AI/LLM for Networking

LFN Strategy document

THE LINUX FOUNDATION

Introduction

A Decade of Networking

Last decade: Completed & in-deployment

- Disaggregation/SDN/NFV
- Open Source based stacks (data plane to network functions)
- Cloud Native (Telecom and Cloud alignment) migration
- Intent Based and SONs
- Secure End to End Interop across OSS projects (SBP) - LFN as umbrella, Nephio, Sylva, ORAN, CAMARA

2023-24

- Edge, Open RAN
- Multi-Cloud
- AI Enabled Networking

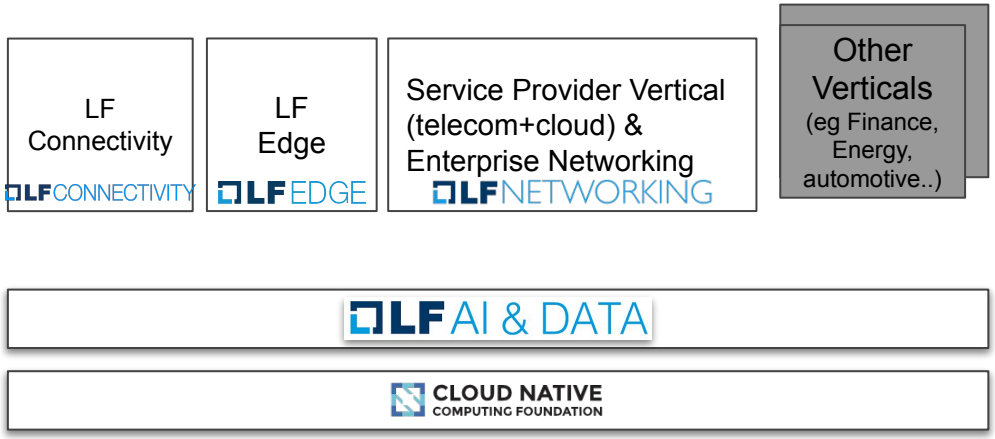
What do we focus on?

- There are many uses for AI in the context of Networking
  - Many of them are related to customer relationship management, technical support, marketing, etc. - this is not our focus
  - The above mentioned use cases might rely on data from the network
- The use cases that are relevant to our communities
  - **AI for Network** - Improve the Network - How AI can be used to optimize the Network and Edge (for any type of workload)?
  - **Network for AI** - New revenue streams - AI/LLM based services for Enterprise & Service Providers
- What are the requirements on LFN projects to “use” the AI/LLM disruption?
- Connecting/Educating Market on how LLM Technologies will shape work done so far in LFN (eg Intent based Networking, NLP, Closed Control, SONs, self-healing, Autonomous Networks…) - terms that have been used loosely in the past few years

Why AI is relevant to Networking?

- Because of the tremendous momentum
  - So many other industries find uses to it, there is probably something in it for us
  - Make developing network and edge technology cool again
- Because it can offer a fresh approach for old problems
  - Can it give us much needed breakthroughs for problems that have been bogging down our industry for years?
- It can create new business opportunities
- We can benefit from work done in other domains to accelerate networking innovation
  - Some of the work in other domains can surely be applied to networking

LF AI + Data and LFN built on each other



1. Vertical Specific Use cases & applications, data sharing governance (CDLA based), OSS/BSS/NMS integration with AI systems and Infrastructure
2. End to end solution testing and interop

1. LF AI & Data focuses on core Enabling Technologies in relation to ML, Models and Data
2. CNCF is a horizontal layer for all Cloud Native software specifically K8s irrespective of vertical integration

Notes:

1. OpenSSF is also another horizontal sub-foundation that helps with security across all umbrellas.
2. LF AI & Data focuses on any Horizontal (cross domain, cross vertical AI and Data open source software)
3. LF Connectivity and AI focuses on Access (RAN, Broadband, Satellite etc) alternative and enhanced access layer
4. LF Edge and AI includes Edge and IOT specific use cases and solutions specific to Edge Verticals like Manufacturing, Automotive, Industrials etc…

