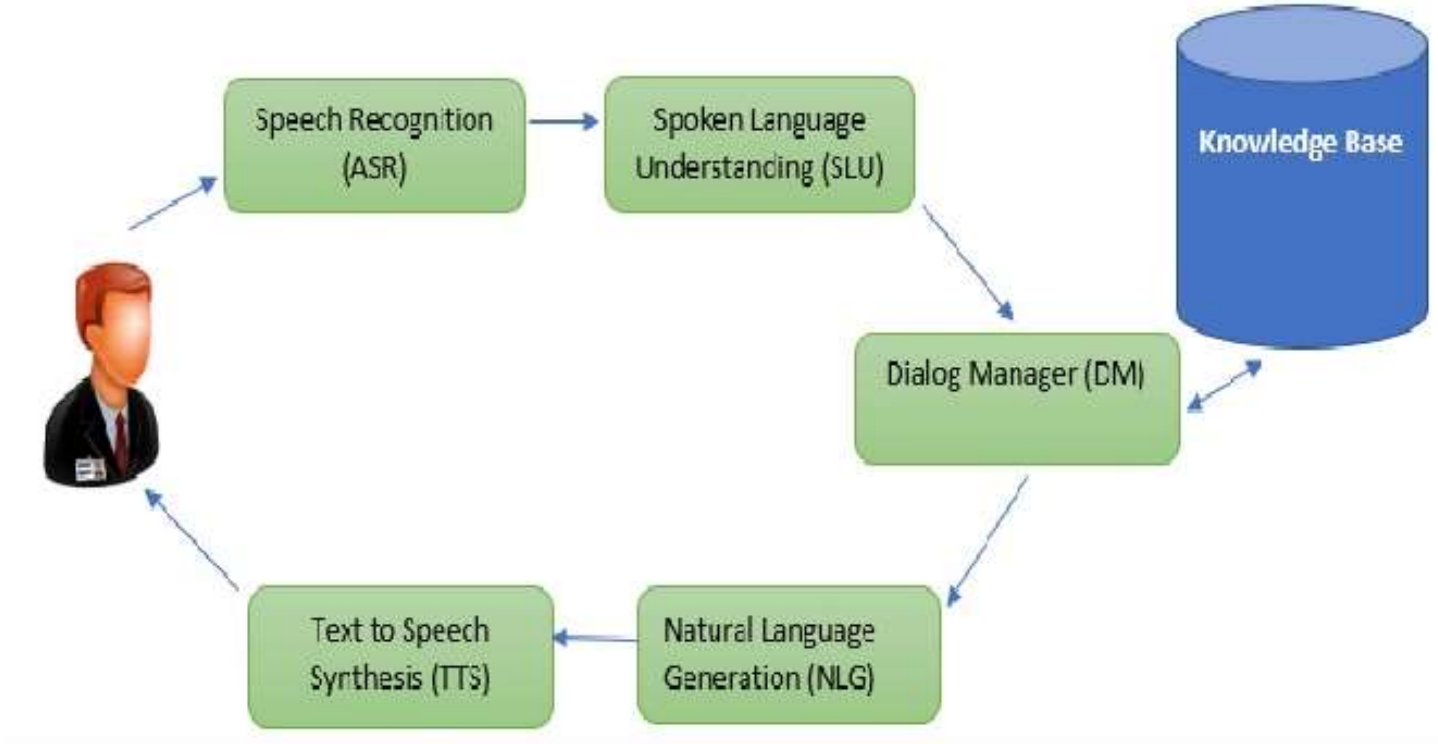
Emotions are found to have the single greatest impact on customer decisions & customer experience

