

tsdsi
India's Telecom SDO

Telecommunications
Standards Development
Society, India

TECH DEEP DIVE

TTDD 2022 CONFERENCE (5th EDITION)

STANDARDS FOR SUSTAINABLE DEVELOPMENT

 **Date:** 7-10 November 2022

**Session 1: AI/ML-based enablers and applications for 6G >
8th November 2022**

**<Enablers, Architecture, Applications of AI/ML in Beyond 5G & 6G
– An Operators view>**

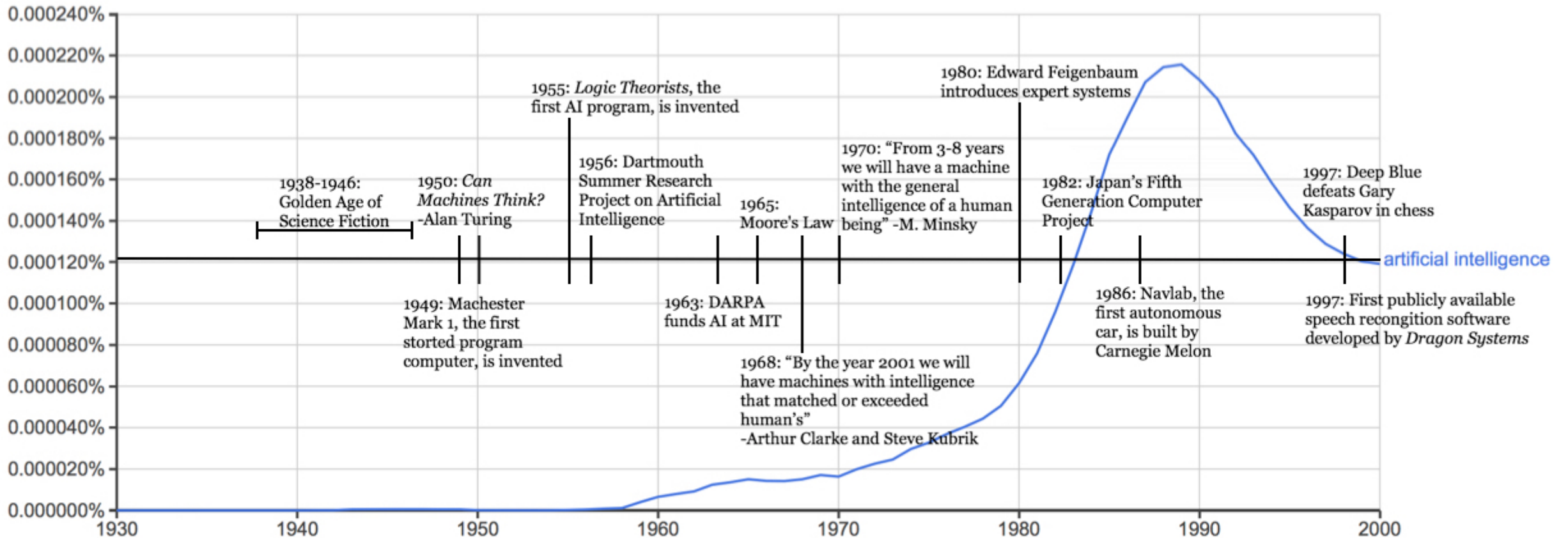
by

**<Mahesh Nayaka Mysore Annaiah>
(AVP: Reliance Jio, Chair: TSDSI SGSS)**

AI is not new or latest technology?

AI – As the name indicates, it provides intelligence and aids in complete automation to wireless networks, by imitating human brain, his/her thought process and his/her intelligent behaviors.

