MoroccoAI AI Summer School

جامعة الأخوين
ⵜⴰⵙⴷⴰⵡⵉⵜ ⵏ ⵓⵅⵡⵉⵏ
AL AKHAWAYN UNIVERSITY

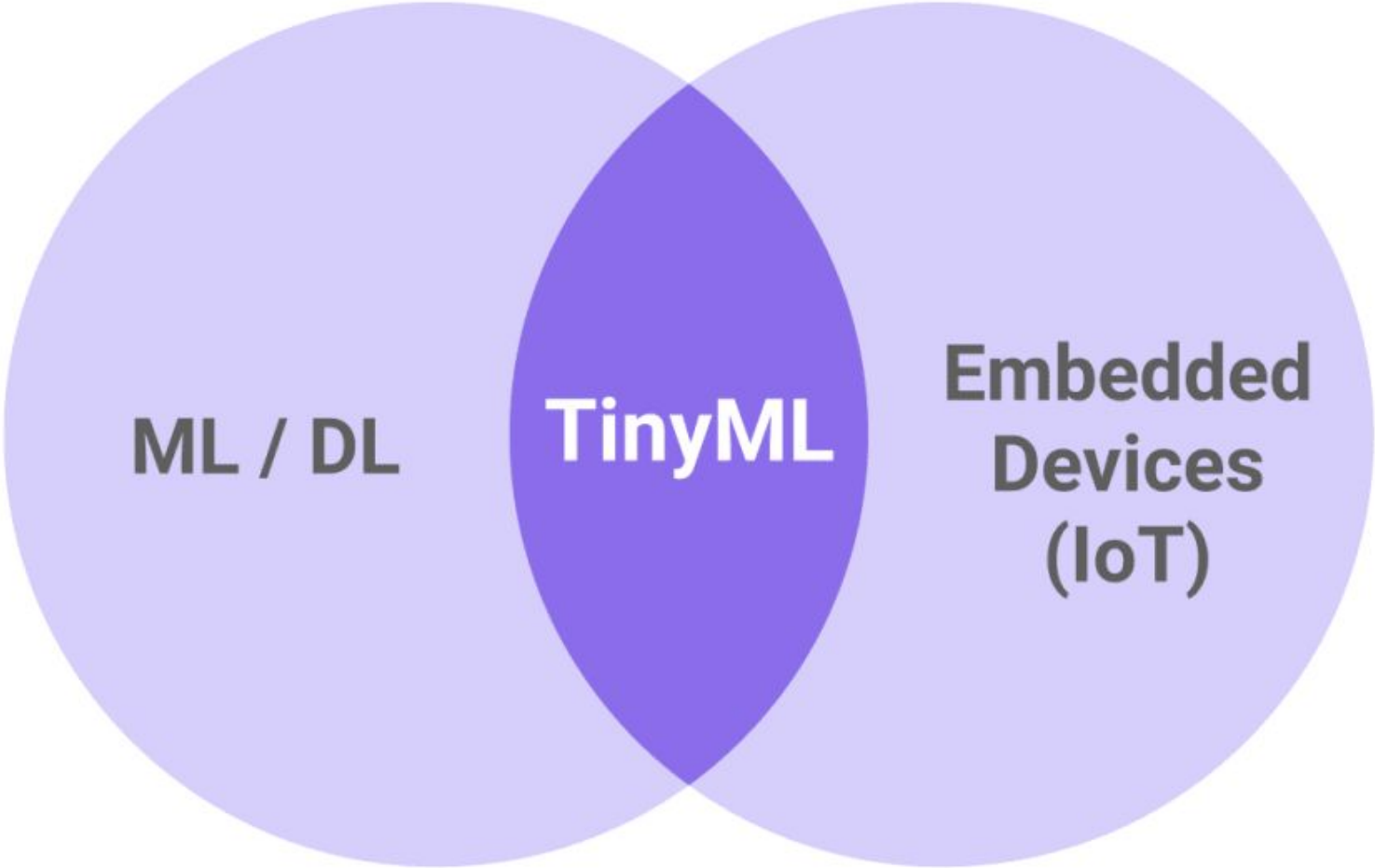
The Big Impact of TinyML:
Embedded Machine Learning at the Extreme Edge

Brian Plancher
Barnard College, Columbia University
brianplancher.com

The Big Impact of TinyML:
Embedded Machine Learning at the Extreme Edge

1. What is **Embedded (Tiny)** Machine Learning?
2. **Beneficial Applications** of Embedded Machine Learning
3. **Challenges** of Deploying ML at the Extreme Edge
4. **Sustainability & Privacy:** MLSensors at Scale

What is Embedded (Tiny) Machine Learning?



Beneficial Applications of Embedded Machine Learning

Zero Hunger & Good Health and Well-Being
(SDG #2 & #3)

Credit: PlantVillage Nuru

Nuru, an ML app more accurate than humans at detecting plant diseases. Increased a farmer's sales by 55% & **yields by 146%**.

Credit: Crop Angel Ltd

Tiny drones can provide targeted pesticide applications that **reduce use to 0.1%** of conventional blanket spraying.

Credit: Sinhyu/Getty Images

Using Edge Impulse, a system was prototyped to identify mosquitoes by wing beats sounds with **88.3% accuracy**.

Zero Hunger & Good Health and Well-Being
(SDG #2 & #3)

https://data.mendeley.com/datasets/zp4nf2dxbh/1

Poultry Vocalization Signal Dataset for Early Disease Detection

Published: 12 June 2023 | Version 1 | DOI: 10.17632/zp4nf2dxbh.1
Contributors: Halleluyah Aworinde, Segun Adebayo, Akinwale Akinwunmi, Olufemi Alabi, Adebamiji Ayandiji, Olaide Oke, Abel Oyebamiji, Adetoye Adeyemo, Aderonke Sakpere, Kizito Echetama

Description

The aim of generating this dataset is to deliver open, accessible and quality machine learning dataset for Poultry farm management. The dataset contains a total of 346 audio signal files which are divided into three folders: healthy, noise and unhealthy. The healthy folder contains 139 audio files, noise folder has 86, while unhealthy folder has 121 files. The audio files are saved in .wav file format.

Download All 869 MB

Life on Land & Below Water

(SDG #14 & #15)

Credit: Rainforest Connection

Rainforest Connection uses **recycled smartphones** for **solar-powered** listening devices to warn of **deforestation** efforts

Credit: RESOLVE and Bivash Pandav

RESOLVE's AI camera transmits notifications of elephant detection and can **run for more than 1.5 years** on a single battery.

Credit: Tim Cole

To prevent collisions with whales in busy waterways, Google deployed a TinyML model on hydrophones to alert ships.

Life on Land & Below Water

(SDG #14 & #15)

Open Access Article

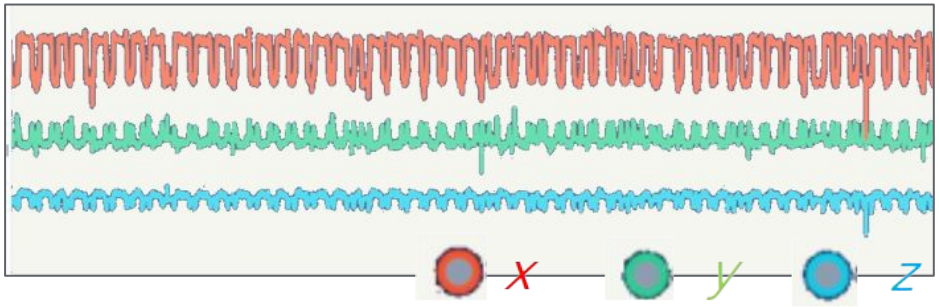
Design and Development of a Family of Integrated Devices to Monitor Animal Movement in the Wild

by Laila Daniela Kazimierski [1,*,†], Andrés Oliva Trevisan [2,3,†], Erika Kubisch [4], Karina Laneri [1] and Nicolás Catalano [2,3]

[1] Consejo Nacional de Investigaciones Científicas y Técnicas, Centro Atómico Bariloche (CONICET), Comisión Nacional de Energía Atómica (CNEA), San Carlos de Bariloche R8402AGP, Argentina
[2] Centro Atómico Bariloche, Comisión Nacional de Energía Atómica (CNEA), San Carlos de Bariloche R8402AGP, Argentina
[3] Instituto Balseiro, Universidad Nacional de Cuyo, Comisión Nacional de Energía Atómica (CNEA), San Carlos de Bariloche R8402AGP, Argentina
[4] Instituto de Investigaciones en Biodiversidad y Medioambiente, Consejo Nacional de Investigaciones Científicas y Técnicas (INIBIOMA, CONICET-Universidad Nacional del Comahue), San Carlos de Bariloche R8400AGP, Argentina
* Author to whom correspondence should be addressed.
† These authors contributed equally to this work.

https://www.mdpi.com/1424-8220/23/7/3684

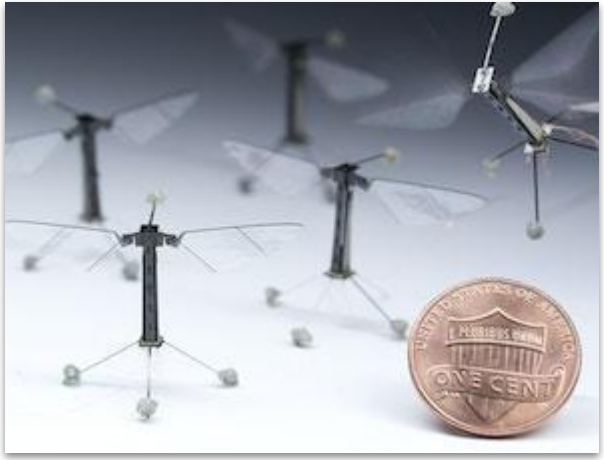
Example of accelerometer signal of a female digging a nest to lay eggs:



Climate Action
(SDG #13)

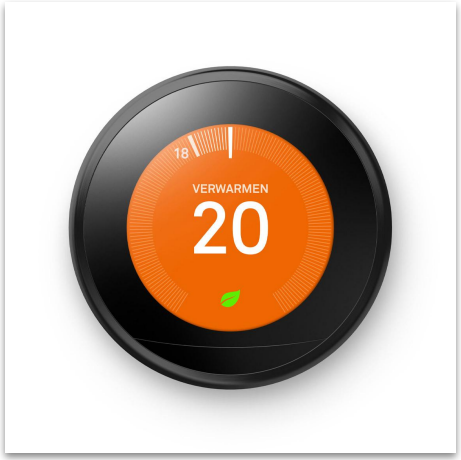
Credit: Ribbit Network

Ribbit Network is **crowdsourcing world's largest greenhouse gas emissions dataset** through distributed intelligent sensors



Credit: Wyss Institute at Harvard University

TinyML can help provide intelligence to **tiny robots like the Robobee** that can be used as artificial pollinators.



Credit: Google Nest

Smart HVAC systems show a **20-40% reduction in building energy** usage.

And many more!

