



# EdgeAI / TinyML

Bringing intelligence to sensors

Prof. Marcelo José Rovai  
[rovai@unifei.edu.br](mailto:rovai@unifei.edu.br)

UNIFEI - Universidade Federal de Itajubá, Brazil



**“Edge AI** is a truly complete technology. As a topic, it makes use of knowledge from everything from the physical properties of semiconductor electronics all the way up to the engineering of high-level architectures that span devices and the cloud. It demands expertise in the most cutting-edge approaches to artificial intelligence and machine learning along with the most venerable skills of bare-metal embedded software engineering. It makes use of the entire history of computer science and electrical engineering, laid out end to end.”



Situnayake, Daniel; Plunkett, Jenny  
**AI at the Edge** (pp. 215-216)  
O'Reilly Media

**Marcelo Rovai** was born in São Paulo and holds a Master's degree in Data Science from the Universidad del Desarrollo (UDD) in Chile and an MBA from IBMEC (INSPER) in Brazil. He graduated in 1982 as an Engineer from UNIFEI, Federal University of Itajubá, with a specialization from Escola Politécnica de Engenharia of São Paulo University (USP), both institutions located in Brazil.

Rovai has experience as a teacher, engineer, and executive in several technology companies such as CDT/ETEP, AVIBRAS Aeroespacial, SID Informática, ATT-GIS, NCR, DELL, COMPAQ (HP), and more recently at IGT as a VP and a Senior Advisor for Latin America.

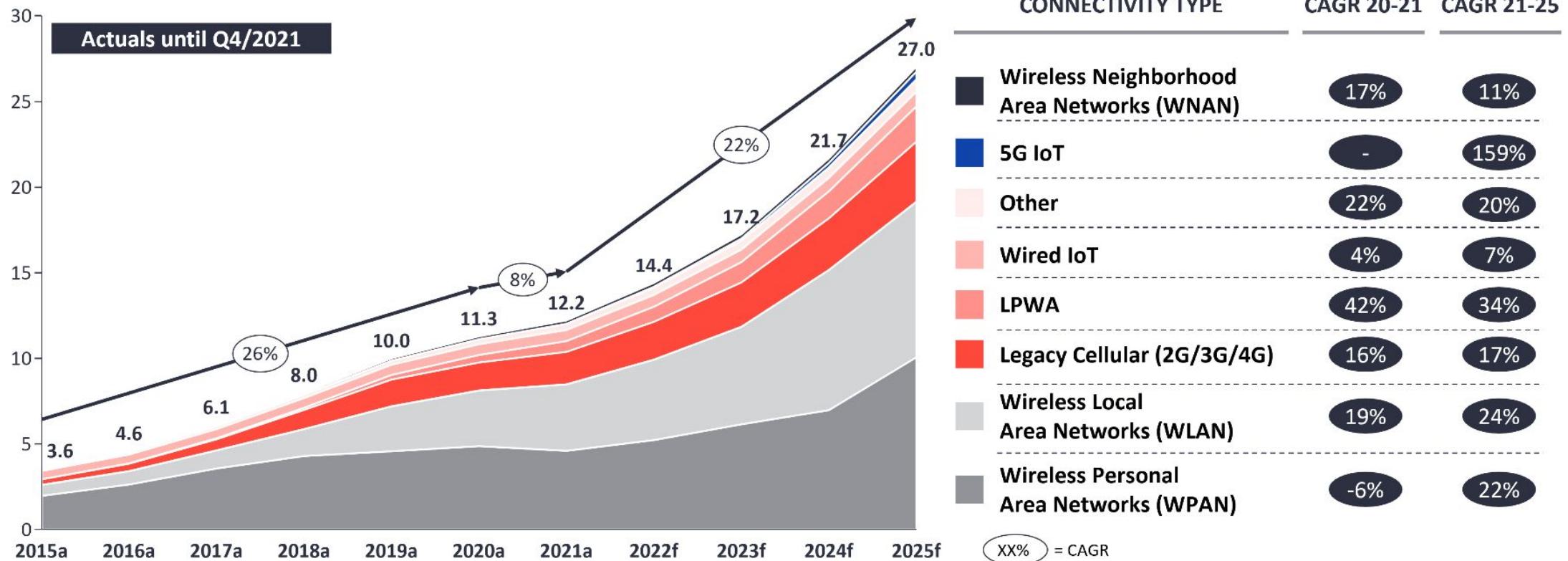
Marcelo Rovai publishes articles about electronics on websites such as [MJRoBot.org](#), [Hackster.io](#), [Instructables.com](#), and [Medium.com](#). Furthermore, he is a volunteer Professor at the UNIFEI in Brazil and a lecturer at several Congresses and Universities on IoT and TinyML. He is an active member and a Co-Chair of the [TinyML4D](#) group, an initiative to bring TinyML education to developing countries.



# Internet of Things (IoT)

# Global IoT Market Forecast [in billion connected IoT devices]

Number of global active IoT Connections (installed base) in Bn

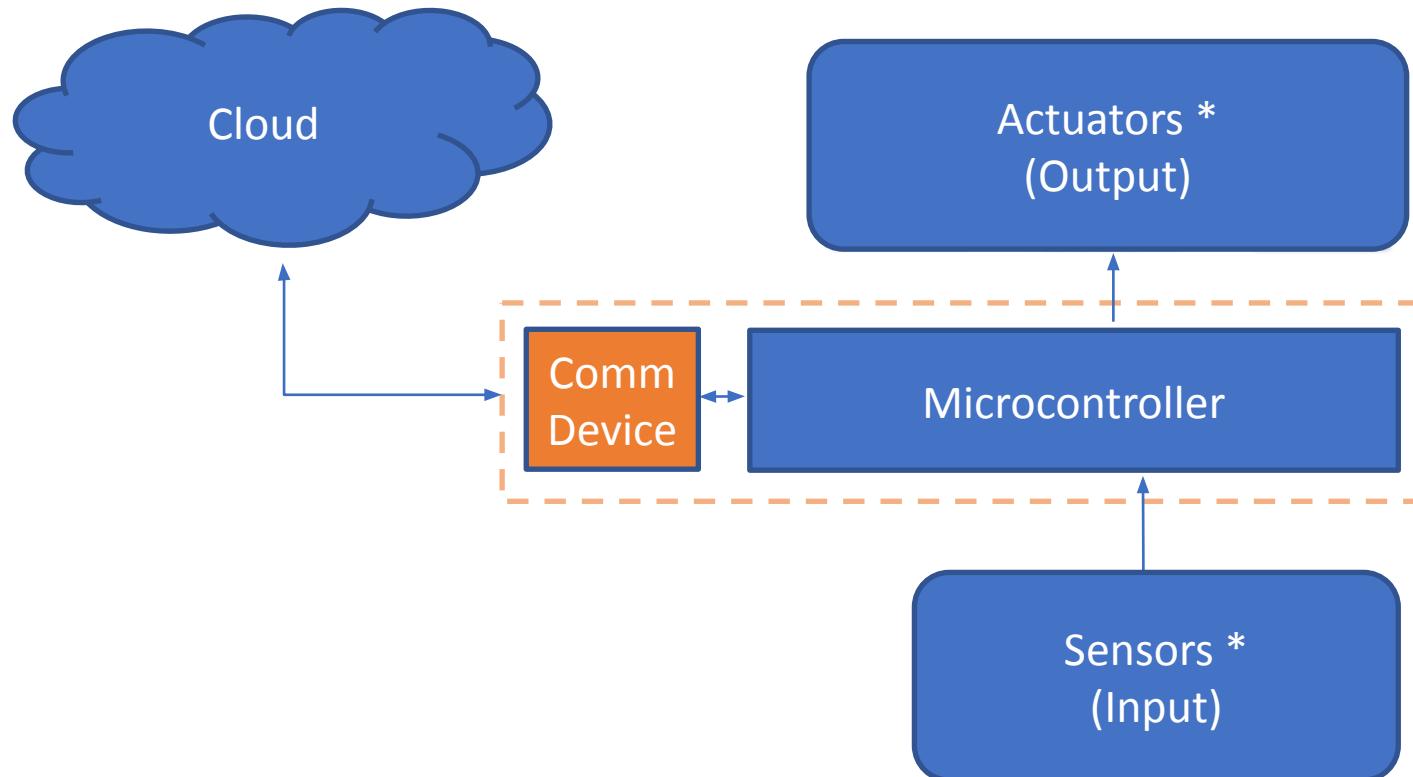


**Note:** IoT Connections do not include any computers, laptops, fixed phones, cellphones or tablets. Counted are active nodes/devices or gateways that concentrate the end-sensors, not every sensor/actuator. Simple one-directional communications technology not considered (e.g., RFID, NFC). Wired includes Ethernet and Fieldbuses (e.g., connected industrial PLCs or I/O modules); Cellular includes 2G, 3G, 4G; LPWAN includes unlicensed and licensed low-power networks; WPAN includes Bluetooth, Zigbee, Z-Wave or similar; WLAN includes Wi-fi and related protocols; WMAN includes non-short range mesh, such as Wi-SUN; Other includes satellite and unclassified proprietary networks with any range.

**Source:** IoT Analytics Research 2022. We welcome republishing of images but ask for source citation with a link to the original post and company website.

<https://iot-analytics.com/number-connected-iot-devices>

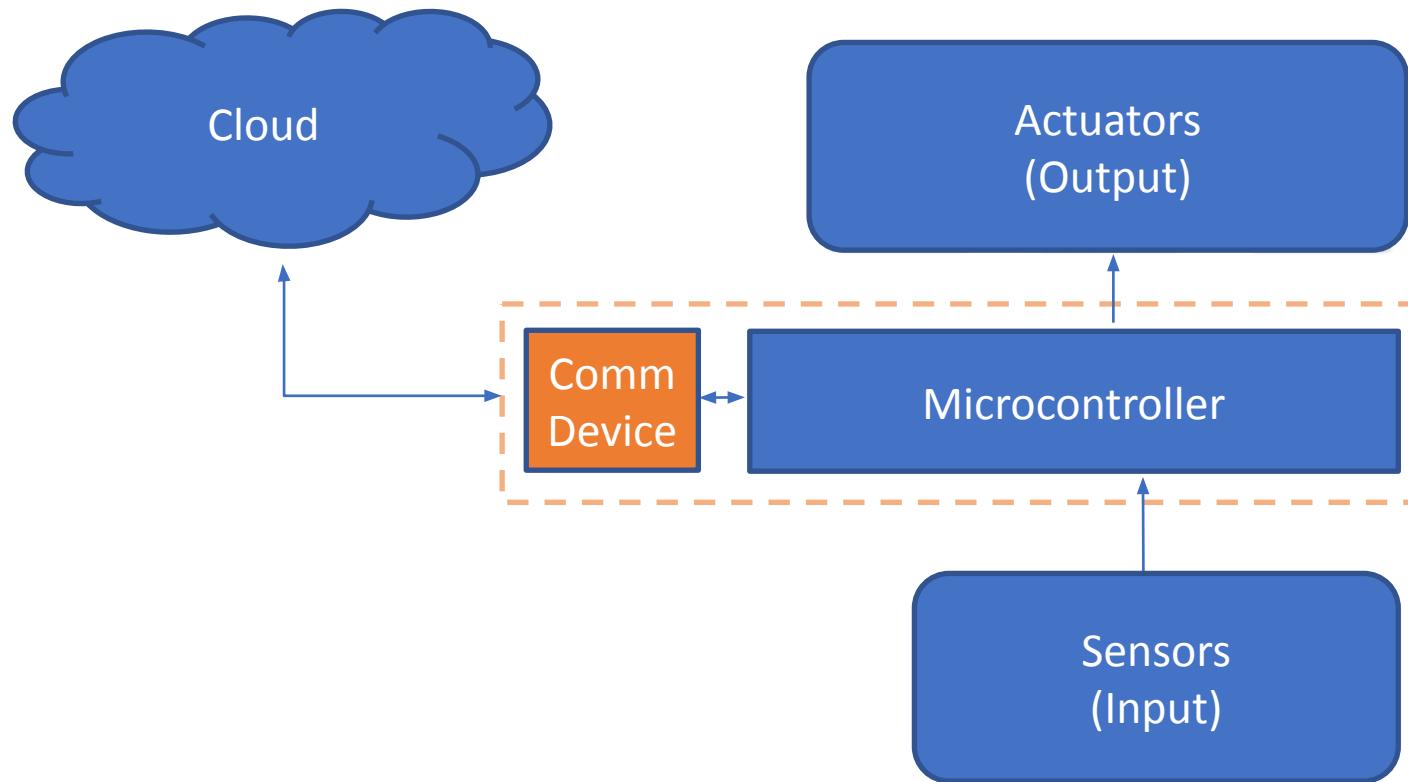
# Typical IoT Project



\* “Things”



# Typical IoT Project



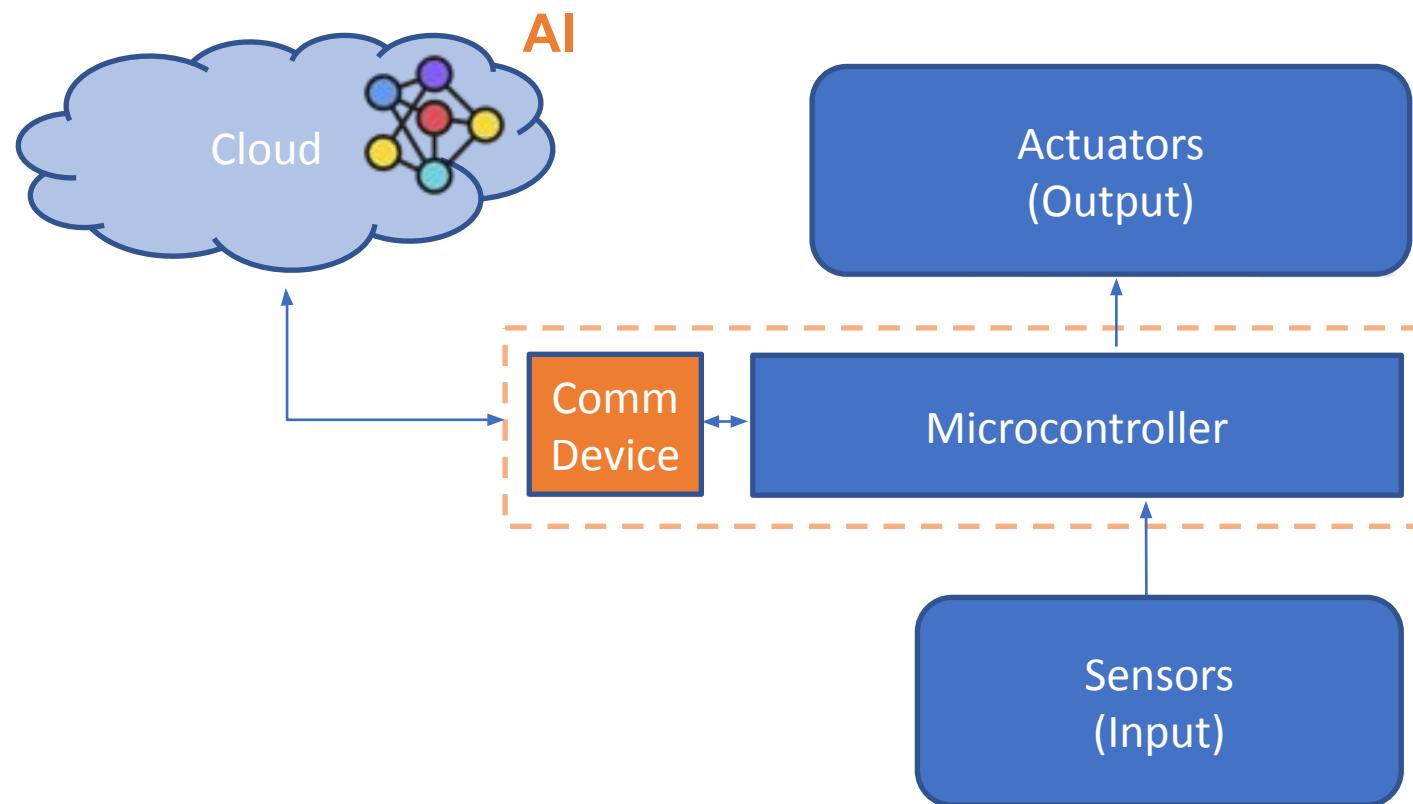
**5 Quintillion**  
bytes of data produced  
every day by IoT

**<1%**  
of unstructured data is  
analyzed or used at all

Source: Harvard Business Review, [What's Your Data Strategy?](#), April 18, 2017

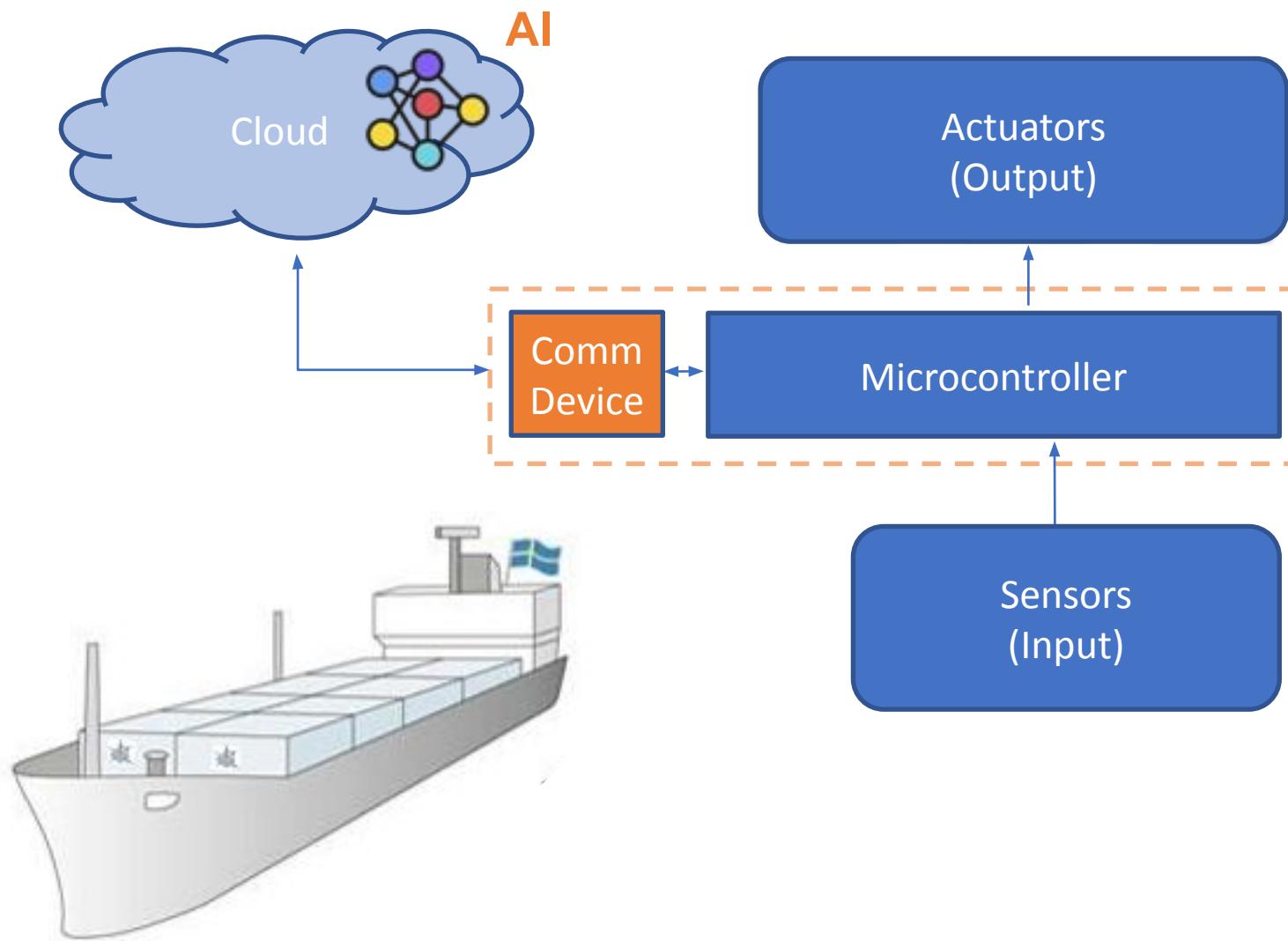
Cisco, [Internet of Things \(IoT\) Data Continues to Explode Exponentially. Who Is Using That Data and How?](#), Feb 5, 2018

