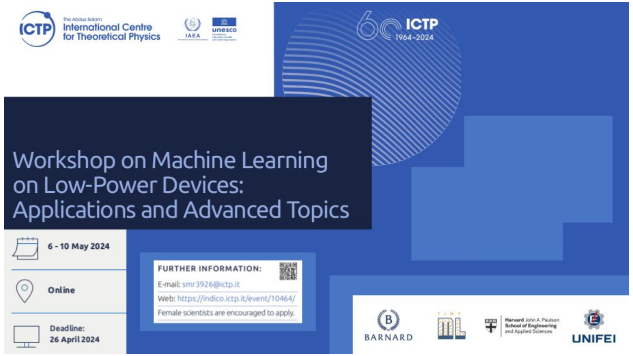
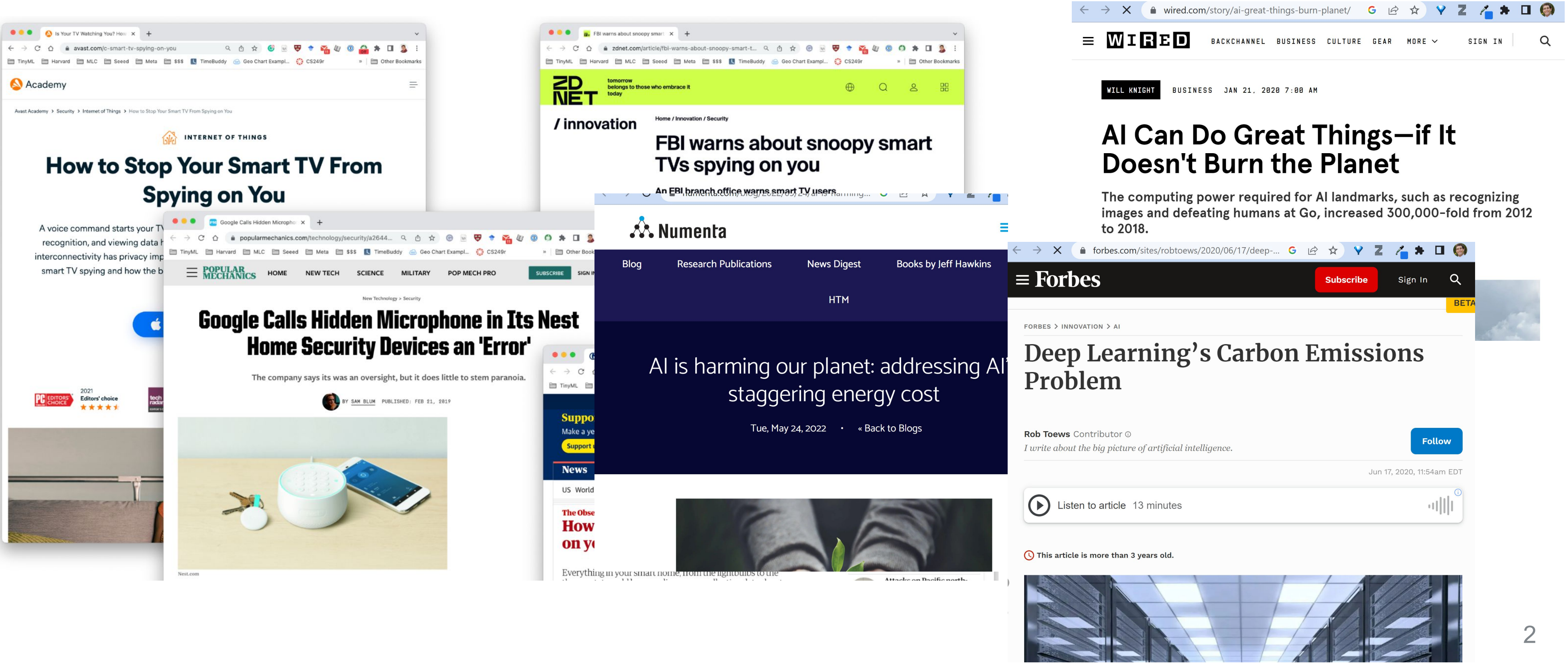