Tiny Robot Learning: Challenges and Directions for Machine Learning in Resource-Constrained Robots

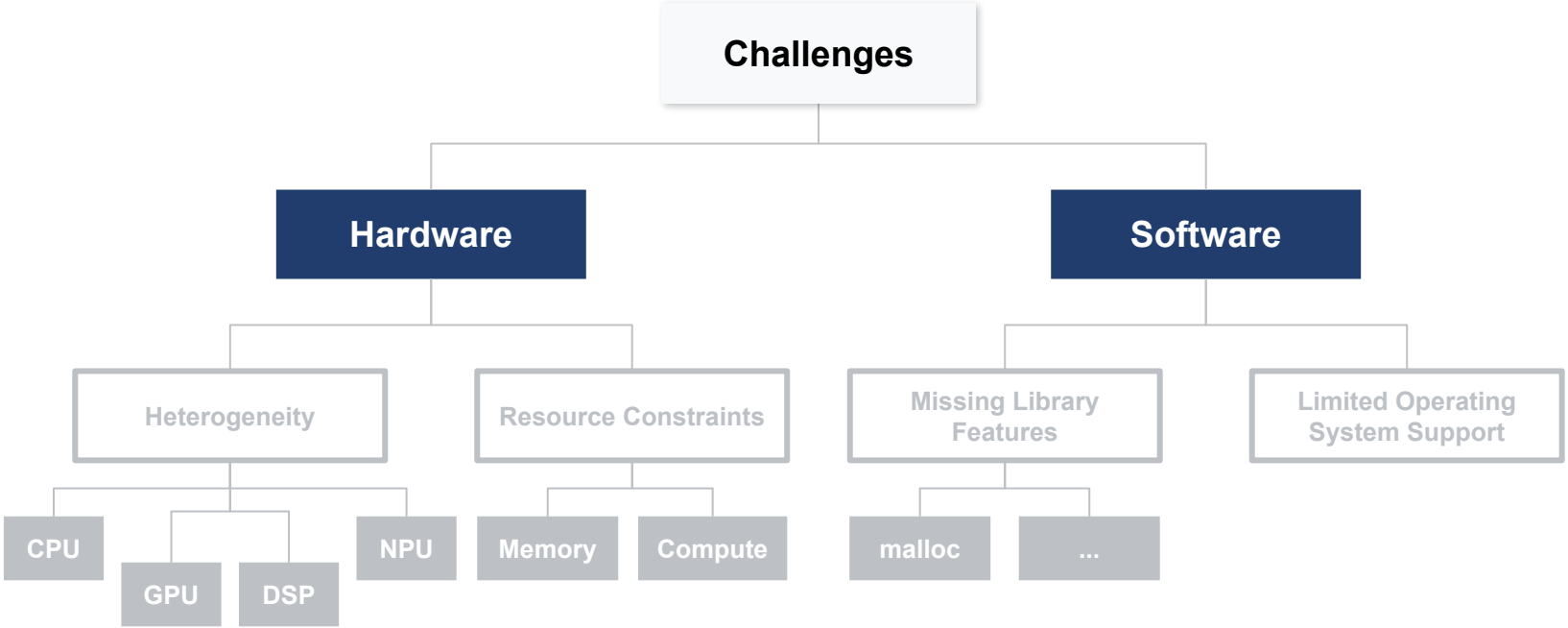
Sabrina M. Neuman[1], Brian Plancher[1], Bardienus P. Duisterhof[2], Srivatsan Krishnan[1], Colby Banbury[1], Mark Mazumder[1], Shvetank Prakash[1], Jason Jabbour[3], Aleksandra Faust[4], Guido C.H.E. de Croon[5], and Vijay Janapa Reddi[1]

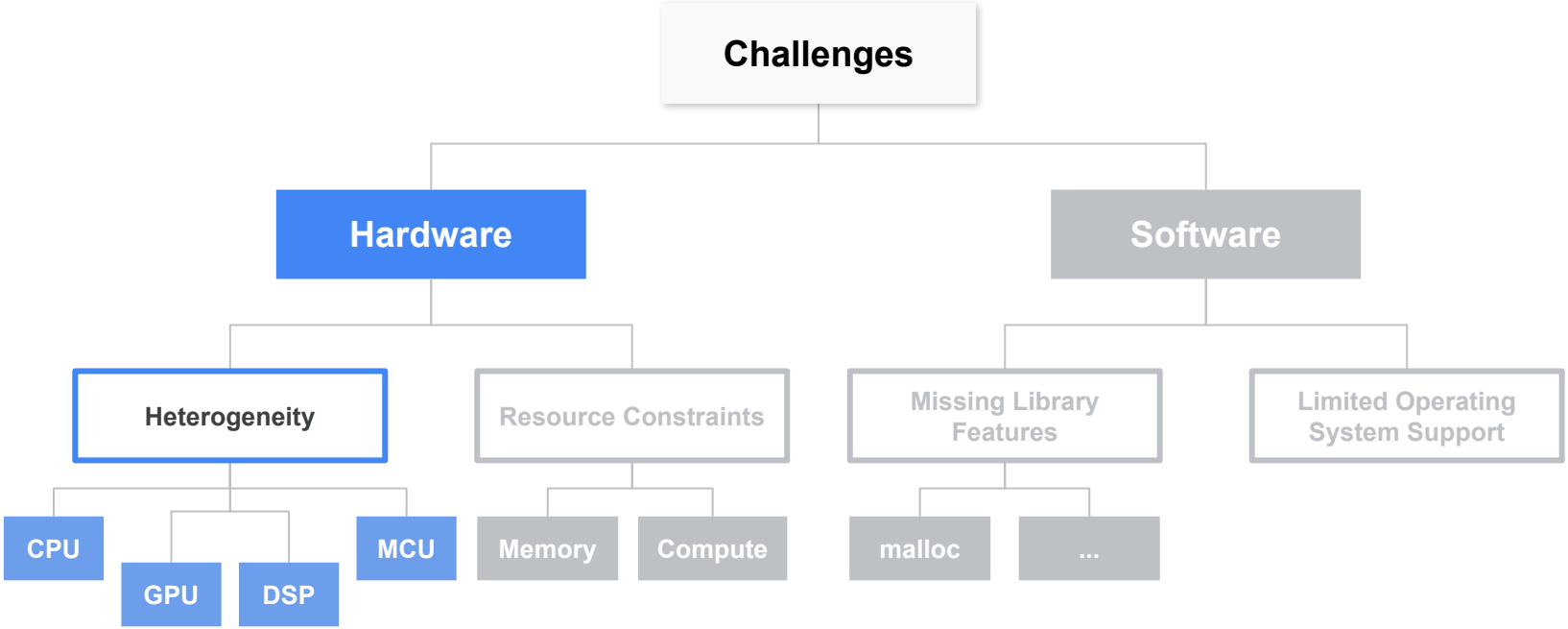
Harvard University[1], CMU[2], University of Virginia[3], Google Brain[4], Delft University of Technology[5]

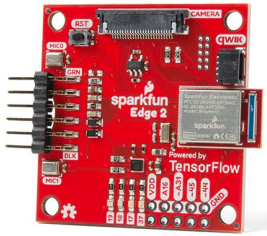
{sneuman@seas, brian_plancher@g, srivatsan@g, cbanbury@g, markmazumder@g, sprakash@g, vj@eecs}.harvard.edu, bduister@andrew.cmu.edu, jjj4se@virginia.edu, aleksandra.faust@gmail.com, g.c.h.e.decroon@tudelft.nl

