NEW AI SERVICES IN TELECOM — REALITY OR WISHFUL THINKING?



Benedek Kovács

Expert – Edge Computing, Ericsson
Head of technology and Innovation, Ericsson Hungary

2026

Call For Action: Unique Value-Creation Window

Network Design

High Performance

New network capabilities

New services for AI (agents)

New services

AI for RAN, Core, Telephony
Spectrum efficiency, network automation, translation

AI to create new network services
Integrated Sensing, Smart Location, Fraud prevention

AI domain, AI as a Service
New network services for sovereign AI applications



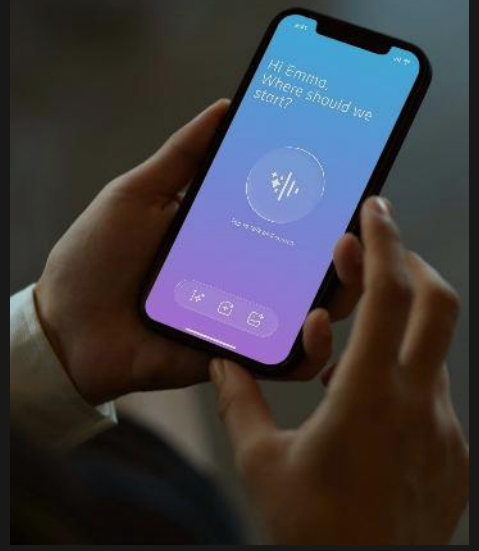
Use Case: Consumers

Emerging use cases, embracing AI

ERICSSON

Timing to scale: 2026/27

Multi-modal & premium AI experience

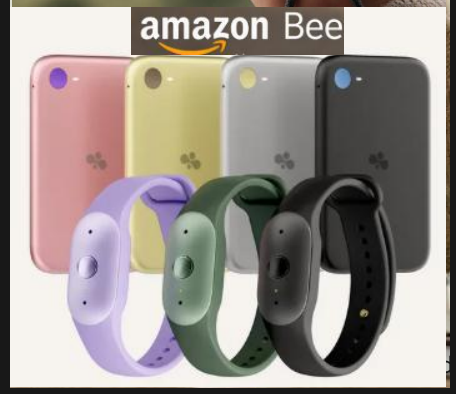


- Voice / conversation
- Live video
- Camera/video input
- Agents

Meta Muse



amazon Bee



- Meta Muse
  (11M downloads in 50 days)
- Charm HW (cellular), Amazon Bee (ship Q4/now)
- Listen/see, always on

2027+

AI glasses & Augmented reality



- Head-mounted displays (AR/AI glasses)
  (Ship >10M 2026, ~30M 2030)
- AR 2027+
- Always on, proactive AI

Vehicles & Drones



- UAV/drones AI assist
  (Ship ~30M 2026,~40M `30)
- Autonomous Vehicles
  (Ship: ~18k 2026, ~600k '30, ~8M 2035)

2030+

Physical AI



- Humanoids
  (Ship ~50k 2026, ~1M `30, ~3.2M `35)

Sources: TechInsight, Gartner, Omdia Robots Hardware Market 2026

Telco offered AI services		User & Description	Market example of operators offering service
In-call services	Instant language translation	Realtime translation for call centers, work meetings Private phone calls live translation	T-Mobile, SK telecom, 中国移动 China Mobile
	Fraught/spam detection	**Using AI to detect voice phishing/scams and alert user**	airtel, SK telecom, LG U+, vodafone, 中国电信 CHINA TELECOM
	AI transcription	AI providing transcriptions of calls	Telefónica, 中国移动 China Mobile
	Call summaries	AI providing summaries of calls, suggesting follow-up tasks etc.	Telefónica, SK telecom, 中国电信 CHINA TELECOM
	Pre-call vetting	**AI automatically answers calls and grasps caller's purpose ahead of user answering**	LG U+
	AI answering	AI as a service e.g., AI health assistant hotline for medical advice AI answers phone allowing user to answer in text when relevant	Telefónica, 中国移动 China Mobile, 中国电信 CHINA TELECOM
	AI video ringback tone	AI replaces the standard ringing screen with tailored video content, e.g., a restaurant call displaying a promotional video	*Unknown Chinese operator*
AI enablement	Token plans / subscriptions	**AI usage sold as traditional mobile data bundles; Sometimes bundled with other connectivity packaging**	Singtel, 中国移动 China Mobile, 中国电信 CHINA TELECOM, China unicom中国联通
	Model-aggregation platforms	Aggregating access to AI models under unified token-management	orange, 中国移动 China Mobile, 中国电信 CHINA TELECOM, China unicom中国联通
AI-eSIM		AI-powered eSIM combining connectivity, identity/security, AI compute platform (incl. models, agents, tools), and packages bytes & tokens	中国移动 China Mobile
Quality compute on demand (priority compute)		Priority compute offering e.g., like AWS Amazon Bedrock "Priority Tier"	*[No CSP example found]*

Link to all reports used

Example from China: tokens, calls, compute

ERICSSON

What they sell today

- AI token plans on the phone bill
- Model hubs: ~300 models in one place
- AI calling: live subtitles, translation
- Faster uplink, lower latency for AI glasses
- AI compute for enterprises

How it is likely built

- Telco billing + identity in front
- Model aggregation and routing
- Own AI data centres, national compute network
- AI on top of IMS calling and 5G slicing

How they earn

- Tokens: ¥9.9 (≈ 500 HUF) = 10 M, up to ¥49.9 (≈ 2,530 HUF) = 80 M
- ¥9.9 ≈ 20 % of a Chinese mobile ARPU (≈ 2,500 HUF),
- Business ¥39.9–299.9 (≈ 2,000–15,200 HUF) incl. uplink + security
- Compute: China Mobile intelligent computing +130 % (H1 2026)

