



Advancing AI: Networks that self-operate

Self-operation via digital and physical AI agents

White paper

Authors: Anne Lee, Ahmet Akyamac, Dan Kushnir, Daniele Quercia, Edyta Bogucka, Huseyin Uzunalioglu, Jin Cao, Johann Daigremont, Lisa Zhang, Matthew Andrews, Mohamed Trabelsi, Sara Ayoubi and Sean Kennedy

Networks that self-operate (NSOs) are set to revolutionize communications via agentic AI agents in combination with AI, digital twins, spatial intelligence, and semantic communications. They will enable autonomous, adaptive networks that reduce bandwidth, enhance operational efficiency, and simplify management through conversational interfaces. While facing challenges like model complexity and data quality, NSOs promise proactive, intelligent systems that could potentially extend beyond network operations, driving Industry 5.0. Strategic, robust implementation is critical to overcoming barriers and achieving a truly self-operating infrastructure. In this white paper, we examine the technologies at the heart of NSOs, their potentials and limitations, as well as the challenges and opportunities that they present to the networking industry.

Contents

Introduction	3
Agentic AI for NSOs	3
Network agents	6
World agents	8
Semantic communications	10
Foundational goals for self-operating networks	12
State of the business	13
Autonomous systems	13
Large world models (LWMs) and large world agents (LWAs)	14
Semantic communications	15
Lessons learned: limitations and future directions	16
The road to self-operating networks	17
Abbreviations	18
References	19

Introduction

Communications networks have been evolving for over a century, first to enable direct human voice interactions and then to drive data communications and the internet revolution with its endless appetite for network connectivity. The introduction of the public cloud for media streaming, e-commerce and XaaS services reshaped the role of the network once again. Now, with the introduction of AI, with its requirements for flexible cloud processing from the far edge to the core, the network is being challenged to provide a highly responsive and agile connectivity layer for a cloud continuum of compute and storage in support of AI.

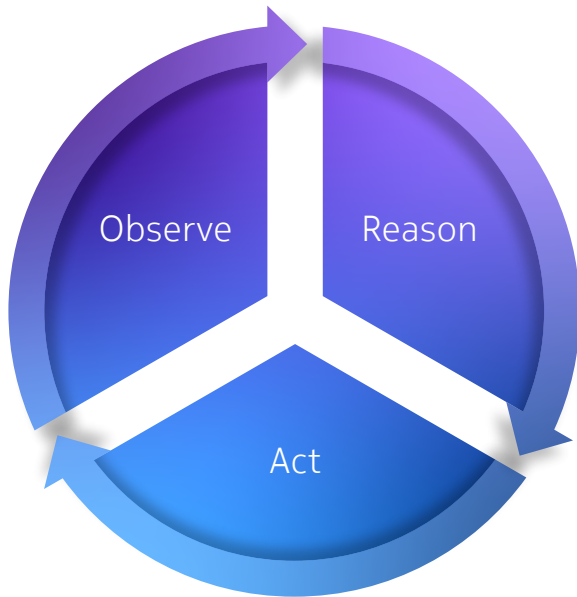
This highly distributed cloud layer will be increasingly called on to support AI-driven applications requiring discrete services over end-to-end network slices. These services need to be dynamically nailed up and taken down in near-real time while meeting the diverse performance requirements of an uncountable number of new use cases and applications.

These AI-driven dynamic service requirements usher in a new stage in the evolution of network operations. These demands now exceed the capabilities of yesterday's mostly human-operated networks. Network automation will be essential to meeting the agility and resilience required to support the new AI-driven use cases, many of which will be business and mission critical. Just as these new requirements are being driven by the application of AI for meeting every conceivable use case, AI could potentially extend beyond network operations driving Industry 5.0 applications and adoption.

Agentic AI for NSOs

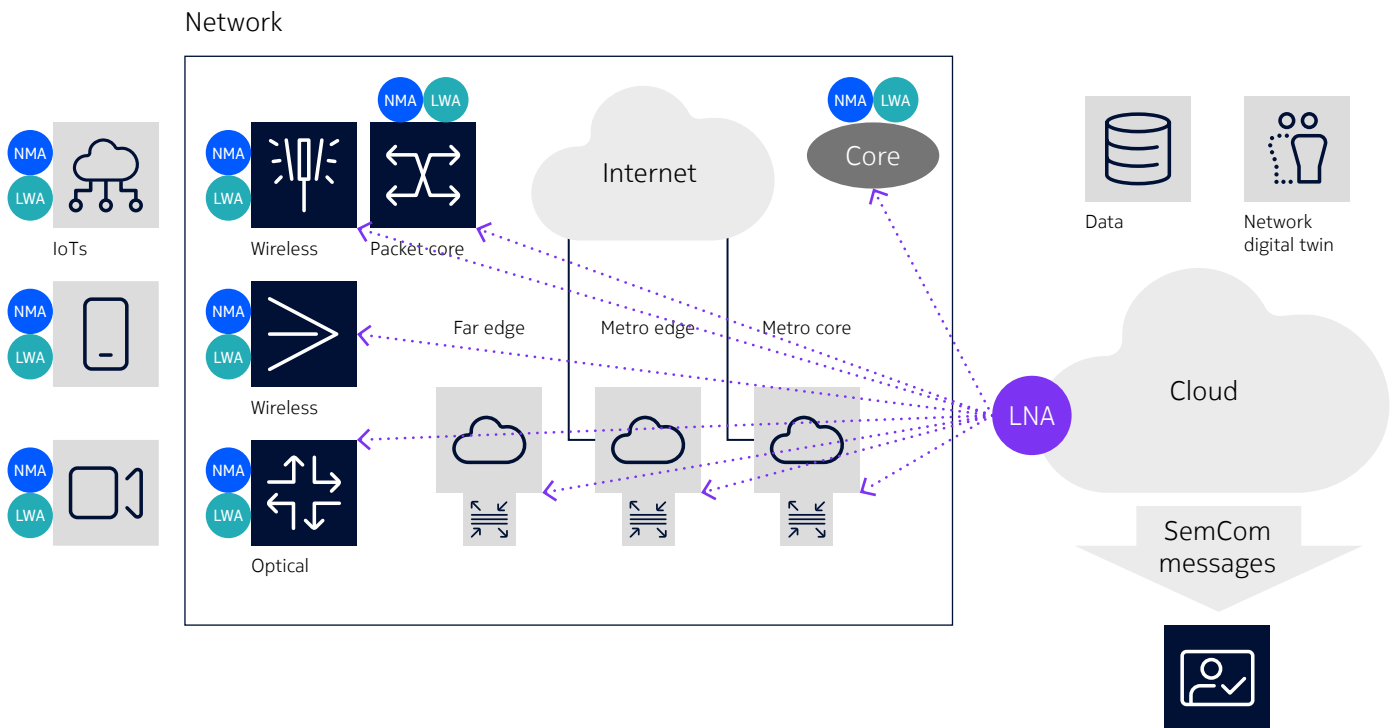
AI chatbots have evolved into AI systems capable of observing, reasoning and acting through function calls and tools. A general AI agent is a system that uses AI models as its core to perceive or observe its environment and reason about the information it processes, including making decisions and taking actions to achieve specific goals through function calls and tools. From general AI agents emerged agentic AI agents, a special type of AI agent that exhibits autonomous decision-making and self-directed behavior with the capacity to set its own goals, within limits, often with a high degree of independence. Agentic AI agents are more advanced, with the ability to modify their objectives, strategies and workflows without direct human input. Both AI agents and agentic AI agents are poised to reshape communications networks.