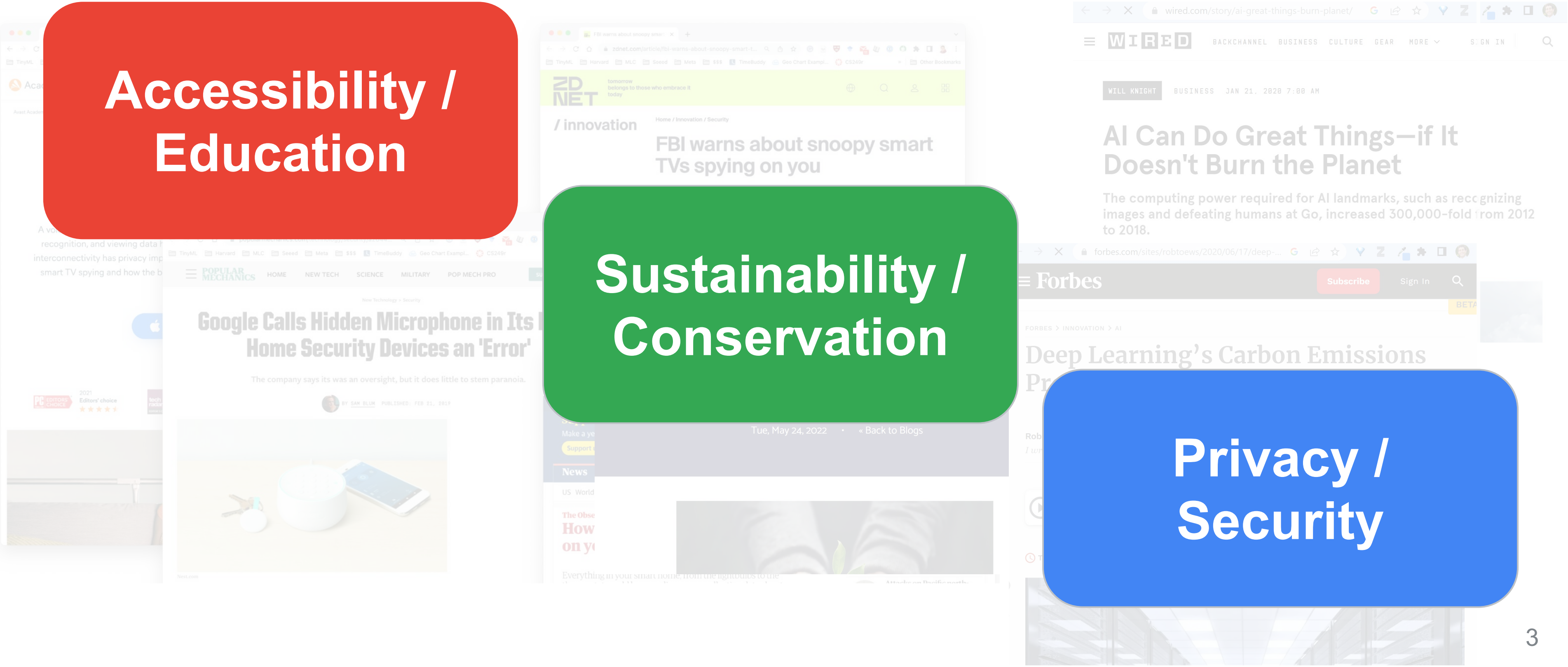
Responsible TinyML

Brian Plancher
Barnard College, Columbia University
brianplancher.com

How can TinyML support **Responsible AI**?



How can TinyML support **Responsible AI**?



Accessibility / Education

Sustainability / Conservation

Privacy / Security

Accessibility / Education

Promoting **Accessibility / Education**

Full Courses

Organization	Course Name	Date of Course	Target Audience	Language of Instruction	Language of Materials	Links
	edX tinyML Specialization by Harvard University	Launched 2020-2022	Everyone	English	English	Course 1-3 Website Course 4 Website All Materials All Colabs Arduino Library
	Embedded Machine Learning on Coursera by Edge Impulse	Launched 2021-2022	Everyone	English	English	Course 1 Course 2 All Materials
	ESE3600: Tiny Machine Learning by the University of Pennsylvania	Fall 2022	Undergraduate and Graduate Students	English	English	Website and Materials
	MIT 6.S965 TinyML and Efficient Deep Learning	Fall 2022	Graduate Students	English	English	Website Materials
	UNIFEI IESTI01 TinyML - Machine Learning for Embedding Devices	Jan 2021 - Present	Undergraduate Students	Portuguese	English	2022.1 Website and Materials 2021.2 Website and Materials 2021.1 Website and Materials
	Harvard CS249r Tiny Machine Learning	Sept 2020 - Present	Graduate Students	English	English	2022 Website and Assignments 2020 Website 2020 Assignments

Workshops

Lead Organizers	Workshop Name	Date of Workshop	Target Audience	Language of Instruction	Language of Materials	Links
	Morocco AI Summer School 2023	July 2023	Everyone	English	English	Website TinyML Part 1 TinyML Part 2
	EdgeMLUP 2023 Workshop on Widening Access to TinyML Network by Establishing Best Practices in Education	July 2023	Everyone	English	English	Website and Materials
	SciTinyML 2023 Scientific Use of Machine Learning on Low-Power Devices	April 2023	Everyone	English	English	Website and Materials
	TinyML at AAU A Workshop at Addis Ababa University	March 2023	Everyone	English	English	Materials
	Artificial Intelligence and its Integration with Everyday Life An Introduction to TinyML by Edwin Marte at	November 2022	Everyone	Spanish	Spanish	Materials

TinyMLedu.org

Machine Learning Systems

FRONT MATTER
Preface
Dedication
Acknowledgements
Contributors
Copyright
About the Book

MAIN
1 Introduction
2 Embedded Systems
3 Deep Learning Primer
4 Embedded AI
5 AI Workflow
6 Data Engineering
7 AI Frameworks
8 AI Training
9 Efficient AI
10 Model Optimizations

FRONT MATTER > Preface

Machine Learning Systems

with TinyML

ABSTRACT
Machine Learning Systems with TinyML offers readers an entry point to understand comprehensive machine learning systems by grounding concepts in accessible TinyML applications. As resource-constrained edge computing sees rapid expansion, the ability to construct efficient ML pipelines grows crucial. This book aims to demystify the process of developing complete ML systems suitable for deployment - spanning key phases like data collection, model design, optimization, acceleration, security hardening, and integration. The text touches on the full breadth of concepts relevant to general ML engineering across industries and applications through the lens of TinyML. Readers will learn basic principles around designing ML model architectures, hardware-aware training strategies, performant inference optimization, benchmarking methodologies and more. Additionally, crucial systems considerations in areas like reliability, privacy, responsible AI, and solution validation are also explored in depth. In summary, the book strives to equip newcomers and professionals alike with integrated knowledge covering full stack ML system development, using easily accessible TinyML applications as the vehicle to impart universal concepts required to unlock production ML.

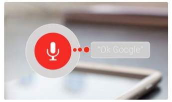
Table of contents
Preface
Why We Wrote This Book
What You'll Need to Know
Book Conventions
Want to Help Out?
Get in Touch
Contributors

Edit this page
Report an issue
View source

Foundations of TinyML

Focusing on the basics of machine learning and embedded systems, such as smartphones, this course will introduce you to the "language" of TinyML.

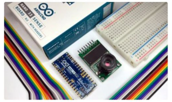
Take the Course on edX



Applications Of TinyML

Get the opportunity to see TinyML in practice. You will see examples of TinyML applications, and learn first-hand how to train these models for Tiny applications such as keyword spotting, visual wake words, and gesture recognition.

Take the Course on edX



Deploying TinyML

Learn to program in TensorFlow Lite for microcontrollers so that you can write the code, and deploy your model to your very own Tiny microcontroller. Before you know it, you'll be implementing an entire TinyML application.

Take the Course on edX

MLOps for Scaling TinyML

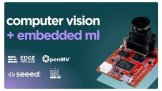
This course introduces learners to Machine Learning Operations (MLOps) through the lens of TinyML (Tiny Machine Learning). Learners explore best practices to deploy, monitor, and maintain (tiny) Machine Learning models in production at scale.

Take the Course on edX

Introduction to Embedded Machine Learning

This course will give you a broad overview of how machine learning works, how to train neural networks, and how to deploy those networks to microcontrollers using the Edge Impulse Platform.

Take the Course on Coursera



Computer Vision with Embedded Machine Learning

This course, offered by a partnership among Edge Impulse, OpenMV, Seeed Studio, and the TinyML Foundation, will give you an understanding of how deep learning with neural networks can be used to classify images and detect objects in images and videos.

