EURECOM
Sophia Antipolis

Four Forums Later:
How GenAI at the Edge Has Evolved

Roberto Morabito
roberto.morabito@eurecom.fr

The Generative Edge AI Working Group

The Generative Edge AI Working Group

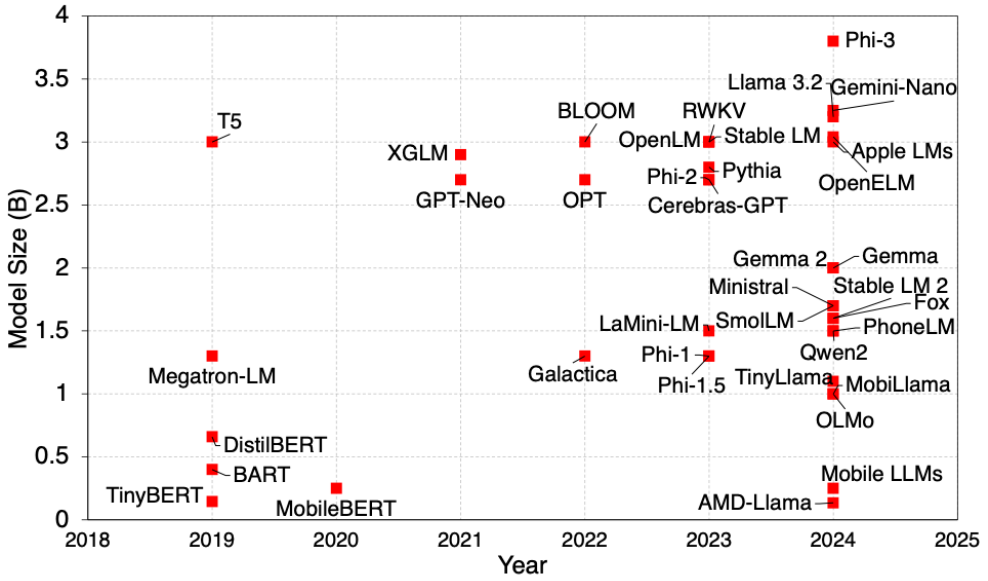
Mission Statement

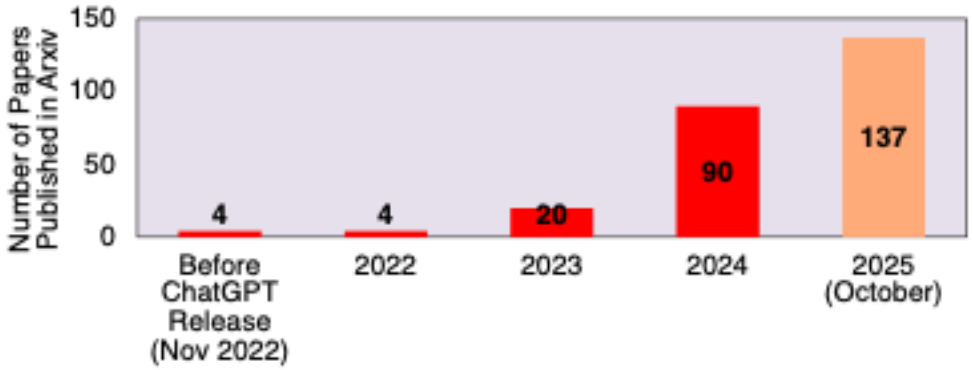
The **Generative Edge AI Working Group** empowers and connects ***academia***, ***industry***, and ***individuals*** to advance knowledge, collaboration, and innovation in Edge AI through education, community engagement, and recognition of groundbreaking achievements.

Read about the WG mission

The Feasibility Shift

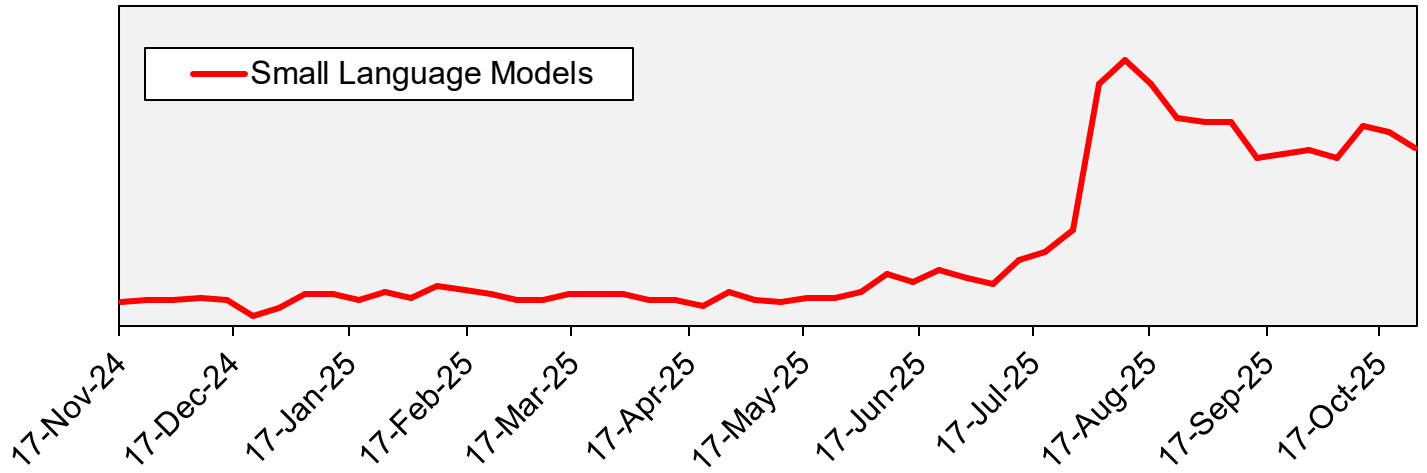
- Growth of **sub-4B parameter language models** over recent years
- Increasing trend toward **compact models optimized** for edge and mobile deployment



