Hexa-X

AI-driven communication & computation co-design

NOF, EAB, ATO, BCO, CEA, EBY, EHU, SZT, INT, NXW, NOG, ORA, OUL, UPI, WIN

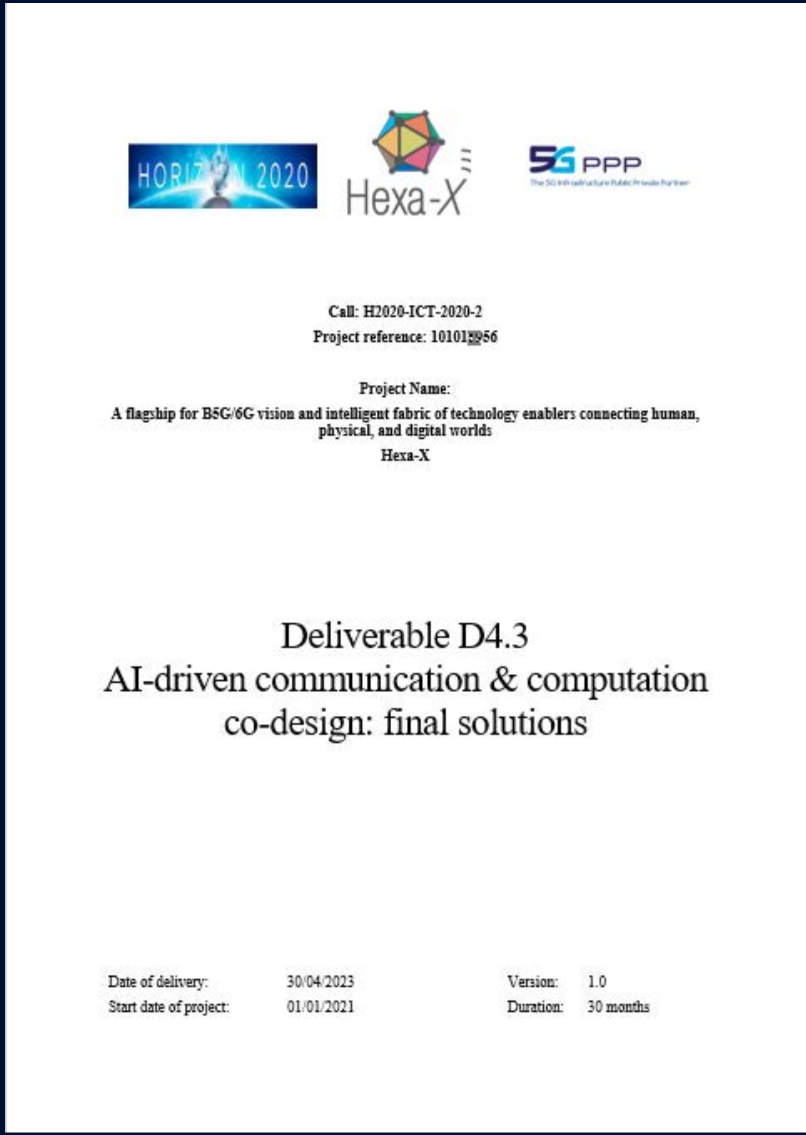
Nandana Rajatheva, OULU

hexa-x.eu

Mission and Scope

- Hexa-X WP4 (*AI-driven communication and computation co-design*) develops concepts for **AI-based air-interface design** and aims to deliver a **secure and sustainable 6G distributed learning platform** able to optimally support and address distributed edge workloads and learning/ inferencing mechanisms

- Technical areas of focus are:
  - **Network performance enhancement using AI/ML in 6G**
  - **6G network as an efficient AI platform**
  - **AI/ML as an enabler for 6G network sustainability**
  - **Privacy, security & trust in AI-enabled 6G**
  - **Demonstration activities - Federated eXplainable AI (FED-XAI) demo**


