

# 2024 White Paper for FINAL REVIEW - CLOSED FOR EDITING

Title: Intelligent Networking, AI and Machine Learning for Telecommunications Operators

Subtitle: Significant industry adoption progress, but challenges remain

Authored by Members of the Linux Foundation AI Taskforce

[Beth Cohen](#) Beth Cohen (Verizon)

[Lingli Deng](#) Lingli Deng (CMCC)

[Andrei Agapi](#) Andrei Agapi (Deutsche Telekom)

[Ranny Haiby](#) Ranny Haiby (Linux Foundation Networking)

[Sandeep Panesar](#)

[Muddasar Ahmed](#) Muddasar Ahmed (Mitre)

[LJ Illuzzi](#) LJ Illuzzi (Linux Foundation Networking)

[Hui Deng](#) Hui Deng (Huawei)

[ChangJin Wang](#) Chang Jin Wang (ZTE)

[Jason Hunt](#) Jason Hunt (IBM)

[Beth Cohen](#) can we use drivers for AI side bars using following :

Key drivers for Intelligent Networking:

- **Network Planning and Design:** Generative AI for precise small cell placement, MIMO antennas, beamforming, and optimized backhaul connections
- **Self-Organizing Networks (SON):** Harnessing AI-based algorithms for autonomous optimization and network resource management
- **Shared Infrastructure:** Leveraging 5G RAN infrastructure resources for training and inference, enhancing AI capabilities and network efficiency
- **Network Protocol Code Generation:** Enabling co-pilot functionality for generating network protocol software
- **Capacity Forecasting:** Utilizing AI to optimize network capacity to avoid unnecessary upgrades or poor network performance due to overloaded circuits

Key drivers for AI/ML for Operations:

- **Network AIOps:** Implementing AIOps methodologies to automate, streamline and improve overall network efficiency.
- **Predictive Maintenance:** Utilizing AI to forecast equipment failures and improve maintenance efficiency
- **Technical Assistant/Customer Service:** Real-time guidance from LLM trained tech assistants

Table of Contents:

- 1 [Executive Summary - Overview and Key Takeaways](#)
- 2 [Background and History](#)
- 3 [Challenges and Opportunities](#)
  - 3.1 [Common Telecom Pain Points](#)
  - 3.2 [Challenges to Achieving Full Autonomy](#)
  - 3.3 [Emerging Opportunities](#)
- 4 [Brief Overview of AI Research in Context](#)
- 5 [Intelligent Networking Differences](#)
- 6 [Projects and Research](#)
  - 6.1 [3GPP Intelligent Radio Access Network \(RAN\)](#)
  - 6.2 [Mobile Core Network Transformation](#)
  - 6.3 [Thoth project - Telco Data Anonymizer](#)
- 7 [How could Open Source Help?](#)
  - 7.1 [Related Open Source Landscape](#)
  - 7.2 [Common Vision: Intelligence plane for XG Networks](#)
- 8 [Call for Action](#)
- 9 [Final Thoughts](#)
- 10 [References](#)

## Executive Summary - Overview and Key Takeaways

0.25 page – [Beth Cohen](#) will fill in when the paper is mostly complete.

Since LFN published the seminal [Intelligent Networking, AI and Machine Learning White Paper](#) in 2021, the telecom industry has seen tremendous growth in both interest and adoption of Artificial Intelligence and Machine Learning (AI/ML) technologies. While it is still early days, the industry is now well past the tire kicking, lab testing phases that was then the state of the art. Intelligent networking is coming into its own as telecoms increasingly use it for operational support; whether that means deploying intelligence into their next generation networks, or for automation of network management tasks such as ticket correlation and predictive maintenance. LFN and Open Source have a pivotal role to play in fostering and developing intelligent networking technologies through the continued support of key projects, ranging from building a common understanding of the [underlying data models](#) to [developing infrastructure models](#) and [integration blueprints](#).

The future of Intelligent Networking and AI is in the hands of the individuals and organizations who are willing and able to contribute to new and existing projects and initiatives. If you are involved in building and operating networks, developing network technology or consuming network services, consider getting involved. Engaging with the LFN projects and communities can be an educational and rewarding way to shape the future of Intelligent Networking.

## Key Takeaways

- Intelligent networking is rapidly moving out of the lab and being deployed directly into production
- Operational maintenance, and service assurance are still a priority, but there is increasing interest in using AI/ML to drive network optimization and efficiency
- More research and development is needed to establish industry wide best practices and a shared understanding of intelligent networking to support interoperability.
- There has been some work on developing common or shared data sources and standards, but it remains a challenge
- LFN and the Open Source community are key contributors to furthering the development of intelligent networking now and in the future

## Background and History

Beth Cohen

*0.5 page*

At their core, telecoms are technology companies driven by the need to scale their networks to service millions of users, reliably, transparently and efficiently. To achieve these ambitious goals, they need to optimize their networks by incorporating the latest technologies to feed the connected world's insatiable appetite for ever more bandwidth. To do this efficiently, the networks themselves need to become more intelligent. At the end of 2021, a bit over 2.5 years ago, LFN published its first [white paper on the state of intelligent networking in the telecom industry](#). Based on a survey of over 70 of its telecom community members, the findings pointed to a still nascent field made up of mostly research projects and lab experiments, with a few operational deployments related to automation and faster ticket resolutions. The survey did highlight the keen interest its respondents had in intelligent networking, machine learning and its promise for the future of the telecom industry in general.

