Smart field monitoring with AIoT
AIoT在監測的發展與應用

Presenter : Yin Jeh Ngui, **Jason 魏殷哲**

National Yang Ming Chiao Tung University

30th September 2021

Short intro

DPWE 防災與水環境研究中心
Disaster Prevention & Water Environment Research Center

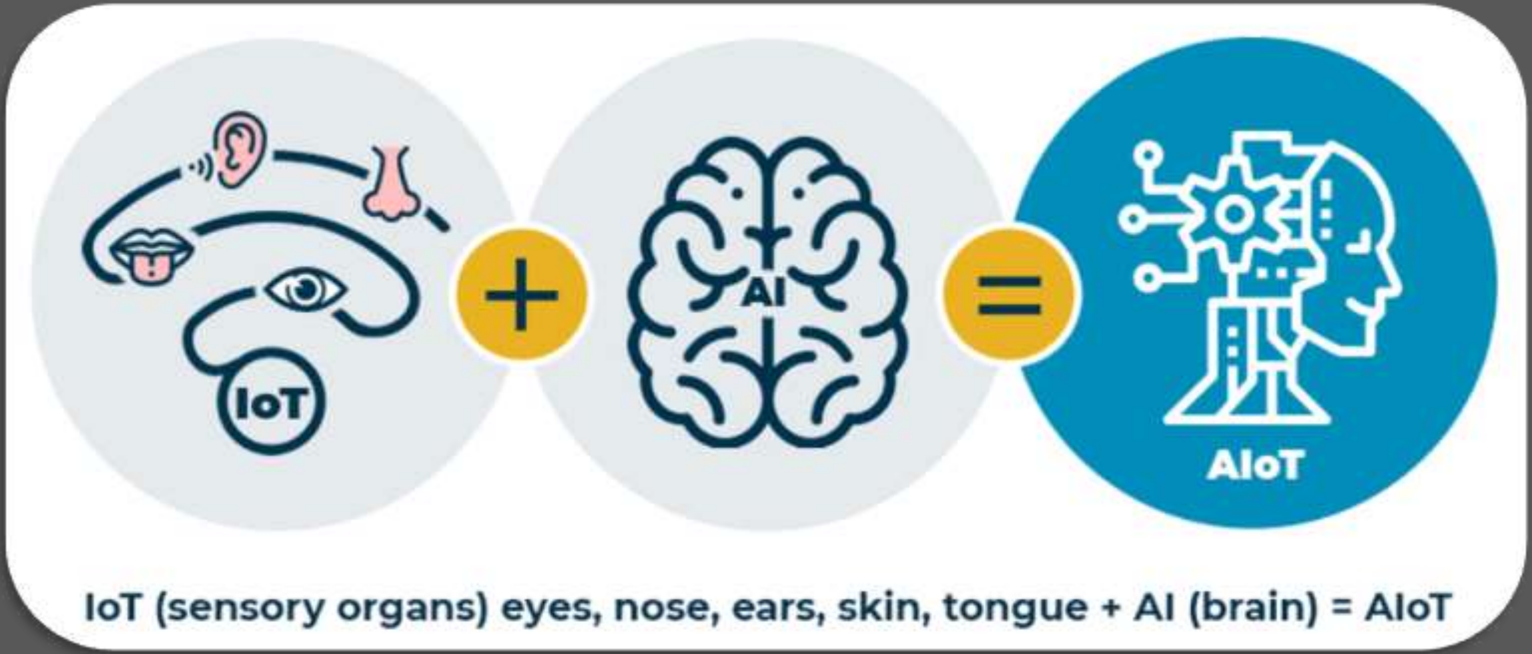
- 2019 – Current : Post-doctoral researcher at DPWE, NYCU
  - Integrated slope monitoring with low-powered long-range IoT devices
  - Dielectric spectroscopy using TDR
  - Suspended sediment concentration (SSC) monitoring in reservoir, river basin
  - Engineering geophysical exploration (borehole geophysics, surface seismic, ERT)
- 2014 – 2019 : PhD in Civil Engineering, NCTU
  - Advisor : Professor Chih-Ping Lin
  - Research group: Geo-Imaging and Geo-Nerve Research Group
- 2010 – 2013 : BEng (Hons) in Civil Engineering, Hong Kong PolyU
- Coding experience
  - MATLAB, Python, Node-red, C and C++
  - Software-hardware integration, Raspberry Pi, Arduino, LoRa
  - PCB design @ KiCAD, easyEDA

Node-RED

Roadmap

- What is AIoT?
- Why AIoT?
- From IoT
  - Architecture
  - Sensors
  - Transmission
  - Presentation
- To AIoT
  - Server / Cloud side AIoT
  - Edge AI + IoT
- Prospects
  - Applications in various fields
  - Smart field monitoring
- Discussions

What is AIoT?



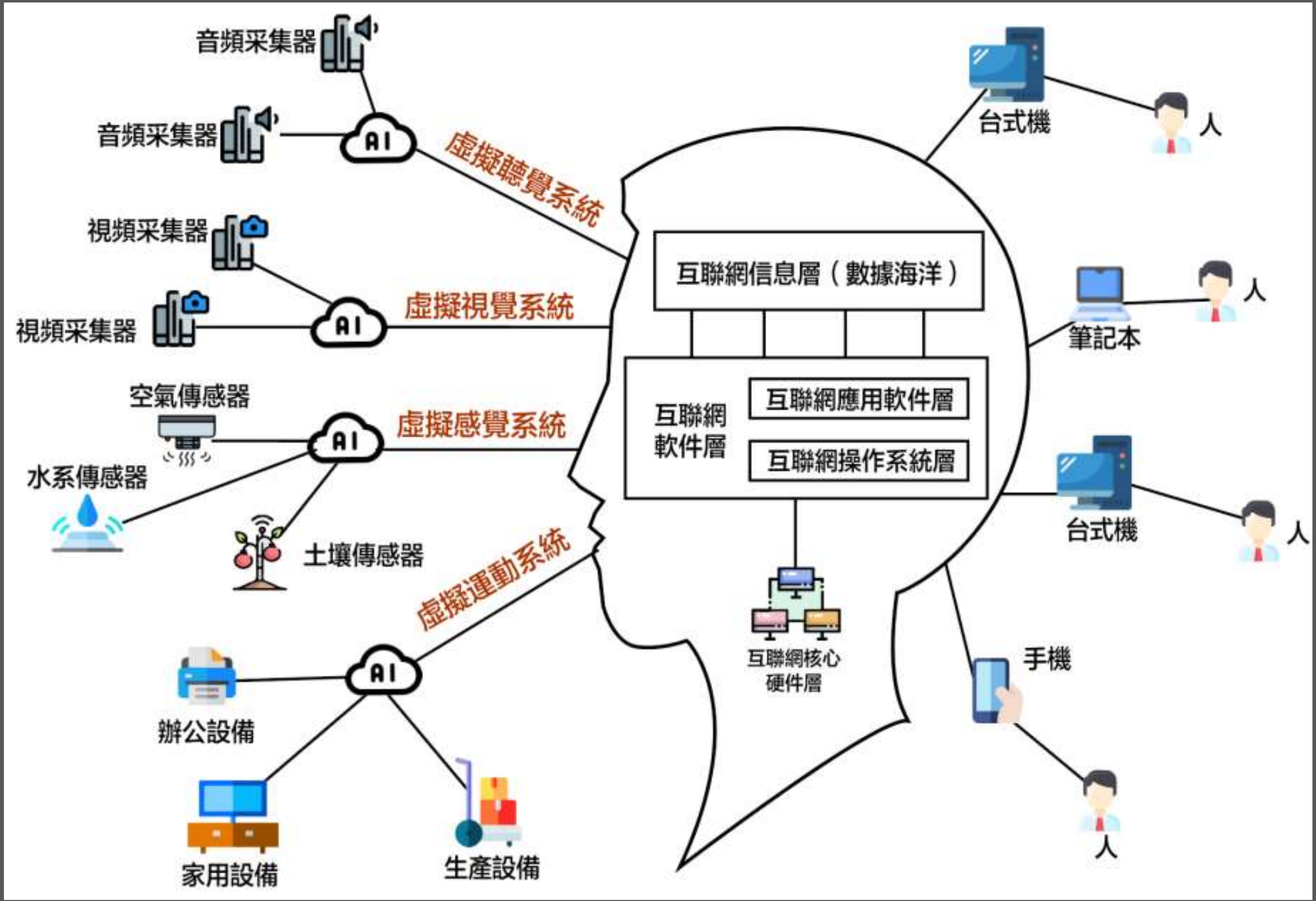
(Anant Desai, 2020)