Figure 1. An agentic AI agent observes, reasons, then takes action



To meet the demands of the AI era, as we've argued, networks will need to self-operate. They will also need to leverage AI, especially agentic AI agents such as large network agents (LNAs), network module agents (NMAs) and large world agents (LWAs), as illustrated in figure 2, to meet the connectivity needs of the AI era and operate with little or no human intervention required.

Figure 2. Networks that self-operate (NSOs) leverage LNAs, NMAs and LWAs

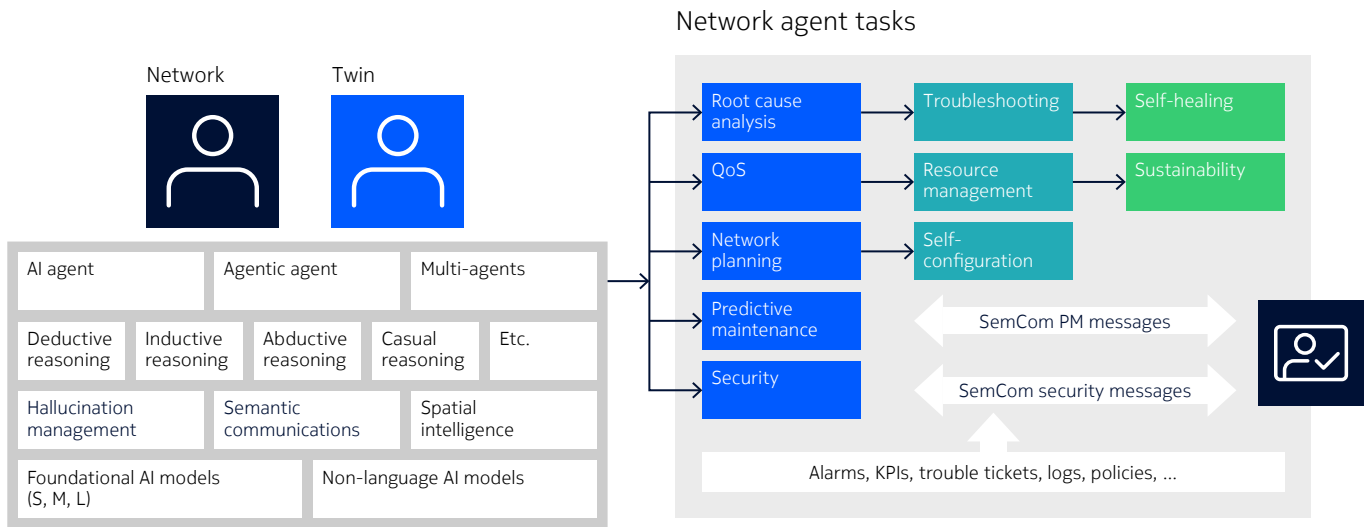


LNAs, NMAs and LWAs are agentic AI agents designed and trained for network operations tasks that run autonomously. These agents also include semantic communications (SemCom) capabilities for enhanced, contextual, intent-based, efficient messaging for the network system. LWAs include spatial intelligence, which extends to the physical world in 3D [1]. Refer to table 1 and figure 3 for an overview of the scope of AI agents used in network operations.

Table 1. Agents at a glance

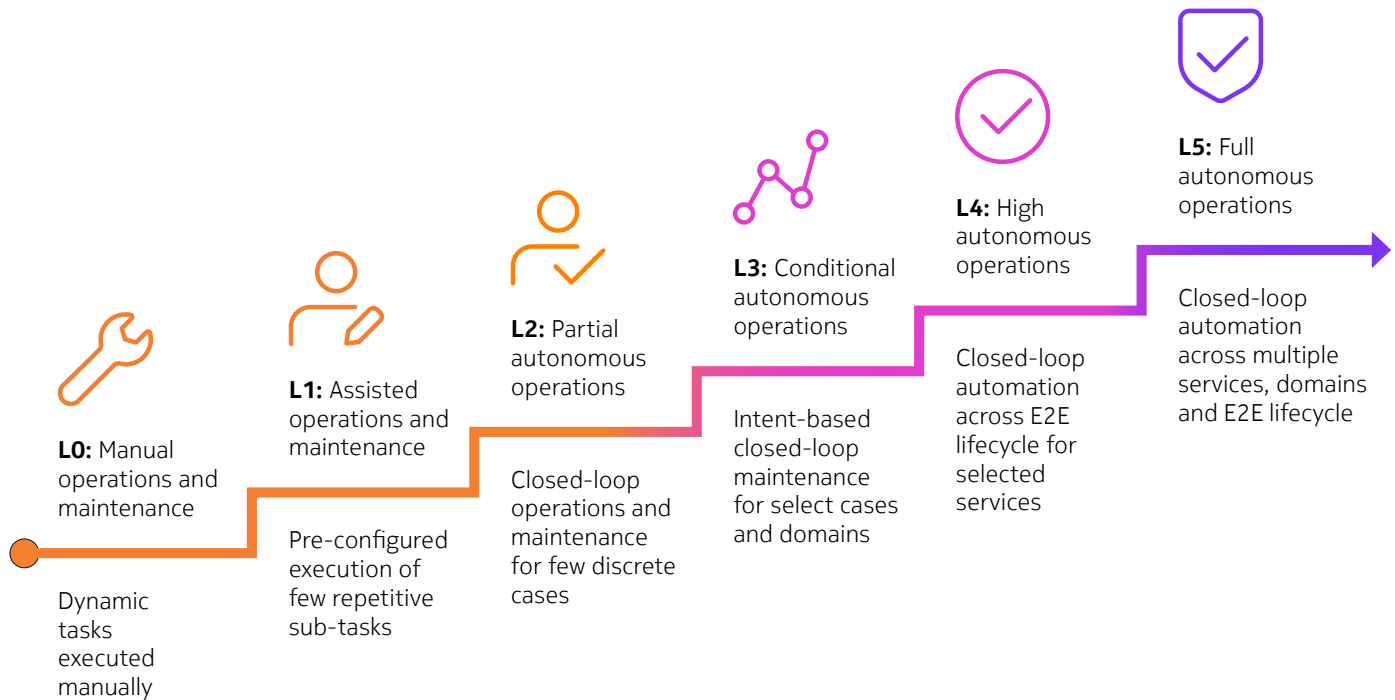
Agent	Runs	Inputs	Decides	Example action
NMA	Node/module (RAN, core, optical)	KPIs, alarms, local logs	Per-function (e.g., RCA)	Parameter tune, self-heal
LNA	Domain/service layer	Aggregated NMAs, topology, twin	Cross-module orchestration	Autonomous troubleshooting, re-configuration
LWA	Edge/device or cloud	Video, sensors, 3D scene	Physical tasks and navigation	Guide inspection drone

Figure 3. Scope of AI agents for network operations



The TM Forum has defined five levels of network automation, with Level 5 being fully autonomous operations—providing closed-loop automation across multiple services, domains and the end-to-end lifecycle (figure 4).