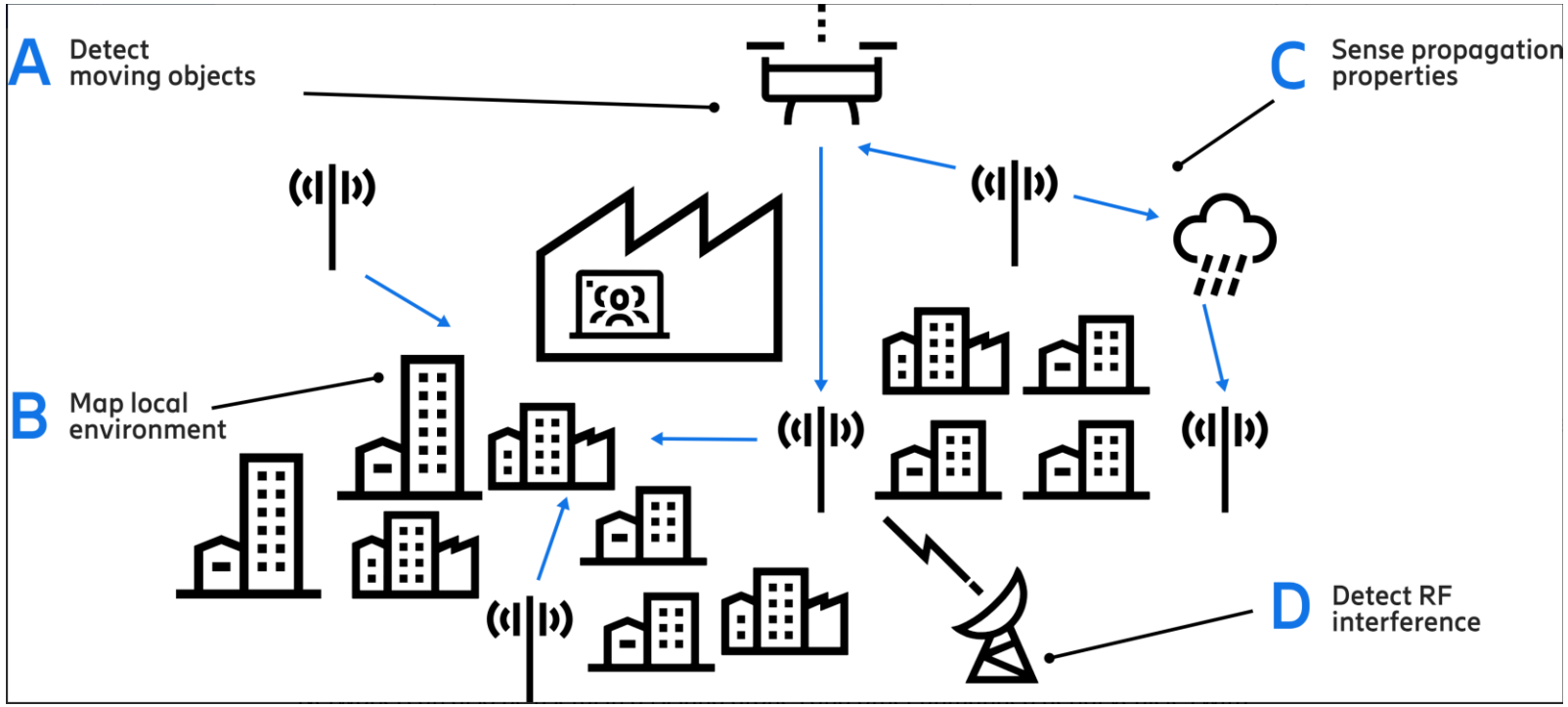
What you see, monetization of AI services:
China Mobile sells AI like it sells data: bundles on the phone bill.

Sources: TelecomTV 05/2026 · Light Reading 06/2026 · China Mobile H1 2026 results · Ericsson 05/2026


ERICSSON
Use Case:
Sensing

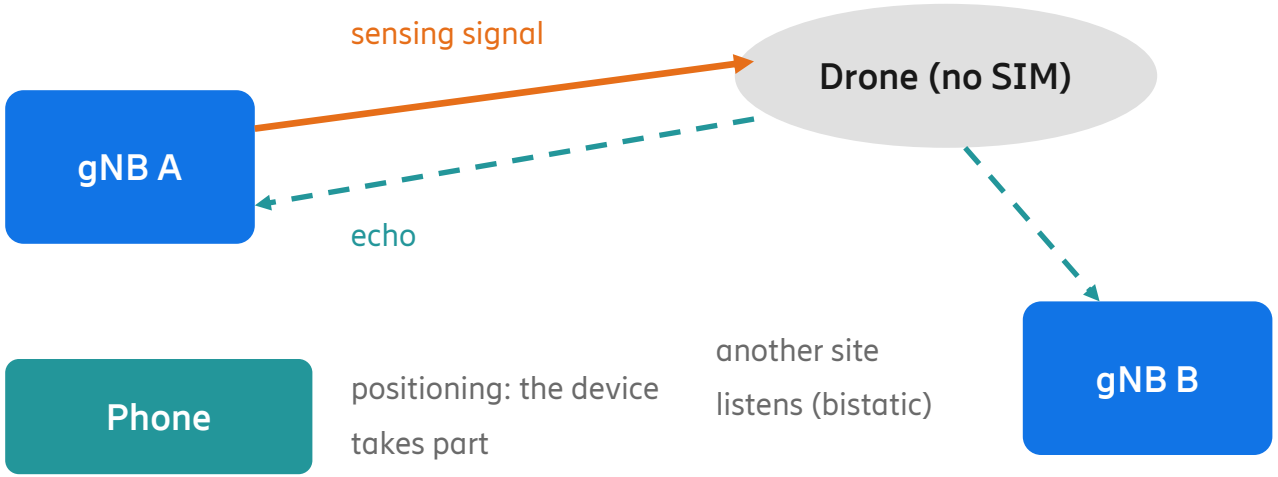
Multiple use cases

- Detecting moving objects
- Mapping local environment
- Sense propagation properties
- Detect RF interference