# Typical AIoT Project



# Typical AIoT Project ...

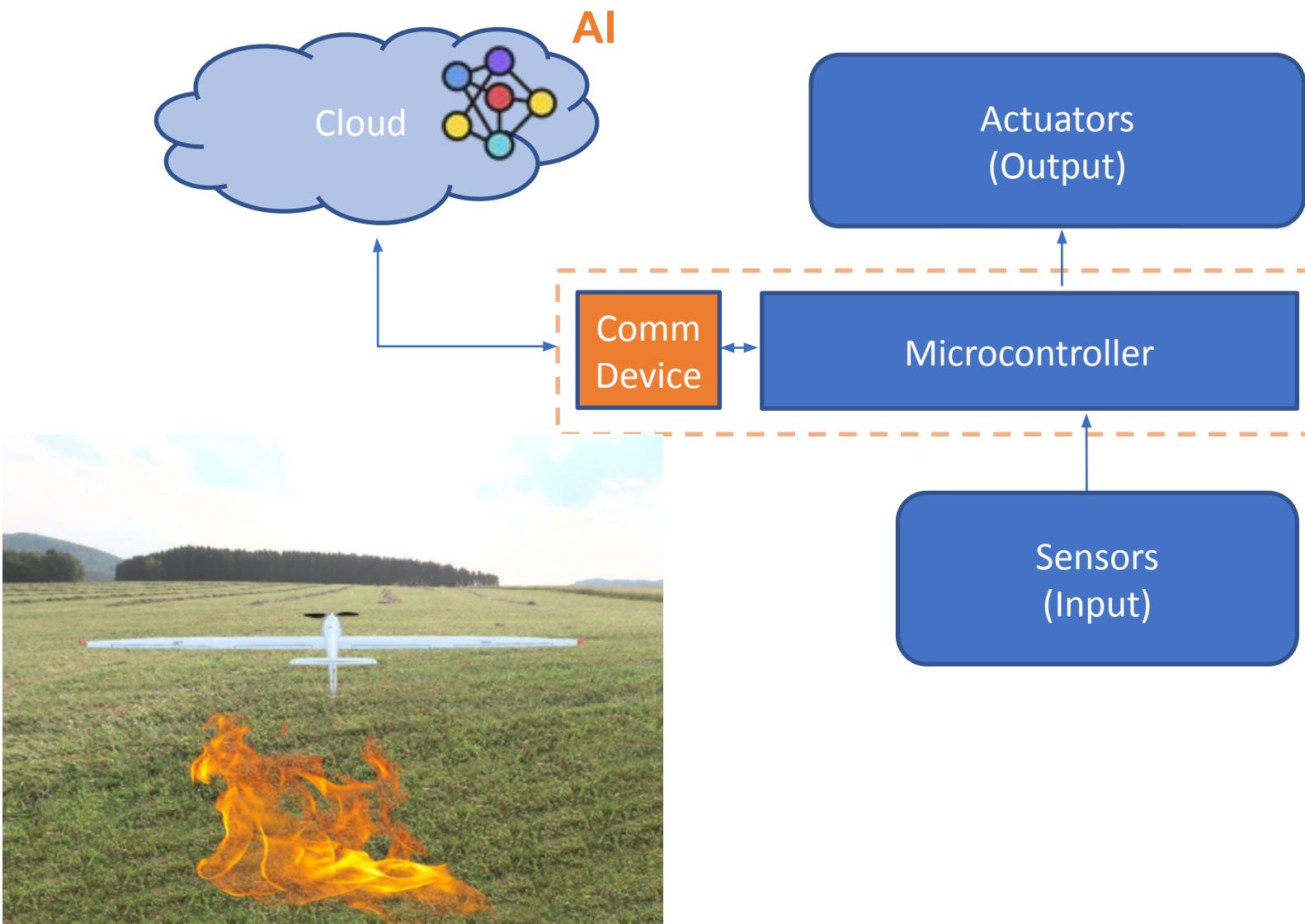
... Issues



Bandwidth

# Typical AIoT Project ...

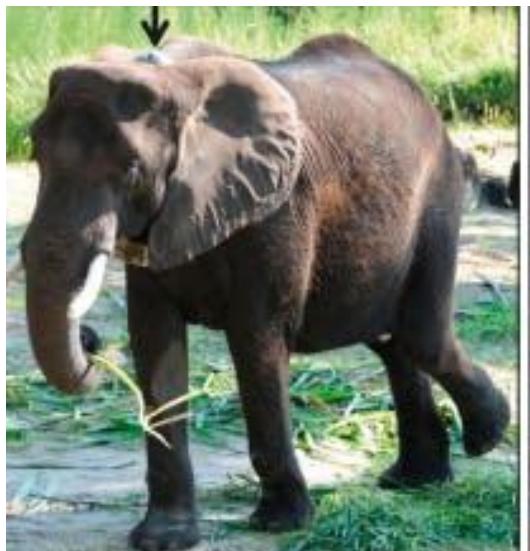
... Issues



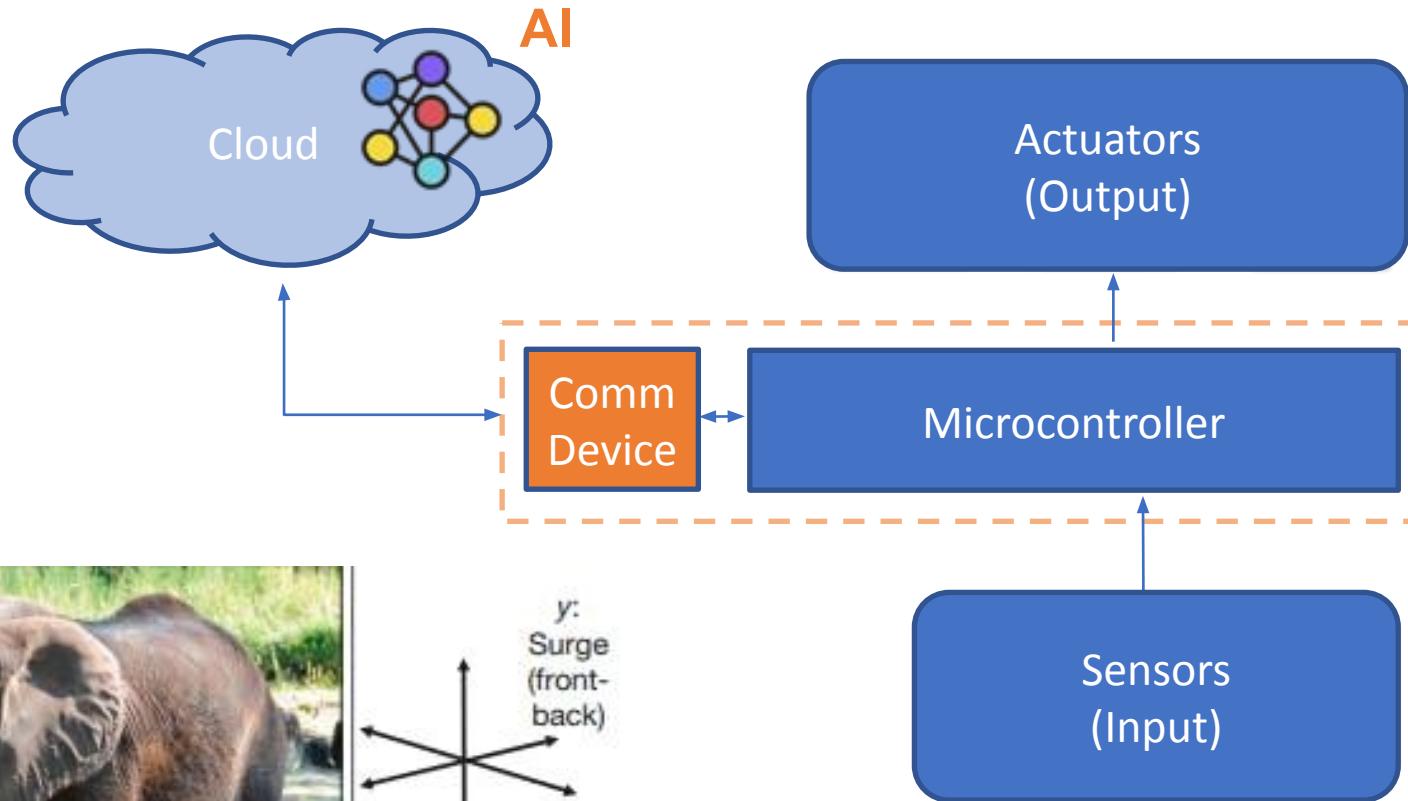
Bandwidth  
Latency

# Typical AIoT Project ...

# ... Issues



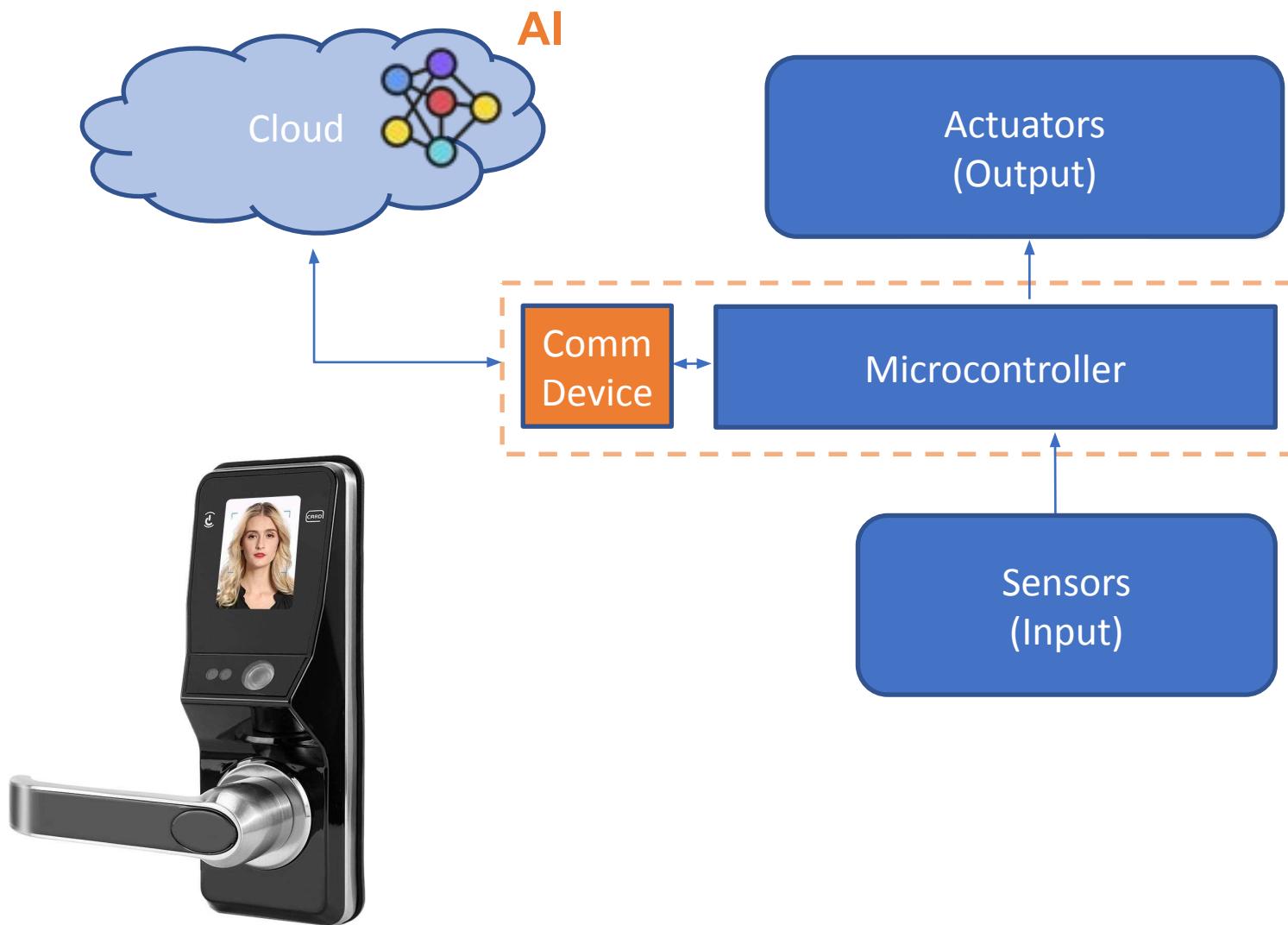
$y$ : Surge (front-back)  
 $x$ : Sway (lateral)  
 $z$ : Heave (up-down)



Bandwidth  
Latency  
Energy

# Typical AIoT Project ...

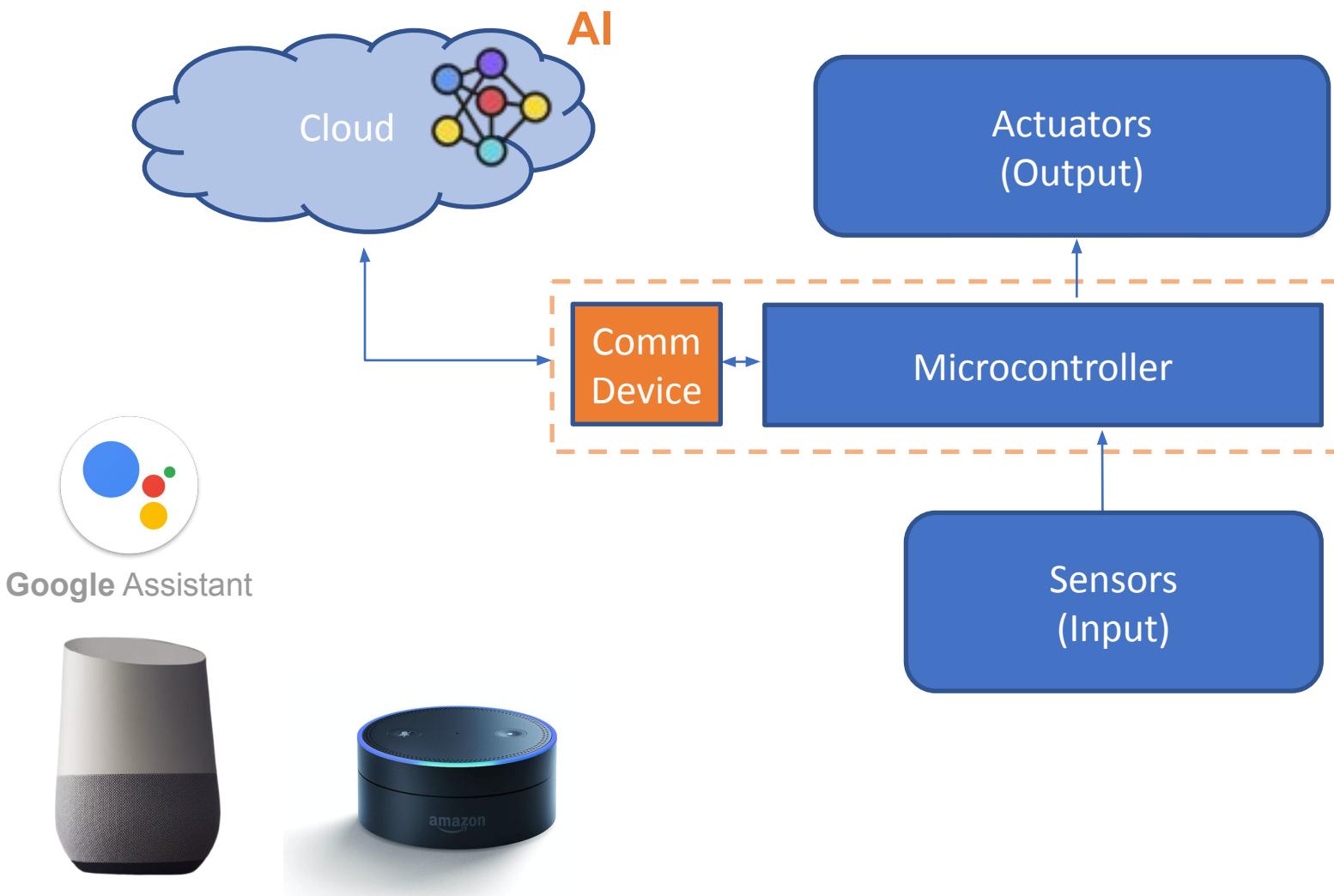
# ... Issues



Bandwidth  
Latency  
Energy  
Reliability

# Typical AIoT Project ...

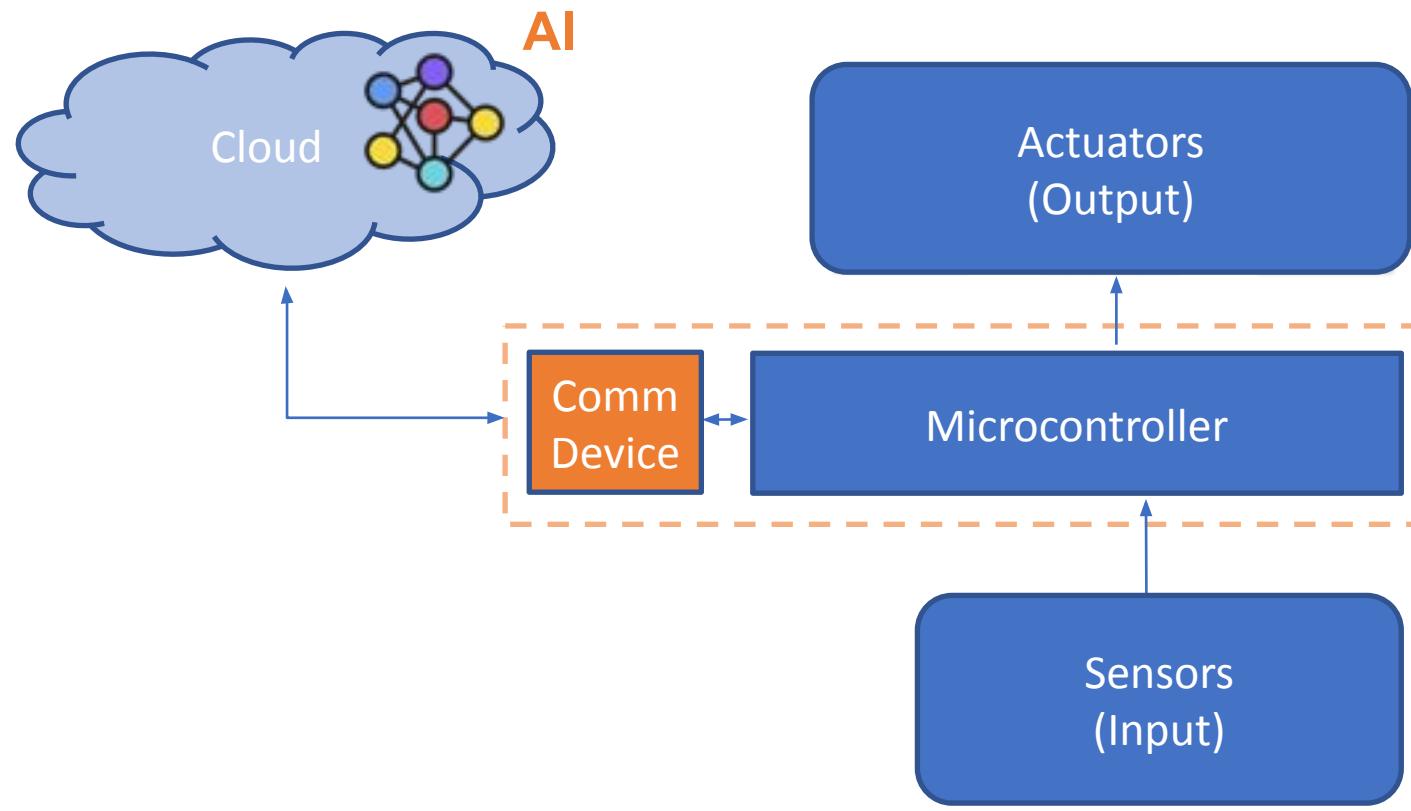
# ... Issues



Bandwidth  
Latency  
Energy  
Reliability  
Privacy

# Typical AIoT Project ...

... Issues

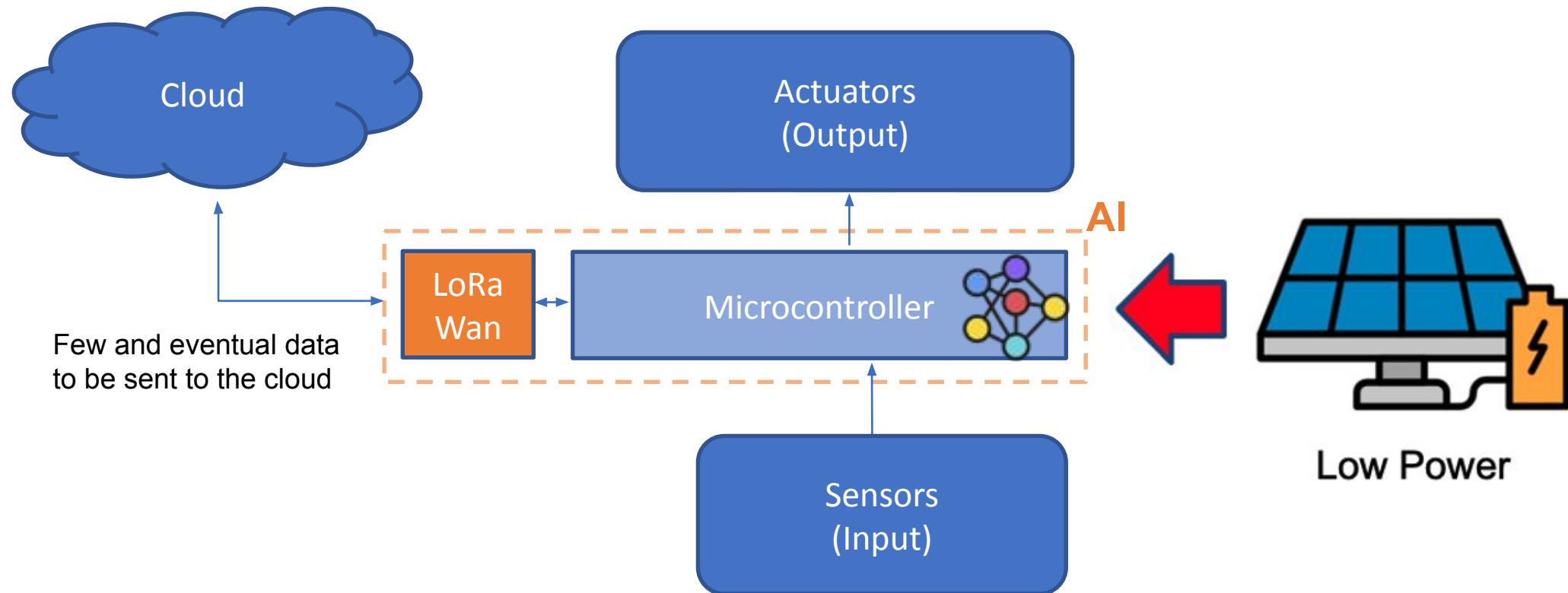


Bandwidth  
Latency  
Energy  
Reliability  
Privacy

... Solution ?

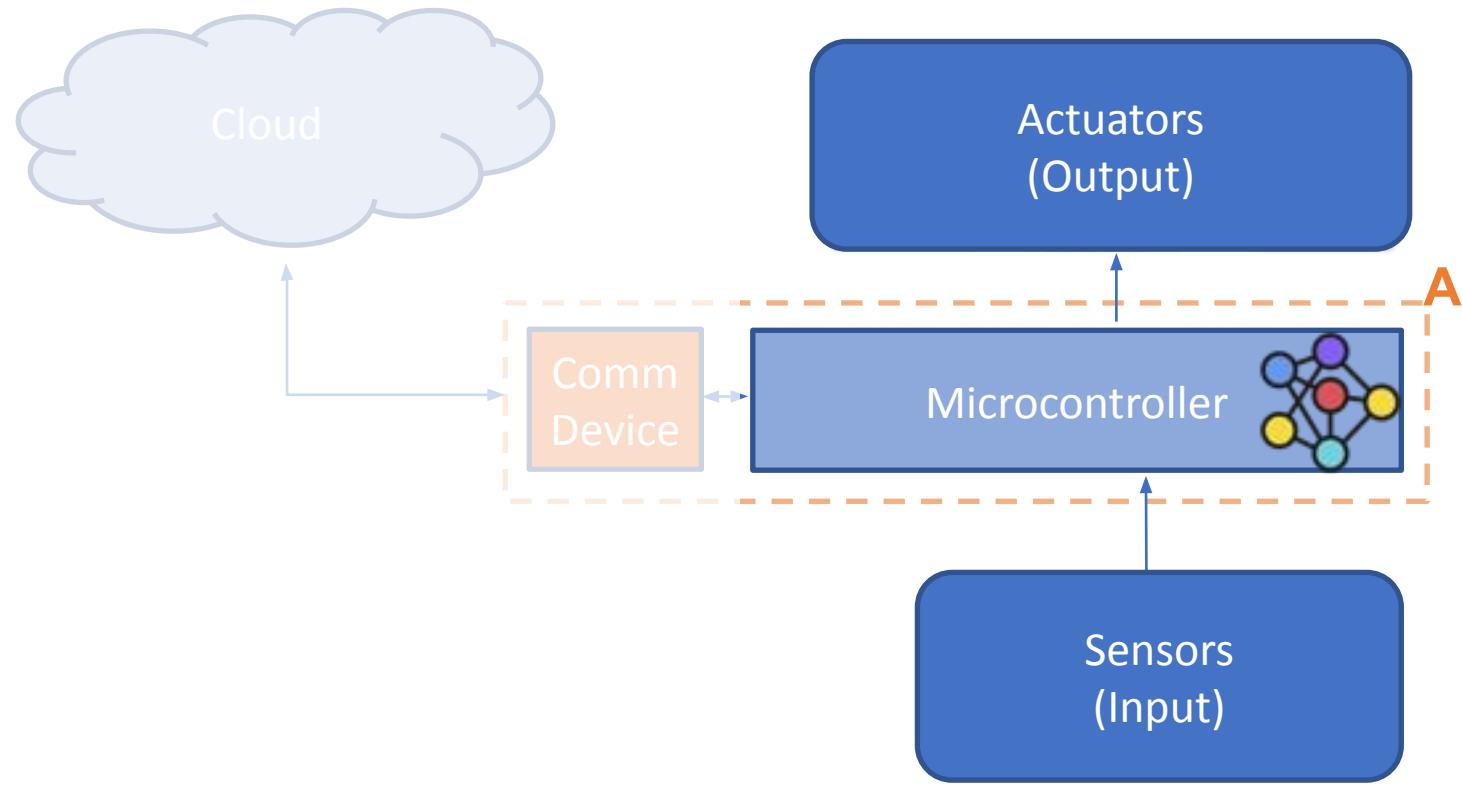
# IoT 2.0 \* – Edge AI/ML

\* Intelligence of Things

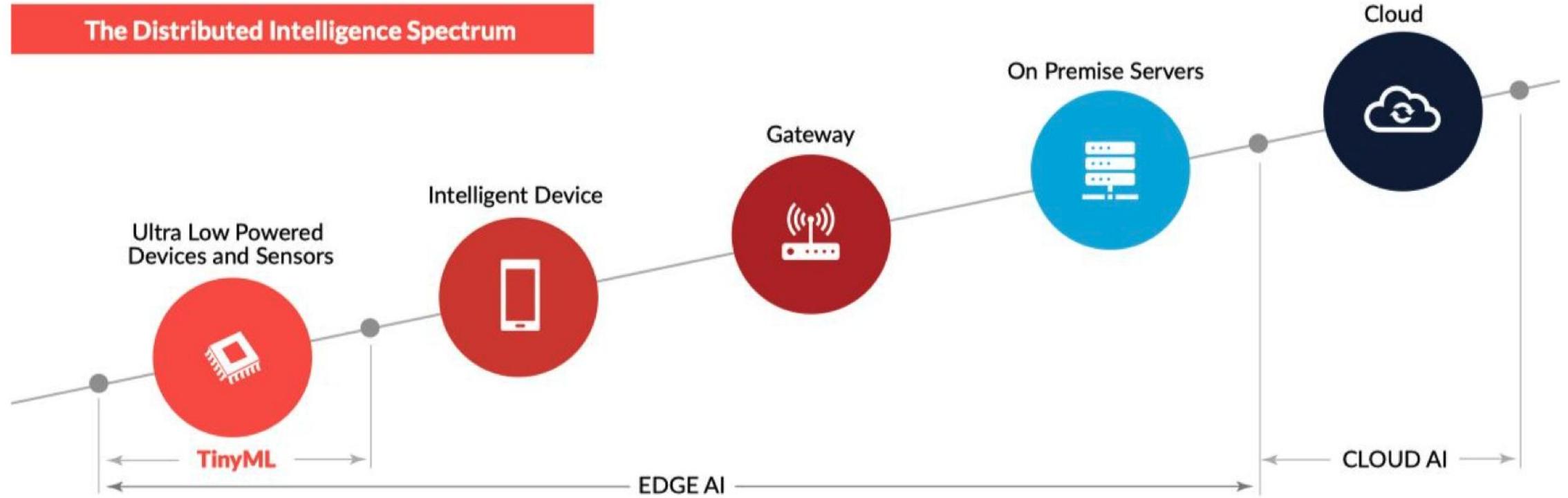


... Solution -> ML goes close to data

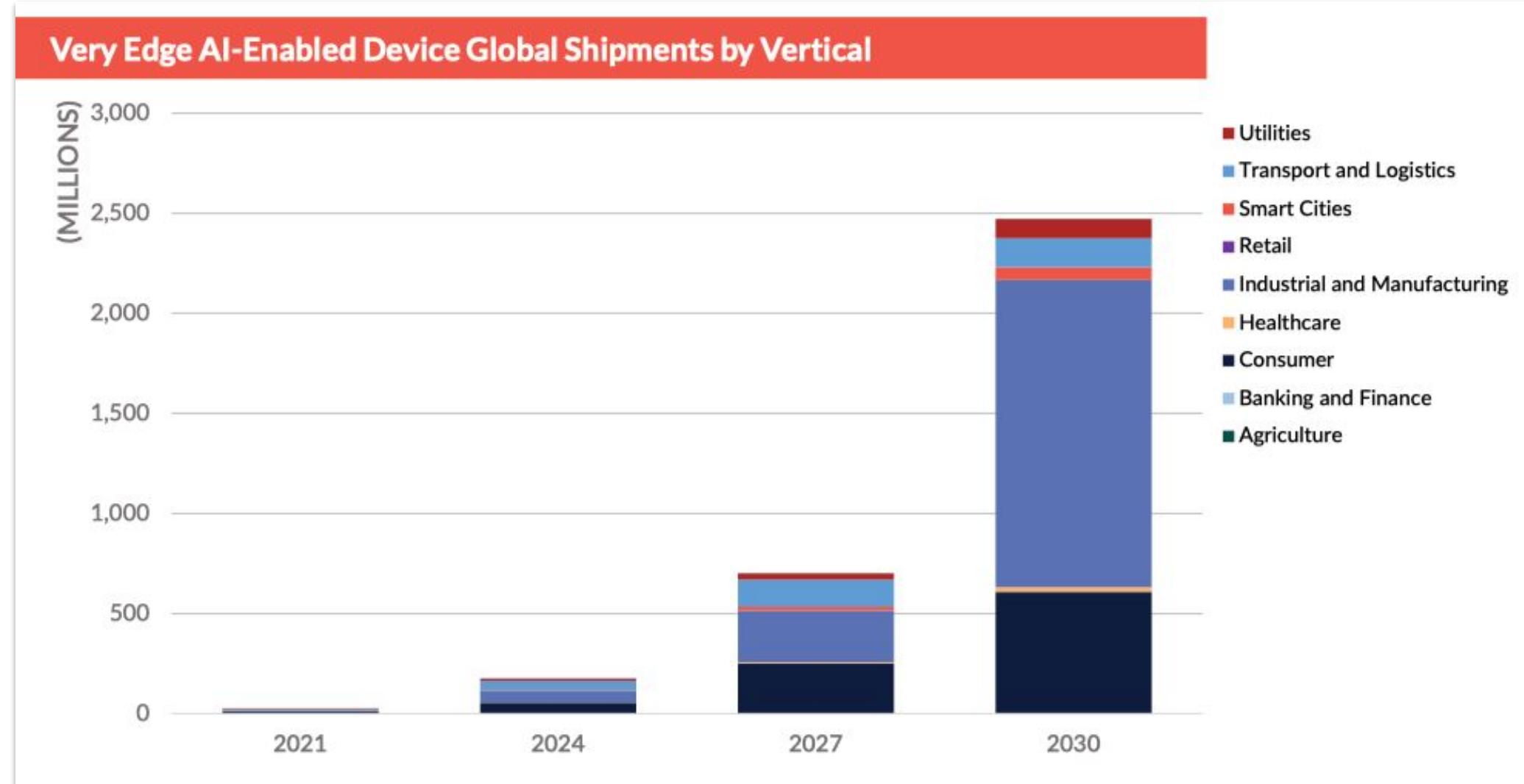
# When to use an Edge AI/ML approach:



**B**andwidth  
**L**atency  
**E**nergy  
**R**eliability  
**P**rivacy

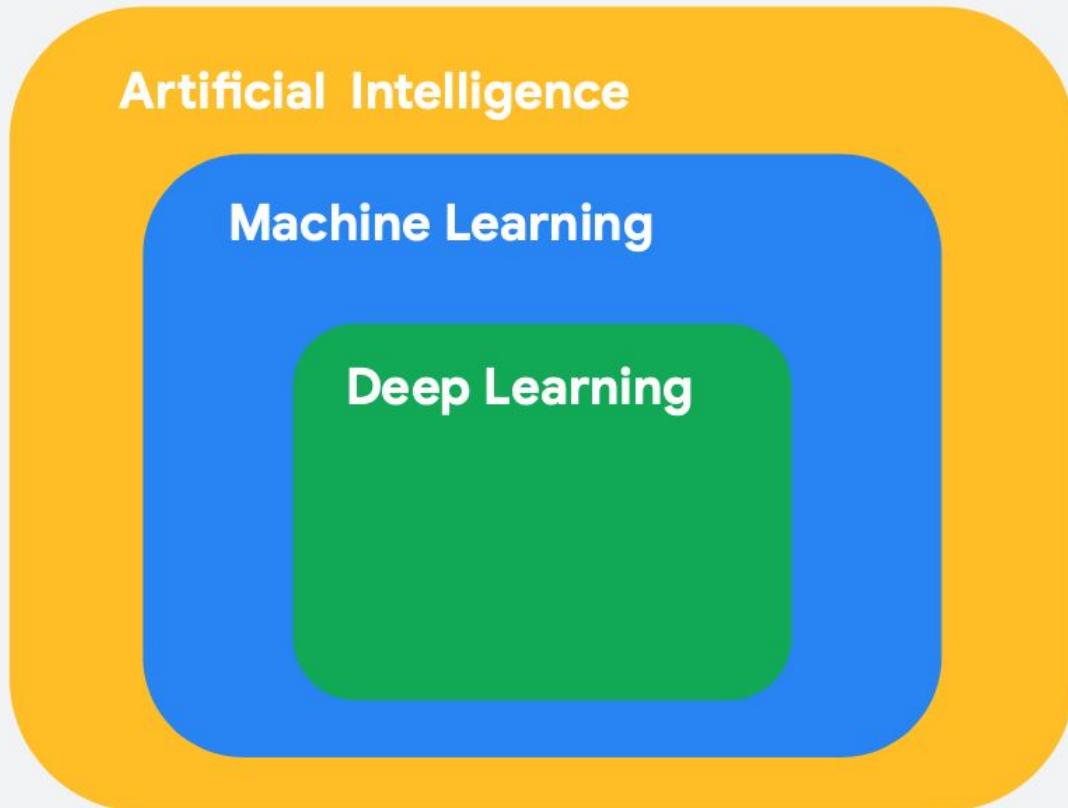


# Market Forecast



# Embedded ML (TinyML)

## Introduction



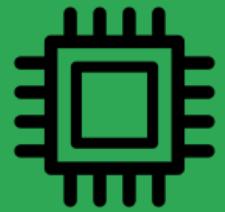
**AI:** Any technique that enables computers to mimic human behavior

**ML:** Ability to learn without explicitly being programmed

**DL:** Extract patterns from data using neural networks

**EdgeAI/ML**

**TinyML**



**Edge AI (or Edge ML)** is the processing of Artificial Intelligence algorithms on edge, that is, on users' devices. The concept derives from **Edge Computing**, which starts from the same premise: data is stored, processed, and managed directly at the Internet of Things (IoT) endpoints.