Take the Course on Coursera

Promoting **Accessibility / Education**

Full Courses

TinyMLedu.org

Widening Access to Applied Machine Learning with TinyML	Vijay Janapa Reddi, Brian Plancher, Susan Kennedy, Laurence Moroney, Pete Warden, Anant Agarwal, Colby Banbury, Massimo Banzi, Matthew Bennett, Benjamin Brown, Sharad Chitlangia, Radhika Ghosal, Sarah Grafman, Rupert Jaeger, Srivatsan …	Harvard Data Science Review
TinyML in Africa: Opportunities and Challenges		
Bridging the Digital Div… the Promising Impact o… TinyML for Developing Countries		
TinyML: Applied AI for Development	… Reddi	… Development Goals
TinyMLedu: The Tiny Machine Learning Open Education Initiative	Brian Plancher, Vijay Janapa Reddi	ACM Technical Symposium on Computer Science Education (SIGCSE)

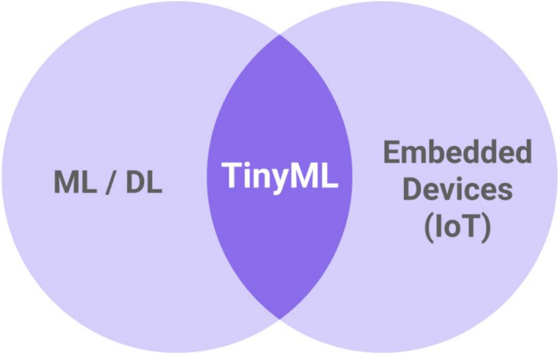
TinyML4D: Scaling Embedded Machine Learning Education in the Developing World	Brian Plancher, Sebastian Buttrich, Jeremey Ellis, Neena Goveas, Laila Kazimierski, Jesus Lopez Sotelo, Milan Lukic, Diego Mendez, Rosdiadee Nordin, Andres Oliva Trevisan, Massimo Pavan, Manuel Roveri, Marcus Rüb, Jackline Tum, Marian Verhelst, Salah Abdeljabar, Segun Adebayo, Thomas Amberg, Halleluyah Aworinde, José Bagur, Gregg Barrett, Nabil Benamar, Bharat Chaudhari, Ronald Criollo, David Cuartielles, Jose Alberto Ferreira Filho, Solomon Gizaw, Evgeni Gousev, Alessandro Grande, Shawn Hymel, Peter Ing, Prashant Manandhar, Pietro Manzoni, Boris Murmann, Eric Pan, Rytis Paskauskas, Ermanno Pietrosemoli, Tales Pimenta, Marcelo Rovai, Marco Zennaro, Vijay Janapa Reddi	AAAI-24 Sprign Symposium on Increasing Diversity in AI Education and Research

Foundations of TinyML
Applications Of TinyML
Deploying TinyML
MLOps for Scaling TinyML
Introduction to Embedded Machine Learning
Computer Vision with Embedded Machine Learning

Global Embedded ML Education Opportunities:

1 **Low Resource** Requirements

2 **Interdisciplinary** Focus



Low Power

Low Cost

Low Connectivity

Global Embedded ML Education Opportunities:

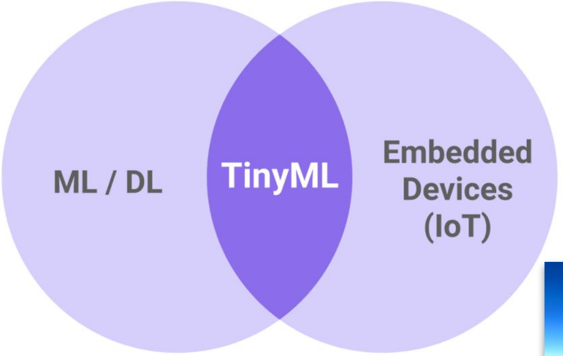
1 **Low Resource** Requirements

2 **Interdisciplinary** Focus and **Applied** Learning

Low Power

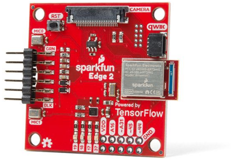
Low Cost

Low Connectivity



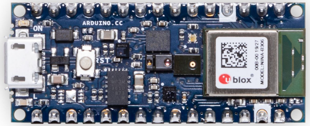
Global Embedded ML Education ~~Opportunities:~~ **Challenges**

1 Software and Hardware Fragmentation



250 Billion

MCUs today



Global Embedded ML Education ~~Opportunities:~~ Challenges

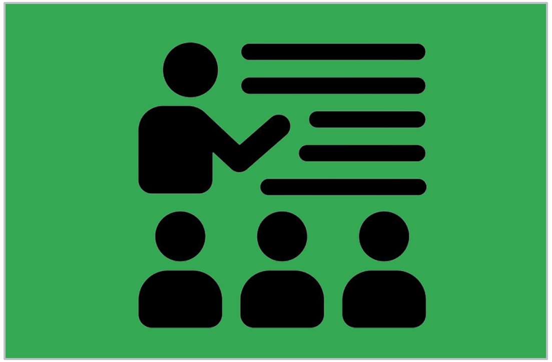
1 Software and Hardware Fragmentation

2 Affordability Barriers and Localization Roadblocks

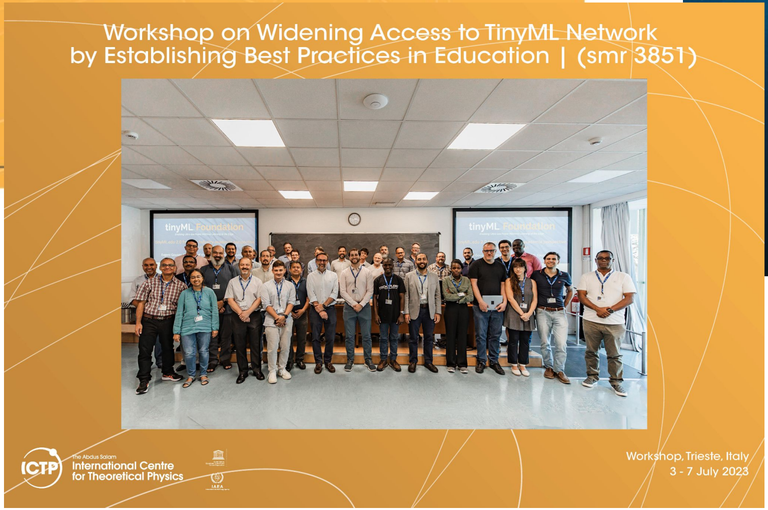
Language and Local Relevance

Global Embedded ML Education ~~Opportunities:~~ Challenges

1. Software and Hardware Fragmentation
2. Affordability Barriers and Localization Roadblocks
3. Educator Readiness and Research Incentives