Figure 4. TM Forum's five levels of network automation



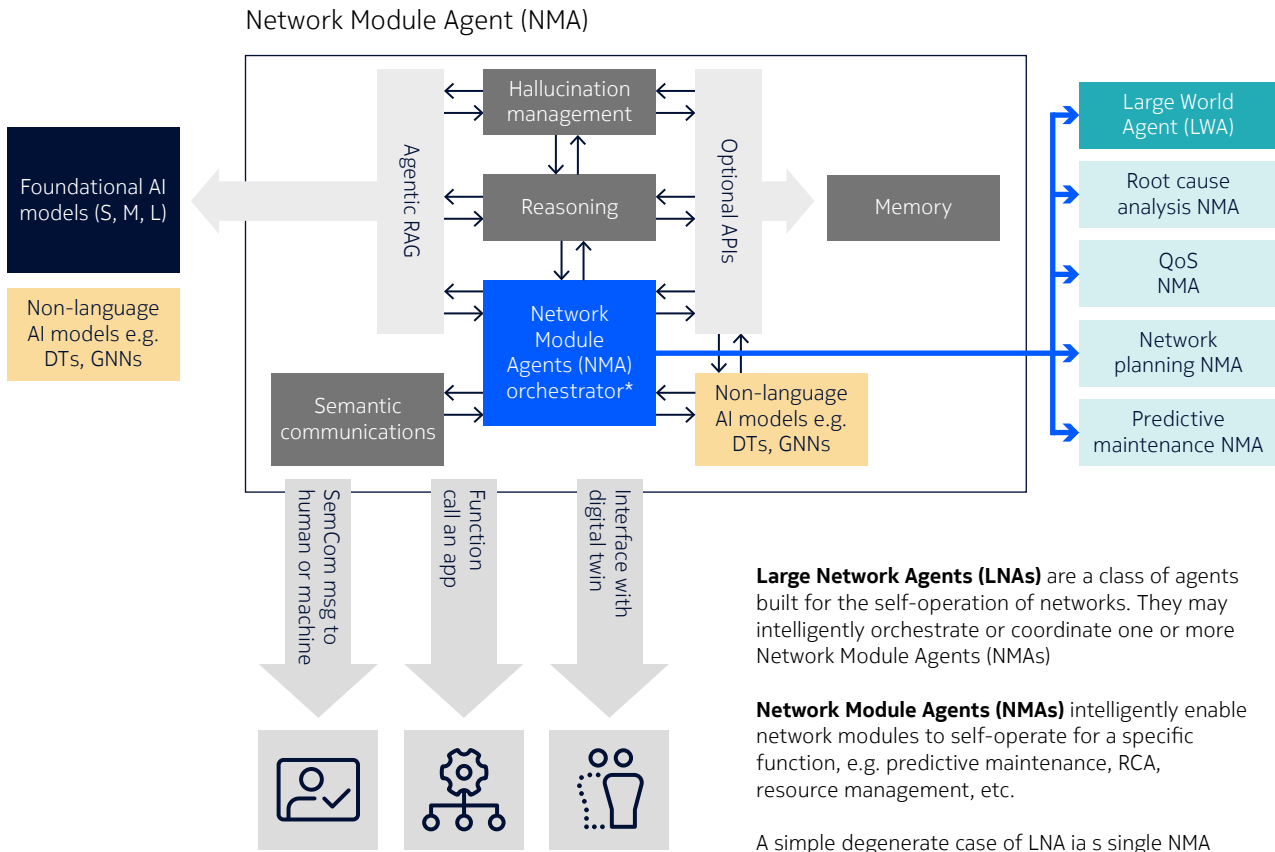
Network agents

LNAs, NMAs and LWAs can help realize the TM Forum's goal of fully autonomous operations at Level 5 and beyond [2]. LNAs can reside in network nodes, e.g., the RAN, core, packet core, or their sub-components, and act as NMAs performing specific operation tasks such as root cause analysis (RCA), predictive maintenance (PM), self-healing, etc. LNAs can reside at a higher-level, interacting and managing a number of NMAs. NMAs can also interact with each other directly.

An LNA is an agentic AI agent, with an LNA orchestrator at its core, coordinating between its components to collectively accomplish a specific task, such as managing multiple NMAs. Included in the LNA are components for (see figure 5):

1. Reasoning
2. Hallucination management
3. Memory
4. SemCom
5. RAG (retrieval augmented generation) functions to retrieve data
6. API calls to foundational AI models
7. Non-language AI models
8. Connections to other NMAs and LWAs, function calls to third-party apps/tools, digital twins, and humans via SemCom utilizing protocols such as Model Context Protocol (MCP), Agent2Agent (A2A) and Agent Communication Protocol (ACP).

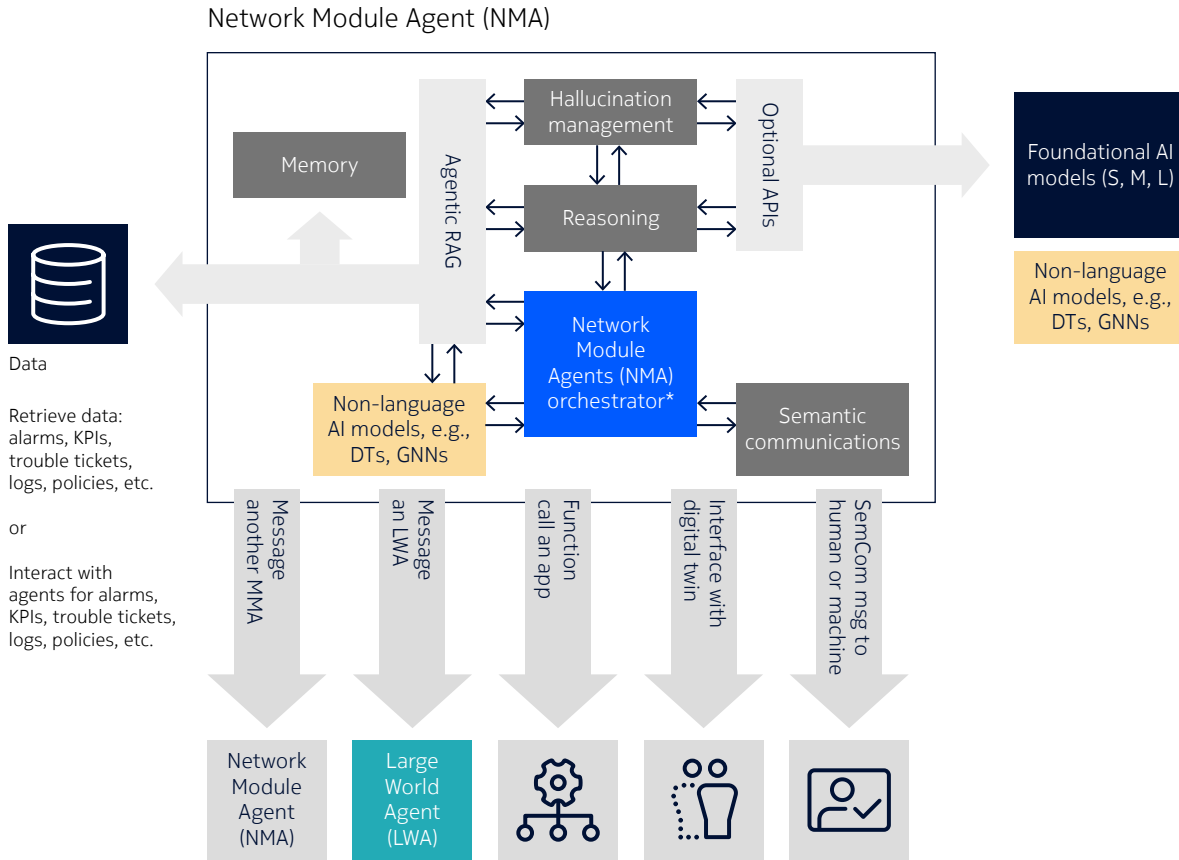
Figure 5. Large network agents (LNAs)



* NMA orchestrator, e.g., RCA, RM, Resource management, etc.

NMAs are agentic AI agents that intelligently enable network modules to self-operate for a specific function or perform a specific task, e.g., PM, RCA, resource management, etc. A simple, degenerate case of an LNA is a single NMA (see figure 6).