Generic AI Framework LF Networking + LF AI and Data

LFNETWORKING

1- Applications/AI Use Cases in Networking

The new functionality that is made available using AI

LF AI & Data

2- AI Models (Generic and **Domain** specific)

The AI capabilities, such as prediction, content generation, anomaly detection, etc. Generic or network specific

3- Data and AI infrastructure (computing elements) (Sharing, Governance, Processing)

How data is collected and stored. The resources used for processing, running and training the models

LFNETWORKING

4- Network Infrastructure (Open Source Projects + Vendor solutions)

The network itself and the data it provides and acts on the learnings from the above layers

Process to get AI going in Domains

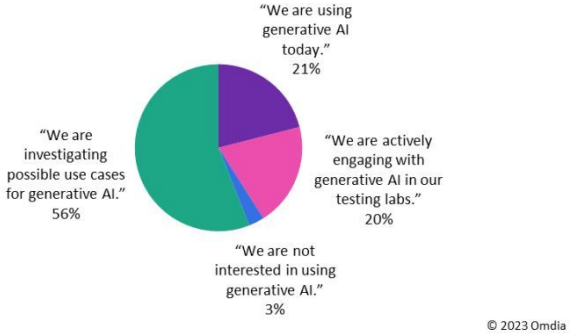
1. Focus on Domain specific Data
   a. LF will help with Data Governance, Processing and sharing
2. Sub foundations can create domain specific models AND use any generic external models as they need

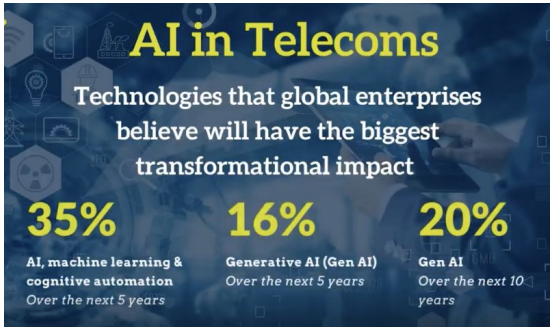
Telecom View - Current Status

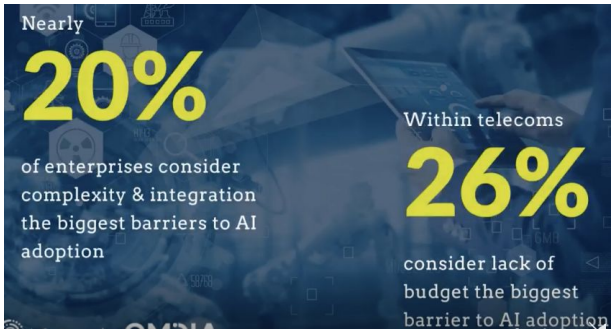
1. AT&T: Uses GenAI for document translation, network optimization, improving customer care effectiveness, HR query resolution, and meeting summaries.
2. Korea Telecom: Developed its own LLM based on OpenAI GPT-3 as a commercial service for enterprise agents.
3. Orange: Investigates GenAI for call transcription, summarizing customer service interactions, and suggesting agent follow-up actions.
4. SK Telecom: Launching a GenAI-based solution to enhance its customer service app.

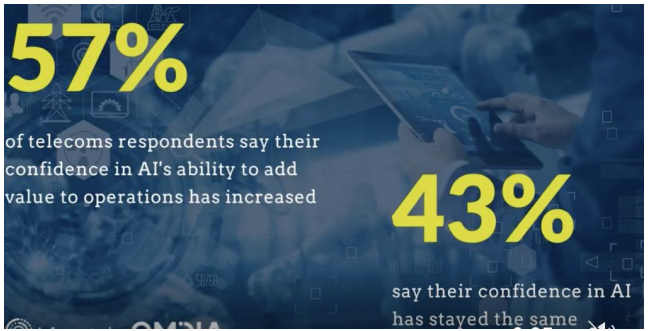
From policy driven control loops to Intent based to autonomous networks (AI enabled)