**TinyML** is a subset of **EdgeML**, where sensors are generating data with ultra-low power consumption (batteries), so that we can ultimately deploy machine learning continuously ("always on devices")

# What is Tiny Machine Learning (**TinyML**)?

**TinyML**



Fastest-growing field of **ML**



# What is Tiny Machine Learning (**TinyML**)?

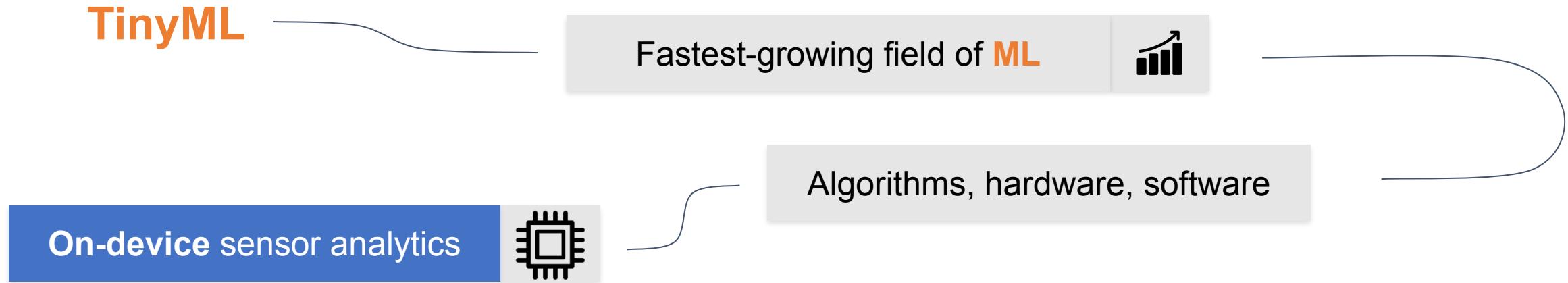
**TinyML**

Fastest-growing field of **ML**

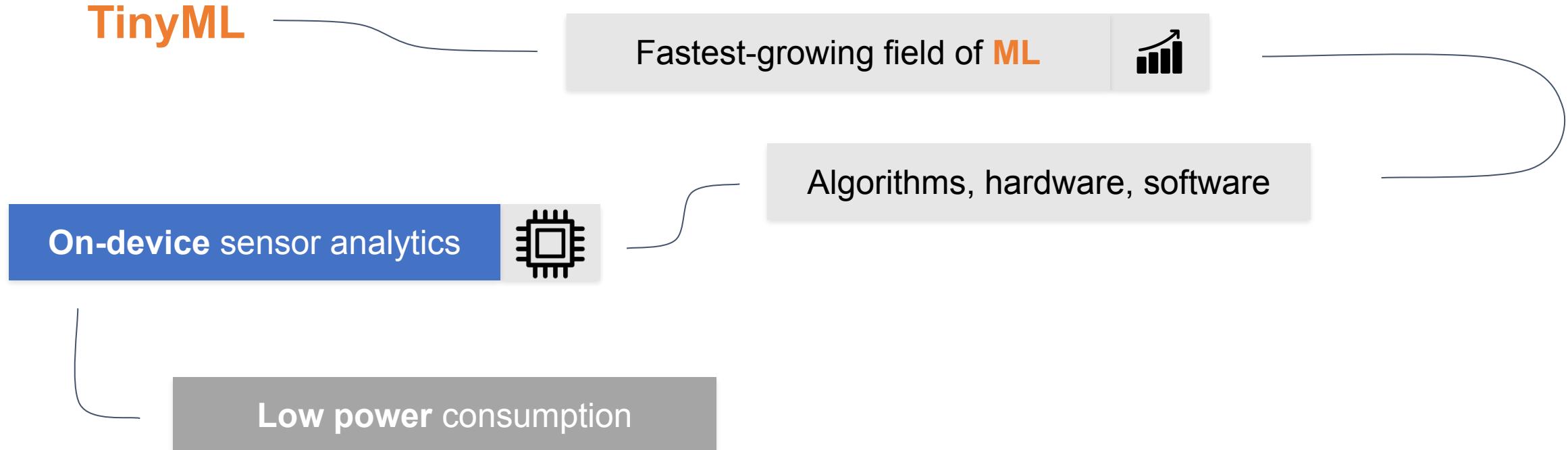


Algorithms, hardware, software

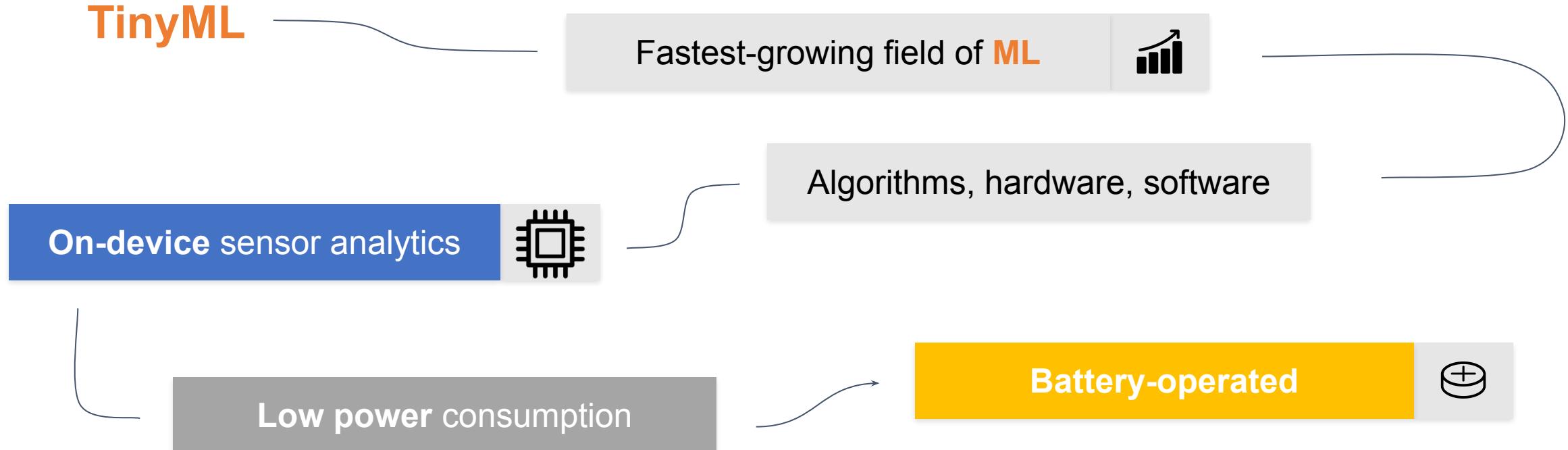
# What is Tiny Machine Learning (**TinyML**)?



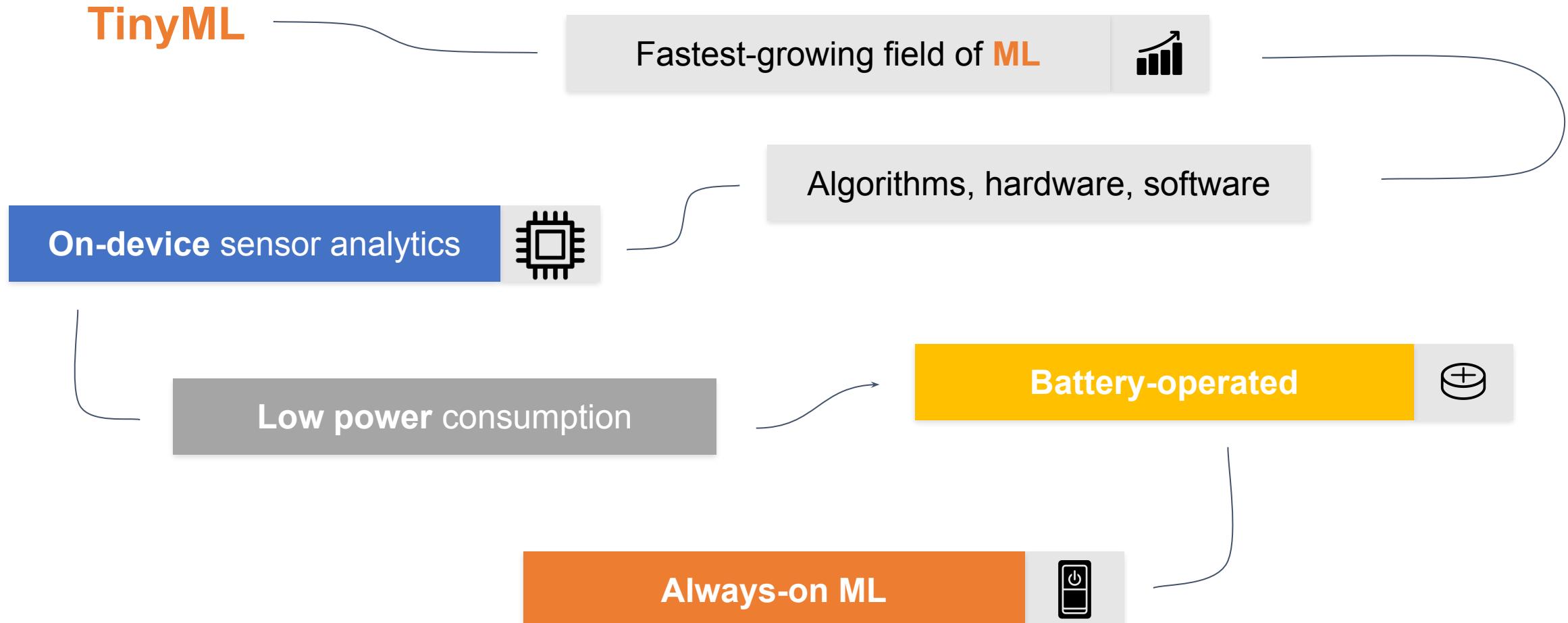
# What is Tiny Machine Learning (**TinyML**)?



# What is Tiny Machine Learning (**TinyML**)?



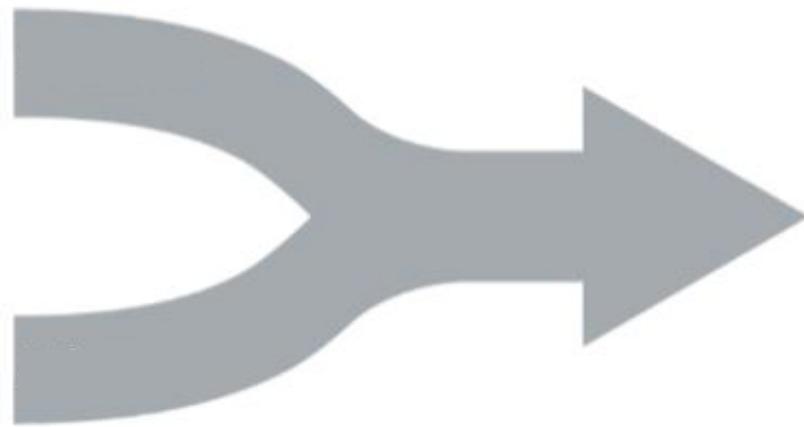
# What is Tiny Machine Learning (**TinyML**)?



# What Makes **TinyML** ?

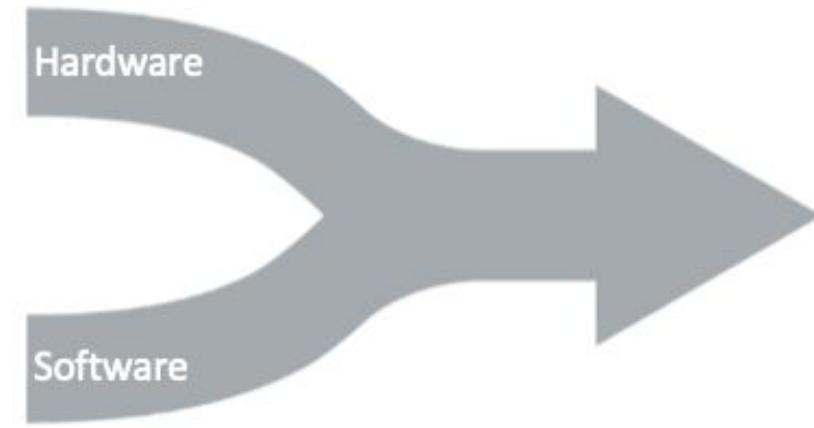
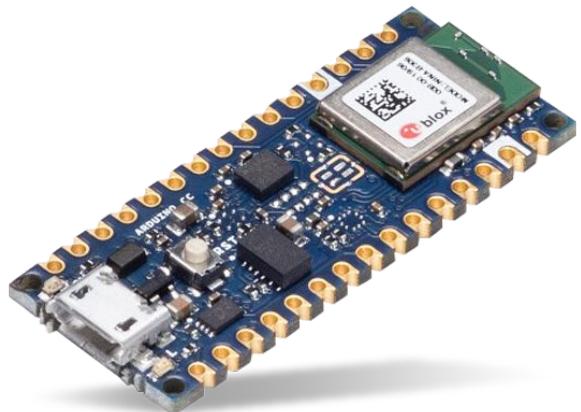
Embedded  
Systems

Machine  
Learning



**TinyML**

# What Makes **TinyML** ?

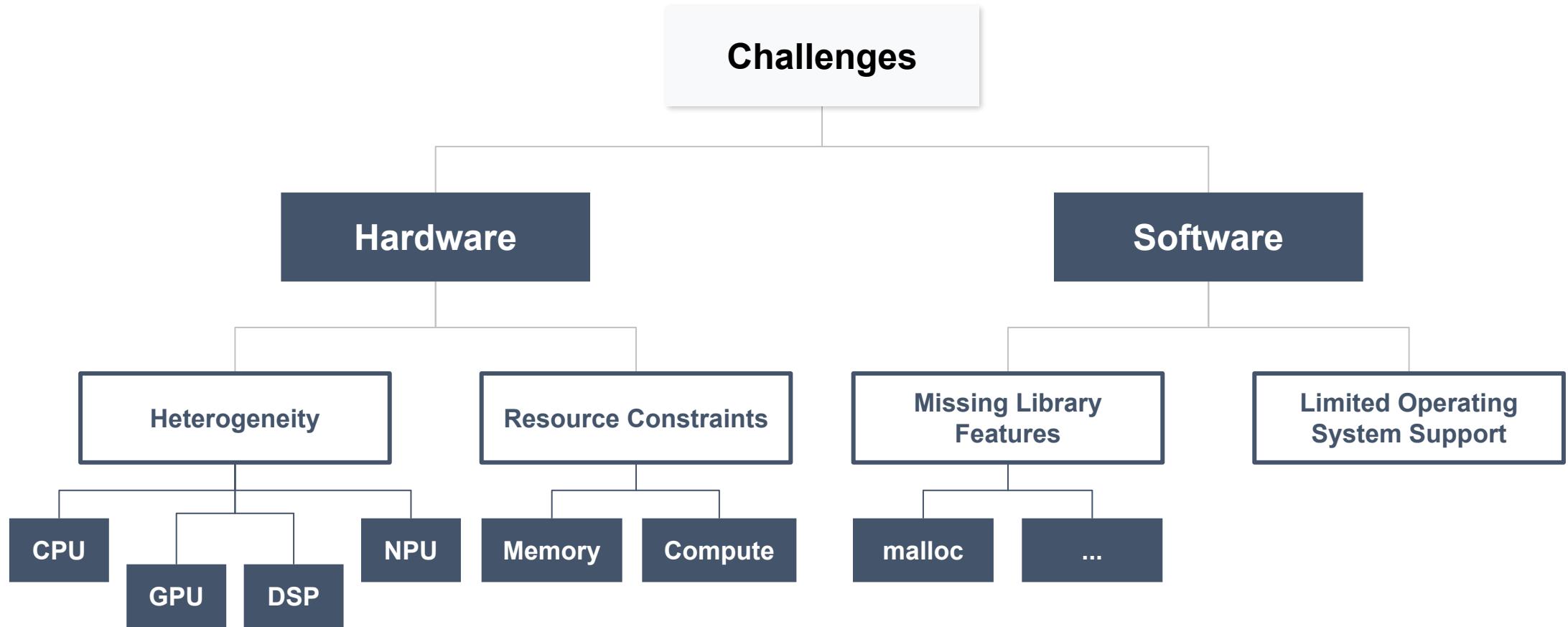


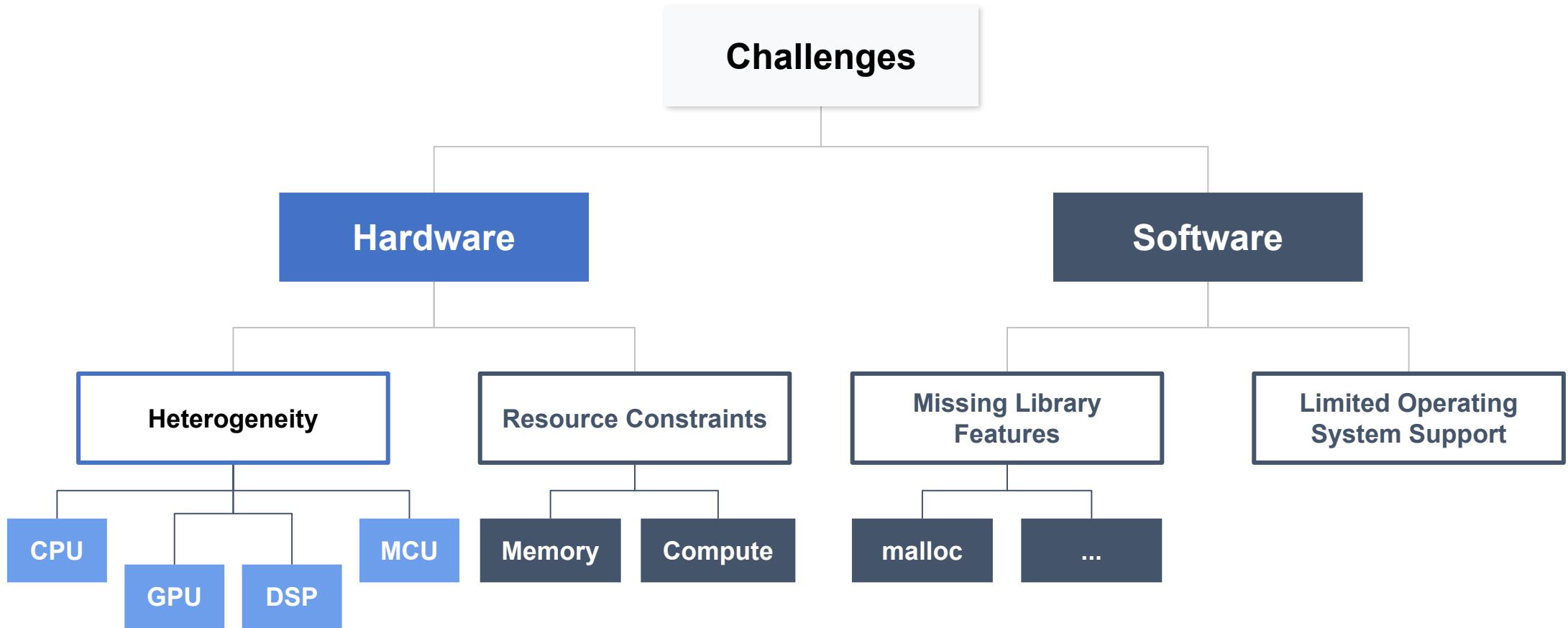
**TinyML**



**TensorFlow Lite**

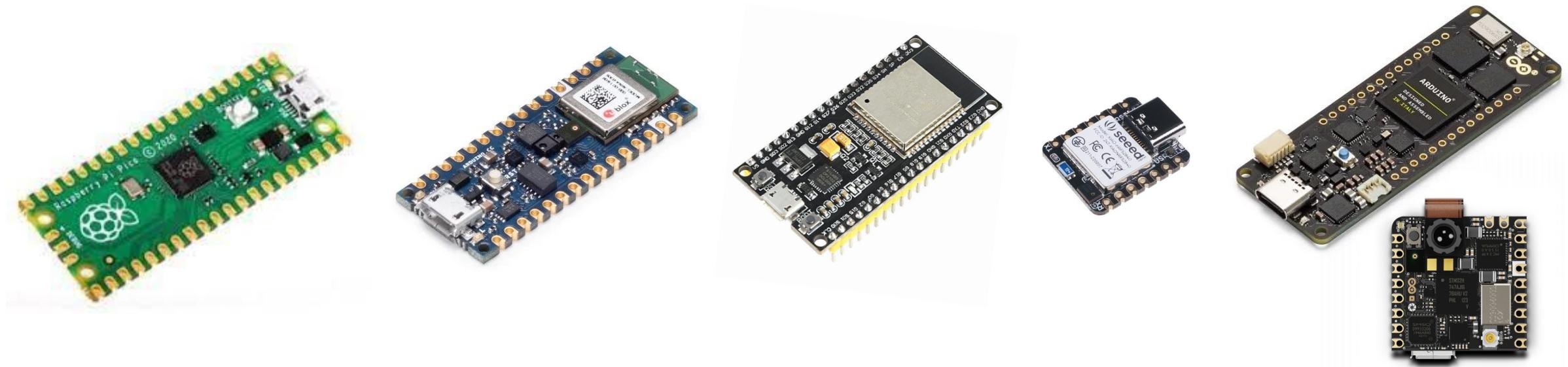
# TinyML Challenges



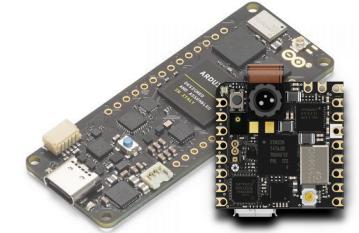


**250 Billion**  
*MCUs today*

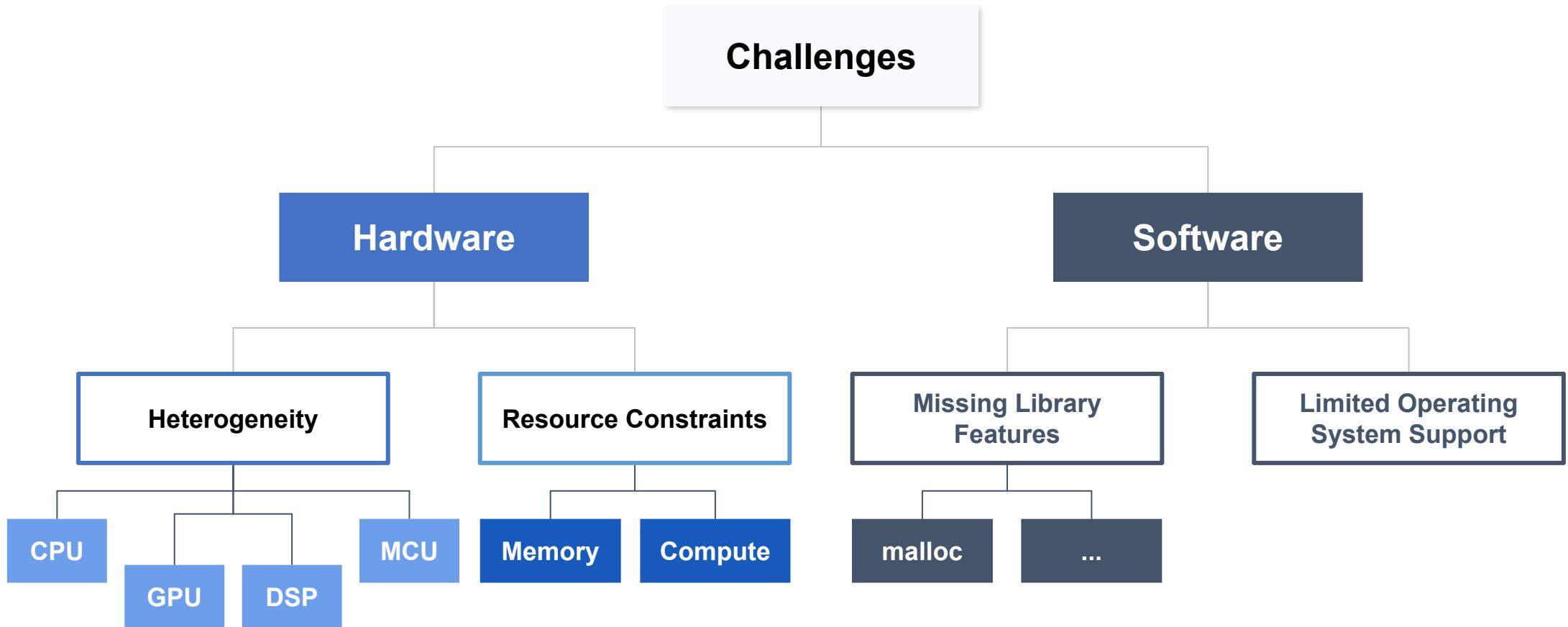
# Hardware



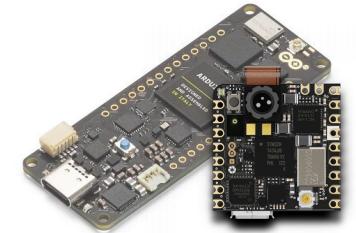
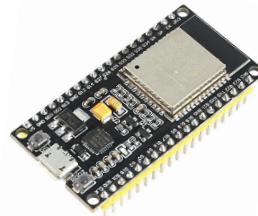
# Hardware