Fast forward to 2023 when at the request of the LFN Governing Board and Strategic Planning Committee, the LFN AI Taskforce was created to coordinate and focus the efforts that were already starting to bubble up in both new project initiatives (Nephio, Network Super Blueprints, just to name a few) and within existing projects (Anuket, ONAP). The Taskforce was given the charter by the Governing Board to look at and make recommendations on what direction LFN should take with this exciting emerging field of research and technology. Some of the areas that the Taskforce looked into included:

- How to create and maintain public Networking data sets for research and development of AI applications? (Ranked #1 in GB member survey)
- What are some feasible goals (short term) in creation of AI powered Network Operations technologies
- Evaluate the existing Networking AI assets coming from member company contributions
- Analyze generic base AI models and recommend the creation of Network specific base models (Ranked high in GB member survey)
- Recommended approaches and the potential for open source projects to contribute to the next generation of intelligent networking tools.

## Challenges and Opportunities

Beth Cohen

*1 page - Focus on Telco pain points only*

Ironically, as generative AI and LLM adoption becomes more widespread in many industries, telecom has lagged somewhat due to a number of valid factors. As was covered in the [previous white paper](#), the overall industry challenges remain the same, that is the constant pressure to increase the efficiency and capacity of operators' infrastructures to deliver more services to customers for lower operational costs. The complexity and lack of a common standard understanding of network traffic data remains a barrier for the industry to speed the adoption of AI/ML to optimize network service delivery. Some of the challenges that are motivating continued research and adoption of intelligent networking in the industry include:

### Common Telecom Pain Points

- **Operational Efficiency:** The continuing need to reduce costs and errors, potentially increase margins
- **Network Automation:** Right-sizing network hardware and software, optimizing location placement
- **Availability:** Identifying single points of failure in systems to improve equipment maintenance efficiency
- **Capacity Planning:** Avoid unnecessary upgrades or poor network performance from overloaded nodes.

## Challenges to Achieving Full Autonomy

*0.5 page* [Lingli Deng](#) [Andrei Agapi](#)

Deploying large models such as LLMs in production, especially at scale, also raises several other issues: 1) Performance, scalability and cost of inference, especially when using large context windows (most transformers scale poorly with context size), 2) Deployment of models in the cloud, on premise, multi-cloud, or hybrid, 3) Issues pertaining to privacy and security of the data, 4) Issues common to many other ML/AI applications, such as ML-Ops, continuous model validation and the need for continuous and costly re-training.

- **Need for High Quality Structured Data:** Telecommunications networks are very different from other AI human-computer data sets, in that a large number of interactions between systems use structured data. However, due to the complexity of network systems and differences in vendor implementations, the degree of standardization of these structured data is currently very low, causing "information islands" that cannot be uniformly "interpreted" across the systems. There is a lack of "standard bridge" to establish correlation between them, and it cannot be used as effective input to incubate modern data-driven AI/ML applications.
- **AI Trustworthiness:** In order to meet carrier-level reliability requirements, network operation management needs to be strict, precise, and prudent. Although operators have introduced various automation methods in the process of building autonomous networks, organizations with real people are still responsible for ensuring the quality of communication and network services. In other words, AI is still assisting people and is not advanced enough to replace people. Not only because the loss caused by the AI algorithm itself cannot be defined as the responsible party (the developer or the user of the algorithm), but also because the deep learning models based on the AI/ML algorithms themselves are based on mathematical statistical characteristics, resulting in sometimes erratic behavior, meaning the credibility of the results can be difficult to determine.
- **Uneconomical Margin Costs:** According to the analysis and research in the [Intelligent Networking, AI and Machine Learning While Paper](#), there are a large number of potential network AI technology application scenarios, but to independently build a customized data-driven AI/ML model for each specific scenario, is uneconomical and hence unsustainable, both for research and operations. Determining how to build an effective business model between basic computing power providers, general AI/ML capability providers and algorithm application consumers is an essential prerequisite for its effective application in the field.
- **Unsupportable Research Models:** Compared with traditional data-driven dedicated AI/ML models in specific scenarios, the research and operation of large language models have higher requirements for pre-training data scales, training computing power cluster scale, fine-tuning engineering and other resource requirements, energy consumption, management and maintenance, etc. It is difficult to determine if it is possible to build a shared research and development application, operation and maintenance model for the telecommunications industry so that it can become a stepping stone for operators to realize useful autonomous networks.
- **Contextual Data Sets:** Another hurdle that is often overlooked is the need for the networking data sets to be understood in context. What that means is that networks need to work with all the layers of the IT stack, including but not limited to:
  - **Applications:** Making sure that customer applications perform as expected with the underlying network
  - **Security:** More important than ever as attack vector expand and customers expect the networks to be protected
  - **Interoperability:** The data sets must support transparent interoperability with other operators, cloud providers and other systems in the telecom ecosystem
  - **OSS/BSS Systems:** The operational and business applications that support network services
- **Community Unity and Standards:** Although equipment manufacturers can provide many domain AI solutions for professional networks/single-point equipment, these solutions are limited in "field of view" and cannot solve problems that require a "global view" such as end-to-end service quality assurance and rapid response to faults. Operators need to aggregate management and maintenance data in various network domains by building a unified data sharing platform, and based on this, further provide a unified computing resource pool, basic AI algorithms and inference platform (i.e. cross-domain AI platform) for scenario-specific AI, end-to-end scenarios, and intra-domain scenarios using an applied reasoning platform.