AI and Telecom View - Market Research









From policy driven control loops to Intent based to autonomous networks (AI enabled)

China Mobile & China Telecom: AI and Networking

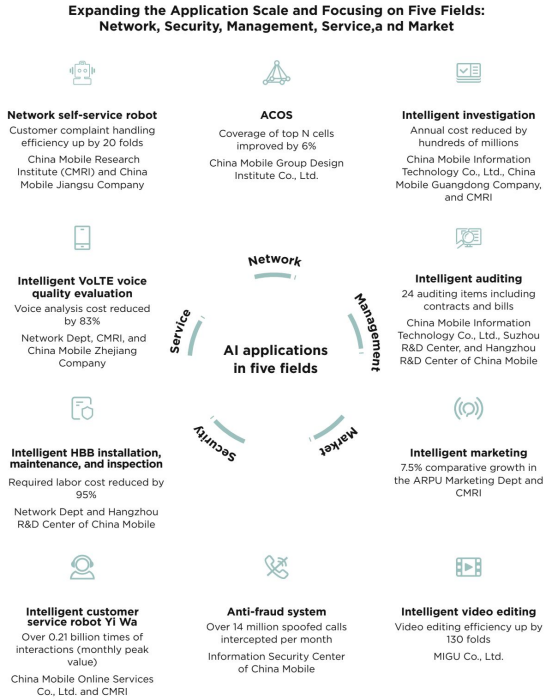


Figure 1-2 AI application planning of China Mobile

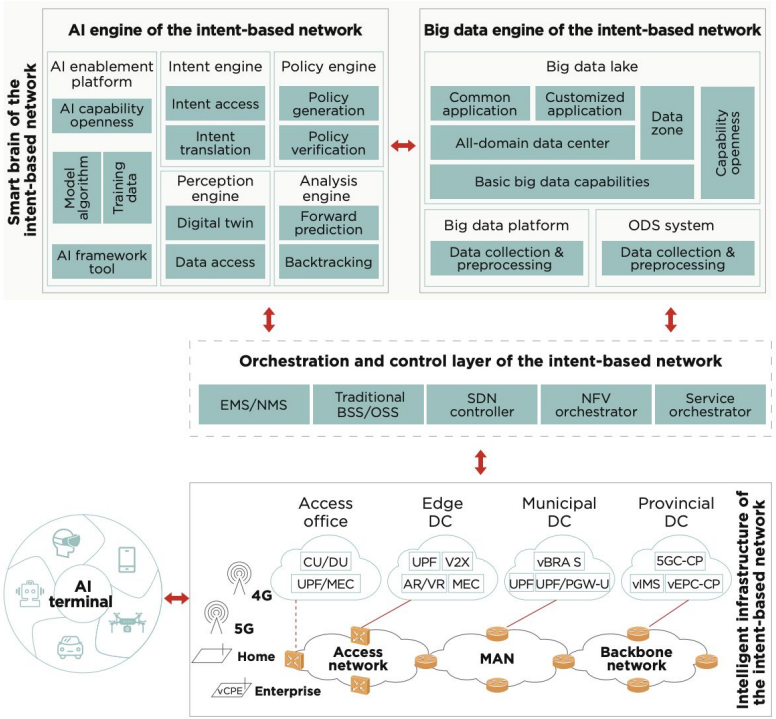
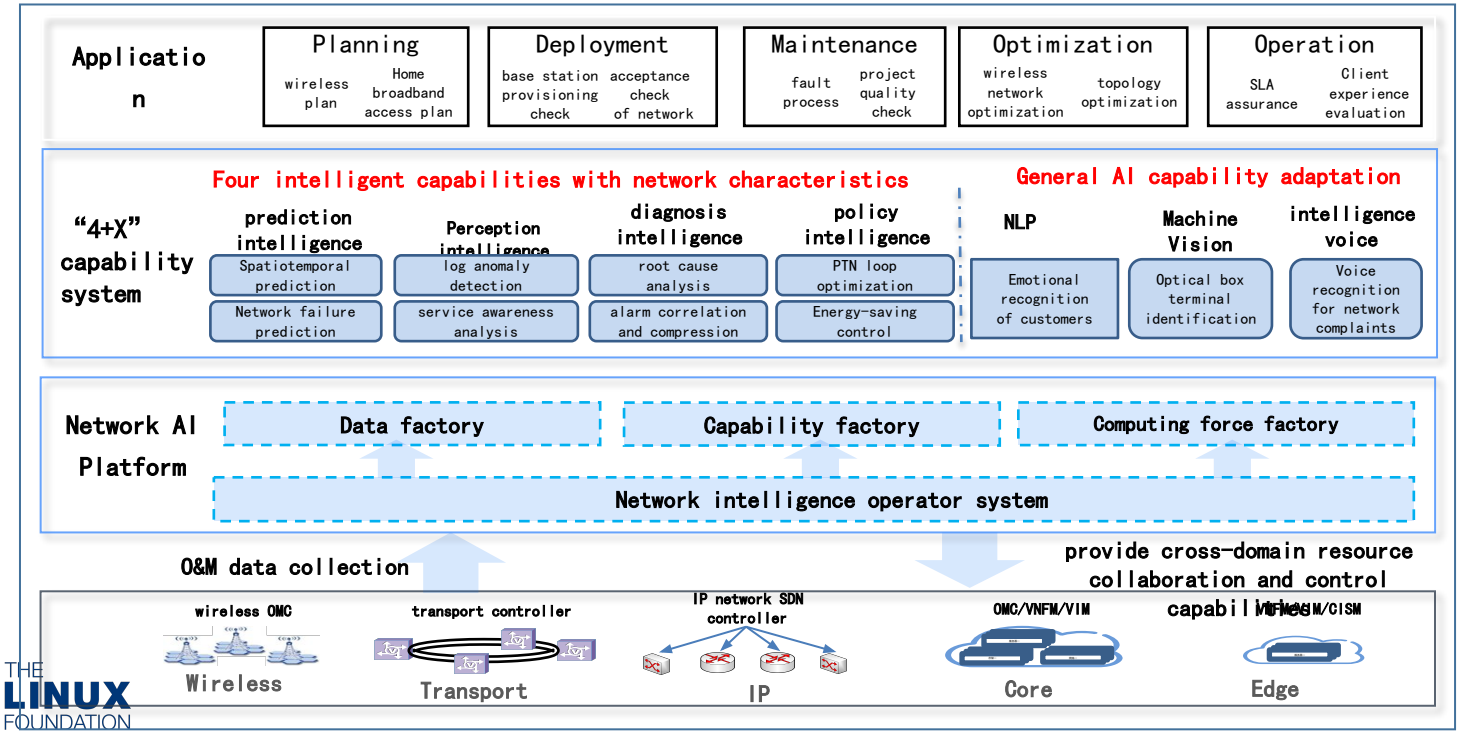


Figure 1-4 Target architecture of China Telecom's intent-based network

Source GSMA

China Mobile - AI is an integral part of Networking

- Build a solid foundation to provide AI computing power, AI capability, AI modelling to the Network Management System(NMS).
- Proposed and constructed the '4+X' network intelligence capability system, deployed more than 2600 capabilities. Multiple KPIs are promoted with more than 20% and 32 papers has been published.



AI capabilities
2600+
Cost reduction
3000+ person year
applications
2200+
daily call
3.8B+
Empowering Value
3Billon per year

10 Use cases suggested by the market

Use Case	Description	Relevance to LFN	New Project Potential	Mentions
1. Network planning and design	Using generative AI for Placement of small cells, MIMO antennas, beamforming, and backhaul connections	M	V	• Generative AI in Telco – Revamping 5G Network Planning and Optimization • The AI-native telco: Radical transformation to thrive in turbulent times | McKinsey
2. Predictive Maintenance	Using AI for equipment failure predictions	M		• Generative AI Solutions in Telecom Industry - Real Time AI Company • Potential Of Generative AI for Enterprises: Statistics, Use Cases, Top Business Examples
3. Automated closed loop	Network assurance using AI models trained on operational data	H		• Telcos, Stop Debating Generative AI and Just Get Going | Bain & Company • Microsoft dangles generative AI for telcos and slams 'DIY' clouds | Light Reading
4. Network AIOps	Use AIOps methodologies to automate and streamline network operations	H		• Amdocs Launches amAIz, a Cutting-Edge Enterprise-Grade Generative AI Framework • Omdia Telco Network and Service Automation Market Report
5. SON	Using AI based algorithms for Self Organizing Networks	H		• Generative AI Solutions in Telecom Industry - Real Time AI Company