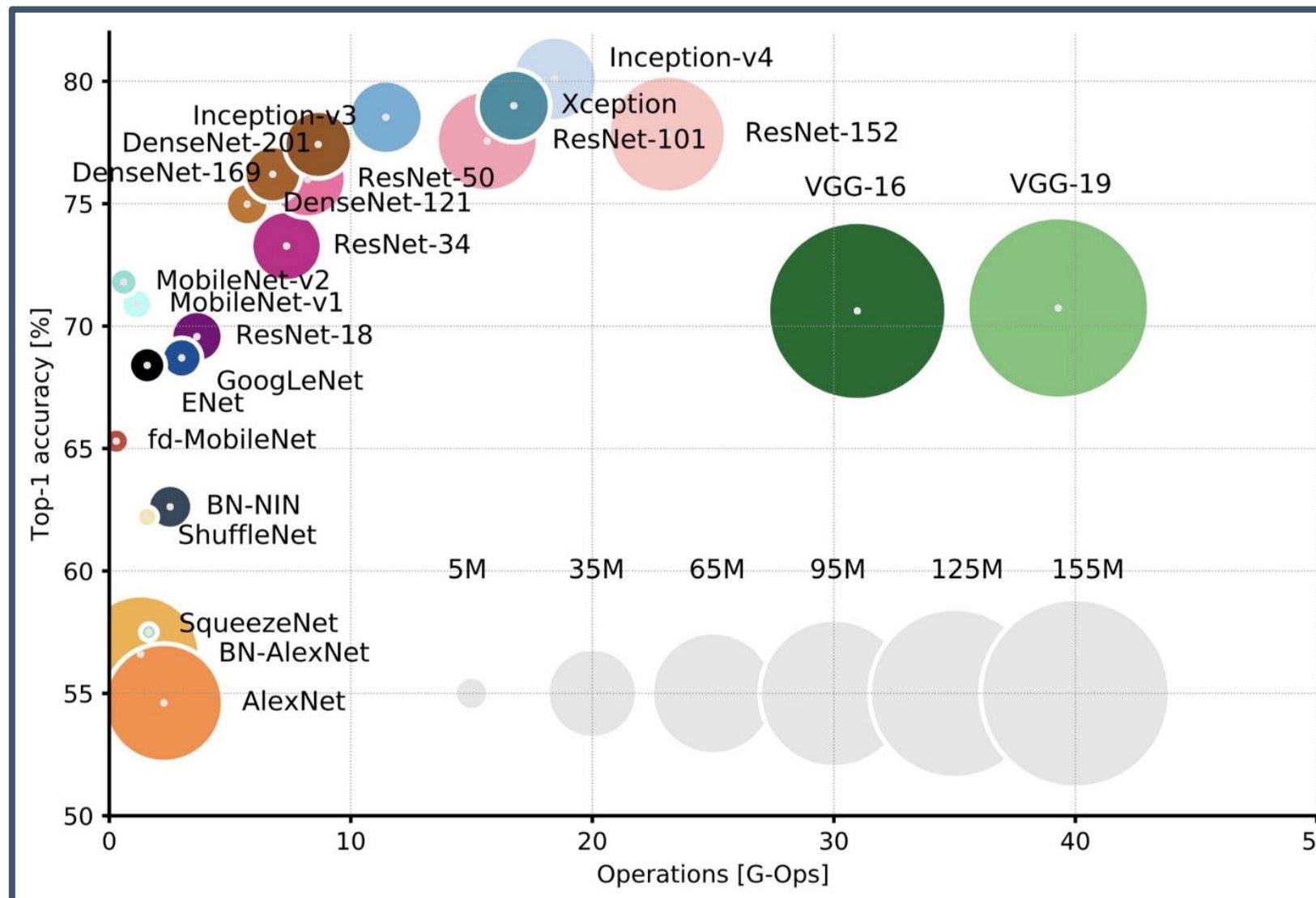
	Raspberry Pico (W)	Arduino Nano Sense	ESP 32	Seeed XIAO BLE Sense	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64MHz	480/240MHz
RAM	264KB	256KB	520KB	256KB	1MB
ROM	2MB	1MB	2MB	2MB	2MB
Radio	(Yes for W)	BLE	BLE/WiFi	BLE	BLE/WiFi
Sensors	No	Yes	No	Yes	No (Portenta) Yes (Nicla)
Price	\$	\$\$\$	\$\$	\$\$	\$\$\$\$



# Hardware

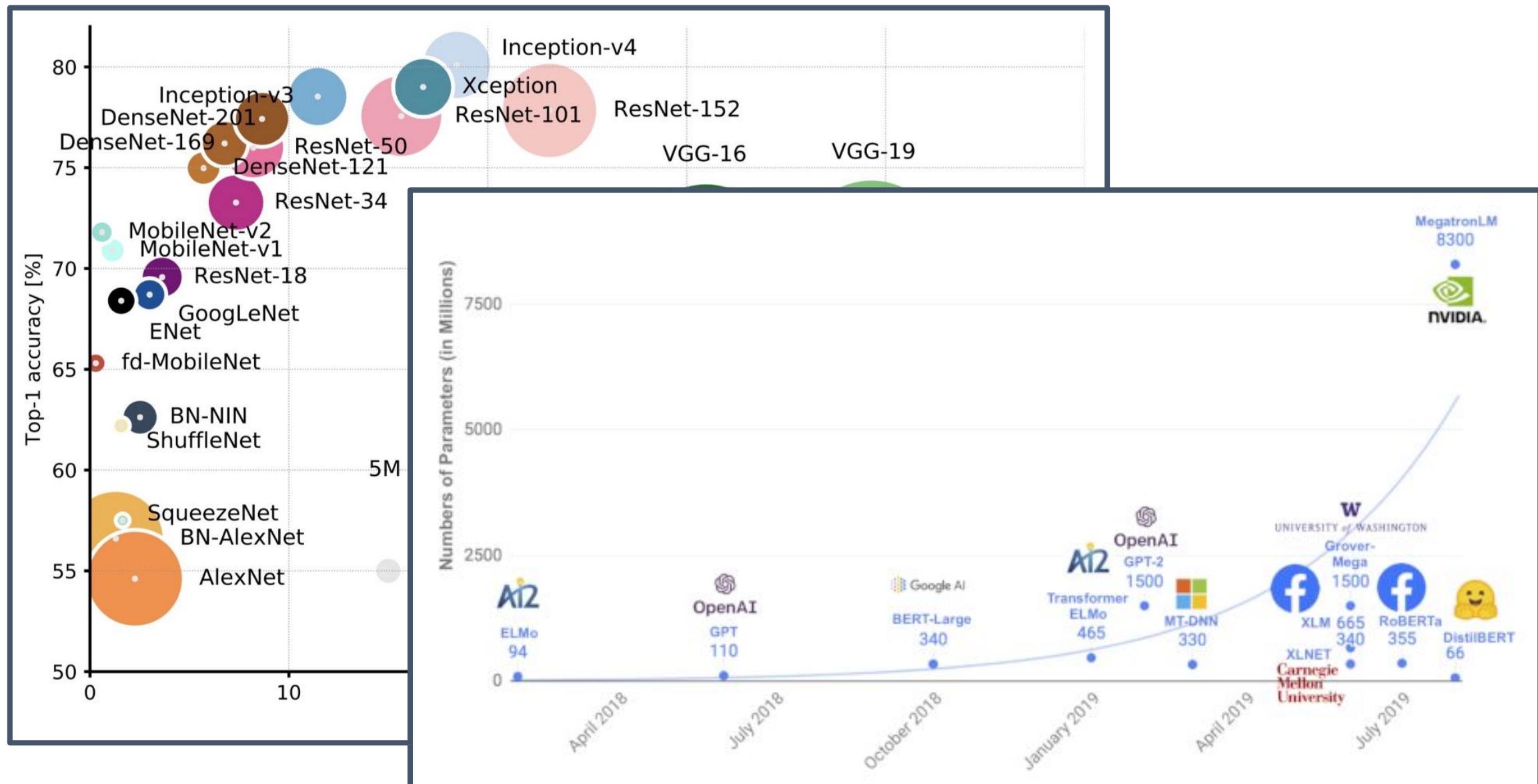


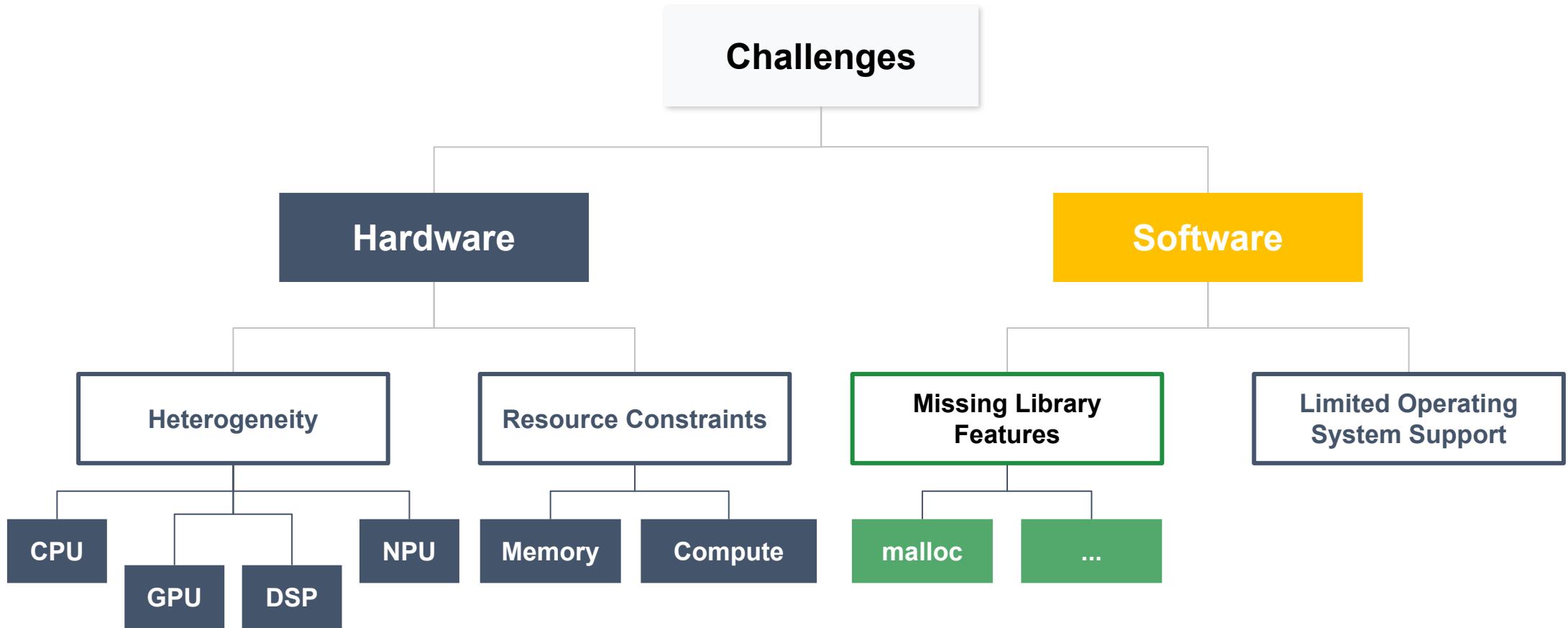
Raspberry Pico (W)		Arduino Nano Sense	ESP 32	Seeed XIAO BLE Sense	Arduino Pro
32Bits CPU	Dual-core Arm Cortex-M0+	Arm Cortex-M4F	Xtensa LX6 Dual Core	Arm Cortex-M4F	Dual Core Arm Cortex M7/M4
CLOCK	133MHz	64MHz	240MHz	64MHz	480/240MHz
RAM	264KB	256KB	520KB	256KB	1MB
ROM	2MB	1MB	2MB	2MB	2MB
Radio	(Yes for W)	BLE	BLE/WiFi	BLE	BLE/WiFi
Sensors	No	Yes	No	Yes	No (Portenta) Yes (Nicla)
Price	\$	\$\$\$	\$\$	\$\$	\$\$\$\$



<https://arxiv.org/pdf/1910.01108.pdf>

<https://towardsdatascience.com/neural-network-architectures-156e5bad51ba>





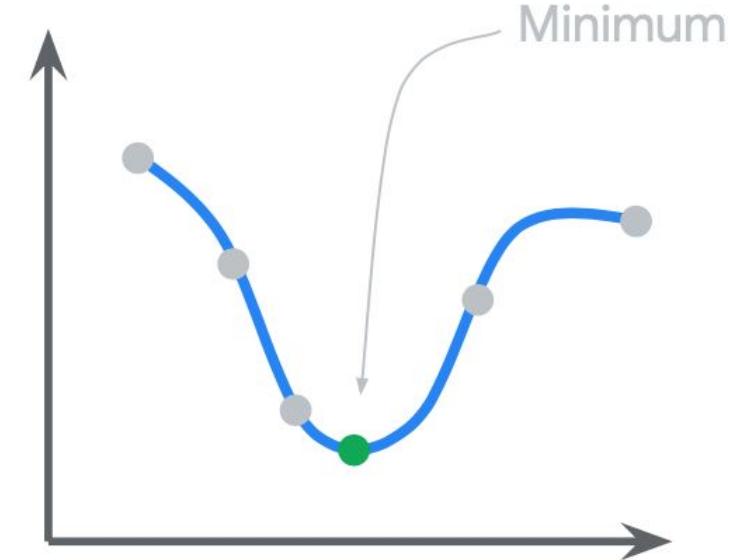
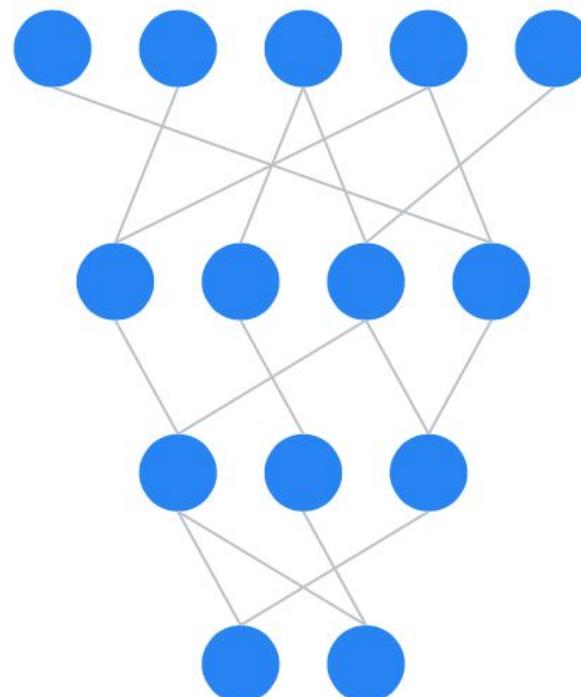
# Sensors Metrics

**Acoustic Sensors**  
Ultrasonic, Microphones,  
Geophones, Vibrometers

**Image Sensors**  
Thermal, Image

**Motion Sensors**  
Gyroscope, Radar,  
Accelerometer

# Models



End-to-end **TinyML** application design

## Datasets Preprocessing

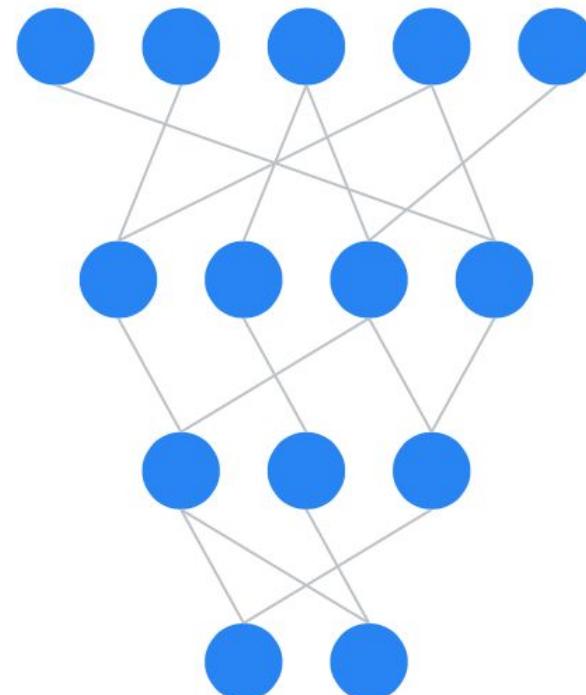
## Quantization Pruning

## Resource constraints

Sound

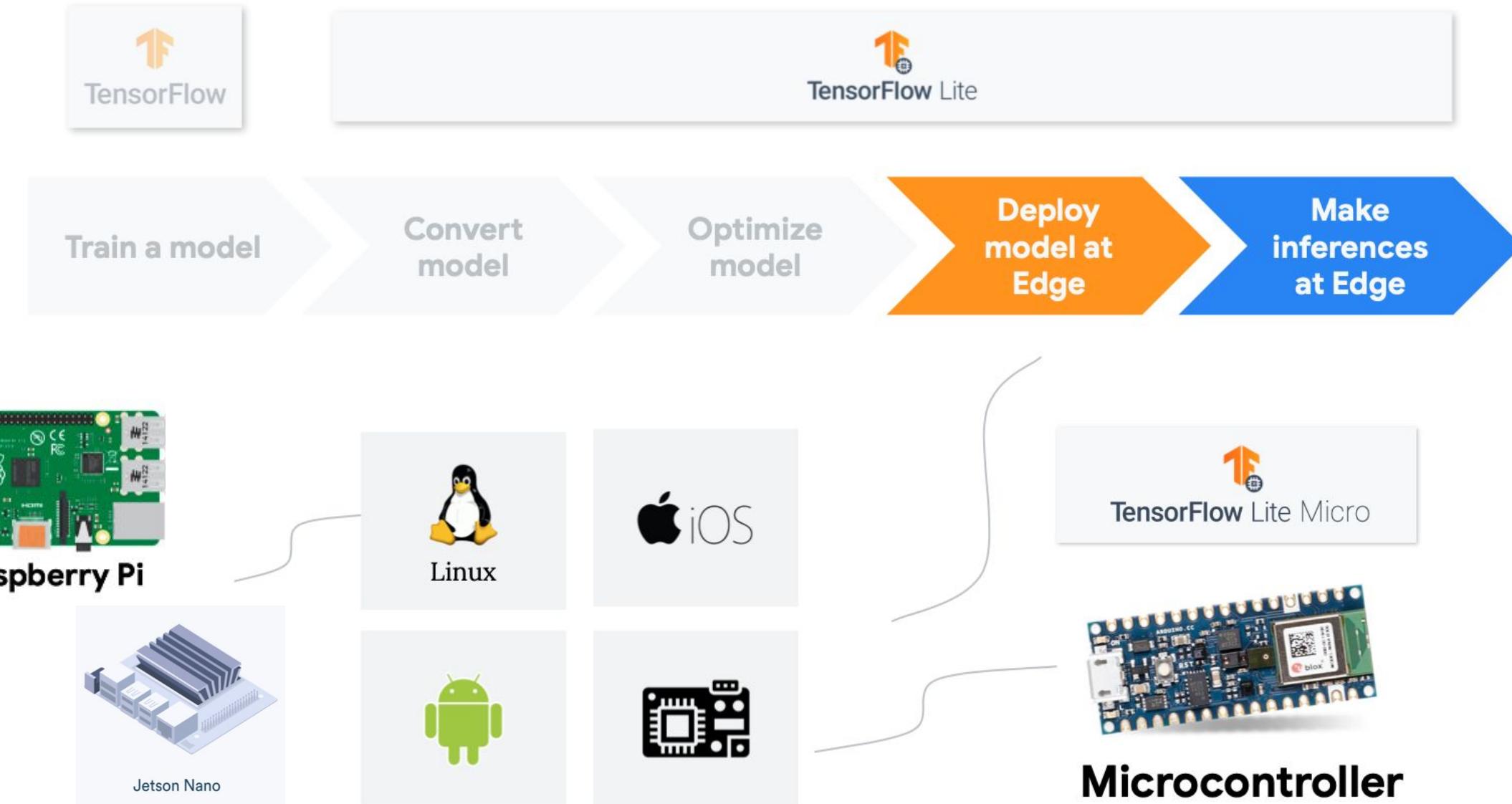
Vision

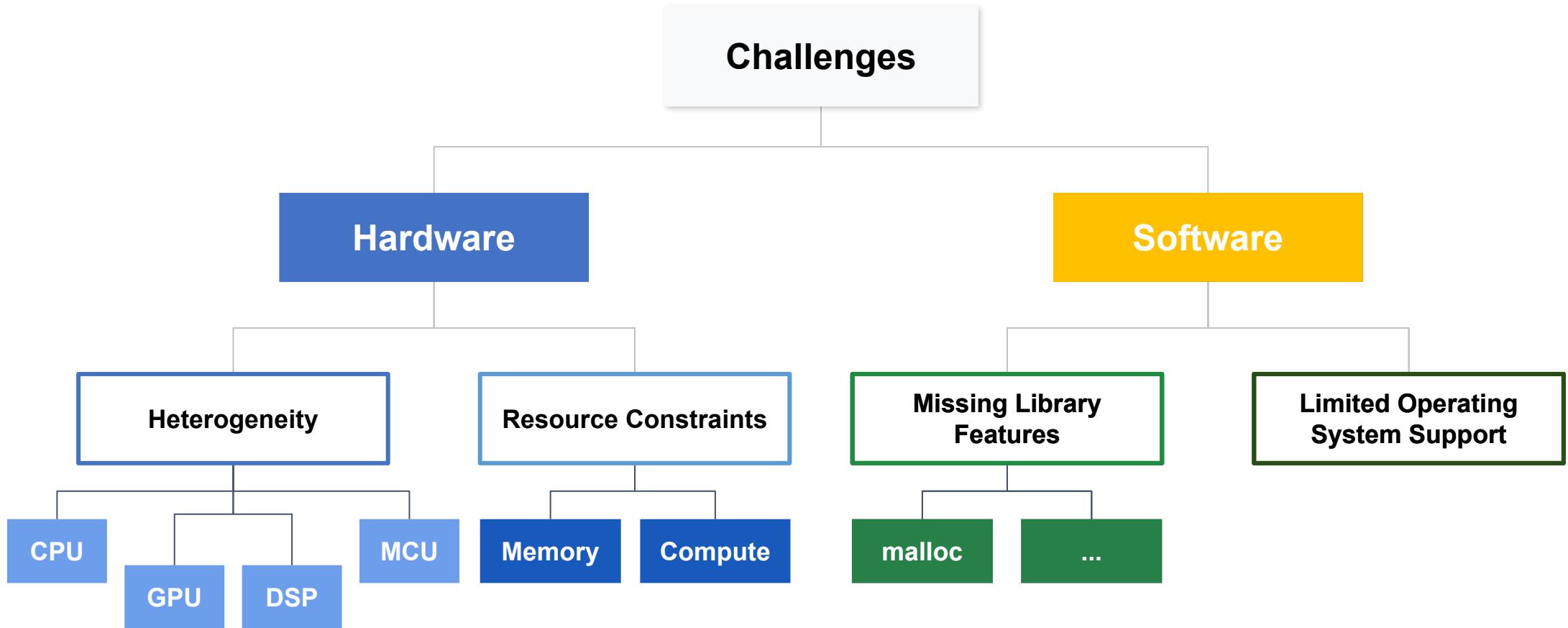
Vibration



End-to-end **TinyML** application design

# Software





# Application Complexity vs. HW

Power



# EdgeML

## TinyML



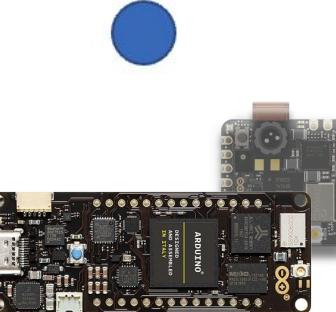
Anomaly Detection  
Sensor Classification  
20 KB



Rpi-Pico  
(Cortex-M0+)



KeyWord Spotting  
Audio Classification  
50 KB



Arduino Pro  
(Cortex-M7)

Image  
Classification  
250 KB+



## TinyML

Object Detection  
Complex Voice  
Processing  
1 MB+



RaspberryPi  
(Cortex-A)



SmartPhone  
(Cortex-A)



Jetson Nano  
(Cortex-A + GPU)