Workshop on Widening Access to TinyML Network by Establishing Best Practices in Education
seeed
ARDUINO
ICTP
arm
3 - 7 July 2023
An ICTP Meeting
Trieste, Italy


Workshop on Widening Access to TinyML Network
by Establishing Best Practices in Education | (smr 3851)
tinyML Foundation
tinyML Foundation
ICTP
The Abdus Salam
International Centre
for Theoretical Physics
Workshop, Trieste, Italy
3 - 7 July 2023

Towards a Modular Curriculum

Optional Modules

Software Focus
- Embedded Software Engineering Deep Dive
- ML Compilers and Optimizers
- Neural Network Architecture and Design

Hardware Focus
- Electronics and IoT Deep Dive
- Sensor Paradigms and Design
- Device Design and Deployment

Domain-Specific Focus
- Conservation
- Predictive Maintenance
- Smart Cities
- The Future of Work
- Climate Change
- Healthcare

Core Modules

Machine Learning and Deep Learning Fundamentals
(E.g., Models, Training, Overfitting, Regression vs. Classification, Neural Networks)

Data Centric AI
(E.g., Data Collection, Pre- and Post-Processing)

Embedded Systems
(E.g, Microcontrollers, Embedded Programming, Basic Electronics & IoT)

Responsible AI
(E.g., Bias, Privacy, Security)

Canonical Hands-On Examples

From data collection to model training to deployed inference

- Audio Keyword Spotting
- Image Classifi-cation
- IMU Anomaly Detection

Course Lengths

- **3-5 day short courses** include one hands-on example from each core module and a high level overview of the theory. Time permitting they include optional modules.
- **Micro-credential courses** dive deeper into the theory of each core module and select optional modules.
- **Full semester-long courses** explore a full track of optional modules and the core modules in detail with multiple hands-on examples and theoretical derivations and explorations.

Towards a Modular Curriculum

[TinyMLedu.org](https://tinymledu.org)

Please feel free to **remix** our materials and please consider **sharing back** your materials for the community!

Optional Modules

Software Focus
- Embedded Software Engineering Deep Dive
- ML Compilers and Optimizers
- Neural

Hardware Focus
- Electronics and IoT Deep Dive
- Sens Paradi and De

Domain-Specific Focus
- Conservation
- Predic Mainte
- The Future of Work
- Clim Char

Core Modules

Machine Learning and Deep Learning ...entals (...rfitting, Regression vs. ...etworks)

...edded Systems (... Microcontrollers, ...dded Programming, ... Electronics & IoT)

...AI (...curity)

Canonical Hands-On Examples

From data collection to model training to deployed inference

- Audio Keyword Spotting
- Image Classifi-cation
- IMU Anomaly Detection

Course Lengths

- **3-5 day short courses** include one hands-on example from each core module and a high level overview of the theory. Time permitting they include optional modules.
- **Micro-credential courses** dive deeper into the theory of each core module and select optional modules.
- **Full semester-long courses** explore a full track of optional modules and the core modules in detail with multiple hands-on examples and theoretical derivations and explorations.

Calls to Action

1. Assessing Our Educational Programs
2. Maintaining Open-Source Software and Courseware
3. Embedded ML Model and Data Zoo
4. Improving Accessibility of Hardware
5. Growing a Research Community
6. Increased Outreach and Diversity Efforts

Calls to Action

1. Assessing Our Educational Programs
2. Maintaining Open-Source Software and Courseware
3. Embedded ML Model and Data Zoo
4. Improving Accessibility of Hardware
5. Growing a Research Community
6. Increased Outreach and Diversity Efforts

Underrepresentation of Women in Robotics Research

Calls to Action

ieeexplore.ieee.org/document/10474552

TABLE 1. FAR for CS and engineering subfields based on prior work and including our result for robotics [1], [3], [4] (data from 2017 to 2023).

FIELD	FAR (%)
CS education	42
Human–computer interaction	26
CS overall average	**16–26**
Knowledge systems	19
Software engineering and languages	14
Artificial intelligence	12
Robotics	**11–12 (our analysis)**
Computer systems	10
Theory and algorithms	8

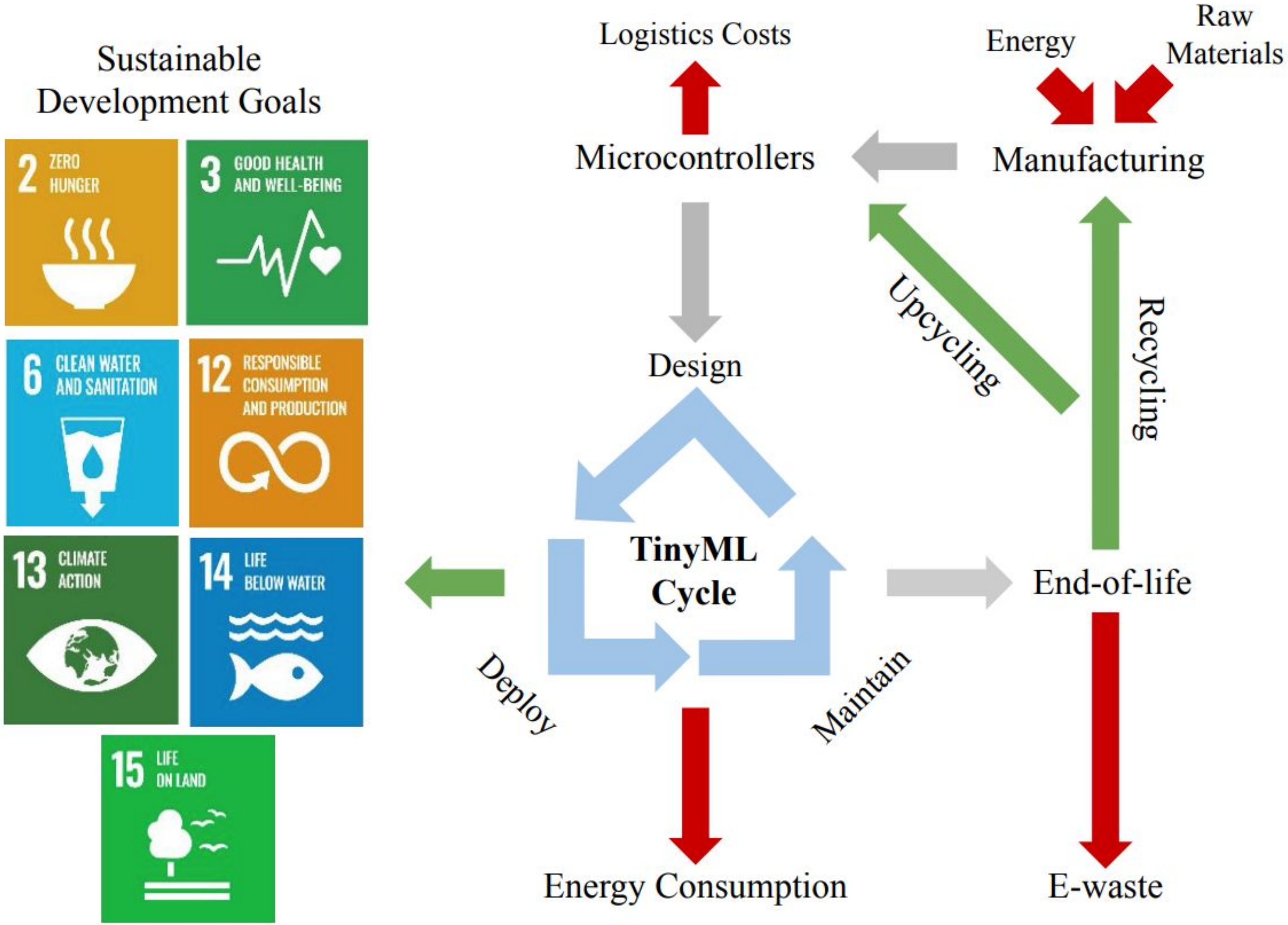
As has been noted in related works, this kind of methodology has many flaws and does not take into account much of the nuance in gender, including issues of bias, misperception, and nonbinary identities [7], [8]. However, we hope that this initial study will help add to the robotics community's understanding of the current state of gender diversity and, at a minimum, provide directionally correct data to help with future diversity, equity, and inclusion efforts.

