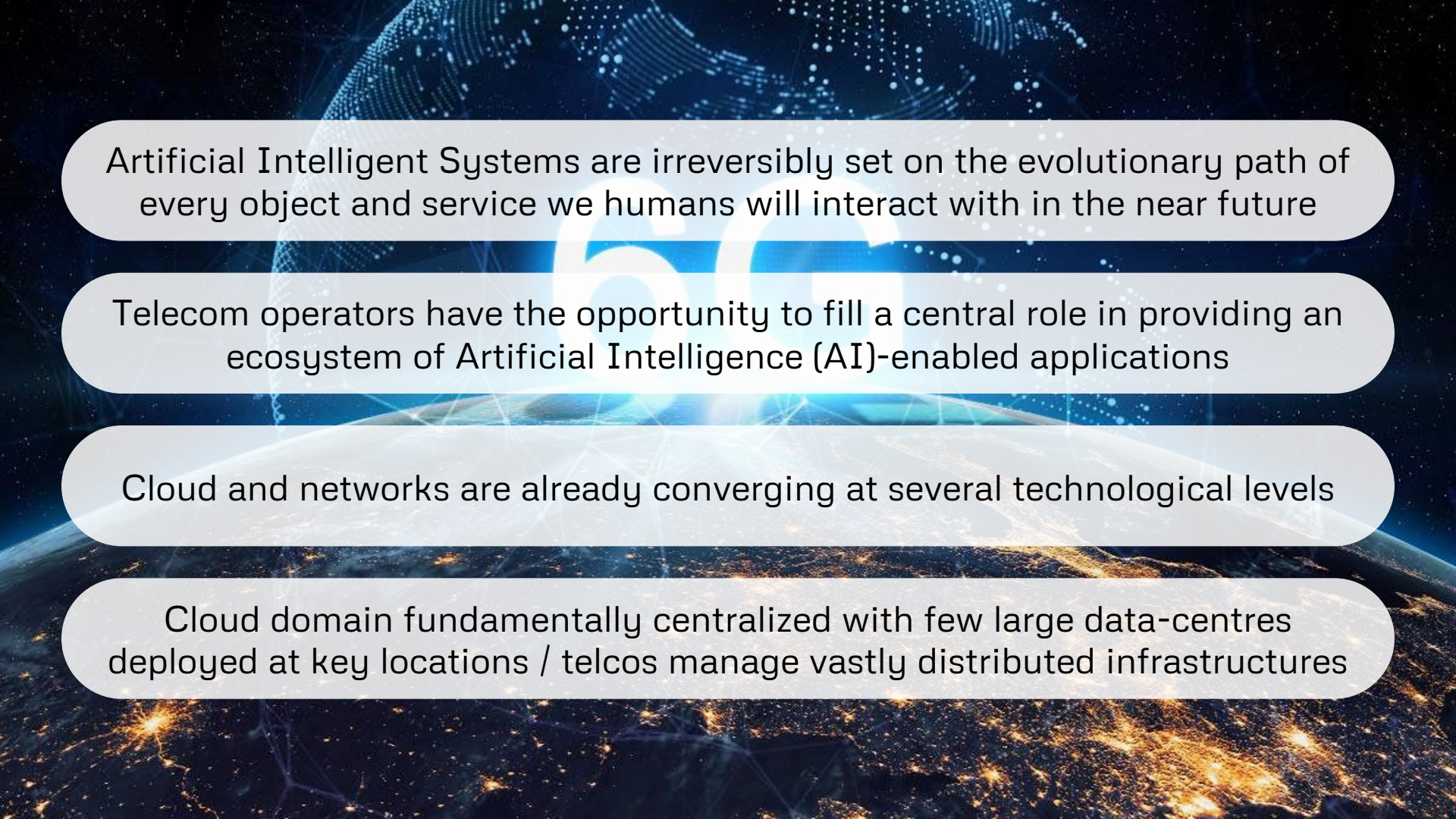


AI@EDGE: A Secure and Reusable Artificial Intelligence Platform for Edge Computing

Roberto Riggio
Associate Professor
Polytechnic University of Marche Italy

1st Open Annual Workshop on Future ICT

- Motivation
- AI@EDGE Objectives
- AI@EDGE Breakthroughs
- AI@EDGE Concept & Architecture
- Use Cases
- AI@EDGE Key Figures
- Conclusions