THE LINUX FOUNDATION

- Improved Network Design/Operation (AI for Network)
- New revenue stream (Network for AI)

10 Use cases suggested by the market (cont'd)

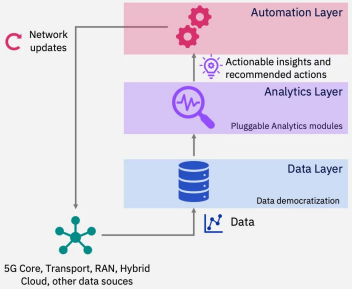
Use Case	Description	Relevance to LFN	New Project Potential	Mentions
6. Tech Assistant	In-field AI based tech assistant providing real-time guidance based on AI models	L	V	• Telcos, Stop Debating Generative AI and Just Get Going | Bain & Company
7. Code generation for network protocols	Co-pilot functionality for generating software implementation of network protocols specifications (3GPP, ETSI, Etc.)	L	V	• Microsoft dangles generative AI for telcos and slams 'DIY' clouds | Light Reading
8. GAN for network anomaly detection	using GANs is to build and test Network anomaly detection/prediction and traffic classification solutions	H	V	• The Thoth project
9. Traffic management	As utilization increases in a particular location or at a certain time of day, traffic can be re-routed to available capacity and resources.	H		• DataRobot - AI in Telecommunications
10. Edge LLM training/inference over SD-WAN	Leverage unused 5G RAN infrastructure resources for LLM training and inference. Bring training data to training infrastructure over SD-WAN. AI infrastructure may or may not be shared with 5G RAN(Nvidia/Softbank)	M	V	• NVIDIA Collaborates With SoftBank Corp. to Power SoftBank's Next-Gen Data Centers Using Grace Hopper Superchip for Generative AI and 5G/6G

- Improved Network Design/Operation (AI for Network)
- New revenue stream (Network for AI)

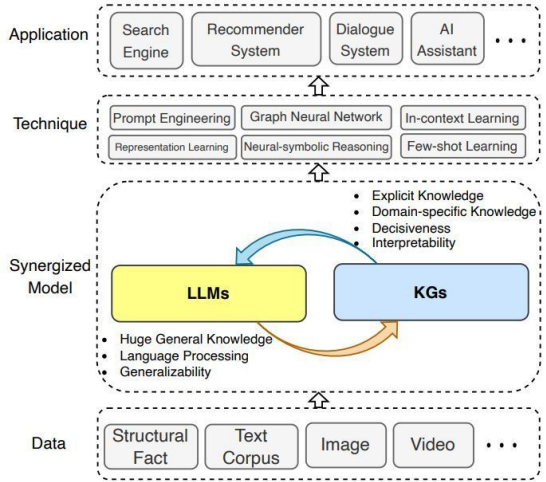
THE LINUX FOUNDATION

What has changed since previous attempts?