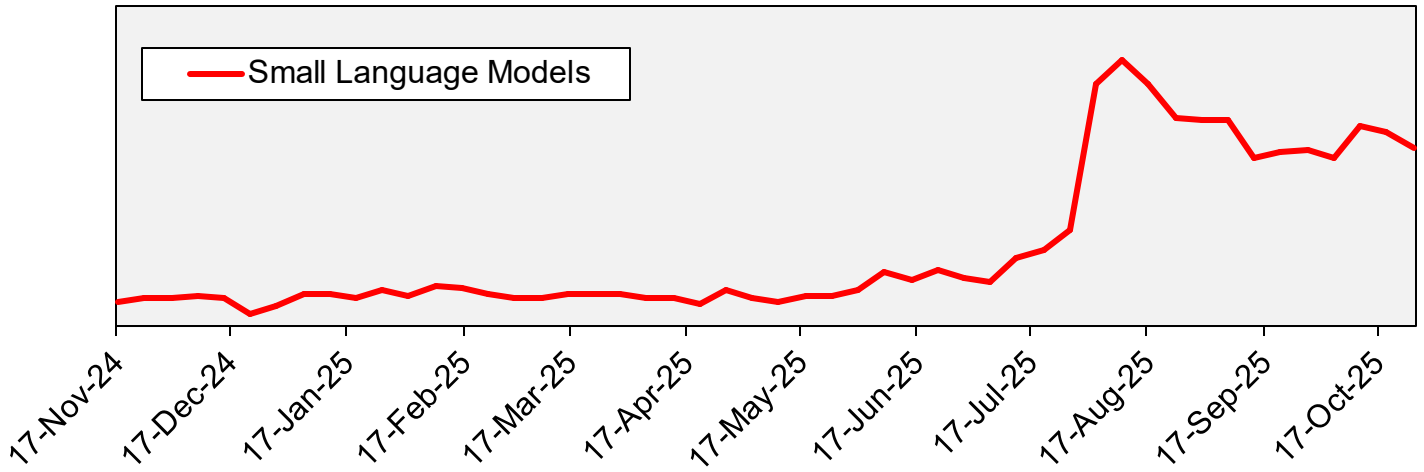


- Research papers published on arXiv mentioning 'Small Language Models' (SLMs) in their title before and after the release of ChatGPT
- Rising research interest in edge-optimized AI models.