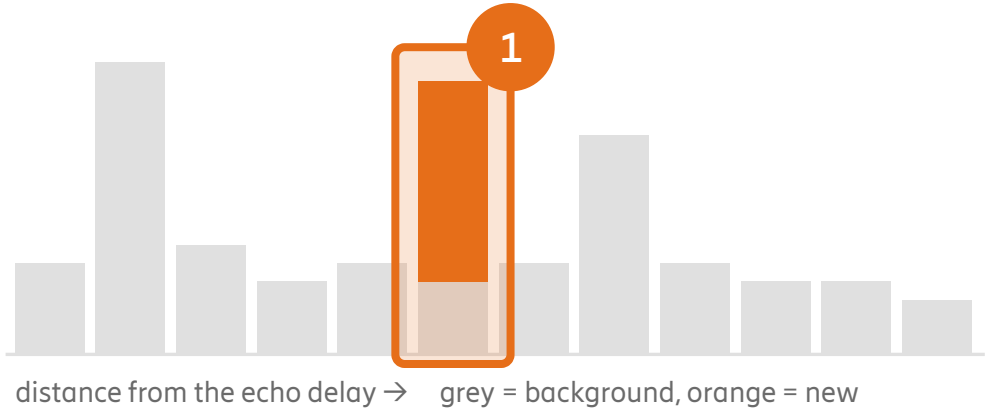


Lesson: what is network sensing?

Match day: a drone appears over the stadium. It has no SIM and sends nothing.



Compare with the learned background

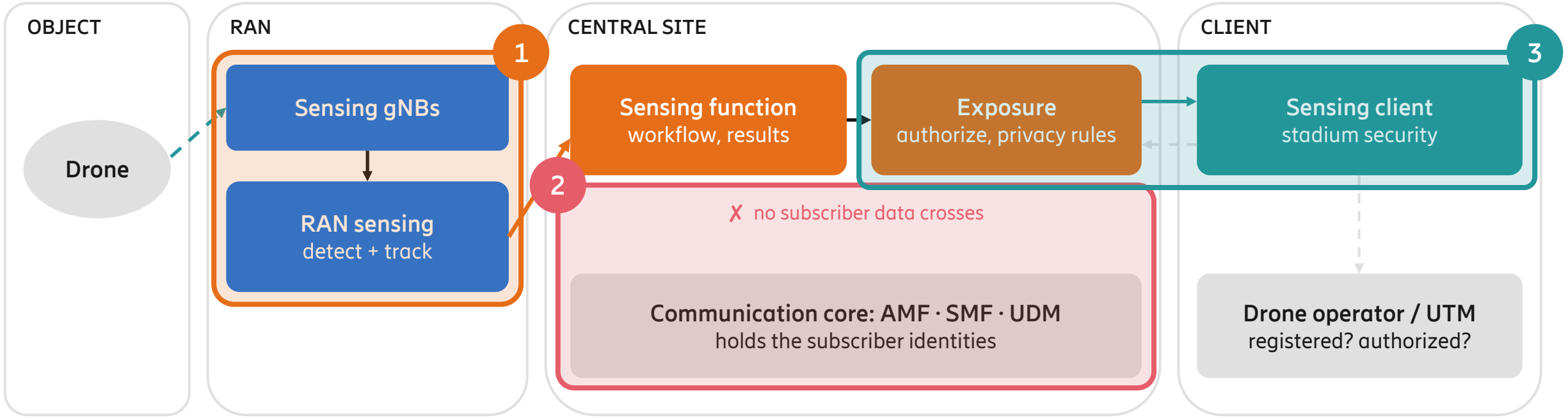


Sensing = detecting, locating and tracking passive objects by analyzing the echoes of radio signals. Unlike positioning, the object takes no part.

Modes: monostatic (same site) · bistatic / multistatic (other sites listen) · UE-assisted

How?
The drone shows up as a new echo against the background: that is a detection.

Sensing flow: request to result



1 The RAN detects and tracks; only objects go to the core.

2 Sensing runs apart from the communication core.

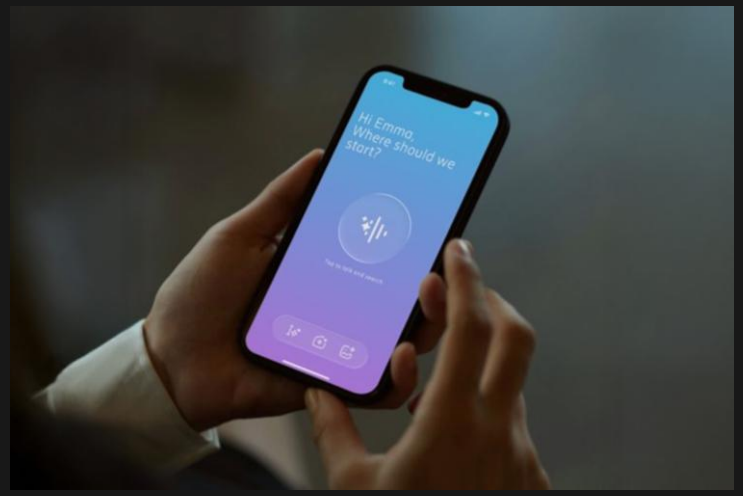
3 Results go only to an authorized client, for an allowed area.