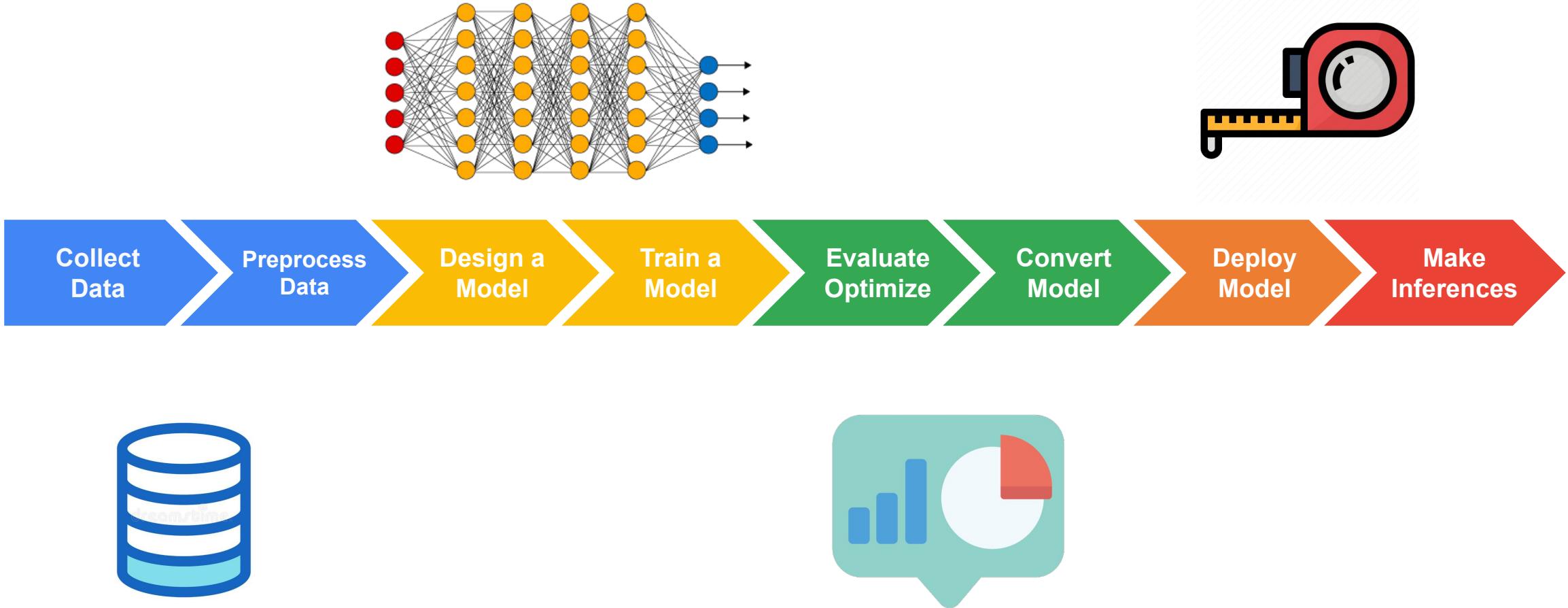
Application Complexity ↑

CPU Power / Memory →

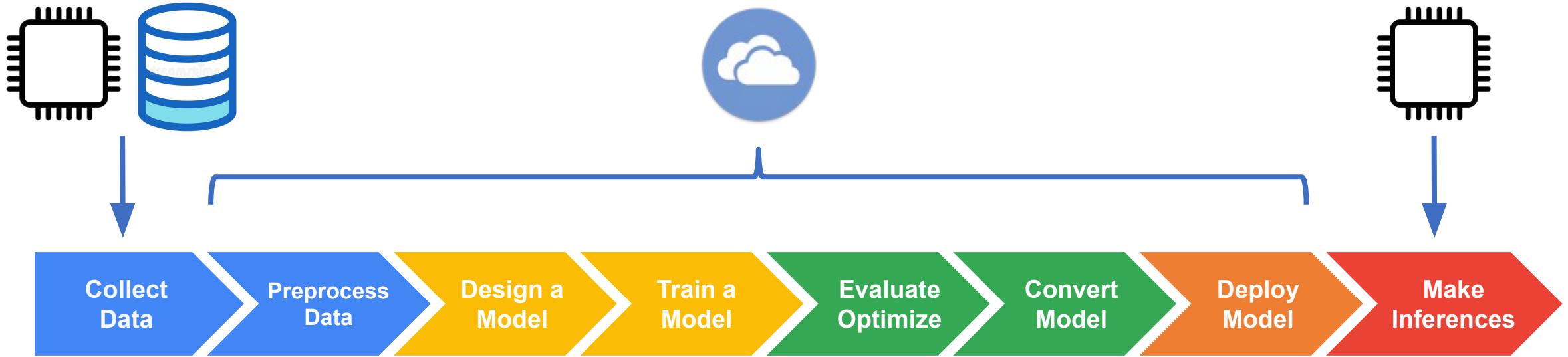
Video  
Classification  
2 MB+

# How to Train a ML Model?

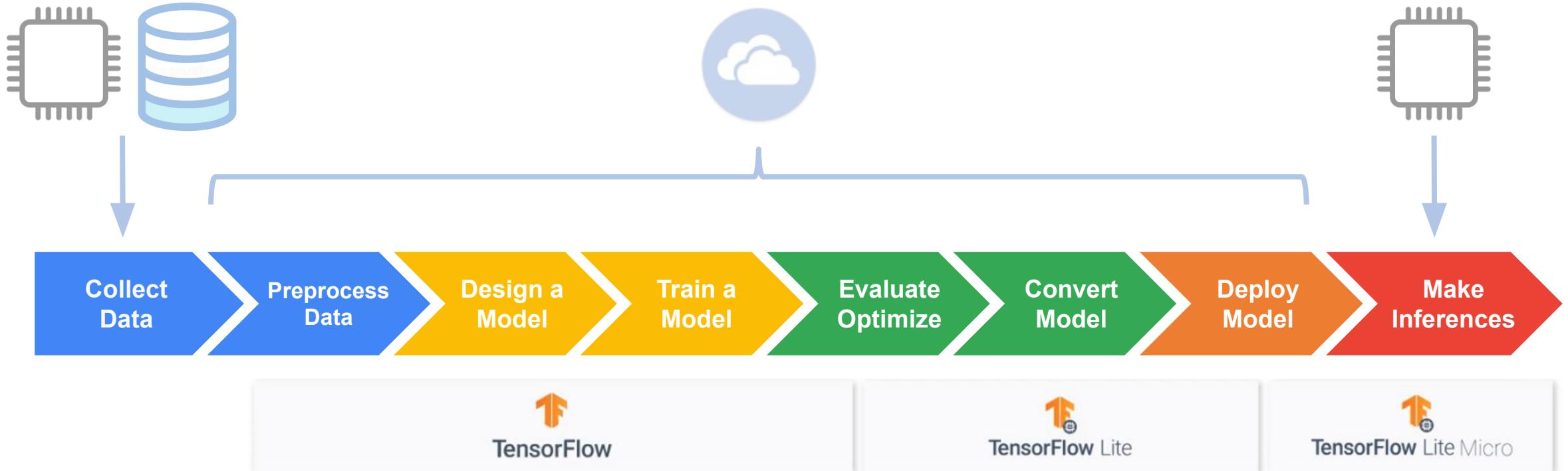
# Machine Learning Workflow (“What”)



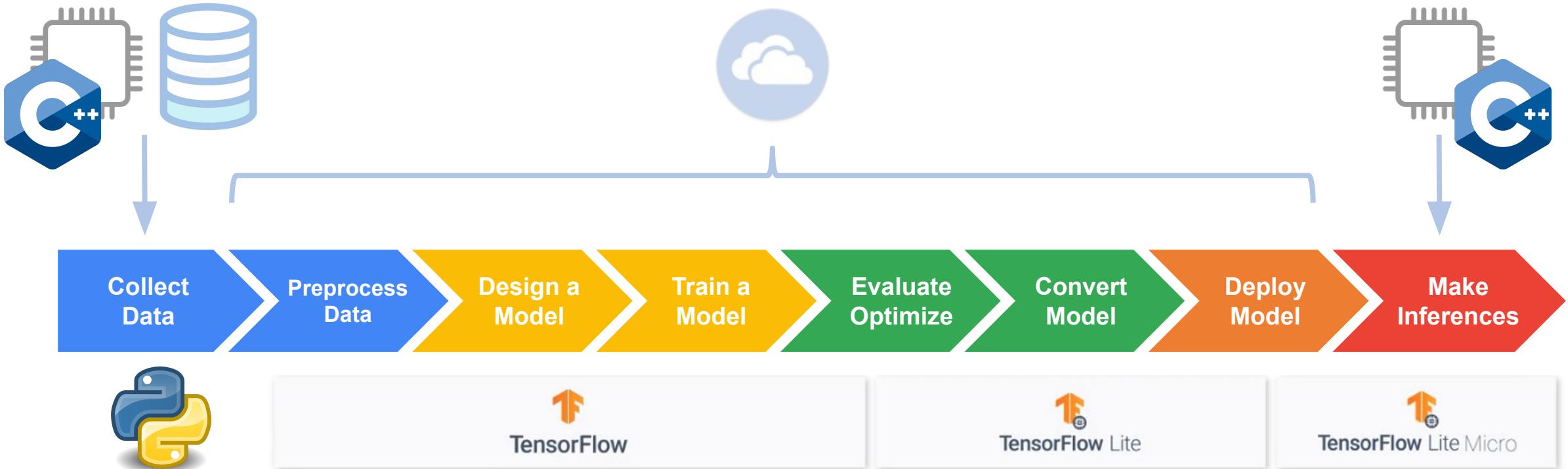
# Machine Learning Workflow (“Where”)



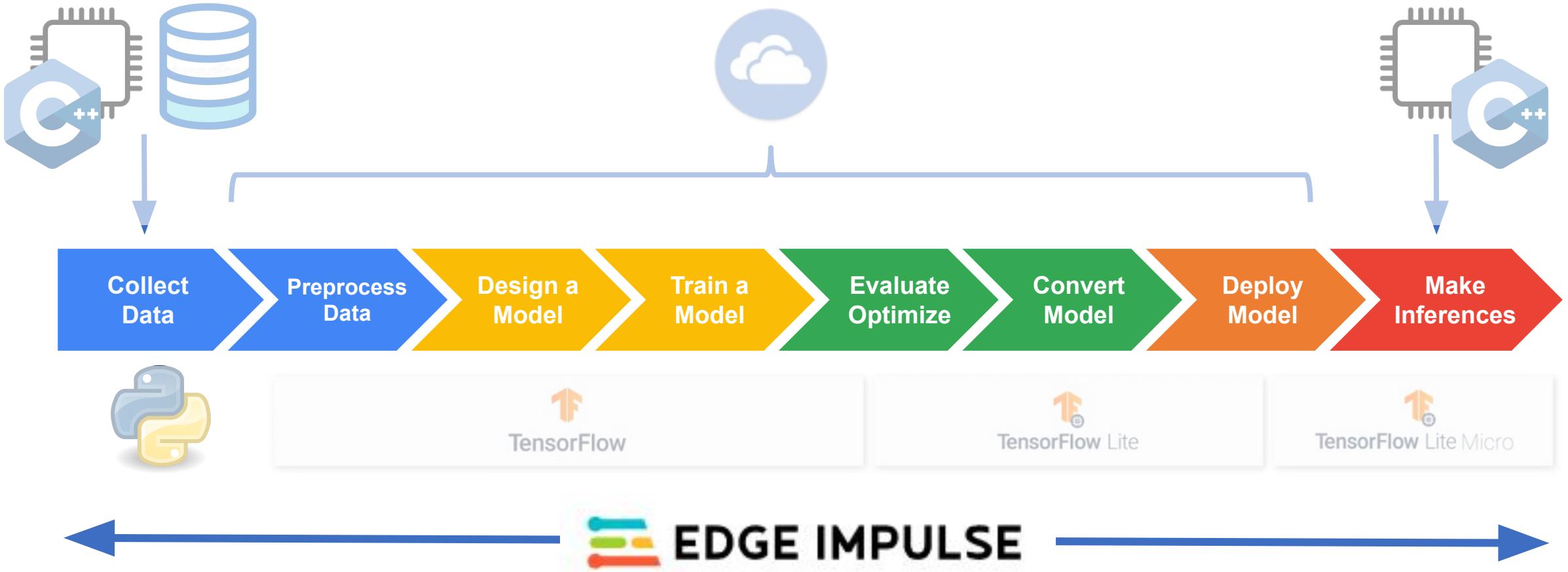
# Machine Learning Workflow (“How”)

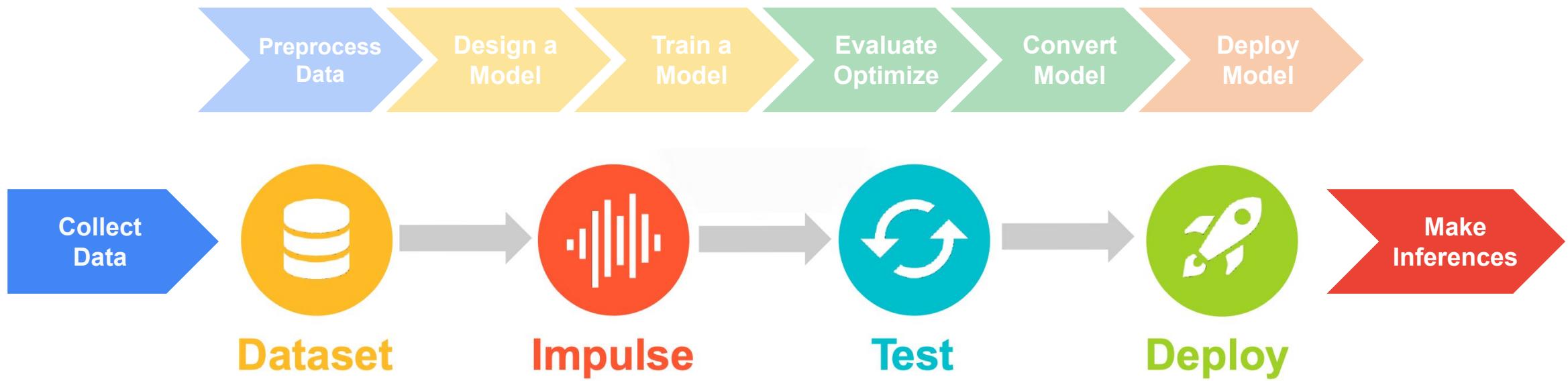


# Machine Learning Workflow (“How”)

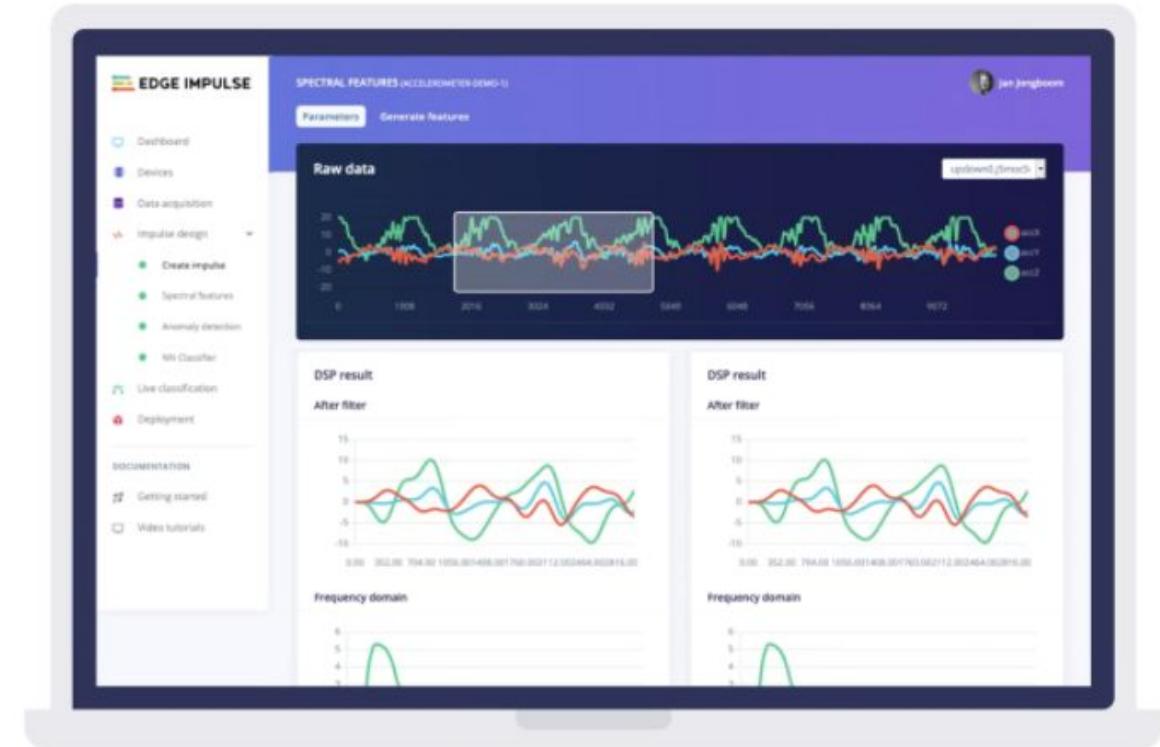
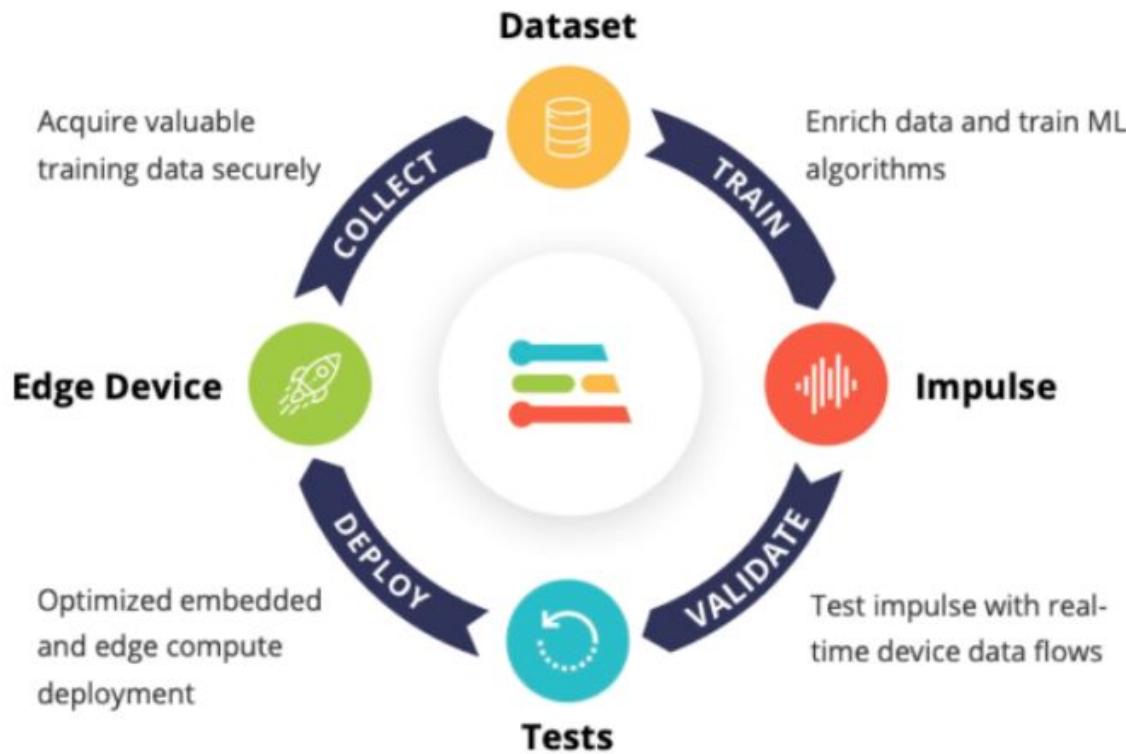


# Machine Learning Workflow (“How”)





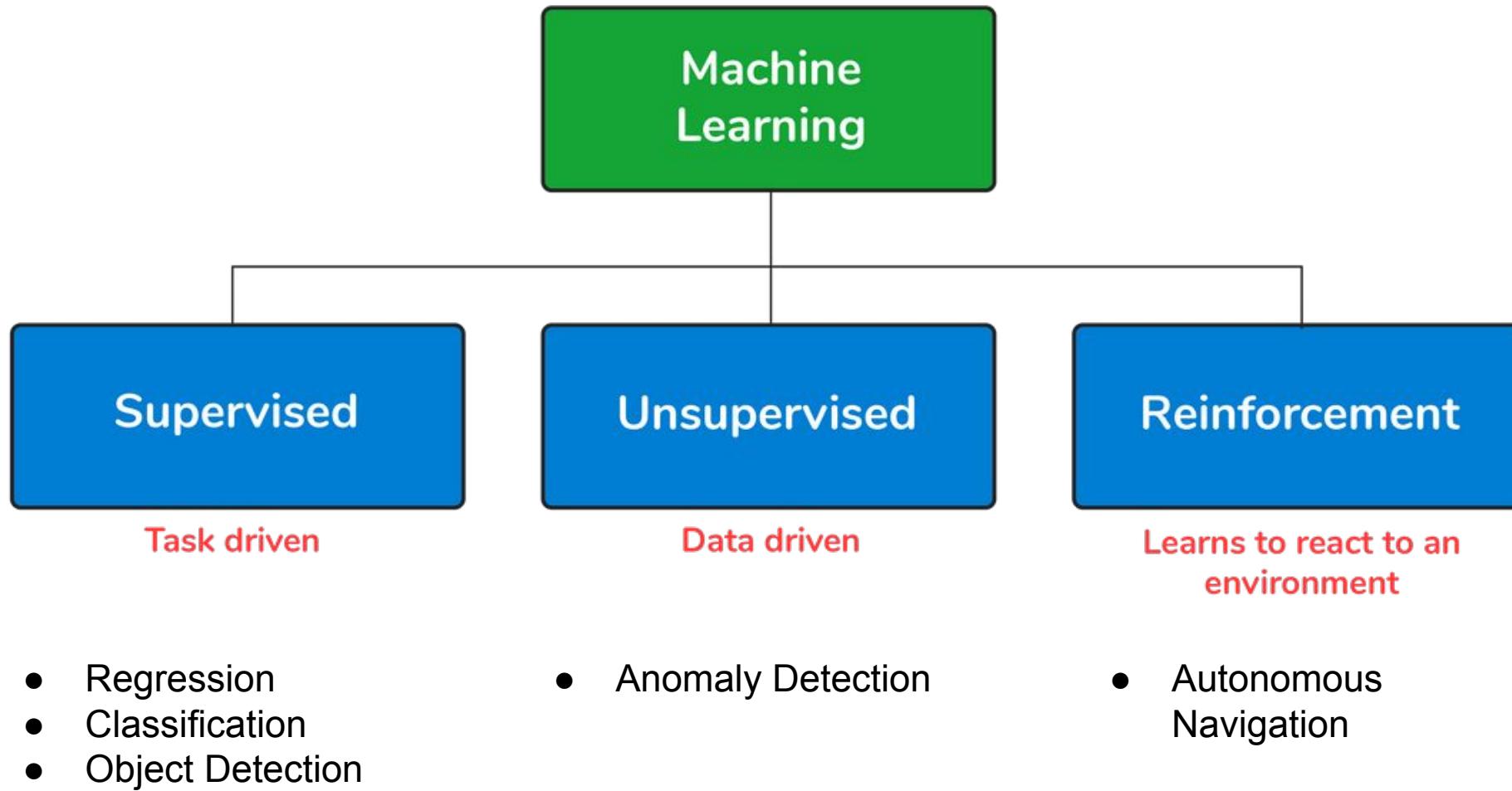
# EI Studio - Embedded ML platform (“AutoML”)



Learn more at <http://edgeimpulse.com>



# TinyML Application Examples



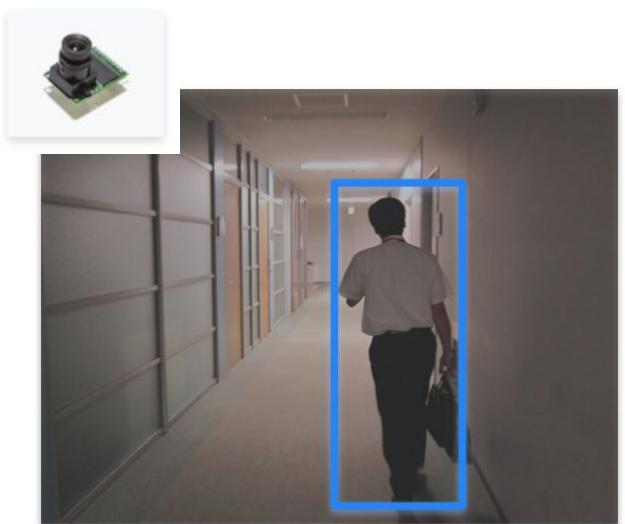
# Sound



# Vibration



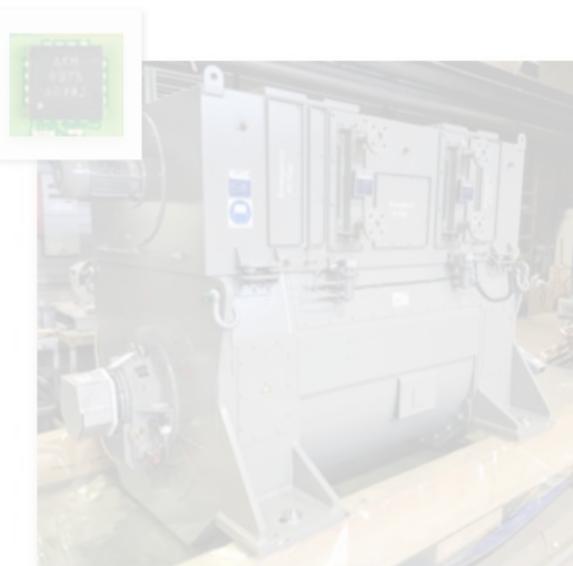
# Vision



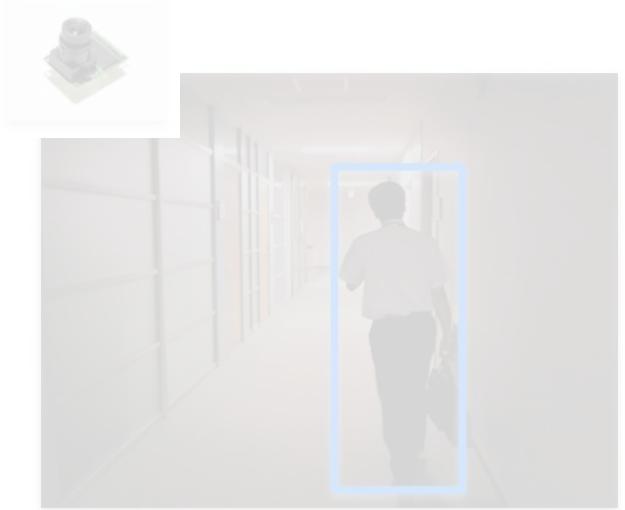
# Sound



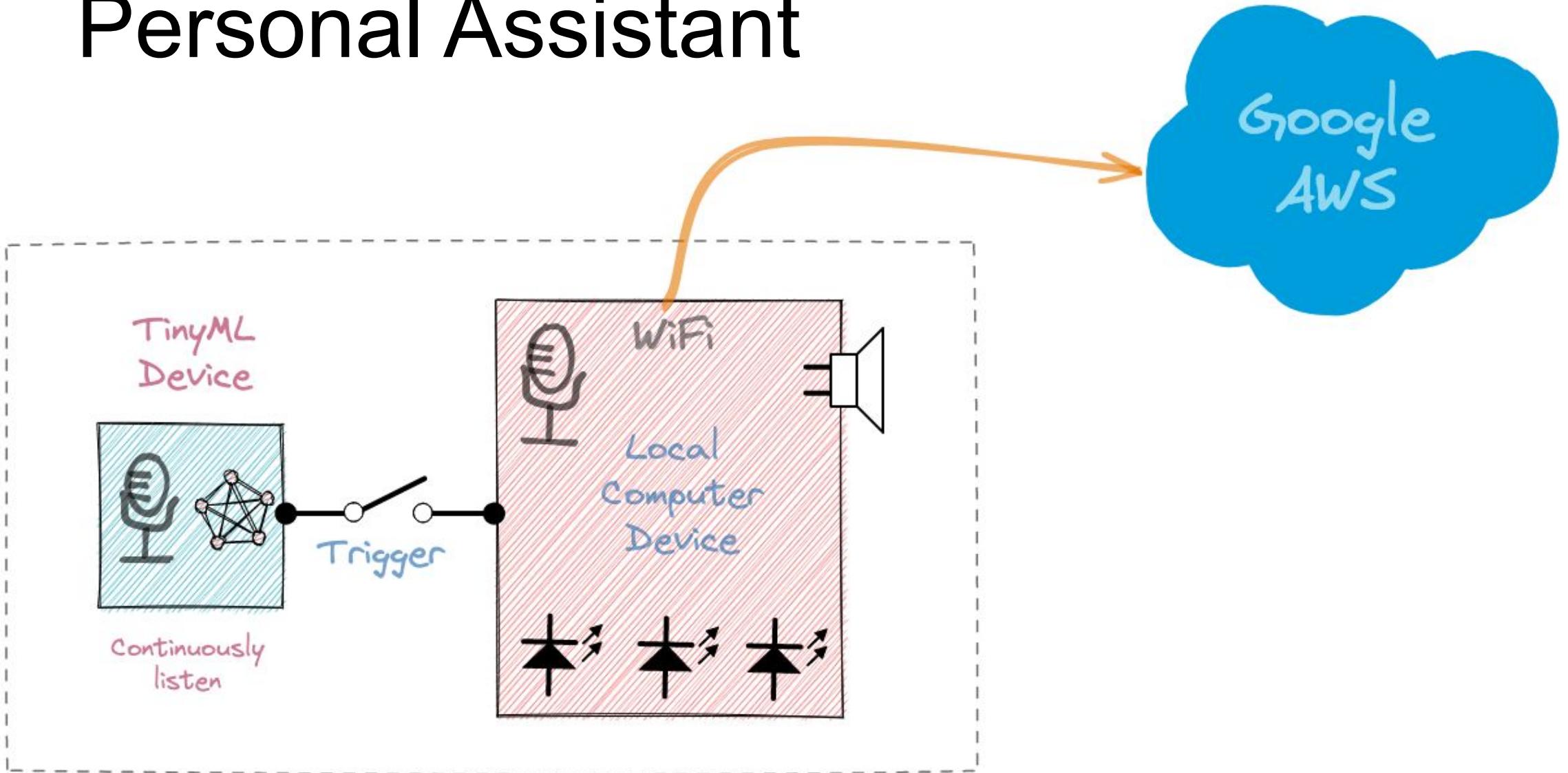
# Vibration



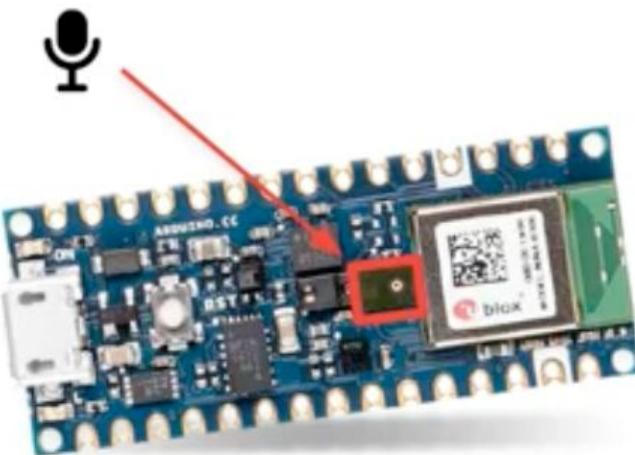
# Vision



# Personal Assistant



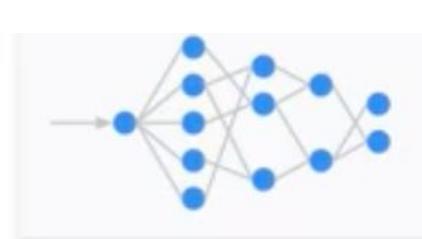
# “Cascade” Detection: multi-stage model