Internet of Things + Artificial Intelligence

||

Artificial Intelligence of Things

IoT



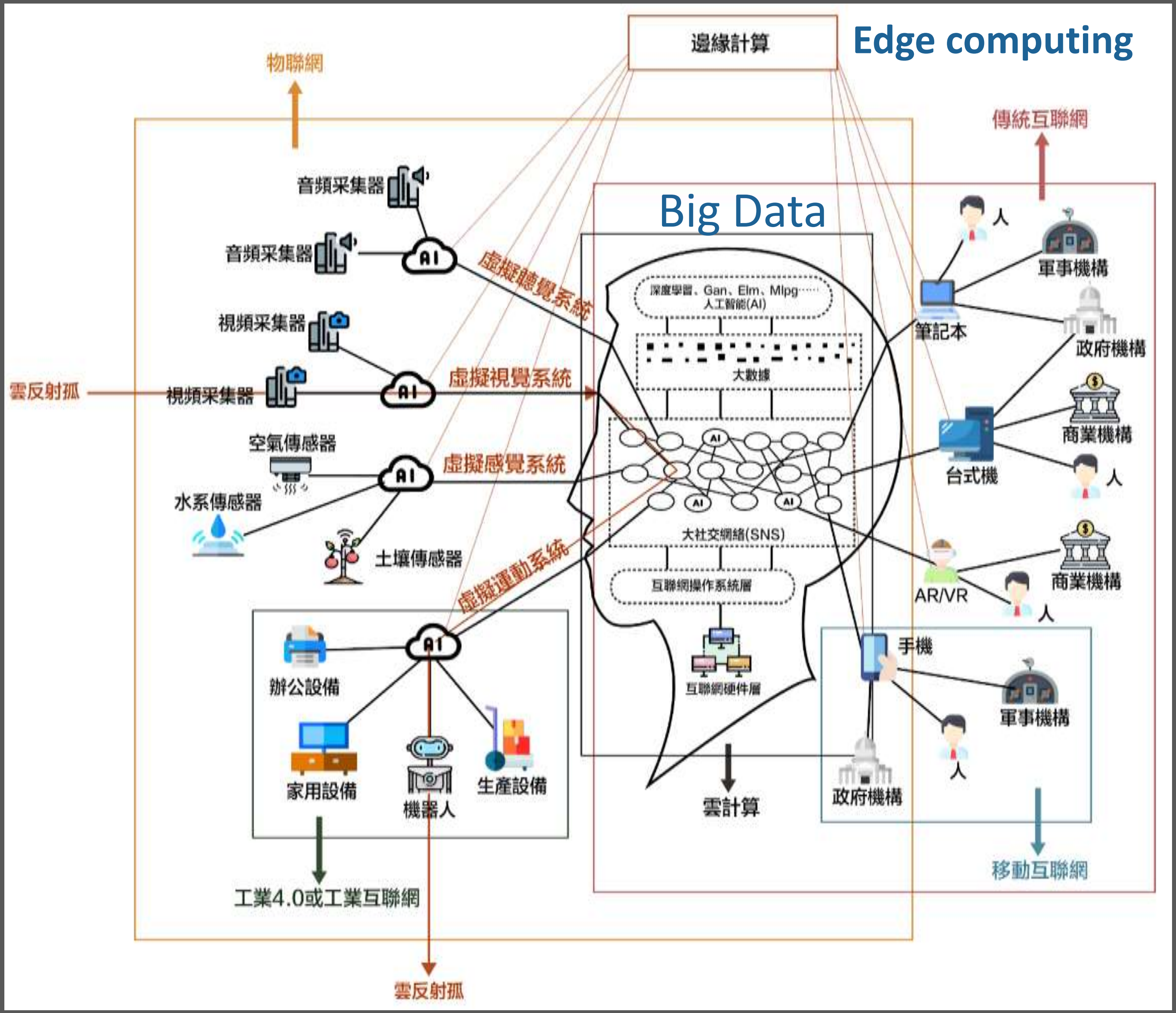
Conventional Internet

Mobile Internet

(zi-media, 2010)

IoT

Industry 4.0



Conventional Internet

Mobile Internet

(zi-media, 2021)

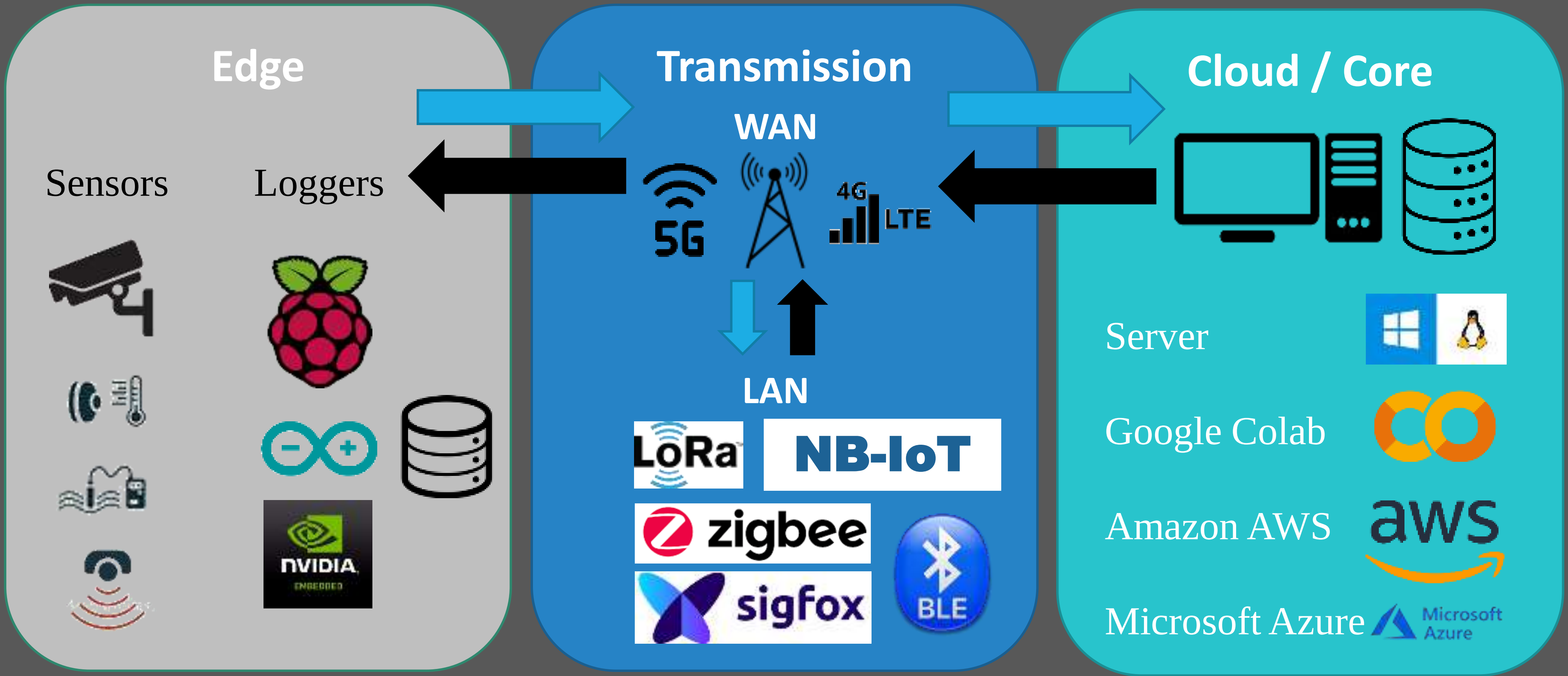
Why AIOT?

- Progression of computing technology allowed rapid data reduction and even artificial intelligence (AI) inference
- Full automation, less human effort required
  - Increase frequencies of measurement, data reduction, information interpretation
  - Reduce delay in data interpretation
  - Early detection and 24/7 monitoring
  - Cost-effectiveness in mass deployment
- Embrace unknowns through AI

Year	Papers published IoT	Papers published AIoT
2010-2015	23000	0
2016-2017	86300	645
2018-2019	164000	1290
2020	68400	1020
2021	44700	788

From IoT

IoT – Architecture

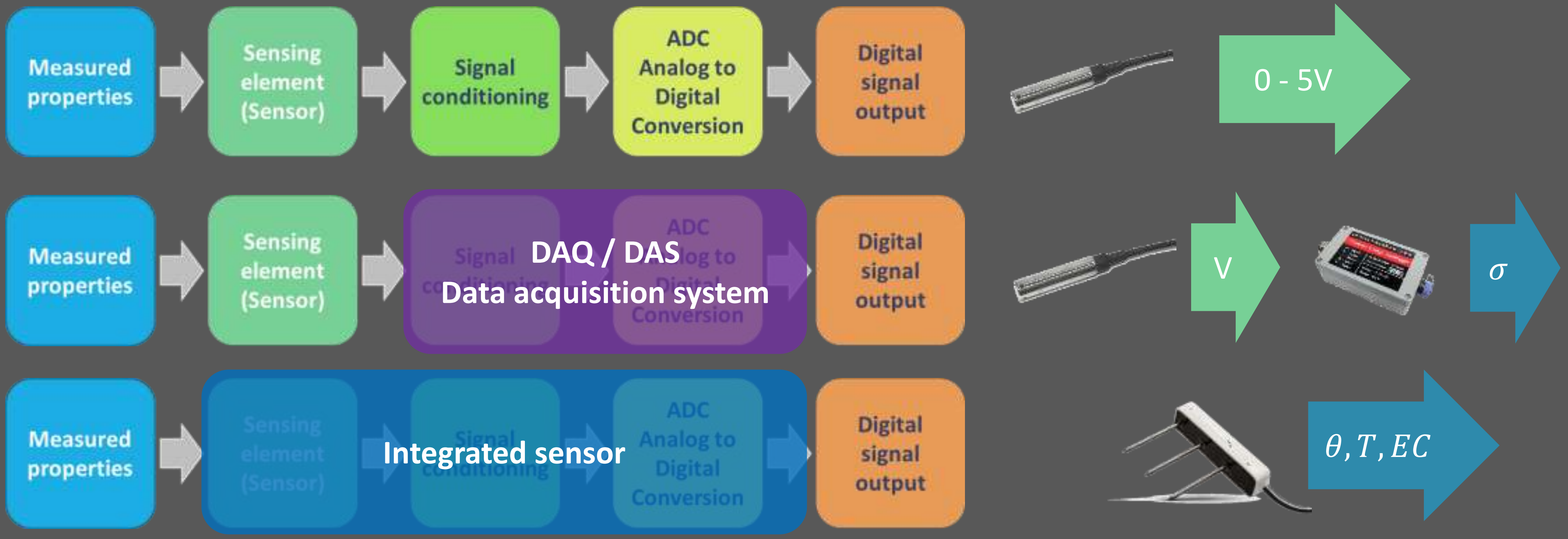


(TaiwanIoT, 2020)

IoT – Sensors

- Geotechnical applications
  - Surveillance image
  - Inclination angle
  - Water level
  - Soil moisture
  - Pressure sensor
    - Overburden / back pressure of soil
  - GPS/GNSS
  - Temperature/humidity
  - Precipitation/rainfall
- Civil engineering applications
  - Vibration sensor (structural)
  - Inclination sensor
  - Flow rate
  - Turbidity

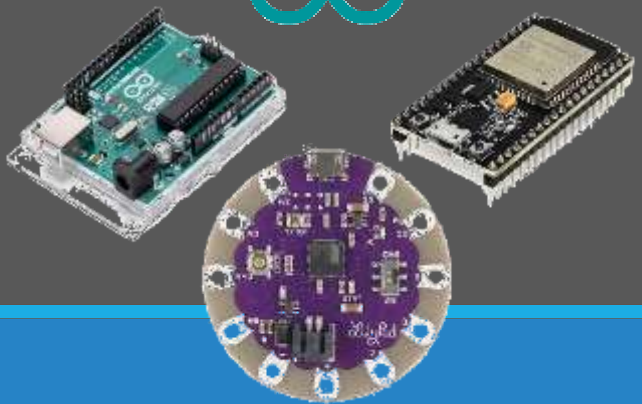
IoT – Sensors + Loggers



IoT – Data loggers

- Data logger is required to store/send acquired data
  - Micro-controllers (μC)
  - Single-board computer (SBC)
  - Embedded system (PC form)
- Ruggedness, small form factor
- Low power consumption
  - Usually 0.1W-10W

