



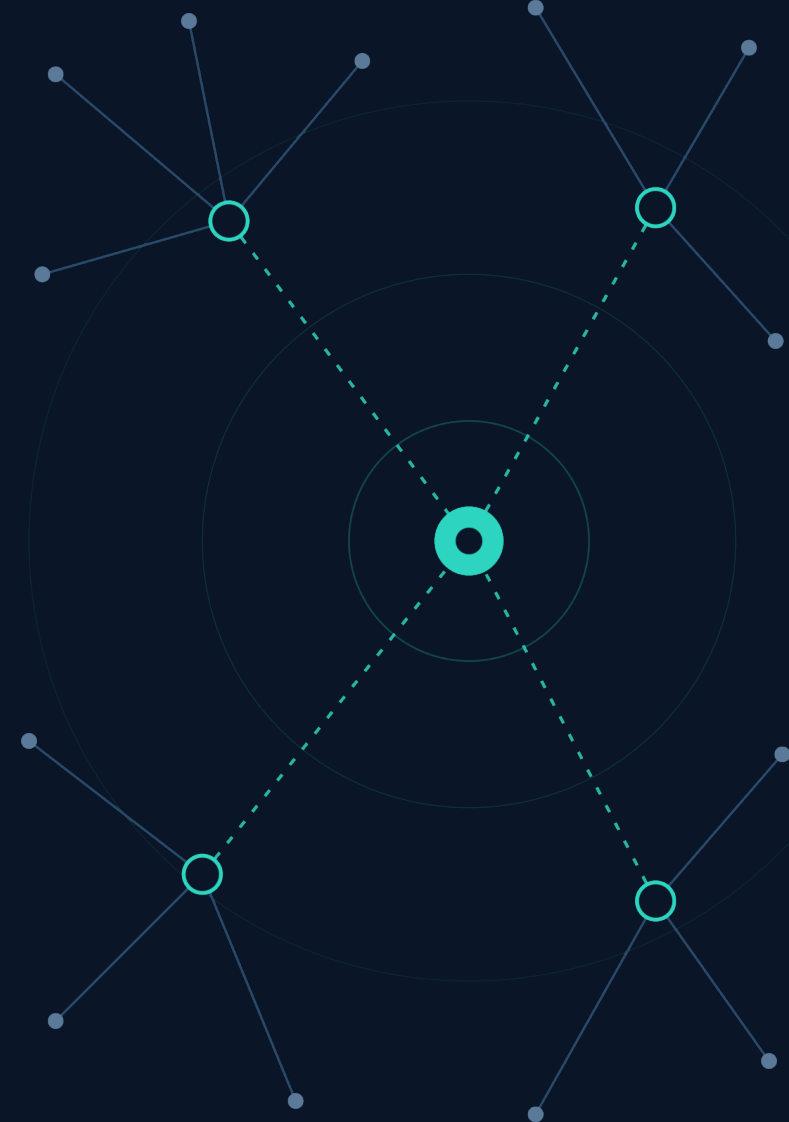
RITE 2026 · BUCHAREST · 27 OCTOBER 2026

Trustworthy edge AI for resilient IoT infrastructure

Autonomous cyber defence with small language models across federated edge environments

Cosmina Stalidi

PhD Candidate in Cybersecurity · National and Kapodistrian University of Athens





Cosmina Stalidi

RESEARCHER · STEAM FOUNDER ·
· COMMUNITY BUILDER

PhD Candidate in Cybersecurity,
National and Kapodistrian University of
Athens

Cybersecurity Researcher, EXARCH

Women4Cyber România · HackShield
România

ABOUT THE SPEAKER

Trust is built in labs and in classrooms classrooms

10+

EU research and innovation projects

Research, proposal writing and dissemination across Horizon Europe, MSCA, DIGITAL Europe and EDF

AIProHealth 2025

Finalist, pitched an AI-for-health tool to the
EIT Health jury

COST Action fellowship

Research fellow, BEiNG-WISE Training
School on cybersecurity

FEPS YAN

Member, FEPS Young Academics Network

STEAM founder

AI and cyber projects for schools ·
European Quality Label



AGENDA

Ten minutes, five steps

01

The problem

A Romanian wake-up call, and why central security falls short

02

The framework in practice

Small language models at the federated edge, in a hospital

03

Trust

Detect, explain, withstand

04

Evaluation

Six dimensions, measured on established IoT datasets

05

Join us

Partnering with CyberSea Festival 2027 and HackShield



FEBRUARY 2024 · ROMANIA

100

hospitals knocked offline by a single single ransomware attack

The attack hit a shared hospital information system: 26 hospitals encrypted, 74 more disconnected as a precaution. Staff went back to paper records.

 One shared dependency. **One point of failure.**



Cloud-centric security reaches its limits at the edge edge



Response latency

Round-trips to a central cloud slow detection and containment when seconds matter.



Scalability

Growing fleets of heterogeneous devices strain centralised analysis pipelines.



Data governance

Moving raw operational data off-site collides with privacy and sovereignty rules.



Unseen threats

Defences tuned to known attacks are of limited use against novel ones.



01 · THE CATCH

The defender becomes a target

AI-enabled security can itself be manipulated. An attacker doesn't need to beat the network, only to fool the model guarding it.