- 1 Continuously listen on the microcontroller

2

- Process the data with **TinyML** at the edge



3

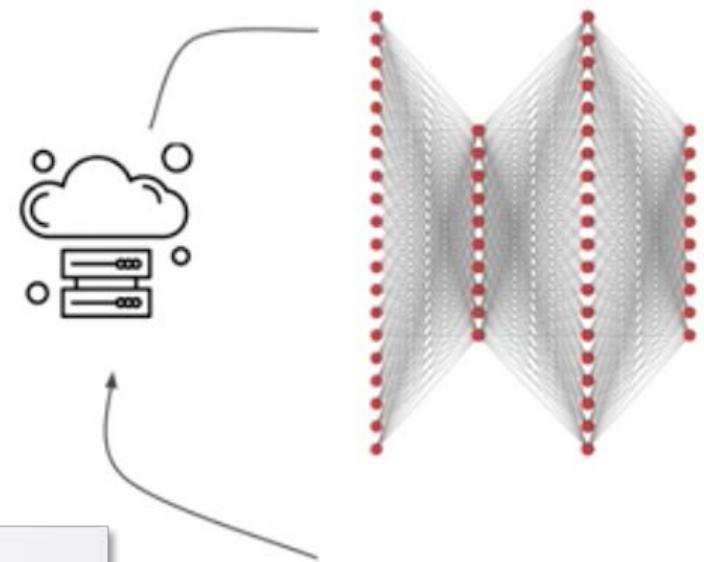
- Process on a secondary larger model on a larger local device



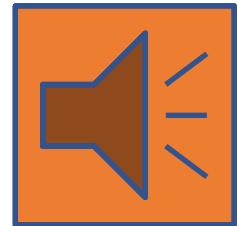
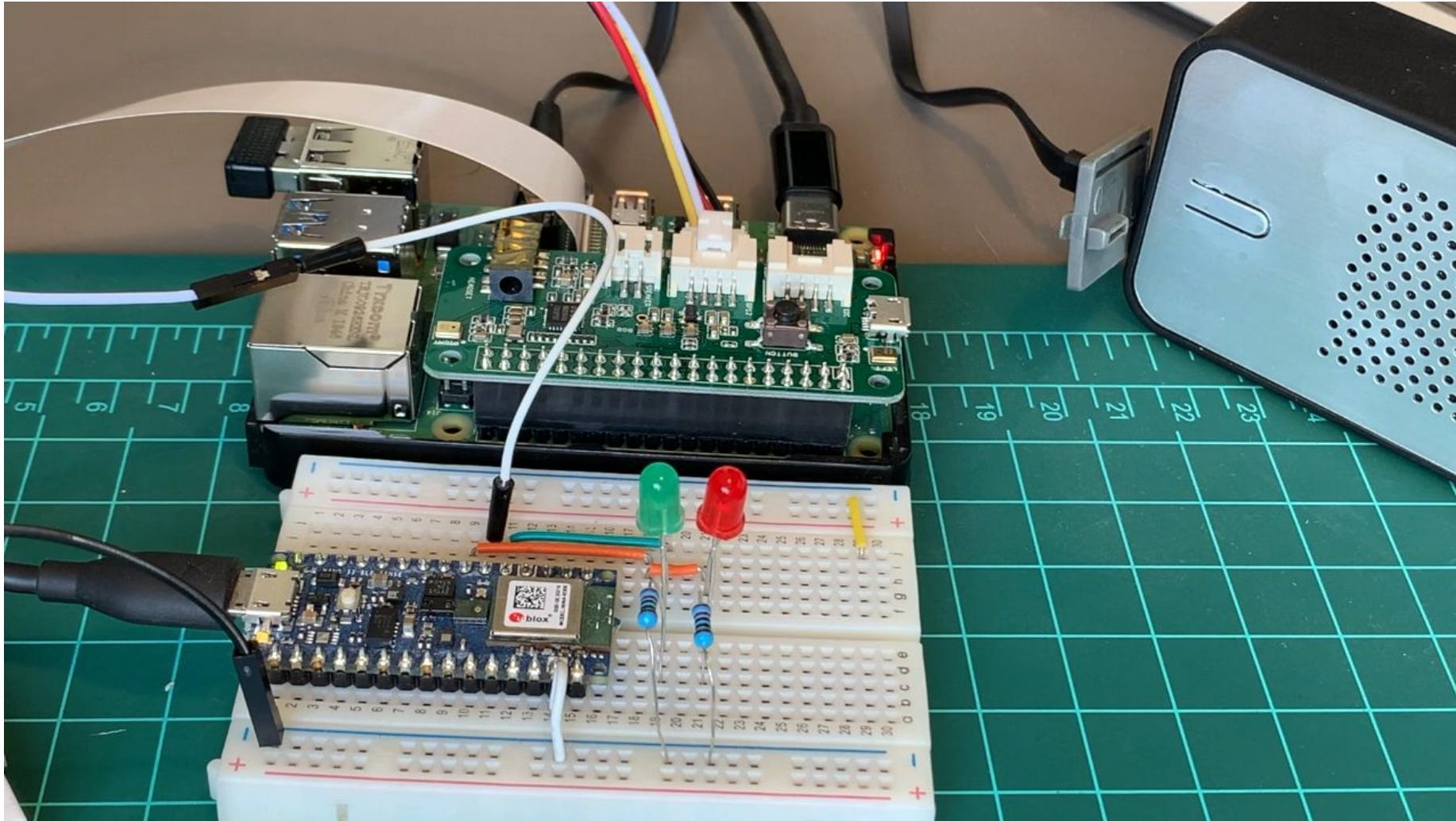
4

- Send the data to the cloud when triggered

- 5 Process the full speech data with a large model in the cloud

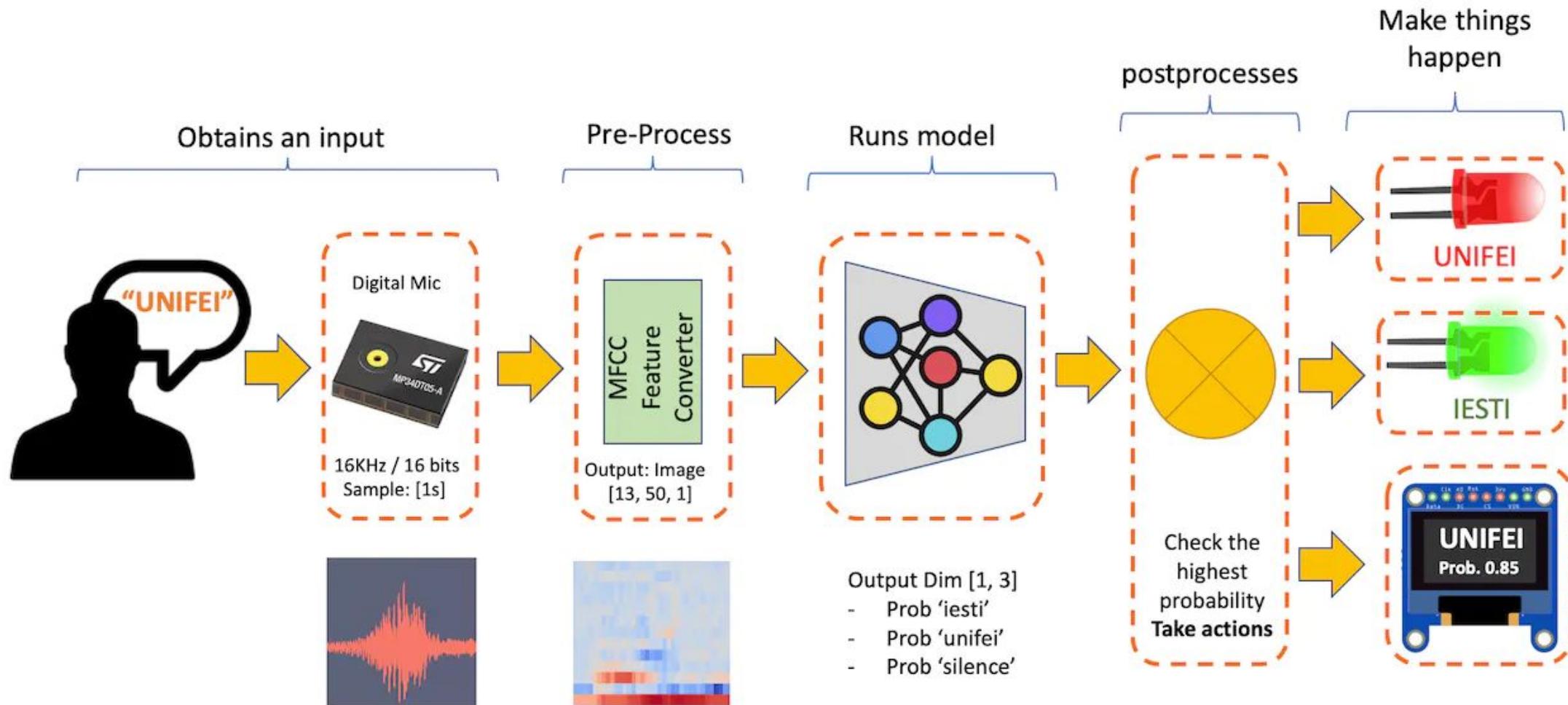


# KeyWord Spotting (KWS)



<https://mijrobot.org/2021/01/27/building-an-intelligent-voice-assistant-from-scratch/>

# KeyWord Spotting (KWS) - Inference



[TinyML Made Easy: Sound Classification \(KWS\)](#)

<https://youtu.be/1ex88hSqqyl>



Moez Altayeb  
University of Khartoum, Sudan  
ICTP, Trieste, Italy  
mohedahmed@hotmail.com

## ABSTRACT

Every year more than one billion people are infected and more than one million people die from vector-borne diseases including malaria, dengue, zika and chikungunya. Mosquitoes are the best known disease vector and are geographically spread worldwide. It is important to raise awareness of mosquito proliferation by monitoring their incidence, especially in poor regions. Acoustic detection of mosquitoes has been studied for long and ML can be used to automatically identify mosquito species by their wingbeat. We present a prototype solution based on an openly available dataset on the Edge Impulse platform and on three commercially-available TinyML devices. The proposed solution is low-power, low-cost and can run without human intervention in resource-constrained areas. This insect monitoring system can reach a global scale.

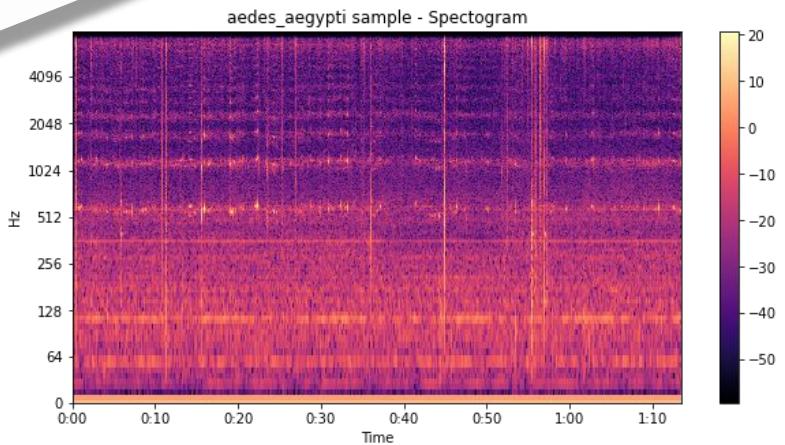
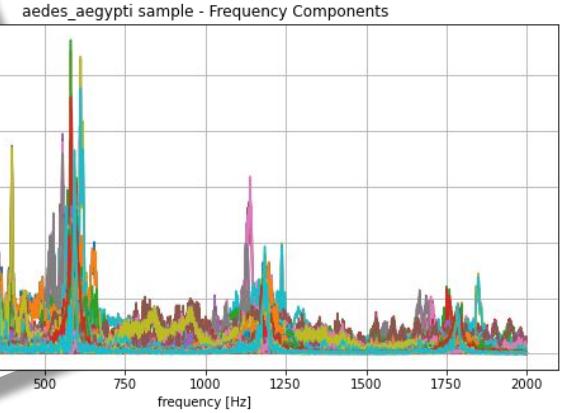
# Classifying mosquito wingbeat sound using TinyML

Marcelo Rovai  
Universidade Federal de Itajubá  
Itajubá, Brazil  
rovai@unifei.edu.br

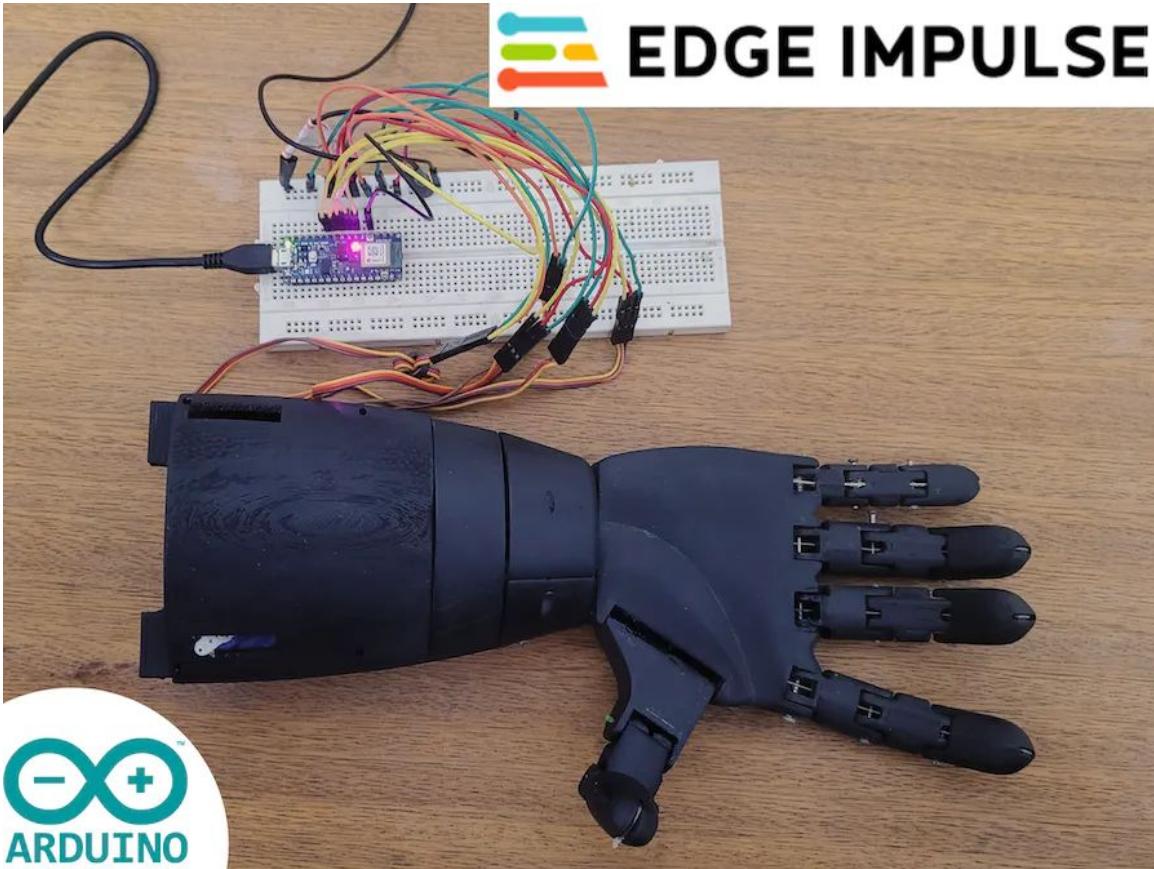
Marco Zennaro  
ICTP  
Trieste, Italy  
mzennaro@ictp.it

affected. People from poor communities with little access to health care and clean water sources are also at risk. Although anti-malarial drugs exist, there's currently no malaria vaccine. Vector-borne diseases also exacerbate poverty. Illness prevent people from working and supporting themselves and their families, impeding economic development. Countries with intensive malaria have much lower income levels than those that don't have malaria.

Countries affected by malaria turn to control rather than elimination. Vector control means decreasing contact between humans and disease carriers on an area-by-area basis. It is therefore of great interest to be able to detect the presence of mosquitoes in a specific area. This paper presents an approach based on TinyML and on embedded devices.



# Bionic Hand Voice Commands Module



EDGE IMPULSE

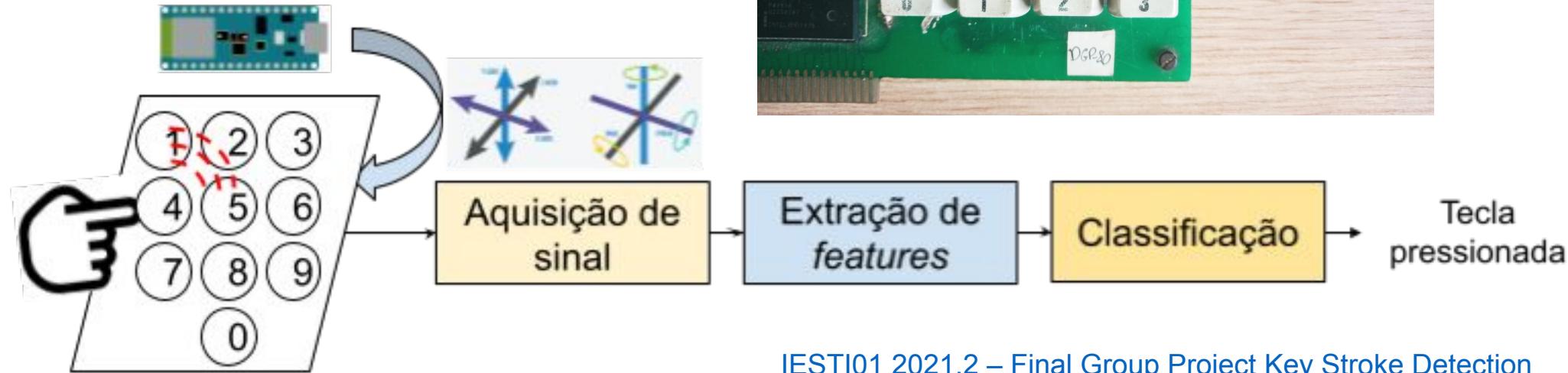


VIDEO



<https://www.hackster.io/ex-machina/bionic-hand-voice-commands-module-w-edge-impulse-arduino-aa97e3>

# Keystroke **Sound** Detection



[IESTI01 2021.2 – Final Group Project Key Stroke Detection](#)



**Renam Castro**  
Professor IFESP

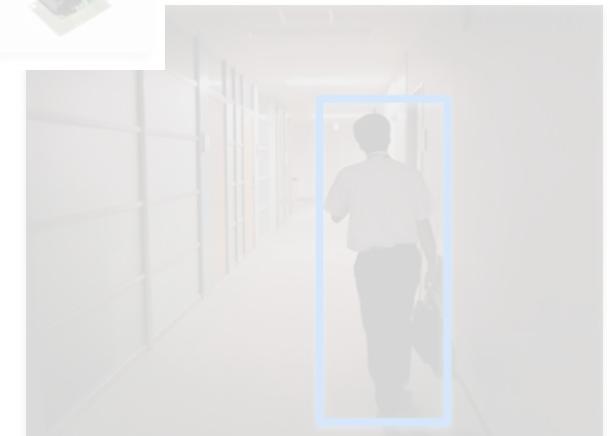
# Sound



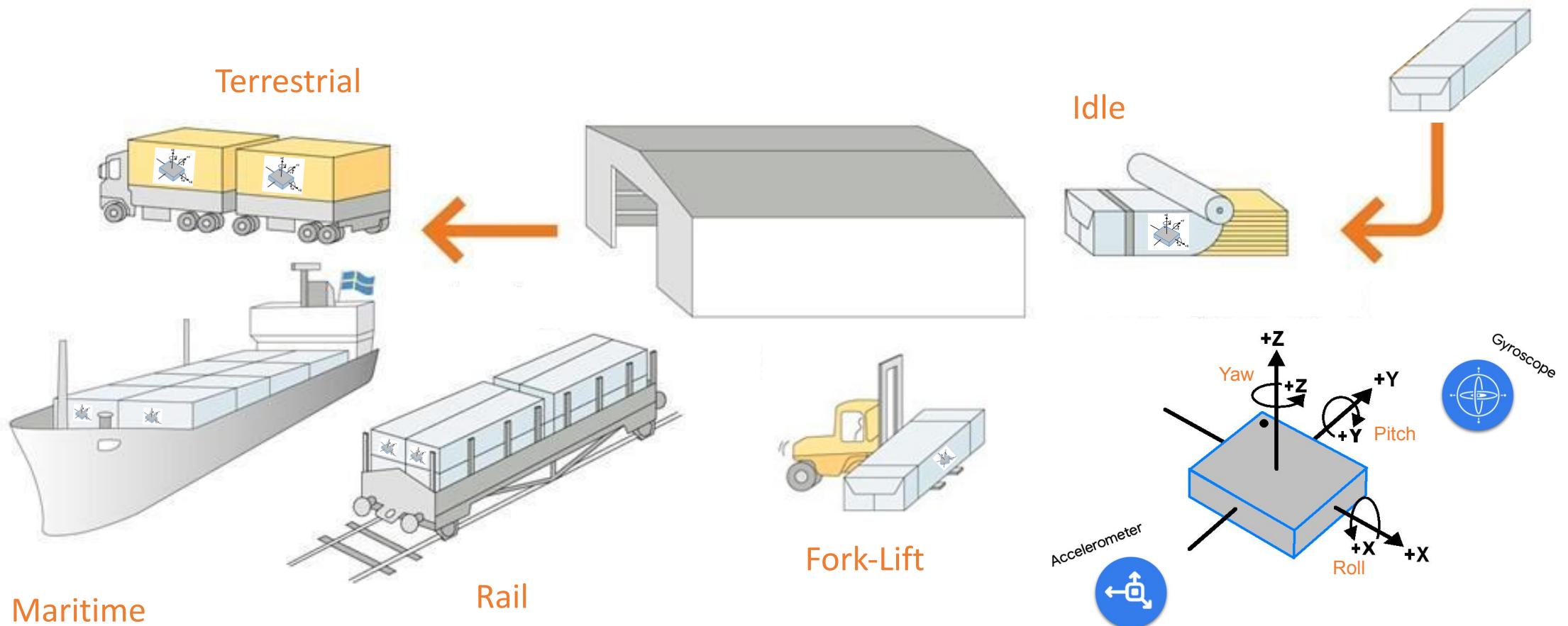
# Vibration



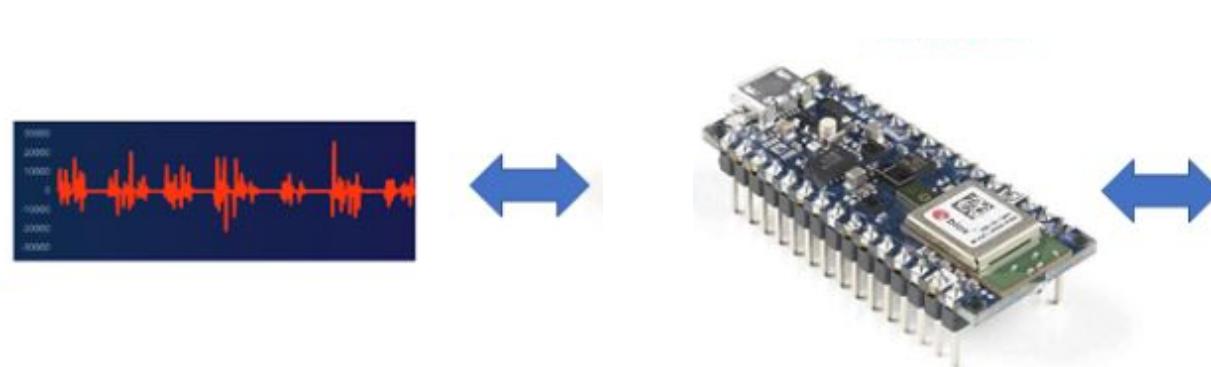
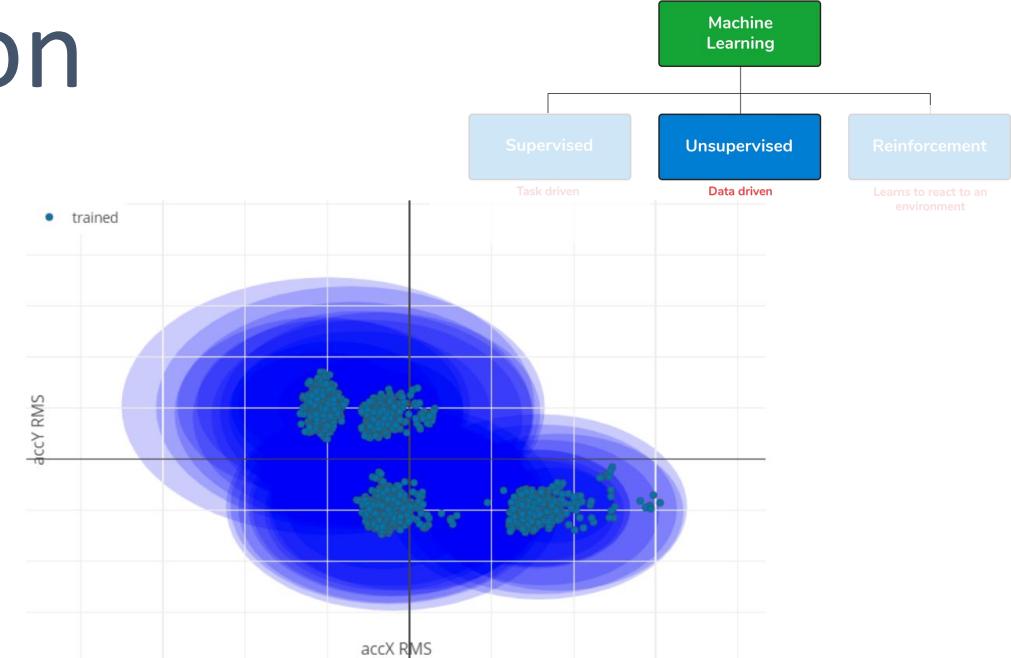
# Vision



