

Neurosystem- An AI Evolution for NextG Services and Auto Drive Connectivity

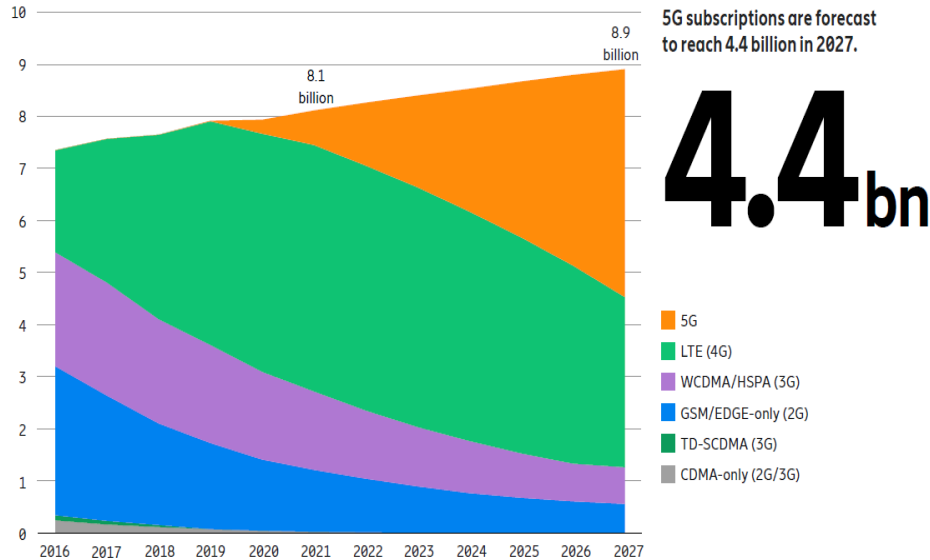


Ravi Sinha (Jio)

5G: Deployments Worldwide (Ericsson Mobility Report)



Figure 3: Mobile subscriptions by technology (billion)



¹ GSA (October 2021).

² A 5G subscription is counted as such when associated with a device that supports New Radio (NR), as specified in 3GPP Release 15, and is connected to a 5G-enabled network.

Up to 2027 5G Subscribers will be close to 50% of the total subscribers. 660 million 2021
5G account 90% subscription at the end of 2027 in North America.

5 Introduction Ericsson Mobility Report | November 2021

Figure 1: Comparison of forecasts vs. up-to-date data

	Forecasted for 2016 (Predictions made in November 2011)	Actual 2016	Forecasted for 2021 (Predictions made in November 2015)	Estimates for 2021 (End of year figures)
Mobile subscriptions	8.4 bn	7.4 bn	9.1 bn	8.1 bn
Mobile PC/tablet subscriptions	550 m	180 m	350 m	300 m
Smartphone subscriptions	2.6 bn	3.7 bn	6.4 bn	6.3 bn
4G subscriptions	510 m	2.0 bn	4.1 bn	4.7 bn
5G subscriptions	-	-	150 m	660 m
Average traffic per smartphone	800 MB/m	1.9 GB/m	8.5 GB/m	11.4 GB/m
Total mobile traffic	4.6 EB/m	6.7 EB/m	51 EB/m	65 EB/m

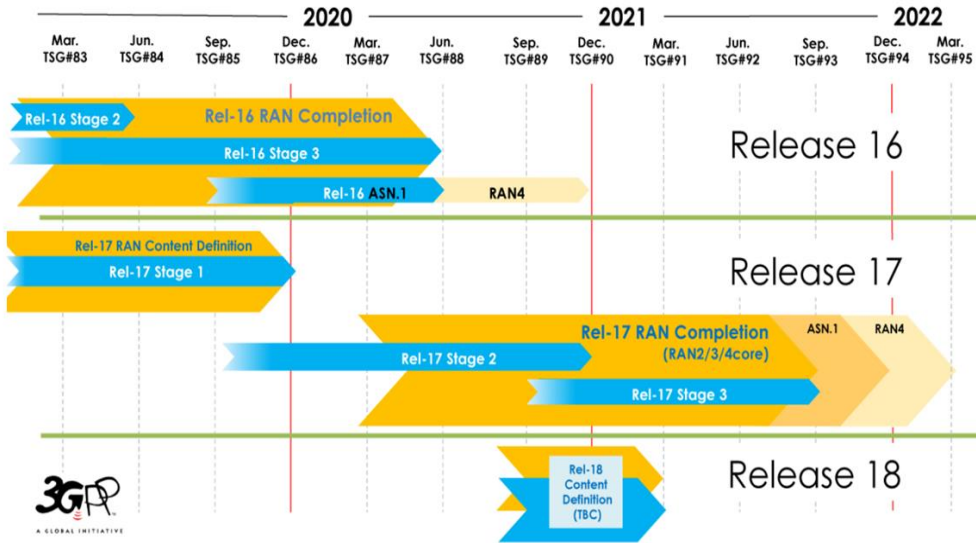
Up to 2027 5G Subscribers will be close to 50% of the total subscribers. 660 million 2021