## Emerging Opportunities

Sandeep Panesar

The Telecoms have been working on converged infrastructure for a while. Voice over IP has long been an industry standard, but there is far more than can be done to drive even more efficiencies with network and infrastructure convergence.

**Converged infrastructure** is needed to support the growth and sustainability of AI. This is particularly important in the need for a single solution designed to integrate compute, storage, networking, and virtualization. Data volumes are trending to grow beyond hyperscale data, and with such massive data processing requirements the ability to execute is critical. The demands on existing infrastructure are already heavy, so bringing everything together to work in concert will be key, to maintain and support the growth in demand for resources. In order to do that the components will need to work together efficiently and the network will play an important role in linking specialized hardware accelerators like GPUs (Graphics Processing Units) or TPUs (Tensor Processing Units) to accelerate AI workloads beyond current capability. Converged infrastructure solutions will lead to the ability to deploy AI models faster, iterate them more efficiently, and extract insights faster. This will pave the way for the next generation of AI.

**Converged Service and integrated solutions** that combine AI with traditional services have the potential to deliver enhanced services to end customers, but more importantly these services need to leverage AI-driven insights, automation, and personalization to optimize user experience, improve efficiency, and drive innovation across industries. There are many existing industry use cases for this already, which include healthcare, legal, retail, telecommunications, networking, and incident tracking. The analytics delivered by a converged service provide automated insights and tools that can provide real-time analysis, tracking, and remediation/response—either with or without human intervention.

**Business Innovation - New Revenue Streams:** Data monetization encompasses various strategies, including selling raw data, offering data analytics services, and developing data-driven products or solutions to customers. Organizations can monetize their data by identifying valuable insights, patterns, or trends hidden within their datasets that no group of human resources can identify quickly. These insights can then be used to create new products and services that will better serve customers and organizations. This is a new strategic business opportunity for organizations looking to monetize anonymous data and leverage it for increased business efficiency and determine product direction and new go to market strategies.

**Data Privacy and Security:** The ability to monetize the data comes with a big caveat, which is that the use of customer data must be handled with care to ensure data privacy, security and regulatory compliance. This requires clear policies and security procedures to ensure anonymity, safety and privacy *at all times*. The good news is that AI can be used to address the growing threat of network vulnerabilities, zero day exploits and other security related issues with predictive analytics and threat analysis.

**Data Model Simplification:** Large language models can be used to understand large amounts of unstructured operation and maintenance data (for example, system logs, operation and maintenance work orders, operation guides, company documents, etc., which are traditionally used in human-computer interaction or human-to-human collaboration scenarios), from which effective knowledge is extracted to guide further automatic/intelligent operation and maintenance, thereby effectively expanding the scope of the application of autonomous mechanism.

# Brief Overview of AI Research in Context

Lingli Deng Andrei Agapi Sandeep Panesar

## LLM (Large Language Models)

This breakthrough in AI research is characterized by vast amounts of easily accessible data, extensive training models, and the ability to generate human-like text, in almost an instant. These models are trained on enormous datasets from sources particular to a given area of research. LLMs have changed the way natural language processing tasks are interpreted. Some of the areas that have been particularly fruitful including: text generation, language translation, summarization, automated chatbots, image and video generation, and routine query/search results.