https://arxiv.org/pdf/2205.05748.pdf

Challenges of Deploying ML at the Extreme Edge

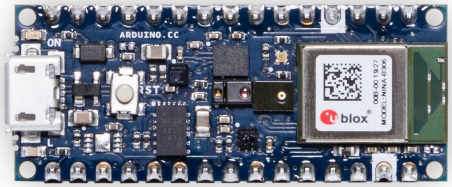


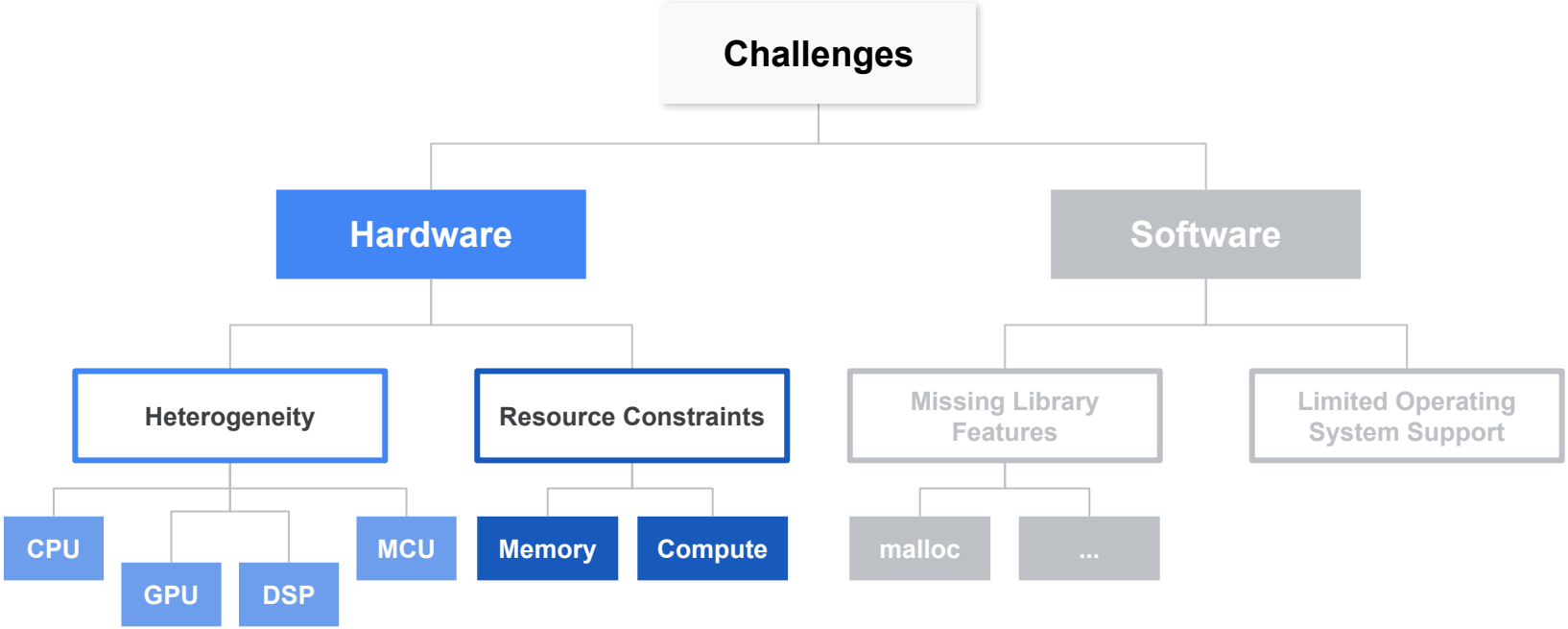




250 Billion

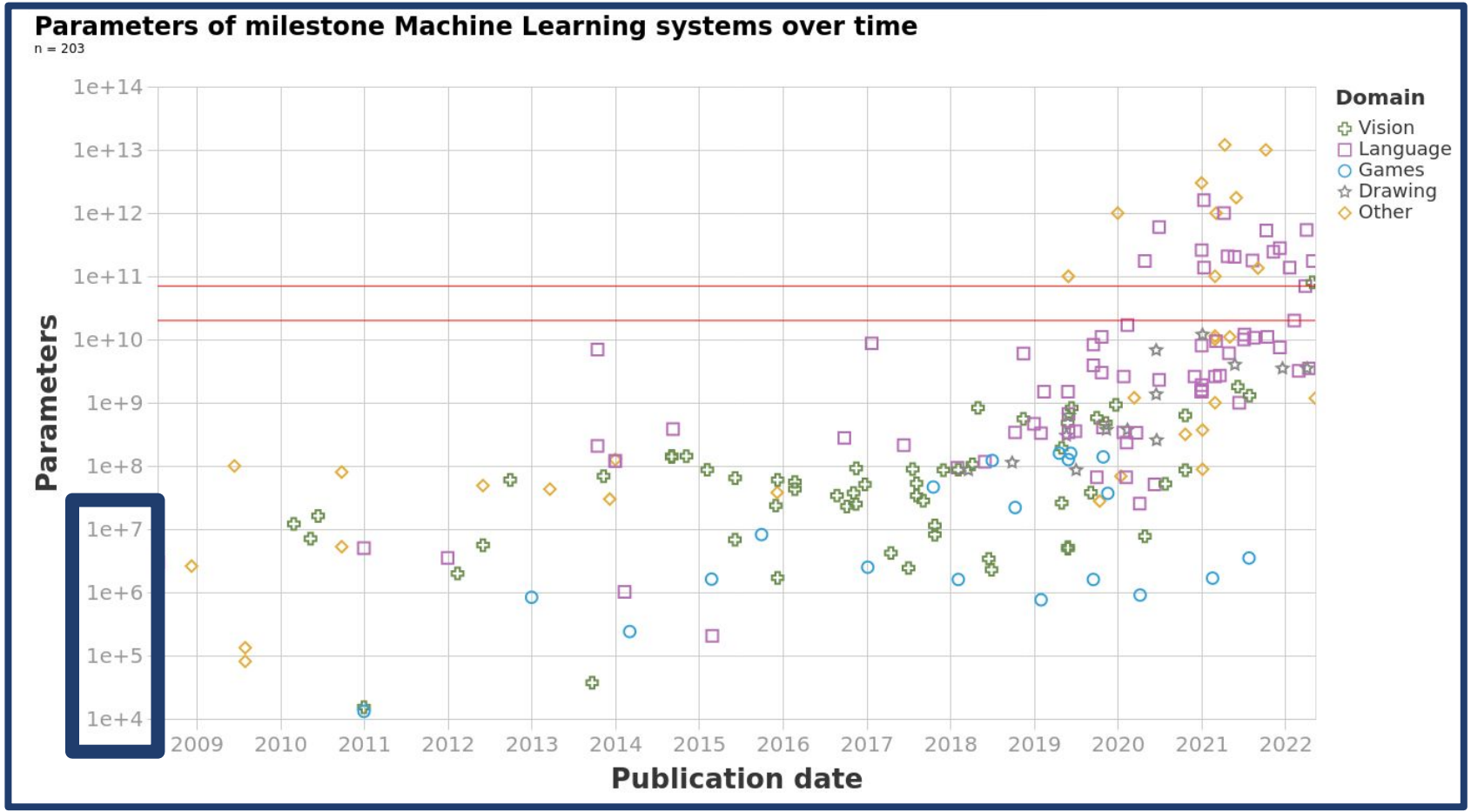
MCUs today

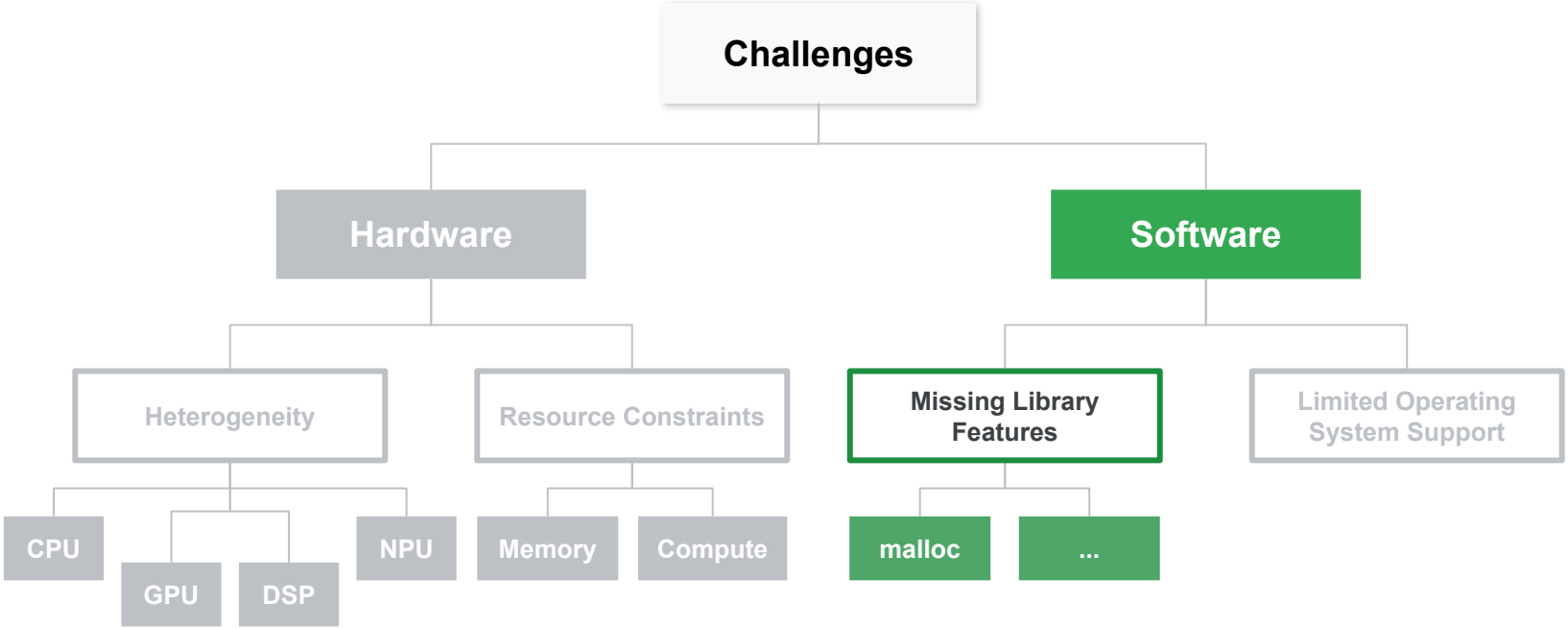


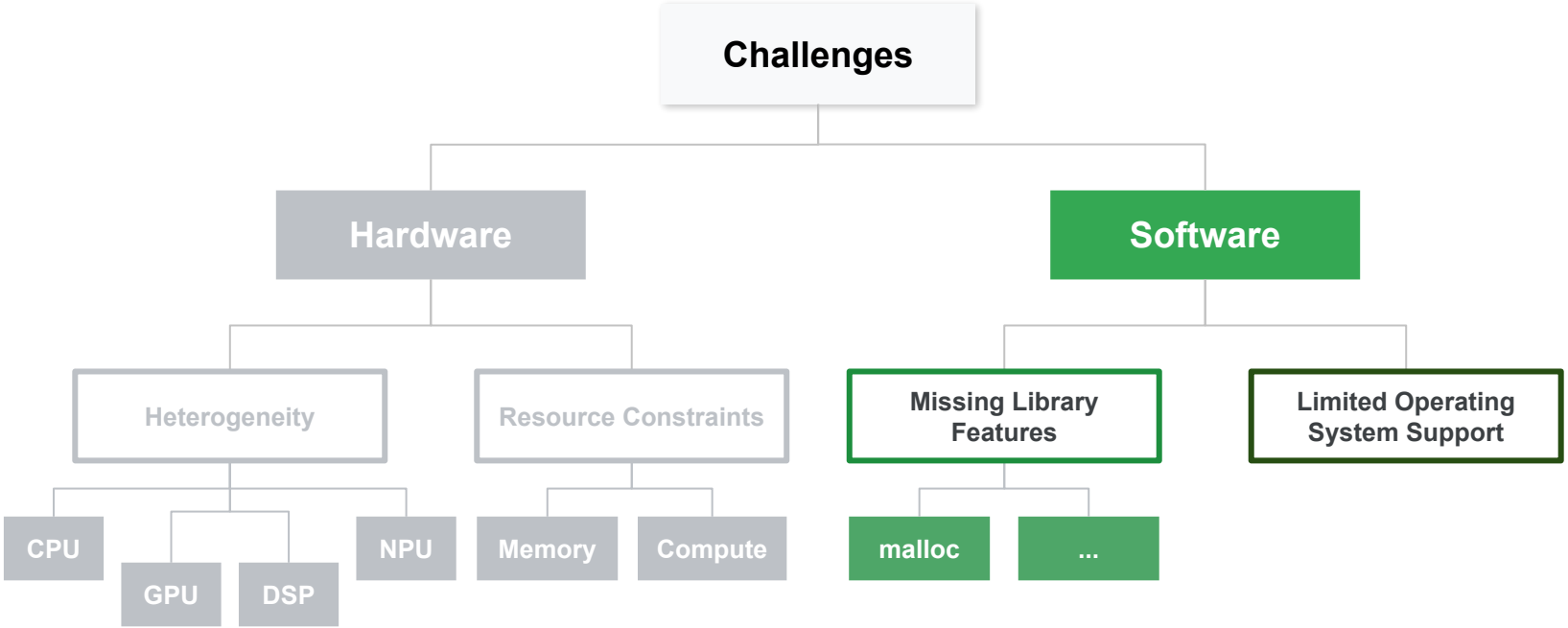


	Board	MCU / ASIC	Clock	Memory	Sensors	Radio
	Himax WE-I Plus EVB	HX6537-A 32-bit EM9D DSP	400 MHz	2MB flash 2MB RAM	Accelerometer, Mic, Camera	None
	Arduino Nano 33 BLE Sense	32-bit nRF52840	64 MHz	1MB flash 256kB RAM	Mic, IMU, Temp, Humidity, Gesture, Pressure, Proximity, Brightness, Color	BLE
	SparkFun Edge 2	32-bit ArtemisV1	48 MHz	1MB flash 384kB RAM	Accelerometer, Mic, Camera	BLE
	Espressif EYE	32-bit ESP32-D0WD	240 MHz	4MB flash 520kB RAM	Mic, Camera	WiFi, BLE