# Mechanical Stresses in Transport



# Industrial – Anomaly Detection



IESTI01 2021.2 - Final Group Project: Bearing Failure Detection

# Predict and classify common Elephant behavior



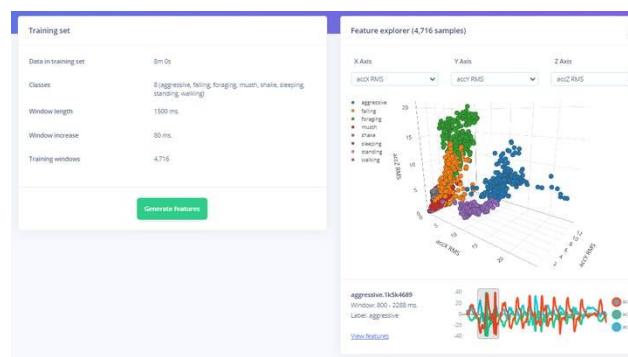
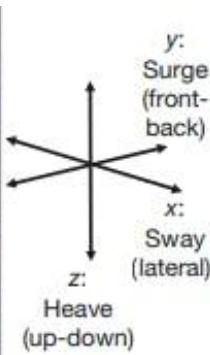
Aggressive



Standing



Sleeping



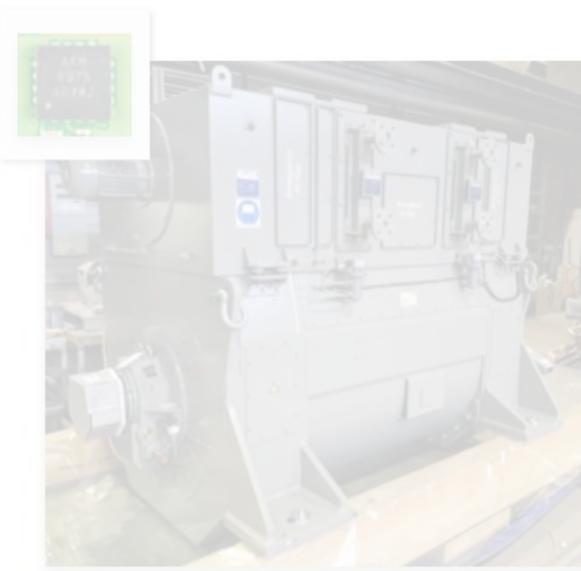
[https://www.hackster.io/dhruvsheth\\_electet-tinyml-and-iot-based-smart-wildlife-tracker-c03e5a](https://www.hackster.io/dhruvsheth_electet-tinyml-and-iot-based-smart-wildlife-tracker-c03e5a)



# Sound



# Vibration



# Vision



# Computer Vision Main Types

## Image Classification (Multi-Class Classification)

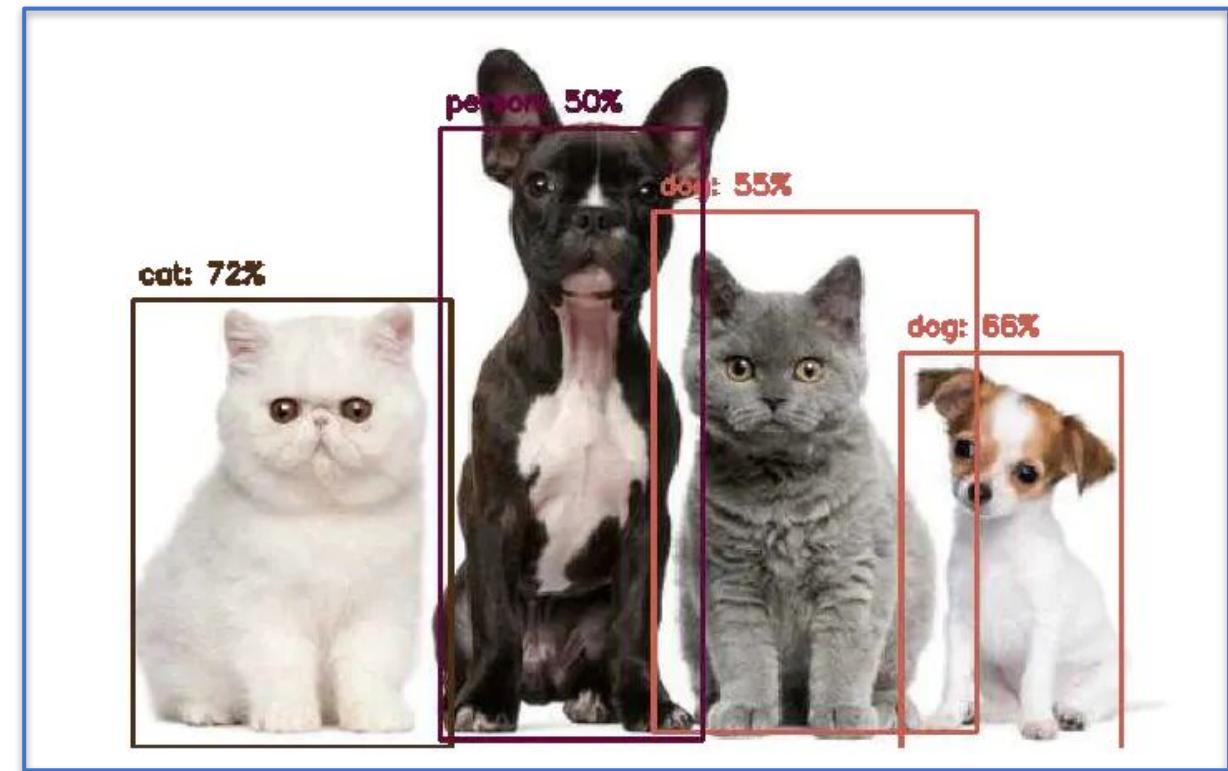


Cat: 70%



Dog: 80%

## Object Detection Multi-Label Classification + Object Localization



# Computer Vision Main Types

## Image Classification (Multi-Class Classification)

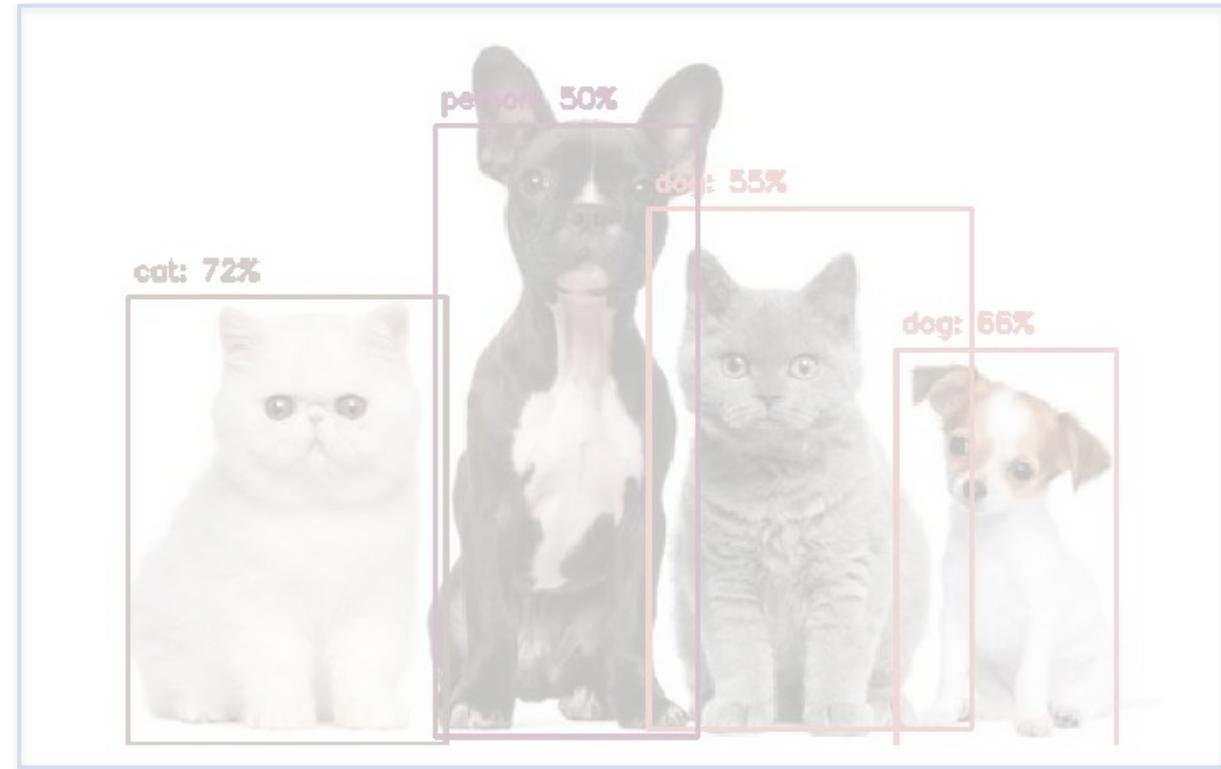


Cat: 70%



Dog: 80%

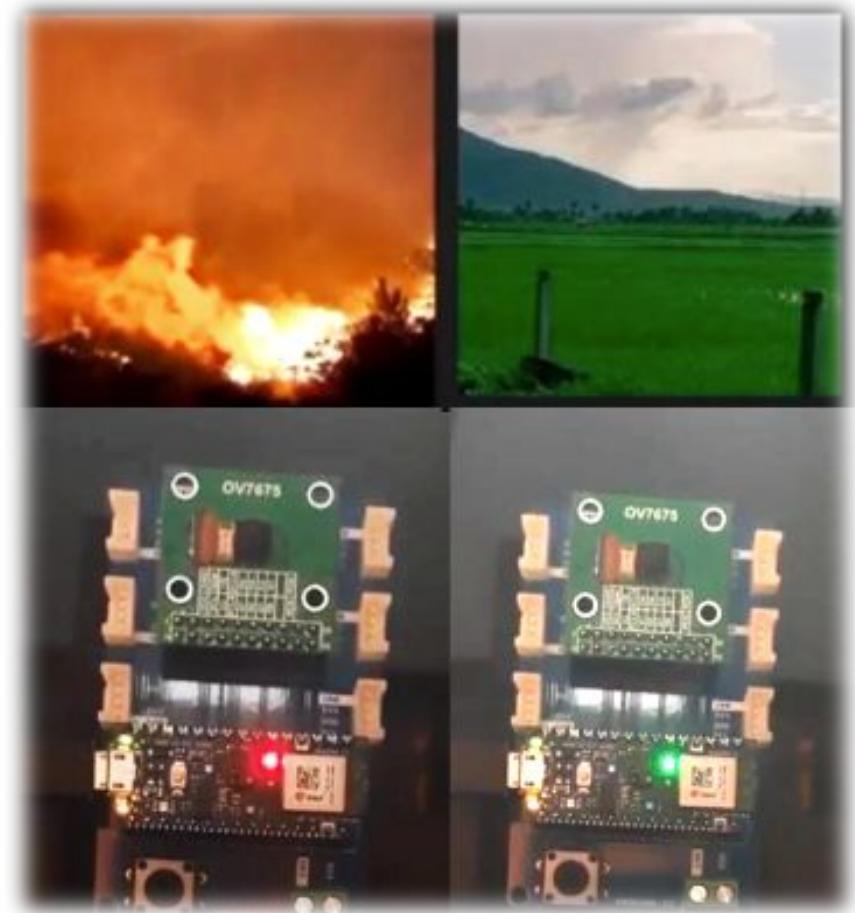
## Object Detection Multi-Label Classification + Object Localization



# Forest Fire Detection



[TinyML Aerial Forest Fire Detection](#)



[IESTI01 - Forest Fire Detection – Proof of Concept](#)

# Coffee Disease Classification



<https://www.hackster.io/Yukio/coffee-disease-classification-with-ml-b0a3fc>

**Introdução**

O Brasil é responsável por 50% do café exportado globalmente, o que é uma atividade importante para o país; geralmente a análise e classificação de doenças em plantas é feita manualmente, que não são acessíveis para pequenos produtores.

Com o aumento do poder de processamento das placas-mãe microcontroladas e processadores dedicados ao machine learning, a tarefa de embarcar todos morais tem-se tornado positiva em diversas áreas.



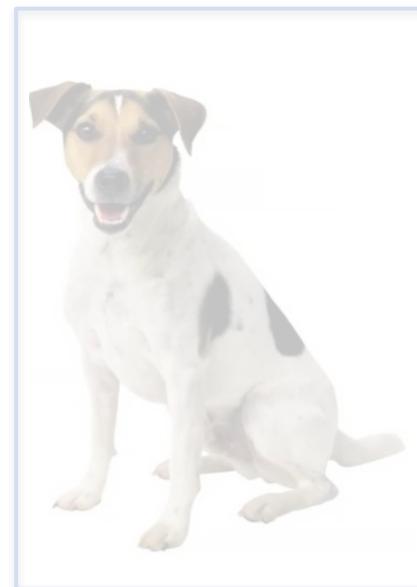
**João Vitor Yukio Bordin Yamashita**  
Graduando em Engenharia Eletrônica pela UNIFEI

# Computer Vision Main Types

**Image Classification**  
(Multi-Class Classification)

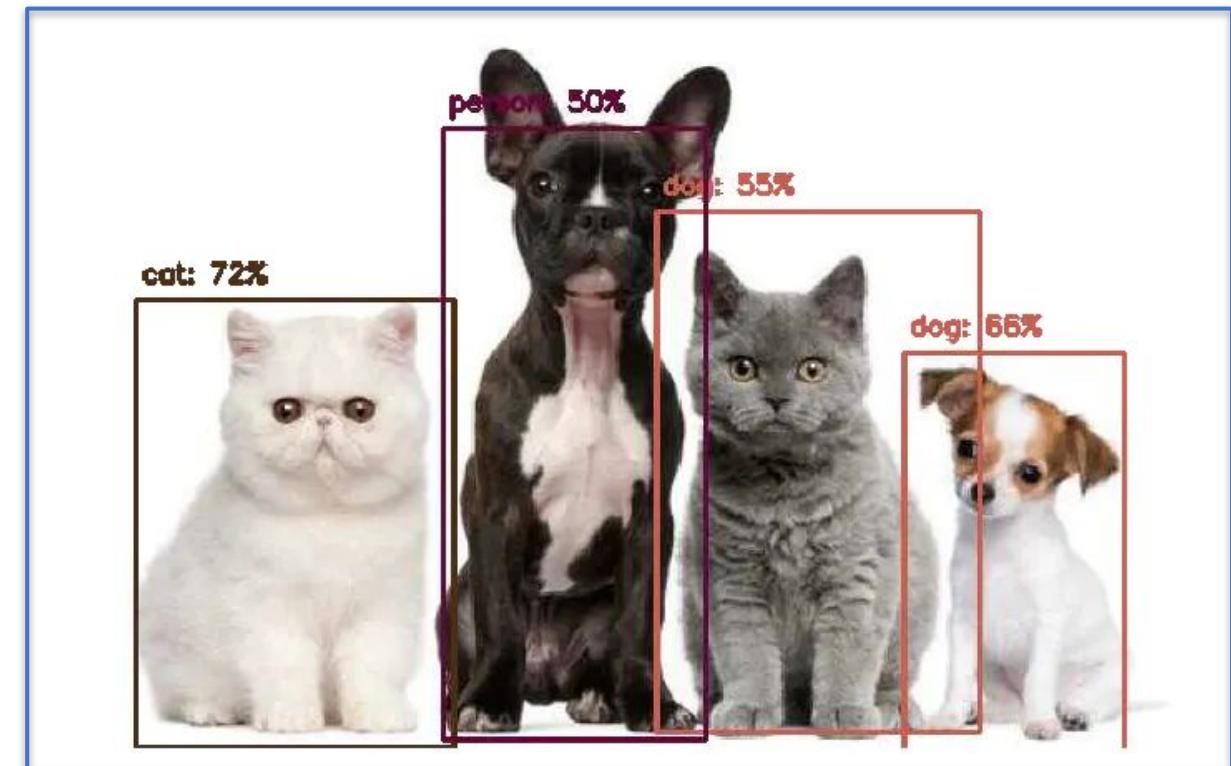


Cat: 70%

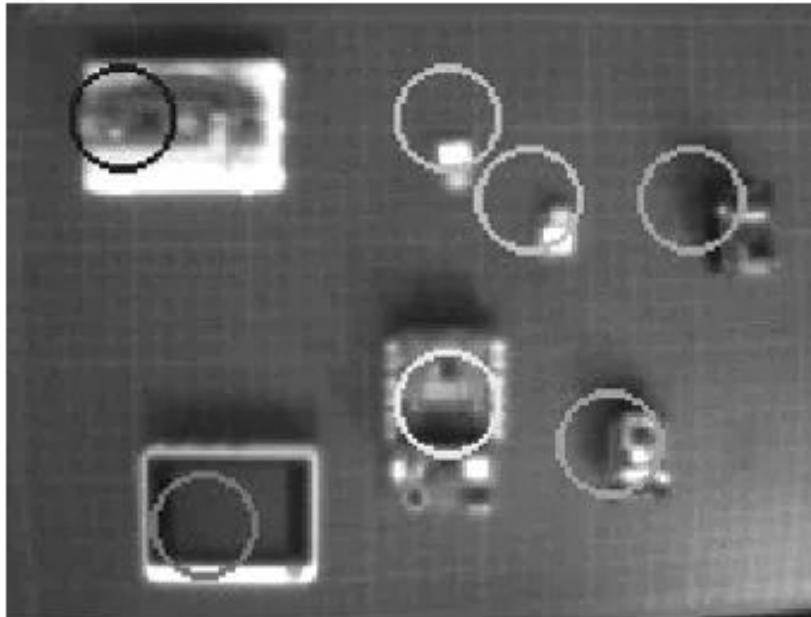


Dog: 80%

**Object Detection**  
**Multi-Label Classification + Object Localization**



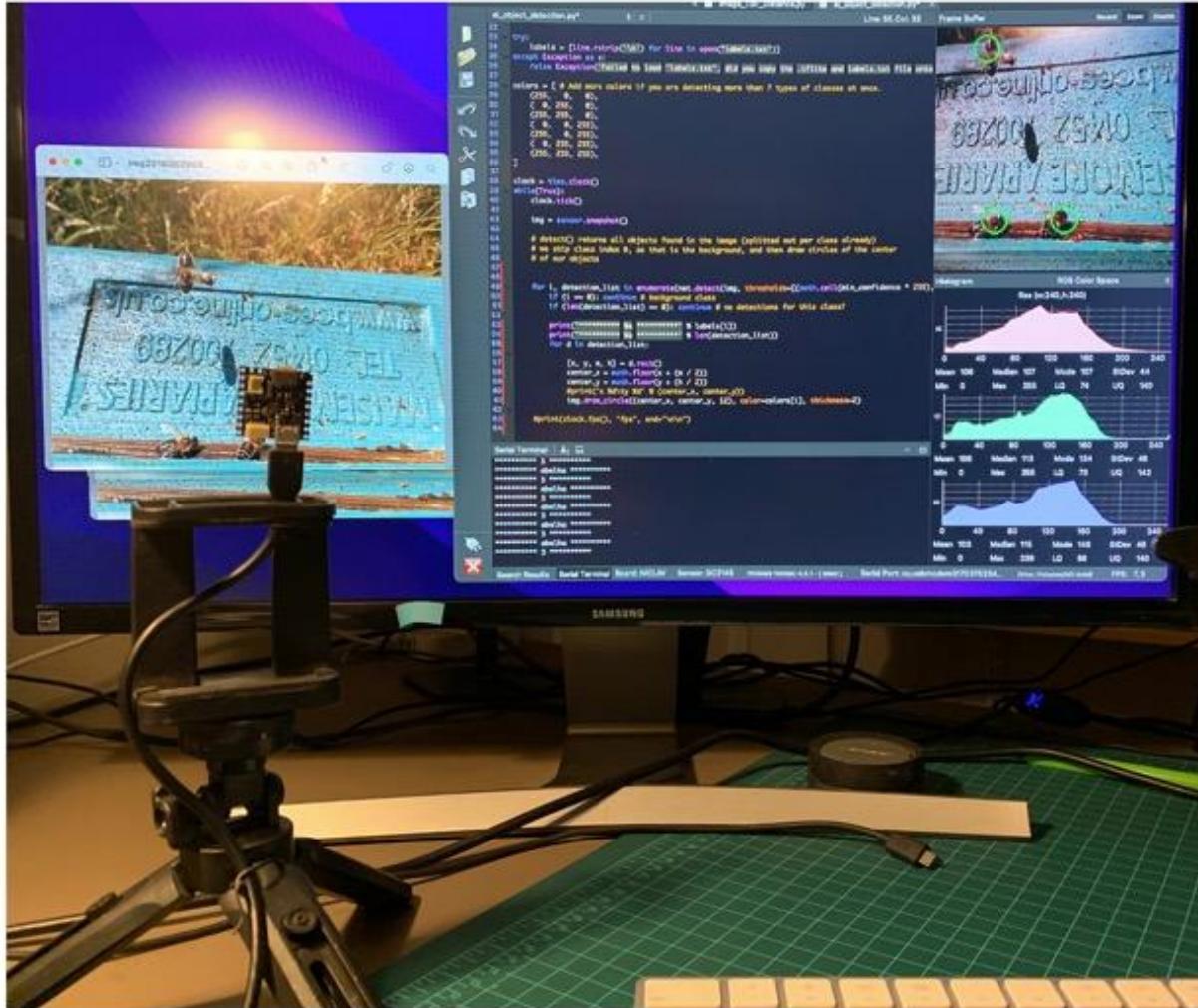
# Detecting Objects using TinyML (FOMO)



```
***** espcam *****
x 70  y 150
x 130  y 170
*****
***** nano *****
x 70  y 110
*****
***** pico *****
x 150  y 30
*****
***** wio *****
x 50  y 50
*****
***** xiao *****
x 150  y 110
x 130  y 130
6.97512 fps
```

[EdgeAI made simple - Exploring Image Processing \(Object Detection\) on microcontrollers with Arduino Portenta, Edge Impulse FOMO, and OpenMV](#)

# Detecting Objects using TinyML (FOMO)

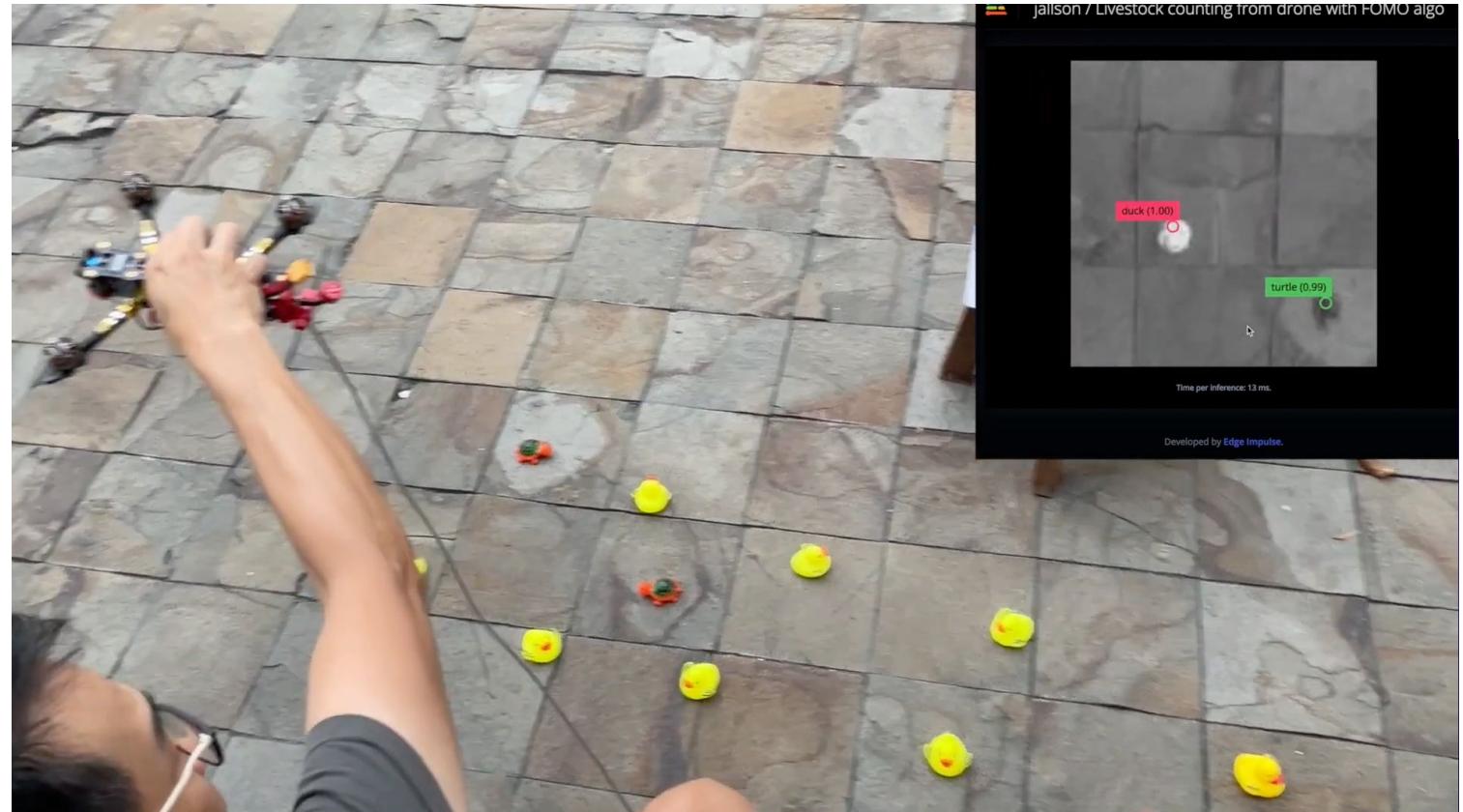
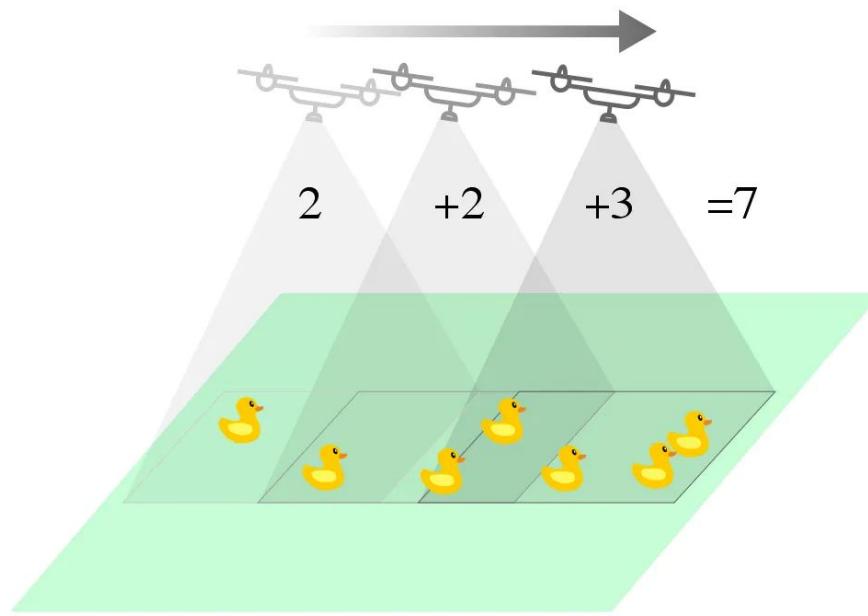