Gen AI @ Edge is Trending

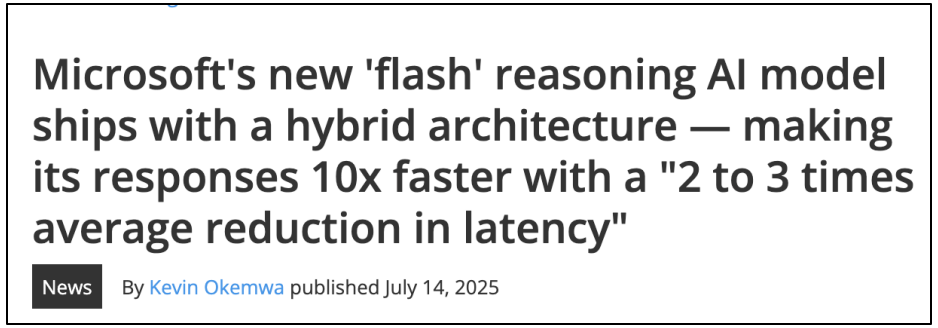


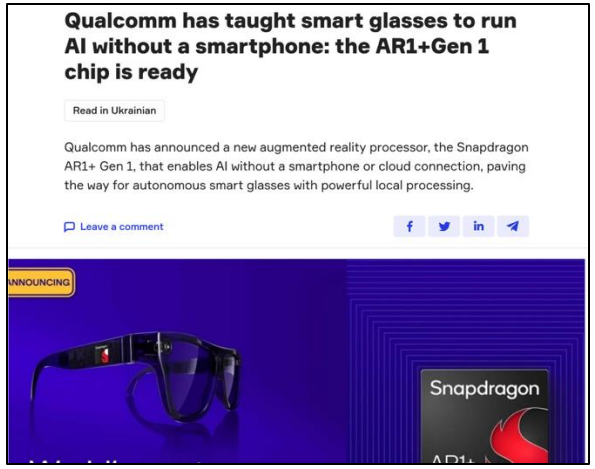
Gen AI @ Edge is Trending

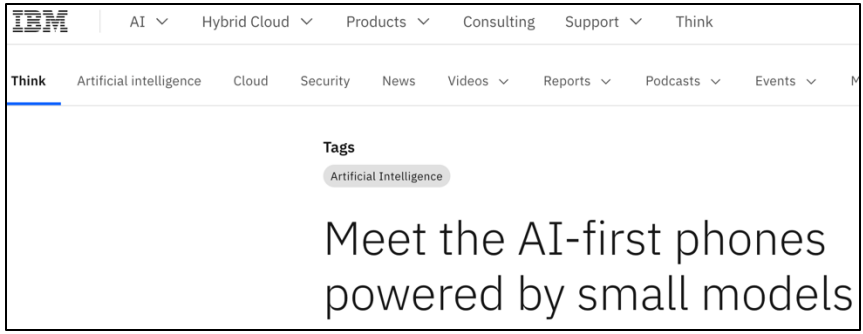


Google I/O

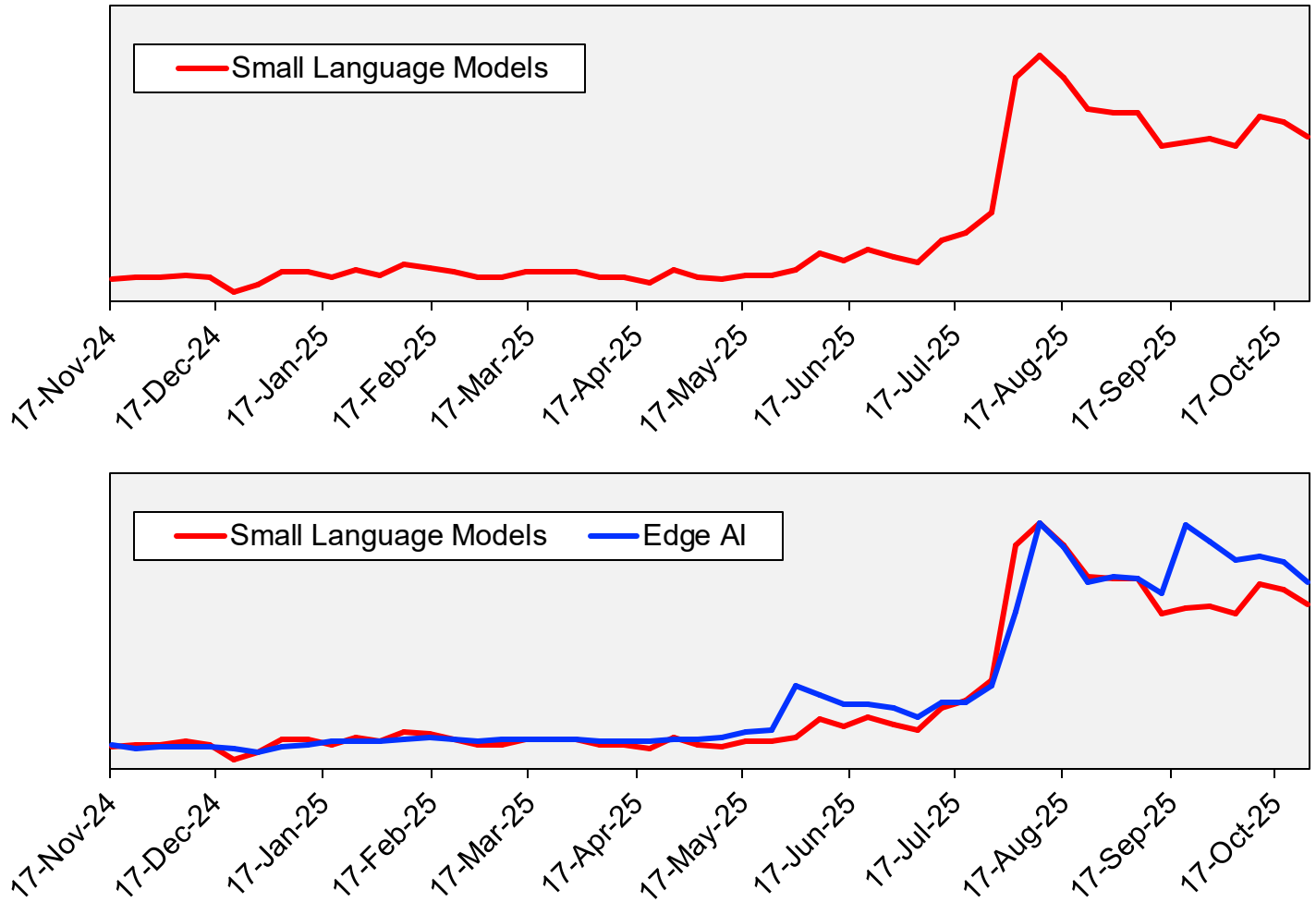
Small language models with Google AI Edge