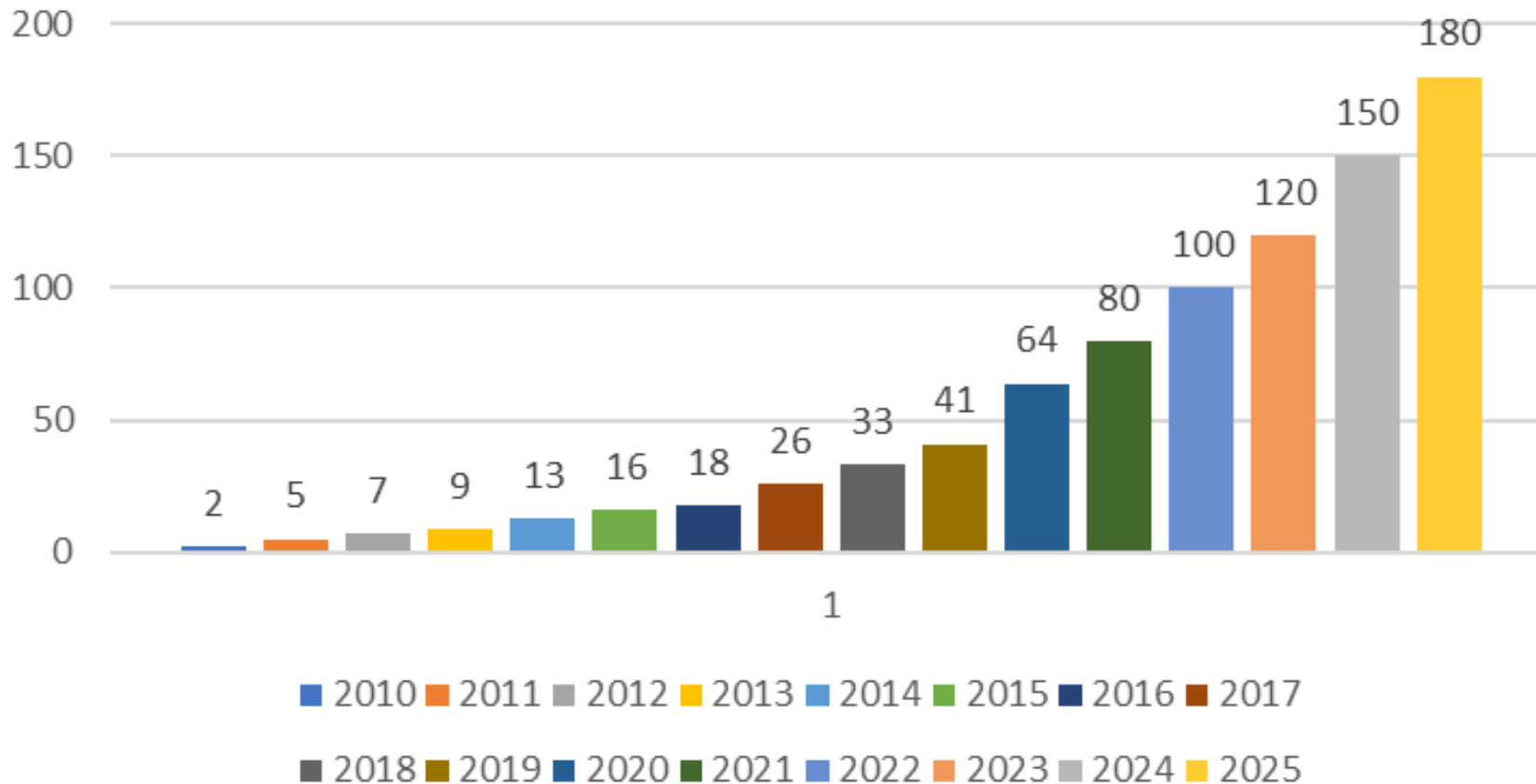
ARTIFICIAL INTELLIGENCE TIMELINE



Enablers for Artificial Intelligence

Data Availability

Data volume in zettabytes.



Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 (in zettabytes)

Src: Statista

Enablers for Artificial Intelligence

High Processing Compute and Storage availability.

	Training		Inference		Generality ⁸⁸	Inference accuracy ⁸⁹
	Efficiency	Speed	Efficiency	Speed		
CPU	1x baseline				Very High	~98-99.7%
GPU	~10-100x	~10-1,000x	~1-10x	~1-100x	High	~98-99.7%
FPGA	-	-	~10-100x	~10-100x	Medium	~95-99%
ASIC	~100-1,000x	~10-1,000x	~100-1,000x	~10-1,000x	Low	~90-98%

Src: Statista

Enablers for Artificial Intelligence

Few Top performing Chips for AI.

Alphabet

- Tensor Processing Unit[TPU]
- ASIC type.
- For Google Tensor Flow framework.
- Edge TPU/ Cloud TPU variants.
- Supports PetaFLOPS of compute.
- Used for ML, DL, Quantum Computing Simulations.

Apple

- A15 Bionic Processor/ M2 processor
- 5nm chip, 2 high-performance cores, 4 high efficiency cores.
- GPU – 5 cores, 10 cores[M2].
- Neural engine – 16 cores.
- Supports 15.5 trillion operations per second.
- Transistors – 15 billion, 20 billion[M2]

ARM

- Neoverse V-series processor
- Cloud-to-Edge purpose
- 32 – 128+ cores
- Used for HPC, AI/ML Accelerated workloads.
- Ethos Neural Processing unit [NPU]
- ML Inference processor.

Intel

- Gaudi-2 Processor
- It is a Training processor.
- 24 Tensor Flow cores.
- 7nm chip
- For DL, Comp. Vision, NLP.
- Cloud-to-Edge purpose.
- Greco Processor
- It is an Inference processor.
- Similar to Gaudi.

Enablers for Artificial Intelligence

Few Top performing Chips for AI.

AMD

- AMD INSTINCT MI250X Accelerator.
- For ML/DL, it supports up to 383 TFLOPS.
- For HPC workloads, it supports up to 95.7 TFLOPS.

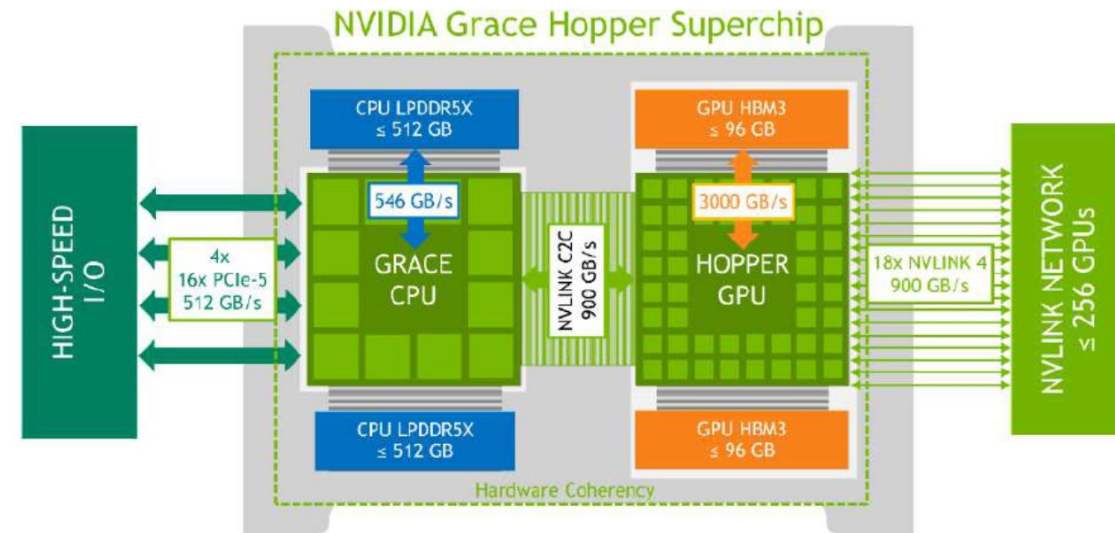
IBM

- Telum Processor [An Inference chip]
- 7nm chip, 8 Processor cores
- Supports 5+ GHz clock speed.
- Integrated AI Accelerator supports 6+ TFLOPS compute capacity.
- Used in Transaction inference in real time and at scale.