Key question:
What is the value of sensing?


ERICSSON
Use Case:
AI domain

The evolution of mobile network



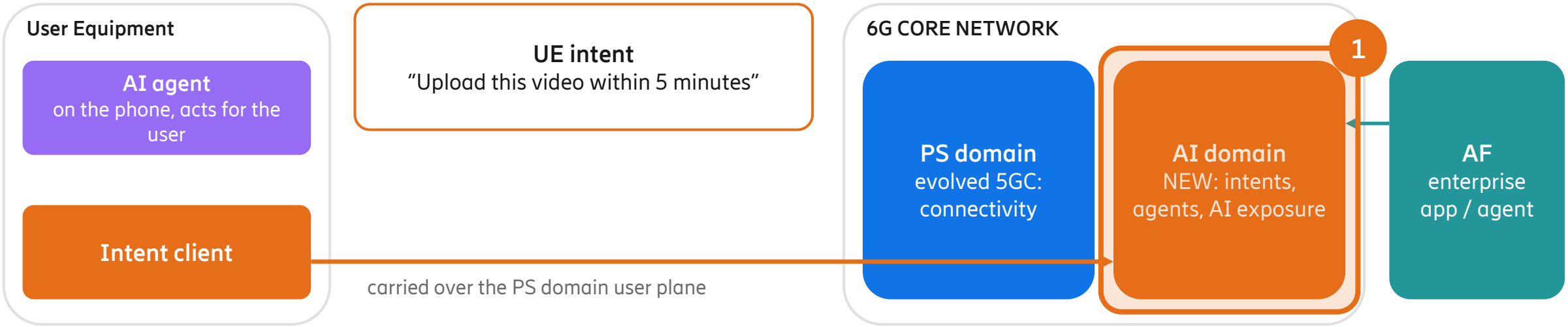




Human to human → Human to machine, Machine to machine → Human to agent, Agent to agent

How to enable agentic communication?

An influencer on a tram wants to upload a video to Instagram; it can wait up to 5 minutes.



The AI domain is a new, optional part of the 6G core next to the evolved 5G core. Devices and applications state the outcome they want (an intent); the network decides how and stays in control.

The AI domain enables AI agents to communicate with the network

Why?
The application asks for an outcome, not for a QoS class or a server address.

How?
By adding a new domain handling AI agent communications.

Ericsson proposal for the 6G core · published: Ericsson blog, Sept 2026

Different types of intents in the 6G network

OAM Intents, UE Intents, AF intents

OAM Intents:

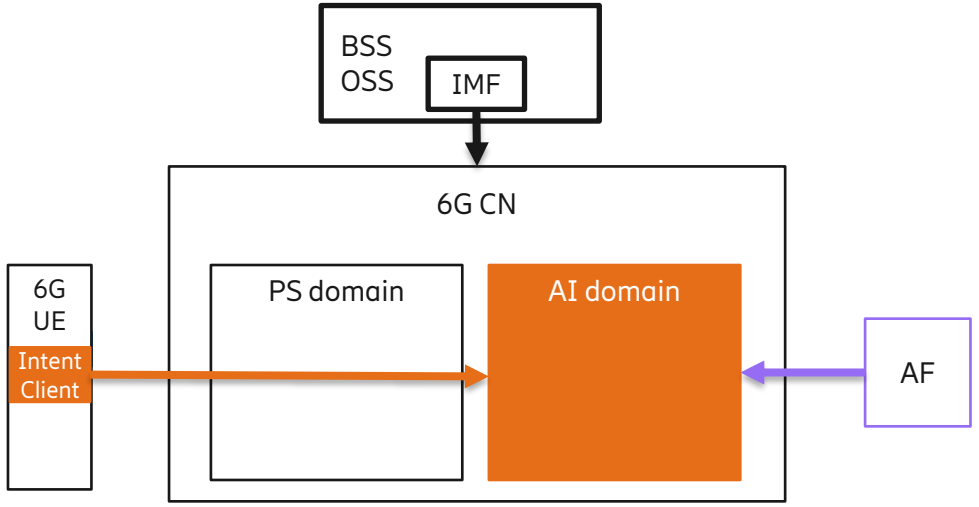
Intents from the OSS/BSS **that change the general network behavior** e.g.,

- Signaling throttling
- IRAT mobility frequency control
- Anomaly detection and logging
- Adding a new connectivity service

UE Intents:

Intents from a UE that aim to **change the specific UE situation** e.g.,

- Connectivity Assistance
- Different QoS than currently used
- Try a new service
- Extend device run time



AF Intents:

Intents from an AF that aims to **change the situation of a single UE or a group of UEs**, e.g. Enterprise based

- Network Assistance
- Different QoS
- Subscription update
- New UE timer values

So what is an intent?

ERICSSON

The rule

- An intent states an outcome; the network decides how.

Monetization

- An intent can be monetized on the outcome

A Keep my video call smooth for 20 minutes	UE intent
B Extend my battery until 18:00	UE intent
C All 50 robots in hall B: ≤ 50 ms tonight	AF intent
D Open a new PDU session with 5Qi 9	✗ not an intent configuration

1

Takeaway
A and B come from the device, C from the robot maker's app.

D says how, not what: that is configuration / call setup

What is the value of a user intent?

