Leveling the Playing Field for Edge AI Research Through High Quality Datasets

Published by EDGE AI FOUNDATION Datasets & Benchmarks Working Group:

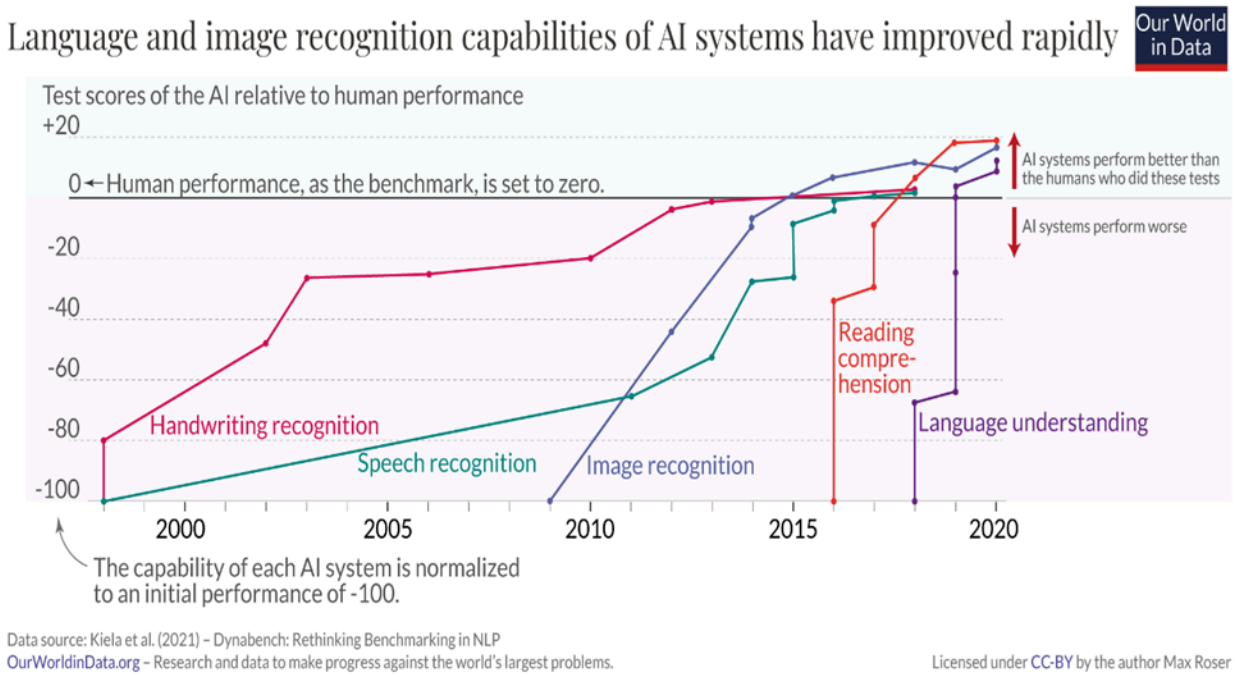
- **Adam Fuks - NXP, Chair**
- **Petrut Bogdan - Innatera**
- **Vijay Janappa Reddi - Harvard University**
- **Eiman Kanjo - Imperial College**
- **Colby Banburry - Harvard University**
- **Sam Al Attiyah - imagimob**
- **Xianghui Wang - Renesas**
- **Emil Jorgenson Njor - Technical University of Denmark**

Introduction

The past decade has witnessed remarkable progress in neural network (NN) techniques—from innovative topologies and training methods to quantization-aware approaches, data augmentation, and model compression. Image recognition, powered by datasets like ImageNet, and natural language processing (NLP), driven by vast internet-scale text corpora, have seen especially significant advances.

These breakthroughs have enabled AI systems to rival and even surpass human performance in specific tasks. A key enabler has been access to large, high-quality datasets used for training, validation, and testing.

However, as model size and complexity continue to grow, this progress raises challenges when deploying AI on edge devices. Unlike cloud servers, edge devices operate under stringent power, memory, and compute constraints—often battery-powered and thermally limited. These constraints demand smaller, more efficient models that preserve accuracy while fitting within tight energy and memory budgets.