Sustainability / Conservation

Promoting **Sustainability / Conservation**

[TinyMLedu.org](https://tinymledu.org)

	How TinyML Can be Leveraged to Solve Environmental Problems: A Survey	Hatim Bamoumen, Anas Temouden, Nabil Benamar, Yousra Chtouki	Innovation and Intelligence for Informatics, Computing, and Technologies
	Design and Development of a		
	Is TinyML Sustainable? Assessing the Environmental Impacts of Machine Learning on Microcontrollers	Shvetank Prakash, Matthew Stewart, Colby Banbury, Mark Mazumder, Pete Warden, Brian Plancher, Vijay Janapa Reddi	Communications of the ACM (CACM)
	Smart Buildings: Water Leakage Detection Using TinyML	Othmane Atanane, Asmaa Mourhir, Nabil Benamar, Marco Zennaro	Sensors
	Classifying Mosquito Wingbeat Sound Using TinyML	Moez Altayeb, Marco Zennaro, Marcelo Rovai	ACM Conference on Information Technology for Social Good

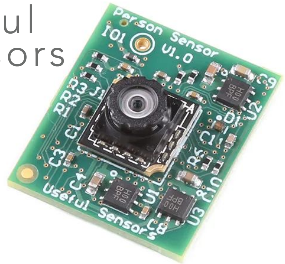


TinyML can support the SDGs but comes with costs. **What is the net impact?**