In 2027 Data consumption is projected to grow up to 288 EB/m which is 65EB/m 2021

Neurosystem 2021



3GPP Roadmap and alignments with AI



FIG, 3GPP Release 16 and 17 timeline

Rel 16

1. Sidelink- V2X, Public Safety, reduced Power Consumption, lower Latency & improved Reliability.
2. Precise Positioning
3. URLLC (Unlicensed)
4. Integrated Access Backhaul- Flexibility of deployment, high capacity backhaul integrated with gNB
5. mMTC NW with Reduced Capacity Data

Rel 17

1. Extended Reality
2. mmWave for Higher spectrum and bandwidths
3. NTN – Non Terrestrial Networks with smart Agriculture and IoT based Satellite NW
4. 5G Multi Cast Broadcast

Rel 18

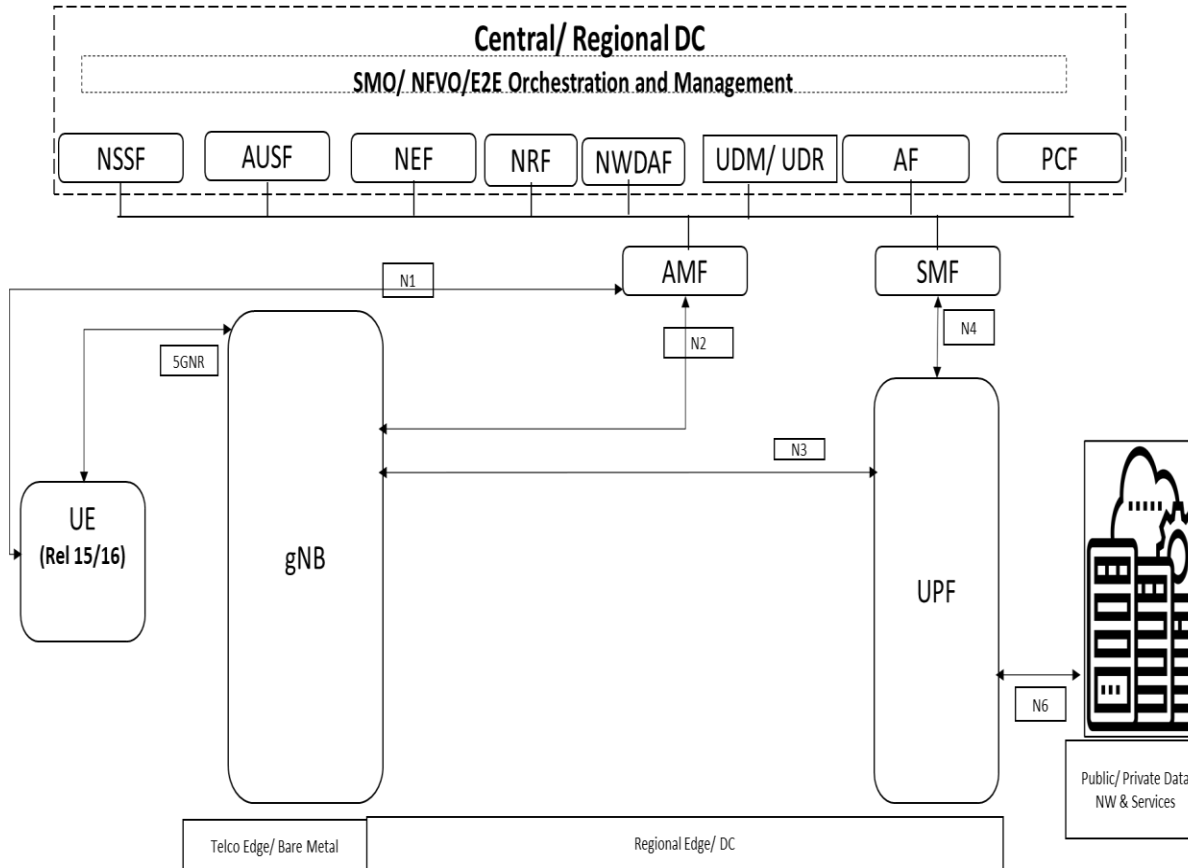
1. Artificial Intelligence and Machine Learning
2. Lower Energy Consumption
3. Full Duplex Communication related to Sub-bands