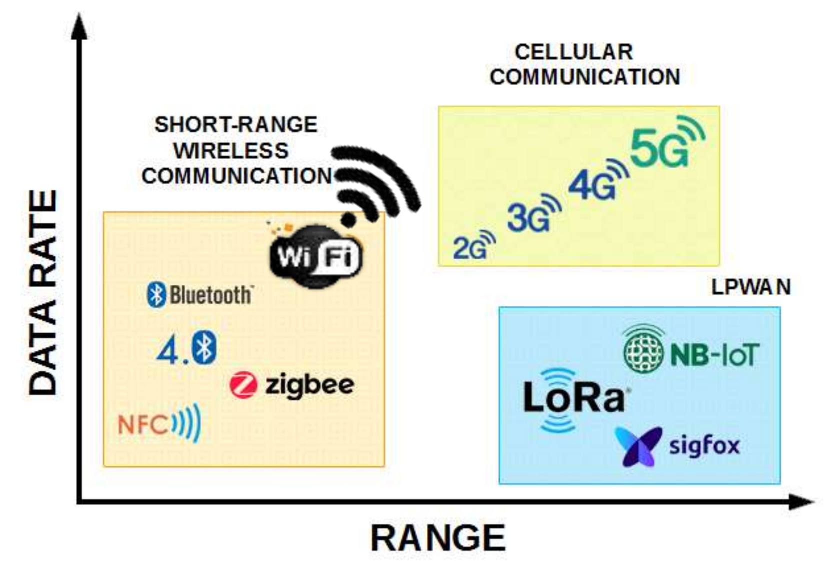
- Rich with GPIO (general purpose input/output)
  - ADC
  - Sensor communication interfaces
    - Synchronous
      - SPI : Faster, needs more wiring
      - I2C : Slower, only needs 2 wire
    - Asynchronous : UART, USB, RS-232, RS-485
      - Needs same baud rate
      - 1-to-1 communication, non-blocking, RX-TX concurrent



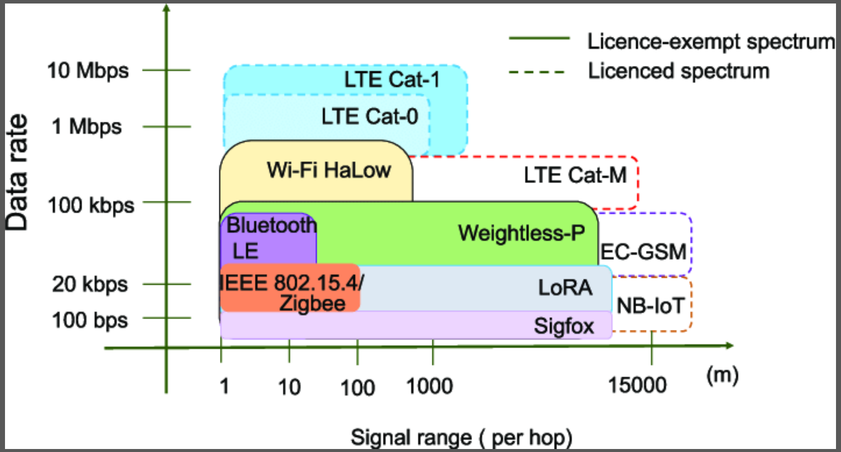
Embedded system

IoT – Transmission

- From data logger to server / cloud service
  - Connect local host/logger to centralized server
  - Involving WAN and LAN
- Some considerations for mass deployment
  - Wireless vs Wired connection
  - Low power consumption
  - Link budget (Transmission distance vs. Data rate)
  - Subscription cost
  - Security

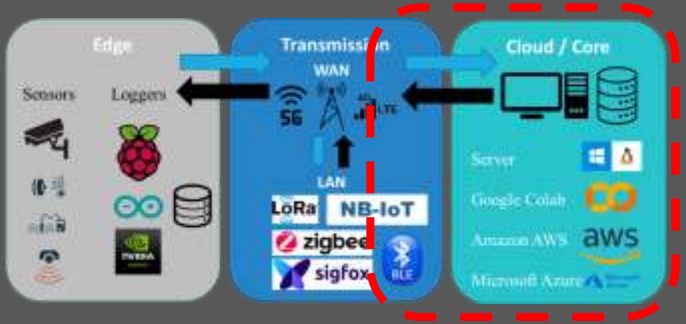


(Arun Kumar V, 2019)

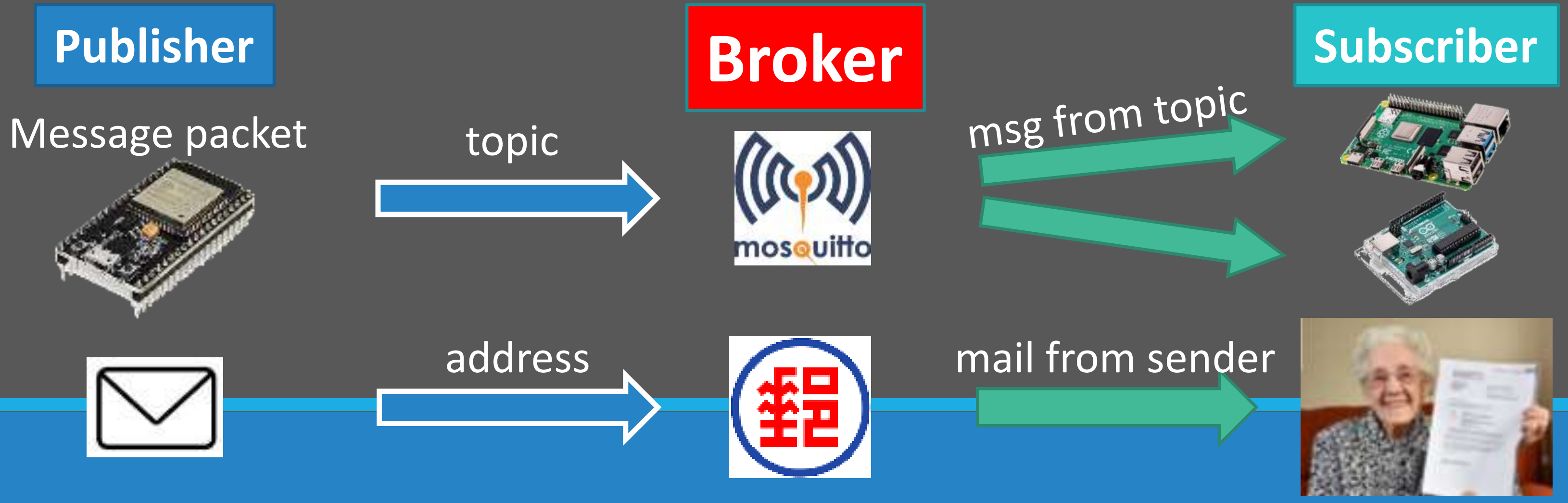


(Nguyen et. al, 2019)

IoT – Interfaces

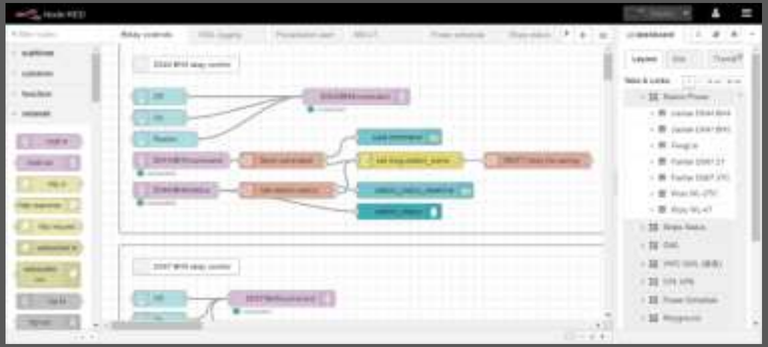


- How to communicate data into database?
- MQTT is the most popular IoT communication protocol
  - Apart from Websocket (http), CoAP, AMQP
  - File synchronization service (Dropbox, Google Drive, OneDrive etc.) is too bulky for IoT
- MQTT is analogous to a post office system



IoT – Presentation

- Presenting IoT data in a meaningful way
- Node-RED
  - Easy, rapid programming tool based on Node.js for wiring IoT components together
  - Hardware devices, APIs and online services
  - Browser-based editor with **flows** that lets user directly visualize data flow directions
  - Easy deployment on local host, device, cloud
- Or other frontend language
  - JavaScript, Python, Java, C++
- Further integration with AI
  - TensorFlow.js, machine learning



To AIoT

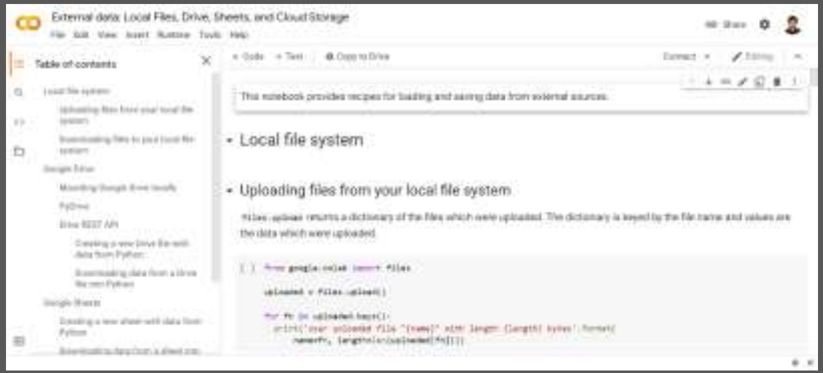
How to AIoT?