Figure 6. Network module agent (NMA) architecture



For each task outlined in figure 3, which details the scope of AI agents for network operations like RCA, PM and troubleshooting, the specifics of the agent components—such as reasoning, hallucination management and SemCom—will vary between different agents. They will be fine-tuned with appropriate data for a specific task and, in some cases, may also be architected differently from one another. Thus, while the overall framework remains consistent, the specifics of each component may vary.

World agents

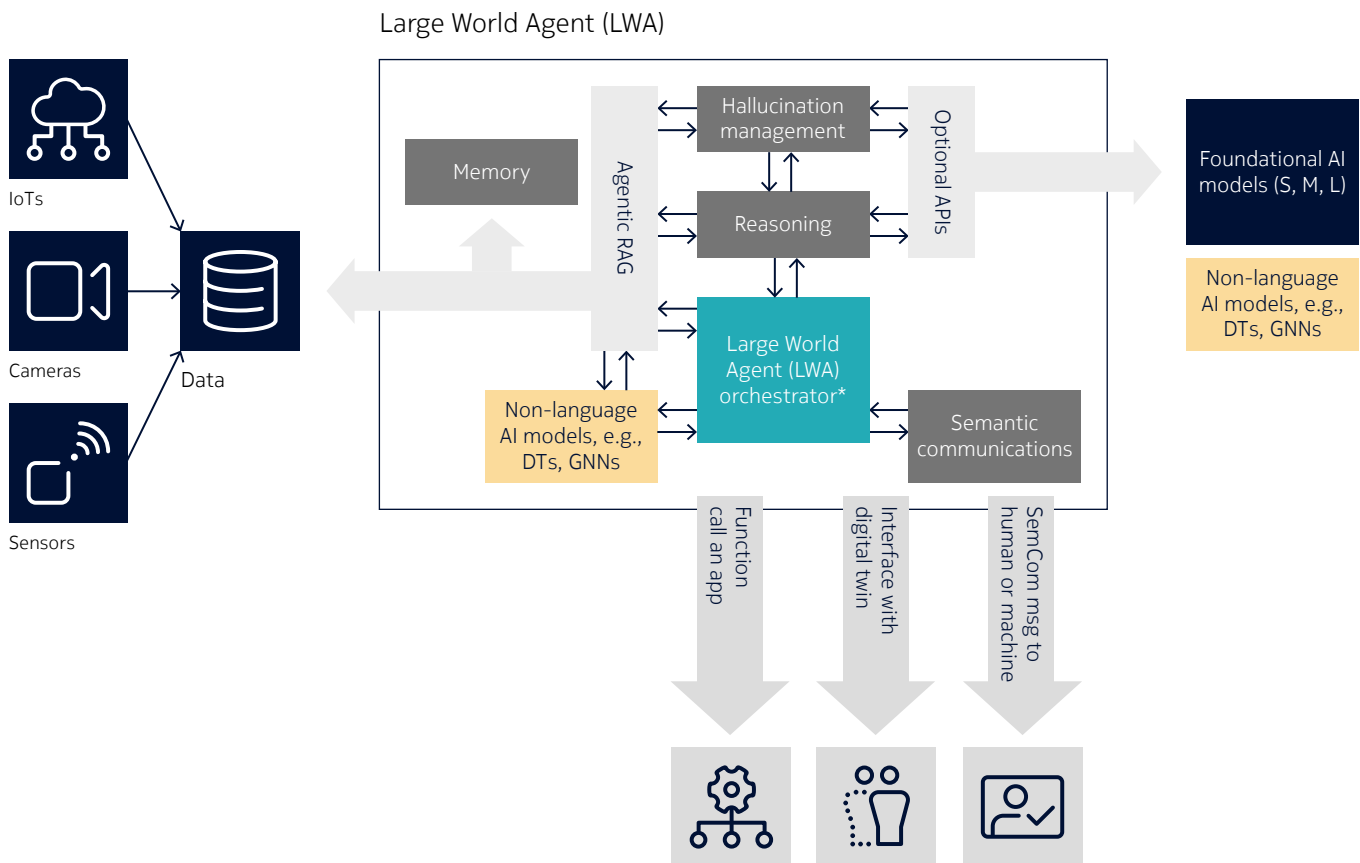
LWAs are agentic AI agents with spatial intelligence that take AI from 2D to 3D. Early AI agents lived primarily in a 2D digital world. LWAs will move AI into a real physical or simulated world—supercharging digital-physical fusion. LWAs and LWMs are AI systems designed to mimic and understand the physical world of a select space in three dimensions [3, 4].

The LWA framework is an agentic AI framework, thus it has similar types of components as the LNA and NMA, namely an orchestrator, memory, reasoning, hallucination management, non-language models, SemCom, RAG, and optional connections to external models, function calls to external apps, and interfaces to digital twins. The difference is in the details of each component and what tasks they have been designed and trained for.

LWAs may utilize large language models (LLMs), non-language models and LWMs—AI models architected and trained specifically with spatial intelligence to understand the 3D physical world.

LWAs and LWMs aim to capture visual, temporal, spatial and physical aspects of the environment. They are trained on a wide range of data, including text, images, video, and audio that can come from sources such as cameras, IoTs and sensors to build a comprehensive understanding of the selected world or scene. For example, LWMs can generate realistic simulations for virtual reality or help robots navigate physical spaces. Input videos for training will include key characteristics of real-world settings and features such as physics and object interaction, thus making possible the extension of AI models into the 3D space—both virtual and physical. The language aspects of AI models will make user interactions with communication network systems natural and intuitive.

Figure 7. Large world agent (LWA) architecture



LWAs and LWMs will reside in one of two places. LWAs and LWMs will either reside in the cloud to oversee one or more devices, or they could be located within the devices themselves. These devices could be tablets, wearables such as AR/VR glasses or headsets, cameras, autonomous land or aerial vehicles, and sensors, actuators and controllers.

Digital twins (DTs) will utilize the data surrounding LWAs and LWMs regardless of their location in the cloud or on devices. DTs aid in the automation of operations. Data collected and used for AI training and inferencing for LWAs and LWMs may also be used by DTs. Additionally, DTs may generate synthetic data to feed into the AI agents when it is appropriate. In the future, LWAs and LWMs may be utilized within DTs as well.

LWAs and LWMs may one day help to achieve a longstanding goal of AI: creating agents that can thoroughly understand complex, constantly changing physical environments, which require ongoing, continuous

learning. With respect to networks, to completely realize the long-term goal of autonomous networking, AI agents for networks must have the capacity to act in the physical domain, not just monitor or analyze it. For instance, if a network anomaly is detected that requires physical inspection or adjustment to hardware, the LNA could dispatch an autonomous system utilizing an LWA to act in the physical space. This necessitates environmental awareness (via LWAs and LWMs) to dispatch the right agent, on the right path, to the right physical location to carry out the task on-site. In this context, a DT of the network and its physical space plays a crucial role by offering a simulated environment where multiple hypotheses about the root cause can be tested safely and efficiently. The LNA can evaluate possible courses of action in the digital space, selecting the most promising one before committing to real-world execution. The ultimate autonomous, self-operating network is based on a tighter integration of logical reasoning, physical embodiment and virtual simulation.

Semantic communications

The integration of SemCom enhances NSOs that utilize LNAs, NMAs and LWAs. SemCom will transform interactions by prioritizing understanding and transmitting meaning versus merely relaying data. This feature applies to machine-to-human (M2H), human-to-machine (H2M), machine-to-machine (M2M), and human-to-human (H2H) interactions. Neural networks (NNs or LLMs) recognize patterns, memorize information and reason through data. This technique makes possible the realization of SemCom through the recognition and conveyance of only key information relevant to the associated interactions, thus minimizing the amount of transmitted data. SemCom improves bandwidth efficiency, latency, robustness in challenging environments, and energy optimization, and it can provide context awareness. But, perhaps, more importantly, SemCom reduces the time to understand a situation, decide on a course of action and execute.