- Acumos (currently Archived) was a very early (maybe too early) look at telecom use cases enabled by AI (mostly in visual recognition eg Base station maintenance etc)
- Technology for domain specific AI has rapidly evolved recently:



source



source

LFN AI Board Questionnaire

Heavy participation and influence from Carriers, Google and Vendors

LFN AI Board Member Questionnaire

LF NETWORKING

(Action from July 19th Board Meeting to Gather your input, requirements, strategy and direction)
Please spare 10-15 min on this so that we can gather and summarize Board input efficiently

Dates

1. Questionnaire release : Jul 21, 2023
2. Request completed document sent back to Kenny Paul by **5 PM PST Aug 4, 2023**
   - Please add your name on the completed document.
3. Special SPC Review (if needed)
4. Discuss Findings in a special LFN governing board meeting Aug 23

Guiding questions framework

The rest of this document contains questions related to the following layers of technology related to AI in networking:

1- Applications/AI Use Cases in Networking
The new functionality that is made available using AI

2- AI Models (Generic and Domain specific)
The AI capabilities, such as prediction, content generation, anomaly detection, etc. Generic or network specific

3- Data and AI infrastructure (Sharing, Governance, processing)
How data is collected and stored. The resources used for processing, running and training the models

4- Network Infrastructure (Open Source Projects + Vendor solutions)
The network itself and the data it provides and acts on the learnings from the above layers

Applicable functional areas for LFN

1. Intent based networking

Organic growth of existing projects

Use cases addressed: #1, #3

- Use LLM to translate natural language intention to network configuration
- Prior work
  - https://wiot.northeastern.edu/news/northeastern-university-launches-fully-automated-and-virtualized-o-ran-private-5g-network-with-ai-automation/
  - ONAP Intent Based Networking - https://wiki.onap.org/display/DW/DEMOS+-+R10+Kohn+Demos?preview=%2F162104332%2F162105361%2FA+General+Implementation+with+Intent-based+network+in+ONAP.pdf
- Use a concept of co-pilot that is now common for source code generation
  - Input could be something like “Create a 5G Core configuration that prioritizes video streams to UEs”
  - The generated output will include all configurations needed in the UPF, SMF, network slices, etc.
- This could be a step towards fully autonomous networks, where the intent itself is auto-generated instead of being created by a human
- Challenge - What training data could be used and will it be publicly available?

2. Standards to code generation

Potential for new projects

Use cases addressed: #7

- Use LLM trained on specification documents to generate the open source implementation
- Prior work
  - OPS-5G TA1 work - Effigy and A5GARD - Archived, but could be used for reference
- Which open source LLM to use?
- Generic LLMs are not trained on relevant data sets. Some necessary training data may not be publicly available (E.g. 3GPP specs)
- Requires collaboration with GSMA, ETSI, other SDOs.
- Potential for announcing a “Standards/specification to code using AI/LLM” project
- Potential for creating a test framework based on test suite code auto-generated from specifications

3. E2E Network planning

Potential for new projects

Use cases addressed: #1, #3, #4, #9

- Placement of small cells, MIMO antennas, beamforming, and backhaul connections is a manual, error-prone and time consuming process. It can be significantly accelerated using generative AI.
- AI based network operations to respond to evolving usage patterns and changing conditions.
- Generate network configuration based on input of current and predicted performance data
- Use of “digital twins” that represent the wireless network and its topology
- Build models and train them on data of production RAN networks
    - Use the LF Connectivity MAVERIC project RDAP for model building

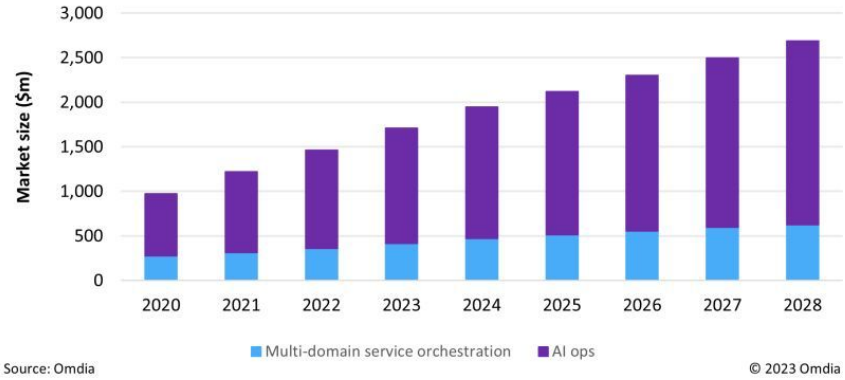
4. AIOps for Telco

Use cases addressed: #4

- AIOps for telecom networking is not a clearly defined product category. The term is often used to describe existing analytics and operational support systems (e.g., service assurance) that have been enhanced in some way through the use of AI. As such, vendors that stake a claim as providing AIOps solutions for telecom networking include companies such as Amdocs, Ericsson, MYCOM OSI, and Nokia.