- AIoT makes IoT even more useful
  - allows user gain understanding quickly
  - deduce key information from big data
- How to incorporate AI into IoT architecture?
  - Value-added analysis at server/cloud side
  - Edge AI
- AIoT at server/cloud side
  - Deep learning/machine learning on accumulated sensor data
  - Useful information is extracted using AI models from big data
- Edge AI
  - Key info is extracted in edge systems before transferred via IoT
  - No internet is needed

(Kavita Char, 2021)

Server/Cloud AIoT

- AI analysis on IoT data stored at server/cloud services
- Train and implement deep learning/machine learning models on measured sensor data
  - Extract data patterns from big data
  - Interpret and identify potential pattern from IoT data
  - Infer possible outcome when new data arrives
- Performed on either self-hosted server or cloud services
  - Google Colab, Amazon Sagemaker, Microsoft Azure
  - Cloud services offer CPU/GPU resources for deep learning
  - Less maintenance required, pay-as-you-use
- Google Colab is popular amongst AI researcher
  - Training data can be accessed from Google Drive directly
  - Access to PyTorch, Keras, TensorFlow, and OpenCV



Edge AI + IoT

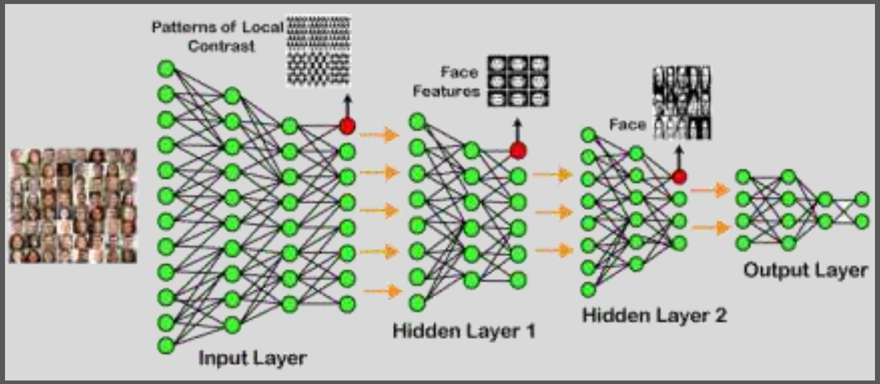
- Most AI applications ran in cloud/serve due to complexity of ML in the past
- **Why Edge AI?**
  - Transmission bandwidth for real-time image / video is too demanding
  - Requires real-time response and interpretation
  - Demand low network latency (low ping)
  - Low power, lower cost
  - Concern to data privacy and security
- Why is it possible now?
  - Higher computational capability on edge devices
  - GPU/ASIC/Neuron sticks available to speed up AI computation at the edge

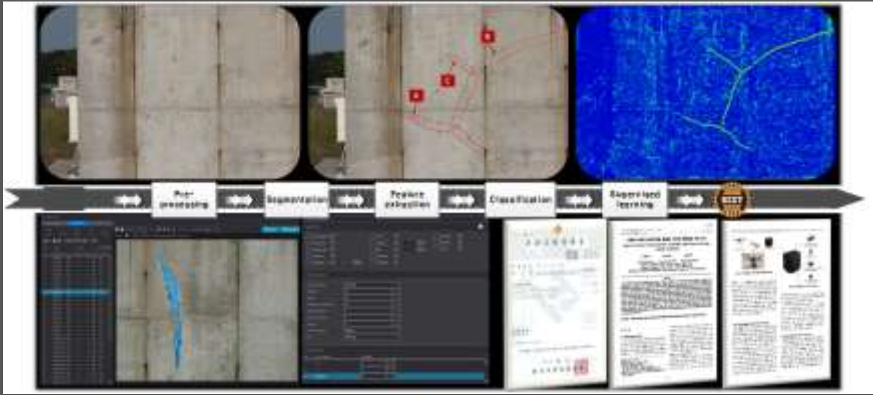
- Common applications
  - Image classification
    - Face recognition
    - Traffic control
  - Autonomous vehicle
  - Vibration analysis
  - Voice processing
  - Computer vision


Prospects

What happens from AIoT?

- Increased operational efficiency
  - AIoT process and detect patterns in real-time data that are not visible to the human eye
  - Instantaneous pattern deduction optimizes production processes and improve workflow
  - Increased efficiency and reduced operational costs
- Improved risk management
  - Risks identification in a timely manner
  - Increase safety and reduce loss
  - E.g. early detection on mechanical faults on airlines and safety risks in machineries
  - Allows for predictive maintenance
  - Reduced unplanned downtime





What happens from AIoT?

- New products and services
  - Process and draw insights from large data
  - New techniques
    - voice recognition, face recognition and predictive analysis
  - New services
    - Autonomous delivery services, smart video doorbells, voice based virtual assistants
    - Predictive maintenance for vehicles or building automation systems
    - Disaster search and rescue operations

- Enhanced / targeted customer experience
  - In retail, AIoT tailors shopping experience and gives personalized recommendations
  - Based on customer behavior, demographic information and customer

Applications

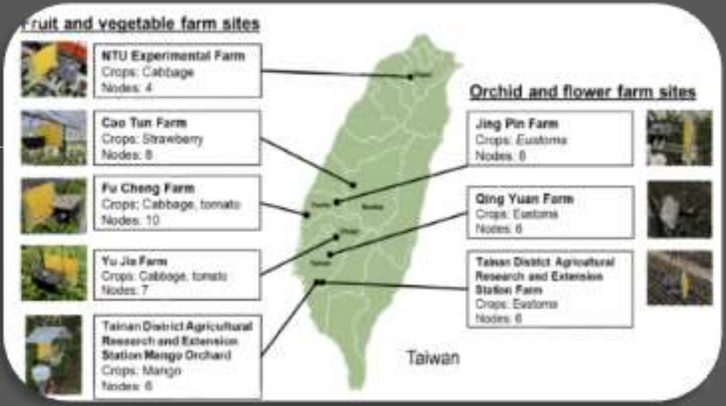
- Intelligent agriculture
- Smart home
- Crowd control
- Traffic detection
- Autonomous vehicle (self-driving cars)
- Healthcare
- Power generation
- Sediment monitoring…

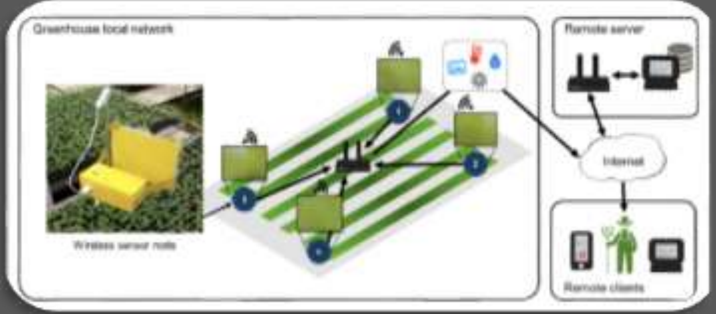
Traffic detection using Yolo v3

Tesla AutoPilot CV

Intelligent agriculture

- Agriculture is one of the earliest sector with IoT involvement, so naturally is AIoT
- Intelligent agriculture system
  - Adjustments based on collected sensor data
  - Weather, water usage, temperature and crop/soil conditions
  - From fuzzy logic to machine learning based action
- AIoT in agriculture
  - Smart management on irrigation, fertilization, pest control
  - Assist in resources utilization, yield enhancement, seasonal forecasting, crop planning
- AI + computer vision (CV) to monitor crops and large farmlands
  - Early detection of pest, intruder, hazard and so forth

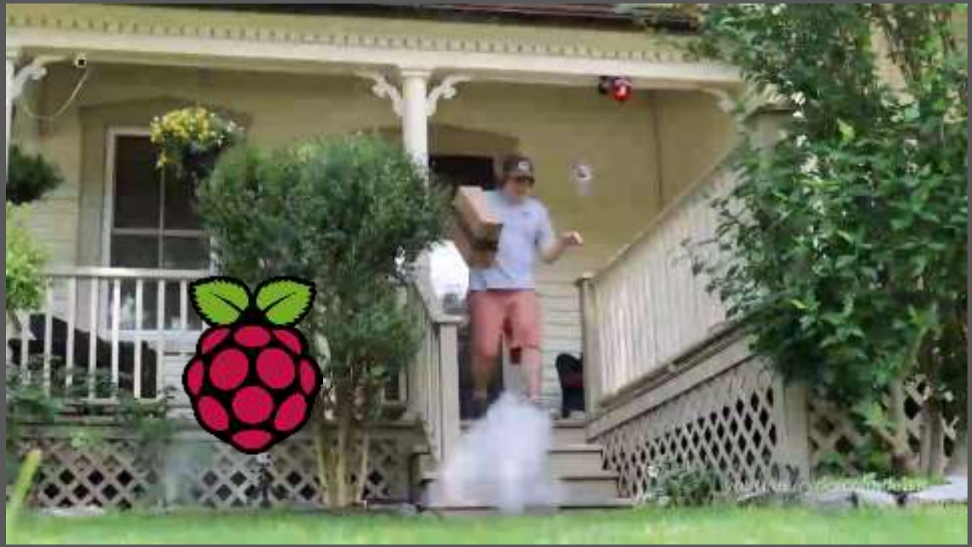




(Lin and Rustia, 2019)