Ericsson blog: the 6G AI domain

ericsson.com · blog, Sept 2026

Application intents
Evolution of exposure functionalities

Where are we?
Different proposals in 3GPP to enable network agents.

Exercise constructed by the speaker · intent examples follow Ericsson's published blog


ERICSSON
Use Case:
Physical AI

Networks for AI



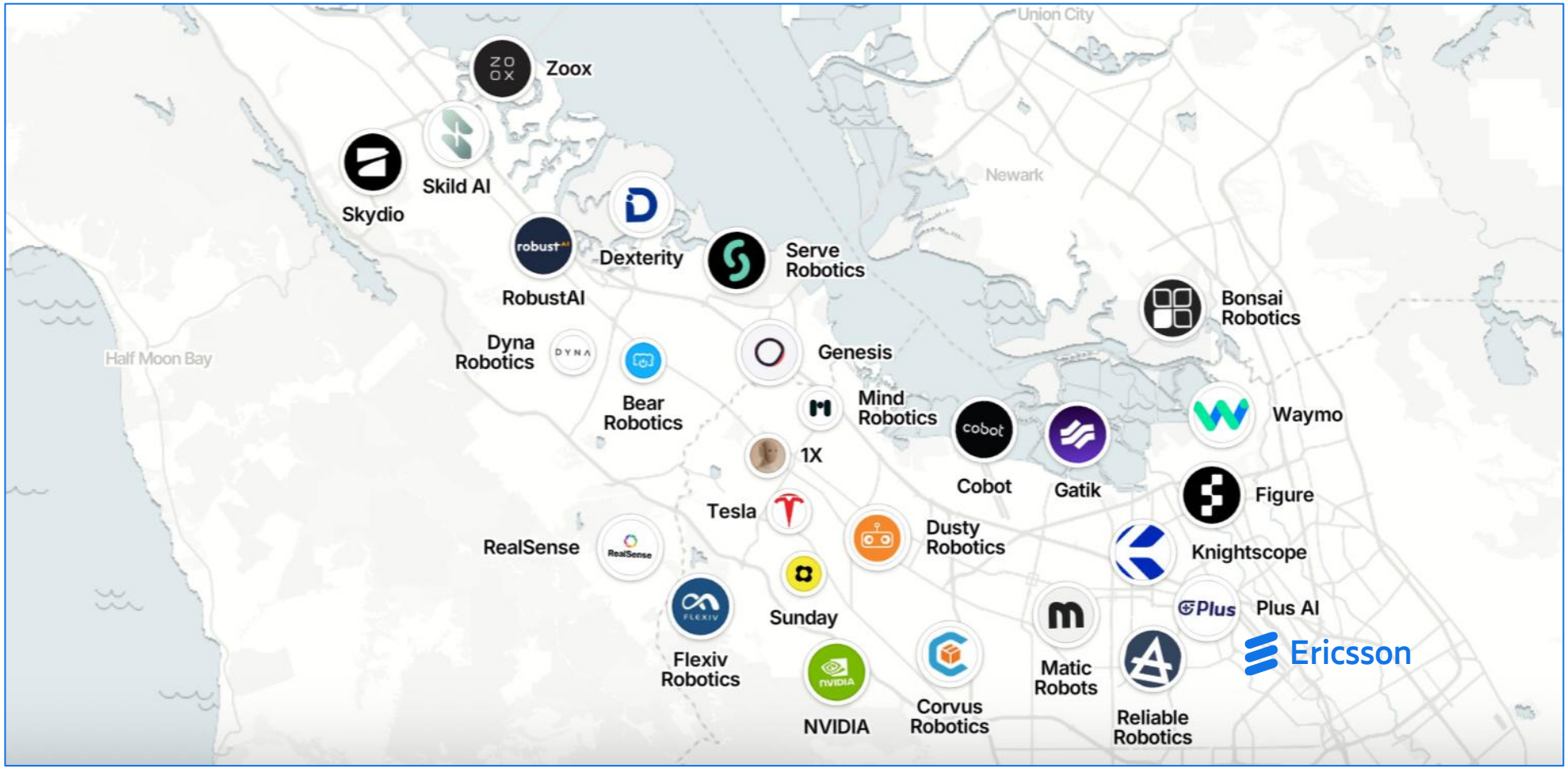
- AI needs to be connected because...



Serve Robotics — Conversational Robot
https://www.instagram.com/reel/DW4bOuPx6U-/

Ecosystem: Exploding Silicon Valley Physical AI

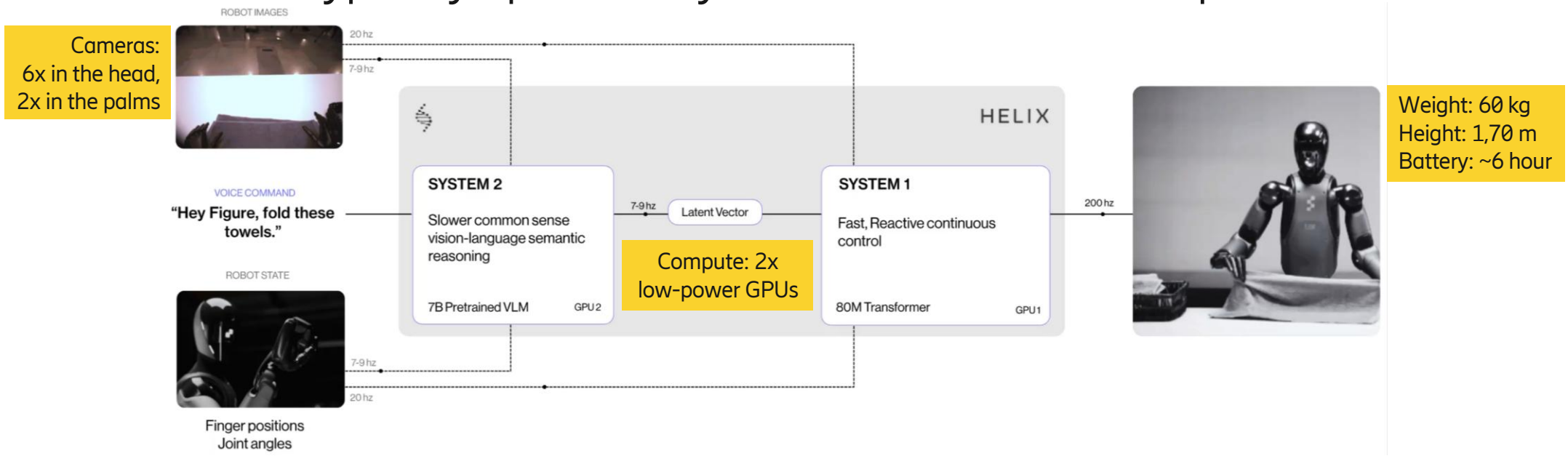
Market cap of our next-door neighbors is $6.78 trillion (> Germany GDP!)