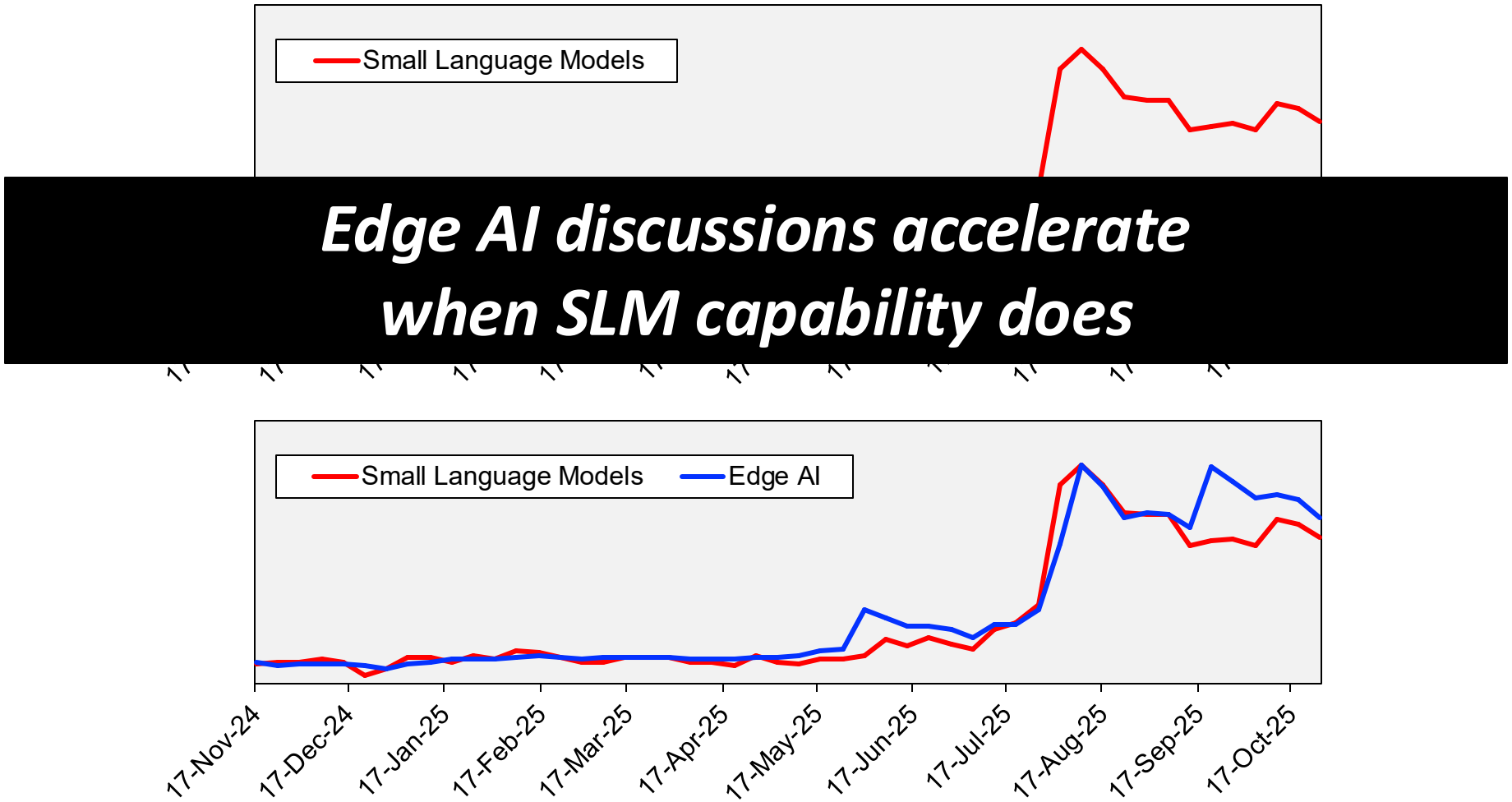




Gen AI @ Edge is Trending



Gen AI @ Edge is Trending



Edge AI discussions accelerate when SLM capability does

EURECOM
Sophia Antipolis

Four Forums Later:
How GenAI at the Edge Has Evolved

Generative Edge AI Forums

1st Forum March 2024

2nd Forum October 2024

3rd Forum May 2025

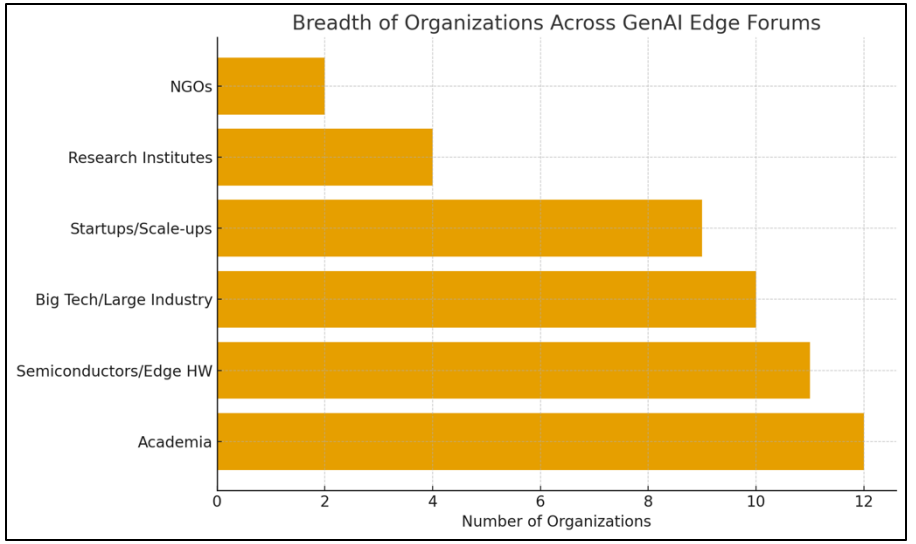
4th Forum November 2025

Generative Edge AI Forums

*Across four forums, we brought together **48 organizations** spanning **universities, semiconductor leaders, startups, major tech companies, research labs, and even NGOs**, highlighting the global and cross-disciplinary nature of Generative Edge AI.*