Source: Omdia

Organic growth of existing projects



How can we address the challenges of observability with AI?

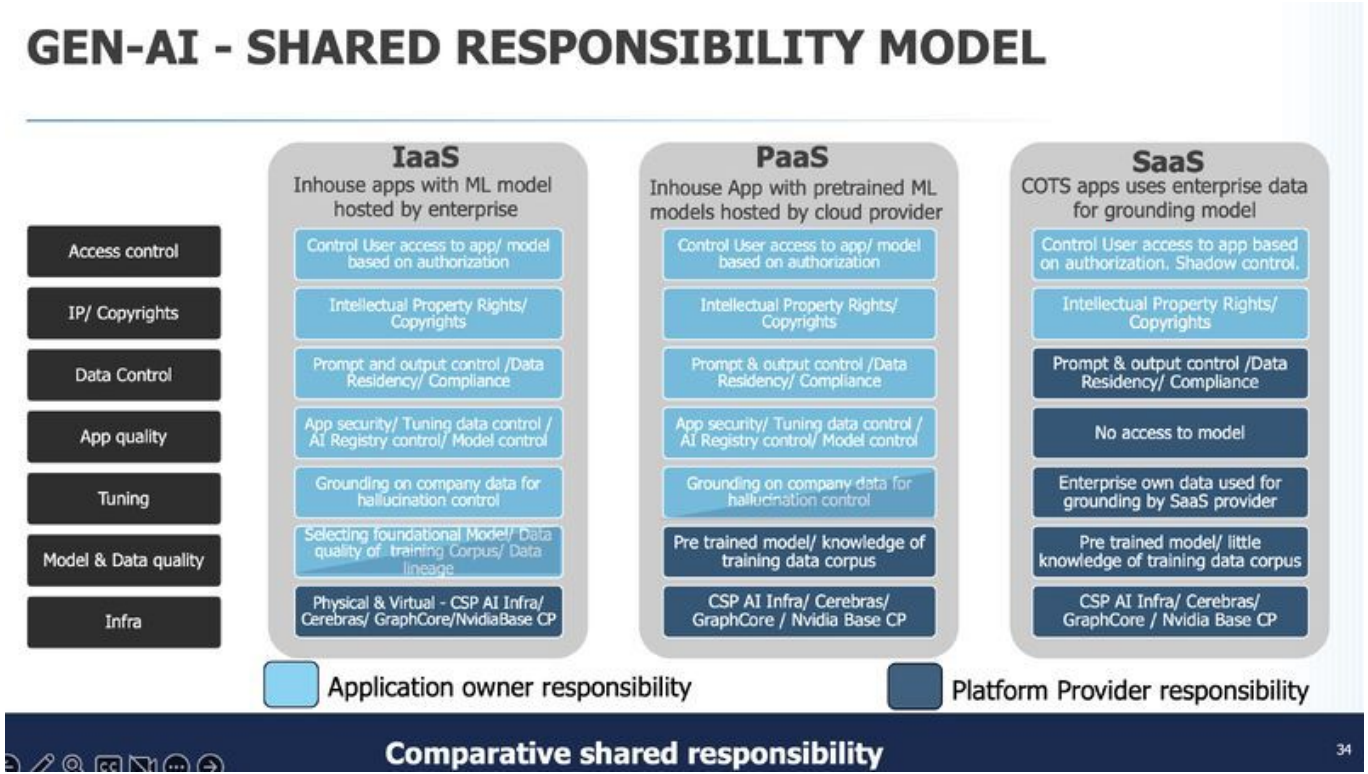
5. Edge LLM over SD-WAN

Potential for new blueprints

Use cases addressed: #10

- Training of LLMs using Enterprise owned data is becoming one of the most important use cases in generative AI
  - E.g. Azure openAI On Your Data
- However the current state of the art requires the Enterprise to upload their training data to the public cloud
  - This is not always desired, and raises concerns about data privacy and compliance
  - Inference is another concern, where prompt data is also required to be sent to the public cloud
- By running the training/inference on Edge infrastructure that is securely connected to the enterprise over SD-WAN, the privacy and compliance issues are addressed.
- Edge computing and SD-WAN providers are best positioned to offer this kind of service, as well as the traditional Tier-1 CSPs.
- Option for Edge resource sharing
  - Considering the bursty nature of AI LLM training workloads, the same Edge infrastructure could be optionally leveraged for other workloads such as O-RAN CNFs

5. Edge LLM over SD-WAN - continued



Source

How? Next steps

Updated post distribution

What should LFN do next?

- Clarify and Market LFN position for AI for Network and Network for AI messaging
- Blueprint to showcase 1-2 use cases based on expanding existing projects
  - Edge LLM over SD-WAN
- New projects to address use cases. Potential domains:
  - Network planning
  - Network protocol code generation
- Domain specific LLMs
  - Trained on public or private domain data sets
  - Created based on open source LLMs (Free Dolly, Future AI Data LLM)
  - Evaluate: Design the evaluation system for network-oriented-LLMs, carry out evaluation on various network operation and maintenance LLMs
- LangChain for network
  - Providing an advanced “LangChain+LLM” solution for network operation and maintenance to seamlessly integrate LLM with network management system
- Competition
  - Organize competition for network-oriented-LLM , invite developers and college students to participate, outstanding solutions can be contributed to LFN open source project

Dr Feng input

- AI for Network (cost and efficiency)
  - Ops, planning, different phase of managing network, autonomous network (Level 4 by 2025 with metrics) 70% AI (training using network data) and 30% Automation (LFN)
