KEYNOTE · IEEE 12TH WORLD FORUM ON INTERNET OF THINGS

Nobody Needs to See It

Making AI for IoT private by design with MPC and FHE



Ajith Suresh

Lead MPC Researcher · Cryptography Research Center · Technology Innovation Institute (TII)

Abu Dhabi · 28 October 2026

ABOUT ME

I build systems that compute on data nobody gets to see

Today

Lead MPC Researcher, Cryptography Research Center, Technology Innovation Institute, Abu Dhabi

Before

PhD, Indian Institute of Science, Bengaluru · Postdoc, TU Darmstadt

Focus

Secure multi-party computation, private machine learning, private federated learning, and lately homomorphic encryption

THIS WEEK AT WF-IOT

We heard about AI that senses, interprets and reasons about the physical world.

Who gets to see what it perceives?

THE ANSWER TODAY

Too often: whoever runs, or breaks into, the cloud

Feb 2026 · Robot vacuums

Live camera, microphone and 3D floor plans

A cloud authentication weakness exposed thousands of DJI Romo owners’ homes.

2026 · Cameras and baby monitors

About 1.1 million devices in 118 countries

One key from a white-label vendor’s app reached live feeds and stored photos.

2025 · Smart-home platform

2.7 billion IoT records left open

An unprotected database of a grow-device maker, with device and network details.

Sources: AOL/Business Insider on DJI Romo (Feb 2026); PetaPixel on Meari (May 2026); Infosecurity Magazine on Mars Hydro (Feb 2025)

THE FRAGILE ASSUMPTION

“To get value from the data, someone has to see it.”



What if nobody had to?

WHY USERS CAN’T BE THE DEFENCE

For the end user, privacy is invisible until it is gone

01

The flow is hidden

A thermostat, a vacuum or a car does not show where its data goes or who reads it.

02

Consent is a click

Long policies and “accept” buttons hand over data without real understanding.

03

The harm comes later

Leaks, profiling and misuse surface years after the data left the device.

So privacy has to be a property of the system, not a decision we push onto the user.

THE TRUST LADDER

Most “private” systems move the trust; few remove it

Trust the provider

Policies, contracts, access control. Data is in the clear while it is used.

Trust the hardware

Trusted execution environments. Apple’s Private Cloud Compute now also runs on Google Cloud (2026): the trust spans more vendors.

Trust mathematics

MPC and FHE. The data stays encrypted or split while it is computed on. No single party ever sees it.

Source: InfoQ, “Apple Private Cloud Compute expands to Google Cloud”, July 2026

THE GAP

We encrypt data when it moves and when it rests, but not while it is used



In transit

TLS, VPNs, encrypted radio links

Protected



At rest

Disk and database encryption

Protected



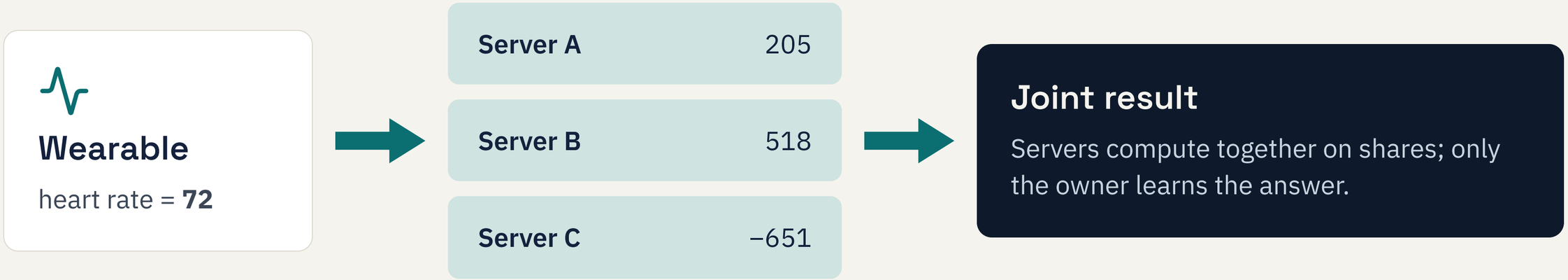
In use

Analytics, training, AI inference

Decrypted, exposed

SECURE MULTI-PARTY COMPUTATION (MPC)