NVIDIA

- NVIDIA GRACE HOPPER Superchip.
- 72 Arm Neoverse V2 CPU cores.
- NVIDIA Grace CPU and an NVIDIA Hopper GPU into a single superchip.
- 900 GB/s C2C Interconnect.
- 546 GB/s of memory bandwidth per CPU

These are just few AI Chip vendors. There are many more such vendors in the eco system, who have designed chips, which can be used for various AI purposes, covering majority of the AI algorithms.



Are the today's cellular networks lack Intelligence? tsdsi India's Telecom SDO

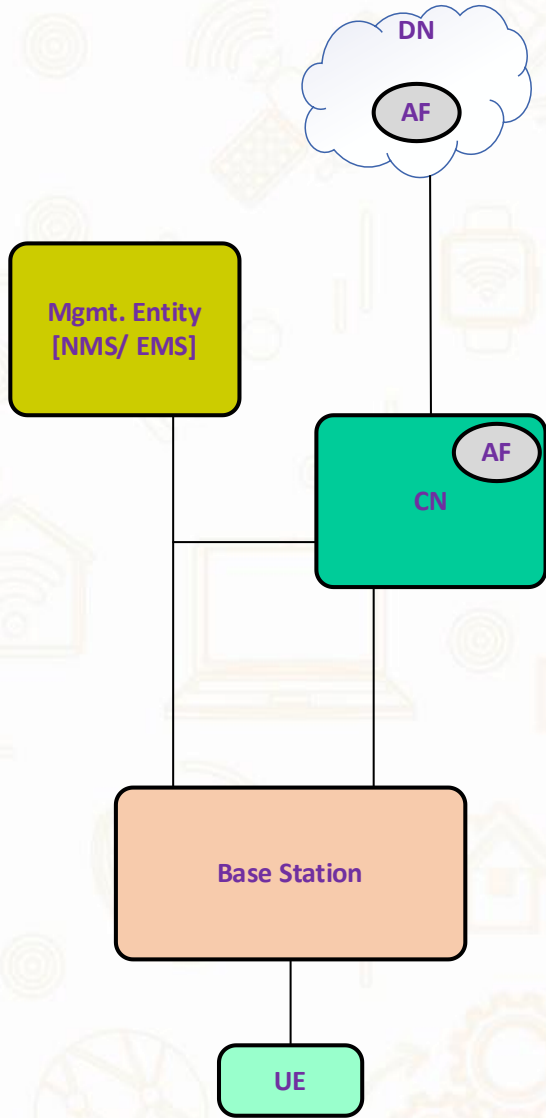


Fig-A: Classical Network

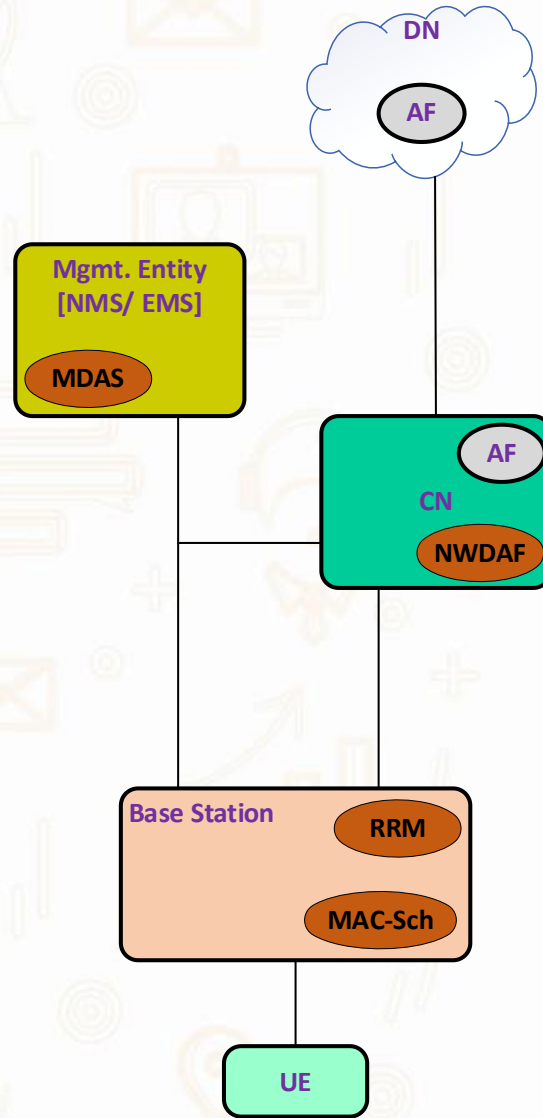


Fig-A1: With Active Entities.