3GPP

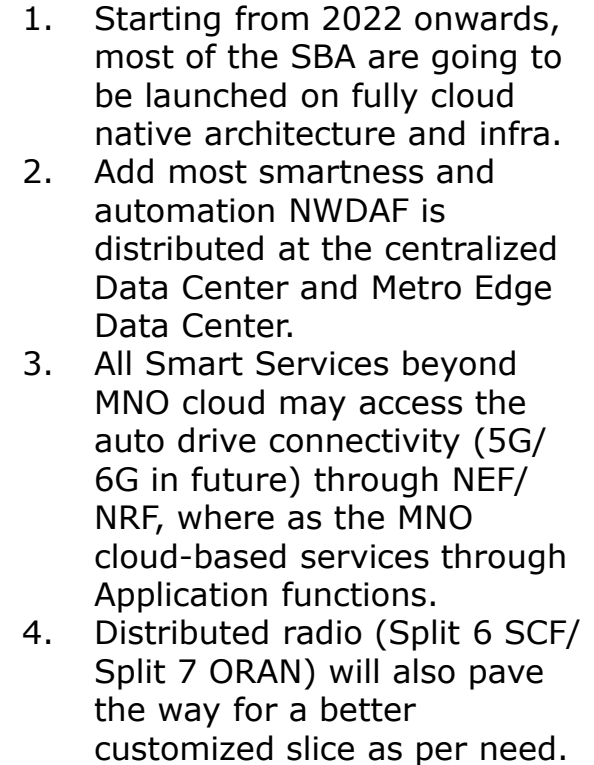
1. Touchbase with AI/ML in 2023 onwards. Some of the basics will start in 2022.
2. This is the time when Stack level interaction with AI will go into specs.
3. Before that only Radio Intelligent Controller level interaction for AI will be in development



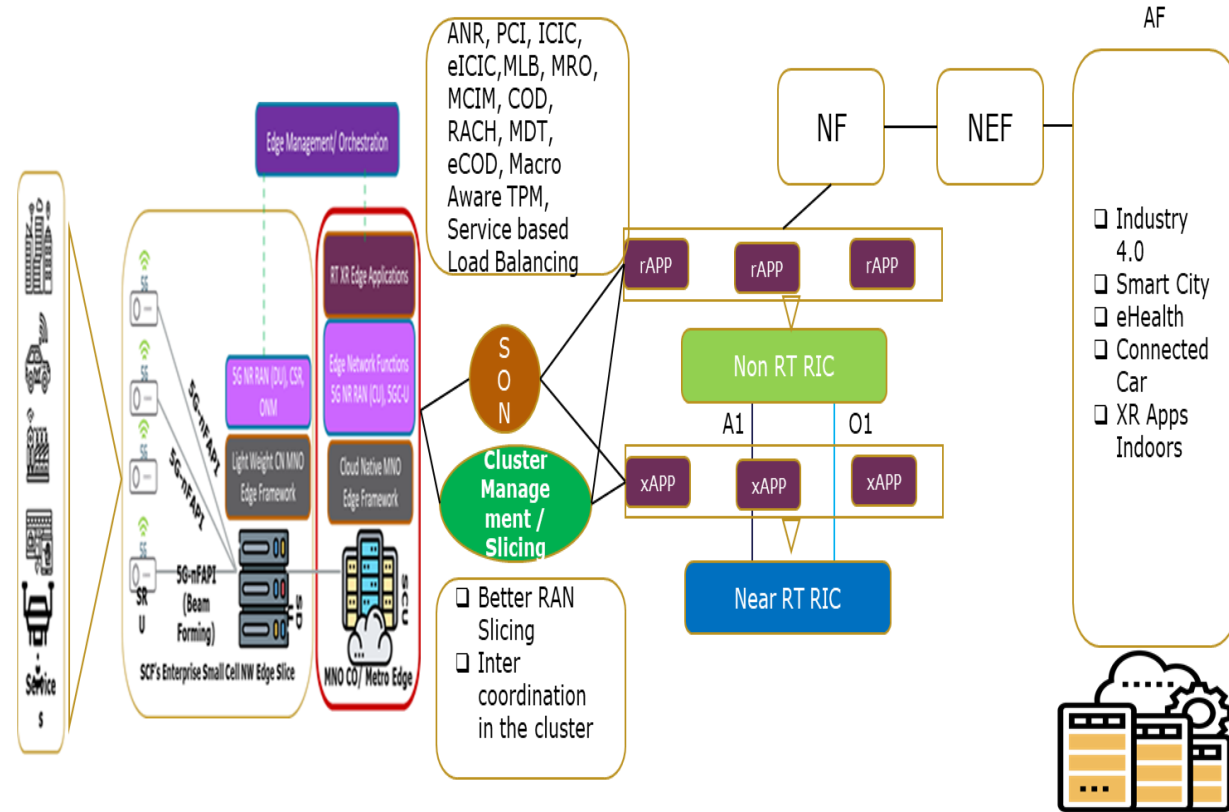
5G: Service Based Architecture, Smart Services