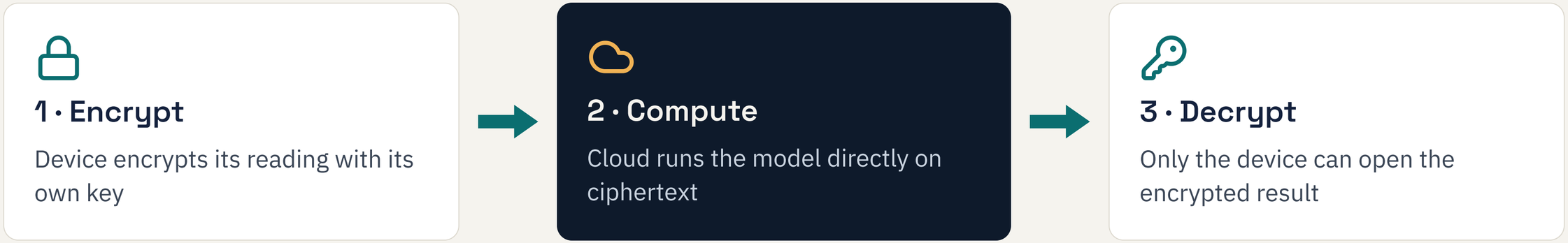
Split the secret, compute on the pieces



Each share alone looks like random noise. Privacy holds as long as the servers do not **all** collude.

FULLY HOMOMORPHIC ENCRYPTION (FHE)

Lock the data, let the cloud work through the glovebox



The cloud never holds the key. One server is enough; the price is heavy computation.

TWO TOOLS, DIFFERENT COSTS

MPC and FHE fit different IoT shapes, and work best combined

	MPC	FHE
Who computes	Several independent servers	A single server
Trust assumption	Servers do not all collude	Only the key holder can decrypt
Main cost	Communication and rounds	Computation and ciphertext size
Natural IoT fit	Many devices, joint analytics and training	One device asking one cloud

Hybrid protocols use FHE where it is cheap and MPC where it is fast.

PATTERN 1 · DEVICES THAT LEARN TOGETHER

Devices share a place, so they could share intelligence

- Thermostat
- Camera
- Wearable
- Voice assistant
- Smart meter
- Robot vacuum

TODAY

One cloud collects everything and coordinates the devices. It sees all of it.

POSSIBLE

Devices learn and decide together, and the data never leaves them in the clear.

FEDERATED LEARNING ON TINY DEVICES

Train where the data lives; share only model updates

1. Each device trains on its own readings
2. It sends a model update, not the raw data
3. An aggregator averages updates into a global model
4. The improved model goes back to every device

ALREADY PRACTICAL

On-board training on IoT devices, combining federated and transfer learning, reached **86.48%** accuracy in a recent study.

Ficco et al., Information Fusion, 2024

M. Ficco et al., “Federated learning for IoT devices: Enhancing TinyML with on-board training”, Information Fusion 104 (2024) · github.com/iotresearchunisa/FL_TinyML

THE CATCH

Federated is not private: the aggregator still sees every update



Updates leak data

Gradients can be inverted to reconstruct training samples, such as images, from a single device's update.