© Lukas Ziegler

Example: Figure AI's Droid Using Helix

Humanoid droid is typically operated by three robotic control loops



	Description	Frequency	Model Size	Offloading?
System 2	Semantic reasoning and producing goals	10Hz (100ms)	7B	Desirable
System 1	Joint-level control of the entire robot	200Hz (5ms)	80M	Potentially
System 0	Execution of control signals and ensuring stability	1 kHz (1ms)	10M	No

Rational for 5G connectivity: (1) Telemetry; (2) Model offloading; (3) Distributed intelligence.

© Figure 03 | AI Wiki

New Opportunity: 10x Value With Offload

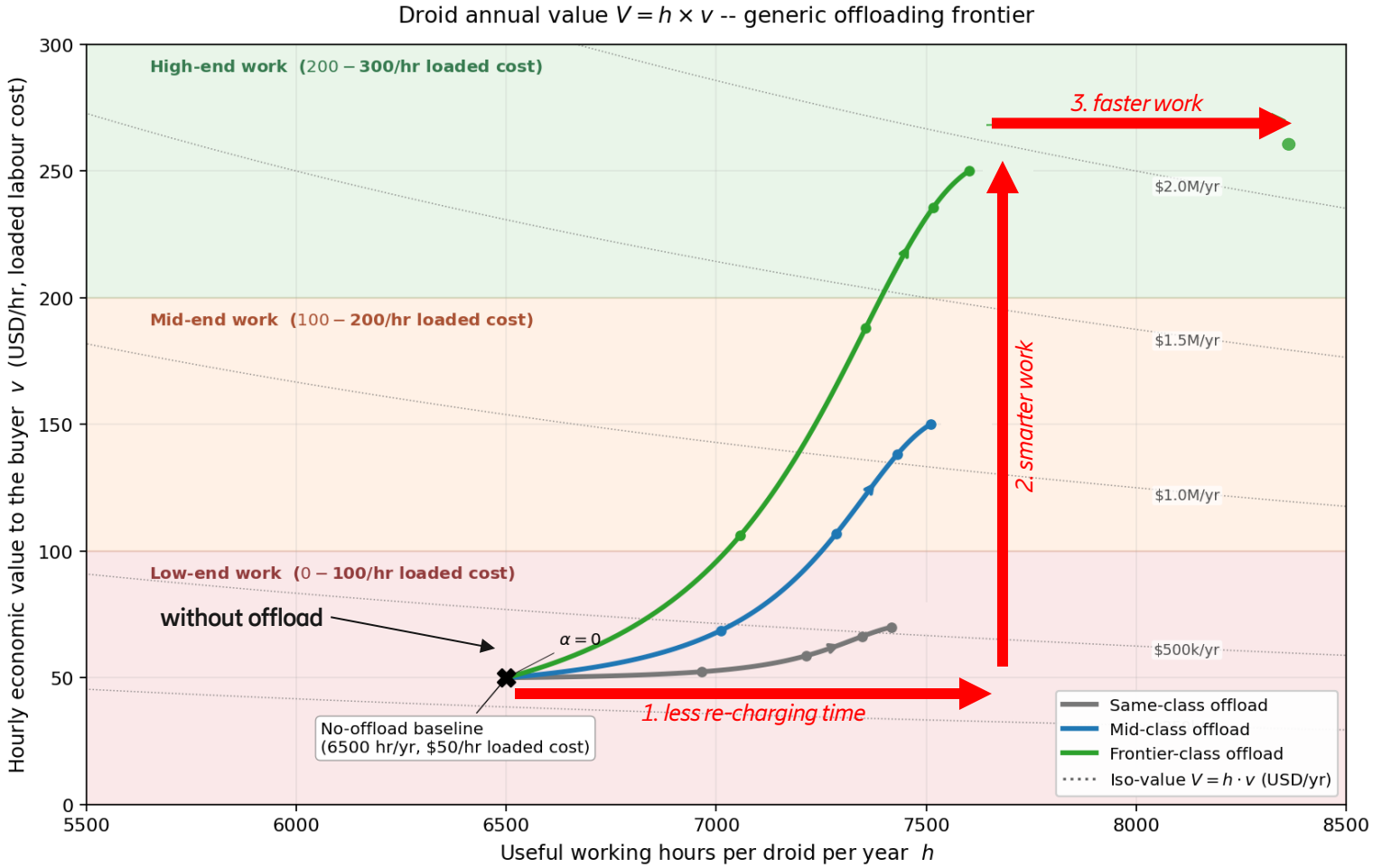
Opportunity for our telco industry to shift to outcome-based pricing

Offloading allows for:

1. Saving operational energy (thus less charging and more work time);
2. larger AI models (thus increases labor value); and
3. increased token rate (thus higher speed of work).