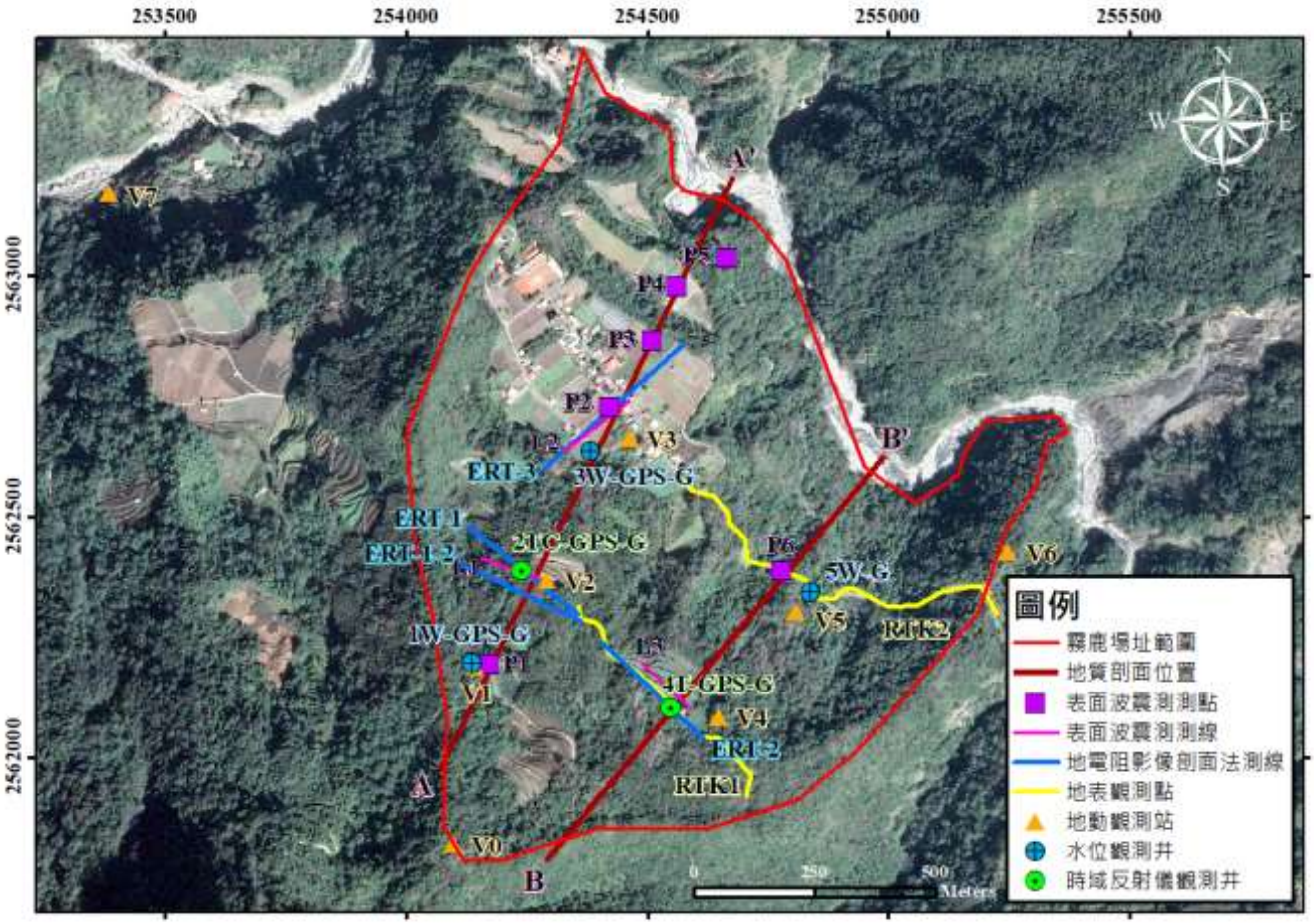
Smart home

- Home assistant
  - Open source system to home automation
  - Rich integration with node-red, MQTT, Zigbee, BLE, IKEA, Google, AWS, so much more
  - **Presence detection**, intruder alert, temperature control, power consumption …
- Closed-source/ proprietary home automation
  - HomeKit (Apple), MiJia (XiaoMi), Amazon Echo, SmartThings
- Interesting example
  - Raspberry Pi controlled intruder alert
  - Identify thieves with AI and CV
  - Custom TensorFlow model => recognize package
  - TF + Python => signal the alarm system

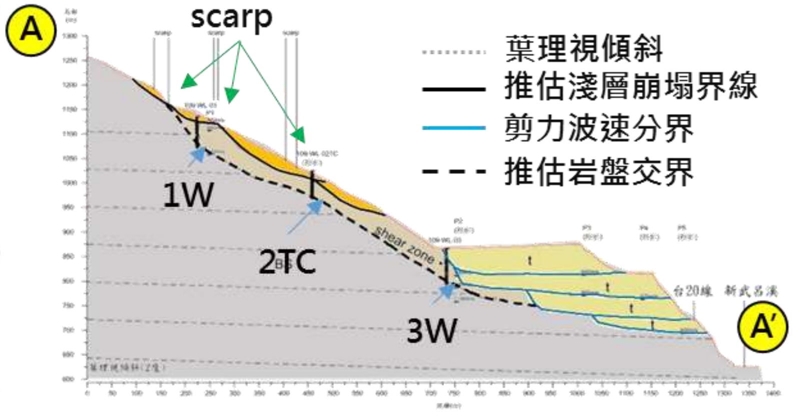


Smart field monitoring with IoT

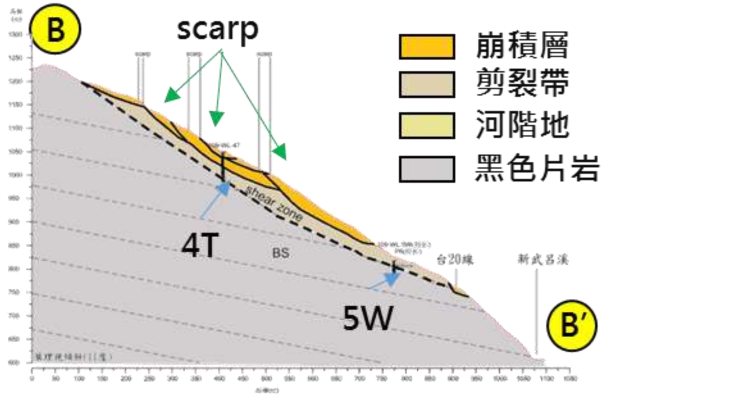
Background



253500
254000
254500
255000
255500
2563000
2562500
2562000
V7
P5
P4
P3
P2
L2
V3
ERT-3
3W-GPS-G
ERT-1
2TC-GPS-G
ERT-1-2
V2
P6
5W-G
V6
V5
RTK2
1W-GPS-G
P1
V1
L3
4T-GPS-G
V4
ERT-2
RTK1
V0
A
A'
B
B'
Meters
圖例
霧鹿場址範圍
地質剖面位置
表面波震測測點
表面波震測測線
地電阻影像剖面法測線
地表觀測點
地動觀測站
水位觀測井
時域反射儀觀測井



A
scarp
葉理視傾斜
推估淺層崩塌界線
剪力波速分界
推估岩盤交界
1W
2TC
shear zone
3W
台20線
新武呂溪
A'



B
scarp
崩積層
剪裂帶
河階地
黑色片岩
4T
BS
shear zone
5W
台20線
新武呂溪
B'

Subsurface monitoring system

- TDR : 2 monitoring hosts
- Ground water level (GWL) : 3 real-time monitoring stations (LoRa-based)
- Volumetric water content : 1 station