AI-Driven Automation in Retail

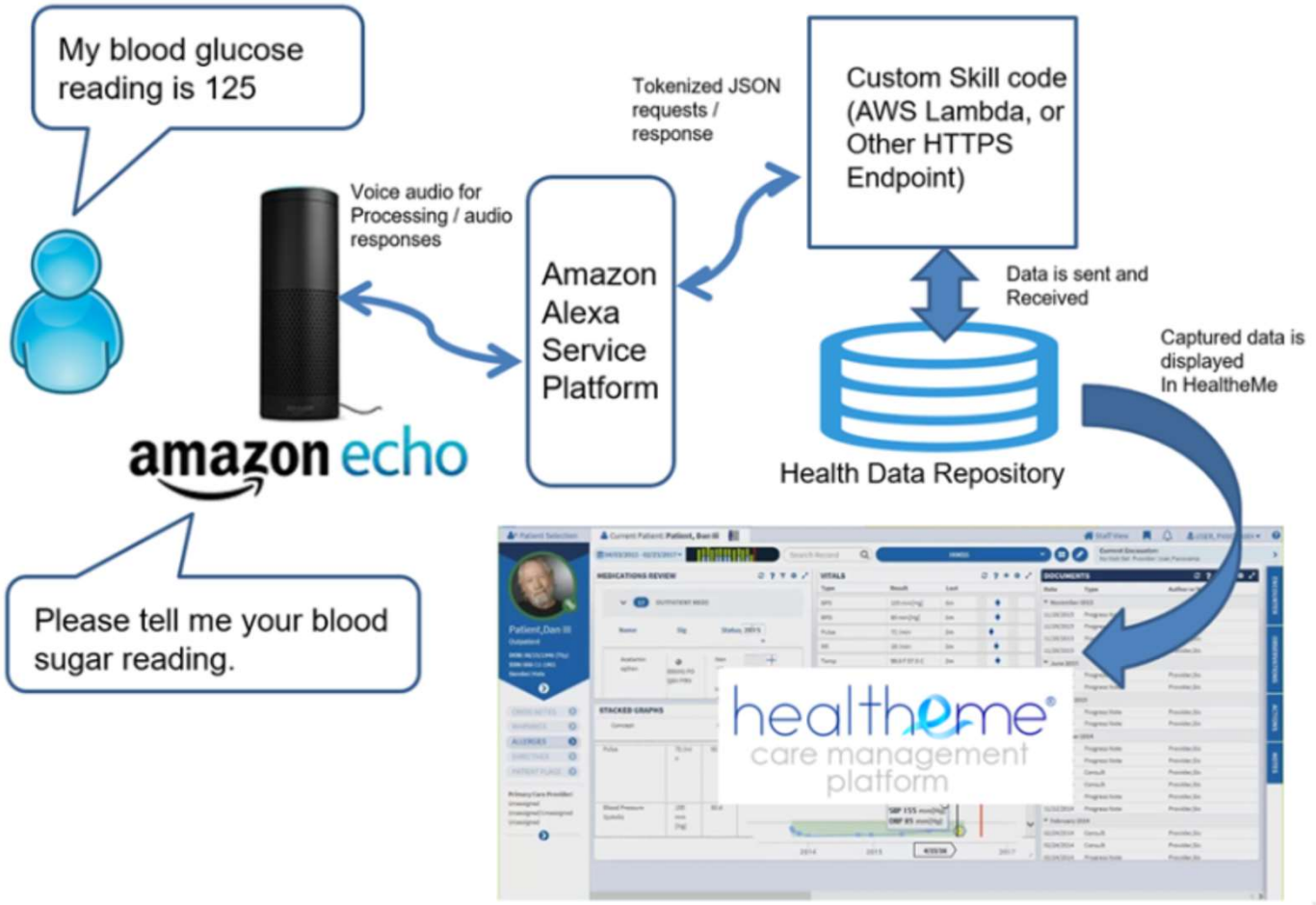


Sensors, sensors, sensors, …..