  - LFN can only provide frameworks/platforms for AI training (not data availability) - eg CMCC opened 7 data sets for research (Not proposing NetworkingGPT for all Telcos)
    - Possibly move projects/frameworks/datasets to LFN (with CDLA)
- Network + Cloud + AI (increase revenue)
  - Better integrate cloud and edge to serve AI computing for training and inference (increase revenue for community)
  - AI service in a package (big growth)
    - Including educating industry and enterprise
- Defining a layer of AI which is suitable for LFN community


THE
LINUX
FOUNDATION

From the press

Recent articles related to telco use of AI, with focus on uses for the network infrastructure

- Generative AI sparks telco industry debate | TMForum
  - Telefonica chairman warns about uncontrolled use of generative AI
  - Concerns about power usage and data governance
- Potential Of Generative AI for Enterprises: Statistics, Use Cases, Top Business Examples
  - Equipment failure prediction
  - Generative AI for subscriber chatbots and virtual assistant, Fraud detection and prevention
- Generative AI Solutions in Telecom Industry - Real Time AI Company
  - Network planning and design
  - Predictive Maintenance
  - SON
- Telcos, Stop Debating Generative AI and Just Get Going | Bain & Company
  - Short term - In-field tech assistant, Automated closed loop, Network assurance
  - Longer term - Real time tech assistant, Supply chain digital twin
- The AI-native telco: Radical transformation to thrive in turbulent times | McKinsey
  - Tactical recommendations for network buildout to maximize customer experience
- Microsoft dangles generative AI for telcos and slams 'DIY' clouds | Light Reading
  - Co-pilot for assisting in telco specific programming, using telco specific training data
  - Analysis of network data, cross-site network automation
- Generative AI tools like ChatGPT game changers for telecom industry - Channel Post MEA
  - Customer care
  - Technical support
- Amdocs Launches amAIz, a Cutting-Edge Enterprise-Grade Generative AI Framework, Set to Propel the Telecom Industry into the Gen AI era
  - Network provisioning
  - Zero touch operations
- Generative AI in Telco – Revamping 5G Network Planning and Optimization
  - 5G Network planning
- LLM data stack


THE
LINUX
FOUNDATION