109-WL-2TC

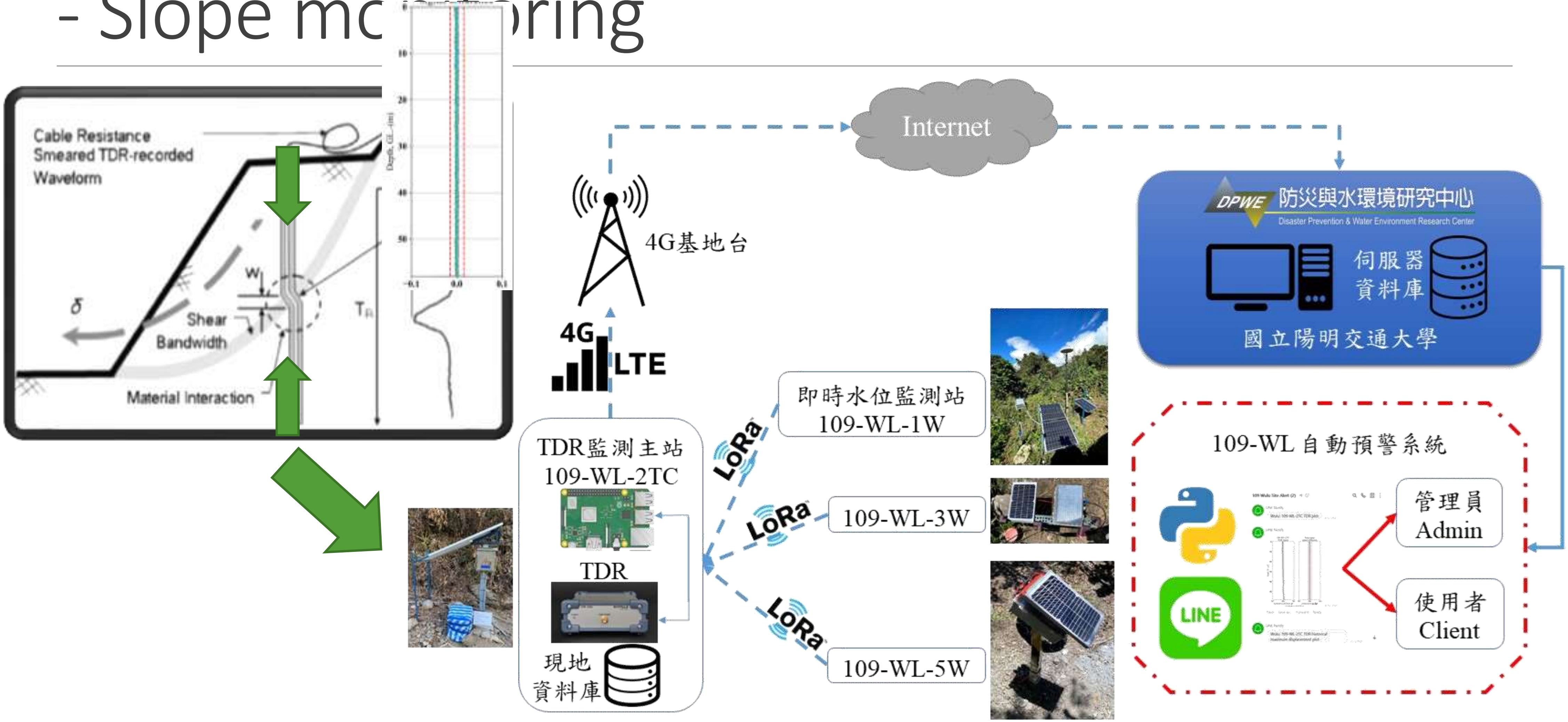
109-WL-4T

109-WL-1W

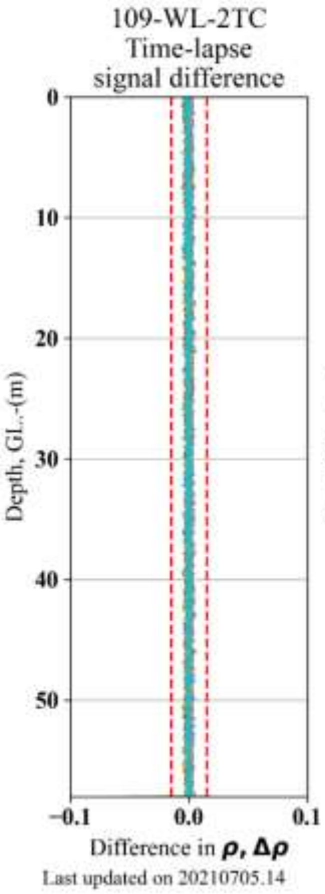
109-WL-3W

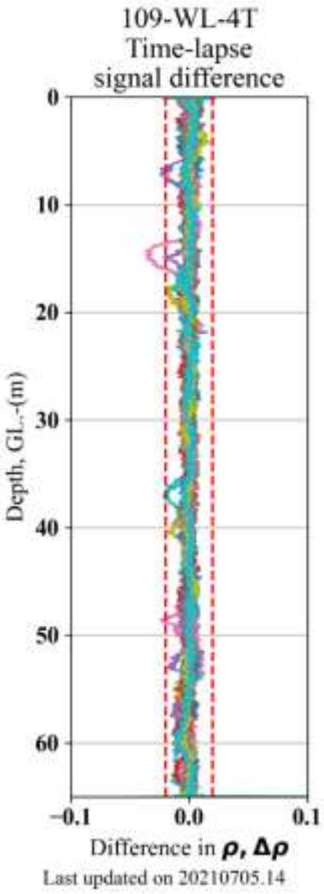
109-WL-5W

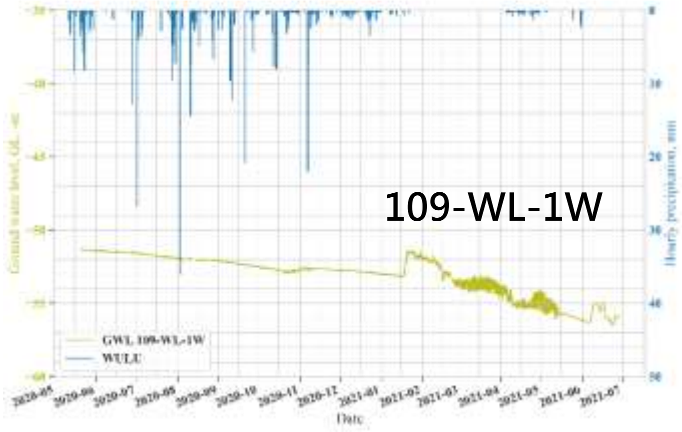
IoT architecture
- Slope monitoring

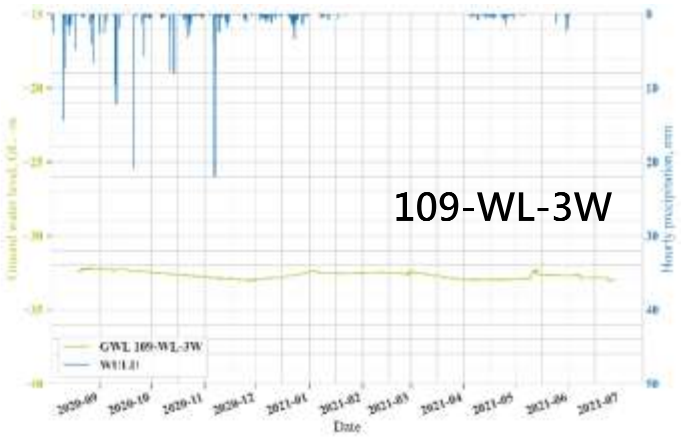


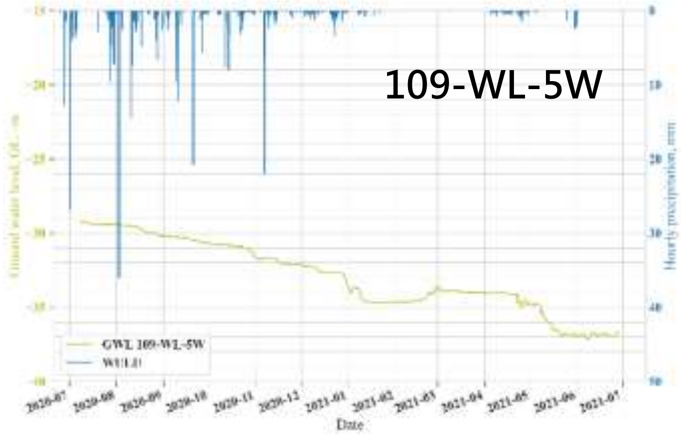
Monitoring data

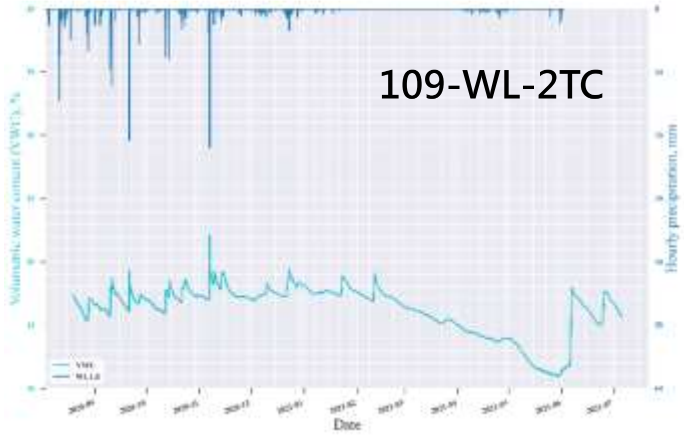




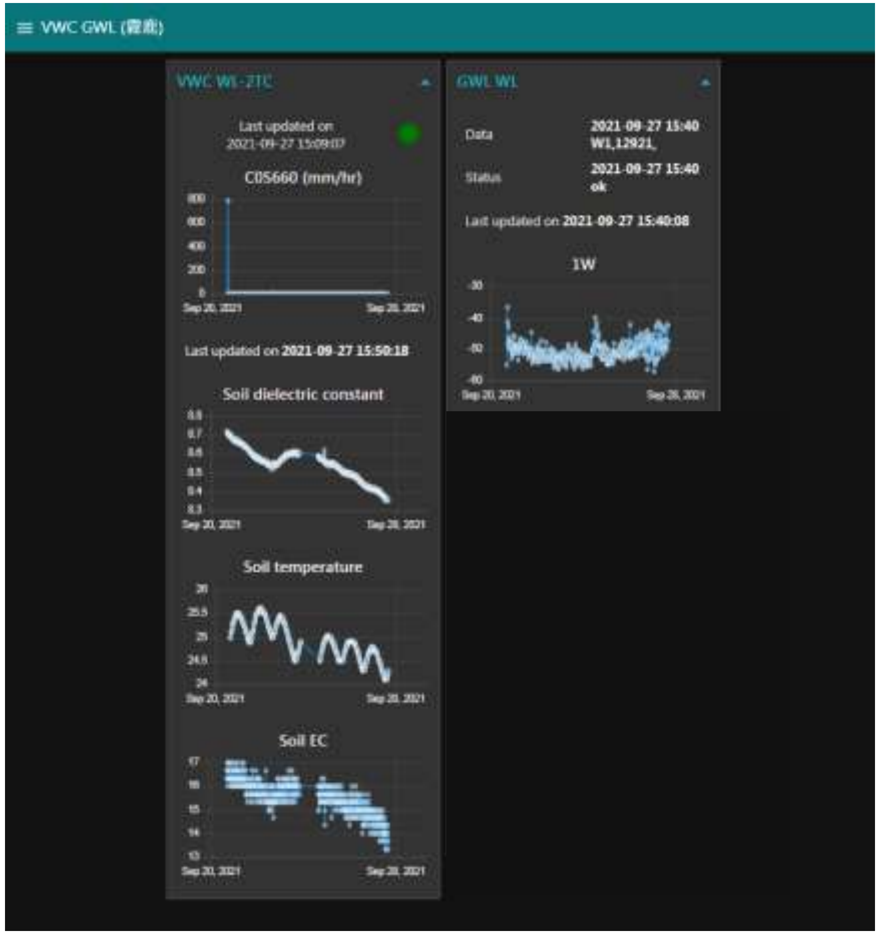








Data visualization
- Node-red flow based programming

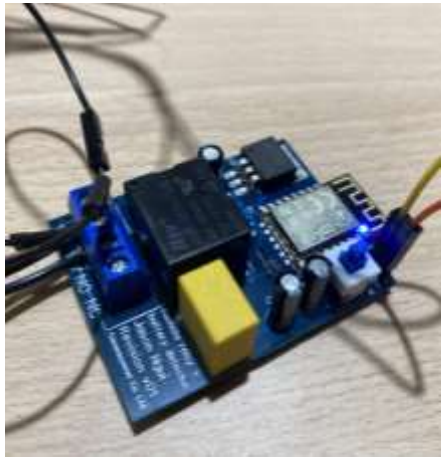


IoT architecture
- Power control

Remote control the power supply system for field stations

- Through MQTT + node-red
- Switch power physically without requiring physical presence
- Power control and status monitoring

From prototyping to custom PCB

AIoT in smart field monitoring

Based on big data gathered from precipitation, ground water level, volumetric water content, TDR signal …

- Perhaps measured physical parameters are not sufficient
- Require extra dimension of data
- Build deep learning model based on statistical model
- Neural network may see **unobserved pattern** from big data
- Currently under progress

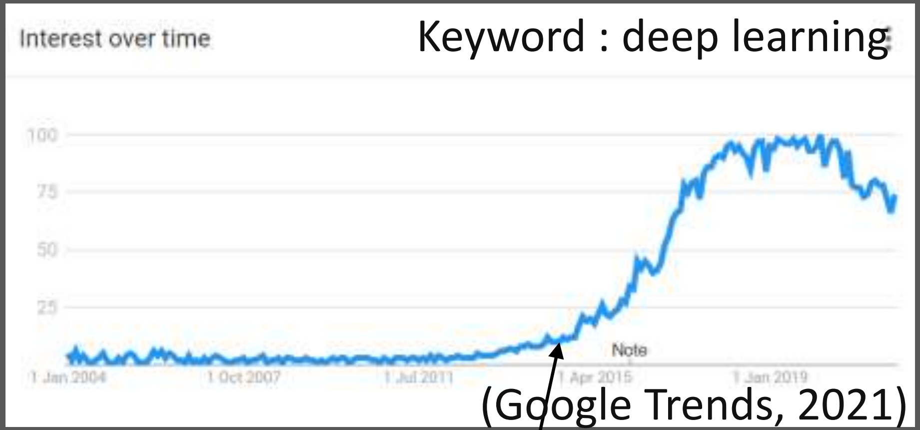
Power control for stations

- Implement power adjustment based on power generation and consumption
- **Predictive maintenance** for the health of battery / solar panel

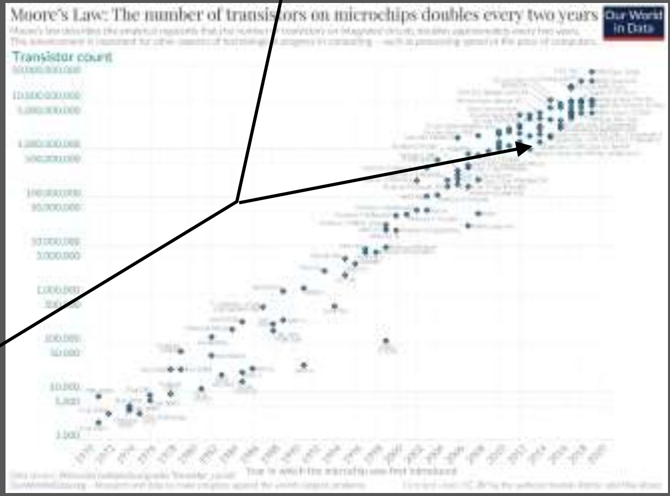
To AI or not to AI?