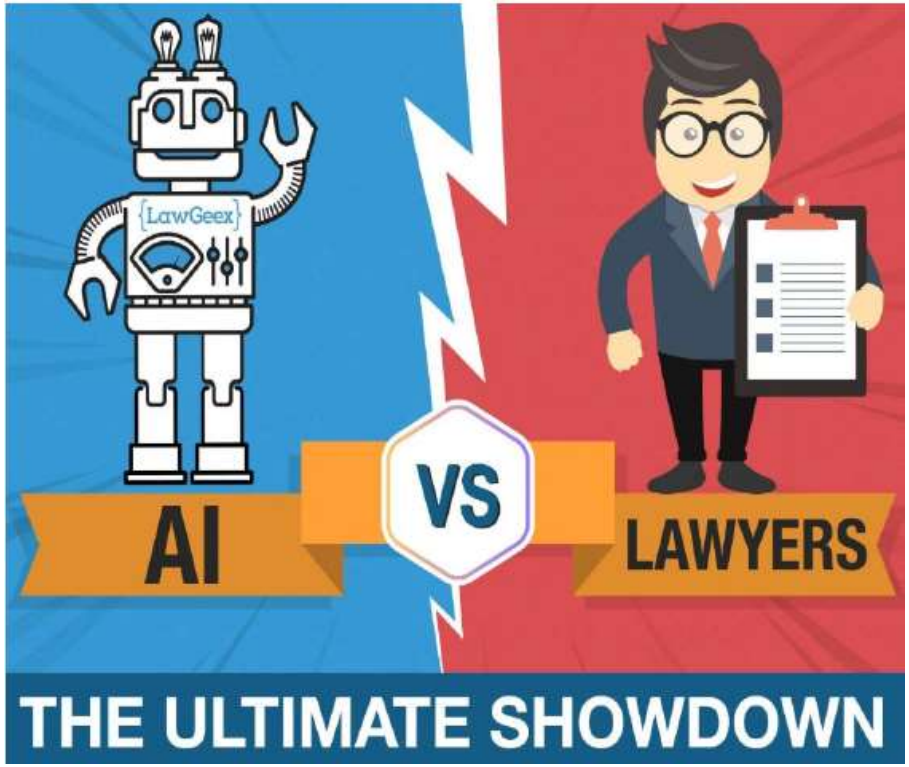
Conversational AI – AI Assistants / Agents



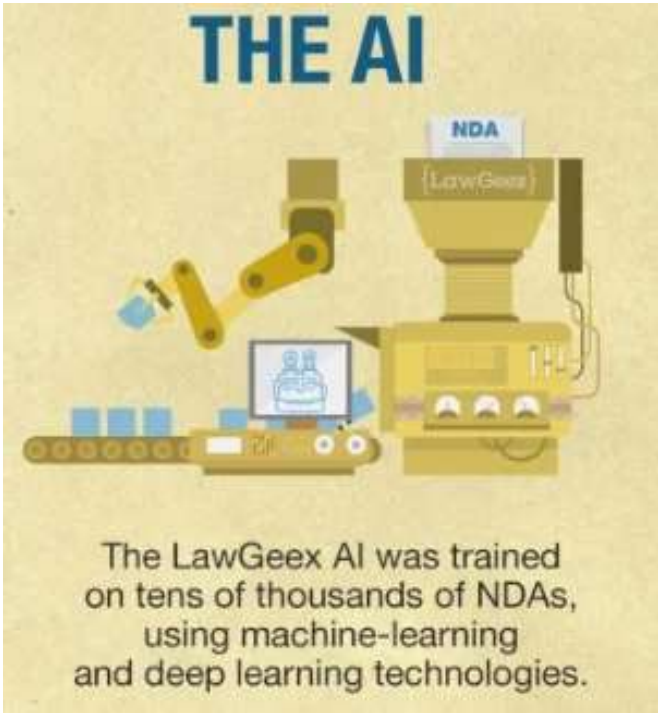
Moving Beyond FAQ's



AI Advisors



https://blog.lawgeex.com/ai-more-accurate-than-lawyers/



- 20 lawyers vs LawGeex AI
- Review 5 NDAs in 4 hours.
- 3213 clauses

	AI	Lawyers
Accuracy	94%	Avg 85%
Time taken to review all NDAs	26 seconds	Avg 92 minutes