Call: H2020-ICT-2020-2
Project reference: 101015956

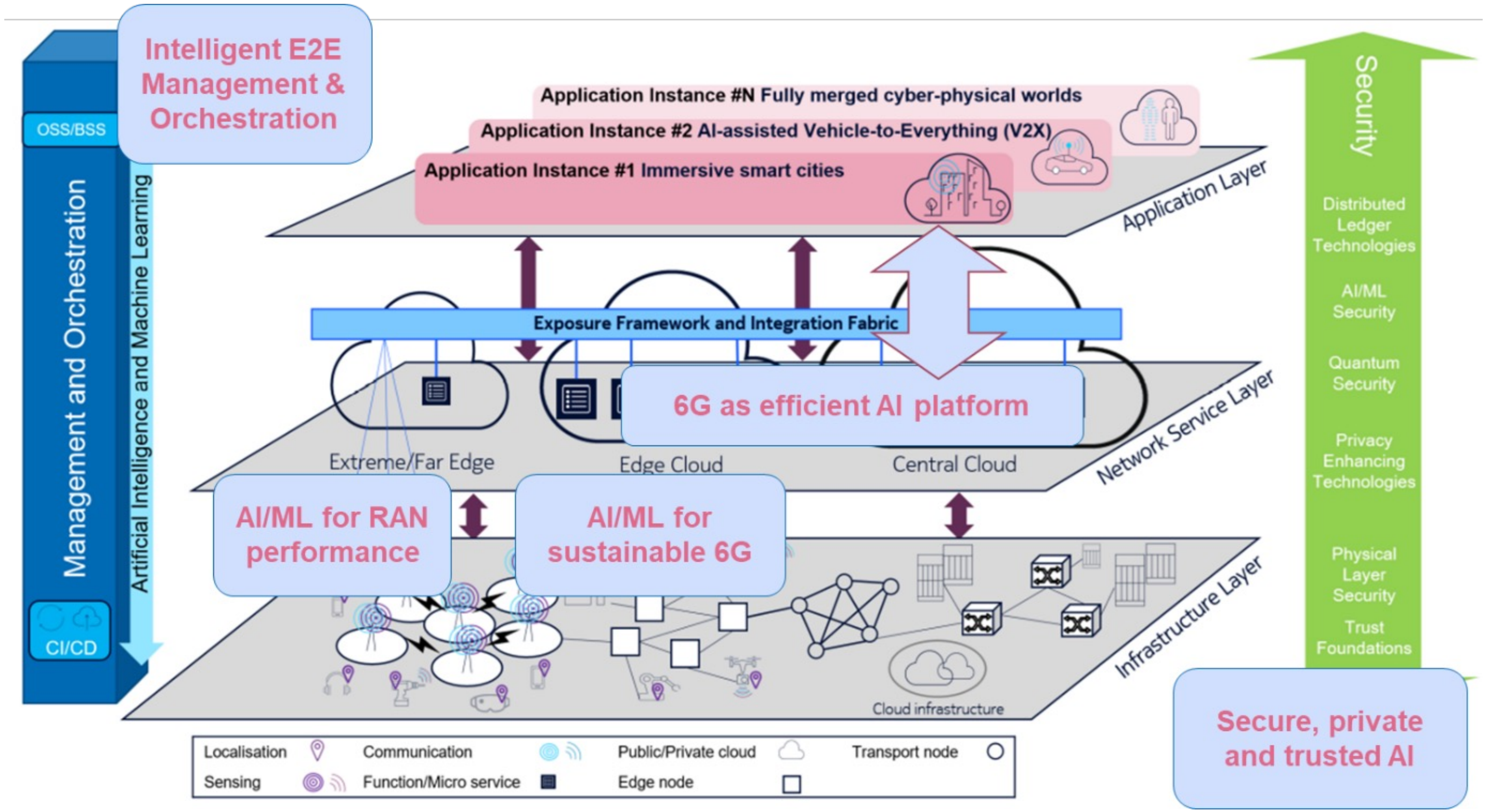
Project Name:
A flagship for B5G/6G vision and intelligent fabric of technology enablers connecting human, physical, and digital worlds
Hexa-X

Deliverable D4.3
AI-driven communication & computation co-design: final solutions

Date of delivery:	30/04/2023	Version:	1.0
Start date of project:	01/01/2021	Duration:	30 months

AI-driven communication and compute solutions

- Future 6G network functions and use cases will be intertwined with various forms of learning and intelligence in many aspects including air interface design, data management, optimality of compute and processing functions, network automation & service availability.
- The right-hand illustration shows how Hexa-X WP4 addresses all the above domains with technical enablers, algorithms, joint solution proposals for communication and computation to help fulfilling Hexa-X Connecting Intelligence research challenges.