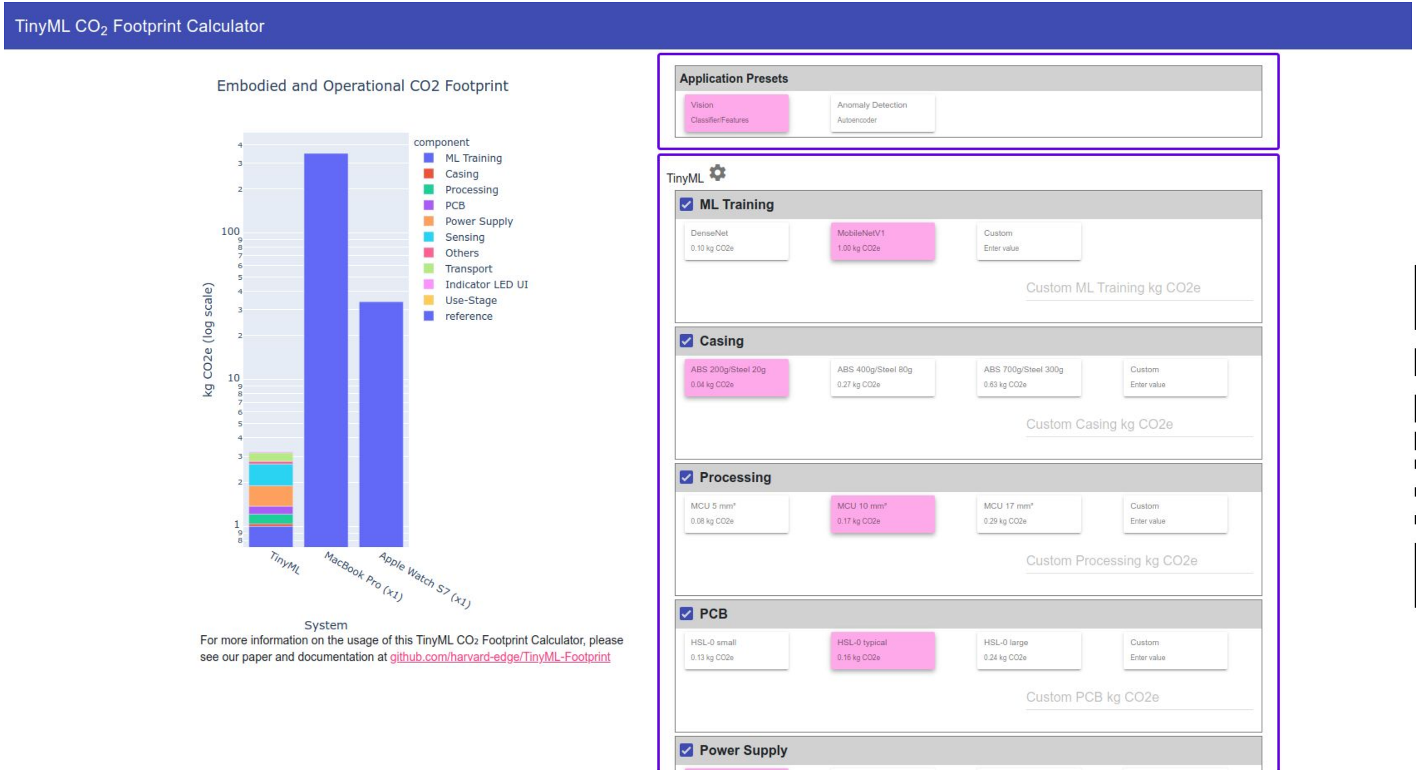
Building Representative Systems

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

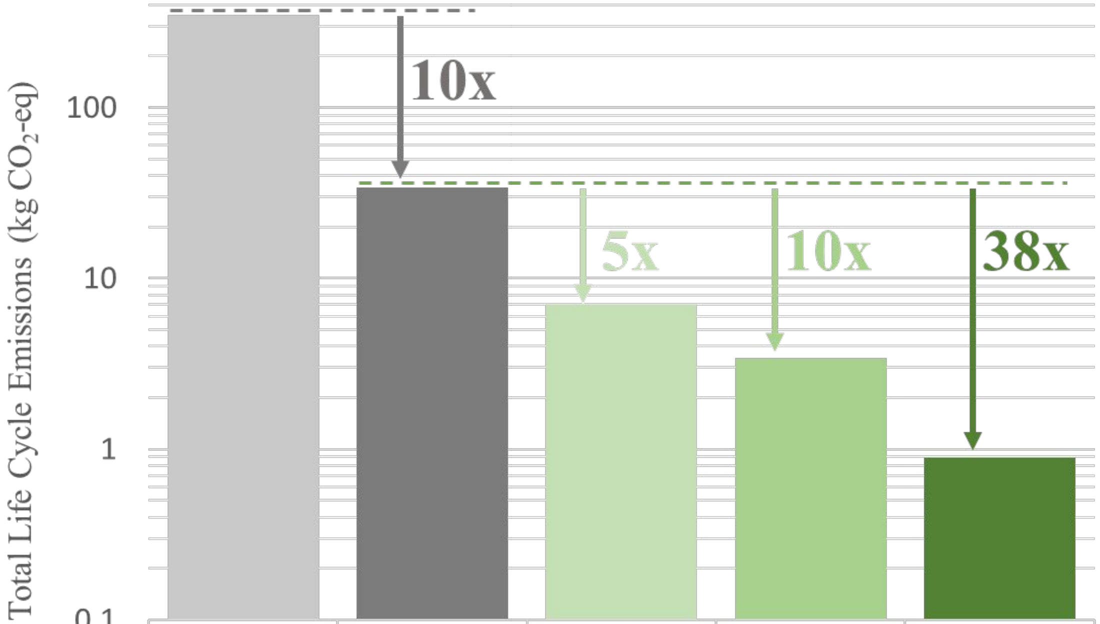
useful
sensors



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



TinyML Systems in Context



16-inch MacBook Pro	Apple Watch Series 7	TinyML High-Cost	TinyML Medium-Cost	TinyML Low-Cost
349	34	7.1	3.4	0.9

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

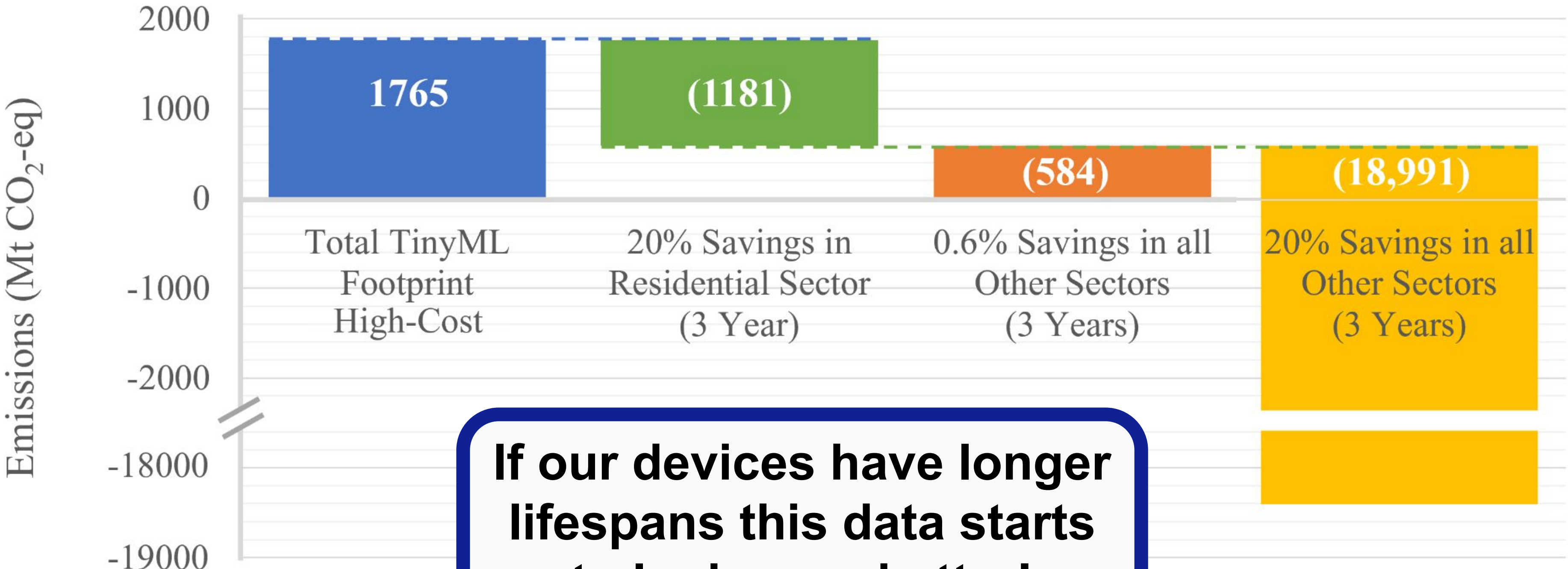
What if we scale to 250bn devices?

There are around **250bn MCUs** deployed today and around **15bn IoT** devices

IoT Device Growth					
	~15 Billion	>50 Billion	>100 billion	>250 Billion	>1 Trillion
Linear	2023	2041	2067	2144	2531
Exponential	2023	2032	2036	2043	2053

https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

What if we scale to 250bn devices?



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

Privacy / Security

TinyML will soon be everywhere!