Generative Edge AI Forums

Across four forums, we brought together ***48 organizations*** *spanning* ***universities, semiconductor leaders, startups, major tech companies, research labs, and even NGOs****, highlighting the global and cross-disciplinary nature of Generative Edge AI.*



Generative Edge AI Forums

1st Forum March 2024 | 2nd Forum October 2024 | 3rd Forum May 2025 | 4th Forum November 2025

Five Waves of Evolution (2024 → 2025)

1. **On-device feasibility** — "Can we run SLMs?"
2. **Edge-optimized pipelines** — compilers, Arm optimizations, toolchains.
3. **Distributed edge intelligence** — multi-device inference, agentic systems.
4. **Verticalization** — automotive, healthcare, industrial, wearables.
5. **Multimodal + agentic edge AI** — sensors, gestures, audio, biosignals.

Generative Edge AI Forums – Wave 1

From "Small LLMs at the Edge" → To a Full Generative Edge AI Ecosystem

Then

Focus was on **porting**, **optimizing**, and **running** small LLMs on devices. *Examples:*

- *"**Running** an LLM on a Raspberry Pi"*
- *"MobileLLM: Optimizing Sub-billion Parameter Models"*
- *"Optimizing LLM Inference for ARM CPUs"*

Now

The conversation shifted from just "running models" to **full generative AI stacks**: toolchains, pipelines, multi-modal models, agent frameworks. Examples:

- *"**Accelerating** LLMs at the Edge"*
- *"AI Backbone Toolchains for GenAI"*
- *"Edge AI Suites"*

Generative Edge AI Forums – Wave 2

From Device-Level Optimization → To Distributed, Multi-Device Collaboration

Then

Heavy emphasis on **single-device inference** and **on-device optimization**.

- *"**On-Device** Generative AI"*
- *"LLM Pipelines on Embedded Devices"*

Now

Emergence of **distributed** and **agentic** edge systems.

- *"**Distributed** SLM-based Agentic AI for the Edge"*
- *"Agent Systems on the Edge"*
- *"**E2EdgeGenAI**"*

Generative Edge AI Forums – Wave 3

From General Demonstrations → To Vertically Specialized Applications

Then

Most talks were ***generic***: “GenAI on the edge”, “LLM pipelines”, “Running an LLM on a Pi”.

Now

Clearly shifting toward **industry-grade verticals**:

- **Automotive**: *“Generative Edge AI in Automotive”*
- **Healthcare/Biosensing**: *“GenAI for Biosensors/Cardio”*
- **Industrial**: *“Industrial Edge: Old Meets New”*
- **Wearables** → AVs: *“GenAI at the Edge: Wearables→AVs”*

Generative Edge AI Forums – Wave 4

From Language Models → To Multimodal and Cross-Modal Edge Intelligence

Then

Mainly **LLMs** and textual use cases.

Now

Strong presence of **multimodality**:

- *"Multimodal Hand Gesture Modeling"*
- *"Artificial Sensor Intelligence & Health"*
- *"Real-Time Audio Denoising (aTENNUate)"*
- *"Visual Language Models for Edge 2.0"*

Generative Edge AI Forums – Wave 5

From Research Vision → To Concrete Engineering: Toolchains, Workflows, Code

Then

Big visions:

- "Toward a Foundation Model for Efficient Damage Assessment"
- "Solve Edge AI Problems with Foundation Models"

Now

Talks focus on workflows, software architectures, and complete toolchains:

- "AI Backbone Toolchains for GenAI"
- "Proposal of workflow and software architecture for complex EdgeAI apps"
- "Advancing LLMs in Resource-Constrained Environments"

Generative Edge AI:
Mission, Vision, and Insights from Industries

Expanding The Horizons of Generative Edge AI: Mission, Vision, and Insights From Industries

Roberto Morabito
Communication System Department
EURECOM
Valbonne, France
roberto.morabito@eurecom.fr

Riccardo Adorante
System Research and Applications
STMicroelectronics
Agrate, Italy
riccardo.adorante@st.com

Hajar Mousannif
Department of computer science
Cadi Ayyad University
Marrakesh, Morocco
hajar.mousannif@gmail.com

Danilo Pietro Pau, FIEEE
System Research and Applications
STMicroelectronics
Agrate, Italy
danilo.pau@st.com

SCAN ME

Generative Edge AI:
Mission, Vision, and Insights from Industries

Expanding The Horizons of Generative Edge AI: Mission, Vision, and Insights From Industries

Roberto Morabito
Communication System Department
EURECOM
Valbonne, France
roberto.morabito@eurecom.fr

Riccardo Adorante
System Research and Applications
STMicroelectronics
Agrate, Italy
riccardo.adorante@st.com

Hajar Mousannif
Department of computer science
Cadi Ayyad University
Marrakesh, Morocco
hajar.mousannif@gmail.com