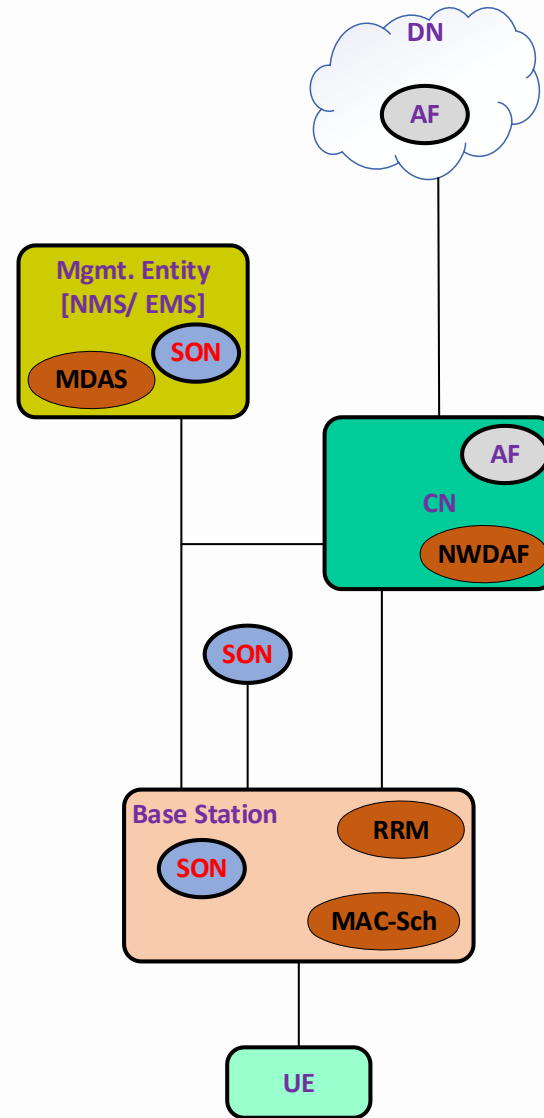


Fig-A2: Enabled with SON.

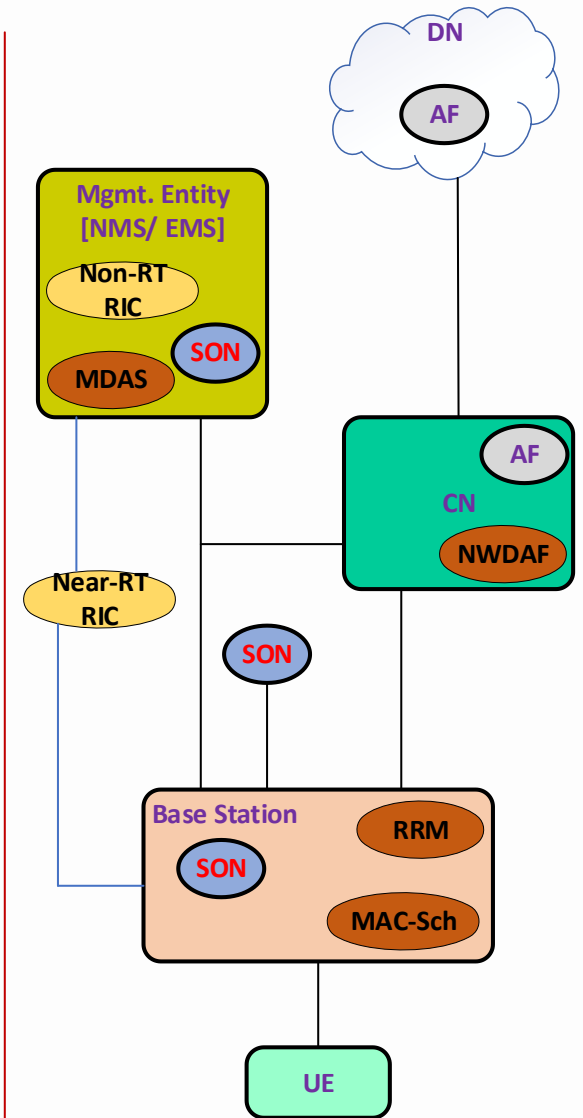


Fig-A3: Enabled with O-RAN.

Functional Architecture for enabling AI.

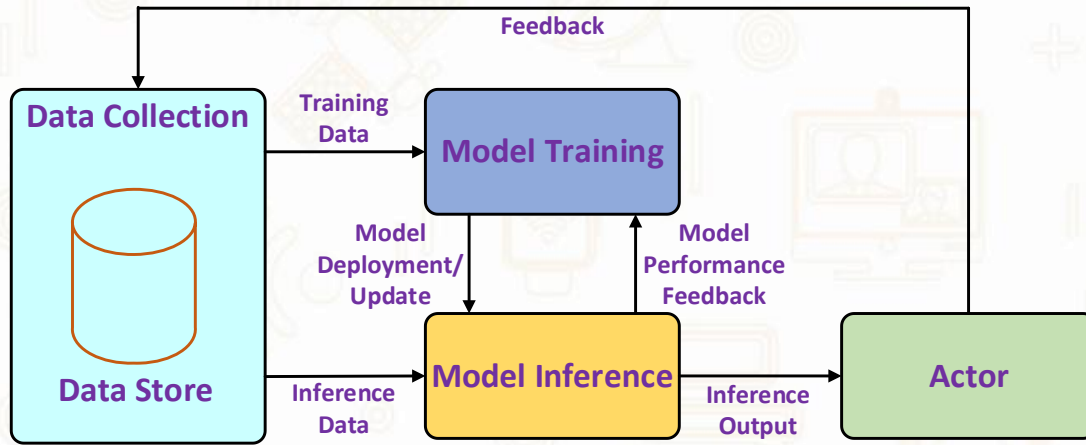


Fig: Functional Architecture for enabling AI.

The target is to achieve Network intelligence level of L5, by enabling AI into the Network Architecture.