## Generative AI (Gen AI)

Gen AI is a much broader category of artificial intelligence systems capable of generating new content, ideas, or solutions autonomously based on a human text, video, image or sound-based input. This includes LLMs as resource data for content generation. As such, Gen AI seems to be able to produce human-like creative content in a fraction of the time. Content creation for websites, images, videos, and music are a few of the capabilities of Gen AI. The rise of Gen AI has inspired numerous opportunities for business cases, from creating corporate logos to corporate videos to saleable products to end-consumers and businesses to creating visual network maps based on the datasets being accessed. Further, even being able to provide optimized maps for implementation to improve networking either autonomous or with human intervention are useful areas for further exploration.

The two combined open the question as to what should Gen AI be used for, and more importantly how is it distinguishable from human work. Many regulatory bodies are looking at solutions around the identification of decisions and what content has been generated to solve a particular problem and solution. The foundation of this combination is to ensure security and safety, mitigate biases, and identify which behaviors were illustrated and acted upon by Gen AI, and which were not.

The advent of transformer models and attention mechanisms and the sudden popularity of ChatGPT and other LLMs, transfer learning, and foundation models in the NLP domain have all sparked vivid discussions and efforts to apply generative models in many other domains. Interestingly, let us not forget that word embeddings, sequence models such as LSTMs and GRUs, attention mechanisms, transformer models and pre trained LLMs have been around long before the launch of the ChatGPT tool in late 2022. Pretrained transformers like BERT in particular (especially transformer-encoder models) were widely used in NLP for tasks like sentiment analysis, text classification, extractive question answering. long before ChatGPT made chatbots and decoder-based generative models go viral. That said, there has clearly been a spectacular explosion of academic research, commercial activity, and ecosystems that have emerged since ChatGPT was launched to the public, in the area of both open and closed source LLM foundation models, related software, services and training datasets.

Beyond typical chatbot-style applications, LLMs have been extended to generate code, solve math problems (stated either formally or informally), pass science exams, or act as incipient agents for different tasks, including advising on investment strategies, or setting up a small business. Recent advancements to the basic LLM text generation model include instruction finetuning, retrieval augmented generation using external vector stores, using external tools such as web search, external knowledge databases or other APIs for grounding models, code interpreters, calculators and formal reasoning tools. Beyond LLMs and NLP, transformers have also been used to handle non-textual data, such as images, sound, and arbitrary sequence data.

# Intelligent Networking Differences

Andrei Agapi

A natural question arises on how the power of LLMs can be harnessed for problems and applications related to Intelligent Networking, network automation and for operating and optimizing telecommunication networks in general, at any level of the network stack. Datasets in telco-related applications have a few particularities unique to the industry. For one, the data one might encounter ranges from fully structured (e.g. code, scripts, configuration, or time series KPIs), to semi-structured (syslogs, design templates etc.), to unstructured data (design documents and specifications, Wikis, Github issues, emails, chatbot conversations).

Another issue is domain adaptation. Language encountered in telco datasets can be very domain specific (including CLI commands and CLI output, formatted text, network slang and abbreviations, syslogs, RFC language, network device specifications etc.). Off-the-shelf performance of LLM models strongly depends on whether those LLMs have actually seen that particular type of data during training (this is true for both generative LLMs and embedding models). There are several approaches to achieve domain adaptation and downstream task adaptation of LLM models. In general these either rely on:

1. In-context-learning, prompting and retrieval augmentation techniques;
2. Fine tuning the models
3. Hybrid approaches.

For fine tuning LLMs, unlike for regular neural network models, several specialized techniques exist in the general area of PEFT (Parameter Efficient Fine Tuning), allowing one to only fine tune a very small percentage of the many billions of parameters of a typical LLM. In general, the best techniques to achieve domain adaptation for an LLM heavily depends on:

1. The type of data and how much domain data is available,
2. The specific downstream task
3. The initial foundation model

In addition to general domain adaptation, many telcos will have the issue of multilingual datasets, where a mix of languages (typically English + something else) will exist in the data (syslogs, wikis, tickets, chat conversations etc.). While many options exist for both generative LLMs [11] and text embedding models [12], not many foundation models have seen enough non-English data in training, thus options in foundation model choice are somewhat restricted for operators working on non-English data. A solution to work around this issue is to use automated translation and language detection models on the data as a preprocessing step.

Beyond these approaches, an emerging technique for pre-training foundation models on network data holds potential promise. With this technique, network data is essentially turned into a language via pre-processing and tokenization, which can then be used for pre-training a new "network foundation model." [34] Initial research has successfully demonstrated this approach on Domain Name Service (DNS) data [35] and geospatial data [36]. As this area of research matures, it could allow for general purpose network foundational models that can be fine-tuned to answer a variety of questions around network data or configurations without having to train individual models for bespoke network management tasks.

In conclusion, while foundation models and transfer learning have been shown to work very well on general human language when pre-training is done on large corpuses of human text (such as Wikipedia, or the Pile [29]), it remains an open question to be answered whether domain adaptation and downstream task adaptation work equally well on the kinds of domain-specific, semi-structured, mixed modality datasets we find in the telecom industry. To enable this, telecoms should focus on standardization and data governance efforts, such as standardized and unified data collection policies and developing high quality structured data across a common definition and understanding.