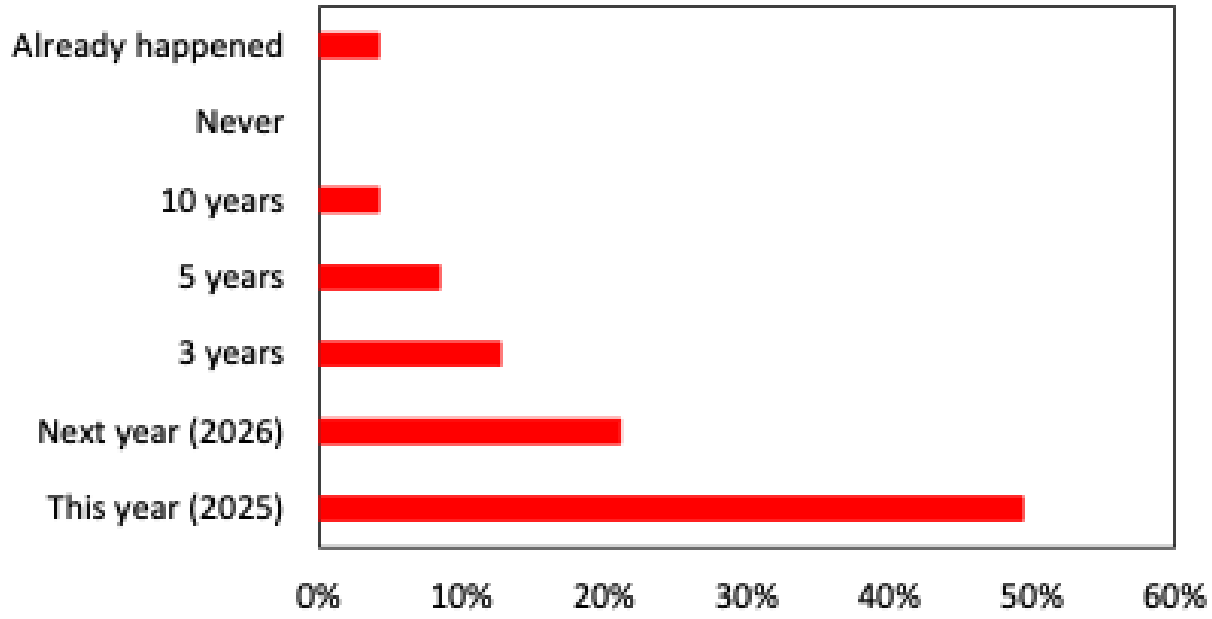
Danilo Pietro Pau, FIEEE
System Research and Applications
STMicroelectronics
Agrate, Italy
danilo.pau@st.com

SCAN ME

We ran a survey among the Edge AI Foundation partners.

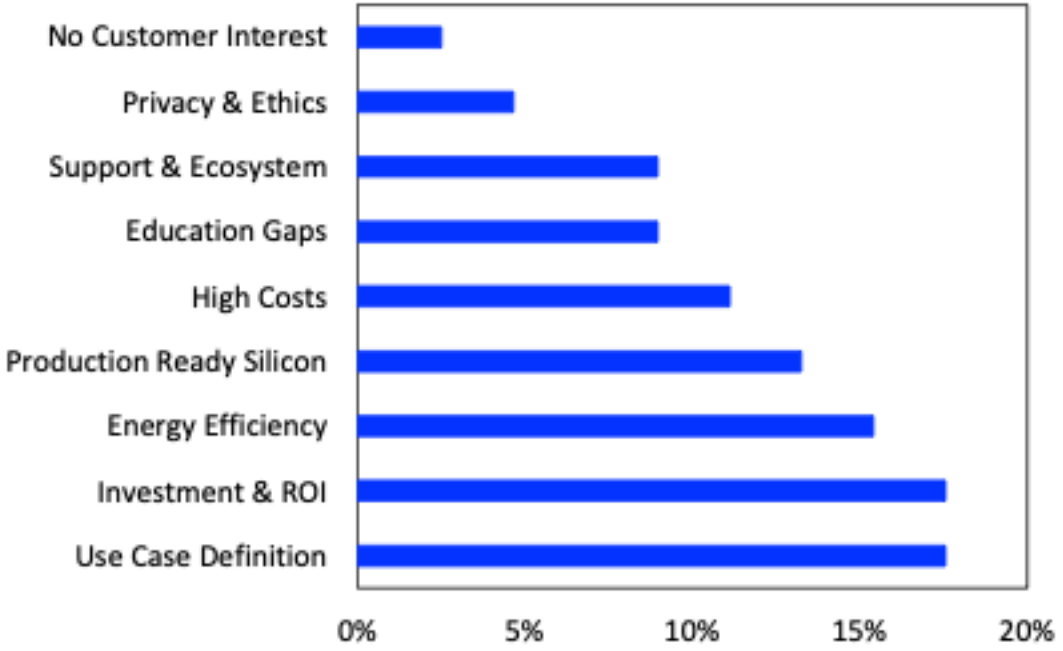
Generative Edge AI
Market Reachability Timeframe

When do you expect Generative Edge AI solutions to start reaching the market?