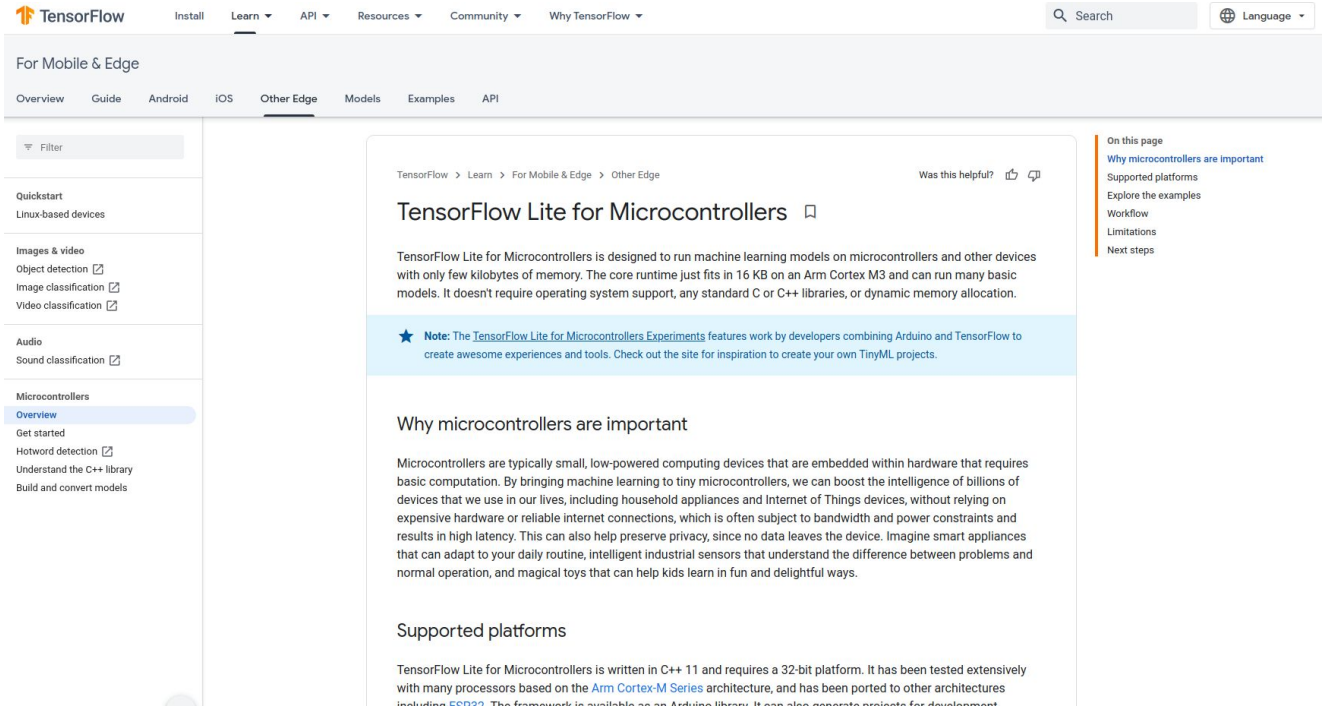
https://epochai.org/blog/machine-learning-model-sizes-and-the-parameter-gap







TensorFlow Lite Micro



EDGE IMPULSE

Create library

Turn your impulse into optimized source code that you can run on any device.

C++ library

Arduino library

Cube.MX CMSIS-PACK

WebAssembly

TensorRT library

Run your impulse directly

Run this impulse directly on your mobile phone or computer, no app required.

Computer

Mobile phone

Build firmware

Get a ready-to-go binary for your development board that includes your impulse.

ST IoT Discovery Kit

Arduino Nano 33 BLE Sense

Espressif ESP-EYE (ESP32)

Raspberry Pi RP2040

SiLabs Thunderboard Sense 2

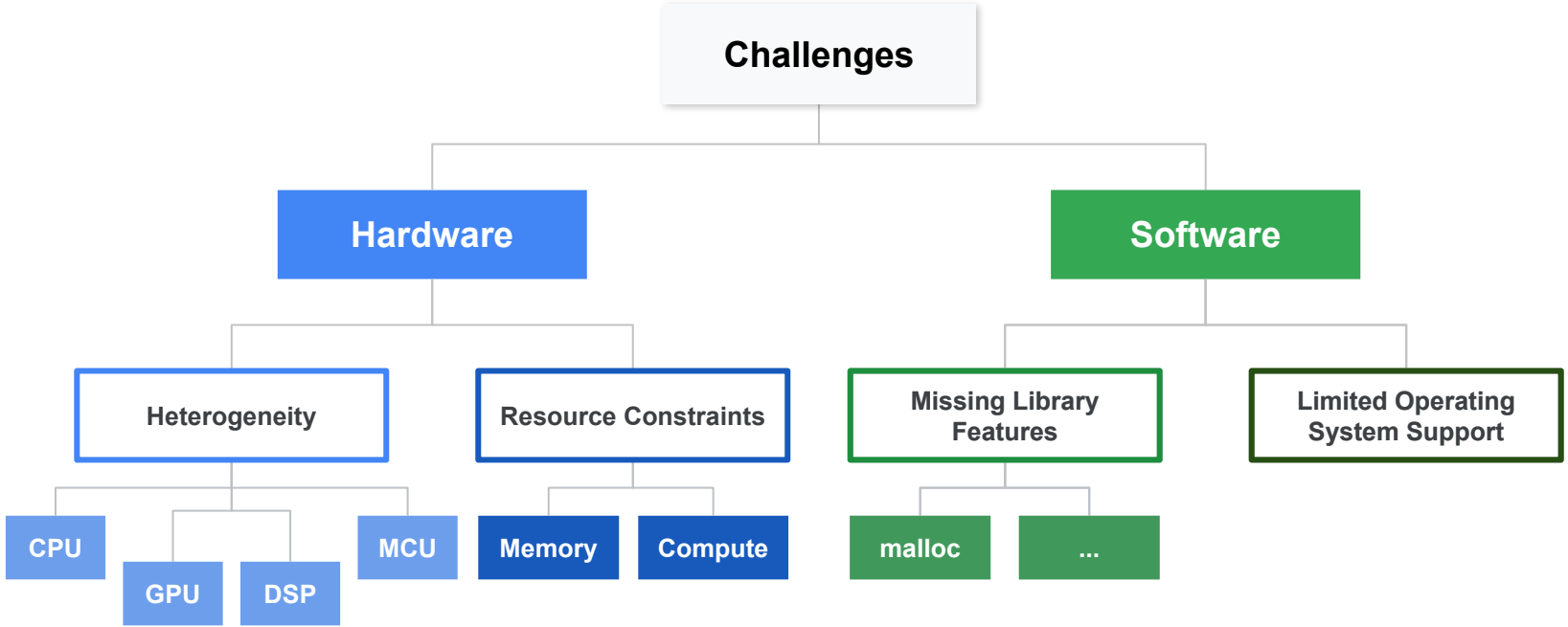
BETA

SiLabs xG24 Dev Kit

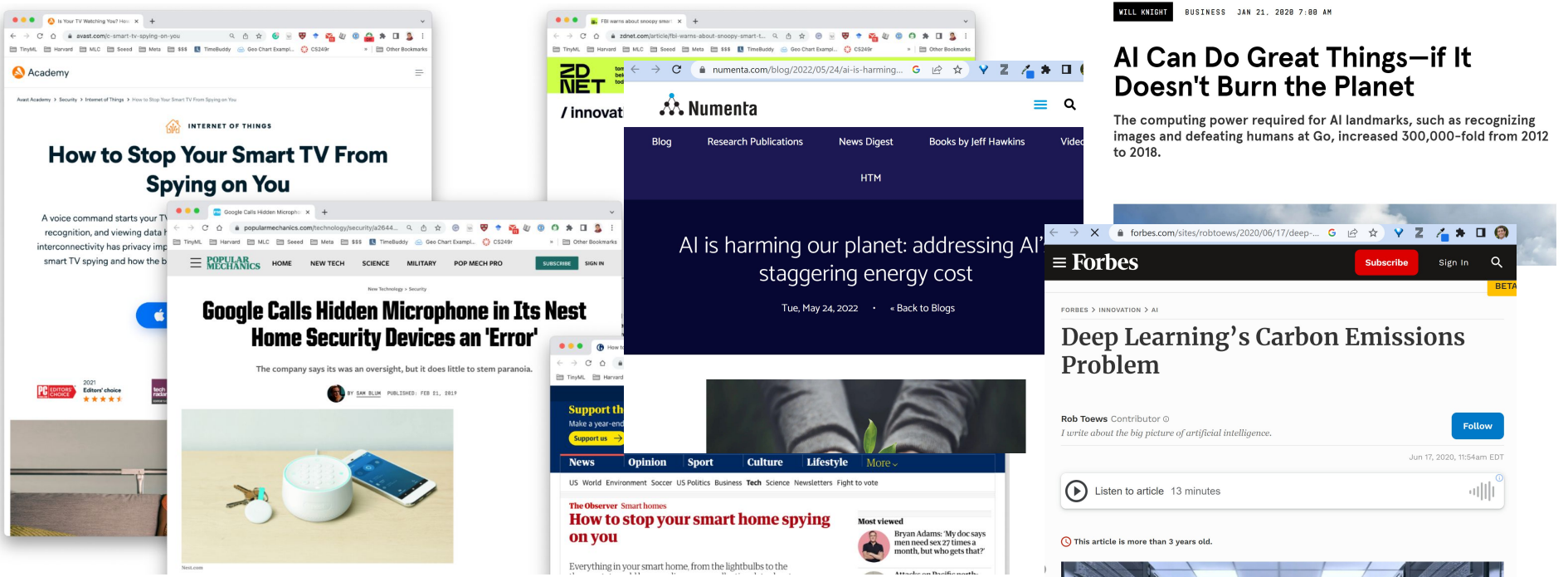
Himax WE-I Plus

Nordic nRF52840 DK + IKS02A1

Nordic nRF5340 DK + IKS02A1

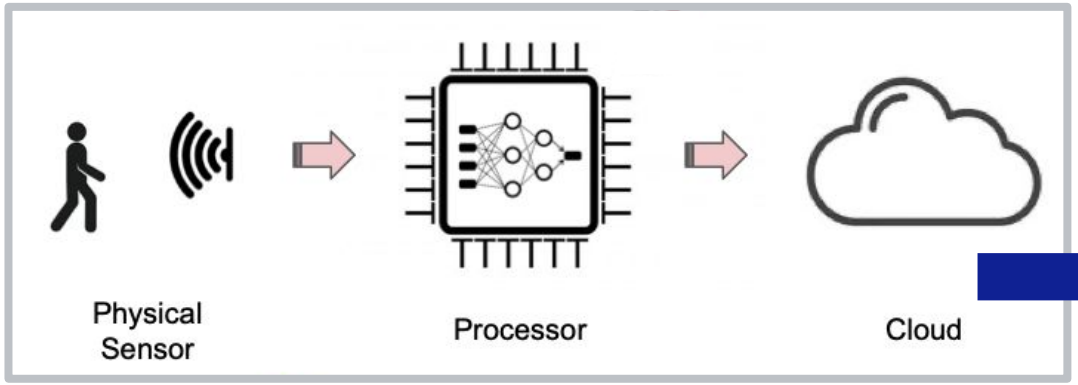


We need to adapt our approach to Edge ML to support a **Responsible AI** future!