Network intelligence level		Dimensions				
		Action implementation	Data collection	Analysis	Decision	Demand mapping
L0	Manual network operation	Human	Human	Human	Human	Human
L1	Assisted network operation	Human and System	Human and System	Human	Human	Human
L2	Preliminary intelligence	System	Human and System	Human and System	Human	Human
L3	Intermediate intelligence	System	System	Human and System	Human and System	Human
L4	Advanced intelligence	System	System	System	System	Human and System
L5	Full intelligence	System	System	System	System	System

NOTE 1 – For each network intelligence level, the decision process has to support intervention by human being, i.e., decisions and execution instructions provided by a human being have the highest authority.

NOTE 2 – It is to be noted that this table may be used to determine the network intelligence level for each dimension (and not the overall network intelligence level).

Ref-1 : 3gpp

Ref-2 : ITU-T Rec Y.3173

Enabling AI for SON: An Illustration.

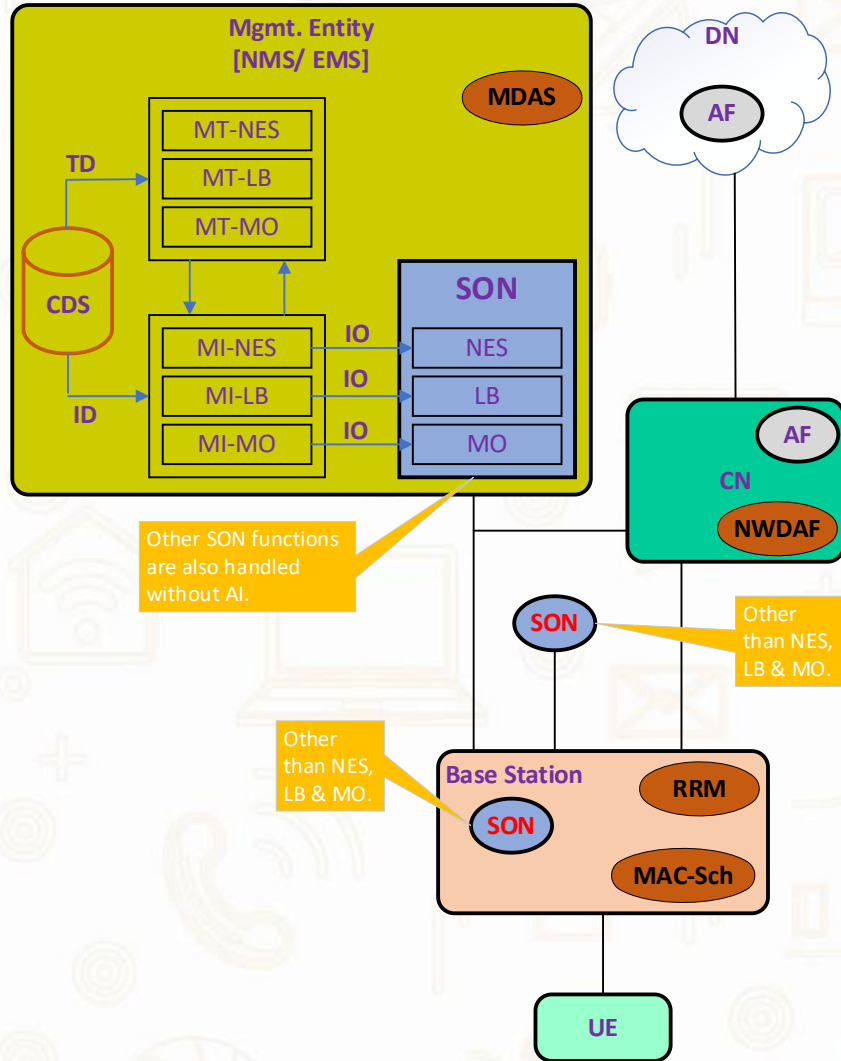


Fig-B: AI Enabled for SON with both Model Training and Model Inference @ the Mgmt. Entity .