One server, one target

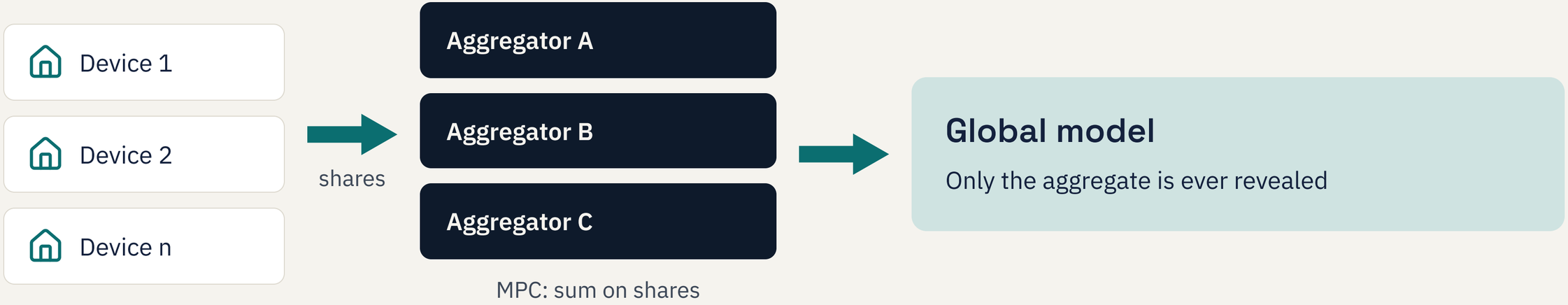
A single aggregator is the same honeypot as before, and it can also hand out a wrong model.

What we need: the **sum** of the updates, without anyone seeing any **single** update.

L. Zhu, Z. Liu, S. Han, "Deep Leakage from Gradients", NeurIPS 2019

FROM TII · PETALGUARD

Replace the one aggregator with several that never see an update



Built on the Flower framework

Handles device dropouts

GPU-accelerated

Scales to large models

PATTERN 2 · WEAK DEVICES BORROWING INTELLIGENCE

Small devices can’t run big models, so every query goes to the cloud in the clear

Health sensors

Skin images, ECG, glucose sent for diagnosis

Voice and assistants

Speech and prompts sent to large language models

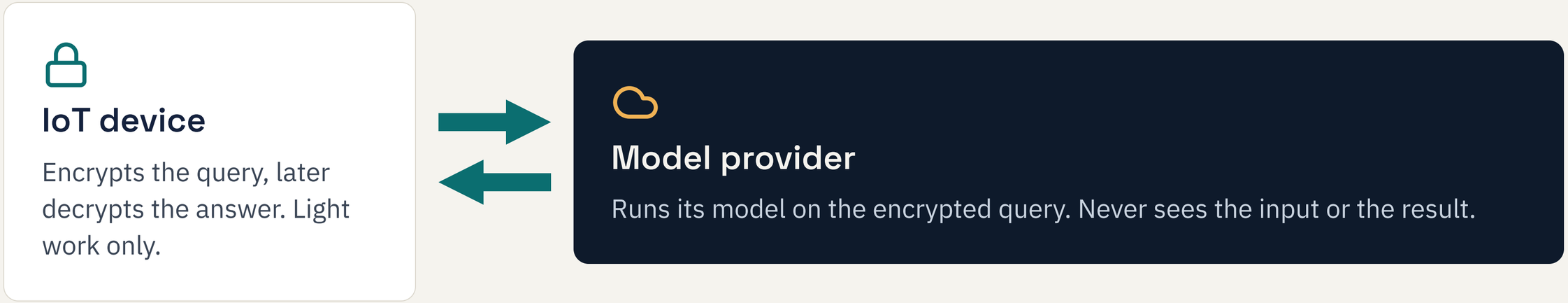
Cameras

Video frames sent for detection and recognition

The provider promises privacy. The breaches show what a promise is worth.

PRIVATE INFERENCE

The device asks in ciphertext and gets the answer in ciphertext



- User data protected from the provider
- Provider’s model stays on its side

FROM TII · VEIL

One private-inference library, from small classifiers to large language models

SMALL MODELS · FHE ONLY

One device, one server, no extra parties

Use cases built so far:

melAdvise · melanoma skin-cancer detection

PhiShark · spam detection

LARGE MODELS · FHE + MPC

Hybrid protocols that scale up to LLMs

FHE for the linear layers, MPC for the operations FHE finds expensive, each where it is cheapest.

THE PRICE OF PRIVACY

The barrier is no longer theory; it is the overhead

IoT budget	What private computation spends
Compute and energy	FHE arithmetic is far heavier than plain arithmetic
Bandwidth	Ciphertexts are much larger than data; MPC parties exchange many messages
Latency	MPC needs rounds of interaction between servers
Memory	Keys and ciphertexts can exceed what a microcontroller holds
Model design	Non-linear layers such as ReLU and softmax are the most expensive part

Someone pays the extra cost: the provider, the device, or the user’s patience.