Contribution of WP4 technical enablers in network architectural blocks

Hexa-X WP4 quantifiable targets

Quantifiable target #	Title	WP4 task of relevance (*)
T1	Increased AI algorithm robustness to system parameter volatility, lower complexity and significant Bit Error Rate (BER)/ BLock-Error Rate (BLER) gain, as compared to classical approaches	T4.2
T2	Increased AI algorithm robustness to system parameter volatility, lower complexity and efficient resource utilisation and rate gain as compared to classical approaches	T4.2
T3	Resilient communication and compute network services for distributed AI applications in large scales	T4.3
T4	The accuracy of an XAI model within (<10%) of “black box” solutions	T4.3
T5	Energy reduction of a factor of (>10) at the infrastructure level and a factor of (>100) at the user devices’ side, as a result of (network & application) workload offloading and learning/ inferencing task delegations	T4.3
T6	Increased trustworthiness of AI through privacy and security enhancing technologies and AI network intrusion detection capability	T4.3

(*) Task 4.2: AI-driven air interface design
Task 4.3: Methods and algorithms for sustainable and secure distributed AI

Network performance enhancement using AI/ML in 6G

Network performance enhancement using AI/ML in 6G

Hexa-X

- **Main emphasis:** how can AI/ML-based solutions enhance the network performance in a quantifiable way?
- The first part focuses on **radio access network performance improvements over classical design methods**
  - Communication reliability improvements
  - Bit-rate and spectral efficiency improvements
  - Designs accounting for nonlinear distortion
- In the latter part, the focus shifts to **improvements in E2E network operation & management**
  - AI/ML-based predictive orchestration
  - Distributed AI for automated UPF scaling in low-latency network slices

- **KPIs:**
  - Bit Error Rate (BER)/ Block Error Rate (BLER)
  - Channel estimation error
  - Complexity
  - Bit rate or spectral efficiency
  - Flexibility
  - Mobility support
  - Latency
  - Network energy efficiency
  - Inferencing accuracy

AI/ML as an enabler for 6G network sustainability

Improving 6G energy efficiency with low-complexity AI solutions

Proposed solutions in 3 main categories

1. Universal functional approximation property of AI/ML
2. Moving complexity from inference into offline training
3. AI/ML to acquire more additional data from network for energy saving