With respect to communications networks, SemCom can be applied to a multitude of use cases. The following examples illustrate each type.

M2H

Predictive maintenance is an excellent use case example. Network equipment has alarm indicators and, sometimes, sensors that continuously monitor various attributes, such as temperature and vibration. AI agents at the edge can process data locally and semantically communicate only the meaningful insights to human field engineers, instead of sending raw data from all the respective sources; for example, they might send a message like, “Cell Tower A’s component needs replacement within 24 hours”. Instead of transmitting temperature and vibration measurements at a constant and continuous rate, AI agents process the data at the edge and send only the insights.

H2M

For H2M scenarios, a human engineer could reconfigure a base station by sending a SemCom message, “Optimize Base Station B for highest QoS levels over maximizing capacity”. Receiving the message, the AI agent for Base Station B will set new values for the appropriate parameters.

M2M

One example of M2M SemCom messaging is in the area of autonomous predictive maintenance. An AI agent that monitors GPU cards could predict a failure in the coming week and send a message to the supply chain AI agent, stating, “Order 10 XYZ vendor GPU cards for Data Center 3 Cabinet 51 Rack 2 Slots 5, 6 and 7”. The supply chain AI agent then orders the cards and sends another message to the data center technician task assignment AI agent, saying, “Schedule engineer to swap out GPU cards for Data Center 3 Cabinet 51 Rack 2 Slots 5, 6, 7 on May 11”.

H2H

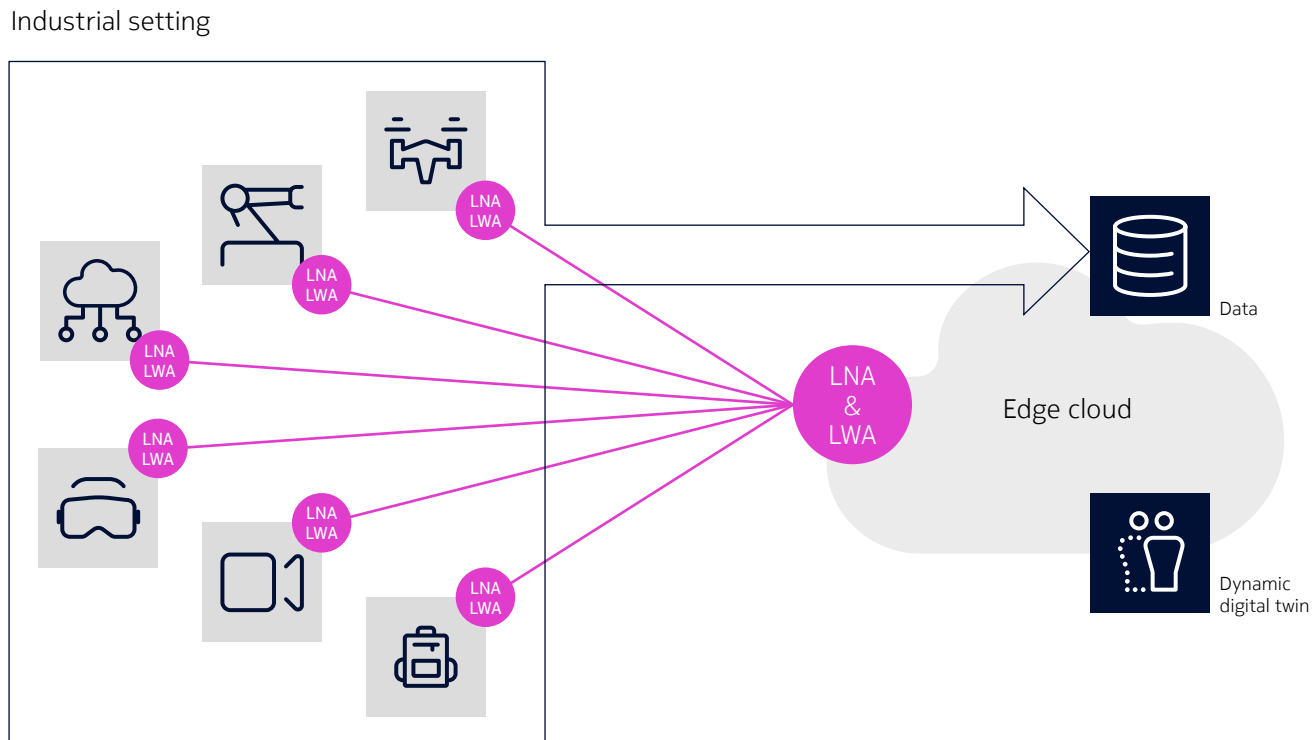
For an H2H SemCom example consider a remote expert assistance use case leveraging LWAs. Let's say a field engineer encounters an issue with the DSL cross-connect box in a customer's backyard that requires specialized knowledge to resolve. The field engineer may use a smart tablet device powered by an AI agent with semantic communication capabilities to initiate a video call with a remote expert. The AI agent on the tablet analyzes the video feed from its camera and extracts relevant information, such as the type of box, wiring, and environmental conditions, that it sends to the remote expert. The remote expert receives this SemCom contextual information and provides precise, step-by-step guidance to the field engineer, possibly using augmented reality overlays to highlight specific components and actions.

NSOs will consist of LNAs, NMAs, and LWAs automating operation tasks throughout the entire system and communicating with each other and human engineers as needed. NSOs will also allow human engineers to query the network for information about the status as well as send commands to it through a conversational interface. In short, NSOs will support the ability for human engineers to talk to the network, semantically or otherwise.

NSOs can be either public networks operated by communications service providers (CSPs) or private wireless networks managed by industries, enterprises or other entities managing networks on behalf of enterprises (see figure 2).

For private wireless networks that operate in factories or warehouses, LNAs, NMAs and LWAs have the potential to be extended to support operations beyond the network to that of the factories and warehouses themselves. Refer to figure 8 for further details. Expanding the capabilities of these AI agents to manage the operational tasks of industrial settings could potentially propel the next industrial revolution, Industry 5.0 [5].

Figure 8. Private wireless networks extended into the industrial setting



Foundational goals for self-operating networks

Networks that self-operate (NSOs) incorporate digital twins with three goals (1) agentic AI agents for network operations tasks, e.g., LNAs and NMAs; (2) agentic AI agents to extend into the third dimension with spatial intelligence, e.g., LWAs; and (3) semantic communications.

Critical to this transformation are AI agents incorporating AI models with the following goals:

- Reasoning
- Planning
- Autonomy
- Real-time, continuous data collection from a diverse array of sources
- Self-checking and self-correction
- Problem solving
- Hallucination mitigation.

Achieving these goals will lead to a future where engineers will be able to talk to their network, troubleshooting will become autonomous, and AI agents at the edge will make possible semantic communications with network nodes and modules.

There are enormous benefits in automating networks. The intelligent fusion of digital twins with the three goals of network operation tasks, spatial intelligence, and semantic communications for networks will take us there. This integration promises to enhance operational efficiency and reduce downtime significantly. By leveraging these advanced technologies, organizations can expect more resilient networks that proactively adapt to challenges and ensure seamless connectivity.

The guiding principles are:

1. Multimodal integration and contextual awareness—seamless integration and reasoning across multiple modalities to provide context-aware insights or actions
2. Explainability, safety and compliance—prioritization of safety and adherence to industry standards ensuring explainable decision-making and incorporating human-in-the-loop validation for critical actions
3. Robustness and reliability—high reliability for edge cases and implementation of fail-safe mechanisms
4. Scalability and real-time performance—support for high-data throughput with real-time inferencing available on devices or the edge cloud as well as the ability to scale to multiple facilities and tasks
5. Data privacy and security—protection of sensitive network data and network AI models.