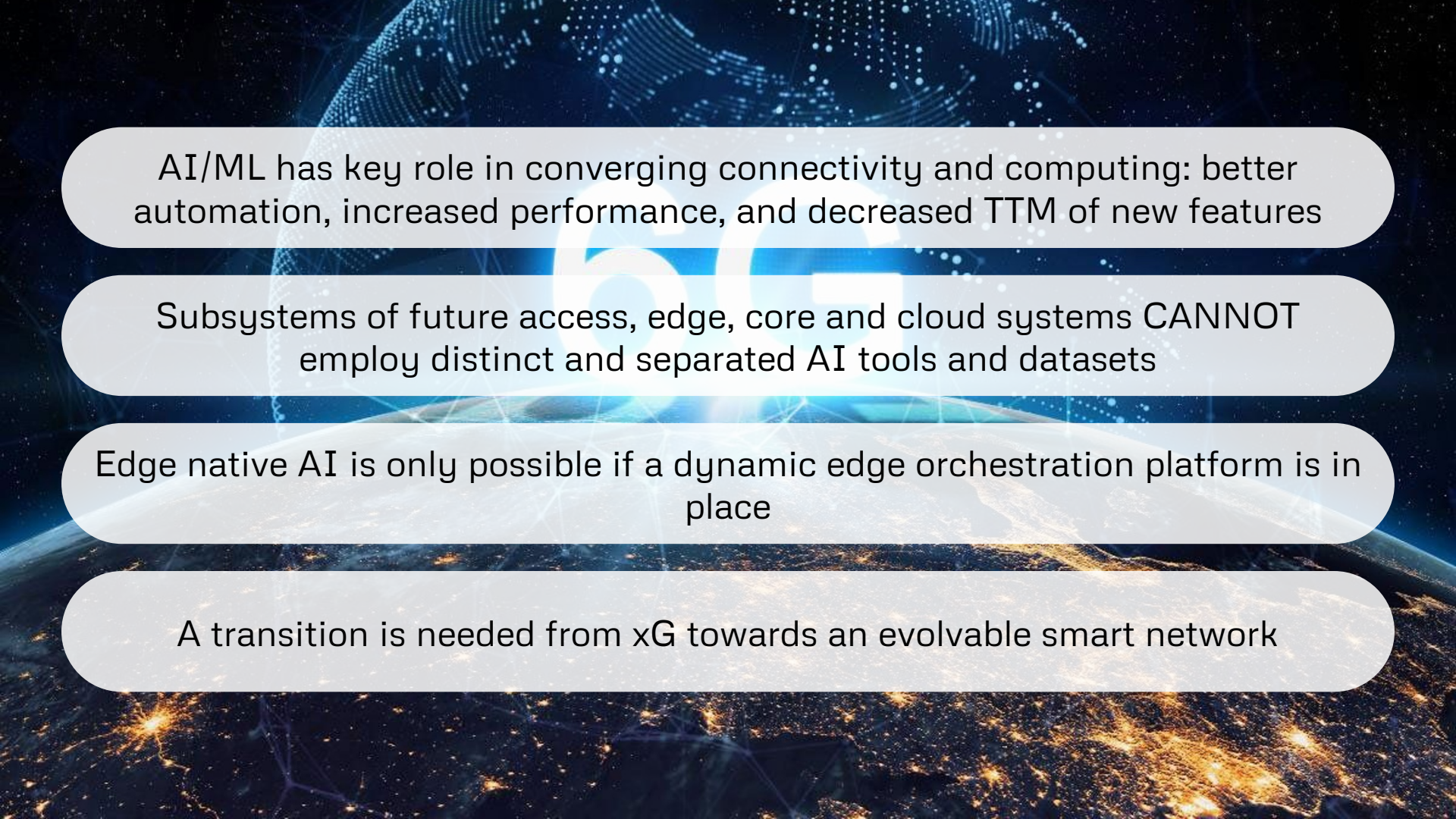


Artificial Intelligent Systems are irreversibly set on the evolutionary path of every object and service we humans will interact with in the near future

Telecom operators have the opportunity to fill a central role in providing an ecosystem of Artificial Intelligence (AI)-enabled applications