CLOSING THE GAP

Four levers are bringing the cost down

Smarter protocols

Do the heavy, data-independent work ahead of time; keep the live phase fast.

Hybrid designs

Combine FHE, MPC and trusted hardware, each where it is cheapest.

Dedicated hardware

FHE accelerators are moving from research papers to working silicon.

Standards

ISO/IEC 28033 for FHE is in its final draft stages; NIST runs a call for multi-party threshold schemes.

HARDWARE LEVER · INTEL HERACLES, ISSCC 2026

1,074–5,547×

faster than a 24-core Xeon CPU on core FHE operations, on a single PCIe accelerator card

Source: Tom’s Hardware, “Intel’s Heracles chip computes fully-encrypted data without decrypting it”, March 2026 (vendor-reported figures)

ADOPTION HAS STARTED

Industry and public institutions are already moving

IN PRODUCTS

Apple, 2024: homomorphic encryption for Live Caller ID Lookup and Enhanced Visual Search in Photos

Microsoft Edge, 2021: encrypted password-breach checking

UK–US pilot, 2025: cross-border rare-cancer statistics with PETs; data-sharing approval cut from over a year to about two months

IN EU POLICY

Eurostat: MPC-based shared computing prototype for the European Statistical System (2024–2026)

EU statistics law: amended to reference PETs (2024)

Data Union Strategy, Nov 2025: names federated learning, HE and MPC for EU data labs

Sources: Apple ML Research; J. Kun, “FHE in production”; UK GDS blog, Oct 2025; Eurostat PET4OS / JOCONDE; COM(2025) 835

THE MISSING PIECE

Three communities, each optimising alone

IoT designers

Optimise energy, memory, radio and cost

“Privacy is the cloud’s job.”

AI engineers

Optimise accuracy, model size and latency

“The data will be in the clear.”

PETs researchers

Optimise security and server-side overhead

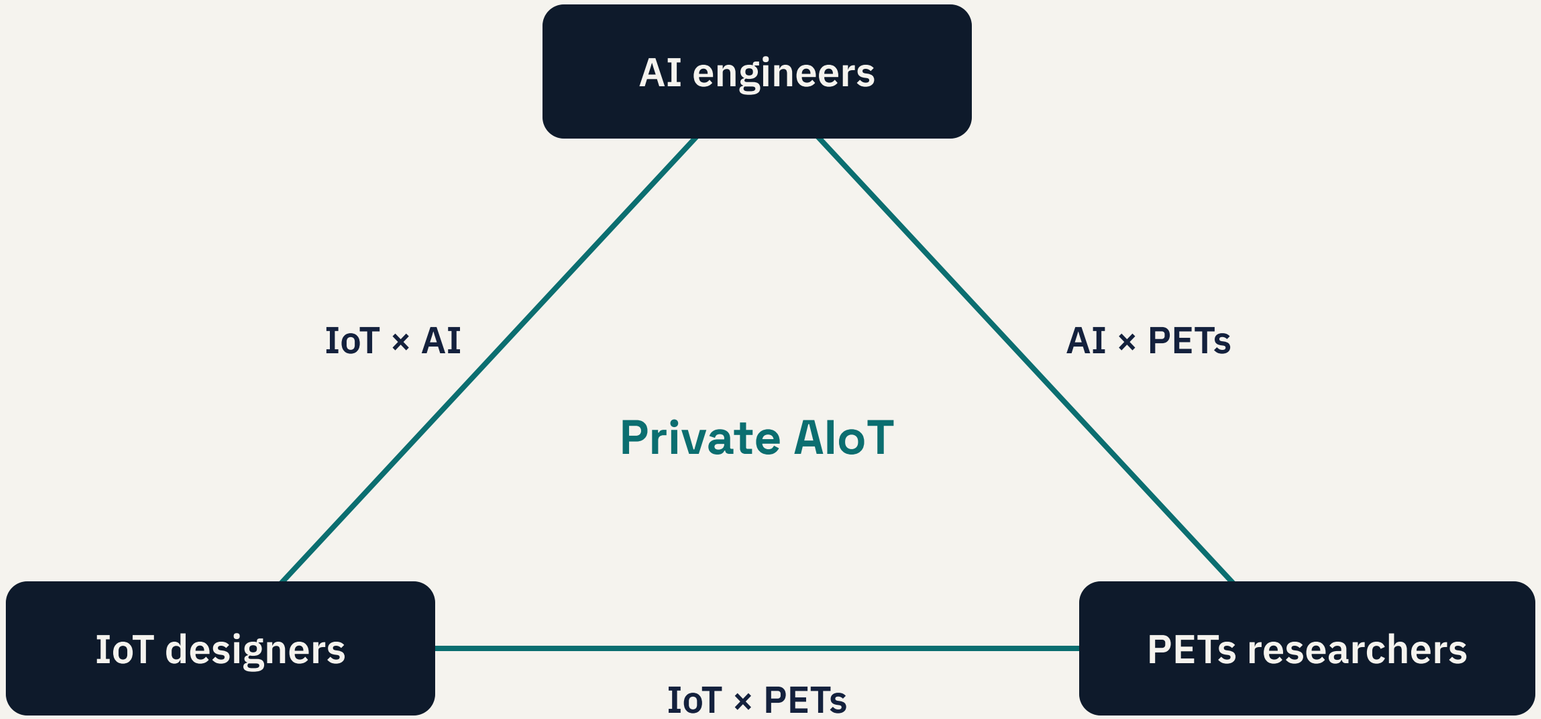
“The model and the device are given.”