1. SBA architecture mainly suitable for Stand alone 5G Core and NW slicing to better serve specific use cases based on business SLA coordinated between MNO and Service provider orchestrators (BSS+OSS+ SMO+PCF).
2. Most of the Tier I vendors are moving to SA/Slice based service support connectivity in 2023/ 2024.
3. UDM, UDR, NWDAF, NSSF, PCF are some of the major modules added to enable these specialized services through NEF/ NRF/ AF

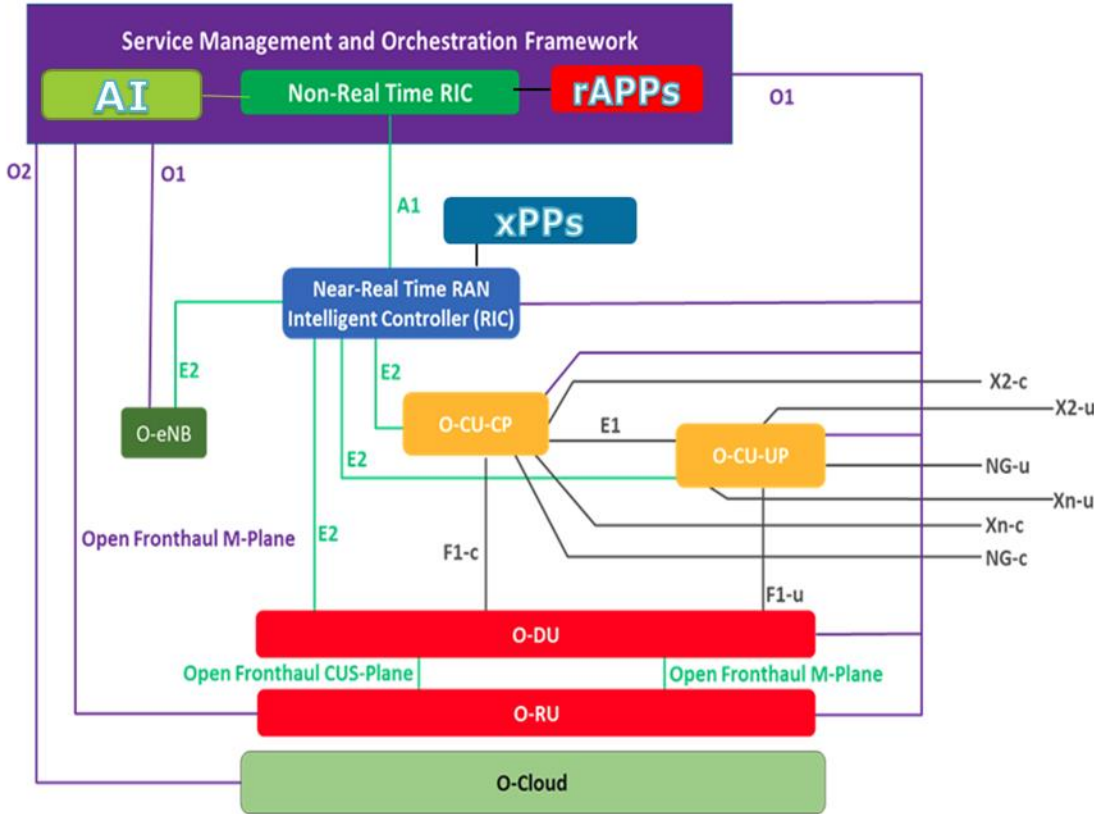


AI with 5G Networks: 3GPP Rel 16 and Open RAN



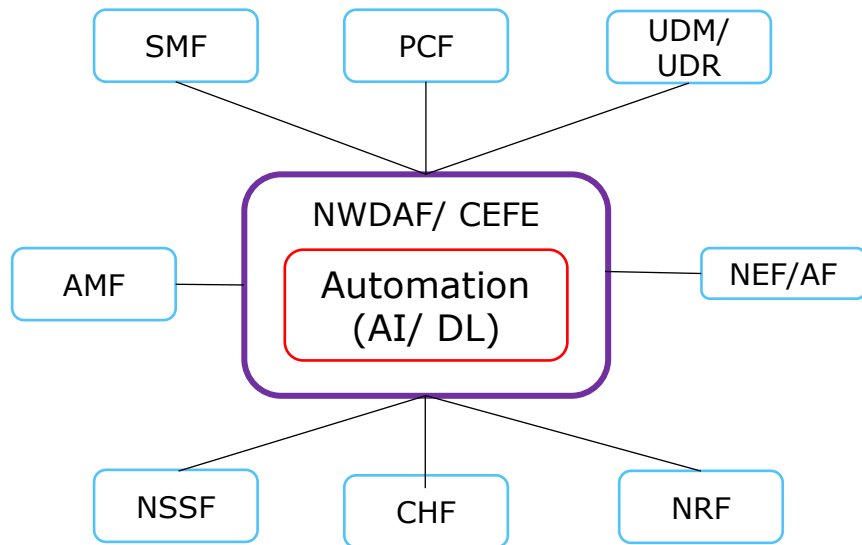
1. Cognitive and Automation approach started with 4G Networks with a concept called SoN (Self optimization Networks) to add prescriptive capacities to the NW along with automation, where the distributed architecture was integrated with smart Radio Resource Manager and managed at a centralized architecture, so that Interference, Auto Neighbor relations, Rach optimization, Power optimizations, Load balancing and several other features can be added.
2. 5G has taken the existing modules at a cloud native platform and orchestration along with distribution of lots of analytics at Edge and CDC.
3. Architecture moved from EMS to SMO

AI architecture with Open RAN solutions (Only for connectivity)



1. O-RAN platform one of the leading Open RAN platforms proposed the architecture, which will revolutionize the Predictive, prescriptive and Cognitive nature of the 5G network along with provides a wholistic view of how AI will be consumed with 5G/ 6G .
2. Here the xAPPs will add all set of automation and application-based smartness to 5G NR and the part of MNO secured Nw connected to Near RT RIC latency ≥ 10 ms < 1 s