Cloud and networks are already converging at several technological levels

Cloud domain fundamentally centralized with few large data-centres deployed at key locations / telcos manage vastly distributed infrastructures

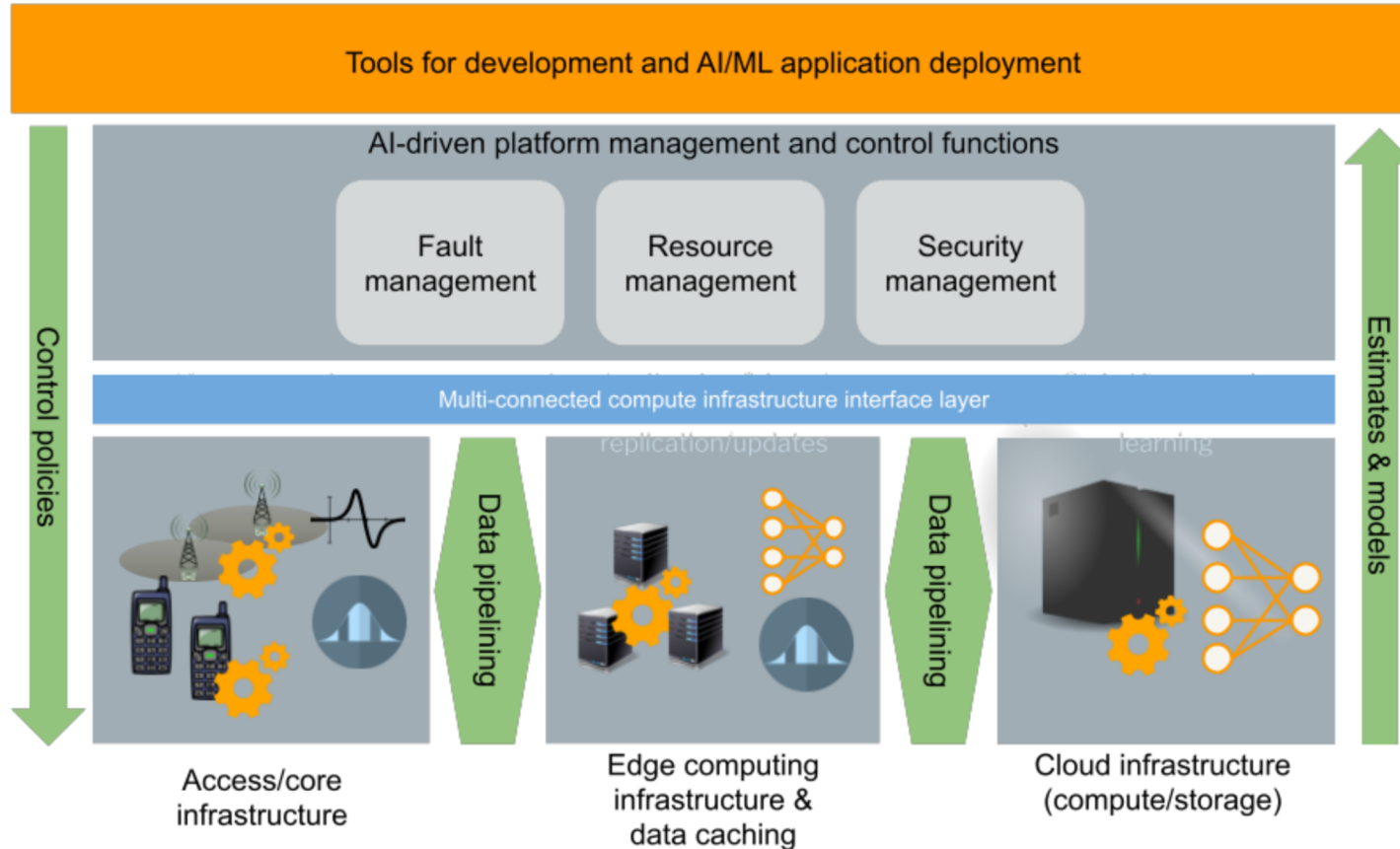


AI/ML has key role in converging connectivity and computing: better automation, increased performance, and decreased TTM of new features

Subsystems of future access, edge, core and cloud systems CANNOT employ distinct and separated AI tools and datasets