6G network as an efficient AI platform

6G network as an efficient AI platform

Hexa-X

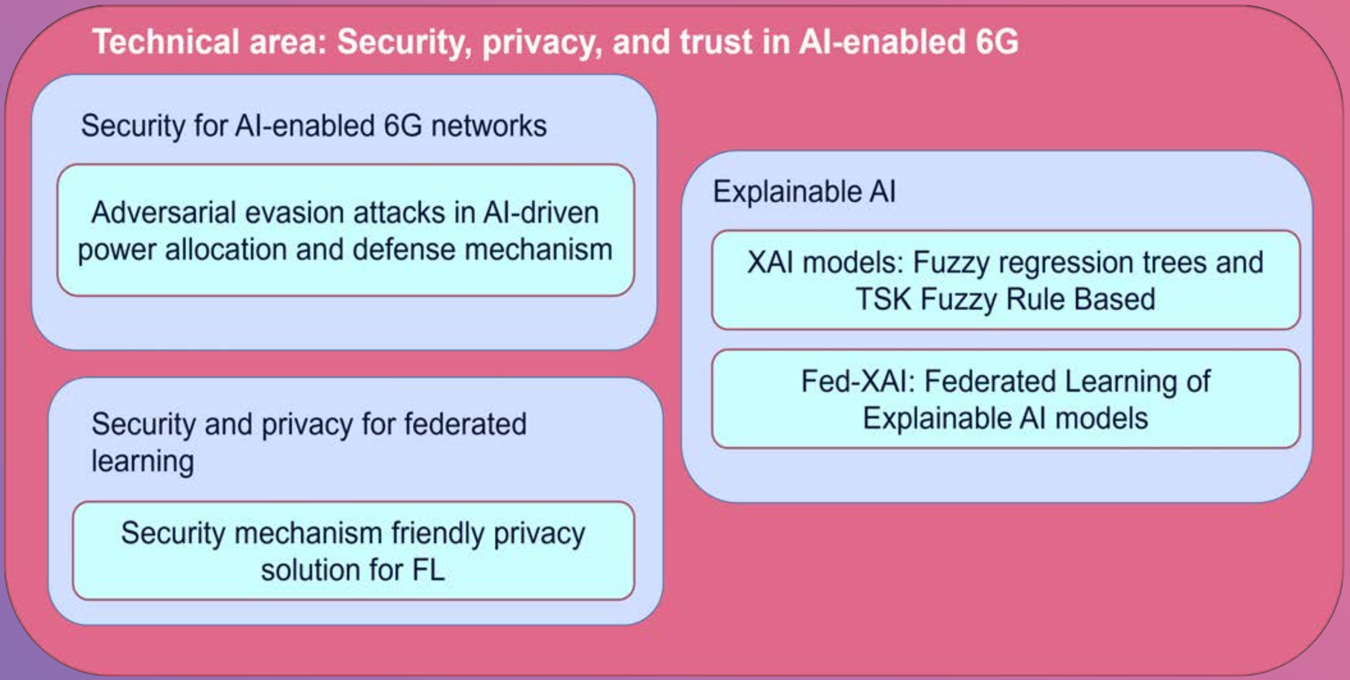
Enable and enhance the global operation of AI services, with **computing as a native part of future networks**

- Proposed solutions address the following problem:
  1. **Network services and data structures for AI applications**
     - AIaaS - seamless exploitation of network knowledge
     - Flexible compute workload assignment, CaaS
     - AI workload placement for energy, knowledge sharing and trust optimisation
  2. **Efficient inference for distributed AI**
     - Scalable and resilient deployment of distributed AI
     - Joint communication and computation orchestration for edge inference
     - Goal-oriented communication approach for edge inference
     - Network impairment resilience of autonomous agents
  3. **Efficient training for distributed AI**
     - Centralized training and decentralized execution (CTDE) approach to multi-cell multi-user MIMO
     - Federated ML model load balancing at the edge
     - Frugal Federated Learning
- Relevant KPIs/KVIs
  - AI agent availability, reliability, latency
  - Network and UE energy reduction, i.e., energy efficiency
  - Inferencing accuracy
  - Resource efficiency and complexity

Privacy, security & trust in AI-enabled 6G

Privacy, security & trust in AI-enabled 6G

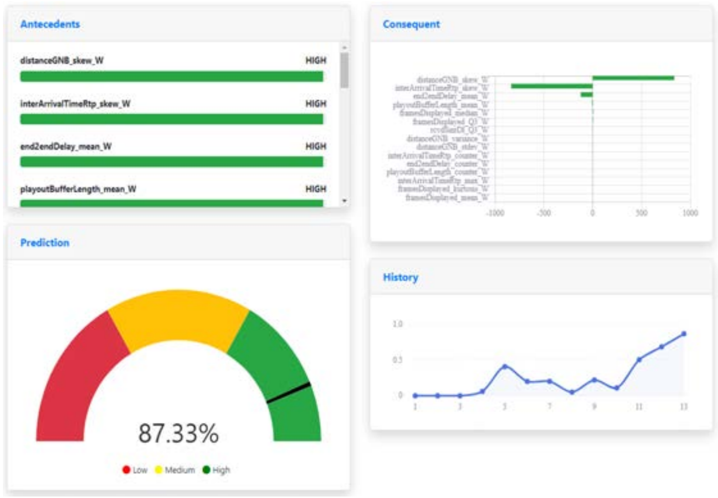
- The use of ML on massive amount of data is steadily increasing in time
- Cyber-attacks can be detected thanks to in-network AI/ML functionality
- Trustworthiness in AI/ML becomes critical for AI-pervasive 6G because AI/ML-based decisions are done for autonomy of communication and detection of cyber-attacks
- D4.3 focuses on possible adversarial attacks to AI/ML and mitigation technique to increase the robustness of AI model, also focus on how privacy of AI/ML can increase, and how to better interpret the AI/ML via explainable AI.