NSOs, LNAs, NMAs, and LWAs must also be adaptable and fine-tunable to specific operation tasks. This adaptability ensures that as operational needs evolve or new tasks are introduced, the systems can seamlessly integrate changes without compromising performance.

State of the business

Digital twin (DT) work is ongoing across the different network domains: 6G, fixed wireless access, and optical networks. These DTs are intended to be virtual, real-time replicas of the physical networks with the following key attributes:

1. Real-time replication—at any snapshot in time, the DT must always be an accurate reflection of the state of the network
2. Modeling hypothetical scenarios—the DT should be able to simulate network performance, predict behavior based on changes, and answer “what if” questions
3. Synthetic data production—DTs can augment data that is required to train models that would run in the physical network.

The NextG Alliance defines three types of DTs: network DTs (NDTs), industrial automation DTs (IADTs), and smart cities DTs (SCDT) [6].

The following subsections provide the state of the business for each pillar (i.e., network operations, spatial intelligence, and semantic communications).

Autonomous systems

2025 is being called the year of the agent. The excitement surrounding the potential of agentic AI agents is palpable, as foundational AI model vendors, enterprises across various industries, and startups are actively advancing their projects.

Autonomous systems are envisioned to seamlessly integrate into human environments, enhancing productivity, safety and quality of life. This involves software solutions and devices that are not only capable of performing complex tasks with high precision but are also efficiently learning, adapting and collaborating with humans and other systems in dynamic settings.

Agentic AI agents are essential for developing flexible autonomous systems, which include networks that are capable of self-operation. Agentic AI agents, regardless of goal(s), consist of the following components and capabilities:

- Orchestrator
- Reasoning
- Guardrails
- Memory
- Interfaces to other Agentic AI Agents, AI models, function calls, and more.

With respect to the interfaces’ capabilities, de facto open standard protocols for interactions between AI agents as well as AI agents with external data or tools are emerging. They include:

1. MCP—an open-source framework developed by Anthropic and released in November 2024 to enable any AI model to interface with any compliant tool or data source [7]
2. A2A—an open protocol launched by Google in April 2025 with support and contributions from more than 50 technology partners, A2A allows multi-vendor, multi-framework AI agents to coordinate and collaborate on tasks, complements MCP, and is governed by the Linux Foundation [8]

3. ACP—an open protocol standard originally released by IBM in May 2025, ACP enables AI agents to communicate with each other, complements MCP, and is also governed by the Linux Foundation [9]
4. ANP (Agent Network Protocol)—an open protocol framework designed for the agentic web, ANP implements decentralized identity authentication, allowing any two agents to connect, and designs an agent description specification for more efficient data exchange and collaboration among agents.

Protocols for AI agent interaction are evolving rapidly, and the number of agentic AI agents and frameworks appearing in the market is growing exponentially.

One agentic AI agent framework that could be of interest to the communications network industry is from Nvidia. In January 2025, Nvidia released AI Blueprints, which enables enterprises to build agentic AI applications capable of performing multi-level task executions [10]. One of those blueprints is the “AI Agent for Telecom Network Configuration Planning” [11]. This agentic AI agent automates and optimizes the configuration of RAN parameters.

Other frameworks are also emerging, such as the agent development kit (ADK), which is an open-source framework for designing agents built on the same framework that powers [Google Agentspace](#) and the [Google Customer Engagement Suite](#) (CES) agents.

Large world models (LWMs) and large world agents (LWAs)

While foundational LWMs are emerging, researchers such as Yann LeCun estimate that it may take ten years before LWMs can fully “understand the world; [machines] that can remember things, that have intuition, have common sense—things that can reason and plan to the same level as humans” [12][13]. Nevertheless, practical deployable LWMs or LWAs may be possible earlier even if they are not at adult human-level capabilities.

Several companies are developing LWMs for various applications and industries, including, in varying degrees, to support robotics and autonomous vehicle (AV) development [14]. They range from start-ups such as Odyssey and World Labs to Big Tech players such as Google/DeepMind and Nvidia [15]. Additionally, UC Berkeley researchers have also developed an LWM, a multimodal autoregressive model capable of processing sequences up to one million tokens and supporting capabilities such as video question-answering over hour-long YouTube videos [16, 17].

In contrast to today’s LLMs—which can read and understand text, images, and sound—LWMs are designed to utilize continuous data from cameras and IoT sensors, enabling them to understand the real world in motion. Many of these LWMs are built leveraging existing foundational, transformer-based models [18].

Note that most LWMs today focus on virtual worlds versus real physical environments, at least, initially.

1. Nvidia offers an LWM platform called Cosmos that provides data curation tools, pre- and post-guardrails, and a family of pretrained multimodal models that developers can use out of the box for world generation and reasoning [19, 20].
2. DeepMind offers a foundational world model called Genie 3 that can generate an endless variety of action-controllable, playable 3D environments for training and evaluating embodied agents [21]. Additionally, DeepMind has a couple of other generative tools that support robotics and AV. They are AutoRT (a system that harnesses large foundation models and large visual models to better train robots) and ALOHA Unleashed (an AI system to help robots learn to perform complex tasks requiring dexterous movement, such as tying shoelaces or hanging shirts).

3. Odyssey's generative LWM is called Explorer. Explorer is an "image-to-world model" that can convert "any" 2D images into highly detailed 3D virtual worlds. Explorer can generate highly photorealistic virtual worlds, complete with live-action motion. They have not yet announced support for robotics or AV.
4. World Labs is initially focusing on developing spatially intelligent LWMs that can understand and reason about the 3D world using images and other modalities. World Labs' technology generates highly interactive virtual worlds from a single input photograph. World Labs has not announced support for robotics or AVs and plans to launch their first product formally in 2025 [22, 23].

Nokia Bell Labs is currently developing a solution for scene representation of the physical world to accurately represent the spatial relationships between different objects. This applies to real-world environments, in contrast to the virtual ones that other vendors are initially focused on.

Semantic communications

The history of SemCom traces back to Bell Labs' Claude Shannon and his information theory. A recurrent statement that shows up across the literature is that semantic communications make it possible to "move beyond Shannon" or "escape the Shannon trap". This quote comes from Weaver's article published in 1949 that discusses the implications of Shannon's theory for a general audience [24]. In this article, Weaver mentions that, regarding communication, there are problems at three levels:

- Level A: How accurately can the symbols of communication be transmitted?
(The technical problem.)
- Level B: How precisely do the transmitted symbols convey the desired meaning?
(The semantic problem.)
- Level C: How effectively does the received meaning affect conduct in the desired way?
(The effectiveness problem.)

Weaver states that Shannon's information theory addresses the problem at Level A. Moving beyond Shannon refers to the fact that semantic communication aims to address the problem at Levels B and C.

Since the article was published, researchers have evolved the technologies to realize and implement SemCom in industrial settings. In 1994, the Open Platform Communications (OPC) consortium was formed. They released the OPC Unified Architecture (OPC UA) standard in 2006 [25].

OPC UA is a cross-platform, open-source data exchange standard for industrial automation. It's designed to provide secure and reliable communication between various devices, applications and systems, where the devices include sensors, actuators and other automation systems.

OPC UA incorporates semantic communication through its information model, which defines the structure and semantics of the data being exchanged. OPC-UA also comes with a language model that allows the expression of semantics and context in data. Through the information model, clients can replace data pulling with data querying, for example, "Give me the list of sensors with temperatures above 50 degrees". Siemens is an example company that incorporates OPC-UA in some of its programmable logic controllers (PLCs).