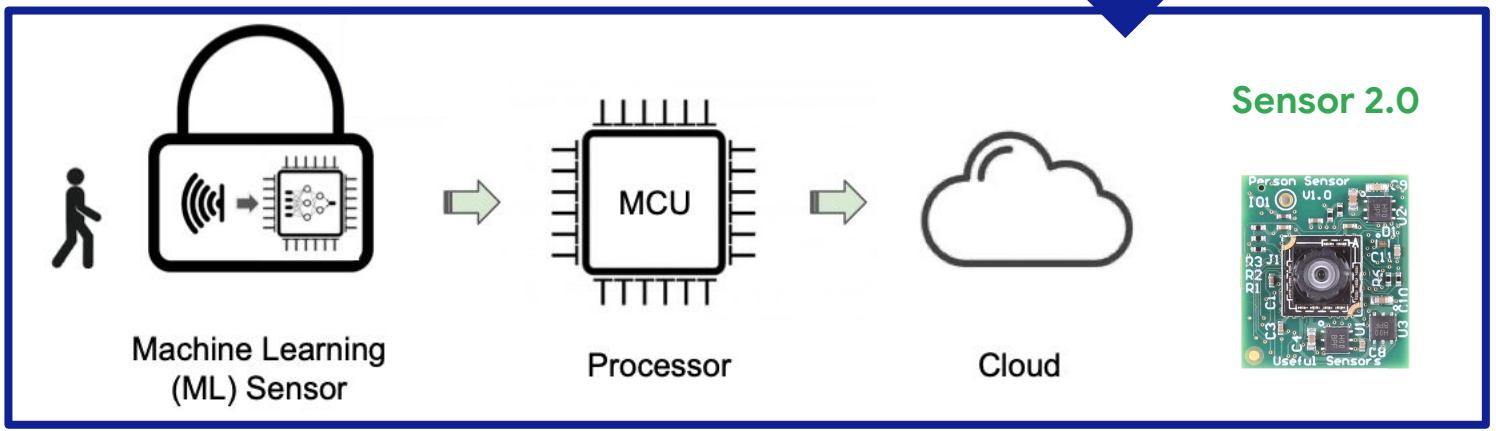


Sustainability & Privacy: MLSensors at Scale

What is a Machine Learning Sensor?



Privacy by Design



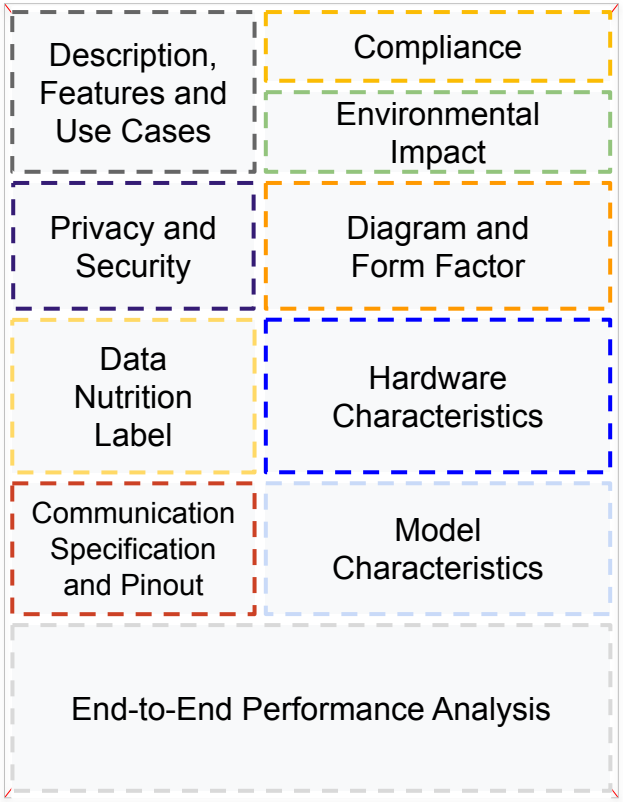
We suggest **transparency** as a core value to overcome these challenges.

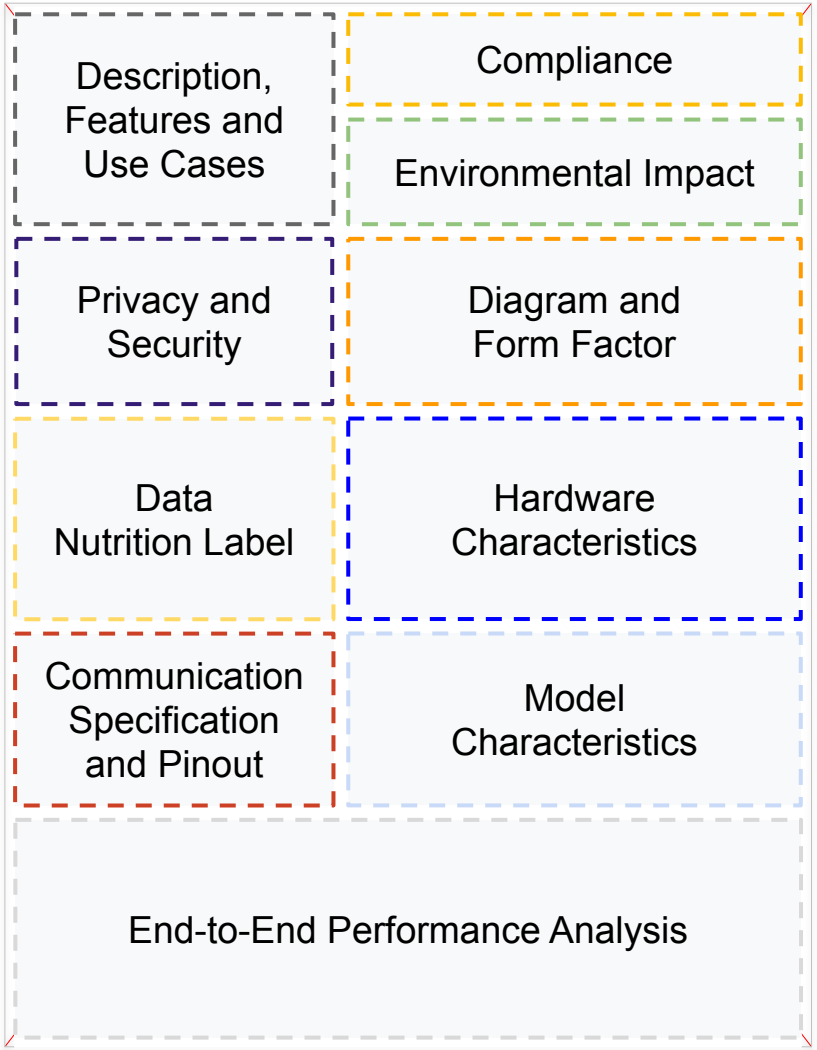
Datasheets for Machine Learning Sensors

Matthew Stewart[1*] Pete Warden[2,5] Yasmine Omri[1] Shvetank Prakash[1] Joao Santos[1]
Shawn Hymel[4] Benjamin Brown[1] Jim MacArthur[1] Nat Jeffries[5] Brian Plancher[3]
Vijay Janapa Reddi[1]

[1]Harvard University [2]Stanford University [3]Barnard College, Columbia University
[4]Edge Impulse [5]Useful Sensors

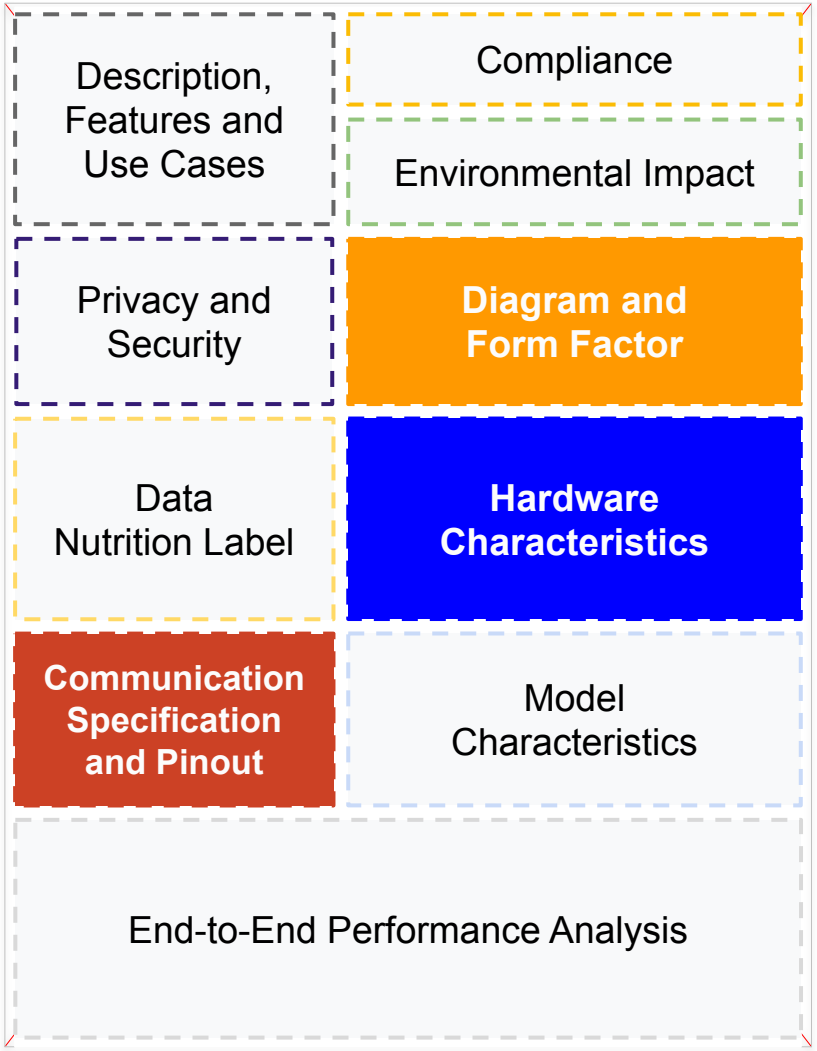
arxiv.org/abs/2306.08848





ML Sensor Datasheets Have 3 Goals:

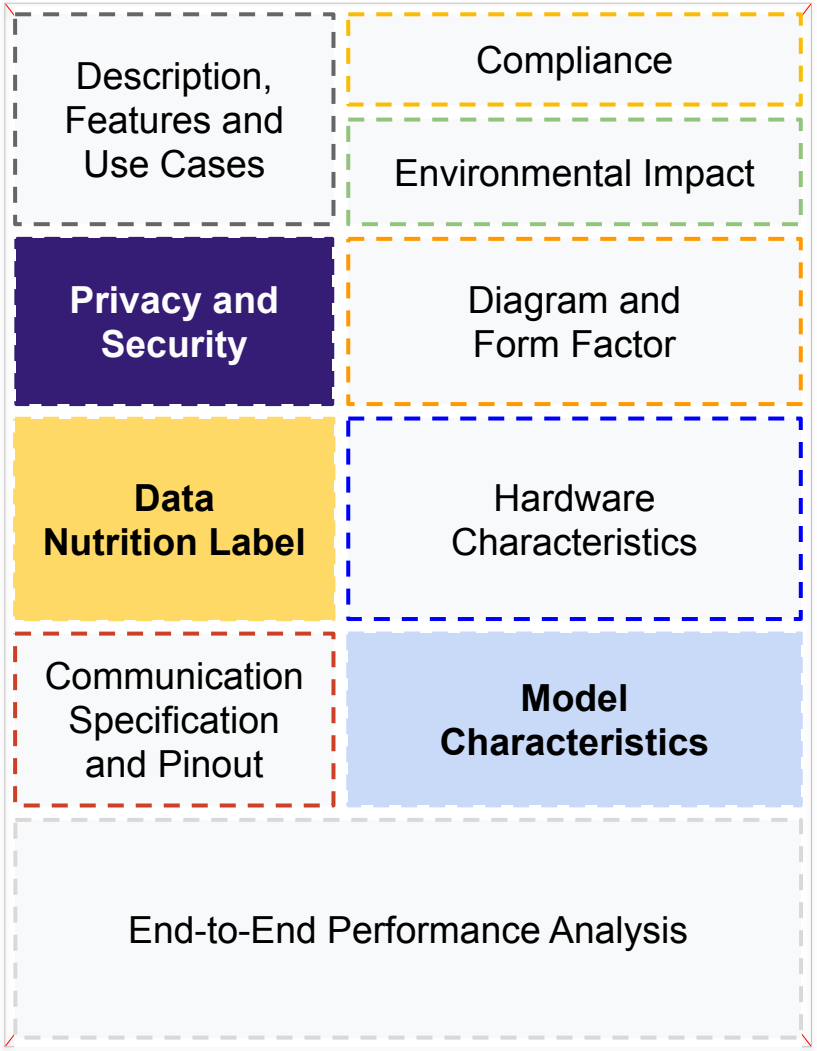
1. Raise the level of **abstraction**
2. Transparent at the **hardware, data, model, and end-to-end layers**
3. Support **Responsible** Use



ML Sensor Datasheets Have 3 Goals:

1. Raise the level of abstraction
2. Transparent at the **hardware,** data, model, and end-to-end layers
3. Support Responsible Use

Standard Sensor Datasheet

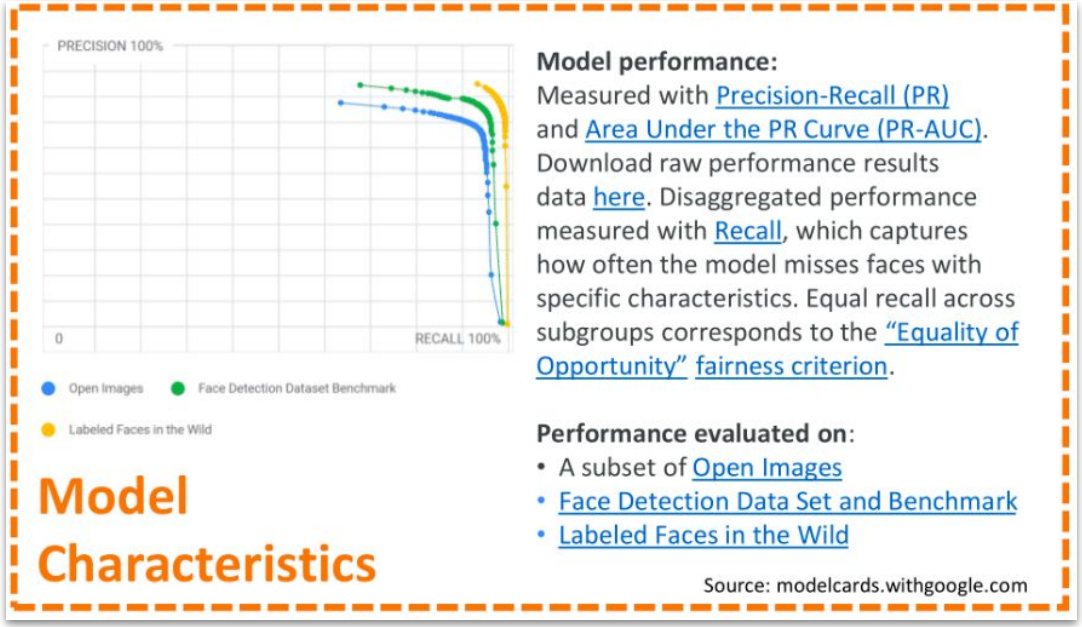


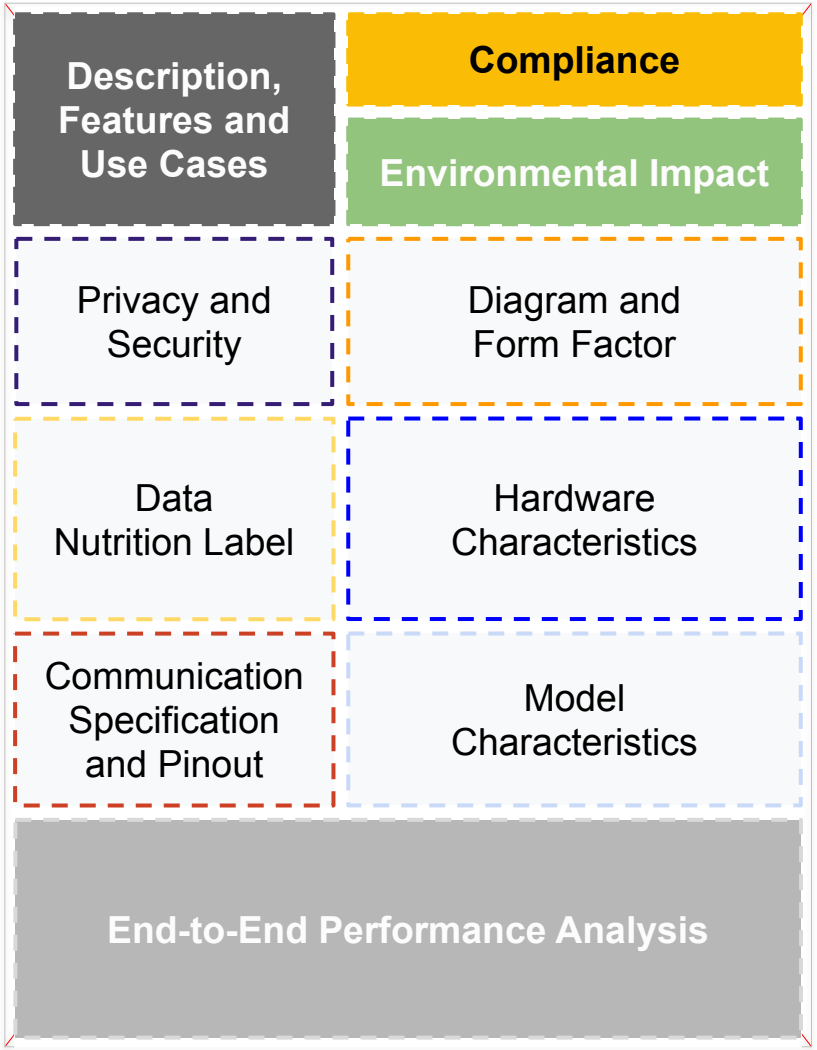
ML Sensor Datasheets Have 3 Goals:

1. Raise the level of abstraction
2. Transparent at the hardware, **data, model,** and end-to-end layers
3. Support **Responsible** Use

Responsible Machine Learning Analysis

Responsible Machine Learning Analysis

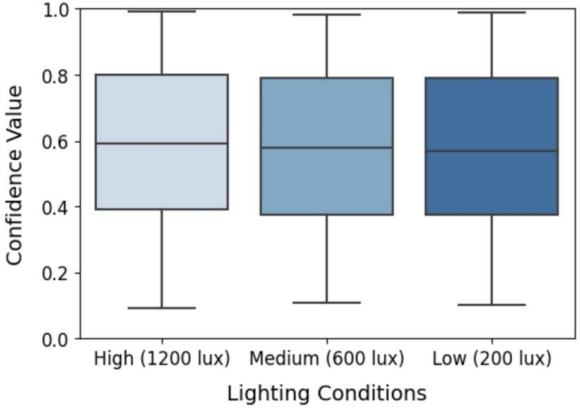


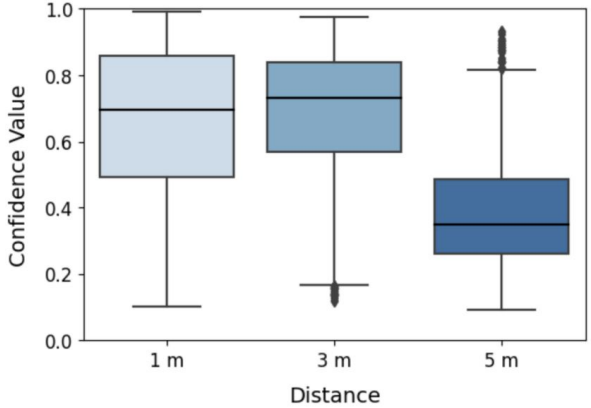


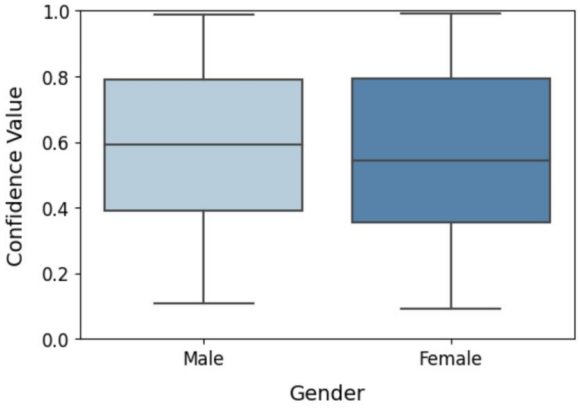
ML Sensor Datasheets Have 3 Goals:

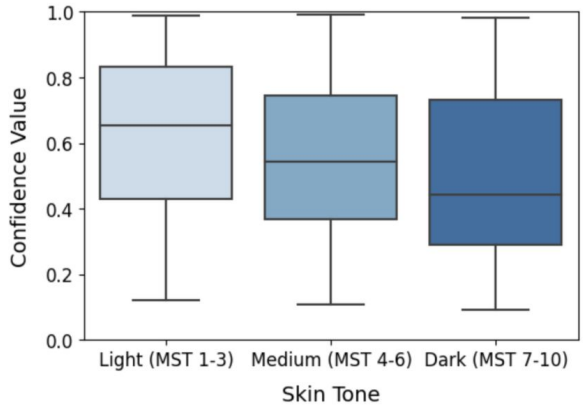
1. Raise the level of **abstraction**
2. Transparent at the hardware, data, model, and **end-to-end layers**
3. Support **Responsible** Use

Overall System Analysis



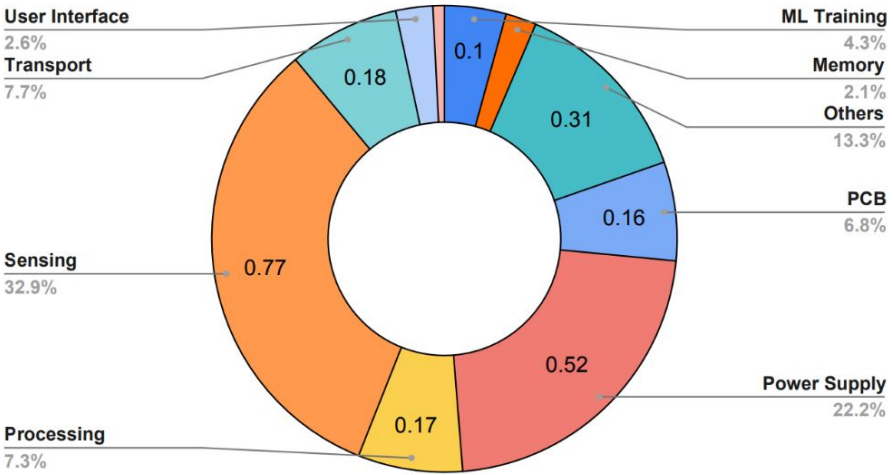


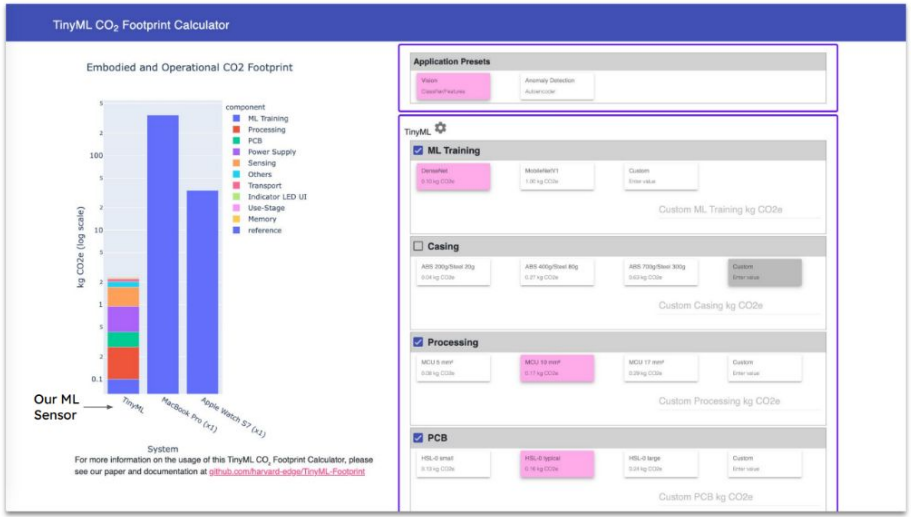




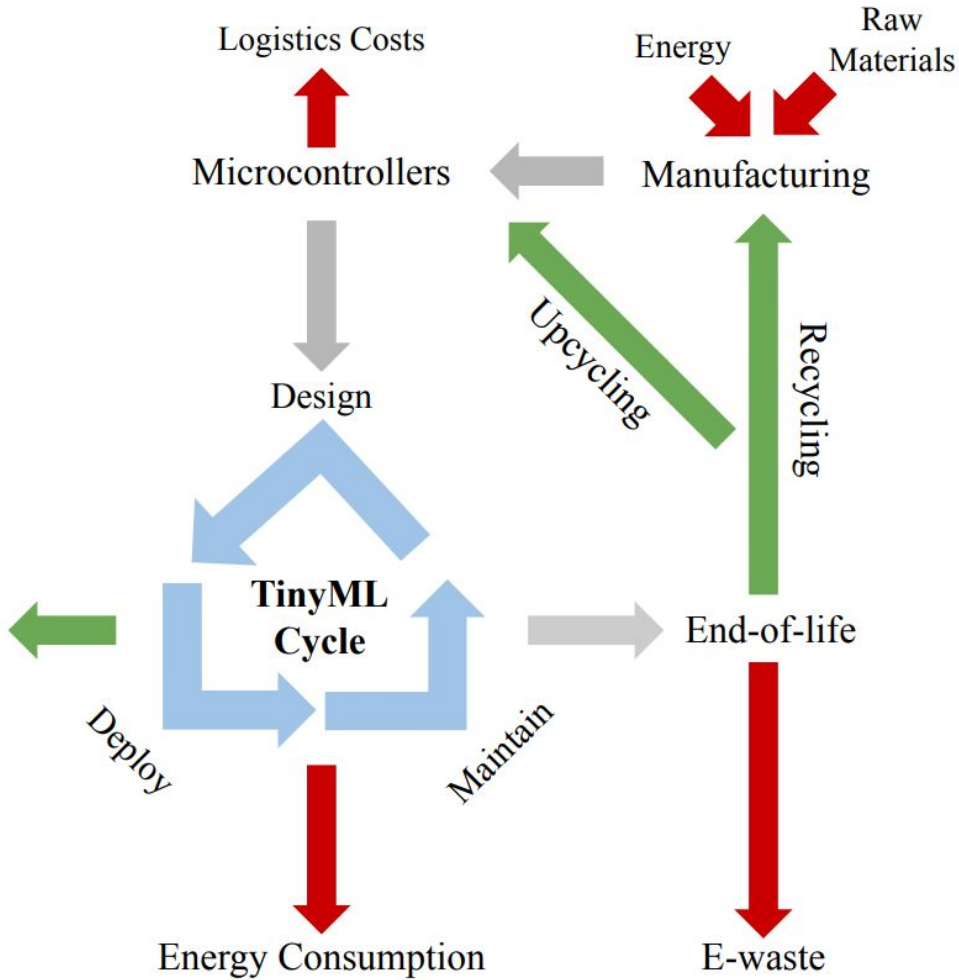
End-to-End
Responsible
Performance
Analysis

Environmental Impact





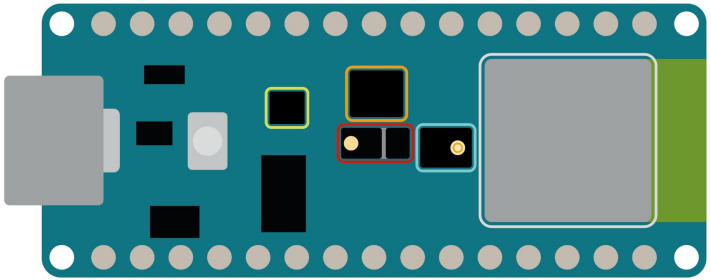
Lets Explore this Impact in More Detail



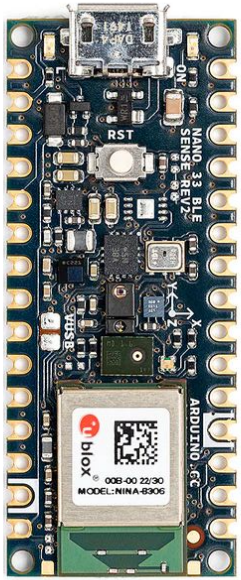
TinyML can support the SDGs but comes with costs. **What is the net impact?**

arxiv.org/abs/2301.11899

Real TinyML Systems are more than just an MCU!

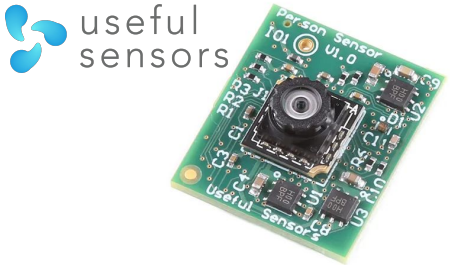


- Color, brightness, proximity and gesture sensor
- Digital microphone
- Motion, vibration and orientation sensor
- Temperature, humidity and pressure sensor
- Arm Cortex-M4 microcontroller and BLE module

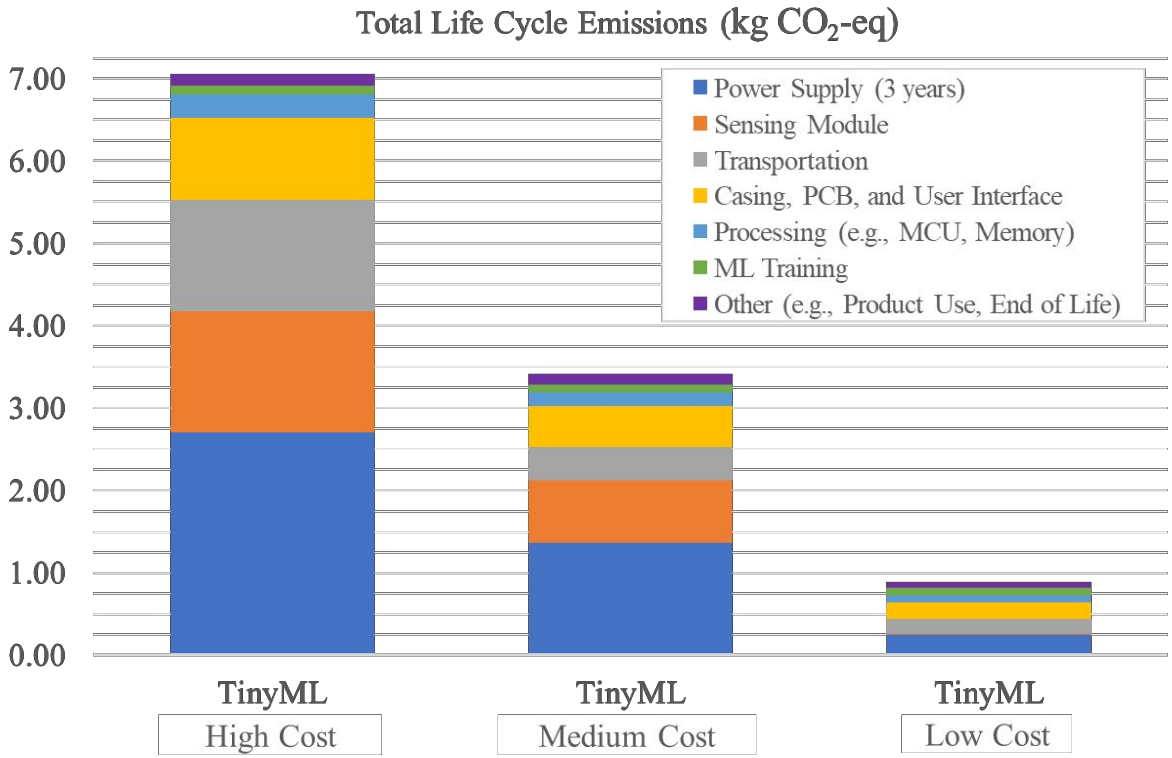


Building Representative Systems

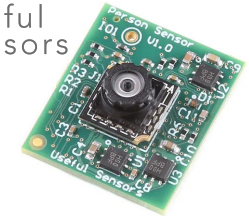
Cost Level	High Cost	Medium Cost	Low Cost
Application	Image Classification		Keyword Spotting
Size	Large	Compact	Compact