Edge native AI is only possible if a dynamic edge orchestration platform is in place

A transition is needed from xG towards an evolvable smart network



AI@EDGE Breakthroughs

AI/ML for closed loop automation (AI4Networks - WP3)

Privacy preserving, machine learning for multi-stakeholder environments (AI4Networks - WP3)

Distributed and decentralized connect-compute platform (Networks4AI - WP4)

Provisioning of AI-enabled applications (Networks4AI - WP4)

Hardware-accelerated serverless platform for AI/ML (Networks4AI - WP4)

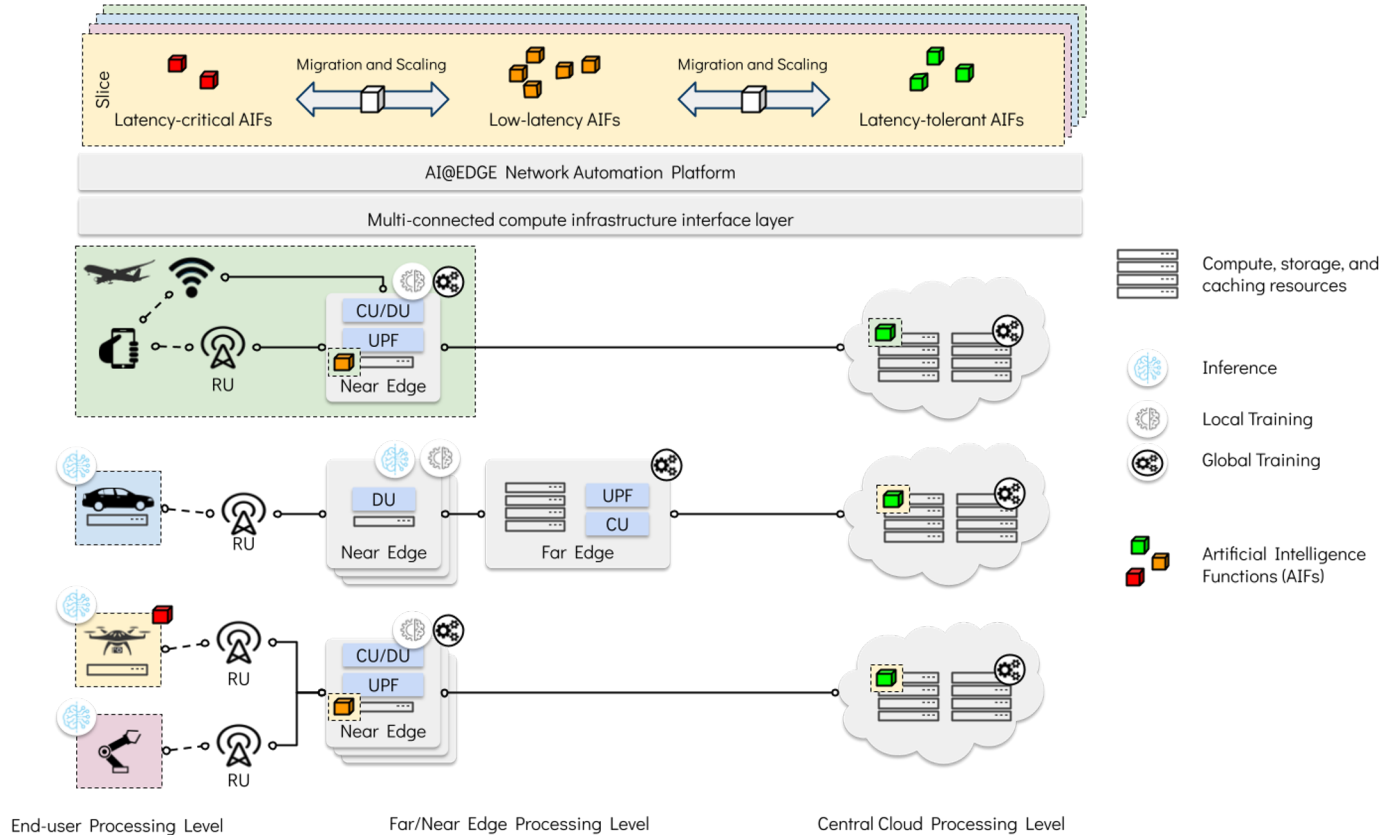
Cross-layer, multi-connectivity and disaggregated radio access (Networks4AI - WP4)