SemCom is also being incorporated into holographic communications. Holoconnects is one example of a company that uses semantic communications [26, 27]. Holographic communication is data-intensive, making it expensive. Thus, capabilities to reduce the required bandwidth are welcome.

In addition, standards are currently being defined in 3GPP for SemCom as part of 6G. A key area in 6G SemCom is the development of universal deep learning-based joint source-channel coding (DJSCC), which would use neural networks to simultaneously encode (compress) input data and add redundancy for error correction to assure resiliency and robustness in the transmission of information, especially in noisy environments.

Companies that are working on 6G SemCom include Nokia, Ericsson, Orange, Samsung, Huawei, NTT Docomo and KT.

Semantic communication has transformative potential in revolutionizing human and machine interactions by prioritizing the transmission of meaning over mere data. Nokia Bell Labs is focusing on three key areas: SemCom at the physical (PHY) layer, SemCom in industrial environments, and SemCom for human-to-human communication. By integrating SemCom into network operations, enhancing automated monitoring systems, and leveraging generative AI for richer human interactions, this technology has the potential to drastically reduce bandwidth requirements, improve communication in constrained environments, and enable seamless human-machine collaboration. There are also challenges and strategic priorities in advancing SemCom, which Nokia Bell Labs is solving, including the development of universal DJSCC models, achieving general scene understanding, and fine-tuning vision language models (VLMs) for specialized tasks.

Lessons learned: limitations and future directions

The potential for NSOs leveraging digital twins with LNAs, NMAs, LWAs, and SemCom is enormous. However, it also presents significant challenges that require attention. About autonomous systems, for instance, Gartner has stated that over 40% of agentic AI projects are expected to be canceled by the end of 2027 due to escalating costs, unclear business value or inadequate risk controls [28].

“Most agentic AI projects right now are early-stage experiments or proof of concepts that are mostly driven by hype and are often misapplied,” said [Anushree Verma](#), Senior Director Analyst, Gartner. “This can blind organizations to the real cost and complexity of deploying [AI agents at scale](#), stalling projects from moving into production. They need to cut through the hype to make careful, [strategic decisions](#) about where and how they apply this emerging technology” [28].

In addition to Gartner’s perspectives, some of the key hurdles for autonomous networks are:

1. Decision-making and adaptability—autonomous networks will be required to make intelligent decisions in complex, multi-task environments, and their AI models will need to be capable of planning in real time, prioritizing tasks, adapting to new tasks and handling unexpected events
2. Data and learning—high-quality, diverse datasets are required that may be difficult to obtain or synthetically generate.

Some of the key hurdles for LWAs are:

1. Computational scale—training and running LWAs will demand immense amounts of computational resources and, with it, high-energy requirements; alternatively, pre-trained LWMs could be used and fine-tuned, thereby dramatically reducing the compute and energy needs
2. Model complexity—LWAs must be able to integrate multimodal data and generalize across diverse scenarios

3. Causal reasoning and interpretability—LWAs cannot depend solely on correlation but must have causal understanding for explainability, or interpretability, which will not only create users’ trust but can also be used for debugging and adjusting the LWAs
4. Data and real-time adaptation—LWAs need vast, diverse and high-quality datasets, and although, over time, there will be more foundational LWMs available to fine tune, fine tuning for a particular “world” will still require a large amount of data, especially where data represent a dynamic “world” that is continuously changing, meaning the LWAs will have to continuously take new data into account and learn online
5. Benchmarks—creating benchmarks to measure the accuracy of LWAs is a parallel endeavor that can be quite complex
6. Ethical and safety concerns—guardrails must be created and applied to ensure that LWAs are not misused or cause harm unintentionally and, because LWAs also have the potential to hallucinate, mitigating hallucinations is critical.

Some key hurdles for SemCom are:

1. Complexity of semantic extraction—networks consist of a diverse set of data from many different sources, which means that extracting meaningful semantic information from this data for specific tasks may require advanced AI models that can generalize across different machines and processes
2. Real-time processing requirements—there are some network tasks, such as those being defined in 6G, that require ultra-low latency responses, meaning semantic communications must be able to process and transmit data in real time, which could be hard to achieve due to delay introduced by the computation needed to process the data
3. Reliability in harsh environments—wireless environments can be noisy and prone to interference, thus transmitting semantic information will require robust error protection methods for resilience in the most adverse conditions
4. Data privacy and security—SemCom scenarios may involve sensitive network user data, which must be protected against any attacks
5. Scalability—large networks may support millions of devices, thus SemCom must be able to handle high volumes of data while maintaining efficiency and accuracy
6. Quantifying accuracy—if the communication is task-oriented, accuracy can be assessed by determining whether the task was successfully completed; however, in the absence of a specific task, it becomes difficult to ensure that the transmitted message has conveyed the intended semantic information.

The road to self-operating networks

NSOs represent a transformative vision for the future of communications infrastructure, leveraging the intelligent fusion of DTs with three foundational pillars: LNAs and NMAs for network operations, LWAs for spatial intelligence, and SemCom for enhanced efficiency and understanding.

The integration of these technologies promises to revolutionize network operations by enabling:

- Autonomous decision-making—agentic AI agents that can observe, reason, and act with bounded autonomy will enable networks to achieve unprecedented levels of self-operation and adaptability

- Enhanced operational efficiency—by incorporating semantic communications that prioritize meaning over raw data transmission, NSOs will dramatically reduce bandwidth requirements while improving communication effectiveness across M2H, H2M, M2M, and H2H interactions
- Spatial intelligence integration—LWAs will extend network capabilities into the physical 3D world, enabling autonomous systems to navigate and interact with real-world environments for tasks such as predictive maintenance and physical infrastructure management
- Conversational network interfaces—engineers will be able to naturally communicate with their networks through semantic interfaces, making network management more intuitive and accessible.

The impact of NSOs could extend beyond traditional network operations, particularly in private wireless networks, where these capabilities could drive the fifth industrial revolution (Industry 5.0) by managing not just network infrastructure but entire industrial operations.

Realizing this vision, however, requires addressing significant challenges, including computational scale requirements, model complexity, causal reasoning capabilities, data quality and availability, and the need for robust frameworks. As we've seen, Gartner predicts that over 40% of agentic AI projects may face cancellation by 2027 due to escalating costs and unclear business value, which emphasizes the importance of strategic and measured implementation approaches.

Despite these challenges, the convergence of advanced AI agents, semantic communications, and spatial intelligence represents a critical evolution toward truly autonomous network operations. Success in this domain will require continued innovation in multimodal integration, explainable AI, robustness and reliability, scalability, and data privacy and security.

NSOs mark a fundamental shift from reactive network management to proactive, intelligent systems that can anticipate, adapt and autonomously resolve issues while seamlessly integrating with both digital and physical operational environments. This transformation will ultimately enable networks to move beyond the TM Forum's Level 5 automation goals, creating a truly self-operating communications infrastructure that can support the requirements of the AI era.