Frontier-class offload can lift a single droid past the $2M/year iso-value contour — this is 10x value add to their customers!

For telcos: This delta above baseline is monetizable through outcome-based pricing, rather than per-SIM pricing.



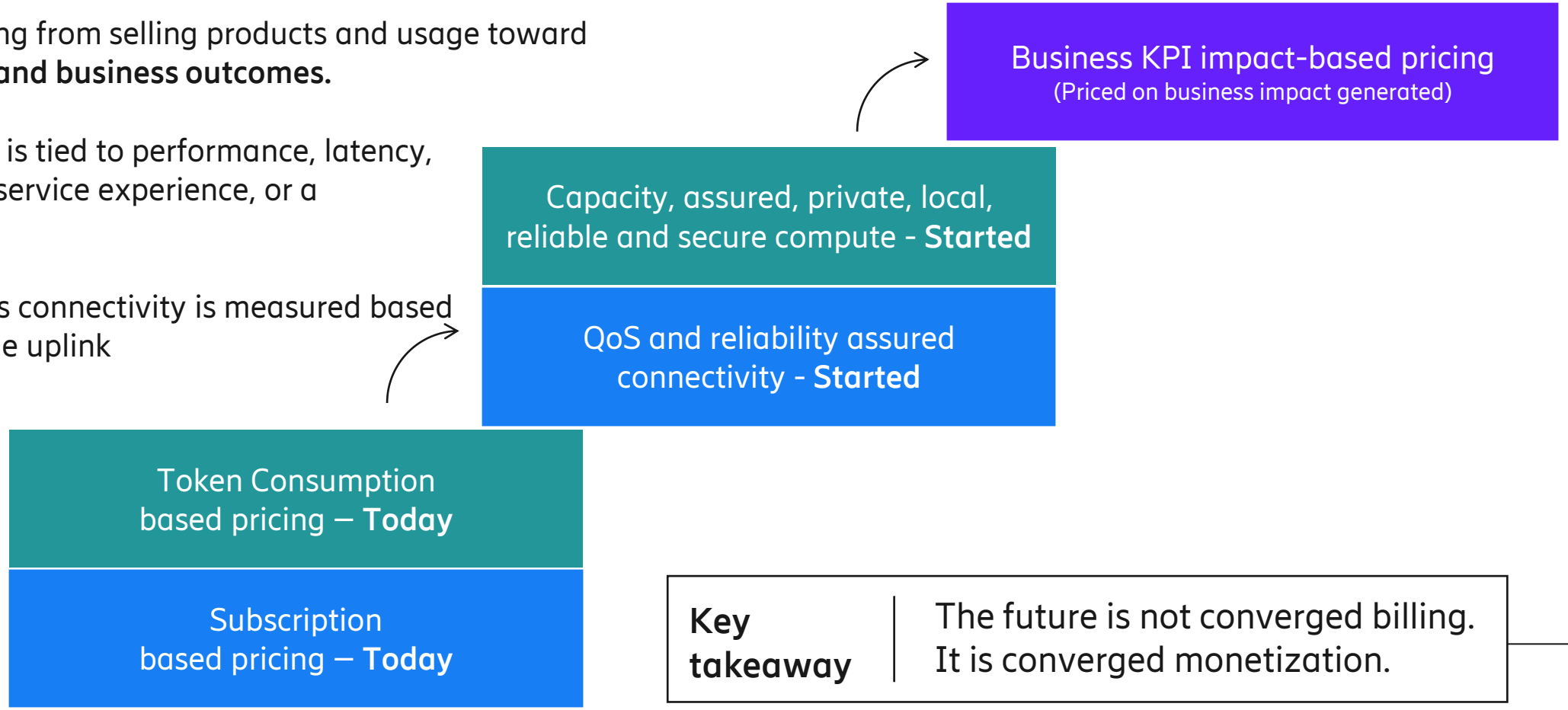
Monetization: Outcome-Based Monetization

Unique opportunity for operators to become part of the value chain

The industry is moving from selling products and usage toward **selling experiences and business outcomes.**

Commercial promise is tied to performance, latency, quality, availability, service experience, or a business outcome.

A surveillance robot's connectivity is measured based on the always reliable uplink



Key takeaway	The future is not converged billing. It is converged monetization.

How? - Kinetic tokens

Task given to a robot waiter: "bring the cup to table 4"



A kinetic token is an AI output that results in a direct action or goal of a device in the physical world: "rotate the arm 75° south" (direct) or "reach position 27.1751, 78.0421" (indirect).

Where kinetic tokens come from

- LLM / SLM output
- Agents acting on model output
- AI apps steering devices
- 6G RAN sensing (measured)

What you see

Kinetic tokens, as tokens for complex actions are the value, not the text

Concept: Ericsson work on Physical-AI monetization · example is illustrative

Example: Price a robot waiter

One month, ILLUSTRATIVE

- A restaurant chain rents 20 robot waiters from a robot maker.
- The robot maker runs its AI app on the operator's edge.
- Each robot serves 300 dishes a day, about 5 kinetic tokens per dish (pick, carry, place).
- 20 × 300 × 5 × 30 days = 900,000 kinetic tokens.