AI4Network

Network4AI

Enter Artificial Intelligence Functions

AI/ML for closed loop automation (AI4Networks - WP3)





Intermedia Architecture





Virtual validation of
vehicle cooperative
perception



Secure and resilient
orchestration of
(I)IoT networks



Edge assisted
monitoring of linear
infrastructures



Content curation for
in-flight media
services



Breakthrough	UC1	UC2	UC3	UC4
Distributed and decentralized serverless connect-compute platform	X	X	X	X
AI-enabled application provisioning	X	X	X	X
Cross-layer, multi-connectivity radio access				X
Hardware accelerated serverless platform for AI/ML			X	X
Network and service automation platform	X	X	X	
Secure, reusable, and resilient machine learning for multi-stakeholder environments		X		

- ICT-52 Smart Connectivity beyond 5G
- 19 Partners from 8 EU Countries
- Total project budget: 8 ME
- Requested Funding: 8 ME
- Project Duration: 36 Months



Conclusions



Reusable, secure, and Privacy aware AI/ML Layer

Automation platform for zero touch management of network and services

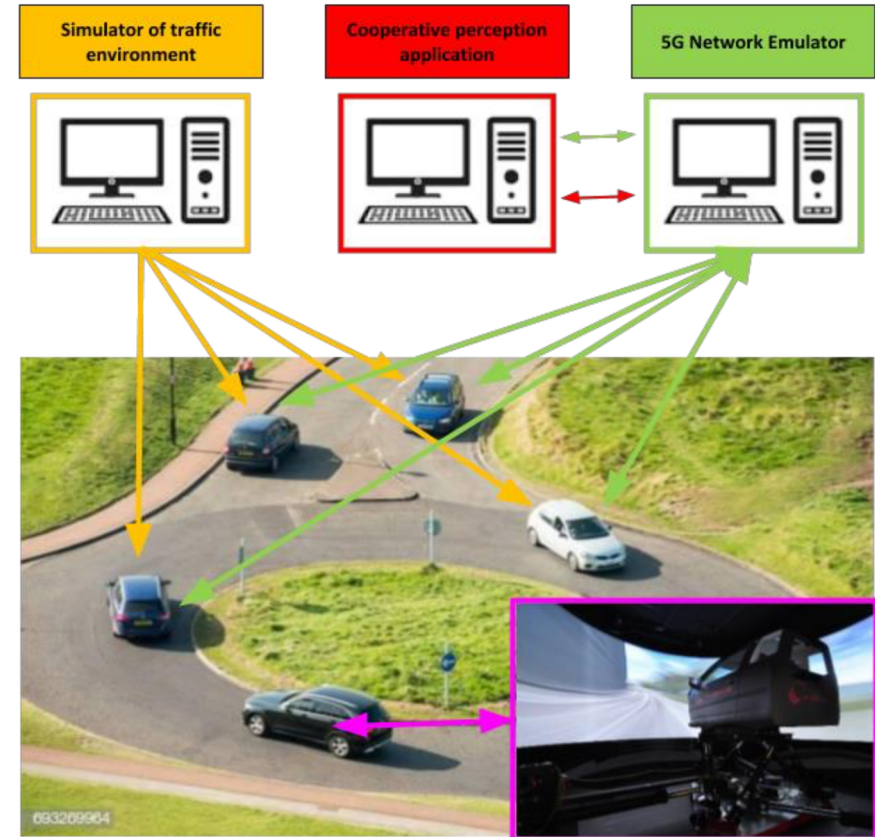
AI-enabled application as composition of AIFs