## Projects and Research

1.5 page

### 3GPP Intelligent Radio Access Network (RAN)

ChangJin Wang

The intelligent evolution of wireless access networks is in a phase of rapid evolution and continuous innovation. In June 2022, 3GPP announced the freezing of R17 and described the process diagram of an intelligent RAN in TR37.817, including data collection, model training, model inference, and execution modules, which together form the infrastructure of an intelligent RAN. This promotes the rapid implementation and deployment of 5G RAN intelligence and provides support for intelligent scenarios such as energy saving, load balancing, and mobility optimization.

- **AI and Machine Learning Drive 5G RAN Intelligence**

Artificial intelligence and machine learning technologies are playing an increasingly important role in 5G RAN intelligence. The application of these technologies enables the network to learn autonomously, self-optimize, and self-repair, thereby improving network stability, reliability, and performance. For example, by using machine learning algorithms to predict and schedule network traffic, more efficient resource allocation and load balancing can be achieved. By leveraging AI technologies for automatic network fault detection and repair, operation and maintenance costs can be greatly reduced while improving user experience. The intelligence of 5G wireless access networks also provides broad space for various vertical industry applications. For instance, in intelligent manufacturing, 5G can enable real-time communication and data transmission between devices, improving production efficiency and product quality. In smart cities, 5G can provide high-definition video surveillance, intelligent transportation management, and other services to enhance urban governance. Additionally, 5G has played a significant role in remote healthcare, online education, and other fields.

- **Challenges Facing 5G RAN Intelligence Industrialization**

However, despite the remarkable progress made in the 5G wireless access network intelligence industry, some challenges and issues remain to be addressed. For example, network security and data privacy protection are pressing issues that require effective measures to be implemented. The energy consumption issue of 5G networks also needs attention, necessitating technological innovations and energy-saving measures to reduce energy consumption. In the future, continuous efforts should be made in technological innovation, market application, and other aspects to promote the sustainable and healthy development of the 5G wireless access network intelligence industry.

### Mobile Core Network Transformation

Hui Deng ChangJin Wang

Mobile core networks can be thought of as the brains of mobile communication. In recent years, these networks have experienced a huge transformation from legacy proprietary hardware to telecom cloud native systems. Today, the majority of mobile core networks are deployed based on telco cloud architectures supported by NFV technologies. Intelligent networking most benefits the packet core networks such as 5GC and UPF which are responsible for packet forwarding, IMS which support delivery of multimedia communications such as voice, message and video, and operational functions that manage the core network itself including telco cloud infrastructure and 5G network applications. There have been three areas where intelligent networking have shown benefits:

- **Network Intelligence Enables Experience Monetization and Differentiated Operations**

For a long time, operators have strived to realize traffic monetization on MBB(Mobile Broadband) networks. However, there are three technical gaps: non-assessable user experience, limited or no dynamic optimization, and no-closed-loop operations. To bridge these gaps, there is a need for an Intelligent Personalized Experience solution, designed to help operators add experience privileges to service packages and better monetize differentiated experiences. Typically in the industry, the user plane on the mobile core network processes and forwards one service flow using one vCPU. As heavy-traffic services increase, such as 2K or 4K HD video and live streaming, microbursts and extremely large network flows become the norm, it becomes more likely that a vCPU will become overloaded, causing packet loss. To address this issue, Intelligent AI supported 5G core networks need to be able to deliver ubiquitous 10 Gbps superior experiences.

- **Service Intelligence Expands the Profitability of Calling Services**

In 2023, New Calling was put into commercial use, based on a 3GPP specification, it can enhance intelligence and data channel (3GPP specification)-based interaction capabilities; it enables users to use multi-modal communications, and helps operators construct more efficient service layouts. In addition, the 3GPP architecture allows users to control digital avatars through voice during calls, delivering a more personalized calling experience. One example where this can be seen as a business opportunity might be an enterprise using the framework to customize their own enterprise ambassador avatar to promote their brand.

- **O&M Intelligence Achieves High Network Stability and Efficiency**



Empowered by the multi-modal large model, the Digital Assistant & Digital Expert (DAE) based AI technology could reduce O&M workload and improve O&M efficiency. It can reshape cloud-based O&M from "experts+tools" to intelligence-centric "DAE+manual assistance". Using the DAE, it is possible that up to 80% of telecommunication operator trouble tickets can be automatically processed, which is far more efficient than the manual processing it is for the most part today. DAE also enables intent-driven O&M, avoiding manual decision-making. Before, it commonly took over five years to train experts in a single domain, with the multi-modal large model it is now possible for it to be trained and updated in mere weeks.