Background of AI

- Back in 1958, Frank Rosenblatt at Cornell design the first artificial neural network
  - Described presciently as “**Pattern-recognizing device**”
  - Era of mainframe computers filled rooms and ran on vacuum tubes
  - Inspired by the interconnections between neurons in the brain
- Limitation of computing hardware soon overcame
  - **Moore's Law** and other improvements in hardware
  - Yielded a roughly 10-million-fold increase in the number of computations that a computer could do in a second
  - Inclusion GPU in computation
- Interest in artificial intelligence (AI) is revisited in the late 2000s
  - Tools available up to the computing challenge
  - Renamed as “**deep learning**"
  - Extra layers of neurons is introduced



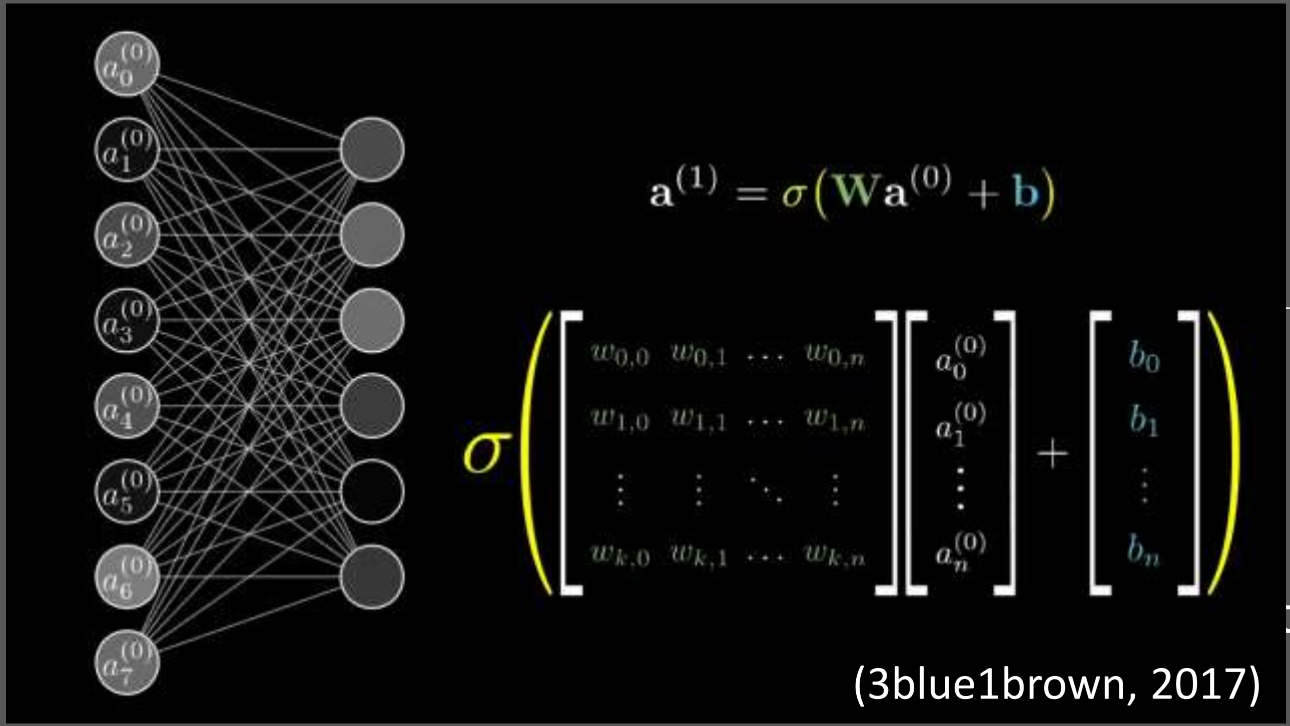
(Google Trends, 2021)



In 2014, Intel launched an even smaller, more powerful 14nm chip

Observation that the transistors amount in a dense integrated circuit (IC) doubles about every two years

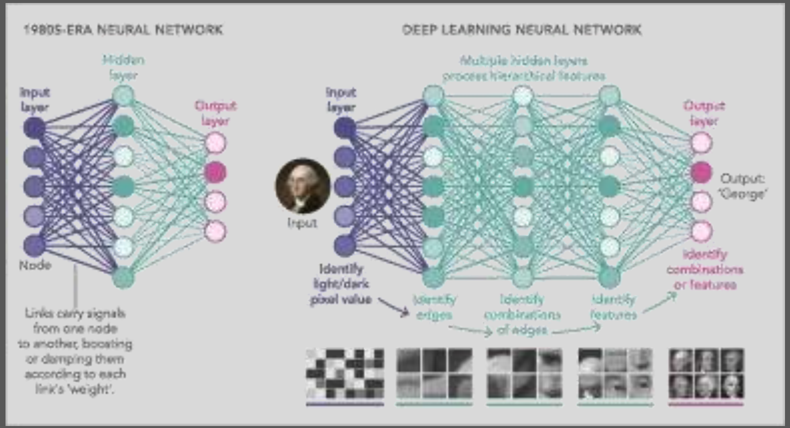
(Thompson et al., “Deep Learning’s Diminishing Returns.”)

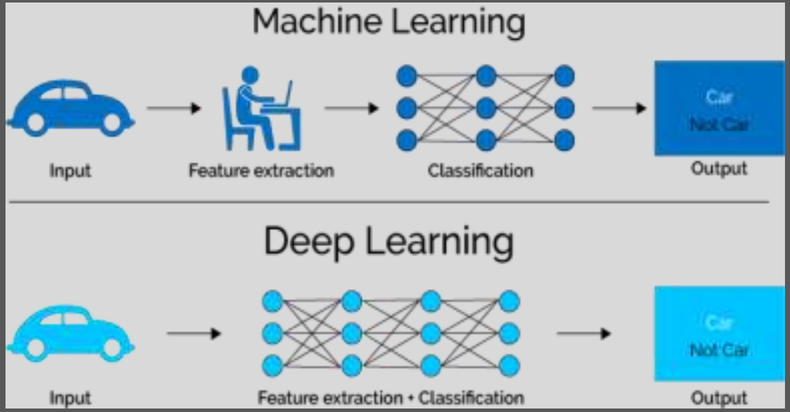


(3blue1brown, 2017)

s

t





(Khan, 2019)

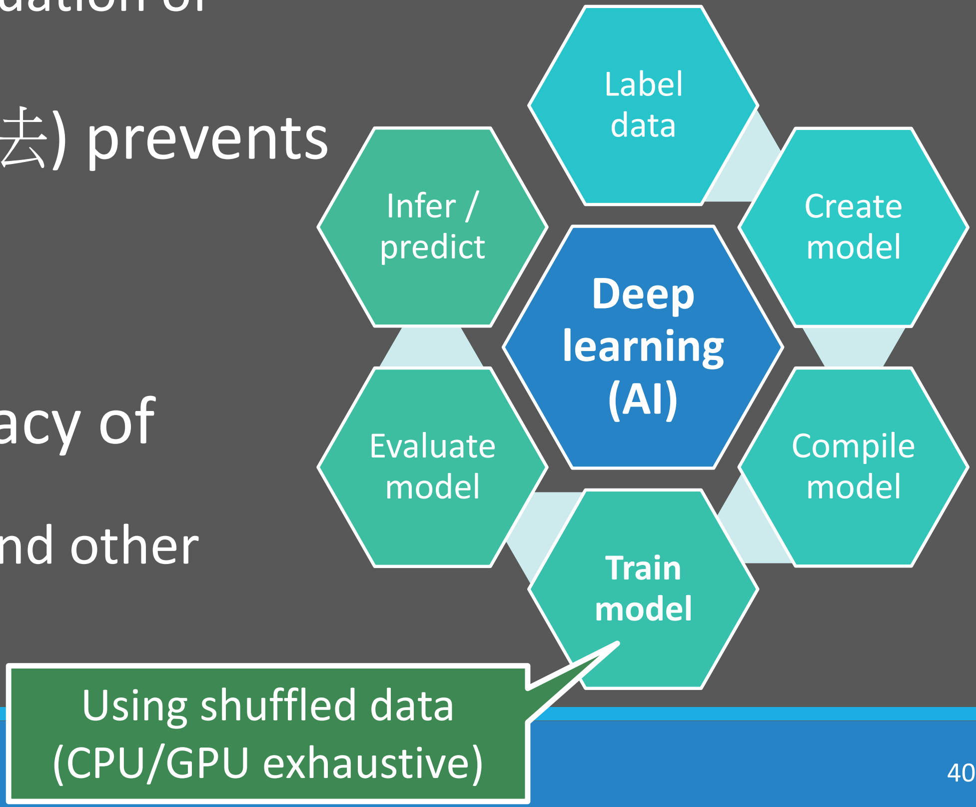
- **Flexible statistical models (most AI/DL)**
  - Require myriad combinations of priori, activation functions, bias, weighting and networks
  - Use numerous neurons (in neural networks) to train suitable AI model
  - Provide outcomes with probability [87% cat]

From expert-based to flexible model

- Key variables are pre-established in **expert-based models**
  - Priori of possible parameter value is provided
  - Done with limited computation amount, yet with reasonable accuracy
  - Popular early on, but learning ability of these models stalls if not all variable is correctly specified by the expert
- **Flexible models** (deep learning) are less efficient
  - Needs vastly more computation to match expert models
  - **But!** with enough computation (and data), **flexible models** can outperform properly established **expert-based models**