High Impact, Industry relevant Use Cases

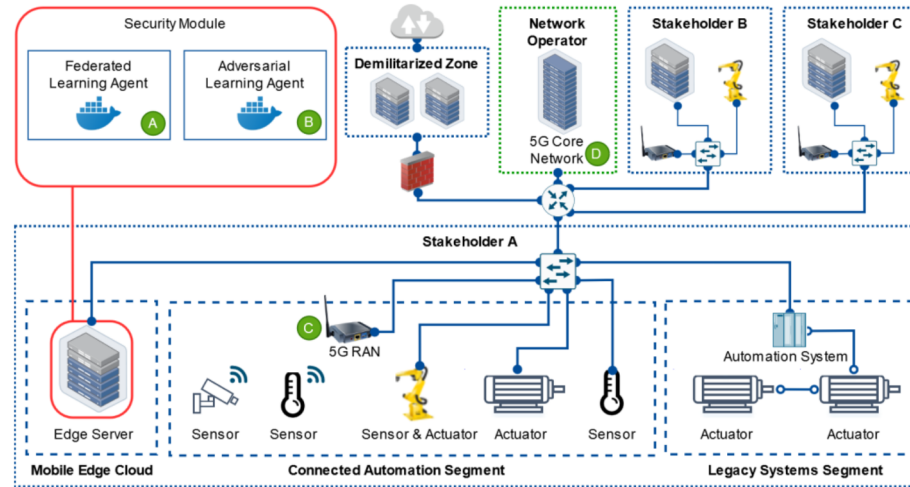
Thank You!

UC1: Virtual validation of vehicle cooperative perception

KPI	Goal
Vehicle density	1,200 vehicles/km ² at 20km/h
Sensors fusion latency	Total from sensor detection to Vehicle including sensor fusion on edge < 10 ms
Communication reliability	from 99.9% to 99.9999%
Range	Up to 500m



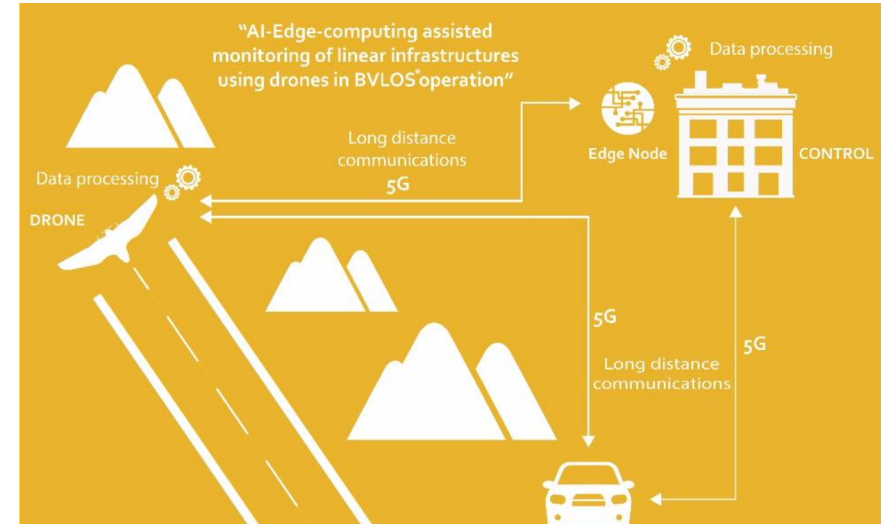
UC2: Secure and resilient orchestration of large (I)IoT networks

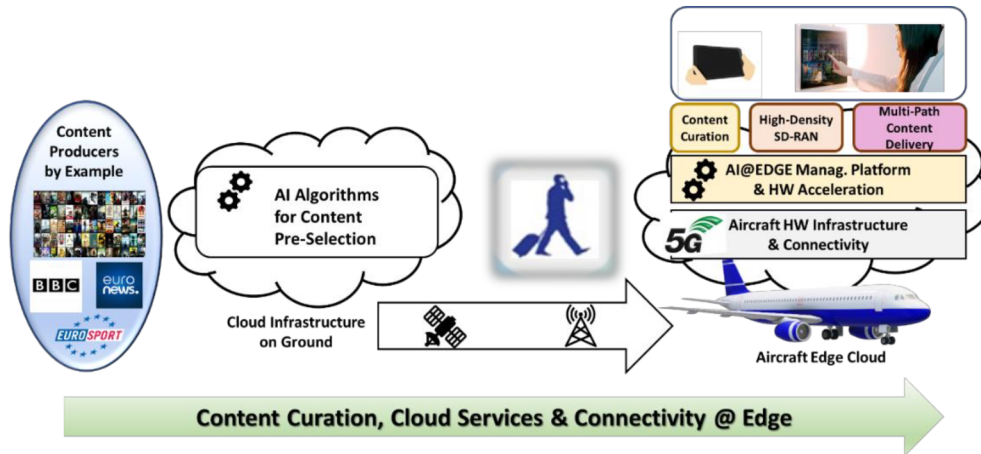


KPI	Goal
Zero-day detection	Capability to detect unknown attacks based on detection models trained on edge nodes
End-to-end latency	Minimal to no effect on the monitored communication
Fast detection	Local within 1s, global within 1m
Robust AIs	< 5% detection rate decrease against adv. samples