## Thoth project - Telco Data Anonymizer

Sandeep Panesar Beth Cohen

The [Thoth project](#), which is a sub project under the [Anuket infrastructure project](#), has recently focused on a major challenge to the adoption of intelligent networks, the lack of a common data set or an agreement on a common understanding of the data set that is needed. AI has the potential for creating value in terms of enhanced workload availability and improved performance and efficiency for NFV use cases. Thoth's work aims to build **machine-Learning models** and **Tools** that can be used by Telecom operators (typically by the operations team). Each of these models is designed to solve a single problem within a particular category. For example, the first category we have chosen is Failure prediction, and the project plans to create 6 models - failure prediction of VMs, Containers, Nodes, Network-Links, Applications, and middleware services. This project also will work on defining a set of **data models** for each of the decision-making problems, that will help both providers and consumers of the data to collaborate.

## How could Open Source Help?

Ranny Haiby

2-3 pages

It is important to understand the current landscape of Open Source Software (OSS) projects and initiatives and how they came into existence to see how OSS is well positioned to address the challenges of Intelligent Networking. Several such initiatives have already laid down the groundwork for building Network AI solutions, or are actively in the process of creating them. Building on these foundations, it is easy to envision the critical role open source software can play in unleashing the power of AI for the future generations of networking. Some of the required technologies are unique to the Networking industry, and will have to be addressed by existing OSS projects on the landscape, or by the creation of additional ones. Some pieces of the technologies are more generic, and will need to come from the broader community of AI OSS. Here is a rough outline of the different layers of Networking AI and the sources for the required technology:

[blocked URL](#)

In addition, successful development of AI models for use in Networking relies on the availability of data that could be shared under a common license. The Linux Foundation created the [Community Data License Agreement \(CDLA\)](#) for this purpose. Using this license, end users can share data and make it available for researchers, who in turn can further develop the necessary models and applications that benefit the telecom ecosystems and ultimately the end users.

## Related Open Source Landscape

1-2 pages

The legal and technical framework for sharing research is important, but what is really essential for driving innovations is the ability for Open Source to provide a forum for creating communities with shared purpose. These communities by allowing people from various companies, diverse skill levels and different cultural perspectives to collaborate, have the potential to spark real innovation in ways that are not really possible in any other context.

- Network Communities

Open Source Software (OSS) communities have been successfully building projects that provide the building blocks for commercial networks for over a decade. OSS projects provide the underlying technology for all layers of the network, including the data/forwarding plane, control plane, management and orchestration. A vibrant ecosystem of contributing companies exists around these projects, consisting of organizations that realize the value of OSS for speeding the development of networking technology:

- Shared effort to develop the foundation layers of the technology, freeing up more resources to develop the value-added layers
- A neutral platform for innovation where individuals and organizations can exchange ideas and develop best of breed technology solutions
- An opportunity to demonstrate thought leadership and domain expertise
- A collaboration space where producers and consumers of technology can freely interact to create new business opportunities

Many of the networking OSS projects are hosted by the Linux Foundation. Most of them are part of [LF Networking](#), [LF Connectivity](#) and [LF Broadband](#).

- AI Communities

The same principles are now being applied to the shared development of Network AI technologies, where the open source community fosters innovation and potentially stimulates business growth. AI innovation has been strongly propelled by OSS projects that were initiated following the same principles of open collaboration. It is hard to imagine doing any modern AI development without heavily relying on OSS. OSS AI and ML projects range from anything from the frameworks for development to libraries and programming tools. OSS work means that data scientists who develop domain specific models such as networking and telecommunications, can focus on innovation by leveraging OSS projects to jump start their work. It would be almost impossible to mention all the relevant OSS AI projects here as there are already so many, and the list only keeps growing. The Linux Foundation AI & Data maintains a useful dynamic landscape [here](#).

- Open AI Models

There is a lot of debate on the definition of what an "Open AI model" really means. While it is out of the scope of this paper to try and settle any of those debates, it is obvious that there is a clear need to create the definition of "open LLM". The sooner such definitions are created and blessed by the industry, the faster innovation can happen.

In the area of open source LLMs, both with respect to generative models [11] as well as more specialized, discriminative models such as text classifiers, QA, summarization, and text embedding models [12], has been particularly vibrant and have evolved rapidly over the past 5 years. A number of global platforms are being widely used for sharing open models, code, datasets and accompanying research papers, have been particularly instrumental in democratizing access to cutting edge technologies and fostering an environment of global collaboration. Among these platforms, Huggingface [30] has played a particularly pivotal role. At the time of this writing, HuggingFace hosts [31] over 350K models, 75K datasets and 150K demo apps (Spaces), in more than 100 languages. It also maintains Transformers, a popular open source library that facilitates integrating, modifying and performing downstream task adaptation for thousands of foundation models from this vast repository. It also provides the Datasets library, as well as several widely used benchmarks and leaderboards [11][12] that are very instrumental for researchers and developers implementing LLM solutions. Other important platforms used by the AI/ML open source community in general (not necessarily LLM-focused) are Kaggle [32] (used for public datasets and high profile ML competitions in all areas) and Paperswithcode [33] (this platform links academic research papers to their respective code and implementation, as well as providing benchmarks and leaderboards comparing different competing solutions for a wide area of ML tasks).