Building Representative Systems

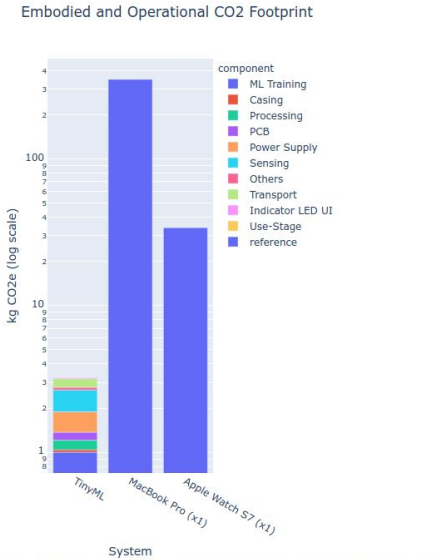


useful sensors

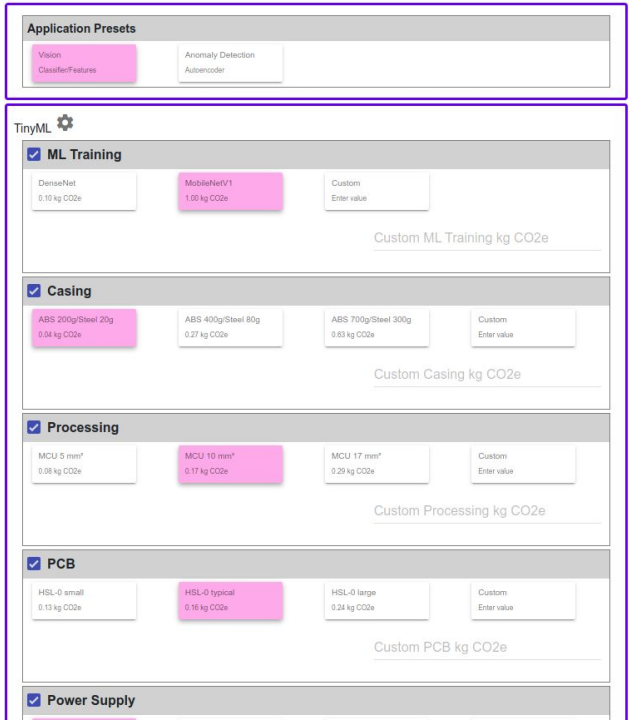


harvard-edge.github.io/TinyML-Footprint/

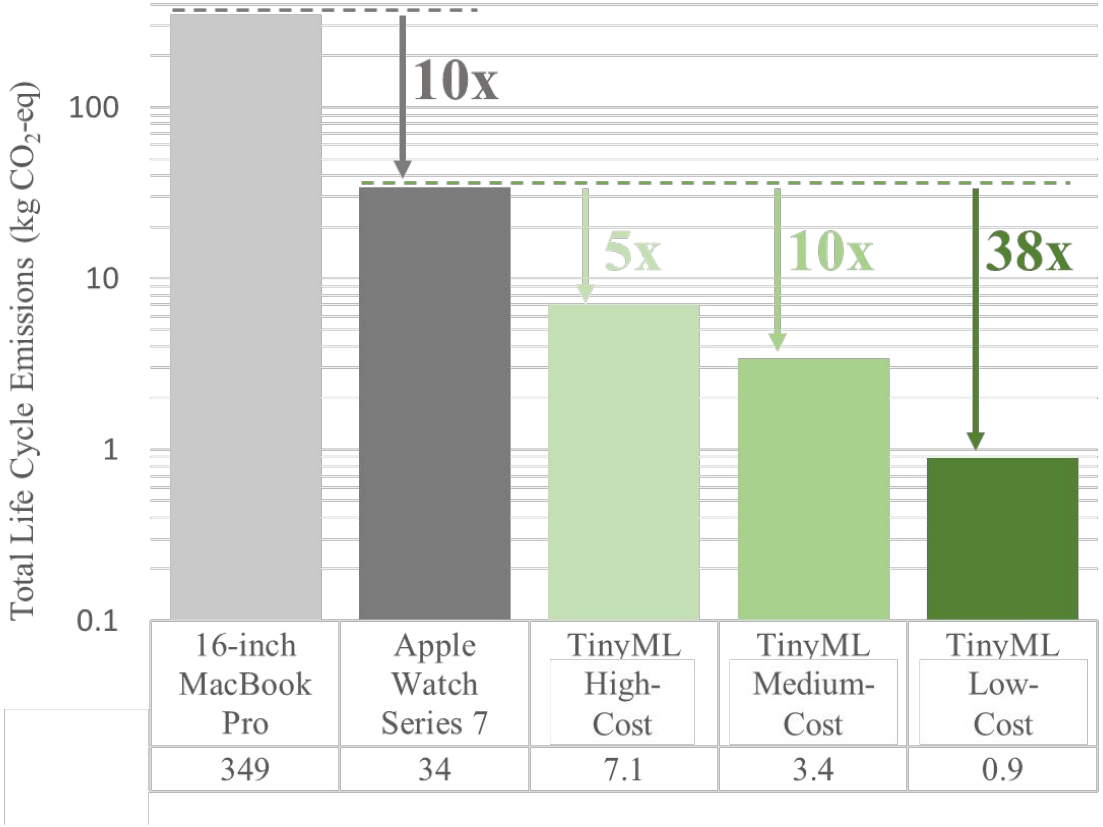
TinyML CO_2 Footprint Calculator



For more information on the usage of this TinyML CO_2 Footprint Calculator, please see our paper and documentation at github.com/harvard-edge/TinyML-Footprint

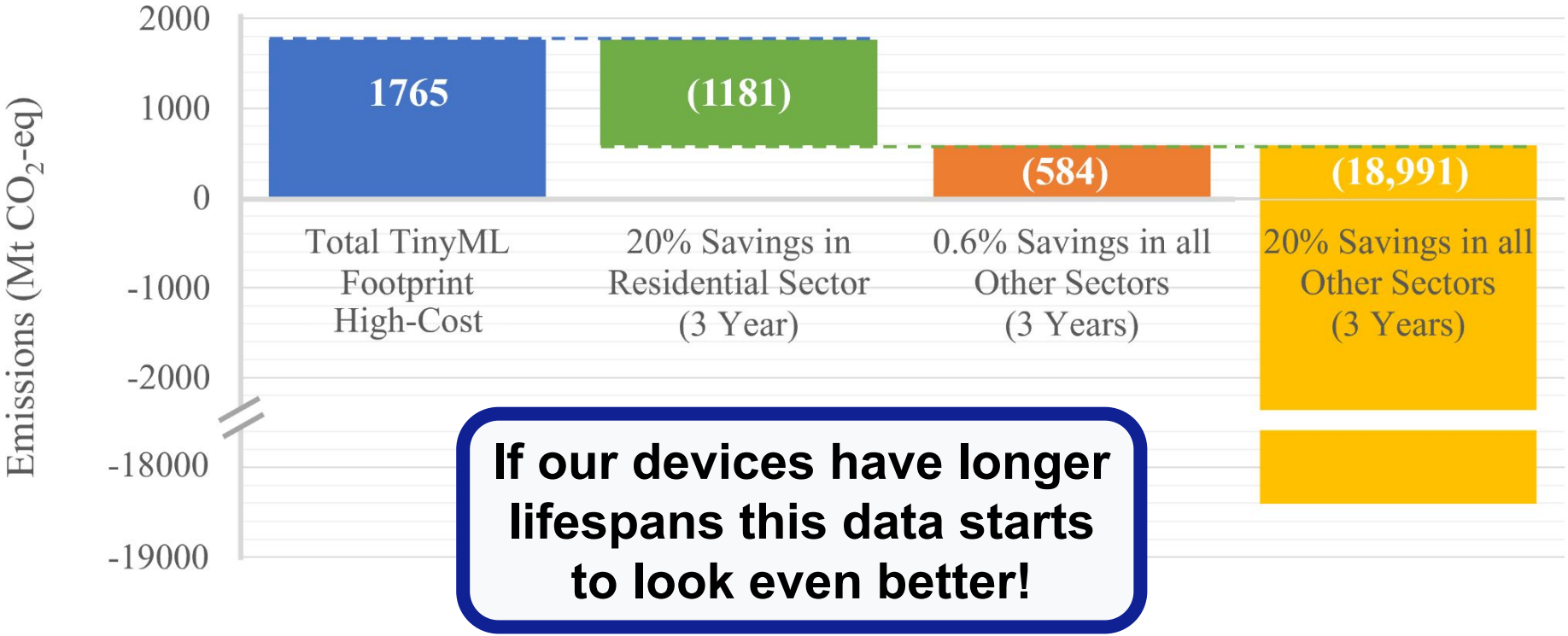


TinyML Systems in Context



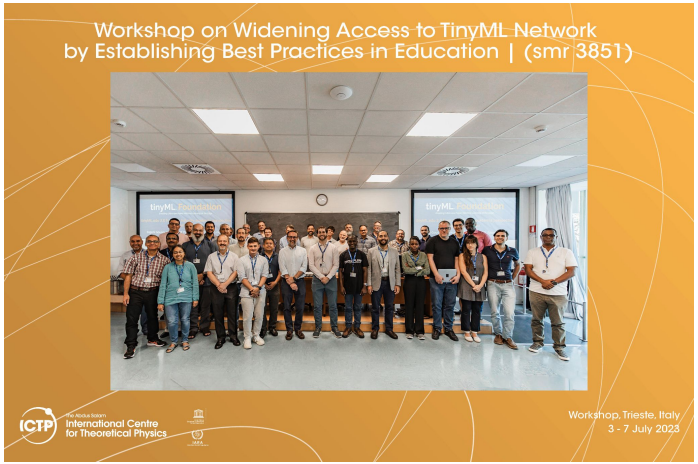
5x to 38x Savings over a 3-year lifespan!

What if we scale to 250bn devices?



If our devices have longer lifespans this data starts to look even better!

The Big Impact of TinyML:
Embedded Machine Learning at the Extreme Edge



tinyMLedu.org

Brian Plancher
Barnard College, Columbia University
brianplancher.com

MoroccoAI AI Summer School
جامعة الأخوين
AL AKHAWAYN UNIVERSITY