Each takes the others’ constraints as fixed, so nobody optimises the whole system.

CO-DESIGN

Efficient private AIoT lives where the three meet

Each edge is a shared design space. Decisions made there remove overhead that no single community can remove alone.



EDGE 1 · AI × PETS

Design the model for encrypted execution, not just for accuracy

Fewer non-linearities

Remove or replace the ReLUs that dominate private-inference cost, trained in from the start.

Crypto-friendly activations

Low-degree polynomial or co-designed activations that FHE and MPC evaluate cheaply.

Fewer bits

Quantised models and updates shrink ciphertexts and messages, e.g. quantised secure aggregation.

Encryption-friendliness should be a design target, like latency or model size.

e.g. Diaa et al., co-designed activations, USENIX Security 2024 · Ben-Itzhak et al., ScionFL, IEEE SaTML 2024

EDGE 2 · IOT × PETS

Shape the protocol around how devices actually live

Precompute while idle

Match the offline/online split to duty cycles: prepare while charging, keep the live phase light.

Reuse existing parties

Home gateways, network operators and vendors can act as the independent MPC servers.

Measure what IoT pays

Report joules, bytes over the air and RAM, not only server seconds.

Expect intermittency

Tolerate dropouts, sleep cycles and lossy links by design, not as an afterthought.

EDGE 3 · IOT × AI

Make on-device AI ready to plug into private computation

Split the work

Decide what runs on the device, the gateway and the cloud with privacy in the cost model.

Budget for it

Leave room in memory, energy and airtime for keys, shares and ciphertexts.

Expose the hooks

TinyML and federated stacks with secure aggregation and encrypted inference built in.

Edge AIoT gateways are a natural place to start.

ONE GUARDRAIL

Faster must never quietly mean weaker

- State the trust assumption: who must not collude, and who could?
- Ask for security proofs and independent review of every shortcut
- Build on vetted libraries and emerging standards, not home-made schemes
- Remember outputs leak too: protect what the result itself reveals

CALL TO ACTION

One ask for each of us

IoT designers

Treat privacy as a hardware and system budget, and tell us your real constraints.

AI engineers

Make encryption-friendliness a design target alongside accuracy, size and latency.

PETs researchers

Benchmark on real devices and IoT workloads, and keep every speed-up provably secure.

Shared benchmarks, shared testbeds, shared venues like this one.

THANK YOU

Nobody needs to see it.

Let’s build AIoT where nobody has to look.



Ajith Suresh · ajith.suresh@tii.ae

Cryptography Research Center, Technology Innovation Institute · linkedin.com/in/ajithsuresh

petalguard.tii.ae