- Integration blueprints

In recent years The Linux Foundation Networking launched a set of Open Source Networking "[Super Blueprints](#)" that outline architectures for common networking use cases and are built using OSS technology. Those blueprints consist of collections of OSS projects and commercial products, integrated together by the open source community and documented for free use by any interested party. Several of these blueprints started incorporating AI technologies and that trend is expected to accelerate, as existing blueprints inspire additional AI-driven solutions. One promising area is intent-based network automation. There is current work on blueprints that use NLP and LLM to translate user intent, expressed in natural spoken language to network requirements, and generate full network configuration based on those requirements. Such an approach can significantly simplify existing network operation processes, and enable new automated services that can be directly consumed, commonly known as Network-as-a-Service (NaaS).

- Verification programs

Experience with OSS in other domains shows that whenever there is an OSS technology that powers commercial products or offerings, there is a need to validate the products to make sure they are properly using the OSS technology and are ready to use in a predictable manner. Such validation/verification programs have existed as part of OSS ecosystems for a while. The Cloud Native Computing Foundation (CNCF) has a successful "Certified Kubernetes" program that helps vendors and end users ensure that Kubernetes distributions provide all the necessary APIs and functionality. A similar approach needs to be applied to any OSS Networking AI projects. Users should have a certain level of confidence, knowing that the OSS based AI Networking solution they use will behave as expected.

## Common Vision: Intelligence plane for XG Networks

[Ranny Haiby Muddasar Ahmed](#)

*1 page*

In the dynamic realm of communication and networking technologies, the fusion of artificial intelligence (AI) with networks promises to redefine connectivity, ushering in an era of unprecedented intelligence, efficiency, and adaptability. As we embark on the journey towards 6G and embrace the vision outlined by the International Telecommunication Union (ITU) for IMT-2030, it becomes clear that AI can and will play a pivotal role in reshaping network operations.

At the heart of this transformation lies the concept of the intelligence plane, where AI-driven systems leverage natural intent interaction to seamlessly bridge the gap between users and networks. Smart orchestration ensures optimal resource allocation and dynamic adaptation to meet evolving demands, thereby enhancing network performance and user experience. Real-time meta network verification, powered by AI algorithms, ensures continuous monitoring and validation of network behavior, preempting potential issues and maintaining operational integrity. Additionally, built-in knowledge open-loops enable networks to autonomously learn and adapt, fostering resilience and responsiveness.

The concept of ubiquitous intelligence envisaged by IMT-2030 underscores the pervasive presence of AI across the telecommunications community. From autonomous network management to context-aware devices, AI will touch every aspect of network operations with enhanced intelligence, efficiency, and adaptability. AI-enabled interfaces, coupled with distributed computing, will pave the way for end-to-end AI applicability, fostering convergence across the telecommunications and computing domains.

Drawing from the IMT-2030 vision, the integration of AI into mobile networks unlocks transformative capabilities. Advanced human-machine interfaces, such as extended reality (XR) displays and haptic sensors, offer immersive experiences, blurring the boundaries between physical and virtual realms. Intelligent machines, equipped with machine perception and autonomous capabilities, facilitate seamless interactions and drive digital economic growth and societal changes.

## Call for Action

[Beth Cohen](#)

*0.25 page*

The Linux Foundation's role in joining the AI revolution underscores the importance of open-source collaboration in advancing network capabilities. By harnessing open-source technologies, network operators can leverage the collective expertise of the community to accelerate innovation and adoption of AI-driven solutions. This collaborative approach not only democratizes access to advanced AI capabilities but also fosters interoperability and scalability across diverse companies and network environments.

The future of Intelligent Networks and AI adoption in the telecom industry is in the hands of the individuals and organizations who are already contributing to projects and initiatives, and those who will join them. If you are involved in building and operating networks, developing network technology or consuming network services, you are most heartily encouraged to get involved. Engaging with OSS communities is a way to shape the future of networking. Your contribution could be small or large, and does not necessarily involve writing code. In fact the community is very much in need of contributors of white papers such as this one, evangelists and big thinkers who want to drive the realization of some really cool and useful leading edge technologies. Some of the ways to contribute include:

- Attending Project meetings
- Attending Developer events
- Joining approved Projects
- Proposing a new Project
- Writing documentation
- Contributing use cases
- Analyzing requirements
- Defining tests / processes
- Reviewing and submitting code patches
- Building upstream relationships
- Contributing upstream code
- Start or join a User Group
- Hosting and staffing a community lab
- Answering questions
- Giving a talk / training
- Creating a demo
- Evangelizing the projects