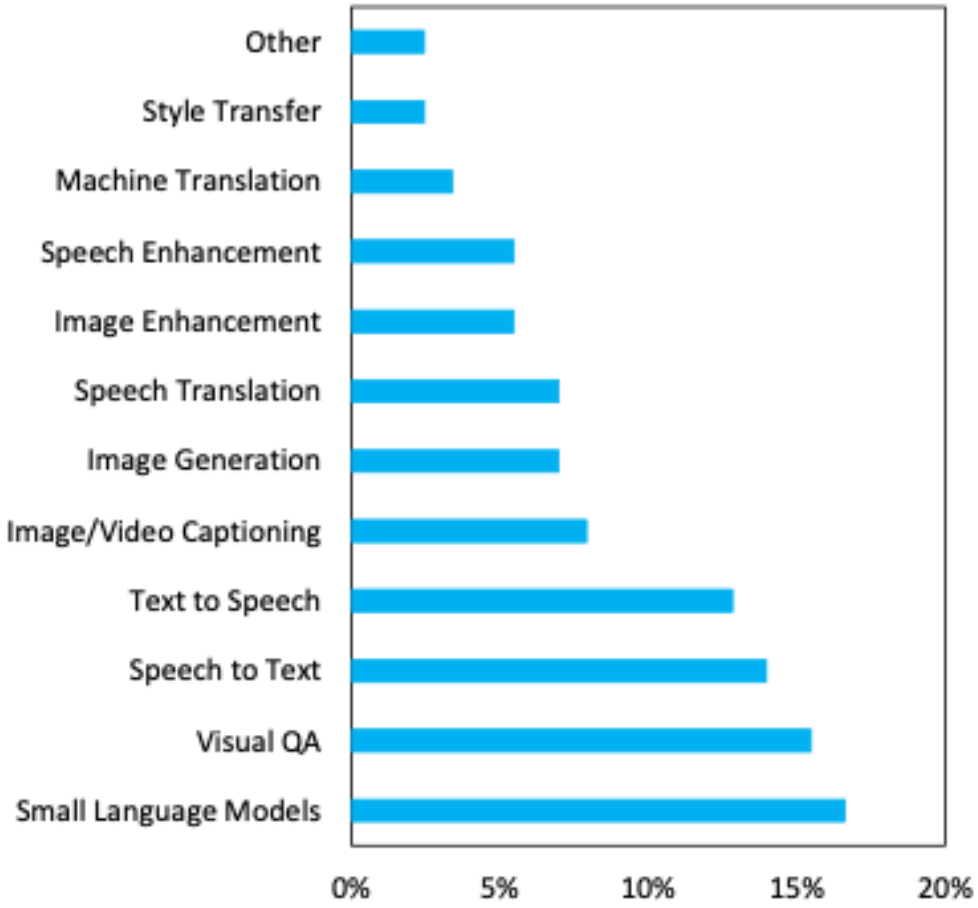
Generative Edge AI Adoption Challenges

What are the main barriers to adopting Generative Edge AI in your organization?



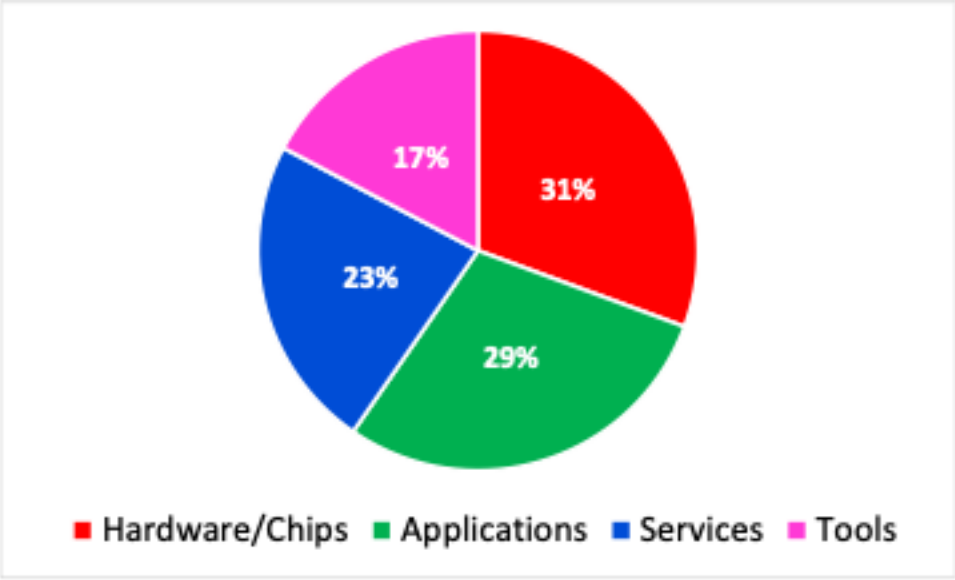
Generative Edge AI Use Cases

Which use cases for Generative Edge AI are most relevant or promising?