This is the domain of **edge AI**. Success here requires tailored techniques and thoughtful benchmarking. Yet, a major barrier remains: a fragmented landscape and a lack of standardized, high-quality datasets that reflect real-world edge use cases.

While many publications tout efficiency gains in "tiny" or edge ML, they often rely on toy examples that fail to translate into production-ready applications. What's missing is a shared, credible benchmark for comparing performance in realistic environments.

Creating a Level-Playing Field for tinyML and Edge AI Research

The **EDGE AI FOUNDATION Datasets & Benchmarks Working Group** was formed to address this gap. Our goals are to:

- Curate realistic, appropriately sized datasets—expanded through community collaboration
- Support open research into performance trade-offs (power, memory, accuracy, etc.)
- Foster shared learning and optimization across the ecosystem

A public repository of datasets specific to edge AI and tinyML use cases will empower researchers, developers, and companies to evaluate their models with confidence. It will also help align innovations with real-world deployment scenarios.

Importantly, the Working Group does not judge submission quality or run official benchmarks. Instead, our role is to **enable honest, community-driven comparison**—creating the infrastructure for innovation.

Choosing Use Cases Thoughtfully

Selecting the right datasets—and corresponding use cases—is essential to developing meaningful benchmarks. The edge AI community spans a wide range of applications, each with unique technical requirements.

To ensure broad and relevant coverage, we propose organizing use cases along several dimensions:

- **Real-time vs. batched processing:** Tasks like fall detection require instantaneous response, whereas others may benefit from batch analysis.
- **Energy constraints:** Devices powered by batteries (e.g., wearables, sensors) require ultra-low power consumption. Wall-powered devices are more flexible.
- **Always-on operation:** Applications such as health monitoring or predictive maintenance require continuous inference and pose challenges for durability and efficiency.
- **Task nature:** Classification tasks differ significantly from regression or transformation tasks, influencing model architecture and evaluation metrics.

- **Data modality:** Whether processing time-series, images, or audio, edge solutions must be evaluated based on their specialized input types.

By mapping benchmarks to these categories, we aim to highlight systems' **capabilities**, not just their performance on isolated tasks.

Improving Datasets, Together

High-quality data is the foundation of trustworthy ML. It's not just about training data—it's about ensuring **testing and validation** reflect the complexity of the real world.

A poor test set can misrepresent a model's performance. A great model might perform poorly if tested on unrealistic scenarios, and a weak model might appear impressive under narrow conditions.

That's why **continually updated, diverse, and well-labeled datasets** are vital. The EDGE AI FOUNDATION is committed to creating and maintaining datasets that reflect evolving needs and innovations in the field.

For each dataset, we will:

- **Build on existing work:** Where possible, enhance and expand proven datasets
- **Ensure variety:** Capture a wide range of real-world scenarios and edge cases
- **Provide rich metadata:** Label data accurately and comprehensively
- **Evolve continuously:** Update test sets to stay aligned with state-of-the-art models

Our first focus is **Visual Wake Words**, with future expansions into other modalities. Each dataset will be vetted for generalization across edge use cases and equipped with the metadata needed for effective benchmarking.

Your Role in Building the Future

The success of this initiative depends on the community. Here's how you can contribute:

- **Suggest datasets or base sets** you believe are valuable starting points or worth improving
- **Provide feedback** on case coverage and diversity

- **Contribute new test cases** that better reflect real-world usage
- **Assist with labeling and annotations** to improve quality and usability
- **Expand edge case scenarios** by contributing niche or underrepresented data

Together, we can build a robust foundation that supports honest comparisons, accelerates development, and unlocks new possibilities for edge AI.

A Call to Action

To truly advance edge AI and tinyML, we must move beyond toy benchmarks and embrace community-led, production-grade testing environments.

We call on the EDGE AI FOUNDATION community to:

- **Share the challenges you're facing** and help us prioritize use cases
- **Contribute datasets and evaluation techniques** that align with your goals
- **Collaborate on establishing optimization best practices** for meaningful benchmarking

Let's build a shared, open, and inclusive ecosystem that drives edge AI forward—together.

JOIN US
joinus@edgeaifoundation.org