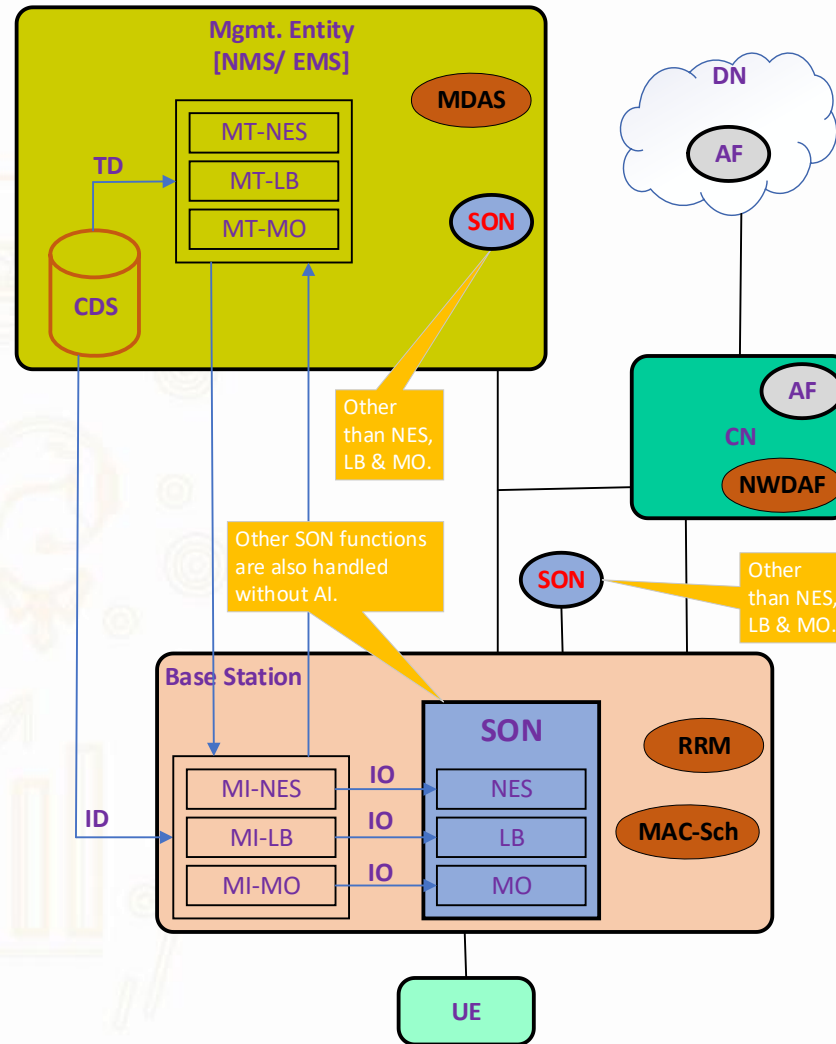


Fig-B1: AI Enabled for SON with Model Training @ Mgmt. Entity and Model Inference at the BS.

- 3GPP has already initiated the study as part of Rel-17 study items.
- The initial focus of this study is only on Network Energy Saving[NES], Load Balancing[LB] and Mobility Optimization[MO].
- TD – Training Data.
- ID – Inference Data.
- MT – Model Training.
- MI – Model Inference.
- CDS – Central Data Store.

Enabling AI for SON: Model Training at all data stores.

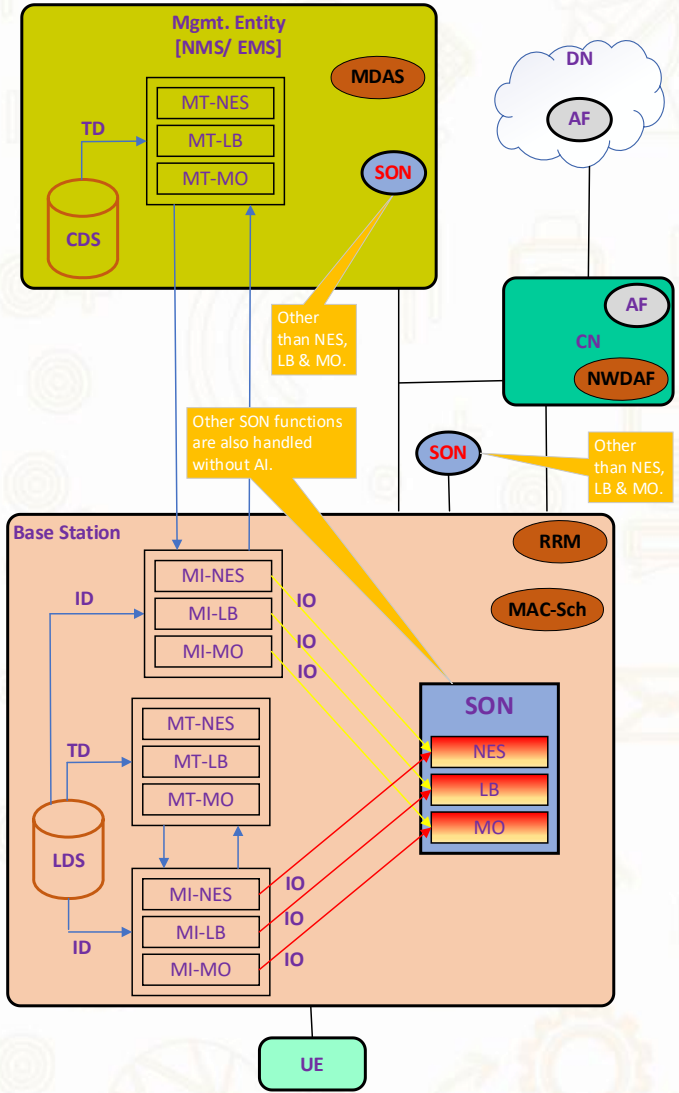


Fig-C: AI Enabled for SON with two separate trained models, one Trained with CDS data and other with LDS data .