MicroPython



<https://youtu.be/MYuc3QISquw>

# Livestock / Wildlife Counting from Drone with FOMO

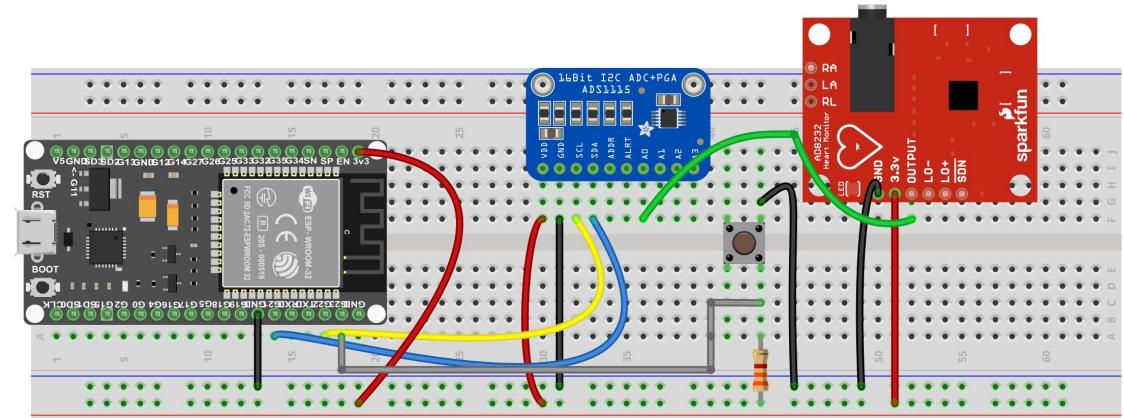


<https://www.hackster.io/jallsonsuryo/livestock-wildlife-counting-from-drone-with-fomo-algorithm-a2f734>

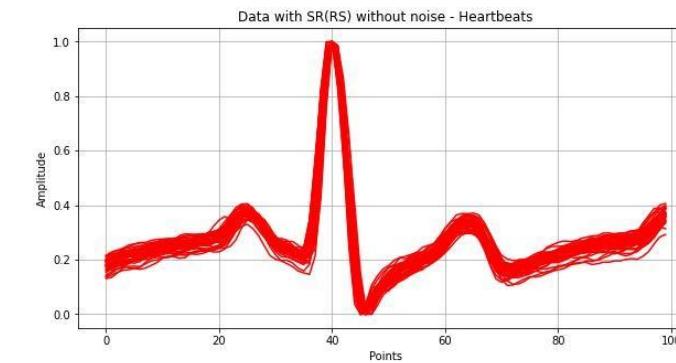
# Other Sensors / MCUs / Models

## Examples

# AD8232 - Single Lead Heart Rate Monitor



fritzing



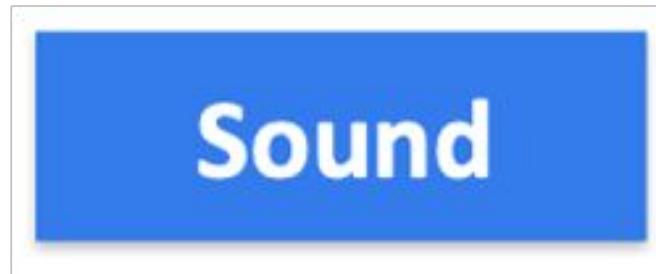
**Guilherme Silva**  
Engenheiro - UNIFEI

[Atrial Fibrillation Detection on ECG using TinyML](#)  
Silva et al. UNIFEI 2021

# Other TinyML / MCUs Project Examples



- Image Classification with [ESP32-CAM](#) [\[Doc\]](#)
- Image Classification with [Portenta H7](#) [\[Doc\]](#)
- Object Detection with [Portenta H7](#) [\[Doc\]](#)

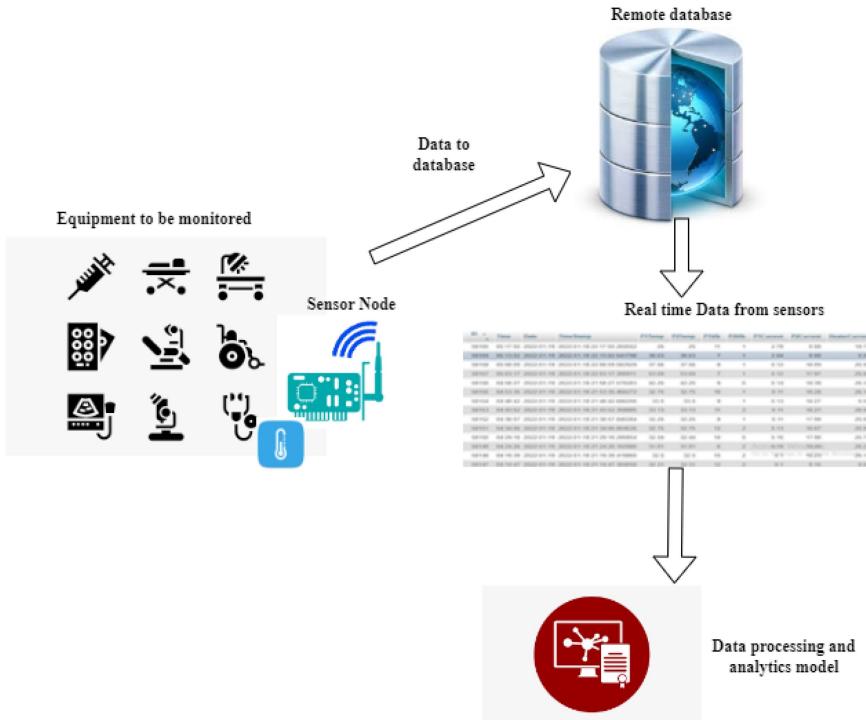
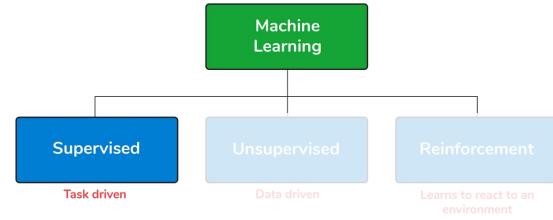


- Listening Temperature with [Nano 33](#) [\[Doc\]](#)
- COPD Detection with [Nano 33](#) [\[Doc\]](#)
- Sound Classification with [XIAO BLE Sense](#) [\[Doc\]](#)

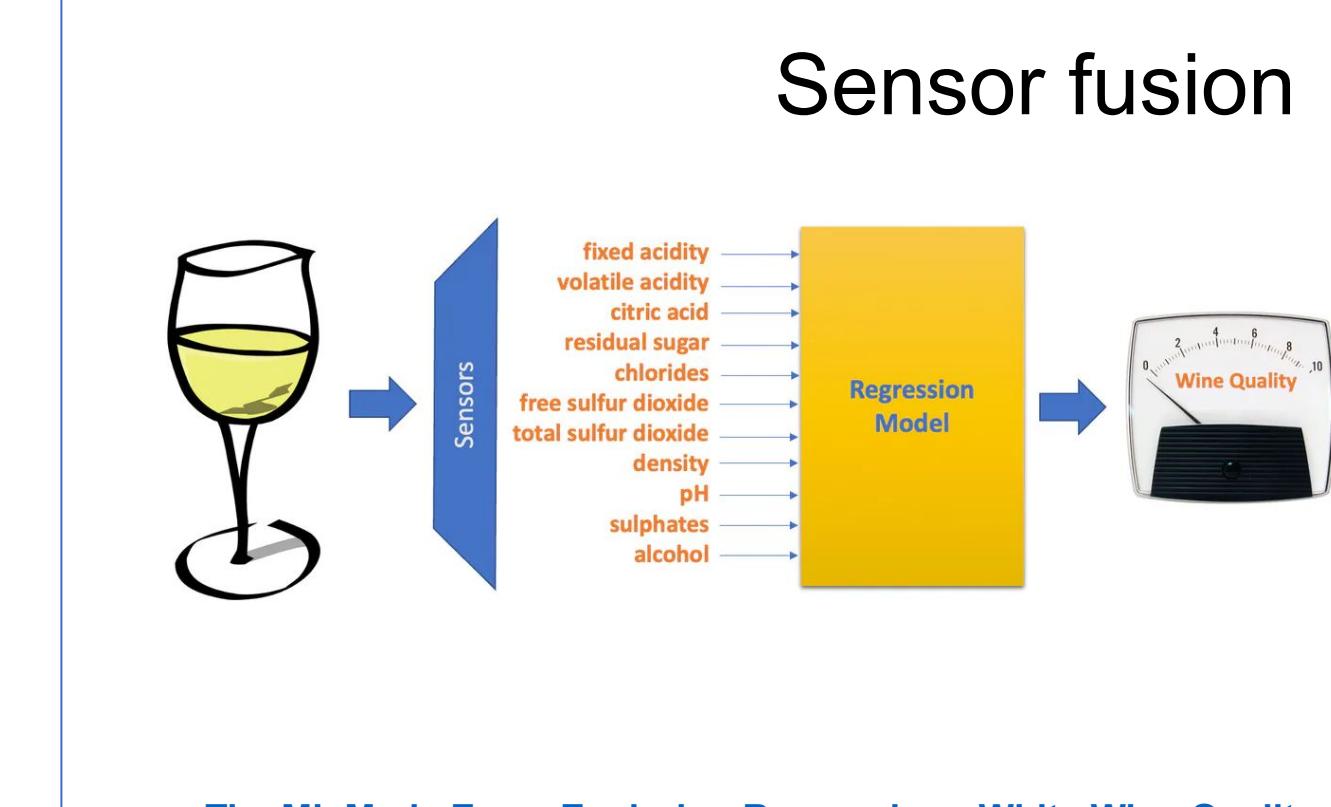


- Motion Recognition with [RPi Pico](#) [\[Doc\]](#)
- Gesture Recognition with [Wio Terminal](#) [\[Doc\]](#)
- Anomaly Detection with [XIAO BLE Sense](#) [\[Doc\]](#)

# Regression on TinyML

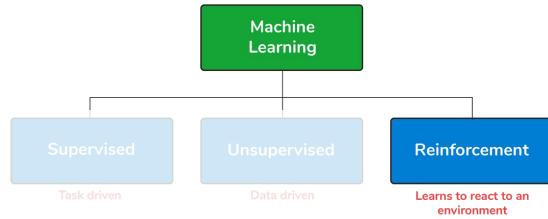


[On-Device IoT-Based Predictive Maintenance Analytics Model: Comparing TinyLSTM and TinyModel from Edge Impulse](#)



[TinyML Made Easy: Exploring Regression - White Wine Quality](#)

# Reinforcement on TinyML



## Deep Reinforcement Learning for Autonomous Source Seeking on a Nano Drone

Bardienus P. Duisterhof<sup>1,3</sup> Srivatsan Krishnan<sup>1</sup> Jonathan J. Cruz<sup>1</sup> Colby R. Banbury<sup>1</sup> William Fu<sup>1</sup>  
Aleksandra Faust<sup>2</sup> Guido C. H. E. de Croon<sup>3</sup> Vijay Janapa Reddi<sup>1,4</sup>

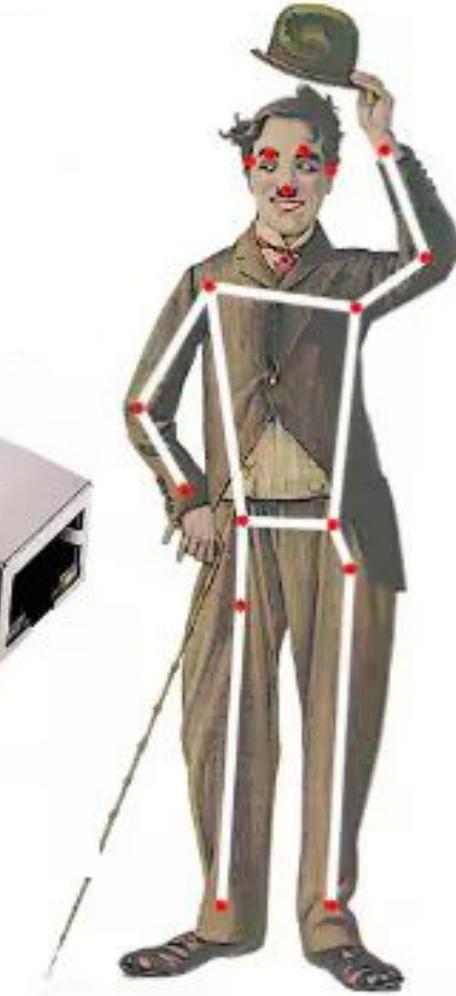
<sup>1</sup>Harvard University, <sup>2</sup>Robotics at Google, <sup>3</sup>Delft University of Technology, <sup>4</sup>The University of Texas at Austin



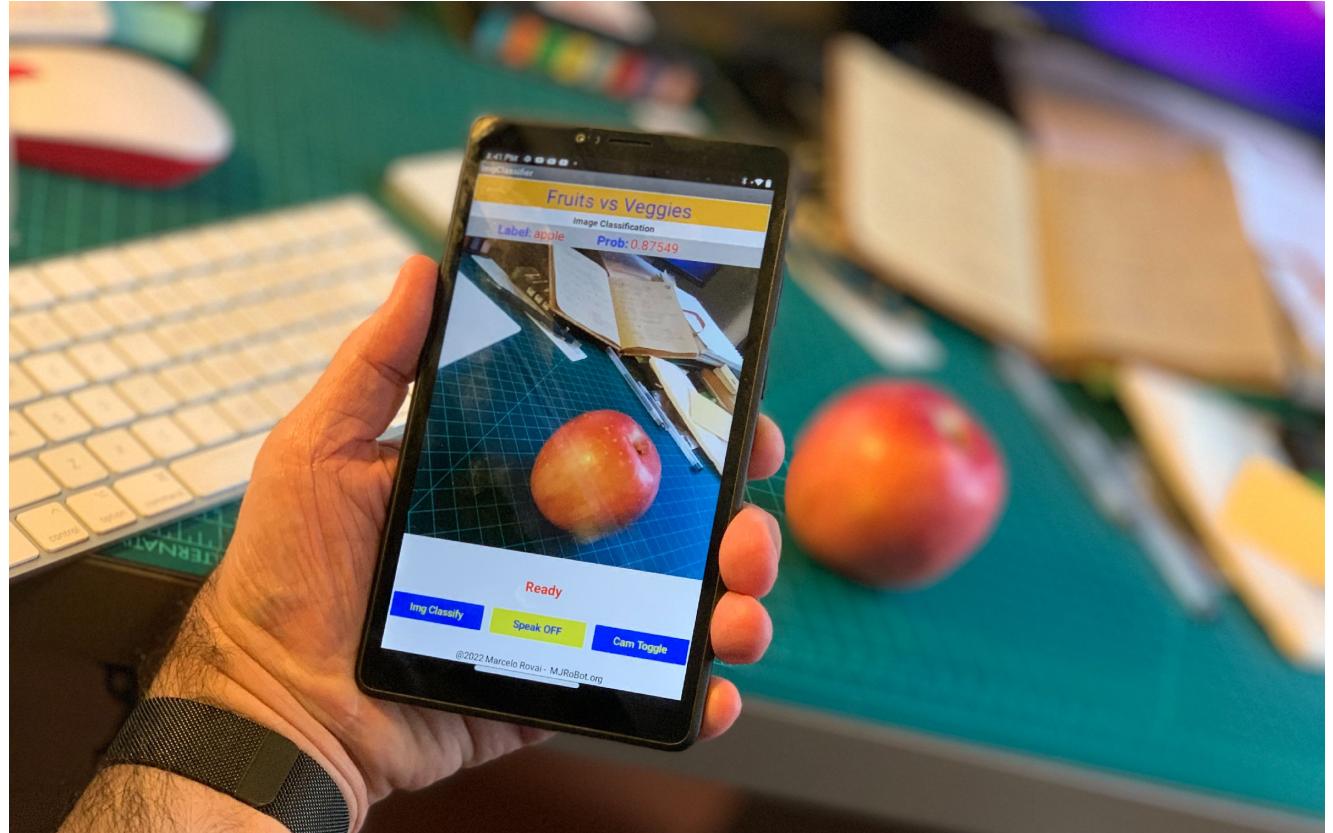
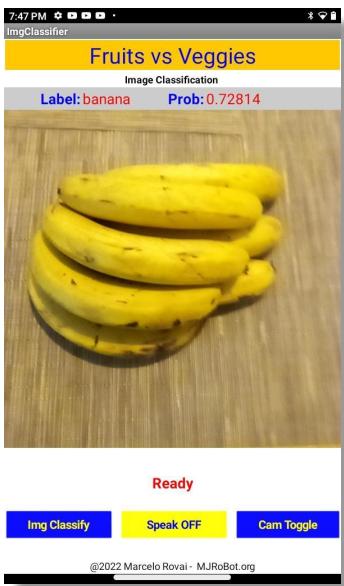
<https://arxiv.org/abs/1909.11236>

<https://youtu.be/wmVKbX7MOnU>

# Exploring AI at the edge (Computer Vision)



# Classifying Images using Smartphones



<https://www.hackster.io/mjrobot/app-inventor-edgeml-image-classification-fruit-vs-veggies-b671da>

# Projects by Students (UNIFEI – IESTI01)

- **Sound:**
  - Earthquake detection
  - Covid Detection (cough)
  - Key Detection
  - Pulmonary Disease
  - Snore Detection
  - Bionic Hand Control
- **Image:**
  - Mask Detection
  - Forest Fire Detection
  - Helmet Detection
  - Water Consumption (hydrometer)
  - Sign Language
  - Coffee Disease Classification
  - Bee Counting
- **Other Sensors:**
  - Bionic Hand – Finger Detection
  - Electric Charges
  - ECG – Fibrial Atrilation detection
- **Vibration:**
  - Personal Trainer
  - Bearing – Anomaly Detection

# TinyML4D Show & Tell Presentations

Date	Thread	Video
March 30th, 2023	Thread here when ready <a href="#">thread here</a>	Video here when ready
Feb 23rd, 2023	<a href="#">thread here 17</a>	Video here when ready
January 26th, 2023	<a href="#">thread here 2</a>	<a href="https://youtu.be/-0xRZ-5UYUc">9</a>
December 1st, 2022	<a href="#">thread here 2</a>	<a href="https://youtu.be/e49pkjnIMIQ">8</a>
October 27th, 2022		<a href="https://youtu.be/s8_hKpOWUwY">1</a>

## TinyML4D Academic Network Show and Tell Main Index.

The TinyML4D Academic Network Students should use this form to sign up for the latest presentations.

[2](https://forms.gle/ic52HZMqVv4pBrkP7)

The Show and Tell are typically held at 2 pm UTC on the last Thursday of each month and will take place in this Zoom room.

[1](https://zoom.us/j/95229860797)

Meeting ID: 952 2986 0797

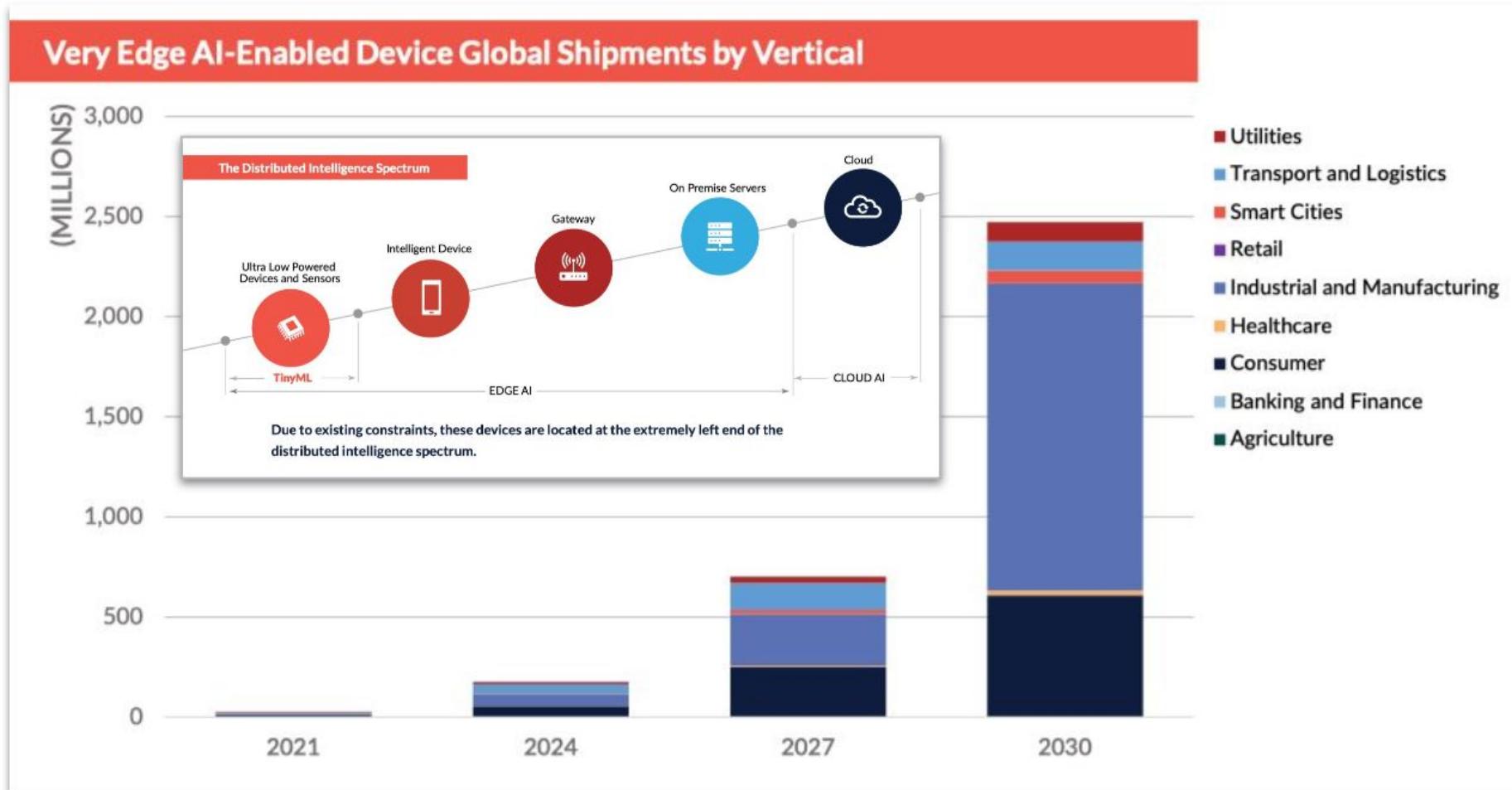
Passcode: 141278

# To learn more about Edge AI

- IESTI01 TinyML - Machine Learning for Embedding Devices (Videos: Pt)
- WALC 22 – Applied AI - TinyML (Videos in Spanish)
- Professional Certificate in Tiny Machine Learning (TinyML) – edX/Harvard
- Introduction to Embedded Machine Learning - Coursera/Edge Impulse
- Computer Vision with Embedded Machine Learning - Coursera/Edge Impulse
- "Deep Learning with Python" book by François Chollet
- "TinyML" book by Pete Warden, Daniel Situnayake
- "TinyML Cookbook" by Gian Marco Iodice
- "AI at the Edge" book by Daniel Situnayake, Jenny Plunkett

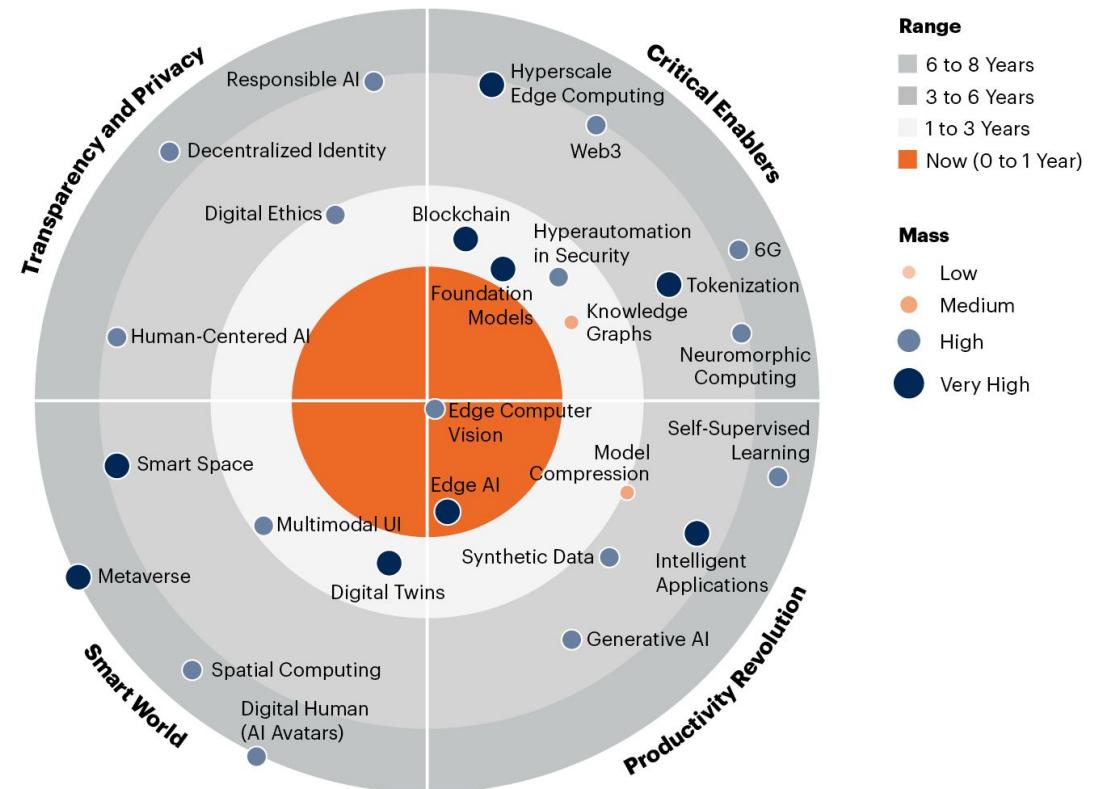
# The Future of the EdgeAI

# Massive Potential for Impact



Source: ABI Research: TinyML

# 2023 Gartner Emerging Technologies and Trends Impact Radar



gartner.com

Note: Range measures number of years it will take the technology/trend to cross over from early adopter to early majority adoption. Mass indicates how substantial the impact of the technology or trend will be on existing products and markets.

Source: Gartner  
© 2023 Gartner, Inc. All rights reserved. CM\_GTS\_2034284

Gartner®

Edge AI has a very high impact potential, and it is for now!

# Conclusion

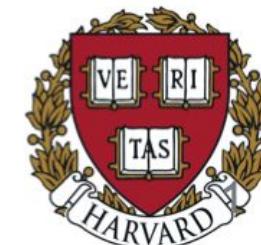
---



---

## The Future of ML is Tiny and Bright

*Vijay Janapa Reddi, Ph. D. | Associate Professor |  
John A. Paulson School of Engineering and Applied Sciences | Harvard University |*



# Thanks



# UNIFEI