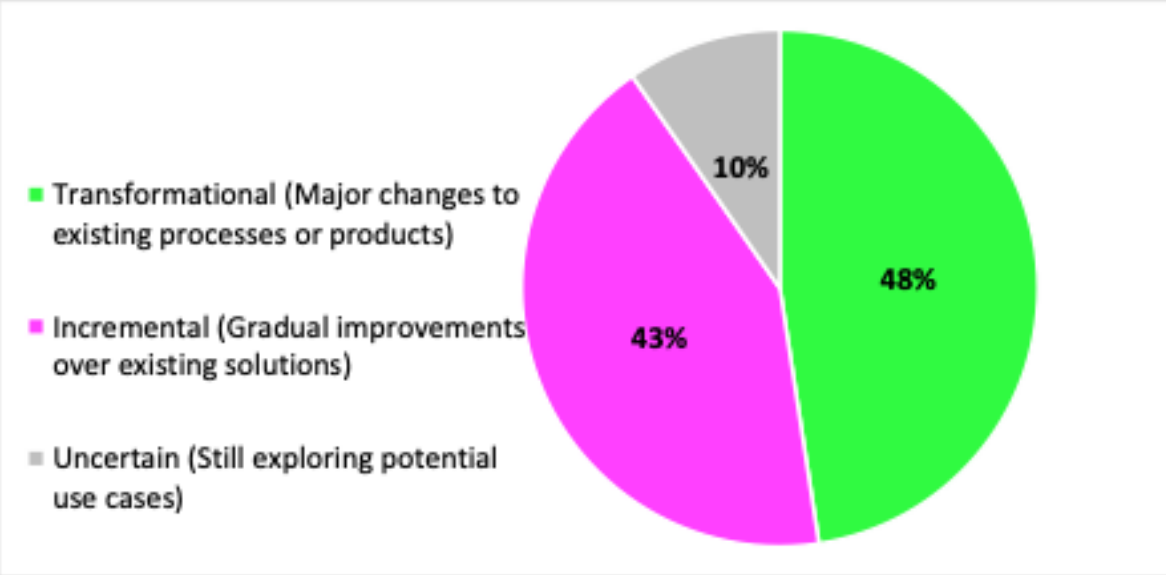


Generative Edge AI Impact

Which types of Generative Edge AI solutions are most likely to emerge?

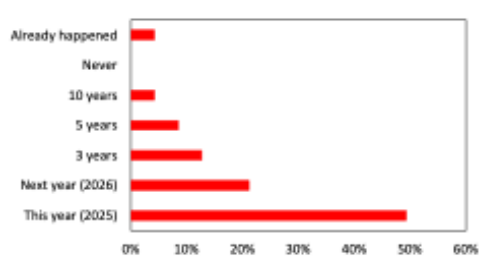


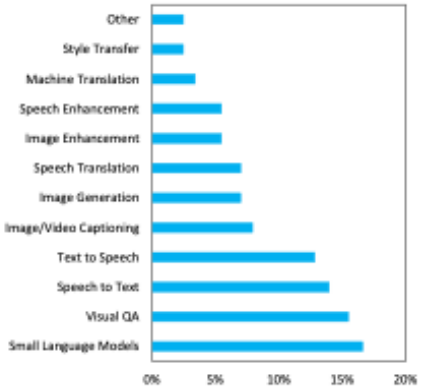
How significant do you expect the impact of Generative Edge AI to be in your industry?

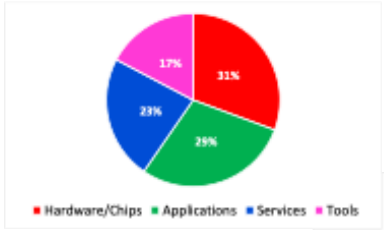


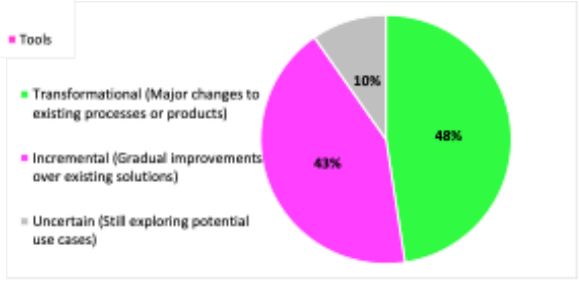
What These Trends Tell Us

1. Adoption is Near-Term, Not Long-Term
2. SLMs + Cross-Modal Models Lead the Pack
3. Innovation Spans Hardware → Tools → Apps
4. Impact Will Be Transformational









What's Next for Generative Edge AI?

- The next phase is **not about new models** but operationalizing what we already have.
- **Agentic workflows** across devices and **coordination** of small models running on **heterogeneous hardware**.
- **Increasing domain-specific foundation models** at the edge (healthcare, industrial, automotive, wearables).
- **Tooling and reproducibility** will become the real competitive advantage: **Whoever builds the easiest pipelines wins**.

What's Next for Generative Edge AI?

- The next phase is **not about new models** but operationalizing what we already have.
- **Agentic workflows** across devices and **coordination** of small models running on **heterogeneous hardware**.
- **Increasing domain-specific foundation models** at the edge (healthcare, industrial, automotive, wearables).
- **Tooling and reproducibility** will become the real competitive advantage: **Whoever builds the easiest pipelines wins**.

The Edge AI Foundation community will be the catalyst

We now have the breadth, the forums, and the people to drive this evolution.

EURECOM
Sophia Antipolis

Four Forums Later:
How GenAI at the Edge Has Evolved

THANKS!

Roberto Morabito
roberto.morabito@eurecom.fr