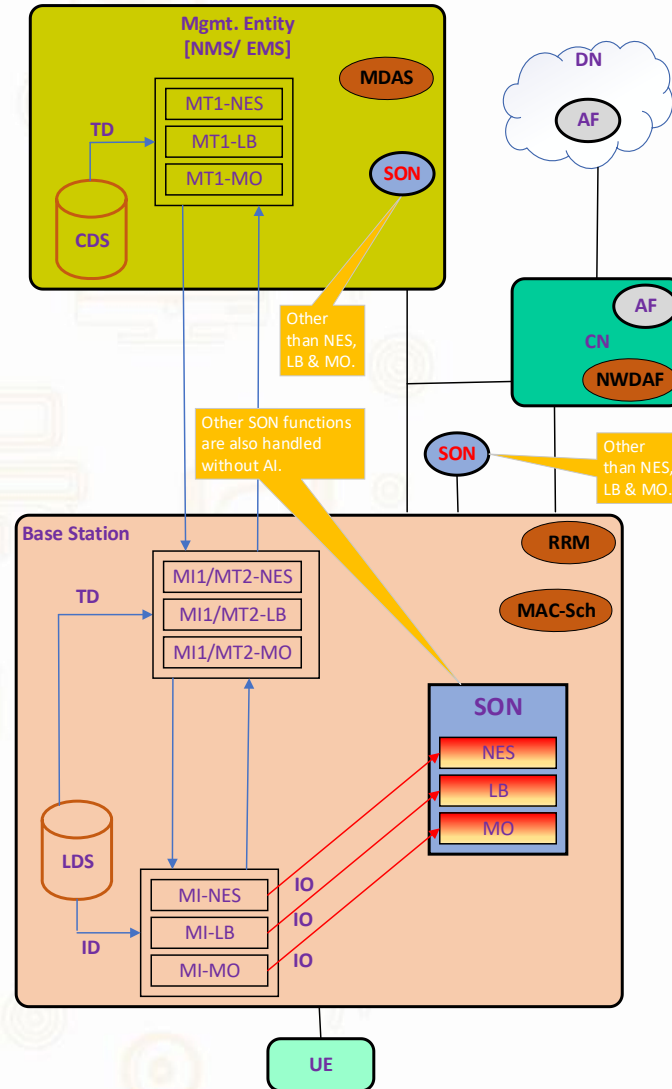
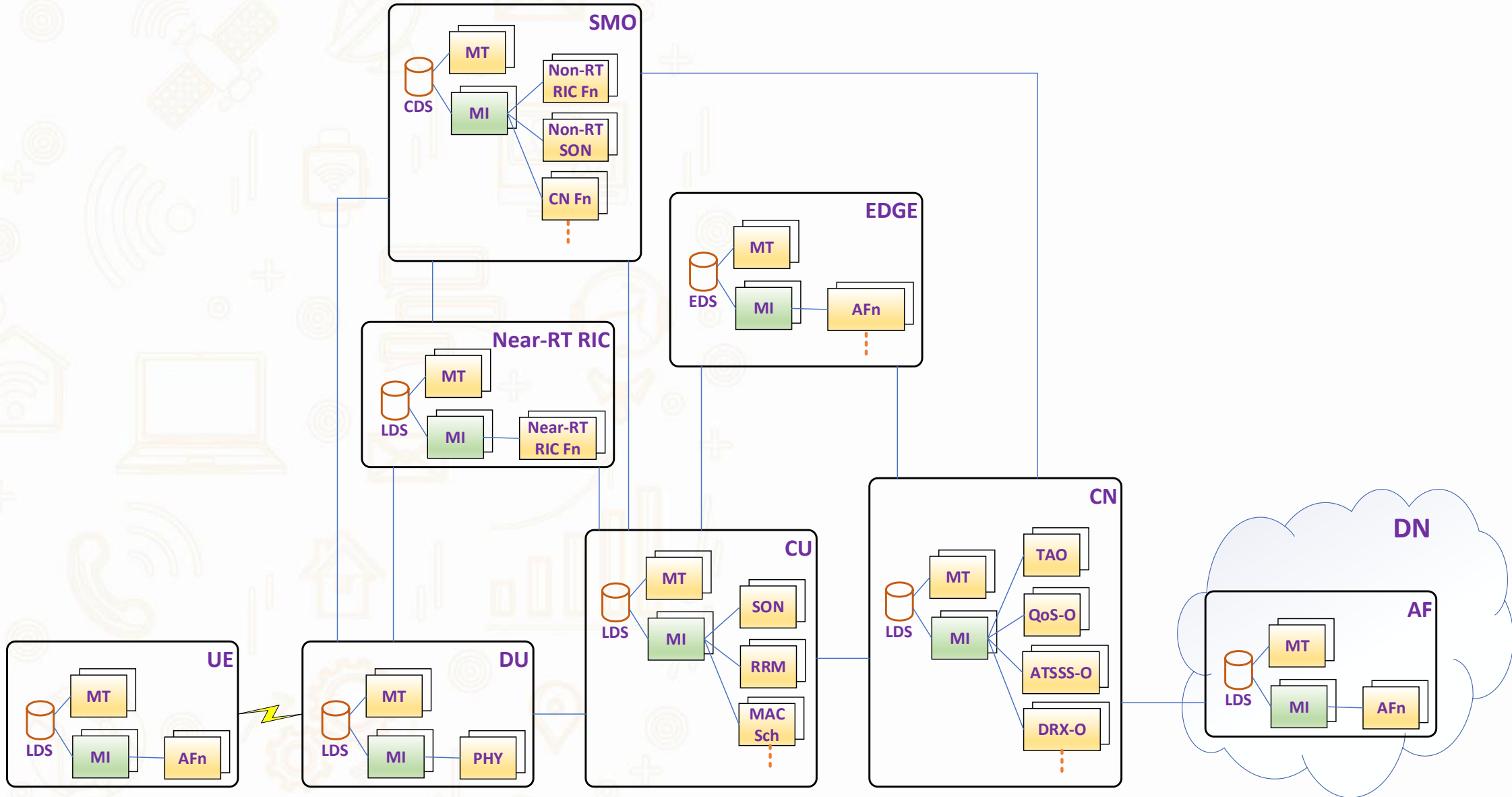


Fig-C1: AI Enabled for SON with one trained model, but trained with both CDS and LDS data separately.

- LDS – Local Data Store.
- As HPC and High storage is affordable today, Base Station can also store huge data in its LDS.
- In Fig-C, the Model gets trained with CDS data at Mgmt. Entity, followed by another model gets trained by LDS data at Base Station. The Inference output from both models will be used by the SON actor to select the optimum one for further decisions.
- In Fig-C1, the Model gets trained with CDS data at Mgmt. Entity, followed by the same model gets trained by LDS data at Base Station. The Inference output from this final model will be used by the SON actor for further decisions.
- It is a kind of Federated Learning.

Enabling AI for the end-to-end network optimization.



Tentative AI Applications in 6G: Network related.