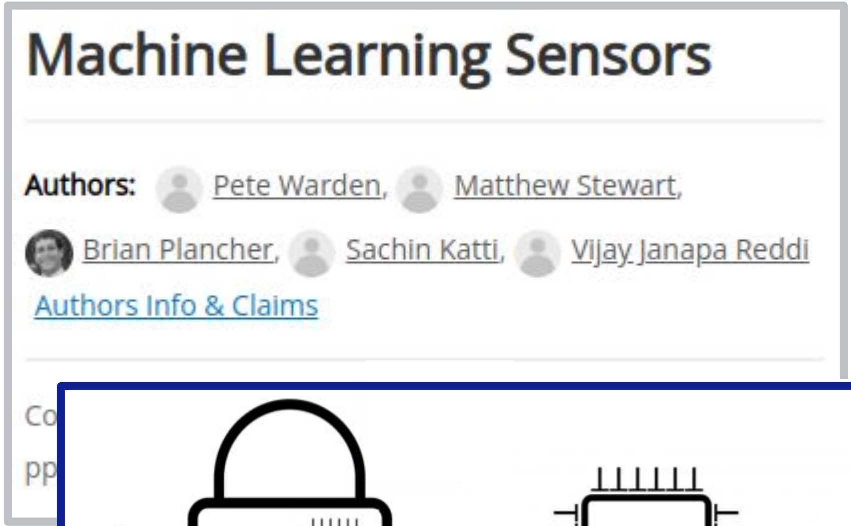
IoT 1.0:
Internet
of Things

↓

IoT 2.0:
Intelligence
on Things

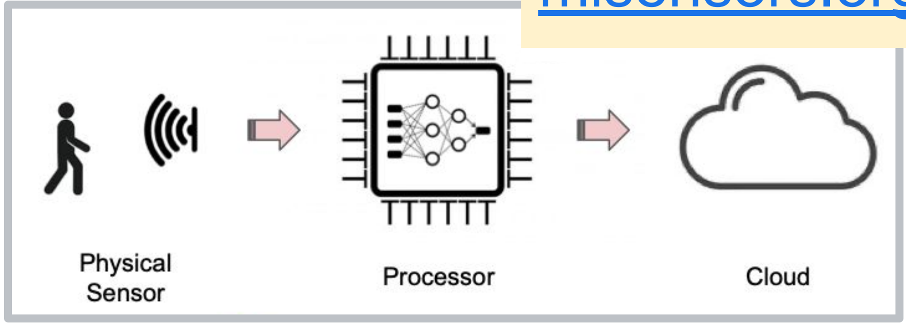
What is a Machine Learning Sensor?

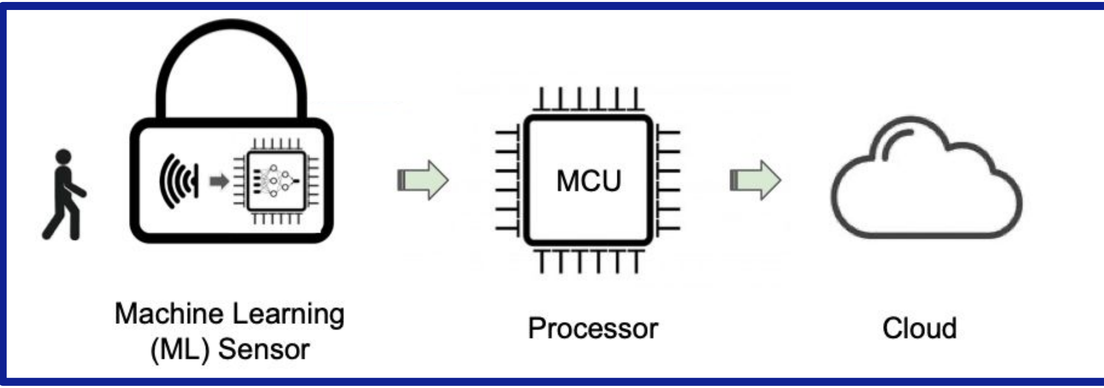
mlsensors.org


Machine Learning Sensors

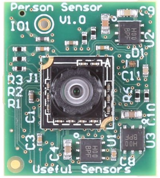
Authors: Pete Warden, Matthew Stewart, Brian Plancher, Sachin Katti, Vijay Janapa Reddi

Authors Info & Claims





Sensor 2.0



Privacy by Design

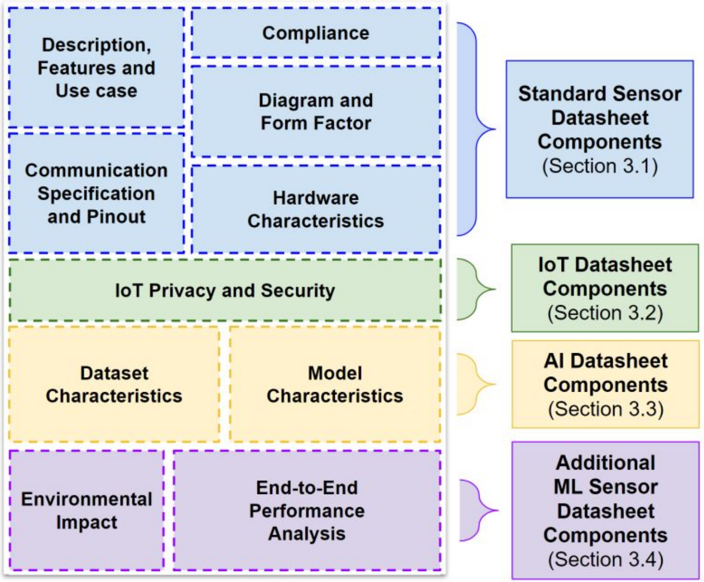
We suggest **transparency** as a core value

Datasheets for Machine Learning Sensors:
Towards Transparency, Auditability, and Responsibility for Intelligent Sensing

MATTHEW STEWART, Harvard University,
PETE WARDEN, Stanford University, Useful Sensors,
YASMINE OMRI, Harvard University,
SHVETANK PRAKASH, Harvard University,
JOAO SANTOS, Harvard University,
SHAWN HYMEL, Edge Impulse,
BENJAMIN BROWN, Harvard University,
JIM MACARTHUR, Harvard University,
NAT JEFFRIES, Useful Sensors,
SACHIN KATTI, Stanford University,
BRIAN PLANCHER, Barnard College, Columbia University,
VIJAY JANAPA REDDI, Harvard University,