Ways to get involved in Intelligent Networking and AI:

- Collaborate on Network Super Blueprints or develop new ones: <https://wiki.lfnetworking.org/x/ArAZB>
- Work on the [Thoth](#) project - Telco Data Anonymizer Project
- Join LFN AI mailing list: <https://lists.lfnetworking.org/g/lfn-ai-taskforce>

## Final Thoughts

In conclusion, the future of networks in the era of 6G and beyond hinges on the transformative power of AI, fueled by open-source collaboration. By embracing AI-driven intelligence, networks can enhance situational awareness, performance, and capacity management, while enabling quick responses to address undesired states. As we navigate this AI-powered future, the convergence of technological innovation and open collaboration holds the key to unlocking boundless opportunities for progress and prosperity in the telecommunications landscape.

## References

- [1] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- [2] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018).
- [3] Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." *arXiv preprint arXiv:1301.3781* (2013).
- [4] Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." *Advances in neural information processing systems* 26 (2013).
- [5] Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. "Glove: Global vectors for word representation." *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 2014.
- [6] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- [7] Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." *arXiv preprint arXiv:1412.3555* (2014).
- [8] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." *arXiv preprint arXiv:1409.0473* (2014).
- [9] Jigsaw Multilingual Toxic Comment Classification Kaggle Competition: <https://www.kaggle.com/competitions/jigsaw-multilingual-toxic-comment-classification>
- [10] TensorFlow 2.0 Question Answering Kaggle Competition: <https://www.kaggle.com/competitions/tensorflow2-question-answering>
- [11] Huggingface Open LLM Leaderboard: [https://huggingface.co/spaces/HuggingFaceH4/open\\_llm\\_leaderboard](https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard)
- [12] Huggingface Massive Text Embedding Benchmark (MTEB) Leaderboard: <https://huggingface.co/spaces/mteb/leaderboard>
- [13] <https://chat.openai.com>
- [14] <https://gemini.google.com>
- [15] <https://grok.x.ai>



- [16] Li, Raymond, et al. "StarCoder: may the source be with you!." *arXiv preprint arXiv:2305.06161* (2023).
- [17] Lozhkov, Anton, et al. "StarCoder 2 and The Stack v2: The Next Generation." *arXiv preprint arXiv:2402.19173* (2024).
- [18] Nijkamp, Erik, et al. "Codegen: An open large language model for code with multi-turn program synthesis." *arXiv preprint arXiv:2203.13474* (2022).
- [19] Nijkamp, Erik, et al. "Codegen2: Lessons for training llms on programming and natural languages." *arXiv preprint arXiv:2305.02309* (2023).
- [20] Azerbayev, Zhangir, et al. "Llemma: An open language model for mathematics." *arXiv preprint arXiv:2310.10631* (2023).
- [21] Kaggle LLM Science Exam Competition: <https://www.kaggle.com/competitions/kaggle-llm-science-exam>
- [22] BabyAGI: <https://github.com/yoheinakajima/babyagi>
- [23] Ouyang, Long, et al. "Training language models to follow instructions with human feedback." *Advances in neural information processing systems* 35 (2022): 27730-27744.
- [24] Borgeaud, Sebastian, et al. "Improving language models by retrieving from trillions of tokens." *International conference on machine learning*. PMLR, 2022.
- [25] Izacard, Gautier, and Edouard Grave. "Leveraging passage retrieval with generative models for open domain question answering." *arXiv preprint arXiv:2007.01282* (2020).
- [26] Yao, Shunyu, et al. "Webshop: Towards scalable real-world web interaction with grounded language agents." *Advances in Neural Information Processing Systems* 35 (2022): 20744-20757.
- [27] Gao, Luyu, et al. "Pal: Program-aided language models." *International Conference on Machine Learning*. PMLR, 2023.
- [28] Wang, Ruoyao, et al. "Behavior cloned transformers are neurosymbolic reasoners." *arXiv preprint arXiv:2210.07382* (2022).
- [29] The Pile Dataset: <https://pile.eleuther.ai/>
- [30] Huggingface platform: <https://huggingface.co/>
- [31] Huggingface platform statistics: <https://originality.ai/blog/huggingface-statistics>
- [32] Kaggle platform: <https://www.kaggle.com/>
- [33] Paperswithcode platform: <https://paperswithcode.com/>
- [34] Le, Srivatsa, et al. "Rethinking Data-driven Networking with Foundation Models" *HotNets '22, November 14–15, 2022, Austin, TX, USA*
- [35] Le, Franck, et al. "NorBERT: NetwOrk Representations Through BERT for Network Analysis & Management." *2022 30th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*. IEEE, 2022.
- [36] Geospatial Foundation Model: <https://research.ibm.com/blog/geospatial-models-nasa-ai>