<p>6G network related.</p>	<p>Network Layer:</p> <ul style="list-style-type: none">- Supervised Learning can improve resource allocation, fault prediction, etc.,- Unsupervised learning algorithms can help in routing, traffic control, parameter prediction, resource allocations, etc.,- Reinforcement learning can help in traffic prediction, packet scheduling, multi-objective routing, security, and classification, etc., <p>Physical Layer:</p> <ul style="list-style-type: none">- Unsupervised learning can help in interference cancellation, optimal modulation, channel-aware feature extraction, channel estimation, etc.,- Deep reinforcement learning can help in link preservation, scheduling, transmission optimization, on-demand beamforming, energy harvesting, etc.,
<p>6G network autonomy.</p>	<p>Detect and Resolve network issues.</p> <ul style="list-style-type: none">- Monitoring network status in real-time.- Keep network healthy.- Can provide intelligence at the edge devices and edge computing, to learn to solve security problems autonomously. <p>Aids in autonomous applications like autonomous aerial vehicles, autonomous robots, etc.,</p>

Tentative AI Applications in 6G: Smart Health care.

Smart Health care.	Real-time health status monitoring remotely.
	Health diagnosis remotely. ML model can detect abnormal health status, can predict possible health issues/ diseases, can decide on the accurate health solutions. Models on the cloud can analyse the physiological data and send a warning message if the diseases occur suddenly.
	Medical treatment remotely, with high accuracy and efficiency.
	AI enables Hospital-to-Home service.
	AI enabled AR, enables doctors to observe the inside of the patient's body clearly without making any incision.
	AI enabled VR, can enable a simulated environment, which helps doctors to practice medical operations.
	Medical Robots can act as Virtual Doctors/ Nurses, to take care of the patients in Hospitals. - Can perform accurate surgeries. - Mini/Micro/Nano Medical Robots can help the doctors to control them to enter the human's body to take pictures, deliver drugs, or remove diseased tissues.
	AI enables the privacy and security of the Health case systems through Edge computing, Block chains and Federated Learning.

Smart Manufacturing:

Smart manufacturing refers to a IoT-connected manufacturing system that applies a variety of control and data analytics approaches to improve manufacturing performance.

- AI enabled Robots will manage the complete Manufacturing process, enabling minimum or nullifying the human intervention.
- AI enables precise and smart manufacturing.
- AI enabled Block chains and Edge computing improves the performance of the manufacturing systems.
- AI enables Inter-Robotic communications in Manufacturing systems for mutual cooperations.
- Intelligent robots on the edge side can take videos of the industrial process and then upload the data to the cloud, while the learning algorithms on the cloud will make decisions to control the robots.
- Nano-Robots are used in Manufacturing processes which are hazardous but still require high precision, such as nuclear power plants and oil pipelines.

Tentative AI Applications in 6G: Smart Home.

Smart Home.	AI enables the remotely controlling of the house hold systems like lighting, furniture, thermostats, etc.,
	AI enables adaptive real-time control without much human intervention.
	Federated learning enables the protection of residents' safety and privacy.
	AI enables fall detection of the elderly and also predict the fall and notify. - If an old man suddenly falls down in a smart home, the data collected by embedded intelligent sensors and video surveillance will be sent to the cloud by 6G wireless network immediately. The well-trained prediction model on the cloud side should detect this emergency by analysing the data and then send distress signals to the man's relatives and the ambulance.
	AI enables the detection of emergencies like forced entry, fire, etc.,
	AI Enables privacy protection where data storage and processing will be done within the residents' premises.

Tentative AI Applications in 6G: Intelligent Transportation Systems.

AI Function	Use-cases
Non-linear prediction	Traffic demand modeling
Control functions	Signal control, dynamic route guidance
Pattern recognition	Automatic incident detection, image processing for traffic data collection and crack identification in pavements or bridge
Clustering	Identification of specific class of drivers based on behavior
Planning	AI based decision support systems for transportation planning
Optimization	Designing an optimal transit network, developing an optimal work plan for maintaining pavement network, developing an optimal timing plan for a group of traffic signals

Tentative AI Applications in 6G: Energy sector.

Smart Energy.

AI enables accurate fault prediction of the power plants.

- monitor equipment and detect failures before they happen.

AI aids real-Time maintenance of the power plants.

AI aids in identification of ideal maintenance schedules for power plants.

AI aids to improve the efficiency and reliability of the power plants.

AI aids in Self-healing of the power plants.

AI enabled UAVs helps to monitor the wires all along the lines from generation, transmission and distribution networks to detect and prevent any breaks, losses, thefts, etc., using image based data analysis.

AI to optimize energy consumption and storage both at power plants and premises.

AI aids to forecasts of electricity demand and generation, improving production decision making.

AI will facilitate decision making about optimal times for distributed generation to contribute to the grid, rather than draw from it.

AI plays a critical role in disaster recovery. Helps to improve assessment of damages and optimization of decision making, using images and information of the damaged sites.

AI aids in prevention of losses due to Informal connections.

Conclusions

- Time has ripened for the practical realization of the AI algorithms, for more accurate results.
- 5G-Advanced may partially utilize the AI techniques, but the 6G and beyond shall be completely enabled with AI at all possible opportunities.
- Every decision shall be AI based in 6G and beyond.
- W.r.t RAN, SON has already optimized the results to the maximum in majority of the use cases at additional cost. The AI should achieve better results with no or less extra cost.
- During initial phases of AI deployment, both SON and AI may coexist, which shall not deteriorate on the optimized results.
- AI shall target to achieve complete 6G automation with null or very minimum human intervention.
- AI automation shall be vendor agnostic.

Thank You

<Mahesh.Mysoreannaiah@ril.com>