3. rAPPs are the marketplace APPS having all the access to xAPPs. These rAPP may always use unified AI engine for analytics, predictive and Cognitive actions.
4. Still the architecture only provides specs to do all AI activities related to the smart management and automation of Radios cognitive processes.

5G:Where the smart service meets auto drive connectivity



NWDAF: Network Data Analytics Function

CEPE: Content Extraction and Profiling Engine

SMF: Session Management Function, PCF: Policy Management Function

UDM/ UDR: Unified Data Management and Unified Data Repository

AMF: Access and Mob Mangment Function,

NEF/NRF/ AF: Network Exposure Function/ NW Repository Function/ Application Function

NSSF: NW Slice Selection Function CHF: Charging Function

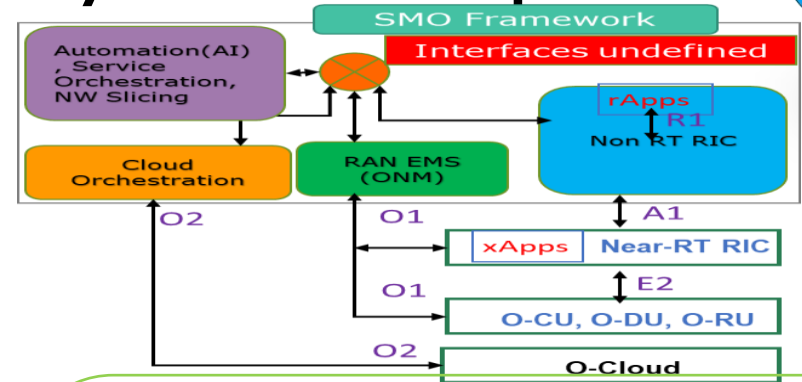
1. 5G Core (SA) with Rel 15/16 defines how Smart services meets Auto drive Connectivity.
2. SA based 5G deployment for slice-based use cases are rare in 2021, hopefully with more focus on SA/slice-based approach in 2022/ 2023 we will see heavy automation and analytics around it, where all service requests coming from Hyperscaler NW or MNO NW gets a better and cognitive treatment by 5G NW with more AI/ ML/DL capabilities being orchestrated using NWDAF and Context Extraction and profiling Engines.
3. This area is still evolving and still how AI/ML/DL engines will be used to interact with all interfaces defined and uniform AI/ML platform designed in way to be a good fit for most of the Next Generation Services.
4. Hyper scaler awareness in this area is also missing heavily.