Abbreviations

A2A	Agent2Agent Protocol
ACP	Agent Communication Protocol
ADK	Agent development kit
ANP	Agent Network Protocol
API	Application programming interface
AV	Autonomous vehicle
IADT	Industrial automation digital twins
CES	Customer engagement suite (Google)
CSP	Communications service providers
DJSCC	Deep learning-based joint source-channel coding
DT	Digital twin

H2H	Human to human
H2M	Human to machine
LLM	Large language model
LNA	Large network agents
LWA	Large world agents
LWM	Large world models
M2H	Machine to human
M2M	Machine to machine
MCP	Model Context Protocol
NMA	Network module agent
NDT	Network digital twins
NN	Neural network
NSO	Networks that self-operate
OPC	Open Platform Communications
OPC UA	OPC Unified Architecture
PHY	Physical layer
PLC	Programmable logic controllers
PM	Predictive maintenance
RAG	Retrieval augmented generation
RAN	Radio access network
RCA	Root cause analysis
SemCom	Semantic communications
SMDT	Smart cities digital twin
VLM	Vision language model

References

1. World Labs, “Hello World Labs,” World Labs website on spatial intelligence AI, viewed 14 Oct 2025. Online: <https://www.worldlabs.ai>
2. Nokia, “Autonomous Networks: What is the current state and how to move forward,” Nokia web site, 30 Sep 2024. Online: <https://www.nokia.com/bell-labs/bell-labs-consulting/articles/autonomous-networks-what-is-the-current-state-and-how-to-move-forward/>
3. Germanidis, A., “Introducing General World Models,” Runway, 11 Dec 2023. Online: <https://runwayml.com/research/introducing-general-world-models>

4. Moiola, F., "The Next Leap In AI: From Large Language Models To Large World Models?" Forbes, 23 Jan 2024. Online: <https://www.forbes.com/councils/forbestechcouncil/2024/01/23/the-next-leap-in-ai-from-large-language-models-to-large-world-models/>
5. Malayil, J., "World's first humanoid robot swarm tackles complex factory tasks at China's car plant," Interesting Engineering, 30 April 2025. Online: <https://interestingengineering.com/innovation/china-robots-tackle-car-factory-tasks>
6. NextG Alliance, "6G Digital Twins Use Cases and Requirements," NextG Alliance web site, Mar 2025. Online: https://nextgalliance.org/white_papers/6g-digital-twins-use-cases-and-requirements/
7. Anthropic, "Introducing the Model Context Protocol," Anthropic web site, 25 Nov 2024. Online: <https://www.anthropic.com/news/model-context-protocol>
8. Surapaneni, R., et al., "Announcing the Agent2Agent (A2A) protocol," Google for Developers web site, 9 Apr 2025. Online: <https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/>
9. Besen, S. and Gutowska, A., "What is agent communication protocol (ACP)?" IBM web site, viewed 14 Oct 2025. Online: <https://www.ibm.com/think/topics/agent-communication-protocol>
10. Boitano, J., "NVIDIA and partners launch agentic AI blueprints to automate work for every enterprise," Nvidia web site, 6 Jan 2025. Online: <https://blogs.nvidia.com/blog/agentic-ai-blueprints/>
11. Nvidia, "AI Agent for Telecom Network Planning," Nvidia web site, Jan 2025. Online: <https://build.nvidia.com/nvidia/telco-network-configuration/blueprintcard>
12. Noreika, A., "Meta's AI Chief Envisions 'World Models' as the Path to Human-Level AI, But the Goal Remains a Decade Away," Technology.Org, 18 Oct 2024. Online: <https://www.technology.org/2024/10/18/metas-ai-chief-envisions-world-models-as-the-path-to-human-level-ai-but-the-goal-remains-a-decade-away/>
13. Hao, S., et al., "Reasoning with Language Model is Planning with World Model," arXiv:2305.14992, 24 Oct 2023. Online: <https://doi.org/10.48550/arXiv.2305.14992>
14. Hall, C., "Large World Models & Their Importance to Robotics & AVs, Cutter Consortium, 22 Jan 2025. Online: <https://www.cutter.com/article/large-world-models-their-importance-robotics-avs>
15. Hall, C., "AI's Next Frontier: Advancing Large World Models for Robotics & AVs," Cutter Consortium web site, 19 Feb 2025. Online: <https://www.cutter.com/article/large-world-models>
16. Rodriguez, J., "Inside Large World Model: UC Berkeley Multimodal Model that can Understand 1 Hour Long Videos," Medium, 26 Feb 2024. Online: <https://jrodthoughts.medium.com/inside-large-world-model-uc-berkeley-multimodal-model-that-can-understand-1-hour-long-videos-d1a97c5c7fa0>
17. AI Base, "UC Berkeley Researchers Propose Large World Model (LWM) Comparable to Gemini 1.5 Pro," AI News, 20 Feb 2024. Online: <https://www.aibase.com/news/5353>
18. Wiggers, K., "What are AI 'world models,' and why do they matter?" TechCrunch, 14 Dec 2024. Online: <https://techcrunch.com/2024/12/14/what-are-ai-world-models-and-why-do-they-matter/>
19. Nvidia, "NVIDIA Announces Major Release of Cosmos World Foundation Models and Physical AI Data Tools," Nvidia newsroom, 18 Mar 2025. Online <https://nvidianews.nvidia.com/news/nvidia-announces-major-release-of-cosmos-world-foundation-models-and-physical-ai-data-tools>
20. Nvidia, "NVIDIA Glossary/World Foundation Models," Nvidia web site, viewed 14 Oct 2025. Online: https://www.nvidia.com/en-us/glossary/world-models/?_bt=738313126223&_bk=world%20model&_bm=b&_bn=g&_bg=180166076110

21. Parker-Holder, J. and Fruchter, S., "Genie 3: A new frontier for world models," Google DeepMind web site, 5 Aug 2025. Online: <https://deepmind.google/discover/blog/genie-3-a-new-frontier-for-world-models/>
22. Strickland, E., "AI Godmother Fei-Fei Li Has a Vision for Computer Vision," IEEE Spectrum, 12 Dec 2024. Online: <https://spectrum.ieee.org/fei-fei-li-world-labs>
23. Fei-Fei, L., "Fei-Fei Li says understanding how the world works is the next step for AI," The Economist, 20 Nov 2024. Online: <https://www.economist.com/the-world-ahead/2024/11/20/fei-fei-li-says-understanding-how-the-world-works-is-the-next-step-for-ai>
24. Weaver, W., "Recent contributions to the mathematical theory of communication," from Shannon C. E., and Weaver W., The Mathematical Theory of Communication, The University of Illinois Press, Jul 1949. Online: https://pure.mpg.de/rest/items/item_2383164_3/component/file_2383163/content
25. Siemens, "OPC UA – Structured data up to the cloud," Siemens Xcelerator Marketplace, viewed 14 Oct 2025. Online: <https://www.siemens.com/global/en/products/automation/industrial-communication/opc-ua.html>
26. Cheng, R., et al., "Enriching Telepresence with Semantic-driven Holographic Communication," Association for Computing Machinery, HotNets '23 conference proceedings, 28 Nov 2023. Online: <https://dl.acm.org/doi/10.1145/3626111.3628184>
27. Holoconnects, "Next Gen Hologram Company," Holoconnects web site, viewed 14 Oct 2025. Online: <https://www.holoconnects.com/>
28. Gartner, "Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027," Gartner press release, 25 Jun 2025. Online: <https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>

About Nokia

At Nokia, we create technology that helps the world act together.

As a B2B technology innovation leader, we are pioneering networks that sense, think and act by leveraging our work across mobile, fixed and cloud networks. In addition, we create value with intellectual property and long-term research, led by the award-winning Nokia Bell Labs, which is celebrating 100 years of innovation.

With truly open architectures that seamlessly integrate into any ecosystem, our high-performance networks create new opportunities for monetization and scale. Service providers, enterprises and partners worldwide trust Nokia to deliver secure, reliable and sustainable networks today – and work with us to create the digital services and applications of the future.

Nokia is a registered trademark of Nokia Corporation. Other product and company names mentioned herein may be trademarks or trade names of their respective owners.

© 2025 Nokia

Nokia Oyj
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: Ref1701800 (October) CID215133