arxiv.org/abs/2306.08848

mlsensors.org

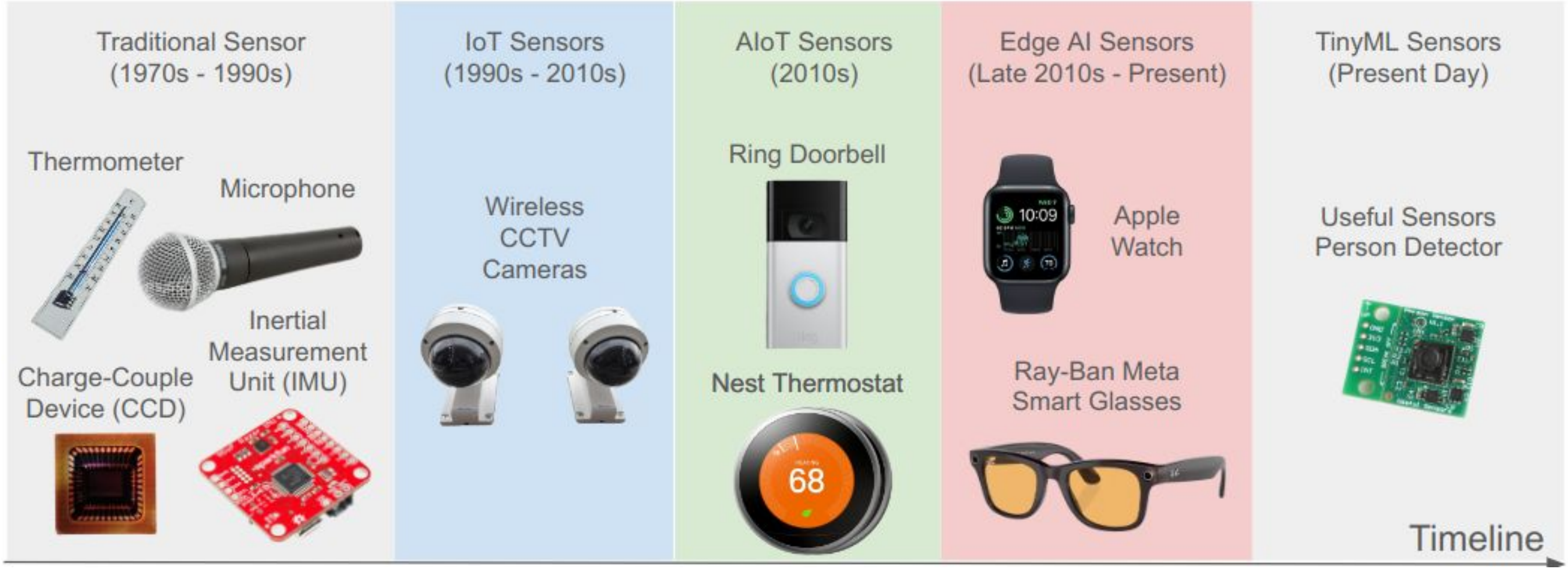


Materiality and Risk in the Age of Pervasive AI Sensors

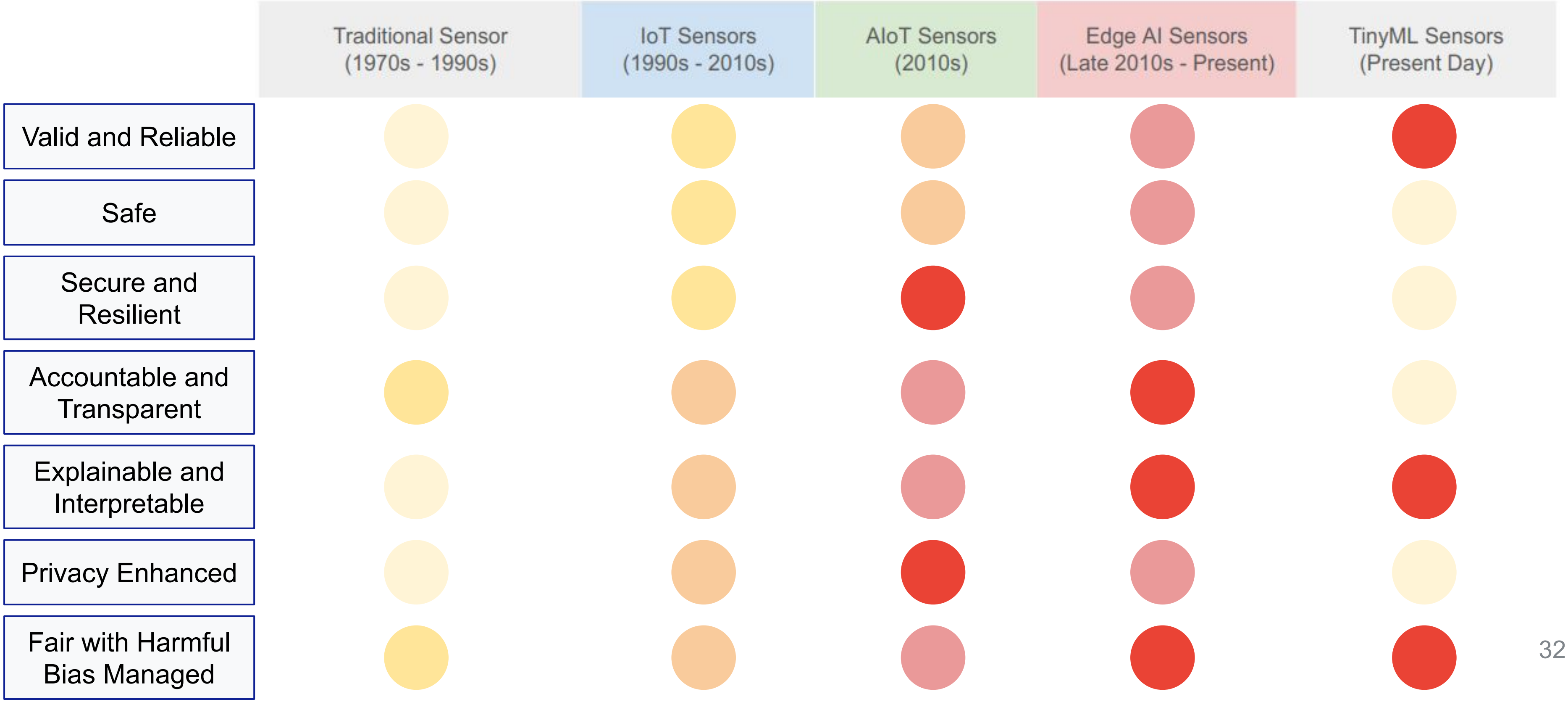
arxiv.org/abs/2402.11183

Brian Plancher
Barnard College, Columbia University
brianplancher.com

Evolution of Sensors…



… and their impact on Responsible AI



How can TinyML support **Responsible AI**?

Accessibility / Education

Sustainability / Conservation

Privacy / Security

Brian Plancher
Barnard College, Columbia University
brianplancher.com