What is missing with 5G connectivity for AI consumption



Stack

1. Real AI involvement to the Stacks of 5G transition to 6G starts with 3GPP Rel 18, expected to take shape in 2023.
2. Since eMBB is the priority for Sub 6GHz and mmWave connectivity of 5G, where stacks are configured at the time of gNB start. uRLLC and mMTC deployments are still at trial level with NSA 5G Core.
3. With SA 5G core deployments and full support of Network Slicing will open the door for lots of automation and optimizations, where some of the proprietary solution may consume some of the AI based scheduling and some tweaking with PHY layer. Some of the companies with DPU (CPU+GPU) concepts are the front runners.
4. At present AI is not used across full stack optimization or automation.
5. Different protocol layers are so tightly couple on stack that they are not good fit for AI based scheduling.
6. RAN is in Sync with Core, but transport is still static and no combined orchestration deployed at scale.



1. Service and Management layers doesn't come under 3GPP, so there is no uniform approach.
2. Major interfaces among Automation, Service orchestration, Cloud Orchestration layers interfaces to Non-Real time RIC is not defined.
3. How a common AI engine will be used to cross sync all modules are not defined.
4. If real time Schedulers are unhooked from L2 and brought to Near RT, how exactly AI engine will get engaged not defined.
5. E2E orchestrator dealing with predictive, Prescriptive and Cognitive Connectivity and Services is not even at its base line.
6. Common AI engine to talk to 5G Rel 16 not available.



AI and Data Science with 5G transition to 6G



Is the approach, right??

1. 5G transition to 6G will change the paradigm a lot.
2. The peak rate is going to move from 20 Gbps to Tbps, Latency is going to move from 1 ms to 100 microseconds and mobility will comply to 1000 km/h.
3. In all these scenarios, a highly distributed Telco Edge infrastructure with distributed NWDAF processing at the local edge will be a necessity.
4. 5G advance and 6G use cases like Cognitive NW, Connected Sustainable World, Programable physical world, Limitless connection, internet of senses, connected intelligent machines will need heavy engagement of AI to the connectivity as well as services.
5. A well aligned Orchestrator to deal with both Advance connectivity and Smart Services will also be a need for the future.
6. Space-Air-Ground-Sea integrated communication, Photonics defied Radio, Synesthesia-Networks, Nano Networks also need heavily integrated unified and federated AI engines.

1. Utilization of AI to Next Generation Telco based services are very much driven by Silicon Vendors like NVIDIA, Qualcomm, AMD where the industry is consuming AI on GPU,DPU and FPGA platforms.
2. Due to the business model and development of their next generation processors are also attached with their business cases.
3. Service industry is also not thinking out of the box for a distributed Service architecture which is negotiating and coordinating with fully distributed DWDEF and CEPE.
4. Standard interfaces and APIs to access these distributed module by the AI engines with Inference and Data Model Training at a distributed fashion is still to come.
5. On the AI front major limitation is their obsession to accuracy, lack of reasoning, extrapolation, imagination, brittle, Energy hungry, curve fitting approaches.



Areas for Neurosystem to focus

1. NW Automation and Orchestration to integrate distributed Context aware connectivity support going to a unified platform.
2. Promoting Silicon vendors to join for an Open AI architecture where different distributed HW platforms of Edge can leverage a flexible AI architecture and vast API related to it.
3. Distributed AI approach to integrate distributed DWDAF and CEPE platforms on Regional and Metro Edge. This also provides a huge opportunity for virtualized memory management to enable shared data environment over container and K8S based PaaS/ IaaS to break vendor lockin.
4. A distributed and open architecture of Event Creation, Ingestion, contextual decision intelligence and at the end analytics and Predictive and Prescriptive reporting.
5. Define the boundaries for next generation Networks in terms configuration, Data Ingestion, Data Verification, Machine resource management, Process management and Monitoring.



Thanks