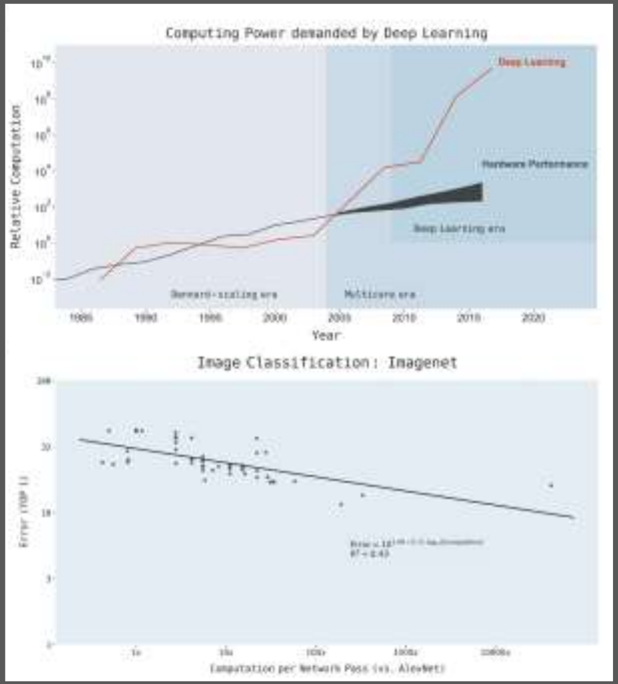
Current issue on AI implementation

- Current AI/DL are mostly over-parametrized
  - Parameter amount (unknown variables) > data amount
  - Classically, this would lead to overfitting
  - Model learns general trends and the random fluctuation of trained data
- Stochastic gradient descent (隨機梯度下降法) prevents this issue by
  - Randomly initialize parameters
  - Iteratively adjust parameter sets
- Surprisingly, this enhanced the overall accuracy of trained AI model
  - Proven in machine translation, object detection, and other general AI models



Cost to higher accuracy

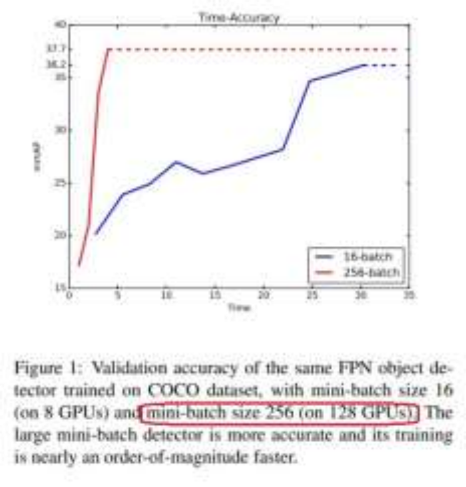
- Improvement cost for deep learning is high
  - Lower error rate by half, expect at least **500x** the current computational resources
  - 2012 AlexNet : 2 GPU @ 5-6days
  - 2018 NASNet-A : Half the error rate of AlexNet, but more than 1,000 times computing power
    - In 1,000-fold difference, only 6-fold improvement came from better hardware
    - The rest : more processors / running longer, incurring higher costs
- Training a model with 5% error rate would require 10^{19} billion floating-point operations
  - Cost US $100 billion (NT$2,781,720,000,000=2781億元) and produce carbon emissions equal to New York City in a month
- Google subsidiary **DeepMind** trained its system to play **Go**
  - Estimated cost $35 million



(Thompson et al., 2020)

Silver lining

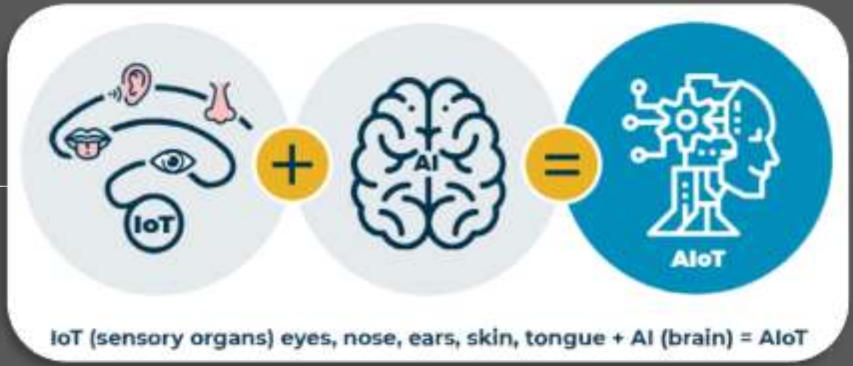
- **Increasing computing power**: Hardware accelerators
  - Already in effect: FPGA, ASIC (Google's TPU), GPU instead of CPU
  - Fundamentally, they sacrifice generality of the computing platform for efficiency of increased specialization
  - Longer-term gains will require adopting wholly different hardware frameworks
  - Perhaps hardware based on analog, neuromorphic, optical, or quantum systems
- **Reducing computational complexity**: Network Compression and Acceleration
  - Pruning away weights, quantizing the network, or using low-rank compression
  - Reduce floating point operations in evaluation
- **Integrated expert-model + AI model**
  - Or other under-appreciated machine learning models



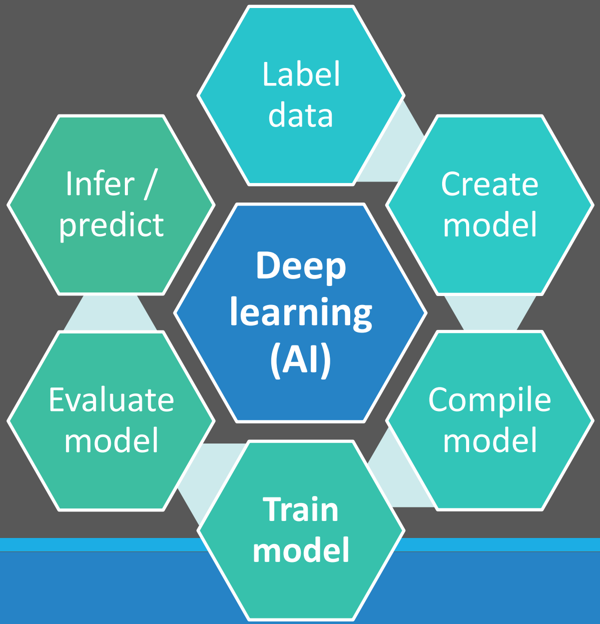
(YOLOv4)

(IBM quantum computer)

Takeaways



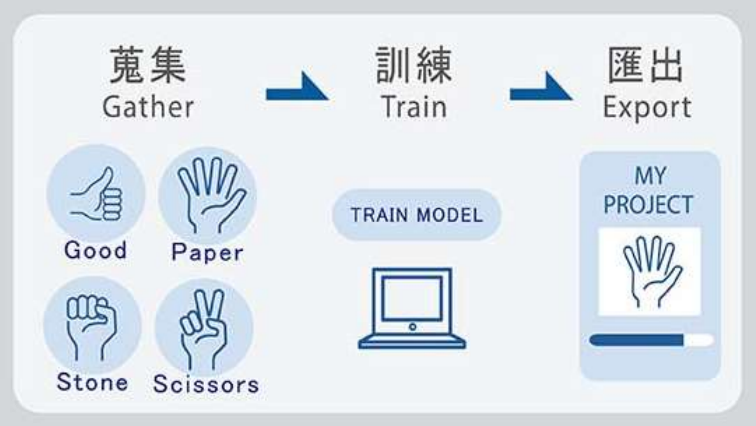
- Embrace AI in future researches
  - AI is merely a large statistical model
  - Addition of expert knowledge into AI model is the **know-how**
- Go towards open source community for robust programming
- TPU > GPU >> CPU



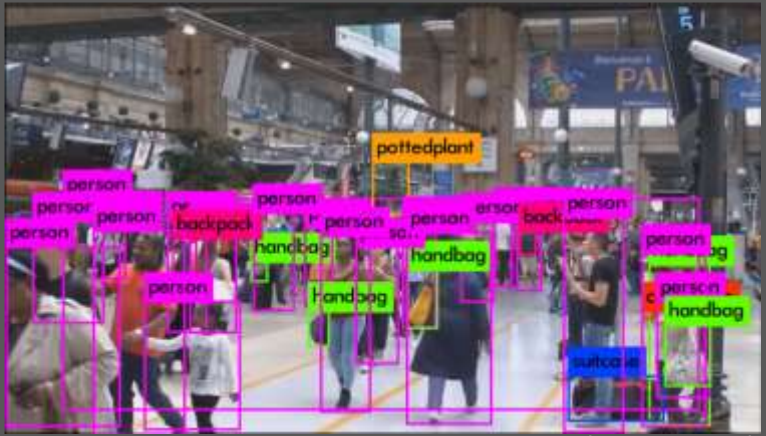
(Mythbusters, 2009)

Interested?

- Start from IoT first!
  - Node-red + MQTT + Arduino
- Check out
  - Teachable machine : hosted by Google
  - Tensorflow : Handwriting recognition
  - YOLOv3/v4 (You Only Look Once, Version 3/4)
    - Real-time object detection algorithm that identifies specific objects in videos, live feeds, or images.
    - Identifies 80 object types
  - TinyML
    - suitable for Edge AI in small MCU
    - Arduino Nano 33 BLE Sense, ESP32…



(Teachable machine)



(YOLOv4)

(Bochkovskiy et al., 2020)

End of presentation

THANK YOU!

Bibliography

- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv:2004.10934 [Cs, Eess]*. http://arxiv.org/abs/2004.10934
- Lin and Rustia (2019, November 29). *Trends of AIoT Application in Smart Agriculture*. FFTC Agricultural Policy Platform (FFTC-AP). https://ap.fftc.org.tw/article/1636
- *How to Build a Raspberry Pi Alarm that Sprays Porch Pirates | Tom's Hardware*. (n.d.). Retrieved September 30, 2021, from https://www.tomshardware.com/how-to/raspberry-pi-porch-pirate-alarm
- Strubell, E., Ganesh, A., and McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. *ArXiv:1906.02243 [Cs]*. http://arxiv.org/abs/1906.02243
- Thompson, N. C., KRISTJAN GREENEWALD, KEEHEON LEE, and GABRIEL F. MANSO. (2021, September 24). *Deep Learning's Diminishing Returns*. IEEE Spectrum. https://spectrum.ieee.org/deep-learning-computational-cost

Discussion

ngui@nycu.edu.tw