Item	Rule	€ / month
Network slice	€5,000 + €50 per robot	6,000
Edge GPU	flat	1,000
AI service	10 M info tokens included	500
Kinetic tokens	€2 per 100 actions	18,000
SLA credit	3 h breach → those tokens free	−150
Total	**≈ €0.14 per dish served**	**25,350**

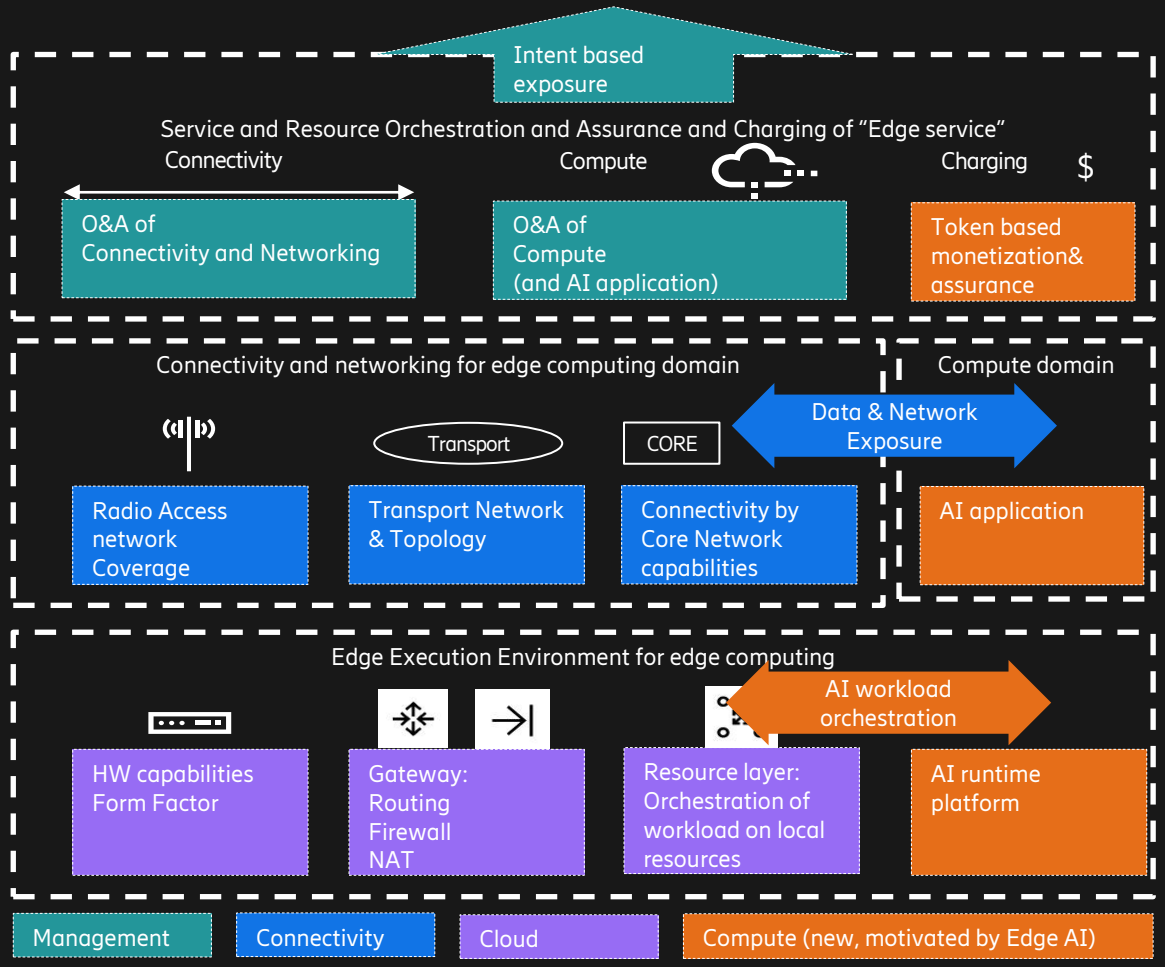
4

Note:
About 70% of the bill is counted in actions, not in GB.

All numbers are illustrative · structure adapted from Ericsson token-monetization work

Ericsson Edge AI target architecture *overview of activities*

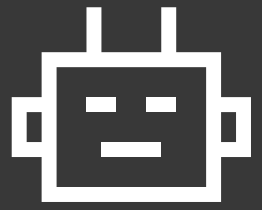
- Multi-domain e2e service orchestration, management and assurance, supporting integration with 3rd party compute platforms, multi-vendor environments, with intent based exposure
- Full connectivity solution supporting data sovereignty, private network connections, device onboarding, quality of service, network location and additional network services for application-network interaction via data and network exposure
- Platform and resource orchestration layer supporting heterogenous HW acceleration depending on the AI workload demand (optimized choices for the AI applications)
- Network and data exposure



Call To Action: Address Opportunities Jointly

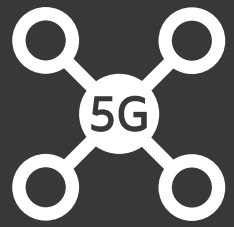
High-performing networks enable AI-centric use cases

Emerging AI-centric Use-Cases and Devices



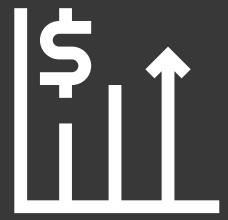
- Smart/AR Glasses and Real-Time Cloud AI
- Real-Time Gen-AI-agents
- Physical AI use-cases

High-Performing 5G SA Networks



- High-performing, secure networks and reliable uplink ; towards 6G
- Differentiated configuration per service
- Compute and AI capabilities

Key Monetization Enablers



- Slicing framework
- API Exposure (e.g. QoD, location and analytics)
- Token- and outcome-based monetization



ERICSSON