EVASION

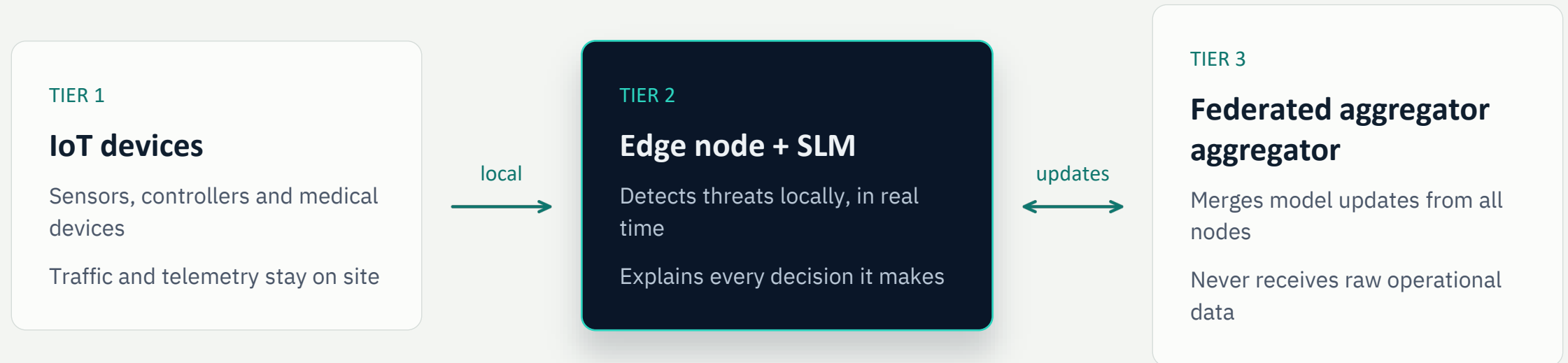
Crafted inputs slip past the detector

POISONING

Corrupted updates degrade what the model learns



Intelligence moves to the edge



Federated learning: nodes improve one shared detection model while raw operational data stays where it was produced.



When a bedside monitor turns attacker

- 1 Detect on site**
The edge SLM flags the anomaly locally, with no cloud round-trip.
- 2 Explain**
A plain-language rationale the operator can check in seconds.
- 3 Contain**
Only the suspicious device is isolated. Care continues.
- 4 Learn together**
A model update, not patient data, protects every hospital in the federation.

• EDGE NODE · ICU-2 · ALERT

DEVICE

Patient monitor #14

NORMALLY

Sends vitals to the nurse station

NOW

Probing 40 internal hosts over SMB

ASSESSMENT

Likely lateral movement, a typical ransomware precursor

ACTION

Isolate segment · alert security team



Trust rests on three interdependent requirements

01

Detect

Both known attacks and previously unseen threats.

02

Explain

Transparent, interpretable decisions that operators can verify.

03

Withstand

Reliable performance under adversarial inputs and compromised federated participants.

A detector that cannot explain itself, or that fails under attack, cannot be trusted — [however accurate it is.](#)



Six dimensions of evaluation

Measured on established IoT cybersecurity datasets

EFFECTIVENESS

Detection performance

Does it catch known and novel attacks?

False-positive rate

Does it avoid flooding operators with false alarms?

EFFICIENCY

Response latency

Can it act in real time at the edge?

Computational efficiency

Does it fit constrained edge hardware?

TRUST

Explanation stability

Do similar events get consistent explanations?

Adversarial robustness

Does it hold up under deliberate attack?



KEY TAKEAWAYS

Three ideas to take home

01

Push detection to the edge edge

Small language models flag threats next to the devices, with no single central point to fail.

02

Learn together, share no raw raw data

Federated updates spread defensive knowledge between sites while sensitive data stays local.

03

Engineer trust, don't assume it

Detection, explanation and robustness are designed together, and measured together.

Distributed · Adaptive · Trustworthy

Starting with connected healthcare, extending to critical IoT infrastructure.



05 · JOIN US

Trust starts with people

Two Women4Cyber România programmes building the next generation of cyber defenders

CYBERSEA FESTIVAL 2027

Romania's summer cybersecurity festival, on the Black Sea coast

CTF & hackathons · Red vs Blue live demos · Career Lab · CyberStartup Zone

2026 partners included ECCC, DNSC, Cisco, Palo Alto Networks and EY

HACKSHIELD ROMÂNIA

Cyber heroes aged 7 to 12

A serious game that teaches children to spot phishing, fake news and online risks, through play, not fear.

722K+

children reached

8

countries

10M

kids safe online by 2030 (goal)

Sponsor · Technology partner · Mentor & speaker · School network

Let's talk after the session →



RITE 2026 · TRUSTWORTHY INTERNET

Thank you

Questions and discussion

Resilient infrastructure runs on trust: in our systems, and in the people who defend them.

Cosmina Stalidi

National and Kapodistrian University of Athens

cosminastalidi15@gmail.com · Cosmina Stalidi on LinkedIn

This research is carried out within the EXARCH project, funded by the European Union's Horizon Europe programme under the Marie Skłodowska-Curie Actions [grant agreement No. 101236244.