UC3: Edge AI assisted monitoring of linear infrastructures using drones in BVLOS operation

KPI	Goal
Drone operation	Drone capable of integrating developed functionalities and running operations according to regulations for the pilot tests
Range	Prepared to reach at least 20 km (according to the state of 5G technology and deployment at the trials)
Integration of AI-assisted drone framework on the 5G network	Detecting incident through AI analysis processed on-board and at edge-node to generate response action from centralized control station





KPI	Goal
Data rate/client	At least 25 Mbit/s per passenger for 4K video streaming
Aggregate in-cabin throughput density	At least 20 Mbit/s/sqm.
Service deployment time	Less than 180 seconds
Curated content delivery time (production and distribution)	Less than 180 seconds
Service recovery time	Less than 30 seconds
Number of served passengers	At least 12 for demonstration



Inference

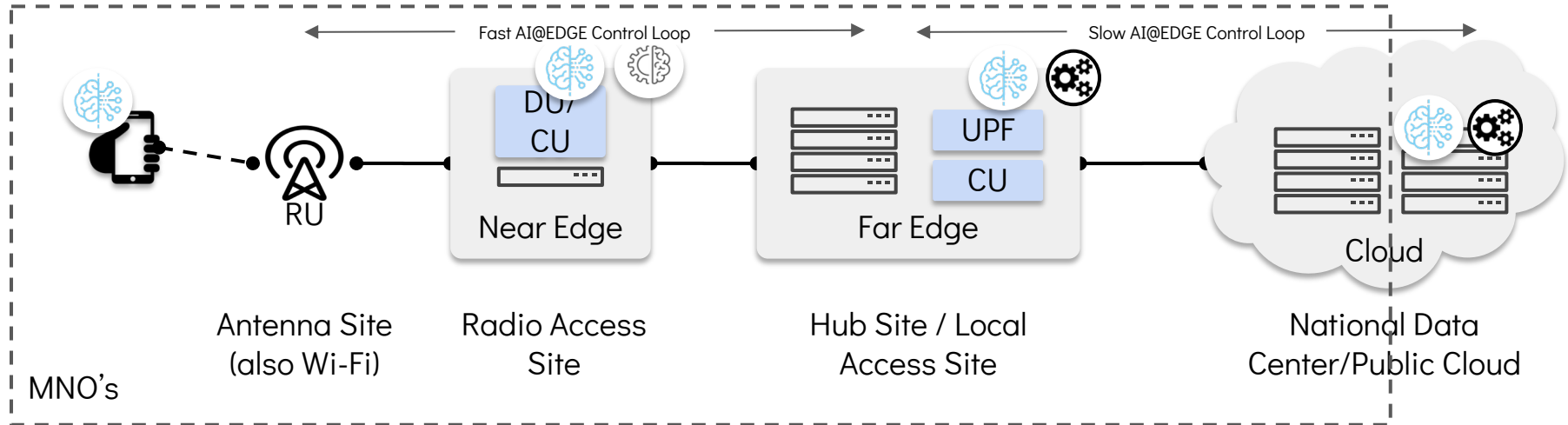


Local Training



Global Training

- Current AI-applications do most of the heavy lifting in public clouds leaving the user terminals in charge of UI and visualization (many app will still work like this)
- Latency/privacy/bitrate bound applications can benefit from a decomposition of the ML tasks across user, edge, and cloud.





Artificial Intelligence
Functions (AIFs)



Inference



Local Training



Global Training

- AI@EDGE introduces the concept of Artificial Intelligence Functions (AIF) to refer to the AI-enabled end-to-end application sub components that can be deployed
- AIFs will be realized using a combination of classical (VM) and cloud-native (Containers, FaaS) virtualization techniques. Hardware acceleration is considered.