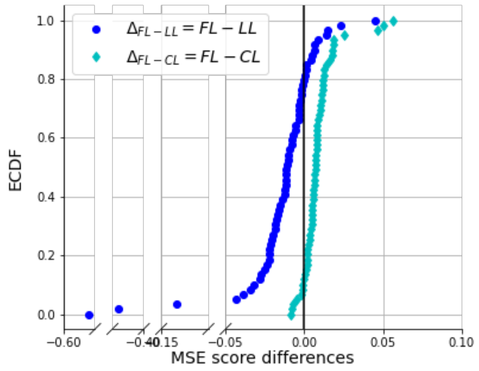
Demonstration activities - Federated eXplainable AI (FED-XAI) demo

Demo: Federated eXplainable AI (FED-XAI)

- **Problem/ challenge to be addressed**
  - Fed-XAI models target at forecasting QoE (regression problem)
    - Several instances of vehicular User Equipment (UE), connected to a B5G/6G network, receive (or send) a video stream, whose perceived quality is crucial for the availability of advanced driving assistance systems, such as see-through (or tele-operated driving)
- **Final proposed solution**
  - Development of a framework for Fed-XAI:
    - Intel OpenFL library for FL process, extended for Fed-XAI support
    - Container as de-facto standard for lightweight virtualization
    - Messages are exchanged via RESTful APIs over HTTPS for security
- **Evaluation towards 6G KPIs/ KVIs**
  - Inference accuracy:
    - Federated Learning (FL) compared with Local Learning (LL) setting
    - Fed-TSK compared with standard FL of NNs
  - Explainability: high level of interpretability ensured by the adoption of inherently interpretable models
- **Quantifiable “Connecting Intelligence” targets**
  - **T4:** accuracy of an XAI model within (<10%) of “black box” solutions



Snapshot of Fed-XAI dashboard: predictions and related explanations



ECDF of the differences of test MSE scores between FL and LL (dark blue) and between FL and Centralized learning (light blue).

FL - TSK		FL - NN	
MSE	R^2	MSE	R^2
0.066	0.559	0.059	0.606

Average values of test MSE scores and R^2 scores obtained with the FL setting by the TSK and NN models

Conclusions on tackling the Connecting Intelligence research challenge

Summary of the AI-driven solutions of the document by their domain of application

Hexa-X

- The technical solutions described in this report contribute to various network domains.
- A corresponding breakdown to technical solutions in the illustration on the right-hand side.

AI/ML for network performance enhancement

RAN performance enhancements

- ML E2E learning for RIS-assisted communication
- NN/ML channel (de)coding for constrained devices
- Enhanced AI-based beam selection in D-MIMO
- AI-empowered receiver for PA-nonlinearity compensation
- AI-based enhancements for sub-THz
- Channel charting based beamforming

Intelligent E2E management & orchestration

- AI/ML based predictive orchestration
- Distributed AI for automated UPF scaling

Enablers for sustainable 6G

- Low complexity resource allocation in cell-free massive MIMO
- Channel estimation for RIS with mobility
- Low complexity channel estimation with NN
- Deep unfolding for channel estimation
- Hybrid model for channel charting

Secure, private and trusted AI

- Adversarial attacks in AI-driven power allocation
- Robustness of AI-driven power allocation
- Security mechanism friendly privacy solutions for FL
- XAI models
- Fed-XAI: FL of Explainable AI models

6G as efficient AI platform

Network services and data structures

- AI as a Service (AIaaS)
- Flexible compute workload assignment, CaaS
- AI workload placement

Efficient distributed inference

- Resilient distributed AI
- Joint edge communication and compute orchestration
- Goal oriented communication for edge inference
- Network impairment resilience

Efficient distributed learning

- Federated ML load balancing at the edge
- Multi-agent ML for multi-cell MU-MIMO
- Frugal Federated Learning

Thank you!

HEXA-X.EU

Hexa-X

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101015956.